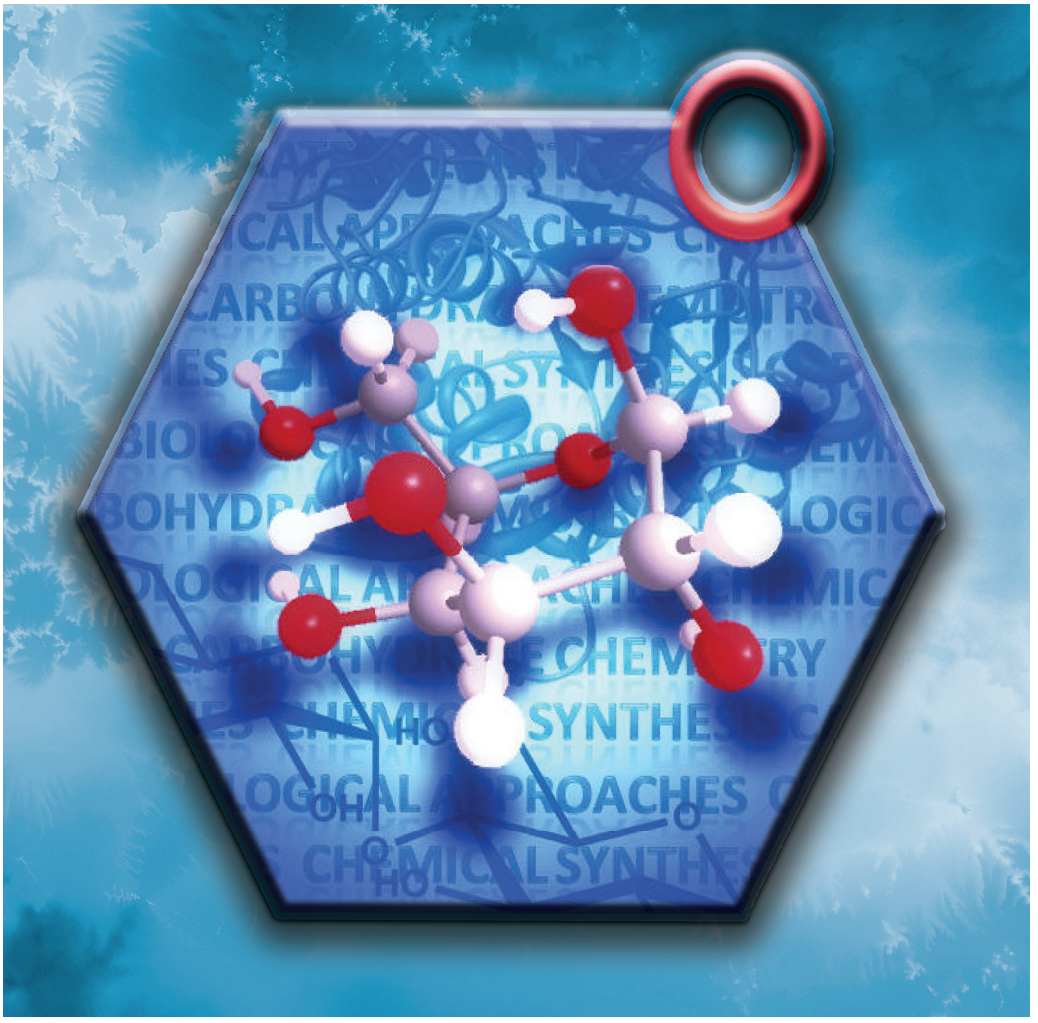
Specialist Periodical Reports

Edited by A Pilar Rauter

Carbohydrate Chemistry

Chemical and Biological Approaches

Volume 38



RSC Publishing

Carbohydrate Chemistry
Chemical and Biological Approaches

Volume 38

A Specialist Periodical Report

Carbohydrate Chemistry
Chemical and Biological Approaches
Volume 38

A Review of the Literature Published between January 2011 and February 2012

Editors
Amelia Pilar Rauter, *Universidade de Lisboa, Portugal*
Thisbe K. Lindhorst, *Christiana Albertina University of Kiel, Germany*

Authors
Tiina Alamäe, *University of Tartu, Estonia*
Marta M. Andrade, *Faculdade de Ciências da Universidade de Lisboa, Portugal*
Ana Ardá, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
Juan M. Benito, *Instituto de Investigaciones Químicas, CSIC - Universidad de Sevilla, Spain*
M. Álvaro Berbís, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
Nele Berghmans, *Rega Institute for Medical Research, Belgium*
Davide Bini, *University of Milano-Bicocca, Milan, Italy*
Pilar Blasco, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
Karin Bodewits, *Ludwig-Maximilians-Universität, Munich, Germany*
Paz Briones, *CSIC, IBC, Sección de Errores Congénitos del Metabolismo, Barcelona, Spain*
Vasco Cachatra, *University of Lisbon, Portugal*
Valérie Calabro, *Aix-Marseille University, France*
Fernando Calais, *Hospital São José, Lisboa, Portugal*
Angeles Canales, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
F. Javier Cañada, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
Alice Capitoli, *University of Milano-Bicocca, Milan, Italy*
Mylène A. Carrascal, *Universidade Nova de Lisboa, Portugal*
Laura Cipolla, *University of Milano-Bicocca, Milan, Italy*
Fabio Dall'Olio, *Università di Bologna, Italy*
Anthony De Soyza, *Newcastle University, UK*
Flaviana Di Lorenzo, *Università di Napoli Federico II, Italy*
Sandrine Donadio-Andréi, *Aix-Marseille University, France*
Nassima El Maï, *Aix-Marseille University, France*
Amalia M. Estévez, *CIQUS, University of Santiago de Compostela, Spain*
Juan C. Estévez, *CIQUS, University of Santiago de Compostela, Spain*
Ramón J. Estévez, *CIQUS, University of Santiago de Compostela, Spain*
Mª Carmen Fernández-Alonso, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*

José G. Fernández-Bolaños, *Universidad de Sevilla, Spain*
Vanessa Ferreira, *Portuguese Association for CDG and other Rare Metabolic Diseases, Portugal*
Luca Gabrielli, *University of Milano-Bicocca, Milan, Italy*
José M. García Fernández, *Instituto de Investigaciones Químicas, CSIC - Universidad de Sevilla, Spain*
Ana M. Gómez, *IQOG-CSIC, Madrid, Spain*
Alejandro González-Benjumea, *Universidad de Sevilla, Spain*
Chloé Iss, *Aix-Marseille University, France*
Jesús Jiménez-Barbero, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
Rosa Lanzetta, *Università di Napoli Federico II, Italy*
Sandra Li, *Rega Institute for Medical Research, Belgium*
J. Cristóbal López, *IQOG-CSIC, Madrid, Spain*
Óscar López, *Universidad de Sevilla, Spain*
Cristina Lupo, *University of Milano-Bicocca, Milan, Italy*
Andres Mäe, *University of Tartu, Estonia*
Filipa Marcelo, *Centro de Investigaciones Biológicas, CSIC, Madrid, Spain*
Karin Mardo, *University of Tartu, Estonia*
Sergio Martos, *Universidad de Sevilla, Spain*
Inés Maya, *Universidad de Sevilla, Spain*
Penélope Merino-Montiel, *Universidad de Sevilla, Spain*
Antonio Molinaro, *Università di Napoli Federico II, Italy*
Francesco Nicotra, *University of Milano-Bicocca, Milan, Italy*
Ana Oliete, *Universidad de Sevilla, Spain*
Ghislain Opdenakker, *Rega Institute for Medical Research, Belgium*
Carmen Ortiz Mellet, *Universidad de Sevilla, Spain*
Stefan Oscarson, *University College Dublin, Ireland*
José M. Otero, *CIQUS, University of Santiago de Compostela, Spain*
Amélia P. Rauter, *University of Lisbon, Portugal*
Catherine Ronin, *Aix-Marseille University, France*
Laura Russo, *University of Milano-Bicocca, Milan, Italy*
Paulo F. Severino, *Universidade Nova de Lisboa, Portugal and Università di Bologna, Italy*
Alba Silipo, *Università di Napoli Federico II, Italy*
Mariana Silva, *Universidade Nova de Lisboa, Portugal*
Raquel G. Soengas, *University of Aveiro, Portugal*
Markus Sperandio, *Ludwig-Maximilians-Universität, Munich, Germany*
Francesca Taraballi, *University of Milano-Bicocca, Milan, Italy*
Clara Uriel, *IQOG-CSIC, Madrid, Spain*
Jo Van Damme, *Rega Institute for Medical Research, Belgium*
Paula A. Videira, *Universidade Nova de Lisboa, Portugal*
Maria-Antonia Vilaseca, *Guia Metabólica, Esplugues de Llobregat, Spain*
Triinu Visnapuu, *University of Tartu, Estonia*
Ulrika Westerlind, *Leibniz Institute for Analytical Sciences, Dortmund, Germany*
Alina D. Zamfir, *National Institute for Research and Development in Electrochemistry and Condensed Matter, and Aurel Vlaicu University of Arad, Romania*

RSC Publishing

If you buy this title on standing order, you will be given FREE access to the chapters online. Please contact E-mail: sales@rsc.org with proof of purchase to arrange access to be set up.

Thank you.

ISBN: 978-1-84973-439-4
ISSN: 0306-0713
DOI: 10.1039/9781849734769

A catalogue record for this book is available from the British Library

© The Royal Society of Chemistry 2012

All rights reserved

Apart from fair dealing for the purposes of research or private study for non-commercial purposes, or for private study, criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988 and the Copyright and Related Rights Regulations 2003, this publication may not be reproduced, stored or transmitted, in any form or by any means, without the prior permission in writing of The Royal Society of Chemistry, or in the case of reproduction in accordance with the terms of the licences issued by the Copyright Licensing Agency in the UK, or in accordance with the terms of the licences issued by the appropriate Reproduction Rights Organization outside the UK. Enquiries concerning reproduction outside the terms stated here should be sent to The Royal Society of Chemistry at the address printed on this page.

Published by The Royal Society of Chemistry,
Thomas Graham House, Science Park, Milton Road,
Cambridge CB4 0WF, UK

Registered Charity Number 207890

For further information see our web site at www.rsc.org

Preface

DOI: 10.1039/9781849734769-FP005

Carbohydrate research plays a remarkable role in chemistry and biology owing to the unique molecular features of the saccharides. They are multifunctional and stereochemically enriched molecules offering a superb structural diversity to serve as molecular scaffolds or key intermediates in the development of new drugs and novel carbohydrate-based or -functionalized materials for a variety of applications. In addition, carbohydrates are the molecular basis for a multitude of biological processes that occur in states of health as well as disease. Understanding the chemistry and biology of this class of molecules will facilitate new options for the treatment of medical conditions for which no cure exists to date. Thus in this volume, glycochemistry and glycobiology topics have been combined to document the latest findings in carbohydrate research and demonstrate the contributions of organic chemistry, modern analytics, biological and biochemical expertise to the increasingly important field of glycomics.

Firstly a modified polysaccharide, chlorite-oxidized oxyamylose, is described as an immunomodulator with therapeutic implications for acute and chronic inflammation, and also cancer. A chapter that focuses on lipopolysaccharide in cystic fibrosis (CF)-related pathogens, namely *Pseudomonas aeruginosa* and *Burkholderia* spp. reports on the structural investigation of these biomolecules, their structural adaptation to the host tissues as well as the importance of their structural features in the clinical management of CF.

Development of carbohydrate vaccines against infections and cancer continues to be a challenge for glycoexperts all over the world. Synthesis of inner core lipopolysaccharides for vaccines against Gram-negative bacterial infections and that of MUC1 glycopeptides conjugated to different immunostimulants with promising results regarding the antibodies induced with these synthetic vaccines against cancer will be highlighted here.

The immune response to invading pathogens during inflammation is crucial in cell biology. Thus, the role of carbohydrate decorations in leukocyte recruitment will be reviewed, putting a strong focus on post-translational modification by sialic acids. New findings emphasizing the influence of carbohydrate decoration on the regulation of inflammatory response will be discussed. New interesting therapeutic approaches in the treatment of acute and chronic inflammatory diseases are being offered.

Recent progress on glycoengineering, an advanced technology based on a glycosylation strategy to optimize protein drugs, is reported. Impact of N-glycosylation on therapeutic proteins, namely that of glycans on drug bioavailability, glycoprotein biopotency, drug immunogenicity, protein folding and epitope expression as well as production, safety and regulatory aspects are discussed. Major advances in glycoengineering are reported with an emphasis on N-glycosylation control to identify the conditions that promote an optimal glycoform profile and reproducibility of glycomodification. In addition, identification and cloning of glycosyltransferase genes

© The Royal Society of Chemistry 2012

are described as new tools for manipulation of expression systems in order to further improve the glycoform profile, in particular engineering of core fucosylation and sialylation.

Glycosylation has significant implications in health and disease. Hence, one chapter is dedicated to congenital disorders of glycosylation, which are a group of disorders of abnormal glycosylation of N-linked oligosaccharides caused by deficiency in 29 different enzymes in the N-linked oligosaccharide synthetic pathway. The relevance and key aspects of the glycosylation changes associated with bladder cancer are highlighted in another chapter of this volume.

Interdisciplinarity of the glycosciences is also demonstrated by a contribution on novel approaches for the production of levansucrases, bacterial extracellular enzymes that act on sucrose producing β-2,6-linked fructans. Biochemical characterization of this protein encoded in the genome from *Pseudomonas* bacteria, and an innovative mass spectrometric study of the reaction products permits to identify the produced potentially prebiotic fructooligosaccharides from sucrose or raffinose.

NMR is currently a potent methodology to analyse sugar conformation and to study binding interactions. New advances from the NMR methodological viewpoint for analysis of saccharide conformation are described in a chapter, in which examples are given for oligo- and polysaccharides, glycopeptides, glycomimetics, and also carbohydrate-protein, carbohydrate-carbohydrate, and carbohydrate-nucleic acid interactions.

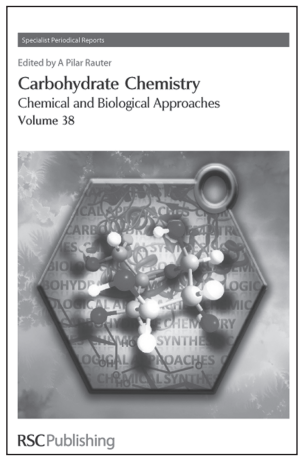
The contribution of glycochemistry to innovation in glycosciences is shown in chapters 10–17. Here imino sugar glycosidase inhibitors, carbasugars, multivalent glycoconjugates, including glycodendrimers, glyconanotubes, and glyconanoparticles, and their uses in medicinal chemistry, as well as artificial saccharide-based and saccharide-functionalized gene delivery systems are presented.

Highly functionalized exo-glycals used for the generation of molecular diversity in a chemoselective manner, namely for the preparation of furanose-based libraries with three or more sites for molecular diversity are discussed. A chapter on siderophores that are based on monosaccharides that have proven effective for Gram-negative bacteria and mycobacteria, and the chapter on biomaterials, in particular on the so-called smart materials, that can modulate and control cell behaviour, complete the volume.

Volume 38 of Carbohydrate Chemistry – Chemical and Biological Approaches contains contributions ranging from glycochemistry to glycobiology. They have been authored mostly by scientists that are members of the European Science Foundation Network Euroglycoforum. This collection demonstrates in a meaningful way how the interdisciplinary approach of an international glyconetwork can advance the field of carbohydrate research in Europe and worldwide.

We hope you will enjoy the beauty of the ''sweet'' chemistry and biology presented herein!

Amélia Pilar Rauter and
Thisbe K. Lindhorst



Cover

Tetrahydropyran-enclosed ball-and-stick depiction of a glucose molecule, and (in the background) part of an α-glycosyl-(1-4)-D-glucose oligosaccharide and a glycosidase, all representative of the topics covered in Carbohydrate Chemistry – Chemical and Biological Approaches. Cover prepared by R. G. dos Santos.

Preface **vii**

Amélia Pilar Rauter and Thisbe K. Lindhorst

Applications of glycobiology: biological and immunological effects of a chemically modified amylose-derivative **1**

Ghislain Opdenakker, Sandra Li, Nele Berghmans and Jo Van Damme

1 Introduction 1
2 Historical breakthroughs and examples 2
3 Polysaccharides and derivatives 5
4 An historical finding in virology? 5
5 COAM does not induce interferon 6
6 COAM is an immunomodulator 7
7 Therapeutic implications for acute and chronic inflammation and cancer 8
8 Conclusions and future perspectives 9
Abbreviations 10
Acknowledgements 10
References 10

Lipopolysaccharide structure and biological activity from the cystic fibrosis pathogens *Burkholderia cepacia* complex **13**

Anthony De Soyza, Flaviana Di Lorenzo, Alba Silipo, Rosa Lanzetta and Antonio Molinaro

1 Introduction 13

© The Royal Society of Chemistry 2012

2 Basic Lipopolysaccharide structure 16
3 The detailed chemical structure of the *P. aeruginosa* LPS 20
4 *Burkholderia cepacia* complex and Cystic Fibrosis 23
5 General structural features of the core region of BCC LPS 28
6 O-chain of BCC LPS 30
7 Conclusion 33
Acknowledegments 34
References 34

Synthesis of bacterial carbohydrate surface structures containing Kdo and *glycero*-D-*manno*-heptose linkages **40**

Stefan Oscarson

1 Introduction 40
2 Kdo-containing structures and Kdo donors 40
3 *Glycero*-D-*manno*-heptose-containing structures 48
References 59

Synthetic glycopeptides in vaccine development and antibody epitope mapping **61**

Ulrika Westerlind

1 Introduction 61
2 Protein conjugate vaccines 62
3 Build-in adjuvant vaccines 65
4 Dendrimer vaccines 68
5 Antibody epitope mapping 69
6 Conclusions 72
References 73

Posttranslational sialylation and its impact on leukocyte recruitment during inflammation **75**

Karin Bodewits and Markus Sperandio

1 Leukocyte recruitment cascade 75
2 Posttranslational modifications 78
3 Sialic acids and sialyltransferases (ST) 82
4 Sialylation-dependent functions of signalling and adhesion relevant molecules 88
5 Conclusion and outlook 91
References 91

Glycoengineering of protein-based therapeutics **94**

Sandrine Donadio-Andréi, Chloé Iss, Nassima El Maï, Valérie Calabro and Catherine Ronin

Introduction 94
1 Impact of N-glycosylation on therapeutic proteins 95
2 Glycosylation optimization 105
3 Conclusions 113
References 115

Congenital Disorders of Glycosylation (CDG): from glycoproteins to patient care **124**

Vanessa Ferreira, Paz Briones and Maria-Antonia Vilaseca

1 Introduction 124
2 Basic principles of protein glycosylation 124
3 Congenital disorders of glycosylation (CDG) 128
4 Concluding remarks and future perspectives: CDG as a challenge for glycobiomedicine in the next decade! 144
5 Highlights 145
6 Further information 146
Abbreviations 146
Acknowledgements 147
References 147

Bladder cancer–glycosylation insights **156**

Paulo F. Severino, Mariana Silva, Mylène A. Carrascal, Fernando Calais, Fabio Dall'Olio and Paula A. Videira

1 Introduction 156
2 Bladder cancer: epidemiology and current treatment 157
3 General aspects of glycosylation 158
4 Final considerations/conclusions 168
Abbreviations 169
References 169

Levansucrases of *Pseudomonas* bacteria: novel approaches for protein expression, assay of enzymes, fructooligosaccharides and heterooligofructans **176**

Tiina Alamäe, Triinu Visnapuu, Karin Mardo, Andres Mäe and Alina D. Zamfir

1 Levansucrase genes and proteins 176
2 Heterologous expression of levansucrases from *Pseudomonas syringae* pv. tomato with two different expression systems 177

3 Biochemical properties of the levansucrases: substrate specificity and kinetic parameters 178
4 Predicted catalytic triad residues and 3D structures of Lsc1, Lsc2, Lsc3 and LscA proteins 179
5 Lsc2, Lsc3 and LscA produce not only levan, but also FOS that can be detected by a novel mass spectrometric method 180
6 Isolation and novel screening methods for levansucrase mutants 186
7 Concluding remarks and future perspectives 189
Acknowledegements 190
References 190

Recent advances on the application of NMR methods to study the conformation and recognition properties of carbohydrates **192**

Ana Ardá, M. Álvaro Berbís, Pilar Blasco, Angeles Canales, F. Javier Cañada, M[a] Carmen Fernández-Alonso, Filipa Marcelo and Jesús Jiménez-Barbero

1 Introduction 192
2 The access to new NMR parameters and methodological developments 192
3 Applications. Saccharides in solution 194
4 Applications. The interaction of saccharides with other natural and synthetic molecules 202
Acknowledgments 210
References 210

Glycosidase inhibitors: versatile tools in glycobiology **215**

Óscar López, Penélope Merino-Montiel, Sergio Martos and Alejandro González-Benjumea

1 Introduction 215
2 Polyhydroxylated pyrrolidines 216
3 Six-membered ring-based imino sugars 219
4 Azepanes and azetidines 226
5 Bicyclic imino sugars 231
6 Thio, seleno and carbasugars as glycosidase inhibitors 245
7 Miscellaneous 250
8 Pharmacological activities of imino sugars 251
9 Concluding remarks 256
Acknowledgements 256
References 256

An overview of key routes for the transformation of sugars into carbasugars and related compounds **263**

Raquel G. Soengas, José M. Otero, Amalia M. Estévez, Amélia P. Rauter, Vasco Cachatra, Juan C. Estévez and Ramón J. Estévez

1 Introduction 263
2 Strategies for the synthesis of carbasugars and pseudo-carbasugars 265
Acknowledgements 296
References 297

Multivalent glycoconjugates in medicinal chemistry **303**

José G. Fernández-Bolaños, Inés Maya and Ana Oliete

1 Introduction 303
2 Glycoclusters with a flexible core 303
3 Glycoclusters with a rigid core 308
4 Multivalent glycopeptides in immunotherapy 313
5 Glycoconjugates on tubular scaffolds 318
6 Glycodendrimer chips 318
7 Glyconanoparticles 320
8 Conclusions 331
Acknowledgements 331
References 331

Glycotransporters for gene delivery **338**

Carmen Ortiz Mellet, José M. García Fernández and Juan M. Benito

1 Introduction 338
2 Carbohydrate-grafted cationic polymers and lipids 341
3 *De novo* designed carbohydrate-based polymers 346
4 Preorganized glycomaterials for gene delivery 354
5 Outlook and perspectives 368
Acknowledgments 368
References 369

Furanose-based templates in the chemoselective generation of molecular diversity **376**

Ana M. Gómez, Clara Uriel and J. Cristóbal López

1 Introduction 376
2 Design and synthesis of functionalized furanose-based precursors 377

3 Reactivity of functionalized furanose-based precursor 4a 380
4 Reactivity of the alkenyl halide moiety in functionalized furanose-based precursors 4b,c 387
5 Three components assembly of functionalized furanose derivatives 4b,c 388
6 Generation of furanosidic libraries possessing three – or more – sites for molecular diversity 391
7 Conclusions 392
Abbreviations 393
Acknowledgements 393
References 393

Synthesis of carbohydrate-based artificial siderophores and their biological applications **398**

Marta M. Andrade and Amélia P. Rauter

1 Introduction 398
2 Artificial siderophores - background 400
3 Synthesis of Carbohydrate-based artificial siderophores 402
4 Study of siderophore activity 407
5 Conclusions and perspectives 412
References 414

Smart biomaterials: the contribution of glycoscience **416**

Laura Cipolla, Laura Russo, Francesca Taraballi, Cristina Lupo, Davide Bini, Luca Gabrielli, Alice Capitoli and Francesco Nicotra

1 Introduction: biomaterials and tissue engineering 416
2 Glycoscience for biomaterials design 424
Conclusions 440
References 440

Applications of glycobiology: biological and immunological effects of a chemically modified amylose-derivative

Ghislain Opdenakker,* Sandra Li, Nele Berghmans and Jo Van Damme

DOI: 10.1039/9781849734769-00001

Carbohydrate chemistry, oligosaccharide sequencing, synthesis technologies and glyco-engineering have helped to establish glycobiology alongside molecular biology. However, examples of therapeutic implications of glycobiology are limited to oligosaccharides. These include engineered antibodies containing sialic acids, inhibition of leukocyte rolling effect in the interactions between mucins and selectins and antiviral imino sugars. The best example of a successful glycodrug is oseltamivir, Tamiflu®, as an inhibitor of the influenza virus neuraminidase. Polysaccharides, although at first sight structurally less complex and biologically less challenging, are interesting molecules with unexplored possibilities in biology and medicine. Polysaccharides may form protease-resistant scaffolds for tissue engineering and chemically modified polysaccharides (CMPs) have potent immunomodulating activities. This is illustrated here with chlorite-oxidized oxyamylose (COAM), a broad-spectrum antiviral agent. COAM interacts with the chemokine system and thus can be used to modulate leukocyte compartmentalization *in vivo*. This knowledge has been used to alter the clinical course of acute viral infections, cancer and experimental autoimmune encephalomyelitis (EAE), an animal model of multiple sclerosis. The latter example is proof-of-concept that CMPs constitute important probes to study immune functions and are novel drugs applicable for specific disease entities.

1 Introduction

From semantic viewpoint the term molecular biology is a pleonastic expression and does not correspond with the present day interpretations. Indeed, all biological processes are driven by molecules. What is meant by molecular biology are processes driven by nucleic acids. Thus, molecular biologists are *stricto sensu* researchers using DNA and RNA tools for sequencing and synthesis, for practical uses in genetic engineering and for therapy with recombinant drugs. Glycobiology (the biology of glycans) is a more strict term in that it defines, by its name, the molecules under study, namely carbohydrates. Why does the field of glycobiology then still need promotion? Why do the DNA/RNA biologists generate more appreciation and the glycobiologists less recognition? This is certainly not because DNA and RNA are sugar-phosphate polymers. Maybe it is because of discovery of the beauty of the linear connection between codons and amino acids.

Rega Institute for Medical Research, Minderbroedersstraat 10, B-3000 Leuven, Belgium. E-mail: ghislain.opdenakker@rega.kuleuven.be

© The Royal Society of Chemistry 2012

Maybe it is also because functional aspects of a branched glycan tree are not yet understood. As long as the interconnections between RNA/DNA biology with glycobiology are not yet hold by a solid stem, scientists need to invest time and resources in finding important examples in glycobiology. This is simply to state that the real basis of glycobiology, for instance relating structures to functions, is a barrier for many scientists and needs further promotion. The original concepts that glycoforms and glycotypes confer different functions to proteins must be seen as eye-openers [1–2]. These concepts, well elaborated for immunoglobulin-G (IgG) and tissue-type plasminogen activator (t-PA), are becoming apparent for many other glycoproteins. Aside effects of glycosylation on the specific activity of t-PA and functional aspects of the glycans on the Fc part of IgG (*vide infra*), glycoprotein folding, trafficking, recognition and involvement in the immune system demonstrate that oligosaccharides are endowed with multiple functions [3–6].

The successes in technology progresses in nucleic acid and protein sequencing and synthesis are evident and their impacts on medicine, industry and society are obvious. The field of RNA/DNA biology has the intrinsic advantage of a limited number of building blocks, with limited natural chemical modifications of the five bases. Protein biology, based on about 20 building blocks, is already more complex and subject to many more natural chemical modifications: proteolytic processing, phosphorylation and dephosphorylation, citrullination, isoprenylation, acetylation/deacetylation, to name a few. Glycosylation is nowadays recognized as the most diverse and complex posttranslational modification of proteins.

Glycobiology forms the next, and perhaps the most challenging level. Intrinsically, carbohydrate biochemistry is also based on a limited number of natural monosaccharides, but their chemical modifications are multifold. In addition, the glycosidic bonding of monosaccharides is subject to various linkages and the involved carbon atoms form stereochemical centers. The branching of oligo- and polysaccharides into arborized structures often supersedes our imagination capacity, which is too much trained for linear DNA, RNA and protein sequences. However, in practice in mammalian systems a limited set of monosaccharides are used. Even so the linkage variation creates an extraordinary diversity of structures.

2 Historical breakthroughs and examples

If glycobiology is to succeed as the science of the 21st century, then one needs to define new breakthroughs. Happily one can walk in the footsteps of those who prepared and enhanced the field with exoglycosidase oligosaccharide sequencing [1, 2], with mass spectrometry [7] and NMR analysis of sugars [8], with lectin studies [9] and with glyco-engineering. The discovery of congenital disorders of glycosylation [10] not only points to future therapeutic applications, but also has yielded a bridging function between glycobiology and nucleic acid biology.

Oligosaccharide sequencing evolved over the last 30 years from an extremely labor-intensive discipline towards an efficient analytical method in the hands of experts. Pioneering work by the groups of Akira Kobata

(University of Tokyo) [11] and of Raymond Dwek (Oxford Glycobiology Institute) [12] led to the understanding of the proinflammatory effects of agalactosyl IgGs in rheumatoid arthritis [2]. This knowledge was complemented with the findings of anti-inflammatory sialic acids in IgGs [13]. Together, these breakthroughs can be summarized in one sentence. Oligosaccharides attached to proteins fine-tune the biological activities of these molecules. This should be not compared with an on/off switch (as is often the case with protein phosphorylation), but rather with the selection of a specific program or tuning. Presently, oligosaccharide exoglycosidase sequencing has been perfected to such a level [14, 15] that broad-spectrum and high-throughput analysis and glycomics are within reach. One next breakthrough will come from coupling of this type of sophisticated analysis with proteomics, genomics and transcriptomics.

A second aspect of structural analysis is mass spectrometry and NMR analysis. Whereas mass spectrometry *per se* is nowadays routinely used in proteomic analysis [16], for oligosaccharide analysis the identification of structural details by mass spectrometry remained more complicated for the simple reason that, for example, the various hexoses maintain the same molecular mass. Nevertheless, sophisticated mass spectrometry technology for sugars has been developed, is presently the expertise of few specialists [7, 17], but is becoming common practice, because the combination of high-performance liquid chromatography and mass spectrometry of oligosaccharides is very powerful and rapid. Considerable growth is expected in this area and basic mass spectrometry analysis is already becoming classical technology for structural analysis. Unfortunately and as is the case for other applications of mass spectrometry, high-end equipment is rather expensive and a limited number of specialized centers, homing dedicated specialists and equipped with such instruments, can answer easily any exotic demands for the analysis of complex glycoconjugates. Similar economic considerations may apply for NMR analysis, with the additional practical point that the amounts of purified glycoconjugates needed for NMR analysis supersede considerably those for mass spectrometry identification.

A third aspect is related to practical aspects, which we best place under the term glyco-engineering. Agalactosyl [2] and sialyl IgGs [13] have already been mentioned as disease promoting and disease-limiting factors, respectively, in inflammation. Such knowledge is successfully addressed by glyco-engineering companies making optimized recombinant glycoprotein preparations [18]. Another example is related to the classical ABO blood group system. Nobel Prize laureate Karl Landsteiner was a glycobiologist *avant-la-lettre*. Present-day glyco-engineers have managed to modify the ABO blood group antigens [19], heralding novel possibilities in transfusion medicine. Because the immunological insights with blood transfusions formed the basis of organ transplantation medicine, it is not excluded that the glyco-engineering of the ABO blood group antigens forms the basis of new applications in xenotransplantation, where galactosylation is a major hinder in the application of pig donor organs [20].

A fourth aspect to be recognized as historical in glycobiology is the study of lectins as probes for sugars. Plant lectins, recognizing in defined ways oligosaccharides, have been widely used in medical diagnostics, often before

their mechanism of action was known [21, 22]. Lymphoblast transformation with phytohaemagglutinin for karyotyping purposes and definition of chromosomal aberrations was used long before the cytokines that are induced in this process were identified as interleukin-2, interleukin-6 and others. Lectins are also used for disease diagnosis in histochemical analysis of cancer and, more recently, with the use of lectin microarrays [23]. The literature on this topic is vast, exemplary reviews are available for detailed and general information [24, 25]. In contrast, the information about lectins as therapeutics is still rather sketchy. After the discovery of selectins, the mammalian lectins that regulate rolling of leukocytes at an early step in the inflammatory response, selectin knockout mice were developed to study immunophenotypes (how these mice differ from wildtype animals with reference to immunological parameters) and attempts were made to develop novel anti-inflammatory agents on the basis of these lectin-glycan interactions [26–28]. A second example about glycotherapy relates to viral infections. Human immunodeficiency virus is enveloped in a glycocoat, limiting access to protein (peptide) epitopes. One way to flag this virus is with mannose-recognizing lectins, a strategy that is investigated with increasing intensity [29].

An important way to enhance the field of glycobiology is to define new medical applications. These may involve deficiencies, resulting from aberrant glycosylation, in need for substitution therapy or storage diseases with a pathogenic role of accumulated saccharides, in which reduction forms a treatment strategy. An important class of rare diseases is the family of congenital diseases of glycosylation (CDG) [10]. However, a critical aspect to recognize is the fact that CDGs create the link between nucleic acid and protein biology and glycobiology. Identified CDGs, resulting from the mathematics of evolution (combinations of random mutations that result in life offspring with clinical phenotypes), demonstrate that glycosylation is a biological process of vital importance. In addition, the creation of knockout mouse models of CDGs is an important step to understand better the biological importance of sugars, but may also form a solid basis to define treatments of these rare diseases, both in cases of deficiencies [30] and of substrate reduction therapy in storage diseases, like Gaucher disease [31].

Finally, we come to the level of commonly used drugs. One viral target enzyme stands out as a glycobiological example: the influenza virus neuraminidase with its inhibitors oseltamivir (Tamiflu®) and zanamivir (Relenza®). The concept is simple, but the efforts done are considerable: definition of a microbial target enzyme (sufficiently different from sister enzymes in the host) and screening for inhibitors [32]. It does not take much imagination that future antivirals and other antimicrobial agents may be developed on the basis of other microbial sugar modifying enzymes. In virology, such examples of interesting glycotarget enzymes may be restricted because of the limitation of the coding capacity of viral genomes and, therefore, the influenza neuraminidase will keep the lead for a while. However, in the world of bacteria, fungi and parasites an enormous potential exists to discover target enzymes, generate inhibitors and define new medicines.

Most of the examples mentioned so far relate to oligosaccharides. What is the status about polysaccharides?

3 Polysaccharides and derivatives

Although changes are emerging with the recognition of the importance of polysaccharide vaccines, in biomedical education the time spent to study polysaccharides is inversely correlated with their abundance in nature. Cellulose [polyβ(1-4)D-glucose], amylose [polyα(1-4)D-glucose] and branched amylopectin in starch form most biomass and energy supply for most organisms. Intestinal resorption of starch glucose (from amylose and amylopectin) as monosaccharides and disaccharides and conversion into glycogen is basic medical knowledge that can be linked to rare glycogen storage diseases and to metabolic disease number one, *i.e.* diabetes. Deficiencies of amylase and disaccharidases lead to alterations of the gut microbiomes and to gastrointestinal disorders as is the case with specific polysaccharide derivatives. Indeed, dextran sodium sulphate (DSS) is commonly used to induce inflammatory bowel disease in mice (used as an animal model for Crohn's disease or ulcerative colitis). In other words, polysaccharide derivatives may be harmful or not, depending on their degradability and their location.

4 An historical finding in virology?

After the discovery of interferon (IFN) in 1957, great hope was generated for its use as therapy for viral diseases and mass production lines were explored ending in the recombinant expression of IFN and use in the clinic [33]. IFNs were not developed into first choice antivirals. Eventually, IFN-α from leukocytes and IFN-β from fibroblasts became first choice anti-inflammatory drugs for the treatment of multiple sclerosis. Recombinant IFN-γ, also dubbed immune interferon, because it is produced by lymphocytes and natural killer cells, found its way to the treatment of chronic granulomatous disease [34].

Piet De Somer (1917–1985), founder of the Rega Institute (Fig. 1) was an insightful man because he thought about and tried to solve the IFN production problem (before its cloning and expression [35]) by finding ways to induce endogenous IFN. He admired and followed his scientist example, Karl Landsteiner, by trying to use simple but practical rules. Karl Landsteiner, who discovered the ABO blood group system, used simple serological tests for discovery.

The aftermath of the second world war was characterized by discoveries of many new mold-derived antibiotics, when antiviral IFN was discovered upon its inducibility by viruses. IFNs protect neighbouring cells by inducing an antiviral state and the virologists in the early 1960s started to search for inducers that could replace viruses. In those booming days of DNA and RNA research, screening of mold or bacterial extracts and synthetic molecules for antiviral activity was a commonly used technique. *Penicillium* products and synthetic polyanionic polymers were found to possess broad spectrum antiviral activity and this incited De Somer to search for and test other polyanions for antiviral activity. De Somer explored synthetic

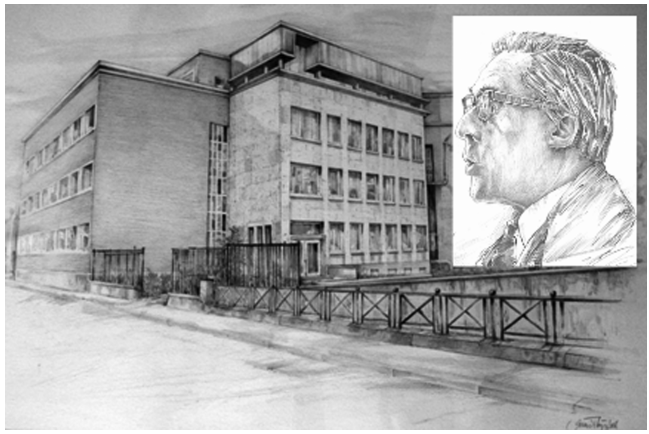


Fig. 1 Rector Piet De Somer (1917–1985) and the Rega Institute for Medical Research, founded in 1954. Drawings are by Gerard Thijs, graphic artist, Leuven.

polyanions, such as polymethacrylates and polyacrylates, as antiviral agents [36] and as inducers of interferon with some success [37]. Later it was found that nucleic acids, such as synthetic double stranded RNA, induce the production of IFN (most probably by mimicking the replicative intermediate forms of RNA viruses). Double-stranded RNA is a polyanionic molecule and like polyriboinosinic acid:polyribocytidylic acid (poly IC), synthetic polyanions were found to be toxic, presumably by their non-degradability. De Somer then gave the chemists of the Rega Institute the task to manufacture a biodegradable polyanionic molecule that could be used as an antiviral agent and/or interferon inducer: chlorite-oxidized oxyamylose or COAM was born [38]. The chemists used amylose as starting material in a stepwise oxidation reaction with openings of the glucose rings (Fig. 2) to yield the polycarboxylic acid structure, now known as COAM. Quickly, the product was tested against viruses and found to be a potent and broad-spectrum antiviral agent [38–48]. In those days, this was often not done with cell-based assays and permissive viruses, but instead by *in vivo* testing (Table 1). COAM was supposed to induce IFN, although from the time of the first studies onwards it was not clarified that the observed serum IFN levels in mice treated with COAM and infected with virus were indeed derived from induction by either the virus or by COAM. Because of some inconclusive discrepancies in the COAM literature [39, 49], the question whether COAM itself induces IFN or enhances IFN production by other inducers remained open. In fact, data obtained already in 1976 from *in vivo* experiments, corroborating earlier studies with COAM [42], remained hidden in our laboratory notebooks till publication in 2010 [50].

5 COAM does not induce interferon

About 35 years after its synthesis [38], we reiterated the COAM project in 2006 using the now available powerful techniques of genetic engineering. First, we applied a simple *in vitro* virus infection model with the picornavirus mengovirus on mouse fibroblastoid L929 cells. COAM protected

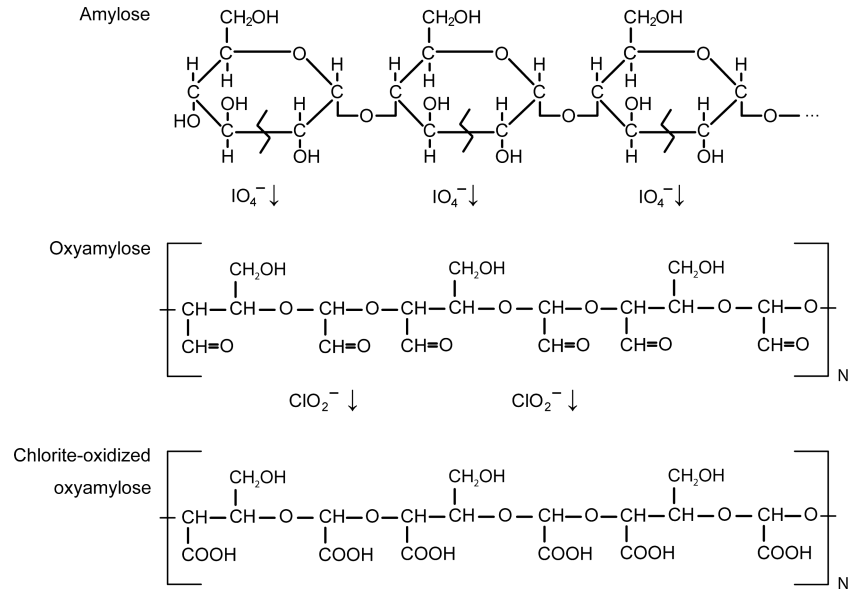


Fig. 2 Chemical synthesis of COAM from amylose. The steps of glucose ring opening by periodate and further oxidation with chlorite into a polycarboxylate structure is illustrated (Adapted from P. Claes *et al*., 1970, reference 38).

dose-dependently the cell cultures against mengovirus, but only poly IC induced an antiviral state by induction of IFN-β. We did not observe induction of IFN bioactivity or of IFN mRNAs by COAM. By performing time-of-addition experiments with COAM, we deduced that COAM acted at the stage of virus entry. It was also noticed that rather high concentrations (1 mg/ml) of COAM were needed to exert an *in vitro* antiviral effect [51].

6 COAM is an immunomodulator

The *in vivo* efficacy of COAM against many viruses (Table 1) was much better than what was expected on the basis of the *in vitro* antiviral profile. For those reasons immune cells and molecules were thought to be involved and were studied. The conclusions that were deduced from these studies included (i) COAM acts mainly on myeloid cells (neutrophils and macrophages) rather than on lymphocytes and (ii) intraperitoneal injection of COAM leads to massive neutrophil influx, hinting to an interaction with neutrophil chemokines. Since the most potent neutrophil chemokine in the mouse is granulocyte chemotactic protein-2 (GCP-2) [52], the induction and binding of GCP-2 by COAM was investigated and the link with chemokines was established. It was concluded that COAM induces and binds GCP-2 and thus establishes a potent chemokine gradient for myeloid cells to the peritoneal cavity where the local virus infection is contained [50]. The information about a CMP drug that protects against lethal virus infections after prophylactic parenteral administration is perhaps not so important for common acute virus infections. However, because rescue (more than 90

Table 1 Antiviral effects of COAM

Virus[a]	Species (treatment)[b]	Dosis/ concentration[c]	Effect[d]	Reference
Coxsackie B4	Mice (profyl. 24 h)	2 mg/mouse ip.	Protection	53
Foot-and-mouth ip.	Mice (profyl. 24 h)	5–60 mg/kg ip.	Protection	40
Foot-and-mouth ip.	Swine (profyl. 24 h)	120 mg/kg ip.	No effect	40
Friend leukemia	Mice (profyl.)	0.5 mg/mouse	Improvement	46
Hog cholera	Swine (profyl. 24 h)	120 mg/kg ip.	Protection	40
HSV-1 in eye	Rabbit (profyl. 4 h)	50 mg/ml topic.	Protection	40
HSV-1 in.	Mice (profyl. 24 h)	0.8–4 mg/mouse ip.	Protection	47
Influenza in.	Mice (profyl. 24 h)	2 mg/mouse ip.	Protection	39
Influenza ip.	Mice (profyl. 24 h)	2 mg/mouse ip.	Protection	41
Mengo ip.	Mice (profyl. 24 h)	2 mg/mouse ip.	Protection	40
Mengo ip.	Mice (profyl. >16 h)	2 mg/mouse ip.	Protection	42
Mengo ip.	Mice (profyl. 24 h)	2 mg/mouse ip.	Protection	50
Mengo	L929 cells (*in vitro*)	1 mg/ml	CPE inhibition	51
Moloney mouse sarcoma im.	Mice (profyl. 24 h)	0.5 mg/mouse ip.	Protection	43
Newcastle disease iv.	Mice (profyl.)	1–10 mg/mouse ip.	Protection	44
PVX leaf discs	Plant (posti.)	1% solution	No effect	48
Rabies im.	Mice (3 h posti.)	0.3 mg/mouse	Worsening	45
Semliki Forrest ip.	Mice (profyl. 24 h)	2 mg/mouse ip.	Protection	39
TMV leaf discs	Plant (posti.)	1% solution	Protection	48
Vaccinia iv.	Mice (profyl. 24 h)	2 mg/mouse ip.	Protection	38
Vaccinia iv.	Mice (profyl. 24 h)	>55 µg/mouse iv.	Protection	39
Vesicular stomatitis	MEF (*in vitro*)	10 µg/ml	CPE inhibition	38
Vesicular stomatitis in.	Mice (profyl.)	10 mg/mouse	No effect	44

[a] HSV-1: Herpes simplex virus type 1; TMV: Tobacco mosaic virus; PVX: Potato virus X.
[b] profyl.: prophylactic, before virus; posti.: postinfection.
[c] iv.: intravenous; ip.: intraperitoneal; in.: intranasal; im.: intramuscular.
[d] CPE: cytopathic effect.

percent survival without disease sequels) was observed after lethal virus challenge by pretreatment with a purified COAM preparation, the CMP preventive action may be crucial to contain clusters of lethal virus infections, including those with Ebola and other deadly viruses [50]. In another model of Coxsackievirus B4 pancreatic infection it was corroborated that COAM has potent antiviral effects, most probably mediated by myeloid inflammatory cells [53].

7 Therapeutic implications for acute and chronic inflammation and cancer

On the basis of the mentioned experimental data, the question was asked whether the polysaccharide derivative COAM may be applicable in other diseases. De Somer's wish of using IFN-β as a potent antiviral agent did not really materialize. Instead, IFN landed as a real drug for the treatment of multiple sclerosis [33].

At the time when it was still controversial whether COAM induces IFN and not yet clear that COAM acted in a different way, we investigated whether COAM might protect against experimental autoimmune encephalomyelitis (EAE), an animal model of multiple sclerosis. It is worth

mentioning that these experiments were also intended to investigate the clinical effects of COAM in virus-free conditions.

The major conclusions from studies with a hyperacute model of EAE formed the basis of a doctoral dissertation (Dr. N. Berghmans, 2008, University of Leuven, Belgium, promoter Prof. H. Heremans). COAM did not induce IFN-α, IFN-β or IFN-γ *in vivo*, but possessed a significant protective effect against hyperacute EAE. Intraperitoneal injection of COAM was locally chemotactic for leukocytes with such potency that it diminished the leukocyte numbers significantly in the central nervous system (brain and spinal cord). Because of these observations, we proposed a new immunotherapeutic strategy to treat acute and chronic CNS inflammation and called this ''cell deviation therapy with an immunostimulating glycomimetic''. This research is presently continued and firm evidence has been established that COAM also protects against autoimmune EAE, induced by myelin oligodendrocyte glycoprotein-derived peptide in EAE-susceptible IFN-γ knockout mice [54] and [Berghmans *et al*., in preparation].

Another type of application was investigated in tumor biology. As already described in Table 1, COAM protected against Moloney mouse sarcoma virus infections [43]. This was interpreted as an antiretroviral effect, and also as an antitumoral effect, because both tumor incidence and lethality caused by tumor growth were reduced by COAM. Moreover, by now it is clear that inflammation is a critical component in tumor biology [55] and that myeloid cells, alongside lymphocytes, are key players in tumor biology [56]. For these reasons, we took the simple approach of testing COAM against melanoma in a syngeneic mouse model. It was found that COAM reduced significantly the tumor burden of dermal melanomas when locally injected early in tumor development [57].

8 Conclusions and future perspectives

In comparison with many studies on the biology of complex oligosaccharide structures, polysaccharides and derivatives have not yet attracted much interest from industry. Nevertheless several interesting principles may be demonstrated with the example of the amylose-derivative COAM. First, old literature and hypotheses (about polysaccharides and derivatives) should be tested with all novel technologies at our disposal. For instance, structural analysis can now be executed with the tools of exoglycosidase sequencing and mass spectrometry and biological analyses may rely on the use of RT-PCR, transcriptional micro-arrays and *in vivo* animal models. Second, on the basis of biological discoveries new therapies may be invented. For instance, on the basis of the discovery that COAM (induces and) binds chemokines and thus may create unseen potent endogenous chemokine gradients, we discovered that it is possible to call leukocytes out of the CNS back into the circulation and the (immunological) periphery to cure mouse EAE. Such proof of principle may pave the way to new treatments of acute and chronic neuroinflammations and should definitely stimulate basic and applied research on polysaccharides.

Abbreviations

CMP	chemically modified polysaccharide
COAM	chlorite-oxidized oxyamylose
CDG	congenital disorder of glycosylation
CPE	cytopathic effect
EAE	experimental autoimmune encephalomyelitis
HSV-1	herpes simplex virus type 1
IgG	immunoglobulin-G
IFN-	interferon-
im.	intramuscular
in.	intranasal
ip.	intraperitoneal
iv.	intravenous
poly IC	polyriboinosinic acid:polyribocytidylic acid
t-PA	tissue-type plasminogen activator

Acknowledgements

This overview is a tribute to the late Professor Piet De Somer, founder of the Rega Institute for Medical Research and an eminent scientist-entrepreneur and university rector in Belgium. To commemorate his passing away, 25 years ago, the University of Leuven honored De Somer with a symposium and the City of Leuven renamed its central place into ''Piet De Somer Square''. The authors thank Professor Alfons Billiau (Leuven) and Professor Raymond A. Dwek (Oxford) for inspiration and critically reading and commenting on the manuscript and Professor Hubertine Heremans (Leuven) for expertise in EAE and promotion of the COAM-project. This work was funded by the ''Geconcerteerde OnderzoeksActies'' (GOA2012/17) and by the Fund for Scientific Research of Flanders (FWO-Vlaanderen). GO is a Steering Committee Member of the European Science Foundation (ESF, Strasbourg, France) Research Network Programme ''European GlycoScience Forum'' (EGSF).

References

1 T. W. Rademacher, R. B. Parekh and R. A. Dwek, *Annu. Rev. Biochem.*, 1988, **57**, 785.
2 R. B. Parekh, R. A. Dwek, B. J. Sutton, D. L. Fernandes, A. Leung, D. Stanworth, T. W. Rademacher, T. Mizuochi, T. Taniguchi, K. Matsuta, F. Takeuchi, Y. Nagano, T. Miyamoto and A. Kobata, *Nature*, 1985, **316**, 452.
3 P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson and R. A. Dwek, *Science*, 2001, **291**, 2370.
4 G. Opdenakker, P. M. Rudd, C. P. Ponting and R. A. Dwek, *FASEB J*, 1993, **7**, 1330.
5 P. M. Rudd and R. A. Dwek, *Crit. Rev. Biochem. Mol. Biol.*, 1997, **32**, 1.
6 P. Van den Steen, P. M. Rudd, R. A. Dwek and G. Opdenakker, *Crit. Rev. Biochem. Mol. Biol.*, 1998, **33**, 151.
7 D. J. Harvey, *Mass. Spectrom. Rev.*, 2012, **31**, 183.
8 J. C. Paulson, J. Weinstein, L. Dorland, H. van Halbeek and J. F. Vliegenthart, *J. Biol. Chem.*, 1982, **257**, 12734.

9 H. Lis and N. Sharon, *Annu. Rev. Biochem.*, 1986, **55**, 35.
10 J. Jaeken and G. Matthijs, *Annu. Rev. Genomics. Hum. Genet.*, 2007, **8**, 261.
11 M. Arakawa, S. Ogata, T. Muramatsu and A. Kobata, *J. Biochem.*, 1974, **75**, 707.
12 P. M. Rudd and R. A. Dwek, *Curr. Opin. Biotechnol.*, 1997, **8**, 488.
13 Y. Kaneko, F. Nimmerjahn and J. V. Ravetch, *Science*, 2006, **313**, 670.
14 P. M. Rudd, G. R. Guile, B. Küster, D. J. Harvey, G. Opdenakker and R. A. Dwek, *Nature*, 1997, **388**, 205.
15 K. Mariño, J. Bones, J. J. Kattla and P. M. Rudd, *Nat. Chem. Biol.*, 2010, **6**, 713.
16 C. Choudhary and M. Mann, *Nat. Rev. Mol. Cell. Biol.*, 2010, **11**, 427.
17 A. Dell and H. R. Morris, *Science*, 2001, **291**, 2351.
18 S. R. Hamilton, R. C. Davidson, N. Sethuraman, J. H. Nett, Y. Jiang, S. Rios, P. Bobrowicz, T. A. Stadheim, H. Li, B. K. Choi, D. Hopkins, H. Wischnewski, J. Roser, T. Mitchell, R. R. Strawbridge, J. Hoopes, S. Wildt and T. U. Gerngross, *Science*, 2006, **313**, 1441.
19 Q. P. Liu, G. Sulzenbacher, H. Yuan, E. P. Bennett, G. Pietz, K. Saunders, J. Spence, E. Nudelman, S. B. Levery, T. White, J. M. Neveu, W. S. Lane, Y. Bourne, M. L. Olsson, B. Henrissat and H. Clausen, *Nat. Biotechnol.*, 2007, **25**, 454.
20 M. S. Sandrin and I. F. McKenzie, *Immunol. Rev.*, 1994, **141**, 169.
21 N. Sharon, *Biochem. Soc. Trans.*, 2008, **36**, 1457.
22 W. J. Peumans and E. J. Van Damme, *Crit. Rev. Biochem. Mol. Biol.*, 1998, **33**, 209.
23 J. Zhao, T. H. Patwa, D. M. Lubman and D. M. Simeone, *Curr. Opin. Mol. Ther.*, 2008, **10**, 602.
24 R. Saldova, M. R. Wormald, R. A. Dwek and P. M. Rudd, *Dis. Markers.*, 2008, **25**, 219.
25 C. D. Rillahan and J. C. Paulson, *Annu. Rev. Biochem.*, 2011, **80**, 797.
26 T. A. Springer, *Cell*, 1994, **76**, 301.
27 K. J. Woollard and J. Chin-Dusting, *Curr. Pharm. Des.*, 2010, **16**, 4113.
28 B. Lepenies, J. Yin and P. H. Seeberger, *Curr. Opin. Chem. Biol.*, 2010, **14**, 404.
29 J. Balzarini, *Nat. Rev. Microbiol.*, 2007, **5**, 583.
30 A. Schneider, C. Thiel, J. Rindermann, C. Derossi, D. Popovici, G. F. Hoffmann, H. J. Gröne and C. Körner, *Nat. Med.*, 2011, **18**, 71.
31 T. D. Butters, R. A. Dwek and F. M. Platt, *Chem. Rev.*, 2000, **100**, 4683.
32 E. De Clercq, *Nat. Rev. Drug. Discov.*, 2006, **5**, 1015.
33 E. C. Borden, G. C. Sen, G. Uze, R. H. Silverman, R. M. Ransohoff, G. R. Foster and G. R. Stark, *Nat. Rev. Drug. Discov.*, 2007, **6**, 975.
34 S. Roy-Ghanta and J. S. Orange, *Clin. Rev. Allergy Immunol.*, 2010, **38**, 39.
35 R. Derynck, E. Remaut, E. Saman, P. Stanssens, E. De Clercq, J. Content and W. Fiers, *Nature*, 1980, **287**, 193.
36 P. De Somer, E. De Clercq, A. Billiau, E. Schonne and M. Claesen, *J. Virol.*, 1968, **2**, 878.
37 P. De Somer, E. De Clercq, A. Billiau, E. Schonne and M. Claesen, *J. Virol.*, 1968, **2**, 886.
38 P. Claes, A. Billiau, E. De Clercq, J. Desmyter, E. Schonne, H. Vanderhaeghe and P. De Somer, *J. Virol.*, 1970, **5**, 313.
39 A. Billiau, J. Desmyter and P. De Somer, *J. Virol.*, 1970, **5**, 321.
40 J. Leunen, J. Desmyter and P. De Somer, *Appl. Microbiol.*, 1971, **21**, 203.
41 A. Billiau, J. J. Muyembe and P. De Somer, *Appl. Microbiol.*, 1971, **21**, 580.
42 A. Billiau, J. J. Muyembe and P. De Somer, *Nat. New. Biol.*, 1971, **232**, 183.
43 E. De Clercq and P. De Somer, *Eur. J. Cancer.*, 1972, **8**, 535.
44 E. De Clercq, W. E. Stewart and P. De Somer, *Infect. Immun.*, 1973, **7**, 167.
45 M. W. Harmon, B. Janis and H. B. Levy, *Antimicrob. Agents. Chemother.*, 1974, **6**, 507.
46 O. S. Weislow and E. F. Wheelock, *Infect. Immun.*, 1975, **11**, 129.

47 E. De Clercq and M. Luczak, *J. Infect. Dis.*, 1976, **133** (Suppl:A), 226.
48 S. Kluge, *Antiviral. Res.*, 1985, **5**, 291.
49 H. B. Levy, J. Duenwald and C. E. Buckler, *Infect. Immun.*, 1973, **7**, 457.
50 S. Li, S. Starckx, E. Martens, C. Dillen, N. Lamerant-Fayel, N. Berghmans, M. Gouwy, M. van Pel, H. Heremans, C. Kieda, W. E. Fibbe, A. Billiau, J. Van Damme and G. Opdenakker, *J. Leukoc. Biol.*, 2010, **88**, 1017.
51 S. Li, E. Martens, C. Dillen, P. E. Van den Steen and G. Opdenakker, *Biochem. Pharmacol.*, 2008, **76**, 831.
52 J. Van Damme, A. Wuyts, G. Froyen, E. Van Coillie, S. Struyf, A. Billiau, P. Proost, J. M. Wang and G. Opdenakker, *J. Leukoc. Biol.*, 1997, **62**, 563.
53 A. M. De Palma, H. J. Thibaut, S. Li, I. Van Aelst, C. Dillen, M. Swinnen, E. Verbeken, J. Neyts and G. Opdenakker, *J. Pathol.*, 2009, **217**, 633.
54 G. Opdenakker and J. Van Damme, *Cytokine Growth Factor Rev.*, 2011, **22**, 359.
55 D. Hanahan and R. A. Weinberg, *Cell*, 2011, **144**, 646.
56 H. Piccard, R. J. Muschel and G. Opdenakker, *Crit. Rev. Oncol. Hematol.*, 2012, **82**, 296.
57 H. Piccard, N. Berghmans, E. Korpos, C. Dillen, I. V. Aelst, S. Li, E. Martens, S. Liekens, S. Noppen, J. V. Damme, L. Sorokin, G. Opdenakker, *Int. J. Cancer*, 2012, **131**, E425.

Lipopolysaccharide structure and biological activity from the cystic fibrosis pathogens *Burkholderia cepacia* complex

Anthony De Soyza,*[a] Flaviana Di Lorenzo,[b] Alba Silipo,[b] Rosa Lanzetta[b] and Antonio Molinaro[b]
DOI: 10.1039/9781849734769-00013

1 Introduction

This chapter will focus on lipopolysaccharide in cystic fibrosis (CF) related pathogens.

Cystic Fibrosis is the most common lethal single gene disorder in Europe. The condition arises due to an autosomal recessive gene defect with an incidence of approximately 1/2000 live births in the Western hemisphere. CF has multiple organ system manifestations including liver, gastrointestinal, reproductive and pulmonary involvement. Although a multisystem disease the predominant morbidity and premature mortality relates to recurrent lung infections that lead to respiratory failure. The basic genetic defect in Cystic fibrosis (CF) reflects mutations in a gene encoding for an epithelial chloride channel, named Cystic Fibrosis Transmembrane Regulator (CFTR). There are now over 1000 described mutations. These cause abnormal electrolyte flow across the mucosal surfaces resulting in abnormally viscid airway secretions. This in turn prevents normal mucociliary function and predisposes to chronic bacterial lung infections. Persistent bacterial infection is a hallmark of the condition with episodic worsening ''exacerbations'' leading to multiple courses of antibiotics, hospitalization and precipitous terminal decline. By adulthood the majority of patients with CF will have chronic low grade infection often termed ''colonisation'' despite the significant morbidity associated with this chronic infection state. Importantly colonisation should be seen as a misnomer as there are serious consequences of these infections with evidence of ongoing inflammation. We will use persistent infection hereafter this occurs in most patients with CF by adulthood. Whilst the genetic defect for CF has been understood for over 20 years as yet gene therapy has not emerged as a clinically useful therapy.[1] Gene therapy is felt most likely to be successful in early stage disease in the absence of significant pulmonary infection. Therefore there will still be significant cohorts of patients with CF with established lung infection requiring improvements beyond our current therapeutic regimens. Mortality is improving; CF was previously fatal in childhood. Survival is better but is still limited. The US Cystic Fibrosis foundation notes the mean age at death for those with cystic fibrosis remains less than 45 years. Even in the modern era over 80% die of respiratory complications.[2]

[a] *Transplantation and Immunobiology group, Institute of Cellular Medicine, Newcastle University, UK, The Freeman Hospital, High Heaton, Newcastle-upon-Tyne, NE7 7DN, England, UK. E-mail: anthony.de-soyza@ncl.ac.uk*
[b] *Dipartimento di Chimica Organica e Biochimica, Università di Napoli, Federico II, Napoli, Italy.*

© The Royal Society of Chemistry 2012

1.1 Cystic Fibrosis pathogens

Characteristic pathogens include *Haemophilus influenza*, *Staphylococcus aureus*[3] and in adults the predominant pathogen is *Pseudomonas aeruginosa*.[4] Whilst other pseudomonads can infect patients with CF herein *Pseudomonas aeruginosa* will be referred to as *Pseudomonas* for simplicity. Important but less common bacterial pathogens include the *Burkholderia cepacia* complex (BCC),[5] non tuberculous *Mycobacteria* (NTM), *Acinetobacter* spp,[6] *Stenotrophomonas* spp,[7] *Achrombacter* spp[8–10] and *Pandorea*. Though some of these can occasionally infect normal human lungs these pathogens are predominantly opportunistic pathogens. Despite this predominance of opportunistic pathogens, a surprising variety of bacterial toxins, proteases, lipases and other virulence factors are produced by the pathogens of interest.[11,12] A common clinical pattern is a sequential replacement of pathogens over time from *Haemophilus* and S. *aureus* with *Pseudomonas* usually becoming the predominant pathogen by adulthood.[13] When compared to other CF pathogens persistent *Pseudomonas* chronic infection has been associated with more rapid decline in health status.[4,14–16]

An individual will usually harbour a unique clonal strain over a period of many years. There are however reported outbreaks of epidemic strains within CF clinics particularly noted for *Pseudomonas*,[17–19] *Burkholderia*[20] and *Staphylococcus aureus*.[21] Strict infection control mechanisms appear to greatly reduce these outbreaks.

Selective pressures on bacteria in persistent chronic CF infection are plausible with exposure to repeated antibiotics, continual exposure to host defence molecules and changes in the local environment caused by worsening lung disease all likely forces.[22,23] There is a widely held but poorly defined concept that ''early'' stage strains isolated soon after acquisition may have different characteristics to those seen in persistent ''late'' stage strains.[24,25] The definition of early stage and late stage is complex and somewhat arbitrary. A recent excellent review covers this area well.[23] It is still somewhat unclear if the adaptation takes years or could be occurs within months although some data from infants with CF suggest that changes occur very rapidly.[23] It is biologically plausible that the characteristics required for acute infection such as adhesion and invasion are likely to be different to those required to maintain persistent infection. Certain themes emerge in *Pseudomonas aeruginosa* strains from patients with advanced CF lung disease; antimicrobial resistance rates including pan-resistant strains are highly prevalent. Furthermore development of alginate and the ''mucoid phenotype'' and a hypermutator phenotype are also common during chronicity of infection.[26] Biofilm forming capacity is common in strains isolated cystic fibrosis particularly those who have persistent infection for years.[27,28] There is usually an increased rate of antimicrobial resistance in biofilms[28,29] thereby further compounding the antimicrobial resistance profiles seen in planktonic state.

1.2 Cystic Fibrosis as an inflammatory disorder

CF is characterised by an inflammatory airway syndrome with mucus retention, neutrophilic airway wall inflammation and permanent dilatation of the airways (bronchiectasis). Retention of secretions further promotes

bacteria persistence and prevents mucociliary clearance of virulence factors. These processes collectively promote further airway damage and inflammation leading to the vicious cycle paradigm. Recently CFTR has been shown to be a negative regulator of a central pro-inflammatory controller Nuclear Factor kappa B (NF-κB).[30] Hence mutations in CFTR may be in themselves pro-inflammatory. Cytokines of relevance in CF include interleukin 8 (IL-8), IL-1, IL-6, IL-17 and TNF-alpha.[31–33] These cytokines have a pro-inflammatory effect through vasodilatation, direct chemokinesis, neutrophil priming and angiogenesis. Recently more novel pro inflammatory moieties that have been associated with inflammation in CF include endogenous damage associated molecular patterns (DAMPS) such as high mobility group box 1 (HMGB1)[34] and lipid metabolites such as ceramide.[35] Despite elevated numbers of airway innate immune cells and evidence of microbe specific adaptive immunity there is clearly impaired immunity.[36] There are data noting neutrophil dysfunction in CF with evidence of neutrophil derived peptides (HNP) inhibiting neutrophil activity.[37] This may appear counterintuitive that neutrophils release a peptide that inhibits their own function- the observed effect may reflect the inhibitory activity at high local concentrations. There are also studies suggesting bacteria may cause human receptor dysfunction such as inhibiting CXC chemokine receptor biology.[38] Hence CF is characterised by high levels of cytokines and chemokines that promote vascular dilatation, influx of inflammatory cells and release of noxious defence molecules that may have bactericidal activity but at high concentrations lead to the perpetuation of host damage.

1.3 Cystic Fibrosis as a complex microbial community disease

The conventional paradigm for acute infections arising from a single pathogenic organism still stands *e.g.* acute pneumonia caused by *Streptococcus pneumonia*. In contrast chronic infection in cystic fibrosis may need a newer paradigm. The poor reproducibility of CF lung disease using agar beads and *Pseudomonas* in animal models in part supports this; a non-acute infection can be established but rarely persists with CF type lung pathology. Increasing evidence now supports a concept of polymicrobial infection as evidenced by non-culture based diagnostics and markers of biofilm infection. Non culture based techniques have recently shown a high diversity of organisms in cystic fibrosis.[39–41] Many pathogens or potential pathogens have been identified through such techniques though the direct clinical relevance is unclear. Certainly many of those identified in this manner are inherently un-cultureable either through their obligate intracellular requirements or need for anaerobic conditions[42] which cannot be replicated in a routine diagnostic microbiology lab. Biofilms are complex communities of bacteria with physical and functional aspects. Biofilms can be seen in the CF lung that contributes to pathogenesis by increasing resistance to phagocytosis, increased rates of antibiotic resistance[28,43,44] and is likely to contribute to persistence of organisms. Good *in vitro* data supports many CF pathogenic organisms can form biofilms and that they can do so in monoculture and in complex cultures.[29] The importance of multispecies biofilms is that supposedly low pathogenicity bacteria such as oral "commensal" can induce virulence factor upregulation in *Pseudomonas*

aeruginosa. The effect of complex biofilm growth on cell surface virulence determinants is to date still an emerging field.

2 Basic Lipopolysaccharide structure

Lipopolysaccharides (LPS) are the major components of the outer membrane of Gram-negative bacteria and exert, in isolated form, a variety of biological activities in animals and plants.[45,46] From the structural point of view, LPSs are amphiphilic macromolecules that show the same general chemical architecture in all of the Gram-negative bacteria, that is, it does not depend on the bacterial activity, (pathogenic, symbiotic, commensal), nor on its ecological niches: (human, animal, soil, plant, water) or on its growth conditions. They are indispensable for growth and survival of Gram-negative bacteria and comprise about 10–15% of the total molecules in bacterial outer membranes covering about 75% of bacterial surface area, thus they are exposed to the bacteria's external environment. LPS represent a defensive barrier which helps bacteria to resist to antimicrobial compounds and environmental stresses and are involved in many aspects of host–bacterium interactions as recognition, adhesion, colonization. Once LPS molecules are released from dead bacteria they play a key role in the pathogenesis of Gram-negative infections potentially causing fever or circulatory shock. LPS has additional key roles in virulence, tolerance for commensal bacteria and symbiosis.[45] From both genetic and chemical viewpoints, LPSs have a common structural architecture consisting in three domains, (Fig. 1): a polysaccharide (known as O-side chain, O-specific polysaccharide or O-antigen) which is covalently linked to an oligosaccharide part (core), in turn, linked to a glycolipid portion (Lipid A).

The lipid A component of LPS is part of the outer leaflet of the external membrane whereas the sugar moiety is oriented outwards.[47] Bacterial colonies can be catalogued by gross colony morphology as *rough* or *smooth* which is related to the occurrence of the O-side chain, being absent in the former and present in the latter. The terminology presently used to designate the different LPS types, namely with or without the O-polysaccharide

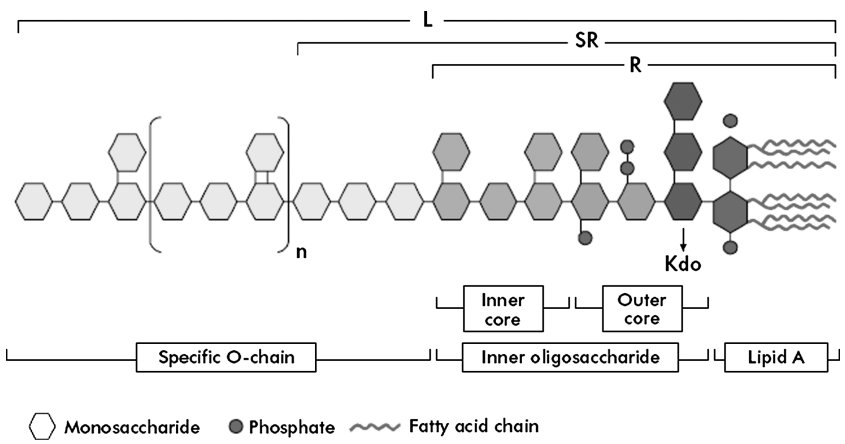


Fig. 1 Scheme of a general chemical structure of bacterial lipopolysaccharides.

antigenic portion, is S-LPS or R-LPS (or lipooligosaccharide, LOS), respectively.

In all cases, the LPS fraction of an individual bacterial strain is characterised by an intrinsic size and structural heterogeneity. Hence it is more appropriate to view LPS as a blend of various LPS molecules (and similarly so for LOS).

2.1 Structural investigation of the glycolipid (LPS or LOS)

LPSs (or LOSs) may be extracted from intact cells by a variety of procedures, but currently those commonly used exploit the phenol/chloroform/light ether and/or phenol-water extractions. Depending on the polarity of glycolipid it can be isolated from aqueous, phenol or the organic phase.

Enzymatic treatments that remove contaminating proteins and nucleic acids and chloroform washings that remove lipid and phospholipids are frequently used. The amphiphilic nature of these molecules is responsible for their low solubility in both aqueous and organic solvents and renders the purification by any chromatographic technique problematic. In contrast electrophoretic analysis using a denaturing agent is highly informative and allows chemotyping of the extracted material: a ladder like profile is diagnostic of LPS molecules reflecting the O-polysaccharide. In contrast the presence of material at bottom of the gel profile, without the ladder profile, denotes the smaller molecule LOS lacking O-polysaccharide. This information is valuable and affects the subsequent procedures for the structural investigation.

However, in both cases, the study of the lipid A moiety is carried out after its cleavage from the rest of the molecule by mild acid hydrolysis of the acid-labile glycosydic linkage of Kdo. During this treatment, the lipid A part precipitate whereas the saccharide moiety remains soluble in the supernatant.

The primary structure determination of very complex molecules such as LPSs (LOSs) currently exploits state-of-art 1D and 2D NMR experiments and soft-ionization mass spectrometry techniques, together with compositional and linkage chemical analyses, usually carried out on intact products (when possible) and *ad hoc* selectively degraded derivatives.

2.2 Lipid A structure elucidation

The lipid fraction obtained by precipitation, after mild acid treatment of LPS (or LOS) consists of a mixture of intrinsically heterogeneous lipid A molecules. The structural approach makes an extensive use of mass spectrometry analysis by MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization – Time Of Flight mass spectrometry) and ESI MS (Electrospray Ionization mass spectrometry) techniques both on the native and selectively degraded lipid A fractions. The MALDI-TOF and ESI-MS data allow gaining insights into the number of lipid A species present in the fraction, the presence of polar heads and the distribution of acyl residues on each GlcN units of the disaccharide backbone.[48] This latter issue is of particular importance given that fatty acid composition and distribution strongly affects lipid A/LPS endotoxic activity, *i.e.* agonist/antagonist activity.

Through the chemical and spectroscopical study of several lipid As and LPSs in our laboratory, novel general approaches to gain structural insights

on such molecules have been developed. In particular, combining MALDI-MS or ESI-MS and selective chemical degradation by ammonium hydroxide hydrolysis, we have developed a general and practical methodology to obtain the secondary fatty acid distribution, which is one of most demanding issues in the structural determination of lipid A.[49,50,51] These approaches must be supported by the classical chemical analyses of the lipid A, which imply the removal of ester-linked fatty acid moiety by mild alkaline treatment, or complete removal by strong alkaline treatment and GLC-MS analyses of fatty acids as methyl ester derivatives.[52,53]

The NMR study of LPS is less frequently used but is equally important[54,55]; the amphiphilic nature of lipid A entails a very poor solubility of the sample in whichever solvent system, however, when possible it is invaluable in obtaining diagnostic information. NMR is highly useful for studies on the saccharide backbone, such as the anomeric configuration of the carbohydrate residues and their sequence, the occurrence and position of the polar heads.

2.3 O-chain polysaccharide structure

The O-chain of LPSs, recovered in supernatants after mild acid treatment, usually consists of a regular long chain polysaccharide with the core oligosaccharide still at the reducing end of the chain, but, since this oligosaccharide moiety represents a small percentage of the whole polymer, it is generally barely detectable from the chemical and spectroscopical point of view. The O-chain structure determination aims to define the composition in monosaccharides, the relative, the anomeric and the absolute configurations of each monosaccharide, their ring sizes, the linkage of each residue and their sequence within the repeating unit. In some cases, the non-stoichiometric presence of some substituents can determine heterogeneity and hide the structural regularity.[56]

The monosaccharide composition is achieved by the classical analytic approaches consisting in acid hydrolysis and/or methanolysis followed by further chemical modification (such acetylation, for example) to perform GLC-MS analyses. The reaction sequence: methylation, acid hydrolysis, reduction and acetylation are carried out to know the glycoside substitution sites and the ring sizes of each monosaccharide. For the absolute configurations, GLC analysis of convenient glycoside derivatives are usually exploited. The anomeric configurations and the sequence of the monosaccharide residues are determined by 1D and 2D ^{1}H and ^{13}C-NMR experiments (COSY, TOCSY, NOESY, ROESY, HSQC, HMBC). In detail, TOCSY and DQF-COSY spectra allow the correct identification and assignment of all ring protons, and on the basis of these data, the assignment for all ^{13}C resonances usually follows from the analysis of an $^{1}H,^{13}C$-HSQC spectrum; the *intra*-residue pattern of dipolar correlations in the 2D NOESY or ROESY spectra gives the proof for specific configurations. Information concerning the sequence of the monosaccharides within the oligosaccharide (or the repeating unit of the polysaccharide) are then deduced from the observation of the *inter*-residual dipolar correlations in 2D NOESY and ROESY spectra and from the existence of scalar long

range correlations, in the $^{1}H,^{13}C$-HMBC spectrum, among the ^{1}H and ^{13}C nuclei involved in glycosidic linkages.

2.4 Core oligosaccharide structure

The target of the structural elucidation of a LOS core oligosaccharide is generally much more complex than that of the O-chain because of the higher monosaccharide number, commonly up to 15, with respect to the components of one O-chain repeating unit, and the presence of labile groups and other substituent that are often present in non stoichiometric amounts. In addition, the acid hydrolysis used to remove the lipid A from LOS occurs by hydrolysis of Kdo that in its reducing form usually appears in different anomeric and cyclic forms, thus determining an ''artificial'' increase of the core heterogeneity. Therefore, LOS sample is generally treated under alkaline conditions that do not cleave the Kdo linkage thereby allowing the analysis of the complete sugar backbone including the two GlcN of lipid A. Notably the LOS sample should be first treated under mild alkaline conditions in order to selectively remove the lipid A ester-bound fatty acids. This more soluble *O*-deacylated LOS is *N*-deacylated under more strong alkaline conditions affording completely deacylated backbone oligosaccharide, which represents the intact sugar skeleton of LOS suitable for sequence determination by NMR analysis.[57] The direct strong de-*O* and de-*N*-acylation is usually not performed in order to avoid the formation of alkali resistant unsaturated N-linked fatty acids. In order to assign phosphate groups and their possible further attached substituents 1D and 2D ^{31}P-NMR experiments are also carried out.

Since alkaline treatment frequently determines the loss of base-labile functions, the oligosaccharide moiety released under acid conditions is also studied as a complemetary and informative approach. More recently, the full NMR study of the O-deacylated sample has been carried out in a water solution containing a denaturing agent. Every step of the selective degradation is usually accompanied by MALDI MS spectrometry analysis that is a very valuable and complementary approach in the study of such molecules. Our collaborators have pioneered the development of new MALDI MS approaches through development of new methods of preparation and experimental study. It is worth emphasising that when such experiments are carried out on the intact sample they are often of very high value as they yield a ''picture'' of the intact endotoxin without any loss of structural information.[58]

2.5 Relevance of structure

The importance of LPS structure is made eloquently clear in the clinical management of CF. A commonly used therapy includes polymixin E or colistin; this cationic antimicrobial peptide has significant microbicidal activity against many CF pathogens and attacks the bacterial cell membrane through a physicochemical interaction with LPS. This interaction is determined by the LPS composition and surface charge. Hence the majority of CF isolates of *Pseudomonas* have negligible rates of clinically relevant colistin resistance whilst *Burkholderia* species are to date almost always resistant. Further pathogens in CF such as the enterobacteriae *Serratia* and *Proteus* are also often resistant to colomycin which again reflects structural

composition of the lipid A as discussed below. These resistance patterns dictated by lipid A composition are reproducible enough to have some diagnostic value: If clinical diagnostics laboratory detects a *Pseudomonas* isolate with high level resistance to colistin- a repeat microbial species-identification step is often undertaken as there such low rates of clinically significant colistin resistance in *Pseudomonas*. Often on closer inspection the organisms has proved to have been misidentified and is not a *Pseudomonas* species. Eloquent experiments in knockout models have demonstrated that the structure of LPS particularly the lipid A component has a marked effect on the pro-inflammatory activity *in vitro*. For example the loss of a lipid A single acyl side chain in *Neisseria*, a non CF pathogen, can reduce pro- inflammatory activity *in vitro* models by up to 50%. Similar data for *E. coli* suggests the importance of lipid A acylation patterns across a diverse range of pathogens.[59] Hence LPS structure may dictate two important clinical end points: resistance to antimicrobial therapy and pro- inflammatory activity.

Recent reviews have comprehensively studied *Pseudomonas* LPS and the reader is directed to these.[60,61] We will briefly highlight important points.

3 The detailed chemical structure of the *P. aeruginosa* LPS

Pseudomonas aeruginosa has a typical Gram-negative bacterial LPS, with a basic lipid A structure containing an *N*- and *O*-acylated diglucosamine bisphosphate backbone [4-*P*-α-D-Glc*p*NII-(1→6)-α-D-Glc*p*NI-(1→*P*] with chemical variation in the number of primary acyl groups and the types of fatty acids substituting the primary and secondary acyl groups (Fig. 2). Most of the laboratory derived *Pseudomonas* strains studied so far synthesize a penta-acylated (75% of the molecules) LPS, with some proportion made as a hexa-acylated LPS (25% of the molecules). The difference between the 2 isoforms is the lack of an O-linked 3-hydroxy decanoic acid (10:0 (3-OH)) group at position 3 of the first glucosamine in the penta-acylated isoform (Fig. 2). Certain Growth conditions *e.g.* low magnesium levels, can result in altered *P. aeruginosa* lipid A acylation patterns. Pseudomonas isolated from chronically infected CF patients are often unable to synthesise O-antigen side chains; these strains have a predominant hexa-acylated LPS. In certain cases a hepta-acylated lipid A has been isolated, containing an additional palmitoyl (16:0) group (Fig. 2) linked to the primary 3-hydroxy decanoic acid group at position 3′ of glucosamine 2.[62] The hexa- and hepta-acylated lipid A moieties also contain cationic 4-amino-4-deoxy-L-arabinose sugars. Bound to the lipid A is a relatively conserved inner-core structure which contains two D-*manno*- oct-2-ulosonic acid residues (KdoI and KdoII) and two L-*glycero*-D-*manno*-heptose residues (HepI and HepII).[63] Bound to the second heptose residue, HepII, is a 7-*O*-carbamyl group,[64] which is found in the LPS of other pseudomonads.[65] The 2 heptose residues are often phosphorylated at positions 2 and 4 of HepI and position 6 of HepII.[66,67] Phosphate substituents can be mono-, di- or even tri-phosphates, with most analyzed *P. aeruginosa* LPS having some tri-phosphate, which, to date, has only been detected in the LPS of this bacterial species. Ethanolamine mono- or diphosphate is present in some

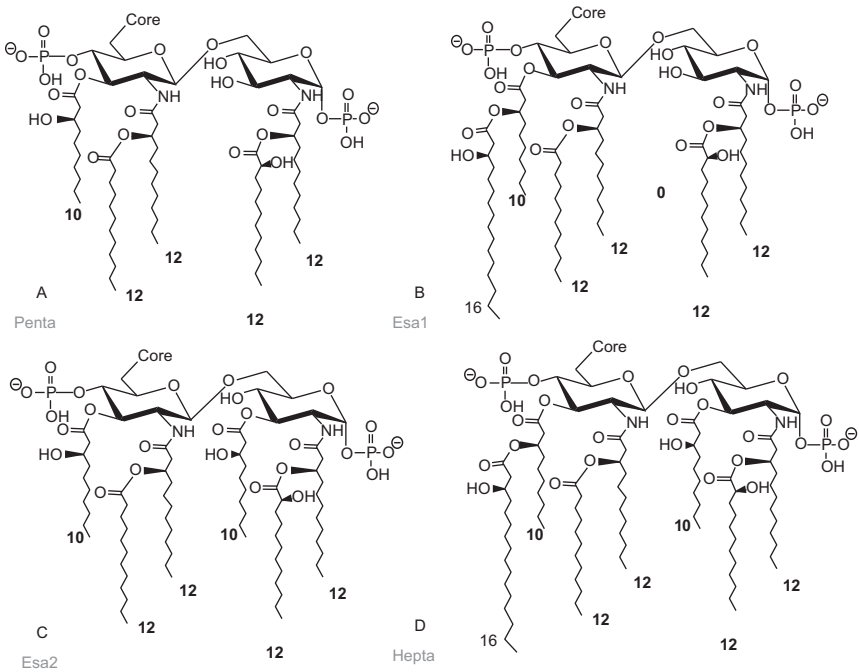


Fig. 2 Chemical structure of *P. aeruginosa* lipid A from clinical strains isolated at the onset of chronic colonization (**A** is the structure after 6 months of colonization) and after years of chronic colonization (**B**, **C** and **D** are after 7.5 years of colonization by a mucoid strain and a non mucoid strain respectively).[77]

Pseudomonas LPS inner cores in non stoichiometric amounts. Notably a mutation in a *waaP* gene that phosphorylates HepI groups at position 4 is lethal strongly suggesting that phosphorylation is critical for bacterial survival.[68] The outer core of the *P. aeruginosa* LPS is usually synthesized as 2 different isoforms or glycoforms by an individual strain.[69] Both outer-core glycoforms contain an N-alanylated galactosamine residue, 3 D-glucose residues, and one L-rhamnose residue the position of which differs in the 2 glycoforms. There is also some structural variation among the different *P. aeruginosa* strains in glycoform I, with a less common variant having four glucose residues in the outer core.[69] There is extensive O-acetylation of the hydroxyl groups in the outer-core sugars, but the acetates are not present in high amounts at any one position, making a clear structural determination of these substituents difficult.

As O-acetyl groups are relatively labile under mild acid or basic conditions they may be lost during LPS purification. If the terminal rhamnose is not substituted with another monomonosaccharide it is often acetylated. With the exception of O-antigens from O14 and O15 serogroups most *P. aeruginosa* LPS-smooth strains studied to date have glycoform II fully substituted with an O-antigen. LPS-rough strains, usually isolated from CF patients, have multiple mutations within the biosynthetic genes for the O-antigen[70] and do not have any O-antigen on glycoform II.[71] The O-antigen portion of the *P. aeruginosa* LPS is responsible for conferring serogroup specificity, which is defined by antibodies specific to the different

variants of this antigen. A variety of serotyping systems have been proposed in the past; most of these can be reconciled when cross-referenced with prior data and strains.[70] In general, most investigators now refer to serogroups using the International Antigenic Typing System (IATS) schema. Chemically the O-antigens can be categorized into at least 11 structural variants,[66] although within these variants there are minor differences among related structures in properties such as sidegroup substituents, linkages between sugars or conformations of different monosaccharides.[66] *P. aerguinosa* serogroups O14 and O15 strains both lack detectable O-antigens or LPS containing even a single O-antigen repeat linked to the glycoform 2 outer core (referred to as the SR-type LPS). This may point to the fact they probablyy LPS-rough bearing strains in spite of the availability of serogroup-specific antisera. It is pluasible that the antisera to the serogroup O14 and O15 antigens are cross reacting with LPS core epitopes or other components in these serogroup strains. Typical sugars within the *P. aeruginosa* LPS O side chains include N-acyl derivatives of different amino sugars along with rhamnose.[71] The structures of some of these monosaccharides were first described in *P. aeruginosa* LPS. Rhamnose is also commonly found in some of the O-antigens. The monosaccharides are arranged in repeat units containing 3 to 4 individual monosaccharides, except for serogroup O7, which is a disaccharide repeat unit. The linkage of the monosaccharide in the first repeat unit to the rhamnose residue in glycoform II of the LPS outer core is usually a 2-*N*-acetyl derivative of a 6-deoxy-D-hexosamine, usually D-quinivosamine or D-fucosamine, although for serogroup O3 it is the 4-amino, 4-deoxy derivative of D-quinivosamine.

3.1 Structural adaptation of LPS to the host CF airway

Selected Gram-negative bacteria have evolved mechanisms to modify the structure of lipid A in different environments, including in host tissues. There is evidence that bacterial resistance mechanisms, such as lipid A modifications may have evolved to protect against host cationic antimicrobial peptides (CAMP).[72] The mucosal epithelial surface produces a number of antimicrobial peptides. Hence alteration of the surface charge of bacteria provides a defence mechanism for the bacteria against these host derived CAMPs.[73] The systems for evoking these changes are only partly understood and in *Pseudomonas* appear to involve a two component magnesium dependent sensor system PhoP/PhoQ.[74–76] This mechanism does not appear present in all CF pathogens with no gene homologues easily identified in BCC, for example (De Soyza, unpublished observations). Selected CF pathogens are associated with specific chemical alterations in their outer membrane when the strain has become chronically established. *P. aeruginosa* clinical bacterial isolates from CF have a unique lipid A structure.[62] These lipid A molecules are structurally different from those that are found in environmental isolates and isolates from other human diseases including acute urinary, blood or eye infections.[66] Lipid A from *P. aeruginosa* CF isolates, in contrast to isolates from other human infection sites, markedly stimulate human Toll receptor 4 (TLR4) signalling and

might contribute to the inflammatory response that is seen in patients with early stage CF.[66] Early adaptation of *P. aeruginosa* to the CF airway by synthesis of a modified lipid A, that includes the addition of aminoarabinose and palmitate, leads to resistance to host innate immune defences (such as cationic antimicrobial peptides) and increased pro-inflammatory signalling. These modifications appear stable and were not lost after continual passage in rich medium *in vitro*. These results suggest that lipid A modifications are constitutively and stably expressed in *P. aeruginosa* isolates from CF patients after growth *in vitro*.[66] In general, production of fully hexa-acylated lipid A as seen in CF is associated with a more strong TLR4 mediated inflammatory response while lipid A with lower levels of acylation triggers reduced cellular responses. in contrast to early stage CF infections there is a switch to penta- and hexa-acylated lipid A in later stages of CF. Therefore the LPS from a *P. aeruginosa* strain isolated from a patient with late stage CF lung disease may induce less inflammation than LPS from an early isolate from the same individual. Such LPS modifications in CF may be a form of host immune evasion.[77] Importantly there is a species specific effect whereby human TLR4 signalling is altered by these chronic strains ''CF lipid A'' changes whilst mouse TLR4 did not respond differentially to early stage and late stage CF lipid A.[78] This has led to the concept that the bacteria adapts specifically to the host TLR.[77] Recent data has suggested CF strains not only differ in their potency to induce inflammatory responses as compared to environmental strains but they also have differing cytokine induction profiles. Divergent responses were seen when comparing ''early stage'' CF infection strains to ''late stage'' strains and an non clinical strain PAO1: The LPS from strain PAO1 was a more effective inducer of IL-8 than any of the early or late stage CF strains but elicited less TNF-alpha from epithelial cells.[77]

Previously reviews have covered important aspects of the important virulence determinants BCC lipid A structure, O polysaccharide and link these to biological activity.[46,79] We will highlight relevant newer data and where contextually important re-address selected elements of data previously covered.

4 *Burkholderia cepacia* complex and Cystic Fibrosis

''*Burkholderia cepacia*'' has long been identified as a plant pathogen where it causes onion rot.[80] More recently it has been recognized as an opportunistic human pathogen for individuals with cystic fibrosis (CF)[81] and also those with chronic granulomatous disease.[82] Previously, infections with ''*Burkholderia cepacia*'' were attributed to a single species. Indeed, such infections are caused by a variety of bacterial species, grouped into the ''*Burkholderia cepacia* complex (BCC)''.[20] Currently this complex comprises over 10 related species named *genomovars* which are phenotypically similar but genotypically different. The use of formal species names is now preferred although to date there are no phenotypic tests that separate each genomovar.

The most prevalent clinical isolates from CF patients are *B. cenocepacia* strains (genomovar III) and *B. multivorans* strains (genomovar II)[83]

and infection with the former is most often associated with patient-to-patient spread and fatal invasive infection (the so-called "*cepacia syndrome*").[11]

Lung transplantation, whilst intensive and aggressive, is to date the only management strategy that offers improved quantity and quality of life of CF patients with advanced lung disease; BCC's impact on the survival of CF patients following lung transplantation is viewed as so detrimental that infection with these organisms is considered a relative or absolute contra-indication for transplantation.[84,85] Post-transplant survival rates are 15% to 63% lower than those of CF patients not infected with BCC,[86] with death usually resulting from BCC sepsis in the early postoperative period.[84] Not all BCC species cause poor outcomes following lung transplantation.[86–89] Furthermore, there are great differences *in vitro* among different isolates from the same genetic lineage *e.g.* for the renowned *B. cenocepacia* epidemic strain ET-12 clonal strains J2315 and K56-2 strains[90] behave differently when used *in vitro* and *in vivo* models. Hence this heterogeneity illustrates the importance of examining multiple clinical bacterial isolates and not generalizing the results from studies of a single strain.

Molecular differences among the BCC family has been suggested as the cause of their variable pathogenicity, and their LPS has been proposed as potentially important in this. Studies conducted soon after genomovar designation have assessed LPS chemotypes within the BCC.[91] So far in our labs all strains tested belonging to *B. cepacia* genomovar I and *B. multivorans* had smooth LPS. In contrast, isolates belonging to *B. cenocepacia*, expressed either rough or smooth LPS chemotype.[92,93] Importantly two clones of the *B. cenocepacia* ET 12 lineage differed with J2315 and K56-2 having a rough and a smooth chemotype respectively. Subsequent studies in clinical BCC strains assessing chemotype have demonstrated both rough and smooth LPS chemotypes are found. Other genomovars studied had predominantly smooth LPS though *B. multivorans* and *B. vietnamiensis* strains expressing rough LPS were noted.[94] Based upon historical studies using rabbit antisera the LPS core of the strain *B. cenocepacia* J2315 appears antigenically related to some but not all other *B. cenocepacia* strains (reviewed in[46,79]).

Interestingly very limited cross reactivity, between *B. cepacia* genomovar I LPS and no cross reactivity with *Pseudomonas* LPS was observed when probed with antisera taken from patients suffering from *B. pseudomallei* infection using a microbead-LPS conjugate based assay. The anti-sera from *B. pseudomallei* infected patients did however cross react with more phylogenetically related bacteria such as *B. mallei* and *B. thailandensis* (one previously called arabinose-negative *B. pseudomallei*).[95] Collectively these data suggest distinct antigenic clustering within the *Burkholderia* genus with similarities between BCC genomovars but not other *Burkholderia* species or *Pseudomonas*. Readers should be aware of the importance of growth conditions *in vitro*: Variation in antigenic cross reactivity and LPS phenotypes has been reported in a comparative study of BCC bacteria grown in either solid agar or broth. Hence certain BCC strains may differentially express LPS dependent on the local environment. This is clinically important as airway mucus may serve as a semi-solid growth media

within CF lungs. As most published studies including our own have extracted LPS from bacteria grown in broth a careful appraisal of these prior studies is needed.

4.1 General structural features of the lipid A of BCC LPS

According to early compositional data of Cox *et al.*, all the lipid A structures of the *Burkholderia* strains elucidated contain a general architectural in which the following elements are always present: the two 2-amino-2-deoxy-glucopyranose (GlcN) carbohydrate backbone and an additional carbohydrate element, the 4-deoxy-4-amino-arabinopyranose (Ara4N); the phosphate groups; the $C_{16(3\text{-}OH)}$ and $C_{14(3\text{-}OH)}$ as primary fatty acids and the $C_{14:0}$ as secondary fatty acid. These elements are very often present in a non stoichiometric fashion giving rise to a wide array of lipid A primary structures that varies on the strain and the environment from which the bacteria has been isolated. As for fatty acids, the above moieties are the most abundant component found in the BCC lipid As, but in selected strains and in particular conditions other minor components have been found, such as $C_{12:0}$ (rather than $C_{14:0}$) or $C_{14:(2\text{-}OH)}$ fatty acids.[46] As yet the only isolated structures of BCC lipid A are penta- or tetra-acylated with no formal identification of hexa-acylated structures. All these data confirmed previous studies noting a resistance of BCC strains to host defence peptides different to that of *Pseudomonas* which suggested divergence in the lipid A and LPS structure between BCC strains from that of *Pseudomonas* species. Indeed, various BCC strains, including three CF isolates and two reference strains, were more resistant to killing by protegrin-1 (PG-1), a potent antibiotic peptide able to prevent infection *via* attack on the bacteria membrane surface. Furthermore, binding/diffusion and surface plasmon resonance assays revealed that isolated lipopolysaccharides and lipid As from the sensitive *P. aeruginosa* strains bound PG-1 more effectively than LPS and lipid A of resistant BCC strains.[46]

Although a body of publications related to the biology and biochemistry of lipid A/LPS from BCC has appeared in recent years, scanty data are available on the molecular structure of these important biomolecules.

In this chapter we report the general features of BCC strains lipid A characterized to date.

4.2 *Burkholderia cepacia* genomovar I

The primary structure of the lipid A of *B. cepacia* genomovar I has been recently elucidated[96] by chemical degradations, mass spectrometry and NMR spectroscopy. These analyses revealed a mixture of species differing in the acylation and in the phosphorylation patterns. The most frequent species was found to be a penta-acylated tetrasaccharide backbone with two phosphoryl-arabinosamine residues in addition to the archetypal glucosamine disaccharide[Ara*p*4N-L-β-1-P→4-β-D-Glc*p*N-(1→6)-α-D-Glc*p*N-1→P-1-β-L-Ara*p*4N].[96] Lipid A fatty acids substitution was also noted, with two 3-hydroxytetradecanoic acids 14:0 (3-OH) in ester linkage, and two 3-hydroxyhexadecanoic acids 16:0 (3-OH) in amide-linkage, one of which was substituted by a secondary a 14:0 residue at its C-3.

4.3 *Burkholderia multivorans* genomovar II and *B. vietnamiensis* genomovar V

Both *B. multivorans* and *B. vietnamiensis* possess an equally heterogeneous lipid A in which tetra-acylated and penta-acylated species (lacking or not a primary 14:0 (3-OH) in ester linkage to the GlcNI) coexist in nearly equal amounts. The lipid A in both these genomovars can bear up to two Ara4N residues with phosphate groups also present in non stoichiometric amount.

4.4 *Burkholderia cenocepacia* genomovar III

The *B. cenocepacia* genomovar III ET-12 clone (J2315) lipid A has a relatively marked heterogeneity in fatty acid content with a consistent proportion of tetraacylated species and only one Ara4N residue at anomeric position of GlcNI.[93] This contrasts with the strain *B. cenocepacia* LMG12614, another ET-12 clone, which possesses a much more homogeneous lipid A in which there is a major penta-acylated species with one or two Ara4N residues. Interestingly our early data noted many shared lipid A species between *B. cenocepacia* LMG12614 (an ET-12 clone) and *B. multivorans* LMG14273 though there was a greater proportion of penta–acylated species in the *B. cenocepacia* strain.[93]

Thus far, the fully established lipid A structures of BCC LPS consist of a blend of pentacylated and tetracylated structures (lacking a primary 14:0 (3-OH)) in which two phosphoryl-arabinosamine residues can be present in non stoichiometric amount.[46] These chemical structures have strong similarities with other non BCC *Burkholderia spp*. lipid A, as for example the plant pathogen *B. caryophylli* LPS[46,97] in which the lipid A is nearly identical to the one of *B. cepacia* LPS (Fig. 2). Collectively these data suggest the biosynthetic pathways for lipid A may not be unique to a particular genomovar but differential lipid A acylation and glycosylation by Ara4N may depend on environmental pressures and possibly genomovar specific responses to such pressures (Fig. 3).

4.5 Importance of lipid A structure of BCC strains in CF

Both lipid A acylation and substitution by aminoarabinose (Ara4N) are pivotal in the survival of Gram negative bacteria against host response.[98] Ara4N substitution in lipid A and LPS inner core contributes to the overall net charge surface of the external membrane and plays a key role in pathogenesis. In fact, LPS from the *Burkholderia* genus are frequently positively charged or in an isoelectric state. This is closely related to the abundance of Ara4N residues present in the lipid A-inner core and confers resistance to antibiotic compounds particularly host cationic antimicrobial peptides. In physiological conditions Ara4N moieties are positively charged which reduces the net surface charge on the bacterial external membrane weakening the ionic attraction for respiratory tract antimicrobial defensins. The inherent resistance of the BCC micro-organisms to polymyxin B, a cyclic polycationic antibiotic peptide that has highly affinity for negatively charged bacterial LPS, is likely to be Ara4N dependent. These residues prevent the antimicrobial action of polymyxin B, namely increasing permeability in bacterial outer membrane. Interestingly, the other frequent cationic binding groups (*e.g.*, ethanolamine) have not been reported in lipid

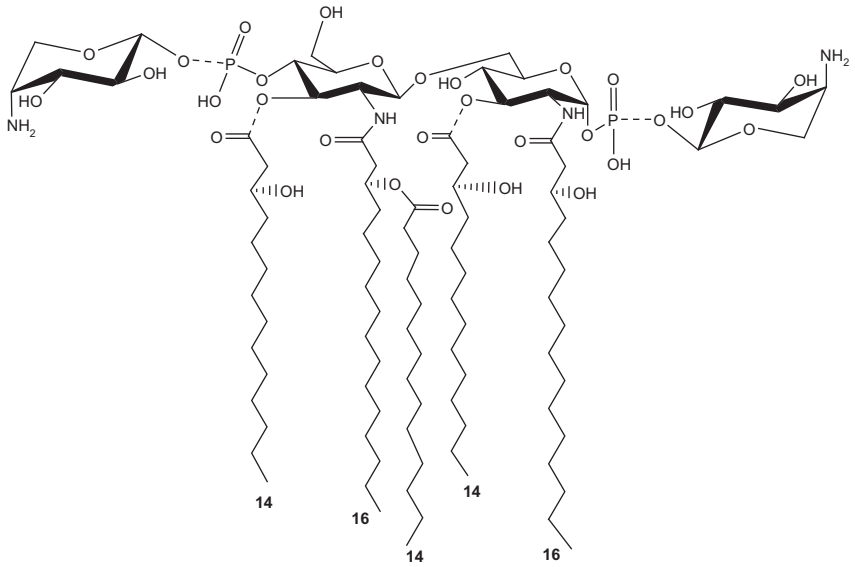


Fig. 3 Lipid A from *B. cepacia* genomovar I. The structure of the different lipid A molecules from *B. cepacia* LPS sketched in a single formula. The dotted lines indicate non stoichiometric substitutions

A from *B. cepacia*, suggesting a unique role for L-Ara4N.[98] Hence, the transferase enabling the attachment of L-Ara4N residues to lipid A and whichever earlier implicated enzymatic step may be a potential novel target for the future developments of drugs against *B. cepacia*. The gene cluster encoding aminoarabinose (Ara4N) biosynthetic enzymes is essential for *Burkholderia cenocepacia* viability. In a conditional mutation strategy using rhamnose supplemented media, marked alteration of bacterial cell permeability, osmotic sensitivity, gross colony morphology and ultrastructure was seen when aminoarabinose biosynthesis was disrupted. These data suggest a fundamental role for Ara4N in maintaining normal cell membrane structure and function in BCC. The biological activity at inducing cytokines from human cells of the different BCC species is also likely determined by the lipid A acylation pattern. Elegant work on *E. coli* lipid A[59] demonstrates the highest biological activity correlates with the higher degrees of acylation degree and to the asymmetrical distribution of the fatty acid alkyl chains with respect to the two glucosamine residues. Collectively these factors leads to conformational changes in the lipid A, that are presumed to affect PAMP receptor engagement. These lipid A changes are closely related to cytokine induction from host immune cells. The activation of host innate immunity system is, in all likelihood, elicited by natural blends of variously acylated bacterial lipid A species[99]; in case of BCC, clinically pathogenic strains are likely to express more heavily acylated forms *e.g.* penta-acyl lipid A as compared to tetra-acylated forms, and this may be one of the reasons of the elevated response observed in the colonized organisms.[46] Similarly, *Porphyromonas gingivalis* a Gram negative oral pathogen produces a LPS with lipid A structural variants differing in the acylation pattern, with penta-acylated and tetra-acylated forms, that can induce

diametrically opposed effects on the immune system, *i.e.*, by altering the relative amount of the two lipid A structures, *P. gingivalis* modulates innate host response.[100]

5 General structural features of the core region of BCC LPS

To date there are few studies published of BCC LPS core region structure. This is surprising as the core region of BCC LPS plays a key role in conferring physico-chemical and biological properties to the outer membrane. The limited data set means it is impossible to form definitive viewpoint on the chemical composition and structure of all BCC species.[46] A unifying and characterising feature of BCC strains tested so far is the presence of the disaccharide D-*glycero*-D-*talo*-oct-2-ulosonic acid-(2-4)-3-deoxy-D-*manno*-oct-2-ulosonic acid, (Ko-(2-8)-Kdo) in the inner core of BCC LPS.[46] Commonly though not universally in BCC strains the Ko can be also glycosylated at position O-8 by a Ara4N residue, therefore up to, three Ara4N residues can be present within the lipid A-inner core region of BCC LPS. Within the outer core, heptoses are present, which can be in L,D or the D,D configuration. To date, no phosphate or other anionic residues have yet been identified within the outer core of the tested BCC strains.[46] The presence of a complete BCC LPS inner core oligosaccharide is required for *in vitro* and *in vivo* bacterial survival.[101] The hldA and hldD genes in *B. cenocepacia* ET-12 strain K56-2, allow smooth LPS expression and knockdown of this reduces bacterial persistence. These genes encode enzymes directing the modification of heptose sugars prior to incorporation into the LPS core oligosaccharide. A double knockdown mutant, defective in both hldA and hldD expression, confirmed that both these *B. cenocepacia* genes were required for synthesis of a complete LPS core oligosaccharide. The LPS produced by $hldA^-/hldD^-$ mutant strains consisted of a short lipid A-core oligosaccharide devoid of the O-antigen. Limited $hldA^-/hldD^-$ bacterial persistence was noted in a rat lung infection model *in vivo*. This may be due to increases sensitivity to antimicrobial peptides polymyxin B and human neutrophil peptide 1 (HNP-1) *in vitro*.[101] In contrast, another *B. cenocepacia* mutant strain that produced complete lipid A-core oligosaccharide but lacked polymeric O-antigen was not sensitive to polymyxin B or melittin.[101] These data are confirmatory to the early data showing that rough LPS bearing "*Pseudomonas cepacia*" isolates were serum sensitive; collectively these confirm the important virulence role for core oligosaccharides. Most recent data showed that it is possible to obtain heptose-less LPS strains of *B. cenocepacia* with high resistance to polymyxin B, and suggests that this may occur through LPS-independent changes that stabilize the outer membrane in some way.[102]

Here we report all available data on the limited number of BCC LPS core region structure published.

5.1 *Burkholderia cepacia* genomovar I

In *B. cepacia* genomovar I an incomplete core structure has been shown in which core oligosaccharide, a heptosyl trisaccharide, is present and directly linked to O-5 of Kdo, that in turn bears at O-4 the disaccharide Ara4N-(1-8)-Ko as explained above. The outer core also includes glucose and

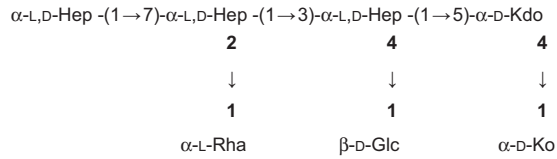


Fig. 4 Core oligosaccharide from *B. cepacia*.

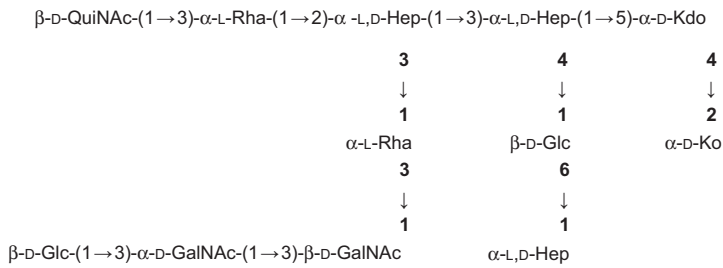


Fig. 5 Core oligosaccharide from *B. multivorans*.

rhamnose. The substitution of Ara4N in the LPS of *B. cepacia* genomovar I to date has been found to always occurs on the Ko residue linked to the Kdo located in the inner core region[103] (Fig. 4).

5.2 *Burkholderia multivorans* genomovar II

B. multivorans LPS core region consists of a much longer and complicate oligosaccharide in which the inner part matches the structure of *B. cepacia* genomovar I whilst the outer consists of a branched oligosaccharide with residues present in non-stoichiometric amount (Fig. 5) such as the presence of a [α-D-Glc-(1→3)-α-D-GalNAc-(1→3)-α-D-GalNAc] trisaccharide sitting on a rhamnose residue.[49]

5.3 *Burkholderia cenocepacia* genomovar III

The core region of *B. cenocepacia* has slight differences to those of *B. multivorans* above. It consists of a branched oligosaccharide where the second heptose residue is a 3,7-Hep that carries at O-7 a α-Hep-(1-7)-α-Hep-(1)- disaccharide.[93,94] The outer core region includes galactose, glucose, a further heptose, that is infrequently found as a component of the outer core of LOSs (Fig. 6) and, interestingly, a non-stoichiometric terminal disaccharide composed by quinovosamine and rhamnose.[93] The hypothesis that this latter disaccharide is a remnant of the O-polysaccharide chain cannot be ruled out. Furthermore it could potentially be a first repeating unit of the O-chain, thus giving rise to a semi-rough type LPS.

5.4 *Burkholderia pyrrocinia* genomovar IX

In a clinical *B. pyrrocinia* isolate, a rough chemotype LPS was found. Together with the classical BCC lipid A it possessed a complex oligosaccharide in which the Kdo unit was substituted at O-5 by the

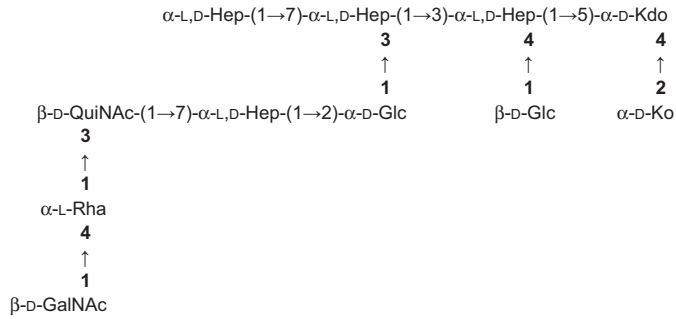


Fig. 6 Core oligosaccharide from *B. cenocepacia*.

pentasaccharide α-Hep-(1-7)-[α-Rha-(1-2)]-α-Hep-(1-3)-[β-Glc-(1-4)]-α-Hep-(1-, a saccharide sequence common to the LOS found in *B. cepacia*.[46] As substantial difference, the last heptose residue of this portion was however further substituted by a heptose trisaccharide, giving rise to a heptan pentasaccharide that is the non reducing part of the outer core.[46]

As for other non BCC core oligosaccharides of *Burkholderia* LPS, the only complete data are deriving from our own studies on *B. caryophylli* LPS.[104] Most strikingly we identified for the first time, that this core region possesses a unique two L-α-D-Hep*p*-(1-5)-α-Kdo-2- moieties, one of which is linked to lipid A with the other linked to Kdo.[104]

6 O-chain of BCC LPS

There exist several O-antigen structures isolated and characterized from BCC strains. It is known that BCC species produce many structurally distinct LPS O-antigens.[79,105] In fact, 16 O-antigen serotypes have been identified among BCC isolates recovered from CF patients.[79,106,107] These serotypes do not correlate with BCC species.[108,109] As a consequence, O-antigen serotype cannot be used to determine the specific BCC species causing an infection. Nevertheless, all the available data show that the O-antigens belonging to BCC are rather simple structures with few sugars constituting the repeating unit, very often containing 6-deoxy sugar (typically L-rhamnose), and, in the case of the trisaccharide repeating unit, with at least one sugar present twice. From the immunological point of view, the presence of rhamnose residues in the O-chain of several *Burkholderia* LPSs has been deemed to play a key role in the bacterial interaction with host-recognition systems, that is, innate plant immunity.[110]

Several studies elucidated the presence of different O-chain polysaccharides within the same BCC strain (*B. multivorans*, *B. vietnamensis*)[79] in which is unusual in Gram-negative bacteria. Furthermore it has been demonstrated that bacteria can adopt a strategy during host invasion in which they undergo ''phase variation'' involving structural variations in antigenic determinants;[111] thus this heterogeneity in O-antigen structure within the same strain could be conceived as a particular kind of phase variation that the bacteria use to augment immunological specificity of the

antigenic portion of its LPS, overall contributing to the variety of antigenic types between BCC species and strains. Analogously, LPS O-acetylation helps *Haemophilus influenza* during infection in the human respiratory tract, facilitating resistance to host immune clearance mechanisms.

As the O-antigen is the most prominent and immunogenic component of the bacterial cell surface, the structure and composition of the O-antigen can be used as the basis of a serotype.

We are aware of few studies determining the relationship between BCC infection in humans and serotype. Rabkin[109] *et al.* isolated BCC from patients with hospital-acquired infections and the environmental source of these infections. The serotype of the strains was determined, but, as mentioned above, no apparent correlation between the serotype and the source or clinical outcome was observed.[109] More recently, the genomovar status of the strains used in the prior study of Rabkin *et al.*[109] was determined: strains from multiple genomovars had the same serotype indicating that genomovar status and serotype did not correlate.[108] There is an obvious need to develop vaccines to prevent BCC infections in susceptible people, particularly those living with CF. As the O-antigen is a strong immunogenic molecule, it may be a potential vaccine candidate for BCC, as has been proposed for other Gram-negative bacteria, including *P. aeruginosa*.[108] However, as there does not appear to be a correlation between serotype and disease state, any O-antigen-based vaccine would need to be multivalent. On the other hand, BCC is being developed for release into the environment for bioremediation or agricultural uses. If there is concern that the commercially useful BCC strains may induce disease in susceptible hosts, a vaccine directed to the O antigen of these strains could be developed and given to high-risk patients.[79]

Here we reported several to date BCC O-antigen structures elucidated.

6.1 *Burkholderia cepacia* genomovar I

The structure of the O-antigen produced by a clinical isolate of *Burkholderia cepacia* isolated from a patient with fibrocystic lung disease has determined as $[\rightarrow 3)\alpha\text{-D-Rha}p\text{-}(1\rightarrow 3)\alpha\text{-D-Rha}p\text{-}(1\rightarrow 4)\text{-}\alpha\text{-D-Gal}p\text{-}(1\rightarrow]_n$.[112] Also, an additional O-antigen from a *B. cepacia* strain varies greatly; the structure was determined to be a disaccharide repeat of a glucopyranose followed by a mannopyranose.[113] The O-antigenic polysaccharides of the lipopolysaccharides produced by three isolates of *Burkholderia cepacia* serogroup A were analysed and showed the repeating trisaccharide unit $[\rightarrow 3)\text{-}\alpha\text{-D-GalpNAc-}(1\rightarrow 3)\text{-}\beta\text{-D-GalpNAc-}(1\rightarrow 4)\text{-}\alpha\text{-L-Rhap-}(1\rightarrow]_n$.[114] Moreover, an isolate of serogroup O7 strain ATCC 17759 has been showed to possess a the disaccharidic unit $[\rightarrow 4)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow 3)\text{-}\alpha\text{-D-Man}p\text{-}(1\rightarrow]$.[115]

6.2 *Burkholderia multivorans* genomovar II

Burkholderia multivorans LPS possesses a mixture of two O-chains with different repeating units: $[\rightarrow 2\text{-}\alpha\text{-D-Man-}(1\rightarrow 2)\text{-}\alpha\text{-D-Rha-}(1\rightarrow 3)\text{-}\alpha\text{-D-Man-}1\rightarrow]_n$ (species X, 60%) and $[\rightarrow 2\text{-}\alpha\text{-D-Man-}(1\rightarrow 2)\text{-}\alpha\text{-D-Aco-}(1\rightarrow 3)\text{-}\alpha\text{-D-Rha}1\rightarrow]_n$ (species Y, 40%).[94] As mentioned above, the two trisaccharide repeating units are both presenting α-*manno*-configured 6-deoxy sugars; in both cases a sugar is repeated twice (α-mannose in the first polysaccharide and α-rhamnose in the second type). Some other important and unusual features

relevant peculiarities can be found in these structures; they both present D-configured rhamnose, which within BCC LPS was only previously found in the O-polysaccharide of *B. cepacia*.[116] Furthermore, in species X the 2 substituted α-Rha is stoichiometrically substituted by a methyl group at O-3 (acofriose). The presence of non-carbohydrate appendages, such as a methyl group, is common to several antigens isolated from BCC strains, often presenting O-acyl decorations (for example, in strain CIP8240 there is a non-stoichiometric acylation at O-6 of α-Galf, in CCUG7246 at O-3 of α-Man, in ATCC 17 759 at O-2 of α-Man) although the O-methyl substitution has been found for the first time to be decorating an O-polysaccharide from *Burkholderia* LPS.

6.3 *Burkholderia cenocepacia* genomovar III

B. cenocepacia strains have been shown to have a [α-D-GalNAc-(1→3)-α-D-GalNAc-(1→4)-α-L-Rha]$_n$ trisaccharide sitting on a quinovosamine residue. Importantly the O polysaccharide has been demonstrated to contribute to immune responses or immune evasion.[117,118] Medicinal chemistry approaches have led to the organic synthesis of the trisaccharide mimicking the repeating unit of an O-polysaccharide portion of LPS present in a clinical isolate of *Burkholderia cepacia*. This may offer a possible polysaccharide based vaccine. However these approaches need to be extended to *B. cenocepacia* and *B. multivorans* as the most clinically prevalent strains in CF.

6.4 *Burkholderia stabilis* genomovar IV

Burkholderia stabilis strain CIP 8235 have two different O-chain structures characterised by a disaccharidic repeating unit: a major component is [→4)-α-D-Glc*p*-(1→3)-α-L-Glc*p*NAc-(1→] and a minor component is [→4)-α-D-Glc*p*-(1→3)-α-L-Rha*p*-(1→].[119,120]

6.5 *Burkholderia vietnamiensis* genomovar V

Burkholderia vietnamiensis strain LMG 6998 possesses the linear trisaccharide repeating-unit [→3)-α-L-Rha*p*-(1→3)-β-D-Man*p*-(1→4)-α-D-Man*p*-(1→]$_n$[76] where the 4-substituted mannose residue carries an O-acetyl group at position 3. Whilst lipopolisaccharide isolated from strain LMG 6999 of *Burkholderia vietnamiensis* is characterised by the presence of two polymeric fractions based on the following trisaccharide repeating units: I: [→3)-α-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→]$_n$; II: [→3)-α-D-Gal*p*NAc-(1→3)-β-D-Gal*p*NAc-(1→4)-α-L-Rha*p*-(1→]$_n$.[121]

6.6 *Burkholderia anthina* genomovar VIII

The structure of the O-antigen from a strain of *B. anthina* is [→3)α-L-Rha*p*-(1→2)α-L-Rha*p*-(1→2)-α-D-Gal*p*-(1→]*n*.[122] This structure is similar to an O-antigen that was identified in a *B. cepacia* clinical isolate: [→3)α-D-Rha*p*-(1→3)-α-D-Rha*p*-(1→4)-α-D-Gal*p*-(1→]$_n$.

Searching for a rationale in the aforementioned heterogeneity of the O-antigen subunit structures from BCC, there are three things to notice. First, many of the O-antigen repeating unit are simple structures, having only two or three sugar residues per subunit. Those subunits with three sugars have at

least one sugar residue repeated: for example, the O antigen of the serotype K strain, IMV 3181, is a polyrhamnose. Second, there are some unusual sugar residues found in the O-antigen structures from BCC; for example, L,D-heptose is usually found only in the core oligosaccharide of other Gram-negative bacteria, and the D forms of rhamnose and fucose are found in O-antigen subunits of BCC and other bacteria of the Pseudomonadaceae family as opposed to the L forms of these sugars.[121,119,123] Third, there are two O-antigen subunit structures of BCC that are in common with *P. aeruginosa* O-antigen subunits: the O-antigen subunits of BCC serotype O3 and O5 are identical to the O-antigen subunits of *P. aeruginosa* serogroup O15 and O17, respectively.[124]

7 Conclusion

Cystic Fibrosis provides a unique clinical challenge. The complex environment within the lung and the changing milieu is associated with strong local bacterial environmental pressures that subsequently alter many aspects of the bacteria virulence. The inflammatory processes within the lung can be traced at least in part to the biological activities of the lipopolysaccharide moieties. Specific changes in LPS occur in two key pathogen groups *Pseudomonas aeruginosa* and *Burkholderia* spp. To date these changes seem specific to CF and have not been seen in other human infections. ''CF specific'' changes in LPS appear to occur relatively rapidly after these organisms are acquired in CF and are maintained *in vitro*. These confer an advantage for persistence of the bacteria but the mechanisms involved may also offer opportunities for manipulation therapeutically. Novel data arising from lung transplantation strains suggest that environmental changes *in vivo* are associated with further changes in LPS and lipid A from the chronic CF type LPS. The potential to ''revert'' CF specific changes in LPS composition is interesting though the consequences need a comprehensive analysis. Unfortunately the lack of good animal models and the host specific interaction between LPS and host TLR demonstrated for *Pseudomonas* makes such studies problematic. The use of clinical strains in research from both early acquisition and later stages of CF lung disease is mandatory to reflect the spectrum of disease. Environmental *Pseudomonas* strains have different virulence and LPS characteristics based on the available data. Hence studies for vaccine development or therapeutic manipulation of LPS biosynthesis must avoid the over-representation of environmental strains as compared to clinically derived strains. Concerted efforts across traditional scientific boundaries of chemistry, microbiology, immunology and medicine will be required to make major advances in our understanding of these areas. Notably several collaborative have formed to this end *e.g.* http://www.cost.eu/domains_actions/bmbs/Actions/BM1003. Future studies should carefully consider the growth conditions used *in vitro* and where possible consideration of testing more than one strain from a patient may yield more generalisable data. The ability to access a standardised panel of CF *Pseudomonas* strains such as is available for the international *Burkholderia* reference panel may well allow more rapid advances to occur.

Acknowledegments

We all acknowledge Cost Action BM1003 ''Microbial cell surface determinants of virulence as targets for new therapeutics in Cystic Fibrosis''. ADS is a Higher Education Funding Council for England (HEFCE) Senior lecturer. We acknowledge funding from the Newcastle Hospitals Special Trustees (ADS), The Italian Cystic Fibrosis Foundation (FFC; grant 09/16 (ADS and AS), FFC 11#2010 with the contribution of Pastificio Giovanni Rana s.p.a. to AM). A.M, A.S.,and R.L acknowledge PRIN MIUR 2008/2009 (A.M. and R.L.).

References

1 F. Ratjen and H. Grasemann, *Current Pharmaceutical Design*, 2012.
2 C. S. Halliburton, D. M. Mannino and R. S. Olney, *Archives of Pediatrics & Adolescent Medicine*, 1996, **150**, 1181–1185.
3 T. G. Liou, F. R. Adler, S. C. Fitzsimmons, B. C. Cahill, J. R. Hibbs and B. C. Marshall, *American Journal of Epidemiology*, 2001, **153**, 345–352.
4 M. I. Gomez and A. Prince, *Curr. Opin. Pharmacol.*, 2007, **7**, 244–251.
5 E. Mahenthiralingam, T. A. Urban and J. B. Goldberg, *Nat. Rev. Microbiol.*, 2005, **3**, 144–156.
6 T. Coenye, J. Goris, T. Spilker, P. Vandamme and J. J. LiPuma, *Journal of Clinical Microbiology*, 2002, **40**, 2062–2069.
7 J. S. Brooke, *Clinical Microbiology Reviews*, 2012, **25**, 2–41.
8 J. Emerson, S. McNamara, A. M. Buccat, K. Worrell and J. L. Burns, *Pediatric Pulmonology*, **45**, 363–370.
9 L. Spicuzza, C. Sciuto, G. Vitaliti, G. Di Dio, S. Leonardi and M. La Rosa, *Eur. J. Clin. Microbiol. Infect. Dis.*, 2009, **28**, 191–195.
10 C. Ronne Hansen, T. Pressler, N. Hoiby and M. Gormsen, *J. Cyst. Fibros.*, 2006, **5**, 245–251.
11 J. R. Govan and V. Deretic, *Microbiological Reviews*, 1996, **60**, 539–574.
12 S. M. Bianchi, L. R. Prince, K. McPhillips, L. Allen, H. M. Marriott, G. W. Taylor, P. G. Hellewell, I. Sabroe, D. H. Dockrell, P. W. Henson and M. K. Whyte, *American Journal of Respiratory and Critical Care Medicine*, 2008, **177**, 35–43.
13 M. Rosenfeld, R. L. Gibson, S. McNamara, J. Emerson, J. L. Burns, R. Castile, P. Hiatt, K. McCoy, C. B. Wilson, A. Inglis, A. Smith, T. R. Martin and B. W. Ramsey, *Pediatric Pulmonology*, 2001, **32**, 356–366.
14 E. Kerem, M. Corey, R. Gold and H. Levison, *Journal of Pediatrics*, 1990, **116**, 714–719.
15 N. Lechtzin, M. John, R. Irizarry, C. Merlo, G. B. Diette and M. P. Boyle, *Respiration; International Review of Thoracic Diseases*, 2006, **73**, 27–33.
16 N. J. Simmonds, S. J. Macneill, P. Cullinan and M. E. Hodson, *Eur. Respir. J.*, 2011, **36**, 1277–1283.
17 K. Cheng, R. L. Smyth, J. R. Govan, C. Doherty, C. Winstanley, N. Denning, D. P. Heaf, H. van Saene and C. A. Hart, *Lancet*, 1996, **348**, 639–642.
18 A. Ashish, M. Shaw, C. Winstanley, M. J. Ledson and M. J. Walshaw, *J. Cyst. Fibros.*
19 M. Al-Aloul, J. Crawley, C. Winstanley, C. A. Hart, M. J. Ledson and M. J. Walshaw, *Thorax*, 2004, **59**, 334–336.
20 E. Mahenthiralingam, B. Bischof, S. K. Byrne, C. Radomski, J. E. Davies, Y. Av-Gay and P. Vandamme, *Journal of Clinical Microbiology*, 2000, **38**, 3165–3173.
21 L. Saiman, *Current Opinion in Infectious Diseases*, 2012, **24**, 390–395.

22 L. Eberl and B. Tummler, *Int. J. Med. Microbiol.*, 2004, **294**, 123–131.
23 A. R. Hauser, M. Jain, M. Bar-Meir and S. A. McColley, *Clin. Microbiol. Rev.*, 2011, **24**, 29–70.
24 G. Doring, I. G. Parameswaran and T. F. Murphy, *FEMS Microbiology Reviews*.
25 E. E. Smith, D. G. Buckley, Z. Wu, C. Saenphimmachak, L. R. Hoffman, D. A. D'Argenio, S. I. Miller, B. W. Ramsey, D. P. Speert, S. M. Moskowitz, J. L. Burns, R. Kaul and M. V. Olson, *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103**, 8487–8492.
26 D. J. Waine, D. Honeybourne, E. G. Smith, J. L. Whitehouse and C. G. Dowson, *Journal of Clinical Microbiology*, 2008, **46**, 3491–3493.
27 M. T. Rybtke, P. O. Jensen, N. Hoiby, M. Givskov, T. Tolker-Nielsen and T. Bjarnsholt, *Inflammation & Allergy Drug Targets*, **10**, 141–157.
28 N. Hoiby, O. Ciofu and T. Bjarnsholt, *Future Microbiology*, **5**, 1663–1674.
29 S. Elias and E. Banin, *FEMS Microbiology Reviews*, 2012.
30 M. J. Hunter, K. J. Treharne, A. K. Winter, D. M. Cassidy, S. Land and A. Mehta, *PloS One*, 2010, **5**, e11598.
31 L. S. Nixon, B. Yung, S. C. Bell, J. S. Elborn and D. J. Shale, *American Journal of Respiratory & Critical Care Medicine*, 1998, **157**, 1764–1769.
32 P. S. Salva, N. A. Doyle, L. Graham, H. Eigen and C. M. Doerschuk, *Pediatric Pulmonology*, 1996, **21**, 11–19.
33 E. Raoust, V. Balloy, I. Garcia-Verdugo, L. Touqui, R. Ramphal and M. Chignard, *PloS One*, 2009, **4**, e7259.
34 S. M. Rowe, P. L. Jackson, G. Liu, M. Hardison, A. Livraghi, G. M. Solomon, D. B. McQuaid, B. D. Noerager, A. Gaggar, J. P. Clancy, W. O'Neal, E. J. Sorscher, E. Abraham and J. E. Blalock, *American Journal of Respiratory and Critical Care Medicine*, 2008, **178**, 822–831.
35 M. Brodlie, M. C. McKean, G. E. Johnson, J. Gray, A. J. Fisher, P. A. Corris, J. L. Lordan and C. Ward, *Am. J. Respir. Crit. Care Med.*, 2011, **182**, 369–375.
36 N. Regamey, L. Tsartsali, T. N. Hilliard, O. Fuchs, H. L. Tan, J. Zhu, Y. S. Qiu, E. W. Alton, P. K. Jeffery, A. Bush and J. C. Davies, *Thorax*, 2012, **67**, 170–176.
37 S. Voglis, K. Quinn, E. Tullis, M. Liu, M. Henriques, C. Zubrinich, O. Penuelas, H. Chan, F. Silverman, V. Cherepanov, N. Orzech, A. A. Khine, A. Cantin, A. S. Slutsky, G. P. Downey and H. Zhang, *American Journal of Respiratory and Critical Care Medicine*, 2009, **180**, 159–166.
38 D. Hartl, P. Latzin, P. Hordijk, V. Marcos, C. Rudolph, M. Woischnik, S. Krauss-Etschmann, B. Koller, D. Reinhardt, A. A. Roscher, D. Roos and M. Griese, *Nature Medicine*, 2007, **13**, 1423–1430.
39 G. B. Rogers, C. A. Hart, J. R. Mason, M. Hughes, M. J. Walshaw and K. D. Bruce, *Journal of Clinical Microbiology*, 2003, **41**, 3548–3558.
40 G. B. Rogers, M. P. Carroll, D. J. Serisier, P. M. Hockey, G. Jones, V. Kehagia, G. J. Connett and K. D. Bruce, *Journal of Clinical Microbiology*, 2006, **44**, 2601–2604.
41 M. S. Doud, M. Light, G. Gonzalez, G. Narasimhan and K. Mathee, *Human Genomics*, 2010, **4**, 147–169.
42 M. M. Tunney, T. R. Field, T. F. Moriarty, S. Patrick, G. Doering, M. S. Muhlebach, M. C. Wolfgang, R. Boucher, D. F. Gilpin, A. McDowell and J. S. Elborn, *American Journal of Respiratory and Critical Care Medicine*, 2008, **177**, 995–1001.
43 T. R. Field, A. White, J. S. Elborn and M. M. Tunney, *Eur. J. Clin. Microbiol. Infect. Dis.*, 2005, **24**, 677–687.
44 N. Bagge, O. Ciofu, L. T. Skovgaard and N. Hoiby, *Apmis*, 2000, **108**, 589–600.

45 A. Silipo, G. Erbs, T. Shinya, J. M. Dow, M. Parrilli, R. Lanzetta, N. Shibuya, M. A. Newman and A. Molinaro, *Glycobiology*, 2010, **20**, 406–419.
46 A. De Soyza, A. Silipo, R. Lanzetta, J. R. Govan and A. Molinaro, *Innate Immunity*, 2008, **14**, 127–144.
47 C. R. Raetz and C. Whitfield, *Annu. Rev. Biochem.*, 2002, **71**, 635–700.
48 N. L. Que, S. Lin, R. J. Cotter and C. R. Raetz, *The Journal of Biological Chemistry*, 2000, **275**, 28006–28016.
49 T. Ierano, A. Silipo, L. Sturiale, D. Garozzo, H. Brookes, C. M. Khan, C. Bryant, F. K. Gould, P. A. Corris, R. Lanzetta, M. Parrilli, A. De Soyza and A. Molinaro, *Glycobiology*, 2008, **18**, 871–881.
50 S. Sforza, A. Silipo, A. Molinaro, R. Marchelli, M. Parrilli and R. Lanzetta, *J. Mass Spectrom.*, 2004, **39**, 378–383.
51 L. Sturiale, D. Garozzo, A. Silipo, R. Lanzetta, M. Parrilli and A. Molinaro, *Rapid Commun. Mass Spectrom.*, 2005, **19**, 1829–34.
52 S. Hase and E.T. Rietschel, *Eur. J. Biochem.*, 1976, **63**, 101–7.
53 Y. A. Knirel, J. H. Helbig and U. Zähringer, *Carbohydr. Res.*, 1996, **283**, 129–39.
54 L. Brecker, *Chem. Phys. Lipids*, 2003, **125**, 27–39.
55 A. Silipo, C. De Castro, R. Lanzetta, A. Molinaro and M. Parrilli, *Glycobiology*, 2004, **14**, 805–15.
56 Y. A. Knirel, N. A. Kocharova, S. O. Hynes, G. Widmalm, L. P. Andersen, P. E. Jansson and A. P. Moran, *European Journal of Biochemistry/FEB*, 1999, **266**, 123–131.
57 O. Holst, *Methods Mol. Biol.*, 2000, **145**, 345–53.
58 A. Silipo, R. Lanzetta, M. Parrilli, L. Sturiale, D. Garozzo, E. L. Nazarenko, R. P. Gorshkova, E. P. Ivanova and A. Molinaro, *Carbohydrate Research*, 2005, **340**, 1475–1482.
59 P. Rallabhandi, A. Awomoyi, K. E. Thomas, A. Phalipon, Y. Fujimoto, K. Fukase, S. Kusumoto, N. Qureshi, M. B. Sztein and S. N. Vogel, *J. Immunol.*, 2008, **180**, 1139–1147.
60 S. M. Moskowitz and R. K. Ernst, *Sub-cellular Biochemistry*, 2010, **53**, 241–253.
61 J. S. Lam, V. L. Taylor, S. T. Islam, Y. Hao and D. Kocincova, *Frontiers in Microbiology*, 2011, **2**, 118.
62 R. K. Ernst, A. M. Hajjar, J. H. Tsai, S. M. Moskowitz, C. B. Wilson and S. I. Miller, *Journal of Endotoxin Research*, 2003, **9**, 395–400.
63 O. V. Bystrova, A. S. Shashkov, N. A. Kocharova, Y. A. Knirel, U. Zahringer and G. B. Pier, *Biochemistry. Biokhimiia*, 2003, **68**, 918–925.
64 F. Beckmann, H. Moll, K. E. Jäger and U. Zähringer, *Carbohydr. Res.*, 1995, **267**, C3–7.
65 Y. A. Knirel, J. H. Helbig and U. Zähringer, *Carbohydr. Res.*, 1996, **283**, 129–39.
66 O. V. Bystrova, A. S. Shashkov, N. A. Kocharova, Y. A. Knirel, B. Lindner, U. Zahringer and G. B. Pier, *European Journal of Biochemistry/FEB*, 2002, **269**, 2194–2203.
67 Y. A. Knirel, O. V. Bystrova, A. S. Shashkov, B. Lindner, N. A. Kocharova, S. N. Senchenkova, H. Moll, U. Zahringer, K. Hatano and G. B. Pier, *European Journal of Biochemistry/FEBS*, 2001, **268**, 4708–4719.
68 A. G. Walsh, M. J. Matewish, L. L. Burrows, M.A. Monteiro, M. B. Perry and J. S. Lam, *Mol. Microbiol.*, 2000, **35**, 718–27.
69 O. V. Bystrova, Y. A. Knirel, B. Lindner, N. A. Kocharova, A. N. Kondakova, U. Zahringer and G. B. Pier, *FEMS Immunology and Medical Microbiology*, 2006, **46**, 85–99.
70 D. J. Evans, G. B. Pier, M. J. Coyne and J. B. Goldberg, *Mol. Microbiol.*, 1994, **13**, 427–34.
71 Y. A. Knirel, *Critical Reviews in Microbiology*, 1990, **17**, 273–304.

72 S. I. Miller, R. K. Ernst and M. W. Bader, *Nat. Rev. Microbiol.*, 2005, **3**, 36–46.
73 M. T. Albrecht, W. Wang, O. Shamova, R. I. Lehrer and N. L. Schiller, *Respir. Res.*, 2002, **3**, 18.
74 J. B. McPhee, S. Lewenza and R. E. Hancock, *Molecular Microbiology*, 2003, **50**, 205–217.
75 J. B. McPhee, M. Bains, G. Winsor, S. Lewenza, A. Kwasnicka, M. D. Brazas, F. S. Brinkman and R. E. Hancock, *Journal of Bacteriology*, 2006, **188**, 3995–4006.
76 E. L. Macfarlane, A. Kwasnicka, M. M. Ochs and R. E. Hancock, *Molecular Microbiology*, 1999, **34**, 305–316.
77 C. Cigana, L. Curcuru, M. R. Leone, T. Ierano, N. I. Lore, I. Bianconi, A. Silipo, F. Cozzolino, R. Lanzetta, A. Molinaro, M. L. Bernardini and A. Bragonzi, *PloS one*, 2009, **4**, e8439.
78 A. M. Hajjar, R. K. Ernst, J. H. Tsai, C. B. Wilson and S. I. Miller, *Nature Immunology*, 2002, **3**, 354–359.
79 A. D. Vinion-Dubiel and J. B. Goldberg, *Journal of Endotoxin Research*, 2003, **9**, 201–213.
80 D. P. Speert, *Paediatric Respiratory Reviews*, 2002, **3**, 230–235.
81 A. Agodi, E. Mahenthiralingam, M. Barchitta, V. Giannino, A. Sciacca and S. Stefani, *Journal of Clinical Microbiology*, 2001, **39**, 2891–2896.
82 J. M. van den Berg, E. van Koppen, A. Ahlin, B. H. Belohradsky, E. Bernatowska, L. Corbeel, T. Espanol, A. Fischer, M. Kurenko-Deptuch, R. Mouy, T. Petropoulou, J. Roesler, R. Seger, M. J. Stasia, N. H. Valerius, R. S. Weening, B. Wolach, D. Roos and T. W. Kuijpers, *PloS one*, 2009, **4**, e5234.
83 A. De Soyza, K. Morris, A. McDowell, C. Doherty, L. Archer, J. Perry, J. R. Govan, P. A. Corris and K. Gould, *Thorax*, 2004, **59**, 526–528.
84 A. De Soyza and P. A. Corris, *J. Heart Lung Transplant.*, 2003, **22**, 954–958.
85 D. Hadjiliadis, M. P. Steele, C. Chaparro, L. G. Singer, T. K. Waddell, M. A. Hutcheon, R. D. Davis, D. E. Tullis, S. M. Palmer and S. Keshavjee, *J. Heart Lung Transplant.*, 2007, **26**, 834–838.
86 B. D. Alexander, E. W. Petzold, L. B. Reller, S. M. Palmer, R. D. Davis, C. W. Woods and J. J. Lipuma, *Am. J. Transplant.*, 2008, **8**, 1025–1030.
87 P. M. Hopkins, T. J. Kidd, C. Coulter, I. H. Feather, P. Derrington and S. C. Bell, *Am. J. Respir. Crit. Care Med.*, 2009, **179**, 257–258.
88 S. Murray, J. Charbeneau, B. C. Marshall and J. J. LiPuma, *Am. J. Respir. Crit. Care Med.*, 2008, **178**, 363–371.
89 A. De Soyza, G. Meachery, K. L. Hester, A. Nicholson, G. Parry, K. Tocewicz, T. Pillay, S. Clark, J. L. Lordan, S. Schueler, A. J. Fisher, J. H. Dark, F. K. Gould and P. A. Corris, *J. Heart Lung Transplant.*, 2010.
90 E. Mahenthiralingam, T. Coenye, J. W. Chung, D. P. Speert, J. R. Govan, P. Taylor and P. Vandamme, *Journal of Clinical Microbiology*, 2000, **38**, 910–913.
91 A. De Soyza, C. D. Ellis, C. M. Khan, P. A. Corris and R. Demarco de Hormaeche, *American Journal of Respiratory and Critical Care Medicine*, 2004, **170**, 70–77.
92 A. Silipo, A. Molinaro, T. Ierano, A. De Soyza, L. Sturiale, D. Garozzo, C. Aldridge, P. A. Corris, C. M. Khan, R. Lanzetta and M. Parrilli, *Chemistry*, 2007, **13**, 3501–3511.
93 A. De Soyza, C. D. Ellis, C. M. A. Anjam, P. A. Corris and R. DeMarco De Hormaeche, *American Journal of Respiratory & Critical Care Medicine*, 2004, **170**, 70–77.
94 T. Ierano, A. Silipo, L. Sturiale, D. Garozzo, C. Bryant, R. Lanzetta, M. Parrilli, C. Aldridge, F. K. Gould, P. A. Corris, C. M. Khan, A. De Soyza and A. Molinaro, *Glycobiology*, 2009, **19**, 1214–1223.

95 H. Iihara, T. Niwa, M. M. Shah, P. H. Nhung, S. X. Song, M. Hayashi, K. Ohkusa, Y. Itoh, S. Makino and T. Ezaki, *Jpn. J. Infect Dis.*, 2007, **60**, 230–4.
96 A. Silipo, A. Molinaro, P. Cescutti, E. Bedini, R. Rizzo, M. Parrilli and R. Lanzetta, *Glycobiology*, 2005, **15**, 561–570.
97 A. Molinaro, C. De Castro, R. Lanzetta, A. Evidente, M. Parrilli and O. Holst, *The Journal of Biological Chemistry*, 2002, **277**, 10058–10063.
98 H. Shimomura, M. Matsuura, S. Saito, Y. Hirai, Y. Isshiki and K. Kawahara, *Infection and Immunity*, 2003, **71**, 5225–5230.
99 G. Erbs, T. T. Jensen, A. Silipo, W. Grant, J. M. Dow, A. Molinaro, M. Parrilli and M. A. Newman, *Microbes and Infection/Institut Pasteur*, 2008, **10**, 571–574.
100 R. A. Reife, S. R. Coats, M. Al-Qutub, D. M. Dixon, P. A. Braham, R. J. Billharz, W. N. Howald and R. P. Darveau, *Cellular Microbiology*, 2006, **8**, 857–868.
101 S. A. Loutet, R. S. Flannagan, C. Kooi, P. A. Sokol and M. A. Valvano, *Journal of Bacteriology*, 2006, **188**, 2073–2080.
102 S. A. Loutet, F. Di Lorenzo, C. Clarke, A. Molinaro and M. A. Valvano, *BMC genomics*, 2011, **12**, 472.
103 Y. Isshiki, U. Zahringer and K. Kawahara, *Carbohydrate Research*, 2003, **338**, 2659–2666.
104 C. De Castro, A. Molinaro, R. Lanzetta, O. Holst and M. Parrilli, *Carbohydrate Research*, 2005, **340**, 1802–1807.
105 M. N. Burtnick, P. J. Brett and D. E. Woods, *Journal of Bacteriology*, 2002, **184**, 849–852.
106 Y. Nakamura, S. Hyodo, E. Chonan, S. Shigeta and E. Yabuuchi, *Journal of Clinical Microbiology*, 1986, **24**, 152–154.
107 A. Soldatkina, A. Knirel' Iu, N. V. Tanatar and I. Zakharova, *Mikrobiologicheskii zhurnal*, 1989, **51**, 32–38.
108 D. T. Kenna, V. A. Barcus, R. J. Langley, P. Vandamme and J. R. Govan, *FEMS Immunology Medical Microbiology*, 2003, **35**, 87–92.
109 C. S. Rabkin, W. R. Jarvis, R. L. Anderson, J. Govan, J. Klinger, J. LiPuma, W. J. Martone, H. Monteil, C. Richard and S. Shigeta, *et al.*, *Rev. Infect. Dis.*, 1989, **11**, 600–607.
110 K. A. Mattos, A. R. Todeschini, N. Heise, C. Jones, J. O. Previato and L. Mendonca-Previato, *Glycobiology*, 2005, **15**, 313–321.
111 M. Lukáčová, I. Barák and J. Kazár, *Clin. Microbiol. Infect.*, 2008, **14**, 200–6.
112 S. Cerantola, A. Lemassu-Jacquier and H. Montrozier, *European Journal of Biochemistry*, 1999, **260**, 373–383.
113 A. D. Cox and S. G. Wilkinson, *Carbohydrate research*, 1990, **195**, 295–301.
114 L. M. Beynon and M. B. Perry, *Biochemistry and Cell Biology = Bichimie et biologie Cellulaire*, 1993, **71**, 417–420.
115 A. Knirel' Iu, N. V. Tanatar, M. A Soldatkina, A. S. Shashkov and I. Zakharova, *Bioorganicheskaia Khimiia*, 1988, **14**, 77–81.
116 S. Cerantola, A. Lemassu-Jacquier and H. Montrozier, *European Journal of Biochemistry/FEB*, 1999, **260**, 373–383.
117 S. Kotrange, B. Kopp, A. Akhter, D. Abdelaziz, A. Abu Khweek, K. Caution, B. Abdulrahman, M. D. Wewers, K. McCoy, C. Marsh, S. A. Loutet, X. Ortega, M. A. Valvano and A. O. Amer, *Journal of Leukocyte Biology*, 2011, **89**, 481–488.
118 M. S. Saldias, X. Ortega and M. A. Valvano, *Journal of Medical Microbiology*, 2009, **58**, 1542–1548.
119 N. A. Paramonov, A. S. Shashkov, A. Knirel Iu, M. A Soldatkina and I. Zakharova, *Bioorganicheskaia Khimiia*, 1994, **20**, 984–993.

120 A. D. Cox, C. J. Taylor, A. J. Anderson, M. B. Perry and S. G. Wilkinson, *European Journal of Biochemistry/FEB*, 1995, **231**, 784–789.
121 D. Gaur and S. G. Wilkinson, *Carbohydrate Research*, 1996, **295**, 179–184.
122 S. Carillo, A. Silipo, V. Perino, R. Lanzetta, M. Parrilli and A. Molinaro, *Carbohydrate Research*, 2009, **344**, 1697–1700.
123 A. Knirel' Iu, M. A. Soldatkina, A. S. Shashkov, N. V. Tanamar and N. A. Paramonov, *Bioorganicheskaia Khimiia*, 1988, **14**, 1678–1683.
124 A. D. Cox and S. G. Wilkinson, *Carbohydrate Research*, 1989, **195**, 123–129.

Synthesis of bacterial carbohydrate surface structures containing Kdo and *glycero*-D-*manno*-heptose linkages

Stefan Oscarson
DOI: 10.1039/9781849734769-00040

1 Introduction

The inherent complexity of oligosaccharides, due to the large amount of hydroxyl groups and the stereochemistry of the glycosidic linkage, is a major drawback in their synthesis. Compared to human structures microbial ones are even more challenging because of a vast variety of possible monosaccharide building blocks as well as substituents. In 2000 we wrote a review on the synthesis of bacterial structures containing two major synthetic challenges, inner core LPS-structures with the unusual sugars Kdo and L-*glycero*-D-*manno*-heptose and CPS structures built up from repeating units linked *via* phosphodiester linkages.[1] Herein an update on the progress in the synthesis of Kdo and *glycero*-D-*manno*-heptose containing structures is presented. Recently this field has been reviewed also by Kosma.[2,3]

2 Kdo-containing structures and Kdo donors

In spite of the biological importance of Kdo-containing oligosaccharide structures, the synthetic interest in Kdo is to a large extent focused on the stereoselective synthesis of the monosaccharide and not in its subsequent use in oligosaccharide synthesis. Thus, annually a number of synthesis of Kdo are published, which have been reviewed.[4,5] However, most of these, although conceptually elegant, are multi-step syntheses. For a quick and reproducible supply of large quantities of Kdo, needed for continued oligosaccharide synthesis, the one-step Cornforth procedure,[6] originally developed for synthesis of sialic acid,[7] is still most often the method of choice. Starting from commercial D-arabinose and oxaloacetic acid, a classical condensation followed by decarboxylation affords Kdo, which usually is isolated as its crystalline acetylated methyl ester.

Apart from not being commercially available there are other problems associated with Kdo in oligosaccharide synthesis, the easy formation of the α,β-unsaturated acid derivative, *i.e.*, a glycal derivative, and the lack of a stereo-directing group in the 3-position. The former is a problem both in protective group manipulation and in glycosylation reactions, since the glycal is formed both under basic and acidic conditions. These problems (and structural features) are similar to the ones encountered with neuraminic acid. In spite of the similarities it is still difficult to transfer experiences between the two sugars, mainly because of the difference in ring conformation, sialic acid commonly being in the $^1C_4$-conformation with

Centre for Synthesis and Chemical Biology, University College Dublin, Belfield, Dublin 4, Ireland. E-mail: stefan.oscarson@ucd.ie

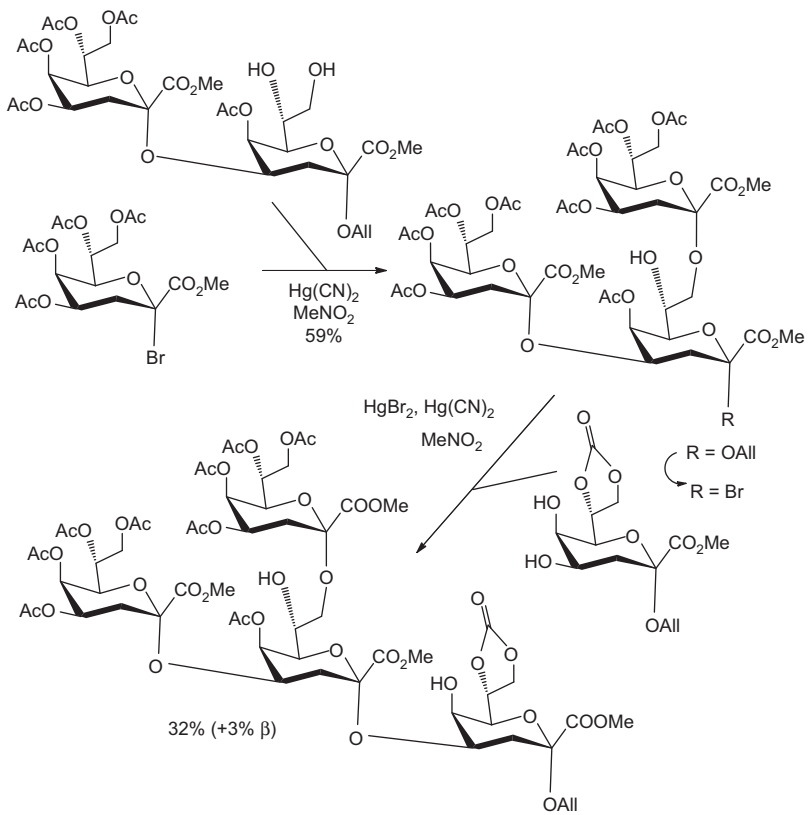
© The Royal Society of Chemistry 2012

an equatorial glycosidic α-linkage, whereas Kdo occurs in Nature in the $^4C_1$-conformation with an axial glycosidic α-linkage.

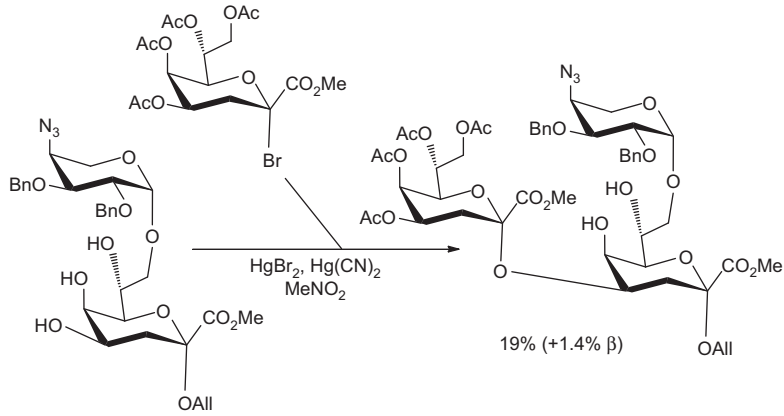
In the earlier review, a survey on the synthesis of Kdo-containing saccharides up to 1999 was made.[1] Several complex structures had been produced especially by the Paulsen group from Hamburg, but sometimes the glycosylation yields were rather low as well as the stereoselectivity. That far mainly halide donors (especially the bromide) had been used in Kdo glycosylation reactions.

Kosma and co-workers have continued their synthesis of *Chlamydia* structures and analogues thereof. Hence the branched trisaccharide motif α-Kdo-(2→4)-[α-Kdo-(2→8)-]-α-Kdo of *Chlamydophila psittaci* was synthesized using acetobromo-Kdo as donor, mercury salts as promoters, and nitromethane as solvent.[8,9] The trisaccharide was synthesized starting either with the (2→4)-linkage[8] or the (2→8)-linkage.[9] The Kdo-trisaccharide was also converted into a bromide donor and used in a coupling to give a (2→4)-linked Kdo tetrasaccharide derivative (Scheme 1).[8] Especially in the latter glycosylation glycal formation was a major side reaction (66%), however, α-stereoselectivity was very high throughout the syntheses.

The Kosma group also, very recently, published the synthesis of a α-Kdo-(2→4)-[β-L-Ara4N-(1→8)-]-α-Kdo structure from *Burkholderia* and



Scheme 1



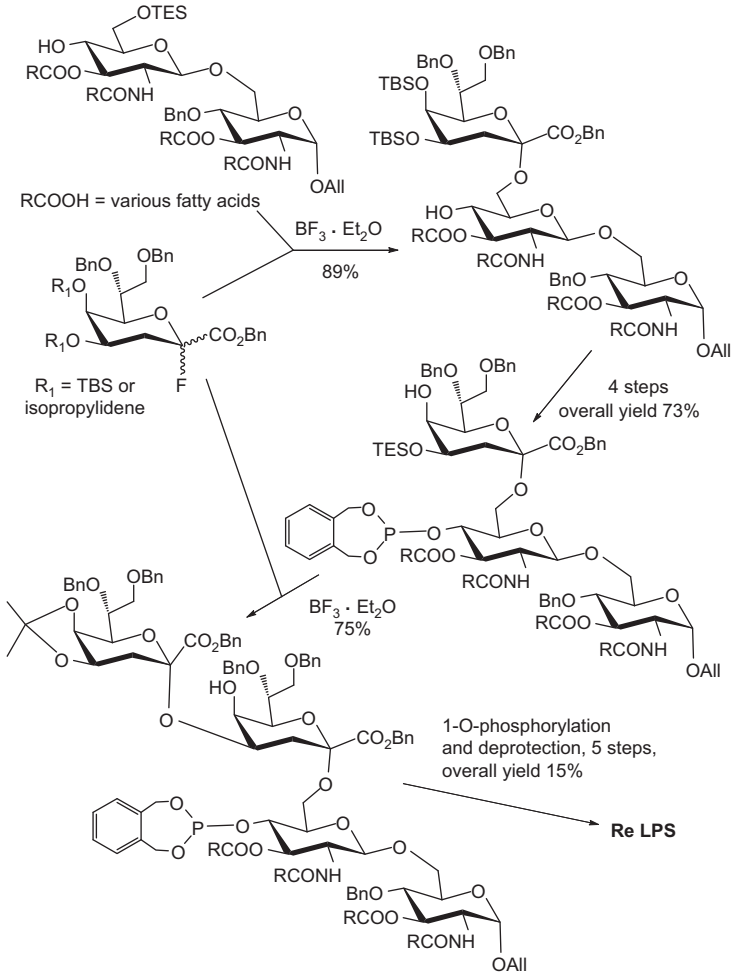
Scheme 2

Proteus LPS.[10] Here, the branching (2→4)-linked Kdo moiety was introduced using a Kdo 4,5,7-triol acceptor. As expected the 5-OH was not involved in glycosylation, but the regioselectivity between the 4- and 7-OH was a problem. In CH_2Cl_2 the (2→7)-linked trisaccharide was dominating over the desired (2→4)-linked product, but in nitromethane this selectivity was reversed. However, only a 19% yield of the α-(2→4)-linked trisaccharide was obtained (together with 1.4% of the β-(2→4)-linked and 5.5% of the α-(2→7)-linked trisaccharide) (Scheme 2).

Kusumoto and coworkers continued their use of isopropylidene acetal protected Kdo fluoride donors as a solution to the stereoselectivity problem, the rationale being that the protecting groups force the ring into a boat-like conformation, which increase the amount of α-side attack and thus formation of the α-glycoside.[11,12] In an elegant synthesis of the complex complete Re-type LPS structure, the authors showed that a 4,5-*O*-isopropylidene group is the important factor, and also that TBDMS protecting groups in the 4,5-positions gives the same effect, probably due to steric crowding causing similar conformational changes in the ring (Scheme 3).[13]

Thus, by employing these fluoro sugars as donors in BF_3-etherate-promoted couplings with Lipid A type disaccharide acceptor, α-(2→6)-linked trisaccharides were obtained in high yields and with high α-selectivity. Subsequent coupling to a Kdo 4,5-diol acceptor using the same donor(s) afforded the α-(2→4)-linked tetrasaccharide. The high yields were accomplished by using a large excess (4 equivalents) of donor to compensate for the high incidence of glycal formation. However, activation of the acceptors as their triethylsilyl derivatives decreased the formation of glycal considerably and consequently afforded increased yields of glycosides, although the excess of donor were brought down to 2 equivalents.

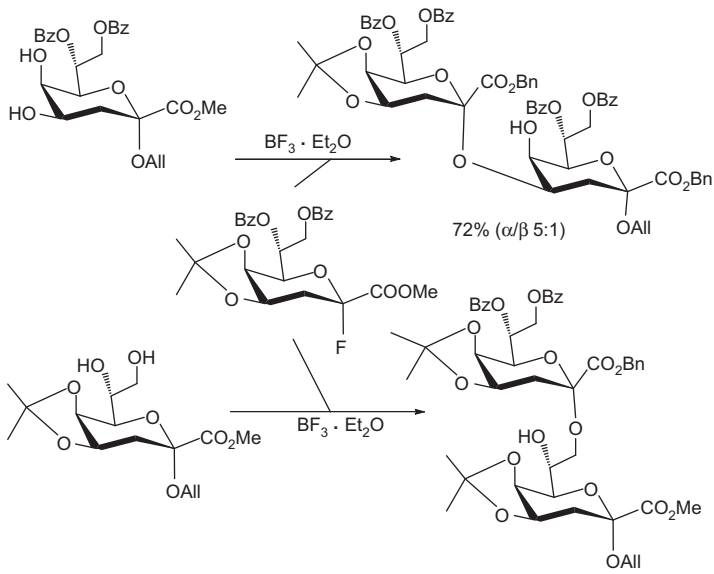
Ichiyanagi *et al.* reported in 2011 on a new Kdo donor, but the only novelty as compared to the derivatives used by Kusumoto and co-workers discussed above was that benzoyl instead of benzyl protecting groups were used in the 7,8-positions.[14] Using (non-silylated) 4,5- or 7,8-diol Kdo acceptors both α-Kdo-(2→4)-α-Kdo and the α-Kdo-(2→8)-α-Kdo disaccharides were produced using BF_3-etherate as promoter, however, the



Scheme 3

stereoselectivity in the former coupling was only α/β 5:1 and the yield in the latter was reported as ''unsatisfactory'' (Scheme 4).

Surprisingly, there was only one report before 2000 on the use of thioglycosides as Kdo glycosyl donors, then with NIS/TfOH as promoter.[15] In contrast, for sialylation reactions, thioglycosides have been used extensively,[16] *e.g.* *O*-xanthates promoted by DMTST[17] or phenylsulfenyl triflate[18] and ethyl thioglycosides promoted by DMTST[19] or IBr.[20] Mannerstedt, Ekelöf and Oscarson reported on an investigation of Kdo-thioglycosides as glycosyl donors testing DMTST, IBr/AgOTf and NIS/AgOTf as promoters, however, the only acceptor investigated was 2-azidoethanol.[21] The Kdo thioglycosides were found to be good donors giving high yields of glycosides. The elimination reaction could be avoided by performing the reactions at low temperature, which was possible through the use of reactive promoters. Stereoselectivity was a problem, as would be expected with a reactive spacer alcohol, α/β-mixtures were obtained often

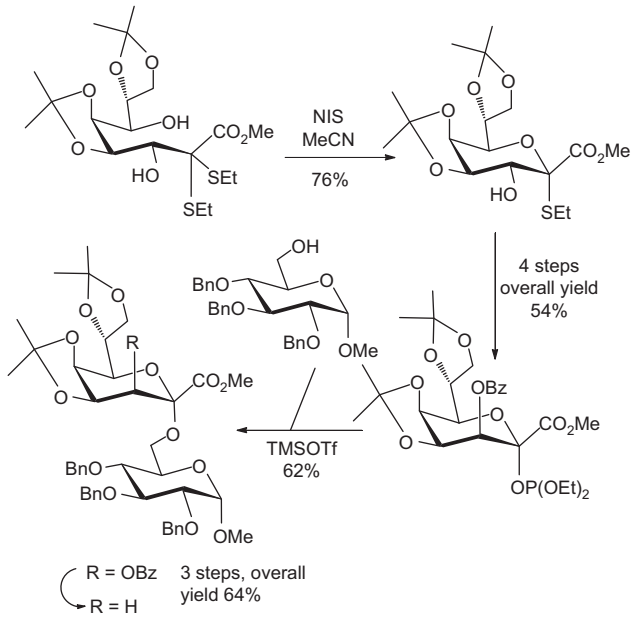


Scheme 4

with the non-natural β-anomer dominating. As discussed above also here a 4,5-*O*-isopropylidene group in the Kdo thioglycoside donor enhanced α-selectivity.

Another type of Kdo donor was investigated by Reiner and Schmidt,[22] a phosphite donor, a type of donor which they had earlier developed and used in formation of fructose glycosides also containing a quaternary aromatic center (Scheme 5).

A 3-hydroxy analogue of Kdo was formed through a reaction between 2,3;5,6-di-*O*-isopropylidene-D-mannofuranose and the anion of methyl glyoxylate diethyl mercaptal. Treatment of the obtained open dithioacetal with NIS yielded the Kdo-analogue through ring formation followed by hydrolysis of the intermediate thioglycoside. If the reaction was performed at low temperature the ethyl thioglycoside could be isolated. The equatorial 3-OH group could be reduced to give the parent Kdo derivative, but was also epimerized to yield a compound with an axial 3-OH group, which was benzoylated, thus introducing a participating group that would ensure α-selectivity in subsequent couplings. According to the authors the thioglycoside was not a good donor, but was hydrolysed and transformed into the anomeric diethyl phosphite derivate. This was used as donor promoted by TMSOTf in couplings with two different 6-OH glucose acceptors to give disaccharides in good yields (56 and 62%), and as expected with complete α-selectivity. The auxiliary 3-hydroxyl group could then be removed, in a three-step sequence with an overall yield of around 60%, to give the corresponding protected Kdo-derivatives. Although α-selective this synthesis suffers from the many transformations performed, giving a rather low overall yield from the cyclized intermediate to the final (protected) Kdo disaccharide.

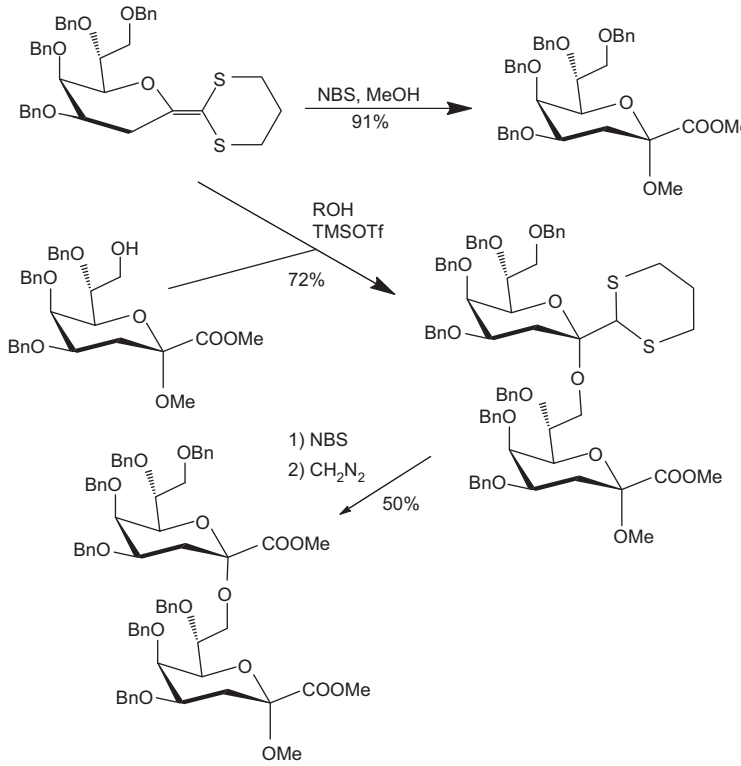


Scheme 5

Another way to Kdo oligosaccharides was described by Mlynarski and Banaszek.[23,24] Starting from D-mannose a 2-deoxy-heptono-1,5-lactone was produced in a seven-step synthesis. The last carbon was introduced by reacting the lactone with a dithiane to form a ketene dithioacetal in a Horner-Emmons reaction. TMSOTf-promoted addition of alcohols to the double bond then afforded α-glycosides both regio- and stereoselectively (Scheme 6). High yields were obtained both with primary heptose- and Kdo acceptors. NBS-promoted hydrolysis of the dithiane followed by treatment with diazomethane then gave the target protected Kdo disaccharides in about 50% yield.

Trichloroacetimidate donors of Kdo seem to be not useful in oligosaccharide synthesis, but recently Fukase and co-workers reported the use of N-phenyltrifluoroacetimidate Kdo donors in the formation of α-Kdo-(2→6)-β-D-GlcNTroc disaccharides and α-Kdo-(2→6)-β-D-GlcNTroc-(1→6)-α-D-GlcNAlloc trisaccharides (Scheme 7) for the synthesis of *Helicobacter pylori* LPS structures.[25,26] The structures synthesized are similar to the ones shown in Scheme 3 (Reference 13) and the donor used is the same except for the anomeric leaving group. After optimization they found that triflic acid as promoter in acetonitrile gave high yields (62–88%) and high stereoselectivity (68:32–100:0 α/β-ratio). Also conditions for performing the glycosylations under microfluidic conditions were developed, in which the α-selectivity of the reaction could be even further improved.

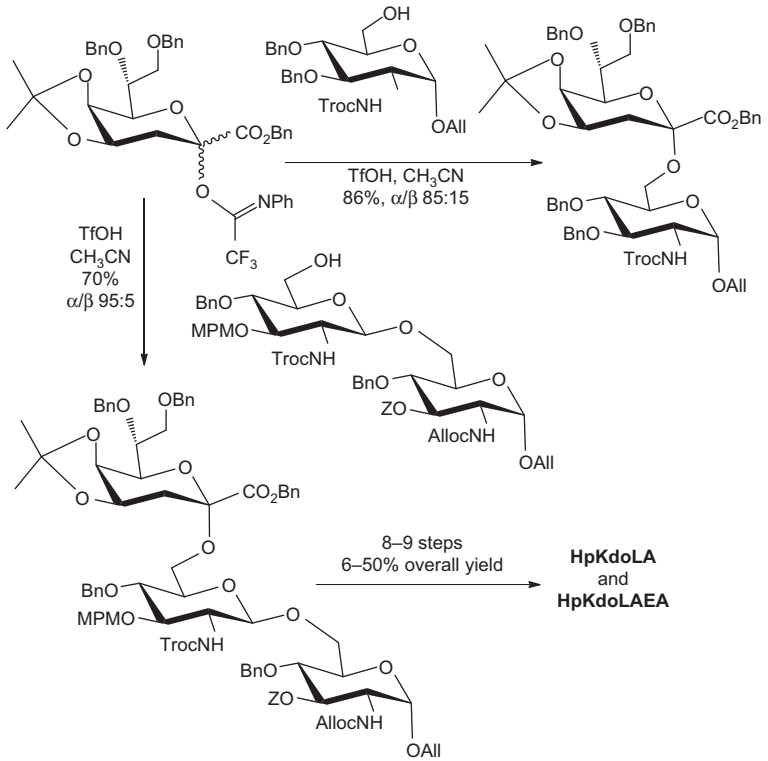
Several methods to form glycosides from easily obtained Kdo glycals by addition to the double bond have been developed.[1] This approach has the advantage that the addition is almost exclusively transdiaxial affording the desired α-glycoside selectively. Several electrophiles have been utilized to



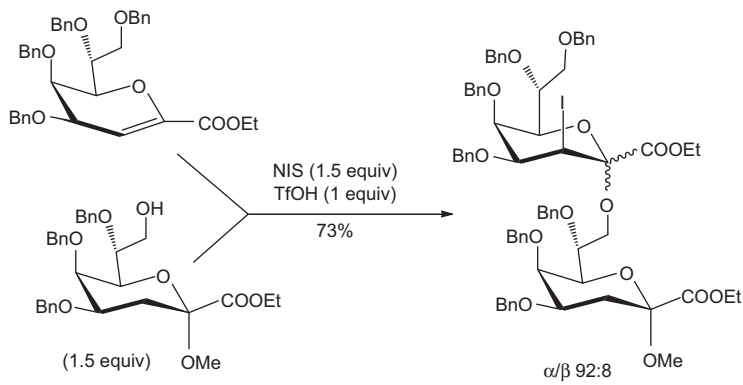
Scheme 6

activate the double bond, sulfenyl and selenyl reagents being the most effective. So far, however, the inreactivity of the conjugated double bond has required the use of quite reactive acceptors, *e.g.* primary spacer alcohols, for which the glycosylation yields are excellent. The stability of the glycal is also showed by the fact that it is not activated by NIS alone.[27] Though, if acetic or triflic acid is added in excess as an activator of NIS, the addition can take place to produce the 3-iodo-glycosides in acceptable yield.[21,28]

Tanaka, Takahashi and Takahashi developed this methodology further to allow efficient synthesis of Kdo oligosaccharides (Schemes 8 and 9).[29] The glycals were formed from open dithianes, which were transformed into sulphoxide/sulfide derivatives. Treatment of these with oxalyl chloride and silver triflate first activated the sulphoxide to afford a cyclisation and then subsequently the formed thioglycoside to yield the eliminated product. The formed glycal derivatives were all benzyl protected thus increasing the reactivity towards electrophiles as compared to earlier used acetylated compounds. This together with the triflic acid activation proved to be an effective combination for addition of even less reactive acceptors to the Kdo-glycal. Employing a variety of monosaccharide acceptors the corresponding 3-iodo-Kdo disaccharides were obtained in high yields and with excellent but not complete α-stereoselectivity. The primary 8-hydroxyl group of Kdo is known to be of surprisingly low reactivity. By using an

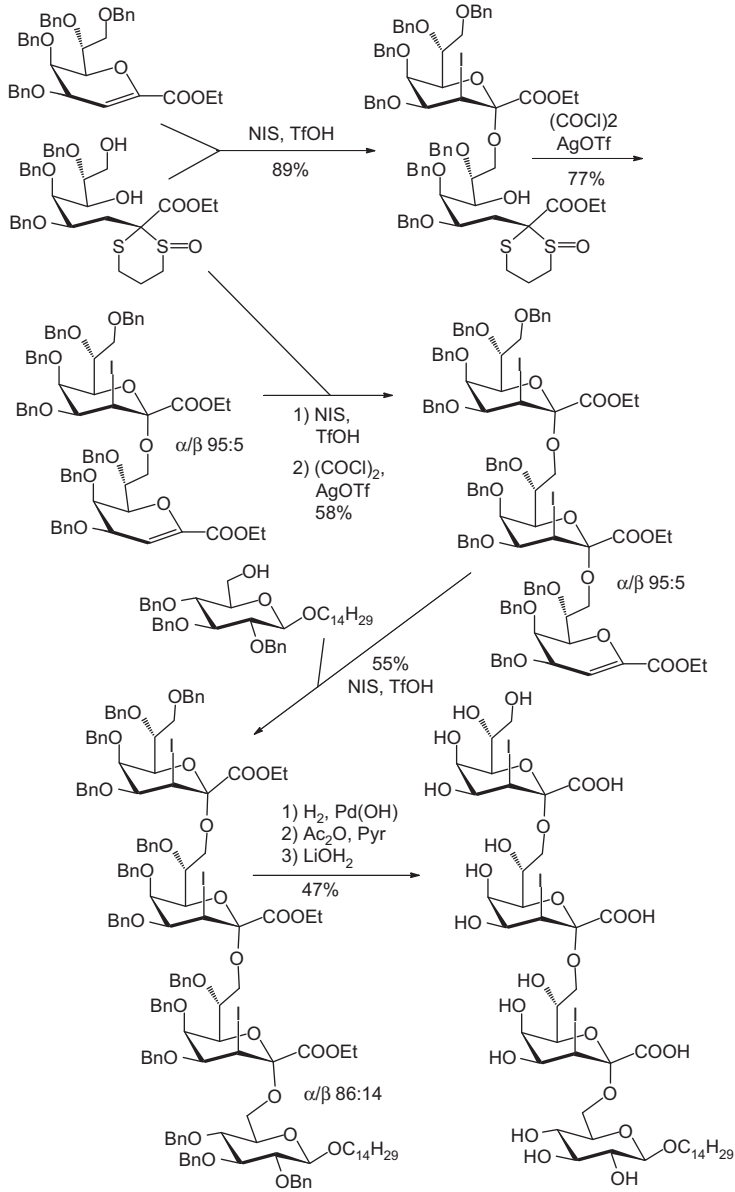


Scheme 7



Scheme 8

open form 8-OH Kdo-precursor acceptor, the reactivity was enhanced to allow an effective addition to the Kdo-glycal to produce an intermediate, which could be transformed into a Kdo-Kdo disaccharide glycal as discussed above. Iteration of this methodology made the synthesis of an α-Kdo-(2→8)-α-Kdo-(2→8)-α-Kdo-(2→6)-β-D-Glc tetrasaccharide possible in high overall yield and stereoselectivity.



Scheme 9

3 *Glycero*-D-*manno*-heptose-containing structures

The methodologies to produce *glycero*-D-*manno*-heptose, found in most inner core LPS structures, most often as the L-*glycero* stereoisomer, were mainly developed during the period 1980–2000 and were summarized in the earlier review.[1] During the last years a methodology described by Dasser *et al.* in 1991 has become popular again.[30] The methodology involves a Grignard reaction on the 6-aldehyde function of a mannodialdo derivative to give the vinyl alcohol which is converted to the corresponding heptose

derivative by dihydroxylation of the double bond, followed by periodate cleavage and reduction. The methodology has a number of advantages, notably complete L-selectivity irrespective of the protecting group pattern in the mannose derivative, the use of a stable and commercial Grignard reagent, which in most cases outweighs the disadvantage of the additional steps required, and the synthesis can be performed on a large scale.[31] Another reinvestigated methodology is the dihydroxylation of 6,7-ene compounds, which as expected from Kishi's empirical rule yields mainly the D-isomer, while the L-isomer is available through a Mitsonobu epimerization. However, the Wittig type ene-formation from the 6-aldehyde precursor is contaminated with formation of 4,5;6,7-dienes and, thus, the methodology competes well mainly when the D-isomer is wanted.[32] Recently a *de novo* synthesis of L-*glycero*-D-*manno*-heptose was published starting from 1,3-*O*-isopropylidene-1,3-dihydroxyacetone and 2,2-dimethoxyacetaldehyde. In an 11-step synthesis involving two stereoselective aldol-condensations a partially protected heptose derivative was produced in about 20% overall yield.[33]

In contrast to Kdo, for L-*glycero*-D-*manno*-heptose there is no added complexity involved regarding their use in oligosaccharide synthesis. Heptose donors behave more or less like hexose donors in glycosylation reactions and the usual type of donors (thioglycosides, halides, and acetimidates) have all been used with good success. Recent synthesis of L-*glycero*-D-*manno*-heptose-containing oligosaccharides have to quite a large extent concerned structures from the LPS inner core of *Haemophilus influenzae* and *Neisseria meningitidis*, with a common branched tetrasaccharide containing two L-*glycero*-D-*manno*-heptose residues, one Kdo moiety, and a branching (1→4)-β-D-Glc residue linked to the first heptose (Figs. 1 and 2).

The 3,4-branched heptose structure is not trivial to synthesise, probably due to steric crowding between the two equatorial substituents. Bernlind and Oscarson circumvented this problem by using a 1,6-anhydro-heptose derivative as acceptor.[34] With the two substituents in diaxial positions the glycosylations to form the branched structure was straight-forward. The synthesis of the anhydro derivative has been improved and is now performed from commercially available methyl α-D-mannopyranoside using

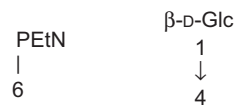


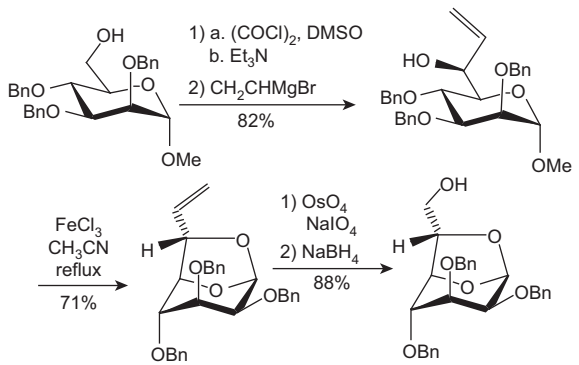
Fig. 1 Structure of the conserved inner core of *H. influenzae* LPS.



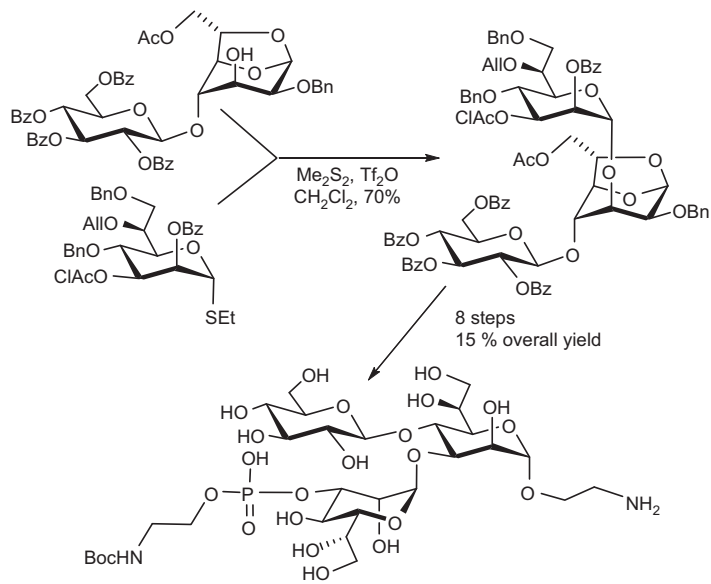
Fig. 2 Structure of the conserved inner core of *N. meningitidis* LPS.

BDA-acetal protection and the vinyl carbon elongation approach discussed followed by formation of the anhydro bridge by heating the methyl glycoside with $FeCl_3$ in THF to produce the anhydro derivative in an overall yield of $\sim$50% (Scheme 10), which then can be processed as earlier described to 3,4-branched structures.

Orthogonally protected heptosyl thioglycoside donor have been synthesised and introduced into the trisaccharide structure affording versatile intermediates for continued synthesis of various *N. meningitidis* and *H. influenzae* structures allowing almost any substituent pattern on the second heptose and also the possibility for elongation at the reducing end.[35,36] These trisaccharide intermediates has been used in the synthesis of spacer trisaccharides with variant phosphorylation pattern, a phosphoetanolamine in the 3′- or 6′-position, corresponding to *Neisseria* and *Haemophilus* structures (Figs. 1 and 2 and Scheme 11).[36]



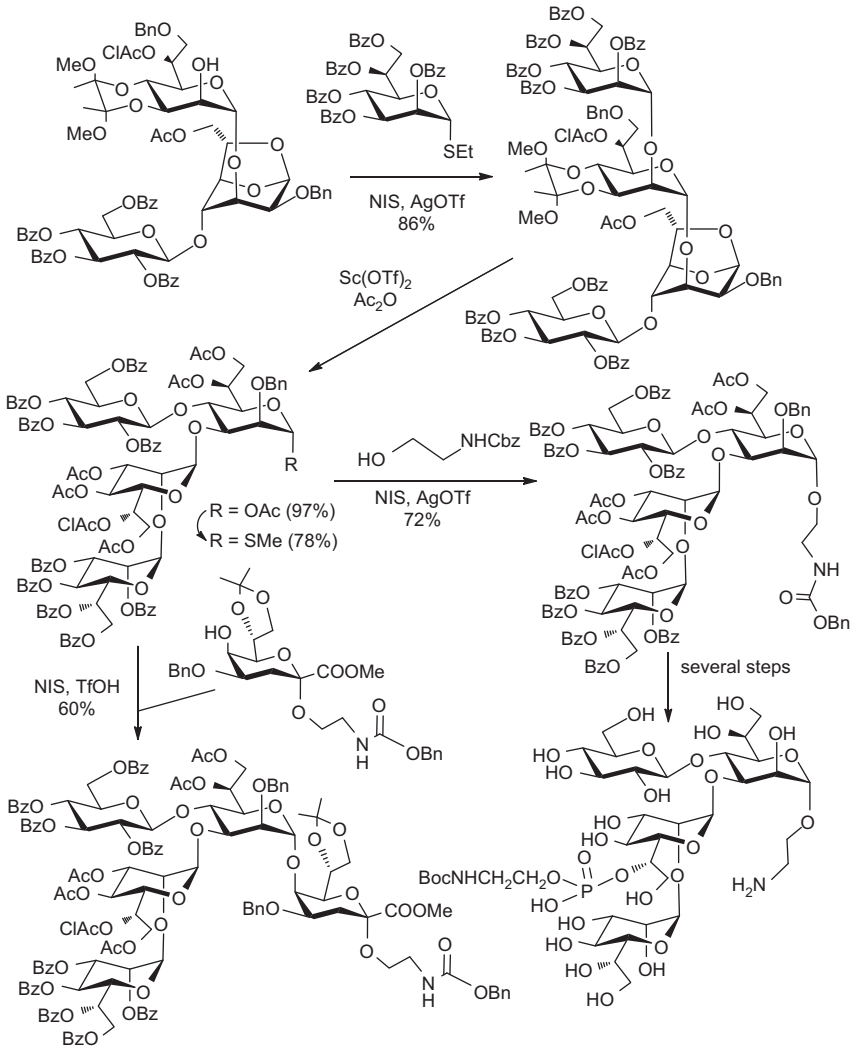
Scheme 10



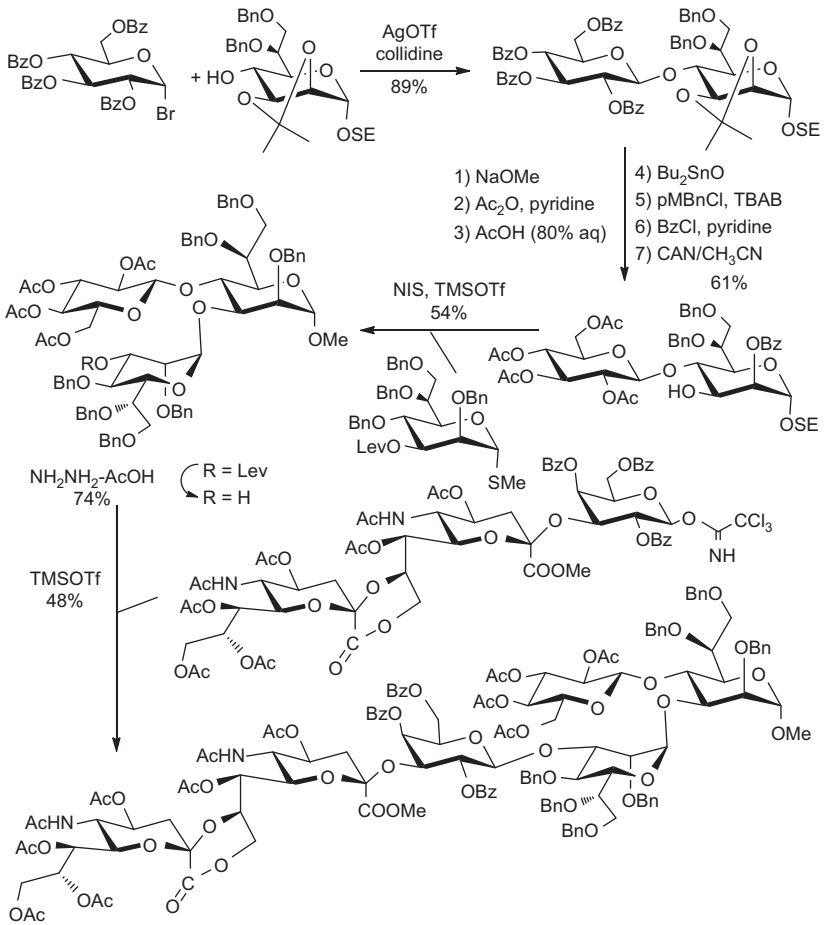
Scheme 11

A third heptose moiety was introduced into the 2′-position to give, after opening of the anhydro bridge and introduction of an ethyl thioglycoside, a tetrasaccharide building block donor corresponding to the outer part of the conserved *Haemophilus* core structure. This tetrasaccharide donor has been transformed into its 6-PEtN spacer derivative and also coupled to a Kdo acceptor (Scheme 12).[37,38]

The corresponding branched trisaccharide (β-D-Glc*p*-(1→4)-[L-α-D-Hep*p*-(1→3)]-L-α-D-Hep*p*) has also been synthesised by two Japanese groups as a precursor for continued synthesis of LPS structures. Both these groups show that it is possible to use a "normal" acceptor with a $^4C_1$-conformation in the production of the 3,4-branched structure, but care has to be taken to find the right glycosylation conditions.



Scheme 12

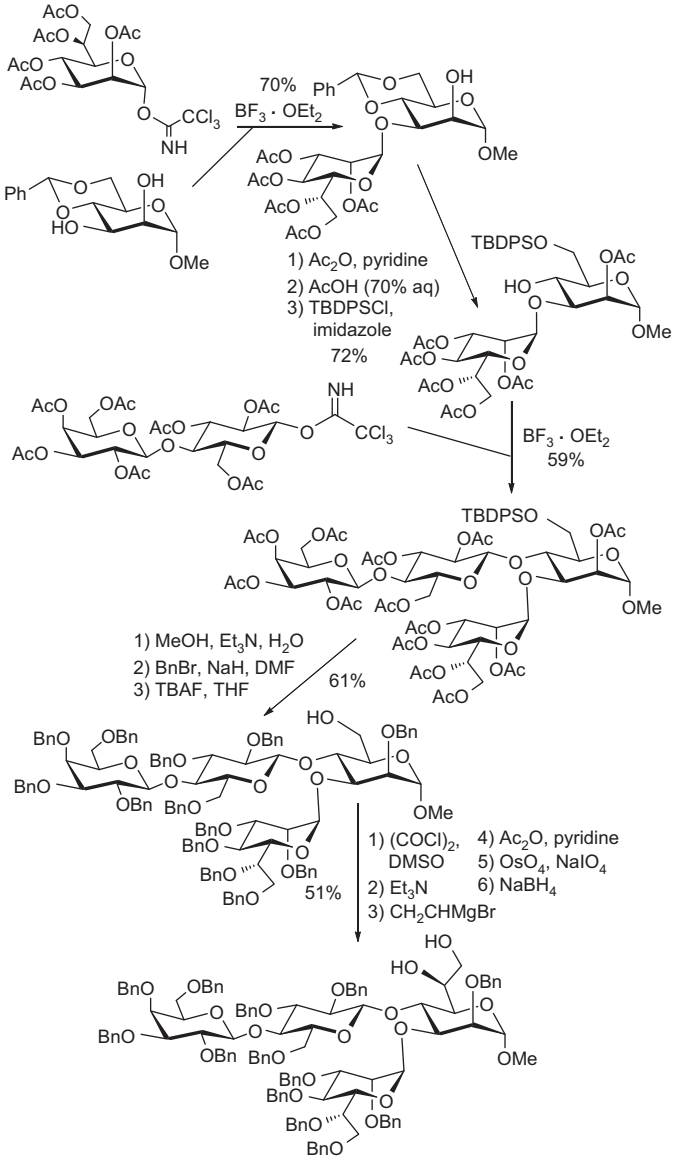


Scheme 13

Thus, Hori *et al.*,[39] when synthesising a structure from *Campylobacter jejuni* LPS, also experienced that it is not possible to introduce the heptose residue in the 3-position when there is a benzoyl-protected glucose moiety in the 4-position. However, if the benzoyl protecting groups were changed to acetates, then the glycosylation proceeded in an acceptable yield (54%, Scheme 13).

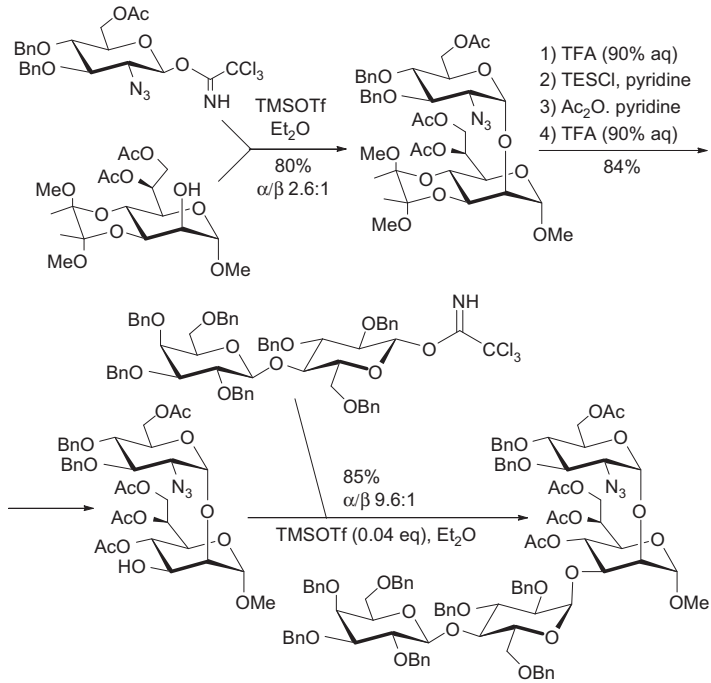
The approach to the 3,4-branched structure used by Yamasaki and co-workers was to use a mannose acceptor and first form the 3,4-branched mannose structure, which was uncomplicated, and then subsequently make the carbon-elongation at the trisaccharide level to produce the branched heptoside (Scheme 14).[40] However, with this late carbon elongation there was a penalty to pay considering massive protecting group manipulation and low-yielding reactions late in the synthesis.

The Yamasaki group has also synthesised 2,3-branched heptose structures corresponding to *N. meningitidis* LPS with an α-D-GlcNAc residue in the 2-position (Scheme 15).[41] The major problem here was not the



Scheme 14

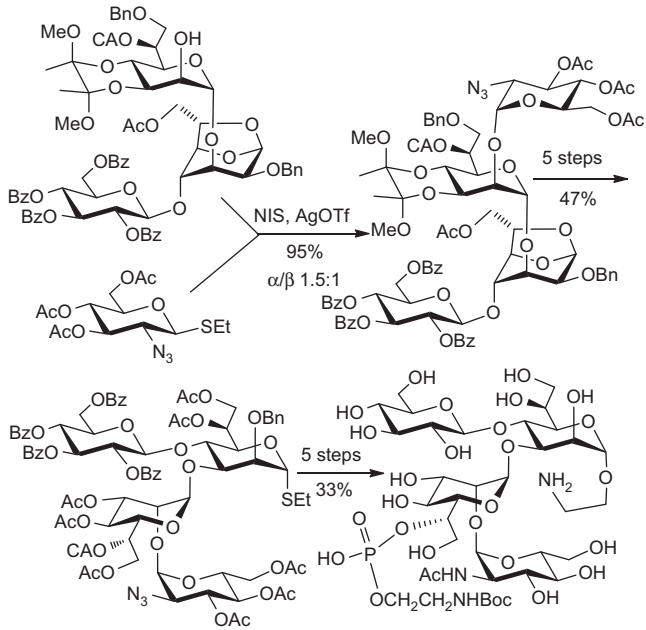
formation of the branched structure but rather the α-stereoselective introduction of the GlcNAc moiety. A 6-*O*-acetyl-2-azido-3,4-di-*O*-benzyl-2-deoxy-β-D-glucopyranosyl trichloroacetimidate donor was utilized, but after optimisation of reaction conditions the best yield of α-anomer obtained was 58% (α/β 2.6/1). After protecting group manipulation a 3-OH acceptor was produced which was glycosylated with a lactosyl trichloroacetimidate donor to yield the target 2,3-branched structure. In this coupling both the yield of the desired α-anomer (77%) and the α/β-ratio (9.6/1) was much better.



Scheme 15

The Oscarson group has also synthesized *Neisseria* structures starting from the 1,6-anhydro heptose intermediates discussed above.[42] A GlcNAc residue was introduced into the 2′-position and the formed tetrasaccharide was transformed, as for the corresponding *Haemophilus* structure, into a tetrasaccharide thioglycoside donor and further into a 6-PEtN spacer derivative (Scheme 16). As opposed to the findings of Yamasaki (Scheme 15), no stereoselectivity problems were encountered in model couplings between a monosaccharide thioglycoside heptose acceptor and an azide trichloroacetimidate donor (see also below, Scheme 17). However, when the obtained disaccharide was tried as a donor in a coupling to the earlier used disaccharide acceptor (Scheme 11) this failed. Hence, the planned [2 + 2] strategy had to be abandoned. Instead a [1 + 3] strategy was tried using the same trisaccharide acceptor as was earlier used for *Haemophilus* structures (Scheme 12). Surprisingly only NIS/AgOTf-promoted glycosylations with the thioglycoside azide donor worked. The yield was excellent but the stereoselectivity low.

Recently, the synthesis of the linear *Neisseria* tetrasaccharide structure α-D-GlcNAc-(1→2)-L-α-D-Hep-(1→3)-L-α-D-Hep-(1→5)-α-Kdo as its 5-aminopentyl glycoside was reported by Seeberger and co-workers.[43] In this case, perhaps due to the fact that the 3,4-branched structure was not targeted, the tetrasaccharide could be obtained through a [2 + 2] block synthesis using a α-D-GlcN$_3$-(1→2)-LD-Hep disaccharide donor in the last glycosylation (Scheme 17). In the synthesis trichloro- and N-phenyl

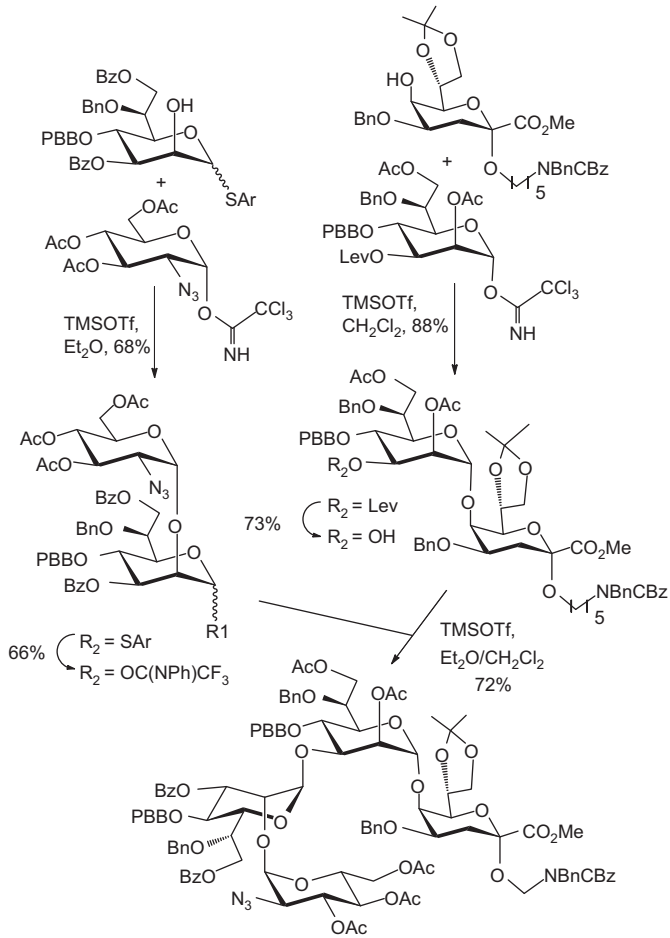


Scheme 16

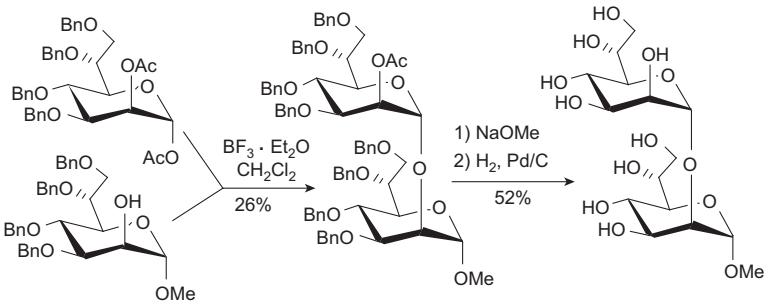
trifluoroacetimidate donors promoted by TMSOTf were employed, since the corresponding thiophenyl donors promoted by NIS/TMSOTf failed to give the desired glycosides. It is interesting to notice that Bernlind and Oscarson, in coupling of ethyl thioglycoside heptose derivatives to a very similar Kdo acceptor, found that this coupling worked well with a 2-*O*-benzyl protecting group in the heptose donor, but failed with a 2-*O*-benzoyl protected donor.[34]

As mentioned, D-*glycero*-D-*manno*-heptose is more rarely found in the core structure of LPS, but some examples reported are LPS structures from *Haemophilus, Klebsiella* and *Plesimonas*, in the latter case as the even more rarely found β-glycoside. Bernlind and Oscarson synthesized a D-*glycero*-α-D-*manno*-heptose-containing pentasaccharide structure from *H. ducreyi* as discussed in the earlier review,[44] and Gurjar and Talukdar have published a synthesis of an α-DD-Hep-(1→2)-α-DD-Hep disaccharide motif found in *K. pneumonia* (Scheme 18).[45] In the latter publication the D-*glycero*-D-*manno*-heptose derivatives were obtained from a 6,7-olefin precursor through a Sharpless asymmetric dihydroxylation reaction affording a 9:1 D/L-mixture, from which the D-isomer could be isolated as its 6,7-isopropylidene derivative in 47%. In this synthesis, a bit surprisingly, a 1-*O*-acetate donor was used, activated by BF_3-etherate, why the yield in the coupling reaction was only 26%.

Crich and Banerjee in their synthesis of *Plesimonas shigelloides* LPS-structures also utilized a 6,7-olefin precursor, but the dihydroxylation was performed using OsO_4 (Scheme 19).[32,46] The anomeric thiophenyl group was not affected under these oxidating conditions. As expected from

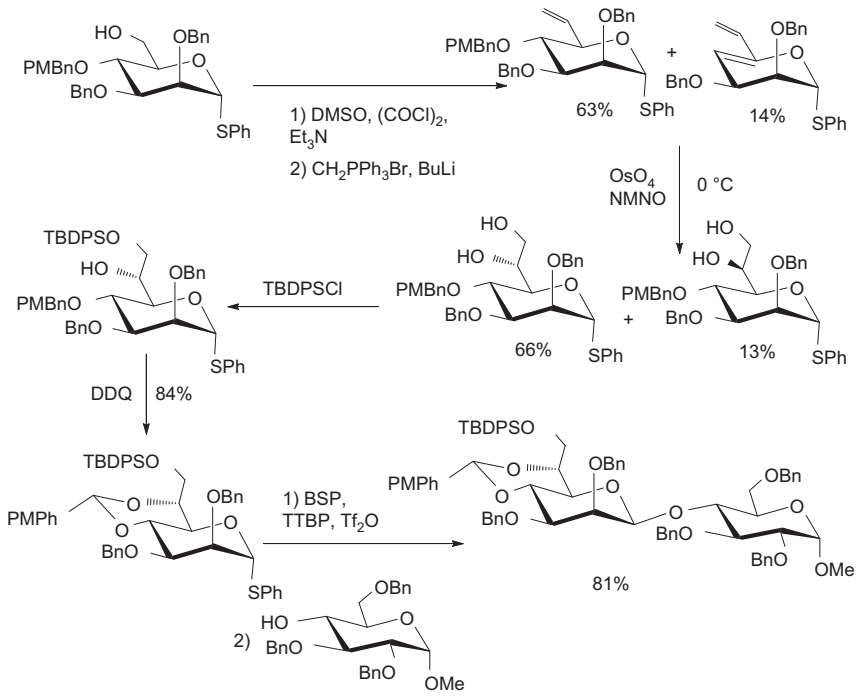


Scheme 17



Scheme 18

Kishi's empirical rule this afforded an excess of the D-isomer (D/L 5:1 at 0 °C). According to the authors asymmetric dihydroxylation did not improve these results. To get good β-selectivity in following coupling reactions with these DD-Hep derivatives, the authors utilized their

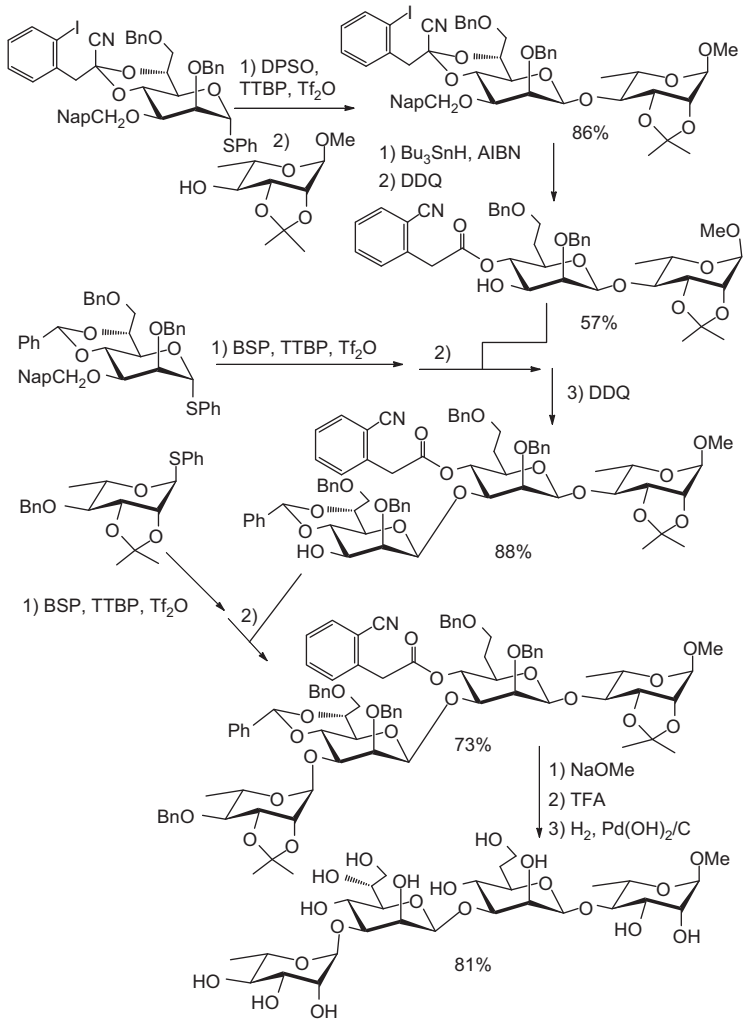


Scheme 19

own methodology originally developed for the stereoselective formation of β-D-mannopyranosides, *i.e.* pre-formation of an α-triflate, stabilized by a 4,6-acetal group, and a subsequent S_N2-reaction to displace the anomeric triflate with an acceptor. DDQ-oxidation of the 4-*O*-*p*-methoxybenzyl group in the heptose formed the necessary 4,6-*O*-benzylidene acetal through participation of the neighbouring free 6-OH group.

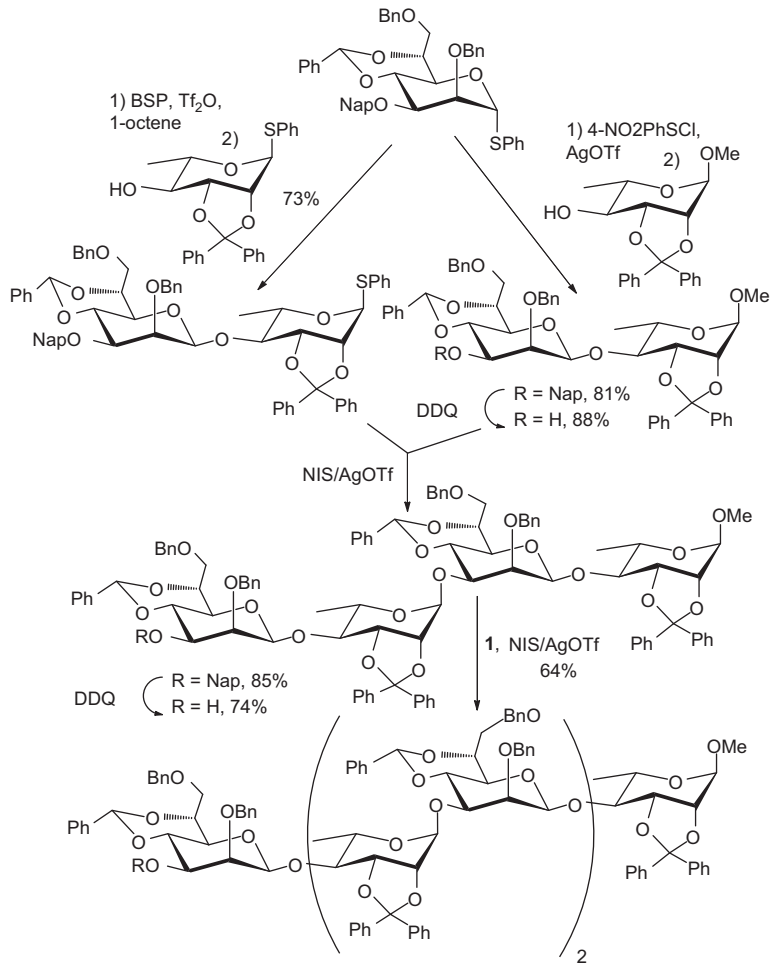
Glycosylation with these donors using 1-benzenesulfinyl piperidine/triflic anhydride as promoter at $-60\,^{\circ}C$ gave, as earlier found with mannose donors, high yields and excellent β-selectivity. The *Plesimonas* structure targeted does not only contain a D-*glycero*-D-*manno*-heptose residue but also a 6-deoxy-D-*manno*-heptose residue, both of them β-linked. To accomplish the latter structural motif a heptose thioglycoside containing a 4,6-*O*-[1-cyano-2-(2-iodophenyl)]ethylidene acetal was synthesized. When used as a donor this 4,6-acetal ensures β-selectivity as discussed and can then subsequently be fragmented by treatment with Bu_3SnH and AIBN to afford a 6-deoxy derivative (Scheme 20). By combined use of these two β-directing D-heptose thioglycoside donors an efficient synthesis of a linear tetrasaccharide from the target LPS structure, α-L-Rha-(1→3)-D-β-D-Hep-(1→3)-6d-β-D-Hep-(1→4)-α-L-Rha, could be accomplished.

In a following paper Crich and Li used the same approach to synthesize di- and trimers of the disaccharide repeating unit of a glycan from a surface-layer glycoprotein of *Bacillus thermophilus*, [→4-α-L-Rha-(1→3)-D-β-D-Hep-(1→]$_n$.[47] The homologation procedure to heptose derivatives was improved by using a 3,4-BDA-acetal in the mannose precursor,



Scheme 20

which prevented the formation of the elimination side product in the Swern oxidation/Wittig reaction sequence and thus improved the yield with about 20% over the two steps. However, the dihydroxylation step with the 3,4-BDA-acetal derivative was found to be much less selective for the D-*glycero* isomer. Following protecting group manipulation on the DD-derivative afforded a suitably protected thioglycoside heptosyl donor with the required 4,6-*O*-benzylidene acetal. Coupling of this with either a methyl rhamnoside or a phenyl thio-rhamnoside acceptor gave a starting disaccharide and an elongating disaccharide building block, respectively, both with complete β-selectivity. These building blocks were then combined through NIS/AgOTf promoted couplings to give the target tetra- and hexasaccharides in high yield and with complete α-selectivity (Scheme 21).



Scheme 21

References

1 J. Hansson and S. Oscarson, *Curr. Org. Chem.*, 2000, **4**, 535.
2 P. Kosma, *Curr. Org. Chem.*, 2008, **12**, 1021.
3 P. Kosma, ''Chemical synthesis of the core oligosaccharide of bacterial lipopolysaccharide'', in *Microbial Glycobiology* (Ed. A. P. Moran), 2009, 429.
4 L.-S. Li and Y.-L. Wu, *Curr. Org. Chem.*, 2003, **7**, 447.
5 A. Banaszek and J. Mlynarski, *Studies in Natural Prod. Chem.*, 2005, **30**, 419.
6 M. A. Ghalambor and E. C. Heath, *Biochem. Biophys. Res. Commun.*, 1963, **10**, 340.
7 J. W. Cornforth and M. E. Firth, *Biochem. J.*, 1958, **68**, 57.
8 P. Kosma, A. Reiter, A. Hofinger, L. Brade and H. Brade, *J. Endotoxin Res.*, 2000, **6**, 57.
9 P. Kosma, A. Hofinger, A. Muller-Loennis and H. Brade, *Carbohydr. Res.*, 2010, **345**, 704.
10 M. Blaukopf, B. Mueller, A. Hofinger and P. Kosma, *Eur. J. Org. Chem.*, 2012, 119.
11 M. Imoto, N. Kusunose, Y. Matsuura, S. Kusumoto and T. Shiba, *Tetrahedron Lett.*, 1988, **29**, 6277.

12 M. Imoto, N. Kusunose, S. Kusumoto and T. Shiba, *Tetrahedron Lett*., 1988, **29**, 2227.
13 H. Yoshizaki, N. Fukuda, K. Sato, M. Oikawa, K. Fukase, Y. Suda and S. Kusumoto, *Angew. Chem. Int. Ed*., 2001, **40**, 1475.
14 T. Isciyanagi, M. Fukunaga, R. Tagashira, S. Hayashi, M. Nanjo and R. Yamasaki, *Tetrahedron*, 2011, **67**, 5964.
15 G. J. P. H Boons, F. L. Delft, P. A. M van Klein, G. A. van der Marel and J. H. van Boom, *Tetrahedron*, 1992, **48**, 885.
16 R. H. Halcomb and M. D. Chappell, *J. Carbohydr. Chem*., 2002, **21**, 723.
17 A. Marra and P. Sinay, *Carbohydr. Res*., 1990, **195**, 303.
18 V. Martichonok and G. M. Whitesides, *J. Org. Chem*., 1996, **61**, 1702.
19 A. Hasegawa, H. Okhi, T. Nagahama, H. Ishida and M. Kiso, *Carbohydr. Res*., 1991, **212**, 277.
20 A. Meijer and U. Ellervik, *J. Org. Chem*., 2002, **67**, 7407–7412.
21 K. Mannerstedt, K. Ekelöf and S. Oscarson, *Carbohydr. Res*., 2007, **342**, 631–637.
22 M. Reiner and R. R. Schmidt, *Tetrahedron Asymm*., 2000, **11**, 319–335.
23 J. Mlynarski and A. Banszek, *Org. Lett*., 1999, **1**, 1709–1711.
24 J. Mlynarski and A. Banszek, *Tetrahedron Asymm*., 2000, **11**, 3737–3746.
25 A. Shimoyama, Y. Fujimoto and K. Fukase, *Synlett*., 2011, **16**, 2359.
26 A. Shimoyama, A. Saeki, N. Tanimura, H. Tsutsui, K. Miyake, Y. Suda, Y. Fujimoto and K. Fukase, *Chem. Eur. J*., 2011, **17**, 14464.
27 M. Brenken, Doctoral Thesis, 1991, Hamburg University
28 P. Kosma, H. Sekljic and G. Balint, *J. Carbohydr. Chem*., 1996, **15**, 701.
29 H. Tanaka D. Takahashi, and T. Takahashi, *Angew. Chemie Int*. Ed., 2006, **45**, 770.
30 M. Dasser, F. Chretien and Y. Chapleur, *J. Chem. Soc. Perkin Trans*., 1990, **1**, 3091.
31 R. Yamasaki, A. Takajyo, H. Kubo, T. Matsui, K. Ishii and M. Yoshida, *J. Carbohydr. Chem*., 2001, **20**, 171.
32 D. Crich and A. Banarjee, *Org. Lett*., 2005, **7**, 1395–1398.
33 T. Ohara, A. Adibekian, D. Esposito, P. Stallforth and P. H. Seeberger, *Chem. Commun*., 2010, **46**, 4106.
34 C. Bernlind and S. Oscarson, *J. Org. Chem*., 1998, **63**, 7780.
35 E. Segerstedt, K. Mannerstedt, M. Johansson and S. Oscarson, *J. Carbohydr. Chem*., 2004, **23**, 443.
36 J. D. M. Olsson and S. Oscarson, *Carbohydr. Res*., 2010, **345**, 1331.
37 K. Mannerstedt, E. Segerstedt, J. Olsson and S. Oscarson, *Org. Biomol. Chem*., 2008, **6**, 1087.
38 K. Mannerstedt, Doctoral Thesis, 2006, Stockholm University
39 K. Hori, N. Sawada, H. Ando, H. Ishida and M. Kiso, *Eur. J. Org. Chem*., 2003, 3752.
40 K. Ishii, Y. Esumi, Y. Iwasaki and R. Yamasaki, *Eur. J. Org. Chem*., 2004, 1214.
41 H. Kubo, K. Ishii, H. Koshino, K. Toubetto, K. Naruchi, and R. Yamasaki, *Eur. J. Org. Chem*., 2004, 1202
42 J. D. M. Olsson and S. Oscarson, *Tetrahedron Asymm*., 2009, **20**, 879.
43 Y. Yang, C. E. Martin and P. Seeberger, *Chem. Sci*., 2012, **3**, 896.
44 C. Bernlind, S. Bennett and S. Oscarson, *Tetrahedron Asymm*., 2000, **11**, 481.
45 M. K. Gurjar and A. Talukdar, *Tetrahedron*, 2004, **60**, 3267.
46 D. Crich and A. Banarjee, *J. Am. Chem. Soc*., 2006, **128**, 8078.
47 D. Crich and M. Li, *J. Org. Chem*., 2008, **73**, 7003.

Synthetic glycopeptides in vaccine development and antibody epitope mapping

Ulrika Westerlind
DOI: 10.1039/9781849734769-00061

Over a number of years, synthetic methods for the preparation of glycopeptide based vaccines have been refined to target the aberrant glycosylated and extensively over-expressed membrane-bound mucin MUC1 glycoprotein. On the epithelial cell surface, MUC1 forms tumor specific epitopes consisting of both the aberrantly short saccharides *e.g.* T_N, T, sialyl-T_N and sialyl-T and the exposed MUC1 tandem repeat peptide backbone. By synthesizing defined glycopeptide structures consisting of the MUC1 tandem repeat region, antibodies selective to the tumor cell surface could be formed. However the tumor-selective MUC1 glycopeptides are only moderately immunogenic and additional stimulation is necessary to generate a humoral immune response. Recent synthesis of MUC1 glycopeptides conjugated to different immuno-stimulants and promising results of evaluated antibodies induced with the synthetic MUC1 vaccines will here be highlighted.

1 Introduction

Specific immune recognition, where both the glycan and the peptide backbone contribute to the binding epitope is of particular interest for development of safe immunotherapy and immunodiagnostics. Since the discovery by Springer, that glycoproteins on the outer cell-membrane of epithelial tumor cells have an altered glycosylation consisting of the Thomsen-Friedenreich (T-) antigen and its precursor the T_N-antigen structure, the synthesis and evaluation of anti-tumor vaccines have been dominating in glycopeptide based vaccine research.[1] Recent advancement in miniaturized bioanlytical techniques and improvements in glycopeptide synthesis have further made it possible to efficiently analyze antibody binding-epitopes in a microarray format.[2]

Mucins are a class of extensively glycosylated proteins expressed on the surface of epithelial cells or secreted to function in mucus. Among them the membrane-bound mucin 1 (MUC1) is expressed on almost all epithelial tissues. Changes of protein glycosylation on the epithelial tumor cell-surface, with concomitant down-regulation of glycosyl transferases, in particular the core 2 β-1,6-N-acetylglucosaminyltransferase, and up-regulation of sialyltransferases together with increased MUC1 protein expression levels result in the formation of tumor specific epitopes.[3] The aberrant under-glycosylation of the mucin extracellular domain further lead to the exposure of its peptide backbone. The specific tumor epitopes formed therefore consist

Gesellschaft zur Förderung der Analytischen Wissenschaften e.V., ISAS – Leibniz Institute for Analytical Sciences, Otto-Hahn-Str. 6b, D-44227 Dortmund, Germany.
E-mail: ulrika.westerlind@isas.de

© The Royal Society of Chemistry 2012

of both the aberrant short saccharides *e.g.* T_N, T, sialyl-T_N and sialyl-T and the exposed MUC1 tandem repeat peptide backbone (PAHGV**TS**APDTR-PAPG**ST**AP) (Fig. 1). Moreover these tumor-associated glycan structures on MUC1 play important roles in adhesion/anti-adhesion binding events promoting tumor progression and tolerance by the immune system.[4] By inducing sufficiently strong immune reactions specific to MUC1 tumor-associated antigens, it should be possible to break the immune tolerance against these structures. A number of synthetic vaccines with the moderately immunogenic MUC1 tumor associated glycopeptide epitopes as target have recently been prepared and employed with different immuno-stimulants. Among the immuno-stimulants, carrier proteins like Tetanus Toxoid protein (T.Tox.) and bovine serum albumin (BSA) and the Toll-like receptor 2 (TLR2) ligand, tripalmitoyl-(S)-glyceryl lipopeptide Pam_3-Cys-Ser-$(Lys)_4$ (Pam_3CSK_4), have shown to be particular promising during application in MUC1 anti-tumor vaccines.[5] Multivalent glycopeptide dendrimer vaccines were recently prepared and are currently employed in immunological studies.[6]

2 Protein conjugate vaccines

Synthetic MUC1-Tetanus Toxoid conjugate vaccines have proven to be particularly interesting. The Tetanus Toxoid protein has already been applied to humans in several vaccines towards bacterial and viral targets. Recent immunizations of MUC1 vaccines in mice have resulted in highly specific and unusually strong immune responses. Furthermore several antibodies induced by these vaccines showed specific recognition of tumor cells.[5b,c]

The MUC1 tandem repeat glycopeptides containing sialyl-T_N-, T- or di-fluoro-T-antigen and a triethylene glycol spacer with a free amine in *N*-terminus were synthesized on solid-phase according to the Fmoc strategy using Fmoc amino acids and Fmoc glycosyl amino acid building blocks in a stepwise fashion. Trityl resin was used in order to avoid diketopiperazine formation. Employing TFA the fully assembled peptides were cleaved from the resin with simultaneous removal of the peptide side chain protective groups. The sialyl-T_N-disaccharide was then deprotected by hydrogenation using palladium on charcoal and subsequent deacetylation using NaOMe in methanol resulting in the formation of the sialyl-T_N-MUC1-glycopeptides **1**, **4** and **5**. The T- and di-fluoro-T-glycopeptides were directly treated with base to give the deacetylated compounds **2** and **3**. The deprotected MUC1-glycopeptides reacted with diethyl squarate by stirring in a mixture of ethanol and water at pH 8 to yield the activated monoamide constructs. The squarate monoamide-glycopeptides were conjugated to the Tetanus Toxoid protein in a sodium phosphate buffer at pH 9 and afforded vaccines **6**–**10** carrying on average more than 20 MUC1-glycopeptides conjugated to the protein (Scheme 1). Immunological evaluation of MUC1 glycopeptide vaccines containing a sialyl-T_N, a T-antigen or a di-fluoro-T-antigen on different positions on the tandem repeat **(6–10)** showed that extraordinary high antibody titers were induced in almost all of the immunized mice. Evaluation of the generated antibodies by ELISA and neutralization experiments as well as by glycopeptide microarrays gave evidence that specific immune responses were elicited towards the MUC1 antigens present

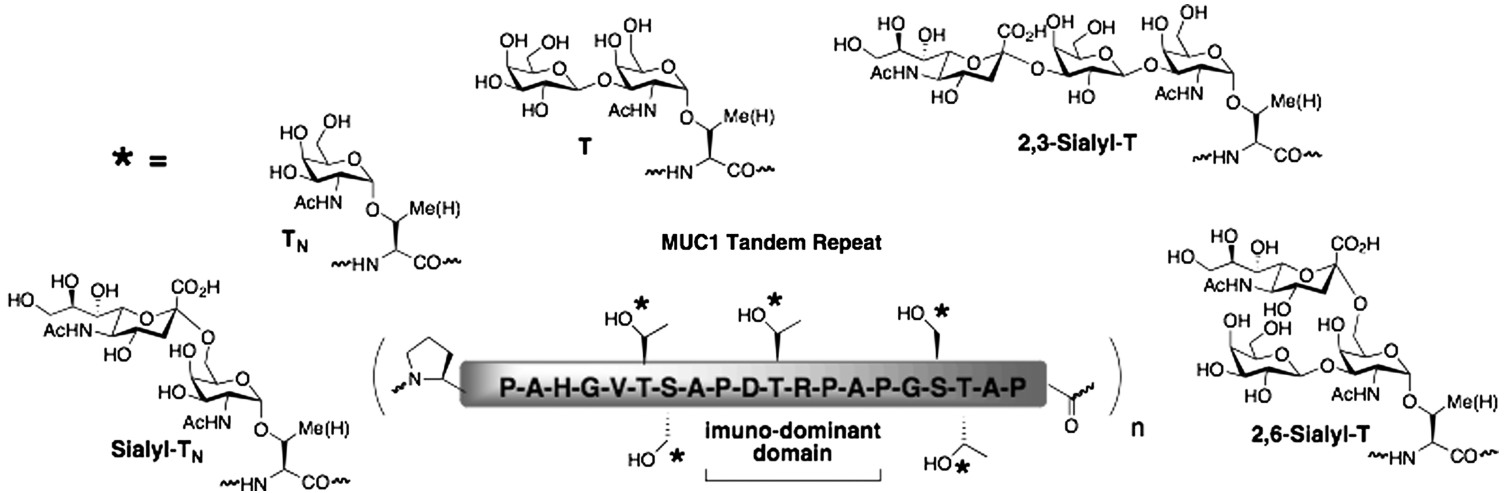
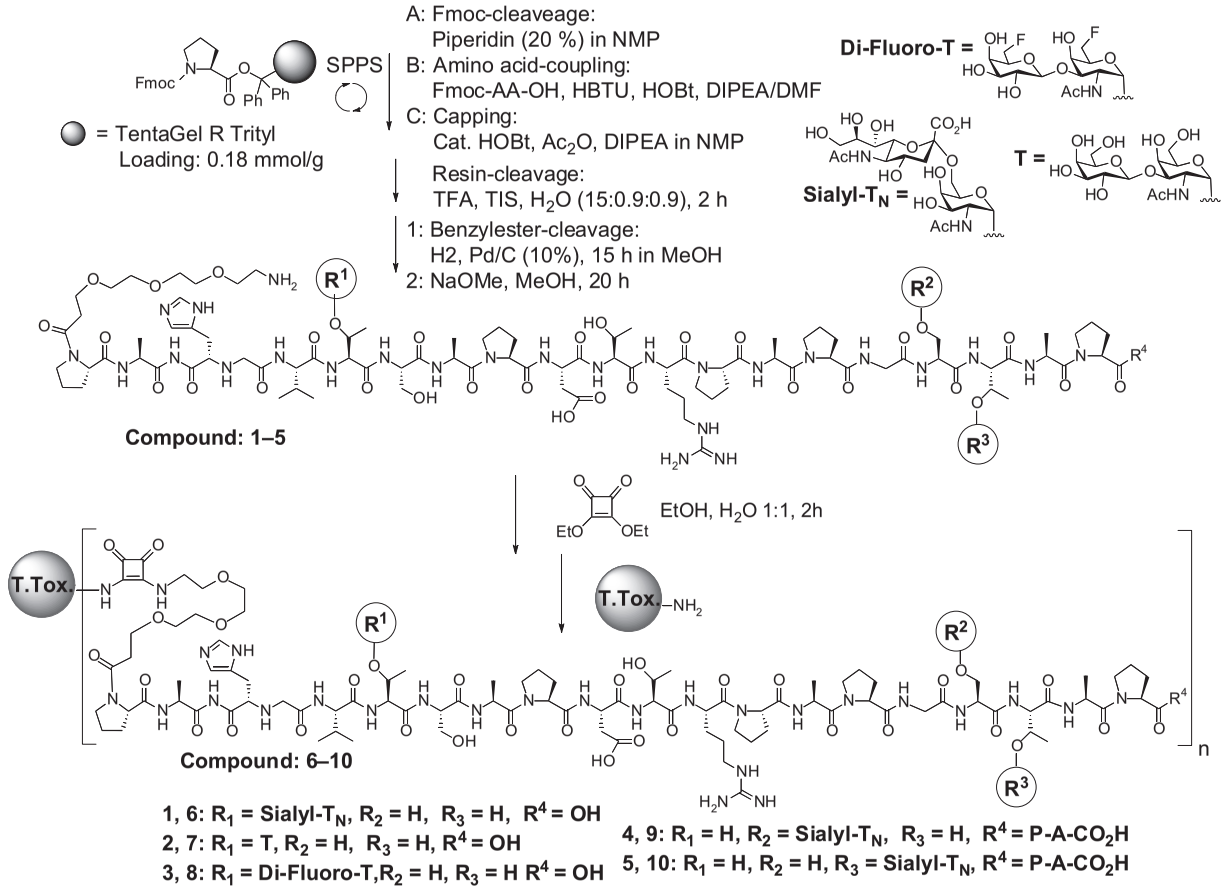


Fig. 1 Tumor-associated glycosylation on MUC1 tandem repeat peptides.



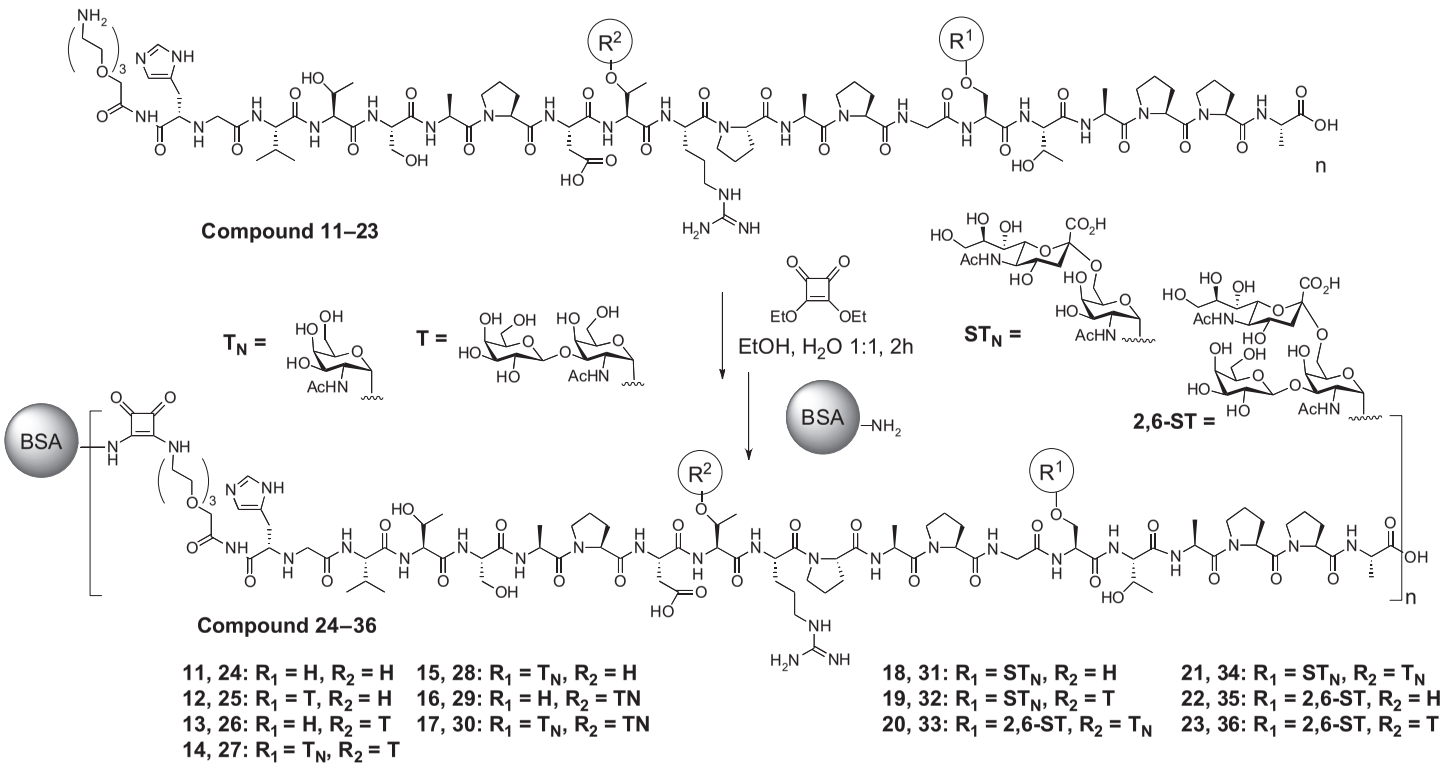
Scheme 1 Synthesis of MUC1 Tetanus Toxoid protein conjugate vaccines.

in the vaccines. FACS analysis of serum antibodies induced by the MUC1 vaccines **6–10** showed all that they recognize MCF-7 breast cancer cells. The serum from mice immunized with vaccine **9** was further evaluated through mammary carcinoma tissue staining experiments. A gradual increase of the tissue staining was found, resulting in a very strong binding to the advanced G3 phase tumor tissue. Taken together these results show that a selective immune response discriminating between healthy and diseased tissue can be efficiently obtained.[5a–c]

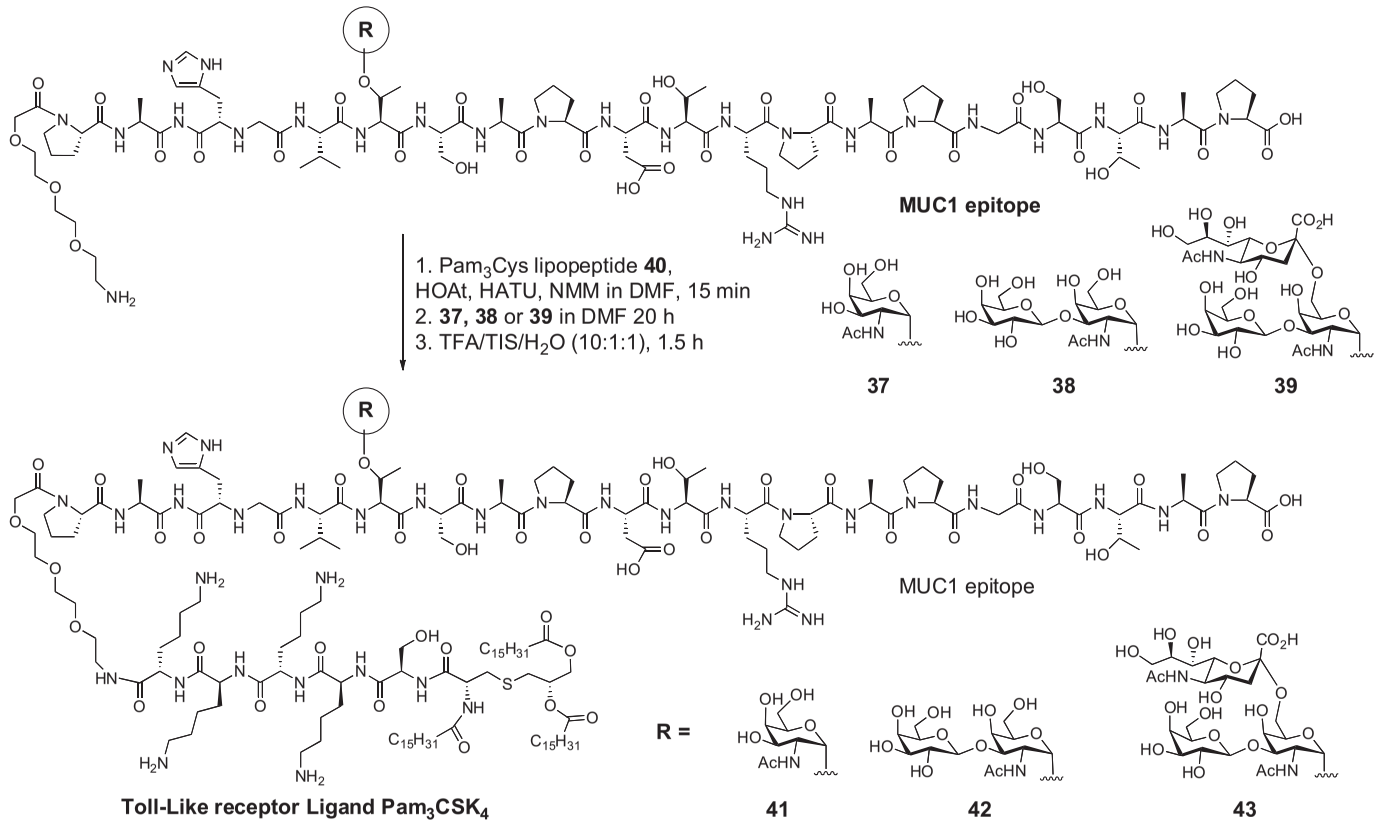
In a different study MUC1 glycopeptides conjugated to a bovine serum albumin (BSA) immune carrier protein were evaluated.[5d,7] In accordance with the above described T.Tox. vaccines, the MUC1 glycopeptides were conjugated to the protein through coupling to diethylsquarate. The MUC1 peptides were glycosylated with the T_N- and T- antigen on Thr9 in the immuno-dominant PDTR region or on Ser15 in the GSTA region of the tandem repeat. Immunological evaluation showed that the elicited antibody titers were depending on the position glycosylated. Vaccines glycosylated with T_N or T **(26, 29, 30)** in the MUC1 PDTR region induced a strong immune response. Glycosylation with the T-antigen or T_N of Ser15 resulted in medium or weak antibody titers **(25, 27, 28)**. No glycosylation of the MUC1 backbone **(24)** resulted in weak immune responses.[7] Triggered by these results, a number of sialyl-T_N and 2,6-sialyl-T glycopeptide BSA constructs **(31–36)** modified on Thr9 in the PDTR domain and on Ser15 in the GSTA region were prepared. Immunization in mice resulted in strong immune responses in all applied sialyl-T_N and 2,6-sialyl-T MUC1 vaccines **(31–36)**.[5d] Studies of antibody isotypes from the antibody serum showed that the IgG_1 isotype is predominant in each case, but some IgM, IgA, $IgG_{2a,b}$ and IgG_3 antibodies are also present. FACS analysis to study binding recognition to MCF-7 breast tumor cells showed that antibodies induced by the non-glycosylated vaccine **24** showed hardly any binding recognition to the tumor cells, vaccine **31** carrying a sialyl-T_N on Ser-15 surprisingly only showed weak binding. Antibody sera induced by vaccines **27, 28, 30** and **32–36** showed a strong binding to the tumor cells. Sera from vaccines **25, 26** and **29** were not tested in the FACS analysis binding studies. From these studies it could be concluded that compared with Tetanus Toxoid conjugates, the BSA vaccines are more dependent on glycosylation and MUC1 backbone conformation in order to efficiently induce a strong immune response. However, the BSA immune carrier is cheaper than the Tetanus Toxoid protein carrier and as shown here, it can in many occasions induce strong and tumor selective immune responses (Scheme 2).

3 Build-in adjuvant vaccines

A number of two- or three-component vaccines containing a Toll-like receptor two (TLR2) ligand have been synthesized and evaluated.[5e,f,g] These self-adjuvanting vaccines avoid invoking an immune response to the immune-carrier, a problem commonly seen employing protein conjugate vaccines. The synthesis of two-component vaccines with a Pam_3Cys TLR2 ligand connected to T_N, T and sialyl T MUC1 glycopeptides was recently described (Scheme 3). The Pam_3CSK_4 lipopeptide fragment **40** was



Scheme 2 Synthesis of MUC1 BSA protein conjugate vaccines.



Scheme 3 T_N, T and Sialyl-T MUC1 Pam3Cys two-component vaccines.

prepared by Fmoc solid-phase synthesis having the amino acid side chain protecting groups still protected after resin cleavage. Different T_N, T and sialyl T MUC1 glycopeptides **(37–39)** were further prepared with both the glycan and peptide protecting groups removed. Peptide synthesis was followed by fragment condensation employing HATU and HOAt and after additional deprotection resulting in the formation of the lipopeptide vaccine constructs **41**–**43**. Immunization of these vaccines in mice showed that a specific immune response was induced in all mice although not with the extraordinary high antibody titers found employing MUC1 Tetanus Toxoid vaccines.[5f] In a later study di- and tri component vaccines were synthesized and compared by immunological evaluation. The three-component vaccines, containing an extra T-cell peptide epitope, showed here a stronger immune response if the MUC1 peptides were glycosylated while the non-glycosylated two- and three-component vaccines did not show any difference in antibody titers.[5e]

4 Dendrimer vaccines

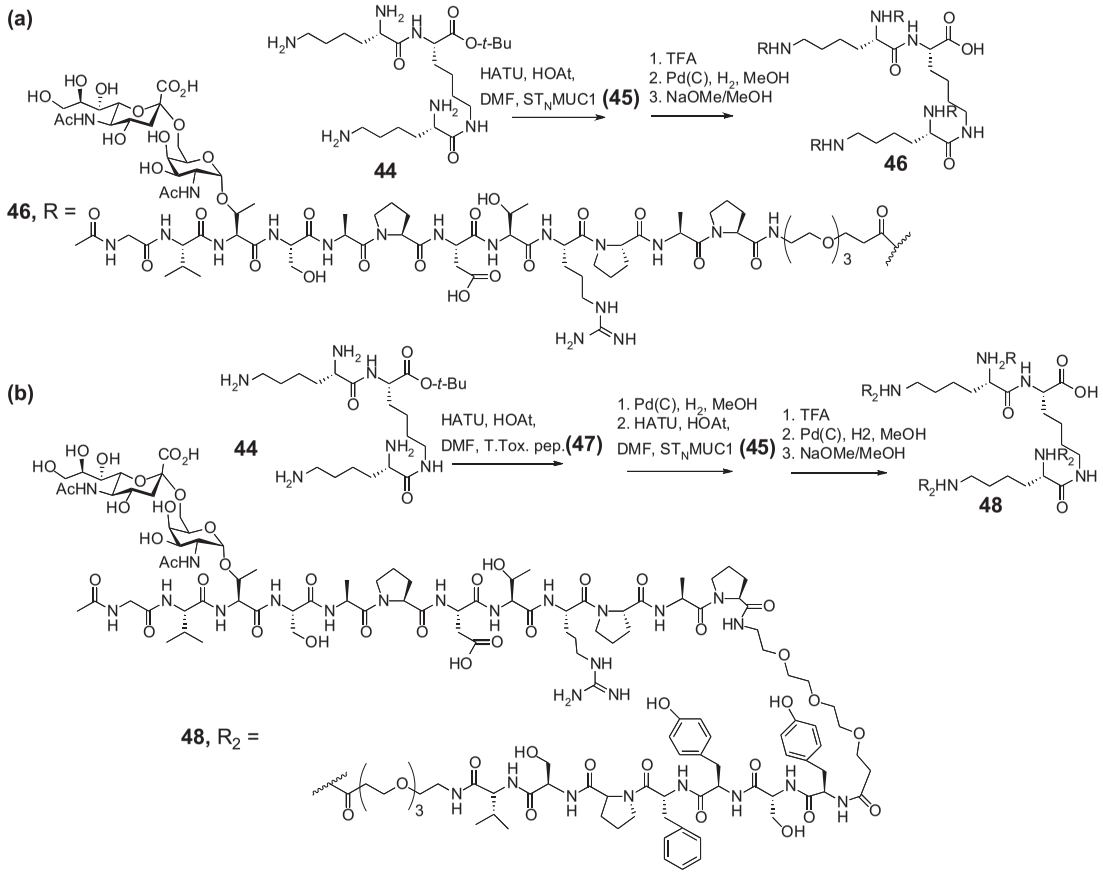
Like the lipopeptide Pam_3CSK_4 vaccines, constructs based on dendrimer multi-epitope presentation can avoid that carrier-induced immune suppression take place during immunization. Vaccines based on multiple antigen presentation may further facilitate antigen uptake and presentation by antigen presenting cells (APCs). Linear peptide backbones, cyclic peptide scaffolds or multi-lysine scaffolds are commonly applied for multivalent presentation of glycans or glycopeptides.[8] Recently MUC1 and MUC4 T_N- and sialyl-T_N-glycopeptide dendrimer two- and three component vaccines based on the oligo-lysine core were reported.[6a,9] A MUC1 sialyl-T_N-glycopeptide dendrimer vaccine was prepared either with or without an extra immunostimulating Tetanus Toxoid T-cell peptide epitope. In between the B- and T-cell epitope and the di-lysyl-lysine core a polar non-immunogenic triethylene glycol spacer amino acid was incorporated. The sialyl-T_N-MUC1-glycopeptide **45** and the Tetanus Toxoid T-cell peptide epitope **47** with a spacer at the *C*-terminal end were prepared by Fmoc-solid-phase peptide synthesis. Starting with aminomethyl-polystyrene resin (AMPS) functionalized with an allylic HYCRON linker the assembly of the peptide backbone were performed according to standard procedures. Palladium(0)-catalyzed cleavage of the allylic linker then gave the MUC1 glycopeptide **45** and the Tetanus Toxoid peptide **47**. For construction of the dendrimers, fragment condensation of sialyl-T_N-glycopeptide **45** to the di-lysyl lysine core **44** was performed employing HATU/HOAt and *N*-methyl morpholine for activation and resulted in the MUC1-glycopeptide dendrimer. Acidolytic removal of the peptide side-chain protecting groups and glycan deprotection by catalytic hydogenation of the benzyl ester and careful deacetylation using catalytic sodium metoxide in methanol generated the dendrimer vaccine **46**. In a similar fashion a glycopeptide dendrimer vaccine including a Tetanus Toxoid T-cell epitope was prepared. Fragment condensation of the Tetanus Toxoid peptide **47** to the solid-phase-linked di-lysine core **44** was carried out. The N-terminal di-benzyloxycarbonyl protecting group was then removed by

hydrogenation employing palladium on charcoal, followed by coupling of the carboxy-activated MUC1-glycopeptide **45**. Subsequent removal of the peptide and glycan protecting groups gave the dendrimer vaccine **48**. The prepared synthetic dendrimer vaccines are currently under investigation in immunization studies[6a] (Scheme 4).

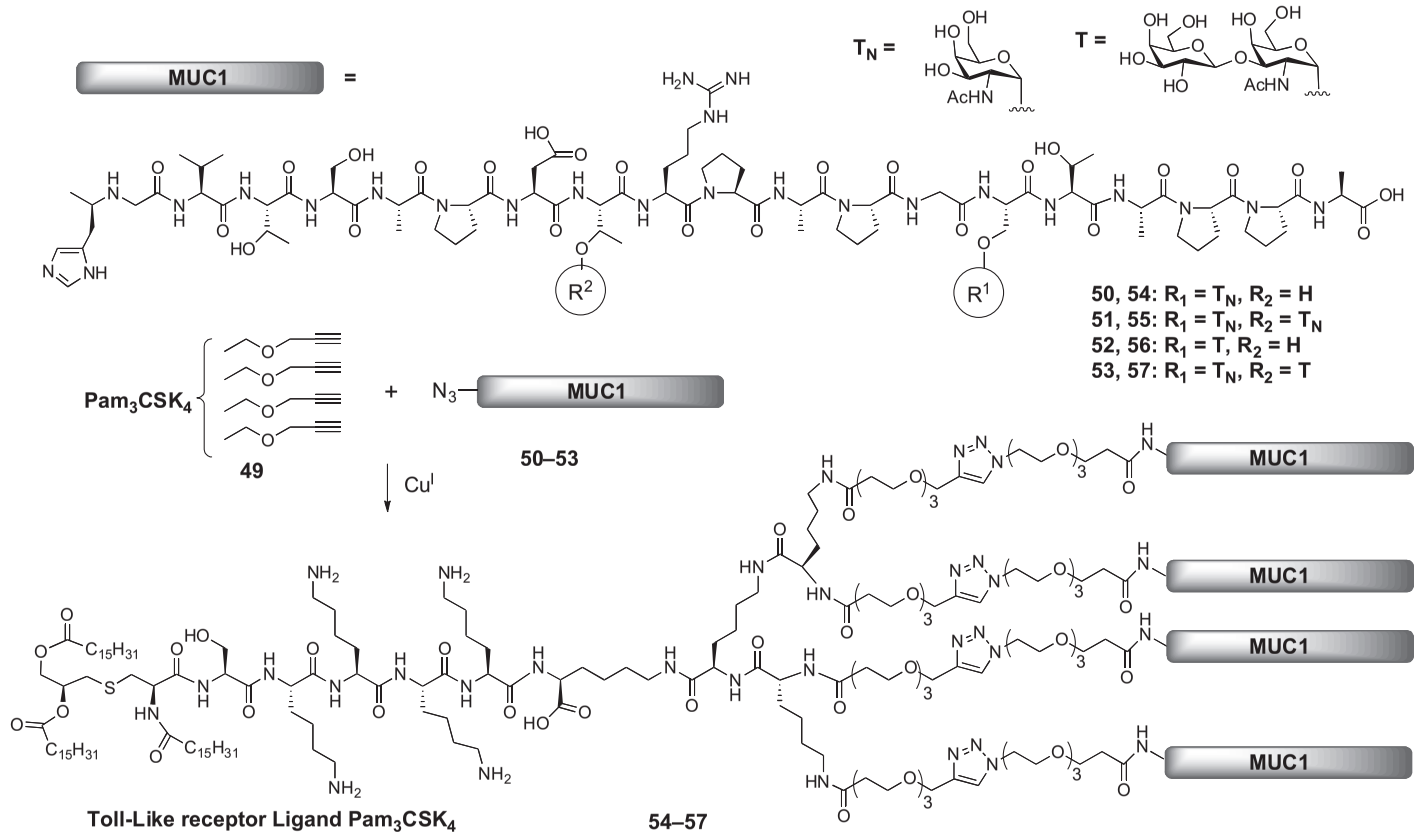
Self-adjuvanting multivalent MUC1 T and T_N glycopeptide vaccine constructs efficiently conjugated through azide alkyne click chemistry to the Pam_3CSK_4 lipopetide immune-stimulant have also been prepared.[6b] On solid-phase, the Pam_3CSK_4 lipopeptide was prepared with an extra *C*-terminal lysine residue with the side-chain protected with a 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methyl-butyl (ivDde) group, which could be selectively removed using hydrazine monohydrate. Repeated coupling with Fmoc-Lys(Fmoc)-OH resulted in a branched di-lysine core structure which was alkyne functionalized by coupling with a triethyleneglycol-alkyne spacer. Glycopeptides **50**–**53** containing a *N*-terminal azide spacer were prepared by standard Fmoc-solid-phase-synthesis. Through copper (I) mediated click chemistry the lipopeptide alkyne construct **49** were then efficiently coupled to glycopeptides **50**–**53** generating the multivalent vaccine constructs **54**–**57**. These vaccines are still to be tested for potential immunogenicity (Scheme 5).

5 Antibody epitope mapping

For development of safe anti-tumor vaccines, it is important that induced antibodies are specific to the tumor-associated MUC1 glycopeptide structures and do not cross-react with glycan structures found on normal cells. Variations between the tumor-associated glycan structures, the positions that are glycosylated and glycosylation density per tandem repeat are seen depending on tumor tissue and stage of disease (stage I-IV). Correlation between antibody binding epitope analysis and binding preferences to different tumor tissue or tumor cell-lines, contribute to conclude which structures that are relevant for future vaccine design. Studies of antibody specificity are also important in order to evaluate immune response differences between individuals. Through antibody neutralization experiments or by immobilization of glycopeptides on microarrays the binding preferences of antibodies induced with the synthetic MUC1 vaccines were further studied.[2a,5a,b] Antibodies induced with vaccines containing sialyl-T_N glycosylation on Thr-6 on the MUC1 glyccopeptide conjugated to T.Tox. protein or to a OVA T-cell epitope, were highly specific to both the sialyl-T_N glycan and to the HGVT glycosylation position. If the sialyl-T_N glycan was moved to a different position on the repeat, then the peptide was not recognized by the antibodies. Glycosylation with T_N, or 2,6-sialyl-T in the HGVT position resulted in moderate binding and glycosylation with T, 2,3-sialyl-T or no-glycosylation resulted in very weak or loss of binding recognition. Evaluation of antibodies induced with T.Tox. conjugated to MUC1 glycopeptides with the T-antigen or a 2,6-di-fluoro-T antigen in Thr-6 in the HGVT region, resulted in completely different specificity pattern. These antibodies showed a broad binding specificity, and the sialyl-T_N, T_N, T, sialyl-T and the non-glycosylated MUC1 tandem repeat peptides were all



Scheme 4 Synthesis of a) sialyl-T_N-MUC1-glycopeptide dendrimer, b) sialyl-T_N-MUC1-T.Tox dendrimer vaccine.



Scheme 5 Multivalent MUC1-Pam_3CSK_4 dendrimer vaccines obtained by copper(I)-mediated click reactions.

recognized by the induced antibody sera. The MUC1 peptide backbone seemed to be particularly important for antibody binding and truncation of the peptide backbone to a 12-mer peptide resulted in that the binding drastically dropped. The MUC1 T-antigen could further not be neutralized by a glycopeptide from a different mucin, MUC4. In contrast to the more specific antibodies initially induced, the antibodies induced with the T-antigen or 2,6-di-fluoro-T antigen MUC1 T.Tox. vaccines with a slightly broader binding recognition did strongly recognize breast cancer tumor cells employing FACS analysis.

Analysis of monoclonal antibodies known to recognize tumor tissue or analysis of auto-antibodies induced in patients with different stages of cancer, may also be valuable in vaccine design. The SM3 monoclonal antibody induced from deglycosylated mucin from human milk, known to bind to tumor tissue, was studied in detail.[2a,10] Through microarray glycopeptide analysis, it was found that the antibody do not only bind to the non-glycosylated MUC1 tandem repeat or MUC1 glycosylated with T_N in the immunodominant PDTR region previously found, but glycosylation with T_N or sialyl-T_N in the GSTA region was also strongly recognized. In later studies employing glycopeptide microarrays, human serum auto-antibodies from cancer patients were found to be directed to MUC1 glycopeptide epitopes glycosylated with sialyl-T_N on the Thr or Ser in the GSTA region.[2b] These findings resulted in the preparation of T.Tox sialyl-T_N MUC1 glycopeptide vaccines glycosylated in these positions **(9, 10)**. As described above, antibodies induced from these vaccines **(9, 10)** recognize both cancer cells from the MCF-7 cell-line and breast cancer tissue.[5c]

Interestingly, in another study sera were analysed from 395 patients with early stage breast cancer.[11] High levels of serum antibodies against the sialyl-T_N MUC1 and also against core 3 glycans were found, which could be associated with a benefit in metastasis free survival. The core 3 glycan and the glycosyltransferase (β3GNT6) responsible for core 3 glycan extension, is however not found in normal or malignant breast tissue. Core 3 antibody induction may therefore be a result of mimicry or due to inflammation in the colon, where these glycans commonly are expressed.[12] In summary it could be concluded that sialyl-T_N glycosylation on MUC1 tandem repeats is a highly relevant breast cancer target both as diagnostic marker and for therapeutic applications.

6 Conclusions

During the recent years, a number of vaccines consisting of tumor-associated MUC1-glycopeptides containing the T_N-, T-, sialyl-T_N- and sialyl-T-saccharide antigens have been synthesized by their conjugation to different immuno-stimulants. Immunological evaluation of these vaccines has shown that antibodies were induced which are specific towards tumor-associated glycopeptides and do recognize the membrane glycoproteins on cancer cells. Several of these synthetic vaccines showed promising effects, in particular the MUC1 Tetanus Toxoid vaccines. Furthermore, extensive synthesis of glycopeptides for antibody analysis in a microarray format make it possible

to identify new potential vaccine candidates as well as to study in detail the antibody binding epitopes induced by the different vaccines.

References

1 (a) G. F. Springer, *Science* 1984, **224**, 1198–206; (b) J. Taylor-Papadimitriou, J. Burchell, D. W. Miles, M. Dalziel, *Biochim. Biophys. Acta, Mol. Basis Dis.* 1999, **1455**, 301–313; (c) T. Becker, S. Dziadek, S. Wittrock, H. Kunz, *Curr. Cancer Drug Targets* 2006, **6**, 491–517.

2 (a) U. Westerlind, H. Schroeder, A. Hobel, N. Gaidzik, A. Kaiser, C. M. Niemeyer, E. Schmitt, H. Waldmann, H. Kunz, *Angew. Chem. Int. Ed.* 2009, **48**, 8263–8267; (b) O. Blixt, E. Clo, E. Nudelman, K. Kildegaard Sö rensen, T. Clausen, H. H. Wandall, P. Livingston, H. Clausen, K. J. Jensen, *J Proteome Res.* 2010, **9**, 5250–5261; (c) H. H. Wandall, O. Blixt, M. A. Tarp, J. W. Pedersen, E. P. Bennett, U. Mandel, G. Ragupathi, P. Livingston, M. A. Hollingsworth, J. Taylor-Papadimitriou, J. Burchell, H. Clausen, *Cancer Res* 2010, **70**, 1306–1313.

3 (a) I. Brockhausen, *Biochim. Biophys. Acta, Gen. Subj* 1999, **1473**, 67–95; (b) I. Brockhausen, J.-M. Yang, J. Bruchell, C. Whitehouse, J. Taylor-Papadimitriou, *Eur. J. Biochem.* 1995, **233**, 607–17; (c) J. M. Burchell, A. Mungul, J. Taylor-Papadimitriou, *J Mammary Gland Biol* 2001, **6**, 355–64; (d) K. O. Lloyd, J. Burchell, V. Kudryashov, B. W. T. Yin, J. Taylor-Papadimitriou, *J. Biol. Chem.* 1996, **271**, 33325–33334.

4 (a) B. J. Swanson, K. M. McDermott, P. K. Singh, J. P. Eggers, P. R. Crocker, M. A. Hollingsworth, *Cancer Res.* 2007, **67**, 10222–10229; (b) L. G. Yu, N. Andrews, Q. Zhao, D. McKean, J. F. Williams, L. J. Connor, O. V. Gerasimenko, J. Hilkens, J. Hirabayashi, Kasai, K.-i., *J Biol Chem* 2007, **282**, 773–81.

5 (a) A. Kaiser, N. Gaidzik, U. Westerlind, D. Kowalczyk, A. Hobel, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* 2009, **48**, 7551–7555; (b) A. Hoffmann-Roeder, A. Kaiser, S. Wagner, N. Gaidzik, D. Kowalczyk, U. Westerlind, B. Gerlitzki, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* 2010, **49**, 8498–8503; (c) N. Gaidzik, A. Kaiser, D. Kowalczyk, U. Westerlind, B. Gerlitzki, H. P. Sinn, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* 2011, **50**, 9977–9981; (d) H. Cai, Z.-H. Huang, L. Shi, Z.-Y. Sun, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Angew. Chem. Int. Ed.* 2012, **51**, 1719–1723; (e) B. L. Wilkinson, S. Day, L. R. Malins, V. Apostolopoulos, R. J. Payne, *Angew. Chem. Int. Ed.* 2011, **50**, 1635–1639; (f) A. Kaiser, N. Gaidzik, T. Becker, C. Menge, K. Groh, H. Cai, Y.-M. Li, B. Gerlitzki, E. Schmitt, H. Kunz, *Angew. Chem. Int. Ed.* 2010, **49**, 3688–3692; (g) V. Lakshiminarayanan, P. Thompson, M. A. Wolfert, T. Buskas, J. M. Bradley, L. B. Pathangey, C. S. Madsen, P. A. Cohen, S. J. Gendler, G. J. Boons, *Proc. Natl. Acad. Sci.* 2012, **109**, 261–266.

6 (a) S. Keil, A. Kaiser, F. Syed, H. Kunz, *Synthesis* 2009, (8), 1355–1369; (b) H. Cai, Z.-H. Huang, L. Shi, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Chem. Eur. J.* 2011, **17**, 6396–6406.

7 H. Cai, Z.-H. Huang, L. Shi, P. Zou, Y.-F. Zhao, H. Kunz and Y.-M. Li, *Eur J Org Chem*, 2011, 3685–3689.

8 (a) A. Schierholt, M. Hartmann, T. K. Lindthorst, *Carbohyd. Res.* 2011, **346**, 1519–1526; (b) I. Bossu, M. Sulc, K. Krenek, E. Dufour, J. Garcia, N. Berthet, P. Dumy, V. Kren, O. Renaudet, *Org. Biomol. Chem.* 2011, **9**, 1948–1959; (c) E. M. V. Johansson, S. A. Crusz, E. Kolomiets, L. Buts, R. U. Kadam, M. Cacciarini, K.-M. Bartels, S. P. Diggle, M. Camara, P. Williams, R. Loris, C. Nativi, F. Rosenau, K.-E. Jaeger, T. Darbre, J.-L. Reymond, *Chemistry &*

biology 2008, **15**, 1249–1257; (d) X. Tian, J. Pai, K.-H. Baek, S.-K. Ko, I. Shin, *Chem. Asian. J.* 2011, **6**, 2107–2113.
9 (a) J. P. Tam, *Proc. Natl. Acad. Sci.* 1988, **85**, 5409–13; (b) T. Becker, A. Kaiser, H. Kunz, *Synthesis* 2009, (7), 1113–1122.
10 (a) J. Burchell, S. Gendler, J. Taylor-Papadimitriou, A. Girling, A. Lewis, R. Millis, D. Lamport, *Cancer Res.* 1987, **47**, 5476–82; (b) J. Burchell, J. Taylor-Papadimitriou, M. Boshell, S. Gendler, T. Duhig, *Int. J. Cancer* 1989, **44**, 691–6.
11 O. Blixt, D. Bueti, B. Burford, D. Allen, S. Julien, M. A. Hollingsworth, A. Gammerman, I. Fentiman, J. Taylor-Papadimitriou, J. Burchell, *Breast Cancer Res* 2011, **13**, R25.
12 (a) B. J. Campell, L. G. Yu, J. M. Rhodes, *Glycoconj* 2001, **18**, 851–858; (b) J. W. Pedersen, O. Blixt, E. P. Bennett, M. A. Tarp, I. Dar, U. Mandel, S. S. Poulsen, A. E. Pedersen, S. Rasmussen, P. Jess, H. Clausen, H. H. Wandall, *Int J Cancer* 2010, **128**, 1860–1871.

Posttranslational sialylation and its impact on leukocyte recruitment during inflammation

Karin Bodewits and Markus Sperandio*
DOI: 10.1039/9781849734769-00075

The immune response to invading pathogens during inflammation is crucial for viability of humans and other mammals. It needs leukocytes to leave the blood vessel system (intravascular compartment) and transmigrate into infected or inflamed tissue to fulfil its function as immune cells. The navigation of immune cells during this process requires a complex network of adhesion, signalling and direction molecules on the immune cells and the blood vessel wall. These molecules include selectins, selectin-ligands, chemokines, chemokine receptors, and integrins. Carbohydrate decorations of some of these adhesion and signalling molecules play a major role in effective leukocyte recruitment, as they can regulate proper function, stability, protease resistance, or quaternary structure. In this chapter, the role of carbohydrate decorations in leukocyte recruitment with a strong focus on posttranslational modification by sialic acids will be reviewed.

1 Leukocyte recruitment cascade

Leukocytes continuously circulate through the vasculature, from one tissue to another and to acute and chronic sites of inflammation. Leukocyte trafficking to tissues where leukocytes are required serves a critical function in the mammalian immune response and is mediated by a large number of adhesion and signalling molecules. The first reports on leukocyte recruitment date back to the early nineteenth century [1, 2]. However, only during the last two to three decades, the molecular mechanisms governing leukocyte recruitment have been elucidated. We have learned that during acute inflammation the recruitment process of mostly neutrophils primarily occurs in postcapillary venules of the affected tissue and follows a well-characterised cascade of events that takes place in the presence of hemodynamic shear forces exerted by the flowing blood (Fig. 1) [3–5].

Neutrophils, also known as polymorphonuclear (PMN) leukocytes, are the most abundant leukocytes in the peripheral blood of humans and are rapidly and efficiently recruited to areas of inflammation. The recruitment process starts with the capture of free flowing neutrophils to the vessel wall, followed by rolling along the endothelium. Capture and rolling of the cells along the endothelium is mediated by a group of glycoproteins, called selectins, which interact with carbohydrate determinants on selectin ligands [6–8]. During rolling, neutrophils slow down and get into close contact with endothelial cells, which enable endothelial-bound chemokines to interact with specific chemokine receptors present on the cell surface of neutrophils.

Walter-Brendel-Center of Experimental Medicine, Ludwig-Maximilians-Universität, Marchioninistr. 15, 81377 Munich, Germany. E-mail: markus.sperandio@lmu.de

© The Royal Society of Chemistry 2012

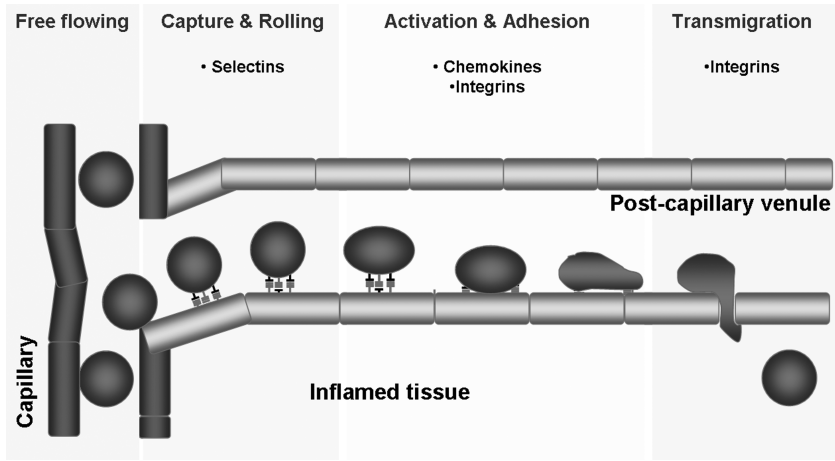


Fig. 1 The multistep leukocyte recruitment cascade. During inflammation, neutrophils or other leukocytes are captured to the luminal surface of inflamed postcapillary venules, where they start rolling along the endothelium. During rolling, chemokine/chemokine receptor interactions trigger the activation of leukocyte-expressed integrins leading to firm arrest of leukocytes on the inflamed endothelium followed by adhesion strengthening and spreading on the endothelium. To find a suitable way out of the vasculature, leukocytes start crawling along the endothelium until the right spot has been found for transmigration.

This in turn triggers the activation of leukocyte-expressed integrins (inside-out signalling), which interact with endothelial-expressed ligands, leading to firm adhesion and ultimately to the arrest of neutrophils on the endothelium [9, 10]. Thereafter, adhesion is further strengthened and neutrophil spreading induced by integrin/integrin ligand interactions (outside-in signalling) [11]. Interestingly, neutrophils do not readily transmigrate through the vessel wall into tissue after the initial firm arrest, but start to slowly crawl on the endothelial luminal surface in search of a good exit point into the inflamed or infected tissue [12]. Upon entering the tissue, gradients of chemokines and other chemoattractants (chemotaxis) direct the further migration of the cells [13].

1.1 Leukocyte capture and rolling: selectins and their ligands

The capture of free flowing leukocytes and rolling along the endothelium is considered a crucial step during the recruitment of leukocytes into inflamed tissue. Leukocyte capture and rolling are mediated by selectins recognising and binding to specific carbohydrate determinants on selectin ligands [14]. Three different types of selectins have been identified in mammals: leukocyte (L)-, platelet (P)-, and endothelium (E)-selectin [8]. L-selectin is constitutively expressed on the surface of most leukocytes and can be shed from the cell surface upon stimulation. P-selectin is stored in Weibel-Palade bodies of endothelial cells and is rapidly mobilised to the cell surface during inflammation or injury. As the name implies, P-selectin is also expressed by activated platelets. Like P-selectin, E-selectin is expressed on endothelial cells. However, E-selectin does not reside in intracellular storage pools, but is newly synthesised after stimulation by different inflammatory mediators. In the mouse, the predominant leukocyte rolling receptor in acutely

inflamed endothelial cells is P-selectin. However, this might be different in humans [15]. P-selectin is specialised in capturing free circulating leukocytes from the blood stream, whereas E-selectin is more liable for stabilising the rolling movement and slowing it down [16].

All three selectins are members of a glycoprotein family called C-type lectins [17], which means they recognise and bind to specific carbohydrate determinants (lectins) and typically need calcium (C-type) for binding. The best-known binding ligand of selectins is the tetrasaccharide sialyl-Lewis-X (sLe^X), which will be discussed later in this chapter. Endothelium-expressed E- and P-selectin bind to selectin ligands expressed on rolling leukocytes. The selectin ligands comprise a heterogeneous group of proteins and glycolipids that are all decorated with specific carbohydrate determinants containing sialic acid modifications. Most of the protein ligands bear the selectin-binding carbohydrate moiety at the N-terminus of the protein away from the cell membrane [18]. P-selectin glycoprotein ligand-1 (PSGL-1), the most prominent member, is expressed on the surface of most leukocytes and functions as a ligand for P-, E- and L- selectin [19, 20]. Besides PSGL-1, many other leukocyte-expressed selectin ligands have been described, among them, the E-selectin ligands CD44, E-selectin ligand-1 (ESL-1), and CD43 [21, 22].

After capturing, neutrophils start to roll on the activated endothelium. The rolling velocity on the vessel wall is $\sim$100–1000-fold slower than the average blood flow velocity [14]. In general, for leukocytes to roll against hemodynamic shear forces, selectin binding to its ligand needs to comply with the following three requirements: 1) rapid bond formation, 2) high tensile strength during binding, and 3) fast dissociation rates [23]. In the case of selectins and their ligands, we know that bond formation requires the presence of shear forces to support rolling. In addition, it has been shown for L-selectin dependent rolling that leukocytes detach from the endothelium when blood flow velocity falls under a critical velocity level [24]. Furthermore, under physiological conditions, bonds become more resistant to shear forces with increasing shear stress (catch bonds) [25]. This characteristic behaviour of selectins is of utmost importance for the recruitment of free flowing leukocytes to the vessel wall and enables subsequent adhesion and transmigration of leukocytes [8].

1.2 Leukocyte adhesion: chemokines, integrins, and their ligands

Before transendothelial migration takes place, firm adhesion of leukocytes to the vessel wall occurs, and leukocytes stop rolling. The transition from rolling to arrest can be triggered by chemokine-or selectin-dependent signalling events leading to the activation of leukocyte-expressed integrins [10].

Chemokines are a large family of small proteins, 8- to 12-kD in size. Some of its members including CXCL1 (keratinocyte-derived chemokine, KC) or CXCL8 (IL-8) can be found on the luminal surface (glycocalyx) of inflamed endothelium in response to pro-inflammatory stimuli [26, 27]. During rolling, chemokine receptors expressed on neutrophils (including CXCR2) sense the repertoire of endothelium-immobilised chemokines on the luminal surface of the vessel wall. Specific interactions between chemokines and their respective chemokine receptors then induce intracellular signalling

events (inside-out signalling), leading to leukocyte integrin activation and firm leukocyte arrest on the endothelium [4].

Integrins are type I transmembrane glycoproteins, composed of an α and a β subunit, and are amongst others expressed on the surface of leukocytes. Integrins that play an important role for neutrophil recruitment include the β_2 integrins leukocyte function-associated antigen 1 (LFA-1, $\alpha_L\beta_2$) and macrophage-1 antigen (Mac-1, $\alpha_M\beta_2$). β_2-integrins are required for slow neutrophil rolling, adhesion strengthening, spreading, crawling towards transmigration sites, and transmigration (diapedesis) through the vessel wall [28]. β_2 integrins are constitutively expressed on free flowing leukocytes, but are mostly kept in their inactive form awaiting activation. They can have three different conformations of their extracellular domain: the bent or inactive form that prohibits binding to its receptors, the extended intermediate-affinity form, and the extended high-affinity conformation [29]. Chemokine-triggered activation causes a rapid conformational change from the inactive (bent) to the extended intermediate-affinity form. This change allows the integrin to bind to its endothelial-expressed ligand, leading to a further conformational shift to the extended high-affinity conformation. This full repertoire of integrin activation is required for leukocyte firm arrest on the vessel wall [28].

Relevant endothelium-expressed integrin ligands for LFA-1 include intercellular adhesion molecules-1 and -2 (ICAM-1, ICAM-2) [30, 31]. Mac-1 binds to the adhesion molecules ICAM-1 and the endothelial receptor for advanced glycation end products (RAGE) [32].

After integrin mediated firm arrest on the vessel wall, neutrophils undergo several morphological changes including adhesion strengthening and spreading which prepare the neutrophil to slowly move (crawling) to an appropriate location where it can exit the vasculature [12].

1.3 Transendothelial migration (diapedesis)

The attachment of leukocytes to the vessel wall and transmigration to inflamed tissue was described more than 150 years ago, and much progress in understanding the diapedesis step has been made since by using electron microscopy (EM), which enabled the observation of morphological changes of both the leukocyte and the endothelium during diapedesis. Two modes of transmigration have been described: the paracellular route and the transcellular route [33]. The paracellular route, which is the most widely accepted route, describes the migration of leukocytes between the junctions of two or more endothelial cells. The transcellular route describes the migration of leukocytes through endothelial cells [34]. The understanding of the molecular mechanism of both diapedesis routes has only just begun to emerge from studies conducted in genetically engineered mice with the identification of different adhesion and signalling molecules (including ICAM-2, PECAM-1, JEM-C, CD99, VE- cadherin) to be involved in transmigration.

2 Posttranslational modifications

The cells of all organisms are composed of four fundamental macromolecular components: nucleic acids, proteins, glycans, and lipids [35].

The term glycan comprises any form of mono-, oligo- or polysaccharide, which can exist in its unbound form or connected to another molecule. Glycans differ from the amino acids and nucleic acid polymers in three ways: 1) they are built-up out of a large repertoire of monomers, 2) they can be heavily branched structures, and 3) the linkage types between the different monomeric subunits are highly variable [36]. The glycome of a single species can consist of hundreds of thousands of distinctive carbohydrate structures, and can by all its diversity carry much more information than the other three building blocks of life [37, 38]. Glycosylation is ubiquitous among eukaryotes and though the historical perception of glycans was that they are mere cellular decorations, we now know that they are of central importance in a wide variety of biological processes, including host-pathogen interactions, cellular trafficking, and intercellular adhesion [39].

Glycosylation can significantly contribute to the biological properties of proteins and lipids including its conformation, solubility, or function as selective ligand. In this context, the influence of carbohydrate decoration on immune function and leukocyte trafficking has gained increasing interest in recent years. Selectin-selectin ligand interactions are among the best-studied carbohydrate-protein interactions known in biology [18, 40]. Furthermore, several studies have shown that posttranslational modifications are functionally relevant for chemokine-dependent arrest of the leukocytes on the vessel wall and integrin-mediated leukocyte adhesion [41, 42].

2.1 Glycans

Glycans decorate the surfaces of cells and are attached to lipids and proteins. Despite their diversity, each class of glycans has a limited number of core structures, and the different capping (outer) carbohydrate groups create the high overall diversity. The terminals, or outer capping carbohydrates, are also often responsible for the biological functions of the carbohydrate decoration [43].

Carbohydrate decorations of cell surface- and secreted proteins are typically covalently attached to proteins and lipids through either a nitrogen atom (N-glycans) or an oxygen atom (O-glycans). The nitrogen atom is supplied by the amino acid asparagine for N-glycans and either a serine or threonine residue of the protein can supply the oxygen atom as attachment point for O-glycans [36]. The different sugar monomers are attached to the underlying protein or glycoconjugate by enzymes called glycosyltransferases.

Selectin ligands carry simple and complex sialylated and fucosylated sequences, in some ligands sulphated sequences can be found as well. A common carbohydrate structure found on selectin ligands is sialyl Lewis X (sLe^X). sLe^X is a tetrasaccharide consisting of galactose (Gal), fucose (Fuc), *N*-acetylglucosamine (GlcNAc), and sialic acids (Fig. 2). It is biosynthesised by several glycosyltransferases in the Golgi complex and can be linked to a diverse variety of glycoconjugates, such as *N*-glycans, *O*-glycans, and glycolipids. On selectin ligands, sLe^X is typically found as a capping group of core 2-decorated O-glycans and is crucial for binding to selectins. Some endothelial selectin ligands which are important in lymphocyte homing carry a sulfated form of sLe^x, 6-sulfo-sLe^X, which is generated by the help of

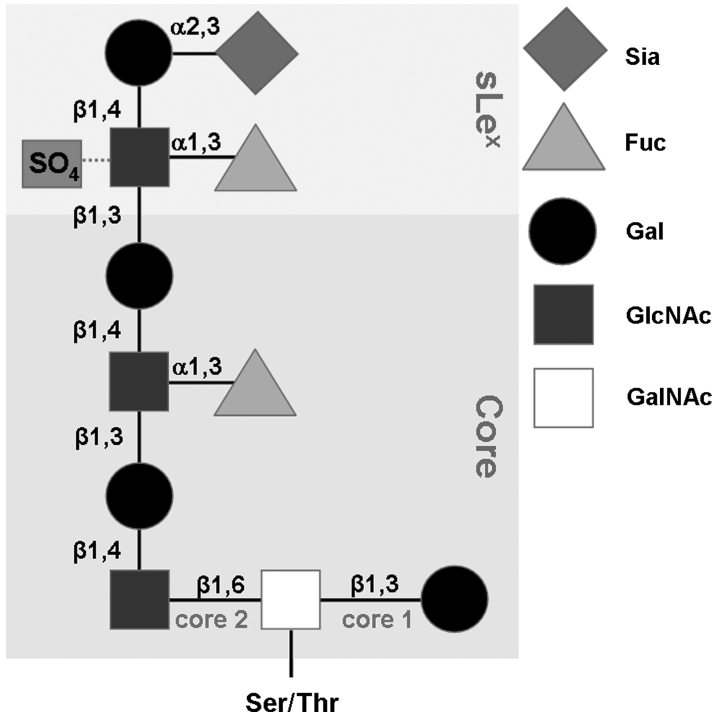


Fig. 2 Key carbohydrate decoration of selectin ligands. The tetrasaccharide sialyl Lewis X (sLe^X) is shown as the capping group (bioactive group) of core 2 decorated O-glycans attached to a serine or threonine residue. The presence of sLe^X on the outside of selectin ligands is critical for the recognition by L-, P-, and E-selectin during leukocyte recruitment. The GlcNAc can be sulfated at C6-position, forming 6-sulfo-sLe^x.

two additional GlcNAc6-sulfotransferases called GlcNAc6ST-1 and GlcNAc6ST-2 [44].

2.2 Glycosyltransferases involved in leukocyte trafficking

The posttranslational modification of proteins and lipids is a dynamic, controlled, and consecutive process. Carbohydrate chains are biosynthesised or biodegraded in a well-organised mode due to the different substrate specificities of glycosyltransferase and glycosidase enzymes [45]. In eukaryotes, post-translational glycosylation mainly takes place in the Golgi apparatus and is dependent on a group of Golgi resident enzymes named glycosyltransferases [46]. Glycosyltransferases form the predominant integral membrane proteins of the Golgi-machinery membranes and over 200 have been identified in the human genome [40]. The sugar substrates of the glycosyltransferases can be biosynthesised and activated in different compartments of the cell, such as cytoplasm and nucleus, before being transferred to the Golgi apparatus.

Most glycosyltransferases belong to the type II transmembrane proteins that have a cytoplasmic tail (CT), a transmembrane-spanning domain (TM), and a stem plus catalytic domain. They have been identified as monomers, homo-dimers, hetero-dimers, and hetero-oligomers [47]. These enzymes specifically transfer activated sugar nucleotide donors to a specific OH-group of a glycoconjugate *via* a α- or β- linkage. Glycosyltransferase activity can be regulated at several levels, which include transcriptional

regulation of the genes encoding the transferases or activation of the enzymes by the presence of their substrates [45].

Several glycosyltransferases involved in the biosynthesis of functional adhesion and signalling molecules essential for immune cell trafficking have been described, including glucosaminyltransferases, galactosyltransferases, fucosyltransferases, and sialyltransferases [18]. The glycosyltransferases involved in selectin ligand function and leukocyte recruitment are summarised in Table 1 and will be discussed briefly in the following subparagraphs.

Table 1 Glycosyltransferases and their contribution to leukocyte recruitment *in vivo*.

Glycosyltransferases	Gene name (mouse)	Function	Phenotype in the knock-out mouse
		Selectin ligand function	
Glucosaminyltransferases			
▪ C2GcT-1	*Gcnt1*	Extension of the core 2 O-glycan branch. Adds GlcNAc to GalNAc.	Impaired P-, L-, and E-selectin mediated rolling of leukocytes along the vessel wall [48–50]
Galactosyltransferases			
▪ β1,4-GalT1	*B4galt1*	Involved in synthesis of polylactosamine.	Leukocyte recruitment defect in several models of acute and chronic inflammation, impaired P-selectin binding to neutrophils [51].
Fucosyltransferases			
▪ FucT-4	*Fut7*	Transfers GDP-fucose to acceptor carbohydrates within the glycan chain.	Increased E-selectin-mediated leukocyte rolling velocities in inflamed venules [59].
▪ FucT-7	*Fut4*	Transfers GDP-fucose to terminal GlcNAc after sialylation of the terminal galactose.	Higher systemic blood leukocyte counts, P- and E- selectin mediated rolling is clearly reduced [58]. (Fut4/Fut7 double knock-out with almost complete loss of leukocyte rolling [57, 59])
Sialyltransferases			
▪ ST3Gal-IV	*St3gal4*	Formation of selectin ligands (sLe^x).	Reduced E-selectin mediated rolling. Increased E-selectin dependent rolling velocity [61]. Reduced L-selectin dependent rolling during inflammation [62].
		Chemokine receptor function	
▪ ST3Gal-IV	*St3gal4*	Putative sialylation of CXCR2	Reduced CXCR2-triggered neutrophil arrest *in vivo* and *in vitro* [41].
		Integrin activation	
▪ ST6Gal-I	*St6gal1*	Sialylation of VLA4's extracellular domain	VLA4 activation [42].

2.2.1 Glucosaminyltransferases. A well-known glucosaminyltransferase relevant for the synthesis of functional selectin ligands is core 2 β1,6*N*-acetylglucosaminyltransferase (core2 GlcNAcT-1, gene name *Gcnt1*). Core2GlcNAcT-1 is one of three glycosyltransferases that build the core 2 backbone and initiate the core 2 extension by adding a N-acetylglucosamine (GlcNAc) to the serine or threonine-bound N-acetylgalactosamine (GalNAc) *via* a β1,6 linkage (Fig. 2). In *Gcnt1*-deficient mice, P- and L- selectin-mediated leukocyte rolling during inflammation *in vivo* is strongly attenuated, while E-selectin dependent rolling is only moderately affected [48–50].

2.2.2 Galactosyltransferase. As Fig. 2 depicts, β-1,4-galactosyltransferase activity is required to form the polylactosamine $(GlcNAc\text{-}Gal)^n$ backbone which is found on selectin ligands. β1,4-GalT-I links galactose to terminal GlcNAc residues *via* a β1,4 linkage and is involved in the synthesis of the terminal capping group sLe^X [51]. *B4galt1* deficient mice showed impaired leukocyte recruitment in several models of acute and chronic inflammation [51]. In addition, *B4galt1* deficient neutrophils exhibited reduced binding to soluble P-selectin. Of note, recently *B4GALT1* deficient patients have been identified. The disease has been termed CDG IId (congenital deficiency of glycosylation-IId), but no severe inflammatory phenotype has been described so far [52, 53].

2.2.3 Fucosyltransferases. Fucosyltransferases transfer fucose from the donor GDP-fucose to acceptor molecules which can be oligosaccharides, glycoproteins or glycolipids. Two fucosyltransferases are important for the biosynthesis of selectin ligands: FucT-4 and FucT-7 [54, 55]. Both enzymes are expressed in the Golgi apparatus of leukocytes. In blood vessel endothelial cells, FucT-4 is only expressed in larger venules, whereas FucT-7 expression is mainly restricted to small postcapillary venules [56]. Either or both genes encoding FucT-4 and FucT-7 [57–59], have been deleted in mice. In *Fut4*-deficient mice the phenotype is rather mild with no defect in P- and L-selectin-dependent rolling. However, E-selectin-dependent leukocyte rolling velocities are significantly increased in inflamed venules, compared to wild-type species, suggesting a defect in E-selectin ligand function [59]. On the contrary, the effects of deleting *Fut7* are dramatic. This is already illustrated by the significantly augmented number of circulating peripheral leukocytes (a hallmark of defective leukocyte adhesion) in *Fut7*-deficient mice [60]. In addition, intravital microscopy experiments revealed that P- and E- selectin-mediated rolling are markedly reduced and L-selectin-mediated lymphocyte homing severely impaired [58, 59]. In *Fut4* and *Fut7*-double deficient mice, P-, E-selectin, and L-selectin mediated rolling of leukocytes are almost completely abolished, which clearly shows the crucial role of fucosyltransferases in the biosynthesis of active selectin ligands [57].

2.2.4 Sialyltransferases. The sialyltransferases will be discussed in section 3 of this chapter.

3 Sialic acids and sialyltransferases (ST)

Nine types of monosaccharides can be identified in the enzymatic process of glycosylation in mammals. Well-conserved biosynthetic pathways for these

monosaccharides can be found in mammals that use sugars and other precursors commonly present in the diet. In this section, the unique features of sialic acids, the biosynthetic pathway, and the three main sialyltransferase families, will be discussed.

3.1 Sialic acids

Sialic acids are mainly found as glycoconjugate decorations on the cell surface of eukaryotes and in some pathogenic microorganisms. The term "sialic acid" is commonly used to refer to the most abundant sialic acid *N*-acetylneuraminic acid (Neu5Ac), although the sialic acids comprise a family of nine-carbon sugars [63]. During evolution the sialic acid family became greatly diversified and more than 50 different forms have been isolated from a variety of natural sources [46, 64, 65]. The family comprises neuraminic acid (Neu) derivatives with a shared nine-carbon backbone (Fig. 3). They are α-keto acids and due to their net negative charge at physiological pH, plus their bulky, hydrophilic chemical structure, they can dramatically change the physical and chemical properties of the glycan structures [66].

The basic structures of sialic acids carry either a *N*-acetyl group (Neu5Ac) or a hydroxyl group at the 5-carbon, named 2-keto-3-deoxy-nononic acid (KDN) (Fig. 3). From the biosynthetic perspective (discussed in Section 3.1.1.), all sialic acids are likely to be derived from either Neu5Ac or KDN [67]. They can be subject to a diverse range of modifications. A common variation is the hydroxylation of the 5-N-acetyl group of Neu5Ac, giving *N*-glycolylneuraminic acid (Neu5Gc), which is, together with Neu5Ac and KDN, a representative member of the sialic acid family (Fig. 3). The negatively charged carboxylate group at the 1-carbon can also form a lactone ester with a hydroxyl group of a neighbouring sugar subunit. Sometimes, the 5-carbon is deacetylated to give Neu. Furthermore, Neu, Neu5Ac, KDN, and Neu5Gc can be further combined with *O*-acetyl,

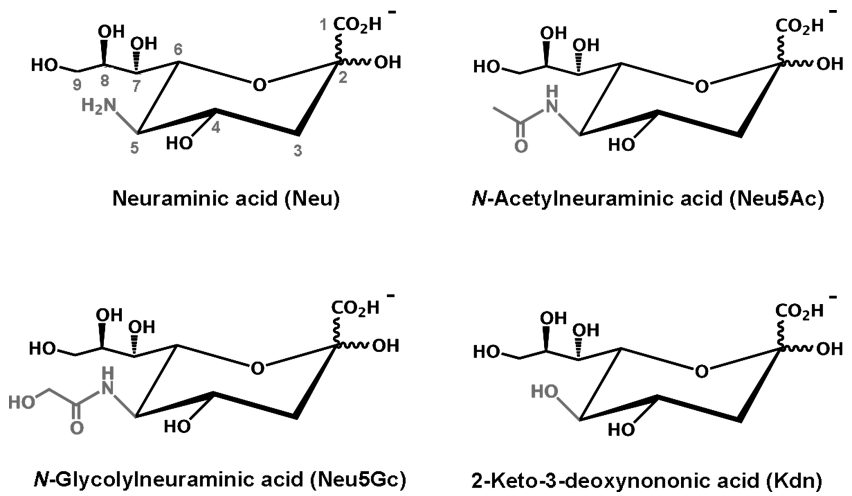


Fig. 3 Natural diversity of sialic acids. The chemical structures of four common sialic acids (*e.g.* Neu, Neu5AC, Neu5Gc, and Kdn) found in nature are shown. All sialic acids are believed to be derived from either Neu5Ac or Kdn.

O-lactyl, *O*-methyl, or *O*-sulfate groups at the different 4-, 7-, 8-, and 9-carbon OH positions [68]. It has to be noted that in mice, only two forms of sialic acids are found, namely Neu5Ac and Neu5Gc. In humans, only the predominant form of sialic acids, Neu5Ac, has been discovered.

Sialic acids are not only complex sugars because of their high diversity and net negative charge, but also because of their ability to form different linkages to the underlying carbohydrate structure. The most common linkages are α-2,3 and α-2,6 to galactose (Gal), α2,6 to *N*-acetylgalactosamine (GalNAc), and α-2,8 linkage to neighbouring sialic acids. In addition, α-2,4 and α-2,9 linkages have been identified.

3.1.1 Sialic acid biosynthesis. As for the majority of sugars in nature the basic building block of sialic acids is glucose. Glucose is converted in several steps (not discussed here) into UDP-N-acetylglucosamine (UDP-GlcNAc), which serves as the precursor of the sialic acid biosynthetic pathway. The biosynthesis of sialic acid (Neu5Ac) involves four enzymatic steps after the formation of UDP-GlcNAc (Fig. 4). The bifunctional enzyme GNE

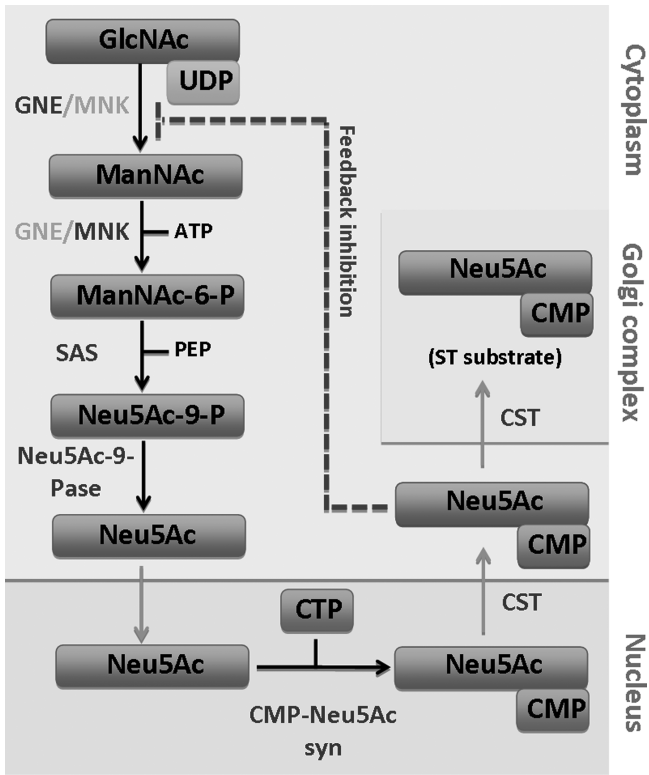


Fig. 4 The Neu5Ac (sialic acid) biosynthesis pathway. After the formation of GlcNAc (derived from glucose), the biosynthesis and activation of Neu5Ac involves five enzymatic steps. The activated sialic acid, CMP-Neu5Ac, is utilised as a substrate by sialyltransferases in the Golgi apparatus. ATP: adenosine triphosphate; CMP: cytidine monophosphate; CST: CMP sialic acid transporter (also known as Slc35A1); CTP: cytidine triphosphate; GNE/MNK: UDP-GlcNAc 2-epimerase/ManNAc-kinase; PEP: phosphoenolpyruvate; SAS: sialic acid phosphate synthase; ST: sialyltransferases; UDP: uridine diphosphate.

(UDP-GlcNAc 2-epimerase)/ MNK (ManNAc-kinase) catalyses the first two committed steps of the pathway. It firstly converts UDP-GlcNAc to N-acetyl-mannosamine (ManNAc) with its UDP-GlcNAc-2-epimerase activity. Thereafter, it catalyses the conversion of ManNAc to ManNAc-6-phospate (ManNAc-6-P) using the kinase catalytic domain of the enzyme [69, 70]. Subsequently, the sialic acid phosphate synthase (SAS) condenses ManNAc-6-P with phosphoenolpyruvate, and Neu5Ac-9-phosphate is formed [71]. Dephosphorylation of Neu5Ac-9-phosphate is catalysed by the Neu5Ac-9-phosphate phosphatase (Neu5Ac-9-Pase) and Neu5Ac is generated [72].

Prior to incorporation of the sialic acid residues into carbohydrate decoration of proteins, sialic acid has to be activated. In contrast to all other vertebrate monosaccharaides that are activated with uridine- or guanosine diphosphates (UDP or GDP, respectively), sialic acid is activated by the cytidine monophosphate (CMP) [67]. The CMP-Neu5Ac synthetase that condenses CMP and Neu5Ac to form CMP-Neu5Ac, is responsible for activating sialic acids. Though all biosynthetic steps are taking place in the cytosol, the activation of sialic acids into CMP-sialic acid takes place in the cell nucleus [46]. After activation, CMP-sialic acid is transported from the nucleus into the Golgi complex *via* a CMP-sialic acid transporter (CST, also known as Slc35A1) [73], where it is utilised as a substrate of sialyltransferases. The amount of sialic acids produced by the cell is regulated by a strong feedback inhibition mechanism of CMP-sialic acid acting on the epimerase catalytic site of the GNE/MNK complex, catalysing the first step of the biosynthesis [69].

Neu5Ac and KDN (derived from the condensation of Man-6-P, as an alternative to ManNAc-6-P for Neu5Ac) are believed to be the precursors for all other members in the sialic acid family. The different enzymatic modifications of Neu5A and KDN to form the other members of the family take place in different cellular compartments.

3.2 Sialyltransferases

The sialyltransferases represent a group of glycosyltransferases that add sialic acid residues to an acceptor glycoconjugate terminating in Gal, GalNAc, or another sialic acid. The different mammalian sialyltransferases identified are membrane-bound enzymes present in the Golgi apparatus, and together with the other glycosyltransferase residents of the Golgi apparatus, they share a type II architecture. All mammalian sialyltransferases characterised today are composed of a short N-terminal cytoplasmic domain, a transmembrane domain, a stem region, and a catalytic domain [74]. The length and amino acid composition of the catalytic domains of the different sialyltransferases are well conserved, which greatly assists in the identification of new sialyltransferases. In contrast, the stem region of the protein is less conserved and the length can vary between proteins, which is believed to influence the substrate specificity of the acceptor sugar [74, 75].

Members of the sialyltransferase family are grouped into subfamilies on the basis of the introduced linkage type (α2,3, α2,6, α2,8) between the donor sialic acid and the accepting sugar. Since 1996, a universal nomenclature is

being used, which encompasses four components. It discloses the protein family, linkage type, the acceptor sugar, and the identification number of the protein within the subgroup. As an example; the β-galactoside α2,3-sialyltransferase-three is called ST3Gal-III, where ST stands for sialyltransferase, 3 represents α2,3 linkage, Gal is the acceptor sugar galactose, and III indicates that it has been the third gene identified within this subfamily [76].

3.2.1 α2,3 sialyltransferase sub-family. The most abundant linkage found in mammalian oligosaccharides is the α2,3 linkage, and to date, six α2,3 sialyltransferases have been identified in mice and humans. The six members of the family can be subdivided into two different groups based on substrate specificity and sequence similarity. The first group consists of the ST3Gal transferases-I and -II that prefer to use core 1 O-glycans (Galβ1,3GalNAc) on glycoproteins or glycolipids as an acceptor substrate [74]. ST3Gal-I is expressed in many human tissues, including the heart, placenta, lungs, and peripheral blood lymphocytes [77]. It is responsible for the sialylation of the core 1 O-glycans on T-cells. *St3gal1*–deficient mice showed increased apoptosis of $CD8^+$ T-cells, resulting in a nearly complete deficiency of this cell type [78]. ST3Gal-II is among others expressed in the heart, liver, and lymphoid tissues of humans, and has a substrate preference for glycolipids over glycoproteins *in vitro* [79].

The second group comprises the ST3Gal transferases-III, -IV, -V, and -VI. ST3Gal-III and -IV can both utilise Galβ1,3GlcNAc and Galβ1,4GlcNAc as acceptor glycans, but their substrate preference differs. ST3Gal-III prefers the type I lactosamine (Galβ1,3GlcNAc) glycan and is expressed in the liver, testis and fetal tissues. ST3Gal-IV prefers type 2 oligosaccharides (Galβ1,4GlcNAc), which can be further processed to the tetrasaccharide sialyl Lewis X (sLe^X), as acceptor glycans and is strongly expressed in the placenta, testis, and ovary [74, 77]. Selectin binding studies in *St3gal3*- and *St3gal4*-deficient mice have shown that ST3Gal-IV is essential for the biosynthesis of active selectin ligands, whereas ST3Gal-III is not. *St3gal4*-deficient neutrophils exhibited a reduction of E- and P-selectin-mediated rolling *in vitro* when compared to wild-type neutrophils, while no difference to wild type was found for rolling of *St3gal3*-deficient neutrophils [80]. ST3Gal-V is expressed in the testis, brain, and skeletal muscles and utilises lactosylceramide (Galβ1,4Glc-Cer) as a substrate and has not been reported to contribute to leukocyte recruitment [75, 77]. In contrast, ST3Gal-VI uses type 2 lactosamines (Galβ1,4GlcNAc) on glycoproteins and glycolipids as an acceptor substrate and is involved in the synthesis of sLe^x. However, the contribution of this enzyme to leukocyte rolling and adhesion is currently unclear.

3.2.2 α2,6 sialyltransferase sub-family. α2,6 linkages are less abundant in nature than α2,3 linkages, and eight mammalian α2,6 sialyltransferases have been identified. Based on substrate specificity, two groups of α2,6 sialyltransferases can be distinguished, the ST6Gal group and the ST6GalNAc group. ST6Gals transfer sialic acids from CMP-sialic acid to the acceptor sugar galactose of glycans. Two enzymes belonging to this group have been discovered in mouse and human, ST6Gal-I and ST6Gal-II.

Both enzymes typically utilise the Galβ1,4GlcNAc structure on carbohydrate decorations of proteins as acceptor substrate, but ST6Gal-I has a broader substrate specificity [74]. The rise in expression of ST6Gal-I during inflammation is believed to be a vital part of the systemic inflammatory response [81].

The second group comprises the ST6GalNAc transferases I-VI that utilise *N*-acetylgalactosamine (GalNAc) as an acceptor sugar. They can add sialic acids to Galβ1,3GlcNAc and Galβ1,4GlcNAc motifs on glycans, but their substrate preference varies [74]. To our knowledge, none of the ST6GalNAcTs has been reported to significantly modulate innate or adaptive immune responses, and therefore ST6GalNacTs will not be further discussed.

3.2.3 α2,8 sialyltransferase sub-family. All six identified α2,8-sialyltransferases in mice and in humans link sialic acids to terminal acceptor sialic acid residues on glycoconjugates. The Siaα2,8Sia sequence is commonly found in various gangliosides, which are sialic acid containing sphingolipids that can play an important role in immunology. Sporadically, the Siaα2,8Sia sequence is seen in carbohydrate decorations of proteins as well [82]. Like the ST3 and ST6 subfamilies, the ST8 subfamily can be segmented into two groups based on substrate specificity and sequence similarity. The first group contains three enzymes, ST8Sia-I, -V, and -VI. ST8Sia-I, better known as GD3 synthase, and ST8Sia-V are both involved in the biosynthesis of gangliosides, but show opposite substrate specificities. ST8Sia-I displays strong activity towards GM3, an intracellular regulator of apoptosis in T lymphocytes, but hardly any activity towards GD1a (co-receptor in Toll-like receptor 2 signalling [83]), GT1b, and GD3. ST8Sia-V on the other hand, uses GD1a, GT1b, and GD3 as acceptor substrates, but shows low activity towards GM3 [82]. In contrast to these two enzymes, ST8Sia-VI utilises O-glycans of glycoproteins as acceptor substrates and shows low activity towards gangliosides [74].

The second group consists of three members, ST8Sia-II, -III, and -IV. ST8Sia-III uses the Siaα2,3Galβ1,4GlcNAc glycan chain on N-glycans of glycoproteins and glycolipids as the sialic acid acceptor. It is also believed to have the ability to biosynthesise Siaα2,8Siaα2,3Gal on certain O-glycans [74]. ST8Sia-II and -IV are polysialic acid synthases that catalyse polysialic acid formation on N-glycans of neural cell adhesion molecules (NCAMs) [74].

3.2.4 Sialidases. Sialidases, which are also termed neuraminidases (although this name is now generally reserved to denote sialidases from viruses [46]) are very common in nature. They can be found in many viral-, bacterial-, invertebrate-, and mammalian species. Sialidases catalyse the hydrolytic cleavage of sialic acids from glycoconjugates. The different enzymes within the sialidases family show different biochemical properties, such as substrate preference, kinetics, and binding affinities. Recently, membrane-bound sialidases on neutrophils have been described to regulate leukocyte adhesion behaviour during inflammation [84].

4 Sialylation-dependent functions of signalling and adhesion relevant molecules

4.1 Selectin ligands

Selectins were the second group of sialic-acid binding proteins discovered. Selectin - selectin ligand interactions are essential for leukocyte recruitment to sites of inflammation. Synthesis of functional selectin ligands is dependent on various glycosyltransferases which were discussed in Section 2 and 3, and are summarised in Table 1. Deficiency of one of these glycosyltransferases leads to more or less pronounced defects in selectin-mediated rolling which in turn can severely affect leukocyte recruitment during inflammation and homing of lymphocytes into secondary lymphatic tissue [18, 50, 57, 58].

To explore the exact role of sialic acids in selectin-selectin ligand interactions, mice deficient in several sialyltransferases have been generated and their phenotypes studied. Since sialic acids found in the sLe^X motif of selectin ligands are α-2,3 linked to the underlying galactose residue, the primary focus has been on studying the role of α-2,3 sialyltransferases on leukocyte recruitment [18, 85]. Three sialyltransferases have been discussed to contribute to the generation of functional selectin ligands: ST3Gal-III, ST3Ga-IV, and ST3Gal-VI. Neutrophils of *St3gal3*-deficient mice showed no difference in binding to either P-selectin or E-selectin suggesting that ST3Gal-III is not relevant in the synthesis of functional selectin ligands on neutrophils [61]. On the contrary, neutrophils of *St3gal4*-deficient mice showed a significant reduction in binding to all three selectins *in vitro* [80]. However, as pre-treatment of *St3gal4*-neutrophils with sialidase further reduced binding of the cells to the different selectins, other sialyltransferases are likely to be involved in the generation of functional selectin ligands. This is supported by additional intravital microscopic experiments in inflamed cremaster muscle venules of *St3gal4*-deficient mice where a much milder E-selectin dependent rolling defect than under *in vitro* conditions was observed with a moderately reduced number of rolling leukocytes on E-selectin and an increase in E-selectin-dependent rolling velocities [80]. P-selectin dependent rolling *in vivo* was not affected at all in the absence of ST3Gal-IV [80]. These results suggest that for E- and P-selectin ligand activity, other sialyltransferases might compensate for the loss of ST3Gal-IV [61]. Besides E- and P-selectin, L-selectin also contributes to leukocyte rolling and leukocyte recruitment during inflammation *in vivo* [86, 87]. Using *St3gal4*-deficient mice, it has been shown that L-selectin-mediated rolling of neutrophils on inflamed cremaster muscle venules was almost completely absent [62]. Since L-selectin-mediated rolling of neutrophils requires leukocyte-expressed PSGL-1, these findings suggest that PSGL-1 binding activity towards L-selectin is strongly dependent on ST3Gal-IV [62].

As described in Section 3, the sialyltransferase ST3Gal-I uses core 1-decorated O-glycans (Galβ1,3GalNAc) as acceptor substrate and therefore is not likely to be involved in the synthesis of the sLe^X motif on core 2-carrying O-glycans [74]. But, interestingly, *St3gal1*-deficient mice showed that the lack of ST3Gal-I activity is accompanied by a strong up-regulation of sLe^X-containing core 2-branched O-glycans, leading to better binding of

the selectins to its selectin ligands which indicates that ST3Gal-I is competing with the glucosaminyltransferase core 2-GlcNAcT-1 for the same acceptor substrate [78].

The sialyltransferase ST3Gal-VI has been connected to contribute to the biosynthesis of sLe^X. However, ST3Gal-VI is not highly expressed in hematopoietic cells, including leukocytes [88]. The exact role of ST3Gal-VI in selectin - selectin ligand interactions *in vivo* is currently unknown, but *St3gal6*-deficient mice have been generated recently. Investigations in those mice will soon elucidate how ST3Gal-VI contributes to selectin ligand activity *in vivo* [46].

4.2 Sialylation and chemokine receptors

Posttranslational glycosylation is not only vital for selectin ligand activity, but also plays a major role in the proper function of chemokine receptors [18, 41]. About a decade ago, the first studies identifying a role of post-translational sialylation on chemokine receptor binding to chemokines have been published [89]. In these first studies, a canine thymocyte cell-line was used that was transfected with either wild-type chemokine receptor 5 (CCR5) or one of the various CCR5 mutants in which specific serine or threonine residues had been exchanged with alanine to remove putative O-glycosylation sites [89]. It was shown that binding of CCR5 to chemokine was strongly dependent on a sialic acid-carrying O-glycan, which was linked to serine 6 at the N-terminus of CCR5 [89].

Subsequently, the important role of posttranslational glycosylation in chemokine-triggered firm leukocyte arrest *in vivo* was published [41]. Using intravital microscopy, the function of ST3Gal-IV on CXCR2-mediated leukocyte adhesion was investigated. These studies revealed that leukocyte arrest in cremaster muscle venules of wild-type mice was rapidly and strongly induced after systemic injection of CXCR2-binding chemokines (CXCL1 or CXCL8), whereas in *St3gal4*-deficient mice leukocyte adherence to the vessel wall was dramatically reduced following systemic injection of CXCL1 or CXCL8. Additional studies showed that neutrophils isolated from *St3gal4*-deficient mice exhibit reduced binding to the chemokines CXCL1 and CXCL8 [41]. These studies strongly suggest that ST3Gal-IV-mediated sialylation is critical for CXCR2-induced firm neutrophil adhesion during inflammation [18,41]. Further studies are warranted aiming to explore the potential role of ST3Gal-IV in regulating the function of other chemokine receptors in neutrophils and other leukocyte subpopulations (monocytes, dendritic cells, eosinophils, and lymphocytes) [85].

4.3 Integrins and their ligands

The importance of integrins in the leukocyte recruitment cascade to inflammatory sites is illustrated by several genetic disorders in humans, in which defects in integrin expression and/or functions occur. Genetic defects in adhesion are collectively called leukocyte adhesion deficiencies (LADs). The most common form of LAD, called LAD-I, is characterised by mutations in the *ITGB2* gene, encoding the leukocyte-expressed β_2 integrin subunit. The different mutations can lead to loss or reduced presence of β_2 integrins on the cell surface and defects in β_2 integrin signalling.

LAD-I patients suffer from impaired wound healing and recurrent bacterial infection [11, 60].

Although posttranslational glycosylation of integrins has been known for quite some time, its functional role for leukocyte adhesion only begins to emerge [90]. Woodard-Grice recently showed that the sialyltransferase ST6Gal-I negatively regulates the activity of the integrin VLA-4 on monocytes by sialylation of the β_1 integrin's extracellular domain. Interestingly, the activity of ST6Gal-I is controlled by the β-site APP-cleaving enzyme 1 (BACE1) secretase, which down-regulates ST6Gal-I through cleavage and shedding of ST6Gal-I from the cell surface. In another report, loss of α2,6-sialylation was demonstrated to cause significant changes in the conformation of β_1 integrin with reduced accessibility of the ligand binding region suggesting an important regulatory function of α2,6 sialylation on integrin function [91].

4.4 Siglecs

At least three classes of lectins involved in leukocyte recruitment have been described in the literature: C-type lectins (selectins), siglecs (I-type lectins), and galectins [92, 93]. Two of them, selectins and siglecs rely on sialic acid containing ligands. The role sialic acids play on proper binding of selectins to its ligands has been described above. In this paragraph the role of siglecs and their ligands will be discussed in the context of leukocyte recruitment.

Siglecs (Sia-recognising immunoglobulin superfamily lectins) belong to the I-type lectin family as they fulfil the two requirements necessary for I-type lectins: 1) they are members of the immunoglobulin-superfamily of proteins and 2) they recognise terminal sugars (sialic acids in case of siglecs) on glycoconjugates. Siglecs are type I transmembrane proteins expressed by all cell types of the innate and adaptive immune system [94–96]. Depending on the presence or absence of a cytoplasmic signalling domain, they can be involved in cell-cell interactions and/or intercellular communication [97]. In addition, siglecs are used as receptors by some pathogenic microorganisms in a process known as molecular mimicry, where the bacterium decorates its glycolipids on the cell surface with Neu5Ac to promote interactions with siglecs that function as immune receptors [95].

Structurally, siglecs contain a transmembrane domain, a variable number of extracellular C2-set immunoglobulin domains, and an N-terminal V-set-immunoglobulin domain that contains a conserved arginine residue for sialic acid binding. On the basis of sequence similarities and evolutionary conservation, siglecs can be divided into two groups. The first group comprises the CD33-related siglecs (CD33/siglec3, siglec-5 till -11, siglec-14 in humans, CD33/siglec-3, siglec-E, -F, -G, and -H in mice) and the second group consists of siglec-1 (sialoadhesin or Sn), siglec-2 (CD22), MAG (siglec-4), and siglec-15. Distinctive siglecs are expressed by different cell-lines of the hematopoietic or immune system. For example, siglec-1 can be found on macrophages and to a lesser degree on granulocytes [98], CD33 (siglec-3) is expressed on neutrophils, monocytes and macrophages [99].

Siglecs from both groups recognise sialic acid residues on *N*- or *O*-linked carbohydrate decorations of proteins or on glycolipids, distinguishing also the sialic acid-linkage type they bind to (α2,3, α2,6, or α2,8). Due to the high

level of sialic acids on cell surfaces, siglecs are mostly bound to sialic acid residues in cis, which leads to a masking effect of the sialic acid binding site on siglecs under homeostatic conditions. This is true for most siglecs. An exception is siglec-1 which has 17 Ig-like domains and extends well beyond the cell surface (roughly 50 nm) [100]. So it can be assumed that the sialic acid binding site of siglec-1 is not masked and might be able to interact with sialic acid containing ligands on other cells (trans-interaction). As siglec-1 does not contain any cytoplasmic signalling domain, it may therefore serve as a mere adhesion molecule. This is indirectly supported in a report by Jiang *et al.* [101]. Using siglec-1-deficient mice, they found an attenuated initial disease progression in an autoimmune uveoretinits model suggesting a pro-inflammatory function of siglec-1 in this T-cell driven inflammation model [101]. Although it is still unclear, how siglec-1 exerts this pro-inflammatory function, there is evidence that siglec-1 binds to glycoproteins (including PSGL-1) containing terminal α2,3-linked sialic acids, which might be an indication that siglec-1 could participate in leukocyte recruitment into inflamed tissue. Future studies including *in vivo* imaging will be needed to further elucidate the exact molecular mechanisms by which siglec-1 regulates immune cell recruitment.

5 Conclusion and outlook

In the last few years, posttranslational sialylation has emerged as an important modification of adhesion molecules and adhesion relevant molecules impacting leukocyte recruitment into tissue during an inflammatory response. While α2,3 sialylation has been well-established as crucial modulator of selectin ligand function and hence leukocyte capture and rolling, recent findings also reveal a role of α2,3 sialylation on chemokine receptor function and chemokine-triggered firm leukocyte arrest. Furthermore, several investigations provide now strong evidence that α2,6 sialylation of β_1 integrin negatively regulates β_1 integrin activation stretching the contribution of posttranslational glycosylation from the initial rolling step to firm leukocyte arrest and postarrest modifications. These new findings emphasize the importance of carbohydrate decoration on the regulation of an inflammatory response and also offer new interesting therapeutic approaches in the treatment of acute and chronic inflammatory diseases.

References

1 H. Dutrochet, Wilhelm Engelmann Verlag, 1824.
2 R. Wagner, Leopold Voss Verlag, 1839.
3 E. C. Butcher, *Cell*, 1991, **67**, 1033.
4 K. Ley, *et al.*, *Nat Rev Immunol*, 2007, **7**, 678.
5 T. A. Springer, *Cell*, 1994, **76**, 301.
6 G. S. Kansas, *Blood*, 1996, **88**, 3259.
7 D. Vestweber and J. E. Blanks, *Physiol Rev*, 1999, **79**, 181.
8 M. Sperandio, *FEBS J*, 2006, **273**, 4377.
9 A. Zarbock and K. Ley, *Am J Pathol*, 2008, **172**, 1.
10 A. Zarbock, *et al.*, *Blood*, 2011, **118**, 6743.
11 C. L. Abram and C. A. Lowell, *Annu Rev Immunol*, 2009, **27**, 339.
12 M. Phillipson, *et al.*, *J Exp Med*, 2006, **203**, 2569.

13 L. Stephens, *et al*., *Curr Biol*, 2008, **18**, R485.
14 M. Sperandio, *et al*., *Methods Enzymol*, 2006, **416**, 346.
15 Z. Liu, *et al*., *J Exp Med*, 2010, **207**, 2975.
16 D. Vestweber, *Immunol Rev*, 2007, **218**, 178.
17 R. D. Cummings and R. P. McEver, C-type lectins, In: *Essentials of Glycobiology*, 2nd edition, 2009.
18 M. Sperandio, *et al*., *Immunol Rev*, 2009, **230**, 97.
19 L. Xia, *et al*., *J Clin Invest*, 2002, **109**, 939.
20 M. Sperandio, *et al*., *J Exp Med*, 2003, **197**, 1355.
21 A. Hidalgo and P. S. Frenette, *Immunity*, 2007, **26**, 753.
22 M. Matsumoto, *et al*., *J Immunol*, 2005, **175**, 8042.
23 P. Sundd, *et al*., *Biorheology*, 2011, **48**, 1.
24 E. B. Finger, *et al*., *Nature*, 1996, **379**, 266.
25 B. T. Marshall, *et al*., *Nature*, 2003, **423**, 190.
26 T. S. Olson and K. Ley, *Am J Physiol Regul Integr Comp Physiol*, 2002, **283**, R7.
27 A. Viola, and A. D. Luster, *Annu Rev Pharmacol Toxicol*, 2008, **48**, 171.
28 N. Hogg, *et al*., *Nat Rev Immunol*, 2011, **11**, 416.
29 F. Ye, *et al*., *Blood*, 2012, **119**, 26.
30 R. Evans, *et al*., *J Cell Sci*, 2009, **122**, 215.
31 M. T. Huang, *et al*., *Blood*, 2006, **107**, 4721.
32 D. Frommhold, *et al*., *Blood*, 2010, **116**, 841.
33 S. Nourshargh, *et al*., *Nat Rev Mol Cell Biol*, 2010, **11**, 366.
34 W. A. Muller, *Annu Rev Pathol*, 2011, **6**, 323.
35 J. D. Marth, *Nat Cell Biol*, 2008, **10**, 1015.
36 R. A. Dwek, *Chem Rev*, 1996, **96**, 683.
37 K. Marino, *et al*., *Nat Chem Biol*, 2010, **6**, 713.
38 H. H. Freeze, *Nat Rev Genet*, 2006, **7**, 537.
39 A. Varki and J. B. Lowe, Biological Roles of Glycans, In: *Essentials of Glycobiology*, 2nd edition, 2009.
40 J. D. Marth and P. K. Grewal, *Nat Rev Immunol*, 2008, **8**, 874.
41 D. Frommhold, *et al*., *J Exp Med*, 2008, **205**, 1435.
42 A. V. Woodard-Grice, *et al*., *J Biol Chem*, 2008, **283**, 26364.
43 G. Reuter and H. J. Gabius, *Cell Mol Life Sci*, 1999, **55**, 368.
44 K. Uchimura and S. D. Rosen, *Trends Immunol*, 2006, **27**, 559.
45 K. Ohtsubo and J. D. Marth, *Cell*, 2006, **126**, 855.
46 A. Varki and R. Schauer, Sialic Acids, In: *Essentials of Glycobiology*, 2nd edition, 2009.
47 L. Tu and D. K. Banfield, *Cell Mol Life Sci*, 2010, **67**, 29.
48 M. Sperandio, *et al*., *J Immunol*, 2001, **167**, 2268.
49 M. Sperandio, *et al*., *Blood*, 2001, **97**, 3812.
50 L. G. Ellies, *et al*., *Immunity*, 1998, **9**, 881.
51 M. Asano, *et al*., *Blood*, 2003, **102**, 1678.
52 B. Hansske, *et al*., *J Clin Invest*, 2002, **109**, 725.
53 M. Guillard, *et al*., *J Pediatr*, 2011, **159**, 1041.
54 B. Ma, *et al*., *Glycobiology*, 2006, **16**, 158R.
55 D. J. Becker and J. B. Lowe, *Glycobiology*, 2003, **13**, 41R.
56 C. M'Rini, *et al*., *J Exp Med*, 2003, **198**, 1301.
57 J. W. Homeister, *et al*., *Immunity*, 2001, **15**, 115.
58 P. Maly, *et al*., *Cell*, 1996, **86**, 643.
59 W. Weninger, *et al*., *Immunity*, 2000, **12**, 665.
60 S. Hanna and A. Etzioni, *Ann N Y Acad Sci*, 2012,
61 L. G. Ellies, *et al*., *Proc Natl Acad Sci USA*, 2002, **99**, 10042.

62 M. Sperandio, *et al*., *Eur J Immunol*, 2006, **36**, 3207.
63 A. Varki, *Glycobiology*, 1992, **2**, 25.
64 A. Varki, *Nature*, 2007, **446**, 1023.
65 S. Kelm and R. Schauer, *Int Rev Cytol*, 1997, **175**, 137.
66 R. Schauer, *Curr Opin Struct Biol*, 2009, **19**, 507.
67 T. Angata and A. Varki, *Chem Rev*, 2002, **102**, 439.
68 R. Schauer, *Glycoconj J*, 2000, **17**, 485.
69 O. T. Keppler, *et al*., *Science*, 1999, **284**, 1372.
70 S. Hinderlich, *et al*., *J Biol Chem*, 1997, **272**, 24313.
71 S. M. Lawrence, *et al*., *J Biol Chem*, 2000, **275**, 17869.
72 P. Maliekal, *et al*., *Glycobiology*, 2006, **16**, 165.
73 M. Eckhardt, *et al*., *Proc Natl Acad Sci USA*, 1996, **93**, 7572.
74 S. Takashima, *Biosci Biotechnol Biochem*, 2008, **72**, 1155.
75 A. Harduin-Lepers, *et al*., *Biochimie*, 2001, **83**, 727.
76 S. Tsuji, *et al*., *Glycobiology*, 1996, **6**, v.
77 H. Kitagawa and J. C. Paulson, *J Biol Chem*, 1994, **269**, 1394.
78 J. J. Priatel, *et al*., *Immunity*, 2000, **12**, 273.
79 V. Giordanengo, *et al*., *Eur J Biochem*, 1997, **247**, 558.
80 L. G. Ellies, *et al*., *Blood*, 2002, **100**, 3618.
81 J. C. Jamieson, *et al*., *Comp Biochem Physiol B*, 1993, **105**, 29.
82 S. Tsuji, *J Biochem*, 1996, **120**, 1.
83 S. Liang, *et al*., *J Biol Chem*, 2007, **282**, 7532.
84 C. Feng, *et al*., *J. Leukoc. Biol*., 2011, **90**, 313.
85 M. Sperandio, *Ann N Y Acad Sci*, 2012,
86 M. L. Arbones, *et al*., *Immunity*, 1994, **1**, 247.
87 K. Ley, *et al*., *Journal of Experimental Medicine*, 1995, **181**, 669.
88 T. Okajima, *et al*., *J Biol Chem*, 1999, **274**, 11479.
89 N. Bannert, *J Exp. Med*., 2001/12/3, **194,** 1661.
90 M. E. Janik, *et al*., *Biochimica Biophysica Acta*, 2010, **1800**, 545.
91 Y. Liu, *et al*., *Proteins*, 2008, **73**, 989.
92 G. R. Vasta, *Nat Rev Microbiol*, 2009, **7**, 424.
93 T. K. Dam and C. F. Brewer, *Glycobiology*, 2010, **20**, 270.
94 L. G. Baum and P. R. Crocker, *Immunol Rev*, 2009, **230**, 5.
95 P. R. Crocker, *et al*., *Nat Rev Immunol*, 2007, **7**, 255.
96 A. Varki and T. Angata, *Glycobiology*, 2006, **16**, 1R.
97 L. Nitschke, *Immunol Rev*, 2009, **230**, 128.
98 P. R. Crocker, *et al*., *EMBO J*, 1994, **13**, 4490.
99 D. Simmons, *et al*., *Nature*, 1988, **331**, 624.
100 A. Varki and P. R. Crocker, I-type lectins, In: *Essentials of Glycobiology*, 2nd edition, 2009.
101 H. R. Jiang, *et al*., *J Immunol*, 2006, **177**, 2258.

Glycoengineering of protein-based therapeutics

Sandrine Donadio-Andréi,† Chloé Iss,† Nassima El Maï, Valérie Calabro and Catherine Ronin*
DOI: 10.1039/9781849734769-00094

Introduction

N-linked glycosylation is widely considered as the most complex modification of proteins in the eukaryotic kingdom [1]. N-glycans play a crucial role in the biosynthesis, folding, stability, biological activity and metabolic clearance of the proteins to which they are attached [2, 3]. They control multiple key biological processes including cell adhesion, receptor activation, signal transduction and endocytosis [3–5]. In addition, they are highly versatile structures that have been shown to be associated with in a wide panel of pathological events, especially in infection, cancer, immunological diseases, metabolic and neurodegenerative disorders [6–9].

The biosynthetic machinery which is in charge of assembling oligosaccharides onto proteins requires a large repertoire of enzymes [10]. Expression of glycosylation enzymes appear to vary among organisms, from cell type to cell type within a single organism as well as with cellular environment and stage. The composition and the linkage(s) of monosaccharide residues in the oligosaccharide chains are thus related to the species. As a result, glycoproteins expressed in animal cells could be therefore strongly immunogenic if administered in humans missing these structures [11]. The nature of the oligosaccharide chain profoundly affects stability and pharmacokinetic properties of most glycosylated therapeutic proteins [3, 12, 13]. N-glycosylation is therefore a major concern for the development of biopharmaceutical proteins.

At present, over 70% of the marked protein-based therapeutics are glycosylated proteins and several hundreds of these products are being developed. They consist in monoclonal antibodies (mAb) and mAbs-based products, proteins for replacement therapies (hormones, growth factors, enzymes) and cytokines. It is important for protein-based therapeutics to be produced with a glycosylation optimal for their function and similar to their human counterparts to minimize risk of immunogenicity that may compromise treatment efficacy. Another issue for the production of biopharmaceutical glycoproteins is the glycoform profile [14]. The inherent properties of the biosynthetic machinery leads to multiple N-glycoprotein variants for which the glycan pattern can undergo substantial variation with changes in cell culture conditions. A major challenge for biopharmaceutical industry is to control this extensive heterogeneity and obtain batch-to-batch consistency for optimum bioefficacy and reliable safety evaluation of their products [15, 16].

Siamed'Xpress, Aix-Marseille Univ., 3 Pl.V.Hugo, 13331-Marseille, France.
E-mail: catherine.ronin@siamedxpress.com
†equal contribution.

© The Royal Society of Chemistry 2012

Protein may display variable glycans and therefore, their exposed surfaces differ among glycoforms depending on the oligosaccharide structure. This may affect antibody recognition [17] and may be an important issue in assaying or tracing a drug candidate. It is highly desirable to generate the right glycoform (when generating either antibodies against a human glycosylated target for diagnostics or for developing therapeutic antibodies). In diagnostics, the desired glycoform(s) of a protein marker should presumably be identical to the circulatory form(s) of the analyte to ensure epitope similarity. Therapeutics antibodies must be selected against the appropriate target glycovariant(s) to display optimal pharmacological activity. Additionally, N-glycosylation changes are associated with diseases such as cancer and inflammation and measuring diseased glycoform(s) could be of major diagnostic and prognostic significance [18, 19].

Currently, therapeutic glycoproteins are predominantly produced in rodent expression systems such as Chinese Hamster Ovary (CHO), mouse NS0 or mouse SP2/0 cell lines because their glycosylation machinery is the closest to that of human cells [20, 21]. However, these systems are not perfect, some sugar components are missing and instead, some rodent-type residues are added and this significantly increases the likelihood of immunogenic reactions. Efforts are currently being made to generate biotherapeutics with ''humanized'' glycosylation to reduce immunogenicity, optimize bioactivity and generate a consistent glycoform profile with limited heterogeneity [22, 23]. Here, we review the issues associated with protein glycosylation for therapeutic and diagnostic applications and the main approaches to date to address these issues.

1 Impact of N-glycosylation on therapeutic proteins

1.1 N-glycosylation achieved by heterologous expression systems

N-glycosylation is not directly encoded by the genome. It is a mechanism conserved across eukaryotes and represents the most widely distributed glycan addition in proteins [1]. It requires the tripeptide consensus site Asn-X-Ser/Thr where X could be any amino acid except proline; the oligosaccharide chain is attached to the asparagine residue. The synthesis occurs co-translationally in the lumen of the endoplasmic reticulum and post-translationally in the Golgi apparatus. It appears tightly regulated and the rules that control this highly complex process are not yet fully understood. First, the glycoprotein sequence and the structure surrounding the consensus sequence is thought to determine glycan addition. Occupied N-glycosylation sites are composed of a common core pentasaccharide constituted by two GlcNAc and three mannose residues and then, depending on the organism, the cells and/or the substrate availability, different sugars are being added to form bi-, tri- or tetraantennary glycans with various terminal residues as shown in Fig. 1. A series of glycosyltransferases and glycosidases are involved in the process. They act sequentially to build up the oligosaccharide chain and compete with each other for both the available protein acceptors and nucleotide donors. Their expression is regulated in function of cell type, development stage, cell environment and culture conditions and their combined actions deliver a wide repertoire of glycoforms within each cell type.

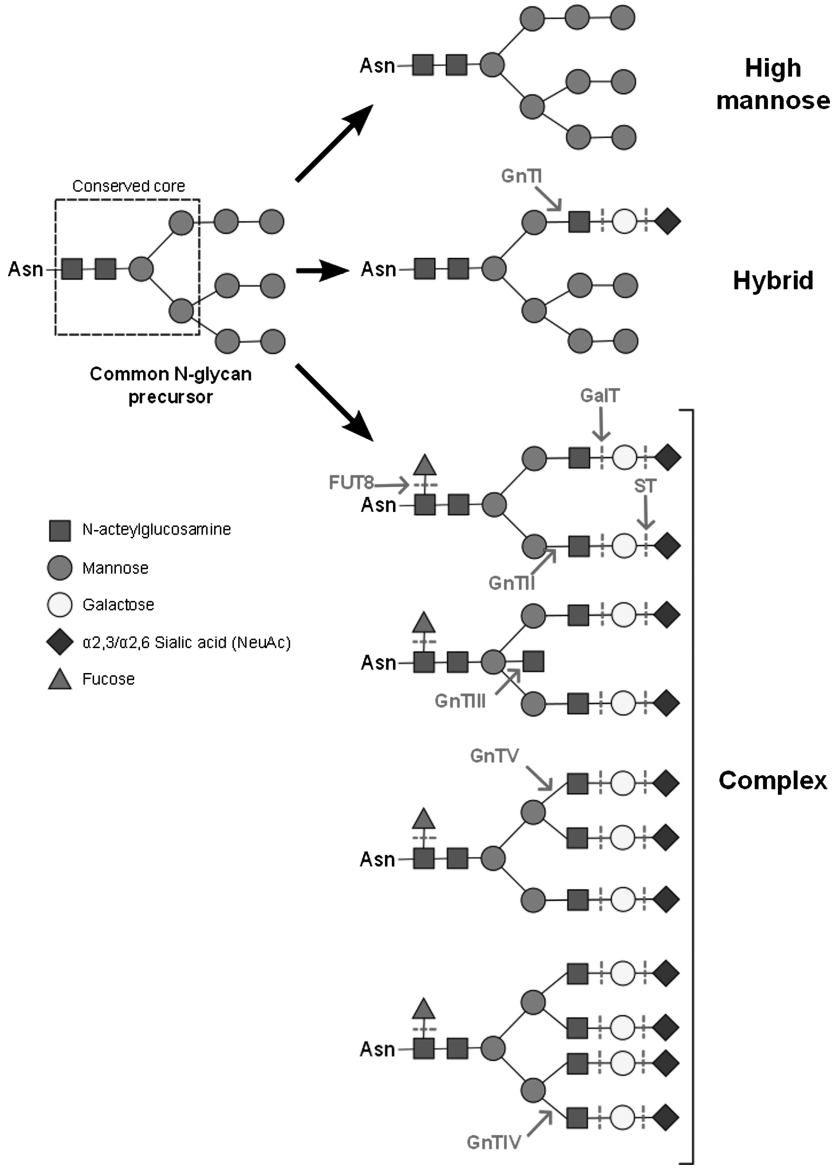


Fig. 1 Typical structure of the common mannose-rich precursor of all human N-glycans and the three classes of N-linked glycans. Dashed bars depict the elongation of the pentasaccharide core with complex antennae, generating a wide panel of N-glycans. Abbrevations: GnT: N-acetylglucosaminyltransferase, FUT8: α1,6 fucosyltransferase, ST: sialyltransferase, GalT: galactosyltransferase.

Several expression systems are available for production of therapeutic proteins: bacteria, yeasts, plants, insect cells, mammalian cells and transgenic animals. So far, mammalian cells remain the system of choice for biotherapeutic glycoprotein production because of its reliability [19]. The use of other systems as an alternative to mammalian cells has been mainly hindered by their limited capacity to produce glycoprotein with human-like

glycosylation [21]. Plant and insect glycans contain β1,2-xylose and core α1,3-fucose which elicit strong IgG and IgE responses when administered to humans [24, 25]. In addition, plant cells poorly transfer outer galactose (Gal) and do not add sialic acid, leaving essentially pauci-mannosidic or hybrid glycans which are rapidly cleared by a serum mannan-binding lectin (MBL) following injection. Yeast cells have not been intensively used for protein drug production because they deliver a strongly immunogenic O-glycosylation. Yeast cells share however similar early biosynthetic pathway for N-glycan synthesis with mammals but are unable to carry out complex glycosylation. Antibodies expressed in yeast show defects in effector function due to altered glycosylation pattern [13]. These cells have being intensively engineered with the knockout of genes to eliminate yeast-specific glycosylation, and introduction of heterologous genes to replicate the branching pattern and terminal N-glycosylation of human proteins [26]. They are probably the most advanced technology that may compete with mammalian cells for glycoprotein production.

While much effort has been invested in engineering human-like glycosylation pathways in plants, insect and avian cells [22], production of therapeutic glycoproteins is still currently performed using mammalian cell systems and is expected to continue over the next decade. Over the years, mammalian cells have been manipulated to produce large quantities of biopharmaceutical proteins and are continuously subject to novel improvements [27]. The most widely used systems are CHO, NS0 and SP2/0 cell lines in various engineered derivatives. Most approved biopharmaceuticals are produced in CHO and CHO-derived cell lines. A quite attractive aspect of CHO cells is their proven track record for producing therapeutic proteins which definitely eases the process of achieving drug approval. Importantly, a wide knowledge and expertise have been accumulated over more than a hundred of gene products. Figure 2 illustrates carbohydrate heterogeneity of such recombinant preparations: branching heterogeneity is illustrated with human TSH **(a)** and terminal heterogeneity with human erythropoeitin (hEPO) (**b**) Indeed, CHO cells display the wide repertoire of N-linked glycans but do not select biantennary structure in TSH as in the native hormone or highly sialylated glycans like in tri-and tetraantennary glycans in native EPO.

In 2011, the complete sequencing of CHO (K1 variant) genome was achieved and this will undoubtedly lead to further optimization for higher productivity and higher product quality [30]. However, CHO glycosylation machinery differs from that of human and most often produces a random branching pattern of glycosylation. More generally, rodent cells usually add terminal Gal α1,3-Gal and provide a low degree of sialylation which may include N-glycolylneuraminic acid (NeuGc); both sugars are absent in human glycoproteins. Nonetheless, since their early approval for the production of recombinant hEPO and human tissue plasminogen activator (tPA), CHO cells have provided several successful therapeutic proteins and will remain the industry's favourite expression system, at least in the near future. Therefore, they are definitively the best cell candidates for testing new glycoengineering strategies for drug optimization.

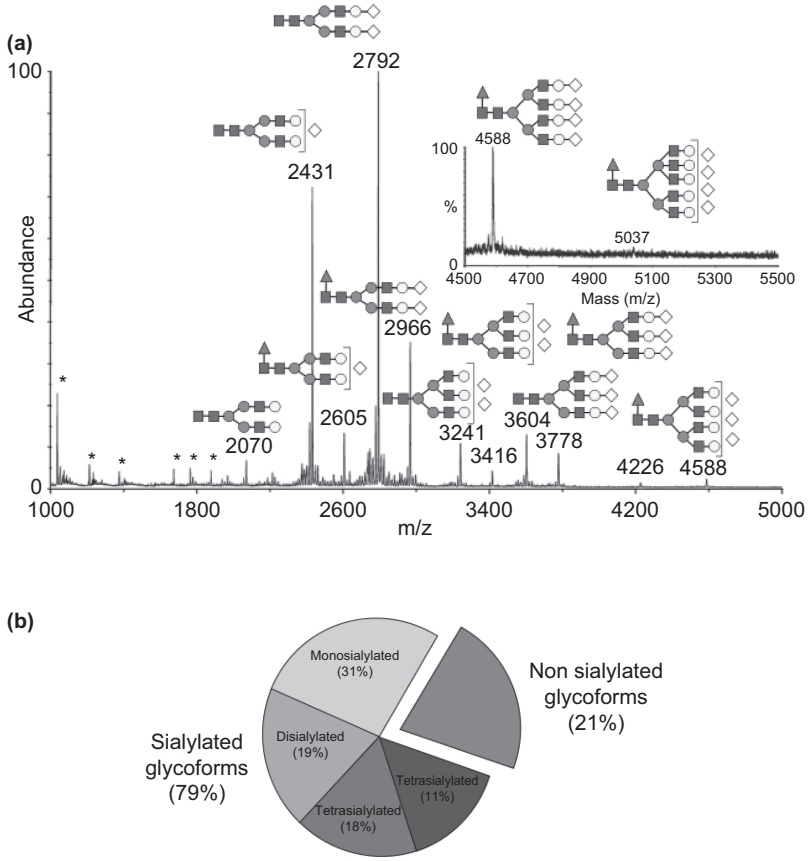


Fig. 2 **(a)** MALDI-TOF–MS of N-glycans released from recombinant human TSH glycans produced in mammalian cells. The composition of the most abundant glycoforms are indicated above each ion. The oligosaccharides are highly heterogeneouss with biantennary tri-, and tetra- as well as pentaantennary oligosaccharides with or without inner fucose and terminal sialic acid. Square represents N-acetylglucosamine (GlcNAc), triangle fucose, circle mannose, circle galactose, diamond sialic acid to symbolize NeuAc/Gc [28]. **(b)** Relative amount of sialylated glycans in recombinant human EPO produced in CHO cells [29].

1.2 Impact of glycans on drug bioavailability

The residence time of proteins in blood is determined by various parameters that characterize each protein and affect its susceptibility to different elimination pathways. These parameters such as protein charge, size, hydrophobicity or presence of epitopes mediate protein removal from the circulation and, therefore, play a major role in therapeutic protein efficacy. N-glycans may affect all these parameters. Oligosaccharide chains are often large and bulky and alter the hydrodynamic radius of the protein which can protect from renal clearance. They also affect the structure, the chemical and physical stability of the protein carrier. Glycosylation has been shown to improve protein resistance to pH, chemical and thermal denaturation, and aggregation [3, 31–33]. There is also ample evidence that oligosaccharides protect from proteolysis by covering amino acids and thus preventing access to proteases [34].

Although presence of oligosaccharide can improve protein half-life in serum, the type of glycans can also promote receptor-mediated metabolic clearance. The metabolic pathways which recycle or eliminate circulatory glycoproteins involve lectin receptors, essentially the liver asialoglycoprotein receptor (ASGP-R) and the mannose receptors *i.e.* the mannan-binding lectin (MBL) and the macrophage mannose receptors. Each of them can promote glycoprotein uptake and degradation within minutes. Mannose receptor can also bind terminal N-acetylglucosamine (GlcNAc) and cause faster protein clearance [35]. For this reason, particular attention has been paid to the presence of high mannose glycoforms during drug production. All heterologous systems ranging from yeast to mammals, including insect and avian cells primarily synthesize mannose-rich glycosylation as it is the precursor form of the glycan repertoire. It often happens during fermentation that those mannose residues are not sufficiently trimmed and induce a liability for drug efficacy.

Asialoglycoproteins with free β-linked Gal residues bind to the ASGP-R the liver receptor and are cleared from serum with remarkably high capacity [36]. Capping with terminal sialic acids actually masks Gal residues, preventing binding to the lectin receptor and liver uptake resulting in prolonged circulatory half-life. Therefore, terminal glycosylation and more especially protein sialylation is a major challenge during bioproduction because it affects the pharmacokinetic properties of the drug. Recently, it has been shown that 6-but not 3-linked sialic acid may also modulate this uptake in mouse [37]. Indeed, loss of activity due to hepatic clearance has been observed for several protein drugs, the most striking example being the case of erythropoeitin [38]. The level of sialylation does not influence circulatory lifespan solely because of the recognition by the ASGP-R but also because of it impacts on protein properties. Sialic acids are highly hydrophilic negatively charged residues that alter the pI and solubility/aggregation of the protein. While the sialic acid content of recombinant hEPO (rhEPO) has been directly to the increase in circulatory residence time [39, 40], such property has also used to design more soluble erythropoietin, Darbepoetin alfa, a hyperglycosylated analogue of hEPO that contains two additional N-linked carbohydrates [41]. This product displayed a greater overall negative charge which extends its circulating half-life by a factor of 3 compare to rhEPO in animal models.

Sialylation therefore appears to be a key parameter to modulate the *in vivo* efficacy of therapeutic glycoprotein by improving pharmacokinetic profiles. Interestingly, a market surveillance conducted by the Swiss Agency for Therapeutic Products (Swissmedic) has recently released a study aiming at comparing the major oligosaccharide profiles of 16 commercially available therapeutic antibodies [42]. They report that N-carbohydrate structures were predominantly truncated biantennary glycans, with core fucose. The extent of galactosylation of these products is limited and they are lacking sialic acid. Hence, enhancing the sialic acid content in therapeutic protein preparations has the potential to reduce drastically the dosage and frequency of administration required in for clinical benefit. The influence of sialic acid on the aggregation propensity of proteins might be important as well for formulation and storage as it would prevent aggregation and permit longer storage.

1.3 Impact of glycans on glycoprotein biopotency

The first therapeutic protein produced in recombinant mammalian cells to obtain market approval was the recombinant human tissue Plasminogen Activator. Recombinant tPA is a serine protease which converts plasminogen into plasmin and induces clot lysis. It is indicated for embolic or thrombotic stroke. tPA is one prominent example in which biological properties of a glycoprotein have been shown to depend on the structural features of its oligosaccharides. Two major glycoforms of recombinant tPA are produced by CHO cells. One glycoform has three glycosylation sites occupied whereas the second one has only two [43]. Both tPA bioactivity and bioavailability vary upon site occupancy. The highly glycosylated glycoform displays lower clot lysis activity and the lack of one site of glycosylation affects clearance rate [44, 45].

Another well documented example of how glycans modulate protein structure, stability and function is represented by the immunoglobulin family, the largest class of therapeutic proteins [13, 46]. Most therapeutic antibodies are of the IgG1 subtype and contain a single N-linked glycan attached to Asn297 in each Fc domain. The nature and composition of the oligosaccharide chains carried by the Fc fragment are known to play a major role in antibody folding and stability. They maintain the spatial organization of the CH2 domains in an "open conformation" of the Fc dimer structure, a conformation promoting interaction with Fc receptors and binding of the C1q protein of the complement [47, 48]. The biological mechanisms that are activated by immune complexes depend on IgG Fc oligosaccharides as well [46]. As in other therapeutic glycoproteins, N-glycosylation also governs biophysical and pharmacokinetic properties of antibodies and can induce immunogenic reactions.

In serum-derived human IgGs, most of the oligosaccharides are of the biantennary complex type with a common pentasaccharide core and variable additions of fucose, bisecting N-acetylglucosamine (GlcNAc) and terminal GlcNAc, Gal, and sialic acid. However, alterations of endogenous IgG glycosylation have been reported in various immune diseases. In patients with rheumatoid arthritis (RA), both IgG Fc galactosylation and sialylation are decreased [49]. IgG Fc glycosylation has been found to modulate inflammatory activities of such IgGs [50]. Even though the underlying mechanism is not yet understood and still under debate, this activity is used in intravenous immunoglobulin (IVIg) treatment for many autoimmune and systemic inflammatory diseases [51]. Several mechanisms could be at the origin of the anti-inflammatory activity. Increased level of asialo, agalacto, core-fucosylated biantennary glycoforms may contribute to inflammation by activating the lectin pathway. Indeed agalactosylated glycoforms of aggregated immunoglobulin G induce association with the MBL and contribute to RA pathology [52]. Likewise, using a mouse model, Kaneko *et al*., has demonstrated that sialylated human IgGs have a greater anti-inflammation activity than non sialylated IgGs [53]. They have shown that α2,6 sialylated N-glycans of IgG can bind a C-type lectin, SIGN-R1 or in a mouse model. This interaction initiates an anti-inflammatory cascade leading to the upregulation of cell surface expression of the inhibitory Fc receptor, FcyRIIb on inflammatory cells and thereby attenuating

autoantibody-initiated inflammation [54, 55]. Recently, it has been suggested that the α2,6 sialic acid residues can be directly involved in the induction of apoptosis of B cells *via* its binding to the sialic acid-binding receptor CD22 thus contributing to the anti-inflammatory activity of IVIg [56].

While native glycosylation of human serum IgGs is sufficient for immune effector functions, marketed recombinant monoclonal antibodies are being optimized to enhance their therapeutic bioefficacy especially in the field of cancer or infectious diseases [22, 57]. Indeed, IgG glycoforms display differences in their ability to trigger effector functions. Various carbohydrates like core fucose, bisecting GlcNAc, sialic acid, and mannose residues affect the binding of IgG to the Fcγ receptors and thereby influence Antibody-Dependent Cellular Cytotoxicity (ADCC) activity [54]. The absence of core fucose as well as the presence of bisecting GlcNAc have been found to significantly increase binding of IgG Fc to FcyRIIIa [58–60]. ADCC is thus increased through longer residence time on FcyRIIIa receptor and out competition with serum IgG. Removal of terminal Gal reduces Complement-Dependent Cytotoxicity (CDC) by approximately 50% relative to CDC function of unmodified mAbs [61, 62]. IgGs with mannose-rich glycans are part of the glycoforms of mAbs produced in CHO cells [63] and show low affinity for the complement C1q protein and low CDC activity compared to complex type glycoforms [64], indicating that their presence in an antibody preparation will decrease the efficacy of the whole product. Alternatively, the presence of terminal sialic acid residues in the IgG1 Fc oligosaccharides reduces binding affinities to the FcyRIIIa receptor, thereby decreasing *in vivo* cytotoxicity [53].

Many new antibody-based therapeutics are currently under development and it is anticipated that these glycoproteins will represent 30% of the marketable drugs in the next five years. Thousands of patents have been issued so far to protect antibody engineering. There is a considerable interest in manipulating both their polypeptide and oligosaccharide chains to maximize their anti-infectious, anti-cancer or anti-inflammatory activities.

1.4 Impact of glycans on drug immunogenicity

Another important cause of low bioavailability resides in antigenic reactions to the drug. Immunogenicity, in addition to pose safety concerns, affects the efficacy of treatment since antibodies raised by the patients against a therapeutic protein may mediate neutralization of the drug and its elimination. Clearance of Infliximab (trade name Remicade®), an anti-TNFα monoclonal antibody, is 2.7 times higher and its half-life 34% weaker in the presence of anti-Infliximab antibody [65].

Immunogenicity first originates from the absence of glycans [11]. It has been observed that lack of glycosylation site occupancy may induce polypeptide misfolding and create new epitopes. For example, recombinant human granulocyte-macrophage-colony-stimulating (GM-CSF) factor has been shown to be more immunogenic when produced in *E. coli* [66]. Anti-recombinant human GM-CSF antibodies react with peptide regions which are normally protected by glycosylation in the natural protein. Indeed, such immunogenicity can lead to serious medical issues as the

protein drugs have a human primary structure. Some patients who were administered recombinant erythropoietin from CHO cells have developed antibody against their own erythropoietin. Studies have shown that in such reactive patients, anti-EPO antibodies were directed against the amino acid and not the glycans [67, 68]. Carbohydrates added on rhEPO by CHO cells are different from native EPO and may alter EPO conformation to create epitopes recognized as foreign by human immune system.

Secondly, non-human glycans elicit an anti-sugar immune response that also greatly reduces blood circulation time. High NeuGc level is found in proteins produced in NS0 and SP2/0 cells and to a much lesser extent in CHO cells. This induces immune responses following administration to patients and causes rapid removal of therapeutic proteins from the circulation [69–71]. The half-life reduction of therapeutic antibodies in clinical studies has likewise been associated with the presence of anti-Gal antibodies. Glycoproteins produced in murine cells mostly carry glycans terminated in Gal-α1,3-Gal residues, which are absent in proteins produced by human and CHO cells [72]. Since anti-α1-3-Gal antibodies are abundant in humans and constitute 1% of circulating antibodies in humans [73], glycoforms carrying this type of residue are rapidly neutralized and cleared. Cetuximab (trade name Erbitux), an anti-EGFR mAb approved by the FDA in 2004 for the treatment of metastatic colorectal cancer, head and neck cancers, has been produced in SP2/0 cell line. 22% of patients treated with Cetuximab displayed severe hypersensitivity reactions. Immunological testing revealed that patient antibodies against Cetuximab were specific for Gal-α1,3-Gal present on the Fab portion of the therapeutic antibody [74]. Alternatively, terminal sialic acid has been shown to prevent recognition by T- and B-cell receptors and offer an opportunity to generate anti-drug neutralizing antibodies [75].

Even though in most instances, the clinical benefit derived from therapies using biologicals outweighs the risk of developing immunogenic reactions, immune recognition of protein drugs is increasingly becoming a key issue for effective long-term treatment. It should be further emphasized that this concern is more particularly relevant for glycoproteins indicated for chronic inflammatory diseases such as RA because such patients already react to their own immunoglobulins [22].

1.5 Impact of glycans on protein folding and epitope expression

Exposure of peptidic epitopes at the surface of a protein may be modulated by glycosylation in different ways. Based on their volume, branching and polarity, glycans can alter protein conformation. Glycoprotein drugs may fail in binding their target *in situ* if their glycosylation state has altered their biological activity. Similarly, changes in glycosylation of a diseased target may significantly diminished recognition by a therapeutic antibody that has been designed toward a recombinant glycoform pattern which fails to represent the native target the natural target. In this respect, lack of epitope similarity between natural and recombinant or even among recombinant preparations of a given glycoprotein antigen has been observed [76]. It is therefore worth producing glycoproteins with a defined glycoprofile adapted to their biomedical application.

However, it has been observed over the various generations of diagnostic tests that routine immunoassays of glycosylated analytes were discordant among manufacturers [77]. Regulatory agencies have been sensitized to current discordances and in the context of cost reduction, now request that assays should be harmonized and standardized to improve accuracy and reproducibility of biological tests. About a hundred of glycoprotein markers are routinely measured in laboratory medicine and there is a clear need for early diagnostic and/or decision-making of personalized drug treatment in many diseases. The International Federation of Clinical Chemistry has elected the measurement of Thyroid Stimulating Hormone (TSH) to initiate this challenging task as this protein-based marker is the 1st protein marker prescribed worldwide. Over the 25 past years, many studies have reported inter-and intra-immunoprocedure discordances in TSH measurements [78–87]. Recent studies from our group revealed that these discrepancies can be attributed to TSH glycosylation. The pituitary hormone exists as a mixture of a dozen of GalNAc-sulfated glycoforms which substantially differ in their biological and immunological properties [88]. In hypothyroid patients, serum TSH is highly enriched in sialic acids as hypothyroidism develops and immunoassays lack accuracy as most antibodies bind the circulating analyte better than the extractive antigen [76, 89, 90]. Essentially, fucose and sialic acid are modulating the expression of epitopes specific for blood TSH: a 2-3 fold increase in antibody binding was observed when α2,6-linkage sialic acid or when fucose is lacking. These findings led us to conclude that glycosylation is not acting on epitope expression by steric hindrance but rather that selective carbohydrates behave as molecular effectors of protein conformation to deliver a variable glycoform pattern. Interestingly, it has been established that new assays can be developed to optimize the measurement of a given glycoprotein in blood whether it is a marker or a drug.

Glycoform specific assays may be of invaluable use in providing accurate tests and assisting clinical trials. Since preparations from human sources are no longer recognized as available sources of protein antigens, it is crucial to be able to produce recombinant glycoforms which resemble as closely as possible to the protein to detect. Glycoengineering is expected to fill the gap currently existing between the recombinant product and its circulating (pathological) forms.

1.6 Impact of glycans on production, safety and regulatory aspects

Living cells react quite sensitively to even the slightest changes in their environment. The combinatorial nature of N-glycosylation provides a molecular and structural basis to these changes. Many factors, from the nutrient solution to the design of the equipment affect glycan composition. For instance, Van Berkel *et al*. have observed changes in N-glycosylation from batch production in shaker to fed-batch production in bioreactor [63]. As a consequence, it is quite difficult to keep a given glycoform profile constant from batch-to-batch. This renders the production highly challenging as it tolerates few to none changes in the manufacturing process during the lifetime of a drug. Regulatory Agencies such as US Food and Drug Administration and European Medicines Agency are concerned about the heterogeneity of preparations and the consistency of bioproduction.

Since each glycoform brings its own pharmacokinetic, pharmacodynamic and efficacy profile, the activity of resulting mixture is hardly predictable. Therefore, glycoprotein heterogeneity soon became a major concern as it poses problems in the quantitative evaluation of the drug principle. The best approach so far thus appears to minimize glycosylation heterogeneity and keep it constant throughout development and manufacturing to ensure safety control and maintenance of both potency and pharmacokinetic properties. The authorities are currently raising new standards that biomanufacturers must follow to address these issues and analyze the glycoprofile of their products [91–93]. Manufacturers are requested to characterize the carbohydrate structure and degree of heterogeneity of their products in quantitative terms in relation to the manufacturing process. It is expected that they demonstrate comparability between preclinical and clinical batches and perform comparability assessment to support changes in manufacturing process during the development of their product and post-approval.

Glycosylation is also an issue in the evolution of biosimilars [15]. Even with full information about the process, it is not easy to replicate the manufacturing process to obtain a product similar if not identical to the original innovator drug. Because of unavoidable differences in bioproduction, regulators request a high level of scrutiny and often additional clinical data proving that the new drugs are at least as safe and effective as the originals. In extreme cases, it can lead to a whole new marketing authorization procedure. Genzyme recently had to file another application to show that Myozyme, a lysosomal glycogen-specific enzyme indicated for the treatment of patients with Pompe disease, manufactured at a location different from the original is as safe and effective than the first batch.

Overall, the past decades have unravelled the remarkable complexity of the glycan repertoire. Extensive characterization of expression systems revealed that their glycosylation machinery is species specific and differ from human cells. When produced in these systems, human protein-based therapeutics often showed a quite variable glycosylation profile that substantially impact their biological performances. In most cells, the same glycosylation enzymes are working on a wide panel of endogenous acceptors as well as on an overexpressed protein with distinct efficiency. We are still lacking a full understanding on the mechanisms which regulate *in vivo* glycosylation so tightly. To date, variable heterogeneity is regarded as a mean for each protein to accomplish several functions with variable intensity. By altering both the folding and/or fate of a given protein, glycans allows a versatile setting of its specific interactions along with its biological activity and may thereby also govern variable response at the target tissue(s). For example, non fucosylated IgG Fc present in human serum have increased ADCC activity while the presence of sialic acid provides anti-inflammatory properties to antibodies. In antibodies and also in other proteins, we know now that both inner and terminal glycosylation appear as key molecular modulators of glycoprotein structure and activity. It is therefore important to integrate this recent knowledge in designing glycoprotein production. Engineering cells to optimize therapeutic proteins and/or bioprocesses should address the manipulation of the glycoform profile to reach better detection, efficacy and/or safety of the desired drug.

2 Glycosylation optimization

Cell lines and bioproduction today are no longer developed solely on the basis of cell growth and production yield but have now integrated the need of controlling N-glycosylation. Culture processes and culture media are evaluated to identify the conditions that promote optimal glycoform profile and reproducibility. Mammalian cell lines may be genetically modified for producing a more homogeneous and non-immunogenic glycosylation. Alternatively, postproduction glycosylation by *in vitro* enzymatic reactions may be used to release undesired terminal sugars and/or complete the oligosaccharide chain enzymatically. Development of new therapeutic glycoproteins will therefore include customization of glycosylation, a more efficient glycan synthesis during bioproduction, and extensive characterization of the glycosylation pattern of the product.

2.1 Tuning cell culture conditions to control glycosylation

Since the approval of the first recombinant protein, biomanufacturing underwent tremendous progress over the past decades [27]. The improvements include optimization of cell lines, development of high-performing media free of animal-derived components as well as advances in bioreactor engineering and purification processes. In addition, fundamental studies on metabolic pathways have shown that culture conditions such as pH and temperature greatly impact the glycosylation profile. Glycoengineering is therefore being accomplished first by selecting the clone that produces the optimal glycoprofile among many overproducer clones for each given protein [63]. Once the clone is selected, tuning of the culture conditions has to be performed to favor production of the appropriate glycoforms. Multiple parameters of cell culture processes and media composition have been shown to influence cell glycosylation machinery and therefore impact glycoform profile [94]. They include medium components for instance metal ions, butyrate, glucose and Gal concentration, cell culture environment as well as growth rate and cell productivity [95]. Screening and tuning of these conditions need to be performed on a case-by-case basis since glycosylation is protein-specific and because a variable glycan pattern cannot be easily predicted based on the use of the expression system [63].

Culture conditions affect altogether occupancy of glycosylation sites, branching and completion of oligosaccharide chains [95–97], presumably by altering the expression of several glycosylation enzymes which consist in various nucleotide sugar synthases and transporters, glycosidases and glycosyltransferases. Expression profiling of 24 N-glycosylation genes in CHO cells has shown that 21 were significantly up- and down- regulated during culture phases inducing change in IFN-γ N-glycosylation [98]. Addition of nutriments to manipulate nucleotide sugar metabolism has also been used to improve complete oligosaccharide chain synthesis. Feeding IFNγ-producing CHO fed-batch cultures with glutamine and glucose increased the content in hybrid and high mannose glycans. Adding galactose was found to facilitate a more galactosylated N-glycan profile [99] and subsequent addition of sialic acid in recombinant antibodies. Supplementation of CHO culture with N-acetylmannosamine (ManNAc), a

precursor of sialic acid, resulted in increased sialylation of IFN-γ [100]. Other components may alter the glycosylation profile by limiting the presence of different sugars. Decrease in NeuGc content of a therapeutic antibody expressed by CHO cells can be achieved by simply adding N-acetylneuraminic acid (NeuAc) to the culture media [71]. Yet, the same experiment using murine myeloma cell lines did not produce similar data, probably because the uptake of NeuAc is poorly efficient in this cell type.

Inhibitors have been also used to reduce activities of glycosyltransferases and favor the synthesis of specific glycans. Zhou *et al.* used the potent inhibitor of α-mannosidase, kifunensine, to produce non-fucosylated, oligomannosidic structures [101]. Scientists at Seattle Genetics modified monoclonal antibody carbohydrates by adding to CHO cultures small fucose analogs. The fucose analogs inhibit GDP-mannose dehydratase (GMD), the first enzyme acting in the *de novo* synthesis of GDP fucose, and therefore extensively deplete the pool of intracellular GDP-fucose. The resulting antibodies showed a significant reduction in their carbohydrate fucosylation [102].

Exoglycosidase activities are other important factors that affect the glycoform profile of a recombinant glycoprotein post-secretion. During bioproduction, cell death and lysis may release cytosolic sialidases and galactosidases which degrade the oligosaccharide chains of newly synthesized proteins [103]. Gu et al have shown a remarkable loss of sialic acids in IFN-γ along with concomitant decrease in CHO cell viability throughout long-term batch cultivation [104]. Neuraminidase degradation can be minimized by adding an inhibitor, a substrate analogue *i.e* 2,3-dehydro-2-deoxy-NeuAc, to the culture medium of cells expressing a recombinant antibody [105].

Modification of processes, optimization of media and addition of supplements can thus exert some control over the protein glycosylation pathway, favor some glycoforms over others and reduce heterogeneity to some extent. It is worth noticing that the control over glycosylation is limited to the pattern constructed by the cell type and there is still a need for further improvement. An increased knowledge on the regulation of N-glycosylation enzymes and cell responses during fermentation will certainly benefit the control of the overall bioproduction process.

2.2 Postproduction modifications

Another approach for remodeling glycosylation has been performed by *in vitro* enzymatic modification. This is a postproduction process using soluble enzymes to reshape glycan structures *in vitro* and obtain a more homogeneous preparation of recombinant glycoprotein. Cloning techniques have made available more than 30 glycosyltransferases from various sources [106] and deletion of their membrane anchor allowed isolation of soluble enzymes and further use in enzymatic reactor [107]. It was observed that *in vitro* completion of glycans definitely improved performances of therapeutic glycoproteins, especially their stability and pharmacokinetics. The process is not cost-effective as it requires quantities of both enzymes and nucleotide sugar donors and necessitates additional purification steps. Yet, commercial therapeutic proteins have been generated through this

method. Cerezyme, a lysosomal glycoprotein enzyme analog of the human enzyme β-glucocerebrosidase prescribed for the treatment of Gaucher disease required postproduction chemical modification for full function. Cerezyme is produced by recombinant DNA technology using mammalian cell culture in CHO cells. The oligosaccharide chains have been modified postproduction to terminate in phosphorylated mannose residues in order to facilitate mannose receptor-mediated uptake by macrophages [108].

N-glycan remodeling *in vitro* has been generally be successfully achieved at the research level [54, 61, 109, 110], especially for terminal glycosylation in order to increase duration in blood. Such synthesis requires sequential addition of these residues by several enzymes and the process further involves rather complex purification procedures to isolate the final product from unreacted material. Raju *et al*. have remodeled the oligosaccharide chain of a Fc domain fused to the extracellular portion of the TNF receptor (TNFR-Fc), produced in CHO cells, by using a β1,4-galactosyltransferase (GalT) and an α2,3sialyltransferase (ST3). Terminal sialylation was increased by approximately 23% preferentially in the TNFR domain rather than in the Fc domain [109]. More recently, Anthony *et al*. could sialylate immunoglobulin-derived Fc *in vitro* based on the use of α2,3 – (ST3) or α2,6 sialyltransferase (ST6) following GalT treatment. Interestingly, ST6 efficiency was found to be more effective than ST3, suggesting that glycan accessibility may affect the efficiency of these sialyltransferases (STs) [54]. Other enzymes such as recombinant N-acetylglucosaminyltransferase III (GnTIII) which increases the content in bisecting GlcNAc have been cloned and tested successfully for *in vitro* enzymatic modification of Fc domain and found to improve effector function [61].

Altogether these studies demonstrated that N-glycans can be remodeled by using engineered glycosyltransferases. These enzymes possess a good potential for achieving glycan customization but the methodology is limited for an industrial biomanufacturing process especially at a GMP scale. The extra purification steps as well as the cost of goods required by this approach are definitely unappealing. Manufacturing transferases and activated sugars at GMP grade is costly and it is probably not a sustainable solution for production at a commercial scale. In addition, sugar addition may also be limited by a lack of accessibility of the glycan in fully folded proteins and can thus never reach completion as during biosynthesis and thus the needed improvement in biopotency.

2.3 Glycoengineering of rodent cells

A good alternative to enzymatic remodelling of glycoproteins resides in the development of knockout and knockin cells with glycosyltransferases. The increasing demand for therapeutic proteins led to the development of genetic approaches to obtain high expressing cell lines [27] and robust bioproduction. Cell engineering is routinely employed to modify cellular processes of host cells and increase protein secretion rate [111]. Identification and cloning of glycosyltransferase genes have provided new tools for further manipulating expression systems [112]. There is currently intensive work in this field to control sugar addition and improve glycoform profile and generate protein with better performances.

2.3.1 Engineering core fucosylation. ADCC enhancement is of a great interest for cancer treatments with antibodies and may have major application in antibody drugs against infectious diseases as well [113]. The level of ADCC depends on the phenotype of the polymorphic FcγIIIa-receptor, limiting the efficiency of the original antibody to about 30% of the patient population. ADCC-enhanced antibodies are expected to treat new patient subpopulations presenting a FcyIIIa receptor mutant. They will also allow a reduction of effective dosage for all patients and thereby reduce side effects. As described above, altering Fc-linked glycans impacts significantly IgG effector functions through core fucose, bisecting GlcNAc and/or branching. The development of glycoengineered cell lines to produce glycoform of antibodies with increased ADCC function is currently under intensive exploitation and the first glycoengineered antibody has gained approval in march 2012.

The first approach to increase ADCC in antibody-based therapeutics is to equip cells with the enzyme that adds bisecting GlcNAc to N-linked glycans. The enzyme GnTIII that catalyzes the transfer of a β1,4-GlcNAc residue to the core β-mannose residue residue is lacking in CHO, NS0 and Sp2/0 cell lines. Davies *et al*. from Idec Pharmaceuticals/Biogen co-expressed rat GnTIII in a CHO/DG44 production cell line along with the anti-CD20 Rituximab (trade name Rituxan® and MabThera®) and showed an increase in up to 20-fold of ADCC activity compared to that of Rituxan® produced in wild type CHO cells. An expression level of the enzyme was sufficient to ensure glycan branching in the antibody but did not affect cell growth in suspension in serum- and protein-free cell culture medium nor antibody production [59]. Scientists at GlycArt overexpressed the rat GnTIII enzyme, engineered with different localization domains of Golgi-resident enzymes in HEK293-EBNA cells [114] (Fig. 3a**)**. The antibodies produced were enriched in bisected non-fucosylated oligosaccharides. GnTIII activity appears to inhibit core fucosylation suggesting that lack of fucose rather than the presence of bisected GlcNAc enhanced ADCC. The authors suggested that the synthesis of bisected oligosaccharides is due to an efficient competition of chimeric GnTIII with endogeneous α1,6-FUT (FUT8) and mannosidase II rather than a high expression level of the branching enzyme itself. The delocalization of GnTIII within early Golgi compartments lead to a more efficient competition with the processing enzymes. The technology named GlycoMAb® was adapted in a CHO cell line to produce the GA101 antibody currently in clinical trials [115].

In contrast to most N-linked glycoproteins, over 90% of antibodies produced in CHO cells are fucosylated. Since non fucosylated IgG1 antibodies have been shown to display enhanced ADCC for anti-cancer antibodies [116], it is highly desirable to produce antibodies without fucose. Mammalian cells including CHO cells attach a fucose residue to the innermost GlcNAc residue of the N-linked oligosaccharide *via* a FUT8 fucosyltransferase, which transfers fucose from GDP-fucose in a α1,6 linkage. Yamane-Ohnuki *et al*. have generated stable CHO/DG44 cells producing non fucosylated antibodies by disrupting both FUT8 alleles. The non fucosylated form showed *in vitro* a $\sim$100 fold increase in the efficacy of FcγRIII-mediated ADCC compared to the fucosylated product [117].

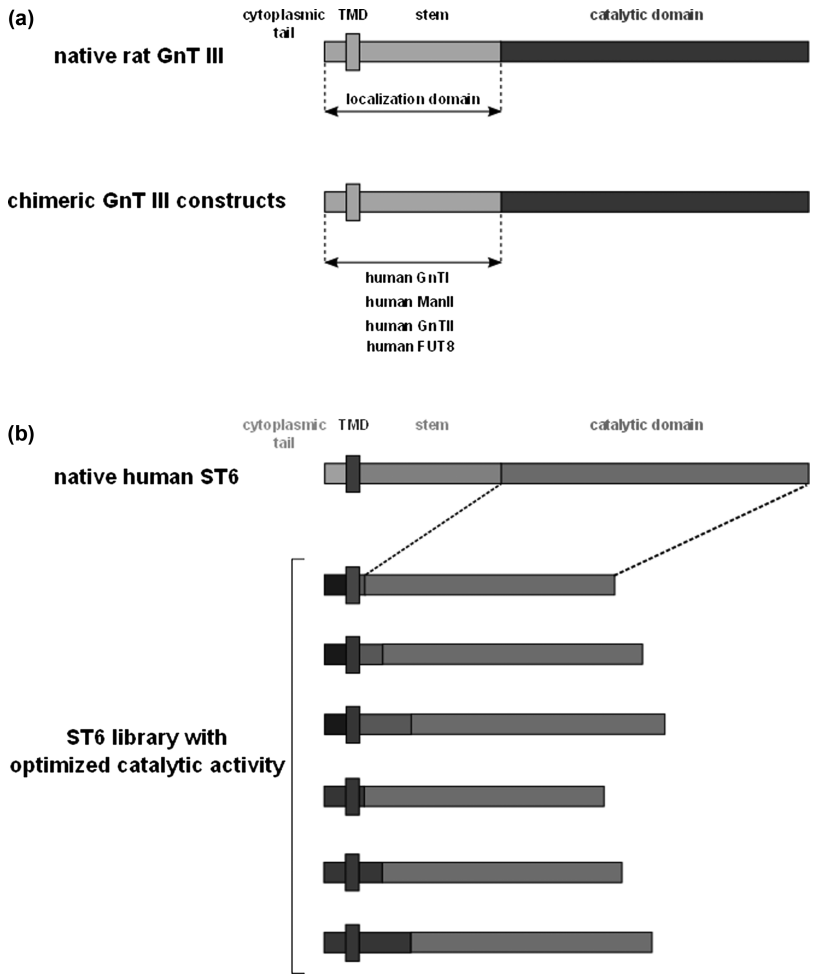


Fig. 3 (**a**) Schema of wild type rat GnTIII and GnTIII chimeric enzymes constructed by substituting the natural N-terminal domain by the localization domains of various Golgi-resident glycosylation enzymes (114] (**b**) Schema of minigenes delivering non natural human ST6 enzymes with optimized catalytic activity. The library is based on the combinatorial assembly of cytoplasmic, transmembrane and stem regions of various glycosyltransferases as reported in the footnote.

As shown by scientists at Genentech, enhanced ADDC activity of the non fucosylated trastuzumab (Trade name Herceptin®) observed *in vitro* translates into *in vivo* efficacy for solid tumor targeting [118].

Other genes to produce non-fucosylated antibodies from mammalian cells could be targeted such as GDP-mannose 4,6-dehydratase (GMD), GDP-6-deoxy-D-lyxo-4-hexulose reductase (RMD) as well as the GDP-fucose transporter (GFT) [119–122]. Scientists at Kyowa have shown that GMD knocked-out CHO/DG44 cells is a viable strategy for the production of very efficient non fucosylated therapeutic antibodies. The GlymaxX technology proposed by ProBioGen AG uses the expression of the prokaryotic enzyme RMD in CHO cells to interrupt the fucose biosynthesis pathway by capturing a reactive sugar intermediate [121]. A team at Chugai

Pharmaceutical disrupts the GFT gene by sequential homologous recombination in CHO/DXB11 cells. The engineered cells produced more than 93% of non-fucosylated anti-CD317 mAb [122].

Conversely, sialylated Fc glycans have been shown to negatively influence ADCC antibody function by decreasing affinity of IgG Fc with the FcγIIIa receptor. Despite all the advantages brought by the presence of sialic acids as terminal glycan residues, it may be, desirable to reduce terminal addition of sialic acid in some particular cases which require blockade of inflammatory cascades. Naso *et al.* have generated a CHO cell line expressing both sialidase A and recombinant antibody. Expression in the condition medium of sialidase A was sufficient to remove sialic acids from secreted antibodies [123].

Glycoengineered therapeutic antibodies with improved efficacy are considered as very promising. Another advantage to the increased potency of these products is the absence of adverse effects since their glycan pattern is very similar to natural human serum IgGs [22]. Engineered CHO cell lines to produce such glycoengineered proteins are now commercially available.

2.3.2 Recent progress in engineering sialylation. Although terminal glycosylation is definitely a key step for the lifespan of glycosylated proteins, little progress has been obtained in engineering sialylation in animal cell lines over the past years. Sialic acid is naturally missing in yeast, algae, plant, insect and avian cells. When expressed in yeasts glycoengineered to achieve complex glycosylation [124], significant transfer of sialic acid has been obtained in selected strains but fully sialylated glycans has not been demonstrated. It is also worth noting that sialic acid is poorly expressed in mammalian cells which terminate their glycans mostly with α-linked Gal. Indeed, CHO cells have originally been selected because their α1,3 galactosyltransferase is mutated in its catalytic domain and cannot add α-linked Gal to glycoproteins . These cells however do express a ST3 activity and are able to transfer sialic acid as NeuAc and NeuGc to complex glycans in the α-2,3-linkage. However, the level of sialic acid in glycoproteins produced in these cells may be artefactually low because of endogeneous sialidases released in the cell culture medium during production. Several groups have tackled this issue by developing CHO cell lines in which the sialidase gene has been silenced [125–127]. A significant reduction in asialoforms has been observed. It is worth noting also that the human retinoblastoma cell line PER C6 has been reported to share a similar glycan pattern as CHO and NSO cells for antibody production [128]. Overall, it therefore appears that to date, no expression system routinely used for drug production is able to achieve α2,6-sialylation.

This type of glycosylation is often regarded as human-like glycosylation because, contrary to other mammals, circulatory human proteins are prominently terminated in α2,6-linked sialic acid. It may indeed be considered as a signal which prevents the protein of interest from uptake from elimination tissues and from recognition by the immune system because sialic acid is crucial for absorption, serum half-life and metabolic clearance. Considering that more than 70% of the marketed proteins are glycosylated,

the availability of cells expressing an efficient α2,6-linked sialyltransferase has long been known as a major goal to achieve.

Engineering efficient sialylation activity is not a straightforward process. Several strategies based on genetic engineering of rodent cell lines have been followed to increase sialic acid content of therapeutic proteins. They include addition of glycan acceptor sites in the protein drug, inhibition of sialidase activities and overexpression of STs with improved efficacy [129].

To increase the number of sialylation acceptor sites, Fukuta *et al*. increased the content in multi-antennary sugar chains on IFN-γ by stable overexpression of N-acetylglucosaminyltransferase IV and V (GnTIV and GnTV) [130]. However, sialylation did not increase proportionally with the number of acceptor sites probably because of insufficient activity of either STs and/or CMP-sialic acid transporters. In further studies, the authors further extended the method by overexpressing mouse ST3 and/or rat ST6 in CHO cells stably transfected with GnTV. Sialylation activities was increased by approximately 80% to reach 61.2% in α2,3-and α2,6-linked sialic acid content, indicating a limited availability of the sialylation enzymes in CHO cells [131].

Increasing sialylation acceptor sites can also be achieved by increasing galactosylation efficiency. Weikert *et al*. have decreased TNK-tPA and TNF-Fc glycoform heterogeneity by increasing sialylated glycoforms using CHO cell lines overexpressing human β1,4-GalT and/or ST3. Over 90% of the resulting glycans were found to be capped with sialic acid [132].

Efficient sialylation does not seem to rely on the presence of a STs only as CHO transfection with ST3 alone hardly changed the content in sialic acid in rhEPO compared to production in CHO cotransfected with β1,4-GalT and ST3 [29]. Rather, increased sialic acid transfer appears also related to the availability of the CMP-sialic acid, the donor substrate for STs within the Golgi lumen. Coexpression of CMP-sialic acid transporter (CMP-SAT), CMP-sialic acid synthase together with human ST3, greatly improved sialylation of rhEPO produced in CHO cells [133]. Another approach to increase availability of CMP-sialic acid is to increase its biosynthesis in the cytoplasm. The enzyme UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase (GNE/MNK) is a key enzyme in the biosynthesis process of CMP-sialic acid but is negatively regulated by CMP-sialic acid [134]. Overexpression of a mutant form of GNE/MNK found in sialuria disorder has been shown to increase the intracellular pool of sialic acids and consequently increase sialylation of rhEPO because the feedback mechanism was abolished [135]. Highly sialylated therapeutic glycoproteins were indeed obtained through overexpression of sialuria-mutated GNE/MNK, ST3 and CMP-SAT [136].

Since the drug approved cell lines are all missing a ST6 enzyme [137], several groups have tried to equip rodent cells with ST6 enzymes. The first introduction of rat ST6GalI in CHO cells was reported by Lee and colleagues in 1989 [137]. Rat or human ST6-expressing CHO cells were successively generated and tested for various therapeutic glycoproteins production [138–141]. In all cases, the content in sialic acid did not rise tremendously and the ratio of α2,6-to α2,3-linked sialic acid never exceeded

50%. It was concluded at the time that the recombinant ST6 enzymes poorly competes with the endogenous ST3 for the same donor and acceptors. Such ratio also indicated a lack of completion in terminal glycosylation and therefore precluded further industrial development.

Earlier work from our group showed that sialic acid transfer in the 6 position may also be limited because of an exquisite specificity of the ST6 enzyme towards its protein acceptors. The wild type ST6GalI was found to exhibit high preference for triantennary glycans and reduced activity on biantennary and tetraantennary glycans. Deleting the enzyme in its N-terminal half was shown to abolish such specificity and improve sialic acid transfer by 30-fold [142]. When transfected into CHO cells, such truncated

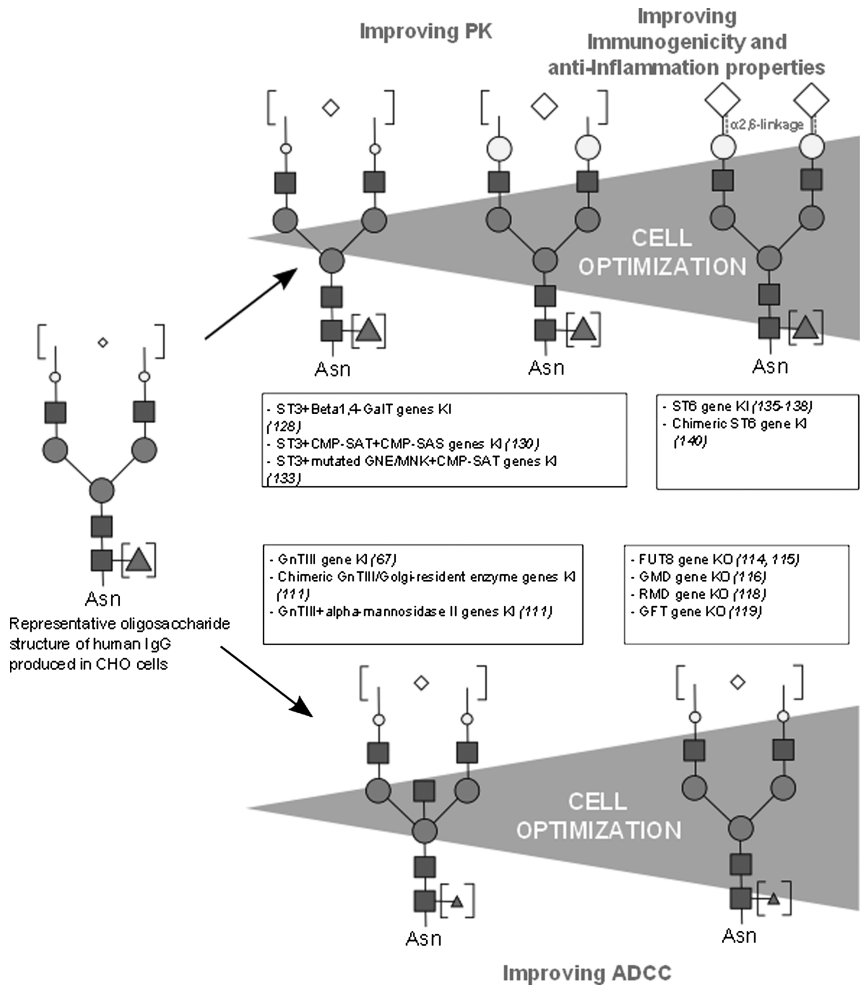


Fig. 4 Overview of the two main glycoengineering strategies in mammalian expression systems for modifying the biantennary oligosaccharide of the Fc fragment of human IgG. The size of each carbohydrate residue is representative of its abundance in the final product. Symbols: Square represents N-acetylglucosamine, triangle core fucose, circle mannose, circle galactose, diamond N-glycolylneuraminic (NeuGc) or N-acetylneuraminic acid (NeuAc).

variants advantageously located in Golgi compartments involved in the secretory pathway, in a place where they can meet protein acceptors trafficking to the cell surface [143]. It therefore appeared that ST6 enzymes require both optimization of their catalytic properties and an appropriate targeting to the secretory compartments of CHO cells to exhibit optimal enzymatic activity. Subsequently, ST6 minigenes have been constructed (Fig. 3b) to combine the optimized catalytic domain with a series of different cytoplasmic, transmembrane and stem regions and obtain non natural enzymes highly active *in vivo* (see footnote). The company SiaMed'Xpress is exploiting the transferase minigenes to equip a wide array of cell lines and significantly enhance sialylation of any glycoprotein of interest with high efficiency. Such optimized glycosyltransferases can be adapted to any drug family to achieve optimal sialic acid transfer under the best cell culture conditions. This presently represents a major advance in manipulating terminal glycosylation in drug approved cell lines.

Finally, engineering terminal sialylation may advantageously be combined with existing methods to achieve for the first time a full panel of glycoengineered cells to provide a consistent and complete pattern of human-like glycosylation.

3 Conclusions

A glycosylation-based strategy to optimize protein drugs is now available and schematized in Fig. 4. For a long time, sialylation engineering has been poorly efficient but the design of highly active transferases has recently opened the wide prospective of tailored glycosylation in drug approved cell lines. Many expression systems including human cells themselves have been extensively characterized to this purpose and each of them presents its own advantages and delivers a human protein with a content in glycoforms significantly different from that of the natural protein. At present, the overall methodology is thus well suited to respond to the current therapeutic needs of many new candidates. Methods have now been developed to produce antibodies with customized glycosylation and improved cytotoxicity and these molecules pioneer the field of glycoengineering.

At present, glycoengineering is an advanced technology which may improve production yield as glycosylation is known to (i) increase secretion [144], (ii) reduce misfolding/aggregation [3, 30] and (iii) enhance bioactivity as sialic acid is a key for prolonged duration in the circulation [33]. Meanwhile, it will help reducing overall costs as bioproduction will be more efficient in delivery products in higher yield, better stability, reduced immunogenicity and improved activity. It will also substantially reduce healthcare costs through enhanced bioavailability and low dosage of the desired drug.

Because glycosylation is species, protein and cell-specific, achieving drug optimization through glycoengineering is an exciting challenge. It is clear

The *"Preparation and Uses of Gene Sequences encoding Chimerical Glycosyltransferases with Optimized Glycosylation Activity" has been patented as* **EP2019864** [C. Ronin and G. Guiraudie-Capraz, inv].

that a glycosylation-based strategy needs to start early in the development stage to optimize safety and efficacy profile of each protein candidate. For that purpose, expression systems should be selected as to provide increased secretion along with proteins with reduced immunogenicity and improved glycoform patterns. Manipulating cell culture conditions will further optimize the glycoform profile to meet the regulatory requirements and give the best opportunity to select lead candidates. Purification will finally select the highly sialylated forms which will ensure enhanced blood duration and improve efficacy during further drug development. Meanwhile, desired glycoform pattern and bioactivity testing will be screened along with the strategy to finally reach the best profile for drug efficacy. It may thus be anticipated that at each stage of development, glycoengineering will take place differently to achieve final optimal improvement in drug performances. Indeed, bioproduction of glycoengineered proteins needs to be evaluated on a case-by-case basis. Such multistep strategy will have to be adapted to each protein from the very beginning and will hopefully overcome most of the difficulties encountered so far, thereby speeding up the scale up production and limiting costs at each step.

Of special interest, accurate estimation of glycosylated biomarker or drug amount is of paramount importance throughout glycoform optimization. Valuable measurement, especially during pharmacokinetic studies, is the key in assessing efficacy improvement. However, a number of current immunoassays heavily discriminate glycoforms because they differ in immunological properties. Quantitative measurement of a human protein produced by recombinant technology may be greatly hampered by epitope non-similarity between preparations of natural and recombinant origin. Consequently, measuring a combination of glycoforms and searching for a gain in bioactivity may readily lack accuracy, especially in blood [76]. While structural analysis is of definite help in characterizing the wide panel of glycans, it does not always estimate the ratio of the various glycoforms accurately. There is thus a clear need to achieve quantitative assessment of glycoform mixture at all steps of drug development, including clinical trials. Development of new assays based on glycoengineered proteins and glycoform-specific antibodies is the most promising advance at the moment.

All the knowledge acquired so far in the field of therapeutic proteins will be of great help to identify the key steps that can be specifically targeted for glycoengineering. To date, terminal sialylation, in most proteins, and core fucosylation, typically in anti-cancer antibodies, are the two major effectors that have been identified among marketed drugs. The remarkable knowledge issued from high throughput glycome analysis will help identifying the proper changes in glycosylation to reach therapeutic targets and/or restore normal function in diseased or age-related states. Focusing glycoengineering on these key issues over a new array of different therapeutic proteins will undoubtedly promote the design of new drugs. Meanwhile, it will also demonstrate that the glycan repertoire can also be advantageously manipulated for the benefit of a wide array of protein targets. Over time, molecular and cellular engineering may be able to design and produce by the key molecules of many different biomedical applications.

References

1 R. G. Spiro, Glycoproteins, *AdvProteinChem*, 1973, **27**, 349–467.
2 H. Lis and N. Sharon, Protein glycosylation. Structural and functional aspects, *Eur J Biochem*, 1993, **218**, 1–27.
3 R. J. Solá and K. Griebenow, Glycosylation of therapeutic proteins: an effective strategy to optimize efficacy, *Bio Drugs*, 2010, **24**, 9–21.
4 P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson and R. A. Dwek, Glycosylation and the immune system, *Science*, 2001, **291**, 2370–6.
5 K. Ohtsubo and J. D. Marth, Glycosylation in cellular mechanisms of health and disease, *Cell*, 2006, **126**, 855–67.
6 J. Wang, C. I. Balog, K. Stavenhagen, C. A. Koeleman, H. U. Scherer, M. H. Selman, A. M. Deelder, T. W. Huizinga, R. E. Toes and M. Wuhrer, Fc-glycosylation of IgG1 is modulated by B-cell stimuli, *Mol Cell Proteomics*, 2011, **10**.
7 J. Zhao, D. M. Simeone, D. Heidt, M. A. Anderson and D. M. Lubman, Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum, *J Proteome Res*, 2006, **5**, 1792–802.
8 A. Ercan, J. Cui, D. E. Chatterton, K. D. Deane, M. M. Hazen, W. Brintnell, C. I. O'Donnell, L. A. Derber, M. E. Weinblatt, N. A. Shadick, D. A. Bell, E. Cairns, D. H. Solomon, V. M. Holers, P. M. Rudd and D. M. Lee, Aberrant IgG galactosylation precedes disease onset, correlates with disease activity, and is prevalent in autoantibodies in rheumatoid arthritis, *Arthritis Rheum*, 2010, **62**, 2239–48.
9 Y. Sato and T. Endo, Alteration of brain glycoproteins during aging, *Geriatr Gerontol Int*, 2010, **10**, 32–40.
10 E. Reynders, F. Foulquier, W. Annaert and G. Matthijs, How Golgi glycosylation meets and needs trafficking: the case of the COG complex, *Glycobiology*, 2011, **21**, 853–63.
11 E. Lisowska, The role of glycosylation in protein antigenic properties, *Cell Mol Life Sci*, 2002, **59**, 445–55.
12 A. M. Sinclair and S. Elliott, Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins, *J Pharm Sci*, 2005, **94**, 1626–35.
13 R. Jefferis, Glycosylation of recombinant antibody therapeutics, *Biotechnol Prog*, 2005, **21**, 11–6.
14 G. Walsh and R. Jefferis, Post-translational modifications in the context of therapeutic protein, *Nature Biotechn*, 2006, **24**, 1241–52.
15 N. Kawasaki, S. Itoh, N. Hashii, D. Takakura, Y. Qin, X. Huang and T. Yamaguchi, The significance of glycosylation analysis in development of biopharmaceuticals, *Biol Pharm Bull*, 2009, **32**, 796–800.
16 H. Li and M. d'Anjou, Pharmacological significance of glycosylation in therapeutic proteins, *Curr Opin Biotechnol*, 2009, **20**, 678–84.
17 U. H. Stenman, Standardization of assays for human chorionic gonadotropin, *Clin Chem*, 2004, **50**, 798–800.
18 J. R. Moskal, R. A. Kroes and G. Dawson, The glycobiology of brain tumors: disease relevance and therapeutic potential, *Expert Rev Neurother*, 2009, **9**, 1529–45.
19 P. M. Drake, W. Cho, B. Li, A. Prakobphol, E. Johansen, N. L. Anderson, F. E. Regnier, B. W. Gibson and S. J. Fisher, Sweetening the pot: adding glycosylation to the biomarker discovery equation, *Clin Chem*, 2010, **56**, 223–36.
20 F. M. Wurm, Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol*, 2004, **22**, 1393–8.

21 Y. Durocher and M. Butler, Expression systems for therapeutic glycoprotein production, *Curr Opin Biotechnol*, 2009, **20**, 700–7.
22 R. Jefferis, Glycosylation as a strategy to improve antibody-based therapeutics, *Nat Rev Drug Discov*, 2009, **8**, 226–34.
23 J. R. Rich and S. G. Withers, Emerging methods for the production of homogeneous human glycoproteins, *Nat Chem Biol*, 2009, **5**, 206–15.
24 M. Bencúrová, W. Hemmer, M. Focke-Tejkl, I. B. Wilson and F. Altmann, Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin, *Glycobiology*, 2004, **14**, 457–66.
25 S. Hino, F. Umeda, S. Inumaru, N. Aoki, C. Sato, T. Okajima, D. Nadano and T. Matsuda, IgG2 dominancy and carbohydrate recognition specificity of C3H/He mouse antibodies directed to cross-reactive carbohydrate determinants (CCDs) bearing beta-(1,2)-xylose and alpha-(1,3)-fucose, *Immunol Lett*, 2010, **133**, 28–34.
26 S. R. Hamilton and T. U. Gerngross, Glycosylation engineering in yeast: the advent of fully humanized yeast, *Curr Opin Biotechnol*, 2007, **18**, 387–92.
27 Y. Lim, N. Wong, Y. Y. Lee, S. Ku, D. C. Wong and M. G. Yap, Engineering mammalian cells in bioprocessing - current achievements and future perspectives, *Biotechnol Appl Biochem*, 2010, **55**, 175–89.
28 W. Morelle, S. Donadio, C. Ronin and J. C. Michalski, Characterization of N-glycans of recombinant human thyrotropin using mass spectrometry, *Rapid Commun Mass Spectrom*, 2006, **20**, 331–45.
29 Y. T. Jeong, O. Choi, H. R. Lim, Y. D. Son, H. J. Kim and J. H. Kim, Enhanced sialylation of recombinant erythropoietin in CHO cells by human glycosyltransferase expression, *J Microbiol Biotechnol*, 2008, **18**, 1945–52.
30 X. Xu, H. Nagarajan, N. E. Lewis, S. Pan, Z. Cai, X. Liu, W. Chen, M. Xie, W. Wang, S. Hammond, M. R. Andersen, N. Neff, B. Passarelli, W. Koh, H. C. Fan, J. Wang, Y. Gui, K. H. Lee, M. J. Betenbaugh, S. R. Quake, I. Famili, B. O. Palsson and J. Wang, The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line, *Nat Biotechnol*, 2011, **29**, 735.
31 G. Vegarud and T. B. Christnsen, Glycosylation of Proteins: a new method of enzyme stabilization, *Biotechnol Bioeng*, 1975, **17**, 1391–7.
32 M. Baudys, T. Uchio, D. Mix, D. Wilson and S. W. Kim, Physical stabilization of insulin by glycosylation, *J Pharm Sci*, 1995, **84**, 28–33.
33 V. Kayser, N. Chennamsetty, V. Voynov, K. Forrer, B. Helk and B. L. Trout, Glycosylation influences on the aggregation propensity of therapeutic monoclonal antibodies, *Biotechnol J*, 2011, **6**, 38–44.
34 T. S. Raju and B. Scallon, Fc glycans terminated with N-acetylglucosamine residues increase antibody resistance to papain, *BiotechnolProg*, 2007, **23**, 964–71.
35 A. J. Jones, D. I. Papac, E. H. Chin, R. Keck, S. A. Baughman, Y. S. Lin, J. Kneer and J. E. Battersby, Selective clearance of glycoforms of a complex glycoprotein pharmaceutical caused by terminal N-acetylglucosamine is similar in humans and cynomolgus monkeys, *Glycobiology*, 2007, **17**, 529–40.
36 A. G. Morell, G. Gregoriadis, I. H. Scheinberg, J. Hickman and G. Ashwell, The role of sialic acid in determining the survival of glycoproteins in the circulation, *J BiolChem*, 1971, **246**, 1461–7.
37 L. M. Steirer, E. I. Park, R. R. Townsend and J. U. Baenziger, The asialoglycoprotein receptor regulates levels of plasma glycoproteins terminating with sialic acid alpha2,6-galactose, *J BiolChem*, 2009, **284**, 3777–83.
38 D. W. Briggs, J. W. Fisher and W. J. George, Hepatic clearance of intact and desialylated erythropoietin, *Am J Physiol*, 1974, **227**, 1385–8.

39 I. C. Macdougall, Optimizing the use of erythropoietic agents--pharmacokinetic and pharmacodynamic considerations, *Nephrol Dial Transplant*, 2002, **17**, 66–70.
40 B. Byrne, G. G. Donohoe and R. O'Kennedy, Sialic acids: carbohydrate moieties that influence the biological and physical properties of biopharmaceutical proteins and living cells, *Drug Discov Today*, 2007, **12**, 319–26.
41 J. C. Egrie, E. Dwyer, J. K. Browne, A. Hitz and M. A. Lykos, Darbepoetin alfa has a longer circulating half-life and greater in vivo potency than recombinant human erythropoietin, *ExpHematol*, 2003, **31**, 290–9.
42 C. Wacker, C. N. Berger, P. Girard and R. Meier, Glycosylation profiles of therapeutic antibody pharmaceuticals, *Eur J Pharm Biopharm*, 2011, **2**.
43 A. W. Guzzetta, L. J. Basa, W. S. Hancock, B. A. Keyt and W. F. Bennett, Identification of carbohydrate structures in glycoprotein peptide maps by the use of LC/MS with selected ion extraction with special reference to tissue plasminogen activator and a glycosylation variant produced by site directed mutagenesis, *Anal Chem*, 1993, **1**, 2953–62.
44 D. P. Beebe and D. L. Aronson, Turnover of tPA in rabbits: influence of carbohydrate moieties, *Thromb Res*, 1988, **51**, 11–22.
45 M. Einarsson, J. Brandt and L. Kaplan, Large-scale purification of human-tissue-type plasminogen activator using monoclonal antibodies, *Biochim Biophys Acta*, 1985, **830**, 1–10.
46 T. S. Raju, Terminal sugars of Fcglycans influence antibody effector functions of IgGs, *Curr Opin Immunol*, 2008, **20**, 471–8.
47 A. Wright and S. L. Morrison, Effect of glycosylation on antibody function: implications for genetic engineering, *Trends Biotechnol*, 1997, **15**, 26–32.
48 J. M. Woof and D. R. Burton, Human antibody-Fc receptor interactions illuminated by crystal structures, *Nat Rev Immunol*, 2004, **4**, 89–99.
49 R. B. Parekh, R. A. Dwek, B. J. Sutton, D. L. Fernandes, A. Leung, D. Stanworth, T. W. Rademacher, T. Mizuochi, T. Taniguchi, K. Matsuta, F. Takeuchi, Y. Nagano, T. Miyamoto and A. Kobata, Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG, *Nature*, 1985, **1–7**, 452–7.
50 Y. Kaneko, F. Nimmerjahn and J. V. Ravetch, Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation, *Science*, 2006, **313**, 670–3.
51 J. Bayry, K. Bansal, M. D. Kazatchkine and S. V. Kaveri, DC-SIGN and alpha2,6-sialylated IgG Fc interaction is dispensable for the anti-inflammatory activity of IVIg on human dendritic cells, *Proc Natl Acad Sci U S A*, 2009, **106**.
52 A. E. Bedwell, C. J. Elson, S. D. Carter, P. A. Dieppe, C. W. Hutton and R. Czudek, Isolation and analysis of complement activating aggregates from synovial fluid of patients with rheumatoid arthritis using monoclonal anti-C3d antibodies, *Ann Rheum Dis*, 1987, **46**, 55–64.
53 Y. Kaneko, F. Nimmerjahn and J. V. Ravetch, Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation, *Science*, 2006, **313**, 670–3.
54 R. M. Anthony, F. Nimmerjahn, D. J. Ashline, V. N. Reinhold, J. C. Paulson and J. V. Ravetch, Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc, *Science*, 2008, **320**, 373–6.
55 R. M. Anthony and J. V. Ravetch, A novel role for the IgG Fc glycan: the anti-inflammatory activity of sialylatedIgG Fcs, *J Clin Immunol*, 2010, **30**, 9–14.
56 J. F. Séïté, D. Cornec, Y. Renaudineau, P. Youinou, R. A. Mageed and S. Hillion, IVIg modulates BCR signaling through CD22 and promotes apoptosis in mature human B lymphocytes, *Blood*, 2010, **116**, 1698–704.
57 D. N. Forthal, J. S. Gach, G. Landucci, J. Jez, R. Strasser, R. Kunert and H. Steinkellner, Fc-glycosylation influences Fcγ receptor binding and cell-mediated anti-HIV activity of monoclonal antibody 2G12, *J Immunol*, 2010, **1**, 6876–82.

58 P. Umaña, J. Jean-Mairet, R. Moudry, H. Amstutz and J. E. Bailey, Engineered glycoforms of an anti neuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity, *Nat Biotechnol*, 1999, **17**, 176–80.
59 J. Davies, L. Jiang, L. Z. Pan, M. J. LaBarre, D. Anderson and M. Reff, Expression of GnTIII in a recombinant anti-CD20 CHO production cell line: Expression of antibodies with altered glycoforms leads to an increase in ADCC through higher affinity for FC gamma RII, *Biotechnol Bioeng*, 2001, **20**, 288–94.
60 C. Ferrara, S. Grau, C. Jäger, P. Sondermann, P. Brünker, I. Waldhauer, M. Hennig, A. Ruf, A. C. Rufer, M. Stihle, P. Umaña and J. Benz, Unique carbohydrate-carbohydrate interactions are required for high affinity binding between Fc{gamma}RIII and antibodies lacking core fucose, *Proc Natl Acad Sci U S A*, 2011, **2**, 12669–74.
61 J. Hodoniczky, Y. Z. Zheng and D. C. James, Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling *in vitro*, *Biotechnol Prog*, 2005, **21**, 1644–52.
62 P. N. Boyd, A. C. Lines and A. K. Patel, The effect of the removal of sialic acid, galactose and total carbohydrate on the functional activity of Campath-1H, *Mol Immunol*, 1995, **32**, 1311–8.
63 P. H. Van Berkel, J. Gerritsen, G. Perdok, J. Valbjørn, T. Vink, J. G. Van de Winkel and P. W. Parren, N-linked glycosylation is an important parameter for optimal selection of cell lines producing biopharmaceutical human IgG, *Biotechnol Prog*, 2009, **25**, 244–51.
64 Y. Kanda, T. Yamada, K. Mori, A. Okazaki, M. Inoue, K. Kitajima-Miyama, R. Kuni-Kamochi, R. Nakano, K. Yano, S. Kakita, K. Shitara and M. Satoh, Comparison of biological activity among nonfucosylated therapeutic IgG1 antibodies with three different N-linked Fc oligosaccharides: the high-mannose, hybrid, and complex types, *Glycobiology*, 2006, **17**, 104–18.
65 D. Ternant, A. Aubourg, C. Magdelaine-Beuzelin, D. Degenne, H. Watier, L. Picon and G. Paintaud, Infliximab pharmacokinetics in inflammatory bowel disease patients, *Ther Drug Monit*, 2008, **30**, 523–9.
66 J. G. Gribben, S. Devereux, N. S. Thomas, M. Keim, H. M. Jones, A. H. Goldstone and D. C. Linch, Development of antibodies to unprotected glycosylation sites on recombinant human GM-CSF, *Lancet*, 1990, **24**, 434–7.
67 S. Porter, Human immune response to recombinant human proteins, *J Pharm Sci*, 2001, **90**, 1–11.
68 N. Casadevall, Antibodies against rHuEPO: native and recombinant, *Nephrol Dial Transplant*, 2002, **17**, 42–7.
69 A. Noguchi, C. J. Mukuria, E. Suzuki and M. Naiki, Immunogenicity of N-glycolylneuraminicacid-containing carbohydrate chains of recombinant human erythropoietin expressed in Chinese hamster ovary cells, *J Biochem*, 1995, **117**, 59–62.
70 A. Zhu and R. Hurst, Anti-N-glycolylneuraminic acid antibodies identified in healthy human serum, *Xenotransplantation*, 2002, **9**, 376–81.
71 D. Ghaderi, R. E. Taylor, V. Padler-Karavani, S. Diaz and A. Varki, Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins, *Nat Biotechnol*, 2010, **28**, 863–7.
72 C. K. Borrebaeck, A. C. Malmborg and M. Ohlin, Does endogenous glycosylation prevent the use of mouse monoclonal antibodies as cancer therapeutics?, *Immunol Today*, 1993, **14**, 477–9.
73 U. Galili, F. Anaraki, A. Thall, C. Hill-Black and M. Radic, One percent of human circulating B lymphocytes are capable of producing the natural anti-Gal antibody, *Blood*, 1993, **15**, 2485–93.

74 C. H. Chung, B. Mirakhur, E. Chan, Q. T. Le, J. Berlin, M. Morse, B. A. Murphy, S. M. Satinover, J. Hosen, D. Mauro, R. J. Slebos, Q. Zhou, D. Gold, T. Hatley, D. J. Hicklin and T. A. Platts-Mills, Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose, *N Engl J Med*, 2008, **13**, 1109–17.
75 G. Antonelli, Development of neutralizing and binding antibodies to interferon (IFN) in patients undergoing IFN therapy, *Antiviral Res*, 1994, **24**, 235–44.
76 S. Donadio, W. Morelle, A. Pascual, R. Romi-Lebrun, J. C. Michalski and C. Ronin, Both core and terminal glycosylation alter epitope expression in thyrotropin and introduce discordances in hormone measurements, *Clin Chem Lab Med*, 2005, **43**, 519–30.
77 C. A. Spencer, M. Takeuchi and M. Kazarosyan, Current status and performance goals for serum thyrotropin (TSH) assays, *Clin Chem*, 1996, **42**, 140–5.
78 W. G. Wood, D. Waller and U. Hantke, An evaluation of six solid-phase thyrotropin (TSH) kits, *J Clin Chem Clin Biochem*, 1985, **23**, 461–71.
79 P. Carayon, E. Martino, L. Bartalena, L. Grasso, C. Mammoli, S. Costagliola and A. Pinchera, Clinical usefulness and limitations of serum thyrotropin measurement by 'ultrasensitive' methods. Comparisons of five kits, *Horm Res*, 1987, **26**, 105–17.
80 B. Y. Chan and R. Swaminathan, Analytical and clinical performance of six sensitive thyrotrophin kits, *Pathology*, 1990, **22**, 11–5.
81 P. Laurberg, Persistent problems with the specificity of immunometric TSH assays, *Thyroid*, 1993, **3**, 279–83.
82 C. A. Spencer and C. C. Wang, Thyroglobulin measurement. Techniques, clinical benefits, and pitfalls, *Endocrinol Metab Clin North Am*, 1995, **24**, 841–63.
83 W. G. Wood, U. Bruns, O. Eber, W. Langsteger, J. Y. Bounaud and M. P. Bounaud, Evaluation of the Abbott IMx ultrasensitive II hTSH immunometric assay in three European centers: a comparison with established commercial immunometric assays for thyrotropin, *Eur J Clin Chem Clin Biochem*, 1996, **34**, 151–8.
84 A. K. Rasmussen, L. Hilsted, H. Perrild, E. Christiansen, K. Siersbaek-Nielsen and U. Feldt-Rasmussen, Discrepancies between thyrotropin (TSH) measurement by four sensitive immunometric assays, *Clin Chim Acta*, 1997, **259**, 117–28.
85 A. Price, C. Burgin, I. Catch and M. Cruise, Functional sensitivity and recovery of thyroid-stimulating hormone, *Clin Chem*, 2001, **47**, 2067.
86 L. M. Demers and C. A. Spencer, Laboratory medicine practice guidelines: laboratory support for the diagnosis and monitoring of thyroid disease, *Clin Endocrinol (Oxf)*, 2003, **58**, 138–40.
87 M. L. Rawlins and W. L. Roberts, Performance characteristics of six third-generation assays for thyroid-stimulating hormone, *Clin Chem*, 2004, **50**, 2338–44.
88 I. Sergi, M. J. Papandreou, G. Medri, C. Canonne, B. Verrier and C. Ronin, Immunoreactive and bioactive isoforms of human thyrotropin, *Endocrinology*, 1991, **128**, 3259–68.
89 S. Donadio, A. Pascual, J. H. Thijssen and C. Ronin, Feasibility study of new calibrators for thyroid-stimulating hormone (TSH) immunoprocedures based on remodeling of recombinant TSH to mimic glycoforms circulating in patients with thyroid disorders, *Clin Chem*, 2006, **52**, 286–97.
90 S. Donadio, A. Pascual, M. Dugas and C. Ronin, Standardization of TSH immunoassays: production of new calibrators and harmonization of assays, *Immunoanalyse & Biologie Spécialisée*, 2006, **21**, 79–90.

91 European Medicines Agency, Pre-authorisation Evaluation of Medicines for Human Use Committee for medicinal products for human use (CHMP) Guideline on development, production, characterization and specifications for monoclonal antibodies and related products BWP/157653 2008 (accessed September 2, 2011 at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003074.pdf)
92 Center for Biologics Evaluation and Research (CBER) Center for Drug Evaluation and Research (CDER) FDA Guidance concerning demonstration of comparability of human biological products, including therapeutic biotechnology-derived products 1996 (accessed September 2, 2011 at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm122879.htm)
93 International conference on harmonization of technical requirements for registration of pharmaceuticals for human use. ICH harmonized Tripartite Guideline. Specifications: Test procedures and acceptance for biotechnological/biological Products Q6B 1999 (accessed September 2, 2011 at http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q6B/Step4/Q6B_Guideline.pdf)
94 K. N. Baker, M. H. Rendall, A. E. Hills, M. Hoare, R. B. Freedman and D. C. James, Metabolic control of recombinant protein N-glycan processing in NS0 and CHO cells, *Biotechnol Bioeng*, 2001, **73**, 188–202.
95 P. Hossler, S. F. Khattak and Z. J. Li, Optimal and consistent protein glycosylation in mammalian cell culture, *Glycobiology*, 2009, **19**, 936–49.
96 M. Gawlitzek, M. Estacio, T. Fürch and R. Kiss, Identification of cell culture conditions to control N-glycosylation site-occupancy of recombinant glycoproteins expressed in CHO cells, *Biotechnol Bioeng*, 2009, **103**, 1164–75.
97 D. C. Andersen, T. Bridges, M. Gawlitzek and C. Hoy, Multiple cell culture factors can affect the glycosylation of Asn-184 in CHO-produced tissue-type plasminogen activator, *Biotechnol Bioeng*, 2000, **70**, 25–31.
98 D. C. Wong, N. S. Wong, J. S. Goh, L. M. May and M. G. Yap, Profiling of N-glycosylation gene expression in CHO cell fed-batch cultures, *Biotechnol Bioeng*, 2010, **15**, 516–28.
99 D. C. Andersen, Cell culture effects on the glycosylation of therapeutic proteins. *Bioprocess International IBC Life Sciences*, 2004.
100 X. Gu and D. I. Wang, Improvement of interferon-gamma sialylation in Chinese hamster ovary cell culture by feeding of N-acetylmannosamine, *Biotechnol Bioeng*, 1998, **58**, 642–8.
101 Q. Zhou, S. Shankara, A. Roy, H. Qiu, S. Estes, A. McVie-Wylie, K. Culm-Merdek, A. Park, C. Pan and T. Edmunds, Development of a simple and rapid method for producing non-fucosylatedoligomannose containing antibodies with increased effector function, *Biotechnol Bioeng*, 2008, **99**, 652–65.
102 P. Taupin, Modulation of GDP-fucose level for generating proteins with reduced rate of fucosylation (WO2010141855), *Expert Opin Ther Pat*, 2011, **21**, 1497–9.
103 N. Munzert, E. Mthing, J. Büntemeyer, H and J. Lehmann, Sialidase activity in culture fluid of Chinese hamster ovary cells during batch culture and its effects on recombinant human antithrombin III integrity, *Biotechnol Prog*, 1996, **12**, 559–63.
104 X. Gu, B. J. Harmon and D. I. Wang, Site- and branch-specific sialylation of recombinant human interferon-gamma in Chinese hamster ovary cell culture, *Biotechnol Bioeng*, 1997, **55**, 390–8.
105 M. J. Gramer, C. F. Goochee, V. Y. Chock, D. T. Brousseau and M. B. Sliwkowski, Removal of sialic acid from a glycoprotein in CHO cell culture

supernatant by action of an extracellular CHO cell sialidase, *Biotechnology (N Y)*, 1995, **13**, 692–8.
106 M. Sarkar, E. Hull, Y. Nishikawa, R. J. Simpson, R. L. Moritz, R. Dunn and H. Schachter, Molecular cloning and expression of cDNA encoding the enzyme that controls conversion of high-mannose to hybrid and complex N-glycans: UDP-N-acetylglucosamine: alpha-3-D-mannoside beta-1,2-N-acetylglucosaminyltransferase I, *Proc Natl Acad Sci U S A*, 1991, **88**, 234–8.
107 J. C. Paulson and K. J. Colley, Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation, *J Biol Chem*, 1989, **264**, 17615–8.
108 Y. Zhu, X. Li, A. Mcvie-Wylie, C. Jiang, B. L. Thurberg, N. Raben, R. J. Mattaliano and S. H. Cheng, Carbohydrate-remodelled acid α-glucosidase with higher affinity for the cation-independent mannose 6-phosphate receptor demonstrates improved delivery to muscles of Pompe mice, *Biochem J*, 2005, **389**, 619–628.
109 T. S. Raju, J. B. Briggs, S. M. Chamow, M. E. Winkler and A. J. Jones, Glycoengineering of therapeutic glycoproteins: *in vitro* galactosylation and sialylation of glycoproteins with terminal N-acetylglucosamine and galactose residues, *Biochemistry*, 2001, **40**, 8868–76.
110 A. Hayashi, T. Chiba, H. Hayashi, S. Sasayama, T. Ishiguro and K. Onozaki, Synthesis of glycosylated human tumor necrosis factor alpha coupled with N-acetylneuraminic acid, *Cancer Immunol Immunother*, 2007, **56**, 545–53.
111 E. Grabenhorst, P. Schlenke, S. Pohl, M. Nimtz and H. S. Conradt, Genetic engineering of recombinant glycoproteins and the glycosylation pathway in mammalian host cells, *Glycoconj J*, 1999, **16**, 81–97.
112 H. Narimatsu, Human glycogene cloning: focus on beta 3-glycosyltransferase and beta 4-glycosyltransferase families, *Curr Opin Struct Biol*, 2006, **16**, 567–75.
113 J. R. Desjarlais and G. A. Lazar, Modulation of antibody effector function, *Exp Cell Res*, 2011, **317**, 1278–85.
114 C. Ferrara, P. Brünker, T. Suter, S. Moser, U. Püntener and P. Umaña, Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1, 4-N-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II, *Biotechnol Bioeng*, 2006, **5**, 851–61.
115 M. S. Czuczman and S. A. Gregory, The future of CD20 monoclonal antibody therapy in B-cell malignancies, *Leuk Lymphoma*, 2010, **51**, 983–94.
116 N. Yamane-Ohnuki and M. Satoh, Production of therapeutic antibodies with controlled fucosylation, *MAbs*, 2009, **1**, 230–6.
117 N. Yamane-Ohnuki, S. Kinoshita, M. Inoue-Urakubo, M. Kusunoki, S. Iida, R. Nakano, M. Wakitani, R. Niwa, M. Sakurada, K. Uchida, K. Shitara and M. Satoh, Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity, *Biotechnol Bioeng*, 2004, **87**, 614–22.
118 T. T. Junttila, K. Parsons, C. Olsson, Y. Lu, Y. Xin, J. Theriault, L. Crocker, O. Pabonan, T. Baginski, G. Meng, K. Totpal, R. F. Kelley and M. X. Sliwkowski, Superior *in vivo* efficacy of afucosylated trastuzumab in the treatment of HER2-amplified breast cancer, *Cancer Res*, 2010, **1**, 4481–9.
119 Y. Kanda, H. Imai-Nishiya, R. Kuni-Kamochi, K. Mori, M. Inoue, K. Kitajima-Miyama, A. Okazaki, S. Iida, K. Shitara and M. Satoh, Establishment of a GDP-mannose 4,6-dehydratase (GMD) knockout host cell line: a new strategy for generating completely non-fucosylated recombinant therapeutics, *J Biotechnol*, 2007, **30**, 300–10.

120 H. Imai-Nishiya, K. Mori, M. Inoue, M. Wakitani, S. Iida, K. Shitara and M. Satoh, Double knockdown of alpha1,6-fucosyltransferase (FUT8) and GDP-mannose 4,6-dehydratase (GMD) in antibody-producing cells: a new strategy for generating fully non-fucosylated therapeutic antibodies with enhanced ADC, *BMC Biotechnol*, 2007, **7**, 84.
121 H. H. Von Horsten, C. Ogorek, V. Blanchard, C. Demmler, C. Giese, K. Winkler, M. Kaup, M. Berger, I. Jordan and V. Sandig, Production of non-fucosylated antibodies by co-expression of heterologous GDP-6-deoxy-D-lyxo-4-hexulose reductase, *Glycobiology*, 2010, **20**, 1607–18.
122 T. Ishiguro, S. Kawai, K. Habu, M. Sugimoto, H. Shiraiwa, S. Iijima, S. Ozaki, T. Matsumoto and H. Yamada-Okabe, A defucosylated anti-CD317 antibody exhibited enhanced antibody-dependent cellular cytotoxicity against primary myeloma cells in the presence of effectors from patients, *Cancer Sci*, 2010, **101**, 2227–33.
123 M. F. Naso, S. H. Tam, B. J. Scallon and T. S. Raju, Engineering host cell lines to reduce terminal sialylation of secreted antibodies, *MAbs*, 2010, **2**, 519–27.
124 S. Wildt and T. U. Gerngross, The humanization of N-glycosylation pathways in yeast, *Nat Rev Microbio*, 2005, **3**, 119–28.
125 J. Ferrari, J. Gunson, J. Lofgren, L. Krummen and T. G. Warner, Chinese hamster ovary cells with constitutively expressed sialidase antisense RNA produce recombinant DNase in batch culture with increased sialic acid, *Biotechnol Bioeng*, 1998, **5**, 589–95.
126 F. A. Ngantung, P. G. Miller, F. R. Brushett, G. L. Tang and D. I. Wang, RNA interference of sialidase improves glycoprotein sialic acid content consistency, *Biotechnol Bioeng*, 2006, **95**, 106–19.
127 M. Zhang, K. Koskie, J. S. Ross, K. J. Kayser and M. V. Caple, Enhancing glycoprotein sialylation by targeted gene silencing in mammalian cells, *Biotechnol Bioeng*, 2010, **105**, 1094–105.
128 D. Jones, N. Kroos, R. Anema, B. van Montfort, A. Vooys, S. van der, Kraats, E. van der, Helm, S. Smits, J. Schouten, K. Brouwer, F. Lagerwerf, P. van Berkel, D. J. Opstelten, T. Logtenberg and A. Bout, *Biotechnol. Prog.*, 2003, **19**(1), 163–168.
129 K. Bork, R. Horstkorte and W. Weidemann, Increasing the sialylation of therapeutic glycoproteins: the potential of the sialic acid biosynthetic pathway, *J Pharm Sci*, 2009, **98**, 3499–508.
130 K. Fukuta, R. Abe, T. Yokomatsu, N. Kono, M. Asanagi, F. Omae, M. T. Minowa, M. Takeuchi and T. Makino, Remodeling of sugar chain structures of human interferon-gamma, *Glycobiology*, 2000, **10**, 421–30.
131 K. Fukuta, T. Yokomatsu, R. Abe, M. Asanagi and T. Makino, Genetic engineering of CHO cells producing human interferon-gamma by transfection of sialyltransferases, *Glycoconj J*, 2000, **17**, 895–904.
132 S. Weikert, D. Papac, J. Briggs, D. Cowfer, S. Tom, M. Gawlitzek, J. Lofgren, S. Mehta, V. Chisholm, N. Modi, S. Eppler, K. Carroll, S. Chamow, D. Peers, P. Berman and L. Krummen, Engineering Chinese hamster ovary cells to maximize sialic acid content of recombinant glycoproteins, *Nat Biotechnol*, 1999, **17**, 1116–21.
133 Y. T. Jeong, O. Choi, Y. D. Son, S. Y. Park and J. H. Kim, Enhanced sialylation of recombinant erythropoietin in genetically engineered Chinese-hamster ovary cells, *Biotechnol Appl Biochem*, 2009, **52**, 283–91.
134 O. T. Keppler, S. Hinderlich, J. Langner, R. Schwartz-Albiez, W. Reutter and M. Pawlita, UDP-GlcNAc 2-epimerase: a regulator of cell surface sialylation, *Science*, 1999, **84**, 1372–6.

135 K. Bork, W. Reutter, W. Weidemann and R. Horstkorte, Enhanced sialylation of EPO by overexpression of UDP-GlcNAc 2-epimerase/ManAc kinase containing a sialuria mutation in CHO cells, *FEBS Lett*, 2007, **581**, 4195–8.
136 Y. D. Son, Y. T. Jeong, S. Y. Park and J. H. Kim, Enhanced sialylation of recombinant human erythropoietin in Chinese hamster ovary cells by combinatorial engineering of selected genes, *Glycobiology*, 2011, **21**, 1019–28.
137 E. U. Lee, J. Roth and J. C. Paulson, Alteration of terminal glycosylation sequences on N-linked oligosaccharides of Chinese hamster ovary cells by expression of beta-galactoside alpha 2,6-sialyltransferase, *J Biol Chem*, 1989, **264**, 13848–55.
138 L. Monaco, A. Marc, A. Eon-Duval, G. Acerbis, G. Distefano, D. Lamotte, J. M. Engasser, M. Soria and N. Jenkins, Genetic engineering of a2,6-sialyltransferase in recombinant ClIO cells and its effects on the sialylation of recombinant interferon gamma, *Cytotechnology*, 1996, **22**, 197–203.
139 X. Zhang, S. H. Lok and O. L. Kon, Stable expression of human alpha-2,6-sialyltransferase in Chinese hamster ovary cells: functional consequences for human erythropoietin expression and bioactivity, *Biochim Biophys Acta*, 1998, **1425**, 441–52.
140 A. Bragonzi, G. Distefano, L. D. Buckberry, G. Acerbis, C. Foglieni, D. Lamotte, G. Campi, A. Marc, M. R. Soria, N. Jenkins and L. Monaco, A new Chinese hamster ovary cell line expressing alpha2,6-sialyltransferase used as universal host for the production of human-like sialylated recombinant glycoproteins, *Biochim Biophys Acta*, 2000, **1474**, 273–82.
141 R. Jassal, N. Jenkins, J. Charlwood, P. Camilleri, R. Jefferis and J. Lund, Sialylation of human IgG-Fc carbohydrate by transfected rat alpha2,6-sialyltransferase, *Biochem Biophys Res Commun*, 2001, **286**, 243–9.
142 P. Legaigneur, C. Breton, A. El Battari, J. C. Guillemot, C. Auge, M. Malissard, E. G. Berger and C. Ronin, Exploring the acceptor substrate recognition of the human beta-galactoside alpha 2,6-sialyltransferase, *J Biol Chem*, 2001, **276**, 21608–17.
143 S. Donadio, C. Dubois, G. Fichant, L. Roybon, J. C. Guillemot, C. Breton and C. Ronin, Recognition of cell surface acceptors by two human alpha-2,6-sialyltransferases produced in CHO cells, *Biochimie*, 2003, **85**, 311–21.
144 Y. Liu, A. Nguyen, R. L. Wolfert and S. Zhuo, Enhancing the secretion of recombinant proteins by engineering N-glycosylation sites, *Biotechnol Prog*, 2009, **25**, 1468–75.

Congenital Disorders of Glycosylation (CDG): from glycoproteins to patient care

Vanessa Ferreira,*[a] Paz Briones[b] and Maria-Antonia Vilaseca[c]
DOI: 10.1039/9781849734769-00124

1 Introduction

Carbohydrates are at the forefront of Congenital Disorders of Glycosylation (CDG), a rapidly growing family of genetic diseases with more than 50 members identified at the molecular and biochemical level since the first clinical description in 1980. They are due to defects in the synthesis of glycans and in their attachment to proteins and lipids.

Glycosylation is one of the most complex post-translational modifications and more than a half of human proteins are glycosylated.[1] A disruption of the glycosylation machinery can lead to a broad range of clinical phenotypes and affect nearly every organ, but particularly the nervous system as most of these disorders have neurological involvement. Although most of the deficiencies observed in CDG patients are only partial, the severity of the clinical manifestations underlines the relevance of glycosylation. Thus, advances in the development of diagnostic tools and research contribute to a better understanding of the causes and mechanisms of CDG, improving (early) diagnosis, increasing the knowledge and contributing to a better quality of life and outcome of children and adults affected by these pathologies.

This chapter will cover developments of the past few years, highlighting several methods and experimental systems that identified them, advantages and limitations of biological model systems, and possible new treatments. Finally, the role of CDG patients in the ''glycobiomedical'' field will be briefly commented.

2 Basic principles of protein glycosylation

Posttranslational modification is a main step in the evolution of proteins towards functional diversity. It extends the range of possible functions of a protein through the introduction of new chemical groups that may, for example, alter the hydrophobicity in membrane proteins or activate or inactivate enzymes. Glycosylation is the most important and complex form of posttranslational modification. Approximately 250 to 500 genes representing $\sim$1–2% of the total number of human genes are involved in this process.[2]

Research over the last decade has conclusively shown that protein-glycosylation is a vital process which plays a critical role in a large number of biochemical and cellular processes including nascent chain folding, cellular targeting, protease resistance, immunogenicity, signal transduction,

[a]*Vanessa Ferreira, Portuguese Association for CDG and other Rare Metabolic Diseases (APCDG-DMR), Rua Manuel da Fonseca n° 46, 2820-381 Charneca da Caparica, Portugal. E-mail: sindromecdg@gmail.com*
[b]*CSIC. IBC, Sección de Errores Congénitos del Metabolismo, Hospital Clínic. Servicio de Bioquímica y Genética molecular. Mejía Lequerica s/n. Edificio Helios III, planta baja. 08028 Barcelona, Spain.*
[c]*Guia Metabólica, Passeig de Sant Joan de Déu n° 2. 08950. Esplugues de Llobregat, Spain.*

© The Royal Society of Chemistry 2012

pharmacokinetics, and protein secretion. The glycan chains regulate enzymatic activity, confer enhanced stability and solubility to secreted proteins, and affect the functionality of immune proteins. However, the exact role of glycans in a multitude of glycoproteins remains unclear. Hypoglycosylation of glycoconjugates functioning as hormones, enzymes or transporters, leads to their impaired bioavailability, decreased activity and/or rapid degradation. Most soluble and membrane-bound proteins are glycoproteins, in contrast with the cytosolic proteins.

Glycosylation takes place both in the ER and Golgi. N-glycosylation takes place cotranslationally in the ER and further processing occurs in the Golgi. Whereas in the N-glycosylation carbohydrates are added as a large preformed oligosaccharide from a carrier, named dolichol, always to asparagine (Asn) as soon as it appears inside the ER, the O-glycosylation occurs posttranslationally and is generally initiated in the cis-Golgi. In this case, groups of the amino acids serine (Ser) or threonine (Thr) are the acceptor sites of the protein.

2.1 N-glycosylation of proteins

N-glycosylation is the most prevalent modification of membrane and secretory glycoproteins in eukaryotes. It is a three part process:

A) Assembly of the lipid-linked oligosaccharide (LLO).

B) En-bloc transfer of the LLO to the protein by the oligosaccharyltransferase (OST).

C) Remodeling of the oligosaccharide (glycan) once attached to the protein by removal and addition of monosaccharides.

A) Assembly of the lipid-linked oligosaccharide (LLO): the dolichol cycle. The biosynthesis of the oligosaccharide starts at the cytoplasmic site of the ER membrane in a step-by-step process on the carrier lipid dolichol phosphate (Dol-P). Dolichol is an isoprenoid compound synthesized by the same metabolic route as cholesterol and it is phosphorylated by a kinase that uses cytidine triphosphate (CTP) to form Dol-P. The availability of Dol-P at the cytoplasmic site of the ER membrane is a rate limiting factor in the dolichol cycle and N-glycosylation.[3] Dol-P is the structure upon which the carbohydrate moieties of N-linked glycoproteins are built. The dolichol cycle has been the focus of intensive research, and its strong conservation among higher systems has been an advantage in the elucidation of novel CDG. Different nucleotide-activated sugar donors synthesized from monosaccharides are required to form the oligosaccharide precursor. In the cytosol, the donors are uridine-diphospho-N-acetylglucosamine (UDP-GlcNAc) as N-acetylglucosamine (GlcNAc) donor and guanosinediphosphate-mannose (GDP-Man) as mannose (Man) donor. GDP-Man is generated from mannose-6-phosphate (Man-6-P), which is derived from the glucose, fructose, and mannose metabolism. Man-6-P can be formed either from fructose 6-phosphate, *via* phosphomannose isomerase (PMI) and the glucose pathway, or from free intracellular mannose phosphorylation by hexokinase (Hk) indicating that the synthesis of LLO is also limited by the pool of free mannose in the cell. Mannose-6-phosphate (Man-6-P) is then isomerized by the action of phosphomannomutase 2 (PMM2) into mannose-1-phosphate (Man-1-P).

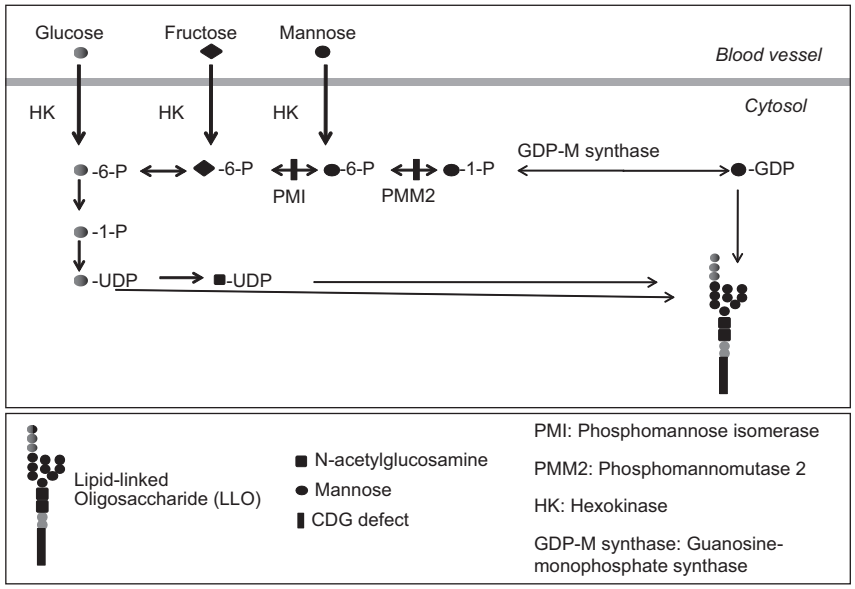


Fig. 1 Metabolic pathway showing the sources and utilization of mannose in N-glycosylation is shown. Glucose, fructose and mannose are three of the major forms of sugar in humans from which all other monosaccharides can be synthesized. Through the action of several enzymes these sugars are converted into various nucleotide sugar donors used to form complex glycans in humans. The synthesis of nucleotide activated sugars is under tight regulatory control. Loss of this control due to enzyme deficiency can significantly impair glycosylation with potential serious clinical manifestations as it causes shortage in GDP-Man and Dol-P-Man, both needed for the proper synthesis of lipid-linked oligosaccharide (LLO). For example, phosphomannomutase (PMM) or phosphomannose isomerase (PMI) deficiency causes PMM2-CDG (CDG Ia) and PMI-CDG (CDG Ib) respectively. Most mannose that is taken up is converted to fructose-6-P *via* the combined actions of hexokinase (HK) and phosphomannose isomerase (PMI). The mannose-6-phosphate (Man-6-P) is isomerised by the action of phosphomannomutase 2 (PMM2) into mannose-1-phosphate (Man-1-P). Note the important position of the PMI and PMM2 enzymes and the entry of mannose into the pathway to understand why treatment is available for PMI-CDG (CDG-Ib).

The latter is subsequently converted to GDP-Man by the action of guanosine-monophosphate transferase[4] (Fig. 1).

The oligosaccharide precursor assembly is initiated by the transfer of two N-acetylglucosamine residues to Dol-P from UDP-GlcNAc. The product is then extended by single and sequential mannose additions through the action of different mannosyltransferases using GDP-Man as the donor substrate. Once the two N-acetylglucosamines (GlcNAc) and five mannoses have been added, the Dol-PP-glycan flips from the cytosol to the ER lumen by the action of a flippase. In the ER lumen, mannosyl- and glucosyltransferases are responsible for further elongation of the oligosaccharide precursor using Dol-P linked sugar donors (Dol-P-Man and Dol-P-Glc).

B) "En-bloc" transfer of the oligosaccharide to the protein by oligosaccharyltransferase (OST). The final 14 monosaccharide glycan ($Glc_3Man_9GlcNAc_2$), is subsequently transferred to an asparagine residue contained in a N-X-S(T) sequon of the polypeptide substrate, where X can be any amino acid except proline (glycosylation consensus site),[5] and releasing Dol-PP. This is catalyzed by the OST complex, a hetero-oligomeric membrane

protein that in humans consists of seven transmembrane proteins: ribophorin I, ribophorin II, OST48, OST4, STT3A/ STT3B, N33/IAP, and DAD1. Studies in yeast showed that OST subunits N33/IAP and OST4 have accessory functions, while the other subunits are proposed to participate directly in N-glycosylation and are essential for yeast viability. Mammalian and other eukaryotic organisms express several OST isoforms.[6] For example, the human variants STT3A and STT3B are co-expressed in different tissues but differ in expression levels; OST isoforms harbouring one or the other STT3 protein differ in catalytic activity and substrate specificity. By expressing different OST isoforms, organisms may respond to alterations of the glycoprotein flux through the secretory pathway and this may contribute to tissue-specific differences in the glycan repertoire. *In vitro* studies have shown that OST has a high affinity for the fully assembled oligosaccharide, even if its abundance is lower, whereas shorter intermediates are transferred inefficiently. For this reason, many different defects leading to incomplete LLO will result in a lack of the complete oligosaccharide on many different glycoproteins at some glycosylation sites. This will affect their functionality and will result in the broad spectrum of symptoms that characterize CDG-I patients. In contrast, recent descriptions of highly specific clinical pictures in patients with mutations in the subunits of OST suggest that OST deficiencies lead to more restricted phenotypes.[7]

C) Remodeling of the oligosaccharide by removal and addition of monosaccharides. Subsequent processing of the core sugar ($Glc_3Man_9GlcNAc_2$) attached to protein starts in the ER by the trimming of the three glucose residues and one mannose. This process is crucial for the quality control of protein folding. High-mannose oligosaccharides interact with the chaperones calreticulin and calnexin which serve as sensors by binding to monoglucosylated proteins, facilitating their folding and assembly. Only glycoproteins that have been correctly folded get packaged into ER-to-Golgi transport vesicles. In the *cis* Golgi, the processing continues with further trimming of mannose residues and addition of one GlcNAc to form branched structures. Addition of N-acetylglucosamine (GlcNAc), galactose (Gal), fucose (Fuc) and sialic acid (N-acetyl-neuraminic acid, NANA) residues in the *medial*-and *trans*-Golgi will result in the mature glycoproteins to be transported to their destination in the cell, in the cell membrane and outside the cell. To this purpose nucleotide sugar transporters are carried into the Golgi[8] (Fig. 2).

2.2 O-glycosylation of proteins

In contrast with N-glycosylation, the O-glycosylation processes do not follow a unique model, and therefore, they produce an immense multitude of more or less branched chemical structures. Seven different types are found in humans that are classified according to the first sugar residue linked to the hydroxyl group of a serine, threonine or hydroxylysine of the protein. O-glycosylation has no processing and thus only consists of assembly. Contrary to N-glycosylation, this assembly mainly occurs in the Golgi. Examples of important O-glycans are *O*-*N*-acetylglucosaminylglycans (mucin-type glycans), *O*-xylosylglycans (glycosaminylglycans), *O*-mannosylglycans, and *O*-fucosylglycans (Table 1).

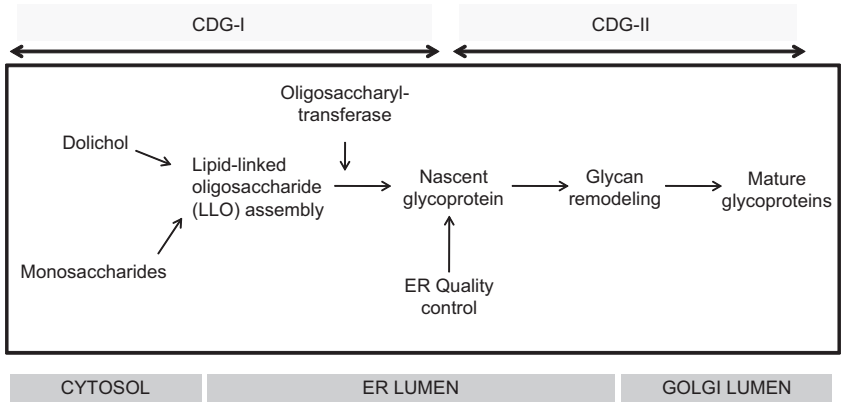


Fig. 2 Overview of the N-glycosylation process and its cellular context. Congenital Disorders of Glycosylation can be classified as type I and type II, depending on the nature and localization of the biochemical defect. Activated sugars are necessary for lipid-linked oligosaccharide assembly in the ER. This is followed by processing of N-linked oligosaccharides, which involves deletions, additions, and rearrangements of sugars starting in the ER and finishing in the Golgi. Upon completion of glycan remodeling, the mature glycoproteins are transported to their destinations. Abnormally glycosylated proteins due to CDGs are removed through specific ER quality controls (Figure adapted from Struewe *et al.*, 2009).

Table 1 Different types of O-linked glycans in humans (modified from Wopereis *et al.*[9]).

Type of O-linked glycan	Glycoprotein
Mucin–type (O-linked GalNAc)	Secreted and plasma membrane
GAG[a]	Proteoglycans
O-linked GlcNAc	Nuclear and cytoplasmic
O-linked Gal	Collagens
O-linked Man	α-dystroglycan
O-linked Glc	EGF protein domains
O-linked Fuc	EGF protein domains; TSR repeats

[a] glycosaminoglycans.

2.3 N- and O-glycosylation

Part of the glycosylation machinery (some glycosyltransferases, sugar-nucleotide transporters, *etc*...) is common to N- and O-glycosylation. In addition, recent studies showed that some proteins are essential for establishing and/or maintaining the structure and function of the Golgi, such as the Conserved Oligomeric Golgi (COG) complex, which is an eight-subunit (Cog 1-8) peripheral Golgi protein involved in retrograde vesicular Golgi trafficking.[10] ATPase 6 is also essential for proper maintenance of the Golgi.[11]

3 Congenital disorders of glycosylation (CDG)

3.1 CDG classification

CDG are due to defects in the synthesis of the glycan moiety of glycoproteins and glycolipids, and in their attachment to proteins and lipids.

There are two main groups of N-glycosylation defects. CDG-I comprises defects in the assembly of the dolichol pyrophosphate oligosaccharide and in the transfer of the oligosaccharide from dolichol to the nascent polypeptides in the ER. CDG-II consists of defects in the processing of the protein-linked glycans in the ER and the Golgi compartment (Fig. 3). O-glycosylation, defects have been identified in the synthesis of O-xylosylglycans, *O*-*N*-acetylgalactosaminylglycans, *O*-xylosyl/*N*-acetylgalactosaminylglycans, *O*-mannosylglycans and of *O*-fucosylglycans (Table 1).

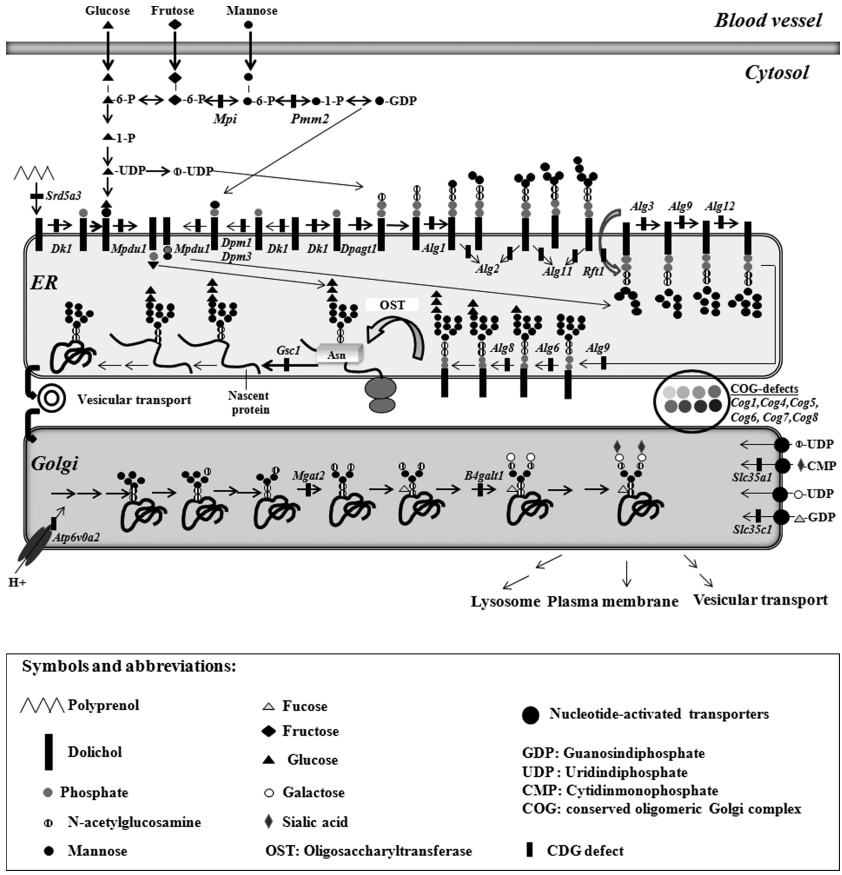


Fig. 3 Overview of glycoprotein biosynthesis and known CDG types. Dolichol is a long chain lipid strongly bound to the ER membrane. On the cytosolic side of the ER membrane, GlcNAc from UDP- GlcNAc, and mannose from GDP-Man are enzymatically linked to dolichol-PP with concomitant cleavage of the nucleotide moiety. As mentioned in the text, the preassembled carbohydrate structure linked to dolichol-PP is commonly referred to as the lipid-linked oligosaccharide (LLO). When the LLO structure is $Man_5GlcNAc_2$–PP-dolichol, the glycan is flipped into the lumen of the ER by a "flippase" and the synthesis is continued by the action of mannosyl and glucosyltransferases until the final structure $Glc_3Man_9GlcNAc_2$ –PP-dolichol. This completed oligosaccharide is transfered "*en bloc*" to the nascent protein by a oligosaccharyltransferase (OST). The processing of its composition proceeds in the ER and later in the Golgi apparatus. The action of various glucosidases, mannosidases, GlcNAc-transferases, galactosyltransferases, sialyltransferases, and eventually fucosyltransferases lead to a final-carbohydrate core. All CDG defects identified so far are symbolized by a black bar and the corresponding gene in italics (Figure adapted from Körner *et al*., 2011).

There is also a growing group of CDG with combined N- and O-glycosylation defects. A number of these are defects in proteins with a broader range of functions besides glycosylation. The first description of this type of combined defects was on COG-CDG. These defects illustrate the importance of the correct environment for glycosylation. COG defects cause alterations in the control of the intracellular trafficking and/or stability of a variety of Golgi-associated glycosylation enzymes.

The subsequent description of dolichol phosphate synthesis defects, ATP6V0A2-CDG, and SEC23B-CDG has led to the distinction of two groups of CDG: the defects in the glycosylation machinery and the group of defects in proteins that are necessary for creating the environment that permits the proper functioning of this machinery.[12]

An overview of glycoprotein biosynthesis and known CDG types is presented in Fig. 3. The disorders are showed according to the gene defect.

A growing number of patients with as yet unclassified CDG (CDG-X) might have similar "environmental" problems, indicating that in the near future other proteins involved in Golgi structure and function may be described and identified as responsible for new CDGs.[13–19]

Of the about 50 reported CDG, 17 affect N-glycosylation, 11 affect O-glycosylation, 3 involve defects in glycosphingolipid and glycosylphosphatidylinositol anchor glycosylation and 19 are caused by defects in multiple glycosylation pathways and other pathways.

The present knowledge of CDG types identified and the approximate number of affected patients are listed in Table 2. The novel nomenclature (introduced in 2009) is used along with the traditional nomenclature presented in parenthesis (for example, Ia for CDG-Ia).[20–22]

Approximately, 1000 patients with N-glycosylation defects have been published, of which PMM2-CDG (CDG-Ia) is by far the most common worldwide. Most CDG subtypes have been described in only a few individuals. Therefore, the phenotypic spectrum of these diseases is incompletely known. Since there are a large number of steps involved in glycosylation and related pathways, it is expected that this group of disorders will expand rapidly in the near future.

3.2 Clinical manifestations of CDG

Since glycosylation has a general importance, it is not surprising, that CDG often result in severe and multisystem disease. In fact, CDG cuts across nearly all medical specialties.

The clinical picture is variable but nearly always comprises dysmorphism, neurological involvement (intellectual impairment, ataxia, seizures, retinopathy), and other organ involvement (hepatic dysfunction, coagulopathies, failure to thrive, cardiomyopathy and pericardial effusion, hydrops fetalis, endocrine abnormalities, renal dysfunction, and skeletal defects). CDG should thus be considered in every patient with an unexplained syndrome. Limited awareness of CDG contributes to underdiagnosis of these disorders. Details about the clinical spectrum of CDG are summarized in Table 2 and are briefly mentioned here because they have been described in numerous excellent reviews (see references in Table 2).

Table 2 Clinical manifestations of human genetic disorders of glycosylation (Adapted from[67–71]).

Defects of protein N-glycosylation			
CDG type Number of reported patients	Protein	Overview of key cinical manifestations	Screening technique/alternative technique
PMM2-CDG (Ia) >800[72]	Phosphomannomutase	**Neurology:** strabismus, ataxia, hypotonia, psychomotor retardation, peripheral neuropathy, epilepsy, stroke-like episodes, (olivoponto)cerebellar hypoplasia **Dysmorphism:** inverted nipples, fat pads **Gastroenterology/hepatology:** failure to thrive, recurrent vomiting, liver involvement, pancreatitis Growth retardation **Neonatology:** hydrops fetalis, floppy baby **Haematology:** thrombosis, increased platelet adhesiveness, clotting and anticlotting factor abnormalities **Endocrinology:** hypothyroidism, hypergonadotrophic hypogonadism, hyperinsulinaemic hypoglycaemia **Orthopaedics:** kyphoscoliosis, joint contractures, osteopenia **Cardiology:** pericardial effusion, cardiomyopathy **Ophthalmology:** retinitis pigmentosa Multi-organ failure[73–97]	Tf-IEF type 1
PMI-CDG (Ib) >25[72]	Phosphomannose isomerase	**Gastroenterology/hepatology:** failure to thrive, protein-losing enteropathy, life-threatening intestinal bleeding, prolonged diarrhoea, congenital hepatic fibrosis, increased serum transaminases **Endocrinology:** hypoglycaemia[34–39,98–101]	Tf-IEF type 1
ALG6-CDG (Ic) >36[72]	Glucosyltransferase I	**Neurology:** moderate to severe psychomotor retardation, seizures, hypotonia, pseudotumor cerebri, epilepsy, strabismus **Gastroenterology/hepatology:** feeding difficulties **Haematology:** deep vein thrombosis **Endocrinology:** hyperandrogenism and virilization **Others:** limb deficiencies resembling brachydactyly type B[102–112]	Tf-IEF type 1

Table 2 (*Continued*)

Defects of protein N-glycosylation			
CDG type Number of reported patients	Protein	Overview of key cinical manifestations	Screening technique/alternative technique
ALG3-CDG (Id) 11[72]	Mannosyltransferase VI	**Neurology:** slowly progressive encephalopathy with microcephaly, severe psychomotor retardation and epileptic seizures **Gastroenterology/hepatology:** failure to thrive **Orthopaedics:** osteopenia[113–116]	Tf-IEF type 1
ALG12-CDG (Ig)	Mannosyltransferase VIII	**Neurology:** microcephaly, psychomotor retardation **Gastroenterology/hepatology:** intestinal malrotation with poor gastrointestinal motility **Neonatology:** short-limb skeletal dysplasia, polyhydramnios, prematurity, and generalized edema **Endocrinology:** small penis with hypospadias, hypoplastic scrotum, and non-palpable testes **Dysmorphic features:** included broad nose, thick ears, thin lips, micrognathia, inverted nipples, ulnar deviation at the wrists, spatulate fingers, fifth finger camptodactyly, nail hypoplasia, and talipes equinovarus **Cardiology:** PDA, VSD, hypertrophic cardiomyopathy, and arrhythmias. **Orthopaedics:** absent ossification of cervical vertebral bodies, pubic bones, knee epiphyses, and tali **Others:** sensorineural deafness, retinal detachment and blindness[117–122]	Tf-IEF type 1
ALG8-CDG (Ih) **9**[72]	Glucosyltransferase II	**Neurology:** multifocal myoclonic seizures, optic nerve atrophy, cerebellar hypoplasia **Gastroenterology/hepatology:** diarrhoea progressing to protein-losing enteropathy with ascites, oedema **Haematology:** bleeding **Cardiology:** pericardial effusion[123–128]	Tf-IEF type 1
ALG2-CDG (Ii) **1**[129]	Mannosyltransferase II	**Neurology:** retardation, seizures, hypomyelination **Gastroenterology/hepatology:** hepatomegaly, **Haematology:** clotting factor abnormalities (similar to PMM2-CDG) **Others:** coloboma of the iris[129]	Tf-IEF type 1

DPAGT1-CDG (Ij) **3**[72]	GlcNActransferase I	In general similar to PMM2-CDG. **Neurology:** mental retardation very severe, severe hypotonia, medically intractable seizures, microcephaly **Ophtalmology:** exotropia[130]	Tf-IEF type 1
ALG1-CDG (Ik) **7**[72]	Mannosyltransferase I	severe multisystem disorder with recurrent nonimmune hydrops fetalis[131–134]	Tf-IEF type 1
ALG9-CDG (Il) 3[135–137]	Mannosyltransferase VII	**Neurology:** psychomotor retardation, esotropia, axial hypotonia, epilepsy delayed myelination, decreased volume of cerebral hemispheres and cerebellum **Dysmorphism:** inverted nipples **Gastroenterology/hepatology:** failure to thrive, hepatosplenomegaly **Cardiology:** pericardial effusion **Others:** cystic renal disease, low levels of multiple serum proteins including antithrombin III, factor XI, and cholesterol[135–138]	Tf-IEF type 1
RFT1-CDG (In) 6[12]	RFT1 protein	**Neurology:** Developmental retardation, poor to absent visual contact, sensorineural deafness, epilepsy, microcephaly, cerebral atrophy and hypotonia **Gastroenterology/hepatology:** feeding problems and failure to thrive **Others:** dysmorphism, respiratory problems and venous thromboses[139,140]	Tf-IEF type 1
ALG11-CDG (Ip) 3[12]	Mannosyltransferase III	Multisystem disorder with muscular hypotonia, seizures, developmental retardation[141,142]	Tf-IEF type 1
MGAT2-CDG (IIa) 4[69]	GlcNActransferase II	Growth retardation, mental retardation, facial dysmorphism[23,143,144]	Tf-IEF type 2. Normal apoC-III IEF. Further confirmation by mass spectrometry
GCS1-CDG (IIb) 1[145]	Glucosidase I (GCS1)	**Neurology:** hypotonia, epilepsy; **Dysmorphism; Hepatic involvement**[145]	Normal Tf- and apoC-III IEF. Urine oligosaccharide analysis
TUSC3-CDG	Oligosaccharyltransferase TUSC3 subunit	Isolated intellectual disability[146–150]	Normal Tf-IEF and Apo C-III IEF.
MAGT1-CDG **4**[41]	Magnesium transporter protein 1	Isolated intellectual disability[150]	Genetic analysis

Table 2 (*Continued*)

Defects of protein O-glycosylation		
CDG type	Protein	Overview of key clinical manifestations
EXT1/EXT2-CDG (Multiple cartilaginous exostoses)	Extosin-1,-2	Multiple cartilage-capped tumors located primarily at the long bones, often a short stature and secondary orthopedic deformities[151]
β4GALT7-CDG (Ehlers-Danlos syndrome) 3[69]	Glucuronyltransferase/ N-acetylglucosaminyl-transferaseβ-1, 4-galactosyltransferase 7	Psychomotor retardation, short stature, osteopenia of all bones, defective deciduous teeth, delayed wound healing, hypotonia, progeroid aspect[152]
GALNT3-CDG (O-N-acetyl-galactosaminyl-glycan defect (familial tumoral calcinosis)	Polypeptide N-acetylgalactosaminyl-transferase 3	Ectopic calcifications around major joints[153]
POMGNT1-CDG (Muscle-eye-brain disease)	O-mannosyltransferase 1 O-mannose β-1,2-N-acetylglucosaminyltransferase	**Neurology, myology:** congenital muscular dystrophy, myoclonic jerks, mental retardation, hydrocephalus **Ophthalmology:** congenital glaucoma and myopia, retinal hypoplasia, pallor of the optic disc[154]
SCDO3-CDG (Spondylocostal dysostosis type 3)	O-fucose-specific β-1,3-N-acetylglucosaminyltransferase	Multiple vertebral segmentation defects and rib anomalies
B3GALTL-CDG (Peter-plus Syndrome)	Beta1,3-glucosyltransferase	**Neurology:** psychomotor retardation **Ophthalmology:** malformations of the eyes that include corneal opaqueness and iridocorneal adhesions (Peters' anomaly) **Dysmorphism:** short stature, short limbs; some have cleft lip/palate[155,156]
SLC35D1-CDG	ER UDP-glucuronic acid/UDP-*N*-acetylgalactosamine	Skeletal dysplasia comprising mainly platyspondyly, extremely short long bones, and small ilia with snail-like appearance[157]

Defects of glycosphingolipid and GPI-anchor glycosylation		
CDG type	Protein	Overview of key clinical manifestations
SIAT9-CDG (Amish infantile epilepsy)	GM3 synthase	**Neurology:** psychomotor retardation, epilepsy

PIGM-CDG	Phosphatidylinositol glycan anchor biosynthesis, class M	**Neurology:** epilepsy **Haematology:** splanchnic vein thrombosis[12]
PIGV-CDG	Phosphatidylinositol glycan anchor biosynthesis, class V	Hyperphosphatasia mental retardation syndrome (Mabry syndrome) **Neurology:** hypotonia, epilepsy **Dysmorphism:** hypertelorism, long palpebral fissures, broad nasal bridge and tip, downturned corners of the mouth, thin upper lip[12]

Defects of multiple glycosylation and other pathways

CDG type Number of reported patients	Protein	Overview of key clinical manifestations	Screening technique/alternative technique
DPM1-CDG (Ie) 8[72]	Dol-P-Man synthase 1	**Neurology:** severe developmental delay, seizures **Dysmorphism**[158,159]	Tf-IEF type 1
MPDU1-CDG (If) 5[72]	Dol-P-Man utilization protein	**Neurology:** psychomotor retardation, epilepsy **Gastroenterology/hepatology:** failure to thrive, **Dermatology:** dry skin and scaling with erythroderma, **Ophthalmology:** impaired vision, night blindness[160,161]	Tf-IEF type 1
DPM3–CDG (Io) 1[72]	Dol-P-Man synthase 3	**Neurology:** stroke-like episode **Cardiology:** cardiomyopathy, elevated serum CK, mildly elevated serum transaminases[72,162–164]	Tf-IEF type 1
GNE-CDG (Hereditary inclusion body myopathy)	Glucosamine (UDP-N-acetyl)-2-epimerase/ N-acetylmannosamine kinase	Myopathy in adults with sparing of quadriceps muscles[165]	
SLC35C1-CDG (IIc) (leukocyte adhesion deficiency type II syndrome-LAD II)	GDP-fucose transporter	**Neurology:** mental retardation Dysmorphism with growth retardation **Others:** recurrent infection, persistent high leukocytosis,[166–172]	Tf-IEF and apo C-III are normal. Alternative assay: Bombay blood group, MS (serum N-glycan profiling)

Table 2 (*Continued*)

Defects of multiple glycosylation and other pathways			
CDG type Number of reported patients	Protein	Overview of key clinical manifestations	Screening technique/alternative technique
B4GALT1-CDG (IId) **1**[68]	Galactosyltransferase	**Neurology:** hydrocephalus, muscular hypotonia, transient cholestatis **Haematology:** clotting factor abnormalities **Others:** myopathy[173,174]	TF-IEF type 2. ApoCIII is normal and MS is next step in diagnostic process
SLC35A1-CDG (IIf)	CMP-NeuAc transporter	Macrothrombocytopenia	TF-IEF and ApoCIII are normal. Alternative assay: CD15s on leukocytes
DK1-CDG (Im) **4**[69]	Dolichol kinase	**Neurology:** muscular hypotonia, nystagmus, **Cardiology:** dilated cardiomyopathy **Dermatology:** hyperkeratosis **Others:** pulmonary infections[175]	TF-IEF type 1
SRD5A3-CDG (Iq) **11**[69]	Steroid5αreductase3	**Neurology:** psychomotor retardation **Ophthalmology:** variable eye malformations, including optic nerve hypoplasia, retinal coloboma, congenital cataract and glaucoma[113,114] **Dermatology:** eczema, ichthyosis[176,177]	Tf-IEF type 1
COG7-CDG (IIe) **7**[69]	Cog7	**Neurology**: progressive, severe microcephaly, hypotonia Growth retardation **Dysmorphism:** wrinkled skin **Gastroenterology/hepatology:** feeding problems by gastrointestinal pseudo-obstruction, failure to thrive, **Cardiology:** septal defects **Others:** episodes of hyperthermia[16,19,178,179]	TF-IEF type 2. ApoCIII is abnormal. MS is next step in diagnostic process.

COG1-CDG (IIg) **3**[69]	Cog1 (COG complex subunit 1)	Cerebrocostomandibular-like syndrome	Tf-IEF type 2. ApoCIII is abnormal. MS is next step in diagnostic process.
COG4-CDG (IIj) **1**[15]	Cog4	**Neurology:** epilepsy, microcephaly, ataxia, brain MRI abnormalities **Haematology:** clotting factor abnormalities **Others:** recurrent infections[15,180]	Tf-IEF type 2. ApoCIII is abnormal. MS is next step in diagnostic process.
COG5-CDG	Cog5	**Neurology:** mild psychomotor retardation with delayed language development [18]	Tf-IEF type 2. ApoCIII is abnormal. MS is next step in diagnostic process.
COG6-CDG	Cog6	**Neurology:** intracranial bleeding with vitamin K deficiency, secondary epilepsy **Gastroenterology:** vomiting, cholestasis Early fatal outcome[181]	Tf-IEF type 2. ApoCIII is abnormal. MS is next step in diagnostic process.
COG8-CDG (IIh)	Cog8	**Neurology:** epilepsy, strabismus, oculomotor apraxia, hypotonia, microcephaly, contractures, peripheral neuropathy, ataxia, brain MRI abnormalities **Dysmorphism:** finger/toe abnormalities **Gastroenterology:** Feeding/intestinal problems, increased serum transaminases **Haematology:** clotting factor abnormalities[14,17]	Tf-IEF type 2. ApoCIII is abnormal. MS is next step in diagnostic process.
ATP6V0A2-CDG (Cutis Laxa type 2)	Vesicular $H^{(+)}$-ATPase subunit a2	**Neurology**: psychomotor retardation **Dysmorphism:** growth delay, joint anomalies, congenital cutis laxa ameliorating with age[11,182–184]	TF-IEF type 2 and apoCIII is abnormal
SEC23B-CDG (Congenital dyserythropoietic anemia type II-CDAII or HEMPAS)	Protein transport protein Sec23B isoform 1	Dyserythropoietic anemia[12]	

Early diagnosis of any genetic disease and thus also of CDG is important for several reasons such as the timely management of clinical problems, genetic counseling, and efficient treatment (PMI-CDG) or partial therapy (SLC35C1-CDG and PIGM-CDG).

3.3 Diagnosis of CDG

3.3.1 Screening techniques. Isoelectrofocusing of serum transferrin Tf (Tf-IEF) is the most widely used method to screen for N-glycosylation defects associated with sialic acid deficiency. Tf is an abundant serum protein. Most serum transferrin is tetrasialotransferrin because most molecules of this protein carry two disialylated biantennary oligosaccharide chains resulting in a total of four terminal sialic acids. In addition, some of these oligosaccharide chains may be triantennary or even tetraantennary, which explains the presence of penta and hexasialo forms. Any defect in the assembly or processing of these glycans results in hyposialylation. Since sialic acid has a negative charge, there is a cathodal shift in an electric field.

Transferrin isoforms are separated by IEF in agarose or polyacrylamide gels containing a pH gradient obtained by addition of ampholytes/ampholynes Serum samples are treated with $Fe3+$ for transferrin iron saturation. After separation, isoforms are immunofixed using anti-Tf antibodies and stained with Coomassie Brilliant Blue. This procedure permits clear identification of the altered isoforms. The altered profiles may be classified in two types. Type 1 corresponds to hypoglycosylation due to defects in the assembly of the glycan in the cytosol or the endoplasmic reticulum and is a consequence of the lack of glycan branches linked to the protein. In this type, there is an increase of asialo and disialo transferrin and a decrease of the tetrasialo band. The type 2 pattern is due to a defect in the procesing of the glycan structure linked to the protein and shows an increase of the less sialylated bands, mainly the tri-, di- and mono- and asialylated ones.[23]

Although widely used, IEF is a laborious and time consuming technique neither suitable for automation nor for accurate quantification. Therefore other more quantitative laboratory techniques such as capillary zone electrophoresis (CZE) and high pressure liquid chromatography (HPLC) have been implemented in some laboratories.

Several analytical procedures have been developed for the analysis of carbohydrate-deficient transferrin (CDT). They are extensively used for the follow-up of chronic alcohol abuse. However they indicate only the increase of hyposialylated isoforms, without identifying the altered isoforms.

3.3.2 Techniques for further investigation of CDG. After an altered pattern has been detected, it is necessary to find the enzymatic deficiency responsible for the defect. First of all, secondary CDG have to be excluded such as galactosemia, hereditary fructose intolerance, chronic alcoholism, sepsis, hepatopathy or immaturity. In these situations, the normal pattern is restored after adequate treatment.

If the pattern is **type 1**, the diagnostic algorithm is clear (Fig. 4). Since PMM2-CDG (previously CDG-Ia) is by far the most frequent CDG the

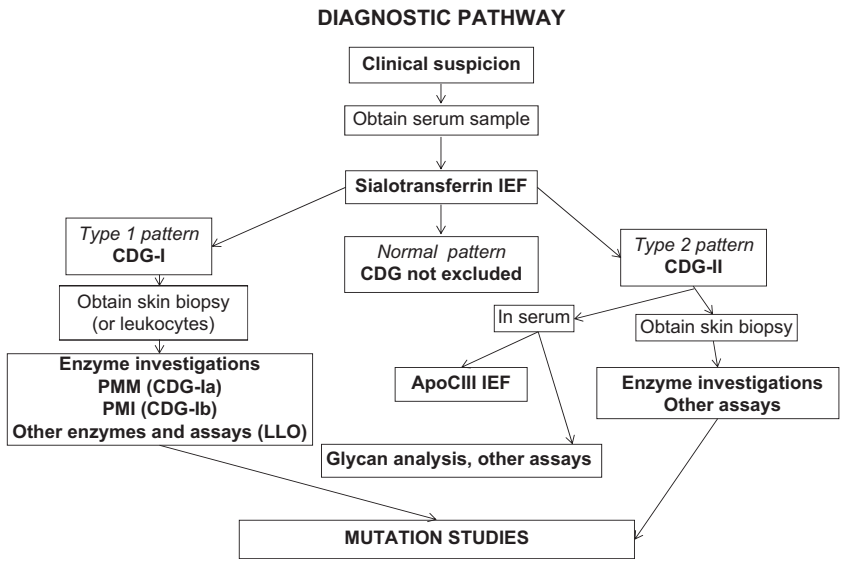


Fig. 4 Algorithm of the diagnostic pathway commonly used from clinical suspicion to reach a definitive diagnosis.

first step is to measure PMM activity in fibroblasts or leucocytes.[20] If the clinical picture suggests phosphomannose isomerase deficiency (PMI-CDG, previously CDG-Ib), PMI activity should be analyzed in leucocytes with urgency because this is a treatable disorder and it is important to introduce mannose supplementation before development of a severe enteropathy.

After PMM and PMI deficiencies have been discarded, the following step is the study of the lipid-linked oligosacharides (**LLO**) in fibroblasts cultured with radioactive mannose. The accumulated abnormal glycan(s) will often point to the precise altered step. This method allows identification of most of the type I defects.[24] In defects before the addition of mannoses to $GlcNAc_2$-PP-dolichol the LLO accumulation is normal or low.

For type 2 patterns, the algorithm is not so well defined (Fig. 4). After excluding secondary causes, the following step is the study of the isoforms of **serum apoCIII by IEF**. This protein is purely O-glycosylated protein. IEF of this protein permits to exclude or confirm core 1 mucin-type O-glycosylation defects. A type 2 transferrin pattern together with an altered apo C-III pattern is seen in some CDG types.[25]

In the presence of type 2 pattern, it is also indicated to analyse the abnormal structures of the protein-linked glycans by electrospray mass spectrometry (ES-MS) or MALDI-TOF, as they are useful tools to suggest the candidate step responsible for the altered glycan synthesis.

Other tools that have been used for the diagnosis of CDG-II are PNA lectin staining which detects galactose residues of O-glycans not protected by a terminal sialic acid residue. Treatment of fibroblasts with Brefeldin A (BFA) is used to assess a possible trafficking defect in the Golgi. BFA is a selective inhibitor of the GTP-exchange factor for ADP ribosylation factor 1 (Arf1) and causes the redistribution of Golgi proteins to the ER *via* retrograde transport. In COG patients, the BFA-induced redistribution is significantly delayed.

3.3.3 Final CDG diagnosis. To establish a definite diagnosis of CDG it is necessary to demonstrate the protein deficiency and/or the genetic mutation(s). Enzymatic procedures are not available for all the enzymes involved in CDG, but besides PMM and PMI some other enzymes are measurable. In all cases, it is indispensable to achieve a precise diagnosis by detection of the disease-causing mutation(s) allowing for a reliable genetic counseling.

As numerous genes are involved in the glycosylation pathways and many CDG-X patients remain undiagnosed, new approaches should be used for their study. At present, whole exome sequencing (targeted exome capture) is being used as a strategy to selectively sequence the coding regions of the human genome to identify functional variations associated with a phenotype or pathology. The exome sequencing is based on a first step of enrichment of the targeted regions of the DNA (mainly exome) by hybridization with probes followed by next generations sequencing. The obtained data are treated and analyzed to try to pinpoint the genetic defect.

3.4 Genetics of CDG

The known CDG are transmitted in an autosomal recessive way except for EXT1/EXT2-CDG (autosomal dominant) and MAGTI-CDG (X-linked).

The most common type of CDG is due to mutations in the *PMM2* gene (MIM#601785), located on chromosome 16p13. It encodes the PMM protein (EC 5.4.2.8, PMM 2). Up to date more than one hundred disease-causing mutations have been described in this gene (HGMD® professional release, http://www.hgmd.cf.ac.uk/ac/index.php). The most common (nearly 85%) are missense changes; patients with this type of changes are usually compound heterozygous for two different missense mutations. PMM2-CDG patients show a strong genetic heterogeneity, although one severe mutation, p.R141H, has been found particularly common in all cohorts of patients analyzed.[26–29] Specific missense changes have been functionally analyzed to improve the understanding of the molecular mechanisms of the disease.[27,30–32]

3.5 Treatment of CDG

3.5.1 Specific treatment. An efficient, specific treatment is available for only one CDG, namely PMI-CDG (CDG-Ib). In this CDG, GDP-Man can still be synthesized from free mannose entering the cell and directly phosphorylated to Man-6-P, overcoming the defect. Thus, the simple daily oral mannose supplementation is an effective treatment for patients with this potentially lethal disorder.[33–39]

Some patients with SLC35C1-CDG or leukocyte adhesion deficiency type II respond to simple supplementation of oral fucose. Finally, in PIGM-CDG that causes histone hypoacetylation, butyrate increases *PIGM* transcription and its supplementation was able to control seizures in the one reported patient.[40]

3.5.2 Symptomatic treatments and preventive measures. Table 3 summarizes symptomatic treatments for CDG. In addition, occupational therapy, physical therapy, and speech therapy are importat. Of note is that before surgery the consult of a hematologist, familiar with CDG, is indicated.

Table 3 Symptomatic management for CDGs without specific treatment.

Clinical feature	
Failure to thrive	Consider nasogastric tube or gastrostomy tube feeding[41]
Strabismus	Glasses, patching, or surgery[41]
Stroke-like episodes	Low doses of aspirin[42,43]
Oral motor dysfunction	Consider consultation with a gastroenterologist/nutritionist[41]
Osteopenia	Bisphosphonate therapy in severe forms[44,45]
Pericardial effusion	Corticosteroids and salicylic acid, pericardial drainage[46]
Scoliosis/kyphosis	Surgical treatment in severe forms[41]
True hypothyroidism	L-thyroxine supplementation[41]

3.5.3 Possible new treatment strategies. Treatment of PMM2-CDG patients with Man1-P from external sources is ineffective due to its high polarity, its inability to penetrate through the cellular membrane and its instability in blood. Different chemically synthesized membrane-permeable derivatives of Man-1-P have been produced in order to increase GDP-Man pools. Some have been found to restore LLO biosynthesis and to increase the level of GDP-Man in PMM2-CDG patients' fibroblasts,[4,47,48] but further work has to be done in order to decrease their toxicity and enhance their stability. Other types of CDG that rely on GDP-Man availability might benefit as well from this approach.

Another study showed that metformin-stimulated D-mannose transport (MSMT) increased mannose uptake by skin fibroblasts from PMM2-CDG (CDG-Ia) patients increasing the LLO precursor pool. However, a therapeutic administration of metformin is questionable because mannose by itself is ineffective and the MSMT effect in other tissues is unknown.[49,50]

Recently, Dahl and colleagues[51] discovered a series of potent inhibitors of PMI with *in vivo* efficacy when combined with the mannose supplementation in disease models, as a strategy to redirect more Man-6-P towards the depleted glycosylation pathway.

In a more recent study,[50] the squalene synthase inhibitor Zaragozic acid was reported to improve dolichol-linked oligosaccharide biosynthesis in DPM1-CDG fibroblasts representing a therapeutic possibility for N-linked glycosylation defects with residual enzymatic activity.

Furthermore, a report showed that functional and immunoreactive PMM protein was obtained using antisense morpholino (AMO) in cultured fibroblasts carrying three different splicing mutations in the *PMM2* gene proving to be a promising mutational-specific therapy option.[32]

The most common disease causing mutations in PMM2-CDG (CDGIa) are missense changes which probably cause misfolding effects leading to protein degradation. Recently, a remarkable study suggested that the loss-of-function of several mutations in PMM2-CDG is based on their increased susceptibility to degradation or aggregation.[52] Therefore, pharmacological chaperones (efficient in diseases such as phenylketonuria) may hold promise for a large group of PMM2-CDG patients.[53]

3.6 CDG model systems

Biological model systems can contribute to a better understanding of the pathophysiology of CDG and provide information needed to devise new

Table 4 Summary of Biological model systems for CDGs.

Biological model systems	Advantages	Disadvantages
Saccharomyces cerevisiae	Rapid molecular diagnostic tool, inexpensive, abundant growth and good genetic analysis	Not useful for complex glycans typical of mammalian cells; some proteins not present in yeast (*i.e.* human LEC 35). Golgi glycosylation pathways differ between human and yeast; unable to predict physiological effects in patients
Mammalian cell lines	Relatively easily manipulated and transformed	Do not reveal the physiological functions of glycans in multicellular organisms
C. elegans	Well characterized; small (1–1.5 mm); easy to cultivate; short life cycle (3 days); genome completely sequenced; self-fertilizing (hermaphrodite)	Biochemistry and pathology are difficult; cells post-reproductive; lower invertebrate; many physiological functions not conserved (immune system)[64]
Drosophila melanogaster	Short generation time (2 weeks); cheap and easy to manipulate; results extrapolable to humans (many homologous disease genes); established methods and tools; suitable to model diseases at the level of tissues or cell communication[65]	Life cycle longer than *S. cerevisiae* and *C. elegans*; not useful for biological processes evolved in vertebrates
Zebrafish (*Danio rerio*)	Good to observe (small and transparent); relatively rapid life cycle, easy breeding, growing list of molecular tools[66]	Knock-out methods not available yet (alternative use of Morpholinos); many genetic tools still lacking
Mouse	Short generation time (8–10 weeks); easy to handle; low cost of maintenance; allows embryonic, physiological and therapeutical evaluation; high identity between mice and human orthologues	Excessive complexity; problems with embryonic lethality or absence of phenotype, which may depend on the strain

treatments that will have an impact on quality of life and life expectancy of these patients. In addition, these models allow researchers to test therapeutic strategies previously mentioned. Mammalian cell lines, *Saccharomyces cerevisiae*, *C. elegans*, zebrafish and mice have been studied in this context (Table 4).

Saccharomyces cerevisiae has been an important model to elucidate the molecular cause of several CDG types because there is a significant conservation of the N-linked glycosylation assembly between yeast and mammalian cells.[54,55] Because Golgi glycosylation differs considerably between yeast and animals, this model has been of little help to study Golgi-related diseases.

Different mammalian cell lines with mutations affecting the biosynthesis of N-linked oligosaccharides such as Chinese hamster ovary (CHO) cells,[56] mouse BW5147 lymphoma cells and baby hamster kidney (BHK) are

currently available to test putative human mutations for which no yeast ortholog exists.[54,57]

In an attempt to better understand CDG pathogenesis, Struwe and colleagues[58,59] studied Caenorhabditis *elegans* to model N-linked glycosylation defects. Several candidate genes have been identified as possibly responsible for disease severity in humans.

Several other animal models have been important for the study of CDG. For example, a study of zebrafish as a model for human SLC35C1-CDG (CDGIIc) where it was shown for the first time that defects in protein fucosylation lead to altered neuronal differentiation, maintenance, axon branching and synapse formation.[60] The gene responsible for this pathology was found to encode a GDP-fucose transporter. It was found that the Drosophila genome encodes at least another GDP-fucose transporter that is involved in the O-fucosylation of Notch signaling.[61]

Mouse models provide an excellent tool to better understand the effects of underglycosylation in an whole organism.[8] In fact, different studies showed that the complete loss of enzymatic activity by ablation of *PMM2* or *PMI* genes is lethal,[54,57,62] highlighting the relevance of glycosylation for the proper function of the organism from the beginning of the embryonic development. Nearly all CDG patients appear to have variable degrees of residual enzymatic activity. Therefore, the generation of hypomorphic mouse models, is more likely to have a positive result by the introduction of specific mutations in the gene of interest rather than by its total ablation. Recently, a remarkable study used frequent mutations causing severe disease in PMM2-CDG (CDGIa) patients. It yielded embryos with a severe phenotype that was reversible by feeding mannose to pregnant dams.[63] Noteworthy, this model mimics the most common human CDG, and thus holds great promise for the development of therapeutic strategies.

3.7 Socio-familial aspects for CDG patients

Non-profit patient organizations are growing rapidly as well as the number of patients and the different types of rare diseases. Their mission mainly comprises: support, education, research, awareness and advocacy. Their participation in the decision-making process is being accompanied by a greater awareness and different converging trends that are helping to intensify their influence.

Problems faced by health professionals and CDG patients include: (1) Limited knowledge of experts and support services; (2) Restricted access to specialized health services or approved therapies; (3) Major lack of information targeted to families about the ongoing basic and clinical research; (4) Difficulties of providing advice, diagnosis and health interventions for these rare diseases; (5) Lack of (national and international) financial resources; and (6) Need for better cooperation between researchers, medical doctors and patients.

CDG patient organizations aim to increase knowledge, awareness and dissemination of CDG and foster cooperation among basic scientists, clinicians and family advocacy groups.

Actions to reach this goal include coordination and organization of scientific and medical communication activities, establishment of collaboration with

key stakeholders such as Guía Metabólica (a powerful and interactive tool targeted to families affected by inborn errors of metabolism); Fostering and encouragement of research, inform and disseminate information amongst families and experts, empowerment of the CDG community by enhancement of understanding, training, awareness and advocacy for better quality of life for patients and their families (at the national and international level).

Interestingly, several initiatives of CDG awareness and dissemination have been established in partnership between CDG organizations, researchers and clinicians such as the : (1) Practical guide targeted to CDG families; (2) Fairy-tale: ''Glycoland and the coloured antennas'' (3) CDG awareness and dissemination tool (4) CDG Online community (5) International CDG logo (6) CDG community store, and others.

Our working model can be used by other Rare Metabolic Disorders.

4 Concluding remarks and future perspectives: CDG as a challenge for glycobiomedicine in the next decade!

About 30 years ago the first patients with a CDG were described by Jaak Jaeken. Following his ground-breaking discovery, a large number of publications and enormous advances have been made in our understanding of CDG.

Several recent reviews cover the discovery of new types of CDG, the clinical spectrum, animal models and advances towards a therapy. Amongst the most relevant are certainly those that shed light on the importance of improving and establishing effective tools for (early) diagnosis aiming for a better quality of life outcomes of children and adults with CDG.

Here we focused on the number of significant tools that are currently available in a growing number of expert clinical centres to screen and validate CDG. The blood test to study the glycosylation status of transferrin continues to be the most frequent indicator of CDG. However, technologies such as carbohydrate microarrays and glycan sequencing profile will probably become more commonly used to dissect the role/function of glycans, glycosylation pathway, and the properties of glycoproteins. Overall, glycans role in biology and in particular in this group of pathologies will be deciphered.

Recent developments in diagnostic sequencing tools to elucidate the molecular basis of genetic diseases, the exome sequencing, will contribute to a faster diagnosis after the CDG clinical suspicion. However, to speed-up diagnosis international collaborations involving multidisciplinary working teams are necessary. They must include cell biologists, clinicians, molecular biologists, biochemists, biostatisticians (rather than informatics) and others. Initiatives such as the exome variant server (http://evs.gs.washington.edu/EVS/), accurate phenotyping and functional validation of results using different approaches such as new animal models, cellular phenotyping (using human cells) and biochemical assays are an added value to a faster and early diagnosis.

Recently, sequencing of individual genomes (whole exome sequencing) became affordable and accessible for research and potentially for clinical applications. Every experimental approach offers advantages and drawbacks and special caution must be used in the attempt to examine data

obtained from this technique: ''accidental'' findings for other possible disease conditions will be obtained. It is important to remember that only few CDGs are completely or partially treatable, as illustrated by the example of PMI-CDG (CDG Ib) in which mannose supplementation is life saving. Other therapeutic possibilities are on the horizon.

Patients suffer from the functional consequences of their mutations. Thus, it is extremely important to understand the mechanisms that underlie the genetic disease, and to devise better specific and symptomatic treatments.

The lack of knowledge on CDG often puts the life of patients at risk and it is likely that many patients with a CDG have a wrong diagnosis leading to multiple medical consultations and prescription of inappropriate or even harmful drugs and treatments. Many patients are not receiving the appropriate treatment from physiotherapists, nutritionists, psychologists, *etc*...

CDG patients will benefit from the data pooled in the International CDG patient registry as it allows a sufficient sample size for epidemiological research, clinical surveillance, future treatment evaluation (efficacy) and monitoring (safety), genotype-phenotype correlation and finally patient quality of life outcomes.

Finally, international collaborations between affected families, clinicians and researchers will continue to foster and enhance research in this field.

There is no doubt that Congenital Disorders of Glycosylation are an exciting and expanding field that will attract the attention of many ''glycobiomedical doctors'' or ''carbohydrate-lovers'' leading to promising discoveries in the next decade.

5 Highlights

- **Congenital (genetic) disorders of glycosylation** are a booming group of diseases in which developments in genetics and glycan disciplines will lead to exciting discoveries in the near future.
- Development of **diagnostic tools:** in particular new CDG screening and validation methods are needed.
- **Conventional** and **symptomatic treatments:** their development should receive special attention from clinicians.
- Congenital Disorders of Glycosylation **cut across many medical specialties:** thus reference centers are urgently needed.
- Development of **animal models** are important in understanding the pathogenesis of these diseases. Recently, a mouse model for the most common form of CDG was developed. Other animal models which still are under-utilized in CDG will probably gain popularity, such as, zebrafish and Drosophila.
- **Awareness and dissemination:** Special attention should be given to initiatives developed by CDG patients' groups in partnership with clinicians and researchers.
- **Patients' registry:** Optimization of the resources currently available and consensus amongst researchers in order to continue pooling and sharing data through the International CDG registry is highly recommended. In addition, patient self-registration for the collection of data might be helpful.

• To **foster and encourage research** more initiatives like ''Horizon 2020'', IRDiRC (International Consortium on Rare Disease Research), E-Rare Call for Proposals 2012 or Sanco-Health funding are urgently need.

6 Further information

CDG patients and their interest groups are highly dynamic worldwide. Their contact information is the following:

Portugal: http://sindromecdg.orgfree.com/
Spain: http://webs.ono.com/aescdg/
France: http://www.lesptitscdg.org/
USA: http://www.cdgfamilynetwork.org/
Canada: http://www.thefog.ca/
http://thelittlefightersfoundation.com/
Germany: https://www.cdg-syndrom.de/
Denmark: http://www.cdgforeningen.dk/
Sweden: http://www.cdgs.se/

If CDG is suspected, a serum transferring IZEF test should be performed. In Europe and United states several referral centers offer diagnosis of CDG.

Abbreviations

AMO	Antisense morpholino
Asn	Asparagine
Arf1	ADP ribosylation factor 1
BFA	Brefeldin A
BHK	Baby hamster kidney
CDG	Congenital Disorders of Glycosylation
CDT	Carbohydrate-deficient transferrin
CTP	Cytidine triphosphate
COG	Conserved Oligomeric Golgi
CHO	Chinese hamster ovary
CZE	Capillary zone electrophoresis
Dol-P	Dolichol phosphate
ER	Endoplasmic reticulum
ES-MS	Electrospray mass spectrometry
Fuc	Fucose
Gal	Galactose
GDP-Man	Guanosinediphosphate-mannose
GlcNAc	N-acetylglucosamine
Hk	Hexokinase
HPLC	High pressure liquid chromatography
Tf-IEF	Isoelectrofocusing of serum transferrin
LLO	Lipid-Linked Oligosaccharide
Man-6-P	Mannose-6-phosphate
Man-1-P	Mannose-1-phosphate
MSMT	Metformin-stimulated D-mannose transport
OST	Oligosaccharyltransferase
PMM2	Phosphomannomutase 2

PMI	Phosphomannose isomerase
Ser	Serine
NANA	N-acetyl-neuraminic acid
Thr	Threonine
UDP-GlcNAc	Uridine-diphospho-N-acetylglucosamine

Acknowledgements

This chapter is dedicated to Dr. Jaak Jaeken (Leuven, Belgium) and to Liliana Ferreira (Lisbon, Portugal). We thank Dr. Jaeken for his ongoing scientific, medical and humane input which has contributed tremendously to expanding CDG knowledge and awareness. His commitment to patient organizations is exceptional and we think he is a successful model to follow. We would like to thank Liliana Ferreira who is an inspiration and a driving force for CDG patient advocacy.

We are also grateful to Dr. Gert Matthjis (Leuven, Belgium) for his dedication to CDG research at the European level: Euroglycan and Euroglycanet greatly improved CDG (early) diagnosis and established successful international collaborations amongst clinical and basic research groups. We would like to express our gratitude also to all families and CDG experts whose encouragement and support for CDG patient activities enabled the development of many projects in a productive partnership. The final goal is to expand CDG research and to improve patients' disease outcome.

References

1 R. Apweiler, H. Hermjakob and N. Sharon, *Biochim Biophys Acta*, 1999, **1473**, 4–8.
2 H. Schachter and H. H. Freeze, *Biochim Biophys Acta*, 2009, **1792**, 925–930.
3 J. S. Rush, N. Gao, M. A. Lehrman and C. J. Waechter, *J Biol Chem*, 2008, **283**, 4087–4093.
4 R. Hardre, A. Khaled, A. Willemetz, T. Dupre, S. Moore, C. Gravier-Pelletier and Y. Le Merrer, *Bioorg Med Chem Lett*, 2007, **17**, 152–155.
5 T. Marquardt and J. Denecke, *Eur J Pediatr*, 2003, **162**, 359–379.
6 D. J. Kelleher and R. Gilmore, *Glycobiology*, 2006, **16**, 47R–62R.
7 E. Mohorko, R. Glockshuber and M. Aebi, *J Inherit Metab Dis*, **34**, 869–878.
8 C. Thiel and C. Korner, *J Inherit Metab Dis*, **34**, 879–889.
9 S. Wopereis, D. J. Lefeber, E. Morava and R. A. Wevers, *Clin Chem*, 2006, **52**, 574–600.
10 W. Vleugels, E. Schollen, F. Foulquier and G. Matthijs, *Biochem Biophys Res Commun*, 2009, **390**, 769–774.
11 M. Mohamed, D. Kouwenberg, T. Gardeitchik, U. Kornak, R. A. Wevers and E. Morava, *J Inherit Metab Dis*, 2011, **34**, 907–916.
12 J. Jaeken, *J Inherit Metab Dis*, 2011, **34**, 853–858.
13 F. Foulquier, E. Vasile, E. Schollen, N. Callewaert, T. Raemaekers, D. Quelhas, J. Jaeken, P. Mills, B. Winchester, M. Krieger, W. Annaert and G. Matthijs, *Proc Natl Acad Sci USA*, 2006, **103**, 3764–3769.
14 F. Foulquier, D. Ungar, E. Reynders, R. Zeevaert, P. Mills, M. T. Garcia-Silva, P. Briones, B. Winchester, W. Morelle, M. Krieger, W. Annaert and G. Matthijs, *Hum Mol Genet*, 2007, **16**, 717–730.
15 E. Reynders, F. Foulquier, E. Leao Teles, D. Quelhas, W. Morelle, C. Rabouille, W. Annaert and G. Matthijs, *Hum Mol Genet*, 2009, **18**, 3244–3256.

16 R. Zeevaert, F. Foulquier, D. Cheillan, I. Cloix, N. Guffon, L. Sturiale, D. Garozzo, G. Matthijs and J. Jaeken, *Eur J Med Genet*, 2009, **52**, 303–305.
17 C. Kranz, B. G. Ng, L. Sun, V. Sharma, E. A. Eklund, Y. Miura, D. Ungar, V. Lupashin, R. D. Winkel, J. F. Cipollo, C. E. Costello, E. Loh, W. Hong and H. H. Freeze, *Hum Mol Genet*, 2007, **16**, 731–741.
18 P. Paesold-Burda, C. Maag, H. Troxler, F. Foulquier, P. Kleinert, S. Schnabel, M. Baumgartner and T. Hennet, *Hum Mol Genet*, 2009, **18**, 4350–4356.
19 X. Wu, R. A. Steet, O. Bohorov, J. Bakker, J. Newell, M. Krieger, L. Spaapen, S. Kornfeld and H. H. Freeze, *Nat Med*, 2004, **10**, 518–523.
20 M. Aebi, A. Helenius, B. Schenk, R. Barone, A. Fiumara, E. G. Berger, T. Hennet, T. Imbach, A. Stutz, C. Bjursell, A. Uller, J. G. Wahlstrom, P. Briones, E. Cardo, P. Clayton, B. Winchester, V. Cormier-Dalre, P. de Lonlay, M. Cuer, T. Dupre, N. Seta, T. de Koning, L. Dorland, F. de Loos and L. Kupers, *et al*., *Glycoconj J*, 1999, **16**, 669–671.
21 J. Jaeken, T. Hennet, H. H. Freeze and G. Matthijs, *J Inherit Metab Dis*, 2008, **31**, 669–672.
22 J. Jaeken, T. Hennet, G. Matthijs and H. H. Freeze, *Biochim Biophys Acta*, 2009, **1792**, 825–826.
23 J. Jaeken, *J Inherit Metab Dis*, 2003, **26**, 99–118.
24 P. Burda, L. Borsig, J. de Rijk-van Andel, R. Wevers, J. Jaeken, H. Carchon, E. G. Berger and M. Aebi, *J Clin Invest*, 1998, **102**, 647–652.
25 S. Wopereis, E. Morava, S. Grunewald, M. Adamowicz, K. M. Huijben, D. J. Lefeber and R. A. Wevers, *Glycobiology*, 2005, **15**, 1312–1319.
26 G. Matthijs, E. Schollen, E. Van Schaftingen, J. J. Cassiman and J. Jaeken, *Am J Hum Genet*, 1998, **62**, 542–550.
27 C. Le Bizec, S. Vuillaumier-Barrot, A. Barnier, T. Dupre, G. Durand and N. Seta, *Hum Mutat*, 2005, **25**, 504–505.
28 S. Kjaergaard, F. Skovby and M. Schwartz, *Eur J Hum Genet*, 1998, **6**, 331–336.
29 P. Briones, M. A. Vilaseca, E. Schollen, I. Ferrer, M. Maties, C. Busquets, R. Artuch, L. Gort, M. Marco, E. van Schaftingen, G. Matthijs, J. Jaeken and A. Chabas, *J Inherit Metab Dis*, 2002, **25**, 635–646.
30 M. Pirard, G. Matthijs, L. Heykants, E. Schollen, S. Grunewald, J. Jaeken and E. van Schaftingen, *FEBS Lett*, 1999, **452**, 319–322.
31 S. Kjaergaard, F. Skovby and M. Schwartz, *Eur J Hum Genet*, 1999, **7**, 884–888.
32 A. I. Vega, C. Perez-Cerda, L. R. Desviat, G. Matthijs, M. Ugarte and B. Perez, *Hum Mutat*, 2009, **30**, 795–803.
33 C. J. Hendriksz, P. McClean, M. J. Henderson, D. G. Keir, V. C. Worthington, F. Imtiaz, E. Schollen, G. Matthijs and B. G. Winchester, *Arch Dis Child*, 2001, **85**, 339–340.
34 H. K. Harms, K. P. Zimmer, K. Kurnik, R. M. Bertele-Harms, S. Weidinger and K. Reiter, *Acta Paediatr*, 2002, **91**, 1065–1072.
35 R. Niehues and M. Hasilik, *MMW Fortschr Med*, 2000, **142**, 171–172.
36 V. Westphal, S. Kjaergaard, J. A. Davis, S. M. Peterson, F. Skovby and H. H. Freeze, *Mol Genet Metab*, 2001, **73**, 77–85.
37 E. Martin Hernandez, A. I. Vega Pajares, B. Perez Gonzalez, M. J. Ecay Crespo, F. Leal Perez, J. Manzanares Lopez-Manzanares, M. Ugarte Perez and C. Perez-Cerda Silvestre, *An Pediatr (Barc)*, 2008, **69**, 358–365.
38 R. Y. Tamminga, D. J. Lefeber, W. A. Kamps and F. J. van Spronsen, *Pediatr Hematol Oncol*, 2008, **25**, 762–768.
39 P. de Lonlay and N. Seta, *Biochim Biophys Acta*, 2009, **1792**, 841–843.
40 A. Almeida, M. Layton and A. Karadimitris, *Biochim Biophys Acta*, 2009, **1792**, 874–880.

41 S. E. Sparks and D. M. Krasnewich, 1993.
42 J. B. Arnoux, N. Boddaert, V. Valayannopoulos, S. Romano, N. Bahi-Buisson, I. Desguerre, Y. de Keyzer, A. Munnich, F. Brunelle, N. Seta, M. D. Dautzenberg and P. de Lonlay, *Mol Genet Metab*, 2008, **93**, 444–449.
43 N. Ishikawa, G. Tajima, H. Ono and M. Kobayashi, *Brain Dev*, 2009, **31**, 240–243.
44 D. Coman, M. Irving, P. Kannu, J. Jaeken and R. Savarirayan, *Clin Genet*, 2008, **73**, 507–515.
45 D. Coman, D. Bostock, M. Hunter, P. Kannu, M. Irving, V. Mayne, M. Fietz, J. Jaeken and R. Savarirayan, *Am J Med Genet A*, 2008, **146**, 389–392.
46 B. J. Feldman and D. Rosenthal, *Pediatr Cardiol*, 2002, **23**, 469–471.
47 S. Rutschow, J. Thiem, C. Kranz and T. Marquardt, *Bioorg Med Chem*, 2002, **10**, 4043–4049.
48 E. A. Eklund, N. Merbouh, M. Ichikawa, A. Nishikawa, J. M. Clima, J. A. Dorman, T. Norberg and H. H. Freeze, *Glycobiology*, 2005, **15**, 1084–1093.
49 J. Shang and M. A. Lehrman, *J Biol Chem*, 2004, **279**, 9703–9712.
50 M. A. Haeuptle, M. Welti, H. Troxler, A. J. Hulsmeier, T. Imbach and T. Hennet, *J Biol Chem*, **286**, 6085–6091.
51 R. Dahl, Y. Bravo, V. Sharma, M. Ichikawa, R. P. Dhanya, M. Hedrick, B. Brown, J. Rascon, M. Vicchiarelli, A. Mangravita-Novo, L. Yang, D. Stonich, Y. Su, L. H. Smith, E. Sergienko, H. H. Freeze and N. D. Cosford, *J Med Chem*, **54**, 3661–3668.
52 A. I. Vega, C. Perez-Cerda, D. Abia, A. Gamez, P. Briones, R. Artuch, L. R. Desviat, M. Ugarte and B. Perez, *J Inherit Metab Dis*, **34**, 929–939.
53 H. H. Freeze, *Biochim Biophys Acta*, 2009, **1792**, 835–840.
54 T. Marquardt and H. Freeze, *Biol Chem*, 2001, **382**, 161–177.
55 M. Aebi and T. Hennet, *Trends Cell Biol*, 2001, **11**, 136–141.
56 M. B. Jones, N. Tomiya, M. J. Betenbaugh and S. S. Krag, *Biochem Biophys Res Commun*, **395**, 36–41.
57 H. H. Freeze, *Biochim Biophys Acta*, 2002, **1573**, 388–393.
58 W. B. Struwe, B. L. Hughes, D. W. Osborn, E. D. Boudreau, K. M. Shaw and C. E. Warren, *Glycobiology*, 2009, **19**, 1554–1562.
59 W. B. Struwe and C. E. Warren, *Methods Enzymol*, **480**, 477–493.
60 Y. Song, J. R. Willer, P. C. Scherer, J. A. Panzer, A. Kugath, E. Skordalakes, R. G. Gregg, G. B. Willer and R. J. Balice-Gordon, *PLoS One*, 2010, **5**, e13743.
61 H. O. Ishikawa, S. Higashi, T. Ayukawa, T. Sasamura, M. Kitagawa, K. Harigaya, K. Aoki, N. Ishida, Y. Sanai and K. Matsuno, *Proc Natl Acad Sci USA*, 2005, **102**, 18532–18537.
62 C. Thiel, T. Lubke, G. Matthijs, K. von Figura and C. Korner, *Mol Cell Biol*, 2006, **26**, 5615–5620.
63 A. Schneider, C. Thiel, J. Rindermann, C. Derossi, D. Popovici, G. F. Hoffmann, H. J. Grone and C. Korner, *Nat Med*, 2011, **18**, 71–73.
64 T. E. Johnson, *Exp Gerontol*, 2003, **38**, 1329–1332.
65 E. Bier, *Nat Rev Genet*, 2005, **6**, 9–23.
66 B. Novoa and A. Figueras, *Adv Exp Med Biol*, **946**, 253–275.
67 P. de Lonlay, V. Valayannopoulos, T. Dupre, S. Vuillaumier-Barrot and N. Seta, *Arch Pediatr*, 2008, **15**, 602–605.
68 J. Jaeken and G. Matthijs, *Annu Rev Genomics Hum Genet*, 2007, **8**, 261–278.
69 J. Jaeken, *Ann N Y Acad Sci*, 2010, **1214**, 190–198.
70 E. Morava and D. Lefeber, *J Inherit Metab Dis*, 2011, **34**, 847–848.
71 D. J. Lefeber, E. Morava and J. Jaeken, *J Inherit Metab Dis*, 2011, **34**, 849–852.
72 M. A. Haeuptle and T. Hennet, *Hum Mutat*, 2009, **30**, 1628–1641.

73 S. E. Sparks and D. M. Krasnewich, 1993.
74 E. Van Schaftingen and J. Jaeken, *FEBS Lett*, 1995, **377**, 318–320.
75 B. Kristiansson, H. Stibler, B. Hagberg and J. Wahlstrom, *Lakartidningen*, 1998, **95**, 5742–5748.
76 H. Carchon, E. Van Schaftingen, G. Matthijs and J. Jaeken, *Biochim Biophys Acta*, 1999, **1455**, 155–165.
77 C. H. van Ommen, M. Peters, P. G. Barth, P. Vreken, R. J. Wanders and J. Jaeken, *J Pediatr*, 2000, **136**, 400–403.
78 T. Dupre, M. Cuer, S. Barrot, A. Barnier, V. Cormier-Daire, A. Munnich, G. Durand and N. Seta, *Clin Chem*, 2001, **47**, 132–134.
79 S. Kjaergaard, M. Schwartz and F. Skovby, *Arch Dis Child*, 2001, **85**, 236–239.
80 S. Kjaergaard, J. Muller and F. Skovby, *Arch Dis Child*, 2002, **87**, 324–327.
81 I. Mader, M. Dobler-Neumann, W. Kuker, H. Stibler and I. Krageloh-Mann, *Childs Nerv Syst*, 2002, **18**, 77–80.
82 G. Sala, T. Dupre, N. Seta, P. Codogno and R. Ghidoni, *Pediatr Res*, 2002, **52**, 645–651.
83 T. Honzik, E. Malonova, H. Hansikova, R. Rosipal, H. Poupetova, P. Martasek and J. Zeman, *Cas Lek Cesk*, 2003, **142**, 276–279.
84 L. M. Neumann, A. von Moers, J. Kunze, O. Blankenstein and T. Marquardt, *Eur J Pediatr*, 2003, **162**, 710–713.
85 K. L. Chu, Y. H. Chien, C. E. Tsai, H. H. Freeze, E. Eklund and W. L. Hwu, *J Formos Med Assoc*, 2004, **103**, 721–723.
86 E. Aronica, A. A. van Kempen, M. van der Heide, B. T. Poll-The, H. J. van Slooten, D. Troost and J. M. Rozemuller-Kwakkel, *Acta Neuropathol*, 2005, **109**, 433–442.
87 I. Giurgea, A. Michel, M. Le Merrer, N. Seta and P. de Lonlay, *Pediatr Neurol*, 2005, **32**, 121–123.
88 D. Coman, J. McGill, R. MacDonald, D. Morris, S. Klingberg, J. Jaeken and D. Appleton, *J Clin Neurosci*, 2007, **14**, 668–672.
89 D. Krasnewich, K. O'Brien and S. Sparks, *Am J Med Genet C Semin Med Genet*, 2007, **145C**, 302–306.
90 D. Wurm, A. Hansgen, Y. J. Kim, A. Lindinger, A. Baghai and L. Gortner, *Eur J Pediatr*, 2007, **166**, 377–378.
91 D. Wurm, G. Loffler, A. Lindinger and L. Gortner, *J Perinatol*, 2007, **27**, 802–804.
92 R. Barone, L. Sturiale, V. Sofia, A. Ignoto, A. Fiumara, G. Sorge, D. Garozzo and M. Zappia, *Am J Med Genet A*, 2008, **146A**, 2103–2108.
93 J. M. Brum, I. M. Rizzo, W. D. Mello and C. E. Speck-Martins, *Arq Neuropsiquiatr*, 2008, **66**, 545–548.
94 S. Grunewald, *Biochim Biophys Acta*, 2009, **1792**, 827–834.
95 E. Jamroz, D. Adamek, J. Paprocka, M. Adamowicz, E. Marszal and R. A. Wevers, *J Child Neurol*, 2009, **24**, 13–18.
96 B. Perez-Duenas, A. Garcia-Cazorla, M. Pineda, P. Poo, J. Campistol, V. Cusi, E. Schollen, G. Matthijs, S. Grunewald, P. Briones, C. Perez-Cerda, R. Artuch and M. A. Vilaseca, *Eur J Paediatr Neurol*, 2009, **13**, 444–451.
97 M. K. Thong, M. Fietz, C. Nicholls, M. H. Lee and O. Asma, *J Inherit Metab Dis*, 2009.
98 J. Jaeken, G. Matthijs, J. M. Saudubray, C. Dionisi-Vici, E. Bertini, P. de Lonlay, H. Henri, H. Carchon, E. Schollen and E. Van Schaftingen, *Am J Hum Genet*, 1998, **62**, 1535–1539.
99 E. Schollen, L. Dorland, T. J. de Koning, O. P. Van Diggelen, J. G. Huijmans, T. Marquardt, D. Babovic-Vuksanovic, M. Patterson, F. Imtiaz, B.

Winchester, M. Adamowicz, E. Pronicka, H. Freeze and G. Matthijs, *Hum Mutat*, 2000, **16**, 247–252.
100 D. F. Kelly, A. Boneh, S. Pitsch, H. Gold, M. Fietz, P. Nelson and M. R. Oliver, *J Paediatr Child Health*, 2001, **37**, 510–512.
101 D. Penel-Capelle, D. Dobbelaere, J. Jaeken, A. Klein, M. Cartigny and J. Weill, *J Inherit Metab Dis*, 2003, **26**, 83–85.
102 M. Al-Owain, S. Mohamed, N. Kaya, A. Zagal, G. Matthijs and J. Jaeken, *Orphanet J Rare Dis*, 2010, **5**, 7.
103 B. S. Miller, H. H. Freeze, G. F. Hoffmann and K. Sarafoglou, *Mol Genet Metab*, 2011, **103**, 101–103.
104 T. Imbach, P. Burda, P. Kuhnert, R. A. Wevers, M. Aebi, E. G. Berger and T. Hennet, *Proc Natl Acad Sci USA*, 1999, **96**, 6982–6987.
105 S. Grunewald, T. Imbach, K. Huijben, M. E. Rubio-Gozalbo, A. Verrips, J. B. de Klerk, H. Stroink, J. F. de Rijk-van Andel, J. L. Van Hove, U. Wendel, G. Matthijs, T. Hennet, J. Jaeken and R. A. Wevers, *Ann Neurol*, 2000, **47**, 776–781.
106 F. Hanefeld, C. Korner, U. Holzbach-Eberle and K. von Figura, *Neuropediatrics*, 2000, **31**, 60–62.
107 T. Imbach, S. Grunewald, B. Schenk, P. Burda, E. Schollen, R. A. Wevers, J. Jaeken, J. B. de Klerk, E. G. Berger, G. Matthijs, M. Aebi and T. Hennet, *Hum Genet*, 2000, **106**, 538–545.
108 V. Westphal, C. Schottstadt, T. Marquardt and H. H. Freeze, *Mol Genet Metab*, 2000, **70**, 219–223.
109 S. Vuillaumier-Barrot, C. Le Bizec, G. Durand and N. Seta, *J Hum Genet*, 2001, **46**, 547–548.
110 J. W. Newell, N. S. Seo, G. M. Enns, M. McCraken, J. F. Mantovani and H. H. Freeze, *Mol Genet Metab*, 2003, **79**, 221–228.
111 V. Westphal, M. Xiao, P. Y. Kwok and H. H. Freeze, *Hum Mutat*, 2003, **22**, 420–421.
112 L. Sun, E. A. Eklund, J. L. Van Hove, H. H. Freeze and J. A. Thomas, *Am J Med Genet A*, 2005, **137**, 22–26.
113 J. Denecke, C. Kranz, J. von Kleist-Retzow, K. Bosse, P. Herkenrath, O. Debus, E. Harms and T. Marquardt, *Pediatr Res*, 2005, **58**, 248–253.
114 E. Schollen, S. Grunewald, L. Keldermans, B. Albrecht, C. Korner and G. Matthijs, *Eur J Med Genet*, 2005, **48**, 153–158.
115 C. Kranz, L. Sun, E. A. Eklund, D. Krasnewich, J. R. Casey and H. H. Freeze, *Am J Med Genet A*, 2007, **143A**, 1414–1420.
116 A. Rimella-Le-Huu, H. Henry, I. Kern, S. Hanquinet, E. Roulet-Perez, C. J. Newman, A. Superti-Furga, L. Bonafe and D. Ballhausen, *J Inherit Metab Dis*, 2008.
117 I. Chantret, T. Dupre, C. Delenda, S. Bucher, J. Dancourt, A. Barnier, A. Charollais, D. Heron, B. Bader-Meunier, O. Danos, N. Seta, G. Durand, R. Oriol, P. Codogno and S. E. Moore, *J Biol Chem*, 2002, **277**, 25815–25822.
118 C. E. Grubenmann, C. G. Frank, S. Kjaergaard, E. G. Berger, M. Aebi and T. Hennet, *Hum Mol Genet*, 2002, **11**, 2331–2339.
119 M. Di Rocco, T. Hennet, C. E. Grubenmann, S. Pagliardini, A. E. Allegri, C. G. Frank, M. Aebi, S. Vignola and J. Jaeken, *J Inherit Metab Dis*, 2005, **28**, 1162–1164.
120 E. A. Eklund, J. W. Newell, L. Sun, N. S. Seo, G. Alper, J. Willert and H. H. Freeze, *Mol Genet Metab*, 2005, **84**, 25–31.
121 C. Kranz, A. A. Basinger, M. Gucsavas-Calikoglu, L. Sun, C. M. Powell, F. W. Henderson, A. S. Aylsworth and H. H. Freeze, *Am J Med Genet A*, 2007, **143A**, 1371–1378.

122 D. Peric, C. Durrant-Arico, C. Delenda, T. Dupre, P. De Lonlay, H. O. de Baulny, C. Pelatan, B. Bader-Meunier, O. Danos, I. Chantret and S. E. Moore, *PLoS One*, 2010, **5**, e11675.
123 I. Chantret, J. Dancourt, T. Dupre, C. Delenda, S. Bucher, S. Vuillaumier-Barrot, H. Ogier de Baulny, C. Peletan, O. Danos, N. Seta, G. Durand, R. Oriol, P. Codogno and S. E. Moore, *J Biol Chem*, 2003, **278**, 9962–9971.
124 E. Schollen, C. G. Frank, L. Keldermans, R. Reyntjens, C. E. Grubenmann, P. T. Clayton, B. G. Winchester, J. Smeitink, R. A. Wevers, M. Aebi, T. Hennet and G. Matthijs, *J Med Genet*, 2004, **41**, 550–556.
125 E. A. Eklund, L. Sun, V. Westphal, J. L. Northrop, H. H. Freeze and F. Scaglia, *J Pediatr*, 2005, **147**, 847–850.
126 M. Jackson, P. Clayton, S. Grunewald, G. Keir, K. Mills, P. Mills, B. Winchester, V. Worthington and E. Young, *J Inherit Metab Dis*, 2005, **28**, 1197–1198.
127 T. Stolting, H. Omran, A. Erlekotte, J. Denecke, J. Reunert and T. Marquardt, *Mol Genet Metab*, 2009, **98**, 305–309.
128 K. Vesela, T. Honzik, H. Hansikova, M. A. Haeuptle, J. Semberova, Z. Stranak, T. Hennet and J. Zeman, *J Inherit Metab Dis*, 2009.
129 C. Thiel, M. Schwarz, J. Peng, M. Grzmil, M. Hasilik, T. Braulke, A. Kohlschutter, K. von Figura, L. Lehle and C. Korner, *J Biol Chem*, 2003, **278**, 22498–22505.
130 X. Wu, J. S. Rush, D. Karaoglu, D. Krasnewich, M. S. Lubinsky, C. J. Waechter, R. Gilmore and H. H. Freeze, *Hum Mutat*, 2003, **22**, 144–150.
131 T. Dupre, S. Vuillaumier-Barrot, I. Chantret, H. S. Yaye, C. Le Bizec, A. Afenjar, C. Altuzarra, C. Barnerias, L. Burglen, P. de Lonlay, F. Feillet, S. Napuri, N. Seta and S. E. Moore, *J Med Genet*, 2010, **47**, 729–735.
132 C. E. Grubenmann, C. G. Frank, A. J. Hulsmeier, E. Schollen, G. Matthijs, E. Mayatepek, E. G. Berger, M. Aebi and T. Hennet, *Hum Mol Genet*, 2004, **13**, 535–542.
133 C. Kranz, J. Denecke, L. Lehle, K. Sohlbach, S. Jeske, F. Meinhardt, R. Rossi, S. Gudowius and T. Marquardt, *Am J Hum Genet*, 2004, **74**, 545–551.
134 M. Schwarz, C. Thiel, J. Lubbehusen, B. Dorland, T. de Koning, K. von Figura, L. Lehle and C. Korner, *Am J Hum Genet*, 2004, **74**, 472–481.
135 C. G. Frank, C. E. Grubenmann, W. Eyaid, E. G. Berger, M. Aebi and T. Hennet, *Am J Hum Genet*, 2004, **75**, 146–150.
136 M. Weinstein, E. Schollen, G. Matthijs, C. Neupert, T. Hennet, C. E. Grubenmann, C. G. Frank, M. Aebi, J. T. Clarke, A. Griffiths, L. Seargeant and N. Poplawski, *Am J Med Genet A*, 2005, **136**, 194–197.
137 W. Vleugels, L. Keldermans, J. Jaeken, T. D. Butters, J. C. Michalski, G. Matthijs and F. Foulquier, *Glycobiology*, 2009, **19**, 910–917.
138 B. E. Baysal, J. E. Willett-Brozick, S. A. Bacanu, S. Detera-Wadleigh and V. L. Nimgaonkar, *Behav Brain Funct*, 2006, **2**, 25.
139 J. Jaeken, W. Vleugels, L. Regal, C. Corchia, N. Goemans, M. A. Haeuptle, F. Foulquier, T. Hennet, G. Matthijs and C. Dionisi-Vici, *J Inherit Metab Dis*, 2009.
140 W. Vleugels, M. A. Haeuptle, B. G. Ng, J. C. Michalski, R. Battini, C. Dionisi-Vici, M. D. Ludman, J. Jaeken, F. Foulquier, H. H. Freeze, G. Matthijs and T. Hennet, *Hum Mutat*, 2009, **30**, 1428–1434.
141 N. Rind, V. Schmeiser, C. Thiel, B. Absmanner, J. Lubbehusen, J. Hocks, N. Apeshiotis, E. Wilichowski, L. Lehle and C. Korner, *Hum Mol Genet*, 2010, **19**, 1413–1424.
142 B. B. de Vries, W. G. van'tHoff, R. A. Surtees and R. M. Winter, *Clin Dysmorphol*, 2001, **10**, 115–121.

143 C. Van Geet, J. Jaeken, K. Freson, T. Lenaerts, J. Arnout, J. Vermylen and M. F. Hoylaerts, *J Inherit Metab Dis*, 2001, **24**, 477–492.
144 P. de Cock and J. Jaeken, *Biochim Biophys Acta*, 2009, **1792**, 844–846.
145 C. Volker, C. M. De Praeter, B. Hardt, W. Breuer, B. Kalz-Fuller, R. N. Van Coster and E. Bause, *Glycobiology*, 2002, **12**, 473–483.
146 M. Garshasbi, K. Kahrizi, M. Hosseini, L. Nouri Vahid, M. Falah, S. Hemmati, H. Hu, A. Tzschach, H. H. Ropers, H. Najmabadi and A. W. Kuss, *Am J Med Genet A*, 2011, **155A**, 1976–1980.
147 M. A. Khan, M. A. Rafiq, A. Noor, N. Ali, G. Ali, J. B. Vincent and M. Ansar, *BMC Med Genet*, 2011, **12**, 56.
148 E. Mohorko, R. Glockshuber and M. Aebi, *J Inherit Metab Dis*, 2011, **34**, 869–878.
149 M. Garshasbi, V. Hadavi, H. Habibi, K. Kahrizi, R. Kariminejad, F. Behjati, A. Tzschach, H. Najmabadi, H. H. Ropers and A. W. Kuss, *Am J Hum Genet*, 2008, **82**, 1158–1164.
150 F. Molinari, F. Foulquier, P. S. Tarpey, W. Morelle, S. Boissel, J. Teague, S. Edkins, P. A. Futreal, M. R. Stratton, G. Turner, G. Matthijs, J. Gecz, A. Munnich and L. Colleaux, *Am J Hum Genet*, 2008, **82**, 1150–1157.
151 M. K. Stancheva-Ivanova, W. Wuyts, E. van Hul, B. I. Radeva, R. V. Vazharova, T. P. Sokolov, B. Y. Vladimirov, M. D. Apostolova and I. M. Kremensky, *J Inherit Metab Dis*, **34**, 917–921.
152 A. De Paepe and F. Malfait, *Clin Genet*, 2012.
153 S. Ichikawa, G. Baujat, A. Seyahi, A. G. Garoufali, E. A. Imel, L. R. Padgett, A. M. Austin, A. H. Sorenson, Z. Pejin, V. Topouchian, P. Quartier, V. Cormier-Daire, M. Dechaux, F. Malandrinou, P. N. Singhellakis, M. Le Merrer and M. J. Econs, *Am J Med Genet A*, 2010, **152A**, 896–903.
154 A. M. Shenoy, J. A. Markowitz, C. G. Bonnemann, K. Krishnamoorthy, A. D. Bossler and B. S. Tseng, *J Clin Neuromuscul Dis*, 2010, **11**, 124–126.
155 T. Y. Heinonen and M. Maki, *Ann Med*, 2009, **41**, 2–10.
156 D. Hess, J. J. Keusch, S. A. Oberstein, R. C. Hennekam and J. Hofsteenge, *J Biol Chem*, 2008, **283**, 7354–7360.
157 S. Hiraoka, T. Furuichi, G. Nishimura, S. Shibata, M. Yanagishita, D. L. Rimoin, A. Superti-Furga, P. G. Nikkels, M. Ogawa, K. Katsuyama, H. Toyoda, A. Kinoshita-Toyoda, N. Ishida, K. Isono, Y. Sanai, D. H. Cohn, H. Koseki and S. Ikegawa, *Nat Med*, 2007, **13**, 1363–1367.
158 T. Imbach, B. Schenk, E. Schollen, P. Burda, A. Stutz, S. Grunewald, N. M. Bailie, M. D. King, J. Jaeken, G. Matthijs, E. G. Berger, M. Aebi and T. Hennet, *J Clin Invest*, 2000, **105**, 233–239.
159 S. Kim, V. Westphal, G. Srikrishna, D. P. Mehta, S. Peterson, J. Filiano, P. S. Karnes, M. C. Patterson and H. H. Freeze, *J Clin Invest*, 2000, **105**, 191–198.
160 C. Kranz, J. Denecke, M. A. Lehrman, S. Ray, P. Kienz, G. Kreissel, D. Sagi, J. Peter-Katalinic, H. H. Freeze, T. Schmid, S. Jackowski-Dohrmann, E. Harms and T. Marquardt, *J Clin Invest*, 2001, **108**, 1613–1619.
161 B. Schenk, T. Imbach, C. G. Frank, C. E. Grubenmann, G. V. Raymond, H. Hurvitz, I. Korn-Lubetzki, S. Revel-Vik, A. Raas-Rotschild, A. S. Luder, J. Jaeken, E. G. Berger, G. Matthijs, T. Hennet and M. Aebi, *J Clin Invest*, 2001, **108**, 1687–1695.
162 M. T. Garcia-Silva, G. Matthijs, E. Schollen, J. C. Cabrera, J. Sanchez del Pozo, M. Marti Herreros, R. Simon, M. Maties, E. Martin Hernandez, T. Hennet and P. Briones, *J Inherit Metab Dis*, 2004, **27**, 591–600.

163 J. Dancourt, S. Vuillaumier-Barrot, H. O. de Baulny, I. Sfaello, A. Barnier, C. le Bizec, T. Dupre, G. Durand, N. Seta and S. E. Moore, *Pediatr Res*, 2006, **59**, 835–839.
164 D. J. Lefeber, J. Schonberger, E. Morava, M. Guillard, K. M. Huyben, K. Verrijp, O. Grafakou, A. Evangeliou, F. W. Preijers, P. Manta, J. Yildiz, S. Grunewald, M. Spilioti, C. van den Elzen, D. Klein, D. Hess, H. Ashida, J. Hofsteenge, Y. Maeda, L. van den Heuvel, M. Lammens, L. Lehle and R. A. Wevers, *Am J Hum Genet*, 2009, **85**, 76–86.
165 A. Broccolini, T. Gidaro, R. Morosetti, C. Sancricca and M. Mirabella, *Acta Myol*, 2011, **30**, 91–95.
166 T. Lubke, T. Marquardt, K. von Figura and C. Korner, *J Biol Chem*, 1999, **274**, 25986–25989.
167 T. Lubke, T. Marquardt, A. Etzioni, E. Hartmann, K. von Figura and C. Korner, *Nat Genet*, 2001, **28**, 73–76.
168 A. Etzioni, L. Sturla, A. Antonellis, E. D. Green, R. Gershoni-Baruch, P. M. Berninsone, C. B. Hirschberg and M. Tonetti, *Am J Med Genet*, 2002, **110**, 131–135.
169 N. Ishida and M. Kawakita, *Pflugers Arch*, 2004, **447**, 768–775.
170 Y. Luo and R. S. Haltiwanger, *J Biol Chem*, 2005, **280**, 11289–11294.
171 L. Sturla, F. Fruscione, K. Noda, E. Miyoshi, N. Taniguchi, P. Contini and M. Tonetti, *Glycobiology*, 2005, **15**, 924–934.
172 C. C. Hellbusch, M. Sperandio, D. Frommhold, S. Yakubenia, M. K. Wild, D. Popovici, D. Vestweber, H. J. Grone, K. von Figura, T. Lubke and C. Korner, *J Biol Chem*, 2007, **282**, 10762–10772.
173 B. Hansske, C. Thiel, T. Lubke, M. Hasilik, S. Honing, V. Peters, P. H. Heidemann, G. F. Hoffmann, E. G. Berger, K. von Figura and C. Korner, *J Clin Invest*, 2002, **109**, 725–733.
174 V. Peters, J. M. Penzien, G. Reiter, C. Korner, R. Hackler, B. Assmann, J. Fang, J. R. Schaefer, G. F. Hoffmann and P. H. Heidemann, *Neuropediatrics*, 2002, **33**, 27–32.
175 C. Kranz, C. Jungeblut, J. Denecke, A. Erlekotte, C. Sohlbach, V. Debus, H. G. Kehl, E. Harms, A. Reith, S. Reichel, H. Grobe, G. Hammersen, U. Schwarzer and T. Marquardt, *Am J Hum Genet*, 2007, **80**, 433–440.
176 V. Cantagrel, D. J. Lefeber, B. G. Ng, Z. Guan, J. L. Silhavy, S. L. Bielas, L. Lehle, H. Hombauer, M. Adamowicz, E. Swiezewska, A. P. De Brouwer, P. Blumel, J. Sykut-Cegielska, S. Houliston, D. Swistun, B. R. Ali, W. B. Dobyns, D. Babovic-Vuksanovic, H. van Bokhoven, R. A. Wevers, C. R. Raetz, H. H. Freeze, E. Morava, L. Al-Gazali and J. G. Gleeson, *Cell*, 2010, **142**, 203–217.
177 E. Morava, R. A. Wevers, V. Cantagrel, L. H. Hoefsloot, L. Al-Gazali, J. Schoots, A. van Rooij, K. Huijben, C. M. van Ravenswaaij-Arts, M. C. Jongmans, J. Sykut-Cegielska, G. F. Hoffmann, P. Bluemel, M. Adamowicz, J. van Reeuwijk, B. G. Ng, J. E. Bergman, H. van Bokhoven, C. Korner, D. Babovic-Vuksanovic, M. A. Willemsen, J. G. Gleeson, L. Lehle, A. P. de Brouwer and D. J. Lefeber, *Brain*, 2010, **133**, 3210–3220.
178 E. Morava, R. Zeevaert, E. Korsch, K. Huijben, S. Wopereis, G. Matthijs, K. Keymolen, D. J. Lefeber, L. De Meirleir and R. A. Wevers, *Eur J Hum Genet*, 2007, **15**, 638–645.
179 B. G. Ng, C. Kranz, E. E. Hagebeuk, M. Duran, N. G. Abeling, B. Wuyts, D. Ungar, V. Lupashin, C. M. Hartdorff, B. T. Poll-The and H. H. Freeze, *Mol Genet Metab*, 2007, **91**, 201–204.
180 B. G. Ng, V. Sharma, L. Sun, E. Loh, W. Hong, S. K. Tay and H. H. Freeze, *Mol Genet Metab*, 2010, **102**, 364–367.

181 J. Lubbehusen, C. Thiel, N. Rind, D. Ungar, B. H. Prinsen, T. J. de Koning, P. M. van Hasselt and C. Korner, *Hum Mol Genet*, **19**, 3623–3633.
182 E. Morava, S. Wopereis, P. Coucke, G. Gillessen-Kaesbach, T. Voit, J. Smeitink, R. Wevers and S. Grunewald, *Eur J Hum Genet*, 2005, **13**, 414–421.
183 M. Guillard, A. Dimopoulou, B. Fischer, E. Morava, D. J. Lefeber, U. Kornak and R. A. Wevers, *Biochim Biophys Acta*, 2009, **1792**, 903–914.
184 C. Noordam, S. Funke, N. V. Knoers, P. Jira, R. A. Wevers, Z. Urban and E. Morava, *Acta Paediatr*, 2009, **98**, 490–494.

Bladder cancer–glycosylation insights

Paulo F. Severino,[a,c] Mariana Silva,[a] Mylène A. Carrascal,[a] Fernando Calais,[b] Fabio Dall'Olio[c] and Paula A. Videira*[a]
DOI: 10.1039/9781849734769-00156

Bladder cancer is a common urologic cancer in Europe, with the highest recurrence rate of any malignancy. Early diagnosis and correct staging of the disease is critical, not only for planning treatment but also for selecting follow-up schedules, due to the high risk of recurrence. Since few biomolecules have reached the clinical practice, an unmet clinical need for biomarkers persists.

Correlations between specific glycan antigens and bladder cancer progression stages, as well as diagnostic approaches based on glycan analysis have been suggested worldwide. This chapter highlights the relevance and key aspects of the glycosylation changes associated with bladder cancer. Moreover, recent studies suggest that glycan-based markers would have widespread applicability in managing bladder cancer treatment and are promising potential targets for novel therapies.

1 Introduction

Bladder cancer has one of the highest incidence rates in Europe, being the fourth most common in men and the ninth in women. Non-muscle invasive bladder cancer (NMIBC), the most frequent type, has a high recurrence rate and represents the most costly common malignancy to manage per patient from diagnosis to death.[1–4]

Bladder cancer signs and symptoms include haematuria, abdominal pain, urinary urgency, pain while urinating, and loss of appetite. The risk factors for the development of the disease are genetic and molecular abnormalities, chronic irritation and infection, and chemical or environmental exposures.[5–7] Cigarette smoking is the most common and also the main preventable risk factor.[8–10] Staging and grading of bladder tumours are carried out mainly based on the cystoscopic and histopathologic analysis.[11,12] However, the current methodologies show poor sensitivity in detecting low grade disease, and sometimes miss prediction of the long-term outcome of tumour lesions.[13]

While some biomarkers are becoming available to aid in the diagnosis and staging, robust and specific biomarkers are still lacking. Besides their usefulness in diagnosis, biomarkers could be also considered as potential therapeutical targets.

One of the phenotypic changes associated with bladder cancer is the expression of an aberrant glycosylation profile. In this chapter, we will review the impact of altered glycosylation in bladder cancer. Consistent with the book's aim on *Carbohydrates in Europe*, in this chapter we will highlight the latest European findings on bladder cancer glycome.

[a]*CEDOC, Departamento de Imunologia, Faculdade de Ciências Médicas, FCM, Universidade Nova de Lisboa, Lisboa, Portugal. E-mail: paula.videira@fcm.unl.pt*
[b]*Grupo Português Génito-Urinário, Hospital São José, Lisboa, Portugal.*
[c]*Dipartimento di Patologia Sperimentale, Università di Bologna, Bologna, Italia.*

© The Royal Society of Chemistry 2012

2 Bladder cancer: epidemiology and current treatment

Bladder cancer is a common malignancy, particularly in southern and western European Countries (more than 30/100 000).[12] Over 90% of bladder cancer cases are originated in the urothelial epithelial cells, being referred as transitional cell carcinoma (TCC), approximately 5% are squamous cell carcinoma, and less than 2% are adenocarcinoma. At diagnosis, more than 75% of TCC cases are non-muscle invasive (NMIBC), including those that are superficial papillary tumours (Ta), invasive tumour of the subepithelial connective tissue (T1), and carcinoma *in situ* (CIS). The invasive forms, which are tumours that invade the muscle (T2), the perivesical tissue (T3) and surrounding organs (T4) are believed to derive from NMIBC.[14]

A decreasing trend in mortality has been evident in most European countries over the past 10 years[15] and presently, bladder cancer is generally considered to have minimal risk of progression to death. However, about 60–70% of the patients develop recurrences, representing the highest recurrence rate among solid tumours and 11% of these recurrences show progression to muscle invasion. Most patients survive but require a lifelong follow-up since more than half of the cases constantly relapse within one year, requiring repeated courses of treatment.[1,15–17]

The treatment strategy of bladder cancer depends largely on its invasive status. While invasive forms are usually treated by cystectomy, NMIBC are treated by transurethral resection (TUR).[6,12] The goal of treatment is twofold: to reduce tumour recurrence and subsequent need for supplementary therapies and to prevent tumour progression and subsequent need for more aggressive therapy. Several adjuvant agents, such as the Bacillus Calmette-Guérin (BCG, an attenuated strain of *Mycobacterium bovis*), or chemotherapeutic agents are essential as complementary therapies to reduce risk for recurrence and progression of NMIBC.[18] In patients with intermediate to high risk for recurrence, BCG therapy is more effective than other agents.[12,14] Although the precise mechanisms responsible for BCG anti-tumour effect have not been fully elucidated, a massive, complex and local activation of the immune system has been well documented.[19,20] After intravesical instillation, BCG is internalised and processed both by professional antigen-presenting cells (dendritic cells and macrophages) and tumour cells. Then, BCG antigens trigger an inflammatory and immune response which targets tumour cells.[21] Recent *in vitro* findings support the notion that the fibronectin attachment protein (FAP) from BCG has an important role in directing BCG into cells.[22,23] The activation of the immune system results predominantly in a local synthesis of pro-inflammatory cytokines and in the recruitment of several leukocyte populations. Owing to its anatomy, the bladder offers advantages to adjuvant therapy because it allows high local concentrations of either BCG or any chemotherapeutic drug.[21]

However, 30–50% of the patients either fail to respond to BCG or recur within the first year after treatment and in 10% of these patients the disease progresses to muscle invasive and metastatic stages. Until now, there are no biomarkers that can predict the patient's prognosis and discriminate those

that will respond to BCG treatment from those who would be best served by more aggressive therapy such as cystectomy.

The diagnosis of bladder cancer depends mainly on cystoscopic examination of the bladder and histological evaluation of multiple tissue biopsies, and also includes urine cytology.[12,24] As a standard procedure, cystoscopy and TUR are performed using white light. However, since it can miss lesions that are present but not visible, photodynamic diagnosis (PDD), which uses violet light after intravesical instillation of 5-aminolevulinic acid or hexaminolaevulinic acid, is also used. It has been confirmed that fluorescence-guided biopsy and resection are more sensitive than conventional procedures for detection of malignant tumours, particularly for CIS. The additional detection rate with PDD was 20% for all tumours and 23% for CIS in a cumulative analysis of prospective trials.[25]

Like in any other cancers, screening of bladder cancer patients for tumour specific biomarkers has been a principal goal of several research groups.[26] The assessment of such biomarkers is of extreme importance to establish non-invasive alternative diagnostic strategies, as well as reliable surveillance markers. Several bladder cancer biomarkers have been described, such as the mutated p53, homeobox protein HOX-1.3, keratins 8 and 13, psoriasin, adipocyte-type fatty acid-binding protein, GST Mu, 15-hydroxyprostaglandin dehydrogenase, tropomyosin 4, MRP-14, CD24, cytochrome oxidase III subunit, the Nuclear Matrix Protein 22 (NMP-22) among others, whose expression strongly correlates with a particular stage of bladder cancer progression.[27–38] Based on these biomarkers, different tissue and urine test kits have been developed, whose sensibilities and specificities have been reviewed by Van Tilborg *et al*., in 2009.[17] Presently, urine tests such as the bladder tumour antigen (BTA) stat test, which measures complement factor H related protein is used in clinical practise.[39] However, it is becoming increasingly clear that no single marker will have the sensitivity and specificity necessary to be used on its own for diagnosis/prognosis of tumours. Interpatient and intratumour heterogeneity provides overwhelming odds against the existence of such an ideal marker. Still, the use of multiple biomarkers can be used for an improvement of the clinical assessment of the disease.[40]

3 General aspects of glycosylation

Glycosylation is one of the most important modifications of proteins and lipids and the majority glycans exist as membrane-bound or soluble glycoconjugates. In eukaryotic cells, glycoconjugates are involved in a variety of important biological mechanisms including cell adhesion, cell-cell interaction, receptor activation and signal transduction.[41] The three main classes of glycoconjugates are glycoproteins, proteoglycans and glycolipids (Fig. 1) and their synthesis occurs mainly in the lumen of the endoplasmic reticulum and in the Golgi apparatus.

The sugar chains of glycoproteins are classified as *N*- or *O*-linked. The *N*-linked chains are attached through the *N*-acetylglucosamine (GlcNAc) residue to asparagine residue of the sequence N-X-S/T (where X is any aminoacid, except proline), while the *O*-linked chains are usually linked

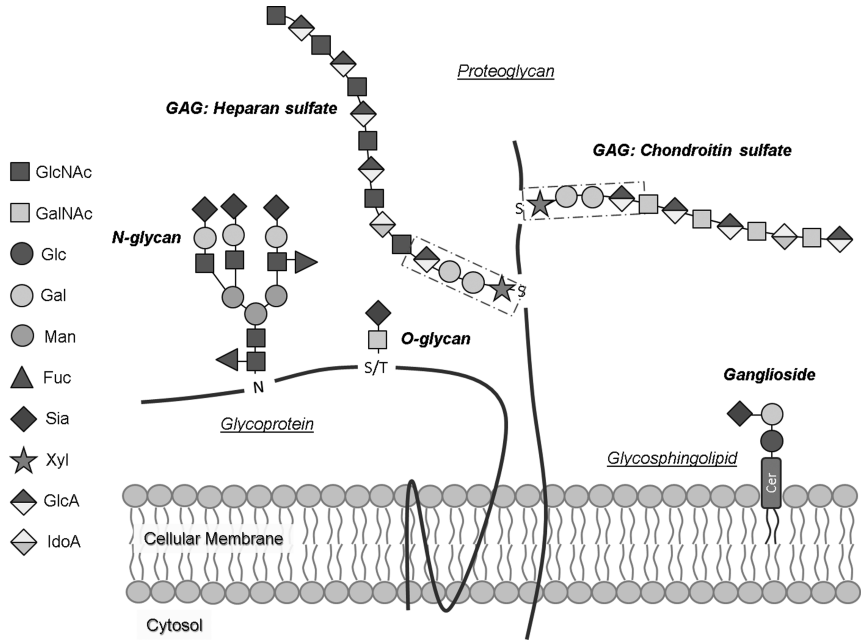


Fig. 1 Glycoconjugates in mammalian cells. Glycans can be attached to proteins as in glycoproteins and proteoglycans. Glycoproteins are decorated with *N*- and/or *O*-glycans and proteoglycan with glycosaminoglycans (GAG). Glycosphingolipids are mainly associated with membranes and are formed by a glycan attached to the lipid moiety ceramide. A ganglioside contains one or more residues of sialic acid (Sia). The glycan structures presented for each *N*-linked, *O*-linked, glycoaminoglycan (GAG) and glycosphingolipids exemplify one of the possible structures.

through an acetylgalactosamine (GalNAc) residue to serine or threonine (Fig. 1). The biosynthesis of *N*-linked chains involves the *en bloc* transfer of a preformed oligosaccharide comprised of two GlcNAc, nine mannoses (Man) and three glucose (Glc) residues to the nascent polypeptide chain.[42] Mature *N*-linked glycans consist of a core structure containing 2 GlcNAc and 3 Man residues and a variable (2–4) number of antennas, usually formed by GlcNAc, Galactose (Gal) and sialic acid (Sia). *O*-linked glycans are attached to the hydroxyl group of serine or threonine on a protein. After the addition of the first residue, the addition of other sugars is a stepwise process which accompanies the maturation of the glycoprotein.[43] In the 'mucin-type' *O*-glycans, the first GalNAc is further extended with Gal, GlcNAc, and Sia (Fig. 1). There are also several types of non-mucin *O*-glycans, including *O*-fucose (Fuc), *O*-xylose (Xyl), *O*-Man and *O*-GlcNAc glycans.[44] In this chapter, however, the term *O*-glycan refers to mucin *O*-glycans.

The basic structure of proteoglycans is comprised of a core protein and one or more covalently attached glycosaminoglycan (GAG) side chain. GAGs are polysaccharides composed of repeating disaccharide units formed by an amino sugar (GlcNAc in heparan sulphate or GalNAc in chondroitin sulphate) linked to D-glucuronic acid (GlcA) or L-iduronic acid (IdoA). The linkage of GAGs to the protein core involves a specific

trisaccharide composed of two Gal and one Xyl sugars (GAG-GalGalXyl-*O*-CH_2-protein). GAGs exist also as free molecules, as is the case of hyaluronan.[45] Negatively charged sulphate groups of the heparan-sulphate chains modulate interactions with different proteins involved in the cell–matrix assemblage, inflammation and cellular division.[45–47] In particular, GAGs could play an important role in the tumour environment by facilitating the storage of important ligands.[46]

In glycolipids, the sugar portion is usually attached through a Glc residue to the hydrophilic portion of a membrane lipid, which is often ceramide. In this case the glycolipid is referred to as glycosphingolipid. Based on their basic glycan structures, glycosphingolipids are classified into four groups according to the sugar types linked to the Gal, namely, ganglio-series, globo-series, paraglobo-series (or neolactoseries) and lacto-series. The glycosphingolipid biosynthesis begins with the enzymatic addition of a Glc to ceramide, forming glucosylceramide, which is then converted into lactosylceramide by the addition of Gal. The successive addition of other sugar residues, mediated by specific glycosyltransferases gives rise to the different series of glycosphingolipids. The glycolipids of the ganglio-series, or gangliosides, may have one, two or three Sia residues linked to the inner Gal unit and are thus termed a-, b- and c-series gangliosides, respectively (Fig. 2). They include GM3, GM2, GM1, GD3 and GD2 and are characterized by the presence of one or more Sia residues.[48] Gangliosides are frequently involved in various physiological and pathological processes, including cell differentiation, migration, intracellular signalling, host-pathogen interactions, and tumourigenesis.[49–51]

Unlike the biosynthesis of nucleic acids and proteins, which are deterministic template-driven processes, glycosylation is a stochastic process, regulated mainly by the relative abundance, cellular localization and specificities of the biosynthetic enzymes (glycosyltransferases) and catabolic enzymes (glycosidases). The stochastic nature of the glycosylation process is at the basis of the phenomenon known as microheterogeneity, which means that the structure of the sugar chains attached to a specific glycosylation site in a given glycoprotein displays a certain degree of variability.[52–55]

The expression of a given glycan structure which is usually cell type-dependent is profoundly modified in cancer.[44] In fact, an aberrant glycosylation of proteins and lipids is frequently observed at the cancer cell surfaces and can be shed into the body fluids of the patients (serum, urine, pleural effusions, *etc*.). Aberrant glycosylation in tumour cells include both a loss or a gain of expression of certain glycan structures, the appearance of truncated structures, as well as of novel structures. Upregulation and/or downregulation of specific glycosyltransferases is often responsible for these changes.[56–59] Tumour-associated carbohydrate structures allow tumour cells to invade and metastasize.[44] In the next sections, we will review some of the most relevant glycosylation changes associated with bladder cancer.

3.1 Carbohydrate antigens in bladder cancer

3.1.1 Blood group antigens. Blood group antigens, which include the ABH and Lewis blood group systems are a group of carbohydrate structures expressed at the terminal part of both glycoproteins (*O*-glycans and

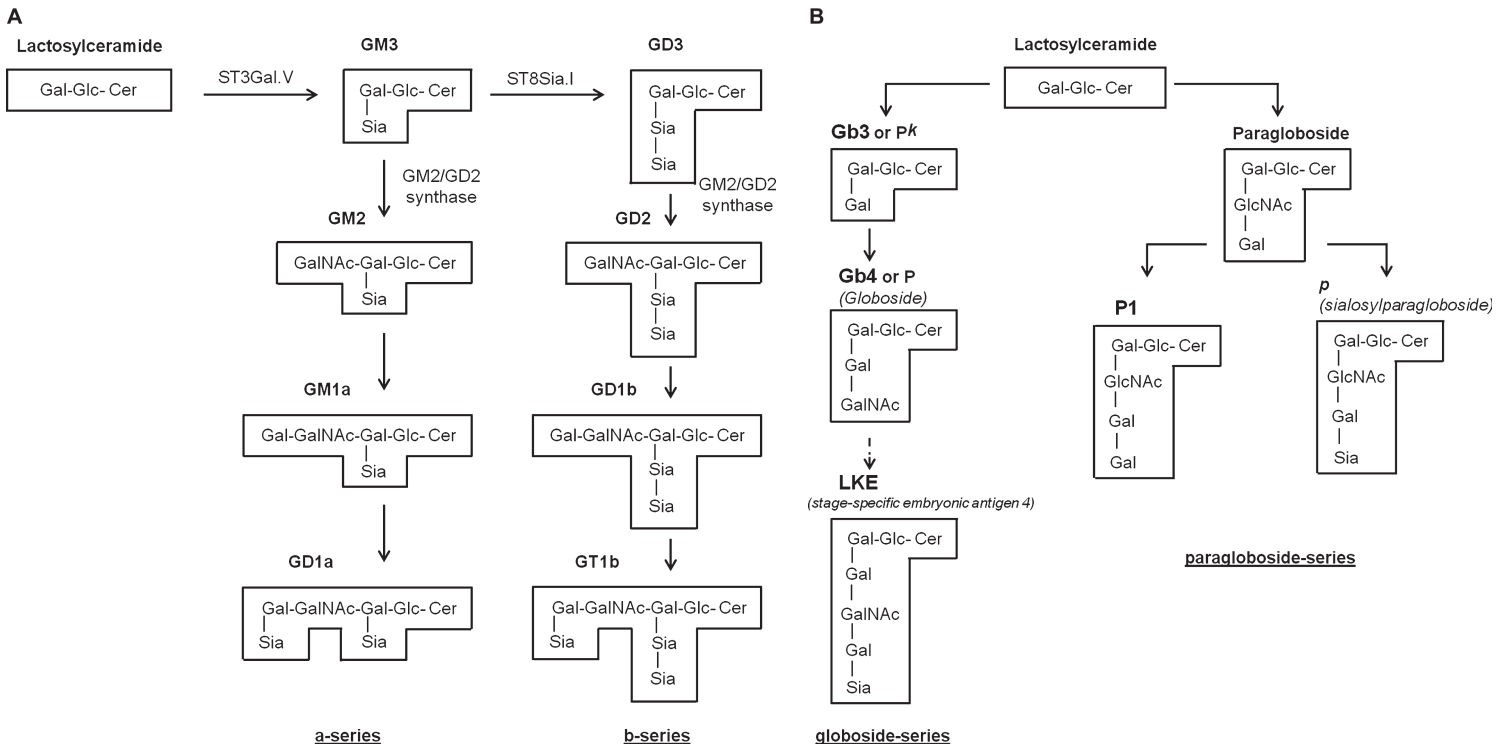


Fig. 2 Biosynthetic pathway of gangliosides and globosides. A: The conversion of lactosylceramide into GM3 and the formation of the main branches, the a- and b-ganglioseries, are shown. B: Lactosylceramide may also be converted into Gb3 (also known as Pk) or paragloboside forming the two main branches, the globoside and paragloboside series. Besides the structures, the enzymes responsible for the initial steps of their biosynthesis are also depicted.

N-glycans) and glycolipids on red blood cells, endothelial cells and epithelial cells, as well as in body fluids and secretions. The ABH and Lewis blood group antigens are genetically and biochemically related, presenting a common basic structure. These antigens are derived from substitutions on type 1 (Galβ1,3GlcNAc) or type 2 (Galβ1,4GlcNAc) oligosaccharide chains. The generation of the different glycan structures requires the contribution of specific glycosyltransferases.[60] The synthesis of the H antigen consists in the transfer of a Fuc residue in α1,2 linkage to the Gal residue of type 1 or 2 oligosaccharide chains (Table 1). The A or B antigens are subsequently formed by the addition of a GalNAc or a Gal residue in α1,3 linkage to the nonreducing terminus of the H antigen. The addition

Table 1 Carbohydrate sequence of ABH and Lewis blood antigens.

Type 1 oligossacharide chain	Galβ1,3GlcNAc
Type 2 oligossacharide chain	Galβ1,4GlcNAc
H antigen	Galβ1,3(4)GlcNAc – R | Fucα1,2
A antigen	GalNAcα1,3Galβ1,3(4)GlcNAc – R | Fucα1,2
B antigen	Galα1,3Galβ1,3(4)GlcNAc – R | Fucα1,2
Lewis[a] antigen	Galβ1,3GlcNAc – R | Fucα1,4
Lewis[b] antigen	Galβ1,3GlcNAc – R | | Fucα1,2 Fucα1,4
Lewis[x] antigen	Galβ1,4GlcNAc – R | Fucα1,3
Lewis[y] antigen	Galβ1,4GlcNAc – R | | Fucα1,2 Fucα1,4
sialyl Lewis[a] antigen	Siaα2,3Galβ1,3GlcNAc – R | Fucα1,4
sialyl Lewis[x] antigen	Siaα2,3Galβ1,4GlcNAc – R | Fucα1,3

of a Fuc residue in α1,4 linkage to the GlcNAc unit of type 1 chains results in the biosynthesis of Lewis^a (Le^a) antigen, while the addition of Fuc residue in α1,3 linkage to type 2 chains results in the biosynthesis of Lewis^x (Le^x) antigen. If the addition of these two sugar residues occurs on the H antigen, either the Lewis^b (Le^b) or the Lewis^y (Le^y) antigens are synthesized (Table 1). If the addition of α1,3/4-linked Fuc is preceded by α2,3-sialylation of type 1 or of type 2 chains, the sialyl Lewis^a (sLe^a) and sialyl Lewis^x (sLe^x) antigens, respectively, are synthesized.[60]

The expression of Lewis antigens is associated with the secretion of the A, B or H antigens. In fact, the α1,2 fucosyltransferase expressed in secretory glands (FucT2) is responsible for both, Le^a to Le^b conversion and the addition of fucose to A, B and H antigens on soluble molecules. Therefore, people with Le^a antigens will not secrete the A, B or H antigens (ABH non-secretors), and the presence of the Le^b antigen is only found in secretors.

The altered expression of blood group antigens is a common feature of many epithelial cancers. The loss of the expression of blood group antigens has been observed in all different tumour stages of TCC and it has been associated with tumour progression, recurrent disease and metastases formation.[61,62] Immunohistochemical analyses revealed that the Le^x antigen is expressed in TCC, regardless the secretory status of the individual and the stage of the tumour,[61,63] while the expression of sLe^x was found to be related with invasive and metastatic potential.[64] Several publications reported that the expression of sLe^a antigen was upregulated in bladder cancer when compared with other urologic diseases.[65–69] Moreover, its expression has been recently correlated with tumour invasion and grade.[70] Accordingly, the expression of either sLe^x or sLe^a is known to allow cancer cells to evade the immune responses and mimic the ability of blood cells to migrate and spread through binding to endothelial selectins.[50,71] Recently, it was reported that Le^y antigen expression was also altered in bladder cancer, with different expression profiles between CIS and non-CIS condition of the bladder.[72] Since CIS lesions are associated with high-risk of progression, but no accurate biomarkers have ever been established, the inclusion of new glycans such as Le^y would be of help in the establishment of a glycoprofile for this type of lesions.

The loss of ABH structures and the increased expression of Lewis antigens are the most evident changes, among the blood group antigens, related with poor prognosis in bladder cancer,[70] suggesting a pathophysiological role in bladder cancer progression.

3.1.2 Thomsen-Friedenreich related antigens. Thomsen-Friedenreich (TF) related antigens are a group of *O*-glycans which includes the Tn and T antigens, and their sialylated forms sialyl Tn (sTn) and sialyl T (sT) antigens (Fig. 3). The Tn antigen (GalNAc-*O*-Ser/Thr) is the precursor of the T antigen (Galα1,3GalNAc-*O*-Ser/Thr), also known as *core 1* structure. sTn is produced by the addition of Sia in α2,6-linkage to GalNAc. The biosynthesis of sT results mainly from the α2,3-sialylation of the Gal residue of T antigen[43] (3-sialyl T antigen, Fig. 3), but it also includes the α2,6-sialylation of the GalNAc (disialyl T antigen) or the Gal residue (6-sialyl T antigen). Sialylation of TF-related glycans blocks the elongation of the

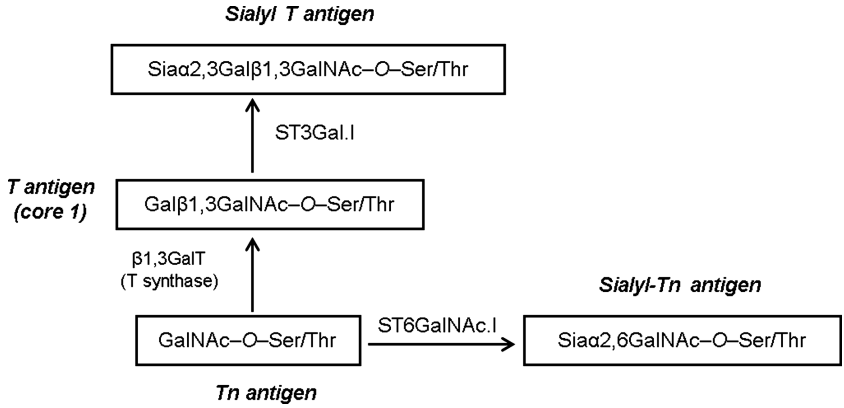


Fig. 3 Structures of Thomsen-Friedenreich related antigens. Both glycan structures and the main enzymes responsible for their synthesis are shown.

sugar chain, resulting in the accumulation of these glycans which, in turn, affect the adhesive properties of cancer cells, their invasive and angiogenic potential as well as their recognition by immune cells.[51,73]

In bladder cancer, the presence and the clinical importance of TF-related antigens is still controversial.[74–78] The earliest evidence of a deregulated expression of TF-related antigens in bladder cancer came from the observation of an enhanced binding of the peanut lectin (PNA), which recognizes the T antigen, to the bladder tumour tissue.[79,80] The PNA staining is a largely used method for the detection of T antigens in human normal and malignant tissues because it requires little amount of samples. The expression of sT antigens can be inferred after sialidase treatment and staining with PNA. In bladder, it was also reported that the PNA binding was correlated with invasion and a higher risk of lymph node metastasis.[79,81] However, the aberrant expression of the sialylated forms of T was also reported.[76,77] The ''cryptic'' T is present in the healthy urothelium but overexpressed in bladder tumors, in accordance with the author's observations. A deeper understanding of the structural nature of these ''cryptic'' glycans may render in the future valuable information regarding bladder cancer Glycobiology.

Besides the TF-related antigens, also the sialyltransferases involved in the biosynthesis of their sialylated forms are potential prognostic markers. In fact, when we analysed the mRNA expression level of the sialyltransferases involved in the synthesis of the sTn and sT antigens in bladder cancer and normal urothelium biopsies, we found that the expression of ST3Gal.I[57] and of ST6GalNAc.I (unpublished data) were augmented in bladder cancer. In the same study, a higher expression of ST3Gal.I and a decreased expression of ST3Gal.II in tumour samples correlated with propensity to tumour recurrence.[57] Concordantly with the increased ST6GalNac.I expression, the analysis of histological sections from bladder cancer patients revealed the presence of sTn antigens apparently related with a low degree of dedifferentiation of the malignant cells (Videira *et al.*, unpublished data). While a correlation has been shown *in vitro* between ST3Gal.I and sT expression, the same was not observed in patient's samples.[57] Further

investigations, involving larger number of patients, are still needed to identify the targets of the above mentioned sialyltransferases and the TF-decorated glycoproteins.

A glycopeptide decorated with a trisaccharide similar, if not identical, with sT was identified in the urine of patients affected by interstitial cystitis, a painful disease caused by thinning or denudation of bladder epithelium. This molecule, which is referred to as antiproliferative factor (APF) exerts a strong anti-proliferative activity on bladder cells and is involved in the pathogenesis of interstitial cystitis. The peptide backbone is related to frizzled 8, a G protein-coupled receptor of the Wnt signalling pathway and the presence of the sugar chain is necessary for the antiproliferative activity of APF.[82] Altogether, these observations are consistent with a role for TF-related glycans in the control of bladder cell proliferation.

3.1.3 Glycosphingolipids. Several reports showed drastic changes in the glycosylation of lipids in bladder tumours and correlated these modifications with tumour invasiveness and progression.[83,84] While NMIBC show an accumulation of GM3 gangliosides and a decrease of Gb3 and Gb4 globo-series structures, invasive tumours show a decreased expression of GM3 and an accumulation of Gb3 and Gb4 structures.[83] The accumulation of GM3 structures in NMIBC may be explained by the upregulation of GM3 synthase and the downregulation of Gb3 and Gb4 synthase.[83] Moreover, it has been demonstrated that the exogenous addition of GM3 inhibits the invasive activity of bladder cancer cell lines. Concordantly, bladder cancer cell lines overexpressing GM3 synthase have reduced proliferation potential, motility and growth as xenografts.[83,85] Todeschini *et al.* reported that GM2/GM3 complexes efficiently inhibit cell motility of two different bladder cancer cell lines.[86] More recently, it was described the involvement of specific glycosphingolipids, including GM3, GM2 and GM1, in the epithelial-to-mesenchymal cell transition (EMT) process of bladder cells, a basic process of cancer progression characterized by loss of cell adhesion and increased cell mobility.[84]

3.2 Glycoproteins

3.2.1 Cadherins. The cadherin cell adhesion molecules are calcium-dependent transmembrane glycoproteins and are the main constituent of the adherens junctions. The cadherins are comprised of extracellular, membrane and cytoplasmic domains with the cytoplasmic domain anchored to the cell cytoskeleton by the catenin family members.[87] Cadherins are implicated in numerous signalling events related to embryonic development, tissue morphogenesis and homeostasis and are the main mediators of cell-cell adhesion in normal and cancer epithelial tissues.[88–90] In several studies, the downregulation of E-cadherin in malignant human urothelium was associated with tumour recurrence, progression and survival.[91–100] *In vitro* studies in bladder cancer cell lines also suggested that low E-cadherin expression was correlated with expansion of CIS.[101–103] On the other hand, higher levels of the proteolytic cleavage product of the E-cadherin, sE-cadherin, can be found particularly in serum of patients with TCC, being detectable also in urine.[104–107] In addition, in bladder cancer it has also been

observed a switch from the production of E-cadherin to that of N-cadherin (typically found in neural and mesodermal tissues) and P-cadherin (typically found in placental tissue). These events seem to be important late processes in the progression of bladder cancer.[108] Cadherins switching has also been reported in several other malignancies.[109] A study on non-malignant and malignant bladder cancer cell lines has revealed a tendency to accumulate β1,6-branched oligosaccharides with poly-*N*-acetyllactosaminic structures and α2,3-linked Sia residues in the latter,[110] which are usually associated with metastasis.[111] However, *N*-acetylglucosaminyltransferase V (GnT-V), a key enzyme in the formation of the β1,6-branched *N*-linked oligosaccharides, was inversely correlated with recurrence events and directly correlated with low malignant potential and good prognosis.[112] Further studies seem to be necessary to understand the pathological role of β1,6-branched *N*-linked glycans on E-cadherins.

3.2.2 Integrins. Integrins are a widely distributed family of heterodimeric cell membrane glycoproteins involved in cell–to–extracellular matrix and cell–to–cell interactions, and mediating important intracellular signalling cascades.[113,114] These glycoproteins consist of sixteen types of α subunits and eight types of β subunits that assemble into twenty two types of cell surface receptors that bind mostly large molecules, usually found in the subendothelial matrix, including fibronectin, vitronectin and collagen.[113,115–118] Integrins are involved in cell growth and differentiation, proliferation and migration, tissue organization, recruiting of lymphocytes and inflammation, cancer invasion and metastasis.[113] Integrin expression and glycosylation are altered upon cell transformation, resulting in altered function of these molecules. In the bladder cancer cells, but not in normal epithelium, it was found the presence of a sialylated tetraantennary *N*-oligosaccharide in both subunits of α3β1 integrin.[119] An aberrant expression of β1-6 branched *N*-linked glycans, as well as high mannose type glycans, has also been detected in α3β1 integrins from bladder cancer cells. Removal of Sia and treatment with swainsonine (an inhibitor of Golgi α-mannosidase II, an important enzyme in *N*-glycan maturation) in T24 bladder cancer cells resulted in a stronger fibronectin-binding activity of α3β1 integrin and in a lower rate of tumour cells migration respectively, suggesting that the glycosylation profile of α3β1 integrin may have an important function in bladder cancer cells migration.[119]

3.2.3 Mucins. Mucins (MUCs) are heavily glycosylated high-molecular-mass glycoproteins that can be present both in membrane-bound and secreted forms and represent a major component of mucus.[120] The abundance of serine and threonine residues in their variable number of tandem repeats provides the backbone necessary for their bulk *O*-glycosylation.[121] In TCC (transitional cell carcinoma), the pattern, intensity and depth of MUC1 staining were significantly associated with tumour stage and grade and metastases.[122] However, larger studies are required to define more precisely the role of MUC1 in the prognosis of bladder cancer. The anti-MUC1, C595, monoclonal antibody labelled with ^{111}Indium or ^{67}Cupper, when intravenously administrated was able to detect bladder cancer cells, including distant metastasis,[123,124] suggesting the possible therapeutic use of

anti-MUC1 antibodies conjugated with cytotoxic isotopes. MUC7 is also atypically expressed in bladder cancer, in the soluble form, being detected into the urine of patients.[125] The glycosylation of both MUC1 and MUC7 varies among different issue and in cancer they have been reported to be more or less glycosylated than the corresponding normal tissues. Until now, the glycosylation profile of both mucins remains unknown. The identification of minute glycosylation changes in these mucins is likely to improve the diagnostic potential of MUC1 and 7 as bladder cancer biomarkers.

3.3 Extracellular matrix

Invasion by tumour cells involves the alteration of cell–matrix interactions.[51] Extracellular matrix is comprised of a complex mixture of macromolecules capable of self-assemble, composed predominantly of collagens, non-collagenous glycoproteins, elastin, hyaluronan and proteoglycans. The extracellular matrix represents not only a scaffold for the cells, but also a reservoir for important biomolecules, such as growth factors and cytokines. In addition, there is increasing evidence that extracellular matrix molecules may exhibit a direct signalling function, either *via* interactions with matrix receptors like integrins or *via* signalling through growth factor receptors themselves.[126,127] These biomolecules may also play important roles in cancer growth.[126] Hyaluronic acid (HA), a nonsulfated GAG comprised of a repeating disaccharide (GlcA-GlcNAc) absorbs significant amounts of water in the extracellular space and confers tissues the ability to resist compression. The expression of HA has been implicated in promoting tumour cell proliferation, invasion and metastasis and represents an accurate diagnostic marker for bladder cancer. On the other hand, hyaluronidase (HAase) degrades HA into proangiogenic fragments that support tumour metastasis. The measurement of HA and of its degrading enzyme hyaluronidase (HAase) in urine has been applied as a commercial urinary test, the HA-HAase test, to the screening of bladder cancer.[128,129]

3.4 Carbohydrate-binding proteins – Lectins

Initial insights into the unique repertoire of glycans expressed on tumour cells emerged from the increased ability of tumours to bind a range of plant lectins.[130] In fact, a large number of the studies on the alterations of glycosylation in cancer has been performed using lectins as probes.[131] Lectins exhibit a typical protein folding that defines the carbohydrate-binding protein families and bind defined monosaccharide or oligosaccharide structures with good specificity[131] in a 'lock-and-key' fashion.[51,132] The carbohydrate-binding proteins are divided into several types depending on their structures and the nature of the interaction they establish.[131] Long before the role of carbohydrate–protein interactions had been explored, many lectins were identified, characterized and applied as useful tools in studying glycoconjugates. Therefore, lectins have been widely used for preparative and analytical purposes in biochemistry, cell biology, immunology and related areas.[131]

Some lectins were found to be useful in differentiating bladder tumours from normal urothelium. Lectins like *Ricinus communis* lectin (RCA, recognizing Galβ1,4GlcNAcβ1-R) and wheat germ lectin (WGA,

recognizing GlcNAcβ1,4GlcNAcβ1,4GlcNAc-R) can bind to all urothelial cells in the normal tissue, while some lectins like PNA, concanavalin A lectin (ConA, recognizing branched α-mannosidic structures and biantennary complex type N-Glycans) and soybean lectin (SBA, recognizing GalNAc) show little binding.[133] Initially, by using cell suspensions from transitional cell carcinoma, it was found that aneuploid cells bound PNA more extensively and were less reactive with WGA than diploid cell population.[77] Drugs or delivery systems modified by lectins have been thought as potential tools for more efficient, local adjuvant intravesical treatment for bladder cancer.[134]

Galectins represent a family of calcium-independent Gal-specific lectins, widely distributed in all living organisms.[135] There are 15 family members reported in humans, each of which plays a distinct role in adhesion, differentiation, growth, apoptosis, *etc*.[136–139] Galectins play a number of important roles in cancer by contributing to tumourigenesis, proliferation, angiogenesis, and metastasis.[140–145] These lectins are the best characterized carbohydrate-binding proteins in bladder cancer.[136,139,146–148] The expression of galectin-1 and galectin-3 was investigated in human bladder transitional cell carcinomas of different histological grade and clinical stage and in normal urothelium samples.[145] Galectin-1 levels were found to be highly increased in most high-grade tumours when compared with normal bladder or low-grade tumours. Galectin-3 mRNA levels were also found increased in most tumours, however with no relevant differences among tumours of different histological grade.[145] In a more detailed study, it was found an increased expression of the galectin-3 mRNA in invasive bladder tumours when compared with NMIBC and an association with poor survival; while galectin-3 protein levels in urine allowed identifying bladder cancer patients.[149] Galectin-3 plays an important role in bladder cancer by inhibiting TNF-related apoptosis-inducing ligand (TRAIL)-induced apoptosis.[150] Furthermore, the development of anti-galectin compounds has been suggested as a potential target for therapeutic intervention in bladder cancer.[151] In addition galectins-1 and -3 have also been considered bladder tumour markers.[145] The expression of galectin-7 had been hypothesised to be related to the sensitivity of bladder cancer to platinum-containing drugs. In fact, galectin-7 levels, which were higher in normal urothelium than in bladder cancer, were found to increase in response to treatment with platinum-containing drugs, but only in p53 wild type tumours. Moreover, transfection of galectin-7 cDNA in p53-mutated cell lines resulted in increased drug sensitivity.[152]

4 Final considerations/conclusions

Owing to its high incidence and percentage of recurrence, bladder cancer represents a major public health concern. Consequently, any effort should be made to improve the survival and the quality of life of the patients. In this light, the development of strategies aimed at early diagnosis and at more effective therapies represents major goals.

Along this chapter, we have provided an overview of how Glycobiology can contribute to pursue these achievements. We have described the main

findings of basic and clinical research, showing how sugar chains, glycosylated molecules and sugar binding molecules can affect the progression of this malignancy. As biomarker discovery becomes more sophisticated, glycobiology is being used to bring more sensitive and specific diagnostic based on cancer-related changes in glycosylation on specific glycoproteins. We are confident that research in the field of glycobiology will offer in the future new diagnostic and therapeutic approaches with a strong positive impact on the life of the patients.

Abbreviations

APF	antiproliferative factor
BCG	*Bacillus Calmette-Guérin*
CIS	carcinoma *in situ*
Fuc	fucose
GAG	glycosaminoglycan
Gal	galactose
GalNac	*N*-acetylgalactosamine
Glc	glucose
GlcA	D-glucuronic acid
GlcNAc	*N*-acetylglucosamine
HA	hyaluronic acid
HAase	hyaluronidase
IdoA	L-iduronic acid
Le^a	$Lewis^a$
Le^b	$Lewis^b$
Le^x	$Lewis^x$
Le^y	$Lewis^y$
Man	mannose
MUC	mucin
NMIBC	non-muscle invasive bladder cancer
PDD	photodynamic diagnosis
PNA	peanut lectin
Sia	sialic acid
sLe^a	sialyl $Lewis^a$
sLe^x	sialyl $Lewis^x$
sT	sialyl T
sTn	sialyl Tn
TCC	transitional cell carcinoma
TF	Thomsen-Friedenreich-related antigens
TUR	transurethral resection

References

1 M. L. Hemani and C. L. Bennett, *Cancer*, 2010, **116**, 3530–3532.
2 T. P. Kresowik and T. S. Griffith, *Immunotherapy*, 2009, **1**, 281–288.
3 M. Racioppi, A. Volpe, R. Falabella, F. Pinto, E. Sacco, G. Gulino, D. D'Agostino and P. F. Bassi, *Arch Ital Urol Androl*, 2007, **79**, 111–117.
4 A. Stenzl, J. Hennenlotter and D. Schilling, *Curr Opin Urol*, 2008, **18**, 488–492.

5 C. Pelucchi, C. Bosetti, E. Negri, M. Malvezzi and C. La Vecchia, *Nat Clin Pract Urol*, 2006, **3**, 327–340.
6 D. S. Kaufman, W. U. Shipley and A. S. Feldman, *Lancet*, 2009, **374**, 239–249.
7 M. J. Thun, S. J. Henley and T. Gansler, *Novartis Found Symp*, 2004, **256**, 6–21; discussion, **22-28**, 49-52, 266–269.
8 B. Anderson, *Br J Community Nurs*, 2009, **14**, 385–392.
9 M. P. Zeegers, F. E. Tan, E. Dorant and P. A. van Den Brandt, *Cancer*, 2000, **89**, 630–639.
10 P. Boffetta, *Scand J Urol Nephrol Suppl*, 2008, 45–54.
11 M. Burchardt, T. Burchardt, A. Shabsigh, A. De La Taille, M. C. Benson and I. Sawczuk, *Clin Chem*, 2000, **46**, 595–605.
12 M. Babjuk, W. Oosterlinck, R. Sylvester, E. Kaasinen, A. Böhle, J. Palou-Redorta and M. Rouprêt, and E. A. o. U. (EAU), *Eur Urol*, 2011, **59**, 997–1008.
13 D. Mitropoulos, A. Kiroudi-Voulgari, P. Nikolopoulos, T. Manousakas and A. Zervas, *J Endourol*, 2005, **19**, 861–864.
14 B. L. Jacobs, C. T. Lee and J. E. Montie, *CA Cancer J Clin*, 2010, **60**, 244–272.
15 D. M. Parkin, *Scand J Urol Nephrol Suppl*, 2008, 12–20.
16 D. M. Latini, S. P. Lerner, S. W. Wade, D. W. Lee and D. Z. Quale, *Urology*, 2010, **75**, 334–339.
17 A. A. Van Tilborg, C. H. Bangma and E. C. Zwarthoff, *Int J Urol*, 2009, **16**, 23–30.
18 K. Hendricksen and J. A. Witjes, *Curr Opin Urol*, 2007, **17**, 352–357.
19 A. Bohle and S. Brandau, *The Journal of Urology*, 2003, **170**, 964–969.
20 P. A. Videira, F. M. Calais, M. Correia, D. Ligeiro, H. J. Crespo, F. Calais and H. Trindade, *Urology*, 2009.
21 R. F. Bevers, K. H. Kurth and D. H. Schamhart, *British Journal of Cancer*, 2004, **91**, 607–612.
22 B. G. Coon, S. Crist, A. M. Gonzalez-Bonet, H. K. Kim, J. Sowa, D. H. Thompson, T. L. Ratliff and R. C. Aguilar, *Int J Cancer*, 2011.
23 H. W. Sinn, B. D. Elzey, R. J. Jensen, X. Zhao, W. Zhao and T. L. Ratliff, *Cancer Immunol Immunother*, 2008, **57**, 573–579.
24 K. H. Kurth, P. F. Schellhammer, E. Okajima, A. Akdas, G. Jakse, H. W. Herr, F. Calais da Silva, S. Fukushima and T. Nagayama, *Int J Urol.*, 1995, **2**(Suppl 2), 8–22.
25 J. A. Witjes, P. M. Moonen and A. G. van der Heijden, *European Urology*, 2005, **47**, 319–322.
26 D. Tilki, M. Burger, G. Dalbagni, H. B. Grossman, O. W. Hakenberg, J. Palou, O. Reich, M. Rouprêt, S. F. Shariat and A. R. Zlotta, *Eur Urol*, 2011, **60**, 484–492.
27 J. E. Celis, M. Ostergaard, B. Basse, A. Celis, J. B. Lauridsen, G. P. Ratz, I. Andersen, B. Hein, H. Wolf, T. F. Orntoft and H. H. Rasmussen, *Cancer Res*, 1996, **56**, 4782–4790.
28 I. Gromova, P. Gromov and J. E. Celis, *Electrophoresis*, 1999, **20**, 241–248.
29 M. Ostergaard, H. Wolf, T. F. Orntoft and J. E. Celis, *Electrophoresis*, 1999, **20**, 349–354.
30 I. Gromova, P. Gromov, H. Wolf and J. E. Celis, *Int J Oncol*, 1998, **13**, 379–383.
31 J. E. Celis, P. Celis, H. Palsdottir, M. Ostergaard, P. Gromov, H. Primdahl, T. F. Orntoft, H. Wolf, A. Celis and I. Gromova, *Mol Cell Proteomics*, 2002, **1**, 269–279.
32 M. Ostergaard, H. H. Rasmussen, H. V. Nielsen, H. Vorum, T. F. Orntoft, H. Wolf and J. E. Celis, *Cancer Res*, 1997, **57**, 4111–4117.

33 G. Ohlsson, J. M. Moreira, P. Gromov, G. Sauter and J. E. Celis, *Mol Cell Proteomics*, 2005, **4**, 570–581.
34 J. M. Moreira, P. Gromov and J. E. Celis, *Mol Cell Proteomics*, 2004, **3**, 410–419.
35 A. Saad, D. C. Hanbury, T. A. McNicholas, G. B. Boustead, S. Morgan and A. C. Woodman, *BJU Int*, 2002, **89**, 369–373.
36 V. Poulakis, U. Witzsch, R. De Vries, H. M. Altmannsberger, M. J. Manyak and E. Becht, *BJU Int*, 2001, **88**, 692–701.
37 K. H. Tsui, S. M. Chen, T. M. Wang, H. H. Juang, C. L. Chen, G. H. Sun and P. L. Chang, *Asian J Androl*, 2007, **9**, 711–715.
38 P. Bassi, V. De Marco, A. De Lisa, M. Mancini, F. Pinto, R. Bertoloni and F. Longo, *Urol Int*, 2005, **75**, 193–200.
39 R. Kinders, T. Jones, R. Root, C. Bruce, H. Murchison, M. Corey, L. Williams, D. Enfield and G. M. Hass, *Clin Cancer Res*, 1998, **4**, 2511–2520.
40 J. M. Moreira, G. Ohlsson, P. Gromov, R. Simon, G. Sauter, J. E. Celis and I. Gromova, *Mol Cell Proteomics*, 2010, **9**, 161–177.
41 A. Varki, *Glycobiology*, 1993, **3**, 97–130.
42 R. Kornfeld and S. Kornfeld, *Annu Rev Biochem*, 1985, **54**, 631–664.
43 I. Brockhausen, *Biochim Biophys Acta*, 1999, **1473**, 67–95.
44 A. Varki, R. Kannagi and B. P. Toole, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, The Consortium of Glycobiology Editors, La Jolla, California, Cold Spring Harbor (NY) Editon edn., 2009, vol. 2nd.
45 J. D. Esko and U. Lindahl, *J Clin Invest*, 2001, **108**, 169–173.
46 M. Bernfield, M. Gotte, P. W. Park, O. Reizes, M. L. Fitzgerald, J. Lincecum and M. Zako, *Annu Rev Biochem*, 1999, **68**, 729–777.
47 J. D. Esko, K. Kimata and U. Lindahl, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, The Consortium of Glycobiology Editors, La Jolla, California, Cold Spring Harbor NYEditon edn., 2009.
48 R. L. Schnaar, A. Suzuki and P. Stanley, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, The Consortium of Glycobiology Editors, La Jolla, California, Cold Spring Harbor NYEditon edn., 2009.
49 C. T. Campbell and K. J. Yarema, *Genome biology*, 2005, **6**, 236.
50 C. A. Reis, H. Osorio, L. Silva, C. Gomes and L. David, *Journal of clinical pathology*, 2010, **63**, 322–329.
51 M. M. Fuster and J. D. Esko, *Nat. Rev. Cancer.*, 2005, **5**, 526–542.
52 G. Lauc and V. Zoldos, *Mol Biosyst*, 2010, **6**, 2373–2379.
53 G. Lauc, I. Rudan, H. Campbell and P. M. Rudd, *Mol Biosyst*, 2010, **6**, 329–335.
54 K. L. Abbott, A. V. Nairn, E. M. Hall, M. B. Horton, J. F. McDonald, K. W. Moremen, D. M. Dinulescu and M. Pierce, *Proteomics*, 2008, **8**, 3210–3220.
55 A. V. Nairn, W. S. York, K. Harris, E. M. Hall, J. M. Pierce and K. W. Moremen, *The Journal of Biological Chemistry*, 2008, **283**, 17298–17313.
56 P. A. Videira, I. F. Amado and M. Correia, F. M. Calais da Silva, F. Calais da Silva, F. Dall'Olio and H. Trindade, *URO*, 2007, **13**, 26–32.
57 P. A. Videira, M. Correia, N. Malagolini, H. J. Crespo and D. Ligeiro, F. M. Calais, H. Trindade and F. Dall'Olio, *BMC Cancer*, 2009, **9**, 357.
58 F. Dall'olio, N. Malagolini and M. Chiricolo, in *Carbohydrate Chemistry*, The Royal Society of Chemistry, Editon edn., 2012.
59 F. Dall'olio, N. Malagolini, M. Trinchera and M. Chiricolo, *Front Biosci*, 2012, **17**, 670–699.

60 P. Stanley and R. D. Cummings, 2009.
61 J. Sheinfeld, V. E. Reuter, A. S. Sarkis and C. Cordon-Cardo, *J Cell Biochem Suppl*, 1992, **16I**, 50–55.
62 C. Cordon-Cardo, V. E. Reuter, K. O. Lloyd, J. Sheinfeld, W. R. Fair, L. J. Old and M. R. Melamed, *Cancer Res*, 1988, **48**, 4113–4120.
63 J. Sheinfeld, V. E. Reuter, W. R. Fair and C. Cordon-Cardo, *Semin Surg Oncol*, 1992, **8**, 308–315.
64 K. Numahata, M. Satoh, K. Handa, S. Saito, C. Ohyama, A. Ito, T. Takahashi, S. Hoshi, S. Orikasa and S. I. Hakomori, *Cancer*, 2002, **94**, 673–685.
65 A. Tizzani, G. Cassetta, A. Cicigoi, P. Piana, A. Cerchier, F. Pecchio and P. Piantino, *Int J Biol Markers*, 1987, **2**, 121–124.
66 O. A. Mansour, S. A. Shouman, M. Khafagy, H. Basma, P. M. Liebmann, K. Schauenstein and O. el-Ahmady, *Lancet*, 1995, **346**, 512–513.
67 K. Noto, M. Fujime, H. Isobe, Y. Wakumato and Y. Kawachi, *Nihon Hinyokika Gakkai Zasshi*, 1997, **88**, 406–413.
68 E. M. Vestergaard, H. Wolf and T. F. Orntoft, *Clin Chem*, 1998, **44**, 197–204.
69 K. Nagao, Y. Itoh, K. Fujita and M. Fujime, *Int J Urol*, 2007, **14**, 795–799.
70 A. Hegele, V. Mecklenburg, Z. Varga, P. Olbert, R. Hofmann and P. Barth, *Anticancer Res*, 2010, **30**, 5195–5200.
71 M. Ugorski and A. Laskowska, *Acta Biochimica Polonica*, 2002, **49**, 303–311.
72 S. Gunia, C. Kakies, M. May, S. Koch and A. Erbersdobler, *J Clin Pathol*, 2011, **64**, 672–676.
73 F. Dall'Olio and M. Chiricolo, *Glycoconjugate Journal*, 2001, **18**, 841–850.
74 Z. F. Dobrowolski, D. Dus, J. Halasa and C. Radzikowski, *International Urology and Nephrology*, 1995, **27**, 395–404.
75 N. C. Langkilde, *Scandinavian Journal of Urology and Nephrology*, *Supplementum*, 1995, **172**, 45–49.
76 N. C. Langkilde and T. F. Orntoft, *Scandinavian Journal of Urology and Nephrology*. *Supplementum*, 1995, **172**, 57–64.
77 N. C. Langkilde, H. Wolf, H. Clausen and T. F. Orntoft, *Cancer Research*, 1992, **52**, 5030–5036.
78 C. B. Pinnock, D. J. Roxby, J. M. Ross, C. H. Pozza and V. R. Marshall, *British Journal of Urology*, 1995, **75**, 461–469.
79 C. Limas and P. Lange, *Cancer*, 1986, **58**, 1236–1245.
80 T. P. Lehman, H. S. Cooper and S. G. Mulholland, *Cancer*, 1984, **53**, 272–277.
81 K. Nakanishi, T. Kawai and M. Suzuki, *Histopathology*, 1993, **23**, 153–158.
82 S. K. Keay, Z. Szekely, T. P. Conrads, T. D. Veenstra, J. J. Barchi, Jr., C. O. Zhang, K. R. Koch and C. J. Michejda, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, **101**, 11803–11808.
83 S. Kawamura, C. Ohyama, R. Watanabe, M. Satoh, S. Saito, S. Hoshi, S. Gasa and S. Orikasa, *Int J Cancer*, 2001, **94**, 343–347.
84 F. Guan, K. Handa and S. I. Hakomori, *Proc Natl Acad Sci U S A*, 2009, **106**, 7461–7466.
85 R. Watanabe, C. Ohyama, H. Aoki, T. Takahashi, M. Satoh, S. Saito, S. Hoshi, A. Ishii, M. Saito and Y. Arai, *Cancer Research*, 2002, **62**, 3850–3854.
86 A. R. Todeschini, J. N. Dos Santos, K. Handa and S. I. Hakomori, *Proc Natl Acad Sci U S A*, 2008, **105**, 1925–1930.
87 F. Nollet, P. Kools and F. van Roy, *J Mol Biol*, 2000, **299**, 551–572.
88 U. Cavallaro, B. Schaffhauser and G. Christofori, *Cancer Lett*, 2002, **176**, 123–128.
89 P. Hulpiau and F. van Roy, *Int J Biochem Cell Biol*, 2009, **41**, 349–369.
90 A. Jeanes, C. J. Gottardi and A. S. Yap, *Oncogene*, 2008, **27**, 6920–6929.

91 A. Clairotte, I. Lascombe, S. Fauconnet, F. Mauny, S. Felix, M. P. Algros, H. Bittard and B. Kantelip, *Am J Clin Pathol*, 2006, **125**, 119–126.
92 P. P. Bringuier, R. Umbas, H. E. Schaafsma, H. F. Karthaus, F. M. Debruyne and J. A. Schalken, *Cancer Res*, 1993, **53**, 3241–3245.
93 R. R. Byrne, S. F. Shariat, R. Brown, M. W. Kattan, R. J. Morton, T. M. Wheeler and S. P. Lerner, *J Urol*, 2001, **165**, 1473–1479.
94 P. K. Lipponen and M. J. Eskelinen, *J Cancer Res Clin Oncol*, 1995, **121**, 303–308.
95 A. Mialhe, J. Louis, S. Montlevier, M. Peoch, D. Pasquier, J. L. Bosson, J. J. Rambeaud and D. Seigneurin, *Invasion Metastasis*, 1997, **17**, 124–137.
96 J. S. Ross, A. D. del Rosario, H. L. Figge, C. Sheehan, H. A. Fisher and H. X. Bui, *Hum Pathol*, 1995, **26**, 940–944.
97 R. L. Hong, Y. S. Pu, T. S. Hsieh, J. S. Chu and W. J. Lee, *J Urol*, 1995, **153**, 2025–2028.
98 K. N. Syrigos, K. Harrington, J. Waxman, T. Krausz and M. Pignatelli, *J Urol*, 1998, **160**, 1889–1893.
99 T. Otto, A. Bex, U. Schmidt, A. Raz and H. Rubben, *Am J Pathol*, 1997, **150**, 1919–1923.
100 Z. Popov, S. Gil-Diez de Medina, M. A. Lefrere-Belda, A. Hoznek, S. Bastuji-Garin, C. C. Abbou, J. P. Thiery, F. Radvanyi and D. K. Chopin, *Br J Cancer*, 2000, **83**, 209–214.
101 E. M. Bindels, M. Vermey, R. van den Beemd, W. N. Dinjens and T. H. Van Der Kwast, *Cancer Res*, 2000, **60**, 177–183.
102 U. H. Frixen, J. Behrens, M. Sachs, G. Eberle, B. Voss, A. Warda, D. Löchner and W. Birchmeier, *The Journal of Cell Biology*, 1991, **113**, 173–185.
103 K. Vleminckx, L. Vakaet, Jr., M. Mareel, W. Fiers and F. van Roy, *Cell*, 1991, **66**, 107–119.
104 S. F. Shariat, K. Matsumoto, R. Casella, W. Jian and S. P. Lerner, *Eur Urol*, 2005, **48**, 69–76.
105 K. Matsumoto, S. F. Shariat, R. Casella, T. M. Wheeler, K. M. Slawin and S. P. Lerner, *J Urol*, 2003, **170**, 2248–2252.
106 T. R. Griffiths, I. Brotherick, R. I. Bishop, M. D. White, D. M. McKenna, C. H. Horne, B. K. Shenton, D. E. Neal and J. K. Mellon, *Br J Cancer*, 1996, **74**, 579–584.
107 A. S. Protheroe, R. E. Banks, M. Mzimba, W. H. Porter, J. Southgate, P. N. Singh, M. Bosomworth, P. Harnden, P. H. Smith, P. Whelan and P. J. Selby, *Br J Cancer*, 1999, **80**, 273–278.
108 R. T. Bryan and C. Tselepis, *J Urol*, 2010, **184**, 423–431.
109 M. J. Wheelock, Y. Shintani, M. Maeda, Y. Fukumoto and K. R. Johnson, *J Cell Sci*, 2008, **121**, 727–735.
110 M. Przybylo, D. Hoja-Lukowicz, A. Litynska and P. Laidler, *Cancer Cell International*, 2002, **2**, 6.
111 M. Granovsky, J. Fata, J. Pawling, W. J. Muller, R. Khokha and J. W. Dennis, *Nat Med*, 2000, **6**, 306–312.
112 H. Ishimura, T. Takahashi, H. Nakagawa, S. Nishimura, Y. Arai, Y. Horikawa, T. Habuchi, E. Miyoshi, A. Kyan, S. Hagisawa and C. Ohyama, *Clin Cancer Res*, 2006, **12**, 2506–2511.
113 D. Cox, M. Brennan and N. Moran, *Nat Rev Drug Discov*, 2010, **9**, 804–820.
114 C. M. Longhurst and L. K. Jennings, *Cell Mol Life Sci*, 1998, **54**, 514–526.
115 J. W. Tamkun, D. W. DeSimone, D. Fonda, R. S. Patel, C. Buck, A. F. Horwitz and R. O. Hynes, *Cell*, 1986, **46**, 271–282.
116 R. Pytela, M. D. Pierschbacher and E. Ruoslahti, *Proc Natl Acad Sci U S A*, 1985, **82**, 5766–5770.

117 D. J. White, S. Puranen, M. S. Johnson and J. Heino, *Int J Biochem Cell Biol*, 2004, **36**, 1405–1410.
118 S. K. Akiyama, *Hum Cell*, 1996, **9**, 181–186.
119 E. Pochec, A. Litynska, M. Bubka, A. Amoresano and A. Casbarra, *European Journal of Cell Biology*, 2006, **85**, 47–57.
120 M. A. Hollingsworth and B. J. Swanson, *Nat Rev Cancer*, 2004, **4**, 45–60.
121 H. S. Silverman, S. Parry, M. Sutton-Smith, M. D. Burdick, K. McDermott, C. J. Reid, S. K. Batra, H. R. Morris, M. A. Hollingsworth, A. Dell and A. Harris, *Glycobiology*, 2001, **11**, 459–471.
122 D. P. Scholfield, M. S. Simms and M. C. Bishop, *BJU Int*, 2003, **91**, 560–566.
123 O. D. Hughes, A. C. Perkins, M. Frier, M. L. Wastie, G. Denton, M. R. Price, H. Denley and M. C. Bishop, *BJU Int*, 2001, **87**, 39–46.
124 O. D. Hughes, M. C. Bishop, A. C. Perkins, M. L. Wastie, G. Denton, M. R. Price, M. Frier, H. Denley, R. Rutherford and P. A. Schubiger, *J Clin Oncol*, 2000, **18**, 363–370.
125 T. Okegawa, M. Kinjo, S. Horie, K. Nutahara and E. Higashihara, *Urology*, 2003, **62**, 182–186.
126 H. Kresse and E. Schonherr, *J Cell Physiol*, 2001, **189**, 266–274.
127 E. Schonherr and R. Heinz-JüHausser, *Clinical and Developmental Immunology*, 2000, **7**, 89–101.
128 V. B. Lokeshwar, G. L. Schroeder, M. G. Selzer, S. H. Hautmann, J. T. Posey, R. C. Duncan, R. Watson, L. Rose, S. Markowitz and M. S. Soloway, *Cancer*, 2002, **95**, 61–72.
129 G. L. Schroeder, M. F. Lorenzo-Gomez, S. H. Hautmann, M. G. Friedrich, S. Ekici, H. Huland and V. Lokeshwar, *J Urol*, 2004, **172**, 1123–1126.
130 A. Raedler and S. Schreiber, *Crit Rev Clin Lab Sci*, 1988, **26**, 153–193.
131 A. M. Wu, E. Lisowska, M. Duk and Z. Yang, *Glycoconj J*, 2009, **26**, 899–913.
132 W. I. Weis and K. Drickamer, *Annu Rev Biochem*, 1996, **65**, 441–473.
133 G. K. Ward, S. S. Stewart, G. B. Price and W. J. Mackillop, *Histochem J*, 1987, **19**, 337–344.
134 L. Neutsch, V. E. Plattner, S. Polster-Wildhofen, A. Zidar, A. Chott, G. Borchard, O. Zechner, F. Gabor and M. Wirth, *J Urol*, 2011, **186**, 1481–1488.
135 H. Lahm, A. Hoeflich, S. Andre, B. Sordat, H. Kaltner, E. Wolf and H. J. Gabius, *Int J Oncol*, 2000, **17**, 519–524.
136 F. T. Liu and G. A. Rabinovich, *Nature reviews.Cancer*, 2005, **5**, 29–41.
137 I. Camby, M. Le Mercier, F. Lefranc and R. Kiss, *Glycobiology*, 2006, **16**, 137R–157R.
138 O. B. Garner and L. G. Baum, *Biochem Soc Trans*, 2008, **36**, 1472–1477.
139 R. D. Cummings and F. T. Liu, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, The Consortium of Glycobiology Editors, La Jolla, California, Cold Spring Harbor NYEditon edn., 2009.
140 N. L. Perillo, M. E. Marcus and L. G. Baum, *J Mol Med (Berl)*, 1998, **76**, 402–412.
141 A. Danguy, I. Camby and R. Kiss, *Biochim Biophys Acta*, 2002, **1572**, 285–293.
142 R. C. Hughes, *Biochimie*, 2001, **83**, 667–676.
143 D. K. Hsu, R. Y. Yang and F. T. Liu, *Methods Enzymol*, 2006, **417**, 256–273.
144 S. Nakahara and A. Raz, *Methods Enzymol*, 2006, **417**, 273–289.
145 L. Cindolo, G. Benvenuto, P. Salvatore, R. Pero, G. Salvatore, V. Mirone, D. Prezioso, V. Altieri, C. B. Bruni and L. Chiariotti, *Int J Cancer*, 1999, **84**, 39–43.
146 S. H. Barondes, D. N. Cooper, M. A. Gitt and H. Leffler, *J Biol Chem*, 1994, **269**, 20807–20810.

147 H. Leffler, S. Carlsson, M. Hedlund, Y. Qian and F. Poirier, *Glycoconj J*, 2004, **19**, 433–440.
148 C. Fred Brewer, *Biochim Biophys Acta*, 2002, **1572**, 255–262.
149 G. Canesin, P. Gonzalez-Peramato, J. Palou, M. Urrutia, C. Cordon-Cardo and M. Sanchez-Carbayo, *Tumour Biol*, 2010, **31**, 277–285.
150 N. Oka, S. Nakahara, Y. Takenaka, T. Fukumori, V. Hogan, H. O. Kanayama, T. Yanagawa and A. Raz, *Cancer Res*, 2005, **65**, 7546–7553.
151 L. Ingrassia, I. Camby, F. Lefranc, V. Mathieu, P. Nshimyumukiza, F. Darro and R. Kiss, *Curr Med Chem*, 2006, **13**, 3513–3527.
152 Y. Matsui, S. Ueda, J. Watanabe, I. Kuwabara, O. Ogawa and H. Nishiyama, *Cancer Res*, 2007, **67**, 1212–1220.

Levansucrases of *Pseudomonas* bacteria: novel approaches for protein expression, assay of enzymes, fructooligosaccharides and heterooligofructans

Tiina Alamäe,*[a] Triinu Visnapuu,[a] Karin Mardo,[a] Andres Mäe[a] and Alina D. Zamfir[b,c]

DOI: 10.1039/9781849734769-00176

Levansucrases are bacterial extracellular enzymes that act on sucrose producing β-2,6-linked fructans of varied chain length. We summarize here our up-to-date results on (i) novel methods for the production of heterologous levansucrase proteins in *Escherichia coli*, (ii) biochemical characterization of the levansucrases encoded in the genome of a tomato and *Arabidopsis* pathogen *Pseudomonas syringae* pv. tomato DC3000, (iii) innovative chip-based mass spectrometric analysis of oligosaccharidic reaction products, (iv) isolation and characterization of the first mutants and structure-function analysis of a levansucrase of *P. syringae* pv. tomato. Data on the levansucrase LscA of *P. chlororaphis* subsp. *aurantiaca* will be presented for comparison. As most important features of the levansucrases from *Pseudomonas* bacteria we emphasize their ability to synthesize potentially prebiotic fructooligosaccharides from sucrose or raffinose and to transfructosylate a variety of alternative acceptor sugars with production of heterooligofructans.

1 Levansucrase genes and proteins

Levansucrases are bacterial extracellular enzymes belonging to family 68 of glycoside hydrolases according to Carbohydrate-Active Enzymes database (http://www.cazy.org/).[1] They use sucrose (α-Glc-1,2-β-Fru; GF) as a substrate to synthesize β-2,6-linked fructans: short-chain fructooligosaccharides (FOS) and also polymeric levan.[2] Levansucrases catalyse following reactions:

(i) hydrolysis of the substrate (sucrose): $\mathbf{GF + H_2O \rightarrow G + F}$;
(ii) polymerization: $\mathbf{nGF + acceptor \rightarrow nG + F_n\text{-}acceptor}$,

where G corresponds to glucose and F to fructose.

Levansucrases are present in Gram-positive bacteria such as *Bacillus subtilis*, *B. megaterium*, *Lactobacillus sanfranciscensis* and *L. reuteri*, but also in Gram-negative bacteria *e.g. Zymomonas mobilis*, *Gluconacetobacter diazotrophicus* and some *Pseudomonas* bacteria. Levansucrase genes have been found in the genomes of *Pseudomonas chlororaphis* subsp. *aurantiaca* (also *P. aurantiaca*, *Pca*), *P. fluorescens* SBW25 and many *Pseudomonas syringae* strains (http://www.cazy.org/GH68_bacteria.html).

[a]*Institute of Molecular and Cell Biology, University of Tartu, Riia 23, 51010 Tartu, Estonia. E-mail: talamae@ebc.ee*
[b]*Mass Spectrometry Laboratory, National Institute for Research and Development in Electrochemistry and Condensed Matter, Plautius Andronescu 1, 300224 Timisoara, Romania.*
[c]*Department of Chemical and Biological Sciences, Aurel Vlaicu University of Arad, 310130 Arad, Romania.*

© The Royal Society of Chemistry 2012

During recent years, we have been focusing on characterization of levansucrases of *P. syringae* pv. tomato DC3000 (*Pst*) and their biosynthesis products. *P. syringae* is a plant pathogenic bacterium, strains of which are assigned to over 50 pathovars according to the host plant. *P. syringae* infects strain-specifically bean, tomato, olive, cucumber and some other economically important plants. *Pst* causes bacterial speck on tomato. This strain has specific position among the *P. syringae* strains because it also infects a popular model plant *Arabidopsis thaliana* with its genome sequence available. Therefore, functional genomics can be applied to contribute to the knowledge of plant-bacterium relationship. Data on completed and ongoing sequencing projects of *P. syringae* strains can be found at http://www.pseudomonas-syringae.org/.

Analysis of completed genomes of *P. syringae* pv. tomato DC3000, *P. syringae* pv. syringae B728a and *P. syringae* pv. phaseolicola 1448A shows that most *P. syringae* pathovars possess three levansucrase alleles, two chromosomal and one plasmid-born. In case of *Pst*, *lsc1* and *lsc2* are chromosomal and *lsc3* resides on a plasmid. Strain B728a has no plasmids and its chromosome encodes two levansucrase proteins. The length of levansucrase proteins deduced from respective open reading frames of *P. syringae* pathovars is either 415 or 431 amino acids (aa). Two types of levansucrases of *P. syringae* differ mostly in their N-termini: the 431 aa proteins have N-terminal extensions of 16 aa that are lacking in 415 aa proteins, otherwise they share a high (89–99%) level of identity.[3] At least one of the three *lsc* genes of *P. syringae* DC3000 must be expressed in native host, because if grown on sucrose-containing medium, it synthesizes levan and has a mucoid phenotype. LscA of *Pca* is 424 aa long and its identity with levansucrase proteins of *Pst* tomato is around 72%.

2 Heterologous expression of levansucrases from *Pseudomonas syringae* pv. tomato with two different expression systems

2.1 Expression from P_{MAL1} using the pHIPMalprom vector

We have shown that P_{MAL1} – the promoter of the maltase gene (*HpMAL1*, AL432586) from a methylotrophic yeast *Hansenula polymorpha* (synonym *Pichia angusta*) functions constitutively in *Escherichia coli*. It is attributed to the presence of two bacterial sigma 70-like boxes, TTGACA-N17-TAAATT and GGTACA-N17-TATTAT in P_{MAL1} with high identity to sigma 70 consensus TTGACA-N17-TATAAT (shown underlined).[3,4] This feature of P_{MAL1} was applied to express levansucrases of *Pst* in *E. coli* using a plasmid-born expression system described in Visnapuu *et al.* (2008).[3] *Lsc* genes were cloned to pHIPMalprom with their native Shine-Dalgarno sequences.

Laboratory strains of *E. coli* do not metabolize sucrose. Thus, levansucrase-synthesizing *E. coli* will acquire a sucrose-positive phenotype – it will grow on sucrose as sole carbon source in liquid culture and produce slimy levan on solid media containing sucrose (see Fig. 1).[3]

All three levansucrase genes (*lsc1*, *lsc2* and *lsc3*) of *Pst* expressed well in *E. coli*. The level of expression depended on the host strain and on the gene to be expressed *i.e.* the highest levansucrase activity (70 U/mg) was recorded for *lsc3* in *E. coli* HB101 (*lac*+) and the lowest (7 U/mg protein) for *lsc2* in

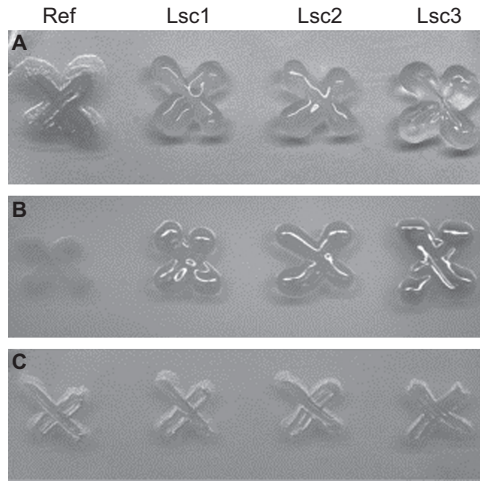


Fig. 1 *E. coli* RA11r transformants harboring pHIPMalprom-lsc1 (Lsc1), pHIPMalprom-lsc2 (Lsc2) or pHIPMalprom-lsc3 (Lsc3) have a mucoid phenotype due to levan synthesis when grown on LB-sucrose (A) and LB-raffinose (B) agar plates, whereas on LB-glucose agar plates they do not produce slimy polyfructan (C). *E. coli* transformed with empty vector pHIPMalprom is analyzed as a reference (Ref).

E. coli RA11r. Lsc1, Lsc2 and Lsc3 were prominent proteins in recombinant *E. coli* cell extracts, whereas Lsc3 had the highest expression level comprising about 20% of the total soluble protein. Using the P_{MAL1}-driven expression system, inclusion body formation was not detected.[3] Thus P_{MAL1} has appropriate strength to produce an adequate amount of soluble catalytically active levansucrase protein in *E. coli*. Applying this expression system, Lsc2 and Lsc3 proteins were purified from *E. coli* HB101 (*lac*+) extracts using size-exclusion chromatography on Sephacryl S-300 column for further biochemical characterization.[3,5]

2.2 Expression from P_{T7} using the pURI3 vector

The pURI3 vector[6] was used to produce wild-type and mutated Lsc3 proteins of *Pst* in *E. coli* BL21(DE3) from the plasmid-based T7 promoter. The pURI family of cloning vectors enables production of recombinant His-tagged fusion proteins in *E. coli* without enzymatic restriction and ligation steps during the cloning.[7] Implementing the pURI3 vector system, levansucrases were produced with N-terminal His_6-tag that enabled their easy purification by Ni^{2+}-affinity chromatography.[8] Extracts of *E. coli* BL21(DE3) expressing His-tagged Lsc3 had total levansucrase activity up to 200 U/mg that is much higher than respective activity achieved when using the pHIPMalprom vector (see above).

3 Biochemical properties of the levansucrases: substrate specificity and kinetic parameters

Levansucrases (EC 2.4.1.10; sucrose:2,6-β-D-fructan 6-β-D-fructosyltransferases) carry out sucrose hydrolysis using water for fructosyl acceptor and transfructosylation of sucrose, FOS or some alternative acceptor molecule. In addition to sucrose, many levansucrases react with a trisaccharide raffinose (α-Gal-1,6-α-Glc-1,2-β-Fru; GalGF) that is also a

widespread sugar in plants.[9–11] Levansucrase protein of *P. syringae* pv. phaseolicola was reported incapable of raffinose hydrolysis.[12] Nevertheless, our experimental data indicated that all three levansucrases of *Pst* use raffinose. It was first suspected from colony phenotype of levansucrase-expressing *E. coli* RA11r growing on LB-raffinose plates (Fig. 1). Regular laboratory strains of *E. coli* possess melibiase (α-galactosidase) splitting raffinose to galactose and sucrose whereas the RA11r strain is melibiase-negative (see references in[3]) and has therefore suitable background to evaluate levansucrase-mediated raffinose utilization. The RA11r colonies expressing either *lsc1*, *lsc2* or *lsc3* of *Pst* produce mucoid colonies on LB agar plates with sucrose or raffinose (Fig. 1). Levansucrase assay of respective bacterial extracts proved release of reducing sugars from both, sucrose and raffinose. Additionally, native polyacrylamide gel electrophoresis of cell extracts indicated levan production from both substrates.[3]

Since substrate specificity data of the LscA levansucrase of *Pca* were absent in the literature, we addressed this issue. LscA, purchased from Sigma-Aldrich, produced reducing sugars not only from sucrose but also from raffinose and a tetrasaccharide stachyose (α-Gal-1,6-α-Gal-1,6-α-Glc-1,2-β-Fru; Gal_2GF).[8] We have shown that all three levansucrases of *Pst* are capable of stachyose splitting ([8] and our unpublished data).

All four *Pseudomonas*-related levansucrases addressed here exhibit some levan-hydrolyzing activity which is less than 1% of respective sucrose-hydrolyzing activity with fructose detected as sole end-product ([8] and our unpublished data). Thus, levan can be considered as extracellular reserve polymer for respective bacteria contributing to their survival in the environment *i.e.* on plant surface.

K_m and k_{cat} values for sucrose splitting and K_i for inhibition of sucrose splitting reaction by raffinose of three levansucrases (Lsc2, Lsc3 and LscA) from pseudomonads are presented in Table 1.

4 Predicted catalytic triad residues and 3D structures of Lsc1, Lsc2, Lsc3 and LscA proteins

Since now, crystallization studies of levansucrases have been focusing mostly on respective proteins of Gram-positive bacteria – *B. subtilis* and

Table 1 Kinetic characteristics of sucrose and raffinose splitting by Lsc2 and Lsc3 from *P. syringae* pv. tomato DC3000 and LscA from *P. chlororaphis* subsp. *aurantiaca*.

Protein	K_m for sucrose (mM)	k_{cat} for sucrose (min^{-1})	k_{cat}/K_m for sucrose ($M^{-1} \times min^{-1}$)	K_i for raffinose (mM)
Lsc2	25.3 ± 1.6	1.64×10^4	6.47×10^5	53.2 ± 4.3
Lsc3	20.6 ± 2.1^a	$1.37 \times 10^{4a,b}$	6.63×10^{5b}	47.6 ± 5.0^a
Lsc3*	18.5 ± 2.5^b	3.03×10^{4b}	1.64×10^{6b}	39.9 ± 6.1
LscA	24.1 ± 1.0^b	4.32×10^{2b}	1.79×10^{4b}	80.8 ± 11.8^b

*His-tagged.
Molecular weights of the proteins used to calculate k_{cat} values: 45.9 kDa (Lsc2), 47.8 kDa (Lsc3), 49.6 kDa (His-tagged Lsc3), 47.0 kDa (LscA).
[a] Data from[3].
[b] Data from[8].

B. megaterium. Only levansucrases of these bacteria have been crystallized in complex with the substrate, sucrose or raffinose.[2,13–15] Crystal structure of LsdA of *G. diazotrophicus* (Protein Data Bank (PDB) code: 1W18) is the sole one to represent levansucrases of Gram-negative bacteria.[16] It should be noted that levansucrases of Gram-positive and -negative bacteria are rather diverse according to the sequence. For example, the sequence identity of Lsc3 of *Pst* with LsdA of *G. diazotrophicus* is 39% whereas that with SacB of *B. subtilis* is merely 19%.[8] Similarly, overall sequence identity between LsdA and SacB proteins is low – 26%.[16] Despite of that, LsdA and SacB have similar structure of five-bladed β-propeller with central acidic pocket harbouring catalytic triad amino acids, two aspartates (Asp) and a glutamate (Glu).[13,16] Putative active site residues of *Pseudomonas*-derived levansucrases predicted as in [8] are presented in Table 2.

In silico 3D structure modelling disclosed five-bladed β-propeller fold for Lsc3 of *Pst* and LscA of *Pca*.[8] The predicted topologies of Lsc3 and LscA were highly similar except for some external loops. One of these loops corresponds to region His90-Asp108 (Lsc3) or His81-Asp99 (LscA) that occurs as insertion between the conserved neighbouring regions in these two proteins (our unpublished data).

Here, we used similar ModWeb server-based modelling[17] to predict 3D structure for Lsc2 of *Pst*. Again, LsdA of *G. diazotrophicus* (PDB: 1W18) was used as a template. Figure 2 shows predicted five-bladed β-propeller structure of the Lsc2 with putative catalytic triad and His305, an equivalent of a polymerization determinant His321 of Lsc3[8] highlighted. As Lsc1 and Lsc3 proteins have the same length (431 aa) and almost identical sequence, their structures are most probably highly identical.

5 Lsc2, Lsc3 and LscA produce not only levan, but also FOS that can be detected by a novel mass spectrometric method

5.1 Characterization of the product spectrum of Lsc2, Lsc3 and LscA

The pattern of polymerization products, fructans, depends on reaction conditions, but also on a particular levansucrase. Up to now, it is not clear why some levansucrases, for example those of *Bacillus* bacteria, synthesize mainly high-molecular levan whereas some other proteins like LsdA from *G. diazotrophicus*, produce mostly FOS (see references in [8]). We hypothesize that polymerization pattern of a bacterial levansucrase may be related to natural environment of the bacterium. For instance, LsdA of a sugar cane endosymbiont *G. diazotrophicus* is solely responsible for nutrition of the

Table 2 Predicted catalytic triad amino acids of levansucrases from *P. syringae* pv. tomato DC3000 and *P. chlororaphis* subsp. *aurantiaca*.

	P. syringae pv. tomato DC3000			LscA of *P. chlororaphis* subsp. *aurantiaca*[a]
Catalytic amino acid	Lsc3[a]	Lsc1	Lsc2	
Nucleophile	Asp62	Asp62	Asp46	Asp53
Transition state stabilizer	Asp219	Asp219	Asp203	Asp210
Acid-base catalyst	Glu303	Glu303	Glu287	Glu294

[a] Data from[8].

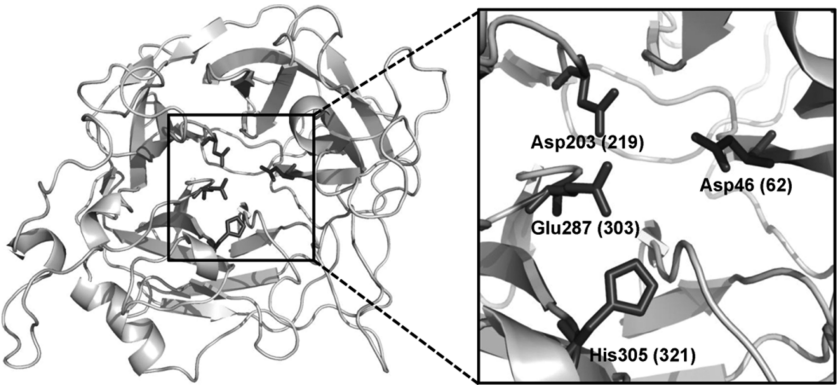


Fig. 2 Predicted 3D structure of Lsc2 from *P. syringae* pv. tomato with catalytic centre zoomed in and catalytically important amino acids indicated. Corresponding amino acid positions of Lsc3 are shown in brackets.

bacterium on sucrose in plant sap. LsdA produces a large amount of short-chain FOS and only a very low amount of high-molecular levan ([16] and references therein). Accumulation of levan in sugar cane vascular tissues is probably disadvantageous for both partners.

E. coli expressing *lsc* genes of *Pst* synthesizes slimy fructan polymer on sucrose-containing agar plates (see Fig. 1). LscA of *P. aurantiaca* S-4380 forms polymer from sucrose that is a typical levan – it has high molecular weight (7×10^5 Da), β-2,6 linkages between fructose residues in the main chain and some branching through β-2,1 linkages.[18] Enzymatic digestion assay of Lsc3-produced levan predicted that levans of LscA and Lsc3 are highly similar from this aspect.[8] Analysis of the polymer produced by Lsc3 protein from sucrose revealed a 99:1 ratio of fructose to glucose that indicates its high degree of polymerization (DP) and molecular weight.[3]

Importantly, when we first analysed transfructosylation products of Lsc2 and Lsc3 by thin layer chromatography (TLC), we did not detect FOS, because considerably short reaction times (up to 30 min) and relatively low substrate concentrations (up to 100 mM) were applied. Still, levan formation from sucrose and raffinose was detected in reactions conducted under these conditions.[3] If we extended reaction time up to 20 h and substrate concentration till 1200 mM, FOS of different migration distance were clearly detectable among reaction products of Lsc3 and LscA with sucrose or raffinose.[5,8] Figure 3 illustrates FOS production from sucrose by Lsc2, Lsc3 and LscA proteins. LscA produces more high-DP FOS than Lsc3, whereas Lsc2 and LscA produce somewhat less levan than Lsc3.

Sugar beet molasses as a cheap sucrose-rich substrate were also shown suitable for FOS and levan production by Lsc3 and LscA. TLC analysis revealed that the spectra of products formed either from molasses or sucrose were highly similar.[8]

Polymerizing activity of levansucrases depends on reaction conditions *e.g.* pH, temperature and substrate concentration. Accordingly, polymerizing activities of Lsc3 and LscA were promoted at high substrate concentrations. At 1200 mM sucrose and 37 °C, transfructosylating

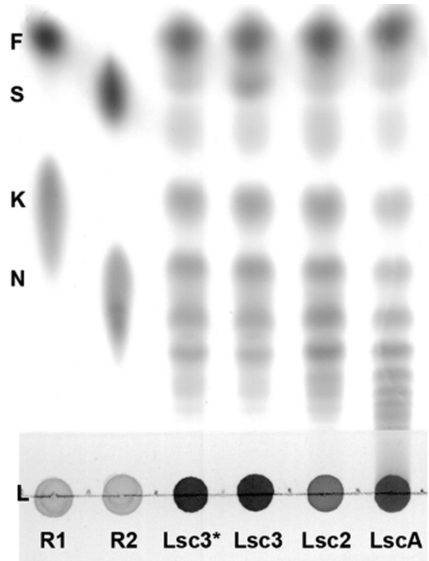


Fig. 3 TLC analysis of reaction products synthesized by levansucrases from 1200 mM sucrose. Reactions and TLC analysis were performed as in Visnapuu *et al.* (2009; 2011).[5,8] Reference sugars R1: fructose – F; 1-kestose – K; levan – L; R2: sucrose – S; nystose – N; levan – L. * His-tagged protein.

activities (TA) of Lsc3 and LscA were similarly high – 76% and 73% of fructose residues originating from reacted sucrose molecules were polymerized.[8] At 300 mM sucrose, Lsc3 had much higher TA (43%) than LscA (15%). TLC analysis of reaction samples indicated that at 300 mM sucrose, LscA produces some kestose (DP 3), nystose (DP 4) and a small amount of levan whereas Lsc3 produces more levan and FOS. Patterns of FOS synthesized from 1200 mM sucrose by LscA and Lsc3 were similar, but LscA was definitely better producer of high-DP FOS (see also Fig. 3). Compared to LscA, Lsc3 is at least twice better levan producer at both, low (300 mM) and high (1200 mM) sucrose concentration.[8] These two proteins are therefore clearly different from LsdA of *G. diazotrophicus* which produces 1-kestose (DP 3) as a major transfructosylation product.[19]

Study of transfructosylating activity at pH values 5.0–7.0 revealed no pH-dependence of this function for Lsc3 and LscA whereas high temperature favoured the hydrolysis reaction.[8]

5.2 Chip-based electrospray mass spectrometry as a valuable tool for glycobiology

The potential of the modern chip-based electrospray ionization (ESI) systems expanded significantly the area of mass spectrometry (MS) applicability in glycobiology. The option for miniaturized, integrated devices for sample infusion into MS was driven by the considerable increase in sensitivity, ionization efficiency, reproducibility and throughput as compared to classical capillary-based electrospray.[20,21] However, because of the lower ionization efficiency as compared to peptides and proteins, glycans required reconsideration of the conditions to promote chip ionization. Moreover,

these conditions needed to be optimized for each class of carbohydrates. The optimization procedures were shown to be complex and dependent on the i) carbohydrate or non-carbohydrate type of labile attachments on the sugar chain such as fucosylation, sialylation, sulfation, phosphorylation, acetylation *etc*., ii) ionizability of the functional groups, iii) hydrophilic and/or hydrophobic nature of molecules, iv) branching of the sugar chains and v) the type of the aglycon in the case of conjugated glycans. Development of effective protocols for system operation in the negative ion mode to detect carbohydrates that form rather anions was another challenging task due to the low ionization efficiency that ultimately leads to decreased sensitivity.

Up to now, two types of chip-based devices for ESI MS were successfully implemented in glycobiology. The first category is represented by the out-of-plane devices, where nozzle-like nanospray emitters are integrated onto a silicon substrate from which electrospray is established perpendicular to the substrate.[21] Incorporated into a robot for automatic sample delivery, the silicon chips yielded the NanoMate platform, uniquely suited to high-throughput glycan analyses by ESI MS at remarkable flow rates (50–150 nl/min) and sensitivities (low picomole to femtomole range). The second chip-ESI system introduced in glycomics consists of a planar polymer microchip embedding a microchannel at the end of which ESI is generated in-plane.[22] These microsprayers provide a high stability of the spray in time, higher tolerance to salts, improved signal-to-noise ratios at various flow rates (200–400 nl/min) and flexibility to different ion source configurations, with the additional advantage of cheaper production costs compared to silicon technologies.

As a result of the very efficient ionization properties, both chips for ESI preferentially form multiply charged ions and minimize the in-source fragmentation of the groups attached to the sugar core through linkages that cleave readily. Due to its versatility and unsurpassed sensitivity NanoMate system is so far dominant chip-ESI system for glycoscreening and sequencing. For glycomic studies, NanoMate has been coupled with a variety of mass spectrometers[21,23–31] such as quadrupole-time-of flight (QTOF), high capacity ion trap (HCT), Fourier-transform ion cyclotron (FTICR), Orbitrap MS and several fragmentation techniques[21,24–27,29,30] like collision-induced dissociation (CID), electron-transfer dissociation (ETD), sustained-off resonance (SORI)-CID and multistage CID MS (MS^n). NanoMate has been applied to almost all glycan and glycoconjugate categories[5,8,32–36] originating from different biological or synthetic matrices and, under properly optimized conditions, provided in all cases highly reproducible and sensitive analyses, detection of minor components in complex mixtures and discovery of biomarker species, often without the need of separation prior to MS analysis.

5.3 Identification of FOS synthesized by the levansucrases by chip-based nanoESI HCT MS

Before our studies, some reports on conventional ESI MS application and optimization for the analysis of levansucrase reaction products were available. It has been used to confirm methyl-fructoside formation from sucrose and methanol by the *Rahnella aquatilis* levansucrase[37] and to study

composition of oligosaccharidic reaction products synthesized from sucrose and its analogs by wild-type and mutant levansucrases of *B. subtilis*.[38]

Although presence of FOS among reaction products of Lsc3 and LscA was detected by us using a traditional TLC, we intended to exert accurate and up-to-date MS method to specify the length of the FOS species. Experimental conditions for fully automated chip-based nanoESI, performed on the NanoMate robot incorporating a 400 nozzle silicon chip for ESI coupled to HCT MS, were optimized.[5] Both, negative and positive ion modes of the NanoMate-HCT MS platform could be used whereas formate or phosphate and sodium additives of FOS were observed, respectively. Those adducts of the saccharidic ions most probably originated from the reaction buffer (McIllvaine's buffer) and the solution for MS sample preparation (it contained methanol and formic acid). Also, hydrated ions were frequently spotted. Though ionization efficiency of neutral underivatized oligosaccharides is limited and relative abundance of larger ions thereby tends to be reduced, we detected FOS with DP up to 5 or 6 formed from sucrose and raffinose by Lsc3 and LscA, respectively. To confirm the saccharidic origin of the isolated ions, MS^n was applied by employing CID as the fragmention technique. Depending on the sequencing conditions we obtained product ion spectra of FOS-derived precursor ions up to MS^3.[5,8] The NanoMate-HCT MS method confirmed our TLC results – at high substrate concentration FOS synthesis prevailed. Regarding the experimental conditions, we obtained best-quality spectra of FOS if levansucrase reactions were carried out at high (1200 mM) substrate concentration. It should be noted that thermally inactivated protein, residual (unreacted) substrate and buffer components of the reaction mixture did not hinder FOS detection. However, a long-lasting steady spray was obtained when reaction samples were dialysed or levansucrase reactions were conducted in deionized water.[5,8]

Our study is the first report on introduction and optimization of novel and highly sensitive chip-based nanoESI HCT MS method for the analysis of underivatized FOS mixtures formed from sucrose and raffinose by a levansucrase protein. Commercial prebiotic inulin-based oligofructose mixtures (Orafti®P95 and Orafti®Synergy1) were analyzed to validate the method. The spectra of commercial FOS were highly similar to those produced from sucrose by Lsc3 and LscA.[5] This suggests potential prebiotic effect of FOS synthesized by levansucrases of pseudomonads.

5.4 Chip-based nanoESI HCT MS revealed synthesis of heterooligofructans by Lsc3 and LscA

Levansucrases and fructosyl transferases from Gram-positive bacteria can besides sucrose transfructosylate also nonconventional acceptors and thereby produce sucrose analogs.[11,38–40] So, for levansucrases of Gram-negative bacteria *P. aurantiaca* and *Z. mobilis*, transfructosylation of lactose yielding bifidogenic lactosucrose has been described.[41,42]

We used above-mentioned rapid and straightforward chip-based nanoESI HCT MS method [5] to screen possible fructosyl acceptors for LscA and Lsc3. D-xylose, D-fucose, L- and D-arabinose, D-ribose, D-sorbitol, xylitol, xylobiose, D-mannitol, D-galacturonic acid, methyl-α-D-glucopyranoside

Table 3 Nonconventional fructosyl acceptors of Lsc3 of *P. syringae* pv. tomato and LscA of *P. chlororaphis* subsp. *aurantiaca*. The degree of polymerization (DP) of mass spectrometrically detected products is shown. Data extracted from Supplementary Table S1 of [8].

Acceptor	Formula of acceptor	DP of Lsc3 products	DP of LscA products
D-arabinose	($C_5H_{10}O_5$)	2-5	2-4
L-arabinose	($C_5H_{10}O_5$)	2-5	2-4
D-fucose	($C_6H_{12}O_5$)	2-5	2-4
D-sorbitol	($C_6H_{14}O_6$)	2-5	2-4
D-xylose	($C_5H_{10}O_5$)	2-5	2-4
D-ribose	($C_5H_{10}O_5$)	2-5	2-4
Xylitol	($C_5H_{12}O_5$)	2-5	2-4
D-mannitol	($C_6H_{14}O_6$)	2-5	2-4
D-galacturonic acid	($C_6H_{10}O_7$)	2-4	2-4
Methyl-α-D-glucopyranoside	($C_7H_{14}O_6$)	2-5	2-4
Xylobiose	($C_{10}H_{18}O_9$)	2-5	2-4

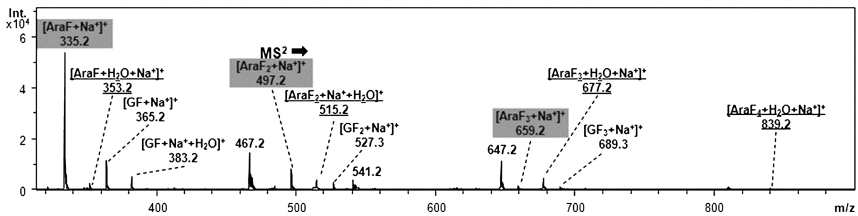


Fig. 4 (+) chip-based nanoESI HCT MS of oligofructans in levansucrase (Lsc3) reaction mixture containing sucrose as a substrate and D-arabinose as an acceptor molecule. Sodiated heterooligofructans are shown on gray background and their hydrated forms are designated as underlined. Cropped figure reproduced from Visnapuu *et al.*, 2011[8] with permission from Elsevier. Ara – arabinose; F – fructose; G – glucose; Na – sodium.

and D-glucosamine were tested for the acceptor function. While all these substrates have molecular masses different from glucose or fructose, hetero-oligofructans (HOF) are clearly detectable if formed in the reaction. The samples for MS analysis were prepared and reactions were conducted as described by us earlier.[8] MS analysis of the samples indicated that all tested nonconventional acceptors except for D-glucosamine were used by Lsc3 and LscA. The chain lengths of HOF detected in case of every acceptor are listed in Table 3.

Importantly, for preliminary screening of the biosynthesis products, reaction mixtures were analyzed by MS without prior purification. Despite of it, the mass spectra were of sufficient quality to specify HOF. We could detect four different series of sodiated oligosaccharidic ions: 1) HOF with 1-4 fructose residues added to the acceptor residue; 2) hydrated forms of HOF with DP up to 5; 3) conventional FOS with DP up to 5 produced from sucrose as a donor and acceptor and 4) hydrated species of conventional FOS. No non-covalent complexes between the acceptor, substrate and fructose were formed showing accuracy of the experiments. As an example, Fig. 4 shows the mass spectra of Lsc3 transfructosylation products with D-arabinose as an acceptor. Saccharidic origin of the detected HOF was concluded by CID MS^n.[8]

Significantly, we showed for the first time that levansucrases can transfructosylate D-sorbitol, D-galacturonic acid, D-mannitol, xylitol, methyl-α-D-glycopyranoside and a disaccharide xylobiose.[8] We presume that use of alternative fructosyl acceptors and synthesis of HOF is probably a common feature of levansucrases. It may be due to relaxed binding properties of the +1 and further subsites of the active centre. According to the ^{1}H and ^{13}C NMR spectroscopy data of previously studied HOF synthesized by levansucrases of bacilli, we assume that the linkage type between a D-isomer of nonconventional acceptor sugar and adjacent fructose residue is most probably the same as between D-glucose and fructose in sucrose molecule *i.e.* α-1,2[11,38] whereas the bond between following fructose residues is most probably β-2,6 as typical for bacterial levans.[38] We consider that chip-based nanoESI HCT MS is a feasible high-throughput method to screen possible acceptor molecules of various glycosyl transferases.

6 Isolation and novel screening methods for levansucrase mutants

6.1 Isolation strategies of Lsc3 mutants

Lsc3 protein was selected to initiate structure-function analysis of levansucrases of pseudomonads using a traditional approach – isolation and characterization of mutant proteins. First, we applied random ethylmethane sulfonate (EMS) mutagenesis of *E. coli* transformants that harboured the *lsc3* gene on a plasmid pHIPMalprom-lsc3. The mutagenized transformants were plated out to obtain a large amount of colonies. The total plasmid pool from these colonies was used as template to amplify the *lsc3* gene fragment for recloning to pHIPMalprom. Then *E. coli* was transformed with this new pool of plasmids and transformants were screened for mucoid phenotype on agar medium with 10% sucrose (similarly as in Fig. 1). Plasmid DNA from selected mutants was isolated and *lsc3* gene was sequenced to identify the mutations. Most of isolated mutants carried multiple amino acid-changing mutations in the levansucrase gene. Based on position conservation analysis, some of the mutations (Asp300Asn and Thr302Pro) were chosen to construct respective His-tagged single mutants site-directedly.[8]

Site-directed mutagenesis of *lsc3* gene, production and purification of respective N-terminally His-tagged mutant Lsc3 proteins was carried out using mutation-introducing oligonucleotides and the pURI3 vector.[8] We substituted histidine (H) at position 321 with arginine (R), lysine (K), leucine (L) and serine (S). His321 of Lsc3 is equivalent to His296 of *Z. mobilis* LevU and Arg360 of *B. subtilis* SacB. These residues are crucial for transfructosylation reaction by these levansucrases.[2,9,43]

6.2 Rapid microplate-based methods for the screening of levansucrase mutants: using permeabilized levansucrase-expressing bacteria as catalysts

Permeabilized microbial cells can be used for the study of intracellular enzyme activities.[44] Earlier we have used microtitre plate-based semi-quantitative enzyme assay in isolation and characterization of yeast mutants. In this procedure, 0.1% CTAB (cetyltrimethylammonium bromide) was used as a permeabilizing agent.[45,46] We have shown that 0.1% CTAB can also be used for permeabilization of *E. coli* cells and that it has

no inhibitory effect on levansucrase protein. Having this information, we introduced a microtitre plate-based levansucrase assay on CTAB-permeabilized levansucrase-expressing *E. coli*.[47]

This type of assay is simple, rapid and can be robotized. All stages of this procedure: precultivation of transformed *E. coli*, induction of the promoter to switch on protein expression, cell permeabilization and evaluation of catalytic activity can be performed on a microplate. In case of efficient expression system, sufficient amount of protein (levansucrase) will be present in 200 µl of bacterial culture in a microplate well to enable enzyme activity analysis. Regarding our pURI3-based expression system, we have used a following protocol to evaluate total (*i.e.* sucrose splitting) levansucrase activity of Lsc3 variants (our unpublished data; see the footnote[†]).

We have shown that permeabilized cells of levansucrase-expressing *E. coli* can also be used for evaluation of FOS- and levan-producing ability of a levansucrase.[47] According to TLC analysis, reaction products synthesized from sucrose by either cell extract of Lsc3-expressing *E. coli* or purified Lsc3 protein were highly similar to those produced by Lsc3-expressing permeabilized *E. coli*. In all cases, FOS with DP up to seven and also nonmigrating products (high-DP FOS and levan) were revealed.[47] Our current protocol for this procedure is presented in the footnote[‡].

Levansucrases of *Pst*, if expressed in *E. coli*, are not secreted from the cell, but remain mostly in the cytoplasm. In case of Lsc3 protein, 88% of the levansucrase activity was located in the cytoplasm and 12% in the periplasm of the host bacterium.[3] Permeabilization by CTAB overcomes the permeability barrier of the cytoplasmic membrane as well as cell envelope's outer membrane. Therefore we could detect a considerable amount of Lsc3 released from intact cells after their treatment with CTAB (our unpublished data).

Some catabolism of sucrose can still take place in the periplasm of untreated (intact) cells because sucrose will diffuse into the periplasm

[†]Assay of total levansucrase activity on permeabilized cells of recombinant *E. coli*. Grow colonies of levansucrase-producing *E. coli* overnight in LB broth on a 96-well microplate. Combine 50 µl of microtitre plate-grown *E. coli* culture and 50 µl of 0.2% CTAB (cetyltrimethylammonium bromide) in McIllvaine's buffer (pH 6.0) in a new microtitre plate well and agitate for 10 minutes at room temperature to permeabilize the cells. Then, add 50 µl of 0.3 M sucrose (final concentration 100 mM) in McIllvaine's buffer (pH 6.0) to start the levansucrase reaction and agitate further at 37 °C. To stop the reaction at suitable time point, *e.g.* in 5 min, withdraw 10 µl of the reaction mixture, combine it in a new well with 30 µl of Tris buffer (200 mM, pH 8.3), heat the plate at 96 °C for 5 min and cool on ice. To visualize glucose released from sucrose, add 160 µl of Glucose Liquicolor reactive (Human GmbH, Germany) to the well and incubate the plate for 5 min at 37 °C. Measure the absorbance of red color at 500 nm (we use Tecan Sunrise™ microplate reader; Tecan Group Ltd., Switzerland, and respective software). If you analyse mutant proteins, use the absorbance value obtained with the culture expressing the wild-type protein for reference. The *E. coli* culture should be diluted in McIllvaine's buffer prior to the assay if catalytic acitvity is above the detection range of the microplate reader.

[‡]Assay of polymerization products of the levansucrases on permeabilized cells of recombinant *E. coli*. Grow colonies of *E. coli* expressing a levansucrase protein overnight in LB broth on a microtitre plate. Combine 40 µl of bacterial culture and 40 µl of 0.2% CTAB in McIllvaine's buffer (pH 6.0) in wells of another microplate and agitate 10 min at room temperature to permeabilize the cells. Then, add 120 µl of 2 M sucrose (final concentration 1.2 M) in McIllvaine's buffer (pH 6.0) to the wells, seal the plate firmly to avoid evaporation and incubate at 37 °C for 20 hours. Heat the plate for 5 min at 96 °C to stop the reaction and apply 0.5 µl of diluted sample (we usually dilute 4 times in distilled water) to a TLC plate and proceed as indicated in[8]. Cultures of transformants should be suitably diluted in McIllvaine's buffer if catalytic acitvity is very high.

through aspecific porins of the outer membrane and can thereby be metabolized by periplasmic fraction of the levansucrase. Accordingly, intact recombinant *E. coli* cells exhibited ~12.5% of total levansucrase activity measured from CTAB-permeabilized cells (our unpublished data). This result agrees well with relative proportion of the levansucrase in the periplasm. Due to periplasmic levansucrase activity, Lsc3-expressing *E. coli* acquires the ability to grow on sucrose minimal medium.[3]

Above-mentioned methods can be used for high-throughput and sensitive screening of large pools of mutants of various sugar-acting enzymes and can also be adjusted for the assay of other enzyme types.

6.3 Characterization of first mutants of the *Pst* Lsc3 protein: His321, Asp300 and Thr302 as residues implicated in polymerization reaction

For the first time, levansucrases of *Pseudomonas* bacteria have been addressed from the structure-function aspect.[8] His321 of Lsc3, an equivalent of Arg360 of *B. subtilis* SacB and His296 of *Z. mobilis* LevU, was predicted as polymerization determinant of Lsc3. The position of this residue is indicated on a 3D model of the Lsc2 protein in Fig. 2. To verify significance of His321, it was site-directedly mutated to Arg, Lys, Leu and Ser. This substitution had expected consequences: K_m values for sucrose splitting reaction increased up to 30 times (the affinity was reduced) and catalytic efficiency decreased by more than 200 times. Highest negative effect on catalysis was determined for H321S mutant. Expectedly, TA of His321 mutants was significantly reduced – from 74% (His-tagged wild-type Lsc3) to 23% (H321L mutant). So, His321 is certainly of catalytic importance for Lsc3 protein.[8]

The mutants D300N and T302P had two to three times reduced affinity for sucrose splitting and k_{cat} values were reduced by one third, compared to wild-type Lsc3. TA of D300N mutant was rather similar to the wild-type enzyme, whereas that of T302P was reduced to 47%. Reactions conducted as in [8] were analysed for the amount of levan and total FOS. Levan production can easily be monitored in a simple turbidity assay on microtitre plates. Our study indicated that levan production by His321 substitution mutants was strongly retarded. When wild-type His-tagged Lsc3 produced 7.2 mg/ml levan, the T302P mutant synthesized less levan (6.2 mg/ml), whereas D300N produced even more polyfructan (8.6 mg/ml) than the wild-type. Total amounts of FOS produced by His321 substitution mutants ranged from 45.3 mg/ml (H321L) to 61.0 mg/ml (H321R). Respective value for T302P mutant was 42.5 mg/ml and for D300N mutant 80.5 mg/ml. The wild-type Lsc3 protein produced 107.9 mg/ml of total FOS under these conditions. Pattern of FOS produced by the mutants revealed FOS with mobility of kestose (DP 3) as major FOS species for H321R mutant while in case of T302P mutant a product with mobility of nystose (DP 4) was also detected. Intriguingly, the FOS pattern of D300N was very similar to that of LscA (see data on LscA in Fig. 3) showing presence of longer-chain FOS species that are not visible among polymerization products of wild-type Lsc3.[8]

According to sequence alignment of levansucrases, Asp300 and Thr302 of Lsc3 belong to highly conserved region D(E/Q)(T/I/V)ER of levansucrases. Position of Asp300 is almost invariant in levansucrases – only the *Clostridium acetobutylicum* protein has Gln (Q) at that position.[16] Unexpectedly, D300N

mutant did not differ much from the wild-type Lsc3 in our assay. Thr302 is located next to predicted acid-base catalyst Glu303 of Lsc3 and is invariant in levansucrases of Gram-negative bacteria. Levansucrases of Gram-positive bacteria have mostly valine (V) or isoleucine (I) at that position.[16] The mutation of a residue equivalent to Thr302 in Lsc3 has not been described for any of the levansucrase proteins. Our results indicate that replacement of Thr302 with Pro significantly decreases transfructosylation, especially synthesis of longer-chain FOS.

The active centre of levansucrases forms a funnel-shaped pocket in the middle of the β-propeller.[13,14,16] Fructose residue of sucrose molecule binds at the bottom of the pocket and glucose residue on top of it.[48] In *B. subtilis* levansucrase, Glu340 and Glu342 of the DEIER motif as well as Arg360 belong to +1 subsite of the pocket. This subsite binds glucose residue of the donor sucrose molecule. Arg360 and Glu340 form tight hydrogen bonds with hydroxyls of the glucose residue fixing it in a proper orientation needed for further reactions.[2,14] We hypothesize that in Lsc3 protein, His321 and Thr302 belong to +1 subsite of the substrate binding cavity. If these residues are mutated, sucrose binding is hindered that is reflected in increased K_m and reduced catalytic constant values of sucrose splitting reaction. The +1 subsite is also involved in acceptor binding. Therefore His321 and Thr302 substitution mutants of Lsc3 exhibit changed pattern of polymerization products: synthesis of levan and long-chain FOS was reduced whereas short-chain FOS were still produced.

7 Concluding remarks and future perspectives

This report highlights our recent results on levansucrases of *Pseudomonas* bacteria. Before we started our research, allelic spectrum of levansucrase genes had been studied in many *P. syringae* pathovars and respective genes of *P. syringae* pv. glycinea and phaseolicola were cloned and expressed in *E. coli*.[49,50] A levansucrase protein of *P. syringae* pv. phaseolicola was purified and biochemically characterized by Hettwer and coworkers.[12]

We have been focusing on enzymology and product spectrum of levansucrases encoded in the genome of *P. syringae* pv. tomato (*Pst*) with LscA from *P. chlororaphis* subsp. *aurantiaca* (*Pca*) studied as a reference. We conclude that levansucrases of *Pseudomonas* bacteria are from many aspects very similar to those of other bacteria – they synthesize not only levan, but also short-chain FOS and predictably have β-propeller fold that is typical for levansucrases. Using a novel rapid and high-throughput chip-based mass spectrometry method, we showed that oligosaccharidic reaction products of levansucrases can be easily analysed without prior derivation and even without purification. Importantly, the screening indicated that transfructosylation of nonconventional acceptors is most probably a common attribute of levansucrases though beforehand demonstrated mostly for the enzymes from bacilli. Our further studies will be focused on the structure-function relationships of levansucrases of *Pst* with first results concerning catalytically relevant positions of Lsc3 protein denoted already in this report. Having in hands a timesaving toolbox for cloning, protein expression, enzymatic assay and reaction product analysis, we hope to get

valuable information on *Pst* levansucrases that can be applied in biotechnology but should also contribute to elucidation of their role for the bacteria.

Acknowledgements

This research was funded by Estonian Science Foundation grants 7528, 9072 and SF0180088s08 to T.A., and Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project PN-II-ID-PCE-2011-3-0047 to A.D.Z.

References

1 B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard and B. Henrissat, *Nucleic Acids Res*., 2009, **37**, D233.
2 W. Lammens, K. Le Roy, L. Schroeven, A. Van Laere, A. Rabijns and W. Van den Ende, *J. Exp. Bot*., 2009, **60**, 727.
3 T. Visnapuu, A. Mäe and T. Alamäe, *Process Biochem*., 2008, **43**, 414.
4 L. Liiv, P. Pärn and T Alamäe, *Gene*, 2001, **265**, 77.
5 T. Visnapuu, A. D. Zamfir, C. Mosoarca, M. D. Stanescu and T. Alamäe, *Rapid Commun. Mass Spectrom*., 2009, **23**, 1337.
6 B. de las Rivas, J. A. Curiel, J. M. Mancheño and R. Muñoz, *Biotechnol. Prog*., 2007, **23**, 680.
7 J. A. Curiel, B. de las Rivas, J. M. Mancheño and R. Muñoz, *Protein Expr. Purif*., 2011, **76**, 44.
8 T. Visnapuu, K. Mardo, C. Mosoarca, A. D. Zamfir, A. Vigants and T. Alamäe, *J. Biotechnol*., 2011, **155**, 338.
9 H. Yanase, M. Maeda, E. Hagiwara, H. Yagi, K. Taniguchi and K. Okamato, *J. Biochem*., 2002, **132**, 565.
10 S. A. F. T. van Hijum, E. Szalowska, M. J. E. C. van der Maarel and L. Dijkhuizen, *Microbiology*, 2004, **150**, 621.
11 J. Seibel, R. Moraru, S. Götze, K. Buchholz, S. Na'amnieh, A. Pawlowski and H. J. Hecht, *Carbohydr. Res*., 2006, **41**, 2335.
12 U. Hettwer, M. Gross and K. Rudolph, *J. Bacteriol*., 1995, **177**, 2834.
13 G. Meng and K. Fütterer, *Nat. Struct. Biol*., 2003, **10**, 935.
14 G. Meng and K. Fütterer, *BMC Struct. Biol*., 2008, **88**, 16.
15 C. P. Strube, A. Homann, M. Gamer, D. Jahn, J. Seibel and D. W. Heinz, *J. Biol. Chem*., 2011, **286**, 17593.
16 C. Martínez-Fleites, M. Ortíz-Lombardía, T. Pons, N. Tarbouriech, E. J. Taylor, L. Hernández and G. J. Davies, *Biochem. J*., 2005, **390**, 19.
17 U. Pieper, N. Eswar, B. M. Webb, D. Eramian, L. Kelly, D. T. Barkan, H. Carter, P. Mankoo, R. Karchin, M. A. Marti-Renom, F. P. Davis and A. Sali, *Nucleic Acids Res*., 2009, **37**, D347.
18 K. H. Jang, E. K. Jang, S. H. Kim, I. H. Kim, S. A. Kang, I. Koh, Y. I. Park, Y. J. Kim, S. D. Ha and H. K. Chul, *J. Microbiol. Biotechnol*., 2006, **16**, 102.
19 Y. Tambara, J. V. Hormaza, C. Perez, A. Leon, J. Arrieta and L. Hernandez, *Biotechnol. Lett*., 1999, **21**, 117.
20 J. Peter-Katalinić, *Methods Enzymol*., 2005, **405**, 139.
21 A. D. Zamfir, L. Bindila, N. Lion, M. Allen, H. H. Girault and J. Peter-Katalinić, *Electrophoresis*, 2005, **26**, 3650.
22 J. S. Rossier, N. Youhnovski, N. Lion, E. Damoc, S. Becker, F. Reymond, H. H. Girault and M. Przybylski, *Angew. Chem. Int. Ed. Engl*., 2003, **3**, 54.

23 A. D. Zamfir, S. Vakhrushev, A. Sterling, H. J. Niebel, M. Allen and J. Peter-Katalinić, *Anal. Chem.*, 2004, **76**, 2046.
24 A. D. Zamfir, C. Flangea, D. G. Seidler, E. Sisu and J. Peter-Katalinić, *Electrophoresis*, 2011, **32**, 1639.
25 A. D. Zamfir, C. Flangea, A. Serb, E. Sisu, L. Zagrean, A. Rizzi and D. G. Seidler, *Methods Mol. Biol.*, 2012, **836**, 145.
26 M. Froesch, L. Bindila, G. Baykut, M. Allen, J. Peter-Katalinić and A. D. Zamfir, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 3084.
27 R. Almeida, C. Mosoarca, M. Chirita, V. Udrescu, N. Dinca, Ž. Vukelić, M. Allen and A. D. Zamfir, *Anal. Biochem.*, 2008, **378**, 43.
28 S. Kim, R. P. Rodgers, G. T. Blakney, C. L. Hendrickson and A. G. Marshall, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 263.
29 C. Flangea, A. Serb, E. Sisu and A. D. Zamfir, *Biochim. Biophys. Acta*, 2011, **1811**, 513.
30 C. Flangea, E. Sisu, D. G. Seidler and A. D. Zamfir, *Anal. Biochem.*, 2012, **420**, 155.
31 A. Ruiz-Aracama, A. Peijnenburg, J. Kleinjans, D. Jennen, J. van Delft, C. Hellfrisch and A. Lommen, *BMC Genomics*, 2011, **12**, 251.
32 Ž. Vukelić, S. Kalanj Bognar, M. Froesch, L. Bîndilă, B. Radić, M. Allen, J. Peter Katalinić and A. D. Zamfir, *Glycobiology*, 2007, **17**, 504.
33 C. Schiopu, A. Serb, F. Capitan, C. Flangea, E. Sisu, Z. Vukelic, M. Przybylski and A. D. Zamfir, *Anal. Bioanal. Chem.*, 2009, **395**, 2465.
34 A. Serb, C. Schiopu, C. Flangea, E. Sisu and A. D. Zamfir, *J. Mass Spectrom.*, 2009, **44**, 1434.
35 E. Sisu, C. Flangea, A. Serb and A. D. Zamfir, *Amino Acids*, 2011, **41**, 235.
36 A. D. Zamfir, C. Flangea, A. Serb, E. Sisu, N. Dinca, P. Bruckner and D. G. Seidler, *Proteomics*, 2009, **9**, 3435.
37 M. G. Kim, C. H. Kim, J. S. Lee, K. B. Song and S. K. Rhee., *Enz. Microb. Technol.*, 2000, **27**, 646.
38 R. Beine, R. Moraru, M. Nimtz, S. Na'amnieh, A. Pawlowski, K. Buchholz and J. Seibel, *J. Biotechnol.*, 2008, **138**, 33.
39 M. Tieking, M. A. Ehrmann, R. F. Vogel and M. G. Gänzle, *Appl. Microbiol. Biotechnol.*, 2005, **66**, 655.
40 J. Seibel, R. Moraru and S. Götze, *Tetrahedron*, 2005, **61**, 7081.
41 W. C. Han, S. H. Byun, J. C. Lee, M. H. Kim, S. A. Kang, C. H. Kim, E. Wha Son and K. H. Jang, *J. Biotechnol.*, 2007, **131**, S113.
42 W. C. Han, S. H. Byun, M. H. Kim, E. H. Sohn, J. D. Lim, B. H. Um, C. H. Kim, S. A. Kang and K. H. Jang, *J. Microbiol. Biotechnol.*, 2009, **19**, 1153.
43 S. Y. Li, M. Chen, G. Li, Y. L. Yan, H. Y. Yu, Y. H. Zhan, Z. X. Peng, J. Wang and M. Lin, *Acta Biochim. Pol.*, 2008, **55**, 201.
44 T. Alamäe and A. Järviste, *J. Microbiol. Methods*, 1995, **22**, 193.
45 T. Alamäe and L. Liiv., *Folia Microbiol.*, 1998, **43**, 443.
46 T. Kramarenko, H. Karp, A. Järviste and T. Alamäe., *Folia Microbiol.*, 2000, **45**, 521.
47 K. Mardo, T. Visnapuu and T. Alamäe, in *Proceedings of the 6th International Conference on Polysaccharides-Glycoscience*, eds. R. Rapkova, J. Copikova, E. Šarka, Praque, Czech Republic, 2010, pp. 36–39.
48 L. K. Ozimek, S. Kralj, M. J. E. C. van der Maarel and L. Dijkhuizen, *Microbiology*, 2006, **152**, 1187.
49 U. Hettwer, F. R. Jaeckel, J. Boch, M. Meyer, K. Rudolph and M. S. Ullrich, *Appl. Env. Microbiol.*, 1998, **64**, 3180.
50 H. Li and M. S. Ullrich. *J. Bacteriol.*, 2001, **183**, 3282.

Recent advances on the application of NMR methods to study the conformation and recognition properties of carbohydrates

Ana Ardá, M. Álvaro Berbís, Pilar Blasco, Angeles Canales, F. Javier Cañada, M[a] Carmen Fernández-Alonso, Filipa Marcelo and Jesús Jiménez-Barbero*
DOI: 10.1039/9781849734769-00192

1 Introduction

This report gathers selected and recent examples of studies of the conformation, structure, dynamics and binding features of saccharides and analogues. It has not been our intention to be exhaustive, since many groups are working in this field either on a regular basis or making frequent contributions. It has been our aim to present different examples of the application of NMR methods and protocols for analysing saccharide conformation.

We will start by presenting new advances from the NMR methodological viewpoint to later describe some key examples of applications of NMR for sugar conformation determination, by discriminating between the free and bound state and also from natural and synthetic sugars (or mimetics thereof).

2 The access to new NMR parameters and methodological developments

The detailed analysis of chemical shifts and coupling constants combined with the use of ^{13}C labeled compounds give rise to important insights in the understanding of carbohydrate conformation. In this sense, it should be mentioned that the amide *cis*/*trans* isomerization in methyl 2-acetamido-2-deoxy-D-glucopyranoside (MeGlcNAc) has been recently quantified by measuring ^{1}H and ^{13}C chemical shifts and ^{1}H-^{1}H and ^{13}C-^{13}C spin-coupling constants [1]. Single ^{13}C-labeled methyl 2-deoxy-2-formamido-D-glucopyranoside anomers were first characterized. Then, double ^{13}C-labeled methyl 2-acetamido-2-deoxy-D-glucopyranoside anomers were investigated, leading to the detection and quantification of *cis* and *trans* amides in this key amino sugar. These studies show that aqueous solutions of GlcNAc-containing oligosaccharides present measurable amounts of both *cis* and *trans* amides, which may influence their biological properties.

The accurate measure of coupling constants also allows developing Karplus equations for conformational analysis of carbohydrates. In this regard, eight new Karplus relationships have been recently determined for use in conformational studies of saccharide N-acetyl side-chains in solution by NMR spectroscopy [2]. Eight NMR vicinal spin-spin coupling constants have been investigated, six that report on the C2-N2 bond ($^{3}J_{\mathbf{H2,NH}}$, $^{3}J_{\mathbf{H2,CO}}$, $^{3}J_{\mathbf{C1,NH}}$, $^{3}J_{\mathbf{C3,NH}}$, $^{3}J_{\mathbf{C1,CO}}$, $^{3}J_{\mathbf{C3,CO}}$) and two focusing on the amide bond ($^{3}J_{\mathbf{NH,CH3}}$, $^{3}J_{\mathbf{C2,CH3}}$) of this side chain. Moreover, from the methodological

Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040, Madrid, Spain. E-mail: jjbarbero@cib.csic.es

© The Royal Society of Chemistry 2012

viewpoint, it has been demonstrated that $^{3}J_{HH}$ strong coupling may strongly affect the reliable measurement of $^{1}J_{CH}$ values. The authors concluded that that spectral simulation yields accurate $^{1}J_{CH}$ with errors as low as 1%, even in the presence of strong coupling [3].

On the other hand, theoretical density functional theory (DFT) methods have been recently used to analyze the structure and the NMR spin-spin coupling constants of a heparin disaccharide [4]. In this work, the effects of solvation and counterions on both ^{1}H-^{13}C and ^{1}H-^{1}H coupling constants have been investigated. The optimized disaccharide geometry showed that the change of counterion (Ca^{2+} instead of Na^{+}) influenced the geometry of pyranose rings and the glycosidic linkage conformation. The DFT-computed three bond proton-proton spin-spin coupling constants agreed well with published experimental data and indicated that the population of the $^{1}C_{4}$ chair form of the 2-*O*-sulfated iduronic acid residue increased in the presence of Ca^{2+} ions compared to the presence of Na^{+} ions. This analysis also showed that the Fermi contact term was not always dominant and that paramagnetic and diamagnetic contributions considerably influenced the magnitudes of the proton-proton spin-spin coupling constants.

It is also worth mentioning that despite the important information that can be obtained from coupling constants and NOE analysis, there is a need of considering additional parameters to completely characterize complex carbohydrates. This is specially required in those cases where the number of long range NOEs that can be detected is very low. In these systems, Residual Dipolar Couplings (RDCs) may provide important structural information. In a recent work, RDCs have been used in combination with proton-proton cross-relaxation rates, trans-glycosidic coupling constants and molecular dynamics simulations to study the complex conformational equilibria of a human milk pentasaccharide, LNF-1[5]. This work also included generalized order parameters, obtained from nuclear spin relaxation experiments. The analysis successfully showed the conformational space which is accessible to LNF-1.

Obviously, carbohydrates are rather flexible molecules. Therefore, NMR observables do not always correlate with a single conformer but with an ensemble of low free energy conformers that can be accessed by thermal fluctuations. In this regard, a novel procedure to identify and weight the contribution of the different conformers based on the comparison of the experimental residual dipolar couplings with those estimated for a family of conformers has been recently described [6]. In this work, a program based on a genetic algorithm was developed that generates the best set of statistical weights for the conformers derived from the conformational searches. This algorithm was applied to derive the subensemble of conformers that appears to be most important in the case of six different complex human milk sugars.

RDC measurements require a partial alignment of the NMR samples, since they average to zero under isotropic motion conditions. The first described examples employed external alignment media, such as phospholipid bicelles, lamellar phases made of cetylpiridinium chloride/hexanol/brine mixtures or phages in aqueous solution. However, none of these media is universally applicable since they are sensitive to pH, salt, and other

experimental conditions. In addition, the media could interact with the target molecule, depending on its chemical nature. In this sense, recent studies have suggested to employ paramagnetic tags directly attached to the carbohydrate for the alignment, to overcome the problems of the external alignment media [7], following methodologies initially developed for proteins. Diamagnetic proteins were converted in paramagnetic compounds through attachment of an EDTA-based paramagnetic tag and this concept has been extended to study carbohydrates and their complexes with receptor molecules.

Furthermore, paramagnetic metals also cause paramagnetic shifts. These so-called pseudocontact shifts (PCSs) arise from dipolar interactions between the unpaired electrons of the metal ion and the nuclei in the vicinity of this ion. They contain important structural information since they depend on the distance between the metal ion and the nuclei. The first conformational studies of diamagnetic carbohydrates bearing a lanthanide binding tag have been recently described showing the potential of this new methodology for analysing both rigid and flexible carbohydrates [8].

NMR techniques have been also developed to be able to study cell extracts. As example, the neural cell metabolism in HT-22 cells that have been turned into a glycolytic state, has been investigated [9]. Two different metabolic pathways could be followed by NMR detection of UDP-GlcNAc, the principal metabolite, in conjunction with ^{13}C labelling of the precursor metabolites, in this case 1-^{13}C glucose. The application of this methodology to other cell lines and primary cell cultures may open new avenues in this field.

3 Applications. Saccharides in solution

3.1 Oligosaccharide and polysaccharides

Oligosaccharides display flexibility in solution, especially around the glycosidic torsion angles and their lateral chains (responsible for the orientations of the hydroxy groups and of the hydroxymethyl chains) [10]. Depending on the chemical natures of the constituent monosaccharides, equilibria between different ring shapes can also be present [11, 12]. In any case, detailed conformational studies based on molecular modelling and NMR spectroscopy [13, 14] have allowed the deduction that some oligosaccharides display major conformations in solution, with local fluctuations around the torsional degrees of freedom [15]. Computational methods, and in particular molecular dynamics (MD) simulations, are essential complementary tools to NMR data to characterise the structures and dynamics of glycans and glycoconjugates in their free and bound states [16].

Nowadays, one relevant item in biomedical research is the synthesis of effective antibacterial or antivirus vaccines; that is, more immunogenic and less toxic. NMR-based conformational analysis has permitted the study of the interaction epitope of oligo- and polysaccharides with antibodies and to dissect the minimum structural requirement for conferring binding affinity to these antibodies. In particular, the O-antigen (O-Ag) repeat units (RUs) of all *S. flexneri* serotypes contain the tetrasaccharide backbone: 2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*(1→3)-β-D-Glc*p*NAc-(1→,

with glycosylation/s and/or O-acetylation defining the different serotypes. The conformational basis of the serotype specificity has been investigated for oligosaccharides representing three RUs of each serotype [17]. The serotype-specific substitutions of the backbone did not induce any new backbone conformation in almost all O-Ags serotype simulations, although the substitutions locally restrained the accessible conformational space. Overall, a branched α-D-glucopyranose influenced the conformational behaviour with flexibility restrictions of the closest backbone glycosidic linkage on the nonreducing side for all the studied oligosaccharides.

The conformational flexibility and dynamics of two (1→6)-linked disaccharides related to glycosyl transferase GnT-V has been investigated [18]. NMR NOE and T-ROE spectroscopy experiments, coupling constants and molecular dynamics (MD) simulations were used in the analyses. The synthesized compounds were α-D-[6-^{13}C]-Man*p*-OMe derivatives, which facilitated the determination of two- and three-bond ^{1}H,^{1}H, ^{1}H,^{13}C and ^{13}C,^{13}C-coupling constants. The population distributions for the glycosidic ω torsion angle in different sugars were determined by this method. The dynamic models, generated with the PARM22/SU01 CHARMM-based force field were in very good agreement with the experimental observations. For instance, β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe is described by an equilibrium of populated states in which Φ has the exo-anomeric conformation, Ψ shows an extended antiperiplanar geometry, and ω shows a distribution of populations between the gauche–trans and the gauche–gauche states. The use of site-specific ^{13}C labelling leads to increased spectral dispersion and the combination with computational procedures opens a promising field for structural studies of complex carbohydrates.

An investigation of the conformational behavior of methyl β-maltoside, methyl α-cellobioside, and methyl β-cellobioside disaccharides using NMR spectroscopy and molecular dynamics (MD) techniques has been presented [19]. Now, emphasis was placed on the validation of a force field for hexopyranose disaccharides followed by elucidation of the conformational properties of two different types of glycosidic linkages, α-(1→4) and β-(1→4). Careful analysis of the data suggested that the intramolecular interactions in the current force fields may be slightly underestimated, leading to slight undersampling of the *syn* conformation. The proposal is that the observed limitations come from the lack of polarization in the potential energy. Ongoing efforts to develop next generation empirical force fields that include electronic polarizability may allow this assumption to be directly tested. Following this idea, the conformational space available to α-D-Man*p*-(1→2)-α-D-Man*p*-OMe, the model for the external disaccharide in N- or O-linked glycoproteins, has been determined using experimental data and simulations combined with a maximum entropy approach [20]. Reparametrization of homo- and heteronuclear Karplus relationships for glycosidic torsion angles in which the importance of electronegative substituents on the coupling pathway was deemed essential resulting in four derived equations, two $^{3}J_{COCC}$ and two $^{3}J_{COCH}$ being different for the Φ and Ψ torsions, respectively.

A novel method for the overexpression of ^{13}C-labeled oligosaccharides has been presented [21] that makes use of genetically engineered *Saccharomyces cerevisiae* cells. Furthermore, ^{13}C labelling at selected positions of the sugar residues can be achieved by using a site-specific ^{13}C-enriched glucose as metabolic precursor. This facilitates NMR spectral assignments and provides the technical basis for NMR analysis of structures, dynamics, and interaction of larger, branches oligosaccharides.

The synthesis of two novel carbasugar analogues of α-L-iduronic acid has been described in which the ring-oxygen has been replaced by a methylene group [22]. In analogy with the conformational equilibrium described for α-L-IdopA, the conformation of the carbasugars was investigated by NMR experiments. The carbasugar bioisosteres do not exhibit the conformational flexibility reported for α-L-iduronic acid in oligo- or polysaccharides, which is of great importance for its interaction with some protein receptors, such as antithrombin.

The glycosaminoglycan (GAG) chondroitin sulfate is essential in human health and disease but exactly how sulfation dictates its 3D-structure at the atomic level is unclear. Homogenous oligosaccharides of unsulfated chondroitin (with and without ^{15}N-enrichment) have been prepared and analyzed by NMR [23]. An experimental 3D-model of the hexasaccharide has been presented. The unsulfated chondroitin bond geometry differs slightly from hyaluronan by rotation about the β(1→3) dihedral angle, while the β(1→4) linkage is unaffected. Furthermore, the comparison shows that this glycosidic linkage geometry is similar in chondroitin-4-sulfate. It seems that both hexosamine OH-4 and OH-6 atoms are solvent exposed in chondroitin, explaining why this molecule is amenable to sulfation, while hyaluronan is not. Moreover, 4-sulfation has little effect on backbone conformation.

α-L-Rha*p*-(1→2)[α-L-Rha*p*-(1→3)]-α-L-Rha*p*-OMe and its site-specifically ^{13}C-labeled(C2,C2)-isotopologue have been studied [24]. Eight conformationally dependent coupling constants were determined, and used together with previously determined trans-glycosidic $^{3}J_{C,H}$, and seven interresidue ^{1}H,^{1}H-T-ROEs, for further conformational analysis of the trisaccharide. This study highlights that different approaches are available in order to obtain all the information needed for the molecular system under investigation. The ABO histo-blood group antigens are best known for their important roles in solid organ and bone marrow transplantation as well as transfusion medicine. The synthesis of the ABO type III and IV antigens with a oct-7-en-1-yl aglycone has been reported [25], along with the NMR conformational study of the ABO type I to VI antigens. These NMR investigations showed very little difference in the ^{1}H chemical shifts, as well as ^{1}H–^{1}H coupling constants, across all compounds, suggesting that these ABO subtypes adopt nearly identical conformations in solution. One exception is the chemical shift for H5 of the fucose residue of the A type V tetrasaccharide, which resonated around 4.65 ppm, a deviation of 0.33 ppm from the 4.32 ppm average. The significance of this observation is currently unclear.

Combination of ^{1}H, ^{13}C and ^{39}K NMR spectroscopy with single-crystal X-ray crystallography have allowed to study the complex formation of

1,6-anhydro-β-maltose and potassium ions [26]. In solution, diffusion coefficients and ^{39}K-T1 data unambiguously indicated complex formation, while ^{13}C NMR studies indicated the importance of the 1,6-anhydro moiety for the complex formation, also supported in the X-ray crystal structure.

The description of the re-orientational dynamics of flexible molecules is a challenging task, in particular when the rates of internal and global motions are comparable. The commonly used simple mode-decoupling models are based on the assumption of statistical independence between these motions. This assumption is not valid when the time scale separation between their rates is small, a situation that was found to arise in oligosaccharides in the context of certain internal motions. To make possible the interpretation of NMR spin relaxation data from such molecules, a comprehensive approach generally applicable to flexible rotators with one internal degree of freedom has been developed [27]. The method has been successfully tested on β-D-Glcp-(1→6)-α-D-[6-^{13}C]-Man*p*-OMe dissolved in DMSO-d_6/D_2O.

Passing to polysaccharides, the capsular polysaccharide of *Escherichia coli* K5 has been hypothesised to promote virulence through its molecular mimicry of host heparan sulphate. To test this hypothesis, pure oligosaccharides from K5 have been produced and their conformational properties investigated with ultra-high-field NMR (900 MHz) [28]. All carbohydrate rings adopt a $^4C_1$ conformation, the amide side chains have a trans orientation and the hydroxymethyl group is freely exposed to bulk solvent. Initial models of the glycosidic linkage conformation based upon simple interpretation of NOE cross-peaks suggests that the β(1→4) linkage adopts a 3D geometry of $\Phi \approx 60^\circ$, $\Psi \approx 0^\circ$ and the α(1→4) linkage prefers $\Phi \approx -30^\circ$, $\Psi \approx -30^\circ$. In this conformation, the overall molecular geometries of K5 polysaccharide, heparan sulphate and even fully-sulphated heparin are remarkably similar. These results substantiate the hypothesis that the K5 capsular polysaccharide confers virulence to *E. coli* K5 by being a 3D molecular mimetic of host heparan sulphate, helping it to evade detection by the mammalian immune system.

The conformational features of hyaluronic acid [4)-D-β-GlcA-(1→3)-D-β-GlcNAc-(1→] have also been investigated using NMR and molecular modelling [29]. A decasaccharide fragment of sodium hyaluronate (HA) was subjected to 3.5 ns of molecular dynamics in explicit water and the approach consisted in the measurement of NMR residual dipolar coupling (RDC), which were used to filter the molecular dynamics data, by retaining those structures which were in agreement with the experimental observations. As a result, this protocol led to the conclusion that hyaluronic acid can adopt two different arrangements, which can be described by three- and four-folded left-handed helix, with the first one as the major conformer. These two different conformers coexist in solution and convert into each other. This information evidences how subtle variations in the dihedral angles are accompanied by a switch from one geometry to another (from a three- to a four-folded helix).

The use of lactic acid bacteria in milk fermentation results in favorable physical and rheological properties due to exopolysaccharide (EPS) production [30]. The hexasaccharide repeating unit EPS from *S. thermophilus* ST1 has been elucidated and presents the following

structure: 3)[α-D-Glc*p*-(1→4)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→4)[β-D-Gal*f*-(1→6)]-β-D-Glc*p*-(1→6)-β-D-Glc*p*-(1→. Thus, the EPS consists of a backbone of four sugar residues with two terminal ones making up two side-chains of the repeating unit. The use of efficient NMR techniques based on parallel acquisition and TILT experiments was very valuable for the structural analysis.

NMR has been employed to study transient hydrogen bonding in a polysaccharide dissolved in water without any cosolvent at room temperature. In this case, the pentasaccharide repeating unit of the *Escherichia coli* O142 lipopolysaccharide has been used as model system [31]. A 105 ns MD predicted transient inter-residue hydrogen bonds from the GalNAc NH groups. To experimentally investigate these predictions, the polysaccharide was uniformly ^{13}C,^{15}N-enriched and the NH, carbonyl, C2, C4, and methyl resonances of the GalNAc and GlcNAc residues assigned. Temperature dependence of the amide NH chemical shifts and one-bond NH J couplings supported that NH groups on two of the GalNAc residues are donors in transient hydrogen bonds. The remaining GalNAc and GlcNAc NHs do not appear to be donors These results raise the biologically important possibility that carbohydrates may utilize hydrogen bonding to assemble three-dimensional scaffolds that would form the basis of biomolecular carbohydrate recognition and function.

Strong efforts have been developed for implementing automatic structural determination of polysaccharides [32], by using the computerized CASPER approach, based on ^{1}H and ^{13}C chemical shifts.

3.2 Glycopeptides

Glycosylation is one of the most widespread post-translational modifications of proteins [33]. Thus, the conformational study of glycopeptides and peptidoglycans continues to be a hot topic of research. The most prevalent type of O-glycosidic linkage involves an α-*N*-acetylgalactosamine (GalNAc) unit and the β-hydroxyl group of either Ser or Thr at the peptide chain [34]. This glycosylation pattern is commonly present in antigenic mucins [35, 36]. In the last few years, the combination of experimental NMR data (chemical shifts, coupling constants and NOEs) assisted by molecular modelling protocols has strongly contributed to disclose the effect of α-*O*-glycosylation on the conformation of mucin-type peptides [37]. The extensive work of Live and co-workers revealed that α-*O*-linkages induce the peptide backbone to adopt an extended conformation [38]. More recently, the work of Corzana *et al*. has demonstrated that the nature and sequence of the amino acid aglycone also modulates the conformation and dynamics effects of Tn antigen's clusters [39, 40]. In addition, Nishimura's group has combined MS and NMR techniques to study the conformational effects of *O*-glycosylation on the peptide backbone structures during enzymatic mucin domain assembly *in vitro* by using an isoform of UDP-GalNAc:*N*-acetyl galactosaminyl transferase-T [41]. Structural analysis, combining NOESY data with structural refinements of the synthetic key intermediates uncovered that the preferential installation of α-GalNAc at Thr-10 of MUC4 is a crucial step for forcing the stability of the peptide backbone structure at this

area. The authors proposed that this event may provide a preferred substrate conformation for the subsequent multiple glycosylation.

Additional investigations have focused on the possibility that the biological activity of the N/OFQ peptide, a native ligand of the pain-related and viable drug target NOP receptor could be modulated by glycosylation. The relationship with conformational effects was also addressed [42]. Hence, the pharmacological analysis of three novel glycopeptides [Thr^5-*O*-α-D-GalNAc-N/OFQ] (**1**), [Ser^{10}-*O*-α-D-GalNAc]-N/OFQ (**2**) and [Ser^{10}-*O*-β-D-GlcNAc]-N/OFQ] (**3**) revealed that glycopeptide **2** exhibited an improved binding affinity, whereas **1** showed a remarkably reduced binding affinity compared to the parent compound and to **3**. NMR data complemented with modelling allowed identifying a certain degree of helical structure for the parent N/OFQ peptide, as well as for **2** and **3** in micellar solutions. In contrast, for **1**, a distinct set of remote NOE contacts were detected, suggesting that this glycopeptide exhibited a significant population of folded hairpin-like structures. Thus, the authors proposed that linear helical structures of nociceptin analogues are more complementary and interact more efficiently with the NOP receptor than folded structures.

The stable incorporation of ^{13}C and ^{15}N into glycoconjugates has allowed to employ NMR methods to characterize the dynamics and interactions of glycoprotein glycans, as recently reviewed [43]. The linewidth of the ^{13}C anomeric peak is inversely proportional to the transverse relaxation rate, T_2, and permits the analysis of the mobility of the IgG-Fc glycans. In particular, the data obtained for the Fc glycoprotein showed that two galactose peaks exhibited extremely different line widths, indicating that these sugar residues have significantly different mobilities in Fc glycans [44]. In addition, the solvent accessibility/hydrogen bonding interactions of the GlcNAc side chains in Fc was monitored by amide exchange experiments. Slow rates for amide proton exchange are associated with shielding from solvent, usually through hydrogen bonding interactions. After selective isotope labelling of the N-glycans on the Fc fragment of IgG, and combining relaxation dispersion-type experiments with temperature-dependent chemical shift perturbation measurements, it was demonstrated that both [α-(1→3)- and α-(1→6)Man] glycan branches of the antibody are accessible to receptors and rather mobile in solution [45].

Conformational studies on unusual phosphorylated *O*-mannosyl glycopeptides from α-dystroglycan have also been presented [46]. It was found that a glycopeptide having a 6-phospho-*O*-mannosyl residue is not an acceptor for POMGnT1, which attaches β(1→2)-GlcNAc to *O*-mannosyl moieties. In addition, the conformational analysis by NMR showed that the *O*-mannosyl modification did not exert a major conformational effect on the peptide backbone. It was then proposed that these residues, introduced at the early stages of glycoprotein glycosylation, have an ability to regulate the loci of subsequent *O*-GlcNAc additions through conformational effects.

One additional example has focused on why PrkC Thr-kinase induces germination of bacterial spores in response to peptidoglycans of the DAP-type but not to those of the Lys-type cell wall. Thus, the structural requirements necessary for recognition and binding has been addressed using STD-NMR and TR-NOESY, together with mutagenesis experiments [47].

In fact, the extracellular region of PrkC senses the peptidoglycan through the DAP-moiety-mediated interactions with an arginine residue of the EC-PrkC receptor.

The conformational properties of the Glc3Man unit has been studied to get information on the chaperon-assisted glycoprotein folding pathway [48]. Calnexin recognises the terminal Glcα1→3Manα linkage, formed by trimming of the Glcα1→2Glcα1→3Glcα1→3Manα (Glc3Man) unit in Glc3Man9GlcNAc2. Thus, the conformation of a series of N-glycans; *i.e.* Glc3ManOMe, Glc3Man4,5,7GlcNAc2 and Glc1Man9GlcNAc2 using 2D NMR NOESY, ROESY, T-ROESY and residual dipolar coupling experiments in a range of solvents, along with MD simulations of Glc3ManOMe. A single conformation for the Glcα1→2Glcα and Glcα1→3Glcα linkages was found, along with two conformers for the Glcα1→3Manα linkage. Modelling of the binding of Glc1Man9GlcNAc2 to calnexin suggests that the minor conformer is recognized by calnexin. This is the first evidence presented on glycoprotein folding that suggests that this process may be optimized to balance chaperone-assisted and chaperone-independent pathways.

As mentioned before, stable-isotope-labelling techniques enhance the quality of NMR data. A stable-isotope-assisted NMR analyses using immunoglobulin G as a model system in order to gain information on the dynamics and interactions of glycoconjugates has been reported [49]. A series of glycopeptides cleaved from the metabolically $^{13}C/^{15}N$-labeled IgG glycoprotein can be used as tools for detailed NMR analyses and for characterisation of sugar-binding modes of lectins.

3.3 Glycomimetics

The design and employment of glycomimetics as possible therapeutic agents remains an attractive area of research [50]. NMR studies have contributed to disclose their conformations in the free and bound state and to provide insights on their interactions with receptors [51]. Some glycomimetics have been employed to disrupt the formation of sugar–protein complexes, including glycosidases and glycosyltransferases [52, 53].

As example, the molecular basis for inhibition of GH84 glycoside hydrolases by substituted seven-membered ring iminoalditols (azepanes) has been reported [54]. This work pointed out the importance to design chemically stable inhibitors, but conformationally flexible. NMR data combined with molecular mechanics (MM) and MD simulations revealed that the polyhydroxylated azepanes adopted a distinct type of conformations in solution than those found by X-Ray crystallography at the enzyme active site. These results showed that the enzyme selects the proper conformation of the conformationally flexible azepane into the enzyme active site to optimize the inhibitor-enzyme interactions. As additional example, the binding and substrate specificity of the recognition of β-D–mannosides and β-D-glucosides by barley (HvBII) and rice (Os3BGlu7) β-D–glycosidases has been investigated by NMR with quantum mechanics/molecular mechanics (QM/MM) calculations and docking protocols [55]. STD-NMR and TR-NOESY experiments performed using competitive thioglycoside mimetics (4NP-S-Glc and 4NP-S-Man) revealed a distinct bioactive conformation for 4NP-S-Glc ($B_{3,O}$ and $^{1,4}B$) and 4NP-S-Man ($^4C_1$) when

bound to HvBII, and the same type of $^4C_1$ conformation when bound to Os3BGlu7. In fact, the experimental and computational data suggested that glycoside recognition and substrate specificity of Os3BGlu7 and HvBII rely on a combination of factors involving (i) the inherent conformational flexibilities of the *gluco*- and *manno*-configured substrates at the active sites, (ii) the subtle differences in the spatial disposition of the active site residues and their capacities to form interactions with specific groups of the substrates, and (iii) the small variations in the charge distributions and shapes of the catalytic sites.

Fragment-based methods have been increasingly adopted in the drug-discovery process, also in the carbohydrate field. The identification of the better way of joining the different molecular fragments to constitute the final lead continues to be a challenge [56]. Thus, techniques that permit a dynamic-system-based drug-design for the identification, synthesis, and evaluation of new leads offer a way to solve this problem. Ramström and coworkers have demonstrated that ^{1}H STD-NMR spectroscopy can be employed for efficient and direct *in situ* identification of β-galactosidase inhibitors within a virtual dynamic hemithiocetal system [57]. The best molecular fragments identified from the dynamic approach were then employed to derive a good β -galactosidase inhibitor using a regular synthetic approach [58].

In the lectin field, different mannosyl glycomimetics have been designed to efficiently target the extracellular lectin domain of DC-SIGN by using a multidisciplinary approach [59, 60]. More recently, the same authors have reported the employment of a α-fucosylamide-based mimic of Lewis X [61]. Interestingly, the subsequent insertion of an aromatic ring to replace the Gal moiety allowed improving the binding affinity, as confirmed by STD competition experiments [62].

An additional example includes a simple change in the chemical nature of one glycosidic linkage to perturb the conformational properties of oligosaccharides and even to derive a moderate competitive inhibitor. In particular, the synthesis of the β(1→3) analogue of pentaacetyl chitopentaose, replacing the natural β(1→4) linkage by the β (1→3) alternative has been reported. NMR and modelling studies permitted to verify that the mimetic is recognized by a lectin (wheat germ agglutinin) and one enzyme (a model chitinase) [63]. The existing subtle differences in the conformational behaviour of both molecules allowed the mimetic still to interact with the two proteins. However, it was not hydrolyzed by the enzyme and acted as a moderate inhibitor of chitin hydrolysis.

One key feature in protein-carbohydrate interactions involves multivalency, which notably improves the affinity and specificity of biomolecular interactions [64]. β-*C*-galactopyranosyl mimetics have been incorporated within a multivalent strategies to design a potent bacterial lectin inhibitor [65]. The multivalent *C*-galacto-conjugated architectures were synthetised and evaluated against PA-IL lectin (Lec A) from pathogenic *Pseudomonas aeruginosa*. NMR diffusion experiments allowed estimating the size and topology of the multivalent compounds, indicating the expected decrease of the diffusion coefficients with increasing diameter, as the number of peripheral epitopes was increased.

A new class of GlcNAc mimics for E-selectin antagonists have been designed and synthesised [66]. The mimic consists of a cyclohexane ring substituted with alkyl substituents adjacent to the linking position of the fucose moiety. By STD-NMR experiments it could be shown that the increase in affinity does not result from an additional hydrophobic contact of the alkyl substituent with the target protein E-selectin, but rather from a steric effect stabilizing the antagonist in its bioactive conformation. When the GlcNAc mimetic with two methyl substituents was used the affinity was regained. As predicted by MD simulations, equatorial alkyl groups at the 2-position in the carbocyclic GlcNAc mimetic restrict the conformational flexibility of the core of the antagonist and entropically improve binding affinities.

A carbohydrate peptide mimic of the O-polysaccharide of *Shigella flexneri* Y, MDWNMHAA, when bound to its complementary antibody mAb SYA/J6, has been studied [67]. Its conformation in the bound state has been characterised as well as the presence of immobilised water molecules in the active site. WaterLOGSY experiments were used in conjunction with STD methods to probe the existence of these immobilised water molecules in the complex. MD simulations were then used to specify the likely locations of the water molecules. Of note, those waters involved in providing complementarity between the peptide and mAb SYA/J6 remained throughout the course of the simulation.

4 Applications. The interaction of saccharides with other natural and synthetic molecules

In the last years, a number of studies on the conformational features of the interaction between proteins and saccharides have been reported, providing new details, at atomic resolution, on these key processes. NMR experiments have been employed to analyse the interaction of sugars with different biomolecules, including proteins (lectins, antibodies and enzymes), nucleic acids and even other saccharides. Synthetic receptors, also dubbed artificial lectins, have been employed to model sugar interactions. Finally, a few cases of the use of cyclodextrins as receptors will be mentioned, as examples of sugars being receptors instead of organic molecules' ligands.

In principle, there are different ways to characterise the complexes: On one hand, in favourable cases, it is possible to measure the NMR parameters of the receptor in the free and bound states, *i.e.*, differences in chemical shift and NOE-contacts, before and after the formation of the complex. On the other hand, it is usually more feasible to observe and detect changes in the ligand NMR signals before and after the binding process. In this latter case, the use of STD (saturation transfer difference) and TR-NOE (transferred NOE) experiments is of paramount importance.

4.1 Carbohydrate-protein interactions

A carbohydrate-binding protein can perform conformer selection, which has been demonstrated in different instances by combinations of NMR and modelling [68]. Combined NMR/modelling approaches have allowed this principle to be explored along with determinations of the specificities,

affinities and structural aspects of many receptor–sugar interactions both with natural sugars and with glycomimetics. Depending on the architecture of the protein's binding site, the major conformation of an oligosaccharide existing in solution might be recognised by the receptor, but in other cases, a conformational selection process can take place with exclusive recognition of one of the conformers [69] which might be drastically different from the major one existing in water solution.

The current state-of-the-art NMR methodologies for the study of protein-carbohydrate interactions have been outlined in a recent review [70]. The article described the most prominent NMR techniques utilised nowadays, both from the ligand and receptor perspectives, as well as their combination with molecular modelling procedures. Among the ligand-detected NMR methods, STD and TR-NOESY experiments are regarded as particularly useful for scoping carbohydrate-protein interactions, thanks to the information that they provide, and to the fact that the optimum kinetic equilibrium for such experiments, lying between the mM to low μM range, are satisfied by the usually moderate to weak carbohydrate-protein interaction affinity regimes. The authors ended up by paying attention to the 3D view of protein-carbohydrate complexes as obtained with the help of computational protocols, such as docking and molecular dynamics, and reported on the advent of more sophisticated computational approaches, such as the application of QM/MM and metadynamics for exploring mechanistic aspects of glycosidases and glycosyltransferases.

STD studies are of pre-eminent importance for investigating protein-ligand interactions. However, accurate quantitative analyses of these interactions have been precluded by the complex dependence of the STD signal magnitude on a variety of factors, which introduce unacceptable biases when dissociation constants (K_D) are attempted to be estimated through STD-titration experiments. A recent paper sought to study the relative impact of such dependencies on the STD signal and to develop a straight-forward protocol to directly measure the K_D of protein-ligand complexes by STD titration curve [71]. The authors identified the accumulation of saturation in the free ligand through ligand rebinding events and the longitudinal relaxation as the main sources of error when estimating affinities through STD data. To overcome these drawbacks, they proposed a novel method, in which the binding isotherms are constructed by using the initial growth rates of the STD amplification factors, taking profit that, at zero saturation time, neither rebinding events nor relaxation occurs. The proposed protocol was validated with two well-studied carbohydrate-protein interaction model systems: the binding of GlcNAc and GlcNAcβ(1-4)GlcNAc (chitobiose) to the WGA lectin. In all instances, the affinities estimated with this novel method were in good agreement with previously reported data obtained with calorimetric studies.

NMR spectroscopy has also been used to characterise the binding properties of 2G12, the only carbohydrate-directed monoclonal antibody against the human immunodeficiency virus (HIV), which has specificity for a cluster of mannose-rich glycans on the viral glycoprotein gp120 [72]. In particular, the authors explored the recognition of mannose-coated gold nanoparticles by 2G12, using SPR and STD-NMR. Since both the antibody

and the nanoparticles are large systems, direct evaluation of the K_D by STD was not possible. Instead, affinity data were obtained by monitoring the competition of the nanoparticles against previously characterized 2G12/oligomannoside systems, revealing affinities in the low micromolar range for the 2G12/nanoparticles recognition. Furthermore, the nanoparticles were shown to compete with gp120 for the antibody, effectively blocking the 2G12-mediated cellular immunity to the virus *in vitro*. These studies are meant to contribute to the development of mannose-gold nanoparticles as safe, potent and versatile agents to promote immunization against the HIV by eliciting 2G12-like antibodies.

Using a similar approach, STD studies have contributed to yield affinity and fine specificity data for chimeric IgE antibodies towards Galα(1-3)Gal, considered of outstanding relevance in the context of xenotransplantation and vaccine design, as part of a study devoted to the controversial anaphylaxis mediated by immunogenic carbohydrate determinants [73].

Among carbohydrate-protein interactions, the specific recognition of sugars by galectins stands among those most well-studied. Although galectins are, by definition, lectins with specificity for β-galactosides, some data have indicated that galectin-1 is able to interact with α-galactosides as well. The recognition of Galα(1-3)Gal, Galα(1-4)Gal and Galα(1-6)Glc (melibiose) by galectin-1 was studied by NMR spectroscopy [74]. STD data revealed a pre-eminent recognition of the non-reducing (Gal) end in all instances, while ^{1}H-^{15}N HSQC experiments revealed that all disaccharides bound to the canonical recognition site, although with affinities more than two orders of magnitude weaker than that of lactose.

Similar studies, involving epitope mapping and affinity evaluations by ^{1}H-^{15}N HSQC titration experiments, have been performed on the cyanobacterial *Oscillatoria Agardhii* agglutinin (OAA) to study its binding capabilities to the mannose-rich clusters found in the HIV gp120 glycoprotein [75]. The lectin was revealed to recognise the branched core unit of Man-9 via two distinct binding sites located at opposite ends of the protein. X-ray studies demonstrated that, unlike most HIV-inactivating lectins, OAA recognised the core structure of the triantennary Man-9, rather than its reducing or non-reducing ends, thus raising the question whether a synergistic action with terminal mannose-recognising lectins could be explored.

Regarding the interaction between enzymes and their carbohydrate substrates, ST6Gal-I, a glycosyltransferase of medical interest, which catalyzes the transfer of sialic acid residues from CMP-NeuAc to galactose residues of glycolipids and glycoproteins on the surface of mammalian cells has been monitored [76]. Its structural determination by X-ray diffraction or NMR has been precluded by the enzyme being glycosylated and membrane-bound, and because of its large size of 40 kDa. Thus, the authors elegantly addressed the issue of how structural information can be obtained by NMR, even without complete structural determination. For this purpose, they investigated several aspects of the enzyme active site from the point of view of its bound ligand, using spin-labelled derivatives of both donor and acceptor moieties. Previously, they had utilised CMP-4-carboxy-TEMPO, a spin-labelled donor mimic, containing a paramagnetic nitroxide group [77]. This moiety significantly increases relaxation of nearby magnetic nuclei in a

r^{-6} dependent manner and can be used to map the position of sixteen ^{15}N-labeled phenylalanines relative to the position of the donor. The authors extended the study to a spin-labelled acceptor mimic, LacNAc-TEMPO, and the results were compared to those obtained with the spin-labelled donor. In addition, the labelled donor was used to perturb the signals of a non-labelled acceptor, providing distance restraints between the donor nitroxide and the acceptor carbon atoms. TR-NOE experiments were performed to determine the conformation of the bound molecules, and STD data gave information on the proximity of the ligands epitopes to the protein surface. Taken together, these data were used to generate a model of the enzyme active site geometry, thus providing key insights into a mechanism of action.

The quest for suitable and selective inhibitors for glycosyltransferases is still challenging. A novel concept that allows the design of inhibitors based on the structure of the donor substrate binding pocket has been presented [78]. As a first step, the design, synthesis and analysis of inhibitors of the human blood group B galactosyl-transferase (GTB) was presented. *In silico* docking of bicyclic heteroaromatic ligands to GTB and experimental verification of their binding affinities by STD gave 9-*N*-pentityl uric acid derivatives as non-ionic mimics of UDP. As demonstrated, pentitol linked uric acid derivatives are very well suited to replace uracil, the ribofuranoside and the pyrophosphate group of UDP, resulting in a similar binding affinity in the presence of Mg^{2+} to GTB.

Following on the design and synthesis of glycosyltransferase inhibitors, a major obstacle has always been the demanding chemistry. Glycosyltransferases usually possesses two distinct binding sites, one for the donor substrate, and one for the acceptor substrate. One approach to glycosyltransferase inhibitors is to chemically link donor site and acceptor site ligands to generate high affinity binders. In this context, a novel approach to identify acceptor site ligands from a fragment library has been described [79]. As model, galactosyltransferase (GTB) was chosen. The approach used a combination of STD NMR, spin-lock filtered NMR experiments and surface plasmon resonance measurements. Molecular fragments from a fragment library that bind to the acceptor site of GTB with affinities of the order of a natural acceptor substrate have been identified. They allow for straightforward chemical modifications and, therefore, could serve as scaffolds for potent GTB inhibitors.

STD-NMR studies for GlfT2 have been reported, using two trisaccharide acceptor substrates, β-D-Gal*f*-(1→6)-β-D-Gal*f*-(1→5)-β-D-Gal*f*-O$(CH_2)_7$CH3 and β-D-Gal*f*-(1→5)-β-D-Gal*f*-(1→6)-β-D-Gal*f*-O$(CH_2)_7$CH3, as well as the donor substrate for this enzyme, UDP-Gal*f* [80]. Competition STD-NMR titration experiments and saturation transfer double difference (STDD) experiments with the two trisaccharides were undertaken to explore the bifunctionality of this enzyme, in particular to answer whether one or two active sites are responsible for the formation of both the β-(1→5)- and β-(1→6)-Gal*f* linkages. It was demonstrated that both bind competitively at the same site. This fact suggests that GlfT2 has one active site pocket capable of catalyzing both β-(1→5) and β-(1→6) galactofuranosyl transfer reactions. The addition of UDP-Gal*f* to GlfT2 in the presence of either one or the other

trisaccharide generated a tetrasaccharide product. This observation means that the enzyme was catalytically active under the conditions at which the STD-NMR experiments were carried out. Moreover, epitope mapping demonstrated a greater enhancement toward the 'reducing' ends of both trisaccharides, and that UDP-galactofuranose (UDP-Gal*f*) made more intimate contacts through its nucleotide moiety. This observation is consistent with the greater flexibility required within the active site of the reaction between the growing polymer acceptor and the UDP-Gal*f* donor.

The oligomerisation profile of chemokines is also a topic of major interest, especially in what concerns their binding to glycosaminoglycans (GAGs). There is sparse information of such events in solution, probably due to the poor solubility of the resulting aggregates. The structure of CCL27, a cutaneous T cell-attracting chemokine, its oligomerisation state and its interactions with GAGs have been investigated by NMR [81]. The structure of CCL27 was solved by using standard procedures, and PFG diffusion and ^{15}N relaxation data revealed that CCL27 displays transitions between monomer, dimer, and tetramer states within the low millimolar range. To characterise the dimer interface, intersubunit NOEs were scoped by utilising a ^{13}C/^{15}N-separated and filtered NOESY experiment on a mixed isotope sample containing 50% ^{13}C-labelled CCL27, and 50% ^{15}N-labelled CCL27, which maximised the detection of intersubunit NOEs between ^{1}H-^{13}C and ^{1}H-^{15}N protons. The oligomeric profile of CCL27 was shown to be more dynamic than that of other members of the family, with more than one type of dimer being present. Direct study of GAG-promoted oligomerisation of CCL27 was impossible due to precipitation of the chemokine upon addition of the heparin oligosaccharides, a fact that was attributed to the formation of insoluble, higher order oligomers. Such effect was then investigated on CCL2, a related chemokine. Addition of a heparin octasaccharide to a CCL2 sample at 1:2 ratio shifted the D_s value of the latter from the monomer value to the expected value for a tetramer, strongly supporting the notion of chemokine association upon interaction with GAG.

The quaternary structure of CCL5/Rantes chemokine has also been investigated in solution by a multidisciplinary approach involving SAXS and NMR [82]. ^{1}H-^{15}N HSQC experiments on oligomeric samples revealed only one set of cross-peaks, suggesting that the oligomeric complexes are symmetric. In order to solve the oligomeric structure, a different strategy was used, which involved residual dipolar coupling (RDC) measurements. Such calculations, supported by SAXS data, evidenced a multimeric assembly of up to eight monomers, formed by a linear arrangement of concatenated dimers. The calculated structure unveiled two solvent-accessible BBXB motifs in each monomer, which become lengthwise-aligned in the oligomer, thus forming sites on which GAG chains would bind. Moreover, this arrangement offers an explanation to the promotion of chemokine oligomerisation exerted by GAG binding, since electrostatic repulsion of positively-charged residues in the binding site would keep the oligomer size small unless they are neutralised by negatively-charged sulphated residues of GAGs. In this context, the characterisation and binding studies of the GAGs [83] (sulfated glycosaminoglycans) are

fundamental from the biological viewpoint because they regulate the activity of, and bind to, a number of different types of proteins. The analysis of the heparin pentasaccharide in the free state and in complex with FGF2 receptor indicated the existence of a conformational selection process. Although there is equilibrium of the 2-*O*-sulfo-α-L-iduronate ring (IdoA2S) in the free state between the $^{1}C_4$ and $^{2}S_0$ conformers (ratio 60:40), FGF2 receptor selects only the unique twisted-boat $^{2}S_0$ conformation of this residue, changing the major conformation of the pentasaccharide in solution.

Turning back to methodological aspects of STD experiments, the last years have witnessed a significant expansion of its applicability beyond the single ligand-receptor system, being used to study protein-ligand interactions in viruses, micelles, or even within whole, living cells. For instance, STD has been used on novovirus-like particles (VLPs) to study their interaction with inhibitors targeting the fucose-binding site [84], and on human embryonic kidney cells to detect the recognition of sweet ligands by the human sweet receptor [85]. The binding of VLPs of a GII.4 norovirus strain to host attachment factors has also been investigated by NMR spectroscopy [86]. STD NMR experiments in conjunction with transfer NOESY experiments have been employed to elucidate the binding epitopes and bioactive conformations of synthetic HBGA fragments. The experiments deliver for the first time a systematic overview of the structural requirements for norovirus attachment to host cell antigens (HBGAs) at atomic resolution.

Mohan *et al*. [87] have reported on the exploitation of the 150-cavity in the active sites of group-1 neuraminidases for the design of new triazole-containing carbocycles related to oseltamivir. Inhibition studies with virus-like particles (VLPs) containing the influenza virus neuraminidase-1 (N1) activity indicate that several candidates are inhibitors, with Ki values in the 10^{-5}–10^{-8} M range. In contrast, a known candidate that preserves the free amino group and a new candidate containing a guanidine function are better inhibitors, with Ki values of 1.5×10^{-9} and 4.6×10^{-10} M, respectively. The most active inhibitor of the N1 enzyme in the triazole series was selective for the N1 class and showed significantly less inhibition (Ki = 2.6 μM *vs* 0.07 μM) of the free influenza virus neuraminidase-2 (N2). STD-NMR spectroscopic studies with this compound and the VLPs show that the entire molecule forms contacts with residues in the active site.

To overcome the challenge posed by the tendency of many cell lines to aggregate and to settle, a methodology combining STD with high resolution-magic angle spinning (HR-MAS) has been developed and applied to detect the interaction of the sugar co-transporter SGTL1 with its ligands [88].

Among the protein-detected methods, those relying on the receptor epitope mapping through the study of resonance shifting in ''finger-print'' spectra (such as ^{1}H-^{15}N HSQC and ^{1}H-^{13}C HSQC) are of paramount importance. ^{1}H, ^{13}C and ^{15}N NMR assignments of several carbohydrate-binding proteins have been recently reported, including those of galectin-1 [89], galectin-7 [90] and a bacterial D-allose binding protein [91], both in its free form and in complex with the sugar. It is foreseen that these works

should pave the way for the study of the interactions between these molecular receptors and their carbohydrate ligands in the near future.

4.2 Carbohydrate-carbohydrate and carbohydrate-nucleic acid interactions

NMR experiments using deuterium isotope shifts can be use for analysis of Lewis X–Lewis X interactions under 100% water conditions, similar to the biological environment [92]. The authors focused on fast exchange O–H groups of the Lewis X glycan because they are expected to be directly involved in stabilizing tertiary structures and interactions with other molecules by hydrogen bonding and/or coordinating the metallic ions. Lewis X–Lewis X interactions were analyzed through measurement of proton-exchange rates under physiological conditions. The technique is able to quantify the proton exchange of hydroxyl groups on glycans in 100% water, in the ten to hundred milliseconds range. Also, it was revealed that Ca^{2+} slows down the rates of H/D exchange of several hydroxyl protons of Lewis X. These groups must be involved in the Lewis X–Lewis X interaction. Not only hydroxyl protons could be observed but also precisely evaluated their dynamics through measurement of the small difference in their exchange rates. This simple ^{13}C-NMR strategy is suitable for analyzing extremely weak interactions including those between carbohydrates, and permits identifying the hydroxyl protons involved in hydrogen bonding and metal coordination.

Regarding amino glycosides, the structure of the complex between apramycin and Cu(II) has been studied by NMR. The paramagnetic relaxation enhancements provided distance restraints that were employed for molecular mechanics calculations to deduce the structure of the complex [93]. On the other hand, the participation of CH/π stacking interactions in the recognition of amino glycosides by RNA has also been reported to be crucial. Thus, by using kanamycin derivatives as simple synthetic models, the role played by CH/π contacts as well as electrostatic forces and solvation of the charged groups have been examined by NMR experiments, and molecular dynamics and quantum mechanics calculations. It has been shown that the recognition phenomenon is extremely dependent on the protonation state of the interacting amino sugars [94].

The interaction of sugar fragments with DNA has also been studied by means of NMR. A series of glyco-oligoamides, with different sugars attached to bi-pyrrol moieties, have been synthesised. The sugar residues modulated the interaction with DNA fragments and contribute to the selectivity of the binding. The 3D structure and binding features of the β-galactose-containing oligoamide with a DNA model sequence was elucidated by standard NMR and modelling methods. The bound galactooligoamide adopts a crescent shaped hairpin conformation, and is located within the minor groove of the DNA fragment, at its central AATT region [95]. Also, the sugar moiety of the anthracyclines seems to be critical for their antitumor activity through its binding to the minor groove of DNA. In this context, the complexes between abarubicin (Men10755) and daunorubicin (Men10749), two anthracycline disaccharide derivatives, with double helix oligonucleotides $d(CGTACG)_2$ and $d(CGATCG)_2$ have also been elucidated by NMR [96].

The importance of the interaction of carbohydrates with DNA has also been studied on carbohydrate-DNA conjugate models, where the oligonucleotide is capped by a mono or a disaccharide attached to the 5′ end of the DNA strand. Chemical shift variations at the DNA between the species conjugated and non conjugated to the sugar, together with NOEs analysis permitted the authors to guess that the polar sugar moieties may stack onto DNA duplexes and stabilise sequences with terminal C-G or G-C base pairs, although with low affinity [97].

4.3 Interactions of carbohydrates with artificial receptors

The development of artificial receptors for carbohydrates has received increasing attention in recent years, being NMR one of the main tools to study the binding process. Very different chemical architectures have been described in order to mimic the molecular recognition process carried out by the carbohydrate binding proteins in Nature. They share however, common features: on one hand, the occurrence of groups capable of participating in hydrogen bonds and, on the other hand, the presence of aromatic moieties being able to form CH-π stacking interactions.

Therefore, NMR has been employed in order to shed light on the structural details as well as to assess binding affinities of the recognition processes. The group of Davis at Bristol has found out that a cage-based receptor, that had demonstrated to recognise glucose in water, was actually a better receptor for β-GlcNAc [98]. The NOESY spectrum of the complex between the synthetic receptor and GlcNAcβOMe allowed identifying an important number of intermolecular NOEs, providing key information for describing the interactions responsible for the binding process at the atomic level. Thus, the recognition process takes place with similar affinity and selectivity to that occurring in Nature, also employing a similar combination of CH-π and hydrogen bonding between β-GlcNAc and the receptor. Indeed, these similarities have led to entitle this type of receptors as 'synthetic lectins' [99]. Further studies using different *O*-glycopeptide fragments permitted to unravel the different contributions for *O*-β-GlcNAc-glycopeptide recognition by these synthetic receptors [100]. Synthetic modifications on the receptor with further introduction of substituents at the biphenyl groups led to a modulation in the affinities and selectivities. In particular, the attachment of alkoxy groups increased affinity for β-glucosyl groups [101], while the introduction of electron donor or withdrawing substituents modified the binding affinities, but to a lesser extent than that expected, evidencing the delicate interplay of different factors for the recognition of carbohydrates in water [102].

Other scaffolds for constructing synthetic carbohydrate receptors have also been investigated. The triethyl benzene structure differently substituted has been widely explored for sugar recognition in polar organic solvents [103]. Selective recognition of β-mannosides has been achieved by a hexaamino dipyrrolic derivative of the triethyl benzene scaffold containing a macrocycle in the structure [104, 105]. Moreover, this new receptor showed chiral discrimination, since the two possible enantiomeric forms of the receptor displayed different affinities for α- and β-mannosides. The 3D structure of all of the complexes was elucidated by a combination of NMR

methods with molecular modelling permitting to explain the observed selectivities [106].

The effect of the substitution of triethyl benzene core-based artificial receptors on their recognition properties has also been explored. NMR chemical shift perturbation methods, with either direct or reverse titrations were employed to monitor the affinity and selectivity towards monosaccharide derivatives in organic solvents. For instance, the substitution of the core scaffold with 8-hydroxyquinoline groups yielded receptors that showed preferred binding toward β-configurations over the α analogues. One of these newly synthesised receptors bearing two 8-hydroxyquinoline fragments was found to effectively bind β-galactosides, while the one bearing just one 8-hydroxyquinoline moiety showed preference for the β-glucoside analogue [107]. In contrast, decoration of the same scaffold with phenantroline groups yielded receptors with binding preference for α-versus β-anomers [108]. On the other hand, the benzene core scaffold has also been decorated with isopropyl- and isobutyl-amino functionalities providing good affinities [109], while the use of larger aromatic surface to host disaccharides has also been explored, by employing dimesitylmethane groups [110].

A new group of artificial receptors has been reported, based on a completely different scaffold. Porphyrin was chosen for providing the aromatic surface, to which various chemical functionalities were attached. In this case, NMR titrations were further supported by other techniques (UV titrations and CD measurements) to demonstrate that these receptors are able to recognise pyranose and furanose rings with affinities in the mM range in methanol/chloroform mixtures, using a combination of CH-π interactions provided by the porphyrin moiety and of hydrogen bonds provided by the polar chemical groups attached to it [111].

Acknowledgments

We thank the Ministry of Economy and Competitiveness (MINECO) of Spain for financial support (Grant CTQ2009-08536). AA, AC, AB, PB, MCFA, and FM thank MINECO (Juan de la Cierva), MINECO (Ramón y Cajal), MINECO (FPI), EU (GlycoHit), CSIC (JAE-DOC), and FCT institutions for funding, respectively.

References

1 X. Hu, W. Zhang, I. Carmichael and A. S. Serianni, *J. Am. Chem. Soc.*, 2010, **132**, 4641.
2 X. Hu, I. Carmichael and A. S. Serianni, *J. Org. Chem.*, 2010, **75**, 4899.
3 B. Yu, H. van Ingen, S. Vivekanandan, C. Rademacher, S. E. Norris and D. I. Freedberg, *Journal of Magnetic Resonance*, 2012, **215**, 10.
4 M. Hricovini, *J. Phys. Chem. B*, 2011, **115**, 1503.
5 E. Säwén, B. Stevensson, J. Östervall, A. Maliniak and G. Widmalm, *J. Phys. Chem. B*, 2011, **115**, 7109.
6 J. Xia, C. J. Margulis and D. A. Case, *J. Am. Chem. Soc.*, 2011, **133**, 15252.
7 M. Erdèlyi, E. d'Auvergne, A. Navarro-Vázquez, A. Leonov and C. Griesinger, *Chem. Eur. J.*, 2011, **17**, 9668.

8 a) A. Mallagaray, G. Domínguez, J. Jiménez-Barbero and J. Pérez-Castells, *Chem Commun (Camb)*, 2011, **47**, 7179. b) S. Yamamoto, T. Yamaguchi, M. Erdélyi, C. Griesinger and K. Kato, *Chemistry*, 2011, **17**, 9280.
9 A. Gallinger, T. Biet, L. Pellerin and T. Peters, *Angew. Chem. Int. Ed.*, 2011, **50**, 11672.
10 A. Imberty and S. Perez, *Chem. Rev.*, 2000, **100**, 4567.
11 B. Mulloy, M. J. Forster, C. Jones and D. B. Davies, *Biochem. J.*, 1993, **293**, 849.
12 M. Sattelle, S. U. Hansen, J. Gardiner and A. Almond, *J. Am. Chem. Soc.*, 2010, **132**, 13132.
13 M. L. DeMarco and R. J. Woods, *Glycobiology*, 2008, **18**, 426.
14 H. Kogelberg, D. Solís and J. Jiménez-Barbero, *Curr. Opin. Struct. Biol.*, 2003, **13**, 646.
15 K. N. Kirschner and R. J. Woods, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 10541.
16 V. Roldós, F. J. Cañada and J. Jiménez-Barbero, *ChemBioChem*, 2011, **12**, 990.
17 F.-X. Theillet, C. Simenel, C. Guerreiro, A. Phalipon, L. A. Mulard and M. Delepierre, *Glycobiology*, 2011, **21**, 109.
18 U. Olsson, E. SäwØn, R. Stenutz and G. Widmalm, *Chem. Eur. J.*, 2009, **15**, 8886.
19 E. Hatcher, E. Sawen, G. Widmalm and A. D. MacKerell, Jr., *J. Phys. Chem. B*, 2011, **115**, 597.
20 E. Säwén, T. Massad, C. Landersjö, P. Damberg and G. Widmalm, *Org. Biomol. Chem.*, 2010, **8**, 3684.
21 Y. Kamiya, S. Yamamoto, Y. Chiba, Y. Jigami and K. Kato, *J Biomol NMR*, 2011, **50**, 397.
22 E. Säwén, M. U. Roslund, I. Cumpstey and G. Widmalm, *Carbohydrate Research*, 2010, **345**, 984.
23 B. M. Sattelle, J. Shakeri, I. S. Roberts and A. Almond, *Carbohydrate Research*, 2010, **345**, 291.
24 K. Hanna, M. Jonsson, R. Pendrill and G. Widmalm, *Magn. Reson. Chem.*, 2011, **49**, 117.
25 P. J. Meloncelli, L. J. West and T. L. Lowary, *Carbohydrate Research*, **346**, 1406.
26 T. Fujimoto, T. Kato, Y. Usui, O. Kamo, K. Furihata, K. Tsubono, T. Kato, T. Machinami and M. Tashiro, *Carbohydrate Research*, 2011, **346**, 1991.
27 M. Zerbetto, A. Polimeno, D. Kotsyubynskyy, L. Ghalebani, J. Kowalewski, E. Meirovitch, U. Olsson and G. Widmalm, *J Chem Phys.*, 2009, **131**, 234501.
28 C. D. Blundell, I. S. Roberts, J. K. Sheehan and A. Almond, *J Mol Microbiol Biotechnol*, 2009, **17**, 71.
29 V. Gargiulo, M. A. Morando, A. Silipo, A. Nurisso, S. Pérez, A. Imberty, F. J. Cañada, M. Parrilli, J. Jiménez-Barbero and C.D. Castro, *Glycobiology*, 2010, **20**, 1208.
30 E. Säwén, E. Huttunen, X. Zhang, Z. Yang and G. Widmalm, *J Biomol NMR*, 2010, **47**, 125.
31 S. E. Norris, J. Landstrom, A. Weintraub, T. E. Bull, G. Widmalm and D. I. Freedberg, *Biopolymers*, **97**, 145.
32 a) M. Lundborg and G. Widmalm, *Anal. Chem.*, 2011, **83**, 1514. b) M. Lundborg, C. Fontana and G. Widmalm, *Biomacromolecules,* 2011, **12**, 3851. c) M. U. Roslund, E. Säwén, J. Landström, J. Rönnols, K. Hanna, M. Jonsson, M. Lundborg, M. V. Svensson and G. Widmalm, *Carbohydrate Research*, **346**, 1311.
33 A. W. Barb, A. J. Borgert, M. Liu, G. Berany and D. Live, *Methods Enzymol.*, 2010, **478**, 365.

34 A. Varki, *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2009.
35 K. G. Ten Hagen, T. A. Fritz and L. A. Tabak, *Glycobiology*, 2003, **13**, 1.
36 M. A. Tarp and H. Clausen, *Biochim. Biophys. Acta*, 2008, **1780**, 546.
37 B. Meyer and H. Möller, *Top. Curr. Chem.*, 2007, **267**, 187.
38 D. M. Coltart, A. K. Royyuru, L. J. Williams, P. W. Glunz, D. Sames, S. D. Kuduk, J. B. Schwarz, X.-T. Chen, S. J. Danishefsky and D. H. Live, *J. Am. Chem. Soc.*, 2002, **124**, 9833.
39 F. Corzana, J. H. Busto, M. Garcia de Luis, J. Jiménez-Barbero, A. Avenoza and J. M. Peregrina, *Chem. Eur. J.*, 2009, **15**, 3863.
40 F. Corzana, J. H. Busto, F. Marcelo, J. L. Asensio, S. Martín-Santamaría, M. Garcia de Luis, J. Jiménez-Barbero, A. Avenoza and J. M. Peregrina, *Chem. Eur. J*, 2011, **17**, 3105.
41 R. Hashimoto, N. Fujitani, Y. Takegawa, M. Kurogochi, T. Matsushita, K. Naruchi, N. Ohyabu, H. Hinou, X. D. Gao, N. Manri, H. Satake, A. Kaneko, T. Sakamoto and S-I. Nishimura, *Chem. Eur. J.*, 2011, **17**, 2393.
42 G. Arsequell, M. Rosa, C. Mayato, R. L. Dorta, V. Gonzalez-Nunez, K. Barreto-Valer, F. Marcelo, L. P. Calle, J. T. Vázquez, R. E. Rodríguez, J. Jiménez-Barbero and G. Valencia, *Org. Biomol. Chem.*, 2011, **9**, 6133.
43 Y. Yamaguchi and K. Kato, *Methods Enzymol.*, 2010, **478**, 305.
44 K. Kato, Y. Yamaguchi and Y. Arata, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **56**, 346.
45 A. W. Barb and J. H. Prestegard, *Nat. Chem. Biol.*, 2011, **7**, 147.
46 K.-F. Mo, T. Fang, S. H. Stalnaker, P. S. Kirby, M. Liu, L. Wells, M. Pierce, D. H. Live and G.-J. Boons, *J. Am. Chem. Soc.*, 2011, **133**, 14418.
47 F. Squeglia, R. Marchetti, A. Ruggiero, R. Lanzetta, D. Marasco, J. Dworkin, M. Petoukhov, A. Molinaro, R. Berisio and A. Silipo, *J. Am. Chem. Soc.*, 2011, **133**, 20676.
48 M. M. Mackeen, A. Almond, M. Deschamps, I. Cumpstey, A. J. Fairbanks, C. Tsang, P. M. Rudd, T. D. Butters, R. A. Dwek and M. R. Wormald, *J. Mol. Biol.*, 2009, **387**, 335.
49 a) Y. Yamaguchi and K. Kato, Methods in Enzymology, **478**. b) K. Kato, Y. Yamaguchi and Y. Arata, *Progress in Nuclear Magnetic Resonance Spectroscopy*, 2010, **56**, 346.
50 J. L. Magnani and B. Ernst, *Discover. Med.*, 2009, **43**, 247.
51 V. Roldos, F. J. Canada and J. Jimenez-Barbero, *ChemBioChem*, 2011, **12**, 990.
52 M. von Itzstein, *Nat. Rev. Drug Discovery*, 2007, **6**, 967.
53 B. J. Gross, J. G. Swoboda and S. Walker, *J. Am. Chem. Soc.*, 2008, **130**, 440.
54 F. Marcelo, Y. He, S. A. Yuzwa, L. Nieto, J. Jiménez-Barbero, M. Sollogoub, D. J. Vocadlo, G. D. Davies and Y. Bleériot, *J. Am. Chem. Soc.*, 2009, **131**, 5390.
55 T. Kuntothom, M. Raab, I. Tvaroska, S. Fort, S. Pengthaisong, J. Cañada, L. Calle, J. Jiménez-Barbero, J. R. K. Cairns and M. Hrmova, *Biochemistry*, 2010, **49**, 8779.
56 P. J. Hajduk and J. Greer, *Nat. Rev. Drug Discovery*, 2007, **6**, 211.
57 R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero and O. Ramström, *Angew. Chem. Int. Ed.*, 2010, **49**, 589.
58 R. Caraballo, M. Sakulsombat and O. Ramström, *ChemBioChem*, 2010, **11**, 1600.
59 J. J. Reina, S. Sattin, D. Invernizzi, S. Mari, L. Martinez-Prats, G. Tabarani, F. Fieschi, R. Delgado, P. M. Nieto, J. Rojo and A. Bernardi, *ChemMedChem*, 2007, **2**, 1030.

60 J. J. Reina, I. Diaz, P. M. Nieto, N. E. Campillo, J. A. Paez, G. Tabarani, F. Fieschi and J. Rojo, *Org. Biomol. Chem.*, 2008, **6**, 2743.
61 G. Timpano, G. Tabarani, M. Anderluh, D. Invernizzi, F. Vasile, D. Potenza, P. M. Nieto, J. Rojo, F. Fieschi and A. Bernardi, *ChemBioChem*, 2008, **9**, 1921.
62 C. Guzzi, J. Angulo, F. Doro, J. J. Reina, M. Thépaut, F. Fieschi, A. Bernardi, J. Rojo and P. M. Nieto, *Org. Biomol. Chem.*, 2011, **9**, 7705.
63 M. Morando, Y. Yao, S. Martín-Santamaría, Z. Zhu, T. Xu, F. Javier Cañada, Y. Zhang and J. Jiménez-Barbero, *Chem. Eur. J.*, 2010, **16**, 4239.
64 Y. M. Chabre and R. Roy, in The Sugar Code: Fundamentals of Glycosciences (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, 2009, pp. 53 –70.
65 Y. M. Chabre, D. Giguère, B. Blanchard, J. Rodrigue, S. Rocheleau, M. Neault, S. Rauthu, A. Papadopoulos, A. A. Arnold, A Imberty and R. Roy, *Chem. Eur. J.*, 2011, **17**, 6545.
66 D. Schwizer, J. T. Patton, B. Cutting, M. Smiesko, B. Wagner, A. Kato, C. Weckerle, F. P. C. Binder, S. Rabbani, O. Schwardt, J. L. Magnani and Beat Ernst, *Chem. Eur. J.*, 2012, **18**, 1342.
67 a) M. G. Szczepina, D. W. Bleile, J. Müllegger, A. R. Lewis and B. M. Pinto, *Chem. Eur. J.*, 2011, 17, 11438. b) M. G. Szczepina, D. W. Bleile and B. M. Pinto, *Chem. Eur. J.,* 2011, **17**, 11446.
68 H.-C. Siebert, J. Jiménez-Barbero, S. André, H. Kaltner and H.-J. Gabius, *Methods Enzymol*, 2003, **362**, 417.
69 H. C. Siebert, S. André, Y. Lu, M. Frank, H. Kaltner, J. A. van Kuik, E. Y. Korchagina, N. Bovin, E. Tajkhorshid, R. Kaptein, J. F. Vliegenthart, C. W. von der Lieth, J. Jiménez-Barbero, J. Kopitz and H. J. Gabius, *Biochemistry*, 2003, **42**, 14762.
70 V. Roldos, F. J. Canada and J. Jimenez-Barbero, *Chembiochem*, 2011, **12**, 990.
71 J. Angulo, P. M. Enriquez-Navas and P. M. Nieto, *Chemistry*, 2010, **16**, 7803.
72 M. Marradi, P. Di Gianvincenzo, P. M. Enriquez-Navas, O. M. Martinez-Avila, F. Chiodo, E. Yuste, J. Angulo and S. Penades, *J. Mol. Biol.*, 2011, **410**, 798.
73 M. Plum, Y. Michel, K. Wallach, T. Raiber, S. Blank, F. I. Bantleon, A. Diethers, K. Greunke, I. Braren, T. Hackl, B. Meyer and E. Spillner, *J. Biol. Chem.*, 2011, **286**, 43103.
74 M. C. Miller, J. P. Ribeiro, V. Roldos, S. Martin-Santamaria, F. J. Canada, I. A. Nesmelova, S. Andre, M. Pang, A. A. Klyosov, L. G. Baum, J. Jimenez-Barbero, H. J. Gabius and K. H. Mayo, *Glycobiology*, 2011, **21**, 1627.
75 L. M. Koharudin and A. M. Gronenborn, *Structure*, 2011, **19**, 1170.
76 S. Liu, L. Meng, K. W. Moremen and J. H. Prestegard, *Biochemistry*, 2009, **48**, 11211.
77 S. Liu, A. Venot, L. Meng, F. Tian, K. W. Moremen, G. J. Boons and J. H. Prestegard, *Chem. Biol.*, 2007, **14**, 409.
78 K. Schaefer, J. Albers, N. Sindhuwinata, T. Peters and B. Meyer, *ChemBioChem*, 2012, **13**, 443.
79 C. Rademacher, J. Landström, N. Sindhuwinata, M. M. Palcic, G. Widmalm and T. Peters, *Glycoconj. J.*, 2010, **27**, 349.
80 a) M. G. Szczepina, R. B. Zheng, G. C. Completo, T. L. Lowary and B. M. Pinto, *ChemBioChem*, 2009, 10, 2052. b) M. G. Szczepina, R. B. Zheng, G. C. Completo, T. L. Lowary and B. M. Pinto, *Bioorganic & Medicinal Chemistry*, 2010, **18**, 5123.
81 A. L. Jansma, J. P. Kirkpatrick, A. R. Hsu, T. M. Handel and D. Nietlispach, *J. Biol. Chem.*, 2010, **285**, 14424.
82 X. Wang, C. Watson, J. S. Sharp, T. M. Handel and J. H. Prestegard, *Structure*, 2011, **19**, 1138.

83 L. Nieto, A. Canales, G. Giménez-Gallego, P.M. Nieto and J. Jiménez-Barbero, *Chem. Eur. J.*, 2011, **17**, 11204.
84 C. Rademacher, J. Guiard, P. I. Kitov, B. Fiege, K. P. Dalton, F. Parra, D. R. Bundle and T. Peters, *Chemistry*, 2011, **17**, 7442.
85 F. M. Assadi-Porter, M. Tonelli, E. Maillet, K. Hallenga, O. Benard, M. Max and J. L. Markley, *J. Am. Chem. Soc.*, 2008, **130**, 7212.
86 B. Fiege, C. Rademacher, J. Cartmell, P. I. Kitov, F. Parra and T. Peters, *Angew. Chem. Int. Ed.*, 2012, **51**, 928.
87 S. Mohan, S. McAtamney, T. Haselhorst, M. von Itzstein and B. M. Pinto, *J. Med. Chem.*, 2010, **53**, 7377.
88 C. Airoldi, S. Giovannardi, B. La Ferla, J. Jimenez-Barbero and F. Nicotra, *Chemistry*, 2011, **17**, 13395.
89 I. V. Nesmelova, M. Pang, L. G. Baum and K. H. Mayo, *Biomol. NMR Assign.*, 2008, **2**, 203.
90 I. V. Nesmelova, M. A. Berbis, M. C. Miller, F. J. Canada, S. Andre, J. Jimenez-Barbero, H. J. Gabius and K. H. Mayo, *Biomol. NMR Assign.*, 2011.
91 D. Castano and O. Millet, *Biomol. NMR Assign.*, 2011, **5**, 31–34.
92 S. Hanashima, K. Kato and Y. Yamaguchi, *Chem. Commun.*, 2011, **47**, 10800.
93 D. Balenci, G. Bonechi, N. D'Amelio, E. Gaggelli, N. Gaggelli, E. Molteni, G. Valensin, W. Szczepanik, M. Dziuba, G. Swiecicki and M. Jezowska-Bojczuk, *Dalton Trans.*, 2009, **7**, 1123.
94 T. Vacas, F. Corzana, G. Jimeénez-Osés, C. González, A. M. Gómez, A. Bastida, J. Revuelta and J. L. Asensio, *J. Am. Chem. Soc.*, 2010, **132**, 12074.
95 P. Peñalver, F. Marcelo, J. Jiménez-Barbero and C. Vicent, *Chem. Eur. J.*, 2011, **17**, 4561.
96 S. Mazzini, L. Scaglioni, F. Animati and R. Mondelli, *Bioorg. Med. Chem.*, 2010, **18**, 1497.
97 R. Lucas, I. Góomez-Pinto, A. Aviñó, J. J. Reina, R. Eritja, C. Gonzáalez and J. C. Morales, *J. Am. Chem. Soc.*, 2011, **133**, 1909.
98 Y. Ferrand, E. Klein, N. P. Barrwell, M. C. Crumps, J. Jiménez-Barbero, C. Vicent, G.-J. Boons, S. Ingale and A. P. Davis, *Angew. Chem. Int. ed.*, 2009, **48**, 1775.
99 S. Kubik, *Angew. Chem. Int. Ed.*, 2009, **48**, 1722.
100 F. Corzana, A. Fernández-Tejada, J. H. Busto, G. Joshi, A. P. Davis, J. Jiménez-Barbero, A. Avenoza and J. M. Perengrina, *ChemBioChem*, 2011, **12**, 110.
101 N. P. Barrwell, M. C. Crumps and A. P. Davis, *Angew. Chem, Int. ed.*, 2009, **48**, 7673.
102 N. P. Barrwell, *J. Org. Chem.*, 2011, **76**, 6548.
103 M. Cacciarini, C. Nativi, M. Norcini, S. Staderini, O. Francesconi and S. Roelens, *Org. Biomol. Chem.*, 2011, **9**, 1085.
104 A. Ardá, C. Venturi, C. Nativi, O. Francesconi, G. Gabrielli, F. J. Cañada, J. Jiménez-Barbero and S. Roelens, *Chemistry*, 2010, **16**, 414.
105 C. Nativi, O. Francesconi, G. Gabrielli, A. Vacca and S. Roelens, *Chemistry*, 2011, **17**, 4814.
106 A. Ardá, F. J. Cañada, C. Nativi, O. Francesconi, G. Gabrielli, A. Ienco, J. Jiménez-Barbero and S. Roelens, *Chemistry*, 2011, **17**, 4821.
107 M. Mazik and C. Geffert, *Org. Biomol. Chem.*, 2011, **9**, 2319.
108 M. Mazik, A. Hatmann and P. G. Jones, *Chemistry Eur. J.*, 2009, **15**, 9147.
109 M. Mazik and C. Sonnenberg, *J. Org. Chem.*, 2010, **75**, 6416.
110 M. Mazik and A. C. Buthe, *Org. Biomol. Chem.*, 2009, **7**, 2063.
111 J.-D. Lee, Y.-H. Kim and J.-I. Hong, *J. Org. Chem.*, 2010, **75**, 7588.

Glycosidase inhibitors: versatile tools in glycobiology

Óscar López,* Penélope Merino-Montiel, Sergio Martos and Alejandro González-Benjumea

DOI: 10.1039/9781849734769-00215

Imino sugars, despite being firstly studied as inhibitors more than forty years ago, are still a relevant target in Medicinal and Bioorganic Chemistry, because of their numerous and substantial biological activities. This review covers the most significant reports concerning the synthesis and evaluation of imino sugars as glycosidase inhibitors in the period 2007–2012. The aim is to present the synthesis of new different families and their biological assays, ranging from *in vitro* tests to animal models.

1 Introduction

Ubiquitous carbohydrate-processing enzymes like glycosidases, glycosyl transferases, and glycogen phosphorylase participate in pivotal biological processes in living organisms such as carbohydrate catabolism, cell wall formation, or molecular recognition, among others.[1]

Such enzymes constitute a therapeutic target of great interest in Medicinal Chemistry, as their anomalous functioning is correlated to the development of numerous diseases, like lysosomal storage disorders, diabetes, or cancer; furthermore they also participate in carbohydrate-carbohydrate recognition processes associated to some microbial infections.[1]

Classical description of glycosidases divides these enzymes into two main categories: retaining or inverting glycosidases, depending on the catalytic activity exerted by the carboxyl/carboxylate moieties at the active site. More recently some other mechanisms have been proposed[2] for small subsets of hydrolases, like a substrate-assisted catalytic mechanism for some hexosaminidases, or a tyrosine residue acting as the catalytic nucleophile for some sialidases and neuraminidases.

In this context, potent and selective enzymatic inhibitors, either synthetic or naturally-occurring, have been postulated as potential pharmacological agents. Among the different families of glycosidase inhibitors, those with a nitrogen atom replacing either the endocyclic oxygen (imino sugars) or a carbon atom (aza sugars) have received great attention. Many of the classical glycosidase inhibitors fall into the following categories: polyhydroxylated pyrrolidines, piperidines or bicyclic templates such as indolizidines, pyrrolizidines or nortropanes,[1] but the number of families has grown up intensively in the last decade. These compounds might act as transition state analogues in terms of geometry and charge; nevertheless, this assumption is highly discussed. Withers claimed[3] that the sp^2-hybridized imino sugars can be good transition state mimics, whereas sp^3-derivatives, despite exhibiting in many cases high affinity to the enzyme, do not fulfill this definition.

Departamento de Química Orgánica, Facultad de Química, Universidad de Sevilla, c/Profesor García González 1, E-41012, Seville, Spain. E-mail: osc-lopez@us.es

© The Royal Society of Chemistry 2012

Glycosidase inhibitors are also suitable as chemical probes for mechanistic or structural studies concerning glycosidases. For instance, Pieter's group reported[4] the synthesis of an amide derived from deoxynojirimycin for specific labeling of lysosomal glucocerebrosidase; this imino sugar exhibited high affinity in the presence of cell lysate. This provides a method for studying structural requirements in the enzyme active site for potent enzyme-inhibitor interactions.

Davies and co-workers made an exceptional contribution[5] to mechanistic studies of glycosidase inhibition by providing a panel of experimental data on the pH dependence of 18 glycosidase inhibitors binding to TMGHa β-glucosidase. Furthermore, a collection of atomic resolution X-ray crystallography of enzyme-inhibitors complexes[6] and titration calorimetry data were also included. Bols and co-workers studied[7] the inhibition profile of ^{13}C or ^{15}N-labeled azafagomine, a cyclic polyhydroxylated hydrazine. For this purpose, magic angle spinning (MAS) solid NMR of inhibitor complexes was used, deducing that title compound bound to β-glucosidase in its protonated state, whereas non-protonated when interacting with α-glucosidase or glucoamylase.

The aim of this review is to carry out a survey through the most striking aspects concerning glycosidase inhibitors in the period 2007-2012, including a coverage of both, synthetic and biological aspects.

2 Polyhydroxylated pyrrolidines

There is a huge number of imino sugar derivatives bearing a pyrrolidine template, many of them exhibiting inhibition constants in the submicromolar range, probably by mimicking the exo-protonated glycoside in the course of the hydrolysis. Usually, pyrrolidine-based inhibitors show reduced selectivity when compared with six-membered ring imino sugars, due to a major conformational flexibility. Like piperidine-based imino sugars, despite the plethora of different structures reported so far, intense research is still kept in the search for better prototypes and mechanistic studies. In this context, the first X-ray crystal structure of a pyrrolidine imino sugar complexed with a β-glycosidase was reported in 2007.[8]

In a recent work,[9] L-DMDP **1** and L-homoDMDP **2** (Fig. 1), the L-enantiomers of DMDP and homoDMDP, have been synthesized *via* sugar-derived cyclic nitrones as well as 3-deoxy-3-fluoro derivatives. Inhibition assays confirmed that DMDP and homoDMDP are potent α-glycosidase inhibitors, whereas C3-fluorinated analogues showed no inhibition against any of the glycosidases tested; these results proved that C3 hydroxyl group is crucial in the interaction with these enzymes. Introducing a branched carbon chain to the pyrrolidine core is a novel approach to obtain glycosidase inhibitors; in this context, isoDAB **17** and its enantiomer isoLAB are positional isomers of DAB and LAB and bear a branched hydroxymethyl group at C2. Their syntheses from D-ribose (Scheme 1) and D-tagatose, together with their biological activities have been studied by Best and co-workers.[10] IsoDAB was found to be the first example of a carbon-branched pyrrolidine acting as a potent and specific inhibitor of α-glucosidases; on the contrary, isoLAB did not inhibit significantly any of the glycosidases tested.

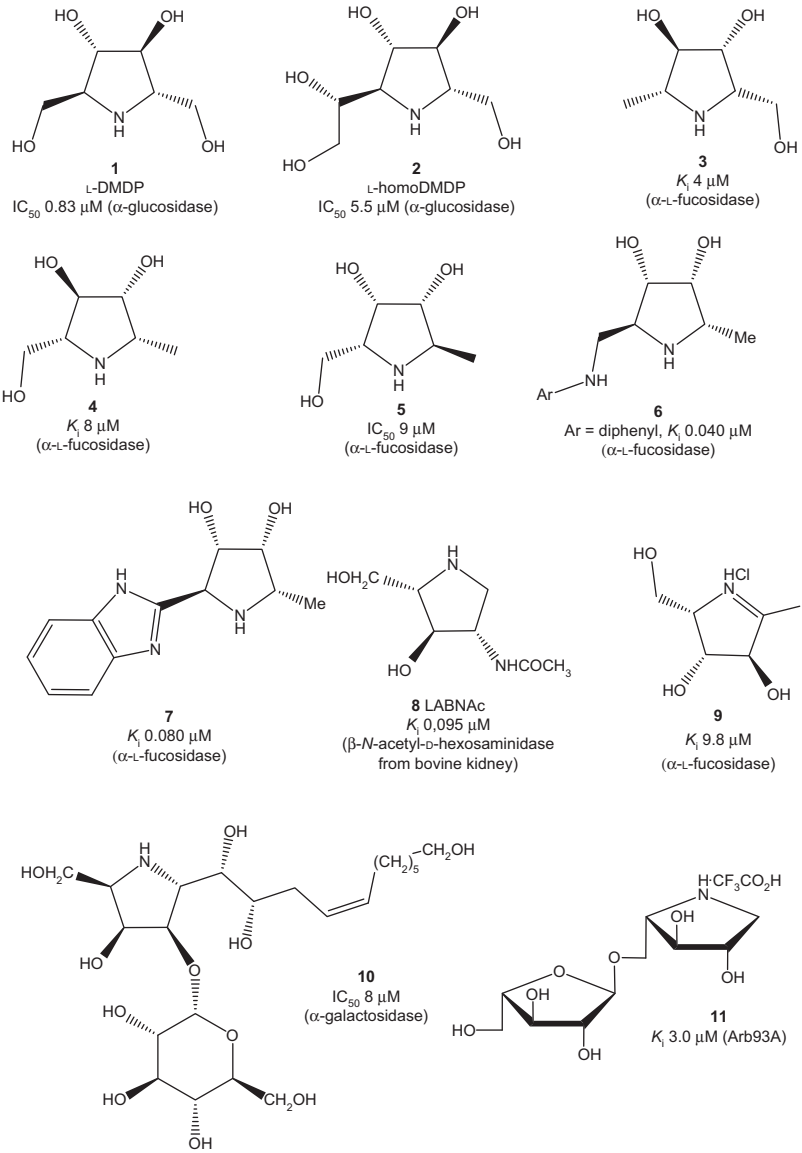
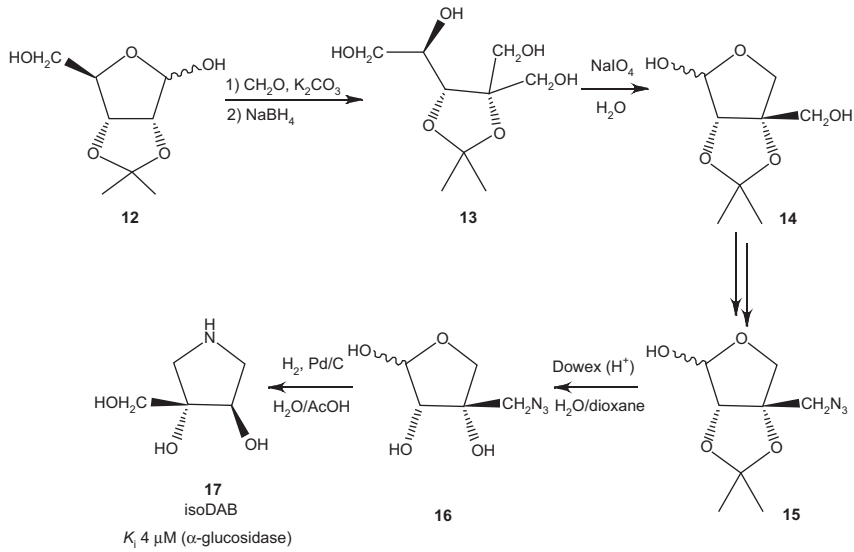


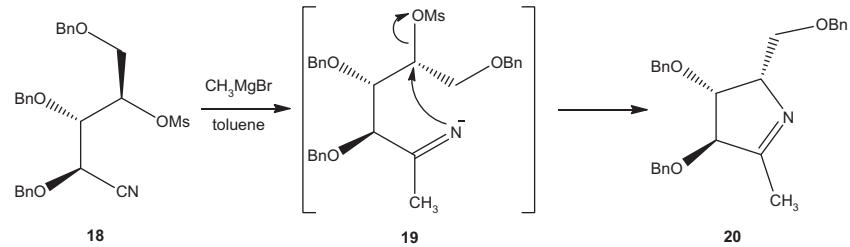
Fig. 1

Izquierdo and co-workers reported[11] a novel synthesis of known trihydroxy methylpyrrolidines **3**–**5** from orthogonally protected pyrrolidines; title compounds incorporated a methyl moiety that might resemble L-*fuco* configuration, and in fact are potent α-L-fucosidase inhibitors with K_i values in the micromolar range (4–9 μM) (Fig. 1).

An approach for improving the activity against α-L-fucosidase is the incorporation of aryl or aminoaryl moieties to the known inhibitor *meso*-pyrrolidine-3,4-diol. Among these compounds, remarkable examples are **6** and **7**, with potency in the submicromolar range.[12]



Scheme 1



Scheme 2

Compounds incorporating an acetamido group on pyrrolidines were also designed as potential hexosaminidase inhibitors; Rountree *et al*. accomplished[13] the synthesis of 2-acetamido-1,4-imino-1,2,4-trideoxy-L-arabinitol (LABNAc, **8**) in 11 steps starting from D-lyxonolactone. This compound was found to be a potent non-competitive inhibitor of β-*N*-acetylhexosaminidases, especially from bovine kidney, with K_i of 95 nM. Its *N*-butyl derivative was also prepared and tested, showing a similar potency.

Behr *et al*. reported[14] the synthesis of ketimine-type imino sugars analogues of nectrisine, an unsaturated 5-membered cyclic imino sugar with an α-glucosidase inhibition even higher than nojirimycin. The key step is a tandem addition-cyclization reaction of a Grignard reagent to ω-methanesulfonylglyconitrile **18** (Scheme 2). For instance, imino sugar **9** (Fig. 1), obtained using this methodology, showed a good inhibition ($K_i = 9.8$ μM) against α-L-fucosidase *via* a mixed mechanism.

Some imino sugars incorporated into oligosaccharide architectures have been found in nature; for instance, α-glucoside **10** (Fig. 1) bearing an unsaturated side chain was isolated from the roots of *Adenophora triphylla*.[15] Biological activity evaluated against several common

glycosidases showed potent and highly specific α-galactosidase inhibition ($IC_{50} = 8$ μM). Considering such results, a smart strategy for obtaining potent glycosidase inhibitors could be the design of oligosaccharides containing imino sugar units, like α-L-arabinofuranosylated pyrrolidines.[16] Using a selective glycosilation methodology,[17] arabinan-like oligosaccharides were prepared, like pseudo-disaccharide **11** (Fig. 1), which was found to be a good inhibitor of the GH93 *exo*-1,5-arabinanase Arb93A from *Fusarium graminearum* ($K_i = 3.0$ μM).

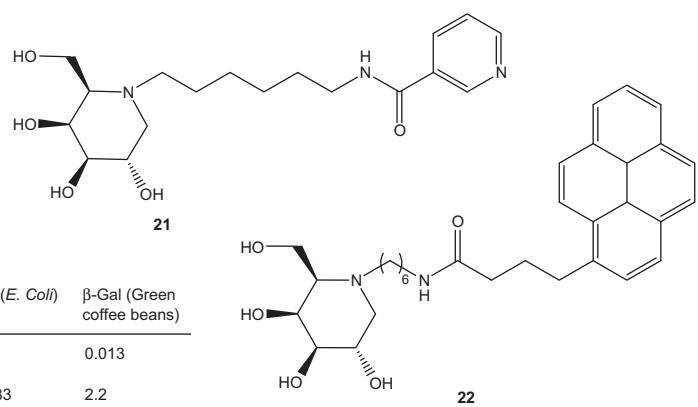
3 Six-membered ring-based imino sugars

3.1 Piperidines

Since the first isolation of polyhydroxy piperidine derivatives from nature, such as nojirimycin and 1-deoxynojirimycin, many other synthetic analogues have been developed using various approaches (*N*-alkylation, substitution of hydroxyl groups, addition of halogens, *etc.*) to improve the spectrum of activity.[1,18]

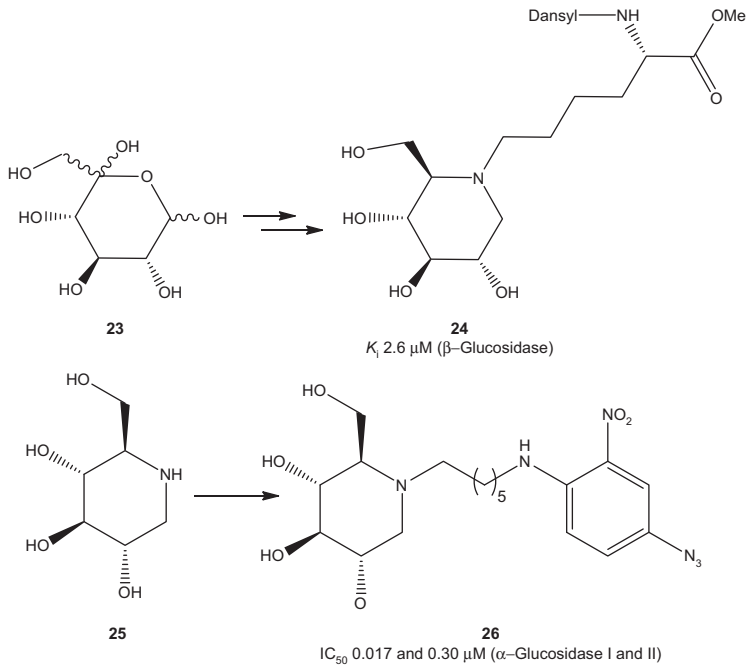
One approach for the synthesis of more potent piperidine-based inhibitors is the increase of the lipophilicity by using fatty chains and aromatic fragments[19,20] as appendages; it is postulated that favorable intermolecular hydrophobic interactions involving lipophilic pockets of the enzyme active site might lead to more potent and selective inhibitors. For instance, Schitter *et al.*[21] have modified 1-deoxygalactonojirimycin by *N*-alkylation with linkers of different lengths bearing terminal aromatic residues. Compounds **21** and **22** (Fig. 2) showed an activity in many cases considerably higher than parent 1-deoxygalactonojirimycin; these results are especially remarkable for β-galactosidase from *E. coli*.

Steiner *et al.*[22] have developed a series of hybrid iminoalditols with lysine, like **24** (Scheme 3), *via* a double reductive amination of D-*xylo*-hexos-5-ulose (**23**) with the terminal amino group of α-methyl ester *N*-Boc-lysine, followed by dansylation of the secondary amine; the corresponding *N*-alkylated derivatives were found to be good inhibitors of β-glucosidase



K_i (μM)	β-Gal (*E. Coli*)	β-Gal (Green coffee beans)
DNJ	13	0.013
21	0.83	2.2
22	0.25	7.0

Fig. 2



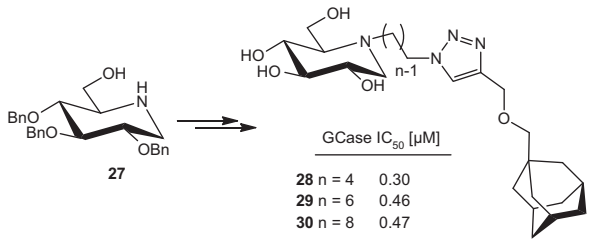
Scheme 3

(*Agrobacterium sp.*), with a potency higher than parent deoxynojirimycin. The same authors reported a similar synthesis starting from L-*arabino*-hexos-5-ulose.[23]

Rawlings *et al.* have also achieved[24] the synthesis of potent α-glucosidase inhibitors based on *N*-alkylated deoxynojirimycin (Scheme 3); in this context, compound **26** was originally designed as a potential photoaffinity probe for endoplasmic reticulum α-glucosidases. It was prepared by reductive amination involving a functionalized aldehyde and the free nitrogen of deoxynojirimycin. This compound proved to be the hitherto best inhibitor of α-glucosidase I, with increased potency compared to *N*-butyl-deoxynojirimycin, one of the most potent inhibitors reported so far for this enzyme.

Ardes-Guisot and co-workers[25] have developed a series of DNJ-derivatives by *click* connection (copper-catalyzed alkyne-azide cycloaddition) to afford triazole derivatives with adamantane appendages, with different linker sizes, like **28**–**30** (Scheme 4). Such compounds behaved as remarkable inhibitors of different α-glucosidases (rice, maltase and sucrase rat intestinal) with IC_{50} values in the low micromolar range (0.030–1.8 μM), all similar or even better than therapeutical inhibitors like miglitol. Furthermore, triazole derivatives also showed a high potency against GCase.[26]

Analogously, Diot *et al.* described[27] the synthesis of some derivatives of deoxynojirimycin bearing triazolyl moieties with aromatic substituents; inhibition of GCase exerted by these compounds proved to be strong (IC_{50} < 5 μM), roughly the similar potency found for *N*-nonyldeoxynojirimycin.



Scheme 4

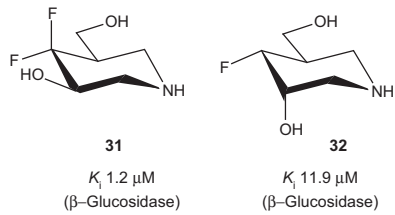
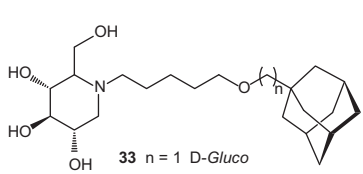


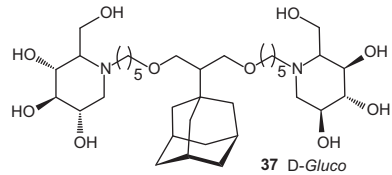
Fig. 3

In recent years, another strategy developed for modifying the structure of imino sugars is the insertion of electron-withdrawing substituents.[28] This idea is based on the fact that strongly electron-withdrawing groups can greatly affect the pK_a of imino sugars, which might have interesting consequences in their mechanism of action against certain glycosidases.[28] Li and co-workers designed[29] the preparation of 4-deoxy-4,4-difluoroisofagomine and analogues from difluoromethylated synthons developed in their research group. The resulting compounds were not very promising, except difluorinated derivative **31** (Fig. 3), which showed a potent and selective β-glucosidase activity ($K_i = 1.2\,\mu M$).[29] Stimulated by these results, some monofluorinated derivatives based on the isofagomine template have been synthesized stereoselectively. Imino sugar **32**[30] (Fig. 3) was obtained, and found to be a strong inhibitor of β-glucosidase (almond) with a K_i of 11.9 μM.

Another approach that should be mentioned concerning inhibition of glycosidases is multivalency; this feature has mainly been studied for carbohydrate-lectin interactions,[31] and currently the number of multivalent imino sugars as glycosidase inhibitors is scarce (*e.g.* Figs. 4 and 5). In principle, considering that glycosidases are monovalent biomolecules, it was not likely that multivalency could be an important fact concerning such enzymes. Overkleeft and co-workers accomplished[32] the preparation of dimeric lipophilic deoxynojirimycins of D-*gluco* and L-*ido* configurations **37** and **38**, respectively, from bis-functionalized adamantanes (Fig. 4). Dimeric derivatives were *in vivo* assayed against glycosylceramide synthase (GCS), a membrane-bound enzyme responsible for the glycosylceramide metabolism, and the results compared with monomeric deoxynojirimycins **33**–**36**. As can be seen, a decreased activity was observed for dimeric architectures. Nevertheless, divalent imino sugars generally kept a comparable potency against glycosidases (*e.g.* membrane-bound β-glucosidase 2, α-glucosidase or sucrose).



33 n = 1 D-*Gluco*
34 n = 1 L-*Ido*
35 n = 2 D-*Gluco*
36 n = 2 L-*Ido*

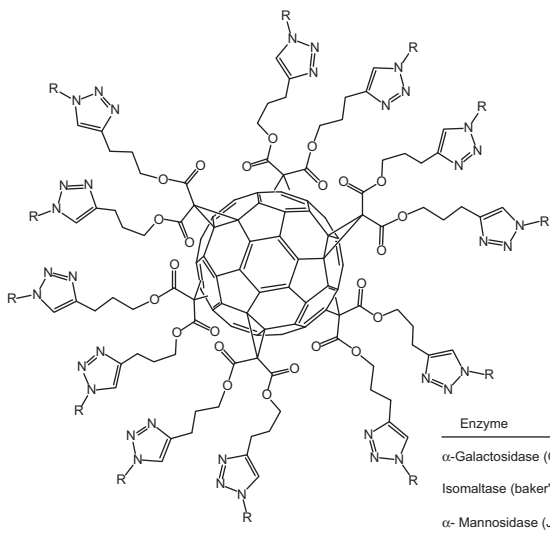


37 D-*Gluco*
38 L-*Ido*

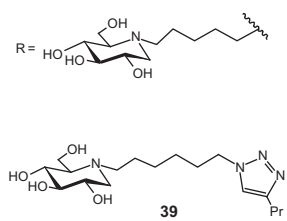
	GCS *in vivo*	GBA2 *in vitro*	Lysosomal α-glucosidase *in vitro*	Sucrase *in vitro*
33	0.2	0.001	0.4	0.5
34	0.1	0.03	>100	>100
35	1	0.005	0.8	0.35
36	1	0.006	>100	200
37	20	0.008	0.35	0.4
38	20	0.150	700	200

IC_{50} in μM

Fig. 4



40



39

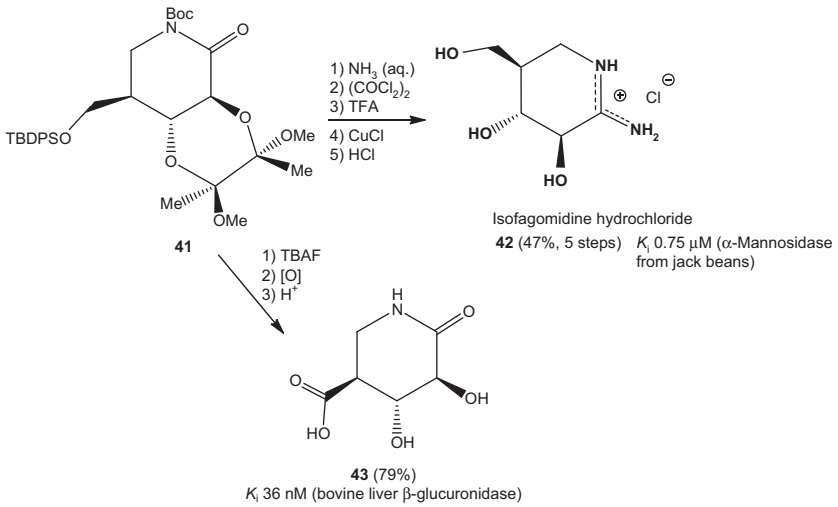
Enzyme	**39**	**40**
α-Galactosidase (Green coffee beans)	N.I. (at 2 mM)	84
Isomaltase (baker's yeast)	943	10.5
α- Mannosidase (Jack bean)	322	0.15

K_i in μM

Fig. 5

Recently, copper(I)-catalyzed alkyne-azide cycloaddition (*click chemistry* approach) has been used to connect azido-containing *N*-substituted deoxynojirimycin to multivalent cores bearing alkynyl residues, like oligo(ethyleneglycol) scaffolds,[33] fullerenes[34] and cyclodextrins;[35] some of these imino sugar derivatives showed strong enhancement of activity against glycosidases as compared to parent monovalent derivatives.

For instance, Fig. 5 depicts fullerene-based dodecavalent derivative **40**, showing roughly 2150 times improvement with respect to monovalent imino sugar **39** against α-mannosidase.



Scheme 5

3.2 Six-membered amidines

Although the use of amidines, amidoximes, and amidrazones as glycosidase inhibitors was developed by Ganem in the 1990's,[1] no previous reports concerning the use of such functional groups at the pseudoanomeric position are found in the literature. In this context, Bols and co-workers reported[36] on the synthesis of isofagomidine **42** (Scheme 5, 16 steps from D-arabinose). The key steps are the introduction of an aminomethyl group at C-4 of D-arabinose, the conversion of C-1 into a nitrile moiety and copper-catalyzed cyclization of an amino group on the nitrile, giving the amidino scaffold.

Surprisingly, isofagomidine hydrochloride behaved as a strong inhibitor of α-mannosidase ($K_i = 0.75$ μM), while being a poor inhibitor of β-glucosidase, regardless the pH of the assay and the incubation time. These results strongly contrast with the inhibitory profile of related isofagomine and isofagomine lactam, low-micromolar inhibitors of β-glucosidase, despite having the three compounds a related structure. Calculated atomic and group charges on isofagomidine, isofagomine and its lactam indicated electronic differences that might support the unexpected different inhibitory profile.

Protected lactam **41** (prepared in 11 steps from D-arabinose) was also efficiently transformed[37] in three steps (silyl group removal, oxidation of the primary alcohol, deprotection) into **43** (Scheme 5). Compound **43** is the uronic acid analogue of isofagomine lactam, and turned out to be a nanomolar glucuronidase inhibitor ($K_i = 36$ nM).

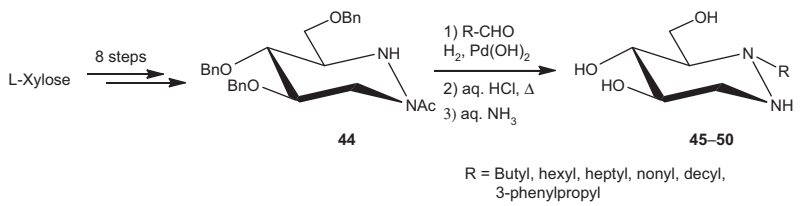
3.3 Azafagomines

Azafagomine **51** (Table 1) is a polyhydroxylated cyclic hydrazine with the D-*gluco* configuration having a hybrid structure between an imino sugar and

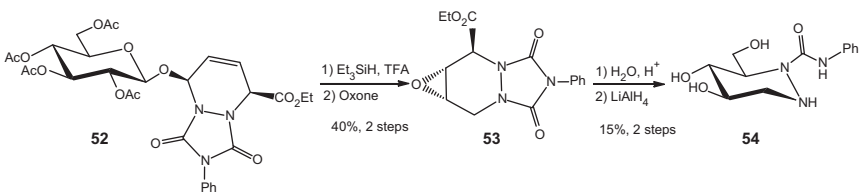
an azasugar.[1] In the search for more potent and selective inhibitors structurally related to azafagomine, López and Bols considered[38] the possibility of a selective alkylation of 1-azafagomine at the 2-*N* position; the N-1 was previously alkylated by Bols' group leading to compounds with reduced activity.[39] Furthermore, alkylation of the pseudoendocyclic position of isofagomine (equivalent to N-2 in azafagomine) was reported to furnish β-glucocerebrosidase inhibitors in the low nanomolar range.[40] Selective N2-alkylation of azafagomine was conducted by reductive amination involving a series of aldehydes and the free nitrogen of partially protected azafagomine **44** (Scheme 6).[41]

Fully unprotected derivatives **45**–**50** were tested as inhibitors of yeast α-glucosidase and β-glucosidase from almonds; inhibitory tests were conducted at pH 6.8 at which all of them should be largely unprotonated.[42] Inhibitory data (selected values are presented in Table 1) indicated that substitution at the 2-*N* position with fatty chains strongly decreased activity against α-glucosidase, suggesting that this kind of substitution disrupts interaction with the active site. Remarkably, a strong improvement in the inhibition of β-glucosidase, especially for the *N*-phenylpropyl derivative **50** ($K_i = 32$ nM), was achieved. These results are probably due to stabilizing interactions with hydrophobic residues located on the active site of β-glucosidase. Therefore the overall effect is a strong selectivity increase towards the latter enzyme.

Alves *et al*. have accomplished[43] a novel preparation of both enantiomers of azafagomine, together with their 2-*N*-phenyl carboxamide derivatives. This synthesis was based on the original preparation of enantiopure 1-azafagomine reported by Bols,[44] in which the key step consisted of a hetero-Diels-Alder reaction between achiral penta-2,4-dien-1-ol and 4-methyl-1,2,4-triazoline-3,5-dione, followed by kinetic resolution of both enantiomeric cycloadducts with lipases. Alves[43] replaced the achiral pentadienol with 2,4-pentanedionates bearing tetra-*O*-acetyl glucosyl chiral auxiliary, which allowed an unprecedented enantiomeric synthesis of azafagomine and its derivatives (Scheme 7).

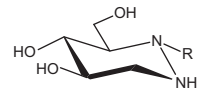


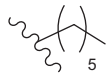
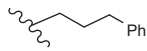
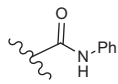
Scheme 6



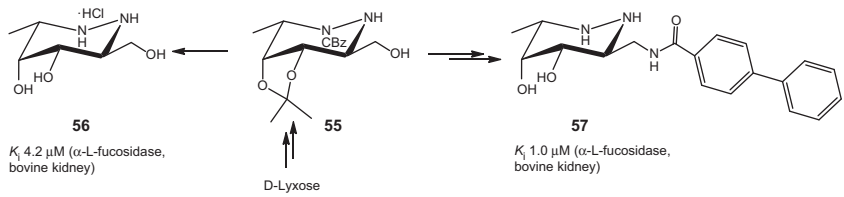
Scheme 7

Table 1 Inhibition profile for azafagomine **51**, *N*-alkylated azafagomines **46**, **48**, **50** and *N*-acylated azafagomine **54** (K_i in μM, pH 6.8).



Compound	R	α-Glucosidase	β-Glucosidase	α/β-Selectvity
51	H	6.90	0.32	22
46		278	0.055	5054
48		169	0.140	1141
50		158	0.032	4938
54		3.36[a]	14.7[a]	0.23

[a] pH 7.0.



Scheme 8

Removal of the chiral auxiliary from cycloadduct **52** (obtained from 4-phenyl-1,2,4-triazole-3,5-dione and the corresponding 2,4-pentadienoate bearing a tetra-*O*-acetyl-β-D-glucopyranosyl moiety), followed by diastereoselective epoxidation, opening of the oxirane and reductive cleavage with $LiAlH_4$ afforded phenyl carboxamide **54** (Scheme 7).

Inhibition assays of **54** revealed[43] (Table 1) that, unlike the 2-*N*-alkylated analogues **45**–**50**, acylation of the same position afforded an improvement of the inhibition against α-glucosidase ($K_i = 3.36$ μM), and an impairment against β-glucosidase as compared with parent azafagomine **51**. Binding of **54** with the active site of yeast α-glucosidase was studied by molecular modelling, suggesting an efficient packing of the carboxamido moiety on a hydrophobic subsite of the enzyme.

Moreno-Clavijo *et al.* have recently reported[45] the first examples of L-*fuco*-configured azafagomines bearing substituents at the C-3, considered as diaza-*C*-glycosyl analogues. A library of such compounds was accessed starting from D-lyxose (Scheme 8), using the reductive hydrazination of a 1-deoxy-ketohexose with *tert*-butyl carbazate as the key step, in a similar

way as in the enantioselective preparation of azafagomine by Bols *via* partially-protected reducing L-xylofuranose.[41] Compounds **56** and **57**, bearing either a hydroxymethyl moiety or a biphenyl-based amide linkage, exhibited potent and selective inhibition against α-L-fucosidase, within the same order of magnitude as parent unsubstituted L-*fuco*-azafagomine.[46]

4 Azepanes and azetidines

Although most imino sugars found in the literature are five- or six-membered ring compounds, some other different ring sizes have been designed and tested as potential glycosidase inhibitors.

Among them, unnatural polyhydroxyazepanes (seven-membered imino sugars) are the most remarkable examples; since pioneering work of Wong[47] in the mid-1990's, a great number of polyhydroxylated azepanes have been reported, including bicyclic derivatives.

One of the most relevant properties concerning azepanes as glycosidase inhibitors is their higher flexibility as compared to pyrrolidines and piperidines; azepanes can adopt more easily a quasi-flattened conformation, and thus might better interact with the active site of glycosidases.[47] Due to the fact that the synthesis of these compounds has been less developed than their five- and six-membered counterparts, the molecular requirements for predicting good inhibitory activities of azepanes still remain unknown.

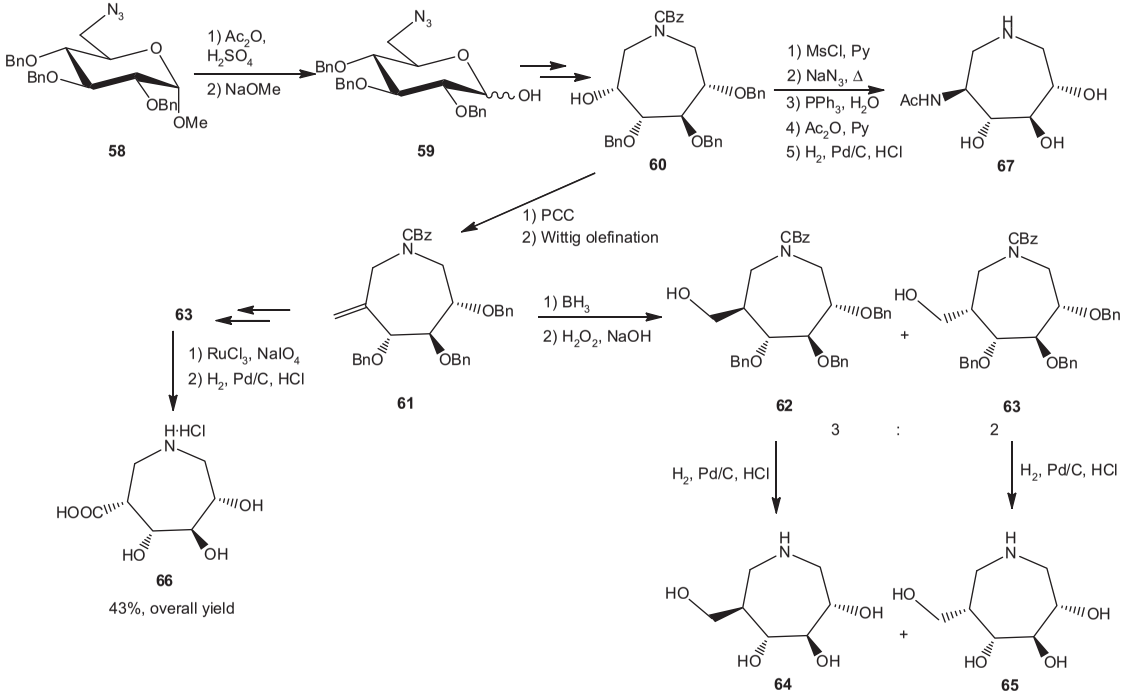
In this context, Blériot's group has accomplished the synthesis and evaluation of a large number of azepanes, by modifying the number of the hydroxyl groups, together with the configurations of the sugar residue, and the substituents on the ring.

For example, benzylated methyl 6-azido-6-deoxyglucopyranoside **58** was converted,[48] *via* a tandem Staudinger-aza-Wittig approach, (Scheme 9) into partially protected azepane **60**. In this approach, the key step is the treatment of azido lactol **59** with triphenylphosphane, followed by intramolecular reaction of the transient iminophosphorane with the latent aldehyde to give an imine which upon reduction with $NaBH_3CN$ furnished the azepane architecture. β-Hydroxyazepane **60** was an efficient building block in the preparation of three different families of seven-membered ring imino sugars bearing either hydroxymethyl, acetamido or carboxy moieties located at the β position to the endocyclic nitrogen (Scheme 9); such compounds are considered as ring homologues of 1-*N*-imino sugars.

The hydroxymethyl moiety present in derivatives **64** and **65** was introduced by oxidation of the free hydroxyl group, followed by Wittig olefination and hydroboration-oxidation process (Scheme 9). Deprotection of C-6 epimers **62** and **63** furnished compounds **64** and **65**.

Azepanes **64** and **65** were found to be rather potent and selective uncompetitive inhibitors (Table 2) of β-glucosidase ($K_i = 4\,\mu M$) and α-mannosidase ($K_i = 42\,\mu M$), respectively.

Blériot's group also attempted the synthesis of stable D-*manno*-configured polyhydroxylated azepanes, which may act as analogues of noeuromycin **138** (Fig. 10). For that purpose, the C-2 epimer of azidolactol **59** was used;[49] in this case, trihydroxylated D-*manno* and L-*gulo* derivatives, bearing a methyl group at the C-6 position were obtained and found to be


1) Ac2O, H2SO4
2) NaOMe
1) MsCl, Py
2) NaN3, Δ
3) PPh3, H2O
4) Ac2O, Py
5) H2, Pd/C, HCl
1) PCC
2) Wittig olefination
1) BH3
2) H2O2, NaOH
1) RuCl3, NaIO4
2) H2, Pd/C, HCl
H2, Pd/C, HCl
H2, Pd/C, HCl
58
59
60
67
61
62
63
3 : 2
66
43%, overall yield
64
65

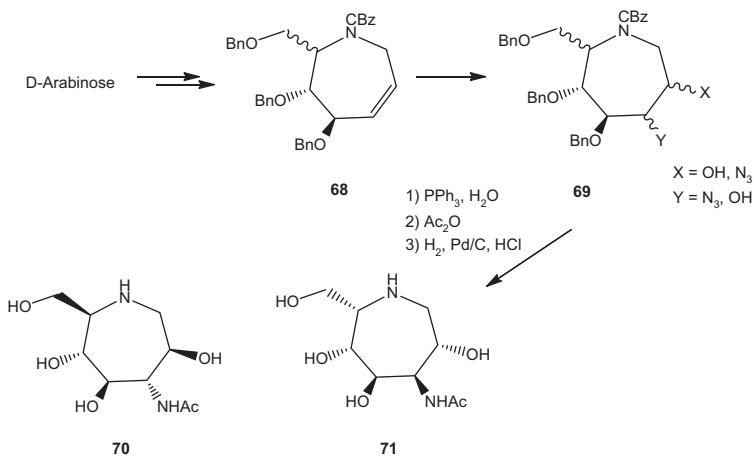
Scheme 9

potent although non-selective inhibitors of α-glucosidase, β-glucosidase and α-L-fucosidase, with a lower potency than the isomers depicted in Scheme 9.

Treatment of derivative **63** with $RuCl_3$ and $NaIO_4$, followed by deprotection, furnished derivative **66**, which incorporates a carboxylic acid moiety; this compound was unexpectedly found to be the most potent azepane-based inhibitor of α-L-fucosidase (41 nM, Table 2),[50] despite lacking both, the CH_3 group and the L-*fuco* configuration. These results clearly indicate how the conformational flexibility of seven-membered ring-based imino sugars might allow strong interactions with enzymes despite even lacking the appropriate configuration on the sugar residue.

Compound **60** was also used as the key intermediate for the preparation of azepanes bearing acetamido groups (Scheme 9); the free hydroxyl group of **60** was transformed into the acetamido moiety upon mesylation, nucleophilic displacement of the mesyloxy group by sodium azide, Staudinger reduction and acetylation of the corresponding amine to give **67**.[51] This compound, with the D-*gluco* configuration, is considered as a ring homologue of 2-acetamido-1-*N*-imino sugars, being a potent and selective inhibitor of β-*N*-acetylglucosaminidases, with K_i values in the submicromolar range (0.4 and 0.7 μM). Moreover, compound **67** was found[52] to be a strong competitive inhibitor of human *O*-GlcNAse ($K_i = 11$ μM) and *Bt*GH84 ($K_i = 8$μM). Crystal structure of **67** bound to *Bt*GH84 revealed a slightly distorted $^{2}B_{5,6}$ conformation, similar to rare ^{4}E and $^{1,4}B$ conformations adopted by complexes of GlcNAc-enzymes on mutant chitinases[53] and β-*N*-acetylglucosaminidases,[54] and quite different to those conformations found in solution ($^{4,5}C_{N,2}$ and $^{5,6}T_{2,3}$ in an almost 1:1 ratio).[52]

Blériot's group also reported[51] the preparation of polyhydroxylated azepanes **70** and **71**, compounds bearing both, an acetamido group at the γ- positions to the endocyclic nitrogen, and a hydroxymethyl moiety. Title compounds were obtained from D-arabinose, using azacycloheptene **68** as



Scheme 10

Table 2 Inhibition profile for azepanes **64**–**67**, **70**, **71** (K_i in μM or inhibition %).

Compound	β-Glucosidase (almonds)	α-Mannosidase (*jack bean*)	α-L-fucosidase (bovine epididymis)	β-*N*-Acetilglucos-aminidase (*jack bean*)	β-*N*-Acetilgluco-saminidase (*bovine liver*)
64	4	33%	N.I. (from bovine kidney)	N.I.	N.I.
65	Not tested	42	Not tested	Not tested	Not tested
66	N.I.	N.I.	0.041	Not tested	Not tested
67	Not tested	N.I.	28%	0.4	0.7
70	Not tested	25%	N.I.	0.5	0.6
71	Not tested	N.I.	N.I.	0.05	0.4

N.I. – no inhibition.

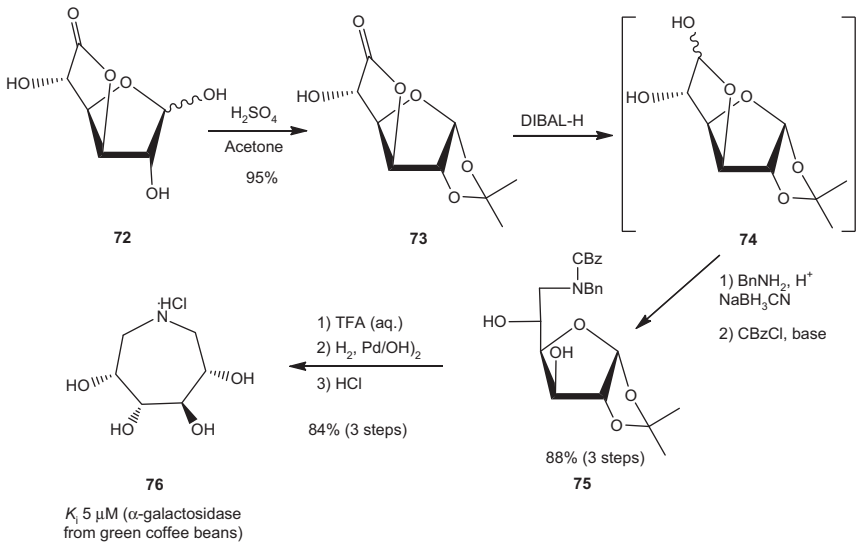
the key intermediate; such derivative was transformed into **69** by epoxidation, followed by azide opening of the oxirane ring (Scheme 10). Access to **70** and **71** was accomplished by Staudinger reduction of the azido group, followed by chemoselective acetylation and deprotection. Azepanes **70** and **71** turned out to be potent hexosaminidase inhibitors (Table 2).

In this case, unlike other related azepanes, there was a clear correlation between the relative configuration of the imino sugar and the target enzyme, as β-D-*gluco*-configured **70** was a much stronger inhibitor against β-*N*-acetyl glucosaminidases than α-L-*ido* and α-D-*manno* counterparts. Nevertheless, this observation was not fulfilled for β-L-*gulo* azepane **71**, which turned out to be an inhibitor ten times stronger than **70** against the same enzyme (Table 2).

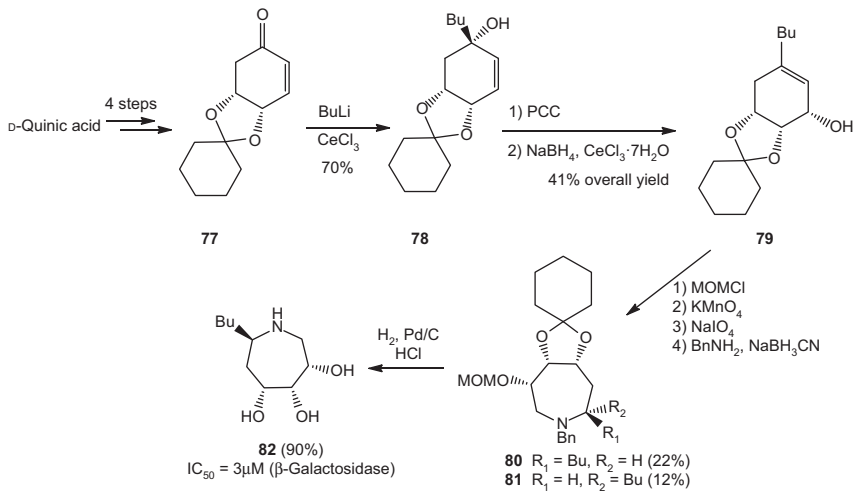
Dhavale's group recently reported[55] the efficient preparation of tri- and tetrahydroxyazepanes starting from inexpensive and readily available D-glucurono-6,3-lactone (**73**). In particular, Scheme 11 depicts the synthetic pathway for preparing tetrahydroxy derivative **76** upon DIBAL reduction of the lactone moiety, reductive amination, *N*-protection, intramolecular cyclization, and deprotection. Azepane **76** was screened against 7 glycosidases, turning out to be a potent and selective α-galactosidase inhibitor ($K_i = 5\,\mu M$).

D-Quinic acid was used by Shih *et al.*[56] as starting material in another procedure for preparing azepanes; Scheme 12 depicts the preparation of trihydroxylated azepane **82**, bearing a *n*-butyl residue at the C-7 (7*R* configuration).[57] Enone **77**, obtained in four steps from D-quinic acid was subjected to 1,2-addition with butyl lithium promoted by $CeCl_3$, giving compound **78** as the major diastereoisomer.

Derivative **78** underwent allylic oxidation with PCC, followed by Luche reduction to afford allylic alcohol **79** with complete stereoselectivity. Protection of the free hydroxyl group, *cis*-dihydroxylation of the double bond, oxidative cleavage and aminocyclization furnished the azepane architecture. Separation of epimeric **80** and **81**, followed by deprotection gave **82**, which turned out to be a potent β-galactosidase inhibitor ($K_i = 3\,\mu M$).

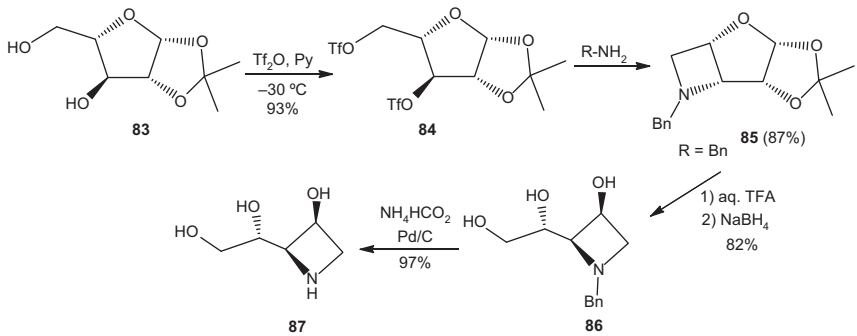


Scheme 11



Scheme 12

Up to date, reports concerning the synthesis of hydroxylated azetidines are scarce; the first search for biological activity of these carbohydrate mimics was reported by Jäger[58] in the late 1990's. More recently, Fleet's group has developed[59] an efficient procedure for preparing azetidine-based imino sugars which behaved from moderate to good inhibitors of several glycosidases. The key building blocks are stable 3,5-di-*O*-triflates derived from pentoses, which undergo a double nucleophilic substitution reaction with amines to afford the azetidine ring. For example, azetidine **87** was found to be a good inhibitor of non-mammalian amyloglucosidase, α-glucosidase and β-galactosidase (Scheme 13).



Scheme 13

5 Bicyclic imino sugars

5.1 Indolizidine, pyrrolizidine, quinolizidine and nortropane analogues

Up to date, many bicyclic imino sugars have been isolated from nature[60] and tested as glycosidase inhibitors; some of them with a strong activity, in many occasions against α-glucosidases. It has been claimed that probably their intrinsic rigid structure is responsible for the biological activity exerted by many of these carbohydrate mimics.[61,62]

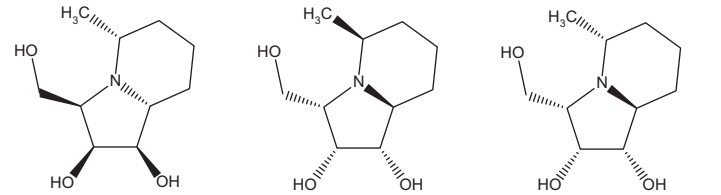
Compounds like lentiginosine, swainsonine, and castanospermine (di, tri and tetrahydroxylated indolizidines), casuarine, and hyacinthacines (polyhydroxylated pyrrolizidines), and quinolizidines are remarkable examples.[1]

A novel indolizidine alkaloid, (−)-steviamine **88**, has recently been isolated from the leaf material of *Stevia rebaudiana*.[63] Its enantiomer **89** and 5-*epi*-steviamine **90** (Fig. 6) have been synthesized from a D-ribose-derived cyclic nitrone.[64] (−)-Steviamine showed weak inhibition of β-glucosidases, but good activity against β-galactosidase ($IC_{50} = 35\,\mu M$), and no inhibition of either α-mannosidase or α-galactosidase, despite its structural resemblance with swainsonine. Interestingly, **88** also inhibited α-galactosaminidase (GalNAcase); although the activity for this latter enzyme was found to be moderate, it was the first case of a natural product inhibiting a GalNAcase.[64]

Despite the potency exhibited by some of the natural imino sugars like deoxynojirimycin and castanospermine, they frequently lack selectivity. For instance, these compounds are potent inhibitors of α-glucosidases at the endoplasmic reticulum, but they also bind lysosomal α- and β-glucosidases and the intestinal glucoamylase, a real drawback in their potential application as therapeutic agents.

The increase of the sp^2-character of the endocyclic nitrogen atom has become an alternative to improve the anomeric selectivity of the corresponding sugar mimic. In this context, Cipolla *et al.* reported[62] the synthesis of novel nojirimycin derivatives containing cyclic carbamates, ureas, and guanidines; nevertheless, the observed inhibitory activities were moderate.

Thiohydantoin-castanospermine analogues, where the pyrrolidine ring of castanospermine is replaced by a thiohydantoin scaffold, were recently prepared as a new family of sp^2-imino sugar-type glycomimetics.[65] It was

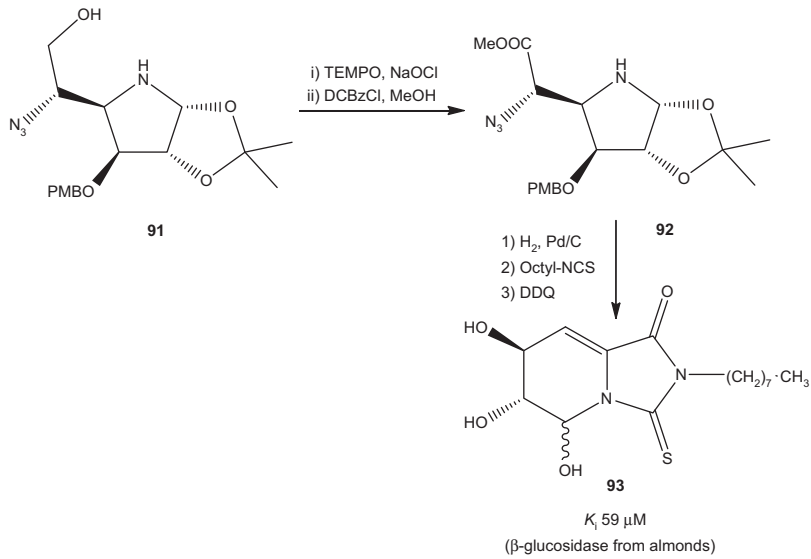


88 (–)-Steviamine
IC_{50} 35 μM (rat β-galactosidase)

89 (+)-Steviamine
IC_{50} 484 μM (α-rhamnosidase)

90 5-*epi*-(+)-Steviamine
IC_{50} 342 μM (α-rhamnosidase)

Fig. 6



91

92

93

K_i 59 μM
(β-glucosidase from almonds)

Scheme 14

claimed that delocalization of the nitrogen electron pair generates a partial positive charge on the nitrogen that might improve mimicry of the transition state of the enzyme-mediated glycoside hydrolysis.[66] Scheme 14 depicts the synthetic pathway for accessing **93**, an example of a thiohydantoin-castanospermine hybrid structure. Precursor **91** was converted into the α-azido ester **92**, followed by subsequent azido reduction, coupling with an alkyl (*e.g.* octyl) isothiocyanate and deprotection to afford glycomimetic **93**. Title compound was found to be a selective and good inhibitor of β-glucosidase.

In a similar fashion, other sp^2-imino sugars resembling castanospermine were prepared; bicyclic iso(thio)urea castanospermine analogues **94–97** (Fig. 7) represent good examples of this approach.[67] Such compounds were accessed starting from 3-*O*-acetyl-5-azido-5-deoxy-1,2-*O*-isopropylidene-α-D-galactofuranose; the key intermediate was a carbodiimide, obtained by treatment of the parent thiourea with HgO.

Regarding compounds **94–97**, the shift of their inhibitory activities from α- to β-glucosidase is remarkable, when compared to castanospermine; such data are even more surprising considering that, due to a strong anomeric effect, derivatives **94–97** exist only as the α-anomers.

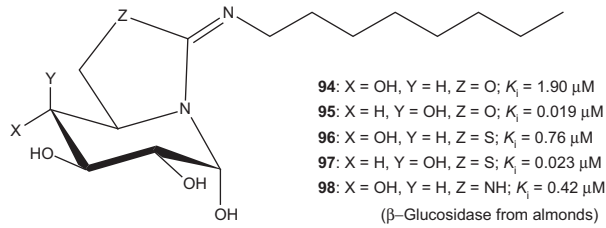


Fig. 7

Furthermore, both, iminooxazolidines **94**, **95** and iminothiazolidines **96**, **97**, with pK_a values of virtually 7.0, were found to be more potent inhibitors at neutral pH than at acidic pH; this fact was explained considering a better interaction of partially protonated imino sugar (roughly 50% at pH 7.3) with fully deprotonated catalytic system at the active site of the enzyme.[67] Thermodynamic studies revealed an entropically-driving binding process with β-glucosidase from *Thermotoga maritime* (*Tm*GH1), quite an uncommon situation.

Additionally, Brumshtein *et al.*[68] accomplished the synthesis of **98** (Fig. 7); this compound incorporated a strongly basic residue (guanidino group) on the castanospermine-based template, in order to study potential favorable coulombic interactions with the active site of glycosidases.

The synthesis started from 5-azido-5-deoxy-1,2-*O*-isopropylidene-α-D-glucofuranose (**99**) and was based on the introduction of a second azido group at the C-6; reduction of both azido moieties yielded diamine **100**, which was transformed into imidazolidine-2-thione **101** by thiocarbonylation with CS_2 and then into the *S*-methyl isothiouronium salt **102**; subsequent nucleophilic displacement of the methylthio group with *n*-octylamine and deprotection gave target compound **98** (Scheme 15).

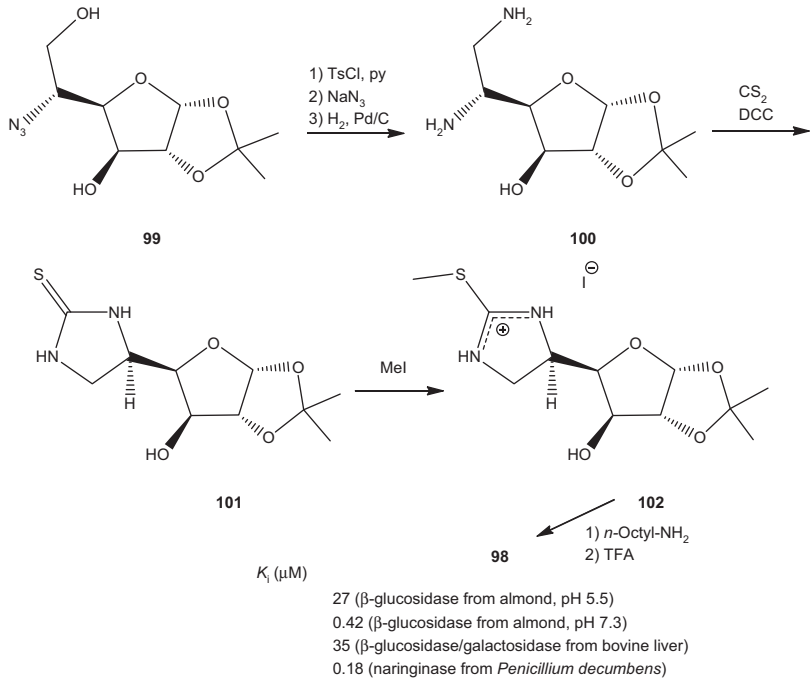
Like derivatives **94**–**97**, compound **98** also showed a high selectivity for β-glucosidase.

Another approach investigated by the same group to improve the selectivity and permeability of imino sugars was the incorporation of axially-anchored alkyl chains at the pseudoanomeric position of carbamate-derived castanospermines (Fig. 8). With that purpose, 5-*N*-octyl (**104**), 5-*S*-octyl (**105**) and 5-*C*-octyl (**106**) derivatives of sp^2-imino sugar-type castanospermine **103** have been prepared and tested.[69]

Title compounds were found to be selective and potent inhibitors of neutral α-glucosidase, and remarkably, none of them was active against the human lysosomal acid α-glucosidase. Encouraging preliminary results on breast cancer cell (MCF-7 cell line) suggested their potential use as antiproliferative agents.

Concerning polyhydroxylated quinolizidines, Pandey *et al.* described[70] a divergent route starting from D-ribose (Scheme 16 displays the retrosynthetic analysis) to get a small library of quinolizidine alkaloids (**111**–**123**). Among them, two compounds had remarkable inhibitory properties: **112**, against α-galactosidase from coffee beans ($K_i = 83.9\,\mu M$), and **123**, against α-glucosidase from yeast ($K_i = 28\,\mu M$).

Among polyhydroxylated pyrrolizidines, casuarine derivatives have become the most investigated structures in this area in the last few years.



Scheme 15

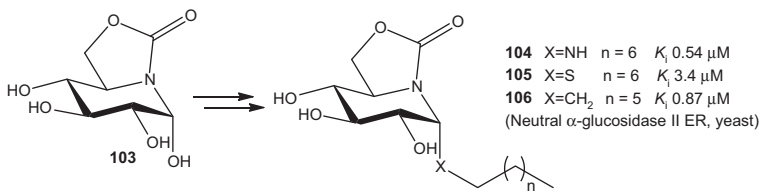
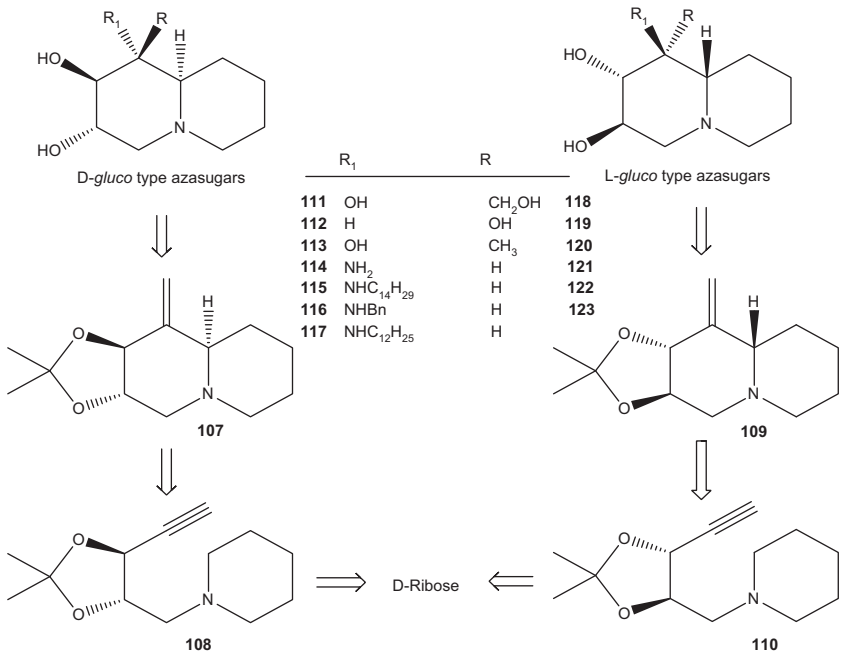


Fig. 8

Casuarine **128** and its 6-*O*-α-glucoside **129** were isolated from the bark of *Casuarina equisetifolia*. Both compounds are potent inhibitors of fungal glucoamylase from *Aspergillus niger* ($IC_{50} = 0.7$ and 1.1 μM, respectively); furthermore, they were recently found to act as strong and selective inhibitors of human *N*-terminal subunit maltase-glucoamylase (NtMGAM), with K_i values of 0.45 and 280 μM respectively.[71] This enzyme is involved in starch digestion, so inhibitors may potentially be used in the treatment of diabetes.

Total syntheses of **128**, **129**[71] (Scheme 17) and other casuarine analogues (**130** and **132**), together with their 6-*O*-α-glucosides (**131** and **133**)[72] have been reported; such syntheses involved stereoselective 1,3-dipolar cycloadditions of a chiral cyclic nitrone and an alkene, Tamao-Fleming oxidation and regioselective glucosylation. Their biological evaluation demonstrated that casuarine derivatives bind tightly and selectively to glucoamylase from *Aspergillus niger*.



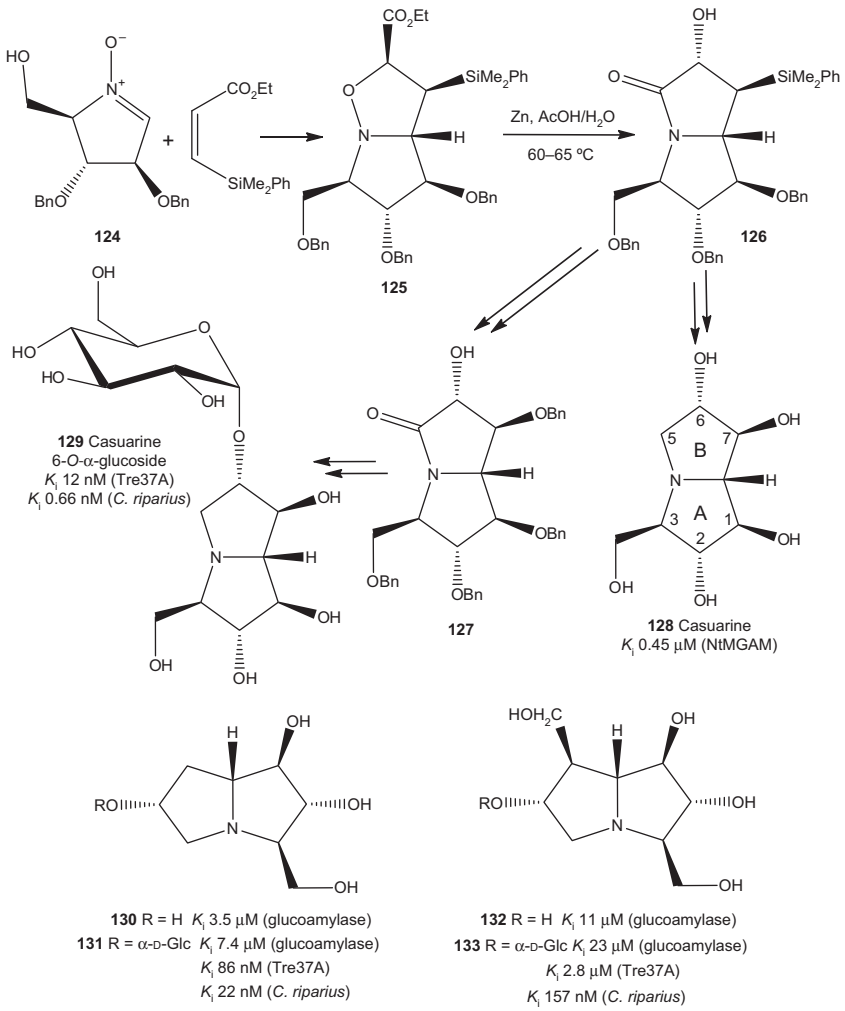
Scheme 16

Casuarine **128** and specially glucoside **129** were also found to be very strong trehalase inhibitors;[71] trehalase is the enzyme that catalyzes breaking down of trehalose, the vital carbohydrate source for insects. Therefore, trehalase inhibitors are claimed to be potential insecticidal agents.[73]

Natural (+)-hyacinthacine A_1 **134** (Fig. 9) was isolated from *Muscari ormeniacum* bulbs.[74] Since then, several syntheses of this compound have been reported[75] due to its good inhibitory properties against β-galactosidase ($IC_{50} = 4.4\,\mu M$), α-L-fucosidase ($IC_{50} = 46\,\mu M$) and amyloglucosidase ($IC_{50} = 2.3\,\mu M$).[74] D'Adamio *et al.*[76] achieved the total synthesis of **134**, (+)-7a-*epi*-hyacinthacine A_1 **135** (Fig. 9), and their 6-hydroxy analogues, like **136**. The synthesis involved a nitrone cycloaddition strategy, with a D-ribose-derived cyclic nitrone as the dipole, and *tert*-butyl acrylate as the dipolarophile.

Another type of relevant bicyclic imino sugars are calystegines; these natural polyhydroxy-nor-tropane alkaloids (named after *Calystegia sepium*) combine both, a pyperidine and pyrrolidine ring in their structure, and are potent and quite selective β-glucosidase inhibitors.[77] This fairly good inhibitory profile results from good mimicry of the transition state of the substrate hydrolysis. It has been observed that, upon interaction with glucosidases, calystegines adopt a conformation that places the endocyclic nitrogen in the same position as the *O*-5 of the glycosides, the natural enzymatic substrates.[78,79]

Further modifications of natural calystegines may provide interesting inhibitors. In this context, Rasmussen and Jensen envisioned noeurostegine **139** as a new stable hemiaminal with a nor-tropane motif,[77] and a hybrid structure between noeuromycin **138**[80] and calystegine B_2 **137** (Fig. 10).



Scheme 17

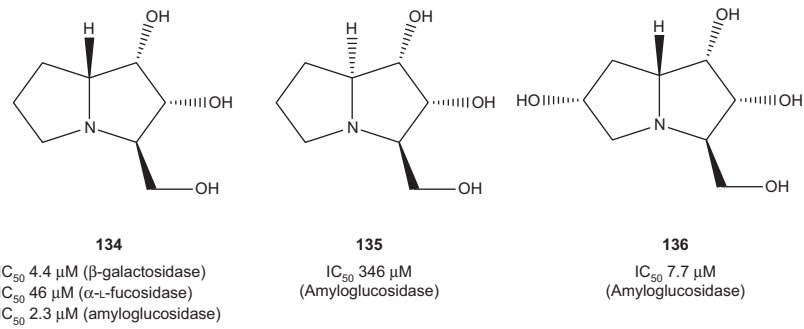


Fig. 9

Noeuromycin is a strong glycosidase inhibitor, but it incorporates an unstable hemi-aminal functionality. Thus, noeurostegine actually combine the stability of calystegines and the potency of noeuromycin against β-glucosidase, while exhibiting a much higher selectivity; derivative **139** is virtually non-active against α-glucosidase, and much less active against other glycosidases.[77]

The synthesis of neurostegine is depicted in Scheme 18; compound **139** was obtained in 22 steps starting from levoglucosan; one of the key steps was the transformation of nonadiene **143** into cycloheptene **144** *via* a ring closing metathesis reaction.

Alternatively, compounds with the structure of 6-oxa-calystegine, where the nitrogen atom has a sp^2-character, were postulated as a good approach to get potent and selective glycosidase inhibitors, in a similar fashion as reported for castanospermine.

For that purpose, thioureas **140** and **141** (6-oxa-*N*-thiocarbamoyl-calystegines, Fig. 10),[81] were prepared from 5-amino-5-deoxy-1,2-*O*-isopropylidene-β-L-idofuranose **146** (Scheme 19). The synthetic pathway involved coupling of **146** with commercially-available *n*-butyl and *n*-octyl isothiocyanates, followed by deprotection of the intermediate thioureas **147** and **148**. 6-Oxa-calystegines **140** and **141** were found to be potent and selective inhibitors of β-glucosidase (*Thermotoga maritima*), with K_i values of 1.1 μM and 0.28 μM, respectively. *n*-Octyl derivative was also quite active against human β-glucocerebrosidase ($K_i = 2.2$ μM), a tenfold better than **140**, demonstrating again that increased lipophilicity considerably improves the potency and selectivity of certain imino sugars. Regarding compound **141**, with a more lipophilic structure, thermodynamic data indicated favorable enthalpic and entropic contributions to binding, probably because such compound displaces more water molecules from the enzyme active site.[81]

Chagnault and co-workers accomplished[82] preparation of 10-azabicyclo[4.3.1]decanes, a different family of bicyclic imino sugars. The best inhibitory profile was observed for the pentahydroxylated nortropane **149** (Fig. 11) against β-glucocerebrosidase ($IC_{50} = 1.2$ μM) and β-glucosidase ($IC_{50} = 2.0$ μM).

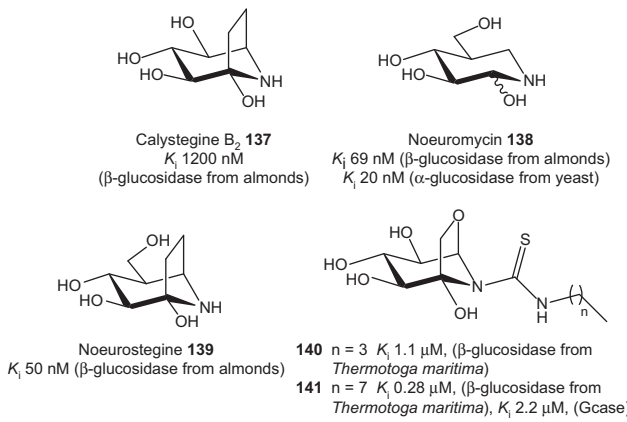
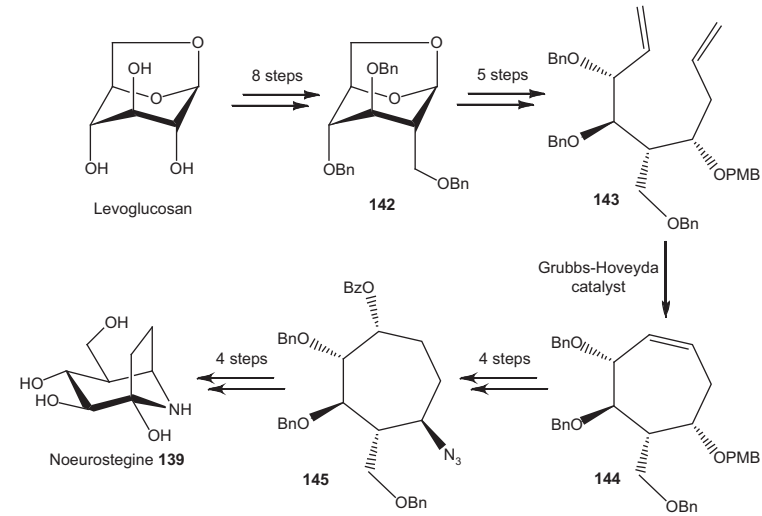
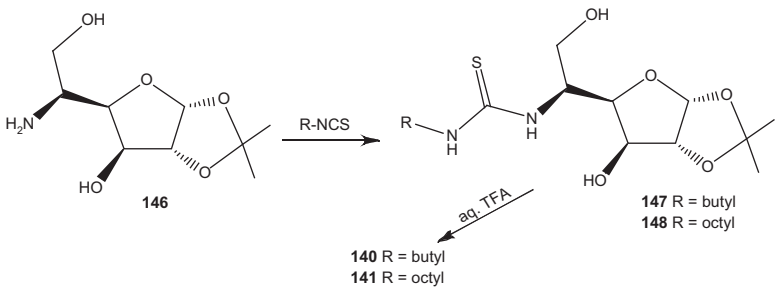


Fig. 10



Scheme 18



Scheme 19

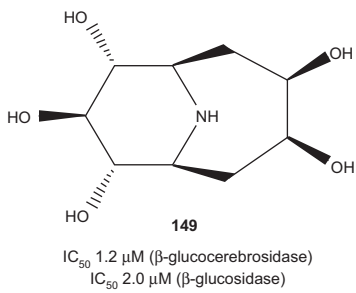


Fig. 11

5.2 Thiazolines

In the mid-1990's, Knapp's and Vocadlo's groups reported the first examples of bicyclic thiazolines as good inhibitors of hexosaminidases, a family of enzymes that exert their catalytic activity trough oxazolinium ion **151** (Fig. 12). Bicyclic thiazoline **153** (GlcNAc thiazoline) was accessed by

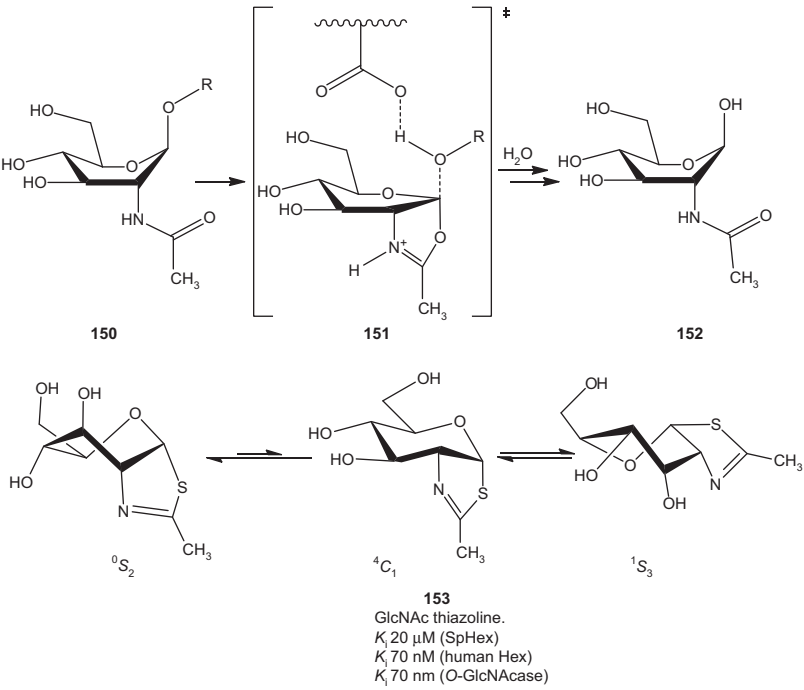
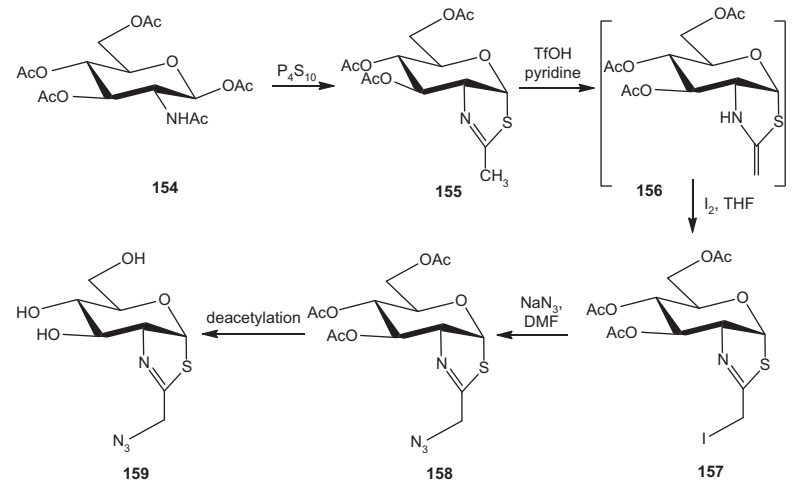


Fig. 12

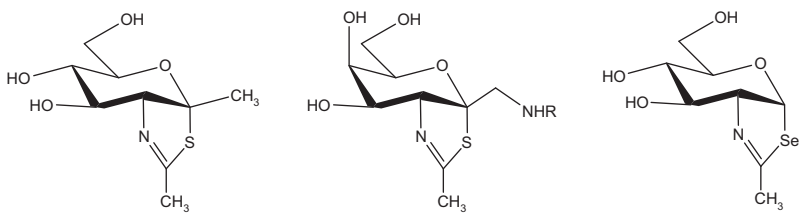
intramolecular cyclization involving the anomeric carbon atom and a thioamide scaffold located at the C-2 of per-*O*-acylated *N*-acetil glucosamine.[83] This compound is a potent inhibitor of retaining *N*-acetylhexosaminidases, for which it is considered as a good transition state analogue;[84] recent conformational studies conducted on GlcNAc thiazoline suggested that **153** binds the active site of the enzyme in a distorted $^4C_1$ chair conformation, the conformation that most closely resembles transition state **151**. NMR data and theoretical calculations demonstrated that, in solution, compound **153** can move among one chair conformation ($^4C_1$) and two skew boat conformations ($^1S_3$ and $^0S_2$, the latter being the most populated one).[85]

Modification of **153**, involving imine-to-enamine tautomeric processes, followed by electrophilic or radical additions (*e.g.* of thiols) has been reported (Scheme 20). Using this methodology, compound **159** was prepared and turned out to possess an excellent selectivity for human *O*-GlcNAcases, and to be almost as potent as parent **153**.[86] Moreover, chitobiose and chitotriose thiazolines, in which the corresponding oligosaccharide bears a GlcNAc thiazoline residue at the terminal unit, have recently been reported as excellent chitinase inhibitors of Chi18A from *Serratia marcescens* ($K_i = 25$ and 0.25 μM respectively).[87]

Other derivatives of GlcNAc and GalNAc thiazolines are the C-1 homologated analogues **160**–**162** (Fig. 13). The inhibition of these compounds against *N*-acetyl-β-hexosaminidases from *Streptomyces plicatus* (SpHex), falls dramatically when increasing the size of the C-1 substituent. This observation suggests strong steric restrictions around the anomeric position for binding the enzyme.[88]



Scheme 20



160
K_i 200 μM (SpHex)

161 R=H K_i >2 mM (SpHex)
162 R=Boc K_i >2 mM (SpHex)

163
GlcNAc selenazoline
K_i 170 μM (SpHex)
K_i 0.7 μM (GlcNAcase)

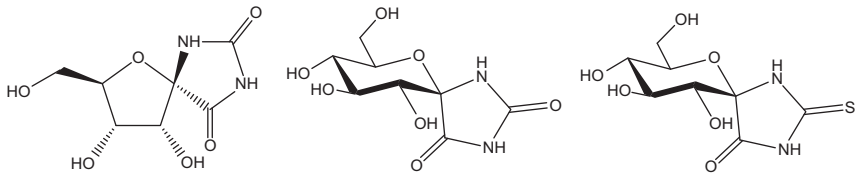
Fig. 13

Replacement of the sulfur atom of thiazoline **153** with selenium led to GlcNAc selenazoline **163**, showing much weaker inhibition of hexosaminidase than **153** (Figs. 12 and 13 for data).[89]

5.3 Carbohydrate-derived inhibitors bearing spiranic motifs

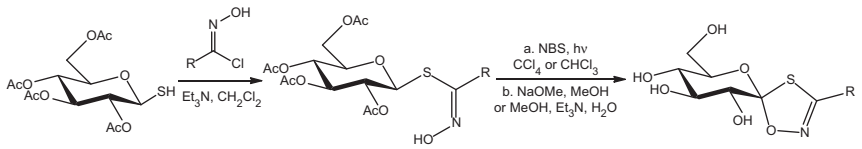
It has been demonstrated that in many occasions there is a close correlation between reduced conformational flexibility and a potent interaction with biomolecules.[90] In this context, natural hydantocidin **164** (Fig. 14), the first spiro compound isolated from nature with a peculiar spiranic hydantoin fragment, showed to be a plant growth regulator and a potent herbicide, without exerting toxicity to humans.[91] The synthesis of C-1 spironucleosides was inspired by hydantocidin **164**.

One of the most targeted enzymes for the developing of drugs has been glycogen phosphorylase (GP), a key enzyme for breaking down glycogen and thus, a therapeutic target in the treatment of type 2 diabetes. Although this carbohydrate-proccessing enzyme is not a glycosidase, it deserves some attention within this review due to its pharmacological importance.



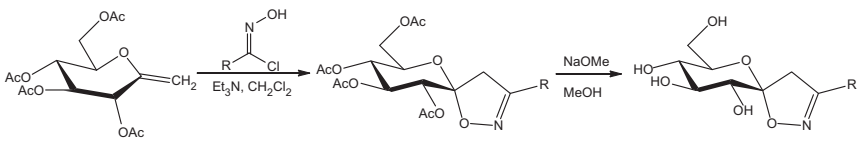
Hydantocidin **164** **165** **166**

Fig. 14



167 **168** **169**

Scheme 21



170 **171** **172**

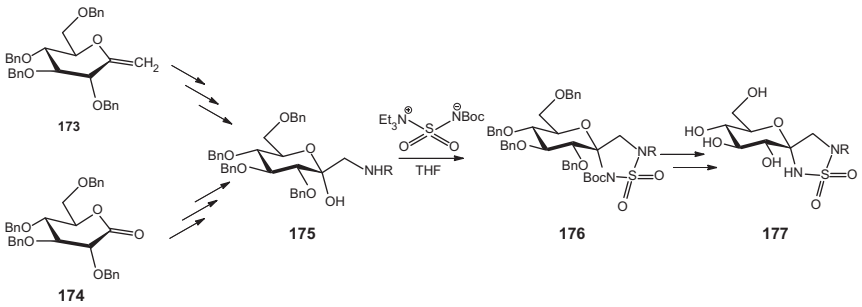
Scheme 22

A new type of GP inhibitors that emerged after the isolation of hydantocidin was D-glupopyranosylidene-spiro-hydantoin[92] **165** (Fig. 14), an extremely potent competitive GP inhibitor (Table 3), but with limited access due to a long synthetic pathway.[92,93]

Somsák *et al*. described[94] the more accessible D-glupopyranosylidene-spiro-thiohydantoin analogue **166** (Fig. 14), which exhibited good inhibitory profile against GP (Table 3).[93] In a recent study,[95] it was observed that **166** caused the inhibition of hepatic glycogen breakdown in rats, which resulted in a rapid decrease of the concentration of glucose in blood.

Crystallographic investigations of GP-inhibitor complexes have recently been accomplished in order to get insight into mechanistic aspects of inhibition.[96] Derivatives like **165** and **166** allowed the development of a series of new spiranic inhibitors.[97] In this context, glycopyranosylidene-spiro-oxathiazoles **169**, bearing aromatic residues (remarkable examples are depicted in Table 3) are noticeable. This family of compounds was obtained[98] by reaction of glucopyranosyl-hydroximothioate derivatives **168** with *N*-bromosuccinimide (NBS) (Scheme 21). Compounds **169** showed from modest to excellent inhibition of muscle GPb (Table 3).

Glucopyranosylidene-spiro-isoxazolines **172** were synthesized from the *exo*-glucal **170** *via* selective 1,3-dipolar cycloadition with a series of aromatic nitrile oxides, generated *in situ* (Scheme 22).[99] Such spiro-isoxazolines showed moderate to good inhibition of muscle GPb (Table 3);



Scheme 23

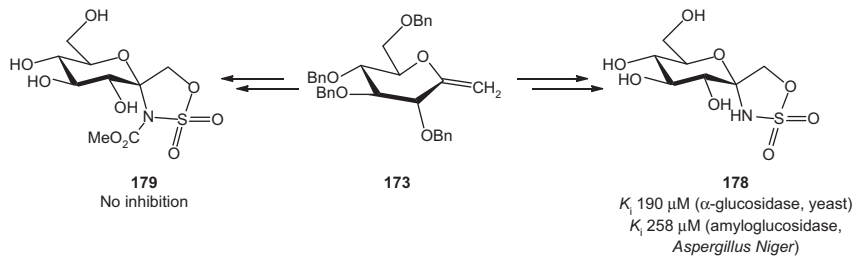


Fig. 15

crystallographic data showed binding of all derivatives to the catalytic site of the enzyme, *via* direct and water-mediated hydrogen bonds.

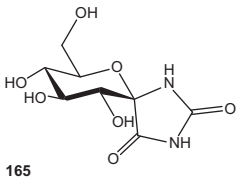
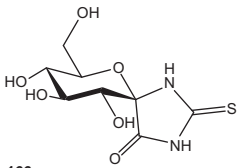
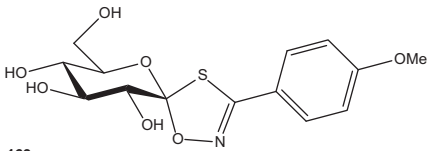
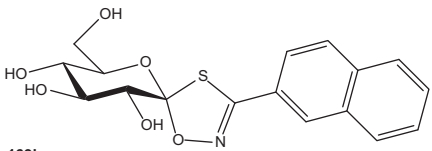
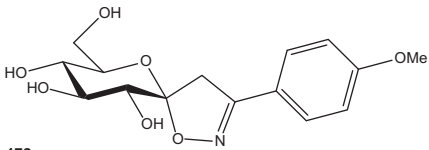
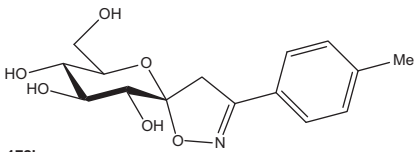
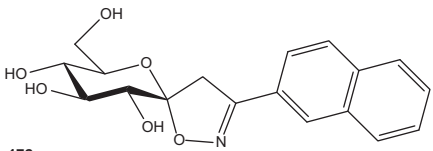
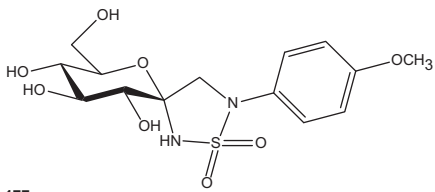
Some glycomimetics that can be obtained from *exo*-glycals are *C*-glycosylated spiro-sulfamides[100] and spiro-sulfamidate glycosides.[101] The spiranic sulfamides were obtained starting either from *exo*-glucal **173**, or from glucoronolactone **174**, which were converted into β-aminoalcohols **175** (Scheme 23); treatment of **175** with a Burgess-type reagent furnished the spiro-sulfamides **176** (Scheme 23).[100] The *p*-anisyl derivative **177a** showed good inhibition against muscle GPb, with a K_i of 15 μM (Table 3).

A similar procedure was used in the synthesis of spiro-sulfamidates (Fig. 15).[101] Compound **178** was tested against α-glucosidase ($K_i = 190\,\mu M$) and amyloglucosidase ($K_i = 258\,\mu M$); such compound turned out to be only a moderate inhibitor of both enzymes. Nevertheless, no inhibition was observed for **179** against any of the tested enzymes, suggesting that the anomeric NH group is essential for the glycosidase inhibition.

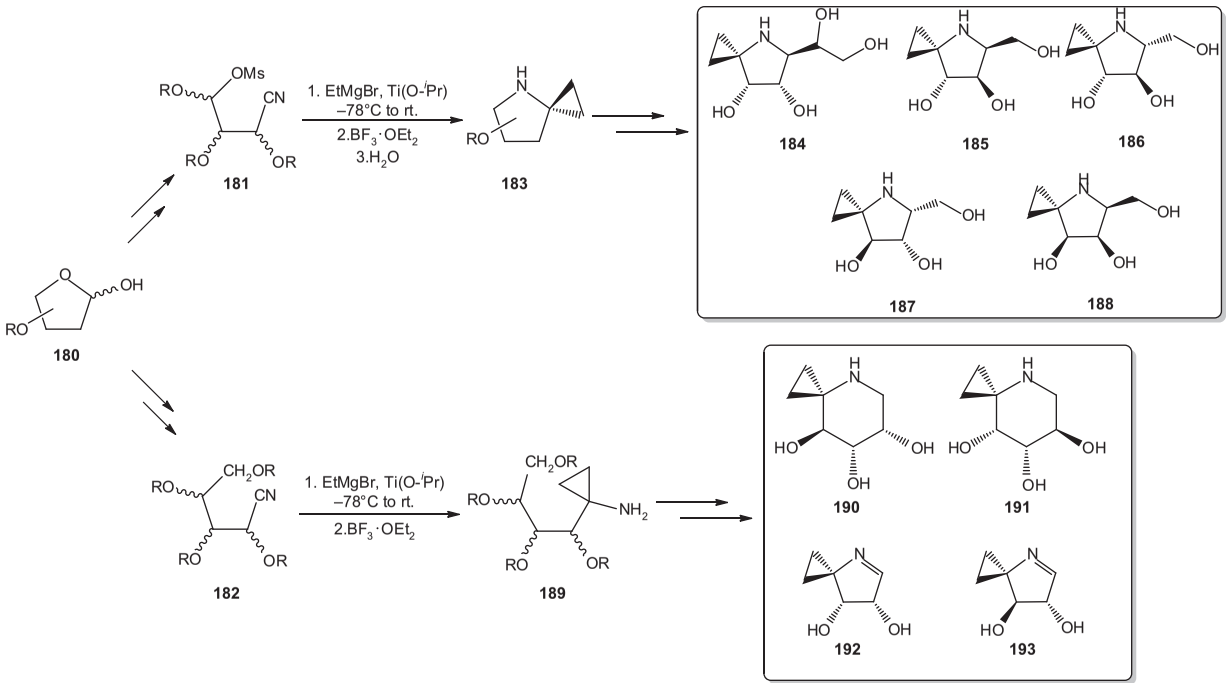
Among all the GP inhibitors described in this chapter, the 2-napthyl oxathiazol **169b** and 2-napthyl isoxazoline **172c** turned out to be the two most potent inhibitors of muscle GPb up to date (Table 3), with inhibition constants in the nanomolar range. X-Ray crystallographic studies demonstrated that the 2-naphtyl moiety interacts with several residues of the enzyme in its less active T-state conformation.[99]

A series of D-*arabino*, L-*fuco* and L-*rhamno*-configured imino sugars bearing a spirocyclopropyl fragment have been described by Behr's group.[102–104] The synthetic pathway involved the formation of an oxime, followed by subsequent dehydration reaction to afford nitriles **181** or **182** (Scheme 24); the next key steps are a titanium-mediated

Table 3 Inhibition of GP (K_i in μM).

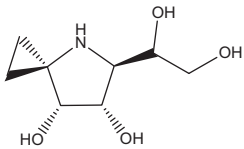
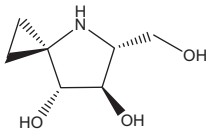
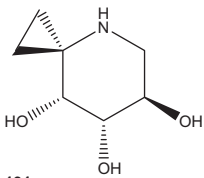
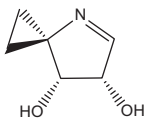
Compound	K_i muscle GPb	K_i liver GPb
 **165**	4.2	12.8
 **166**	5.1	7.0
 **169a**	8.2	—[a]
 **169b**	0.16	—[a]
 **172a**	6.6	—[a]
 **172b**	7.9	—[a]
 **172c**	0.63	—[a]
 **177a**	15	—[a]

[a] Value not available or not tested.


OMs
RO
CN
RO
OR
181
1. EtMgBr, Ti(O-iPr)
–78°C to rt.
2.BF3·OEt2
3.H2O
183
184
185
186
187
188
180
CH2OR
NH2
182
189
190
191
192
193

Scheme 24

Table 4 Selected enzymatic inhibition data for imino sugars **184–188** and **190–193**.[a]

Compound	α-L-Fucosidase (bovine kidney)	α-L-Rhamnosidase (*Penicillium decumbens*)	β-Galactosidase (bovine liver)	Amylogluco-sidase (*Aspergillus Niger*)
 **184**	97% (1.6 μM)	—[b]	9%	N.I.[c]
 **186**	50%	—[b]	24%	82% IC_{50} 100 μM
 **191**	83% (18 μM)	N.I.[c]	—[b]	N.I.[c]
 **192**	76%	87% (57 μM)	—[b]	N.I.[c]

[a] Percentage of inhibition at 1 mM concentration, K_i values in brackets.
[b] Value not available or not tested.
[c] No inhibition detected at 1mM.

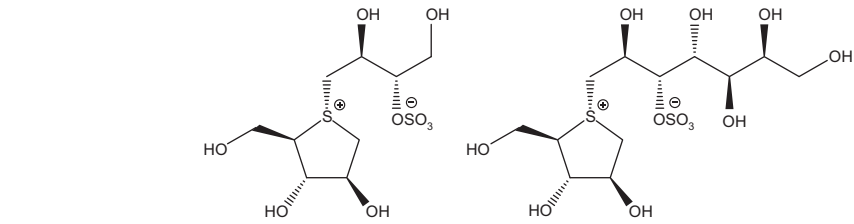
aminocyclopropanation involving the nitrile group and final intramolecular cyclization.

Imino sugars **184–188** and **190–193** were tested against a series of glycosidases, and the most significant results are depicted in Table 4. Compounds **184** and **191** showed to be potent and selective inhibitors of α-L-fucosidase ($K_i = 1.6\,\mu M$ and $18\,\mu M$ respectively), while on the contrary, the L-*fuco*-configured **192** turned out to be only a weak inhibitor of α-L-fucosidase, but exhibited a strong affinity for α-L-rhamnosidase ($K_i = 57$ μM). Curiously, *arabino*-configured **186** showed a moderate inhibition of amyloglucosidase ($IC_{50} = 100\,\mu M$), but was completely inactive against other α-glucosidases.[102]

6 Thio, seleno and carbasugars as glycosidase inhibitors

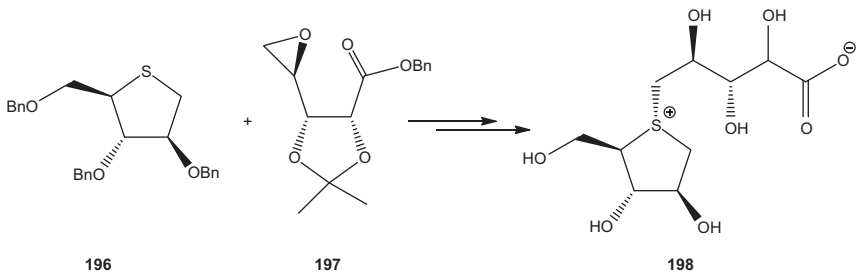
6.1 Thio and seleno sugars

Thio-sugars emerged as interesting scaffolds for glycosidases inhibitors[105] after the discovery of the natural glycosidase inhibitors salacinol[106] **194** and



α-glucosidases (Rat intestinal disaccharidase)	Salacinol **194**	Kotalanol **195**
Maltase	K_i 0.31 μM	K_i 0.23 μM
Succrase	K_i 0.32 μM	K_i 0.18 μM
Isomaltase	K_i 0.47 μM	K_i 1.8 μM

Fig. 16



196 **197** **198**

Scheme 25

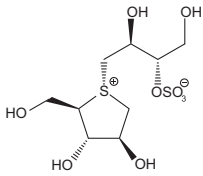
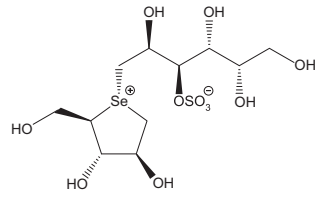
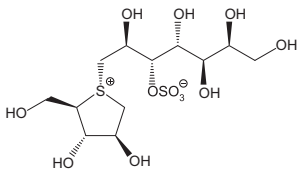
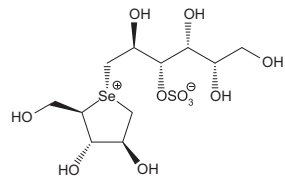
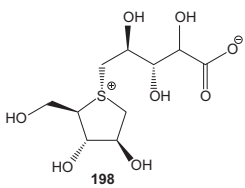
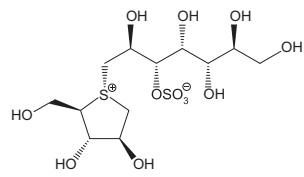
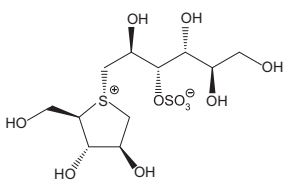
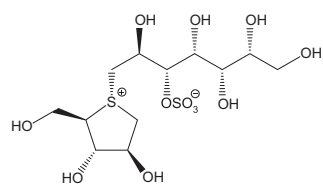
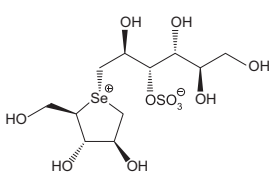
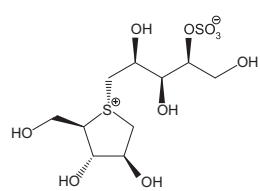
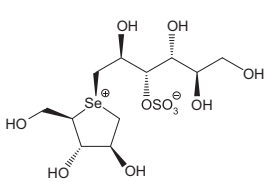
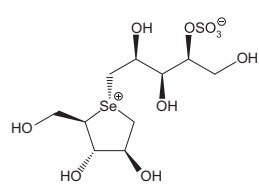
kotalanol **195**[107,108] (Fig. 16); both compounds were isolated from *Salacia reticulata* and have been used in the Indian Ayurvedic tradicional medicine. Salacinol and kotalanol share a common zwitterionic sulfonium-sulfate structure and a polyhydroxylated side chain. These compounds showed potent inhibition against a series of α-glucosidases (Fig. 16 and Table 5), a property that can explain the hypoglycemic activity[109] of *Salacia* species.

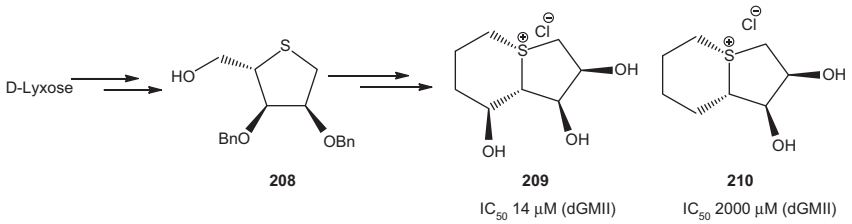
Analogues to salacinol have been synthesized by coupling a thio(seleno)sugar with alkylating species, like epoxide **197** (Scheme 25). In this case, such coupling afforded the zwitterionic derivative **198**,[110] a potent inhibitor of human maltase glucoamylase (MGA) (Table 5).

Following this strategy, several thio(seleno)sugars have been described by Pinto's group; these compounds showed potent inhibition of MGA (Table 5). From these data, it is clear that the presence of a sulfate group at C-3 of the side chain is essential for a good enzymatic affinity, demonstrated by the high K_i values of compounds **198**, **206** and **207**, which lack such group or bear it at a different position. The sulfate moiety present in some of the derivatives was introduced by nucleophilic attack of protected five-membered thio or selenosugars on 1,3,-cyclic sulfates.[111] Furthermore, a change in the configuration of the heteroatom was found to decrease the potency of the corresponding inhibitor (*e.g.* **200** *vs.* **201**).

Some bicyclic thiosugars related to swainsonine and 8-*epi*-lentiginosine have been synthesized from D-lyxose (Scheme 26).[116] The inhibitory

Table 5 Enzymatic inhibition of sulfonium-based thiosugars.

Compound	MGA (K_i in μM)	Compound	MGA (K_i in μM)
 Salacinol **194**	0.19[105]	 **202**	0.10[112]
 Kotalanol **195**	0.19[113]	 **203**	0.10[112]
 **198**	10[110]	 **204**	0.13[114]
 **199**	0.65[111]	 **205**	0.10[114]
 **200**	0.14[111]	 **206**	20[115]
 **201**	10.1[111]	 **207**	53[115]



Scheme 26

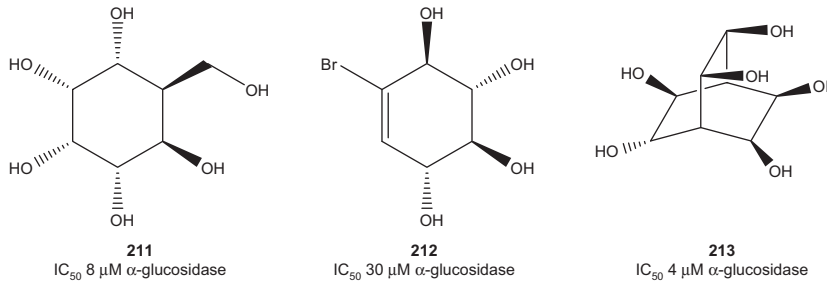


Fig. 17

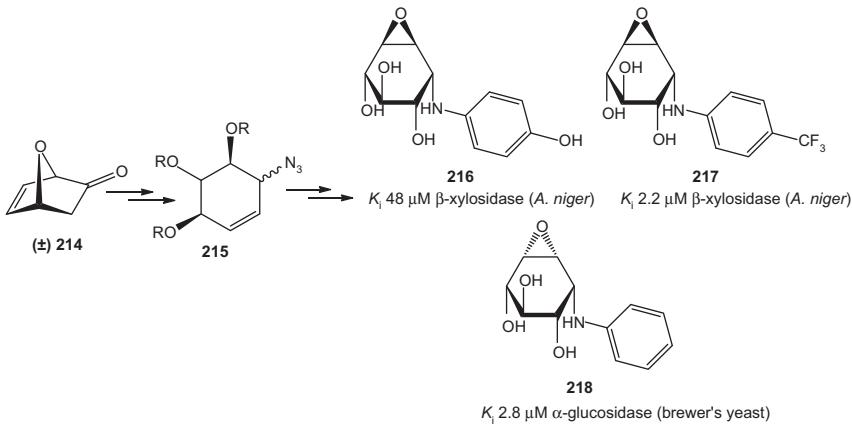
activities of 8-*epi*-thiolentiginosine **209** and di-*epi*-thioswainsonine **210** against *Drosophila* Golgi α-mannosidase II (dGMII) were tested. Compound **209** showed a much stronger inhibition ($IC_{50} = 14\,\mu M$) than derivative **210** ($IC_{50} = 2000\,\mu M$).

6.2 Carbasugars

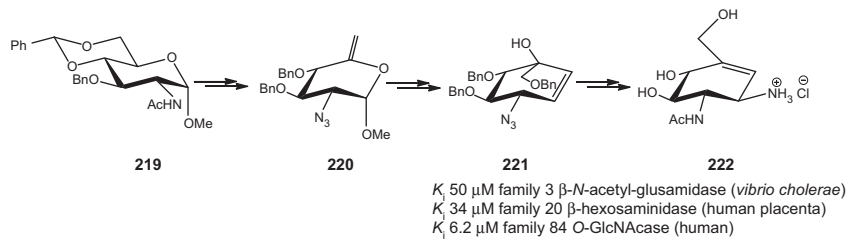
In the literature, the synthesis of carbasugars is considered as a useful source of glycosidase inhibitors,[117] like polihydroxylated derivatives **211**–**213** (Fig. 17).[118–120]

Conduramine F-1 epoxides **216**–**218** (Scheme 27)[121] were synthesized starting from (±)-7-oxanorbornenone **214**, using cyclohexene derivative **215** as the key intermediate, which was subjected to an epoxidation reaction. All compounds behaved as poor inhibitors against β-D-galactosidase (bovine liver) and amyloglucosidase (*Aspergillus niger*); however compounds **216** and **217**, bearing *trans* epoxy and amino moieties, exhibited potent non-competitive inhibition of β-D-xylosidase (*Aspergillus niger*; $K_i = 48\,\mu M$ and $2.2\,\mu M$, respectively). On the other hand, compound **218**, with a *cis* arrangement of the epoxy and amino functionalities, turned out to be a non competitive strong inhibitor of α-glucosidase (brewer's yeast) with a K_i value of $2.8\,\mu M$. It is also remarkable that non-substituted derivatives on the nitrogen atom were found to be quite less active.

The structure of the well known carbasugar valienamine has also been an inspiration for the development of new glycosidase inhibitors.[117] 6-*epi*-Valienamine derivative **222**, bearing an acetamido group at C-2 of the pseudo-sugar ring, was synthesized from protected methyl glucopyranoside **219** in a 14-step synthetic pathway (Scheme 28).[122] Compound **222** showed



Scheme 27



Scheme 28

Table 6 Inhibitory activity of isoureas and guanidines derived from carbasugars (IC_{50} in μM).

Imiglucerase	**223** a. R = C_6H_{13} b. R = C_8H_{17} c. R = C_9H_{19}			**224** a. R = C_6H_{13} b. R = C_8H_{17} c. R = C_9H_{19}			**225** a. R = C_6H_{13} b. R = $(CH_2)_3Ph$	
	a (±)	b (±)	c (±)	a (±)	b (±)	c (±)	a (–)	b (–)
pH 7	1.6	0.08	0.04	0.04	0.01	0.008	0.004	0.01
pH 5.2	6.4	0.60	0.22	0.10	0.02	0.009	0.01	0.04

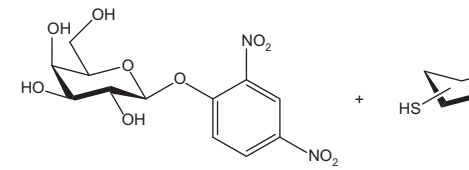
strong inhibition of human β-hexosamidase ($K_i = 34\,\mu M$) and human *O*-GlcNAcase ($K_i = 6.2\,\mu M$).

A series of *N*-alkylated bicyclic isoureas and guanidines **223**–**225** were recently prepared from 1,4-aminoalcohols or 1,4-diamines; the key steps are a HgO-promoted cyclodesulfurization reaction of a thiourea-type intermediate to give a transient carbodiimide, followed by cyclization involving the hydroxyl or amino groups located at the C-4 position.[123] These compounds showed a remarkable inhibition of recombinant β-glucocerebrosidase (imiglucerase) in the nM range, and in most cases in a pH dependent-manner (Table 6).

7 Miscellaneous

Although imino sugars, azasugars and related compounds are the most frequent inhibitors of carbohydrate-processing enzymes, there are some other different architectures that also show valuable inhibitory properties.

In this context, Withers and co-workers accomplished the synthesis of thioglycosides as non cleavable substrate analogues designed for the inhibition of glycosidases and carbohydrate-binding proteins.[124] In particular, a small library of galactosyl thioglycosides **228**–**230** was accessed by coupling 2,4-dinitrophenyl β-D-galactopyranoside (**226**) with thiol-containing carbohydrates of D-*gluco* and D-*galacto* configurations (**227**); such reaction was promoted by a thioglycoligase derived from the *Xanthomonas manihotis* β-galactosidase (BgaX) (Scheme 29). Title compounds were tested against a series of β-galactosidases. Gal-*S*-Glc derivative **228** showed to be a potent



226

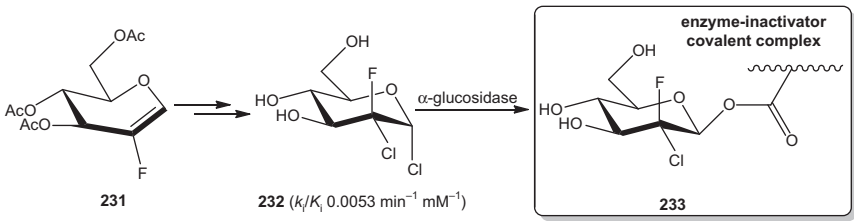


227

thioglycoligase → **228-230**

	β-Galactosidase (K_i)		
	E. Coli	X. maihotis	human lysosomal
228 (85%)	9 μM	700 μM	8 μM
229 (83%)	51 μM	1200 μM	350 μM
230 (80%)	150 μM	1300 μM	1700 μM

Scheme 29



Scheme 30

inhibitor of human and *E. coli* galactosidases. An inversion of the stereochemistry of C-4 of the terminal sugar unit (**229**) was found to decrease about 40 times the inhibition; furthermore, even lower affinity was observed by changing the linkage from β-S-1,3 to β-S-1,4 (**230**) (Scheme 29).[124]

Withers also designed[125] a series of 2-deoxy-2,2-dihalo glucosyl chlorides as a useful class of mechanism-based inhibitors of α-glucosidase; the presence of a leaving group at the anomeric position allowed trapping glycosyl-enzyme intermediates, what can be of interest as potential therapeutics. In this context, 2-chloro-2-deoxy-2-fluoro-α-D-glucosyl chloride derivative (**232**) was synthesized by a chlorination reaction of 2-fluoroglucal **231** (Scheme 30). Compound **232** was found to be a time-dependent inactivator of α-glucosidase (yeast), as covalent glycosyl-enzyme intermediate **233** could be obtained upon interaction with the enzyme.

Cu(I)-Catalyzed alkyne-azide cycloaddition has become a useful reaction for linking sugar and triazolyl scaffolds in a regioselective fashion. In this context, a series of 4-substituted 1,2,3-triazolyl-sugar derivatives have been synthesized as potential glycosidases inhibitors.[126,127] Among the compounds that bear a triazolyl moiety at the anomeric position, it has been demonstrated that α-configured triazolyl derivatives inhibit yeast α-glucosidase preferentially over their β-anomers; however the inhibition observed was found to be modest.[127] By changing the triazolyl scaffold to the positions 3 or 5 of the sugar ring (Fig. 18),[128,129] the corresponding derivatives were found to be moderate to good inhibitors of glycosidases, despite the carbohydrate residues were kept fully protected. Especially noticeable are derivatives **237** and **238**, which were reported to exhibit a potent inhibition of yeast maltase ($IC_{50} = 3.8$ and $5.2\,\mu M$, respectively) (Fig. 18).[129]

8 Pharmacological activities of imino sugars

Imino sugars are characterized by a high pharmaceutical potential; up to date, hundreds of different imino sugars and related compounds have been reported to exhibit strong interactions with an ample number of essential glycosidases and other carbohydrate-processing enzymes. Therefore, such carbohydrate mimics can address a vast number of biological targets, being candidates for the treatment of diseases ranging from cancer to lysosomal storage disorders or microbial infections, among others (Fig. 19).[130,131]

Inspired by naturally-occurring imino sugars,[132] a great number of synthetic imino sugars emerged in the so-called first generation. Although

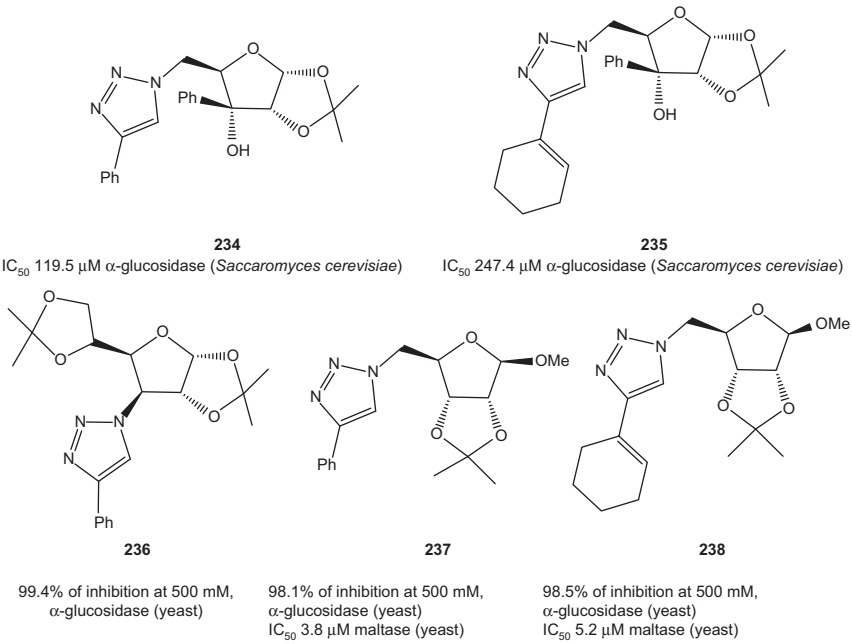


Fig. 18

many of them lacked enzymatic selectivity, two drugs were approved and marketed for clinical use, and some others are in advanced clinical-stages: Miglitol® (*N*-2-hydroxyethyl-deoxynojirimycin) and Zavesca® (*N*-butyl-deoxynojirimycin) are currently used for the treatment of type 2 diabetes and Gaucher/Niemann-Pick type C diseases (substrate reduction therapy), respectively.[130]

8.1 Imino sugars for the treatment of lysosomal storage disorders

Lysosomal storage disorders are diseases caused by the impairment of lysosomal enzymes, leading to the accumulation of partially degraded metabolites in the intra-lysosomal region of several organs. Gaucher (accumulation of glucosylceramide), Fabry (accumulation of globotriaosylceramide), Tay-Sachs, and Sandhoff (accumulation of *N*-acetyl-hexosamine-terminating glycolipids or *N*-glycoproteins, respectively) diseases are remarkable examples.[1,133,134] In some of the lysosomal storage disorders, handicapped proteins are misfolded, but still keep some catalytic activity; nevertheless, they are targeted and degraded by the quality control machinery on the endoplasmatic reticulum. One promising therapy for the treatment of such diseases is called the chaperone therapy, where a glycosidase inhibitor at sub-inhibitory concentrations binds the active site of the enzyme and induces the proper folding.[1]

Gaucher disease is the lysosomal storage disorder most frequently studied; an ample number of imino sugar derivatives have been designed and proposed for chaperoning-based studies, ranging from unsubstituted/*N*-alkylated deoxynojirimycins,[135] bicyclic guanidines,[136] noeurostegines,[137]

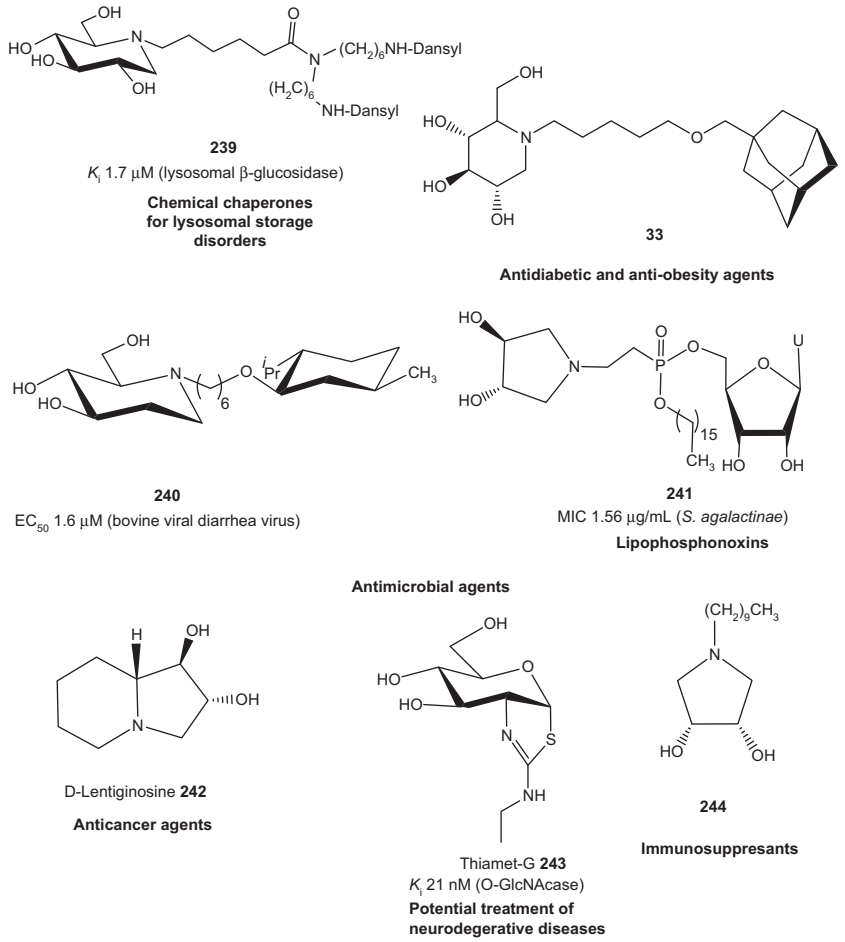


Fig. 19

N-substituted δ-lactams[138] or tetrahydroxyazepanes,[139] among others. Stütz's group has developed[140] a series of fluorescently-tagged deoxynojirimycins bearing dansyl groups for diagnosis purpose in lysosomal storage disorders. For instance, compound **239** (Fig. 19), a potent inhibitor of lysosomal β-glucosidase ($K_i = 1.7\,\mu M$), was effective at enhancing mutant lysosomal glucocerebrosidase activity.

1-Deoxygalactonojirimycin (migalastat) was found[141] to increase α-galactosidase A activity in mice upon oral administration, and also to significantly decrease the concentration of globotriaosylceramide in skin, heart, kidney and even in brain, as this compound was capable of crossing the blood-brain barrier. These results might suggest a potential therapeutic use in Fabry disease.

N-Butyldeoxynojirimycin, already used in clinical treatment of Gaucher disease, was chosen[142] as a model compound for studying Sandhoff disease in mice; surprisingly, it was found that this compound was efficient for lowering the concentration of forebrain ganglioside content, and the combination with a restricted ketogenic diet enhanced its transportation to brain.

8.2 Antidiabetic agents

Type 1 and 2 diabetes mellitus are a worldwide health concern provoking chronic hyperglycemia, which in turn leads to long-term damage in several organs and cardiovascular diseases. Two ways of combating type 2 diabetes using enzymatic inhibitors can be suggeted:[1] to regulate digestive α-glucosidases, or to decrease the hepatic biosynthesis of glucose mediated by glycogen phosphorylase[143] (*e.g.* by using spiranic compounds covered in section 5.3).

In this context, research on imino sugar-based glycosidase inhibitors led to miglitol, an intestinal α-1,4-glucosidase inhibitor, marketed by the Bayer Company in 1996 for use in the United States, South America and some European countries.[144]

There are evidences that correlate obesity and the excess of lipidic matter with diminished response to insulin.[145] Recently, attention has been focused on lipophilic *N*-[5-(adamantan-1-ylmethoxy)pentyl]-1-deoxynojirimycin **33** (AMP-DNM, Fig. 4) which inhibits both, the synthesis of glycosylceramide (the precursor of glycosphingolipids like glangliosides), and carbohydrate assimilation, which in turn enhances hepatic insulin sensitivity.[146,147] Langeveld *et al.* suggested[148] that this compound could be a good candidate for the treatment of obesity and metabolic diseases, as AMP-DNM efficiently reduced the weight of genetically obese mice by increasing fat oxidation and by reducing the animal food intake. Furthermore, it is also a good candidate for improving glucose homeostasis and hepatic steatosis, the first stage of non-alcoholic liver disease.[149]

Moreover, AMP-DNM also provoked a deep decrease in the amount of lipids present in plasma by stimulating reverse cholesterol transportation; this proved to protect mice against atherosclerosis development.[150]

8.3 Imino sugars as antimicrobial agents

It is well-documented that a way of combating viral infections on humans (*e.g.* HIV, hepatitis B and C, or influenza) is to affect the viral *N*-glycan processing pathway; α-glucosidases are one of the enzymes involved in this process.[144] This observation led to the identification of some *N*-alkylated imino sugars as good antiviral agents like *N*-nonyl-deoxynojirimycin, but with usually poor cytotoxic profiles. In this context, Chang *et al.* have improved[151] the therapeutic profile of *N*-alkylated deoxynojirimycins against viruses of the *Flaviviridae* family like BVDV (bovine viral diarrhea virus) and DENV (dengue virus). They found that an oxygenated alkyl side chain bearing a terminal ring, either cycloalkyl or aromatic, was crucial for maintaining a strong antiviral activity, while low citotoxicity. In particular, compound **240** (Fig. 19) was found to inhibit *in vitro* DENV infection with 90% effective concentration at submicromolar concentrations.

Regarding antibacterial derivatives, Krásný and co-workers have recently reported[152] a novel family of imino sugars named lipophosphonoxins **241**, bearing four different molecular modules: an imino sugar, a nucleoside scaffold, a lipophilic alkyl chain module and a phosphonate linker (Fig. 19). The optimized structure-activity relationships of these four structural motifs afforded compounds exhibiting strong antibacterial activity against a panel of Gram-positive species, including multiresistant strains. Title compounds

exhibited good MIC (*Mininum Inhibitory Concentration*) values, while higher cytotoxic concentrations against human cell lines.

8.4 Imino sugars as antitumor agents

It has been reported that one of the processes that accompanies the growth of tumour cells is aberrant glycosylation, where glycosidases are obviously involved.[153] In the imino sugars field, some glycosidase inhibitors have been found to possess antitumoral activity; the potential use of imino sugars as anticancer agents was recently reviewed by Wrodnigg *et al.*[154] The best-known imino sugars exhibiting antitumoral activity are castanospermine and swainsonine; the latter was included in phase I clinical studies[144] for treating advanced tumours, as it prevented metastatic growth. Nevertheless, the significant hepatoxicity of this compound precluded any further study. Therefore, deeper search for new imino sugar-based templates is required for developing new antitumoral agents with better pharmacological profiles. In this context, Xiao's group reported[155] the successful use of a microtiter-plate based method applied to live cells for screening antitumor activity using a library of imino sugar derivatives.

Inspired by the indolizidine-type alkaloids castanospermine and swainsonine, Macchi and co-workers focused attention on lentiginosine, a potent amyloglucosidase inhibitor. This group found[156] that the non-natural enantiomer (D-lentiginosine **242**, Fig. 19) induced caspase-dependent apoptosis on tumor cells at relatively low concentrations, in contrast with its natural enantiomer.

Promising anticancer activity has also been found in pyrrolidine-[157] and piperidine-[157a,158] based motifs. Well-known natural 1-deoxynojirimycin was found to inhibit the metastasis of B16F10 melanoma cells; studies on its antimetastatic mechanism suggested that this compound could be a good candidate as adjuvant of other antimestatic agents, like cisplatin.[158] Padró *et al*. accomplished a structure-activity study for *N*-alkylated D-fagomine and a trihydroxylated pyrrolidine, which revealed a direct correlation between the length of the alkyl residue and the citotoxicity in human cancer cell lines.[157a]

8.5 Treatmet of neurodegenerative diseases

Vocadlo's group reported[159] the synthesis of thiamet-G **243** (Fig. 19), a *cis*-fused bicyclic thiazolidine, the hitherto most potent inhibitor of eukaryotic *O*-GlcNAcases ($K_i = 21$ nM). This compound behaved as a highly selective inhibitor, as had no effect on β-hexosaminidases. Thiamet-G was proved to increase cellular levels of *O*-GlcNAc in neuron-like PC-12 cells, which in turn blocks phosphorylation of microtubule-associated protein tau at Thr231, Ser396 and Ser422; pathological hyperphosphorylation of two of these sites is characteristic in neurodegenerative Alzheimer's disease and associated taupathies.

8.6 Other biological activities of imino sugars

Ye and co-workers described[160] new *N*-alkylated 1,5-dideoxy-1,5-imino-pentitols and *N*-alkylated 1,4-dideoxy-1,4-iminotetritols (*e.g*. **244**, Fig. 19) which inhibited the mouse splenocyte proliferation by suppressing IFN-γ

and IL4-cytokines. As a result, such compounds prolonged the allograft survival in a mouse skin transplantation experiment, that is, they exhibited immunosuppressive activities, the first synthetic imino sugars showing such property in animal models.

9 Concluding remarks

Imino sugars, despite being studied for more than 40 years, are still a main goal in Medicinal Chemistry. Nature is a continuous source of new imino sugar templates, which in turn are the inspiration of synthetic chemists to design novel families with improved activity, bioavailability and selectivity. The broad pharmacological spectrum of this sugar mimics justify the current intense research in this area.

Acknowledgements

We thank the Dirección General de Investigación of Spain (CTQ2008 02813) and Junta de Andalucía (FQM 134) for financial support. S. M. thanks MICINN of Spain for the award of a fellowship.

References

1 M. Bols, Ó. López, and F. Ortega-Caballero, in *Comprehensive Glycoscience-From Chemistry to Systems Biology*, Ed. J.P. Kamerling, Elsevier, 2007, 815.
2 T. M. Gloster and G. J. Davies, *Org. Biomol. Chem.*, 2010, **8**, 305.
3 J. Wicki, S. J. Williams and S. G. Withers, *J. Am. Chem. Soc.*, 2007, **129**, 4530.
4 M. van Scherpenzeel, R .J. B. H. N. van der Berg, W. E. Donker-Koopman, R. M. J. Liskamp, J. M. F. G. Aerts, H. S. Overkleeft and R. J. Pieters, *Bioorg. Med. Chem.*, 2010, **18**, 267.
5 T. M. Gloster, P. Meloncelli, R. V. Stick, D. Zechel, A. Vasella and G. J. Davies, *J. Am. Chem. Soc.*, 2007, **129**, 2345.
6 W. A. Offen, D. L. Zechel, S. G. Withers, H. J. Gilbert and G. J. Davies, *Chem. Commun.*, 2009, 2484.
7 A. C. Sivertsen, M. Gasior, M. Bjerring, S. U. Hansen, Ó. López López, N. C. Nielsen and M. Bols, *Eur. J. Org. Chem*, 2007, 1735.
8 M. E. C. Caines, S. M. Hancock, C. A. Tarling, T. M. Wrodnigg, R. V. Stick, A. E. Stütz, A. Vasella, S. G. Withers and N. C. J. Strynadka, *Angew. Chem. Int. Ed.*, 2007, **46**, 4474.
9 Y.-X. Li, M.-H. Huang, Y. Yamashita, A. Kato, Y.-M. Jia, W.-B. Wang, G. W. J. Fleet, R. J. Nash and C.-Y. Yu, *Org. Biomol. Chem.*, 2011, **9**, 3405.
10 D. Best, S. F. Jenkinson, A. W. Saville, D. S. Alonzi, M. R. Wormald, T. D. Butters, C. Norez, F. Becq, Y. Blériot, I. Adachi, A. Kato and G. W. J. Fleet, *Tetrahedron Lett.*, 2010, **51**, 4170.
11 I. Izquierdo, M. T. Plaza, J. A. Tamayo, F. Sánchez-Cantalejo and D. L. Re, *Synthesis*, 2008, 1373.
12 E. Moreno-Clavijo, A. T. Carmona, Y. Vera-Ayoso, A. J. Moreno-Vargas, C. Bello, P. Vogel and I. Robina, *Org. Biomol. Chem.*, 2009, **7**, 1192.
13 J. S. S. Rountree, T. D. Butters, M. R. Wormald, S. D. Boomkamp, R. A. Dwek, N. Asano, K. Ikeda, E. L. Evinson, R. J. Nash and G. W. J. Fleet, *Chem. Med. Chem.*, 2009, **4**, 378.
14 J.-B. Behr, A. Kalla, D. Harakat and R. Plantier-Royon, *J. Org. Chem.*, 2008, **73**, 3612.

15 A. Kato, J. Hollinshead, Y. Yamashita, S. Nakagawa, Y. Koike, I. Adachi, C.-Y. Yu, G.W.J. Fleet and R. J. Nash, *Phytochem. Lett*, 2010, **3**, 230.
16 E. D. Goddard-Borger, R. Carapito, J.-M. Jeltsch, V. Phalip, R. V. Stick and A. Varrot, *Chem. Commun*., 2011, **47**, 9684.
17 Y. Du, Q. Pan and F. Kong, *Carbohydr. Res*., 2000, **329**, 17.
18 V. H. Lillelund, H. H. Jensen, X. Liang and M. Bols, *Chem. Rev*, 2002, **102**, 515.
19 Z. Yu, A. R. Sawkar, L. J. Whalen, C.-H. Wong and J. W. Kelly, *J. Med. Chem*., 2007, **50**, 94.
20 T. Wennekes, R. J. B. H. N. van der Berg, W. Donker, G. A. van der Marel, A. Strijland, J. M. F. G. Aerts and H. S. Overkleeft, *J. Org. Chem*., 2007, **72**, 1088.
21 G. Schitter, E. Scheucher, A. J. Steiner, A. E. Stütz, M. Thonhoefer, C. A. Tarling, S. G. Withers, J. Wicki, K. Fantur, E. Paschke, D. J. Mahuran, B. A. Rigat, M. Tropak and T. M. Wrodnigg, *Beilstein J. Org. Chem*, 2010, **6**, 21.
22 A. J. Steiner, A. E. Stütz, C. A. Tarling, S. G. Withers and T. M. Wrodnigg, *Carbohydr. Res*., 2007, **342**, 1850.
23 A. J. Steiner, G. Schitter, A. E. Stütz, T. M. Wrodnigg, C. A. Tarling, S. G. Withers, K. Fantur, D. Mahuran, E. Paschke and M. Tropak, *Bioorg. Med. Chem*., 2008, **16**, 10216.
24 A. J. Rawlings, H. Lomas, A. W. Pilling, M. J.-R. Lee, D. S. Alonzi, J. S. Rountree, J. Shane, G. W. J. Fleet, R. A. Dwek, J. H. Jones and T. D. Butters, *ChemBioChem*, 2009, **10**, 1101.
25 N. Ardes-Guisot, D. S. Alonzi, G. Reinkensmeier, T. D. Butters, C. Norez, F. Becq., Y. Shimada, S. Nakagawa, A. Kato, Y. Blériot, M. Sollogoub and B. Vauzeilles, *Org. Biomol. Chem*., 2011, **9**, 5373.
26 H. S. Overkleeft, G. H. Renkema, J. Neele, P. Vianello, I. O. Hung, A. Strijland, A. M. von der Burg, G.-J. Koomen, U. K. Pandit and J. M. F. G. Aerts, *J. Biol. Chem*., 1998, **273**, 26522.
27 J. D. Diot, I. G. Moreno, G. Twigg, C. O. Mellet, K. Haupt, T. D. Butters, J. Kovensky and S. G. Gouin, *J. Org. Chem*., 2011, **76**, 7757.
28 R. W. Wang, X. L. Qiu, M. Bols, F. Ortega-Caballero and F. L. Quing, *J. Med. Chem*., 2006, **49**, 2989.
29 R.-j. Li, M. Bols, C. Rousseau, X.-g. Zhang, R.-w. Wang and F.-L. Qing, *Tetrahedron*, 2009, **65**, 3717.
30 Y. Yang, F. Zheng, M. Bols, L. Marinescu and F.-L. Qing, *J. Fluor. Chem*, 2011, **132**, 838.
31 S. R. S. Ting, G. Chen and Martina H. Stenzel, *Polym. Chem*, 2010, **1**, 1392.
32 T. Wennekes, R. J. B. H. N. van den Berg, K. M. Bonger, W. E. Donker-Koopman, A. Ghisaidoobe, G. A. van der Marel, A. Strijland, J. M. F. G. Aerts and H. S. Overkleeft, *Tetrahedron: Asymmetry*, 2009, **20**, 836.
33 J. Diot, M. I. García-Moreno, S. G. Gouin, C. Ortiz Mellet, K. Haupt and J. Kovensky, *Org. Biomol. Chem*., 2009, **7**, 357.
34 P. Compain, C. Decroocq, J. Iehl, M. Holler, D. Hazelard, T. Mena Barragán, C. Ortiz Mellet and J.-F. Nierengarten, *Angew. Chem. Int. Ed*., 2010, **49**, 5753.
35 C. Decroocq, D. Rodríguez-Lucena, V. Russo, T. Mena Barragán, C. Ortiz Mellet and P. Compain, *Chem. Eur. J*, 2011, **17**, 13825.
36 E. Lindbäck, Ó. López, J. G. Fernández-Bolaños, S.P.A. Sauer and M. Bols, *Org. Lett*., 2011, **13**, 2908.
37 E. Lindbäck, Y. Zhou, Ó. López, J. G. Fernández-Bolaños, C. M. Pedersen and M. Bols, *Tetrahedron Lett*., 2012, **53**, 2045.
38 Ó. López López and M. Bols, *ChemBioChem*, 2007, **8**, 657.
39 A. Lohse, K. B. Jensen and M. Bols, *Tetrahedron Lett*., 1999, **40**, 3033.

40 X. Zhu, K. A. Sheth, S. Li, H. H. Chang and J. Q. Fan, *Angew. Chem. Int. Ed. Engl.*, 2005, **44**, 7450.
41 B. V. Ernholt, I. B. Thomsen, A. Lohse, I. W. Plesner, K. B. Jensen, R. G. Hazell, X. Liang, A. Jakobsen and M. Bols, *Chem. Eur. J*, 2000, **6**, 278.
42 H. H. Jensen, L. Lyngbye, A. Jensen and M. Bols, *Chem. Eur. J*, 2002, **8**, 1218.
43 M. J. Alves, F. T. Costa, V. C. M. Duarte, A. G. Fortes, J. A. Martins and N. M. Micaelo, *J. Org. Chem.*, 2011, **76**, 9584.
44 X. Liang and M. Bols, *J. Org. Chem.*, 1999, **64**, 8485.
45 E. Moreno-Clavijo, A. T. Carmona, A. J. Moreno-Vargas, M. A. Rodríguez-Carvajal and I. Robina, *Bioorg. Med. Chem.*, 2010, **18**, 4648.
46 H. H. Jensen, A. Jensen, R. G. Hazzel and M. Bols, *J. Chem. Soc. Perkin Trans*, 2002, **1**, 1190.
47 F. Morís-Varas, X.-H. Qian and C.-H. Wong, *J. Am. Chem. Soc.*, 1996, **118**, 7647.
48 H. Li, Y. Zhang, P. Vogel, P. Sinaÿ and Y. Blériot, *Chem. Commun.*, 2007, 183.
49 J. Deschamp, M. Mondon, S. Nakagawa, A. Kato, D. S. Alonzi, T. D. Butters, Y. Zhang, M. Sollogoub and Y. Blériot, *Bioorg. Med. Chem.*, 2012, **20**, 641.
50 H. Li, T. Liu, Y. Zhang, S. Favre, C. Bello, P. Vogel, T. D. Butters, N. G. Oikonomakos, J. Marrot and Y. Blériot, *ChemBioChem*, 2008, **9**, 253.
51 H. Li, F. Marcelo, C. Bello, P. Vogel, T. D. Butters, A. P. Rauter, Y. Zhang, M. Sollogoub and Y. Blériot, *Bioorg. Med. Chem.*, 2009, **17**, 5598.
52 F. Marcelo, Y. He, S. A. Yuzwa, L. Nieto, J. Jiménez-Barbero, M. Sollogoub, D. J. Vocadlo, G. D. Davies and Y. Blériot, *J. Am. Chem. Soc.*, 2009, **131**, 5390.
53 D.M.F. van Aalten, D. Komander, B. Synstad, S. Gaseidnes, M. G. Peter, and V.G.H. Eijsink, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**, 8979.
54 I. Tews, A. Perrakis, A. Oppenheim, Z. Dauter, K. S. Wilson and C. E. Vorgias, *Nat. Struct. Biol.*, 1996, **3**, 638.
55 N. B. Klamkar, V. M. Kasture, S. T. Chavan, S. G. Sabharwal and D. D. Dhavale, *Tetrahedron*, 2010, **66**, 8522.
56 T.-L. Shih, R.-Y. Yang, S.-T. Li, C.-F. Chiang and C.-H. Lin, *J. Org. Chem.*, 2007, **72**, 4258.
57 T.-L. Shih, M.-T. Liang, K.-D. Wu and C.-H. Lin, *Carbohydr. Res.*, 2011, **346**, 183.
58 B. Krämer, T. Franz, S. Picasso, P. Pruschek and V. Jäger, *Synlett*, 1997, 295.
59 G. M. J. Lenagh-Snow, N. Araujo, S. F. Kenkinson, C. Rutherford, S. Nakagawa, A. Kato, C.-Y. Yu, A. C. Weymouth-Wilson and G. W. J. Fleet, *Org. Lett.*, 2011, **13**, 5834.
60 J. P. Michael, *Nat. Prod. Rep.*, 2008, **25**, 139.
61 B. La Ferla, L. Cipolla, and F. Nicotra, in *Iminosugars-from Synthesis to Thereapeutic Applications*, Ed. P. Compain and O.R. Martin, John Wiley & Sons, 2007, 25.
62 L. Cipolla, M. R. Fernandes, M. Gregori, C. Airoldi and F. Nicotra, *Carbohydr. Res.*, 2007, **342**, 1813.
63 A. Michalik, J. Hollinshead, L. Jones, G. W. J. Fleet, C.-Y. Yu, X.-G. Hu, R. van Well, G. Horne, F. X. Wilson, A. Kato, S. F. Jenkinson and R. J. Nash, *Phytochem. Lett*, 2010, **3**, 136.
64 X.-G. Hu, B. Bartholomew, R. J. Nash, F. X. Wilson, G.W.J. Fleet, S. Nakagawa, A. Kato, Y.-M. Jia, R. van Well and C.-Y Yu, *Org. Lett.*, 2010, **12**, 2562.
65 M. Aguilar-Moncayo, C. O. Mellet, J. M. G. Fernández and M. I. García-Moreno, *J. Org. Chem.*, 2009, **74**, 3595.

66 M. Aguilar-Moncayo, P. Díaz-Pérez, J. M. García Fernández, C. Ortiz Mellet and I. García Moreno, *Tetrahedron*, 2012, **68**, 681.
67 M. Aguilar-Moncayo, T. M. Gloster, J. P. Turkenburg, I. García Moreno, C. Ortiz Mellet, G. J. Davies and J. M. García Fernández, *Org. Biomol. Chem.*, 2009, **7**, 2738.
68 B. Brumshtein, M. Aguilar-Moncayo, I. García Moreno, C. Ortiz Mellet, J. M. García Fernández, I Silman, Y. Shaaltiel, D. Aviezer, J. L. Sussman and A. H. Futerman, *ChemBioChem*, 2009, **10**, 1480.
69 E. M. Sánchez-Fernández, R. Rísquez-Cuadro, M. Chasseraud, A. Ahidouch, C. Ortiz Mellet, H. Ouadid-Ahidouch and J. M. García. Fernández, *Chem. Commun.*, 2010, **46**, 5328.
70 G. Pandey, D. Grahacharya, K. S. Shashidhara, M. I. Khan and V. G. Puranik, *Org. Biomol. Chem.*, 2009, **7**, 3300.
71 F. Cardona, C. Parmeggiani, E. Faggi, C. Bonaccini, P. Gratteri, L. Sim, T. M. Gloster, S. Roberts, G. J. Davies, D. R. Rose and A. Goti, *Chem. Eur. J*, 2009, **15**, 1627.
72 C. Bonaccini, M. Chioccioli, C. Parmeggiani, F. Cardona, D. Lo, Re, G. Soldaini, P. Vogel, C. Bello, A. Goti and P. Gratteri, *Eur. J. Org. Chem*, 2010, 5574.
73 Ó. López, J. G. Fernández-Bolaños and M. V. Gil, *Green Chem.*, 2005, **7**, 431.
74 N. Asano, H. Kuroi, K. Ikeda, H. Kizu, Y. Kameda, A. Kato, I. Adachi, A. A. Watson, R. J. Nash and G.W.J. Fleet, Tetrahedron, *Asymmentry*, 2000, **11**, 1.
75 a) L. Chabaud, Y. Landais, and P. Renaud, *Org. Lett.*, 2005, **7**, 2587; b) S. Chandrasekhar, B.B. Parida, and C. Rambabu, *J. Org. Chem.*, 2008, **73**, 7826; c) P.V. Reddy, A. Veyron, P. Koos, A. Bayle, A.E. Greene, and P. Delair, *Org. Biomol. Chem.*, 2008, 6, 1170; d) T. J. Donohoe, R. E. Thomas, M. D. Cheeseman, C. L. Rigby, G. Bhalay, and I. D. Linney, *Org. Lett.*, 2008, **10**, 3615.
76 G. D'Adamio, A. Goti, C. Parmeggiani and E. Moreno-Clavijo, I. Robina, and F. Cardona, *Eur. J. Org. Chem*, 2011, 7155.
77 T. S. Rasmussen and H. H. Jensen, *Org. Biomol. Chem.*, 2010, **8**, 433.
78 T. M. Gloster, R. Madsen and G. J. Davies, *ChemBioChem*, 2006, **7**, 738.
79 D. L. Zechel, A. B. Boraston, T. Gloster, C. M. Boraston, J. M. Mcdonald, M. G. Tilbrook, R. V. Stick and G. J. Davies, *J. Am. Chem. Soc.*, 2003, **125**, 14313.
80 H. Liu, X. Liang, H. Søhoel, A. Bülow and M. Bols, *J. Am. Chem. Soc.*, 2001, **123**, 5116.
81 M. Aguilar, T. M. Gloster, I. García Moreno, C. Ortiz Mellet, G. J. Davies, A. Llebaria, J. Casas, M. Egido-Gabás and J. M. García Fernández, *ChemBioBhem*, 2008, **9**, 2612.
82 V. Chagnault, P. Compain, K. Lewinski, K. Ikeda, N. Asano and O. R. Martin, *J. Org. Chem.*, 2009, **74**, 3179.
83 S. Knapp, D. Vocadlo, Z. N. Gao, B. Kirk, J. P. Lou and S. G. Withers, *J. Am. Chem. Soc.*, 1996, **118**, 6804.
84 G. E. Whitworth, M. S. Macauley, K. A. Stubbs, R. J. Dennis, E. J. Taylor, G. J. Davies, I. R. Greig and D. J. Vocadlo, *J. Am. Chem. Soc.*, 2007, **129**, 635.
85 S. Knapp, D. Fash, M. Abdo, T. J. Emge and P. R. Rablen, *Bioorg. Med. Chem.*, 2009, **17**, 1831.
86 S. Knapp, M. Abdo, K. Ajayi, R. A. Huhn, T. J. Emge, E. J. Kim and J. A. Hanover, *Org. Lett.*, 2007, **9**, 2321.
87 J. M. Macdonald, C. A. Tarling, E. J. Taylor, R. J. Dennis, D. S. Meyers, S. Knapp, G. J. Davies and S. G. Withers, *Angew. Chem. Int. Ed.*, 2010, **49**, 2599.
88 B. Amorelli, C. Yang, B. Rempel, S. G. Withers and S. Knapp, *Bioorg. Med. Chem.*, 2008, **18**, 2944.

89 E. J. Kim, D. C. Love, E. Darout, M. Abdo, B. Rempel, S. G. Withers, P. R. Rablen, J. A. Hanover, and S. Knapp, *Bioorg. Med. Chem.*, 2010, **18**, 7058.
90 S. S. Kale, S. T. Chavan, S. G. Sabharwal, V. G. Puranik and G. J. Sanjayan, *Org. Biomol. Chem.*, 2011, **9**, 7300.
91 S. Mio, R. Ichinose, K. Goto, S. Sugai and S. Sato, *Tetrahedron*, 1991, **47**, 2111.
92 C. J. F. Bichard, E. P. Mitchell, M. R. Wormald, K. A. Watson, L. N. Johnson, S. E. Zographos, D. D. Koutra, N. G. Oikonomakos and G. W. J. Fleet, *Tetrahedron Lett.*, 1995, **36**, 2145.
93 E. Ösz, L. Somsák, L. Szilágyi, L. Kovács, T. Docsa, B. Tóth and P. Gergely, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 1385.
94 L. Somsák, V. Nagy, T. Docsa, B. Tóth and P. Gergely, *Tetrahedron: Asymmetry*, 2000, **11**, 405.
95 T. Docsa, K. Czifrák, C. Hüse, L. Somsák and P. Gergely, *Mol. Med. Report*, 2011, **4**, 477.
96 L. Somsák, *C.R. Chimie*, 2011, **14**, 211.
97 L. Somsák, V. Nagy, S. Vidal, K. Czifrák, E. Berzsényi and J.-P. Praly, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 5680.
98 V. Nagy, M. Benltifa, S. Vidal, E. Berzsényi, C. Teilhet, K. Czifrák, G. Batta, T. Docsa, P. Gergely, L. Somsák and J.-P. Praly, *Bioorg. Med. Chem.*, 2009, **17**, 5696.
99 M. Benltifa, J. M. Hayes, S. Vidal, D. Gueyrard, P. G. Goekjian, J.-P. Praly, G. Kizilis, C. Tiraidis, K.-M. Alexacou, E. D. Chrysina, S. E. Zographos, D. D. Leonidas, G. Archontis and N. G. Oikonomakos, *Bioorg. Med. Chem.*, 2009, **17**, 7368.
100 T. Tite, L. Tomas, T. Docsa, P. Gergely, J. Kovensky, D. Gueyrard and A. Wadouachi, *Tetrahedron Lett.*, 2012, **53**, 959.
101 M. Benltifa, M. De Kiss, M. I. Garcia-Moreno, C. Ortiz Mellet, D. Gueyrard and A. Wadouachi, Tetrahedron, *Asymmetry*, 2009, **20**, 1817.
102 C. Laroche, J.-P. Behr, J. Szymoniak, P. Bertus, C. Schütz, P. Vogel and R. Plantier-Royon, *Bioorg. Med. Chem.*, 2006, **14**, 4047.
103 M. S. M. Pearson, R. Plantier-Royon, J. Szymoniak, P. Bertus and J.-B. Behr, *Synhtesis*, 2007, 3589.
104 M. S. M. Pearson, N. Floquet, C. Bello, P. Vogel, R. Plantier-Royon, J. Szymoniak, P. Bertus and J.-B. Behr, *Bioorg. Med. Chem.*, 2009, **17**, 8020.
105 S. Mohan and B. M. Pinto, *Carbohydr. Res.*, 2007, **342**, 1551.
106 M. Yoshikawa, T. Murakami, H. Shimada, H. Matsuda, J. Yamahara, G. Tanabe and O. Muraoka, *Tetrahedron Lett.*, 1997, **48**, 8367.
107 M. Yoshikawa, T. Murakami, K. Yashiro and H. Matsuda, *Chem. Pharm. Bull.*, 1998, **46**, 1339.
108 K. Jayakanthan, S. Mohan and B. M. Pinto, *J. Am. Chem. Soc.*, 2009, **131**, 5621.
109 Y. Li, T.H.-W. Huang and J. Yamahara, *Life Sci.*, 2008, **82**, 1045.
110 W. Chen, L. Sim, D. R. Rose and B. M. Pinto, *Carbohydr. Res.*, 2007, **342**, 1661.
111 R. Nasi, L. Sim, D. R. Rose and B. M. Pinto, *J. Org. Chem.*, 2007, **72**, 180.
112 H. Liu, R. Nasi, K. Jayakanthan, L. Sim, H. Heipel, D. R. Rose and B. M. Pinto, *J. Org. Chem.*, 2007, **72**, 6562.
113 S. Mohanand and B. M. Pinto, *Nat. Prod. Rep.*, 2010, **27**, 481.
114 R. Nasi, B. O. Patrick, L. Sim, D. R. Rose and B. M. Pinto, *J. Org. Chem.*, 2008, **73**, 6172.
115 R. Nasi, L. Sim, D. R. Rose and B. M. Pinto, *Carbohydr. Res.*, 2007, **342**, 1888.
116 N. S. Kumar, D. A. Kuntz, X. Wen, B. M. Pinto and D. R. Rose, *Proteins*, 2008, **71**, 1484.

117 S. Ogawa and M. Kanto, *Curr. Top. Med. Chem.*, 2009, **9**, 58.
118 A. Baran and M. Balci, *J. Org. Chem*, 2009, **74**, 88.
119 S. Cantekin, A. Baran, R. Çalişkan and M. Balci, *Carbohydr. Res.*, 2009, **344**, 426.
120 A. Baran, A. Günel and M. Balci, *J. Org. Chem*, 2008, **73**, 4370.
121 R. Lysek, S. Favre and P. Vogel, *Tetrahedron*, 2007, **63**, 6558.
122 A. Scafdi, K. A. Stubbs, R. J. Dennis, E. J. Taylor, G. J. Davies, D. J. Vocadlo and R. V. Stick, *Org. Biomol. Chem.*, 2007, **5**, 3013.
123 A. Trapero, I. Alfonso, T. D. Butters and A. Llebaria, *J. Am. Chem. Soc*, 2011, **133**, 5474.
124 Y-W. Kim, H.-M. Chen, J. H. Kim, J. Müllegger, D. Mahuran and S. G. Withers, *ChemBioChem*, 2007, **8**, 1495.
125 R. Zhang, J. D. McCarter, C. Braun, W. Yeung, G. D. Brayer and S. G. Withers, *J. Org. Chem.*, 2008, **73**, 3070.
126 K. Slámová, P. Marhol, K. Bezouška, L. Lindkvist, S. G. Hansen, V. Křen and H. H. Jensen, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 4263.
127 S. Dedola, D. L. Hughes, S. A. Nepogodiev, M. Rejzek and R. A. Field, *Carbohydr. Res.*, 2010, **345**, 1123.
128 D. R. da Rocha, W. C. Santos, E. S. Lima and V. F. Ferreira, *Carbohydr. Res.*, 2012, **350**, 14.
129 S. B. Ferreira, A.C.R. Sodero, M.F.C. Cardoso, E. S. Lima, C. R. Kaiser, F. P. Silva and V. F. Ferreira, *J. Med. Chem.*, 2010, **53**, 2364.
130 G. Horne, F. X. Wilson, J. Tinsley, D. H. Williams and R. Storer, *Drug Discov. Today*, 2011, **16**, 107.
131 T. Kajimoto and M. Node, *Curr. Top. Med. Chem.*, 2009, **9**, 13.
132 B. G. Winchester, *Tetrahedron: Asymmetry*, 2009, **20**, 645.
133 T. D. Butters, *Curr. Opin. Chem. Biol.*, 2007, **11**, 412.
134 S. D. Boomkamp, J. S. S. Rountree, D. C. A. Neville, R. A. Dwek, G. W. J. Fleet and T. D. Butters, *Glycoconjugate J.*, 2010, **27**, 297.
135 T. Cox T, R. Lachmann, C. Hollak, J. Aerts, S. van Weely, M. Hrebícek, F. Platt F, T. Butters, R. Dwek, C. Moyses, I. Gow, D. Elstein and A. Zimran, *Lancet*, 2000, **355**, 1481.
136 Z. Luan, k. Higaki, M. Aguilar-Moncayo, L. Li, H. Ninomiya, E. Nanba, K. Ohno, M. I. García-Moreno, C. Ortiz Mellet, J. M. García Fernández, and Y. Suzuki, *ChemBioChem* 2010, **11**, 2453.
137 T. S. Rasmussen, S. Allman, G. Twigg, T. D. Butters and H. H. Jensen, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 1519.
138 G.-N. Wang, G. Reinkensmeier, S.-W. Zhang, J. Zhou, L.-R. Zhang, L.-H. Zhang, T. D. Butters and X.-S. Ye, *J. Med. Chem.*, 2009, **52**, 3146.
139 S. D. Orwing, Y. L. Tan, N. P. Grimster, Z. Yu, E. T. Powers, J. W. Kelly and R. L. Lieberman, *Biochemistry*, 2011, **50**, 10647.
140 R. F. G. Fröhlich, R. H. Furneaux, D. J. Mahuran, B. A. Rigat, A. E. Stütz, M. B. Tropak, J. Wicki, S. G. Withers and T. M. Wrodnigg, *Carbohydr. Res.*, 2010, **345**, 1371.
141 R. Khanna, R. Soska, Y. Lun, F. Feng, M. Frascella, B. Young, N. Brignol, L. Pellegrino, S. A. Sitaraman, R. J. Desnick, E. R. Benjamin, D. J. Lockhart and K. J. Valenzano, *Mol. Ther.*, 2010, **18**, 23.
142 C. A. Denny, K. A. Heinecke, Y. P. Kim, R. C. Baek, K. S. Loh, T. D. Butters, R. T. Bronson, F. M. Platt and T. N. Seyfried, *J. Neurochem.*, 2010, **113**, 1525.
143 J.-P. Praly and S. Vidal, *Mini-Rev. Med. Chem.*, 2010, **10**, 1102.
144 T. M. Cox, F. M. Platt and J. M. F. G. Aerts, in *Iminosugars-from Synthesis to Thereapeutic Applications*, Ed. P. Compain and O. R. Martin, John Wiley & Sons, 2007, 295.
145 R. H. Unger, *Endocrinology*, 2003, **144**, 5159.

146 J. M. Aerts, R. Ottenhoff, A. S. Powlson, A. Grefhorst, M. Van Eijk, P. F. Dubbelhuis, J. Aten, F. Kuipers, M. J. Serlie, T. Wennekes, J. K. Sehti, S. O'Rahilly and H. S. Overkleeft, *Diabetes*, 2007, **56**, 1341.
147 T. Webbejesm, A. J. Meijer, A. K. Groen, R. G. Boot, J. E. Groener, M. Van Eijk, R. Ottenhoff, N. Bijl, K. Ghauharali, H. Song, T. J. O'shea, H. Liu, N. Yew, D. Copeland, R. J. Vn den Berg, G. A. van der Marel, H. S. Overkleeft and J. M. Aerts, *J. Med. Chem*., 2010, **53**, 689.
148 M. Langeveld, S. A. A. van den Berg, N. Bijl, S. Bijland, C. P. Van Roomen, J. H. Houben-Weerts, R. Otterhoff, S. M. Houten, K. W. Van Dijk, J. A. Romijn, A. K. Groen, J. M. Aerts and P. J. Voshol, *Metab. Clin. Exp*., 2012, **61**, 99.
149 N. Bijl, M. Sokolović, C. Vrins, M. Langeveld, P. D. Moerland, R. Ottenhoff, C. P. A. A. van Roomen, N. Claessen, R. G. Boot, J. Aten, A. K. Groen, J. M. F. G. Aerts and M. Van Eijk, *Hepatology*, 2009, **50**, 1431.
150 F. Bietrix, E. Lombardo, C. P. A. A. van Roomen, R. Ottenhoff, M. Vos, P. C. N. Rensen, A. J. Verhoeven, J. M. Aerts and A. K. Groen, *Arterioscler. Thromb. Vasc. Biol*., 2010, **30**, 931.
151 J. Chang, L. Wang, D. Ma, X. Qu, H. Guo, X. Xu, P. M. Mason, N. Bourne, R. Moriarty, B. Gu, J.-T. Guo and T. M. Block, *Antimicrob. Agents Chemother*., 2009, **53**, 1501.
152 D. Rejman, A. Rabatinová, A. R. Pombinho, S. Kovačková, R. Pohl, E. Zborníková, M. Kolář, K. Bogdanová, O. Nyč, H. Šanderová, T. Látal, P. Bartůněk and L. Krásný, *J. Med. Chem*., 2011, **54**, 7884.
153 S. Hakomori, *Cancer Res*., 1985, **45**, 2405.
154 T. M. Wrodnigg, A. J. Steiner and B. J. Ueberbacher, *Anticancer Agents Med. Chem*., 2008, **8**, 77.
155 L. Zhang, F. Sun, Y. Li, X. Sun, X. Liu, Y. Huang, L.-H. Zhang, X.-S. Ye and J. Xiao, *ChemMedChem*, 2007, **2**, 1594.
156 B. Macchi, A. Minutolo, S. Grelli, F. Cardona, F. M. Cordero, A. Mastino and A. Brandi, *Glycobiology*, 2010, **20**, 500.
157 a) M. Padró, J. A. Castillo, L. Gómez, J. Joglar, P Clapés and C. De Bolós, *Glyconj. J*., 2010, **27**, 277. b) C. Bello, M. Cea, G.D. Bello, A. Garuti, I. Rocco, G. Cirmena, E. Moran, A. Nahimana, M.A. Duchosal, F. Fruscione, P. Pronzato, F. Grossi, F. Patrone, A. Ballestrero, M. Dupuis, B. Sordat, A. Nencioni, and P. Vogel, *Bioorg. Med. Chem*., 2010, **18**, 3320.
158 R.-J. Wang, C. H. Yang and M.-L. Hu, *J. Agric. Food Chem*., 2010, **58**, 8988.
159 S. A. Yuzwa, M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Witworth, K. A. Stubbs, E. J. McEachern, G. J. Davies and D. J. Vocadlo, *Nature Chem. Biol*., 2008, **4**, 483.
160 G.-N. Wang, Y. Xiong, J. Ye, L.-H. Zhang and X.-S. Ye, *ACS Med. Chem. Lett*., 2011, **2**, 682.

An overview of key routes for the transformation of sugars into carbasugars and related compounds

Raquel G. Soengas,*[a] José M. Otero,[b] Amalia M. Estévez,[b] Amélia P. Rauter,*[c] Vasco Cachatra,[c] Juan C. Estévez[b] and Ramón J. Estévez*[b]

DOI: 10.1039/9781849734769-00263

This chapter reports an overview of the most relevant routes developed for the transformation of sugars into carbasugars and related compounds.

1 Introduction

Carbasugars, a family of carbohydrate mimics, have attracted great interest in recent years among chemists and biochemists due to their potential biological properties.[1] The ability of carbasugars to mimic the size, structure and polarity of parent sugars and their stability towards hydrolysis makes them potential candidates for the competitive inhibition of glycosidases and glycosyltransferases.

Most of the synthetical strategies developed for the synthesis of carbasugars make use of carbohydrates as starting materials to maintain their high grade of functionalisation in the final product. It is clear that the use of carbohydrates provides important advantages for the preparation of their carbocyclic analogues, since the hydroxyl groups can be maintained throughout the synthetic sequence, with no need for ''hydroxylation'' reactions whereas the enantiomeric purity of the target carbasugars will be guaranteed.

1.1 Carbasugars and pseudo-carbasugars

Carbasugars are carbohydrate analogues in which the endocyclic oxygen has been replaced by a methylene group (Fig. 1). Natural and synthetic carbasugars, either as single molecules or as subunits of more complex molecules, have shown to display interesting biological activities, mainly as enzymatic inhibitors.[2]

The carbasugars were introduced about forty years ago when McCasland and coworkers first synthesised the talopyranose analogue **1**,[3] followed by two new isomeric compounds **2** and **3**[4] (Fig. 2).

McCasland used the term ''pseudosugar'' to designate such analogues but, as time went by, this term started to be employed to designate a large variety of sugar analogues. S. Ogawa proposed the use of the prefix ''carba'', preceded, where considered necessary, by the appropriate locant

[a]*Department of Chemistry, University of Aveiro, Portugal.*
[b]*Center for Research in Biological Chemistry and Material Sciences (CIQUS), University of Santiago de Compostela, Spain. E-mail: ramon.estevez@usc.es*
[c]*Carbohydrate Chemistry Group, Center of Chemistry and Biochemistry, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Lisbon, Portugal. E-mail: aprauterfcul@gmail.com*

© The Royal Society of Chemistry 2012

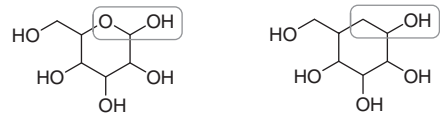


Fig. 1 Structures of a six-membered ring sugar and the corresponding carbasugar.

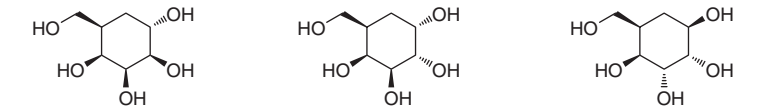


Fig. 2 First three carbasugars synthesised by McCasland's group.

("4a" for an aldofuranose, "5a" for an aldopyranose), followed by the name of the sugar.[5]

However, this definition is sometimes misused by many authors, who include carbocyclic polyols containing a non-substituted carbon, which should be called pseudo-carbasugars because they do not have an associated homomorphic sugar associated.[6]

1.2 Biological significance of carbasugars and pseudo-carbasugars

Carbasugars are not very abundant in Nature. In fact, the 5a-carba-α-D-galactopyranose (**2**), a weak antibiotic against *Klebsiella pneumonia* MB-1264, is the only carbasugar isolated from the fermentation broth of *Streptomyces sp*. MA-4145.[7] However, a large number of carbasugar related highly oxygenated cyclohexanes of biological interest are present in Nature. Representative examples include inositols[8] (components of the cell-membrane and second messengers); glycosidase inhibitors such as conduritols[9] and cyclophellitols;[10] aminocarbasugars such as valienamine **7**,[11] validamine **8**, hydroxyvalidamine **9** and valiolamine **10**,[12] which are components of natural aminoglycoside antibiotics,[13] and (−)-shikimic acid (**11**)[14] and (+)-quinic acid (**12**),[15] two important biochemical intermediates in plant and microorganism metabolism (Fig. 3).

Some of these compounds have relevant biological activities and potential pharmacological applications. For example, (+)–cyclophellitol **6**, isolated from the culture filtrate of a mushroom strain, *Phellinus sp*.[16] is a potent inhibitor of β-glucosidases, with potential inhibition of the human immunodeficiency virus (HIV) and with possible antimetastatic therapeutic activity.[17] The (+)-MK7607 **13**, isolated from the fermentation broth of *Curvularia eragrostidis* D2452,[18] is a patented naturally occurring unsaturated pseudo-carbasugar possessing herbicidal activity, and (−)-shikimic acid **11** is a key intermediate in the biosynthesis of amino acids, which has some pharmacological relevance.[19] It has been used as staring material in the synthesis of many drugs, the most relevant being the antiviral agent oseltamivir (Tamiflu).[20]

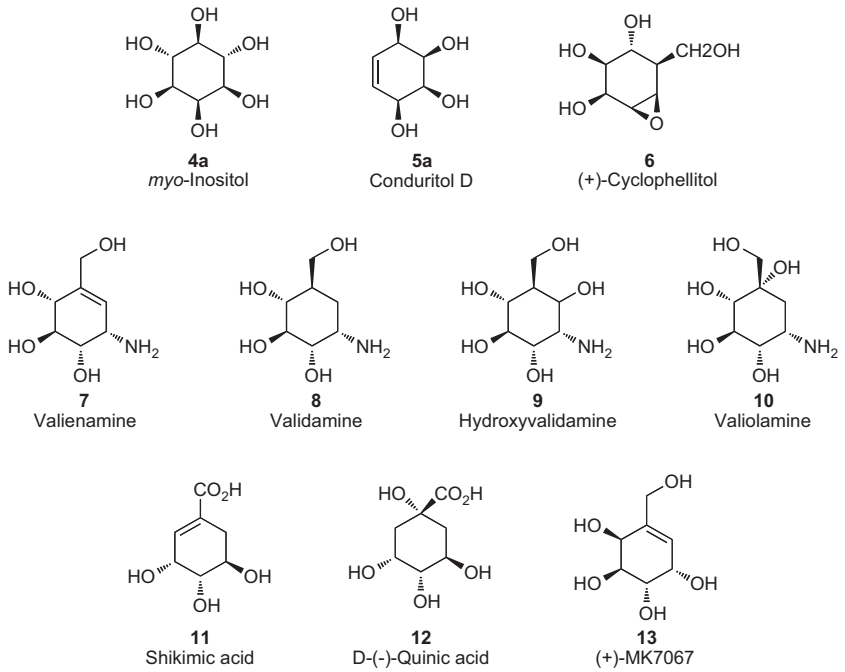


Fig. 3 Examples of naturally occurring pseudo-carbasugars.

Probably, the most important carbasugar derivatives from a biological point of view are the aminocarbasugars, some of which have become clinically successful to combat diseases in humans and in plants. In 1970, researchers at Takeda Chemical Company discovered the agricultural antibiotic Validamycin A,[21] a representative member of the family of Validamycin antibiotics. The core of these antibiotics is composed by valienamine (**7**), an amino pseudo-carbasugar. The pronounced biological activity of valienamine (**7**), as well as validamine (**8**) and hydroxyvalidamine (**9**) as glycosidase inhibitors prompted the scientists to synthesize and evaluate several derivatives.[22] Among them are carbocyclic analogues of glycosylamides, structurally related to glycosphingolipids and glycoglycerolipids, which showed similar potencies as immunomodulators, comparable to those of the corresponding true sugar analogues.[23]

Carbasugar derivatives are also relevant as potential entities for the elucidation and/or control of biological events involving sugar moieties. The construction of enzyme substrate analogues, in which part of their structure has been replaced by a carbasugar unit, is expected to be useful in the establishement of the mode of action of sugar transferases. An area of recent interest is the design of potential substrates and inhibitors of fucosyl and sialyl transferases involved in the last steps of the biosynthesis of Lewis oligosaccharide antigens.[24]

2 Strategies for the synthesis of carbasugars and pseudo-carbasugars

The biological significance of carbasugars has stimulated a great interest in their preparation.[1,25] Several approaches have been developed for the

synthesis of carbasugars and related pseudo-carbasugars, the most common being those that make use of carbanionic and radical cyclizations, ring closing metathesis (RCM), pinacol couplings, cycloaddition reactions and rearrangement of cyclic sugars.

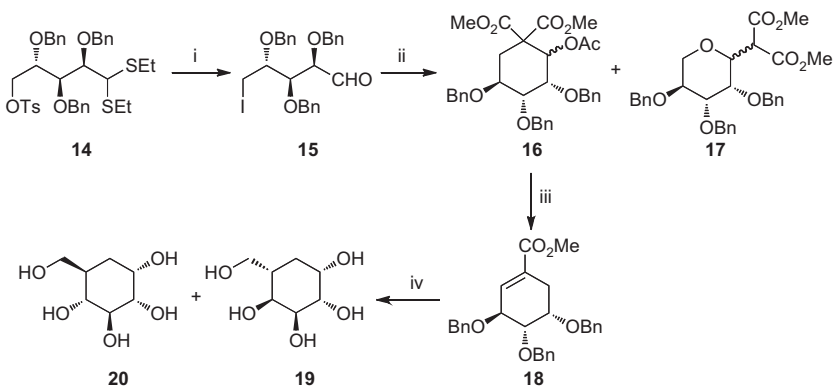
2.1 Carbanion cyclisations

The use of anionic species as nucleophiles in the formation of C-C bonds is widespread in organic chemistry and it has also been extensively applied to the formation of carbasugars. For this purpose, various stabilising groups have gained particular attention, namely nitro groups, phosphorous ylides and aldol precursors.

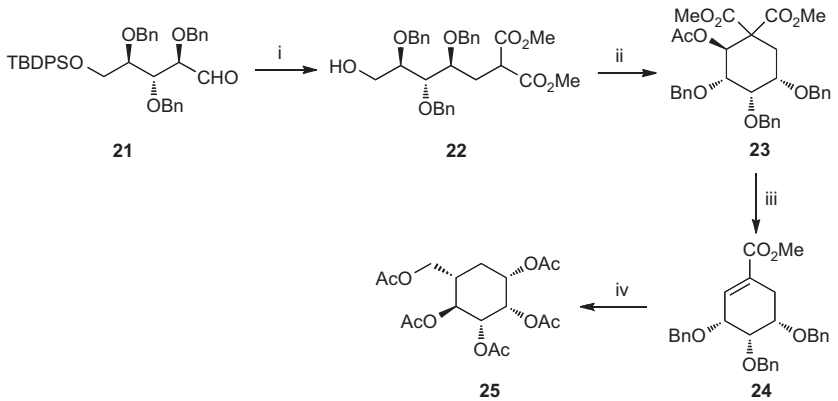
2.1.1 Malonates. The use of malonate anions to synthesise carbocycles from open-chain sugars was first introduced by Suami *et al*.[26] Starting from L-arabinose diethyl dithioacetal tosylate **14**, the key iodo-aldehyde **15** was easily obtained. Reaction of **15** with dimethyl malonate anion, followed by cyclisation and acetylation provided the cyclohexyl derivative **16** and pyranose derivative **17** in ratio 5:3. Formation of compound **17** could be explained in terms of a displacement of the iodide by an alkoxide. A Krapcho decarboxylation of **16**, followed by elimination of an acetoxy group gave the cyclohexene derivative **18**. Reduction of the unsaturated ester with $LiAlH_4$, followed by hydroboration-oxidation and final deprotection gave a mixture of 5a-carba-α-D-glucopyranose (**19**) and 5a-carba-β-L-altropyranose (**20**) (Scheme 1).

In order to avoid the competition between addition to the carbonyl and substitution of the iodide, an alternative pathway incorporating the malonate moiety prior to the transformation of the primary hydroxygroup into a leaving group was developed.[27] This alternative was used, for instance, for the preparation of 5a-carba-β-L-mannopyranose pentaacetate (**25**) (Scheme 2).

2.1.2 Nitro groups. A nitro group is particularly effective at stabilising carbanions on the adjacent carbon. The nitronate anions, resulting from the



Scheme 1 *Reagents and conditions:* (i) a) $HgCl_2$, $CaCO_3$; b) NaI, acetone, reflux, 54%; (ii) a) dimethyl malonate, NaH, DMF; b) Ac_2O, Py, 43% of **16** and 33% of **17** over two steps; (iii) DMSO, NaCl, heat, 75%; (iv) a) $LiAlH_4$, −15 °C, 83%; b) B_2H_6, then H_2O_2, NaOH; c) Ac_2O, py, 69% (two steps); d) Na, liq. NH_3; e) Ac_2O, py; f) NaOMe, MeOH, 65% of **19** and 86% of **20**.



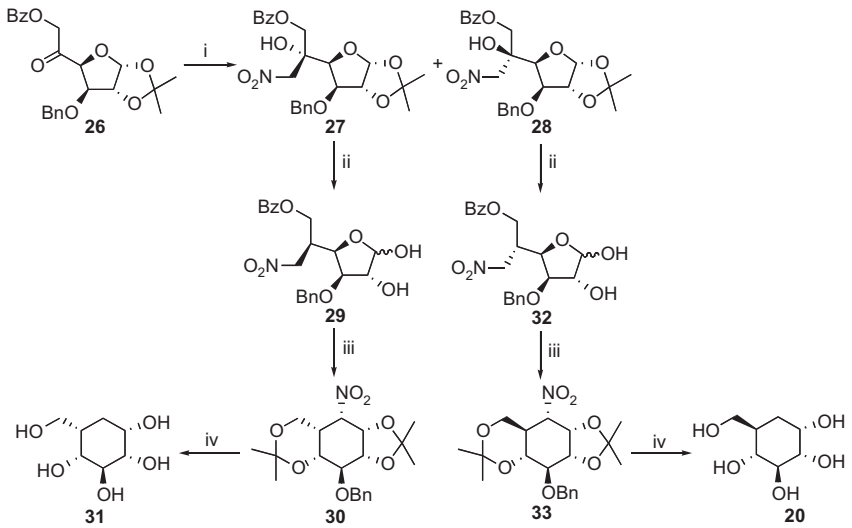
Scheme 2 *Reagents and conditions:* (i) a) dimethyl malonate, py, Ac_2O, 85%; b) H_2, Raney Ni; c) TBAF, 48% (two steps); (ii) a) PCC, CH_2Cl_2; b) Ac_2O, Py, 69% (two steps); (iii) DMSO, NaCl, heat, 70%; (iv) a) $LiAlH_4$, −15 °C, 88%; b) B_2H_6, then H_2O_2, NaOH; c) Ac_2O, py, 66% (two steps); d) Pd-black, MeOH; e) Ac_2O, py, 57% (two steps).

abstraction of the proton in the α position (pK_a 9-10) are bidentate nucleophiles stabilised from resonance. On account of this nucleophilicity, nitroderivatives are highly valuable intermediates for the formation of C-C bonds and have been widely used in carbohydrate chemistry.[28] The most relevant C-C bond formation reactions promoted by the nucleophilic addition of nitronates reactions are the Henry reaction (addition to an aldehyde or ketone),[29] the C-alkylation (addition to a saturated electrophilic carbon)[30] and the Michael addition (addition to an α,β-unsaturated system).[31] In particular, reactions involving cyclisation of nitrosugars are useful processes in the enantioselective preparation of carbasugars.[32]

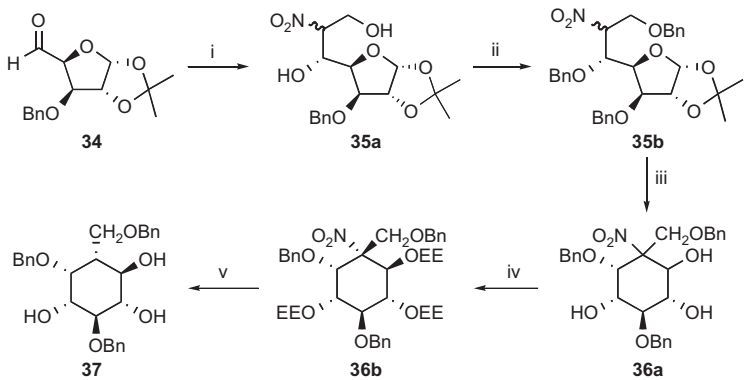
In 1984 Kitagawa and co-workers showed that an intramolecular Henry reaction of a nitrosugar is a useful process for the enantioselective preparation of carbasugars (Scheme 3).[33] Thus, addition of nitromethane to sugar derivative **26** afforded epimeric alcohols **27** and **28**. Removal of hydroxyl group of **27** by reductive deacetoxylation and subsequent removal of the acetal protecting group provided nitro sugar **29**, which in basic conditions cyclised to a pentahydroxylated nitro cylohexane, which was then easily converted into derivative **30**. Finally, radical-mediated denitration and subsequent removal of the di-*O*-isopropylidene groups gave 5a-carba-α-D-glucopyranose (**31**). On the other hand, nitro sugar **28** was similarly converted into 5a-carba-β-L-idopyranose (**20**), *via* compounds **32** and **33**.

Since then, this nitro sugar based protocol has been widely used for the preparation of carbasugar and related compounds.[34]

A recent example is the synthesis of hydroxymethyl-*myo*-inositol **37**, which was easily obtanided from a D-glucose-derived aldehyde **34.** Firstly fluoride catalysed Henry reaction of nitroethanol and aldehyde **34** afforded nitro sugar epimers **35a**. Benzylation of the free hydroxyl groups afforded the nitrosugar mixture **35b**, which, after liberation of their anomeric positions by acidic hydrolysis, were submited to an intramolecular Henry reaction using hydrogen sodium carbonate as a base. The result was the



Scheme 3 *Reagents and conditions:* (i) CH_3NO_2, NaH, 15-crown-5, DMF, 61%; (ii) a) Ac_2O, TsOH, $NaBH_4$, EtOH, 69%; b) 80% aq AcOH, 80 °C, 93% of **29** and 42% of **32**; (iii) a) KF, 18-crown-6, DMF then NaOMe, MeOH; b) 2,2-dimethoxypropane, TsOH, $CuSO_4$, 72% of **30** and 42% of **33**; (iv) a) n-Bu_3SnH, AIBN, Benzene, 80 °C; b) aq AcOH; c) Na, liq NH_3, 30% of **31** and 20% of **34**.

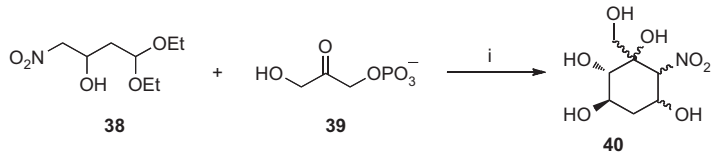


Scheme 4 *Reagents and conditions:* (i) $HO(CH_2)_2NO_2$, TBAF, THF; (ii) BTCA, TfOH, C_6H_{12}, CH_2Cl_2; (iii) a) TFA/dioxane/H_2O; b) $HNaCO_3$ 2% MeOH: (iv) $CH_2{=}CHOCH_2CH_3$, C_6H_{12}, CH_2Cl_2; (v) a) $HSnBu_3$, AIBN, toluene, reflux; b) PPTS, EtOH, 50 °C.

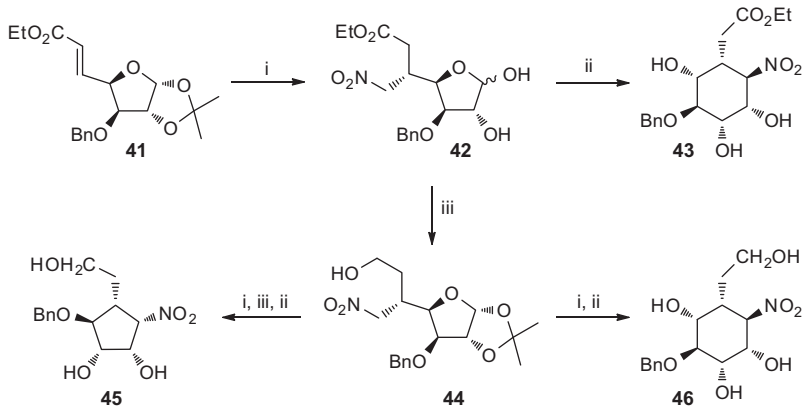
specific formation of nitrocyclohexane **36a**, which was finally transformed into desired inositol **37** by removal of the nitro group (Scheme 4).[35]

Recently, El Blidi *et al*. have publised the synthesis of a nitrocyclohexane derivative by means of a chemoenzymatic variant of the classical method of inter-intramolecular Henry reaction (Scheme 5).[36] Nitroaldol cyclisation between a masked 3-hydroxy-4-nitrobutyraldehyde **38** and dihydroxyacetone phosphate (DHAP) **39** catalysed by fructose-1,6-biphosphate aldolase (RAMA) produced the nitrocyclohexane **40**.

Another useful strategy for the preparation of carbasugars from nitro sugars involves a Michael addition-Henry reaction sequence.



Scheme 5 *Reagents and conditions*: (i) Fructose-1,6-biphosphate aldolase.

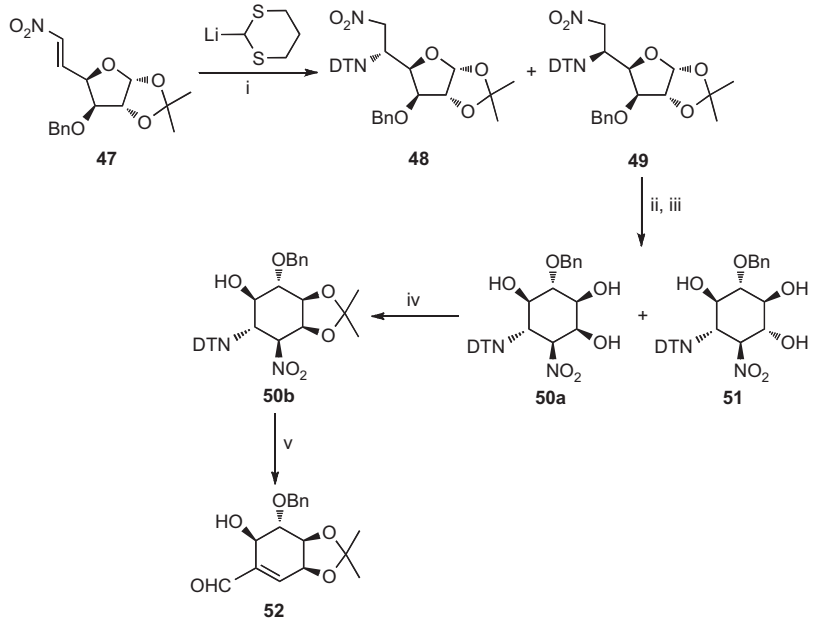


Scheme 6 *Reagents and conditions:* (i) CH_3NO_2, K_2CO_3, EtOH; ii) TFA/H_2O; (iii) DABCO, benzene, reflux; (iv) LAH, THF; (v) $NaIO_4$.

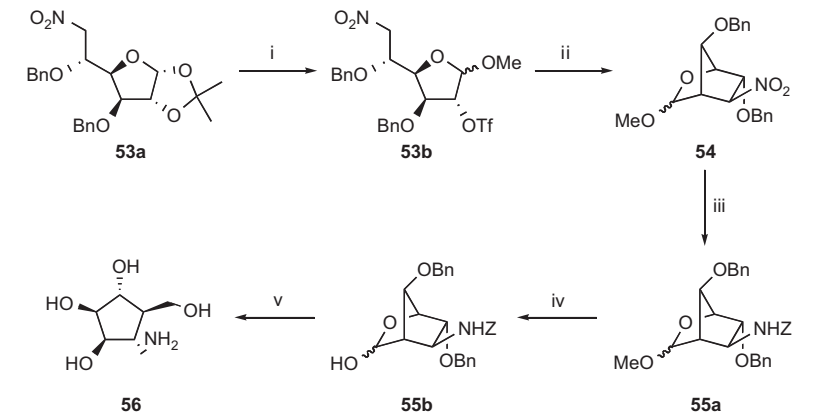
For example, a Michael addition of nitromethane to α,β-unsaturated ester **41** yielded nitroester **42** (Scheme 6). An intramolecular nitroaldol reaction of compound **42** with DABCO in refluxing benzene afforded nitrocyclohexane **43** in high yield, as the only diastereomer. This remarkable diastereoselectivity was also observed in the transformation of the sugar derivative **44** into nitrocyclopentane **45** and nitrocyclohexane **46**, when proceeding as stated in Scheme 6.[37]

On the other hand, nitro olefins are powerful electrophilic species that react with nucleophiles to give 1,2-addition products, which can then react intramolecularly with aldehyde groups to render carbocycles. This strategy was recently used for the preparation of carbasugar rancinamycin (**52**).[38] Michael addition of 1,3-dithiane to sugar nitro olefin **47** provided a separable mixture of epimers **48** and **49**. The major isomer **49** was subjected to a reaction protocol including an intramolecular Henry reaction to give a mixture of epimeric carbocycles **50a** and **51**, which when reacted with 2,2-dimethoxypropane and PTSA produced only its derivative **50b** only. Finally, treatment of **50b** with MeI and aqueous sodium hydrogen carbonate gave compound **52**, as a result of the removal of the carbonyl protecting group followed by a kinetically and thermodynamically favoured $E_{1C}B$ elimination of the nitro group (Scheme 7).

C-Alkylation of nitronates has also been applied to the transformation of nitro sugars into carbasugars. Although *O*-alkylation is usually favoured, *C*-alkylation is the predominant process when it is the thermodynamically



Scheme 7 *Reagents and conditions:* (i) THF, −78 °C; (ii) 75% AcOH, reflux; (iii) 2% aq. $HNaCO_3$, MeOH; (iv) $(CH_2)_3C(OMe)_2$, *p*-TsOH, acetone; (v) MeI, aq. $HNaCO_3$, MeCN, 35 °C.



Scheme 8 *Reagents and conditions:* (i) a) AcCl/MeOH; b) Tf_2O, py, CH_2Cl_2; (ii) TBAF, THF; (iii) a) H_2, Raney Ni, MeOH; b) $NaCO_3H$, $ClCO_2Bn$, AcOEt; (iv) TFA, H_2O; (v) a) $NaBH_4$, MeOH; b) H_2, Pd/C, AcOH.

favoured process. Thus, as stated in Scheme 8 similar to the synthesis of aminocyclopentitol shown in Scheme 6 an stereocontrolled intramolecular alkylation of the nitronate of the sugar trifil derivative **53b**, provided the novel bicyclic compound **54**. Catalytic hydrogenation of **54** afforded the corresponding amino acetal, which was directly reacted with benzyl chloroformate and sodium bicarbonate to obtain its derivative **55a**. Subsequent hydrolysis of this bicyclic glycoside with trifluoroacetic acid was followed by the immediate reduction using $NaBH_4$ of the implicit carbonyl group of the

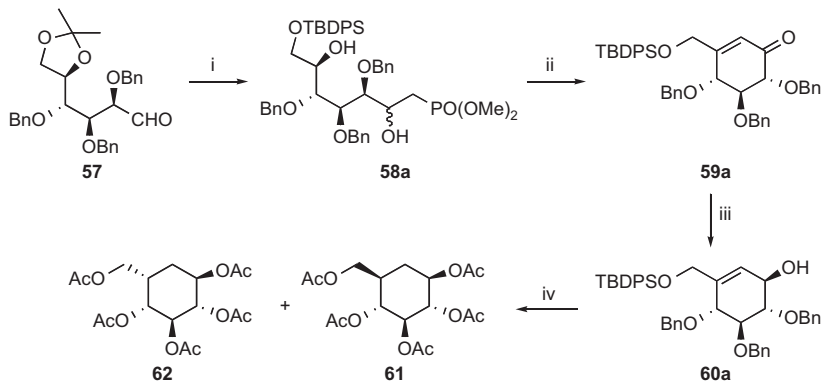
resulting bicyclic hemiacetal **55b** and final deprotection, resulting in the amino carbasugar **56** (Scheme 8).[39]

This synthetic methodology was also applied to the preparation of several polihydroxylated cyclopenthyl β-amino acids.[40]

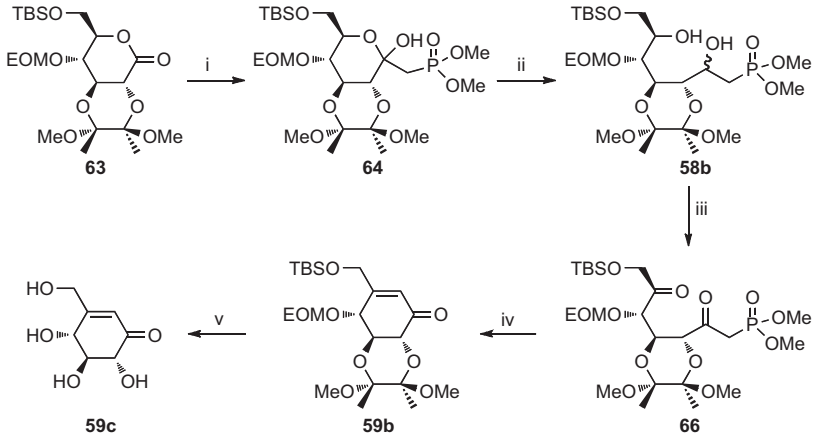
2.1.3 Reactions of phosphorus-stabilised species. The cyclisation of carbanions stabilised by phosphonate neighbouring groups have been extensively used in the synthesis of carbapyranoses from carbohydrates. Special attention was devoted to the intramolecular Horner-Wadsworth-Emmons olefination (HWE)[41] since this strategy for the preparation of sugar-derived carbocycles was introduced by Paulsen *et al*. (Scheme 9).[41a] Reaction of the glucose-derived aldehyde **57** with lithium dimethyl methyl phosphonate and further manipulation of protecting groups afforded the intermediate **58a**. Swern oxidation followed by spontaneous cyclisation yielded enone **59a**, which upon treatment with sodium borohydride produced the alcohol **60a** as the major product. *O*-Desilylation, catalytic hydrogenation, *O*-debenzylation and acetylation converted **60a** into 5a-carba-β-D-glucopyranose pentaacetate (**61**) and 5a-carba-α-L-idopyranose pentaacetate (**62**).

A HWE olefination was the key step in a recent synthesis of gabosine I from D-glucuronolactone described by Shin *et al*.[42] D-Glucuronolactone was differentially protected as a *trans*-diacetal at *C*-2,3, an ethoxymethyl ether at *C*-4 and a *tert*-butyldimethylsilyl ether at *C*-6, to give its derivative **63** (Scheme 10). A nucleophilic addition of lithiated dimethyl methyl phosphonate to the lactone carbonyl afforded phosphonate-lactol **64** in 96% yield. The hemiacetal was reduced by sodium borohydride and the resulting derivative **58b** was subjected to a Swern oxidation to produce the diketone **66**. An intramolecular HWE olefination was effected in the same pot to give enone **59b** in 78% overall yield. Complete deprotection of **59b** left gabosine I (**59c**).

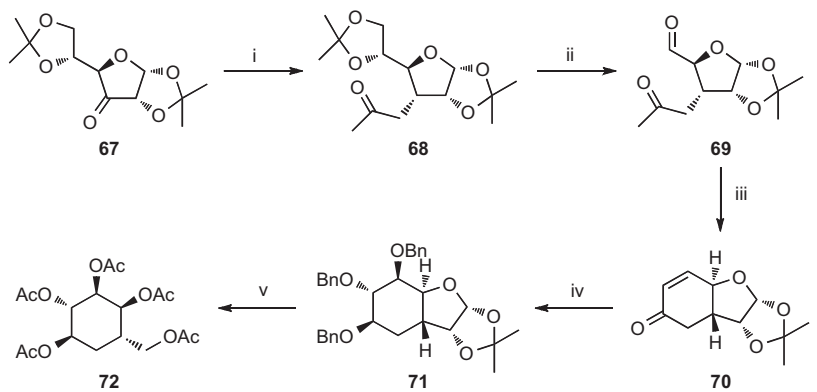
2.1.4 Aldol cyclisations. Aldol condensations are important carbon–carbon bond formation tools in organic synthesis. Tadano *et al*. introduced



Scheme 9 *Reagents and conditions:* (i) a) $CH_3P(O)(OMe)_2$, *n*-BuLi, THF, −78 °C. b) AcOH, 60 °C. c) TBDPSCl, imidazole, 85% (three steps). (ii) $(COCl)_2$, DMSO, Et_3N, −78 °C, 50%. (iii) $NaBH_4$, 87%. (iv) a) TBAF, THF. b) H_2, Ni-Ra, dioxane. c) H_2, Pd/C. d) Ac_2O, py, 54% (four steps).

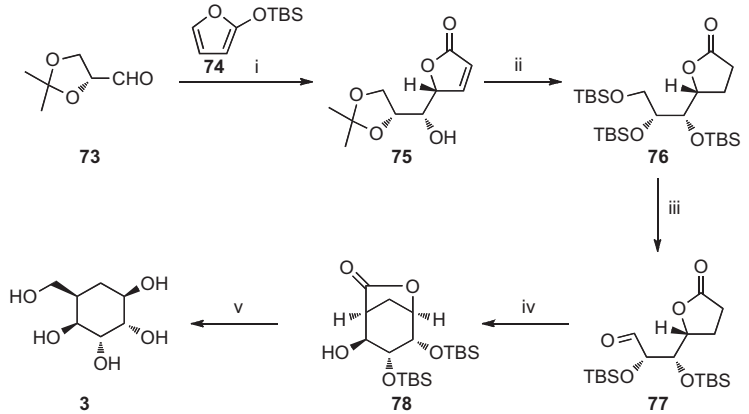


Scheme 10 *Reagents and conditions:* (i) $CH_3P(O)(OMe)_2$, LDA, THF, −78 °C, 96%. (ii) $NaBH_4$, MeOH, 0 °C, 96%. (iii) a) TFAA, DMSO, CH_2Cl_2, −78 °C. b) DIPEA, −78 °C. (iv) TEA, LiCl, r.t., 78%. (v) TFA, H_2O, r.t., 96%.



Scheme 11 *Reagents and conditions:* (i) a) $Ph_3P{=}CHCOCH_3$, benzene, 96%; b) H_2, Raney-Ni, MeOH; c) PCC, CH_2Cl_2, 88%; (ii) a) 60% aq. AcOH, 95%; b) $NaIO_4$, H_2O, MeOH, quant.; (iii) DBU, benzene then Ac_2O, py, 44%. (iv) a) H_2O_2, NaOH, MeOH, 96%; b) $NaBH_4$, EtOH, 84%; c) 2-methoxyethanol, H_2O, NaOAc; d) NaH, BrBn, DMF; (v) a) AcOH, H_2O, dioxane then $NaBH_4$, EtOH; b) $NaIO_4$, H_2O, MeOH; c) Ac_2O, py, 44%.

this methodology for the preparation of several carbasugars (Scheme 11).[43] Thus, a Wittig olefination of compound **67** with acetylmethylene-triphenylphosphorane, followed by catalytic hydrogenation and oxidation provided ketone **68**. Hydrolysis of the 5,6-*O*-isopropylidene moiety, followed by oxidative cleavage of the diol afforded di-carbonyl compound **69**, which on treatment with DBU followed by acetic anhydride and pyridine afforded the intramolecular aldol condensation product **70**. Epoxidation of **70**, reduction with sodium borohydride and epoxide ring opening with a hydroxide anion gave, after benzylation of the free alcohols, the diaxial opening product **71**. Finally, glycol cleavage, reduction, removal of the protecting groups, and acetylation yielded 5a-carba-α-L-altropyranose pentaacetate (**72**).

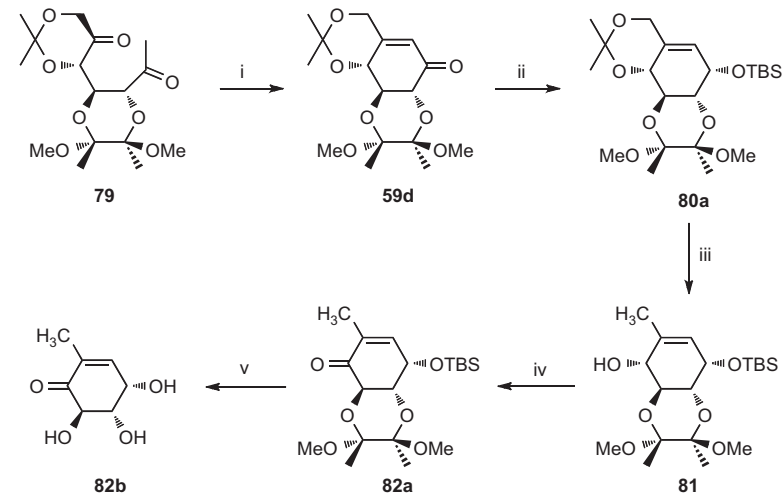


Scheme 12 *Reagents and conditions:* (i) $BF_3 \cdot Et_2O$, CH_2Cl_2, −80 °C, 75%; (ii) a) H_2, Pd/C, THF, 91%; b) TBSOTf, 2,6-lutidine, CH_2Cl_2, 90%; c) 70% aq. AcOH, 50 °C, 96%; d) TBSCl, imidazole, py, 70%; (iii) a) 80% aq. AcOH, 80%; b) $(COCl)_2$, DMSO, Et_3N, CH_2Cl_2, −80 °C, 97%; (iv) LDA, THF, −80 °C, 51%; (v) a) TESOTf, DMAP, py, 95%; b) $LiAlH_4$, THF, 0 °C, 90%; c) 6N HCl/THF/MeOH, 95%.

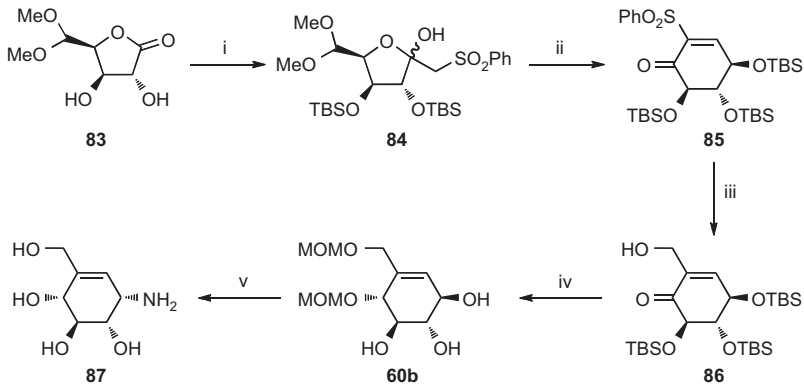
A synthesis of 5a-carba-α-D-gulopyranose (**3**) from D-glyceraldehyde was later reported. As shown in Scheme 12, protected derivative D-glyceraldehyde **73** was reacted *via* a vinylogous cross aldolisation with furan silyloxy diene **74** to selectively give the key alcohol **75**. Reduction of the double bond, followed by silylation of the free hydroxyl, removal of the isopropylidene protecting group and full protection of the free hydroxyls with silyl groups, lead to **76**. Selective desilylation of the primary hydroxyl, followed by Swern oxidation afforded the di-carbonyl compound **77**, which under basic conditions was transformed into the bicyclic lactone **78** as a single stereoisomer. Silylation of the generated alcohol, reduction of the lactone with $LiAlH_4$ and desilylation, gave 5a-carba-β-D-gulopyranose (**3**).[44]

More recently, several gabosines have been synthesised from D-glucose *via* an intramolecular direct aldol reaction of a 2,6-diketone.[45] Thus, an intramolecular aldol cyclisation of D-glucose derived diketone **79**,[46] provided enone **59d**, that was reduced with K-selectride and the resulting alcohol treated with TBSCl and imidazole to give a silyl ether **80a** (Scheme 13). The isopropylidene blocking group in **80a** was selectively removed with 80% aqueous acetic acid to afford the corresponding diol, which on regioselective mesylation followed by treatment with Super hydride ($LiEt_3BH$) afforded alcohol **81**. PDC oxidation of **81** gave enone **82a** in a high yield. (+)-Gabosine A (**82b**) was the final result after acidic hydrolysis of **82a**.

Tatsuta and co-workers described a procedure for the cyclisation of carbohydrates to carbasugar derivatives based on a $SnCl_4$-promoted aldol-like cyclisation of silylenol ethers.[47] Thus, D-xylonolactone derivative **83** was reacted with lithiated methyl phenyl sulfone to give furanose **84** (Scheme 14). Compound **84** was converted into cyclohexenone **85** by ring opening with TBSOTf and $SnCl_4$-promoted intramolecular reaction. A Michael reaction with tributylstannyl lithium followed by the trapping of the produced anion with formaldehyde and subsequent desulfonylation



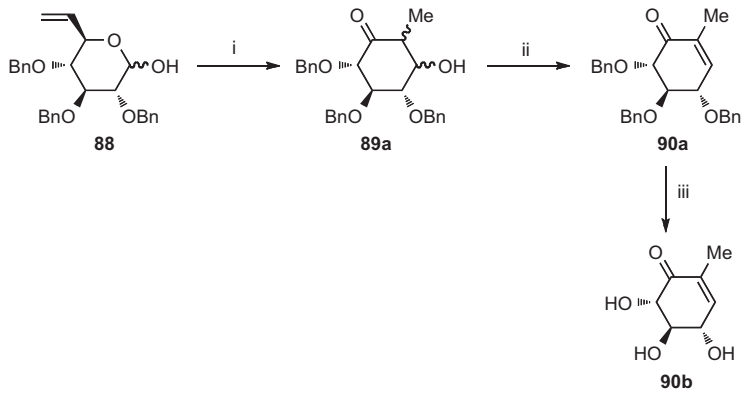
Scheme 13 *Reagents and conditions:* (i) a) L-Proline, DMSO, 83%; b) $POCl_3$, py, 99%; (ii) a) K-selectride, THF, −78 °C, 99%; b) TBSCl, imidazole, DMF, 95%; (iii) a) 80%, AcOH, 88%; b) MsCl, 2,4,6-collidine, CH_2Cl_2, −78 °C, 92%; c) $LiEt_3BH$, THF, −78 °C, 84%; (iv) PDC, 3A, MS, CH_2Cl_2, 92%; (v) TFA, H_2O, CH_2Cl_2, 94%.



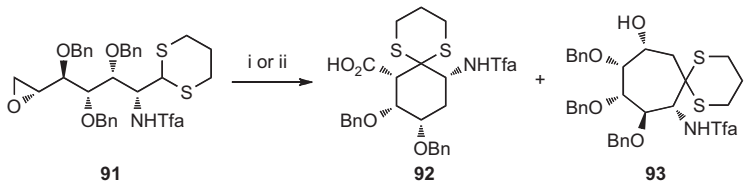
Scheme 14 *Reagents and conditions:* (i) $MeSO_2Ph$, *n*-BuLi/THF, −78 °C, 0.5 h, 94%; (ii) a) TBSOTf, 2,6-lutidine/CH_2Cl_2, 40 °C, 48 h, 92% b) $SnCl_4/CH_2Cl_2$, −78 °C, 3 h, 70%; (iii) n-Bu_3SnLi, HCHO/THF, −78 °C to 40 °C, 72 h, 84%; (iv) a) $Zn(BH_4)_2$/Ether, 0 °C, 1 h, 80%; b) MOMCl, n-Bu_4NI, DIPEA CH_2ClCH_2Cl, 50 °C, 24 h, 85%; c) TBAF/THF, rt, 3 h, 97%; (v) a) HN_3, Ph_3P, DEAD/THF, rt, 1 h, 81%; b) H_2, Raney-Ni, H_2O, 1,4-dioxane, quant; c) 3% HCl, MeOH, 99%.

yielded the α-hydroxymethylcyclohexenone **86**. Stereoselective reduction of the carbonyl group in **86** followed by protection and deprotection steps led to cyclohexene **60b**. Valienamine (**87**) was produced from **60b** by Mitsunobu inversion of the allyl alcohol and deprotection.[47b]

An iron-catalysed tandem isomerisation–intramolecular aldolisation of vinyl pyranoses reaction was developed for the preparation of carbohydrate-derived cyclohexenones. Thus, treatment of D-glucose derived vinyl pyranose **88** with 10 mol% of $Fe(CO)_5$ under irradiation with a HPK 125



Scheme 15 *Reagents and conditions:* (i) $Fe(CO)_5$ (10 mol%), THF, hν, 95%; (ii) MsCl, Et_3N, CH_2Cl_2, 69%; (iii) $FeCl_3$, CH_2Cl_2, 55%.



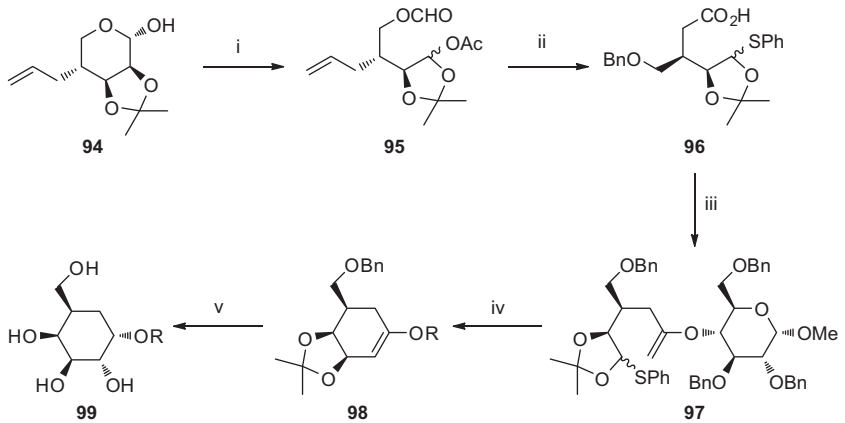
Scheme 16 *Reagents and conditions:* (i) *n*-BuLi, THF, −10 °C; (ii) *n*-BuLi, THF, 1 eq. LiBr, −90 °C to −50 °C.

lamp afforded cyclohexanone **89a** as a mixture of stereoisomers (Scheme 15).[48] Intermediate **90a** was obtained from **89a** through the corresponding mesylate. Then, removal of the benzyl groups with $FeCl_3$ in CH_2Cl_2 afforded the desired 4-*epi*-gabosine A (**90b**).

2.1.5 Sulfur-stabilised carbanions. The Corey–Seebach method was used for the intramolecular anionic cyclisation of sugar derivatives,[49] as in the case of the D-glucosamine derivative **91**. Treatment with 2.5 eq. of n-BuLi in THF, resulted in cyclisation to the carba analogue of 1-amino-1-deoxy-β-L-*galacto*-pyranose derivative **92**, accompanied by a certain amount of the carba analogue of the 1-amino-1-deoxy-α-D-*altro*-septanose derivative **93** (Scheme 16).[50] Under standard Corey–Seebach conditions, compounds **92** and **93** were obtained in 54% and 23% yield, respectively. However, when LiBr was added, the cyclization was speeded up and the ratio of the six-membered product increased significantly, probably *via* a Li-chelated intermediate (62% yield for compound **92** and 8% yield for compound **93**).

2.2 Carbocation cyclisations

Even though cyclisation processes involving carbanions are more common, methods making use of carbocations also proved to be successful in the synthesis of carbasugars. For example, Mootoo *et al*. reported a synthesis of β-carba-*galacto*-disaccharides involving a key step consisting of the



Scheme 17 *Reagents and conditions:* (i) DIB, I_2, cyclohexane, 95%; (ii) a) PhSH, $BF_3 \cdot OEt_2$, −78 °C, then NaOMe, MeOH; b) NaH, BnBr, TBAI, DMF, 70% (two steps); c) O_3, MeOH-CH_2Cl_2, −78 °C, then Ph_3P; d) $NaClO_2$, CH_3CN, 2-methyl-2-butene, 59% two steps; (iii) DCC, DMAP, methyl 2,3,6-tri-O-benzyl-α-D-glucopyranoside, 80%; (iv) a) Tebbe reagent, 80%; b) MeOTf, DTBMP, CH_2Cl_2, 64%; (v) a) $BH_3 \cdot SMe_2$, then H_2O_2, NaOH, 72%; b) HCl, MeOH; c) Pd/C, EtOH, HCO_2H, 72% (two steps).

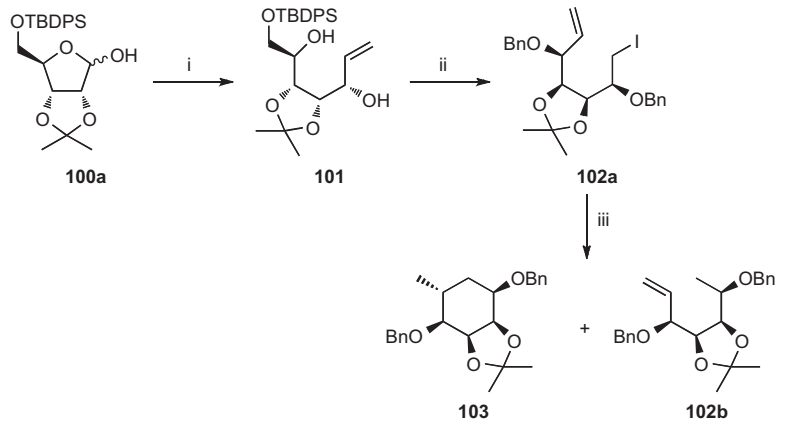
formation of the carbasugar ring *via* an oxycarbenium ion-enol ether cyclisation.[51] Thus, a Suarez fragmentation (Scheme 17) of D-lyxose-derived branched pyranoside **94** resulted in the formation of lactol **95**, which was transformed into the mixture of acids **96**. A DCC-mediated esterification was employed to introduce the glycon moiety and the key enol ether **97** was obtained after a Tebbe reaction. Activation with methyl triflate in the presence of 2,6-di-*tert*-butyl-4-methylpyridine led to the cyclic enol ether **98**. Finally, treatment of **98** with $BH_3 \cdot SMe_2$ and deprotection steps provided the carbadisaccharide **99**.

2.3 Radical cyclisations

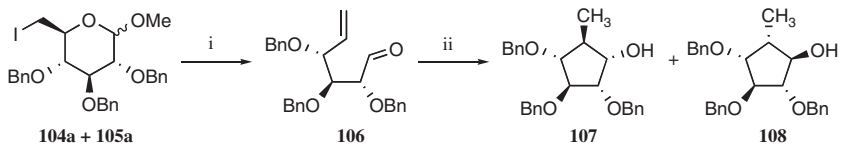
Radical cyclisations have been extensively used to convert carbohydrates into carbocycles.[52] The most extended strategy consists of opening the sugar, transforming the carbonyl into a radical acceptor (alkene or alkyne) and converting one of the hydroxyl groups into a nucleophilic radical. On other hand, the most common radical generators are tributyl tin hydride and samarium diiodide.

An early contribution was a comprehensive study on the application of the radical cyclisation of hept-1-enitols to the preparation of 5a-carbasugars published by Redlich and co-workers in 1992.[53] Starting from D-ribose derivative **100a**, they prepared different substrates and studied the outcome of the *n*-Bu_3SnH/AIBN promoted radical cyclisation reaction. For example, radical cyclisation of iodoalkene **102a** afforded cyclohexane **103**, together with some deiodination product **102b** (Scheme 18).

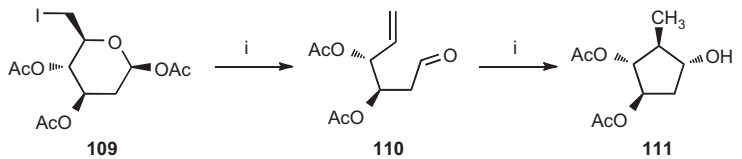
On the other hand, it is known that the reductive dealkoxyhalogenation of 6-deoxy-6-iodoglycosides (Grob fragmentation) gives rise to ring-opened hex-5-enals,[54] which are suitable substrates for ketyl-olefin radical cyclisation. Thus, 5-hexenals obtained by treatment of the corresponding methyl 6-deoxy-6-iodoglycosides with powdered zinc have been employed as



Scheme 18 *Reagents and conditions:* (i) vinylmagnesium bromide, THF, 79%; (ii) a) NaH, BrBn, DMF, 62%; b) TBAF, THF; c) I_2, PPh_3, imidazole, 79%; (iii) Bu_3SnH, AIBN, benzene.



Scheme 19 *Reagents and conditions:* (i) Zn, EtOH; (ii) SmI_2, THF.

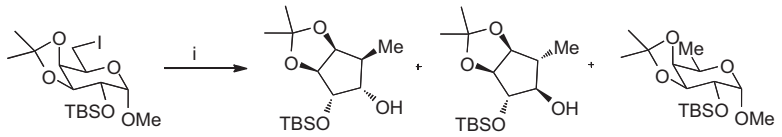


Scheme 20 *Reagents and conditions:* (i) SmI_2, THF.

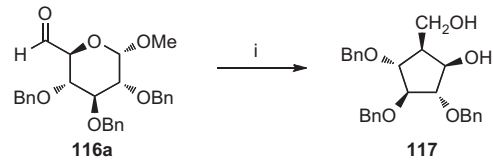
substrates for a SmI_2-mediated radical cyclisation, allowing the overall transformation of pyranoses into cyclopentanes. For example, reaction of D-glucose derived 5-hexenal **106** with SmI_2 under dilute conditions in the presence of a proton source provided the desired cyclopentanols **107** and **108** in a 2:1 ratio and in good yield (Scheme 19).[55]

A modified and more efficient one-pot procedure, in which Grob fragmentation was executed under the action of samarium diiodide, was successfuly applied to the preparation of cyclopentane derivative **111**, as depicted in Scheme 20.[56]

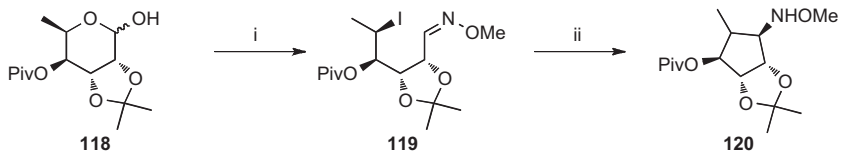
Similarly, Chiara *et al.* reported a cascade reaction of 6-deoxy-6-iodohexopyranosides promoted by samarium diiodide.[57] A series of 6-deoxy-6-iodohexopyranosides with different configurations and substitution patterns were treated with SmI_2 in THF/HMPA to give the desired cyclopentanols. For example, reaction of iodopyranoside **112** with SmI_2 in THF and HMPA afforded cyclopentanols **113** and **114**, together with the dehalogenation product **115** (Scheme 21).



Scheme 21 *Reagents and conditions:* (i) SmI_2, HMPA, THF.



Scheme 22 *Reagents and conditions:* (i) SmI_2, HMPA, THF, *t*-BuOH.



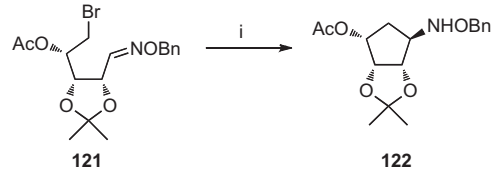
Scheme 23 *Reagents and conditions:* (i) a) NH_2OMe, MeOH; b) I_2, PPh_3, toluene; (ii) SmI_2, HMPA, THF. 72%.

6-Formyl methyl pyranosides undergo a ring contraction induced by samarium(II) iodide that results in the formation of highly functionalised cyclopentanes. Thus, treatment of methyl pyranoside **116a** with samarium(II) iodide in the presence of HMPA and *tert*-butyl alcohol yielded cyclopentane **117** (Scheme 22).[58]

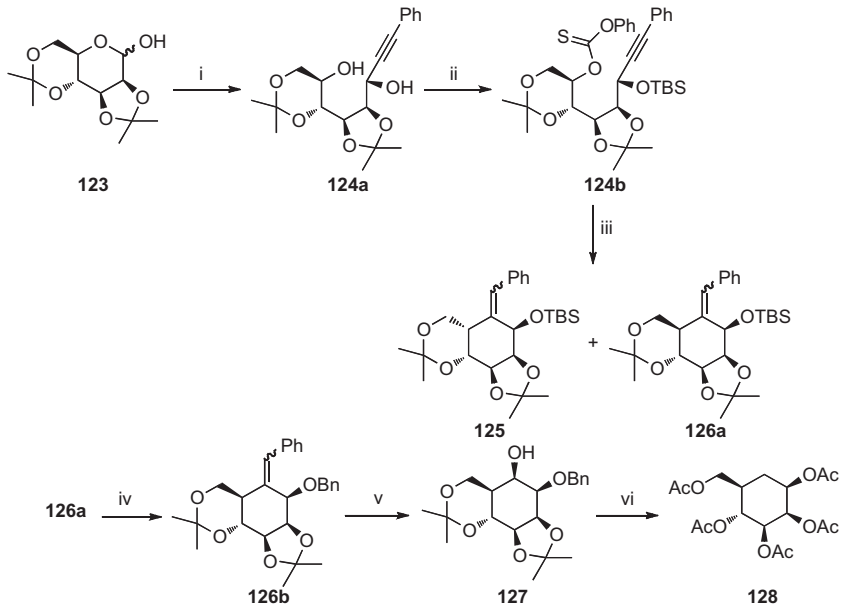
An efficient method for the preparation of highly functionalised aminocyclopentanes from readily available carbohydrate templates was reported in 1997. In this procedure, a radical generated by treatment of an iodoalkane and samarium diiodide is trapped intramolecularly by an oxime ether, thus constructing a cyclopentane ring.[59] For example, carbohydrate hemiacetal **118** was condensed with *O*-methylhydroxylamine to give, after iodination of the released hydroxyl group, iodo oxime **119**. The samarium diiodide-mediated reductive cyclisation of **119** in a THF solution containing HMPA furnished the alkoxy aminocyclopentane **120** in good yield (Scheme 23).

A full paper published by Chiara *et al*. in 1997, describes a series of SmI_2-mediated intramolecular reactions that involve trapping radicals by sugar derived oxime ethers, comparable with the transformation of oxime derivative **121** into amino cyclopentane **122** (Scheme 24).[60]

Particularly relevant is the work carried out by Gómez *et al*., designed to allow a direct transformation of carbohydrates into the corresponding carbasugars.[61] A representative example depicted in Scheme 25 involves a 6-*exo-dig* radical cyclisation involving the alkyne moiety (radical acceptor)



Scheme 24 *Reagents and conditions:* (i) SmI_2, HMPA, THF, −40 °C, 61%.



Scheme 25 *Reagents and conditions:* (i) PhCCLi, THF, −78 °C, 65%; (ii) a) TBSCl, Et_3N, DMAP, CH_2Cl_2, 65%; b) ClC(S)OPh, py, CH_3CN, 80%; (iii) *n*-Bu_3SnH, AIBN, toluene, 90 °C, 95%; (iv) a) TBAF, THF, 91%; b) NaH, BrBn, Bu_4NI; (v) a) O_3, MeOH/CH_2Cl_2 (1:1), −78 °C, then Me_2S; b) $BH_3 \cdot SMe_2$, THF, 75% (3 steps); (vi) a) NaH, CS_2, MeI; b) *n*-Bu_3SH, AIBN, toluene, 85% (2 steps); c) H_2, Pd/C, MeOH; d) AcOH-THF-H_2O (4:2:1), 60 °C; e) Ac_2O, py, 85% (3 steps).

and the thionocarbonate moiety (radical precursor) of the D-mannose derivative **124b**.[59b] This was easily prepared by treatment of D-mannose diacetonide **123** with lithium phenylacetylide and then converted, in two steps, into phenyl thionocarbonate **124b**. Treatment of **124b** with *n*-Bu_3SnH and AIBN resulted in a radical cyclisation leaving a separable 3:2 mixture of *exo* methylenecyclohexanes **125** and **126a**. Cyclohexane **126a** was transformed into its derivative **126b**, which upon ozonolysis followed by reduction of the resulting ketone lead to cyclohexane **127**. Finally, compound **127** was transformed into carbasugar **128** in five steps.

2.4 Pinacol couplings

Pinacol coupling of carbonyl compounds has been extensively studied[62] and can be promoted by a wide range of organometallic reagents and

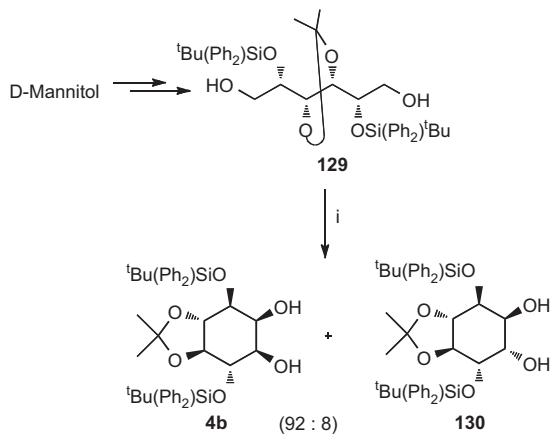
low-valence lanthanide salts such as samarium diiodide, activated manganese, titanium derivatives and zinc powder.[63] This ketyl radical cyclisation seemed to be particularly suitable for the preparation of polyhydroxylated carbocyles from carbohydrates and specifically those promoted by samarium diiodide, which have been extensively used in carbohydrate chemistry.[64] This reducing agent has proven to be particularly useful for promotion of intramolecular pinacol coupling reactions of polyoxygenated substrates under very mild conditions, in high yield and with high diastereoselectivity.

2.4.1 Intramolecular pinacol coupling of aldehydes and ketones. After being reported in 1991 that carbohydrate-derived dialdehydes undergo a pinacol coupling reaction in the presence of SmI_2 to give *cis*-diols as the preponderant products,[65] this intramolecular coupling reaction was extensively applied to the preparation of polyhydroxylated cycloalkanes and derivatives of biological interest, such as inositols,[66] carbasugars[67] and aminocycloalkanes.[68] For example, in 1994 Chiara *et al.* reported that a tandem Swern oxidation/SmI_2-mediated intramolecular pinacol coupling of D-mannitol derivative **129** resulted in a 98:2 mixture of inositols **4b** and **130** (Scheme 26).[66a]

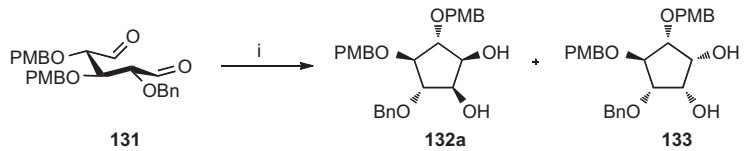
A samarium diiodide-mediated intramolecular pinacol coupling of a carbohydrate derived dialdehyde was also the key step in a route designed for the preparation of cyclopentane-based congeners of *myo*-inositol.[66c] Treatment of dialdose **131** with an excess of samarium diiodide and *t*-BuOH in THF gave a 1:3 mixture of cyclopentitols **132a** and **133** (Scheme 27).

The key step of the synthesis of the natural carbasugar caryose, reported in 1997, was also a samarium diiodide promoted pinacol coupling.[67a] Thus, the SmI_2-mediated cyclisation of ketoaldehyde **135**, derived from 3,4,5-tri-*O*-benzyl-l-deoxy-D-iditol **134**, followed by a mild oxidation of the *cis*-l,2-diol **132b** and a diastereoselective allylation of the ketol **136**, gave **137**, a precursor of caryose (**138**) (Scheme 28).

Recently, a samarium diiodide-mediated intramolecular pinacol coupling of a carbohydrate-derived dialdehyde was also the key step in a route

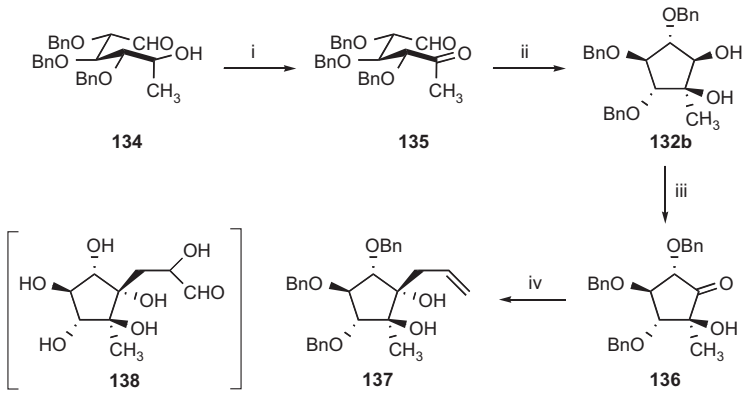


Scheme 26 *Reagents and conditions:* (i) $(COCl)_2$, DMSO, Et_3N, CH_2Cl_2. (ii) SmI_2, *t*-BuOH, THF.



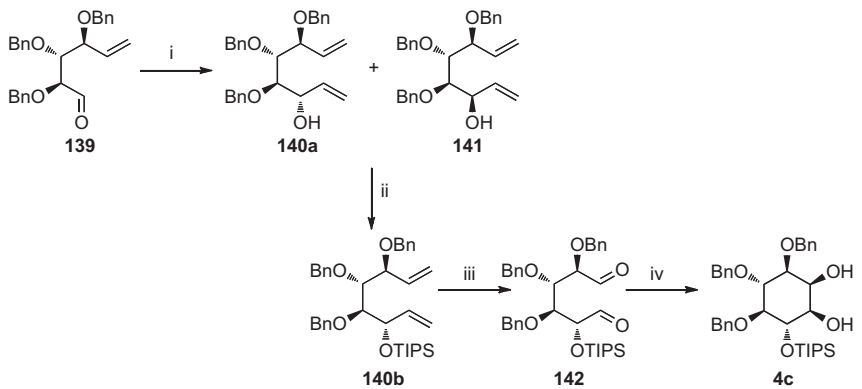
131 132a 133

Scheme 27 *Reagents and conditions:* (i) SmI_2, *t*-BuOH, THF, 43%.



134 135 132b 136 137 138

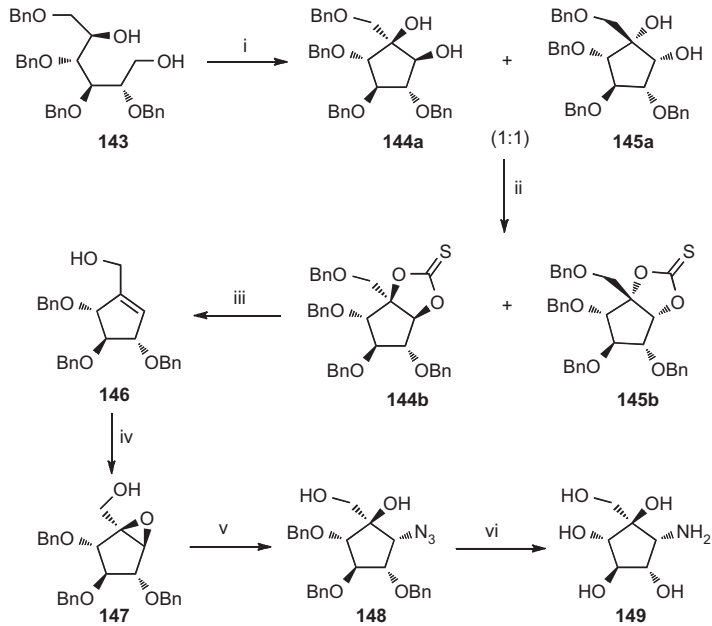
Scheme 28 *Reagents and conditions:* (i) $(COCl)_2$, DMSO, Et_3N, CH_2Cl_2; (ii) SmI_2, *t*-BuOH, THF; (iii) TEMPO, KBr, 0.9M NaOCl, CH_2Cl_2; (iv) AllylTMS, $TiCl_4$, CH_2Cl_2.



139 140a 141 140b 142 4c

Scheme 29 *Reagents and conditions:* (i) CH_2=CHMgBr, THF, −78 °C; (ii) a) TIPSCl, imidazole, DMF; b) column chromatography; (iii) a) O_3, CH_2Cl_2, −78 °C; b) Me_2S; iv) SmI_2, −78 °C; then sat. aq. $NaHCO_3$, −78 °C to 20 °C.

designed for the preparation of *myo*- and *chiro*-inositols from D-xylose.[69] Dienols **140a** and **141**, prepared from D-xylose derivative **139**,[70] were silylated and the silylated mixture was separated by silica gel chromatography to produce pure **140b**. Treatment of **136b** with ozone produced the corresponding dialdehyde **142**, which upon treatment with 6 equivalents of SmI_2 at −78 °C for 20 minutes, followed by dropwise addition of saturated aqueous $NaHCO_3$ at −78 °C and then warming to 20 °C, resulted in the formation of diol **4c** (Scheme 29).

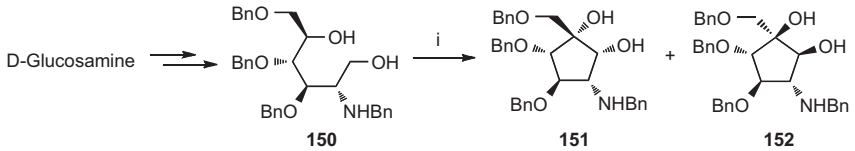


Scheme 30 *Reagents and conditions:* (i) a) $(COCl)_2$, DMSO, THF, NEt_3, −60 °C; b) SmI_2, *t*-BuOH, THF, r.t., 90%. (ii) 1,1-thiocarbonyldiimidazole, toluene, 110 °, 97%. (iii) a) Ac_2O, TMSOTf. b) $(EtO)_3P$, heat. c) NaOMe, MeOH, 82%. (iv) L-DIPT, $Ti(Oi\text{-}Pr)_3$, *t*-BuOOH, CH_2Cl_2, 94%. (v) LiN_3, NH_4Cl, DMF, 125 °C, 89%. (vi) H_2, $Pd(OH)_2$, EtOH/THF/TFA, 67%.

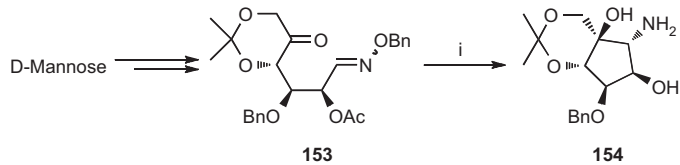
One of the strategies used for the preparation of polyhydroxylated aminocyclopentanes from carbohydrates is based on the intramolecular pinacol coupling of sugar aldehydes and/or ketones and subsequent introduction of the amino group. Thus, Swern oxidation of diol **143**, derived from D-glucose, followed by dropwise addition of the crude mixture into a solution of samarium diiodide in THF containing *t*-BuOH, gave a 1:1 mixture of cyclitols **144a** and **145a**. From this mixture, a sequence involving formation of the cyclopentene from the cyclic thiocarbonate, epoxidation, ring opening with an azide and hydrogenolysis afforded the desired amino cyclopentanol **149** (Scheme 30).[68a]

Another useful procedure for the preparation of polyhydroxylated aminocyclopentanes would be the intramolecular pinacol coupling of glucosamine-derived aldehydes. The oxidation-pinacol coupling sequence was used for the preparation of the 2:1 mixture of amino cyclopentitols **151** and **152** from glucosamine-derived diol **150**. The protocol consisted of a one-pot Swern oxidation followed by SmI_2-mediated pinacol coupling (Scheme 31).[71]

2.4.2 Intramolecular pinacol coupling of aldehydes or ketones and imino derivatives. The tandem oxidation-pinacol coupling reaction has been widely used for the construction of the cycloalkane ring from carbohydrate diols. As the nitrogenated function is not present in the newly created bonds, it either has to be present in the sugar skeleton or it has to be introduced by functionalisation of the cycloalkane ring. However, the process whereby a carbonyl is coupled to an imino derivative would allow



Scheme 31 *Reagents and conditions:* (i) a) $(COCl)_2$, DMSO, THF, NEt_3, $-60\,^{\circ}C$; b) SmI_2, *t*-BuOH, THF, r.t., 82%.



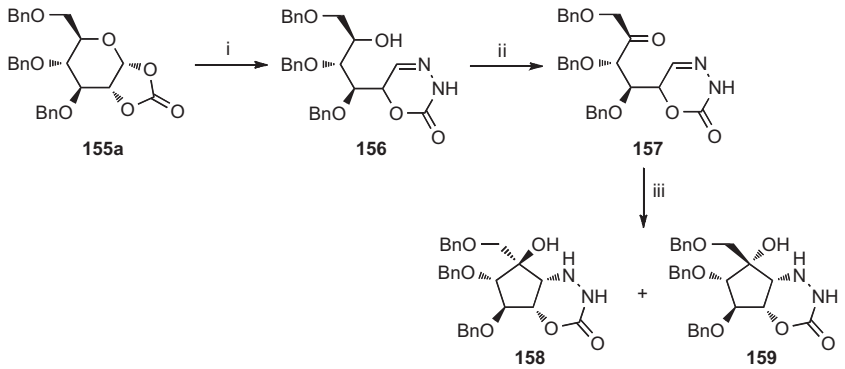
Scheme 32 *Reagents and conditions:* (i) a) SmI_2, *t*-BuOH, $-30\,^{\circ}C$; b) H_2O; c) LiOH, 98% (three steps, one pot).

the creation of the carbocyclic ring with concomitant formation of the 1,2-aminoalcohol functionality and accordingly, the samarium diiodide mediated intramolecular pinacol coupling of oxime ethers and aldehydes or ketones would allow the direct formation of 2-aminocycloalkan-1-ols.[72] Chiara *et al*. extensively studied the application of this strategy to the preparation of aminocyclopentitols from carbohydrates.[73] In a recent example, they synthesised a new trehazolin analogue **154** from carbohydrate precursors, by a highly efficient route based on their previously developed ketone/oxime ether reductive carbocyclisation reaction for the construction of the cyclitol ring.[74] Starting from D-mannose, condensation with *O*-benzylhydroxylamine and subsequent oxidation of the released hydroxyl group, afforded **153**, a suitable substrate for the samarium diiodide-promoted cyclization to cyclopentane **154** (Scheme 32).

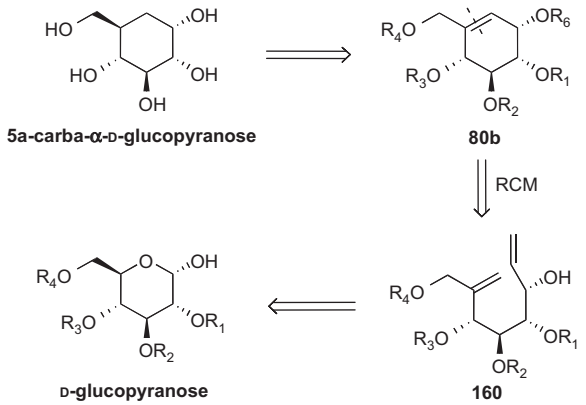
The pinacol coupling is highly diastereoselective, with the *trans*-1,2-aminoalcohols formed preferentially, because formation of *cis*-1,2-aminoalcohols is disfavoured by 1,3 diaxial strain in the transition state. As sometimes the preparation of amino alcohols with *cis* stereochemistry is required, some protocols were developed aimed at changing the stereoselectivity of the reaction in order that *cis*-1,2-aminoalcohols could preferentially be obtained.[75] A recent strategy developed to overcome the 1,3 diaxial strain and hence obtain *cis*-amino alcohols is the intramolecular tethering of the C=N group to the vicinal hydroxyl group.[76] Thus, treatment of 1,2-cyclic carbonate **155a** with hydrazine produced hydrazone **156**, which upon oxidation gave substrate **157**. Treatment of **157** with samarium diiodide gave the expected cyclopentitol in 65% yield as a 7:1 mixture of isomers **158** and **159** (Scheme 33).

2.5 Ring-closing metathesis

Metathesis reactions have been widely used with carbohydrate-related substrates for the synthesis of new carbohydrate derivatives, heterocycles, glycomimetics or glycoconjugates compounds.[77] With regard to carbasugar



Scheme 33 *Reagents and conditions:* (i) NH_2NH_2 HCl, DIEA, EtOH, 68%; (ii) Dess-Martin, CH_2Cl_2, 51%; (iii) SmI_2, *t*-BuOH, THF, −30 °C, 65%.

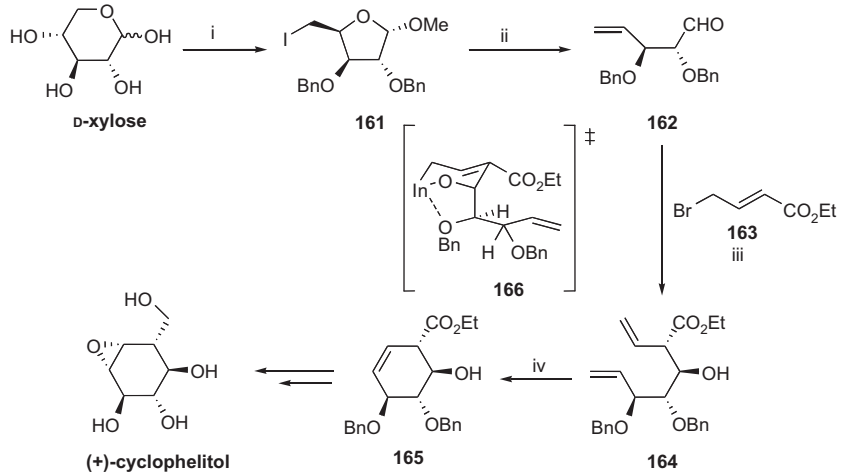


Scheme 34 Retrosynthetic analysis of carbasugar from carbohydrate.

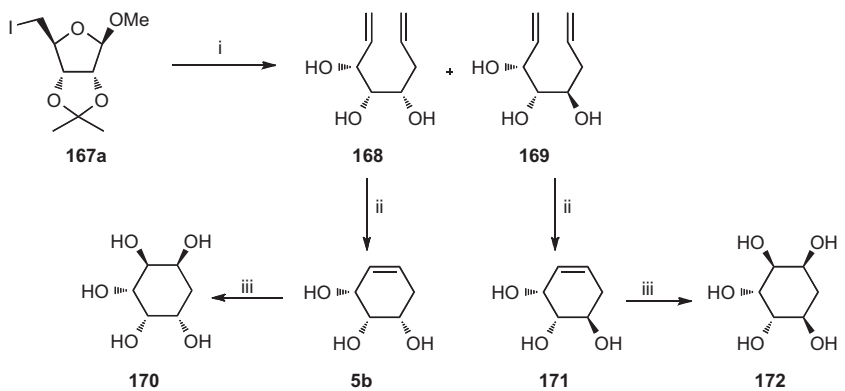
synthesis, the most applied methodology so far, is the ring-closing metathesis (RCM) reaction.[78] This is due to the fact that it permits the easy formation of the carbasugar's ring from a diene acyclic precursor **160**, which is normally obtained by simple modifications to the equivalent carbohydrate derivative (Scheme 34).

Many different diene derivatives have been developed as precursors for the synthesis of carbasugars by means of RCM reactions several general reviews on synthesis of carbasugars have been published in the recent years marking a large emphasis on the applications of this methodology for the preparation of the cycloalkane ring.[79] However, this synthetic strategy is of great scope and today is still very widely used to carry out the synthesis of new carbasugars.

The most challenging step in the preparation of carbasugars by RCM from carbohydrate-based dienes is the introduction of double bonds in the starting structure. We will see below the most actual approaches for the preparation of carbohydrates-based dienes and their application in the synthesis of carbasugar by mean of RCM reactions.



Scheme 35 *Reagents and conditions:* (i) a) MeOH, $HCl_{(aq)}$, 50 °C, b) I_2, PPh_3, imidazole, THF, 65 °C, c) $BnOC(NH)Cl_3$, TfOH, dioxane, rt; (ii) Zn, THF, H_2O, 40 °C, sonication; iii) 163, In, H_2O, $La(OTf)_3$, 48 h; (iv) Grubbs' catalyst II, CH_2Cl_2.



Scheme 36 *Reagents and conditions:* (i) Zn or In, allyl bromide, THF, H_2O, sonication; (ii) Grubb's catalyst, CH_2Cl_2, 91%; (iii) OsO_4, CH_2Cl_2.

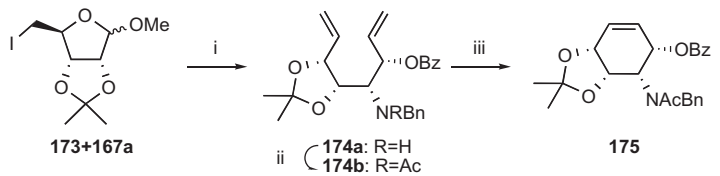
One classical transformation in carbohydrate chemistry to introduce terminal olefins in the non-reducing end is the zinc-mediated fragmentation of primary iodo-derivatives originally developed by Bernet and Vasella.[80] Madsen *et al*. used this protocol in 2005 for the synthesis of (+)-cyclophelitol (Scheme 35).[81] Thus, a Zn-promoted reductive elimination of iodo-derivative **161** provided compound **162**, with the terminal carbon–carbon double bond. Consequently, an indium-mediated allylation at the reducing end of the carbohydrate, produced diene **164**, which after reaction with a 2nd generation Grubbs' catalyst, allowed the construction of the cyclohexene ring of compound **165**.

This protocol has proven to be very useful for the synthesis of carbohydrate-based dienes, and was used for the direct preparation of similar dienes from ω-iodo glycosides (Scheme 36).[82] In this one-pot reaction, the resulting

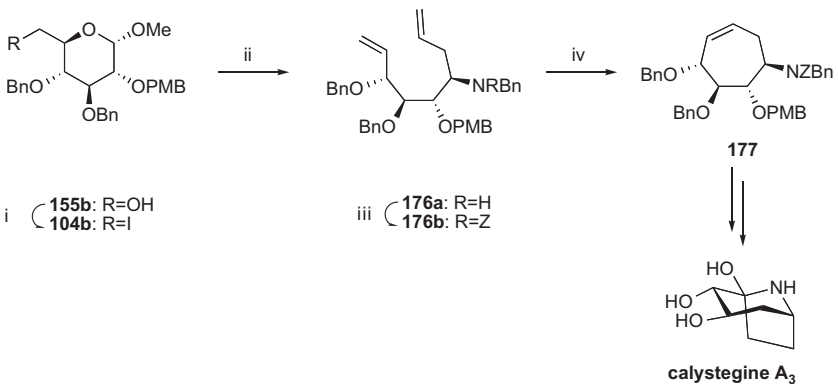
aldehyde was trapped by allyl bromide in a Barbier-type alkylation and zinc serves a dual function, promoting both the reductive elimination and activation of allyl bromide for the allylation. This domino zinc-mediated transformation would convert the iodo glycosides directly into functionalized dienes that can be subsequently converted into carbocycles by RCM. For example, a domino reaction with 5-iodo-ribofuranosides as **167a** in the presence of zinc or indium and allyl bromide under sonication gave full conversion and very high yield of the desired dienes **168** and **169.** A RCM reaction gave cyclohexene **5b** and hydroxylation of its carbon-carbon double bond provided carbasugar **170.** On the other hand, diene **165** was converted into carbasugar **172** *via* cyclohexene **171**.[83]

Moreover, the intermediate aldehyde obtained in this process can be intercepted with an amine prior to the allylation. The allylation then takes place on the formed imine and results in the introduction of an amino group in the final product. Thus, treatment of **173** + **167a** with zinc, benzylamine and allyl benzoate bromide succeeded in giving the amino diene **175** in (Scheme 37).

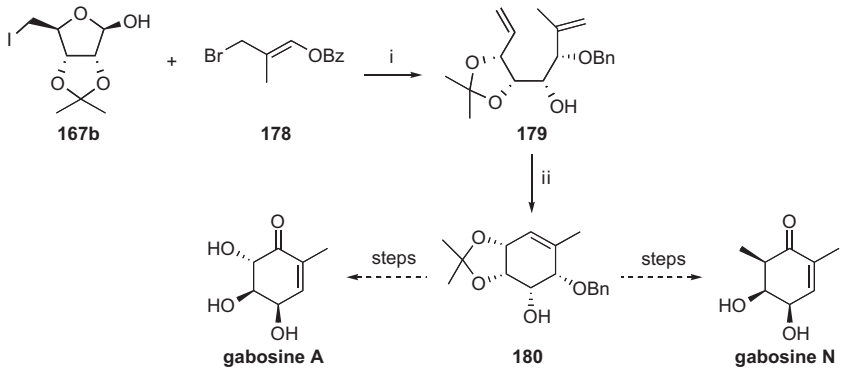
This procedure was recently used by Madsen *et al.* to develop the synthesis of the diene **176a**, which was easily transformed by RCM into the seven-membered ring of carbasugar **177** (Scheme 38), the main intermediate in the



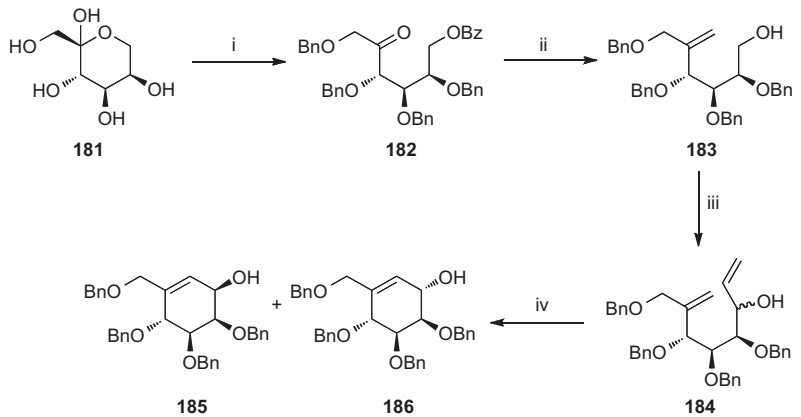
Scheme 37 *Reagents and conditions:* (i) Zn, TMSCl, THF, ultrasound, 40 °C, then $BnNH_2$, 3 Å MS, THF, rt, then Zn, $BzOCHCHCH_2Br$, THF, ultrasound, 40 °C; (ii) Ac_2O, Et_3N, DMAP, CH_2Cl_2, 40 °C; (iii) $(C_3H_4N_2Mes_2)Cl_2Ru{=}CHC_6H_4OC_3H_7$, toluene, 80 °C.



Scheme 38 *Reagents and conditions:* (i) I_2, PPh_3, imidazole, THF; (ii) Zn, TMSCl, THF, sonication, then $BnNH_2$, then $CH_2{=}CHCH_2Br$, sonication; (iii) ZCl, $KHCO_3$, H_2O, CH_2Cl_2; (iv) Grubbs' catalyst II, CH_2Cl_2.



Scheme 39 *Reagents and conditions:* (i) Zn, THF, sonication, then **178**, sonication; (ii) Grubbs' catalyst II, CH_2Cl_2.

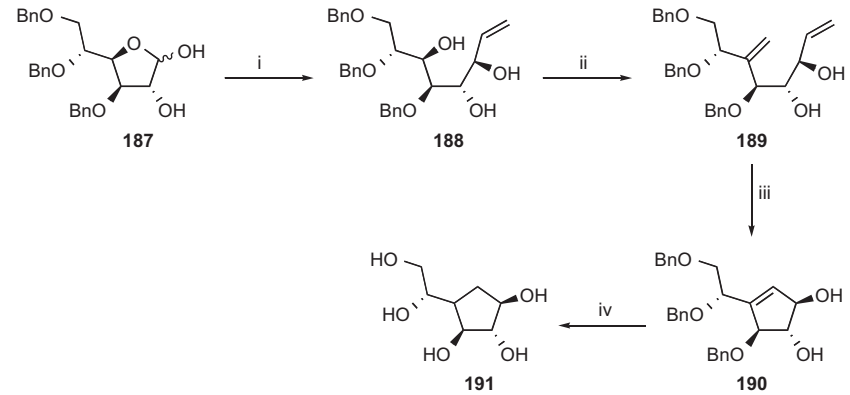


Scheme 40 *Reagents and conditions:* (i) a) 2-bromoethanol, 89%; b) Pd/C, H_2, $NaHCO_3$, H_2O, 50 °C; c) BnBr, NaH, DMF, 74% from **181**; d) AcOH, $HCl_{(aq)}$, 60 °C, 85%; e) BzCl, py, 86%; (ii) a) Ph_3PCH_3Br, *t*-BuOK, PhMe, 0 °C; b) NaOMe, MeOH, 84%; (iii) a) $(COCl)_2$, Me_2SO, CH_2Cl_2, −78 °C, then Et_3N, −78 °C to rt; b) vinylmagnesium chloride, THF, 0 °C, 76%; (iv) Grubbs' catalyst II, toluene, 60 °C; 28% **185** + 59% **186**.

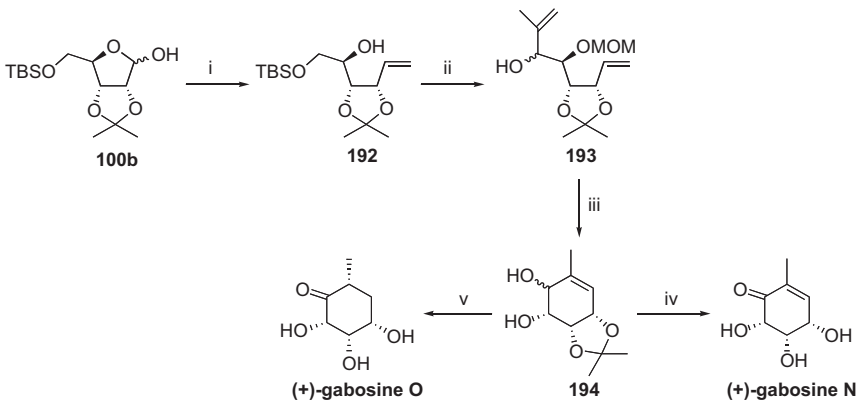
synthesis of calystegine A_3.[84] A similar strategy was also used for the synthesis of diene **179**, the RCM precursor of gabosine A and gabosine N (Scheme 39).[85]

Another typical alternative to introduce two carbon-carbon double bonds into a carbohydrate molecule deals with the employment of the Wittig reaction and the asymmetric addition of proper Grignard reagents to the conveniently functionalised carbohydrate molecule. In this way, an interesting synthesis of two carbocyclic mannose mimics **185** + **186** was developed from D-fructopyranose **181** through a synthetic pathway where the olefinic double bonds were sequentially introduced by a Wittig olefination and an addition of vinylmagnesium chloride (Scheme 40).[86]

A similar protocol was recently used by Frigell *et al*. during the synthesis of compound **189** (Scheme 41), the RCM diene precursor of 4a-carba-α-galactofuranose **191**, from the D-glucose derivative **187**.[87]



Scheme 41 *Reagents and conditions:* (i) vinylmagnesium chloride, THF, rt, 17 h, 87%; (ii) a) 2-methoxypropene, PPTS, CH_2Cl_2, rt, 15 min, 94%; b) DMSO, $(COCl)_2$, CH_2Cl_2, Et_3N, rt, 80%; c) Ph_3PMeBr, *n*-BuLi, toluene, rt, 81%; d) AcOH, H_2O, 2 h, quant.; (iii) Grubbs' catalyst II, toluene, 48 h; 87%; (iv) H_2, Pd/C, EtOAc, EtOH, rt, 1.5 h, quant.



Scheme 42 *Reagents and conditions:* (i) $Ph_3P{=}CH_2$, THF, 4 h, 76%; (ii) a) MOMCl, DIPEA, DMAP, CH_2Cl_2, 12 h, 93%; b) TBAF, THF, 4 h, 95%; c) $(COCl)_2$, DMSO, CH_2Cl_2, Et_3N, 2 h; d) 2-bromopropene, $CrCl_2$, $NiCl_2$, DMF, 12 h, 72% or 2-bromopropene, Mg, THF, 4 h, 85%; (iii) Grubbs' catalyst II, toluene, 12 h, 85%; (iv) a) PDC, CH_2Cl_2, 4 Å MS, 12 h, 82%; b) Amberlyst, THF/H_2O, 5 h, 75%; (v) a) PDC, CH_2Cl_2, 4 Å MS, 12 h, 82%; b) H_2, Pd/C, MeOH, 1 h, 95%; c) Amberlyst, THF/H_2O, 5 h, 85%.

In this process, both carbon-carbon double bonds were sequentially introduced in the glucose unit. First, the asymmetric addition of vinylmagnesium chloride to hemiacetal **187** provided compound **188** in good yield and diastereoselectivity. Then, after proper protection, the oxidation of the hydroxyl group in position C-4 followed by a Wittig olefination produced diene **189**. This compound undergoes RCM with the 2nd generation Grubbs' catalyst providing cyclopentene **190** in excellent yield, which after hydrogenation with Pd yielded 4a-carba-α-D-galactofuranose **191**.

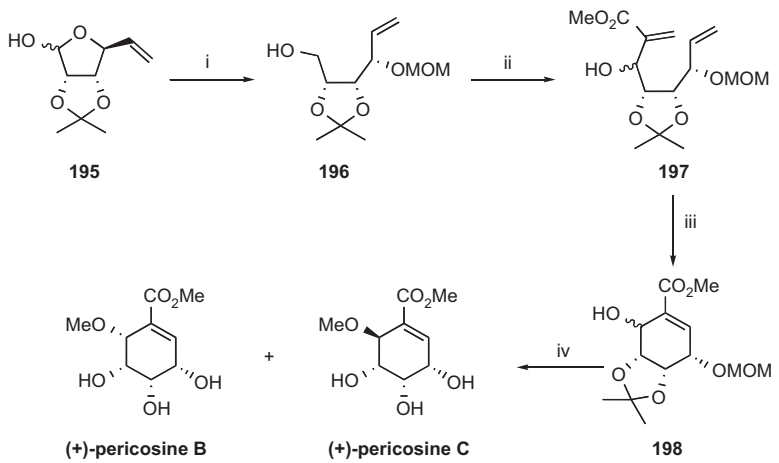
In a similar way, Rao *et al.* obtained from D-ribose the carbasugar **194**, a key compound in the synthesis of natural products (+)-gabosine N and (+)-gabosine O (Scheme 42) from D-ribose.[88] Unlike the previous case, the Wittig olefination was employed first in the transformation of the reducing

end of the furanoside **100b.** The primary hydroxyl group of compound **192** was then oxidised and submitted to an asymmetric addition of an organometallic derivative of 2-bromopropene. The obtained diene **193** was then efficiently transformed in carbasugar **194** by mean of a 2nd generation Grubbs' catalysed RCM reaction.

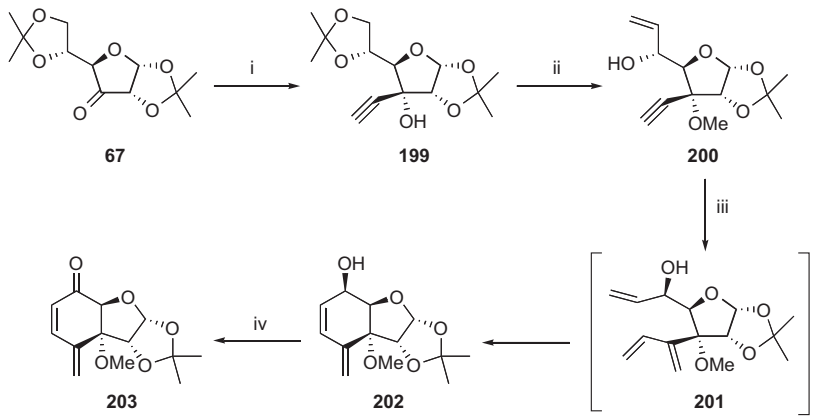
Another different protocol that allows the preparation of complex RCM diene precursors from commercially available carbohydrates is based on the asymmetric aldol reaction of ketones with acrylates. This procedure was recently used by several authors during the synthesis of natural products (+)-pericosine A, (+)-pericosine B and (+)-pericosine C as stated in Scheme 43, which involved a synthetic pathway where the asymmetric addition of methyl acrylate takes place during the transformation of derivative **196** to diene **197**. This compound easily undergoes the formation of a cyclohexene ring *via* RCM reaction.[89]

Finally, another novel alternative for introducing carbon-carbon double bonds into a carbohydrate-based substrate can be carried.[90] In this way, Betkekar *et al*. recently described an elegant approach to the 4-methylene-2-cyclohexenone subunit present in otteliones and loloanolides derivatives, by way of a strategy that involved a tandem enyne/ring closing metathesis as the key reaction (Scheme 44).[91]

Therefore, the addition of lithium acetylene to the carbonyl group of gluco-D-furanose derivative **67** provided in good yield and diastereoselectivity the compound **199**. Then, a terminal double bond was introduced by the epoxidation of the 5,6-diol system and the regioselective opening of the epoxide moiety. This was developed by lithium trimethylsilane, producing enyne **200** which was transformed in cyclohexenol **202** through an efficient tandem enyne/ring closing metathesis with a 2nd generation Grubbs' catalyst in the presence of ethylene. The tandem enyne/ring closing metathesis takes place first by an enyne metathesis reaction of the alquine moiety of **200**



Scheme 43 *Reagents and conditions:* (i) $NaBH_4$, MeOH, 97%; (ii) a) PivCl, Py, CH_2Cl_2, 51%; b) MOMCl, DIPEA, CH_2Cl_2, 89%, c) NaOMe, MeOH, 95%; d) CrO_3, Py, Ac_2O, CH_2Cl_2, e) methyl acrylate, DABCO, DMF; (iii) Grubbs' catalyst II, toluene, 86%; (iv) a) Ag_2O MeI, THF, Me_2S, 91–95%, b) TFA, MeOH, 81–88%.



Scheme 44 *Reagents and conditions:* (i) acetylene, *n*-BuLi, THF, 78%; (ii) a) NaH, MeI, THF 96%, b) AcOH, H_2O, 96%, c) *p*-TsCl, py, d) NaOMe, MeOH, 67%, e) $(CH_3)_3Si$, *n*-BuLi, THF, 95% (iii) Grubbs' catalyst II, ethylene, CH_2Cl_2, 83%; (iv) DMP, CH_2Cl_2 91%.

with one equivalent of ethylene, providing the triene intermediate **201**. This compound thenundergoes an RCM reaction which generates the cyclohexene system of compound **203**.

2.6 Cycloaddition reactions

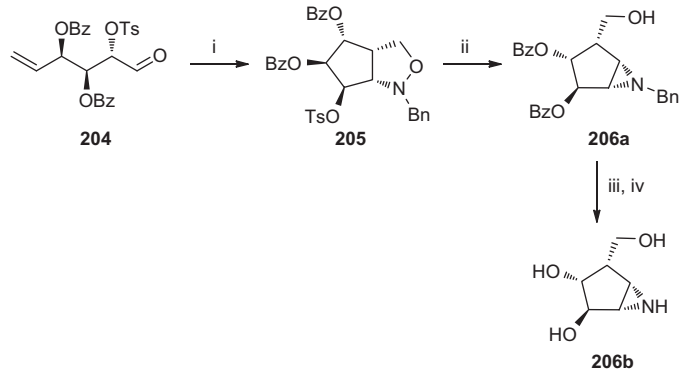
Intramolecular cycloaddition reactions of 1,3-dipoles have been used for the construction of carbocyclic natural products, including carbohydrate derivatives.[92]

Thus, the [3 + 2] cycloaddition of oximes or nitrones bearing a double bond was used for the formation of the cyclopentane ring of several biologically relevant 1-amino-2-hydroxymethyl-cyclopentitols.[93] For example, alk-5-enyl aldehyde **204** was easily transformed into the corresponding oxime, which upon thermolysis resulted in the isolation of hexahydro-1*H*-cyclopent[c]isoxazol **205** in good yield, *via* intramolecular oxime olefin cycloaddition. Catalytic hydrogenation afforded aziridine **206a**, which finally gave **206b** after removal of the protecting groups (Scheme 45).[94]

Recently, an intramolecular ketonitrone-olefin cycloaddition reaction was used for the synthesis of aminocyclopentitol **209** (Scheme 46).[95]

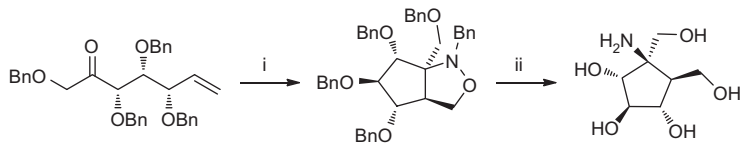
Using a 1,3-dipolar reaction as the key step, Gallos and co-workers synthesised *ent*-gabosine E. Condensation of *N*-methylhydroxylamine with lactol **210a** led to the formation of cycloadducts **211** and **212** in a 2:1 ratio (Scheme 47), *via* the corresponding nitrone.[96] The major diastereoisomer **211** was then converted to *ent*-gabosine E in a five further five steps sequence, in which the isoxazolidine was reductively opened and the amine eliminated to give the olefin in the desired product.

Nitronates and nitrile oxides can also undergo [3 + 2] cycloadditions,[97] allowing the construction of cycloalkane rings when the reaction occurs intramolecularly. This strategy was used for the carbocyclisation of sugar derivatives and for the preparation of pseudo-carbasugars.[98] A representative example is the synthesis of cyclophellitol **219,** reported in 2003.[99] A Henry reaction between D-glucose derived enal **106** and nitromethane



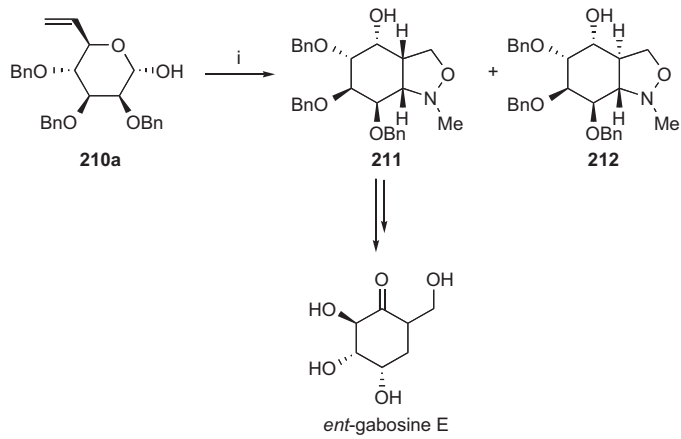
204 205 206a 206b

Scheme 45 *Reagents and conditions*: (i) BnNHOH.HCl, $CaCO_3$, toluene, 50 °C; (ii) H_2, Ni Raney, EtOAc/EtOH 3:1; (iii) NaOMe, MeOH; (iv) 10% Pd/C, H_2, EtOH.



207 208 209

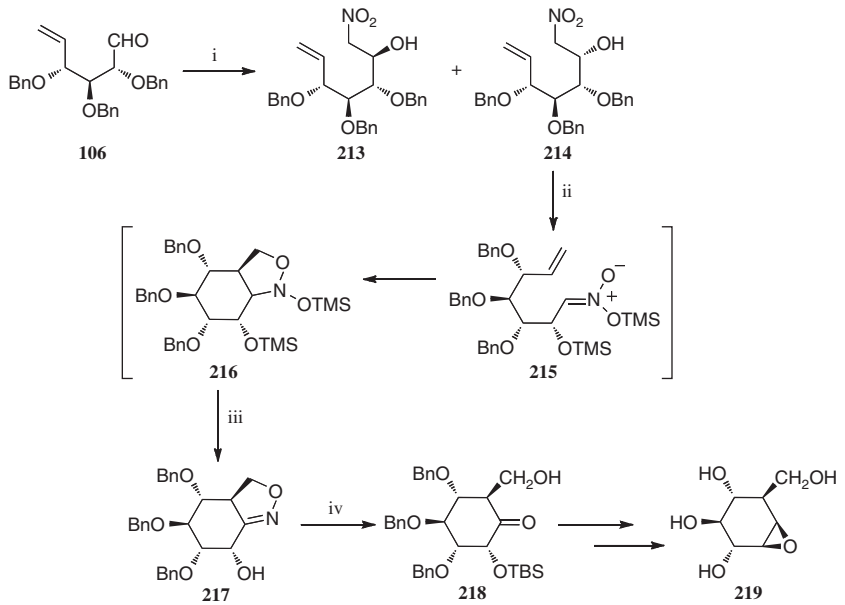
Scheme 46 *Reagents and conditions*: (i) BnNHOH.HCl, $CaCO_3$, toluene, 50 °C; (ii) H_2, Ni Raney, EtOAc/EtOH 3:1; (iii) NaOMe, MeOH; (iv) 10% Pd/C, H_2, EtOH.



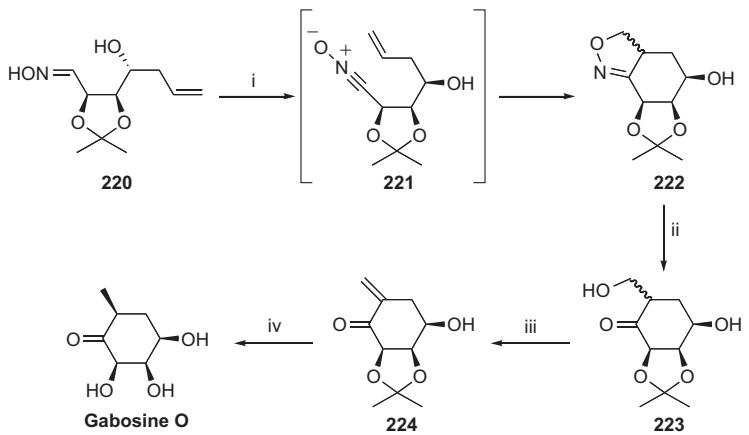
210a 211 212

Scheme 47 *Reagents and conditions:* (i) MeNHOH.HCl, NaOEt, EtOH, 20 °C, 80%.

provided nitro olefins **213** and **214**, which upon treatment with chlorotrimethylsilane may lead to cycloadduct **216** by a sequence including a spontaneous nitronate cycloaddition of the resulting intermediate **215**. *In situ* treatment of **216** with *p*-TsOH finally afforded bicyclic compound **217**. Silylation, and $Mo(CO)_6$-mediated reductive N-O bond cleavage afforded the product **218** was easily converted into cyclophelitol **219** in five steps (Scheme 48).



Scheme 48 *Reagents and conditions:* (i) CH_3NO_2, TMG, THF, 58% of **107a**; (ii) TMSCl, Et_3N, DMAP, THF; (iii) *p*-TsOH, THF, 55%. (iv) a) TBSOTf, Et_3N, CH_2Cl_2, 95%; b) $Mo(CO)_6$, MeCN, 90 °C.



Scheme 49 *Reagents and conditions:* (i) Chloramine T, silica gel, EtOH, 79%, α:β 4.6:1; (ii) H_2, Raney Ni, AcOH, 97%; (iii) Martin's sulfurane, THF, −78 °C; (iv) a) H_2, Raney Ni, EtOH/H_2O/1,4-Dioxane; b) TFA/H_2O, 89% (three steps).

Shing and coworkers reported that oxidation of oxime **220** with chloramine T ($TolSO_2NCl^-Na^+$) provided the intermediate nitrile oxide **221** that spontaneously underwent a cycloaddition to give the bicyclic adduct **222** (Scheme 49).[100] The dihydroisoxazole ring was cleaved with Raney nickel and hydrogen in acetic acid to give the β-hydroxy ketone **223**. Regioselective dehydration of the primary alcohol with Martin's sulfurane afforded enone

224. Catalytic hydrogenation with Raney-Ni followed by hydrolysis of the acetonide gave natural gabosine O.

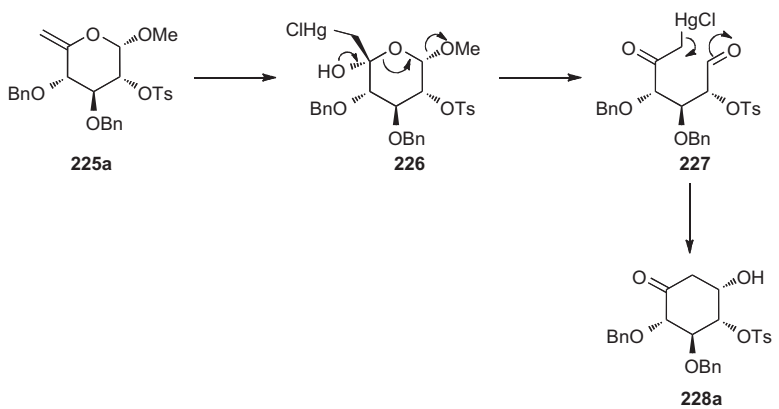
2.7 Rearrangement of cyclic sugars

The rearrangement of cyclic sugars allows their direct conversion into carbocycles. This straightforward approach to carbasugars relies on the structure of sugar enol-ethers, which inherently carry both masked nucleophilic and electrophilic functionalities.[101]

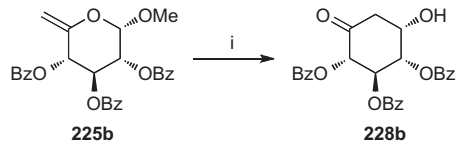
2.7.1 The Ferrier-II rearrangement. The first breakthrough for this class of reaction was the mercury(II)-mediated rearrangement of hex-5-enopyranoside derivatives, known as the Ferrier (II) rearrangement.[102] The representative example included in Scheme 50 show the mechanism for this transformation which involves a regioselective hydroxymercuration of hex-5-enopyranoside **225a** to give an unstable acetal **226** that decomposes to the ketoaldehyde **227** *via* the loss of methanol.[103] A subsequent intramolecular aldol-cyclisation affords cyclohexanone **228a** in high yields, as a single diastereomer (as in most cases). The stereochemistry of the newly formed stereogenic center linked to the hydroxyl group is dependent on the stereochemistry of the *C*-3 substituent: both groups are usually *trans*.

Since then, this reaction has provided a practical route to a large variety of bioactive substances such as aminocyclitols,[104] pseudosugars,[105] inositols[106] and other complex hexitols.[107]

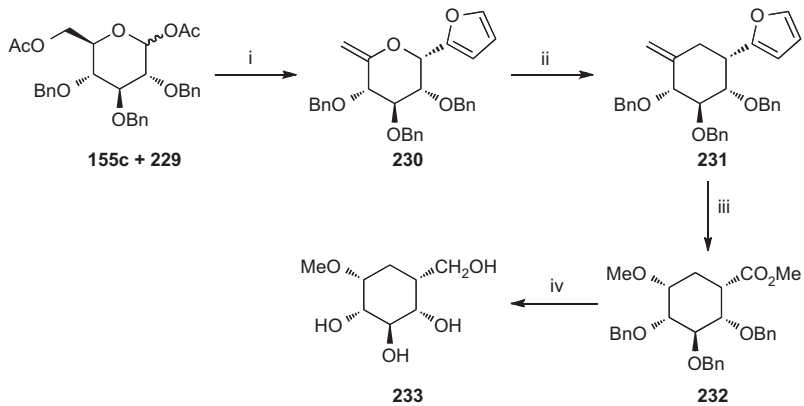
Despite the usefulness of this procedure, an important drawback is the toxicity of mercury, which prompted the development of nontoxic alternatives. Thus, Adam reported a similar rearrangement of amino-6-deoxyhex-5-enopyranoside derivatives mediated by $PdCl_2$ or $Pd(OAc)_2$ in the presence of aqueous sulfuric acid,[108] which was later applied to several 6-deoxyhex-5-enopyranosides.[109] For example, when benzoyl protected gluco-hexenopyranoside **225b** was treated with a catalytic amount of $PdCl_2$, cyclohexane **228b** was obtained in good yield and excellent diastereoselectivity ($\alpha/\beta > 99:1$) (Scheme 51).[110]



Scheme 50 *Reagents and conditions:* $HgCl_2$, acetone, reflux.



Scheme 51 *Reagents and conditions:* (i) $PdCl_2$, dioxane/H_2O 2:1, 60 °C, 68%.



Scheme 52 *Reagents and conditions:* (i) a) Furan, 4 Å MS, TMSOTf, CH_3CN, −40 °C to −20 °C, 71%; b) NaOMe, MeOH, 91%; c) I_2, PPh_3, imidazole, toluene, 70 °C, 89%; d) NaH, DMF, 72%; (ii) TIBAL, toluene, 50 °C, 83%; (iii) a) MeI, NaH, DMF, 96%; b) O_3, CH_2Cl_2, MeOH, −78 °C, then $KHCO_3$, MeI, DMF, 62%; (iv) a) $LiAlH_4$, Et_2O, 76%; b) H_2, Pd/C, MeOH, 75%.

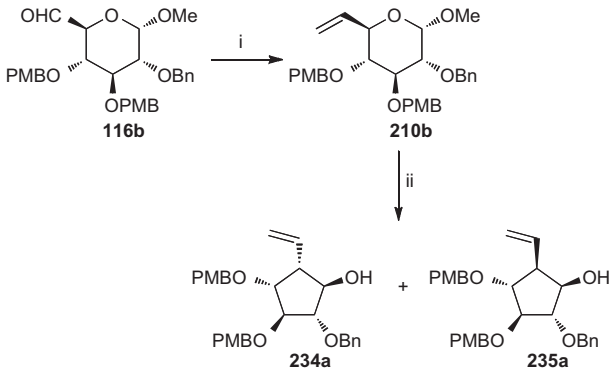
Another drawback of the Ferrier (II) reaction is that, despite being a really efficient way to form six-membered carbocycles, it is unsuitable to form cyclopentitols.

2.7.2 Endocyclic cleavage induced rearrangement. In order to overcome the limitations of the Ferrier rearrangement, different alternatives were investigated.

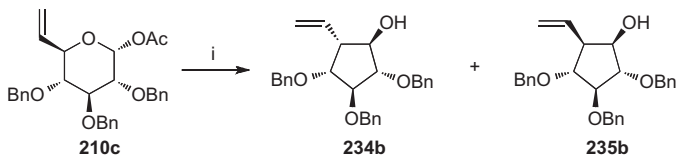
Prof. Pierre Sinay and co-workers reported an alternative which consists of a direct conversion of hex-5-enopyranosides into highly functionalised cyclohexane derivatives by reductive rearrangement induced by triisobutylaluminium (TIBAL).[111] Later on, Sinay and Sollogoub also demonstrated that the titanium-based Lewis acid Ti(O*i*Pr)Cl_3 was also able to mediate the rearrangement of *O*-glycosides under milder conditions.[112] For example *C*-furyl glycoside, derived from D-glucose **(155c + 229)**, was converted to a partially protected carba-α-D-idopyranoside **230** (Scheme 52).[113] The furyl group was used as a masked form of the hydroxymethyl moiety at *C*-5.

A remarkable feature of this type of rearrangements is the retention of the anomeric configuration. Accordingly, this methodology was employed in the synthesis of carbadisaccharides from disaccharides.[114]

On the other hand, modified carbohydrates, where the terminal carbon has been elaborated into a vinyl unit, are useful precursors of cyclopentanes *via* an intramolecular reaction between aldehyde and an allylmetal species generated *in situ*.[115] Thus, a zirconium-mediated ring-contraction is induced



Scheme 53 *Reagents and conditions:* (i) CH_3PPh_3Br, *t*-BuOK, THF; (ii) "Cp_2Zr"/THF then $BF_3 \cdot Et_2O$, −78 °C to rt.



Scheme 54 *Reagents and conditions:* (i) SmI_2, $Pd(Ph_3)_4$, THF, 80 °C, 82% **234b** and 12% **235b**.

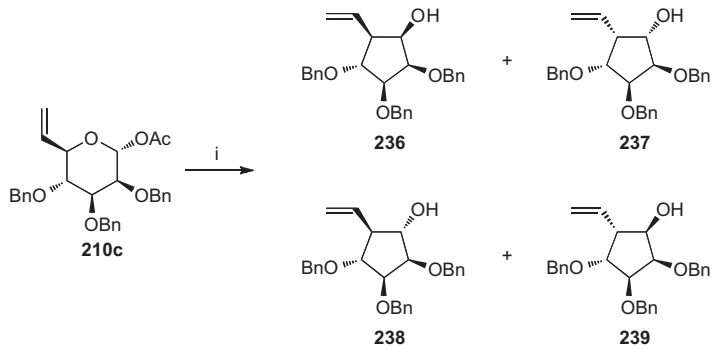
by sequential treatment of vinyl carbohydrate derivatives with "Cp_2Zr" and $BF_3 \cdot OEt_2$.[116] A recent example is the preparation of vinyl cyclopentanes **234a** and **235a** from sugar aldehyde **116b**, in which the key step is a ring contraction carried out by treating **210b** with Cp_2Zr-(*n*-Bu)$_2$ (prepared *in situ* from zirconocene dichloride and 2 equiv of *n*-butyllithium), followed by boron trifluoride etherate in THF (Scheme 53).[117]

Samarium(II) iodide in the presence of Pd(0) can also promote this transformation. An example is the transformation of 5-alkenylpyranoside **210c** into homoallyl cyclopentanols **234b** and **235b** (Scheme 54).[118] The first step is the formation of an allenylpalladium complex, followed by reduction by SmI_2 and subsequent intramolecular allylation of the generated carbonyl-tethered allenylsamarium.

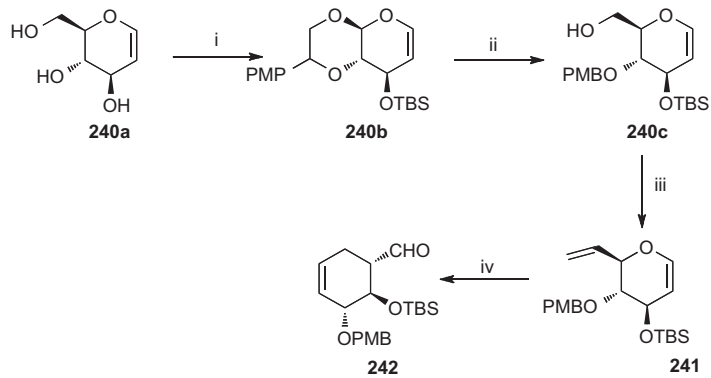
Comparing both processes, they differ in the stereochemical outcomes that are possible in the final cyclopentanes depending on the particular metal employed. Thus, allylzirconium intermediates lead to a *syn* relationship between vinyl and hydroxyl groups in the cyclised products, whereas allylsamarium intermediates preferentially yield the corresponding *anti* products.[119]

Recently, the possibility of using the Pd(0)/Et_2Zn system in this transformation was also investigated.[120] Upon treatment of vinyl carbohydrate **210c** with $Et_2Zn/Pd(PPh_3)_4/ZnCl_2$ the cyclopentane products **236**, **237**, **238** and **239** were obtained in good yields and diastereoisomeric ratio (Scheme 55).

2.7.3 Claisen rearrangement. Recently, a synthesis of antiviral Tamiflu from cheap and abundantly available D-glucal (**240a**) was described,



Scheme 55 *Reagents and conditions:* (i) Et_2Zn, Pd(0), $ZnCl_2$, THF, r.t., 6% of **224**, 6% of **225**, 75% of **226** and 13% of **227**.



Scheme 56 *Reagents and conditions:* (i) a) *p*-anisaldehyde diethyl acetal, PPTS, DMF, 2 h; b) TBSCl, imidazole, DMAP, DMF, 3 h; (ii) DIBAL-H, dichloromethane, −15 to 0 °C, 2 h, 65% (3 steps); (iii) Dess-Martin, dichloromethane, 2 h followed by methyltriphenylphosphonium bromide, *n*BuLi, THF, −78 to 25 °C, 1 h, 67% (2 steps); (iv) diphenyl ether, 210 °C, 2 h, 88%.

featuring a Claisen rearrangement to obtain the cyclohexene backbone (Scheme 56). Installation of a 4,6-benzylidene acetal and silylation of the 3-hydroxyl group gave fully protected D-glucal **240b**, which underwent selective opening of the benzylidene acetal with DIBAL-H in dichloromethane to give the free primary alcohol **240c.** The primary hydroxyl group in **240c** was oxidised to the aldehyde by using Dess–Martin periodinate and then subjected to Wittig methylenation to give terminal olefin **241** in 67% yield. The next step was the critical Claisen rearrangement reaction, which allowed ready access to carbocycle **242**.[121]

Acknowledgements

We thank the Spanish Ministry of Science and Innovation (CTQ2009-08490) and by the Xunta de Galicia (Research Project CN2011/037) for financial support and the University of Aveiro, the Fundação para a Ciência e a Tecnologia (FCT) and the Fundo Europeu de Desenvolvimento Regional (FEDER) for funding the Organic Chemistry Research Unit. R.S.

thanks the FCT for a ''Investigador Auxiliar'' position. A.M.E. thanks the Spanish Ministry of Science and Innovation for an FPU grant.

References

1 (a) O. A. Arjona, M. Gómez, J. C. Lopez, J. Plumet, *Chem. Rev*., 2007, **107**, 1919–2036.
2 (a) E. B. d. Melo, A. d. S. Gomes, I. Carvalho, *Tetrahedron*, 2006, **62**, 10277–10302; (b) P. Vogel, *Chimia*, 2001, **55**, 359–365.
3 G. E. McCasland, S. Furuta, L. J. Durham, *J. Org. Chem*., 1966, **31**, 1516–1521.
4 (a) G. E. McCasland, S. Furuta, L. J. Durham, *J. Org. Chem*., 1968, **33** (7), 2835–2841; (b) G. E. McCasland, S. Furuta, L. J. Durham, *J. Org. Chem*., 1968, **33**, 2841–2844.
5 T. Suami, S. Ogawa, *Adv. Carbohydr. Chem. Biochem*., 1990, **48**, 21–90.
6 (a) P. Sears, C. H. Wong, *Angew. Chem. Int. Ed*., 1999, **38**, 2300–2324; (b) J. Sardinha, A. P. Rauter, M. Sollogoub, *Tetrahedron Lett*., 2008, **49**, 5548–5550.
7 T. W. Miller, B. H. Arison and G. Albers-Schonberg, *Biotechnol. Bioeng*., 1973, **15**, 1075–1080.
8 (a) T. Posternak, *Les cyclitols*, Hermann, Paris, 1962; (b) D. J. Cosgrove, G. C. J. Irving, *Inositol Phosphates*, Elsevier, Amsterdam, 1980; (c) B. Kilbas, M. Balci, *Tetrahedron*, 2011, **67**, 2355–2389.
9 (a) M. Balci, Y. Sutbeyaz, H. Secen, *Tetrahedron*, 1990, **46**, 3715–3742; (b) M. Honzumi, K. Hiroya, T. Taniguchi, K. Ogasawara, *Chem. Commun*., 1999, 1985–1986.
10 (a) K. Tatsuta, 'Carbohydrate Mimics', Wiley-VCH Verlag GmbH, Weinheim, 1998; (b) K. Tatsuta, *Pure & Appl. Chem*., 1996, **68**, 1341–1346.
11 (a) Y. Kameda, N. Asano, M. Teranishi, M. Yoshikawa, K. J. Matsui, *J. Antibiot*., 1981, **34**, 1237–1240; (b) N. Asano, M. Takeuchi, K. Ninomiya, Y. Kameda, K. Matsui, *J. Antibiot*., 1984, **37**, 859–867; (c) S. Ogawa, Y. Miyamoto, A. Nakajima, *Chem. Lett*., 1989, 725–726; (d) S. Ogawa, A. Nakajima, Y. Miyamoto, *J. Chem. Soc., Perkin Trans*., 1991, **1**, 3287–3290.
12 S. Horii, T. Iwasa and Y. Kameda, *J. Antibiot*., 1971, **24**, 57–58.
13 (a) X. Chen, Y. Fan, Y. Zheng, Y. Shen, *Chem. Rev*., 2003, **103**, 1955–1977; (b) T. Mahmud, *Nat. Prod. Rep*., 2003, **20**, 137–166.
14 (a) H. O. L. Fischer, G. Dangschat, *Helv. Chim. Acta*, 1934, **17**, 1196–1200; (b) H. O. L. Fischer, G. Dangschat, *Helv. Chim. Acta*, 1934, **17**, 1200–1204; (c) H. O. L. Fischer, G. Dangschat, *Helv. Chim. Acta*, 1935, **18**, 1204–1206; (d) H. O. L. Fischer, G. Dangschat, *Helv. Chim. Acta*, 1935, **18**, 1206–1213; (e) H. O. L. Fischer, G. Dangschat, *Helv. Chim. Acta*, 1937, **20**, 705–716.
15 A. Barco, S. Benetti, C. D. Risi, P. Marchetti, G. P. Pollini and V. Zanirato, Tetrahedron, *Asymm*., 1997, **8**, 3515–3545.
16 (a) J. Marco-Contelles, *Eur. J. Org. Chem*., 2001, 1607–1618. (b) Y. Kobayashi, 'Glycoscience, Chemistry and Chemical Biology', Springer-Verlag, Berlin, 2001. (c) S. Atsumi, K. Umezawa, H. Iinuma, H. Naganawa, H. Nakamura, Y. Iitaka, T. Takeuchi, *J. Antibiot*., 1990, **43**, 49–53.
17 S. Atsumi, H. Iinuma, C. Nosaka and K. Umezawa, *J. Antibiot*., 1990, **43**, 1579–1587.
18 Y. Nobuji, C. Noriko, M. Takashi, U. Shigeru, A. Kenzon, I. Michaki, *Jpn. Kokai Tokkyo Koho*, Vol. 06306000, JP, 1994.
19 A. M. Estévez, R. J. Estévez, J. C. Estévez, *Mini-rev. Med. Chem*. 2012, *in press*.

20 D. Bradley, *Nat. Rev. Drug Discov.*, 2005, **4**, 945–946.
21 T. Iwasa, H. Yamamoto and Y. Kameda, *J. Antibiot.*, 1970, **32**, 595–602.
22 A. Berecibar, C. Grandjean and A. Siriwardena, *Chem. Rev.*, 1999, **99**, 779–844.
23 H. Tsunoda, S. Sasaki, T. Furuya and S. Ogawa, *Liebigs Ann. Chem.*, 1996, 159–165.
24 (a) S. I. Hakomori, *Adv. Cancer Res.*, 1989, **52**, 257–331; (b) M. Fukuda, O. Hindsgaul, IRL Press, Oxford, New York, Tokyo, 1994.
25 R. J. Ferrier, *Chem. Rev.*, 1993, **93**, 2779–2831.
26 (a) T. Suami, K. Tadano, Y. Kameda, Y. Iimura, *Chem. Lett*. 1984, 1919–1922. (b) K. Tadano, Y. Kameda, Y. Iimura, T. Suami, *J. Carbohydr. Chem.* 1987, **6**, 231–244.
27 K. Tadano, H. Maeda, M. Hoshino, Y. Iimura and T. Suami, *Chem. Lett.*, 1986, 1081–1084.
28 R. G. Soengas, J. C. Estévez, A. M. Estévez, F. Fernández and R. J. Estévez, Recent advances in nitrosugar chemistry, *Carbohydrate Chemistry*, 2009, 173–199.
29 L. Henry, *Bull. Soc. Chim. France*, 1895, **13**, 999–1001.
30 (a) H. B. Hass, M. L. Bender, *J. Am. Chem. Soc.*, 1949, **71**, 1767–1769. (b) S. J. Etheredge, *Tetrahedron Lett.*, 1965, **50**, 4527–4530.
31 P. W. Wade, R. M. Giuliano, *Nitro Compounds Recent Advances in Synthesis and Chemistry*, 1990, ed. by H. Feuer and A. T. Nielsen, VCH, New York.
32 H. H. Baer, L. Urbas, *The Chemistry of the Nitro and Nitroso Groups*, Willey Interscience, New York, 1970, *part 2*, 75.
33 M. Yoshikawa, B. Cha, T. Nakae and I. Kitagawa, *Chem. Pharm. Bull.*, 1984, **36**, 3714–3717.
34 (a) J. Kovar, H. H. Baer, *Can. J. Chem.*, 1973, **51**, 1801–1811; (b) J. Kovar, H. H. Baer, *Can. J. Chem.*, 1973, **51**, 2836–2842; (c) M. Yoshikawa, B. Cha, T. Nakae, I. Kitagawa, *Chem. Pharm. Bull.*, 1984, **36**, 3714–3717. (d) K. Brewster, J. M. Harrison, T. D. Inch, N. Williams, *J. Chem. Soc. Perkin Trans. I*, 1987, 21–26; (e) S. D. Gero, D. Horton, A. M. Sepulchre, D. Wander, *Tetrahedron*, 1973, **29**, 2963–2972; (f) M. Yosikawa, Y. Yokokawa, Y. Inoue, S. Yamagichi, N. Murakami, *Tetrahedron*, 1994, **50**, 9961–9974; (g) E. J. Corey, I. Vlattas, K. Harding, *J. Am. Chem. Soc.*, 1969, **91**, 535–536; (h) S. Tori, T. Inokuchi, R. Oi, K. Kondo, T. Kobayashi, *J. Org. Chem.*, 1986, **51**, 254–256.
35 R. G. Soengas, J. C. Estévez and R. J. Estévez, *Org. Lett.*, 2003, **5**, 4457–4459.
36 L. El Blidi, M. Ahbala, J. Bolte and M. Lemaire, *Tetrahedron: Asymm.*, 2006, **17**, 2684–2688.
37 C. Chakraborty, V. P. Vyavahare, V. G. Puranic and D. D. Dhavale, *Tetrahedron*, 2008, **64**, 9574–9580.
38 J. M. Otero, F. Fernandez, J. C. Estévez and R. J. Estévez, *Tetrahedron: Asymm.*, 2005, **16**, 4045–4049.
39 R. G. Soengas, J. C. Estévez and R. J. Estévez, *Org. Lett.*, 2003, **5**, 1423–1425.
40 (a) R. G. Soengas, M. B. Pampin, J. C. Estévez, R. J. Estévez, *Tetrahedron Asymm.*, 2005, **16**, 205–211. (b) A. M. Estévez, R. G. Soengas, R. Tato, P. Thomas, J. C. Estévez, R. J. Estévez, *Tetrahedron Asymm.*, 2010, **20**, 116–120.
41 (a) H. Paulsen, W. v. Dein, *Liebigs Ann. Chem.*, 1987, 125–131; (b) H. Fukase, S. Horii, *J. Org. Chem.*, 1992, **57**, 3651–3658; (c) S. Cai, M. Stroud, S. Hakomori, T. Toyokuni, *J. Org. Chem.*, 1992, **57**, 6693–6696.
42 (a) T. K. M. Shing, Y. Chen, W.-L. Ng, *Tetrahedron*, 2011, **67**, 6001–6005. (b) T. K. M. Shing, Y. Chen, *J. Org. Chem.*, 2007, **72**, 6610–6613.
43 (a) T. Suami, K. Tadano, Y. Ueno, C. Fukabori, *Chem. Lett.*, 1985, 1557–1560; (b) K. Tadano, C. Ueno, C. Fukabori, Y. Hotta, T. Suami, *Bull. Chem.*

Soc. Jpn., 1987, **60**, 1727–1739; (c) K. Tadano, C. Fukabori, M. Miyazaki, H. Kimura, T. Suami, *Bull.Chem. Soc. Jpn*., 1987, **60**, 2189–2196.
44 (a) G. Rassu, L. Auzzas, L. Pinna, L. Battistini, F. Zanardi, L. Marzocchi, D. Acquotti, G. Casiraghi, *J. Org. Chem*., 2000, **65**, 6307–6318; (b) F. Zanardi, L. Battistini, L. Marzocchi, D. Acquotti, G. Rassu, L. Pinna, L. Auzzas, V. Zambrano, G. Casiraghi, *Eur. J. Org. Chem*., 2002, 1956–1965.
45 (a) T. K. M. Shing, H. M. Cheng, *Org. Biomol. Chem*., 2009, **7**, 5098–5102; (b) T. K. M. Shing, H. M. Cheng, *Synlett*, 2010, 142–144.
46 (a) T. K. M. Shing, H. M. Cheng, W. F. Wong, C. S. K. Kwong, J. Li, C. B. S. Lau, P. S. Leung and C. H. K. Cheng, *Org. Lett*., 2008, **10**, 3145–3148. (b) T. K. M. Shing and H. M. Cheng, *Org. Lett*., 2008, **10**, 4137–4139.
47 (a) K. Tatsuta, M. Takahashi. N. Tanaka, *J. Antibiot*., 2000, **53**, 88–91. (b) K. Tatsuta, H. Mukai, M. Takahashi, *J. Antibiot*., 2000, **53**, 430–435.
48 D. H. Mac, R. Samineni, J. Petrignet, P. Srihari, S. Chandrasekhar, J. S. Yadav and R. Grée, *Chem. Commun*., 2009, 4717–4719.
49 (a) E. J. Corey, D. Seebach, *Angew. Chem., Int. Ed. Engl*., 1965, **4**, 1075–1077; (b) D. Seebach, N. R. Jones, E. J. Corey, *J. Org. Chem*., 1968, **33**, 300–305.
50 Y.-L. Chen, H. Redlich, K. Bergander and R. Fröhlich, *Org. Biomol. Chem*., 2007, **5**, 3330–3339.
51 X. Cheng, N. Khan, G. Kumaran and D. R. Mootoo, *Org. Lett*., 2001, **3**, 1323–1325.
52 For the first references of radical cyclization of carbohydrate derivatives leading to carbasugars see: (a) W. Schmid, G. M. Whitesides, *J. Am. Chem. Soc*., 1990, **112**, 9670–9671; (b) F. O. Andersson, B. Classon, B. Samuelsson, *J. Org. Chem*., 1990, **55**, 4699–4704.
53 H. Redlich, W. Sudau, A. Szardenings and R. Vollerthun, *Carbohydr. Res*., 1992, **226**, 57–78.
54 (a) B. Bernet, A. Vasella, *Helv. Chim. Acta*, 1979, **62**, 1990–2016. (b) B. Bernet, A. Vasella, *Helv. Chim. Acta*, 1979, **62**, 2400–2410. (c) B. Bernet, A. Vasella, *Helv. Chim. Acta*, 1979, **62**, 2411–2431. (d) M. Nakane, C. R. Hutchinson, H. Gollman, *Tetrahedron Lett*., 1980, **21**, 1213–1216. (e) A. Fürstner, D. Jumbam, J. Teslic, H. Weidmann, *J. Org. Chem*., 1991, **56**, 2213–2217. (f) R. J. Ferrier, P. J. Prasit, *Chem. Soc., Chem. Comm*., 1981, 983–985. (g) A. Fürstner, J. Baumgartner, D. N. J. Jumbam, *Chem. Soc., Perkin Trans*., 1993, **1**, 131–138.
55 J. J. C. Grové, C. W. Holzapfel and D. B. G. Williams, *Tetrahedron Lett*., 1996, **37**, 1305–1308.
56 J. J. C. Grové, C. W. Holzapfel and D. B. G. Williams, *Tetrahedron Lett*., 1996, **37**, 5817–5820.
57 J. L. Chiara, S. Martínez and M. Bernabe, *J. Org. Chem*., 1996, **61**, 6488–6489.
58 A. Chenede, P. Pothier, M. Sollogoub, A. J. Fairbanks and P. Sinay, *J. Chem. Soc., Chem. Comm*., 1995, 1373–1374.
59 J. J. C. Grové and C. W. Holzapfel, *Tetrahedron Lett*., 1997, **38**, 7429–7432.
60 J. Marco-Contelles, P. Gallego, M. Rodríguez-Fernández, N. Khiar, C. Destabel, M. Bernabé, A. Martínez-Grau and J. L. Chiara, *J. Org. Chem*., 1997, **62**, 7397–7412.
61 (a) A. Gómez, E. Moreno, S. Valverde, J. López, *Synlett*, 2002, 891–894; (b) A. Gómez, E. Moreno, S. Valverde, J. López, *Eur. J. Org. Chem*., 2004, 1830–1840; (c) A. Gómez, E. Moreno, S. Valverde, J. López, *Tetrahedron Lett*., 2002, **43**, 5559–5562.
62 G. M. Robertson, *Comprehensive Organic Synthesis*, Oxford: Pergamon Press, 1991; vol. 3, 563.
63 (a) J. L. Namy, J. Souppe, H. B. Kagan, *Tetrahedron Lett*., 1983, **24**, 765–766; (b) J. L. Chiara, W. Cabri, S. Hanessian, *Tetrahedron Lett*., 1991, **32**,

1125–1128; (c) F. Machrouhi, J.-L. Namy, *Tetrahedron Lett*., 1999, **40**, 1315–1318; (d) R. Annunziata, M. Benaglia, M. Cinquini, F. Cozzi, L. Raimondi, *Tetrahedron Lett*. 1998, **39**, 3333–3336; (e) R. D. Rieke, S.-H. Kim, *J. Org. Chem*. 1998, **63**, 5235–5239; (f) S. Talukdar, A. Banerji, *J. Org. Chem*., 1998, **63**, 3468–3470; (g) A. Clerici, L. Clerici, O. Porta, *Tetrahedron Lett*. 1995, **36**, 5955–5958; (h) M. P. Dutta, B. Baruha, A. Baruah, D. Prajapati, J. S. Sandhu, *Synlett*, 1998, 857–858.

64 (a) J. L. Chiara, W. Cabri, S. Hanessian, *Tetrahedron Lett*., 1991, **32**, 1125–1128; (b) J. Uenishi, S. Masuda, S. Wakabayashi, *Tetrahedron Lett*., 1991, **32**, 5097–5100; (c) J. L. Chiara, M. Martín-Lomas, *Tetrahedron Lett*., 1994, **35**, 2969–2972; (d) J. P. Guidot, T. Le Gall, Mioskowski, *Tetrahedron Lett*., 1994, **36**, 6671–6672; (e) J. L. Chiara, N. Valle, *Tetrahedron: Asymm*., 1995, **6**, 1895–1898; (f) J. L. Chiara, N. Valle, *Tetrahedron Lett*., 1994, **35**, 2969–2972. (g) D. J. Jenkins, B. V. L. Potter, *J. Chem. Soc. Perkin Trans*., 1998, **1**, 41–50; (h) A. Kornienko, D. I. Turner, C. H. Jaworek, M. d'Alarcao, *Tetrahedron: Asymm*. 1998, **9**, 2783. (i) T. Sawada, R. Shirai, S. Iwasaki, *Tetrahedron Lett*., 1996, **37**, 885–886; (j) M. Adinolfi, G. Barone, A. Iadonisi, L. Mangoni, R. Manna, *Tetrahedron*, 1997, 11767–11780. (k) M. Carpintero, C. Jaramillo, A. Fernández-Mayoralas, *Eur. J. Org. Chem*., 2000, 1285–1296. (l) I. Storch de Gracia, H. Dietrich, S. Bobo, J. L. Chiara, *J. Org. Chem*. 1998, **63**, 5883–5889.

65 J. L. Chiara, W. Cabri and S. Hanessian, *Tetrahedron Lett*., 1991, **32**, 1125–1128.

66 (a) J. L. Chiara, M. Martín-Lomas, *Tetrahedron Lett*., 1994, **35**, 2969–2972. (b) A. Kornienko, D. I. Turner, C. H. Jaworek, M. d'Alarcao, *Tetrahedron: Asymm*., 1998, **9**, 2783–2786. (c) D. J. Jenkins, B. V. L. Potter, *J. Chem. Soc. Perkin Trans. 1*, 1998, 41–49. (d) J. P. Guidot, T. Le Gall, C. Mioskowski, *Tetrahedron Lett*., 1994, **36**, 6671–6672.

67 (a) M. Adinolfi, G. Barone, A. Iadonisi, L. Mangoni, R. Manna, *Tetrahedron*, 1997, 11767–11780. (b) M. Carpintero, C. Jaramillo, A. Fernández-Mayoralas, *Eur. J. Org. Chem*., 2000, 1285–1296.

68 (a) I. Storch de Gracia, H. Dietrich, S. Bobo, J. L. Chiara, *J. Org. Chem*., 1998, **63**, 5883–5889. (b) J. L. Chiara, I. Storch de Gracia, A. Bastida, *Chem. Commun*., 2003, 1874–1875.

69 G. Luchetti, K. Ding, A. Kornienko and M. d'Alarcao, *Synthesis*, 2008, 3148–3154.

70 G. Luchetti, K. Ding, M. d'Alarcao and A. Kornienko, *Synthesis*, 2008, 3142–3147.

71 J. L. Chiara, I. Storch de Gracia and A. Bastida, *Chem. Commun*., 2003, 1874–1875.

72 (a) H. Miyabe, M. Torieda, T. Kiguchi, T. Naito, *Synlett*, 1997, 580–582; (b) H. Miyabe, M. Torieda, K. Inoue, K. Tajiri, T. Kiguchi, T. Naito, *J. Org. Chem*., 1998, **63**, 4397–4407.

73 (a) J. L. Chiara, J. Marco-Contelles, N. Khiar, P. Gallego, C. Destabel, M. Bernabé, *J. Org. Chem*., 1995, **60**, 6010–6011; (b) S. Bobo, I. Storch de Gracia, J. L. Chiara, *Synlett*, 1999, 1551–1554; (c) I. Storch de Gracia, S. Bobo, M. D. Martín-Ortega, J. L. Chiara, *Org. Lett*., 1999, **11**, 1705–1708.

74 J. L. Chiara, I. Storch de Gracia, A. García, A. Bastida, S. Bobo and M. D. Martín-Ortega, *ChemBioChem*, 2005, **6**, 186–191.

75 A. Boiron, P. Zillig, D. Faber and B. Giese, *J. Org. Chem*., 1998, **63**, 5877–5882.

76 J. L. Chiara and A. García, *Synlett*, 2005, 2607–2610.

77 (a) J. C. López, J. Plumet, *Eur. J. Org. Chem*., 2011, 1803–1825; (b) J. Li, D. Lee, *Eur. J. Org. Chem*., 2011, 4269–4287.

78 (a) S. T. Nguyen, L. K. Johnson, R. H. Grubbs, *J. Am. Chem. Soc.*, 1992, **114**, 3974–3975; (b) G. C. Fu, S. T. Nguyen, R. H. Grubbs, J. W. Ziller, *J. Am. Chem. Soc.*, 1993, **115**, 9856–9857; (c) P. Schwab, R. H. Grubbs, J. W. Ziller, *J. Am. Chem. Soc.*, 1996, **118**, 100–110.
79 (a) R. Madsen, *Eur. J. Org. Chem.*, 2007, 399–415; (b) J. Plumet, A. M. Gomez, J. C. López, *Mini-Reviews in Org. Chem.*, 2007, **4**, 201–216; (c) V. B. Kurteva, C. A. M. Afonso, *Chem. Rev.*, 2009, **109**, 6809–6857.
80 (a) B. Bernet, A. Vasella, *Helv. Chim. Acta* 1979, **62**, 2400–2411; (b) M. Nakane, C. R. Hutchinson, H. Gollman, *Tetrahedron Lett.* 1980, **21**, 1213–1216; (c) A. Fürstner, D. Jumbam, J. Teslic, H. Weidmann, *J. Org. Chem.*, 1991, **56**, 2213–2217; (d) R. J. Ferrier, P. J. Prasit, *Chem. Soc., Chem. Comm.*, 1981, 983–984; (e) A. Fürstner, J. Baumgartner, D. N. Jumbam, *J. Chem. Soc., Perkin Trans.*, 1993, **1**, 131–138.
81 F. G. Hansen, E. Bundgaard and R. Madsen, *J. Org. Chem.*, 2005, **70**, 10139–10142.
82 M. Jorgensen, P. Hadwiger, R. Madsen, A. E. Stutz and T. M. Wrodnigg, *Curr. Org. Chem.*, 2000, **4**, 565–588.
83 L. Hyldtoft and R. Madsen, *J. Am. Chem. Soc.*, 2000, **122**, 8444–8452.
84 (a) P. R. Skaanderup, R. Madsen, *J. Org. Chem.*, 2003, **68**, 2115–2122; (b) R. N. Monrad, C. B. Pipper, R. Madsen, *Eur. J. Org. Chem.*, 2009, 3387–3395.
85 R. N. Monrad, M. Fanefjord, F. G. Hansen, N. M. E. Jensen and R. Madsen, *Eur. J. Org. Chem.*, 2009, 396–402.
86 C. Ramstadius, M. Boklund and I. Cumpstey, *Tetrahedron: Asymmetry*, 2011, **22**, 399–405.
87 J. Frigell, L. Eriksson and I. Cumpstey, *Carbohydr. Res.*, 2011, **346**, 1277–1290.
88 J. P. Rao and B. V. Rao, *Tetrahedron: Asymmetry*, 2010, **21**, 930–935.
89 (a) W. H. Kim, J. A. Kang, H. R. Lee, A. Y. Park, P. Chun, B. Lee, J. Kim, J. A. Kim, L. S. Jeong, H. R. Moon, *Carbohydr. Res.* 2009, **344**, 2317–2321; (b) S. Tripathi, A. C. Shaikh, C. Chen, *Org. Biomol. Chem.*, 2011, **9**, 7306–7308; (c) Y. S. Reddy, P. Kadigachalam, R. K. Basak, A. P. J. Pal, Y. D. Vankar, *Tetrahedron Lett.*, 2012, **53**, 132–136; (d) C. MuniRaju, J. P. Rao, B. V. Rao, *Tetrahedron: Asymmetry*, 2012, doi:10.1016/j.tetasy.2012.01.002.
90 S. T. Diver and A. J. Giessert, *Chem. Rev.*, 2004, **104**, 1317–1382.
91 V. V. Betkekar, S. Panda and K. P. Kaliappan, *Org. Lett.*, 2012, **14**, 198–201.
92 J. M. B. Adam and I. Coldham, *Curr. Org. Synth.*, 2010, **7**, 312–331.
93 (a) H. M. I. Osborn, N. Gemmell, L. M. Harwood, *J. Chem. Soc., Perkin Trans.*, 2002, **1**, 2419–2438; (b) A. E. Koumbis, J. K. Gallos, *Curr. Org. Chem.*, 2003, **7**, 585–628; (c) J. K. Gallos, A. E. Koumbis, V. P. Xinpbaki, C. C. Dellios, E. Coutoul-Argyropoulou, *Tetrahedron*, 1999, **55**, 15167–15180; (d) P. J. Dransfield, S. Moutel, M. Shipman, V. Sik, *J. Chem. Soc., Perkin Trans.*, 1999, **1**, 3349–3355.
94 O. L. Lopez, J. G. Fernández-Bolaños, V. H. Lillelund and M. Bols, *Org. Biomol. Chem.*, 2003, **1**, 478–482.
95 Y. S. Reddy, P. Kadigachalam, V. R. Doddi and Y. D. Vankar, *Tetrahedron Lett.*, 2009, **50**, 5827–5830.
96 C. I. Stathakis, M. N. Athanatou and J. K. Gallos, *Tetrahedron Lett.*, 2009, **50**, 6916–6918.
97 N. Ono, The Nitro Group in Organic Synthesis; Wiley-VCH: New York, 2001; pp. 267–274.
98 A. E. Koumbis and J. K. Gallos, *Curr. Org. Chem.*, 2003, **7**, 398–426.
99 T. Ishikawa, Y. Shimizu, T. Kudoh and S. Saito, *Org. Lett.*, 2003, **5**, 3879–3882.

100 (a) T. K. M. Shing, W. F. Wong, H. M. Cheng, W. S. Kwok, K. H. So, *Org. Lett*., 2007, **9**, 753–756. (b) T. K. M. Shing, K. H. So, W. S. Kwok, *Org. Lett*., 2009, **11**, 5070–5073.
101 P. I. Dalko and P. Sina, *Angew. Chem. Int. Ed*., 1999, **38**, 773–777.
102 R. J. Ferrier, *J. Chem. Soc. Perkin Trans*., 1979, **1**, 1455–1458.
103 R. Blattner, R. J. Ferrier and S. R. Haines, *J. Chem. Soc. Perkin Trans*, 1985, **1**, 2413–2415.
104 D. H. R. Barton, J. Camara, P. Dalko, S. D. Gero, B. Quiclet-Sire and P. Stütz, *J. Org. Chem*., 1989, **54**, 3764–3766.
105 (a) D. H. R. Barton, S. Augy-Dorey, J. Camara, P. Dalko, J. M. Delaumény, S. D. Gero, B. Quiclet-Sire, P. Stütz, *Tetrahedron*, 1990, **46**, 215–230; (b) S. Augy-Dorey, P. Dalko, S. D. Gero, B. Quiclet-Sire, J. Eustache, P. Stütz, *Tetrahedron,* 1993, **49**, 7997–8006.
106 (a) S. L. Bender, R. J. Budhu, *J. Am. Chem. Soc*., 1991, **113**, 9883–9885; (b) V. A. Estevez, G. D. Pestwich, *J. Am. Chem. Soc*., 1991, **113**, 9885–9887.
107 S. Amano, N. Ogawa, M. Ohtsuka, S. Ogawa and N. Chida, *Chem. Commun*., 1998, 1263.
108 (a) S. Adam, *Tetrahedron Lett*., 1988, **29**, 6589–6592.
109 (a) D. H. R. Barton, S. Augy-Dorey, J. Camara, P. Dalko, J. M. Delaumény, S. D. Géro, B. Quiclet-Sire, P. Stiitz, *Tetrahedron* 1990, **46**, 215–220. (b) P. Laszlo, A. Dudon, *J. Carbohydr. Chem*., 1992, **11**, 587–593.
110 T. Iimori, H. Takahashi and S. Ikegami, *Tetrahedron Lett*., 1996, **37**, 649–652.
111 S. K. Das, J.-M. Mallet and P. Sinay, *Angew. Chem. Int. Ed*., 1997, **36**, 493–496.
112 M. Sollogoub, J.-M. Mallet and P. Sinaÿ, *Tetrahedron Lett*., 1998, **39**, 3471–3472.
113 M. Sollogoub, A. J. Pearce, A. Hirault and P. Sinaÿ, Tetrahedron, *Asymm*, 2000, **11**, 283–294.
114 A. J. Pearce, M. Sollogoub, J.-M. Mallet and P. Sinaÿ, *Eur. J. Org. Chem*., 1999, 2103–2117.
115 (a) H. Ito, Y. Motoki, T. Taguchi, Y. Hanzawa, *J. Am. Chem. Soc*., 1993, **115**, 8835–8836; (b) Y. Hanzawa, H. Ito, T. Taguchi, *Synlett*, 1995, 299–305; (c) D. J. Jenkins, A. M. Riley, B. V. L. Potter, *J. Org. Chem*., 1996, **61**, 7719–7726; (d) L. A. Paquette, R. M. Arbit, J. A. Funel, S. Bolshakov, *Synthesis*, 2002, 2105–2109; (e) A. Chénedé, P. Pothier, M. Sollogoub, A. J. Fairbanks, P. Sinaÿ, *J. Chem. Soc., Chem. Commun*., 1995, 1373–1374; (f) J. J. C. Grové, C. W. Holzapfel, D. B. G. Williams, *Tetrahedron Lett*., 1996, **37**, 5817–5820; (f) J. L. Chiara, S. Martinez, M. Bernabé, *J. Org. Chem*., 1996, **61**, 6488–6488; (g) J. M. Aurrecoechea, B. López, *Tetrahedron Lett*., 1998, **39**, 2857–2860.
116 (a) H. Ito, Y. Motoki, T. Taguchi, Y. Hanzawa, *J. Am. Chem. Soc*., 1993, **115**, 8835–8836; (b) Y. Hanzawa, H. Ito, T. Taguchi, *Synlett*, 1995, 299–305; (c) D. J. Jenkins, A. M. Riley, B. V. L. Potter, *J. Org. Chem*., 1996, **61**, 7719–7772.
117 D. J. Jenkins, A. M. Riley and B. V. L Potter, *J. Org. Chem*., 1996, **61**, 7719–7726.
118 J. M. Aurrecoechea, B. Lopez and M. Arrate, *J. Org. Chem*., 2000, **65**, 6493–6501.
119 J. M. Aurrecoechea, J. H. Gil and B. López, *Tetrahedron*, 2003, **59**, 7111–7121.
120 J. M. Aurrecoechea, M. Arrate, J. H. Gil and B. López, *Tetrahedron*, 2003, **59**, 5515–5522.
121 J. Ma, Y. Zhao, S. Ng, J. Zhang, J. Zeng, A. Than, P. Chen and X.-W. Liu, *Chem. Eur. J*, 2010, **16**, 4533–4540.

Multivalent glycoconjugates in medicinal chemistry

José G. Fernández-Bolaños,* Inés Maya and Ana Oliete
DOI: 10.1039/9781849734769-00303

This review summarises recent advances in the design of multivalent glycoconjugates, such as glycodendrimers, glyconanotubes, and glyconanoparticles, and their uses in medicinal chemistry including the development of carbohydrate-based antitumor vaccines. Literature is covered since 2007.

1 Introduction

Carbohydrates play a key role in many biochemical processes: fertilization, immune response, inflammation, cell growth, cellular recognition and signalling, viral replication, and parasitic infection, through the interaction with their corresponding protein recognition receptors.[1] Therefore, interference with these recognition processes is a logical route for the development of new drugs.[2]

As the binding between an individual or monovalent carbohydrate and an individual protein is generally weak, in the millimolar range,[3] nature gets strong and specific responses through multiple protein-carbohydrate interactions, a phenomenon known as *multivalency*.[4] Multivalency leads to greater enhancements in binding affinity than predicted from the simple linear additivity, and this enhancement obtained with multivalent ligands compared to their monovalent counterparts is often referred to as the cluster or multivalent effect.[5,6]

A large number of multivalent ligands have been developed to inhibit or promote many biological processes. Most of these multivalent ligands are glycomimetic inhibitors of lectins, in which the valency, the topology of the epitope, the length of the spacers, and the type of the scaffold are key features for a good cluster effect. The affinity of multivalent glycoconjugates for lectins has been extensively described in excellent reviews.[7–14]

As the typical low affinity of carbohydrates for proteins is the major drawback of their use as therapeutics, the use of multivalent carbohydrates can be a good strategy, since they can bind much more strongly to their targets than their monovalent counterparts.[15] Affinities can be increased several orders of magnitude due to multivalency if chelation is possible, binding a multisite target protein.

The research on multivalent glyconjugates has experienced a dramatic development in the last years, and an impressive number of research papers have appeared. Only a limited number of selected references of the last five years have been included in this review for the sake of conciseness.

2 Glycoclusters with a flexible core

A promising strategy to fight the increase in the occurrence of antibiotic-resistant pathogens[16] is to interfere with the adhesion process of the

Departamento Química Orgánica, Facultad Química, Universidad de Sevilla, 41012-Seville, Spain. E-mail: bolanos@us.es

© The Royal Society of Chemistry 2012

bacteria,[7] required for bacteria to establish a local site of infection overwhelming host defenses. Pathogens are attached to specific host cells by protein adhesins on the surface of the bacteria that adhere to complementary carbohydrate receptors on the host cells. One advantage associated with inhibiting the bacterial infection by blocking the adhesions, is reducing the chance of resistance build-up as the pathogen is not killed, another is to avoid problems associated to the release of toxic by-products from dead bacteria.[17]

Cholera bacteria, causing cholera disease, produce the cholera toxin (CT), a protein capable of binding simultaneously to the carbohydrate moieties of five GM1 ganglioside molecules present on the cell surfaces of the intestines. This binding is the key step for the internalization of the toxin. Pieters *et al*. have been able to prepare using *click* chemistry the multivalent conjugated versions of the GM1 oligosaccharides[18] (Fig. 1). The divalent **c**ompound **2a** was an inhibitor of CT 10,000-fold stronger than

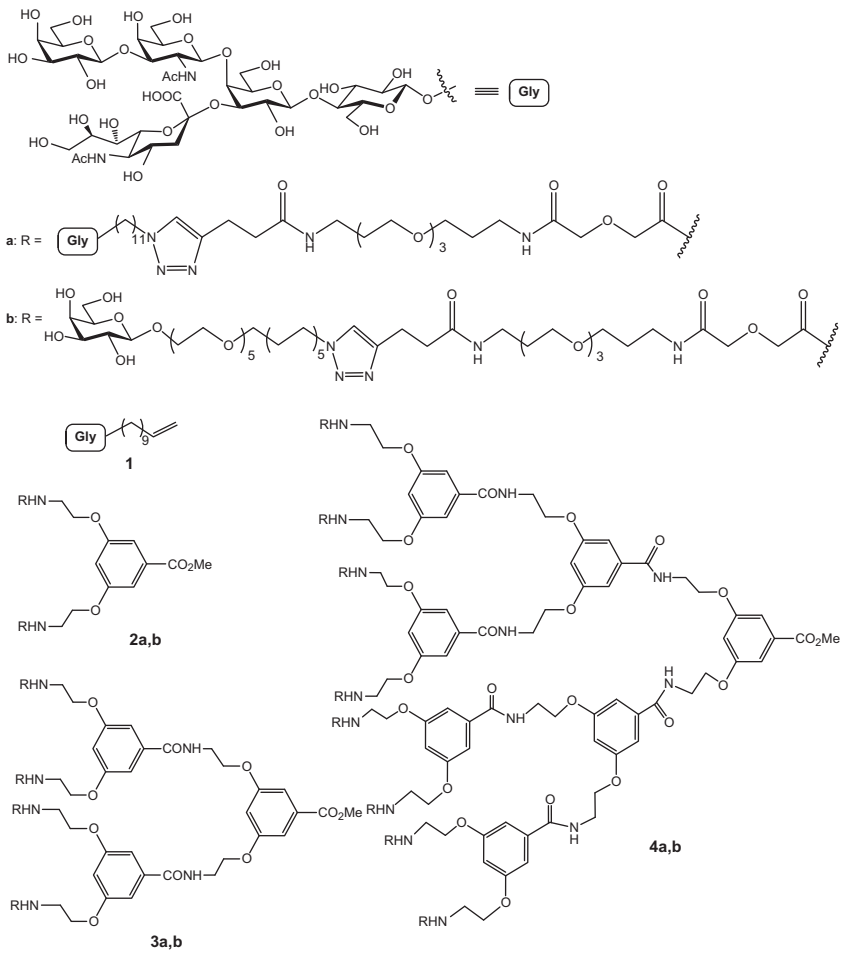


Fig. 1 Multivalent conjugated versions of GM1 oligosaccharides.

the monovalent ($IC_{50} = 19\,\mu M$), whereas the tetravalent and the octavalent dendrimers **3a** and **4a** were 83,000-fold ($IC_{50} = 0.23\,nM$) and 380,000-fold ($IC_{50} = 50\,pM$), respectively, more potent than the monovalent derivative **1**. After correction for valency, the activity of the octavalent compound is 47,500-fold stronger per sugar than the monomer.

In a similar way, that group prepared galactose-containing di-, tetra- and octavalent dendrimers **2b**, **3b**, **4b** with longer spacer arms.[19] The pentasaccharide of ganglioside GM1 was simplified to a β-D-galactopyranosyl unit, the terminal monosaccharide of GM1, while the rest of the pyranose rings were substituted by a PEG fragment. The inhibitory effects of **2b–4b** (**2b**, $IC_{50} = 130\,\mu M$; **3b**, $IC_{50} = 25\,\mu M$; **4b**, $IC_{50} = 12\,\mu M$) were much smaller than those observed for the pentasaccharide derivatives **2a–4a**, however the multivalency effects were strong, as deduced from the relative potency per sugar unit: 923-fold stronger than galactose for **2b**, increased to 2,400-fold for the tetravalent stage and it remained about the same (2,500) for the octavalent derivative.

As bacterial adhesion is the first step in the infection by pathogens, bacterial adhesins are main targets for antibacterial therapeutics. Lectin FimH, a specific adhesin located at the tip of type 1 fimbriae in uropathogenic *Escherichia coli*, binds to glycoproteins with terminally exposed mannose.[20] Gouin *et al.* have described the synthesis by Cu(I)-catalyzed azide-alkyne cycloadditions of a tetravalent heptyl-mannosides on a pentaerythritol skeleton (Fig. 2). This glycocluster inhibits bacterial bladder cell binding in the nanomolar range (12.2 nM, ~6,000- and 64-fold lower than mannose and heptyl mannoside, respectively).[21] The same group has

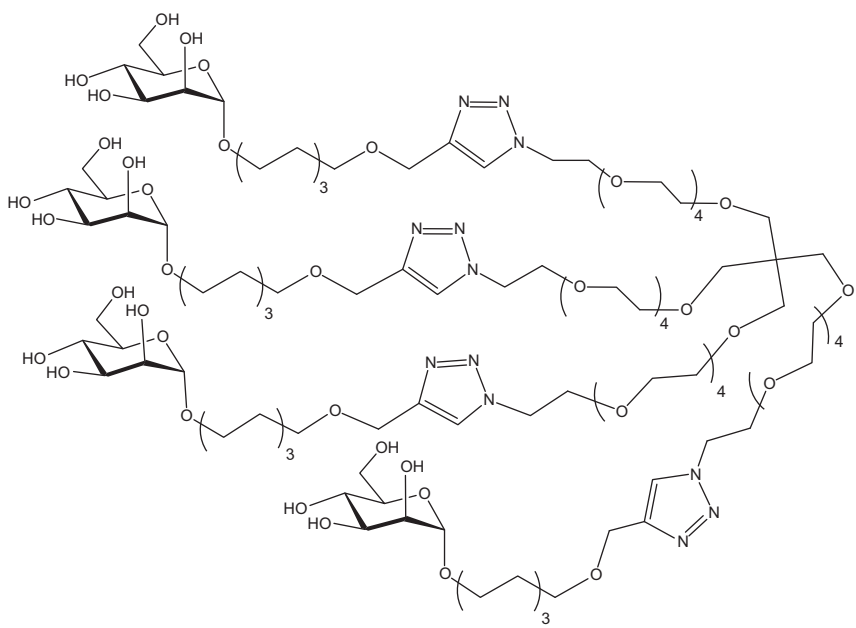


Fig. 2 Tetravalent mannosides with a flexible core.

synthesized, also by click chemistry, multimeric lactosides based on carbohydrate scaffolds with valencies ranging from 1 to 4 and different linker lengths. The binding affinities of the multilactosides towards galectins showed a modest cluster effect.[22]

Heptasaccharidic glycodendrons as mimics of high mannose-type oligosaccharides (Fig. 3) could be obtained using radical addition of mercaptoethanol to double bonds and glycosylation in a repetitive reaction sequence. All of the glycodendrons tested as inhibitors of the mannose-specific adhesion of *E. coli* proved to be better inhibitors than methyl α-D-mannoside (MeMan). The thiahexyl-spacered glycodendrons were generally better inhibitors than the propyl-spacered ones, probably due to increased conformational flexibility of the spacers.[23,24]

In the search of new antiviral drugs able to prevent the infection process, DC-SIGN (dendritic cell-specific ICAM-3 grabbing non-integrin), a mannose-binding lectin present on dendritic cells and involved in the initial stages of HIV infection, is a promising therapeutic target. DC-SIGN is able to recognize highly mannosylated glycoproteins at the surface of pathogens, such as HIV, dengue, Ebola and hepatitis C viruses, and *Mycobacterium tuberculosis* bacterium.[25,26] Therefore, the initial stages of the infection could be prevented by the inhibition of pathogen recognition by DC-SIGN. A tetravalent dendron containing four copies of a linear trimannoside mimic in which the central mannose unit is replaced by a carbocyclic diol (Fig. 4), was shown to inhibit DC-SIGN-mediated HIV *trans* infection of $CD4^+$ T lymphocytes at low micromolar concentration.[27] This dendron inhibits in a dose-dependent manner HIV-1 infection of human cervical explants, even in presence of elevated viral load, and displays a long lasting effect.[28]

Multifunctional dendrimers bearing two surface functionalities, ''Janus dendrimers'', might be useful as drug delivery devices that can combine tissue targeting and imaging, or be directed to a specific cell type. Boons and co-workers[29] have developed a methodology for the assembly of dendrimers having two different functions, based on three sequential azide–alkyne cycloadditions (Fig. 5).

This methodology is compatible with biologically important compounds such as carbohydrates, peptides, or fluorescent labels. In this approach, a strain-promoted azide–alkyne cycloaddition (SPAAC) between polyester dendrons, bearing an azido and a 4-dibenzocyclooctynol moiety at the focal point, furnished dendrimers with terminal and trimethylsilyl-protected alkynes. The terminal alkynes were assembled with azido-modified galactosyl residues by using Cu(I)-catalyzed azide–alkyne cycloadditions. TMS deprotection and a subsequent *click* reaction of the resulting terminal alkyne with azido-containing compounds gave multifunctional dendrimers, such as that shown in Fig. 6, bearing sugars and Arg-Gly-Asp (RGD) peptides at the periphery.[29]

A previously reported multivalent dendrimer with dual function, recognition and detection, containing 16 mannose units and 2 coumarin chromophores, exhibited 240-fold greater potency than monomeric mannose in standard hemagglutination assays using Concanavalin A (Con A), with a relative activity of 15 per sugar moiety when compared to mannose.[30]

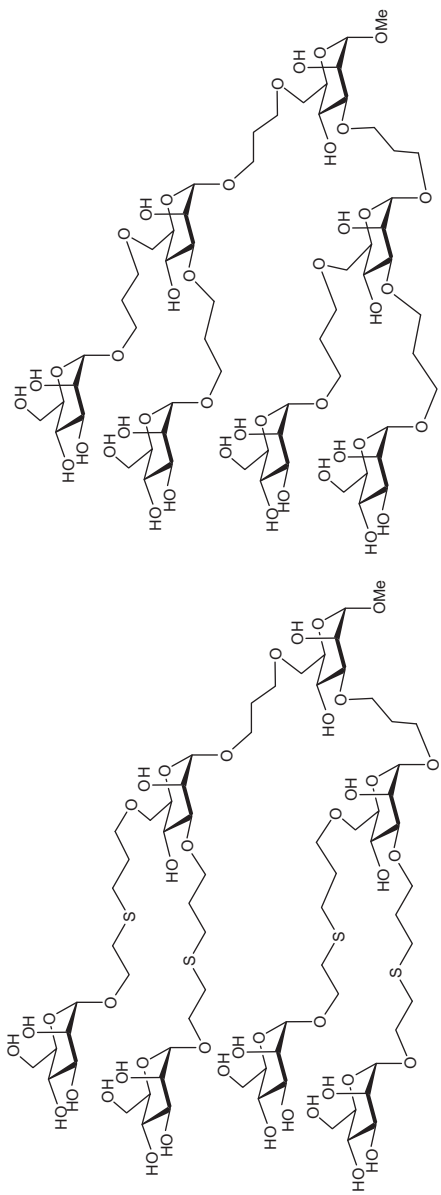


Fig. 3 Glycodendrons as oligomannoside mimetics.

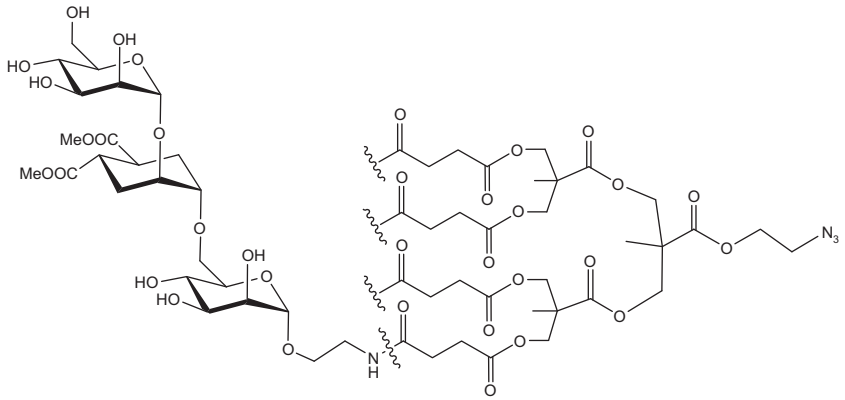


Fig. 4 Linear trimannoside mimic tetravalent glycodendron.

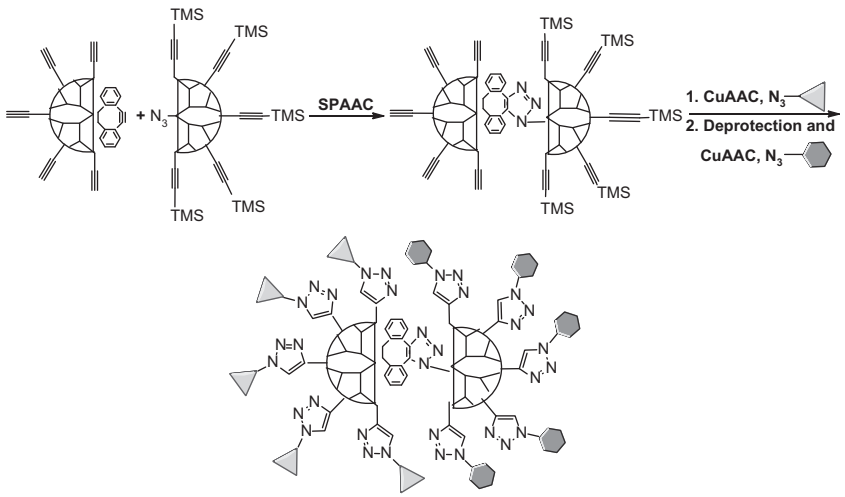


Fig. 5 Multifunctional dendrimer synthesis by three consecutive *click* reactions.

3 Glycoclusters with a rigid core

3.1 Glycocalix[n]arenes

Calix[*n*]arenes are useful cyclic scaffolds[31] with well-defined conformational properties and cavities, which can modulate shape and conformational flexibility of sugar ligands. Calixarene glycoconjugates are effective in binding important lectins such as cholera toxin,[32] and the influenza A viruses.[33] A series of calixarenes substituted with β-GlcNAc units linked by a thiourea spacer has been tested for binding activity to the activation-associated receptor CD69 (cluster of differentiation 69) on human NK (natural killer) cells. A GlcNAc-tetrasubstituted calix[4]arene proved to be the best CD69 ligand identified to date, exhibiting stimulation of natural cytotoxicity of human PBMC (peripheral blood mononuclear cell) in 10^{-4} and 10^{-8} M concentrations, and inducing an increase in the spontaneous death of tumor targets.[34]

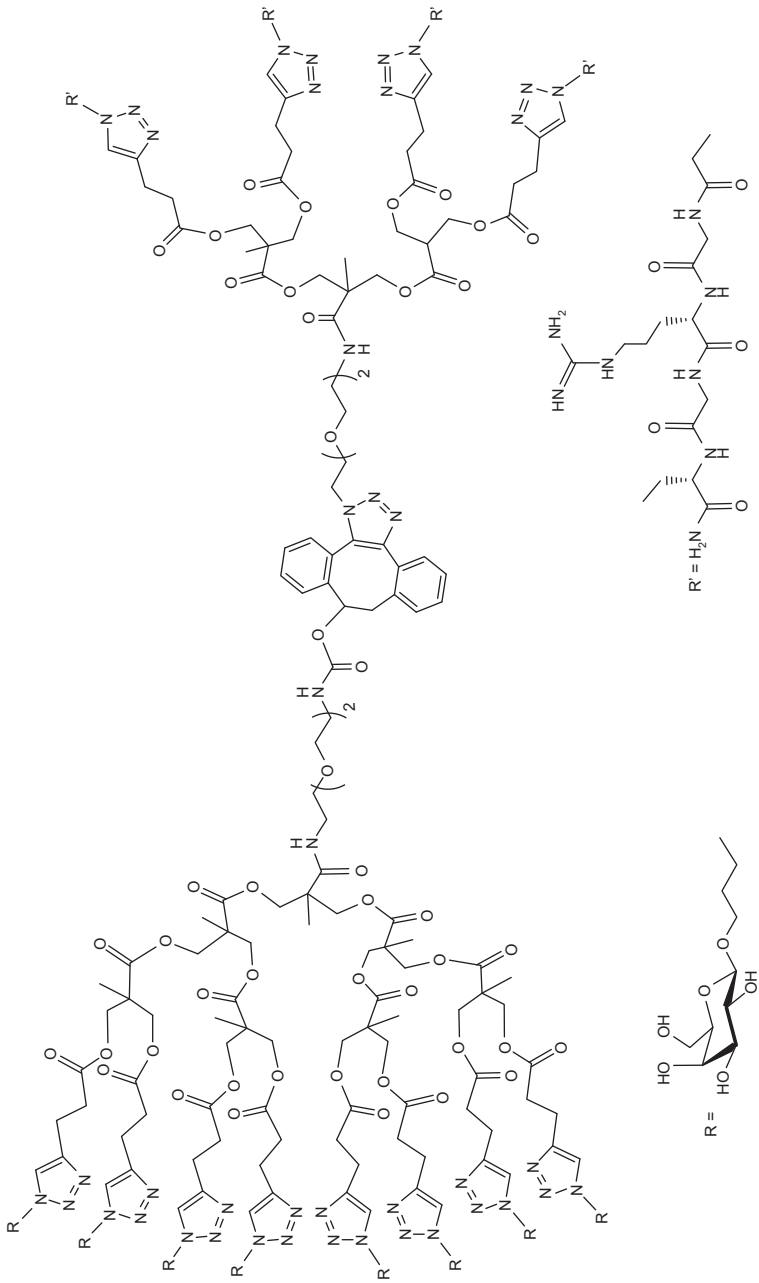


Fig. 6 "Janus dendrimers" bearing two different functions: sugars and RGD peptides.

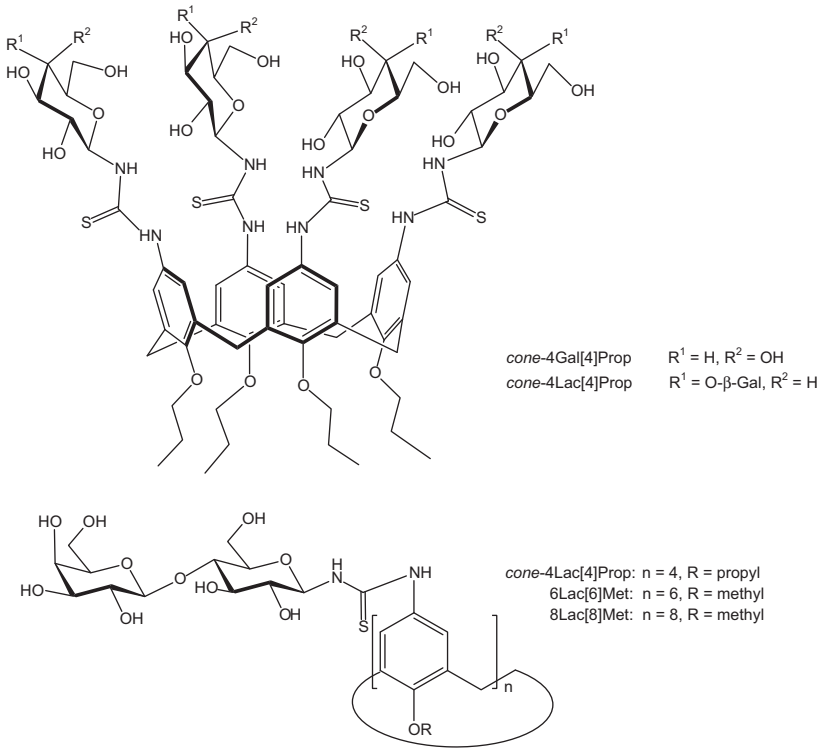


Fig. 7 Glycosylthioureido calix[n]arenes.

Ungaro *et al*. have prepared a series of calix[*n*]arenes (n = 4, 6, and 8) with thiourea-linked galactose or lactose moieties at the upper rim (Fig. 7). The calix[6,8]arenes proved to be especially effective as inhibitors of binding of three human galectins to human tumor cells *in vitro*, and also effective towards a plant toxin (*Viscum album* agglutinin). It was shown that inhibition is related to the conformational properties of the calix[n]-arene scaffold, and to the shape and valency of the glycoclusters. For example, the fixed-cone type, *cone*-4Lac[4]Prop (Fig. 7) could block galectin-3 binding to colon cancer cells, and the flexible calixarenes 6Lac[6]Met and 8Lac[8]Met could avoid galectin-4 binding to pancreatic carcinoma cells.[35] In a similar way, glucosylthioureido calix[6 and 8]arenes showed interaction with plasmid DNA and agglutination with a glucose-specific lectin such as Con A.[36]

A series of calix[4]arene glycoconjugates have been prepared by Vidal and co-workers[37] by coupling alkyne-derivatized calix[4]arene precursors to a galactose azido derivative by microwave-assisted Cu(I)-catalyzed azide–alkyne cycloadditions. The designed glycoconjugates were evaluated as ligands for the galactose-binding PA-IL (*Pseudomonas aeruginosa* agglutinin-I lectin) from the opportunistic bacterium *Pseudomonas aeruginosa*, a major cause of nosocomial lung infections for immuno-compromised patients.[38] The interaction with PA-IL lectin was shown to be strongly dependant on both the valence and the topology, with tetravalent

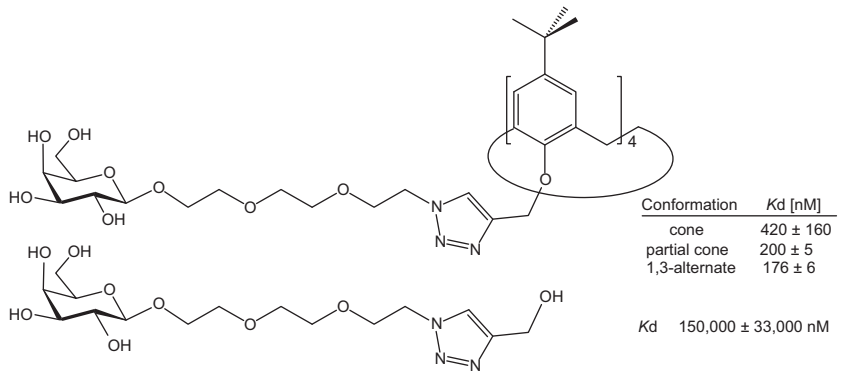


Fig. 8 Three topologically isomeric tetravalent calix[4]arenes.

galactose ligands binding most strongly. Nanomolar affinities were measured by isothermal titration microcalorimetry (ITC).

Three topologically isomeric tetravalent conjugates were compared, the one locked in 1,3-*alternate* conformations, topology 2:2 (Fig. 8) was the most active, with an 800-fold enhancement of affinity compared to the monovalent derivative. The one in the *cone* conformation, topology 4:0, showed only a 360-fold enhancement. Atomic force microscopy has revealed that PA-IL lectin and the 1,3-*alternate* glycocluster create monodimensional filaments according to an aggregative chelate binding mode.[39]

A recent report by Cecioni *et al*.[40] describes the synthesis of a library of optimized tetravalent glycoclusters (Fig. 9) using triethylene glycol spacers (**a**) and three structurally modified analogues that incorporate one or two amide moieties and/or a phenyl ring (**b–d**), in order to study the influence of the flexibility of the linker arm on the interaction of the prepared glycoclusters with PA-IL lectin. The azido-functionalized precursors containing per-*O*-acetylated galactose were coupled with propargyl acetate to get the monovalent references and with tetra-propargylated core scaffolds (Zn-porphyrin **A**, methyl α-D-glucopyranoside **B**, *cone*-calix[4]arenes **C**, partial *cone* conformer **D**, and 1,3-*alternate* conformer **E**) to afford acetylated glycoclusters. The coupling was carried out by CuAAC, under microwave irradiation.

Monovalent *p*-amidophenyl-based ligand **6** (Fig. 9) showed a dissociation constant value (K_D) with PA-IL, determined by ITC, of 5.8 μM, which is an unprecedented affinity for a monovalent glycomimetic ligand toward this lectin, and much higher affinity than that shown by the analogues with more flexible spacers analysed in the article (107–181 μM). The binding of the prepared glycoclusters was also analysed by ITC, and **5bE** with a 1,3-*alternate* calix[4]arene as the core and one amido moiety in the spacer showed the best affinity (K_D value of 90 nM),[40] whereas for the more flexible **5aE** the K_D value was 176 nM.[37]

Other series of tetravalent calix[4]arene glycoclusters incorporating β-lactosyl residues to the upper or to the lower rims of the macrocycle has been prepared by Field *et al*. by using *click* chemistry methodology.[41] These

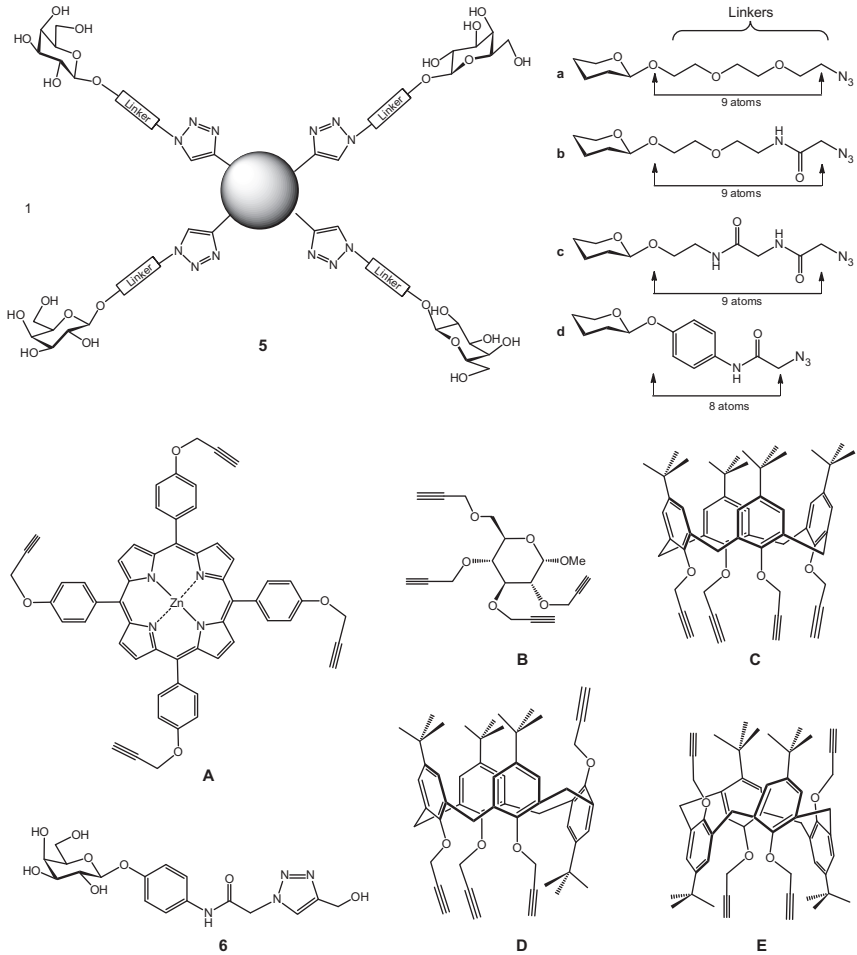


Fig. 9 Structure of the four linker arms (**a**–**d**) and five cores (**A**–**E**), coupled to give tetravalent glycoclusters **5**.

compounds showed trypanocidal activity against *T. cruzi*, with one of them showing similar activity to benznidazole, an established anti-trypanosomal drug.

3.2 Glycofullerenes

Viruses such as retrovirus HIV are spherical particles decorated with envelope glycoproteins.[42] These glycoproteins are responsible to interact with cell surface receptors in a multivalent mode and activate a cascade of processes that lead to cell attachment and entry of the virus into the cells. Therefore, they are important targets for both therapeutic and prophylactic approaches.[43] [60]Fullerenes covered by carbohydrates (''sugar balls'') could mimic the virus surface and could interact in a multivalent manner with cell surface receptors. In this context, fullerene hexakis-adducts bearing 12 or 24 peripheral carbohydrate moieties were prepared using *click*-chemistry methodology, thus obtaining a spherical platform for a globular

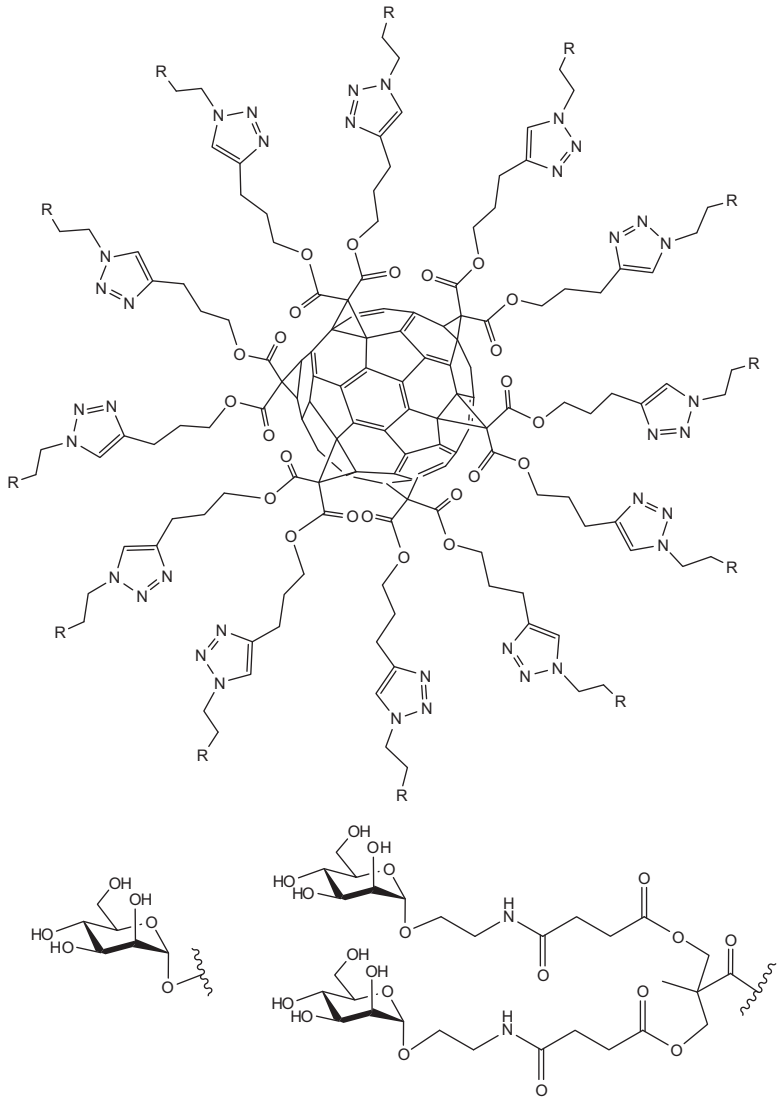


Fig. 10 Fullerene hexakis-adducts bearing 12 or 24 peripheral carbohydrate moieties.

multivalent presentation of ligands.[44,45] The mannosyl derivatives (Fig. 10) were recognised by Con A in a multivalent manner, showing that fullerene is an efficient platform for the globular presentation of carbohydrates for protein–carbohydrates interactions. The ΔG values of the binding process for the glycofullerene with 12 mannose units is apparently more favourable than with 24, due to the higher entropic penalty for the binding process with the latter.

4 Multivalent glycopeptides in immunotherapy

Cancer cells show aberrant glycosylation of cell surface glycoproteins, with many glycosyl epitopes constituting tumor-associated antigens. Although it

remains unclear whether aberrant glycosylation is a result or a cause of cancer, this process is a key event in induction of invasion and metastasis.[46] In epithelial cancers, mucin glycoproteins over-express *O*-linked glycans called tumor-associated carbohydrate antigens (TACAs) such as Tn (GalNAcα1-*O*-Ser/Thr), TF (Thomsen-Friedenreich, Galβ1-3GalNAcα1-*O*-Ser/Thr) and the corresponding sialylated derivatives STn and STF, which are considered a promising target for immunotherapy.[47,48] Glycolipids such as gangliosides are also overproduced on the surface of various cancer types.[49]

The immune responses towards glycoconjugates rely on T-cell triggering mechanisms depending on the nature of the carbohydrate antigens. However these carbohydrate antigens exhibit poor immunogenicity as carbohydrates alone are usually T-cell independent antigens.[50] Carbohydrate antigens activate independent T-cell responses which create short-lived immunoglobulin G (IgG) antibodies. Therefore, they cannot be employed as vaccines alone. In the last few years there has been an impressive development of the design of synthetic multi-component MUC1-based glycopeptide vaccines.[51–59] These vaccines rely on mimicking the natural presentation of TACAs, typically found in clusters, for generating strong antibody responses and immunogenicity.[60,61] In this context, glycodendrimers have shown to be useful in vaccine preparation, due to their immunostimulating properties.[62]

Kunz *et al.*,[63] according to the principle of multiple antigen presentation, prepared a tetrameric vaccine candidate (Fig. 11) consisting of a di-lysyl-lysine core and four branches formed by a T cell epitope from tetanus toxoid and a sialyl Tn-glycododecapeptide of MUC1. In a similar way, the same group prepared by solid-phase procedures tetrameric and octameric MUC1 and MUC4 Tn-glycopeptide vaccine candidates.[64] They have also described the preparation of efficient fully synthetic antitumor vaccines by the combination of tumor-associated mucin glycopeptide antigens with lipopeptide Toll-like receptor 2 ligands, able to induce immune reactions in mice.[65]

Other antitumor vaccine candidates have been developed, consisting of mono-, di-, and tetravalent MUC1 glycopeptides conjugated by *click*

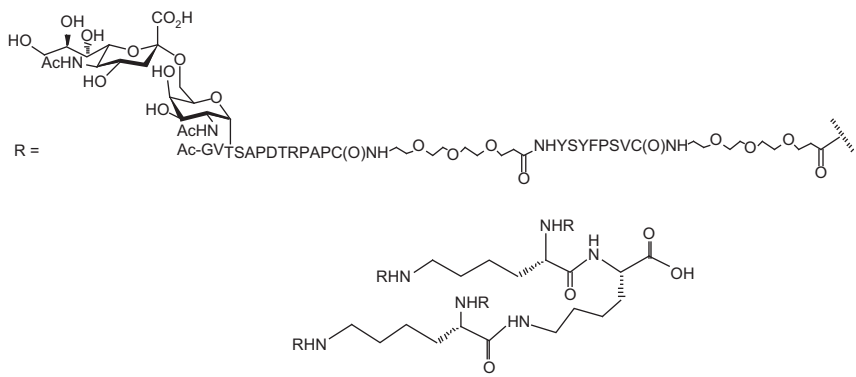


Fig. 11 Tetravalent sialyl-Tn-MUC1-Lys2-Lys1 glycodendron vaccine.

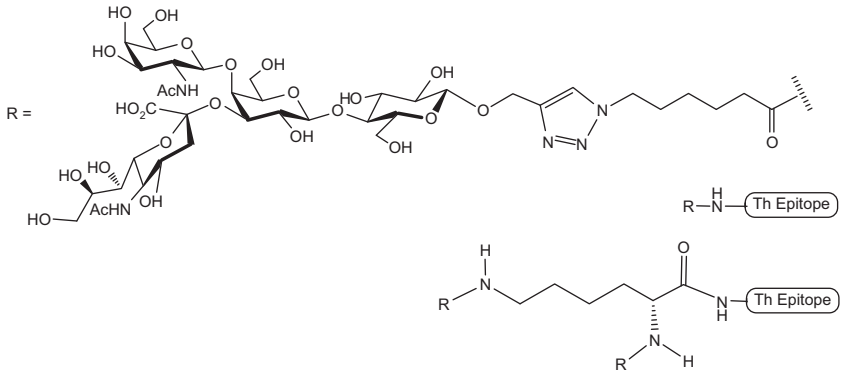


Fig. 12 Synthetic GM2 neoglycopeptide.

chemistry with immunostimulating toll-like receptor2 (TLR2) lipopeptide ligands. Oligovalent glycopeptide–lipopeptide conjugates are considered more immunogenic than their monovalent analogues.[66]

A 40 amino acid MUC1 glycopeptide bearing 10 copies of the Tn antigen has been prepared by Payne and co-workers using the direct aminolysis ligation of a peptide thioester with a deprotected MUC1 glycopeptide.[67] The same group described the synthesis of MUC1-lipopeptide vaccine candidates, consisting of MUC1 glycopeptides, bearing multiple copies of Tn and TF tumor-associated carbohydrate antigens tethered to the lipopeptide immunoadjuvant $Palmitoyl_3CysSer$.[68]

A neoglycopeptide immunogen displaying one or two copies of the tetrasaccharidic moiety of GM2 ganglioside (Fig. 12), an important target for specific anticancer immunotherapy, has been prepared by ligation of alkyne-functionalized GM2 with an azido $CD4^+$ T cell epitope peptide. These synthetic glycopeptides can induce human tumor cell-specific antibodies in mice after immunization. It is noteworthy that the monovalent presentation of GM2 peptide induced antimelanoma antibodies, but failed the multivalent form. This is considered the first report on the induction of human tumor cell-specific antibodies after immunization with a synthetic glycopeptide.[69]

Danishefsky *et al.* have developed glycoconjugated vaccines that are useful for the treatment of cancer and effective at inducing antibodies against all carbohydrate antigens present on the glycoconjugates. For example, vaccine shown in Fig. 13 was prepared by conjugation of a pentavalent glycopeptide, containing five prostate and breast cancer associated carbohydrate antigens (Globo-H, GM2, sialyl Tn, Tn and TF), to maleimide-activated carrier protein KLH (Keyhole Limpet Hemocyanin). These vaccines were able to induce in mice IgM and IgG (immunoglobulins M and G) against each one of those five carbohydrate antigens.[70,71]

In the design of carbohydrate-based vaccines the choice of the right template for multivalent carbohydrate presentation is a key step, because very flexible scaffolds will decrease effective clustering of the antigen. Danishefsky and co-workers have synthesized cyclic peptide scaffolds[72] (Fig. 14) upon which the clustered tumor-associated antigens Tn and STn

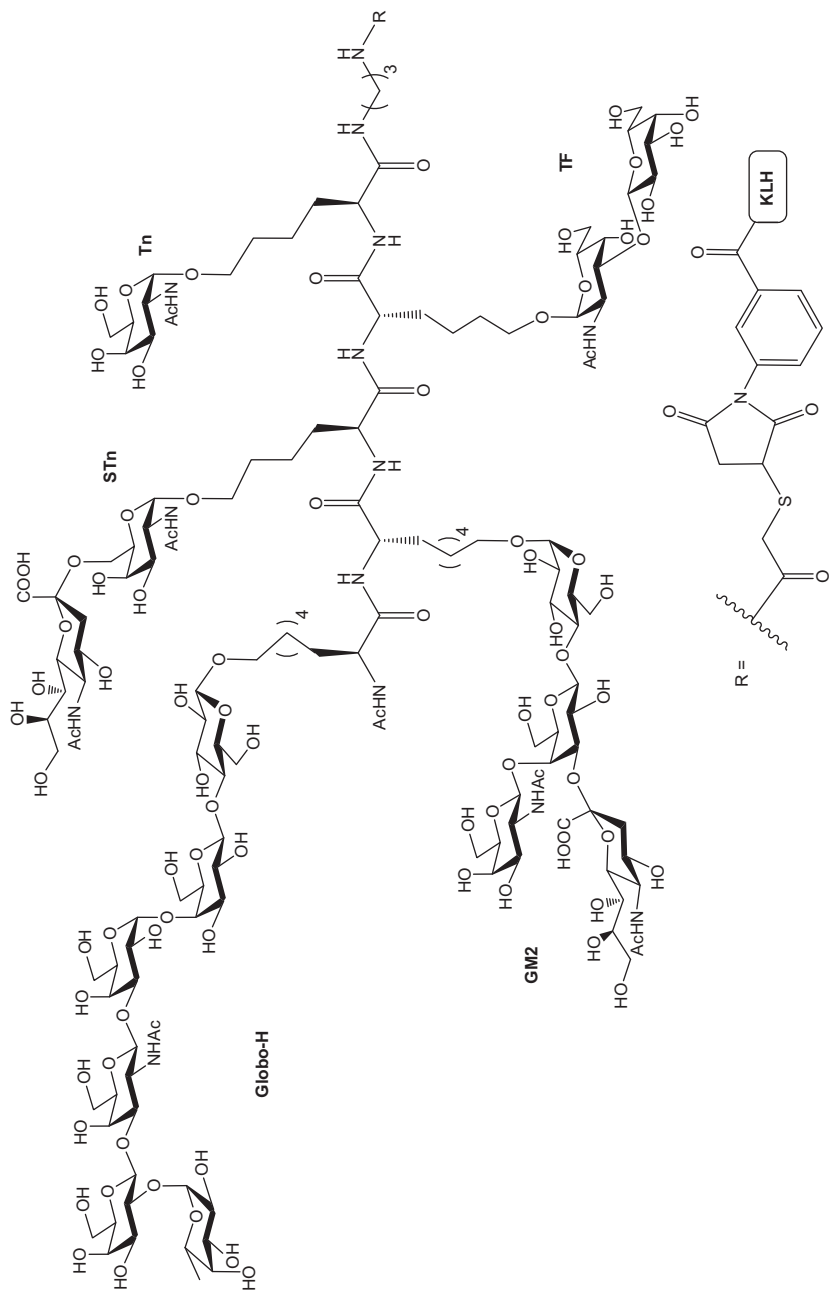


Fig. 13 Pentavalent glycopeptide, containing five carbohydrate antigens (Globo-H, GM2, sialyl Tn, Tn and TF) conjugated to carrier protein KLH.

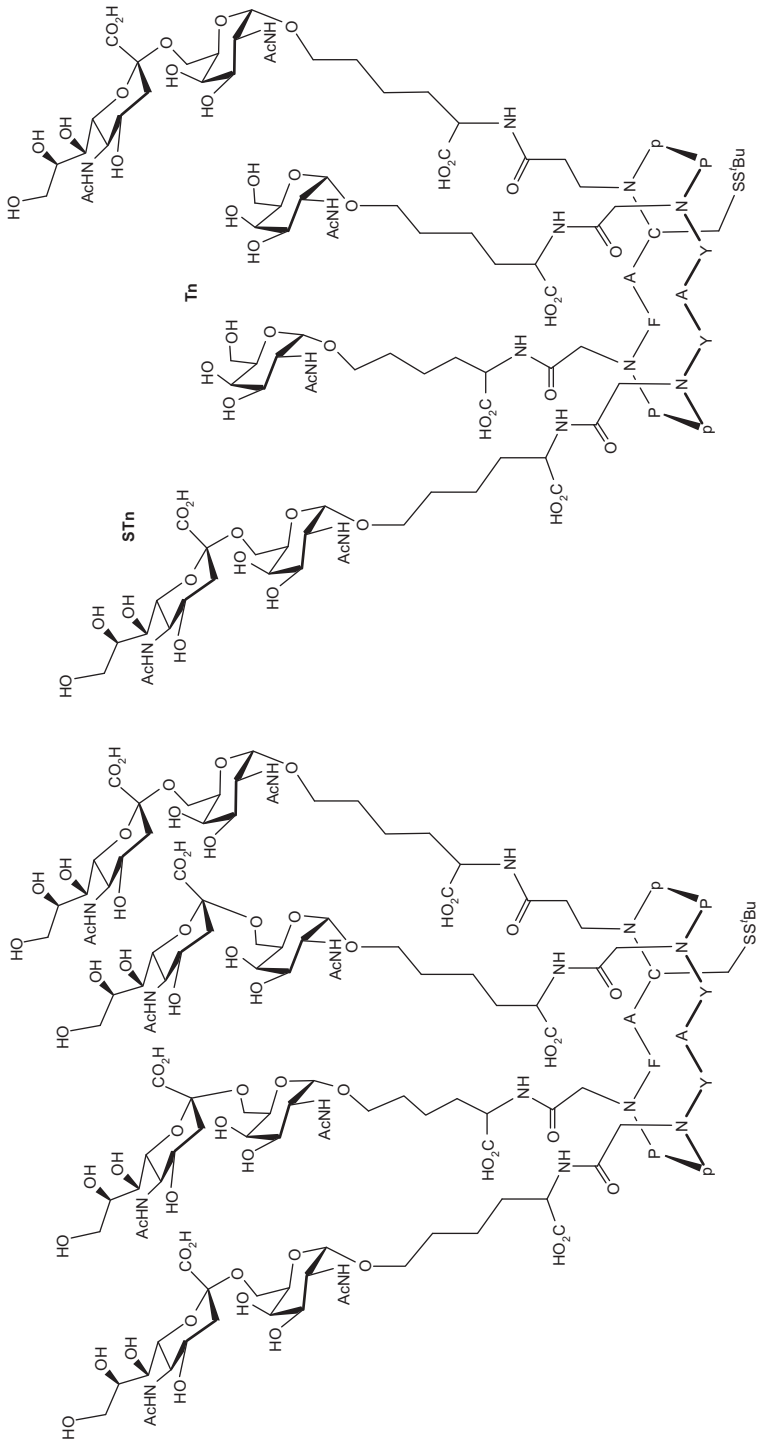


Fig. 14 Cyclic peptide scaffolds as vaccine precursors.

would be displayed in a well-defined orientation. A cross-linker has been introduced by a ring-closing metathesis-mediated route to control the distance between glycans attached to the peptide platform. This type of cyclic peptide scaffolds has been employed by the same group in the development of an HIV vaccine.[73,74] In this context, Wang *et al.* have developed a strategy for the synthesis of highly functionalized oligomannose dendrons in the search of a HIV vaccine,[75] and Kabanova *et al.* prepared polyamidoamine (PAMAM) dendrons displaying oligomannose clusters of HIV-related antigens, able to induce IgG antibodies.[76] Carbohydrate based vaccines against bacteria and parasites have also been developed.[77–79]

5 Glycoconjugates on tubular scaffolds

5.1 Glycocarbon nanotubes

Single-walled carbon nanotubes (SWNTs) can be functionalized with carbohydrates, enabling effective interactions with cells and other biological species that are not available to the free bioactive molecules. For example, SWNTs functionalized with mannose or galactose (Man-SWNT or Gal-SWNT) (Fig. 15) aggregate with *Bacillus anthracis* and *Bacillus anthracis* spores in the presence of divalent cations to result in significant reduction in colony forming units (CFU).[80,81] The effectiveness in the CFU reduction of *B. subtilis* spores by Man-SWNT was in the order of $Mg^{2+} < Ca^{2+} < Ba^{2+}$. It was surprising that the divalent or tetravalent sugar dendron-functionalized SWNTs (Man2-SWNT, Gal2-SWNT, Man4-SWNT, and Gal4-SWNT) were found to be generally ineffective or less effective in binding and aggregating the *Bacillus* spores in the presence of Ca^{2+}.[81] Mannose and galactose dendron-functionalized SWNTs were also evaluated in binding assays with pathogenic *E. coli*.[82]

5.2 Tubular supramolecular architecture

Columnar supramolecular polymers (Fig. 16) have shown to be effective polyvalent scaffolds for detecting and binding two strains of *Escherichia coli* (ORN178 and ORN208).

The self-assembly into columnar polymers of disc-shaped molecules such as shown in Fig. 17, decorated or not with sugars, enhances the potency of the monomers significantly more than expected on the basis of the multivalent effect alone. It was also observed that an effective decreasing of the amount of mannose component of these polymers actually enhanced bacterial aggregation, due to a reduction of steric crowding.[83]

6 Glycodendrimer chips

Carbohydrates can be displayed on microarray chips to be screened by large series of carbohydrate binding proteins and pathogens.[84–86] These studies are useful for the design of new drugs using only a tiny amount of material. Pieters *et al.*[87] have displayed glycodendrimers on a microarray surface in such a way that the number of carbohydrates on each spot is the same while changing the valency (Fig. 18). Spots containing from monovalent α-D-mannopyranoside compounds to octavalent dendrimers were evaluated with fluorescent Con A and GNA (*Galanthus nivalis* agglutinin) lectins, and

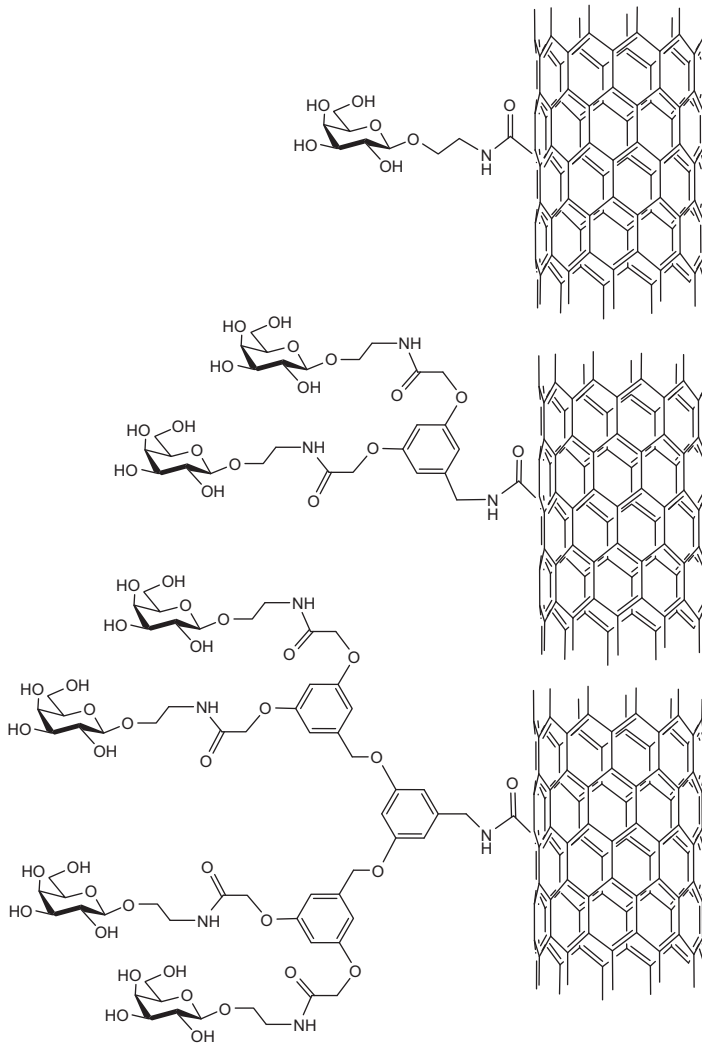


Fig. 15 Galactose-functionalized SWNTs.

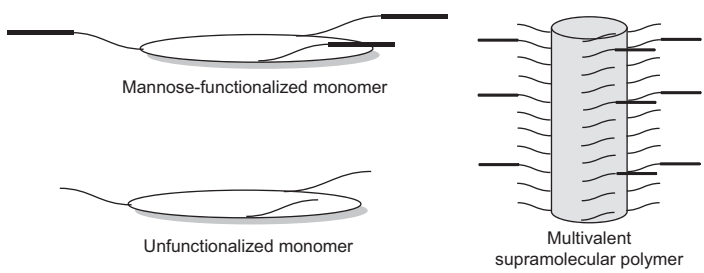


Fig. 16 Multivalent mannose-functionalized and unfunctionalized columnar supramolecular polymers.

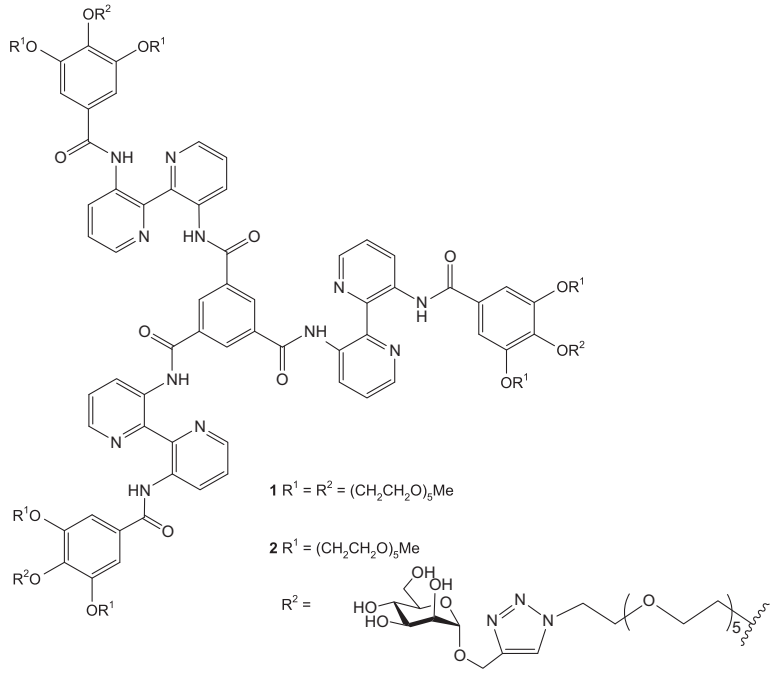


Fig. 17 Mannose functionalized disc-shaped molecules that self-assemble into columnar supramolecular structure.

the binding was observable in real time. In a single experiment it was possible to observe the multivalency enhancement, small for ConA, in agreement with its widely spaced binding sites, and large for GNA, according to its more closely spaced binding sites.

T. cruzi displays host-dependant cell surface diversity dominated by mucin *O*-linked glycoproteins, and sialylation of non-reducing terminal galactose residues of these mucins has been related to evasion of the immune response.[88] Field *et al.* have investigated the role of sialic acid recognition in host infection by *Trypanosoma cruzi* using surface plasmon resonance imaging (SPRi) as a real-time analysis of protein interactions with carbohydrate microarrays.[89] SPRi of an array of 40 sialylated and unsialylated glycans determined the binding preference of hSiglec 7 (human sialic acid binding Ig-like lectins) for α-2,8-linked disialic acid, whereas murine siglec E showed preference for terminal sialic acid.[90] The same technique was used for screening ligands of the ricin surrogate RCA_{120}. A non-natural 6-modified galactoside exhibited up to a 3–4 fold enhanced RCA_{120} binding compared to the parent galactoside.[91]

7 Glyconanoparticles

Carbohydrate-based nanoparticles or glyconanoparticles (GNPs) are structures that mimic carbohydrate presentation of glycocalyx. They consist of a surface covered with multiple carbohydrate residues, which enhances

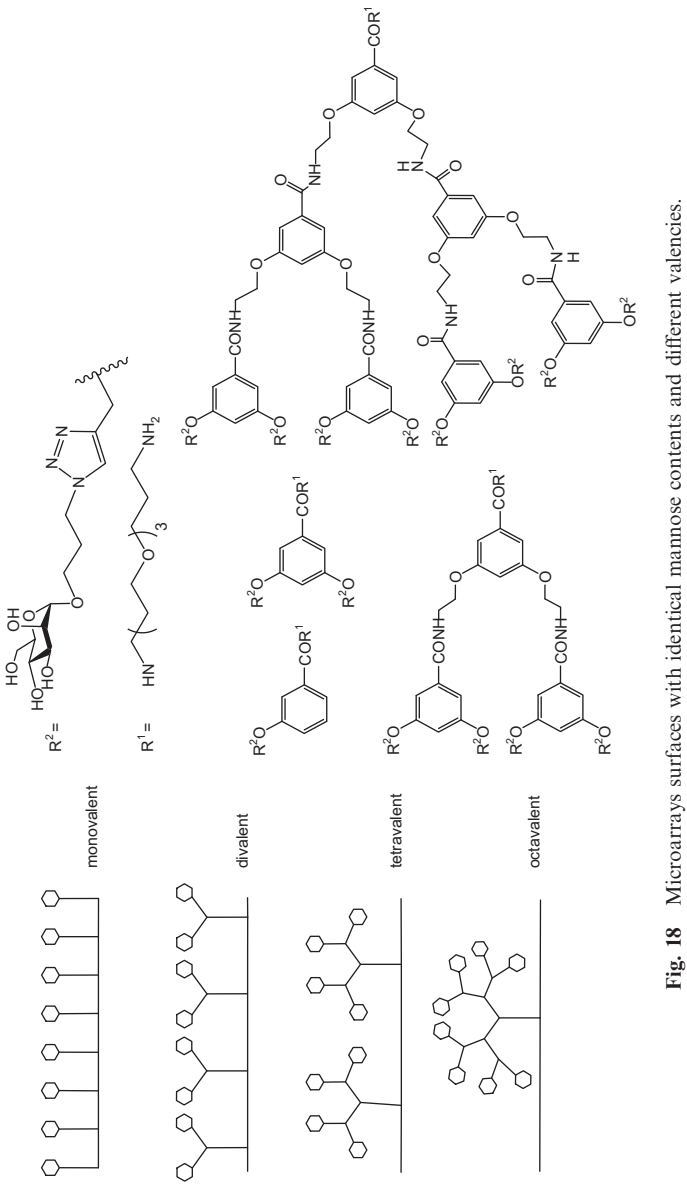


Fig. 18 Microarrays surfaces with identical mannose contents and different valencies.

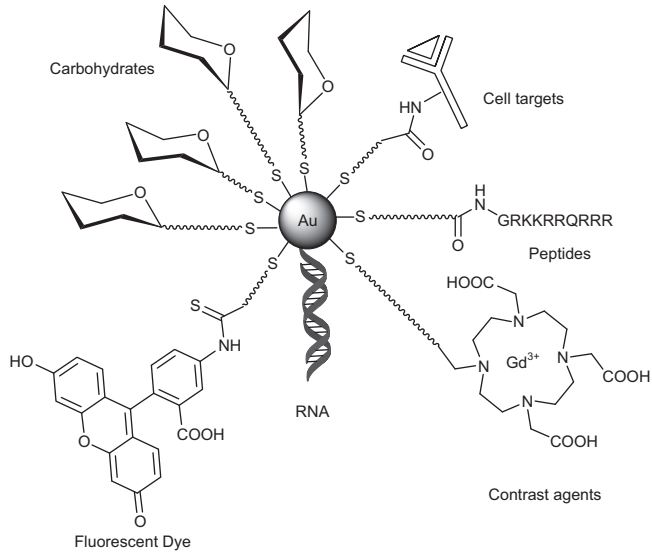


Fig. 19 Multifunctional glyconanoparticles and possible molecules to be attached on the surface: fluorescent probes, peptides, DNA/RNA, metal chelates and cell targets, with linkers of different nature and length.

affinity binding with the corresponding receptor, and an inorganic and nano-sized core which provides useful optical, electronic and magnetic properties depending on the chosen material.[92]

Some features of this model are its globular shape, controllable nanometer size, multivalent display of carbohydrates and easy achievement of multifunctionality on the inorganic surface. The modification of the density of sugar residues and the nature of the linker modulates accessibility to the ligands and their behavior during the molecular recognition events (Fig. 19).[93]

Nanoparticles (NPs) have many applications in clinical medicine as drug carriers[94,95] or biosensors.[96] Among NPs, gold and magnetic NPs are promising tools for medical imaging, diagnosis and therapy.[97–99] Carbohydrate coating confers favorable properties to metallic nanostructures for biomedical applications such as stability and solubility in biological media, biocompatibility, non-toxicity and improved specificity. Carbohydrates facilitate distribution and penetration of glyconanoparticles into cells and thus, GNPs are attracting attention as anti-HIV agents,[100,101] anti-cancer vaccines[102] and magnetic probes in cellular labeling and Magnetic Resonance Imaging (MRI).[92,103]

Recent biomedical applications of glyco-NPs in protein recognition, inhibition of the adhesion of virus, bacteria and toxins, cellular targeting and as probes in fluorescence and MRI will be described in this section. We specially highlight examples of metallic nanoparticles in which the sugar ligands, covalently linked to the metallic nucleus, play a key role in biological recognition processes and not the ones where carbohydrates are used as stabilizing agents.[104] Examples are divided into three subsections according to the NP core composition: gold NPs, magnetic NPs and quantum dots.

7.1 Gold glyconanoparticles

7.1.1 Carbohydrate-carbohydrate interaction

Carbohydrate-carbohydrate interactions may be considered as a first step in cell-cell recognition processes. They are dependent on divalent cations and very specific. However, the low affinity between carbohydrates needs to be compensated with multiple binding events. Multivalent presentation of carbohydrates is therefore a requirement for potential glycocalyx-like surface models.

Gold NPs have been successfully employed in the study of weak carbohydrate-carbohydrate interactions since first reports by Penadés and coworkers.[105,106] Gold NPs coated with the trisaccharide Lewis X (Le^x) were used to study the calcium-dependent Le^x self-aggregation, demonstrating the self-recognition capability of this antigen.[105]

Similarly, Kamerling *et al.* studied carbohydrate-mediated self-recognition of marine sponges. Using gold NPs coated with a sulfated disaccharide and with a pyruvated trisaccharide naturally occurring in the extracellular surface of marine sponge cells, and studying by AFM (Atomic Force Microscopy)[107] its Ca^{2+}-dependent self-recognition, they concluded that the self-interaction between the sulfated disaccharide fragments is stronger than that between the pyruvated trisaccharides. They have also employed NMR spectroscopy[108] [transfer-nuclear Overhauser effect spectroscopy (TR-NOESY) and diffusion ordered (DOSY)-NMR] to obtain structural information related to these interactions.

Russell *et al.* prepared 16 nm gold nanoparticles coated with lactose through thiol-ended chains of different lengths and studied the calcium-induced aggregation by TEM (Transmission Electron Microscopy) and UV-visible spectroscopy.[109] It was shown that the length and nature of the linker had an important effect on the metal ion-mediated carbohydrate-carbohydrate interaction of the lactose nanoparticles.

Aminooxy-functionalized Au nanoparticles lead to glyconanoparticles when reacting with glycosphingolipid (GSL) aldehydes, derived by ozonolysis of natural glycosphingolipid expressed in mouse melanoma B16 cells.[110] The technology named "glycoblotting" (Fig. 20), consisting in glycan enrichment based on chemical reaction, has led to the evaluation by SPR of the interaction between GM3-funtionalized GNPs and self-assembled monolayer (SAM) of Gg3Cer (gangliotriaosylceramide).

7.1.2 Carbohydrate-protein interactions

Studies on carbohydrate–protein interactions using gold GNPs[111] are mostly based on colorimetric methods detecting changes in the surface plasmon resonance (SPR) of Au nanoparticles. Gold GNPs have been developed into a colorimetric biosensor for studying these interactions.[112] The specific recognition by proteins of carbohydrate covering gold GNPs leads to aggregation of nanoparticles causing a color change associated with a shift in the surface plasmon absorption-band of the metal cluster, and this can be monitored using UV-visible spectrophotometry.

Kataoka *et al.* studied the specific and reversible interaction between lactoside-coated gold NPs and *Ricinus communis* agglutinin (RCA_{120}).[113]

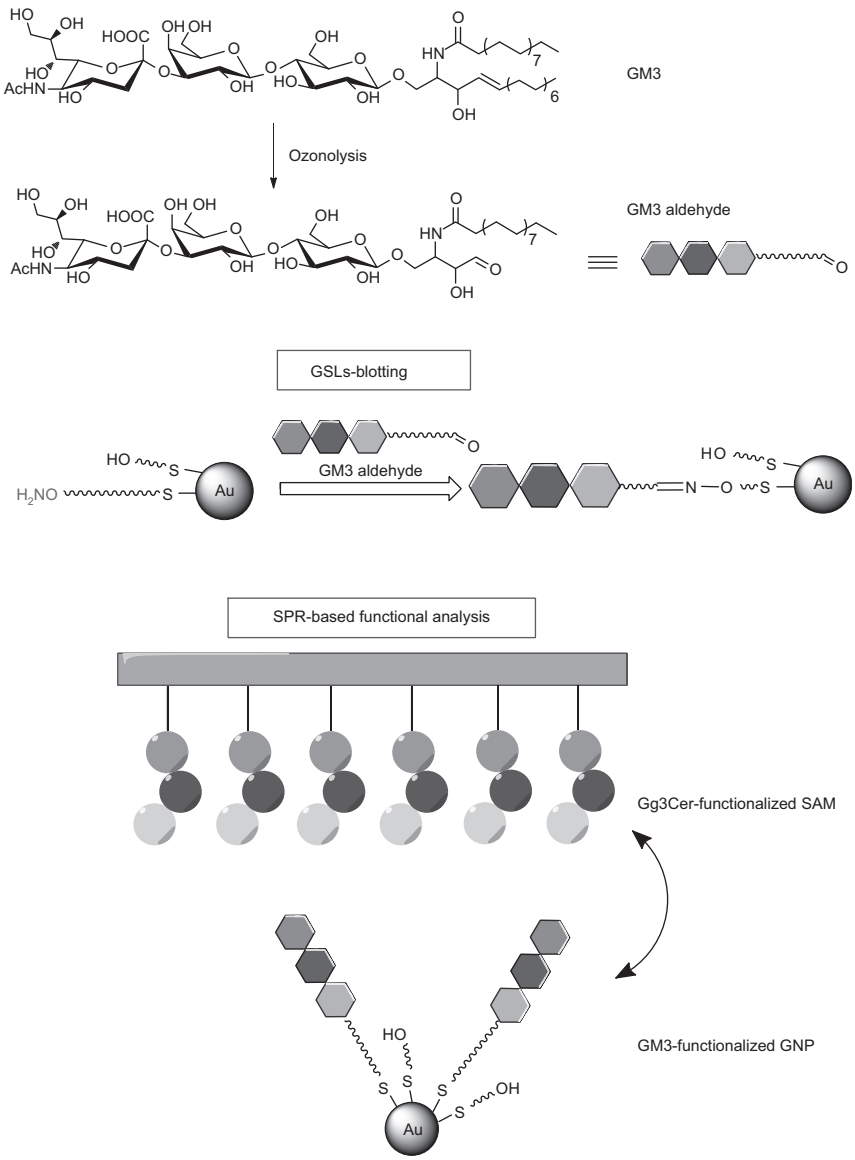


Fig. 20 Glycoblotting methodology and interaction between GM3-functionalized GNPs and Gg3Cer-functionalized SAM.

It was shown that the degree of aggregation is dependent on lectin concentration and on nanoparticle ligand density. Other groups have also worked on colorimetric assays based on aggregation of gold nanoparticles because of their simplicity and versatility. For example, Russell *et al.* synthesized carbohydrate-stabilized gold nanoparticles to detect cholera toxin[114] and RCA_{120},[115] showing that the degree of aggregation is dependent on chain length in addition to the surface ligand density.

Jensen and co-workers tested oxime- and oxyamine-linked (oligo)glucoside-GNPs with Con A, the most broadly used lectin in carbohydrate–protein

interaction studies.[116] The ring-closed tautomer of *N*-glucosyl oximes is recognized by the lectin, whereas the reduced, open chain form, displayed a complete lack of aggregation. In a recent paper Thygesen and Jensen review methods for preparing glyconanoparticles based on chemoselective formation of thiol-ended glycoconjugate intermediates and those ones based on chemoselective reactions between reducing carbohydrates and reactive functional groups on polymer coated nanoparticles.[117]

Interaction between galactose functionalized AuNPs and PA-IL lectin from *Pseudomonas aeruginosa* has been evaluated by hemagglutination inhibition assay (HIA), SPR and isothermal titration calorimetry assays, being Gal-AuNPS the most efficient binding partner for this lectin to date, with a K_d of 50 nM.[118]

Recognition studies of maltose- and maltotriose-NPs of different ligand density with glucoamylase concluded that lower carbohydrate density facilitates accessibility of the sugars to the enzyme.[119] Penadés and co-workers have synthesized lactoside-functionalized NPs and studied their interaction with *Viscum album* agglutinin and their hydrolysis with *E. coli* β-galactosidase.[119] These results proved the influence of the density and presentation of sugar residues toward protein recognition and enzymatic hydrolysis. Thus, selection of correct ligand density is an important factor in order to achieve the desired affinity while escaping degradation by glycosidases. The activity of polysaccharide-degrading enzymes related to pathophysiological alterations in human diseases has also been tested using polysaccharide-GNPs.[120]

Biosensing methods other than colorimetric ones have been developed. Biotinylated carbohydrate-based copolymers containing β-GlcNAc or α-Man residues were conjugated to the surface of gold nanoparticles and immobilized on an avidin-coated DOTLab® sensor chip (Diffraction Optic Technology) to study the specific recognition of carbohydrates with lectins Con A or wheat germ agglutinin (WGA),[121] the first one recognizing α-Man, and the second one recognizing β-GlcNAc residues.

Interaction between sialic acid-coated gold nanoparticles and Alzheimer's amyloid-beta (Aβ) peptides has been electrochemically tested.[122] Yan and co-workers developed a fluorescent competition binding assay method to measure the binding affinity of GNPs-lectin interaction, showing that carbohydrate ligands conjugated to AuNPs exhibited affinities up to five orders of magnitude higher than those of the corresponding monomeric ligands with lectins.[123] Cyanovirin-N, a cyanobacterial lectin which exhibits inhibitory activity against a number of viruses, including HIV, has been recently evaluated by a fluorescence-based competition assay.[124] Furthermore, a method for protein detection based on surface energy transfer (SET) has recently been reported to study the interaction between fluorescent labeled FITC-Con A (fluorescein isothiocyanate-Con A) and mannose-functionalized-dendrimer covering gold nanoparticles.[125]

Among other analytical applications, gold GNPs have been successfully employed to isolate and purificate lectins,[126] and for the quantitative determination by SPR of small molecular weight analytes like galactose, using lactosyl-(polyethylene glycol)-gold NPs adsorbed onto *Ricinus communis* agglutinin (RCA_{120}).[127]

7.1.3 Gold GNPs as non-viral gene delivery agents

Non-viral vectors are attracting much attention in gene therapy in order to overcome the safety problems of their viral counterpart.[128] Gold GNPs may be used as non-viral gene delivery agents as Glc- and Gal-gold nanoparticles were highly efficient DNA binders. Interaction between GNPs and DNA showed binding affinity dependence on the sugar ligand.[129]

7.1.4 Gold GNPs as antiadhesion agents

Many pathogenic infections start through multivalent carbohydrate-dependent adhesion of microorganisms to host cells. Carbohydrate-functionalized nanomaterials could act as antiadhesion agents avoiding the adhesion of microbes to host cells and thus, becoming an alternative to antibiotic strategy. Manea *et al.*[130] have assembled thiolated synthetic analogues of type A *Neisseria meningitides* antigens on the surface of gold nanoparticles to inhibit the binding of this bacterium to specific mouse polyclonal antibodies.

Gold Gal- and Glc-NPs were employed to evaluate their ability to displace HIV envelope glycoproteins gp120 from plate-bound galactosylceramides.[131] Promising results have also been achieved by Penadés and co-workers when using gold GNPs to study the multivalent interactions with gp120. They prepared a set of hybrid high-manno-NPs to mimic the presentation of oligomannosides on the HIV envelope protein gp120 and have studied them as inhibitors of DC-SIGN–gp120 interactions.[100] GNPs containing the disaccharide Manα1-2Manα (Fig. 21) showed more than 20,000-fold increased activity of inhibition compared to the monomeric disaccharide.

These high mannose GNPs, mimicking the cluster presentation of glycoprotein gp120 on the virus surface,[100] were tested in cell-based experiments for evaluating their anti-HIV potential as inhibitors of DC-SIGN-mediated HIV-1 trans-infection of human T cells.[101] Low percentages of infection compared to untreated control were found; therefore

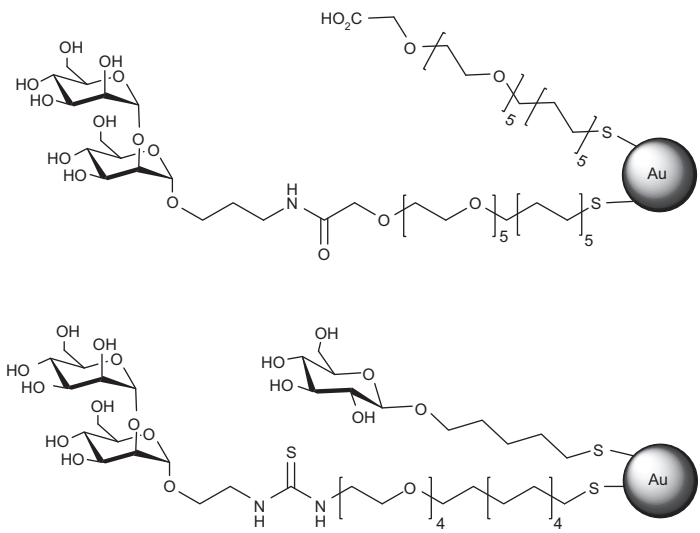


Fig. 21 Hybrid Manα1-2Manα glyconanoparticles.

these GNPs may function as anti-adhesive barrier at an early stage of HIV infection.

Affinity information on the molecular recognition of synthetic linear and branched-oligomannosides with the anti-HIV-1 antibody 2G12 has been reported using NMR techniques, showing that repetition of binding motifs or increased complexity of the oligomannosides do not necessarily involve an enhancement of affinity.[132] AuNPs coated with a linear tetra-mannoside conjugate inhibit 2G12-gp120 binding with IC_{50} values in the micromolar range, and could block the 2G12-mediated neutralization of a HIV infection, under conditions similar to the ones in which normal serum prevents infection.[133] It has been described that glucose used as a second ligand in folate coated-gold nanoparticles enhanced cell recognition of cancer cells as a result of multivalent interactions between both ligands and cells.[134]

7.1.5 Gold GNPs as antitumor agents

The specific anti-metastatic effect of lactose-GNPs on development of lung tumor foci in mice has been demonstrated.[135] Mice were injected with melanoma cells (B16F10) pre-incubated with Lac-GNPs and after three weeks, the lungs were analyzed and compared to the control ones, showing an inhibition up to 70% of metastasis in lung cells of mice. Heparin coated AuNPs exhibited effective inhibition of basic fibroblast growth factor 2 (FGF-2)-induced angiogenesis in a mouse model. Naked Au administrated at the same concentration resulted lethal to the animals, demonstrating that carbohydrates impart biocompatibility to these particles.[136]

7.1.6 Gold GNPs as tool for pathogen and toxin detection

Pathogens often have carbohydrate-binding proteins to recognize and invade host organisms or to deliver toxins. Therefore, GNPs could be used as a new methodology for rapid and sensitive pathogen and toxin detection. For example, globotriose-functionalized gold NPs have been designed for detection and inhibition of verotoxin, a Shiga-like toxin.[137] These AuNPs presenting carbohydrate analogues to the glycolipids on cell surface can effectively compete for these toxins, being the inhibition dependent on the structure and density of the glycans.[138]

Russell and co-workers reported a rapid colorimetric bioassay based on SPR for detection and quantification of cholera toxin using carbohydrate-stabilized gold nanoparticles.[114] Glyconanoparticles were also used by Field and co-workers to elucidate structural and mechanistical information of the role of sialic acid recognition in the host infection by *Trypanosoma cruzi*. These nanoparticles designed as *T. cruzi* mucin glycan mimetics were used in lectin binding assays and the results highlighted the importance of the branched structures of mucin glycans in biological recognition. In addition, sialic acid tethered to gold nanoparticles was used to detect siglec lectin specificity.[89]

7.1.7 Gold GNPs for molecular and cellular imaging

There is an increasing demand for new probes in the field of *in vivo* optical imaging. Quantum dots (QDs), magnetic and gold nanoparticles have

demonstrated to be sensitive and target-specific imaging. Herein we report some examples where carbohydrate-functionalized gold nanoparticles have also been applied for *in vivo* imaging.

Multifunctional gold GNPs coated with hyaluronic acid (HA) residues labeled with a fluorescence dye have been used for *in vivo* imaging of arthritic inflammation and human ovarian carcinoma tumor (OVCAR-3) in mice.[139] *Ex vivo* examination of tissues revealed large amounts of HA-AuNPs accumulated in tumor sites. A strategy to confer magnetic properties to noble-metal GNPs consists in the functionalization of the metal surface. Hybrid gold nanoparticles coated with sugar conjugates and Gd(III) complexes have been prepared as paramagnetic probes for Magnetic Resonance Imaging (MRI) and were used for *in vivo* imaging of glioma in mouse brain.[140] The behavior of these GNPS as *in vivo* contrast agents is dependent of the nature of the sugar and the relative position of the sugar with respect to the gadolinium ion.

7.2 Magnetic glyconanoparticles

Magnetic glyconanoparticles (MGNPs) are materials with potential application in the field of clinical diagnosis and therapeutics and have mainly been used as contrast agents for MRI.

7.2.1 Lectin interaction and bacterial detection with MGNPs

Carbohydrates attached to magnetic NPs could be used in the recognition of lectins such as Con A. For example, fluorescein isocyanate-labeled Con A recognizes Man-MGNPs and could be separated from the medium by a magnetic field.[141]

MGNPs could also serve as a method for pathogen detection and decontamination, such as *E.coli* ORN178 containing mannose binding proteins on surface. Silica-coated iron oxide magnetite nanoparticles, functionalized with mannose or galactose residues by *click* chemistry or peptide coupling were used by the group of Huang[142] to detect *E. coli* strains with a detection limit of 10^4 cells/mL. Taking advantage of their magnetic properties, up to 88% of the bacteria were removed from the medium (Fig. 22). Furthermore, based on the response patterns three *E. coli* strains could be differentiated. Detection of pathogenic *Streptococcus suis* bacteria was achieved using magnetic glyconanoparticles bearing biotinylated sugar and an ATP detection system.[143]

7.2.2 Cell binding with MGNPs

Gold-iron oxide superparamagnetic glyconanoparticles were first synthesized by reduction in water a mixture of gold and iron salts in the presence of thiol-ended glycoconjugates.[144] The interaction of these Fe/Au nanoparticles with human fibroblast cell lines was tested, showing that these cells recognize selectively the sugars on the nanoparticle surface.[145]

Several hyaluronic acid-coated magnetic nanoparticles have been reported as target-specific MR imaging probes in cancer cells[146] and macrophages.[147] Recently, El-Boubbou[141] prepared MGNPs with different monosaccharides as a nanosensor system to detect and differentiate nine types of cancer cells using a statistic method based on different response patterns.

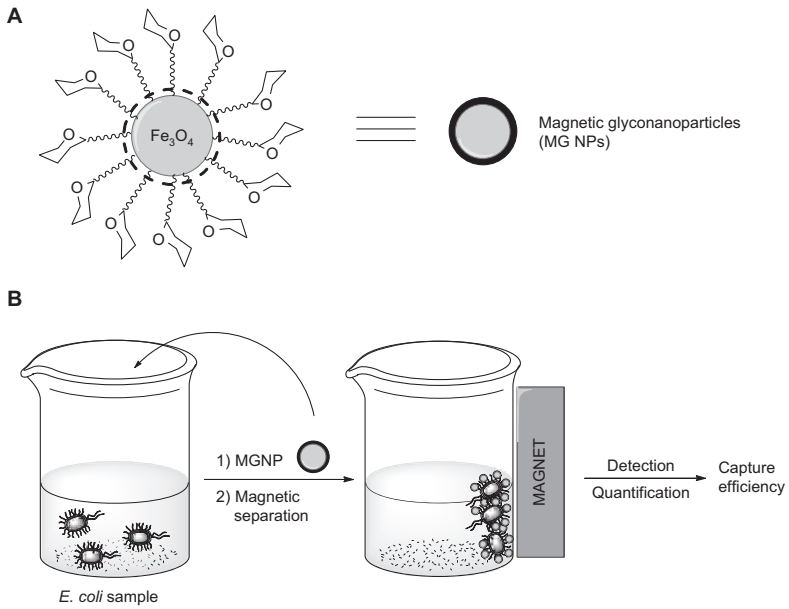


Fig. 22 *Escherichia coli* detection and decontamination with MGNPs. A) Silica coated magnetic glyconanoparticles. B) Pathogen detection.

Some recent examples of multimodal magnetic immuno-glyconanoparticles designed for immunolabeling and imaging of cells have also been reported.[148,149] In these nanoparticles, carbohydrates confer them biocompatibility, stability and resistance to nonspecific adsorption.

7.2.3 *In vivo* applications of MGNPs

Diverse carbohydrate-functionalized magnetic nanoparticles, synthesized by chemical functionalization of previously prepared iron oxide nanoparticles, were design for *in vivo* inflammation detection. sLe^X-MGNP resulted to be a highly sensitive and selective T_2 contrast agent for direct detection of E-/P-selectins, endothelial markers in acute inflammation.[150]

7.3 Glyco quantum dots

NPs can be improved by modifying their surfaces by attaching a fluorescent dye or by modifying the nature of the core with semiconductor nanocrystals like CdSe, CdTe or InP (quantum dots, QDs). A combination of the two methodologies could lead to a bimodal imaging probe.[151] Chemical and physical properties of quantum dots (size-tunable light emission, bright luminescence, broad absorption spectra and resistance to photobleaching) have made this material popular as bioluminescent probes for live-cell and *in vivo* animal imaging.[152]

7.3.1 Protein recognition

QDs, as well as silver and magnetic GNPs, functionalized with dextran have been used to study the interaction with lectins such as Con A.[153] The specific

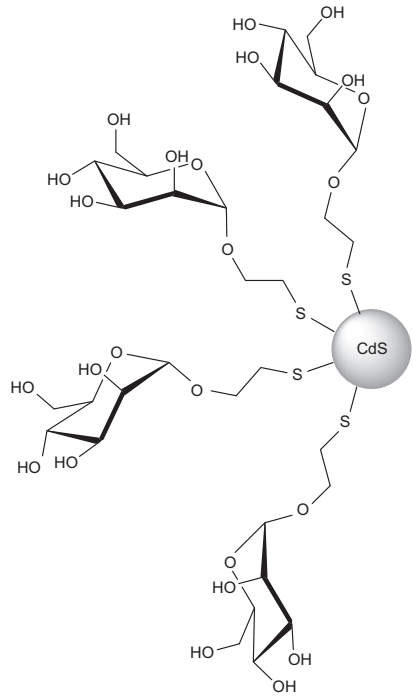


Fig. 23 Mannose-QDs for the detection of *E. coli*.

interactions between soybean agglutinin and Con A with CdSe/ZnS glyco-QDs have been monitored using the light scattered at 600 nm.[154] Mannose-coated CdS quantum dots (Man-QDs) (Fig. 23) induce luminescent aggregates of *E. coli* which can be used to detect bacteria in cell suspensions, containing as few as 10^4 bacteria per mL. The aggregation process is due to *E. coli* cell surface FimH mannose-specific lectins and did not occur either with an *E. coli* strain defective in the FimH lectin or with galactose-coated QDs.[155]

7.3.2 Live-cell and *in vivo* animal imaging

Tri-antennary β-galactoside dendrons attached to CdSe/ZnS nanoparticles were smoothly delivered inside HeLa, kidney, or metastatic lung cancer cells. The transmembrane efficiency is increased with increasing density of asialoprotein receptors on cellular surface, indicating a receptor-mediated endocytosis mechanism.[156] Biocompatible chitosan nanoparticles covalently bound to InGaP/ZnS QDs were designed as fluorescent probes for deep-tissue imaging, improving cellular uptake drastically and allowing deeper tissue penetration in comparison with PEG–QD nanoparticles.[157]

Highly biocompatible mannosylated QDs were prepared from commercially available PEG-coated quantum dots in order to interact with mannose receptor on macrophages thus tracking macrophage migration and fate, what is important since macrophages play crucial roles in many pathophysiological processes.[158] QDs bound to biotinylated polymeric Le^X or to HIV-1 envelope glycoprotein gp120 were employed to monitor in

living cells antigen-binding, entry and fate of gp120-QDs conjugates once uptaken by cells.[159]

Niikura and co-workers studied the uptake of glyco-QDs by digitonin-permeabilized HeLa cells. It was observed than GlcNAc-QDs accumulate in the endoplasmic reticulum in the presence of ATP.[160] When oligosaccharides[161] were attached to the QDs, the fluorescence intensity measured in the nucleus showed different degree of accumulation depending on the sugar covering QDs.

Recently, an electrochemiluminescence (ECL) strategy for monitoring carbohydrate expression on living cells was designed using carbohydrate-functionalized QDs immobilized on carbon nanotube electrodes. Based on the specific interactions between lectins and the corresponding carbohydrates, this method was used to calculate the average number of mannosyl residues on living human gastric carcinoma cell surface. This method has potential application to decode mechanisms of carbohydrate-related biological processes when dynamic changes in cell surface carbohydrates are involved.[162] Carbohydrate-capped QDs can be exploited for *in vivo* targeting as well. Different sugar-PEGylated QDs were tested *in vivo* for specific liver sequestration.[163] Man- and GalN-QDs accumulated selectively in the liver due to the presence of specific receptors on hepatic cells.[164]

8 Conclusions

An impressive development on the synthesis of multivalent glycoconjugates has taken place in the last few years. The election of the multivalent scaffold matched with the proper spacer and the right carbohydrate structure plays a key role in modulating carbohydrate-protein interactions. An overview of different families of structural assemblies presenting multiple carbohydrate moieties such as glycodendrimers, glyconanotubes, glyconanoparticles, show us their multiple biological and medical applications, among which the design of antitumoral vaccines constitutes a landmark.

Acknowledgements

We thank the Dirección General de Investigación of Spain (CTQ2008 02813) and Junta de Andalucía (FQM 134) for financial support. A. O. thanks University of Seville of Spain for the award of a fellowship.

References

1 R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683.
2 B. Ernst and J. L. Magnani, *Nat. Rev. Drug Discov.*, 2009, **8**, 661.
3 H. Lis and N. Sharon, *Chem. Rev.*, 1998, **98**, 637.
4 M. Mammem, S. K. Choi and G. M. Whitesides, *Angew. Chem. Int. Ed.*, 1998, **37**, 2755.
5 Y. C. Lee and R. T. Lee, *Acc. Chem. Res.*, 1995, **28**, 321.
6 D. Deniaud, K. Julienne and S. G. Gouin, *Org. Biomol. Chem.*, 2011, **9**, 966.
7 R. J. Pieters, *Med. Res. Rev.*, 2007, **27**, 796.
8 R. J. Pieters, *Org. Biomol. Chem.*, 2009, **7**, 2013.
9 Y-B. Limand and M. Lee, *Org. Biomol. Chem.*, 2007, **5**, 401.

10 A. L. Baldini, A. Casnati, F. Sansone and R. Ungaro, *Chem. Soc. Rev.*, 2007, **36**, 254.
11 P. Niederhafner, J. Ŝebestík and J. Ježek, *J. Pept. Sci.*, 2008, **14**, 2.
12 P. Niederhafner, J. Ŝebestík and J. Ježek, *J. Pept. Sci.*, 2008, **14**, 44.
13 A. Imberty, Y. M. Chabre and R. Roy, *Chem. Eur. J.*, 2008, **14**, 7490.
14 Y. M. Chabre and R. Roy, *Adv. Carbohydr. Chem. Biochem.*, 2010, **63**, 165.
15 J. J. Lundquist and E. J. Toone, *Chem. Rev.*, 2002, **102**, 555.
16 M. N. Alekshun and S. B. Levy, *Cell*, 2007, **128**, 1037.
17 R. J. Pieters, *Drug Discov. Today Technol.*, 2009, **6**, e27.
18 A. V. Pukin, H. M. Branderhorst, C. Sisu, C. A. G. M. Weijers, M. Gilbert, R. M. J. Liskamp, G. M. Visser, H. Zuilhof and R. J. Pieters, *ChemBioChem*, 2007, **8**, 1500.
19 H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser and R. J. Pieters, *Chem. Commun.*, 2007, 5043.
20 V. Tchesnokova, P. Aprikian, D. Kisiela, S. Gowey, N. Korotkova, W. Thomas and E. Sokurenko, *Infect. Immun.*, 2011, **79**, 3895.
21 S. G. Gouin, A. Wellens, J. Bouckaert and J. Kovensky, *ChemMedChem*, 2009, **5**, 749.
22 S. G. Gouin, J. M. García Fernández, E. Vanquelef, F.-Y. Dupradeau, E. Salomonsson, H. Leffler, M. Ortega-Muñoz, U. J. Nilsson and J. Kovensky, *ChemBioChem*, 2010, **11**, 1430.
23 C. D. Heidecke and T. K. Lindhorst, *Chem. Eur. J.*, 2007, **13**, 9056.
24 M. Hartmann and T. K. Lindhorst, *Eur. J. Org. Chem.*, 2011, 3583.
25 G. Lin, G. Simmons, S. Pohlmann, F. Baribaud, H. Ni, G. J. Leslie, B. S. Haggarty, P. Bates, D. Weissman, J. A. Hoxie and R. W. Doms, *J. Virol.*, 2003, **77**, 1337.
26 L. Balboa, M. M. Romero, N. Yokobori, P. Schierloh, L. Geffner, J. I. Basile, R. M. Musella, E. Abbate, S de la Barrera, M. C. Sasiain and M. Alemán, *Immunol. Cell Biol.*, 2010, **88**, 716.
27 S. Sattin, A. Daghetti, M. Theépaut, A. Berzi, M. Saénchez-Navarro, G. Tabarani, J. Rojo, F. Fieschi, M. Clerici and A. Bernardi, *ACS Chem. Biol.*, 2010, **5**, 301.
28 A. Berzi, J. J. Reina, R. Ottria, I. Sutkeviciute, P. Antonazzo, M. Sánchez-Navarro, E. Chabrol, M. Biasin, D. Trabattoni, I. Cetin, J. Rojo, F. Fieschi, A. Bernardi and M. Clerici, *AIDS*, 2012, **26**, 127.
29 P. A. Ledin, F. Friscourt, J. Guo and G.-J. Boons, *Chem. Eur. J.*, 2011, **17**, 839.
30 P. Wu, M. Malkoch, J. N. Hunt, R. Vestberg, E. Kaltgrad, M. G. Finn, V. V. Fokin, K. B. Sharpless and C. J. Hawker, *Chem. Commun.*, 2005, 5775.
31 B. Mokhtari and K. Pourabdollah, *J. Incl. Phenom. Macrocycl. Chem.*, 2012, **73**, 1.
32 D. Arosio, M. Fontanella, L. Baldini, L. Mauri, A. Bernardi, A. Casnati, F. Sansone and R. Ungaro, *J. Am. Chem. Soc.*, 2005, **127**, 3660.
33 A. Marra, L. Moni, D. Pazzi, A. Corallini, D. Bridi and A. Dondoni, *Org. Biomol. Chem.*, 2008, **6**, 1396.
34 K. Krenek, M. Kuldova, K. Hulikova, I. Stibor, P. Lhotak, M. Dudic, J. Budka, H. Pelantova, K. Bezouska, A. Fiserova and V. Kren, *Carbohydr. Res.*, 2007, **342**, 1781.
35 S. André, F. Sansone, H. Kaltner, A. Casnati, J. Kopitz, H.-J. Gabius and R. Ungaro, *ChemBioChem*, 2008, **9**, 1649.
36 F. Sansone, L. Baldini, A. Casnati and R. Ungaro, *Supramol. Chem.*, 2008, **20**, 161.
37 S. Cecioni, R. Lalor, B. Blanchard, J.-P. Praly, A. Imberty, S. E. Matthews and S. Vidal, *Chem. Eur. J.*, 2009, **15**, 13232.

38 N. Mesaros, P. Nordmann, P. Plesiat, M. Roussel-Delvallez, J. Van Eldere, Y. Glupczynski, Y. Van Laethem, F. Jacobs, P. Lebecque, A. Malfroot, P. M. Tulkens and F. van Bambeke, *Clin. Microbiol. Infect*., 2007, **13**, 560.
39 D. Sicard, S. Cecioni, M. Iazykov, Y. Chevolot, S. E. Matthews, J.-P. Praly, E. Souteyrand, A. Imberty, S. Vidal and M. Phaner-Goutorbe, *Chem. Commun*, 2011, **47**, 9483.
40 S. Cecioni, J.-P. Praly, S. E. Matthews, M. Wimmerová, A. Imberty and S. Vidal, *Chem. Eur. J*., 2012, **18**, 6250.
41 E. Galante, C. Geraci, S. Sciuto, V. L. Campo, I. Carvalho, R. Sesti-Costa, P. M. M. Guedes, J. S. Silva, L. Hill, S. A. Nepogodiev and R. A. Field, *Tetrahedron*, 2011, **67**, 5902.
42 P. Poignard, E. Ollmann Saphire, P. W. H. I. Parren and D. R. Burton, *Annu. Rev. Immunol*., 2001, **19**, 253.
43 J. Grove and M. Marsh, *J. Cell Biol*., 2011, **195**, 1071.
44 A. M. Sánchez-Navarro, A. Muñoz, B. M. Illescas, J. Rojo and N. Martín, *Chem. Eur. J*., 2011, **17**, 766.
45 J.-F. Nierengarten, J. Iehl, V. Oerthel, M. Holler, B. M. Illescas, A. Muñoz, N. Martín, J. Rojo, M. Sánchez-Navarro, S. Cecioni, S. Vidal, K. Buffet, M. Durka and S. P. Vincent, *Chem. Commun*., 2010, **46**, 3860.
46 S. Hakomori, *Proc. Natl Acad. Sci. USA*, 2002, **99**, 10231.
47 M. A. Tarp, A. L. Sørensen, U. Mandel, H. Paulsen, J. Burchell, J. Taylor-Papadimitriou and H. Clausen, *Glycobiology*, 2007, **17**, 197.
48 Tumor-Associated Antigens, eds. O. Gires and B. Seliger, Wiley-VCH, 2009.
49 M. A. Hollingsworth and B. J. Swanson, *Nature Rev*., *Cancer*, 2004, **4**, 45.
50 L. Peña Icart, V. Fernández-Santana, R. C. Veloso, T. Carmenate, S. Sirois, R. Roy and V. Vérez-Bencomo, in *Carbohydrate-Based Vaccines*, ed. R. Roy, *ACS Symp. Ser*., Washington DC, 2008, vol. 989, p. 1.
51 R. M. Wilson, J. D. Warren, O. Ouerfelli and S. J. Danishefsky, in *Carbohydrate-based Vaccines*, ed. R. Roy, *ACS Symp. Ser*., *Washington DC*, 2008, vol. 989, p. 258.
52 Z. Guo and G. J. Boons (Eds), *Carbohydrate-Based Vaccines and Immunotherapies*, John Wiley & Sons, Hoboken, New Jersey, 2009.
53 Z. Guo and Q. Wang, *Curr. Opin. Chem. Biol*., 2009, **13**, 608.
54 A. Liakatos and H. Kunz, *Curr. Opin. Mol. Ther*., 2007, **9**, 35.
55 U. Westerlind and H. Kunz, *Chimia*, 2011, **65**, 30.
56 R. Roy and T. C. Shiao, *Chimia*, 2011, **65**, 24.
57 L. Morelli, L. Poletti and L. Lay, *Eur. J. Org. Chem*., 2011, 5723.
58 U. Westerlind and H. Kunz, in *Carbohydrate Chemistry*, eds. A. P. Rauter, T. Lindhorst, RSC, Cambridge, UK, 2010, vol. 36, 1.
59 P. Niederhafner, M. Reiniš, J. Šebestík and J. Ježek, *J. Pept. Sci*., 2008, **14**, 556.
60 R. D. Atronomo and D. R. Burton, *Nat. Rev. Drug Discov*., 2010, **9**, 308.
61 T. A. Gilewski, G. Ragupathi, M. Dickler, S. Powell, S. Bhuta, K. Panageas, R. R. Koganty, J. Chin-Eng, C. Hudis, L. Norton, A. N. Houghton and P. O. Livingston, *Clin. Cancer Res*., 2007, **13**, 2977.
62 T. C. Shiao and R. Roy, *New J. Chem*., 2012, **36**, 324.
63 S. Keil, A. Kaiser, F. Syed and H. Kunz, *Synthesis*, 2009, 1355.
64 T. Becker, A. Kaiser and H. Kunz, *Synthesis*, 2009, 1113.
65 A. Kaiser, N. Gaidzik, T. Becker, C. Menge, K. Groh, H. Cai, Y. M. Li, B. Gerlitzki, E. Schmitt and H. Kunz, *Angew. Chem. Int. Ed*., 2010, **49**, 3688.
66 H. Cai, Z.-H. Huang, L. Shi, Y.-F. Zhao, H. Kunz and Y.-M. Li, *Chem. Eur. J*., 2011, **17**, 6396.
67 B. L. Wilkinson, C. K. Y. Chun and R. J. Payne, *Biopolymers*, 2011, **96**, 137.

68 B. L. Wilkinson, L. R. Malins, C. K. Y. Chun and R. J. Payne, *Chem. Comm.*, 2010, **46**, 6249.
69 S. Bay, S. Fort, L. Birikaki, C. Ganneau, E. Samain, Y.-M. Coïc, F. Bonhomme, E. Dériaud, C. Leclerc and R. Lo-Man, *ChemMedChem*, 2009, **4**, 582.
70 J. Zhu, Q. Wan, D. Lee, G. Yang, M. K. Spassova, O. Ouerfelli, G. Ragupathi, P. Damani, P. O. Livingston and S. J. Danishefsky, *J. Am Chem Soc.*, 2009, **131**, 9298.
71 S. J. Danishefsky, G. Ragupathi, P. O. Livingston, J. Zhu, K. Iyer, G. Yang and P. Sabbatini, 2011, WO 2011/156774 A2.
72 I. Jeon, D. Lee, I. J. Krauss and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2009, **131**, 14337.
73 I. J. Krauss, J. G. Joyce, A. C. Finnefrock, H. C. Song, V. Y. Dudkin, X. J. Geng, D. Warren, M. Chastain, J. W. Shiver and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2007, **129**, 11042.
74 J. G. Joyce, I. J. Krauss, H. C. Song, D. W. Opalka, K. M. Grimm, D. D. Nahas, M. T. Esser, R. Hrin, M. Feng, V. Y. Dudkin, M. Chastain, J. W. Shiver and S. J. Danishefsky, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 15684.
75 S.-K. Wang, P.-H. Liang, R. D. Astronomo, T.-L. Hsu, S.-L. Hsieh, D. R. Burton and C.-H. Wong, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 3690.
76 A. Kabanova, R. Adamo, D. Proietti, F. Berti, M. Tontini, R. Rappuoli and P. Costantino, *Glycoconjugate J.*, 2010, **27**, 501.
77 M. C. Galan, D. Benito-Alifonso and G. M. Watt, *Org. Biomol. Chem.*, 2011, **9**, 3598.
78 V. Pozsgay, *Curr. Top. Med. Chem.*, 2008, **8**, 126.
79 T. Lipinski, P. I. Kitov, A. Szpacenko, E. Paszkiewicz and D. R. Bundle, *Bioconjugate Chem.*, 2011, **22**, 274.
80 H. Wang, L. Gu, Y. Lin, F. Lu, M. J. Meziani, P. G. Luo, W. Wang, L. Cao and Y.-P. Sun, *J. Am. Chem. Soc.*, 2006, **128**, 13364.
81 P. G. Luo, H. Wang, L. Gu, F. Lu, Y. Lin, K. A. Christensen, S.-T. Yang and Y.-P. Sun, *ACS Nano*, 2009, **3**, 3909.
82 L. Gu, P. G. Luo, H. Wang, M. J. Meziani, Y. Lin, L. M. Veca, L. Cao, F. Lu, X. Wang, R. A. Quinn, W. Wang, P. Zhang, S. Lacher and Y. P. Sun, *Biomacromolecules*, 2008, **9**, 2408.
83 M. K. Müller and L. Brunsveld, *Angew. Chem., Int. Ed.*, 2009, **48**, 2921.
84 T. Horlacher and P. H. Seeberger, *Chem. Soc. Rev.*, 2008, **37**, 1414.
85 S. Park, M.-R. Lee and I. Shin, *Chem. Soc. Rev.*, 2008, **37**, 1579.
86 C. Y. Wu, P.-H. Liang and C.-H. Wong, *Org. Biomol. Chem.*, 2009, **7**, 2247.
87 H. M. Branderhorst, R. Ruijtenbeek, R. M. J. Liskamp and R. J. Pieters, *ChemBioChem*, 2008, **9**, 1836.
88 R. M. de Lederkremer and R. Agusti, *Adv. Carbohydr. Chem. Biochem.*, 2009, **62**, 311.
89 R. A. Field, P. Andrade, V. L. Campo, I. Carvalho, B. Y. M. Collet, P. R. Crocker, M. Fais, R. Karamanska, B. Mukhopadhayay, S. A. Nepogodiev, A. Rashid, M. Rejzek, D. A. Russell, C. L. Schofield and R. M. van Well, in *Petite and Sweet: Glyco-Nanotechnology as a Bridge to New Medicines*, eds. X. Huang, J. J. Barchi, Jr, *ACS Symp. Ser.*, Washington DC, 2011, vol. 1091, p. 143.
90 R. Karamanska, J. Clarke, O. Blixt, J. I. MacRae, J. Q. Zhang, P. R. Crocker, N. Laurent, A. Wright, S. L. Flitsch, D. A. Russell and R. A. Field, *Glycoconj. J.*, 2008, **25**, 69.
91 M. Fais, R. Karamanska, S. Allman, S. A. Fairhurst, P. Innocenti, A. J. Fairbanks, T. J. Donohoe, B. G. Davis, D. A. Russell and R. A. Field, *Chem. Sci.*, 2011, **2**, 1952.

92 M. Marradi, M. Martín-Lomas and S. Penadés, *Adv. Carbohydr. Chem. Biochem.*, 2010, **64**, 211.
93 R. Ojeda, J. L. de Paz, A. G. Barrientos, M. Martin-Lomas and S. Penadés, *Carbohydr. Res.*, 2007, **342**, 448.
94 P. Ghosh, G. Han, M. De, C. K. Kim and V. M. Rotello, *Adv. Drug Deliv. Rev.*, 2008, **60**, 1307.
95 S. C. McBain, H. H. P. Yiu and J. Dobson, *Int. J. Nanomed.*, 2008, **3**, 169.
96 H. Yang, H. Li and X. Jiang, *Microfluid. Nanofluid.*, 2008, **5**, 571.
97 S. Parveen, R. Misra and S. K. Sahoo, *Nanomed. Nanotechnol. Biol. Med.*, 2012, **8**, 147.
98 K. B. Hartman, L. J. Wilson and M. G. Rosenblum, *Mol. Diagn. Ther.*, 2008, **12**, 1.
99 E. Boisselier and D. Astruc, *Chem. Soc. Rev.*, 2009, **38**, 1759.
100 O. Martínez-Ávila, K. Hijazi, M. Marradi, C. Clavel, C. Campion, C. Kelly and S. Penadés, *Chem. Eur. J.*, 2009, **15**, 9874.
101 O. Martínez-Ávila, L. M. Bedoya, M. Marradi, C. Clavel, J. Alcamí and S. Penadés, *ChemBioChem*, 2009, **10**, 1806.
102 J. J. Barchi, Jr., in *Petite and Sweet: Glyco-Nanotechnology as a Bridge to New Medicines*, eds. X. Huang, J. J. Barchi Jr, *ACS Symp. Ser.*, Washington DC, 2011, vol. 1091, p. 161.
103 I. García, J. Gallo, M. Marradi, and S. Penadés, in *Engineered Carbohydrate-Based Materials for Biomedical Applications: Polymers, Surfaces, Dendrimers, Nanoparticles, and Hydrogels*, ed. R. Narain, John Wiley & Sons, 2011, p. 213.
104 Z. Shervani and Y. Yamamoto, *Carbohydr. Res.*, 2011, **346**, 651.
105 J. M. de la Fuente, A. G. Barrientos, T. C. Rojas, J. Rojo, J. Cañada, A. Fernández and S. Penadés, *Angew. Chem. Int. Ed.*, 2001, **40**, 2257.
106 J. M. de la Fuente and S. Penadés, *Glycoconj. J.*, 2004, **21**, 149.
107 A. Carvalho de Souza, D. N. Ganchev, M. M. E. Snel, J. van der Eerden, J. F. G. Vliegenthart and J. P. Kamerling, *Glycoconj. J.*, 2009, **26**, 457.
108 J. L. Santos, A. C. de Souza, F. J. Canada, S. Martín-Santamaría, J. P. Kamerling and J. Jiménez-Barbero, *ChemBioChem.*, 2009, **10**, 511.
109 A. J. Reynolds, A. H. Haines and D. A. Russell, *Langmuir*, 2006, **22**, 1156.
110 N. Nagahori, M. Abe and S. I. Nishimura, *Biochemistry*, 2009, **48**, 583.
111 W.-T. Chien, C.-C. Yu, C.-F. Liang, C.-H. Lai, P.-C. Lin, C. C. Lin, *in Petite and Sweet: Glyco-Nanotechnology as a Bridge to New Medicines*, eds. X. Huang, J. J. Barchi, Jr, ACS Symp. Ser., Washington DC, 2011, vol. 1091, p. 15.
112 X. Wang, L.-H. Liu, O. Ramström and M. Yan, *Exp. Biol. Med.*, 2009, **234**, 1128.
113 S. Takae, Y. Akiyama, H. Otsuka, T. Nakamura, Y. Nagasaki and K. Kataoka, *Biomacromol.*, 2005, **6**, 818.
114 C. L. Schofield, R. A. Field and D. A. Russell, *Anal. Chem.*, 2007, **79**, 1356.
115 C. L. Schofield, B. Mukhopadhyay, S. M. Hardy, M. B. McDonnell, R. A. Field and D. A. Russell, *Analyst*, 2008, **133**, 626.
116 M. B. Thygesen, J. Sauer and K. J. Jensen, *Chem. Eur. J.*, 2009, **15**, 1649.
117 M. B. Thygesen and K. J. Jensen, in *Petite and Sweet: Glyco-Nanotechnology as a Bridge to New Medicines*, eds. X. Huang, J. J. Barchi, Jr, *ACS Symp. Ser.*, Washington DC, 2011, vol. 1091, p. 37.
118 M. Reynolds, M. Marradi, A. Imberty, S. Penadés and S. Pérez, *Chem. Eur. J*, 2012, **18**, 4264.
119 A. G. Barrientos, J. M. de la Fuente, M. Jiménez, D. Solís, J. Cañada, M. Martín-Lomas and S. Penadés, *Carbohydr. Res.*, 2009, **344**, 1474.
120 Z. Ban, C. J. Bosques and R. Sasisekharan, *Org. Biomol. Chem.*, 2008, **6**, 4290.

121 X. Jiang, A. Housni, G. Gody, P. Boullanger, M.-T. Charreyre, T. Delair and R. Narain, *Bioconjugate Chem*., 2010, **21**, 521.
122 M. Chikae, T. Fukuda, K. Kerman, K. Idegami, Y. Miura and E. Tamiya, *Bioelectrochem*., 2008, **74**, 118.
123 X. Wang, O. Ramström and M. Yan, *Anal. Chem*., 2010, **82**, 9082.
124 X. Wang, E. Matei, L. Deng, O. Ramström, A. M. Gronenborn and M. Yan, *Chem. Commun*., 2011, **47**, 8620.
125 N. Bogdan, R. Roy and M. Morin, *RSC Adv*., 2012, **2**, 985.
126 S. Nakamura-Tsuruta, Y. Kishimoto, T. Nishimura and Y. Suda, *J. Biochem*., 2008, **143**, 833.
127 S. Takae, Y. Akiyama, Y. Yamasaki, Y. Nagasaki and K. Kataoka, *Bioconjugate Chem*, 2007, **18**, 1241.
128 A. Ragusa, I. García and S. Penadés, *IEEE Trans. Nanobiosci*., 2007, **6**, 319.
129 P. Eaton, A. Ragusa, C. Clavel, C. T. Rojas, P. Graham, R. V. Durán and S. Penadés, *IEEE Trans. Nanobiosci*., 2007, **6**, 309.
130 F. Manea, C. Bindoli, S. Fallarini, G. Lombardi, L. Polito, L. Lay, R. Bonomi, F. Mancin and P. Scrimin, *Adv. Mater*., 2008, **20**, 4348.
131 B. Nolting, J.-J. Yu, G.-Y. Liu, S.-J. Cho, S. Kauzlarich and J. Gervay-Hague, *Langmuir*, 2003, **19**, 6465.
132 P. M. Enríquez-Navas, M. Marradi, D. Padro, J. Angulo and S. Penadés, *Chem. Eur. J*., 2011, **17**, 1369.
133 M. Marradi, P. Di Gianvincenzo, P. M. Enríquez-Navas, O. M. Martínez-Ávila, F. Chiodo, E. Yuste, J. Angulo and S. Penadés, *J. Mol. Biol*., 2011, **410**, 798.
134 X. Li, H. Zhou, L. Yang, G. Du, A. S. Pai-Panandiker, X. Huang and B. Yan, *Biomaterials*, 2011, **32**, 2540.
135 J. Rojo, V. Díaz, J. M. de la Fuente, I. Segura, A. G. Barrientos, H. H. Riese, A. Bernad and S. Penadés, *ChemBioChem*, 2004, **5**, 291.
136 M. M. Kemp, A. Kumar, S. Mousa, E. Dyskin, M. Yalcin, P. Ajayan, R. J. Linhardt and S. A. Mousa, *Nanotech*., 2009, **20**, 455104.
137 Y.-Y. Chien, M.-D. Jan, A. K. Adak, H.-C. Tzeng, Y.-P. Lin, Y.- J. Chen, K.-T. Wang, C.-T. Chen, C.-C. Chen and C.-C. Lin, *ChemBioChem*, 2008, **9**, 1100.
138 A. A. Kulkarni, C. Fuller, H. Korman, A. A. Weiss and S. S. Iyer, *Bioconjugate Chem*., 2010, **21**, 1486.
139 H. Lee, K. Lee, I. K. Kim and T. G. Park, *Biomaterials*, 2008, **29**, 4709.
140 M. Marradi, D. Alcántara, J. M. de la Fuente, M. L. García-Martín, S. Cerdán and S. Penadés, *Chem. Commun*., 2009, **46**, 3922.
141 K. El-Boubbou, D. C. Zhu, C. Vasileiou, B. Borhan, D. Prosperi, W. Li and X. Huang, *J. Am. Chem. Soc*., 2010, **132**, 4490.
142 K. El-Boubbou, C. Gruden and X. Huang, *J. Am. Chem. Soc*., 2007, **129**, 13392.
143 N. P. Pera, A. Kouki, S. Haataja, H. M. Branderhorst, R. M. J. Liskamp, G. M. Visser, J. Finne and R. J. Pieters, *Org. Biomol. Chem*., 2010, **8**, 2425.
144 J. M. de la Fuente, D. Alcantara, P. Eaton, P. Crespo, T. C. Rojas, A. Fernandez, A. Hernando and S. Penades, *J. Phys. Chem. B*., 2006, **110**, 13021.
145 J. M. de la Fuente, D. Alcántara and S. Penadés, *IEEE Trans. Nanobio*., 2007, **6**, 275.
146 Y. Lee, H. Lee, Y. B. Kim, J. Kim, T. Hyeon, H. Park, P. B Messersmith and T. G. Park, *Adv. Mater*., 2008, **20**, 4154.
147 M. Kamat, K. El-boubbou, D. C. Zhu, T. Lansdell, X. Lu, W. Li and X. Huang, *Bioconjugate Chem*., 2010, **21**, 2128.

148 I. Garciéa, J. Gallo, N. Genicio, D. Padro and S. Penadeés, *Bioconjugate Chem*., 2011, **22**, 264.
149 J. Gallo, I García, N. Genicio, D. Padro and S. Penadés, *Biomaterials*, 2011, **32**, 9818.
150 S. I. van Kasteren, S. J. Campbell, S. Serres, D. C. Anthony, N. R. Sibson and B. G. Davis, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 18.
151 F. Y. Cheng, S. P. H. Wang, C. H. Su, T. L. Tsai, P. C. Wu, D. B. Shieh, J. H. Chen, P. C. H. Hsieh and C. S. Yeh, *Biomaterials*, 2008, **29**, 2104.
152 V. Biju, T. Itoh, A. Anas, A. Sujith and M. Ishikawa, *Anal. Bioanal. Chem*., 2008, **391**, 2469.
153 C. Earhart, N. R. Jana, N. Erathodiyil and J. Y. Ying, *Langmuir*, 2008, **24**, 6215.
154 P. Babu, S. Sinha and A. Surolia, *Bioconjug. Chem*., 2007, **18**, 146.
155 B. Mukhopadhyay, M. B. Martins, R. Karamanska, D. A. Russell and R. A. Field, *Tetrahedron Lett*., 2009, **50**, 886.
156 C. T. Chen, Y. S. Munot, S. B. Salunke, Y. C. Wang, R. K. Lin, C. C. Lin, C. C. Chen and Y. H. Liu, *Adv. Funct. Mater*., 2008, **18**, 527.
157 M. G. Sandros, M. Behrendt, D. Maysinger and M. Tabrizian, *Adv. Funct. Mater*., 2007, **17**, 3724.
158 Y. Higuchi, M. Oka, S. Kawakami and M. Hashida, *J. Control. Release*, 2008, **125**, 131.
159 A. Cambi, D. S. Lidke, D. J. Arndt-Jovin, C. G. Figdor and T. M. Jovin, *Nano Lett*., 2007, **7**, 970.
160 K. Niikura, T. Nishio, H. Akita, Y. Matsuo, R. Kamitani, K. Kogure, H. Harashima and K. Ijiro, *ChemBioChem*., 2007, **8**, 379.
161 K. Niikura, S. Sekiguchi, T. Nishio, T. Masuda, H. Akita, Y. Matsuo, K. Kogure, H. Harashima and K. Ijiro, *ChemBioChem*., 2008, **9**, 2623.
162 E. Han, L. Ding, S. Jin and H. X. Ju, *Biosens. Bioelectron*., 2011, **26**, 2500.
163 M. Ishikawa and V. Biju, *Prog. Mol. Biol. Transl. Sci*., 2011, **104**, 53.
164 R. Kikkeri, B. Lepenies, A. Adibekian, P. Laurino and P. H. Seeberger, *J. Am. Chem. Soc*., 2009, **131**, 2110.

Glycotransporters for gene delivery

Carmen Ortiz Mellet,[a] José M. García Fernández[b] and Juan M. Benito*[b]

DOI: 10.1039/9781849734769-00338

Nucleic acids, from whole genes to short RNA sequences, have emerged as attractive sources of therapeutic and biotechnological tools. Their highly specific and predictable mode of action theoretically allow for either stimulating or silencing the expression of any protein. Due to their poor cellular uptake and rapid degradation in biological media, successful applications critically depend on the development of efficient purpose-conceived carriers that protect and deliver them into their target cells. Because of their natural ability to infect cells, modified viruses have been long time considered as the vehicles of choice. However, given the inherent difficulties and risks associated to virus manipulation, artificial (non-viral) carriers are taking over. Within few years, literally hundreds of non-viral nucleic acid shuttling devices have flourished, based on the most bewildering array of concepts and materials. Carbohydrates have proven particularly valuable components for the bottom-up rational design of many of these artificial nucleic acid carriers. Their structural diversity and biocompatibility, their unique scaffolding features and targeting capabilities, and their highly flexible chemical functionalization have untapped delirious possibilities where chemist creativity is the only limit. This review will survey the recent advances in the design of artificial saccharide-based and saccharide-functionalized gene delivery systems, focussing on the trending topics and strategies implemented to evade the physiological barriers that hamper nucleic acid delivery into cells. Emphasis will be laid on the synergic combination of chemical synthesis and nanotechnology to correlate architectural features of the glycotransporters with nucleic acid delivering capabilities, feeding back newly optimized designs. Though major achievements of glycotransporters in gene delivery are still to come, some of them are already being marketed as investigational carriers or even successfully clinically trialed.

1 Introduction

It is hardly conceivable for the pharmaceutical industry nowadays, investing a vast amount of resources on high-throughput drug discovery technologies, that the strategy to bring up new drugs could rely on the design of a single candidate programmed to develop the required task through a rational set of rules. And that is precisely what made nucleic acids (genes, oligonucleotides, aptamers, ribozymes, DNAzymes, or small interfering RNAs) an attractive source of therapeutic and biotechnological agents. Their intimate structure-activity relationship and highly specific mode of

[a]*Departamento de Química Orgánica, Facultad de Química, Universidad de Sevilla, Profesor García González 1, E-41012 Sevilla, Spain*
[b]*Instituto de Investigaciones Químicas, CSIC - Universidad de Sevilla, Américo Vespucio 49, Isla de la Cartuja, E-41092 Sevilla, Spain. E-mail: juanmab@iiq.csic.es*

© The Royal Society of Chemistry 2012

action theoretically permit, when using appropriate designs, to exploit the cellular machinery in a predictable fashion to either stimulate or silence the expression of virtually any protein, with reduced potential for toxicity and fewer side effects. The far-reaching implications on gene therapy and biotechnology are out of question. The genetic and epigenetic information obtained from biomedical research could lead to novel research tools, diagnostic methods and treatments. However, the nucleic acid poor cellular uptake, due to their large size and anionic charge, and rapid degradation in biological media turns this conceptually simple strategy into a chimera.

While conventional drugs consist of a formulation of a bioactive species and a carrier, accounting the former for most of the sophistication of the design, the success of gene therapy critically depends on the conception of efficient and safe delivery systems enabling the cargo to reach its target. Because of their natural ability to infect cells, viruses were the intuitive choice to deliver genes at the right spot. Several recombinant non-replicating viruses have proved very efficient at promoting genetic material uptake and expression;[1] however, several drawbacks (*e.g.* immunogenicity, toxicity and quality of scaled-up production) have tempered the initial prospects.[2] To date, neither the FDA nor the EMA have yet approved any viral-vector-based gene delivery therapeutic due to these concerns (only a couple of them are approved by the Chinese State Food and Drug Administration).[3,4] Alternatively, non-viral-based carriers have gathered momentum. Synthetic materials can be structured using bottom-up approaches to tailor the desired functional features in terms of bioavailability and biocompatibility, nucleic acid complexation, protecting and packing capabilities, or membrane-crossing and targeting abilities.

As the understanding of the self-assembling processes involved in aggregate formation between nucleic acids and synthetic vectors increases, the range of designed biomaterials for this purpose open wider.[5,6] Most of these non-viral nucleic acid vectors fall within the category of cationic lipids or polymers (Figs. 1A and 1B, respectively), featuring functional groups that electrostatically neutralize nucleic acids and cooperatively promote compaction into colloidal nanoparticles (termed *lipo-* and *polyplexes*, respectively) with increased metabolic stability and membrane permeability. Nanoparticle stability and transfection efficiency largely depend on the structural features of the cationic vector, the type of nucleic acid to be delivered and the relative formulation ratio. Non-viral vectors are, in principle, less prone to elicit the immune system response. However, they may trigger inflammatory response and their delivery efficiency and selectivity, despite few exceptions,[7] are far from that of their viral counterparts. This is not surprising considering that most of the first generation non-viral candidates were not purposely conceived for such task, but just *off-the-shelf* materials, such as poly-L-lysine (PLL) or polyethylenenimine (PEI), for which gene delivery capabilities were incidentally discovered. Their densely-charged nature does not usually furnish them with the optimal biocompatibility and cytotoxicity profiles.[8]

Recent efforts in the field of non-viral vectors have focused on approaches aimed at overcoming the physiological barriers to gene delivery.[9,10] Hypothesis-driven designs have been implemented aimed at regulating

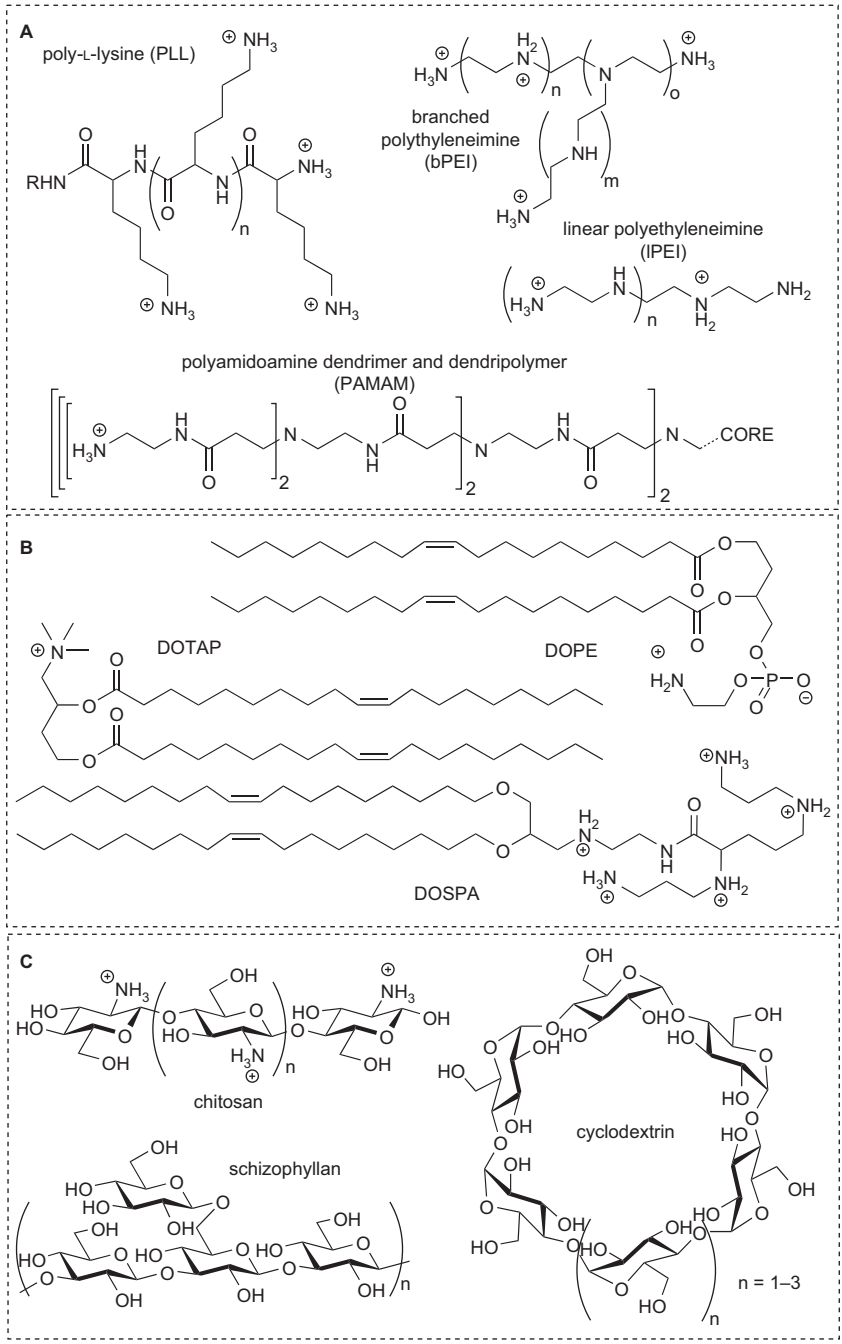


Fig. 1 Structure of some representative (A) cationic polymers, (B) lipids and (C) carbohydrates used for gene delivery. DOTAP: 2,3-dioleoyloxy trimethylammonium propane; DOPE: 2,3-di(oleolyloxy)propyl phosphatidyl ethanolamine; DOSPA: 2,3-dioleoyloxy-*N*-[2-(sperminecarboxamido)ethyl]-*N,N*-dimethyl-1-propaniminium trifluoroacetate.

nucleic acid-carrier condensation,[11] enhancing bioavailability and biocompatibility of their assemblies,[12] shielding aggregates from non-specific interactions and avoiding off-target delivery,[13] promoting selective cellular uptake (*e.g.* receptor-mediated)[14] or inducing stimuli-responsive cargo release.[15] Carbohydrates have proven particularly useful towards these goals. Similarly to polyethyleneglycol (PEG) coating (*pegylation*),[16] glycocoating can sterically shield the positively charged surface of colloidal aggregates by improving solvation, thus preventing from non-specific interactions with intra- and extracellular components. Glycoconjugates can also be specifically recognized by cell membrane receptors and intracellular trafficking machineries, whose expression depends not only in cell type but also in cell developmental state.[17,18] Nature already offers a number of examples that illustrate the potential of carbohydrate polymers as non-viral nucleic acid vectors, chitosan (Fig. 1C) being probably the most exhaustively investigated.[19]

Non-viral gene delivery has benefited from the large structural and functional diversity of carbohydrates, offering broad opportunities to interfere and manipulate gene transfer capabilities of first generation non-viral vectors. Moreover, the inherent flexibility of carbohydrate and glycoconjugate chemistry has enlarged the range of functional materials that can be rationally conceived. Systematic studies on structure-function relationships of such synthetic materials have shed light on the impact that carrier composition, hydrophobic/hydrophilic balance or charge and functional group distribution exert on nucleic acid transfer ability. This review will critically highlight the trending topics on the exploitation of the structural and functional features of carbohydrates in the bottom-up design of non-viral gene delivery systems, focussing on three separate aspects: i) the effect of glycosylation in the performance of first generation non-viral gene vectors, ii) the *de novo* design of carbohydrate-based cationic polymers as a tool to create functional materials with tailored nucleic acid shuttling capabilities, and iii) the design of molecularly well-defined carbohydrate-based conjugates using pre-organized molecular scaffolds as a mean to intimately correlate molecular structure with gene transfer capabilities. Due to space limitations, those approaches based on the use of naturally occurring carbohydrate polymers will only be succinctly commented.[20]

2 Carbohydrate-grafted cationic polymers and lipids

Despite their undisputable investigational utility, manipulation of the functional features of many of the first generation non-viral vectors is not an easy task. Cationic polymers, while structurally diverse (*e.g.* PEI, PAMAM, PLL, or chitosan; see Fig. 1), are usually very restrained in terms of precise structural modification. It is known that the polymer molecular weight and their polydispersity or branching degree significantly influences their physicochemical properties, delivery efficiency, and toxicity profiles.[21] Improved polymerization technologies have rendered low polydisperse materials,[22] but finely tuning these parameters is very rarely sufficient to achieve clinically useful performances.[23] The situation with cationic lipids does not significantly differ.

Alternatively, the gene transfer capabilities of cationic polymer and lipid carriers may be manipulated by decoration, either *via* covalent or supramolecular linkages, with additional functional elements. For instance, pegylation has successfully been exploited to coat transfectious nanoparticles with a hydrophilic shield that enhances stability and biocompatibility by preventing non-specific interactions with biological components.[24] Grafting functional elements driving carrier-nucleic acid condensation,[25] receptor-mediated nanoparticle cell uptake,[26] endosome release[27] or cellular[14,28] and nuclear[29] targeting and trafficking has also been investigated. Vitamins,[30] peptides,[31] proteins[32,33] or antibodies[34] have been used for these purposes, but carbohydrates probably offer the wider structural diversity and functional versatility in this regard. This is illustrated by the large array of glycosylated non-viral carriers reported to date and the diversity of synthetic strategies to prepare them. From the chronological point of view, nucleic acid carrier glycosylation was originally conceived to elicit targeted delivery towards cell membrane specific receptors.[35] However, a number of other potential roles of carbohydrates (*e.g.* nucleic acid complexation modulators, nanoparticle biocompatibilizers or active intracellular transport enhancers) have been also investigated to improve the overall performance of artificial gene vectors.

A Glycomodulating nucleic acid condensation

Upon designing novel synthetic nucleic acid carriers, condensation and packing is probably the aspect that receives the lesser attention, despite the robust evidences of its relevance.[36] Electrostatic (non-hierarchical) interaction usually leads to suboptimal nucleic acid-carrier packing and considerable nanoparticle dispersity. It has been shown that cationic polymer glycosylation may serve to modulate this process. Aigner and co-workers[37] have recently observed that covalent (oligo)maltose grafting of PEI alter pDNA and siRNA condensation thermodynamics and nanoparticle stability, consequently affecting a number of other features (*vide infra*, heading 2B). Strand and co-workers have found that the kinetics of DNA complexation and release can be modulated by re-engineering the glycosyl residues decorating chitosan backbone.[38] Sakurai and co-workers succeeded at accelerating the complex formation kinetics of cholesterol-grafted schizophillan by supramolecularly decorating the cationic polymer with cyclodextrins (Fig. 2).[39] Interestingly, glycomodulation of nucleic acid-carrier interaction is not limited to cationic polymers. Pitard and Lehn have designed a series of cationic lipids inspired in the structure of RNA-binding aminoglycosides (Fig. 3), their lipoplexes exhibiting enhanced colloidal stability as compared with conventional cationic lipids. Remarkably, lipoplex particle features and performance could be related to lipid aminoglycoside structure and siRNA binding capabilities, finely illustrating the potential of rational design in nucleic acid delivery.[40]

B Glycomodulating nanoparticle biocompatibility and bioavailability

Structural modification of a gene carrier does not only affect a particular property, but simultaneously several features. Ascertain the precise effect of each structural change is, thus, complicated. For instance, oligomaltosylation

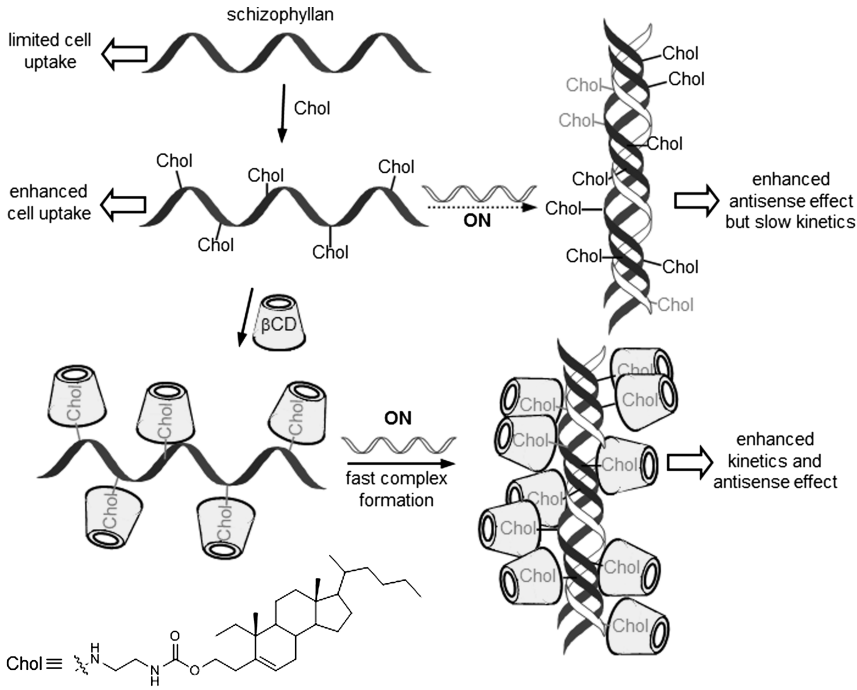


Fig. 2 Schematic representation of βCD-enhanced cholesterol-appended schizophyllan – antisense oligonucleotide (ON) complex formation.[39]

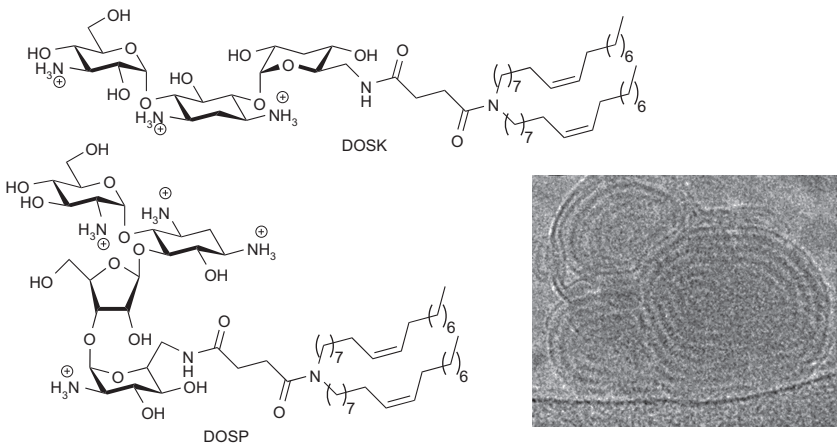


Fig. 3 Structure of some lipidic aminoglycoside derivatives (dioleyl succinyl kanamycin A, DOSK, and dioleyl succinyl paromomycin, DOSP) and a representative Cryo-TEM micrograph of their siRNA complexes (Reproduced from ref. 40a).

of PEI did not only affect nucleic acid binding dynamics, but also physicochemical features of the nanocondensates in a carbohydrate density-dependent fashion.[37] (Oligo)maltosylation render complexes that become largely unaffected by the presence of serum proteins, far less toxic, and exploit different cell uptake mechanisms as compared to non-maltosylated PEI.

The potential of carbohydrates to modulate the suboptimal biocompatibility of PEI has been also demonstrated by Yu and Wang with a strategically distinct approach. By crosslinking short PEI chains (PEI_{600}) using different cyclodextrins (CDs) they built a series of cationic polymers exhibiting better DNA-transfecting capabilities than large PEI polymers (*e.g.* branched PEI, 25 kDa), while maintaining the low cytotoxic profile of the shorter polymers (Fig. 4).[41] Tang and co-workers have implemented a number of modifications on this design to achieve targeted delivery of a gene encoding the epidermal growth factor to folate receptors and enhanced uptake assisted by cell penetrating peptides (*e.g.* octaarginine). Preliminary experiments *in vivo* demonstrate enhanced antitumor effects on mice as compared to bPEI (25 kDa) and non-targeted CD-PEI_{600} polymers.[42] Li and co-workers have exploited the molecular inclusion features of CDs in structurally related polymers to supramolecularly graft PEG chains for improving their stability under physiological conditions. Such stealth nanoparticles were less efficient at delivering pDNA but, remarkably, exhibited higher siRNA-mediated silencing efficiency than commercial formulations.[43]

Arima and Uekama have conducted parallel studies on PAMAM dendrimers, which gene delivery efficiency is circumscribed to dendritic generations for which toxicity is a concern (> G3).[44] They found that covalent grafting of CDs on low generation PAMAM (*e.g.* G2) resulted in conjugates maintaining a low toxic profile while boosting transfection capabilities to rival that of higher generation PAMAM standards,[45] a result explained on the basis of the ability of CDs to interact with membrane components.

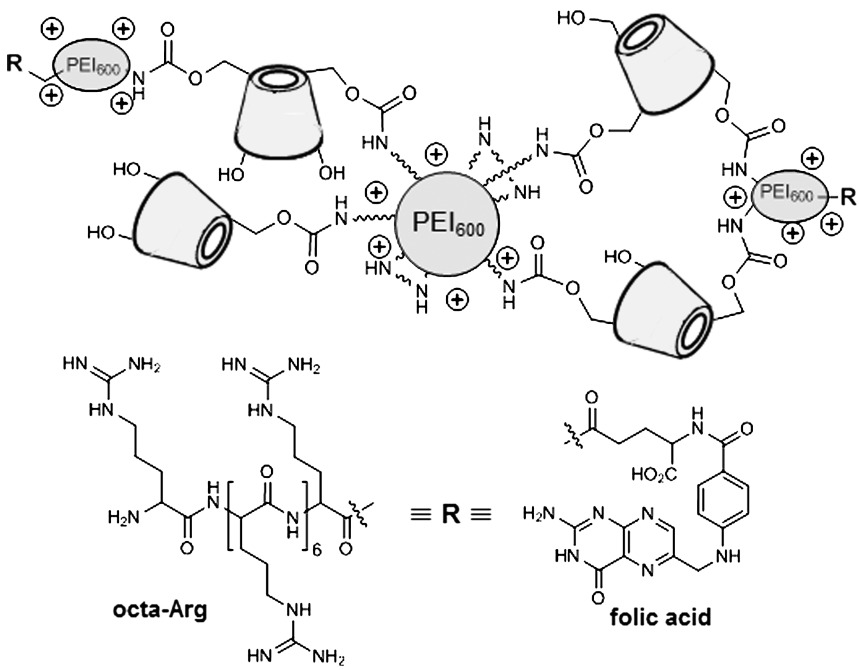


Fig. 4 Schematic structure of the folic acid or octaarginine-grafted βCD-PEI_{600} cationic polymers reported by Tang and Xu.[42b]

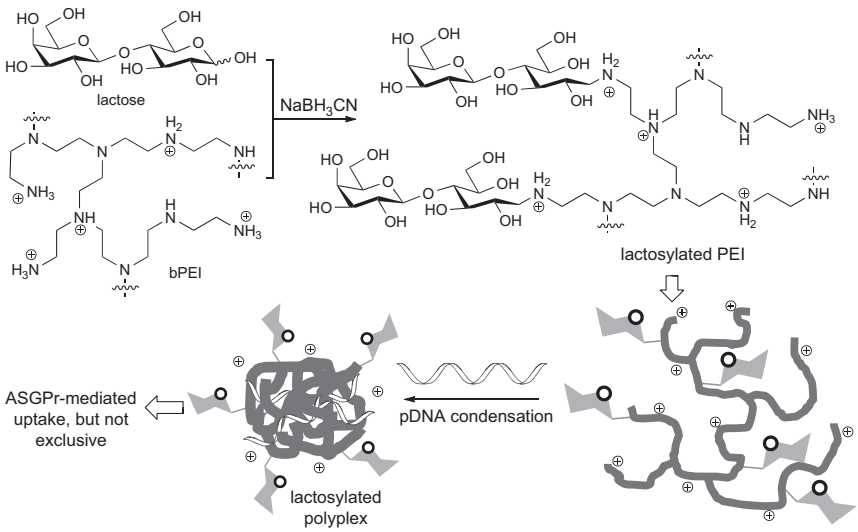
Based in this concept, the authors developed high efficiency carriers for DNA and RNA delivery.[46] Further glycoengineering led Arima and co-workers to decorate their PAMAM-CD conjugates with saccharidic antennae. While the initial goal was selective targeting (*vide infra*, heading 2C), they also found that certain degrees of glycosylation resulted in serum-resistant polyplexes.[47] A similar suppression of serum-induced nanoparticle aggregation was noted by Sato and co-workers upon glycodecoration of chitosan with mannosyl and lactosyl residues.[48]

A number of other polymeric scaffolds have been glycocoated in order to modulate their biocompatibility and bioavailability with gene delivery purposes. Carbon nanotubes,[49] gold nanoparticles,[50] quantum dots,[51] or (pseudo)rotaxanes[52] are some relevant examples that further provide peculiar functionalities to the gene carrier. In addition, glycomodulation of the pharmacodynamics (*e.g.* gene entrapment and controlled release) of self-assembled gene-embedding multilayers has been shown a promising strategy for tissue engineering.[53]

C Glycomodulating nanoparticle receptor-mediated cell uptake and intracellular trafficking

While non-viral carrier glycosylation dramatically influence a number of carrier capabilities, it was certainly carbohydrate specific recognition by biological receptors that first called the attention. In fact, the first ever reported evidence of specific non-viral gene delivery used a galactose-terminated PLL carrier targeted to the asialoglycoprotein receptors (ASGPr) on hepatocyte cell membrane.[35] Since then, the topic has received a sustained interest. Galactose (or lactose) and mannose have been mostly exploited as glycoligands due to their hepatic and immune system cell targeting abilities. But also glucosamine, targeting vimentin,[54] and galactosamine[55] have been recently used. Glycotargeted versions of the most usual cationic polymers (including PEI,[56] PLL,[57] or PAMAM)[58] and lipids[59] have been described. The increase in gene expression when using glycoligands is associated to the improved receptor-mediated uptake of the carriers, but only to a certain extent. As above outlined, many other features of nucleic acid-carrier interaction and nanoparticle pharmacodynamics can be altered upon glycosylation, which makes it difficult to quantify the precise role of the saccharide antennae. For instance, Behr and co-workers demonstrated that although the ASGPr is involved in the internalization of galactosylated PEI by hepatocytes, this is not sufficient to achieve specificity since other uptake mechanisms simultaneously operate (Scheme 1). Thus, relevant transfection levels were measured upon addition of ASGPr competing agents as well as in ASGPr-devoid cells.[56a] In other cases it was observed that even though the glycosylated carriers were recognized by the target receptors, endocytosis did not take place or, if it did, the resulting endosomes were unproductive.[47a]

Glycoligands may also exert a function in controlling the fate of the internalized nucleic acid-carrier nanocomplexes in a way that remain to be unveiled. Monsigny and co-workers showed that cytosolic and nuclear lectins might be involved in intracellular trafficking and nuclear import[60] and latter coined the term *glycofection* to refer to transfection strategies



Scheme 1 Synthesis of lactosylated PEI *via* reductive amination and schematic representation of the condensation of lactosylated PEI with plasmid DNA into lactosylated polyplexes.[56a]

based on the glycomodulation of intracellular barriers for gene delivery.[61] In a number of cases, lactosylated carriers appeared to induce faster escape from the late endosome/lysosome than other glycosylated counterparts and more efficiently mediated nuclear import.[62] Though these observations are still a matter of debate, recent publications by the same group[63] and others[64] highlight the potential benefits of controlling of intracellular trafficking as a prerequisite for efficient non-viral gene delivery.

3 *De novo* designed carbohydrate-based polymers

Efficiency and toxicity of polymeric vectors can be largely (and unexpectedly) affected by subtle structural changes, for example in counterion nature,[65] branching degree[66] or size,[67] which may lead even by batch-to-batch disparities. This unwanted variability has been overcome to a large extent by the implementation of novel polymerization techniques, such as controlled polycondensation, radical polymerization and click chemistry reactions, which offer superior control over functional features.[68] Through the judicious choice of monomer structure and polymerization strategies a number of relevant parameters regarding gene delivery capabilities can be tailored. In this context, the *de novo* design of polymeric gene carriers from carbohydrate-derived building blocks has been very appealing, given the structural diversity and availability of simple saccharides. This section surveys a number of examples that illustrate the nucleic acid delivery capabilities of synthetic carbohydrate-based polymers.

A Polyglycoamidoamine-based polymers

One of the most representative examples of the potential of synthetic glycopolymers at tailoring non-viral gene vectors is the case of

polyglycoamidoamines (PGAAs). Polyamidoamines (PAAs) are among the most accessible entries to structurally controlled polymers with biomedical applications, their properties being intimately related to structural and functional features.[69] The first insight into the gene delivery capabilities of carbohydrate-based PAAs (polyglycoamidoamines, PGAAs) were almost simultaneously reported by Reineke[70] and Guan[71] (Fig. 5). In both cases the diversity, availability and biocompatibility of carbohydrates and their readily conversion to polymerizable building blocks was the main motivation. Reineke and co-workers focused on imitating the structure of linear PEI polymers by co-polymerizing their carbohydrate building blocks with a set of short oligoethyleneimine (OEI) chains. Guan's design aimed at mimicking the structure of poly-L-lysine (PLL) by condensing carbohydrate bricks with lysine segments. Both studies revealed significantly reduced cytotoxicity of the resulting PGAAs as compared to PEI and PLL, respectively, while maintained comparable pDNA transfection efficiencies.

Guan and co-workers later oriented their strategy to the elaboration of protein-resistant biodegradable materials.[72] The Reineke's group on its side has thoroughly analysed the influence of architectural issues of their PEI-inspired PGAAs on nucleic acid complexation, nanoparticle topology and stability, cell uptake, toxicity and gene expression efficiency.[73] By screening a library of cationic PGAAs, (Fig. 6) synthesized from a range of linear OEIs (different length and number of amino groups) and esterified aldaric acid and lactone co-monomers (different stereochemistry and number of hydroxyls), they identified vector candidates that mediated high gene expression levels in several cell lines (HeLa, H9c2(2-1), HepG2, or

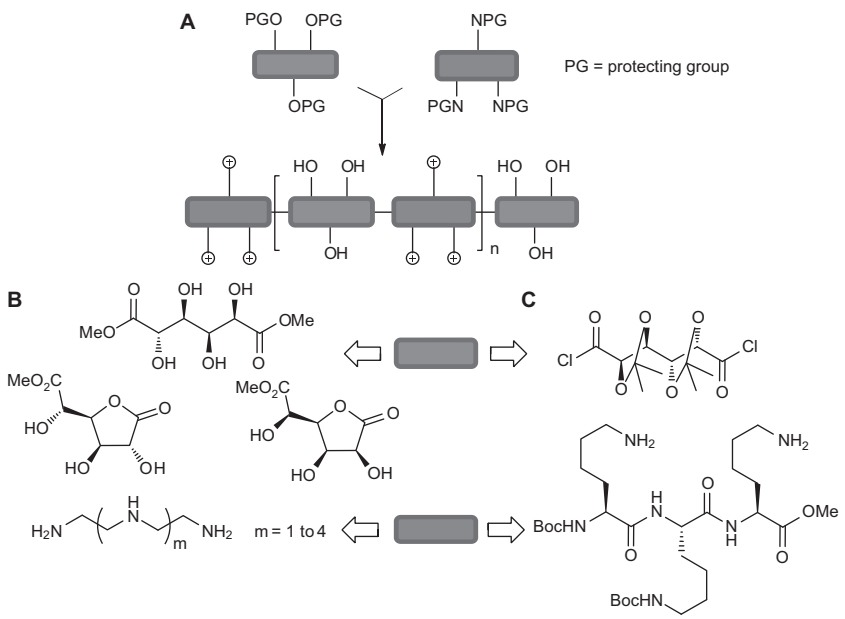


Fig. 5 Schematic representation of the synthesis of polyglycoamidoamines (A) and building blocks used by Reineke's[70] (B) and Guan's[71] (C) group.

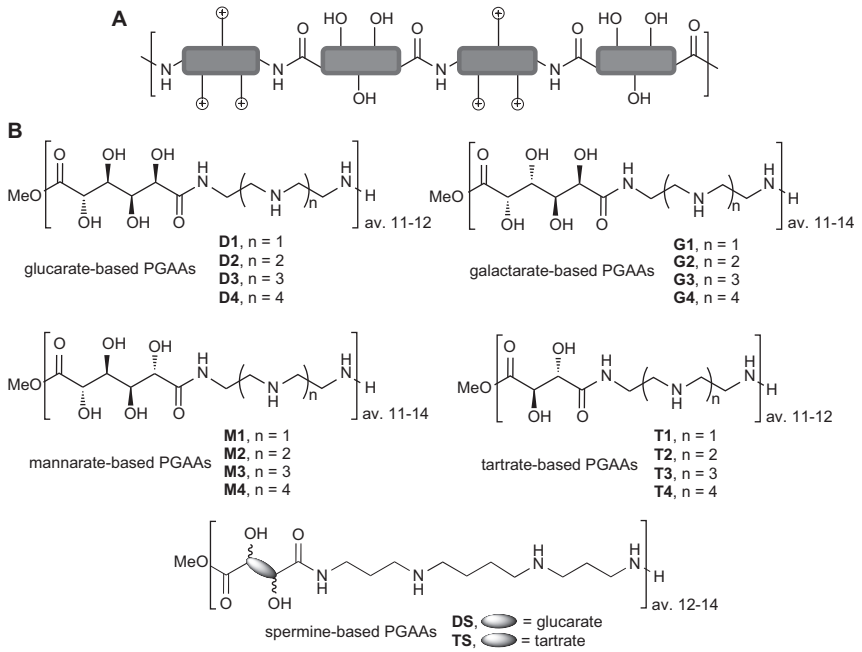


Fig. 6 General structure of PGAAs studied by Reineke and co-workers (A) with indication of their structural diversity (B).

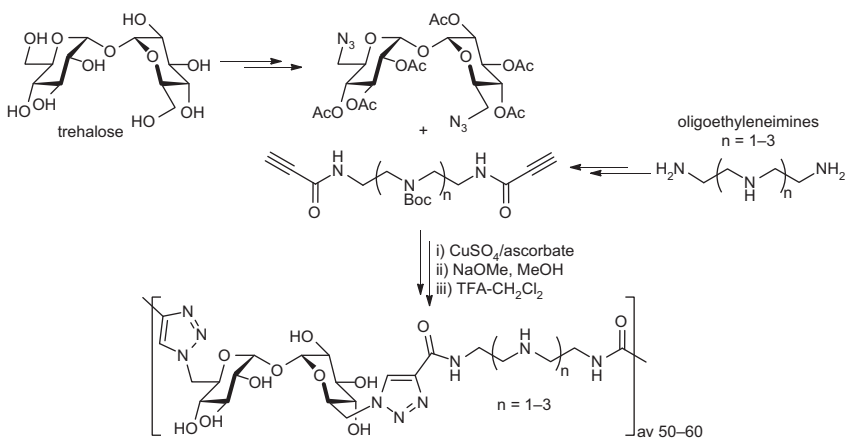
BHK-21). pDNA transfer efficiency peaked for the polymers containing the longer OIE chains tested (four secondary amino groups).[74] Transfection efficiency favourably compared to that achieved with PEI, but was accompanied by a far less toxic profile. They hypothesized that the iterative interruption of OEI charge density by carbohydrate blocks was at the origin of the reduced toxicity as compared to PEI, though their rapid hydrolysis in biological media might also account for this merit.[75]

PGAAs with longer OIE segments did not significantly influenced cell uptake and overall transgene expression, but significantly improved polyplex stability in serum-containing media. Unfortunately, increased toxicity was also observed.[76] Polymers with longer inter-amine spacers (*e.g.* spermine) also exhibited increased toxicity, highlighting a critical relationship between charge density distribution and biocompatibility.[77] The nature of the saccharide segment, though modestly, also influenced polymer performance, apparently by altering the buffering capabilities of the system. Since the response of polyplexes at the increase of pH in the endosomes closely relates to the pDNA release dynamics, the results might indicate certain folding differences upon pDNA:PGAA complex formation. This study let identify PGAA G4, combining pentaethylenehexaamine and galactarate segments (Fig. 6), as the best performing vector candidate, promoting high transgene expression levels in a range of primary cell lines (similar to PEI) without causing relevant toxicity. This novel formulation has been recently licenced to Techulon and is now commercialized as Glycofect™.[78]

B Amidine- and triazole-linked carbohydrate-based polymers

As an alternative to PGAAs, carbohydrate-based cationic polymers with amidine[79] and, more recently, triazol linking tethers[80] have also been explored. The amidine functionality is protonated at physiological pH, which might be favourable regarding nucleic acid complexation and stabilization. However polyamidines usually exhibit severe cytotoxicity. Reineke and Davis observed that this drawback could be overcome by intercalating saccharidic moieties between the cationic elements while maintaining the gene transfer capabilities.[79] The closer the amidine and the carbohydrate units are located in the polymer chain, the more efficiently the cytotoxicity is reduced. Such favourable effect was first observed for trehalose-amidine copolymers and further confirmed for β-cyclodextrin as the sugar building block (*vide infra*, heading 3C).

Trehalose itself has been found to impart some favourable properties to polymeric gene vectors, such as shielding aggregation of the resulting polyplexes with proteins and lipids.[81] To fully benefit from this, Reineke and co-workers have devised a strategy to build a new family of cationic trehalose-embedding polymers that exploits the Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC) reaction.[82] By coupling diazido-functionalized trehalose units with a diverse set of α,ω-dipropargylated OEIs (Scheme 2), a series of polymers were obtained that allowed the systematic investigation of architectural effects (OEI size and polymer length) on pDNA complexation, nanoparticle properties, cellular uptake and transgene expression. Transfection efficiencies that paralleled that of jetPEI-based polyplexes were achieved using trehalose "click" polymers with 3-4 ethylenediamine segments.[80a] Dynamic light scattering revealed that polyplexes formulated with larger polymers exhibited the best tolerance to serum-containing media, but displayed higher toxicity.[80b] This unfavourable feature was alleviated by the fact that "click" polymers required far lower N/P ratios, as compared to PGAAs, to achieve similar gene expression levels.



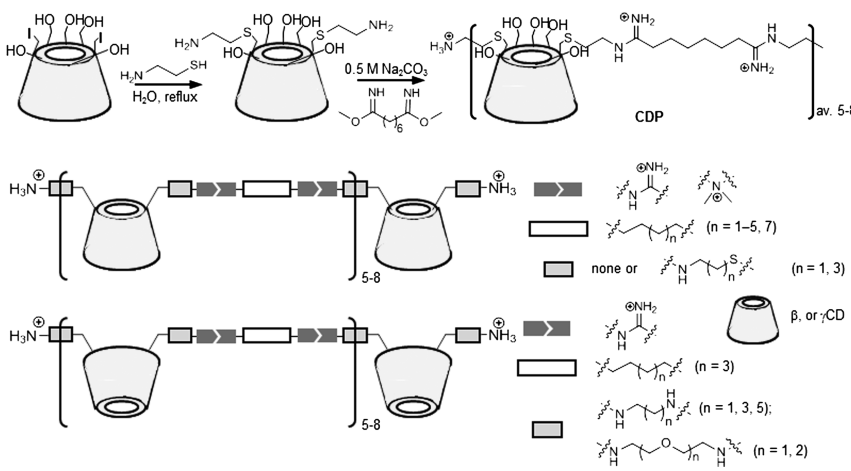
Scheme 2 Synthesis of trehalose-based cationic "click" polymers.[80]

C Cyclodextrin-based polymers

The application of CD-containing polymers to gene delivery was pioneered by Davis and co-workers. At the light of the increasing interest on the biomedical application of CD polymers,[83] they conceived a new class of cationic polymers specifically designed to deliver macromolecular therapeutics.[84] Their synthetic strategy was based on the controlled polycondensation of bifunctional CD monomers and cationic co-monomers, in order to yield linear chains with alternating CD and cationic units (CDP in Scheme 3).[85] Electrostatic-driven complexation of the resulting cationic CD polymers (CDPs) and negatively charged pDNA (~5 kpb) rendered nanometric polyplexes (polyCDplexes; 100–150 nm) featuring *in vitro* cell transfection efficiency comparable to that obtained with polyethyleneimine (PEI) and Lipofectamine™ while preserving a reduced toxicity.

This milestone contribution was followed by a series of reports in which the effect of structural modifications of CDPs on the gene delivery capabilities were investigated.[86–91] The influence of factors such as the CD size (β- or γCD, in comparison with linear saccharides)[87,89] and the distribution and nature of cationic elements[88,90] (their linkages, distances,[86] and relative dispositions),[87] as well as the polymer size and polydispersity,[89] have been examined. Structure-activity relationship (SAR) studies concluded that low molecular weight polymers (*ca.* 10 kDa, degree of polymerization, DP, 5-8) with the CD units sufficiently spaced from amidine cationic centres were the optimal architectures in terms of both high delivery efficiency and low cytotoxicity (Scheme 3).[91]

More recently, the same authors have reported a significant improvement in delivery efficiency by modifying the polymer endings with imidazole groups.[92] Though the precise mechanism leading to this efficiency enhancement is unclear, the authors argued that imidazole moieties should impart buffering capacity to CDP-gene polyCDplexes that would prove

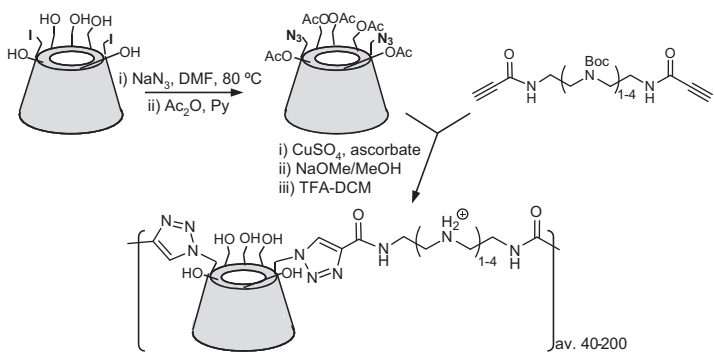


Scheme 3 Synthesis of βCD-based cationic polymers (top) with indication of their structural diversity (bottom).

instrumental for efficient endosomal release by virtue of the ''proton sponge'' effect.[93]

Very recently Srinivasachari and Reineke have investigated a versatile approach towards linear CD-containing cationic polymers for pDNA delivery by using Cu(I)-catalysed azide-alkyne 1,3-dipolar cycloaddition of a diazido βCD derivative and α,ω-dipropargylated oligoethyleneimines (Scheme 4). The gene expression profiles of the resulting ''click'' polymers in HeLa (human epithelial cervical carcinoma) cells were mostly dependent on the oligoethylenemine/CD ratio, with the polymers having longer oligoethylenemine segments being the better performers.[94]

The nanoparticles obtained from CDPs upon oligonucleotide complexation can be supramolecularly modified at their surface. For instance, the inclusion of the adamantane moiety of adamantane-polyethyleneglycol conjugates (Ad-PEG) into CD cavities of CDP-pDNA complexes creates a steric shield around the particles that prevents aggregation and non-specific interactions with biological components.[95] Furthermore, polyCDplexes coated with galactosylated Ad-PEG were shown to exhibit selectivity towards hepatocytes.[96] PolyCDplexes covered with transferrin-modified Ad-PEG (Ad-PEG-Tf)[97,98] transfected luciferase-encoding gene to K562 human myelogeneous leukemia cells with better efficiency than non-targeted particles (Fig. 7).[97] These Tf-targeted gene vectors have shown *in vivo* transfection efficiency and selectivity towards different tumor models (murine[99–101] and primates).[102] In June 2008, a targeted therapeutic based in this concept, CALAA-01 (Calando Pharmaceuticals) entered phase I clinical trials. The active ingredient of CALAA-01 is a small interfering RNA (siRNA). This siRNA inhibits tumour growth *via* RNA interference to reduce expression of the M2 subunit of ribonucleotide reductase (R2). The CALAA-01 siRNA is protected from nuclease degradation within a stabilized polyCDplex termed RONDEL™ (RNAi-oligonucleotide nanoparticle delivery). RONDEL is a three part delivery system: the cationic linear CDP, the AD-PEG surface modification element to increase stability and serum half-life and the Ad-PEG-Tf targeting ligands that assure targeting of the nanoparticles to the tissues of interest. The study is directed to adults with solid tumours refractary to standard-of-cure therapies and is currently



Scheme 4 Synthesis of βCD-based cationic ''click'' polymers.[94]

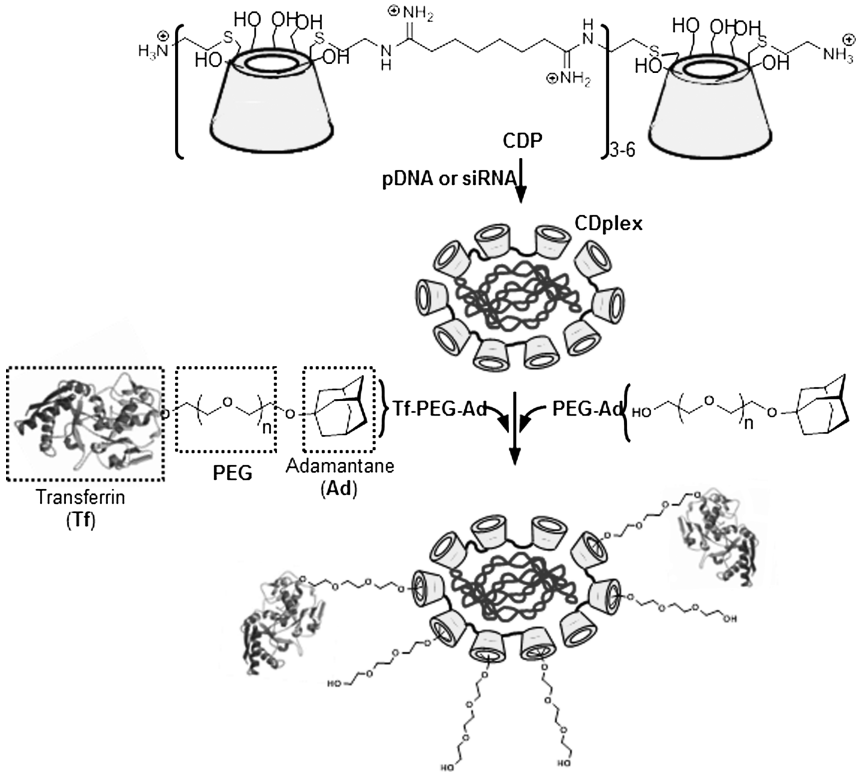
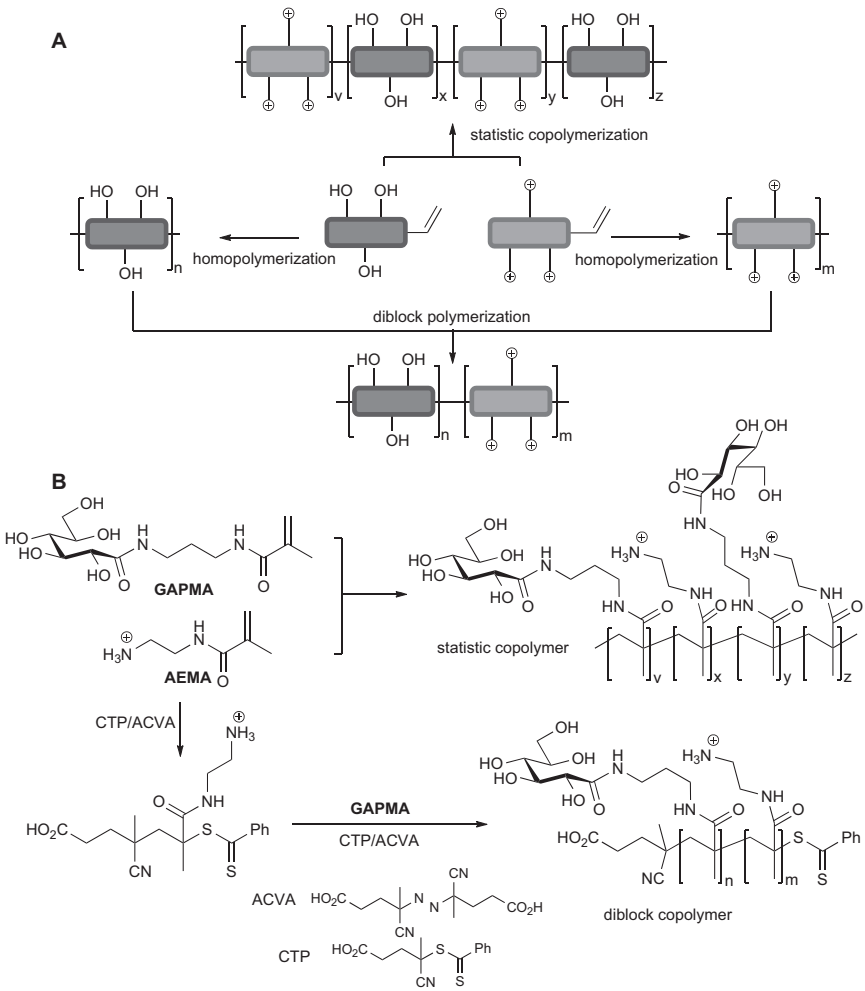


Fig. 7 Schematic representation of the elaboration of the transferrin-targeted nucleic acid-CDP nanoparticles (RONDEL™) developed by Davis and co-workers.[77]

recruiting participants.[103] Alternative targeting ligands (*e.g.* cancer cell markers and antibodies) are being investigated.[7,104,105]

D Diblock glycopolymers

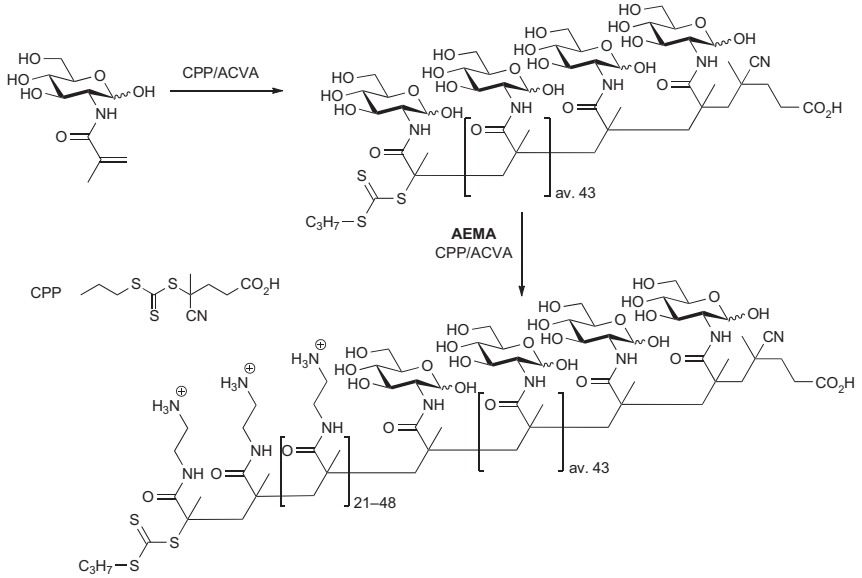
Structural control of polymeric species has experienced a dramatic improvement with the advent of modern polymerization technologies. For instance, living radical polymerization enables the synthesis of polymers, co-polymers or multi-block polymers with highly defined architectures and broad functional group tolerance. Such polymerization techniques have already shown promising perspectives in the elaboration of functional gly-copolymers.[106] Ahmed and Narain have recently used radical addition-fragmentation chain transfer (RAFT) polymerization to synthesize a library of carbohydrate-pendant cationic polymers, where carbohydrate and cationic units can be either statistically distributed (random polymers) or segregated into separate blocks (diblock polymers) (Scheme 5). Precise control on chain length, monomer ratio and polydispersity allowed the authors to rationalize the large influence that polymer architecture exerts on gene delivery capabilities. Statistical copolymers with high molecular weight produced superior gene expression with lower toxicity as compared to the corresponding diblock copolymers, both in the presence or absence of serum.[107] On the other hand, diblock copolymers resulted in nanoparticles



Scheme 5 Schematic synthesis of statistical vs diblock copolymers (A) illustrated with the cationic glycopolymers synthesized by Ahmed and Narain (B).[107]

that were less prone to unspecifically interact with serum proteins and cell membrane components. Such shielding features are comparable to those imparted by conventional PEG grafting on cationic polymers, but probably more easily modulated.

RAFT polymerization has been also exploited by Reineke and co-workers for the synthesis of a series of diblock cationic glycopolymers having the sugar units in the chain (Scheme 6).[108] Interestingly, regardless of the cation-to-carbohydrate balance, diblock polymer-pDNA colloids were found to be stable in the presence of salt and serum over several hours, whereas other formulations (PEI or Glycofect™) tend to aggregate. Cell uptake of the complexes was found to be as efficient as that observed for PEI or Glycofect™. Although pDNA transgene expression was only marginal, siRNA-induced gene knockdown efficacy was comparable to that achieved with Lipofectamine™ 2000.



Scheme 6 Synthesis of diblock cationic glycopolymers by RAFT polymerization described by Reineke and co-workers.[108]

The current examples of diblock glycopolymers cannot compete in terms of gene transfer efficiencies with conventional natural or synthetic glycopolymers. Nevertheless, the potential of controlled polymer manipulation techniques open new opportunities with promising perspectives in the field of non-viral gene delivery. Targeted delivery for instance, where non-specific interactions have to be shielded at the same time that specific uptake is promoted, might be one of those channels.

4 Preorganized glycomaterials for gene delivery

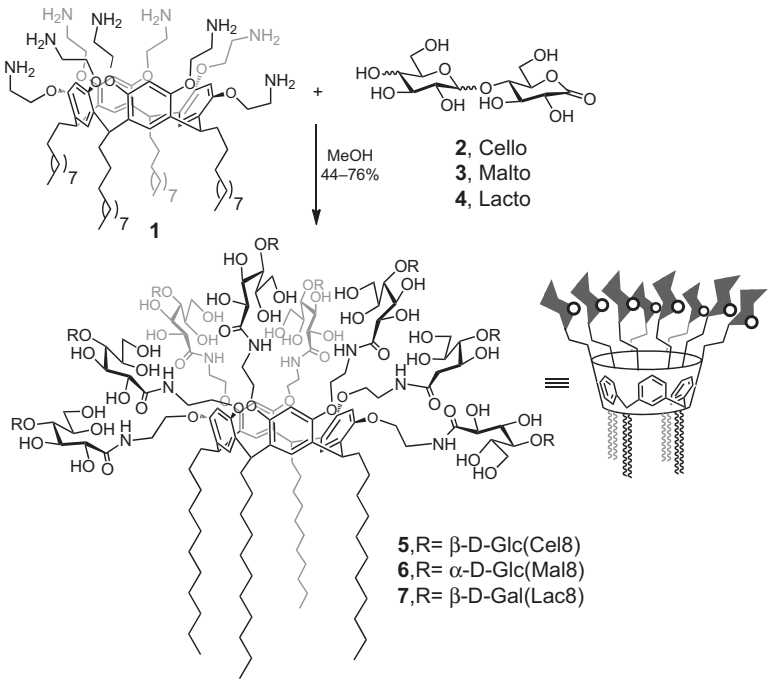
A Non-carbohydrate macrocycle-scaffolded glycoclusters

Appending or inserting carbohydrate moieties into linear and branched polymeric or oligomeric amines has proven a very versatile strategy to access gene delivery systems with improved biocompatibility, furtivity, membrane crossing, and/or targeting abilities.[7,73] The intrinsic polydispersity of these glycomaterials and their random conformational properties make it difficult to undertake a systematic investigation of the influence of structural modifications on the transfecting properties, however. Moreover, their generally flexible character may give rise to self-folding of the polyamine chains,[109] which decreases the binding ability towards DNA and forces the use of higher N/P ratios to achieve full complexation and protection. Preorganization of the cationic functional elements onto macrocyclic platforms has the potential to allow controlling their spatial orientation and, ultimately, the self-assembling behaviour of discrete architectures to produce nanometric objects that can be programmed to complex, compact, deliver and release plasmid DNA in a target cell. Most interestingly, homogeneity can be preserved at the molecular level

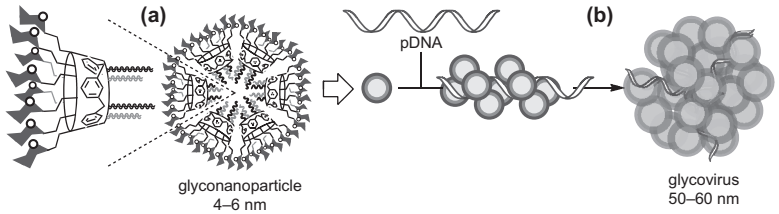
in structurally related series of compounds by implementing selective chemical functionalization methodologies, offering unprecedented opportunities for structure-activity relationship studies.[110]

Elaboration of macrocyclic-based glycomaterials with gene delivery capabilities can be conceived through two general tactics, namely glycodecoration of non-carbohydrate platforms or incorporation of functional elements for oligonucleotide complexation onto cyclooligosaccharide backbones. Several non-carbohydrate macrocyclic scaffolds have been exploited to implement the concept of preorganized gene delivery systems, including macrocyclic polyamines,[111] calixarenes[112] and calix[4]resorcinarenes (resorcarenes).[113] All of them have been largely used for the construction of glycoconjugates that proved extremely useful to assess the influence of architectural parameters in carbohydrate-protein recognition processes,[114] but only in the last case the formation of transfectious glycocomplexes with DNA after installation of the carbohydrate appendages has been demonstrated and thoroughly investigated.

The use of glycoresorcarenes in gene delivery was pioneered by Aoyama, in Japan, in a work originally intended at developing unimolecular mimics of the cell-membrane sphingoglycolipid clusters.[115] For such purpose, the bowl-shaped calix[4]resorcarene framework was chosen.[116] Coupling of octaamine **1** with the disaccharide (cellobiose, maltose and lactose) lactone derivatives **2–4** afforded octavalent glycolipid bundles **5–7** bearing four undecyl tails, with the hydrophilic sugar moieties and the hydrophobic chains located onto opposite rims of the macrocyle in a well-defined geometry (Scheme 7).



Scheme 7 Synthesis of glycoresorcarenes.



Scheme 8 Schematic representation of the process leading from glycoresorcarenes to glycomicelles and to glycoviruses.

Given their amphiphilic character, glycoclusters **5–7** were expected to form micelle-like aggregates in water. Notwithstanding, they were found to be not surface-active. A combined GPC and TEM study revealed that they indeed form micelle-like spherical aggregates, termed *glycocluster nanoparticles* (GNPs), with a diameter of 4-5 nm and an aggregation number of $\sim$6 (Scheme 8, step a). In marked contrast to usual cases, micellization is practically irreversible for those systems. Strong intermolecular hydrophobic entangling of the alkyl groups, possibly coupled with lateral or intercluster hydrogen bonding in the glycocluster domain, may be at the origin of the unusual stability.[117] Interestingly, the GNPs were serendipitously found to undergo agglutination with phosphate ions as a glue, which involves the establishment of polar sugar-to-phosphate interactions, possibly reinforced by inter GNP sugar-sugar interactions. This unexpected phenomenon led Aoyama's group to take up DNA, a natural phosphate polymer, as an interaction partner of GNPs.

The three resorcarene-based glycoclusters **5–7** were able to complex plasmid DNA to afford transfectious nanometric complexes called *glycoviruses* by the authors (Scheme 8, step b). The results evidenced that at a host/phosphate (or base) ratio of $\sim$0.6 saturation is reached and the size of the resulting glycovirus remains nearly constant thereof ($\sim$54 nm). The surface (ζ) potential of the complex, which is negative in the presaturation region, becomes then neutral. The 0.6:1 (*i.e.*, 12:20) glycocluster:nucleobase ratio indicates that two particles of hexameric GNPs accommodate in every helical pitch of the DNA, composed of 10 bp. The glycovirus particles were capable of delivery a luciferase-encoding reporter gene into cultured cells, which is remarkable considering the neutral character of the vector. In general, a strong dependence of the transfecting efficiency on the nature of the carbohydrate moieties was observed, however, which inversely correlated with the tendency of the corresponding glycoviruses to self-aggregate (Mal > Lac >> Cel).

The hierarchical process leading from glycoclusters to glyconanoparticles to glycoviruses warrants that at any molecular or aggregated stage there will be sugar residues exposed to the external medium, therefore potentially available to act as ligands of complementary receptors. Actually, the formal valency must increase with the size of the supramolecular edifice, favouring recognition by lectins. Thus, in the case of the hepatic cell line HepG2, Lac glycoviruses exhibited a much higher activity than expected due to the occurrence of specific interactions between the galactose residues and the

asialoglycoprotein receptors present at the cell surface, which could be exploited for targeted gene delivery. Optimization of the cluster structure was achieved by decreasing the number of lactose residues onto the octaamine scaffold and keeping cationic amino groups at the unsubstituted positions, thereby minimizing the detrimental impact of self-aggregation. Heterogeneous compounds with an overall $(Lac)_3(NH_3^+)_5$ resorcarene structure exhibited the best properties in term of efficiency of gene delivery to hepatocytes and cell viability.

B Cyclodextrin-scaffolded gene delivery systems

Modulating the molecular topology and the finite macromolecular association of preorganized systems by acting not only on the head and tail groups nature and density, but also on the size and shape of the macrocyclic nucleus offers further opportunities to optimize both transfection efficiency and cell viability parameters. The use of cyclodextrins (CDs) as macrocyclic platforms to build preorganized gene delivery systems deserves special consideration in this context. Cyclodextrins are biocompatible, commercially available nano-objects with two well-differentiated and chemically distinguishable faces, one limited by the primary hydroxyls and the other by the secondary ones. Installation of a broad range of functional elements in a controlled, space-oriented manner can then be achieved through position-selective and face-selective functionalization methodologies. This chemical flexibility has been exploited to elaborate macromolecular constructs that can be programmed to recognize biomacromolecular partners, including lectins,[118] toxins[119] and DNA.[120]

In 2004 O'Driscoll and Darcy reported the first examples of monodisperse cyclodextrin-centred polycationic clusters intended for DNA complexation,[121] namely βCD derivatives bearing alkyl and arylamine antennae on their primary rim and unmodified OH-2 and OH-3 hydroxyls (*e.g.* **8**), and showed that they were able to promote transfection *in vitro* to some extent. This work has been followed by others, including Yannakopoulou's,[122] Reineke's[123] and Midoux's groups,[124] who proposed the installation of guanidinium (**9**)/guanidinioalkylamino (*e.g.* **10**), oligoethyleneamine (*e.g.* **11**) or histidine (**12**)/lysine substituents (**13**) at the primary (C-6) positions of CD platforms (Fig. 8). The highly hydrophilic character of these materials someway hampers the self-organization of the monomers at the surface of the oligonucleotide chain, however. Consequently, relatively big excesses are necessary to achieve full protection from the environment. An exception is the histidine-αCD derivative **12**, for which it was hypothesized that the corresponding CD-DNA complexes (CDplexes) are stabilized through intermolecular inclusion involving the imidazole ring and the CD cavity. In any case, all these series the CD-based polycations were able to induce condensation of DNA. In some cases, the resulting CDplexes exhibited remarkable cell-penetrating and transfection capabilities while featuring much less toxic profiles than commercial cationic polymers.

The above examples illustrate the suitability of CDs as molecular platforms for the incorporation of cationic elements with a precise spatial orientation and the potential of the resulting monodisperse polycations as

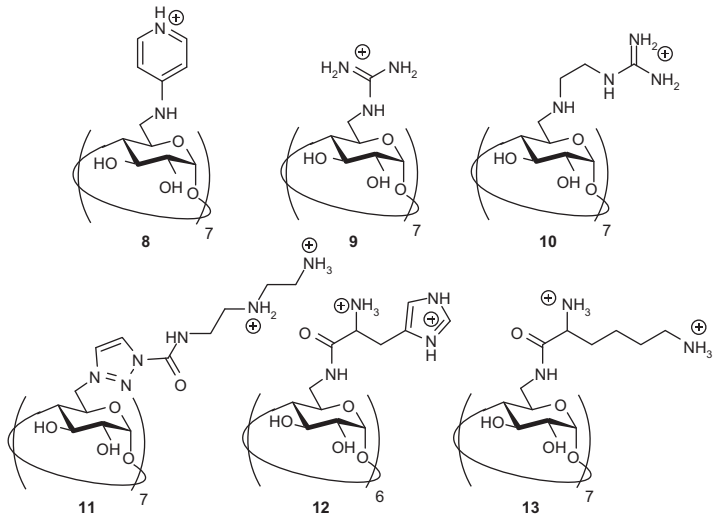
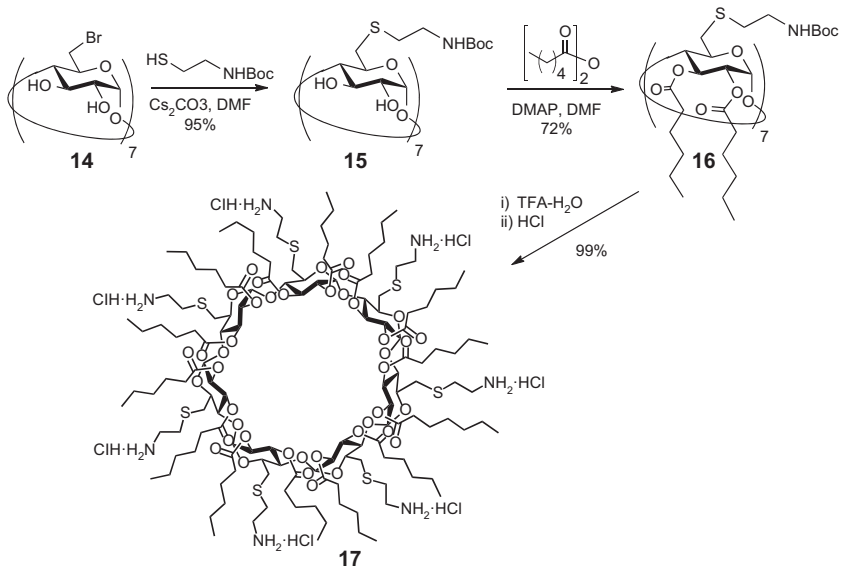


Fig. 8 Examples of non-amphiphilic polycationic cyclodextrins assayed as gene vectors.

gene vectors. Most interestingly, the rim anisotropy of the basket-shaped CD structure further allows accessing compounds with segregated cationic and lipophilic domains for which, according to the "facial amphiphilicity" concept, biomimetic self-assembly and gene delivery properties could be expected. Darcy and co-workers first tested this hypothesis by preparing heterogeneous βCD derivatives bearing thioalkyl chains at the primary rim and amine-terminated oligoethylene branches of variable lengths at the secondary hydroxyl positions. The resulting polycationic amphiphilic CDs (paCDs) were shown to entrap pDNA by forming CDplexes that transfected COS-7 and human hepatocellular liver carcinoma HepG2 cells with efficiencies and toxicity profiles comparable to Lipofectamine™ 2000-based lipoplexes.[125]

The possibility of implementing synthetic methodologies allowing full control of the chemical composition and of the conformational behavior of the vector let envisage the goal of investigating the influence of structural modifications at the molecular level on the self-assembling, DNA complexation, cell internalization and transfection properties. With this idea in mind, a multicollaborative project was settled between several European laboratories, including those of Defaye (Grenoble, France), Vierling and Di Giorgio (Nice, France), Ortiz Mellet and García Fernández (Sevilla, Spain), to develop a new family of monodisperse polycationic amphiphilic CD-based glycomaterials as gene delivery systems. The first candidate was a βCD derivative with seven cysteaminyl groups at the primary face and fourteen hexanoyl chains at the secondary rim, accessible in only three steps in multigram scale after a single column chromatography purification.[126] The synthetic sequence starts with the displacement of the bromo groups in the heptabromo derivative **14** with *N*-Boc-protected cysteamine (→**15**). The fatty acyl chains were put in place by subsequent reaction with hexanoic anhydride in DMF using DMAP as base promoter (→**16**). These acylation conditions warrant full homogeneity of the final compound, avoiding

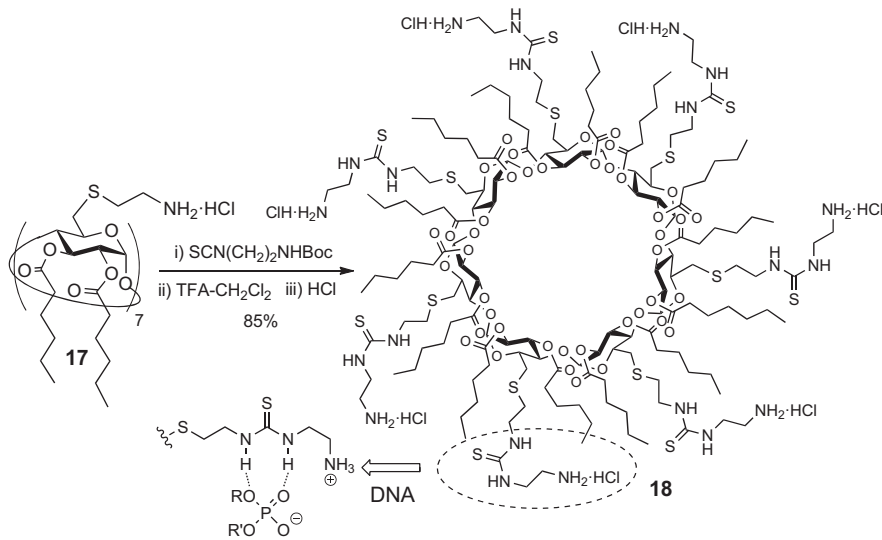


Scheme 9 Synthesis of the cysteaminyl polycationic amphiphilic cyclodextrin (paCD) **17**.

under- and overacylation side products. Final acid hydrolysis of the seven carbamate groups afforded the target paCD derivative **17** in excellent yield (Scheme 9).

Electrophoretic mobility experiments on agarose gel indicated that the amphiphilic cysteamine derivative formed a stable complex with pDNA at N/P 5 that fully protected the plasmid from degradation by nucleases or from binding to the intercalating agent etidium bromide. The plasmid was recovered intact after dissociation of the complex by treatment with sodium dodecylsulfate (SDS). Further transfection experiments in COS-7 cells indicated that it was as efficient as branched PEI in promoting transfection at N/P 5. However, the efficiency was 100-fold lower as compared to bPEI at N/P 10, the optimal charge ratio for the reference cationic polymer, although with a much more favorable toxicity profile.

To improve the transfection efficiency, the authors got inspiration from Nature's mechanisms of phosphate anion recognition,[127] which imply the concerted action of hydrogen bonding and electrostatic interactions. Thus, a new vector prototype incorporating a belt of thiourea segments between the cyclooligosacharide core and the polycationic cluster was designed. The synthesis implies the nucleophilic addition of the heptacysteaminyl derivative **17** to Boc-protected 2-aminoethyl isothiocyanate and further acidic cleavage of the protecting groups, affording the amphiphilic polyaminothiourea adduct **18** in 85% overall yield (Scheme 10). The homogeneity of the compound was confirmed by combustion analysis, MS and NMR spectroscopy, which fully demonstrated that the C_7 molecular symmetry was preserved. The biological results indicated a significant increase in the transfection efficiency for polyaminothiourea **18** as compared with polyamine **17**, paralleling that for polyplexes formulated with PEI at its optimal N/P 10 ratio.[126] Moreover, transfection efficiency was not



Scheme 10 Synthesis of the aminoethylthioureido paCD **18**.

affected in serum-containing media, an important requisite for *in vivo* applications.

The above synthetic strategy had been purposely conceived to be molecular diversity-oriented. In a first series of compounds, the nature of the spacer between the thiourea and the amino groups was systematically modified.[128] Aromatic spacers proved to be detrimental for transfection and, moreover, led to a substantial increase in cytotoxicity. In contrast, aminoalkyl thioureido derivatives were excellent transfection agents, with the two-carbon spacer being the optimal structure. Their comparative evaluation indicated that a dendritic presentation of amino groups separated by ethylene segments provided the best transfection capabilities, with transfection efficiencies 1000- or 10-fold higher as compared with the reference polycationic polymer PEI at N/P 5 or 10, respectively (Fig. 9).

Dynamic light scattering (DLS) measurements and transmission electron microscopy (TEM) images of the CDplexes obtained from amphiphilic polyaminothiourea-CDs revealed a narrow size distribution. The nanocomplexes are rather regular and have about the optimal size (40–50 nm) for systemic transfection. Considering that the plasmid alone has a size of above 1 μm the compaction ratio was about 20-fold, which according to theoretical calculations implies that they are monomolecular in DNA. A snail-like ultrastructure can be distinguished in which the electron-dense dark regions correspond to the DNA chain and the lighter regions to the self-assembled paCDs. Probably the later are arranged in bilayers, with the hydrophobic domains facing each other and the cationic clusters pointing to the oligonucleotide chain, thereby favoring desolvation and condensation of the plasmid (Fig. 10).

The fact that the transfection efficiencies improved with an increase in the number of cationic groups over the initial 7 primary centers in βCD derivatives suggested that reversing the facial polarity, by anchoring the

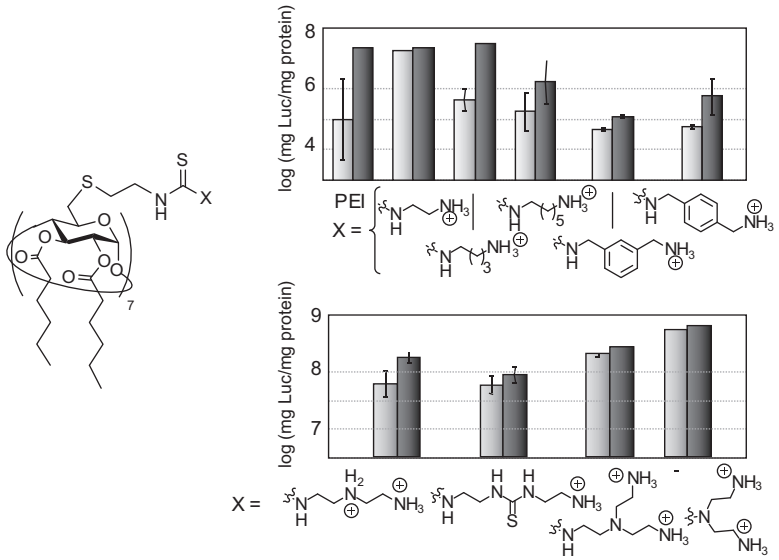


Fig. 9 Comparative transfection efficiencies of structurally diverse thioureido-paCDs.

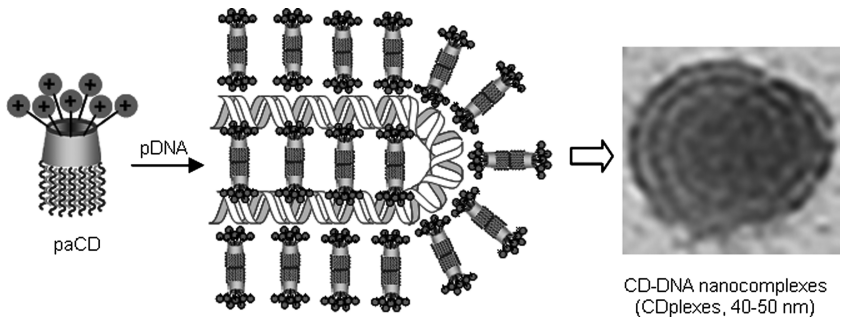
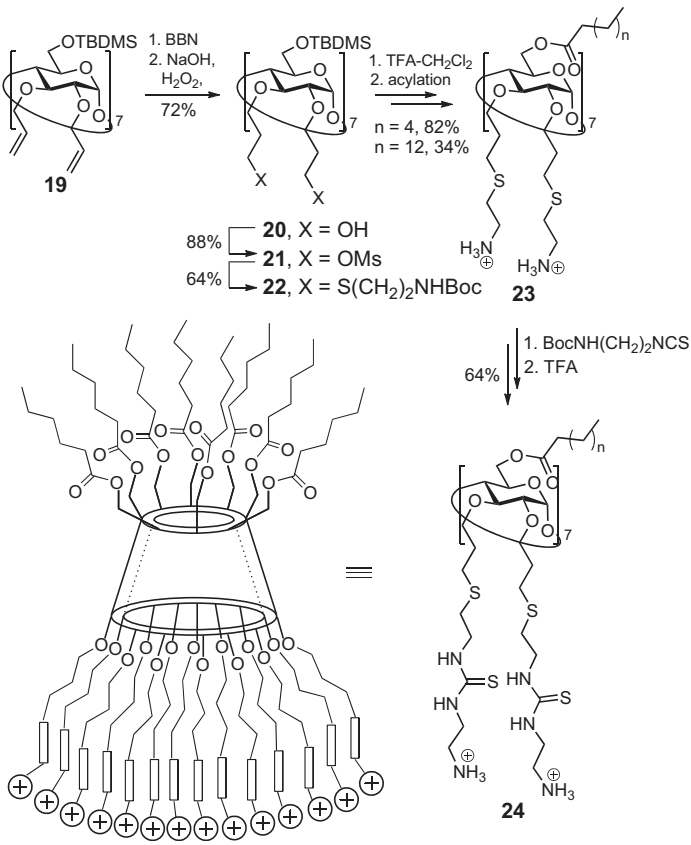


Fig. 10 Schematic representation of the formation of the supramolecular paCD-pDNA nanocomplexes CDplexes and TEM image of a CDplex obtained from compound **18**.

charged functionalities at secondary positions, might be beneficial for gene vector design. However, the difficulties associated with secondary hydroxyl groups reactivity, steric hindrance and positional isomerism make homogeneous functionalization at this face a far more complicated challenge. A general strategy was developed starting from the facially differentiated precursor **19**, having the primary (O-6) hydroxyls protected as the corresponding *tert*-dibutyldimethylsilyl ethers and the secondary (O-2 and O-3) hydroxyls fully allylated. Hydroboration-oxidation (→**20**) generated 14 primary hydroxyls that were subjected to mesylation (→**21**) and nucleophilic displacement by Boc-protected cysteamine (→**22**). Sequential hydrolysis of the silyl ether groups, acylation and Boc hydrolysis afforded a first family of ''reverse'' cysteaminyl paCDs. Optimization of the transfection properties required the installation of myristoyl acyl groups at the primary face (**23**, n = 12) in this case. Alternatively, the cysteaminyl groups

in the hexanoylated derivative (**23**, n = 4) were transformed into aminoethylthioureido functionalities (**24**). Both **23** (n = 12) and **24** (n = 4) condensed pDNA into monodisperse populations of very small-size cationic CDplexes (47 ± 15 nm, ζ-potential 39 ± 1 mV and 40 ± 3 nm, ζ-potential 44 ± 1 mV, respectively, at N/P 10) that exhibited transfection efficiencies similar to that of PEI-based polyplexes (Scheme 11).[129]

As an alternative approach for the construction of monodisperse polycationic amphiphilic CDs, the archetypal click-type Cu(I)-catalyzed azide-alkyne cycloaddition reaction was explored in collaboration with the group of Santoyo-González. The bioorthogonal character of the chemical functionalities at play and the unique property of CuAAC as a quantitative ligation method make this methodology very well suited for the construction of vector libraries. The use of solid-supported Cu(I) catalysts[130] was found to be particularly well suited for multiple CuAAC with polyfunctionalized cyclodextrins while facilitating purification of the C_7-symmetric macromolecular triazol adducts. The density of positive charges, their distance to the CD core and the length of the lipophilic tails are key parameters that can be systematically modified to finely tune the transfection capabilities. The results indicated that adjusting the hydrophilic/hydrophobic



Scheme 11 Synthesis and schematic representation of "reverse" paCDs.

balance is critical to achieve optimal self-assembling properties and stability of the resulting CDplexes.[131] The benefits imparted by amphiphilicity appeared evident when comparing the behaviour within the branched series **25**, **26** and **27** (Fig. 11). Compound **26**, bearing hexanoate tails, gave rise to stable CDplexes that warranted full protection of pDNA from the environment at N/P values ≥ 5. Moreover, these CDplexes showed remarkably high transfection efficiencies, analogous to that obtained using Lipofectamine™ 2000. Increasing hydrophobicity by replacing the hexanoate into tetradecanoate chains did not alter the complexing efficiency, as seen from data for **27**, but drastically decreased the transfection capabilities, indicating that molecular modifications in paCDs do not only affects the formation of CD-DNA nanoparticles, but also has dramatic consequences on the further processes leading to cell internalization and gene expression.

In general, triazole-type "click" paCDs were less efficient in promoting gene transfection than thiourea-type derivatives in saline and serum-containing media, but this could be corrected by implementing a "dual click" approach implying sequential CuAAC and thiourea-forming reactions (*e.g.* **28**, Fig. 12).[132] Broadening the spectrum of paCD structures available by combining these two very robust methodologies opens new opportunities for vector optimization.

Further work, carried out in collaboration with the group of Tros de Ilarduya (Navarra, Spain), indicated that CDplexes formulated from a plasmid codifying interleukin 12 (IL-12), a cytokine with anticancer activity, and the dendritic aminoethylthioureido paCD derivative **29** were able to transfect hepatocarcimoma Hep-G2 cells with significantly higher efficiency than the corresponding polyplexes prepared with PEI.[133] CDplexes loaded with a luciferase-encoding reporter gene were next injected in a group of mice to investigate the expression of the protein in different organs. The nanocomplexes escaped from the lungs and exhibited a certain tropism to the liver, thereby illustrating its potential in cellular hepatocarcinoma-targeted gene therapies (Fig. 13).

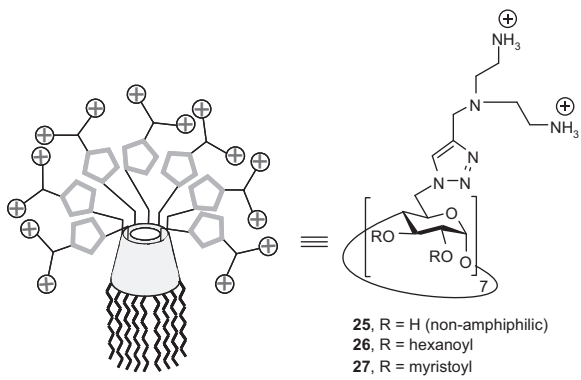


Fig. 11 Schematic representation of "click" paCDs and structures of some representative examples.

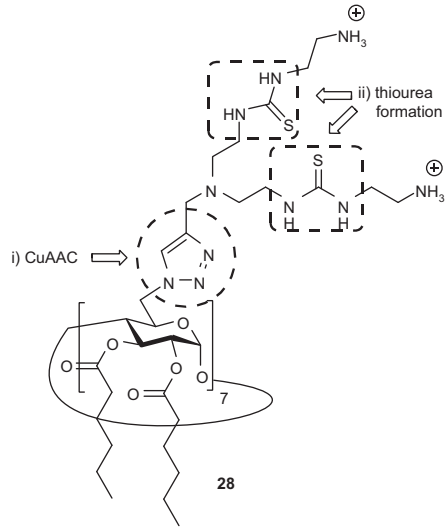


Fig. 12 Representative example of ''dual click'' paCD with highly efficient gene delivery capabilities.

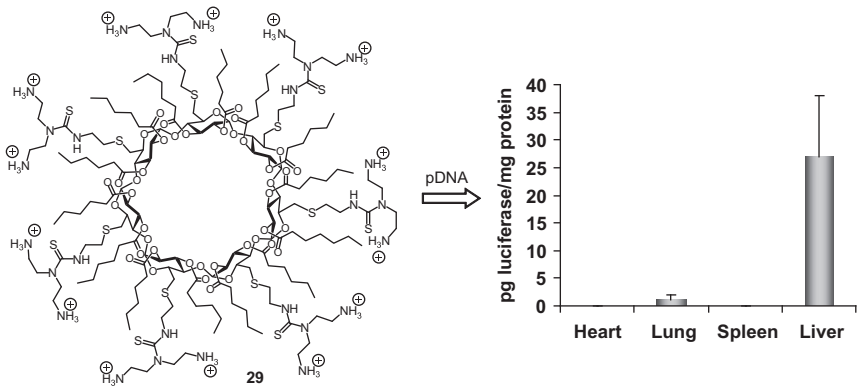


Fig. 13 Biodistribution of CDplexes obtained from paCD **29** after injection in mice.

An important issue in relationship with the possible preferential transfection of certain cell lines is the elucidation of the routes by which CDplexes are internalized and their fate inside the cell. This question has been addressed in collaboration with the group of De Smedt (Ghent, Belgium). The kinetics and the mechanism of internalization of CDplexes formulated with paCD **30**, bearing up to 21 amino groups and labelled with the fluorescent probe lissamine-rhodamine, were investigated in kidney epithelial VERO cells.[134] The uptake of the CDplexes was monitored in real time by confocal laser microscopy, revealing that this is a rather fast process. The resulting endosomes tend to accumulate at the vicinity of the nucleus after a certain time, pointing to the implication of active transport mechanisms. These observations are consistent with internalization *via*

active-mediated endocytosis rather than through passive diffusion. Clathrin-dependent and caveolin-dependent mechanisms were evidenced by using inhibitors of the main internalization routes, but the endosomes arising from the caveolae route were far more productive regarding transfection (Fig. 14). It must be stressed, however, that those results only applied strictly to the system under study. Other mechanisms may be favoured for other paCDs or in other cell types. Thus, it has been recently reported that macropinocytosis is the preferred internalization mechanism for a cysteaminyl paCD derivative in intestinal epithelial Caco-2 cells.[135]

Since caveolae-mediated endocytosis and macropinocytosis operate in many cell types, a main conclusion of the above work is that CDplexes should be considered as broad range transfection systems. Although some tissue selectivity can be achieved upon systemic injection, probably as a function of particle size, cell-specific delivery would need the incorporation of targeting ligands. Given the fundamental role of carbohydrates in cell communication processes, the formulation of glycoCDplexes appeared particularly attractive. Actually, glycoconjugation has been broadly used to modulate the recognition and supramolecular properties of cyclodextrins,[136] including amphiphilic CDs.[137] In a first attempt, the combined use of polycationinc amphiphilic CDs and neutral glycoamphiphilic CDs for pDNA complexation was assayed. However, such formulations failed to provide full DNA protection and mediate transfection, even when using low

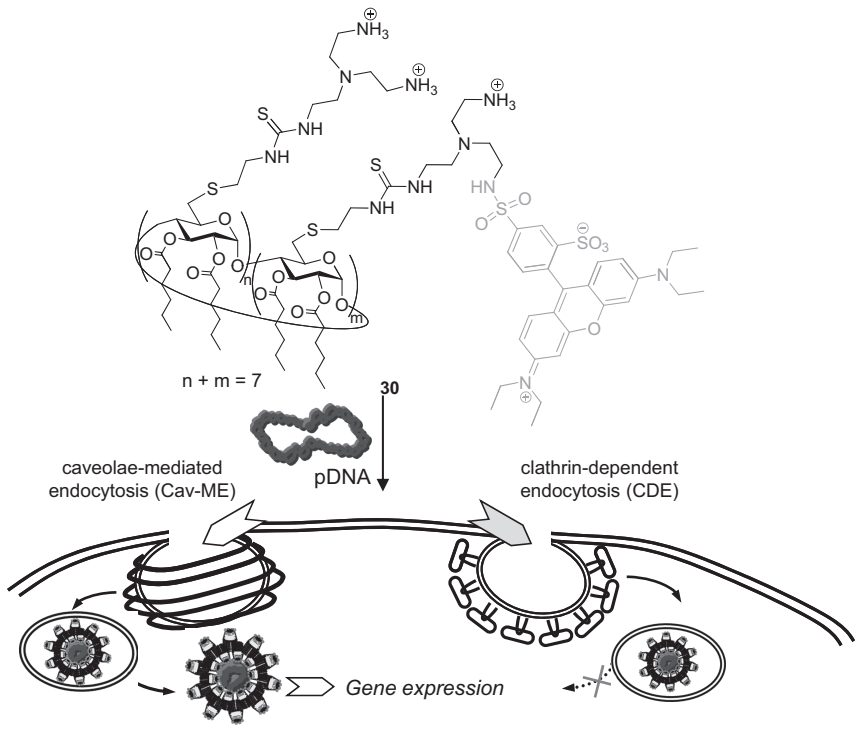


Fig. 14 Schematic representations of the internalization routes at play for CDplexes formulated with the fluorescent paCD **30** in VERO cells.

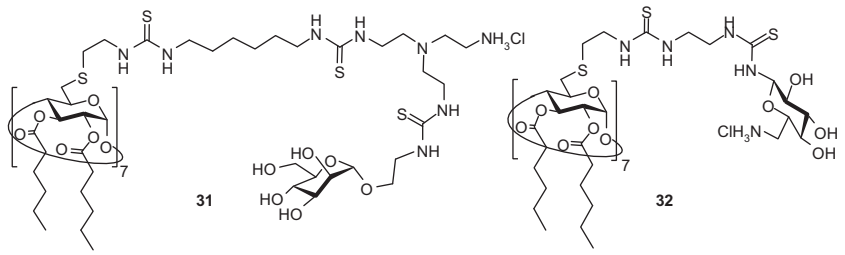
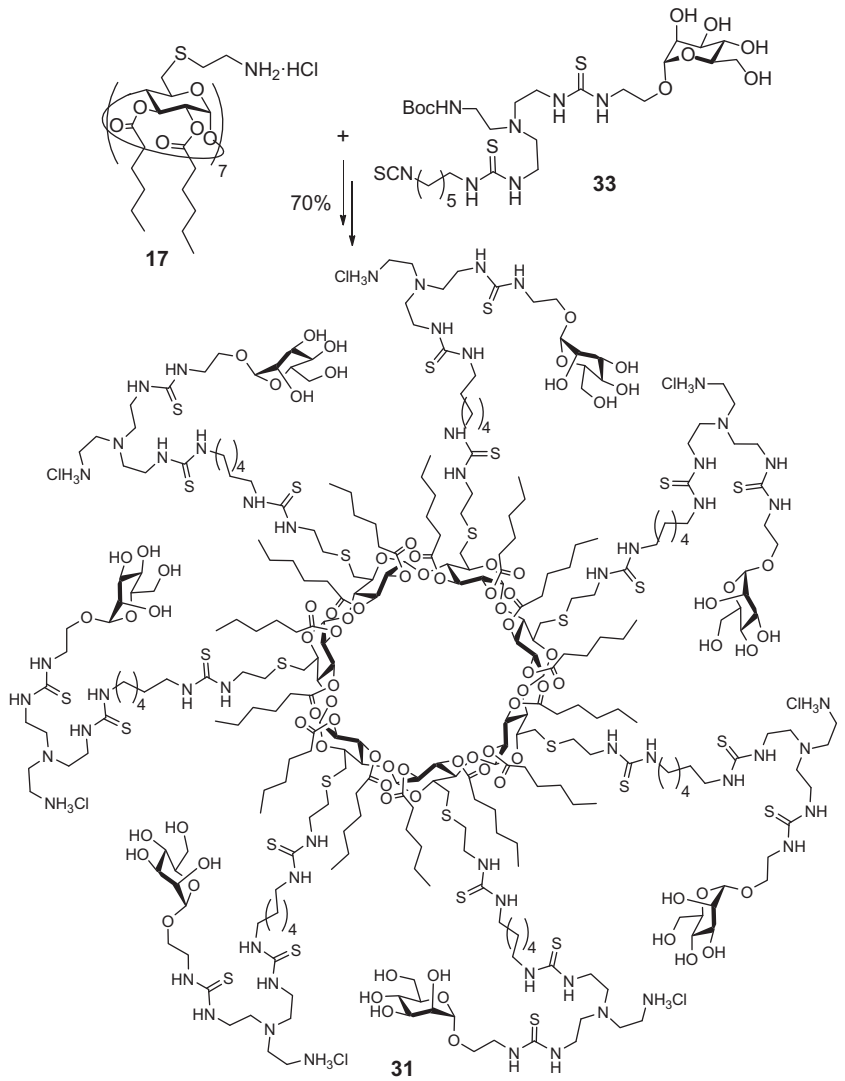


Fig. 15 Structures of the two prototypes of polycationic glycoamphiphilic cyclodextrins (pGaCDs) **31** and **32**.

(2%) proportions of the neutral glycocluster. It was reasoned that the presence of neutral microdomains strongly destabilized the supramolecular nanocomplexes. To overcome this problem, a new generation of polycationic glycoamphiphilic CDs (pGaCDs), with regular arrangements of glyco-cationic moieties, was designed.[138] Two different prototypes were considered, namely a derivative having the sugar moieties (mannosyl residues) and the cationic centres in separate branches (**31**) and another one in which the cationic groups are located in the sugar moieties (**32**). In the second case, the 6-amino-6-deoxyglucopyranosyl motif was chosen since it is present in aminoglycoside antibiotics known to bind to the DNA double-helix (Fig. 15).[40]

The synthesis of these macromolecules followed a convergent scheme in which the heterobifunctional antenna, conveniently armed with a reactive isothiocyanate group, was first prepared and coupled with the amphiphilic cyclodextrin polyamine core **17** in a late step, as illustrated in Scheme 12 for the mannosylated pGaCD **31**. After the final deprotection reaction, the target 7-fold symmetrical pGaCDs were obtained in homogeneous form, as confirmed by MS, NMR and combustion analysis. Both compounds were able to complex pDNA to form nanoparticulate glycoCDplexes that fully protected the plasmid and promoted transfection of COS-7 cells.

The potential of the mannosylated pGaCD derivative for site specific gene delivery was preliminary evaluated by measuring the ability of the corresponding glycoCDplexes to bind the mannose specific lectin concanavalin A (Con A) by enzyme-linked lectin assay (ELLA). Further evidence was obtained by developing a modified ELLA protocol for recombinant human macrophage mannose receptor (rhMMR) and by measuring the binding affinity towards murine macrophages in vitro. CDplexes obtained from the non-glycosylated paCD or from the aminoglucosylated pGaCD were used as negative controls, whereas mannosylated-BSA was employed in competitive assays to confirm the specific mannose receptor-dependent character of the association. The results evidenced that mannosyl-CDplexes binding to macrophages can proceed through non-specific electrostatic interactions with negatively charged proteoglycans at the cell surface or through specific interaction with MMRs (Fig. 16). At N/P 10, both mechanisms simultaneously operate. However, at N/P 5 the MMR-dependent route is exceedingly dominant. CDplexes



Scheme 12 Synthesis of the pGaCD **31**.

formulated at this N/P ratio specifically transfected macrophage RAW264.7 cells versus epithelial COS-7 or hepatic BNL-CL2 cells.[138]

Although the above commented results require validation in vivo, they represent a proof of concept of the suitability of the strategy based on pGaCDs to mimic the site-specific transfection capabilities of viruses. The synthetic effort can be limited by statistic, but controlled, incorporation of the appropriate glycotopes at the requested density. This strategy has been already implemented for the preparation of galactosylated pGaCDs that were able to complex pDNA and mRNA.[139] The corresponding glycoCDplexes were internalized by hepatocytes *via* the asialoglycoprotein receptor. Non-amphiphilic polycationic cyclodextrins exposing mannose,

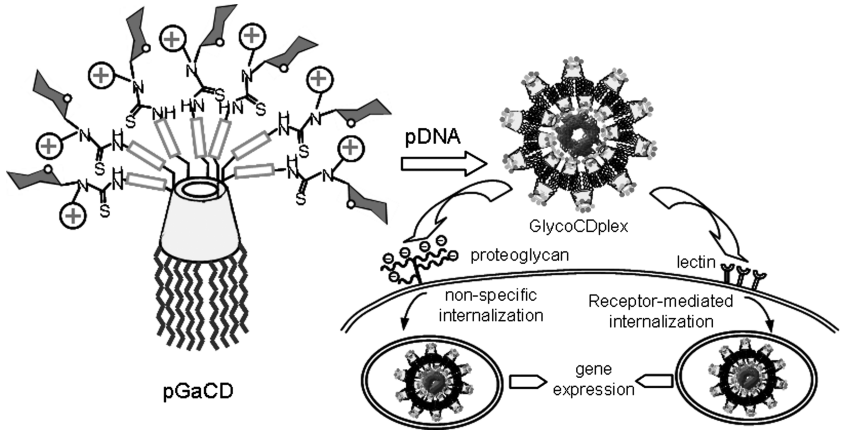


Fig. 16 Schematic representation of the competing cell internalization routes for glycoCDplexes. The lectin-mediated internalization pathway is largely favoured at N/P 5.

glucosamine and galactose residues have also been recently reported, but no DNA complexing capabilities were studied.[140]

5 Outlook and perspectives

In this review, we covered the recent literature regarding the use of glycomaterials in non-viral gene vector design. So far, most of the examples exploit the potential of carbohydrates to improve the transfection capabilities of preformed polymeric vectors by enhancing biodisponibility or imparting targeting capabilities. But there are examples that demonstrate that carbohydrates can play an active role in the interaction with polynucleotides or in providing the right spatial orientation of key functional elements for efficient DNA binding. Cyclodextrins are paradigmatic carbohydrate platforms for these channels. Their suitability to access monodisperse facial amphiphiles that self-associate in the presence of pDNA to form nanocomplexes with transfectious capabilities has been already demonstrated. Moreover, the nanometric cavity remains accessible in the corresponding CDplexes, opening the door to nanoparticle surface coating through supramolecular interactions. The CD commercial availability, easy and relatively inexpensive synthesis of appropriate derivatives, facile purification, robustness and stability, biocompatibility, lack of immunogenicity and safety match important criteria for non-viral vectors. Moreover, the combination of a carbohydrate core and a carbohydrate coating in pGaCDs offers new venues for the design of tailor-made artificial viruses. In all, there should be plenty of room for carbohydrate chemists to develop high level research connecting sugars and gene therapy.

Acknowledgments

The Spanish Ministerio de Economía y Competitividad (contract numbers CTQ2010-15848 and SAF2010-15670; co-financed with the Fondo Europeo de Desarrollo Regional FEDER) and the Junta de Andalucía are thanked for funding.

References

1 N. Cartier, S. Hacein-Bey-Abina, C. C. Bartholomae, G. Veres, M. Schmidt, I. Kutschera, M. Vidaud, U. Abel, L. Dal-Cortivo, L. Caccavelli, N. Mahlaoui, V. Kiermer, D. Mittelstaedt, C. Bellesme, N. Lahlou, F. Lefrère, S. Blanche, M. Audit, E. Payen, P. Leboulch, B. l'Homme, P. Bougnères, C. Von Kalle, A. Fischer, M. Cavazzana-Calvo and P. Aubourg, *Science*, 2009, **326**, 818.

2 S. Nayak and R. W. Herzog, *Gene Ther.*, 2010, **17**, 295.

3 C. C. Gao, *Hum. Gene Ther.*, 2011, **22**, A114.

4 J. K. Räty, J. T. Pikkarainen, T. Wirth and S. Ylä-Herttuala, *Curr. Mol. Pharmacol.*, 2008, **1**, 13.

5 M. A. Mintzer and E. E. Simanek, *Chem. Rev.*, 2009, **109**, 259.

6 R. J. Christie, N. Nishiyama and K. Kataoka, *Endocrinology*, 2010, **151**, 466.

7 M. E. Davis, *Mol. Pharm.*, 2009, **6**, 659.

8 D. J. Glover, H. J. Lipps and D. A. Jans, *Nat. Rev. Genet.*, 2005, **6**, 299.

9 D. Putnam, *Nat. Mater.*, 2006, **5**, 439.

10 B. Demeneix, Z. Hassani and J.-P. Behr, *Curr. Gene Ther.*, 2004, **4**, 445.

11 K. Itaka and K. Kataoka, *Eur. J. Pharm. Biopharm.*, 2009, **71**, 475.

12 J. H. Jeong, S. W. Kim and T. G. Park, *Prog. Polym. Sci.*, 2007, **32**, 1239.

13 G. F. Walker, C. Fella, J. Pelisek, J. Fahrmeir, S. Boeckle, M. Ogris and E. Wagner, *Mol. Ther.*, 2005, **11**, 418.

14 H. Mok and T.-G. Park, *Macromol. Biosci.*, 2009, **9**, 731.

15 S. Ganta, H. Devalapally, A. Shahiwala and M. Amiji, *J. Control. Release*, 2008, **126**, 187.

16 K. Knop, R. Hoogenboom, D. Fischer and U. S. Schubert, *Angew. Chem. Int. Ed.*, 2010, **49**, 6288.

17 C. Biesa, C.-M. Lehra and J. F. Woodley., *Adv. Drug Deliv. Rev.*, 2004, **56**, 425.

18 H. Zhang, Y. Ma and X.-L. Sun, *Med. Res. Rev.*, 2010, **30**, 270.

19 S. Mao, W. Sun and T. Kissel, *Adv. Drug Deliv. Rev.*, 2010, **62**, 12.

20 For a recent survey on the use of naturally-occurring carbohydrate polymers in gene delivery, see: A. Sizovs, P. M. McLendon, S. Srinivasachari and T. M. Reineke, *Top. Curr. Chem.*, 2010, **296**, 131.

21 (a) L. Wightman, R. Kircheis, V. Rossler, S. Carotta, R. Ruzicka, M. Kursa and E. Wagner, *J. Gene Med.*, 2001, **3**, 362; (b) Y. H. Choi, F. Liu, J.-S. Kim, Y. K. Choi, J.-S. Park and S. W. Kim, *J. Control. Release*, 1998, **54**, 39; (c) P. Van de Wetering, E. E. Moret, N. M. E. Schuurmans-Nieuwenbroek, M. J. Van Steenbergen and W. E. Hennink, *Bioconjugate Chem.*, 1999, **10**, 589; (d) K. M. Kitchens, M. E. H. El-Sayed and H. Ghandehari, *Adv. Drug Deliv. Rev.*, 2005, **57**, 2163.

22 S. Aoshima and S. Kanaoka, *Chem. Rev.*, 2009, **109**, 5245.

23 S. M. Moghimi, P. Symonds, J. C. Murray, A. C. Hunter, G. Debska and A. Szewczyk, *Mol. Ther.*, 2005, **11**, 990.

24 T. Merdan, K. Kunath, H. Petersen, U. Bakowsky, K. H. Voigt, J. Kopecek and T. Kissel, *Bioconjugate Chem.*, 2005, **16**, 785.

25 C. K. Chan and D. A. Jans, *Gene Ther.*, 2001, **8**, 166.

26 M. Kursa, G. F. Walker, V. Roessler, M. Ogris, W. Roedl, R. Kircheis and E. Wagner, *Bioconjugate Chem.*, 2003, **14**, 222.

27 T. Kakudo, S. Chaki, S. Futaki, I. Nakase, K. Akaji, T. Kawakami, K. Maruyama, H. Kamiya and H. Harashima, *Biochemistry*, 2004, **43**, 5618.

28 (a) E. Wagner, C. Culmsee and S. Boeckle, *Adv. Genet.*, 2005, **53**, 333; (b) J. Guo, K. A. Fisher, R. Darcy, C. F. Cryan and C. O'Driscoll, *Mol. Biosyst.*, 2010, **6**, 1143.

29 M. A. Zanta, P. Belguise-Valladier and J.-P. Behr, *Proc. Natl Acad. Sci. USA*, 1999, **96**, 91.
30 (a) S. H. Kim, J. H. Jeong, K. C. Cho, S. W. Kim and T. G. Park, *J. Control. Release*, 2005, **104**, 223; (b) X. Zhao, H. Li and R. J. Lee, *Expert Opin. Drug Deliv.*, 2008, **5**, 309.
31 (a) K. Kunath, T. Merdan, O. Hegener, H. Häberlein, T. Kissel, *J. Gene Med.*, 2003, **5**, 588; (b) B. Lebleu, H. M. Moulton, R. Abes, G. D. Ivanova, S. Abes, D. A. Stein, P. L. Iversen, A. A. Arzumanov and M. J. Gait, *Adv. Drug Deliv. Rev.*, 2008, **60**, 517.
32 (a) M. Ogris, P. Steinlein, M. Kursa, K. Mechtler, R. Kircheis and E. Wagner, *Gene Ther.*, 1998, **5**, 1425; (b) R. Huang, W. Ke, Y. Liu, C. Jiang and Y. Pei, *Biomaterials*, 2008, **29**, 238.
33 (a) B. Xu, S. Wiehle, J. A. Roth and R. J. Cristiano, *Gene Ther.*, 1998, **5**, 1235; (b) H. Lee, T. H. Kim and T. G. Park, *J. Control. Release*, 2002, **83**, 109.
34 (a) M. C. Palanca-Wessels, A. J. Convertine, R. Cutler-Strom, G. C. Booth, F. Lee, G. Y. Berguig, P. S. Stayton and O. W. Press, *Mol. Ther.*, 2011, **19**, 1529; and (b) V. Torchilin, *Expert Opin. Drug Deliv.*, 2008, **5**, 1003.
35 (a) G. Y. Wu, C. H. Wu, *J. Biol. Chem.*, 1987, **262**, 4429; (b) G. Y. Wu and C. H. Wu, *J. Biol. Chem.*, 1988, **263**, 14621
36 S. C. De Smedt, J. Demeester and W. E. Hennink, *Pharm. Res.*, 2000, **17**, 113.
37 S. Höbel, A. Loos, D. Appelhans, S. Schwarz, J. Seidel, B. Voit and A. Aigner, *J. Control. Release*, 2011, **149**, 146.
38 (a) S. P. Strand, M. M. Issa, B. E. Christensen, K. M. Vårum and P. Artursson, *Biomacromolecules*, 2008, **9**, 3268; (b) S. P. Strand, S. Lelu, N. K. Reitan, C. de Lange Davies, P. Artursson and K. M. Vårum, *Biomaterials*, 2010, **31**, 975.
39 K. Koumoto, M. Mizu, T. Anada, T. Nagasaki, S. Shinkai and K. Sakurai, *Bull. Chem. Soc. Jpn.*, 2005, **78**, 1821.
40 (a) L. Desigaux, M. Sainlos, O. Lambert, R. Chevre, E. Letrou-Bonneval, J.-P. Vigneron, P. Lehn, J.-M. Lehn and B. Pitard, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 16534; (b) M. Sainlos, M. Hauchecorne, N. Oudrhiri, S. Zertal-Zidani, A. Aissaoui, J.-P. Vigneron, J.-M. Lehn and P. Lehn, *ChemBioChem*, 2005, **6**, 1023.
41 (a) H. Huang, G. Tang, Q. Wang, D. Li, F. Shen, J. Zhoub and Hai Yu, *Chem. Commun.*, 2006, 2382; (b) G. P. Tang, H. Y. Guo, F. Alexis, X. Wang, S. Zeng, T. M. Lim, J. Ding, Y. Y. Yang, S. Wang, *J. Gene Med.*, 2006, **8**, 736.
42 (a) H. Huang, H. Yu, G. Tang, Q. Wang and J. Li, *Biomaterials*, 2010, **31**, 1830; (b) Q.-Y. Jiang, L.-H. Lai, J. Shen, Q.-Q. Wang, F.-J. Xu and G.-P. Tang, *Biomaterials*, 2011, **32**, 7253.
43 Y. Ping, C. Liu, Z. Zhang, K. L. Liu, J. Chen and J. Li, *Biomaterials*, 2011, **32**, 8328.
44 X. Liu, P. Rocchi and L. Peng, *New J. Chem.*, 2012, **36**, 256.
45 (a) H. Arima, F. Kihara, F. Hirayama and K. Uekama, *Bioconjugate Chem.*, 2001, **12**, 476; (b) F. Kihara, H. Arima, T. Tsutsumi, F. Hirayama and K. Uekama, *Bioconjugate Chem.*, 2002, **13**, 1211.
46 T. Tsutsumi, F. Hirayama, K. Uekama and H. Arima, *J. Control. Release*, 2007, **119**, 349.
47 (a) H. Arima, Y. Chihara, M. Arizono, S. Yamashita, K. Wada, F. Hirayama and K. Uekama, *J. Control. Release*, 2006, **116**, 64; (b) Y. Chihara, H. Arima, M. Arizono, K. Wada, F. Hirayama and K. Uekama, *J. Incl. Phenom. Macrocycl. Chem.*, 2006, **56**, 89.
48 (a) M. Hashimoto, M. Morimoto, H. Saimoto, Y. Shigemasa and T. Sato, *Bioconjugate Chem.*, 2006, **17**, 309; (b) M. Hashimoto, M. Morimoto,

H. Saimoto, Y. Shigemasa, H. Yanagie, M. Eriguchi and T. Sato, *Biotechnol. Lett*., 2006, **28**, 815.
49 M. Ahmed, X. Jiang, Z. Deng and R. Narain, *Bioconjugate Chem*., 2009, **20**, 2017.
50 M. Ahmed, Z. Deng and R. Narain, *ACS Appl. Mater. Interfaces*, 2009, **9**, 1980.
51 J.-M. Li, M.-X. Zhao, H. Su, Y.-Y. Wang, C.-P. Tan, L.-N. Ji and Z.-W. Mao, *Biomaterials*, 2011, **32**, 7978.
52 (a) J. Li and X. J. Loh, *Adv. Drug. Deliv. Rev*., 2008, **60**, 1000; (b) J. J. Li, F. Zhao and J. Li, *Appl. Microbiol. Biotechnol*., 2011, **90**, 427.
53 (a) I.-K. Park, H. A. von Recum, S. Jiang and S. H. Pun, *Langmuir* 2006, **22**, 8478; (b) X. Zhang, K. K. Sharma, M. Boeglin, J. Ogier, D. Mainard, J.-C. Voegel, Y. Mély and N. Benkirane-Jessel, *Nano Lett*., 2008, **8**, 2432.
54 S.-J. Kim, H. Ise, M. Goto, K. Komura, C.-S. Cho and T. Akaike, *Biomaterials*, 2011, **32**, 3471.
55 D. B. Rozema, D. L. Lewis, D. H. Wakefield, S. C. Wong, J. J. Klein, P. L. Roesch, S. L. Bertin, T. W. Reppen, Q. Chu, A. V. Blokhin, J. E. Hagstrom and J. A. Wolff, *Proc. Natl. Acad. Sci USA*, 2007, **104**, 12982.
56 (a) M.-A. Zanta, O. Boussif, A. Adib and J.-P. Behr, *Bioconjugate Chem*., 1997, **8**, 839; (b) S. S. Diebold, M. Kursa, E. Wagner, M. Cotten and M. Zenke, *J. Biol. Chem*., 1999, **274**, 19087; (c) T. Bettinger, J.-S. Remy and P. Erbacher, *Bioconjugate Chem*., 1999, **10**, 558.
57 (a) P. Erbacher, A. C. Roche, M. Monsigny and P. Midoux, *Bioconjugate Chem*., 1995, **6**, 401; (b) J. C. Perales, T. Ferkol, H. Beegen, O. D. Ratnoff and R. W. Hanson, *Proc. Natl. Acad. Sci. USA* 1994, **91**, 4086; (c) T. Ferkol, J. C. Perales, F. Mularo and R. W. Hanson, *Proc. Natl. Acad. Sci. USA* 1996, **93**, 101.
58 H. Arima, S. Yamashita, Y. Mori, Y. Hayashi, K. Motoyama, K. Hattori, T. Takeuchi, H. Jono, Y. Ando, F. Hirayama and K. Uekama, *J. Control. Release*, 2010, **146**, 106.
59 (a) K. Motoyama, Y. Nakashima, Y. Aramaki, F. Hirayama, K. Uekama and H. Arima, *J. Drug Deliv*., 2011, Article ID 476137, doi:10.1155/2011/476137; (b) T. Montier, P. Delépine, T. Benvegnu, V. Ferrières, M.-L. Miramon, S. Dagorn, C. Guillaume, D. Plusquellec and C. Férec, *Blood Cells Mol. Dis*. 2004, **32**, 271; (c) M. Hashida, S. Kawakami and F. Yamashita, *Chem. Pharm. Bull*., 2005, **53**, 871.
60 C. Rondanino, M. T. Bousser, M. Monsigny and A. C. Roche, *Glycobiology*, 2003, **13**, 509.
61 (a) A. C. Roche, I. Fajac , S. Grosse, N. Frison, C. Rondanino, R. Mayer and M. Monsigny, *Cell. Mol. Life Sci*., 2003, **60**, 288; (b) M. Monsigny, C. Rondanino, E. Duverger, I. Fajac and A.-C. Roche, *Biochim. Biophys. Acta*, 2004, **1673**, 94.
62 (a) S. Grosse, A. Tremeau-Bravard, Y. Aron, P. Briand, I. Fajac, *Gene Ther*., 2002, **9**, 1000; (b) M. Hashimoto, M. Morimoto, H. Saimoto, Y. Shigemasa and T. Sato, *Bioconjugate Chem*., 2006, **17**, 309.
63 S. Grosse, G. Thévenot, M. Monsigny and I. Fajac, *J. Gene Med*., 2006, **8**, 845.
64 (a) H. Akita, T. Masuda, T. Nishio, K. Niikura, K. Ijiro and H. Harashima, *Mol. Pharm*., 2011, **8**, 1436; (b) T. Masuda, H. Akita, T. Nishio, K. Niikura, K. Kogure, K. Ijiro and H. Harashima, *Biomaterials*, 2008, **29**, 709.
65 M. Breton, J.-F. Berret, C. Bourgaux, T. Kral, M. Hof, C. Pichon, M. Bessodes, D. Scherman and N. Mignet, *Langmuir*, 2011, **27**, 12336.
66 P. Vicennati, A. Giuliano, G. Ortaggi and A. Masotti, *Curr. Med. Chem*., 2008, **15**, 2826.

67 M. Huang, C.-W. Fong, E. Khor and L.-Y. Lim, *J. Control. Release*, 1999, **106**, 391.
68 C. Troiber and E. Wagner, *Bioconjugate Chem.*, 2011, **22**, 1737.
69 C. Lin and J. F. J. Engbersen, *J. Control. Release*, 2008, **132**, 267.
70 Y. Liu, L. Wenning, M. Lynch and T. M. Reineke, *J. Am. Chem. Soc.*, 2004, **126**, 7422.
71 M. Metzke, N. O'Connor, S. Maiti, E. Nelson and Z. Guan, *Angew. Chem. Int. Ed.*, 2005, **44**, 6529.
72 M. Metzke and Z. Guan, *Biomacromolecules*, 2008, **9**, 208.
73 N. P. Ingle, B. Malone and T. M. Reineke, *Trends Biotechnol.*, 2011, **29**, 443.
74 (a) Y. Liu and T. M. Reineke, *J. Am. Chem. Soc.* 2005, **127**, 3004; (b) Y. Liu and T. M. Reineke, *Bioconjugate Chem.*, 2006, **17**, 101.
75 Y. Liu and T. M. Reineke, *Biomacromolecules*, 2010, **11**, 316.
76 C.-C. Lee, Y. Liu and Theresa M. Reineke, *Bioconjugate Chem.*, 2008, **18**, 428.
77 Y. Liu and T. M. Reineke, *Bioconjugate Chem.*, 2007, **18**, 19.
78 http://www.techulon.com/products, last visited March 2012.
79 T. M. Reineke and M. E. Davis, *Bioconjugate Chem.*, 2003, **14**, 247.
80 (a) S. Srinivasachari, Y. Liu, G. Zhang, L. Prevette and T. M. Reineke, *J. Am. Chem. Soc.*, 2006, **125**, 8176; (b) S. Srinivasachari, Y. Liu, L. Prevette and T. M. Reineke, *Biomaterials*, 2007, **28**, 2885.
81 R. D. Lins, C. S. Pereira and P. H. Hünenberger, *Proteins*, 2004, **55**, 177.
82 M. Meldal and C. W. Tornøe, *Chem. Rev.*, 2008, **108**, 2952.
83 M. E. Davis and M. E. Brewster, *Nat. Rev. Drug Discov.*, 2004, **3**, 1023.
84 J. E. Zuckerman, C. H. J. Choi, H. Han and M. E. Davis, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 3137.
85 H. Gonzalez, S. J. Hwang and M. E. Davis, *Bioconjugate Chem.*, 1999, **10**, 1068.
86 S. J. Hwang, N. C. Bellocq and M. E. Davis, *Bioconjugate Chem.*, 2001, **12**, 280.
87 T. M. Reineke and M. E. Davis, *Bioconjugate Chem.*, 2003, **14**, 247.
88 T. M. Reineke and M. E. Davis, *Bioconjugate Chem.*, 2003, **14**, 255.
89 S. R. Popielarski, S. Mishra and M. E. Davis, *Bioconjugate Chem.*, 2003 **14**, 672.
90 M. E. Davis and N. Q. Bellocq, *J. Incl. Phenom. Macrocycl. Chem.*, 2003 **44**, 17.
91 M. E. Davis, S. H. Pun, N. C. Bellocq, T. M. Reineke, S. R. Popielarski, S. Mishra and J. D. Heidel, *Curr. Med. Chem.*, 2004, **11**, 179.
92 S. Mishra, J. D. Heidel, P. Webster and M. E. Davis, *J. Control. Release*, 2006, **116**, 179.
93 J.-P. Behr, *Chimia*, 1997, **51**, 34.
94 S. Srinivasachari and T. M. Reineke, *Biomaterials*, 2009, **30**, 928.
95 S. Mishra, P. Webster and M. E. Davis, *Eur. J. Cell Biol.*, 2004, **83**, 97.
96 S. H. Pun and M. E. Davis, *Bioconjugate Chem.*, 2002, **13**, 630.
97 N. C. Bellocq, S. H. Pun, G. S. Jensen and M. E. Davis, *Bioconjugate Chem.*, 2003, **14**, 1122.
98 D. W. Bartlett and M. E. Davis, *Bioconjugate Chem.*, 2007, **18**, 456.
99 S. H. Pun, F. Tack, C. C. Bellocq, J. Cheng, B. H. Grubbs, G. S. Jensen, M. E. Davis, M. Brewster, M. Janicot, B. Janssens, W. Floren and A. Bakker, *Cancer Biol. Ther.*, 2004, **3**, 641.
100 S. Hu-Lieskovan, J. D. Heidel, D. W. Bartlett, M. E. Davis and T. J. Triche, *Cancer Res.*, 2005, **65**, 8984.
101 D. W. Bartlett and M. E. Davis, *Biotechnol. Bioeng.*, 2008, **99**, 975.
102 J. D. Heidel, Z. Yu, J. Y. Liu, S. M. Rele, Y. Liang, R. K. Zeidan, D. J. Kornbrust and M. E. Davis, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 5715.

103 http://clinicaltrials.gov/ct2/show/NCT00689065
104 M. E. Davis, J. E. Zuckerman, C. H. J. Choi, D. Seligson, A. Tolcher, C. A. Alabi, Y. Yen, J. D. Heidel and A. Ribas, *Nature*, 2010, **464**, 1067.
105 For the most recent updates, see: http://calando.interactive.biz/technology/rondel/ or http://clinicaltrials.gov/ct2/show/NCT00689065.
106 (a) B. A. Lowe and R. Wang, *Polymer*, 2007, **48**, 2221; (b) L. Albertin, M. Stenzel, C. Barner-Kowollik, L. John, R. Foster and T. P. Davis, *Macromolecules*, 2004, **37**, 7530; (c) L. Liu, J. Zhang, W. Lv, Y. Luo and X. Wang, *J. Polym. Sci. Part A: Polym. Chem.*, 2010, **48**, 3350; (d) C. Boyer and T. P. Davis, *Chem. Commun.*, 2009, 6029.
107 M. Ahmed and R. Narain, *Biomaterials*, 2011, **32**, 5279.
108 A. E. Smith, A. Sizovs, G. Grandinetti, L. Xue and T. M. Reineke, *Biomacromolecules*, 2011, **12**, 3015.
109 T. Fujiwara, S. Hasegawa, N. Hirashima, M. Nakanishi and T. Ohwada, *Biochim. Biophys. Acta-Biomembr.*, 2000, **1468**, 396.
110 C. Ortiz Mellet, J. M. Benito and J. M. García Fernández, *Chem. Eur. J.*, 2010, **16**, 6728.
111 (a) C. Li, H. Tian, N. Rong, K. Liu, F. Liu, Y. Zhu, R. Qiao and Y. Jiang, *Biomacromolecules*, 2011, **12**, 298; (b) H. Yan, Z.-F. Li, Z.-F. Guo, Z.-L. Lu, F. Wang and L.-Z. Wu, *Bioorg. Med. Chem.*, 2012, **20**, 801; (c) Q.-D. Huang, H. Chen, L.-H. Zhou, J. Huang, J. Wu and X.-Q. Yu, *Chem. Biol. Drug Des.*, 2008, **71**, 224; (d) L.-H. Zhou, M. Yang, H. Zhou, J. Zhang, K. Li, Y.-Z. Xiang, N. Wang, Y.-F. Tian and X.-Q. Yu, *Chem. Biol. Drug Des.*, 2009, **73**, 216; (e) J.-L. Liu, Q.-P. Ma, Q.-D. Huang, W.-H. Yang, J. Zhang, J.-Y. Wang, W. Zhu and X.-Q. Yu, *Eur. J. Med. Chem.*, 2011, **46**, 4133; (f) W.-J. Yi, Z.-H. Feng, Q.-F. Zhang, J. Zhang, L.-D. Li, W. Zhu and X.-Q. Yu, *Org. Biomol. Chem.*, 2011, **9**, 2413; (g) Q.-D. Huang, G.-X. Zhong, Y. Zhang, J. Ren, Y. Fu, J. Zhang, W. Zhu and X.-Q. Yu, *PLoS ONE*, 2011, **6**, e23134.
112 (a) F. Sansone, M. Dudič, G. Donofrio, C. Rivetti, L. Baldini, A. Casnati, S. Cellai and R. Ungaro, *J. Am. Chem. Soc.*, 2006, **128**, 14528; (b) R. V. Rodik, A. S. Klymchenko, N. Jain, S. I. Miroshnichenko, L. Richert, V. I. Kalchenko and Y. Mély, *Chem. Eur. J.*, 2011, **17**, 5526.
113 (a) Y. Aoyama, *Bull. Chem. Soc. Jpn.*, 2009, **82**, 419; (b) S. Horiuchi and Y. Aoyama, *J. Control. Release*, 2006, **116**, 107.
114 (a) D. A. Fulton and J. F. Stoddart, *Bioconjugate Chem.*, 2001, **12**, 655; (b) Y. M. Chabre and R. Roy, *Adv. Carbohydr. Chem. Biochem.*, 2010, **63**, 165.
115 (a) Y. Aoyama, *Chem. Eur. J.* 2004, **10**, 588; (b) Y. Aoyama, *Trends Glycosci. Glycotechnol.*, 2005, **17**, 39; (c) Y. Aoyama, *Chem. Eur. J.*, 2009, **82**, 419.
116 Y. Aoyama, Y. Tanaka and S. Sugahara, *J. Am. Chem. Soc.*, 1989, **111**, 5397.
117 O. Hayashida, K. Mizuki, K. Akagi, A. Matsuo, T. Kanamori, T. Nakai, S. Sando and Y. Aoyama, *J. Am. Chem. Soc.*, 2003, **125**, 594.
118 (a) M. Lahmann, *Top. Curr. Chem.*, 2009, **288**, 17; (b) C. Ortiz Mellet, J. Defaye and J. M. García Fernández, *Chem. Eur. J.*, 2002, **8**, 1982.
119 (a) E. M. Nestorovich, V. A. Karginov, M. R. Popoff, S. M. Bezrukov and H. Barth, *PLoS ONE* 2011, **6**, e23927; (b) A. Díaz-Moscoso, A. Méndez-Ardoy, F. Ortega-Caballero, J. M. Benito, C. Ortiz Mellet, J. Defaye, T. M. Robinson, A. Yohannes, V. A. Karginov and J. M. García Fernández, *ChemMedChem*, 2011, **6**, 181; (c) V. A. Karginov, E. M. Nestorovich, M. Moayeri, S. H. Leppla and S. M. Bezrukov, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 15075.
120 C. Ortiz Mellet, J. M. García Fernández and J. M. Benito, *Chem. Soc. Rev.*, 2011, **40**, 1586.
121 S.-A. Cryan, A. Holohan, R. Donohue, R. Darcy and C. M. O'Driscoll, *Eur. J. Pharm. Sci.*, 2004, **21**, 625.

122 (a) N. Mourtzis, K. Eliadou, C. Aggelidou, V. Sophianopoulou, I. M. Mavridis and K. Yannakopoulou, *Org. Biomol. Chem.*, 2007, **5**, 125; (b) N. Mourtzis, M. Paravatou, I. M. Mavridis, M. L. Roberts and K. Yannakopoulou, *Chem. Eur. J.*, 2008, **14**, 4188.
123 S. Srinivasachari, K. M. Fichter and T. M. Reineke, *J. Am. Chem. Soc.*, 2008, **130**, 4618.
124 V. Bennevault-Celton, A. Urbach, O. Martin, C. Pichon, P. Guégan and P. Midoux, *Bioconjugate Chem.*, 2011, **22**, 2404.
125 (a) S.-A. Cryan, R. Donohue, B. J. Ravoo, R. Darcy, C. M. O'Driscoll, *J. Drug. Deliv. Sci. Technol.*, 2004, **14**, 57; (b) A. McMahon, E. Gomez, R. Donohue, D. Forde, R. Darcy and C. M. O'Driscoll, *J. Drug. Deliv. Sci. Technol.*, 2008, **18**, 303
126 A. Díaz-Moscoso, P. Balbuena, M. Gómez-García, C. Ortiz Mellet, J. M. Benito, L. Le Gourriérec, C. Di Giorgio, P. Vierling, A. Mazzaglia, N. Micali, J. Defaye and J. M. García Fernández, *Chem. Commun.*, 2008, 2001.
127 (a) A. K. H. Hirsch, F. R. Fischer and F. Diederich, *Angew. Chem. Int. Ed.*, 2007, **46**, 338; (b) E. A. Katayev, Y. A. Ustynyuk and J. L. Sessler, *Coord. Chem. Rev.*, 2006, **250**, 3004.
128 A. Díaz-Moscoso, L. Le Gourriérec, M. Gómez-García, J. M. Benito, P. Balbuena, F. Ortega-Caballero, N. Guilloteau, C. Di Giorgio, P. Vierling, J. Defaye, C. Ortiz Mellet and J. M. García Fernández, *Chem. Eur. J.*, 2009, **15**, 12871.
129 F. Ortega-Caballero, C. Ortiz Mellet, L. Le Gourriérec, N. Guilloteau, C. Di Giorgio, P. Vierling, J. Defaye and J. M. García Fernández, *Org. Lett.*, 2008, **10**, 5143.
130 A. Megia-Fernandez, M. Ortega-Muñoz, J. Lopez-Jaramillo, F. Hernandez-Mateo and F. Santoyo-Gonzalez, *Adv. Synth. Catal.*, 2010, **352**, 3306.
131 A. Méndez-Ardoy, M. Gómez-García, C. Ortiz Mellet, N. Sevillano, M. D. Girón, R. Salto, F. Santoyo-González and J. M. García Fernández, *Org. Biomol. Chem.*, 2009, **7**, 2681.
132 A. Méndez-Ardoy, N. Guilloteau, C. Di Giorgio, P. Vierling, F. Santoyo-Gonzalez, C. Ortiz Mellet and J. M. García Fernández, *J. Org. Chem.*, 2011, **76**, 5882.
133 A. Méndez-Ardoy, K. Urbiola, C. Aranda, C. Ortiz-Mellet, J. Manuel García-Fernández and C. Tros de Ilarduya, *Nanomedicine*, 2011, **6**, 1697.
134 A. Díaz-Moscoso, D. Vercauteren, J. Rejman, J. M. Benito, C. Ortiz Mellet, S. C. de Smedt and J. M. García Fernández, *J. Control. Release*, 2010, **143**, 318.
135 (a) M. J. O'Neil, J. Guo, C. Byrne, R. Darcy and C. M. O'Driscoll, *Int. J. Pharm.*, 2011, **413**, 174; (b) C. Byrne, F. Sallas, D. K. Rai, J. Ogier and R. Darcy, *Org. Biomol. Chem.*, 2009, **7**, 3763.
136 (a) M. Gómez-García, J. M. Benito, R. Gutiérrez-Gallego, A. Maestre, C. Ortiz Mellet, J. M. García Fernández and J. L. Jiménez Blanco, *Org. Biomol. Chem.*, 2010, **8**, 1849; (b) M. Gómez-García, J. M. Benito, A. P. Butera, C. Ortiz Mellet, J. M. García Fernández and J. L. Jiménez Blanco, *J. Org. Chem.*, 2012, **77**, 1273; (c) J. M. Benito, M. Gómez-García, C. Ortiz Mellet, I. Baussanne, J. Defaye and J. M. García Fernández, *J. Am. Chem. Soc.*, 2004, **126**, 10355; (d) D. Grünstein, M. Maglinao, R. Kikkeri, M. Collot, K. Barylyuk, B. Lepenies, F. Kamena, R. Zenobi and P. H. Seeberger, *J. Am. Chem. Soc.* 2011, **133**, 13957; (e) A. Vargas-Berenguel, F. Ortega-Caballero and J. M. Casas-Solvas, *Mini-Rev. Org. Chem.*, 2007, **4**, 1.
137 (a) R. V. Vico, J. Voskuhl and B. J. Ravoo, *Langmuir*, 2011, **27**, 1391; (b) J. Voskuhl, M. C. A. Stuart and B. J. Ravoo, *Chem. Eur. J.*, 2010, **16**, 2790;

(c) S. McNicholas, A. Rencurosi, L. Lay, A. Mazzaglia, L. Sturiale, M. Perez and R. Darcy, *Biomacromolecules*, 2007, **8**, 1851.

138 A. Díaz-Moscoso, N. Guilloteau, C. Bienvenu, A. Méndez-Ardoy, J. L. Jiménez Blanco, J. M. Benito, L. Le Gourriérec, C. Di Giorgio, P. Vierling, J. Defaye, C. Ortiz Mellet and J. M. García Fernández, *Biomaterials*, 2011, **32**, 7263.

139 (a) N. Symens, A. Méndez-Ardoy, J. M. García Fernández, S. C. De Smedt, and J. Rejman, *Human Gene Ther.*, 2011, **22**, A122; (b) N. Symens, A. Méndez-Ardoy, A. Díaz-Moscoso, E. Sánchez-Fernández, K. Remaut, J. Demeester, J. M. García Fernández, S. C. De Smedt and J. Rejman, *Bioconjugate Chem.*, 2012, DOI: 10.10217bc3001003.

140 M. Lampropoulou and K. Yannakopoulou, *J. Incl. Phenom. Macrocycl. Chem.*, 2011, **70**, 345.

Furanose-based templates in the chemoselective generation of molecular diversity

Ana M. Gómez,* Clara Uriel and J. Cristóbal López*
DOI: 10.1039/9781849734769-00376

Highly functionalized 1-*exo*-alkylidene-2,3-anhydro-, and 1'-halo-1-*exo*-alkylidene-2,3-anhydro-furanoses, which are accessible from D-mannose in four or five steps, respectively, can be efficiently transformed in furanose based templates with up to four sites of molecular diversity, by way of a series of chemoselective transformations at their double bond, oxirane, or allylic oxirane, functionalities.

1 Introduction

One of the hurdles associated to drug discovery is related to the spatial (3D) display of pharmacophoric groups to attain a desired biological response.[1] In this context, carbohydrates were soon recognized as privileged platforms for appending pharmacophores with stereo-determined orientations.[2] Monosaccharides are endowed with rigid core-like characteristics and are available at low cost with a variety of spatial orientations of their hydroxyl substituents. They can also enjoy pyranose or furanose structures, which make them excellent platforms to tailor molecular diversity by appending desired substituents at selected positions. An example is shown in Fig. 1, where a β-D-glucopyranoside with five potential pharmacophoric units, one at each of its five functionalized positions, is displayed.

Based on the concept that a bioactive molecule consists on a display of pharmacophoric units, responsible for binding, assembled around an inert, non-binding, structure acting as a scaffold,[3] Hirschmann, Nicolaou, Smith and co-workers developed, in 1992, a peptidomimetic (**2**, Fig. 2) that was able to interact with somatostatin receptors.[4] In their approach, a β-D-glucopyranose structure functioned as the scaffold to which several pharmacophores were attached. In their landmark work, several β-D-glucopyranose derivatives were shown to mimic the active conformation of the cyclic hexapeptide L-363,301 (**1**), a potent somatostatin receptor agonist. Thus, the sugar template replaces the metabolically labile peptide backbone and serves to control the relative positioning and orientation of the binding side chains.

The usefulness of sugar scaffolds as peptide backbone surrogates continued to be exploited by Kessler and co-workers that developed several peptidomimetics, based on β-D-mannopyranose scaffolds, for the integrin family.[5] Even though most examples of sugar-based scaffolds as peptidomimetics are based on pyranosidic structures, furanoses have also been studied by Hanessian *et al.*[6] They described furanose-based scaffolds as

Instituto de Química Orgánica General, IQOG-CSIC, Juan de la Cierva 3, 28006 Madrid, Spain. E-mail: ana.gomez@csic.es (AMG); jc.lopez@csic.es (JCL)

© The Royal Society of Chemistry 2012

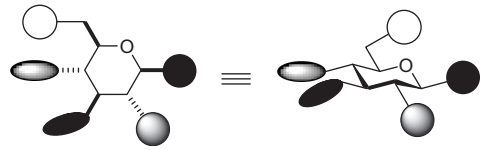


Fig. 1 Spatial arrangement of substituents on a β-D-glucopyranose core.

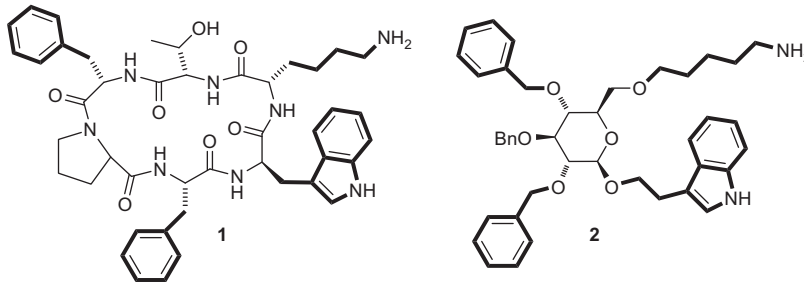


Fig. 2 Hexapeptide L-363,301 (**1**), a potent somatostatin receptor agonist, and a β-D-glucopyranose-based peptidomimetic, **2**.

conformationally constrained analogs of acyclic succinic acid-based inhibitors of metalloproteinases.

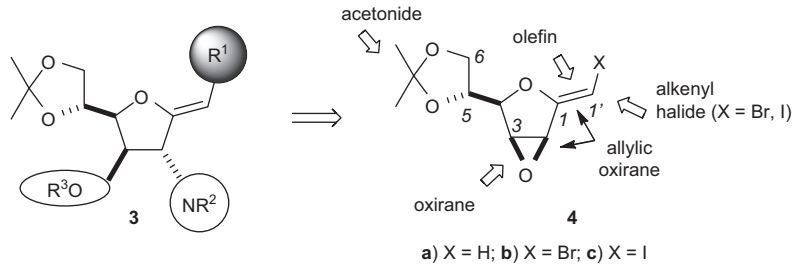
The successful use of carbohydrates as scaffolds in the area of peptidomimetics triggered a considerable research effort on the use of sugar derivatives as templates in bioactive compound discovery.[7] Carbohydrate templates have consequently been generated from pyranoses,[8] furanoses,[9] disaccharides[10] and bicyclic sugar derivatives.[11] Sugar scaffolds also provide an unparalleled opportunity to generate libraries of high functional and structural diversity, with most of the strategies for their construction having been based on the stepwise derivatization of *orthogonally protected* carbohydrate derivatives.[12–14]

2 Design and synthesis of functionalized furanose-based precursors

Based on these literature precedents, we were interested in the design of less-frequented furanose-based templates by chemoselective incorporation of the sought pharmacophoric units on a functionalized furanose-core, rather than a stepwise incorporation of appendages onto an orthogonally differentiated furanose polyol. Our furanose precursor should then be endowed with functionalities other than the current ''hydroxyl groups''. Additionally, in order to make this template appealing, our furanose scaffold should be accessible in few efficient synthetic steps from a readily available starting material.

2.1 Epoxy *exo*-glycals as polyfunctionalized precursors for furanose-based templates

Along these lines, we recently embarked on a project aimed at the synthesis of furanose-based templates **3**, from polyfunctionalized furanose derivatives **4a–c** (Scheme 1). Compound **4a** can be subjected to simple chemical



Scheme 1 Functionalized furanose-based precursors **4** for the preparation of furanose templates **3**.

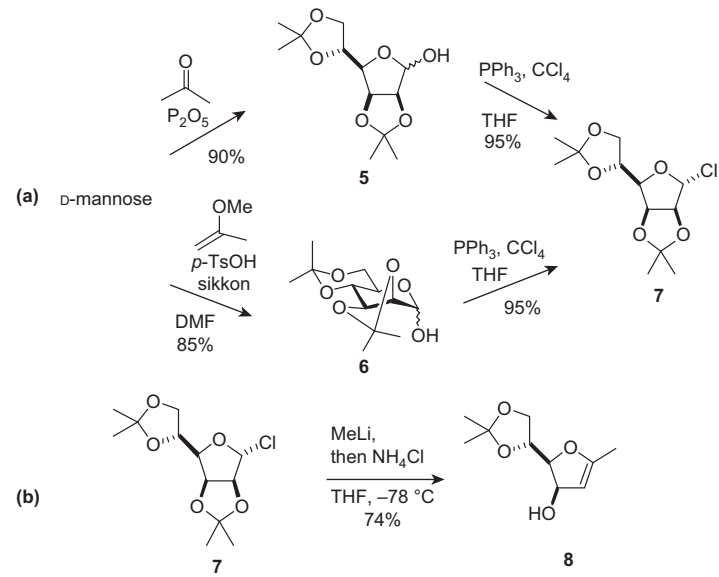
reactions that would lead to diversified platforms, *e.g.* **3**, in which the three different substituents could be incorporated sequentially, and in a completely stereocontrolled manner.

In this context, compound **4a** contains an allylic oxirane moiety, which could be dissected (reactivity-wise) into an oxirane and an exocyclic enolether-type olefin. In derivatives **4b,c**, the exocyclic double bond is further functionalized with the presence of a halogen atom. Particularly appealing to us, was that furanoses **4** could display different sets of reactivity, *e.g.* towards nucleophiles and/or electrophiles. For instance, reaction at the exocyclic olefin in compound **4a** would still leave the oxirane moiety for future transformations, likewise reaction of the oxirane moiety in **4a** would leave the enol ether functionality intact. Besides this postulated reaction pattern, compound **4a** could also display allylic-oxirane reactivity. Furthermore, the presence of a halogen atom in furanoses **4b** or **4c** would still add a supplementary handle for further functionalization at C-1' in these derivatives.

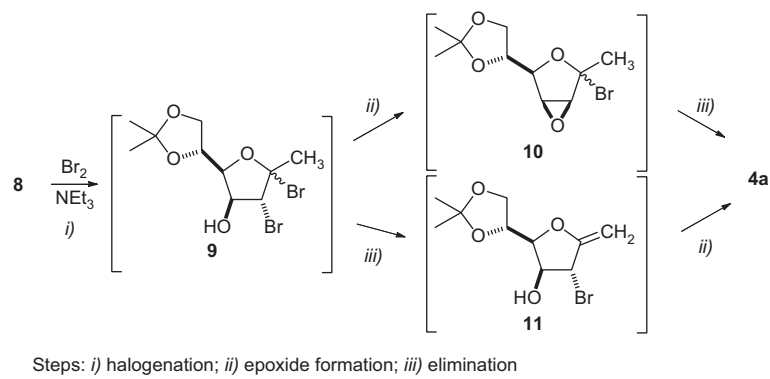
2.2 Synthesis of epoxy-exo glycal 4a, from D-mannose

Our synthetic protocol for the synthesis of compounds **4** took advantage of some transformations described in our laboratory. Thus, reaction of furanosyl chloride **7** with MeLi provided a good route to *C*-1 methyl glycal **8** (Scheme 2b).[15] Access to glycosyl chloride **7** could be done by reaction of either thermodynamic, *i.e.* **5**,[16] or kinetic, *i.e.* **6**,[17] mannose-diacetonides with triphenylphosphane in carbon tetrachloride (Scheme 2a).[18]

The desired epoxy-*exo* glycal **4a** was readily obtained from *C*-1 methyl glycal **8** in just one synthetic operation by treatment with bromine and triethylamine in dichloromethane (**8**→**4a**, Scheme 3).[19] This one-pot transformation involves three widely-known transformations in carbohydrate chemistry, namely: *i)* the Lemieux, Fraser-Reid electrophilic bromination of glycals,[20] *ii)* the base-mediated formation of oxiranes from 1,2-halohydrins,[21] and *iii)* the elimination of anomeric halides mediated by base.[22] Although we have no proof of the actual steps order in the overall transformation, in our opinion, the sequence will likely start with the stereoselective bromination of the electron rich double bond (step *i*, Scheme 3) to furnish dibromides **9**, and since no reaction intermediate (**10** or **11**) has ever been observed (TLC) in these transformations, a pathway **9**→**11**→**4a** (involving a reactive allylic bromide) seems more likely.



Scheme 2 Synthesis of *C*-1 methyl glycal **8**, *via* furanosyl chloride **7**, from D-mannose.

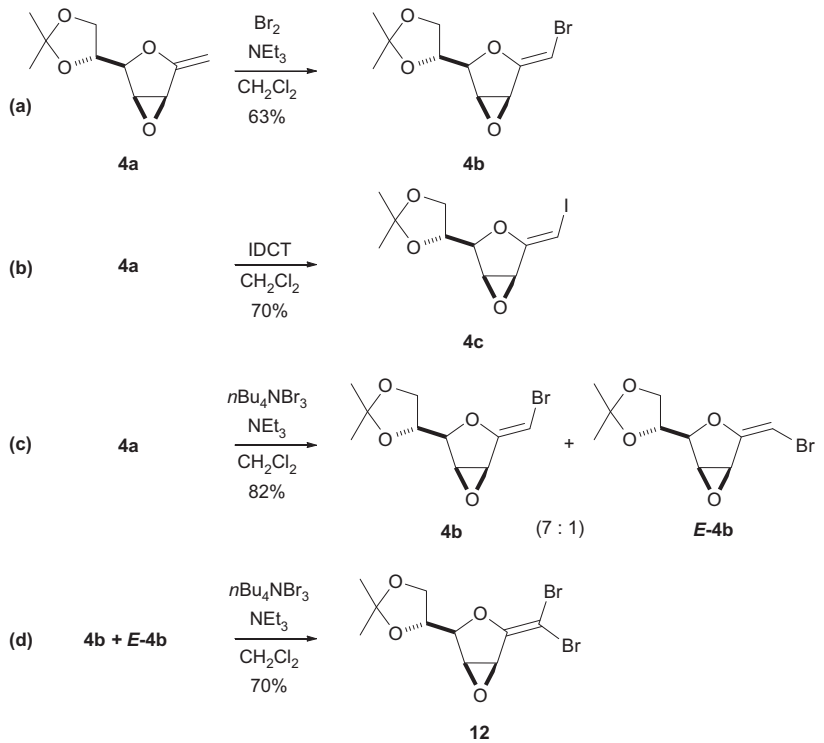


Scheme 3 One-pot transformation of furanosidic *C*-1 methyl glycal **8** into epoxy-exo-glycal **4a**.

2.3 Synthesis of epoxy-1'-halo-*exo* glycal 4b and 4c: Electrophilic halogenation of 4a

Access to halo-*exo*-glycals **4b,c** requires just one synthetic operation starting from **4a**. Accordingly, treatment of the latter with bromine/triethylamine or iodonium dicollidinium triflate (IDCT)[23] furnished halo-*exo*-glycals **4b** and **4c**, in 63% and 70% yields, respectively. The corresponding (*Z*)-derivatives were obtained with complete stereoselectivity (Scheme 4a and 4b).

On the other hand, reaction of **4a** with tetra-*n*-butylammonium tribromide (nBu$_4$NBr$_3$, TBAT)[24] furnished a 7:1 mixture of isomeric (*Z)*- and (*E)*-**4b** (Scheme 4c). The latter mixture could be efficiently transformed into dibromo *exo*-glycal **12**, by further treatment with TBAT (Scheme 4d).The stereochemical outcome of these reactions has already been rationalized[25,26]



Scheme 4 Electrophilic halogenations of *exo*-glycals **4a** and **4b**.

and involves proton abstraction at C-1' followed by base-induced, HBr or HI, elimination from a preferred rotamer, in which the bulky halogen is away from the oxirane ring.

3 Reactivity of functionalized furanose-based precursor 4a

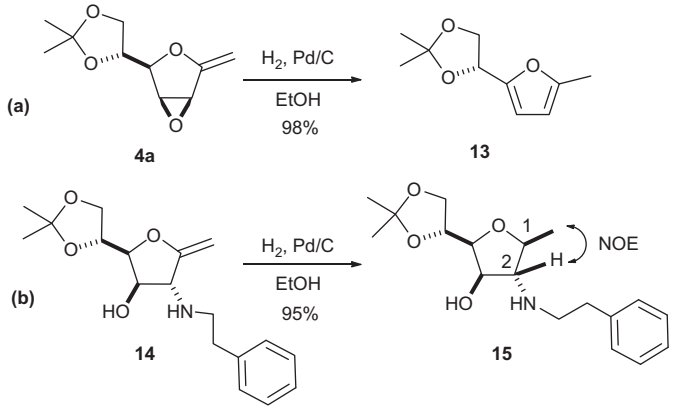
As illustrated in Scheme 1, epoxy-*exo*-glycal **4a** is a polyfunctionalized derivative of potential rich reactivity. In this context, we set up to study its chemistry by first evaluating the chemical behavior of the exocyclic olefin and the oxirane ring as isolated functionalities, to subsequently study the behavior of the vinyl oxirane moiety as a single unit.

3.1 Reactions at the exocyclic olefin

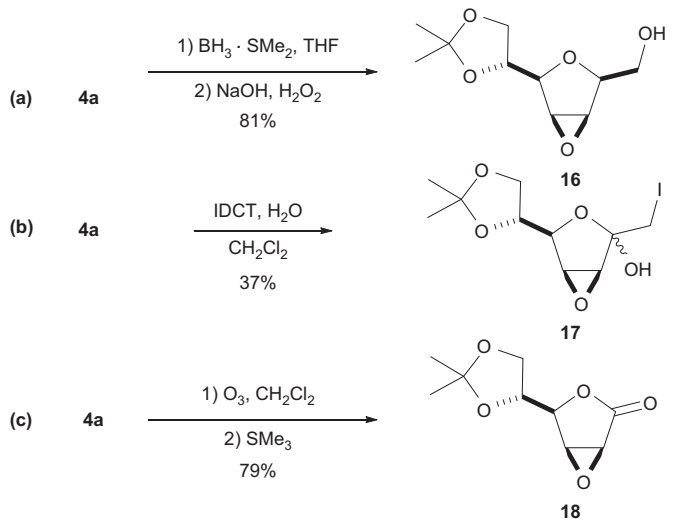
As part of the planned synthesis of **4b,c** and within the context of evaluating the reactivity of the enol ether-type olefin in **4a**, we previously mentioned the electrophilic halogenation of **4a**.

3.1.1 Attempted hydrogenation of the enol ether double bond. The hydrogenation of the exocyclic olefin was also studied since it could grant access to synthetically valuable, functionalized, *C*-glycosyl compounds.[27]

Early attempts at hydrogenation of **4a** (H_2, Pd/C, EtOH, 25 psi) led to the formation of furan derivative **13** (Scheme 5a) under a variety of hydrogenation conditions.[28] These results were in line with previous results from our laboratory that already pointed to the easiness with which aromatization



Scheme 5 Hydrogenation of *exo*-glycals **4a**, and **14**.



Scheme 6 Several transformations of the exocyclic olefin in **4a**.

takes place in these systems.[29] In our opinion the aromatization process is favored by the allylic oxirane moiety. The absence of the oxirane ring in these molecules, *e.g.* **14**, render the hydrogenation process possible (Scheme 5b). The stereochemistry at the newly created stereogenic center in compound **15** was assigned by an observed NOE between H-2 and the methyl group at C-1.

3.1.2 Miscellaneous transformations at the exocyclic double bond. The independent reactivity of the enol ether double bond was further demonstrated by the transformations outlined in Scheme 6.[19a] Thus, hydroboration/oxidation of **4a**, proved to be completely stereoselective to furnish *C*-1 glycoside **16**, in 81% yield (Scheme 6a). Ketose **17** could also be accessed, although in moderate yield, from **4a** by treatment with IDCT in the presence of H_2O (37%, Scheme 6b).[30] Finally, ozonation of **4a** permitted access to synthetically useful epoxy-lactone **18**, in 79% yield (Scheme 6c).

3.2 Reactions at the oxirane ring

Nucleophilic ring opening reactions of furanose derived 2,3-oxiranes are well documented. Lowary and co-workers[31] have studied the regiochemistry of the nucleophilic ring opening of isomeric alkyl 2,3-epoxy furanosides and Hirota and co-workers[32] studied the nucleophilic ring opening of furanosidic allylic oxirane systems, more related to derivatives **4**.

3.2.1 Reaction with oxygen nucleophiles. First, we examined the nucleophilic ring opening of allylic oxiranes **4a–c** with hydroxide anion. The reaction took place smoothly upon treatment with nBu$_4$NOH in refluxing THF/H$_2$O (3:1) to yield the corresponding diols **19a–c**. (or diacetates therefrom, **19b** and **19c**)[19a] in moderate yield with complete regio- and stereoselectivity (Scheme 7).

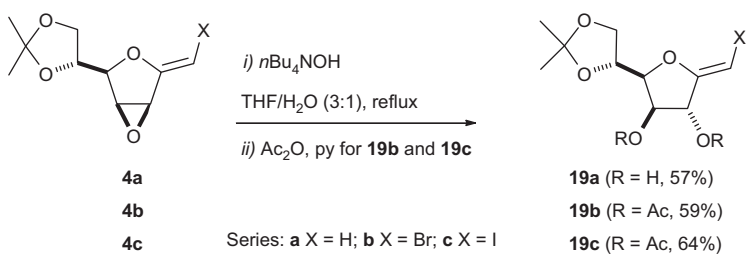
3.2.2 Reaction with amines. The behavior of oxiranes **4** towards amines proved to be similar, and their reactions with a variety of primary and secondary amines in refluxing ethanol provided, in a completely regioselective manner, 2-deoxy-2-amino derivatives in good yields. Displayed in Scheme 8 are examples of the reaction of allylic oxiranes **4a–c** with morpholine (**20**) and 1-(2-methoxyphenyl)piperazine (**21**) to give allylic amines **22a–c** and **23a–c**. In general the reactions were completed in 3–5 h. All the reactions studied took place with complete stereoselectivity by nucleophilic attack at the allylic C-2 position. Deshalo-epoxy-*exo*-glycal **4a** gave better yields of allylic amines than bromo- and iodo-*exo*-glycals **4b** and **4c**, respectively, with the latter giving the lower yields (e.g **22c** and **23c** compared with **22b** and **23b**, Scheme 8).

Compounds **22** and **23** possess allylic amine substructures that have been found in a variety of biologically active compounds.[33]

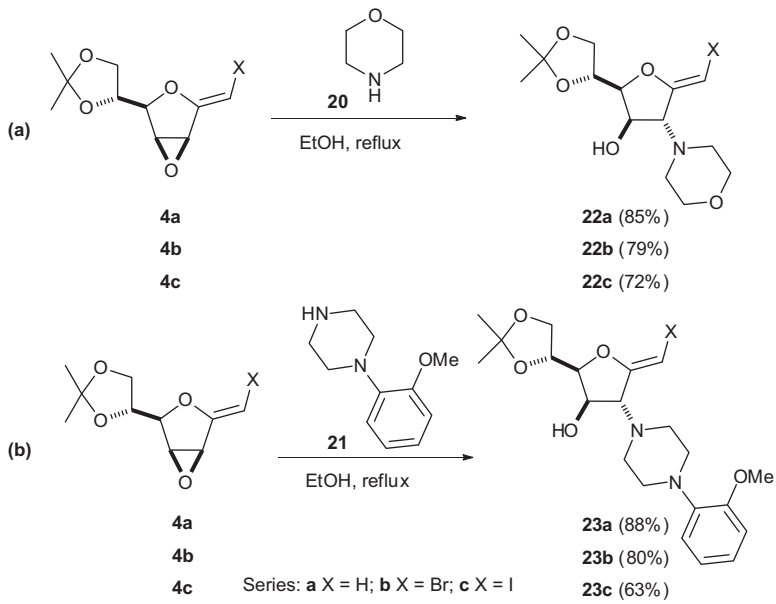
3.3 Pd(0)-Catalyzed reactions at the allylic oxirane moiety

The vinyl oxirane moiety in **4a** made this compound an especially attractive synthon for testing its reactivity in Pd(0)-catalyzed reactions by way of its potential π-allyl palladium intermediate **24**.[34,35] With that in mind, we studied the reaction of **4a** with Pd(0) in the presence of nucleophiles (**4a** → **A**),[19b] electrophiles (**4a** → **B**),[36] and vinyl stannanes (**4a** → **C**)[37] as shown in Scheme 9.

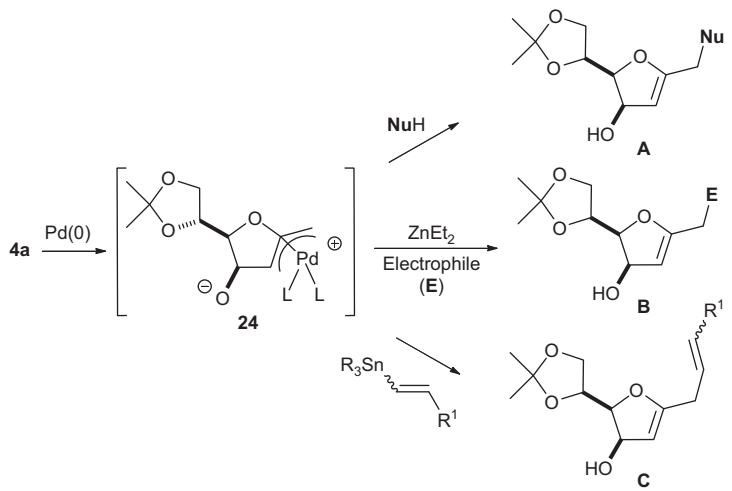
3.3.1 Palladium catalyzed reaction of 4a with nucleophiles. π-Allylpalladium intermediates, *e.g.*, **24**, are known to react with soft carbon



Scheme 7 Reaction of allylic oxiranes **4a–c** with nBu$_4$NOH leading to diols, or diacetates therefrom.

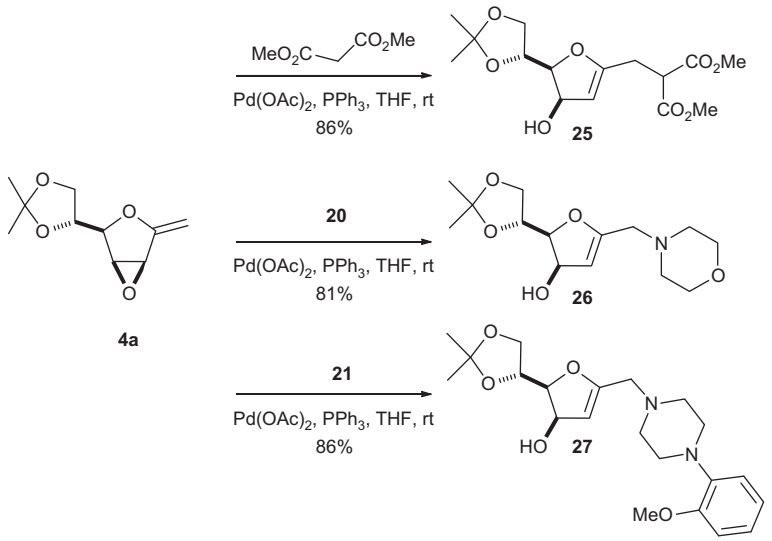


Scheme 8 Reaction of epoxy-*exo*-glycals **4a–c** with amines in refluxing ethanol.

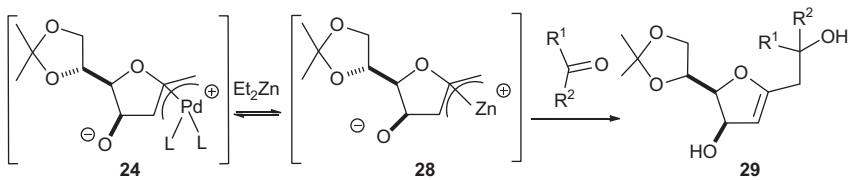


Scheme 9 Proposed Pd(0)-catalyzed reactions of vinyl oxirane **4a** with nucleophiles, electrophiles, and vinylstannanes.

nucleophiles and amines, to give substituted allylic alcohols, *e.g.*, **A** (Scheme 9). In order to test this transformation, we first treated vinyloxirane **4a** with $Pd(OAc)_2$ and PPh_3 in THF at rt, overnight, in the presence of methyl malonate. Compound **25**, which had resulted from 1,4 addition to the vinyloxirane system could then be isolated.[19b] Likewise, treatment of diverse primary and secondary amines yielded allylic derivatives in moderate to good yields, *e.g.* **26**, **27** (Scheme 10).



Scheme 10 Pd(0)-catalyzed reactions of vinyl oxirane **4a** with methyl malonate, morpholine (**20**), and 1-(2-methoxyphenyl)piperazine (**21**).



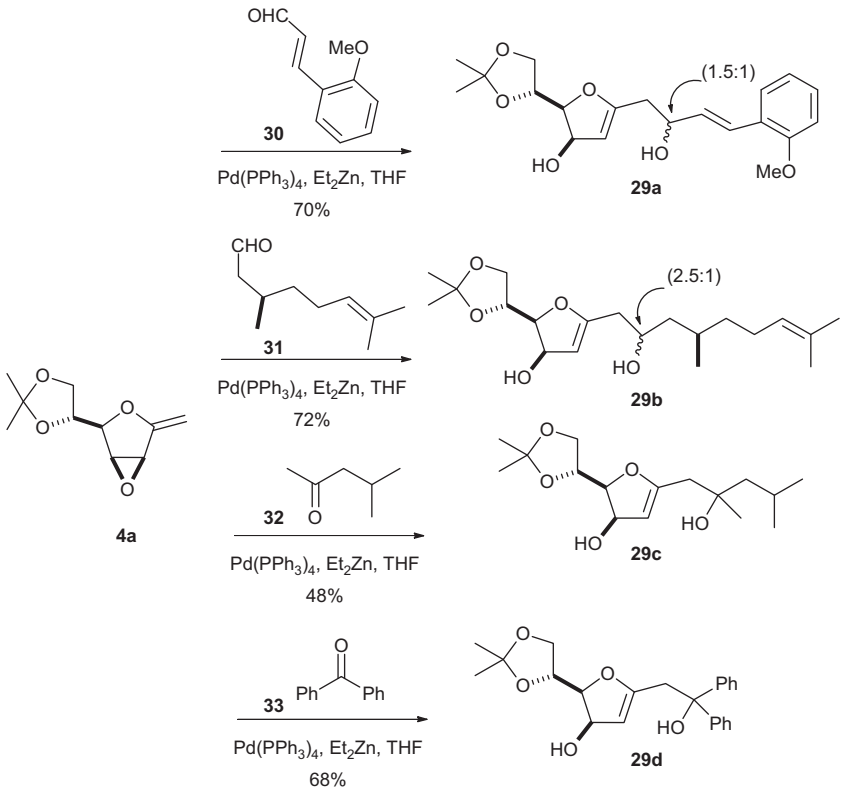
Scheme 11 Proposed reaction pathway for Pd(0)/Et_2Zn-catalyzed reaction between π-allylpalladium species **24** and electrophilic carbonyl compounds.

3.3.2 Palladium catalyzed reaction of 4a with electrophiles. In general, π-allylpalladium complexes are of electrophilic nature, however, several examples of polarity inversion in π-allylpalladium species by transmetallation with low valence metals have appeared.[38] More recently, Et_3B and Et_2Zn have been used in the Pd-catalyzed generation of nucleophilic allyl species from allyl esters and ethers.[39]

In our case, we found that allylzinc derivatives, *e.g.* **28**, readily obtained from π-allylpalladium species **24** by ligand transfer to Et_2Zn, were capable of reacting with aldehydes and ketones to give functionalized *C*-1 glycals **29** (Scheme 11).[40]

From a experimental standpoint the reaction was carried out as follows, a solution containing vinyl epoxide **4a**, $Pd(PPh_3)_4$, and the corresponding carbonyl compound, in THF, at rt, was treated with Et_2Zn. These reactions take place smoothly, and dissapearance of the starting material normally occurred within 1–3 h.

This Pd(0)-catalyzed transformation works well with aldehydes, *e.g.* **30**, **31**, and moderately well with ketones, *e.g.* **32**, **33**, and with excellent 1,4-regioselectivity (Scheme 12). *o*-Methoxycinnamaldehyde (**30**) yielded a 1.5:1 diastereomeric mixture of allylic alcohols **29a** (Scheme 12). A slightly higher



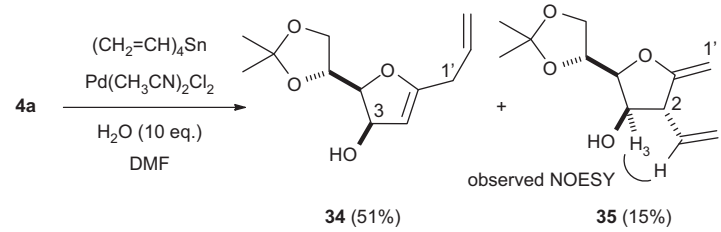
Scheme 12 Reaction of oxirane **4a** with aldehydes and ketones mediated by Pd(0)/Et_2Zn in THF.

diastereomeric ratio (2.5:1) was observed in the reaction of **4a** with citronellal (**31**). Ketones **32** and **33**, gave moderate yields of tertiary hydroxy derivatives **29c** and **29d**.

3.3.3 Palladium catalyzed reaction of 4a with vinyl stannanes. Based on the significant work of Stille's,[41] and Echavarren's[42] groups on the use of stannanes as mild nucleophiles in palladium catalyzed couplings with vinyl epoxides we decided to explore the reactivity of compound **4a**.

Accordingly, the reaction of **4a** with commercially available tetravinyl tin under the conditions recommended by Stille and co–workers (Scheme 13)[41b] took place smoothly and furnished a mixture of regioisomeric derivatives **34** and **35**, resulting from 1,4- and 1,2-addition to the allylic system, respectively. In keeping with literature precedents, compound **35**, generated by 1,2-addition was obtained as the minor isomer (1.4:1 ratio).[37] The structural assignments were based on the study of their ^{1}H- and ^{13}C NMR spectra, and the configuration at C-2 in compound **35**, already inferred by mechanistic reasons, was confirmed by a NOESY experiment that showed the correlation of H-3 with the non-terminal vinylic proton.

The flexibility of this reaction was tested with more complex vinyl stannanes, *e.g.* **36**–**38**, readily available by Et_3B catalyzed radical addition of Bu_3SnH to terminal alkynes (Fig. 3).[43]



Scheme 13 Reaction of vinyloxirane **4a**, with tetravinyl tin catalyzed by $Pd(CH_3CN)_2Cl_2$.

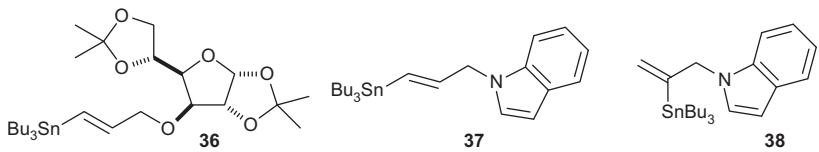


Fig. 3 Vinyl stannanes **36**–**38**, prepared by Et_3B catalyzed radical addition of Bu_3SnH to alkynes.

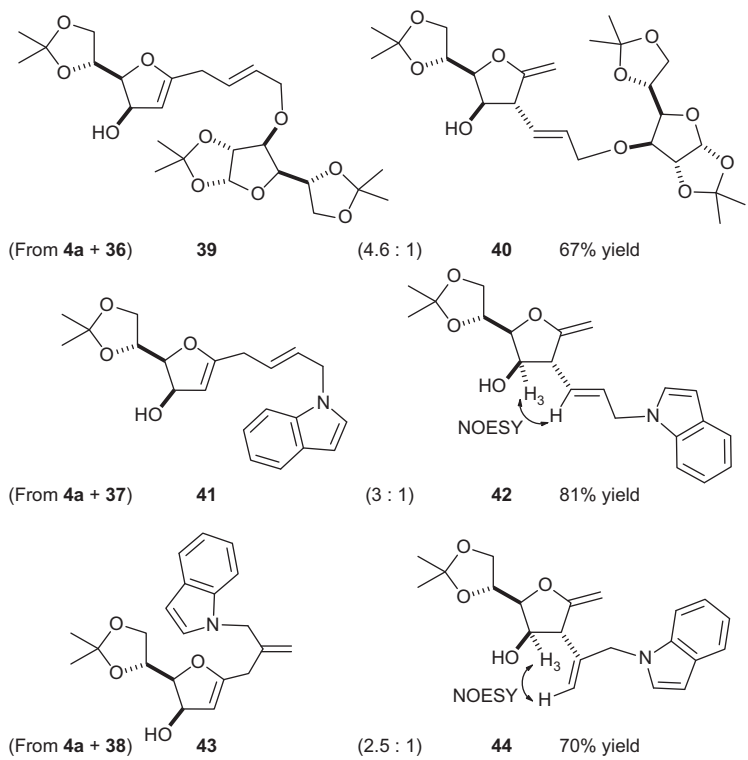


Fig. 4 Reaction products of vinyl oxirane **4a** with vinyl stannanes **36**–**38**.

According to that, alkenyl stannanes **36**–**38** (Fig. 3) react with **4a** under the above mentioned Stille conditions to give the compounds displayed in Fig. 4. In all cases, a mixture of 1,4- and 1,2-regioisomers was obtained. The C-1 glycal derivatives (1,4-isomers) were always obtained as the major

isomers and proved to be very acid sensitive, then leading to the corresponding furan derivatives, however the use of Et_3N (3%) in the chromatography purifications helped to solve that issue.

4 Reactivity of the alkenyl halide moiety in functionalized furanose-based precursors 4b,c

In addition to the functionality present in *exo*-glycal **4a**, compounds **4b** or **4c** possess an alkenyl halide which could function as an additional handle for incorporating molecular diversity at C-1'. In this context we carried out preliminary studies on Pd-catalyzed reactions of structurally related furanose-derived halo-*exo*-glycals **19** (devoid of the oxirane moiety).

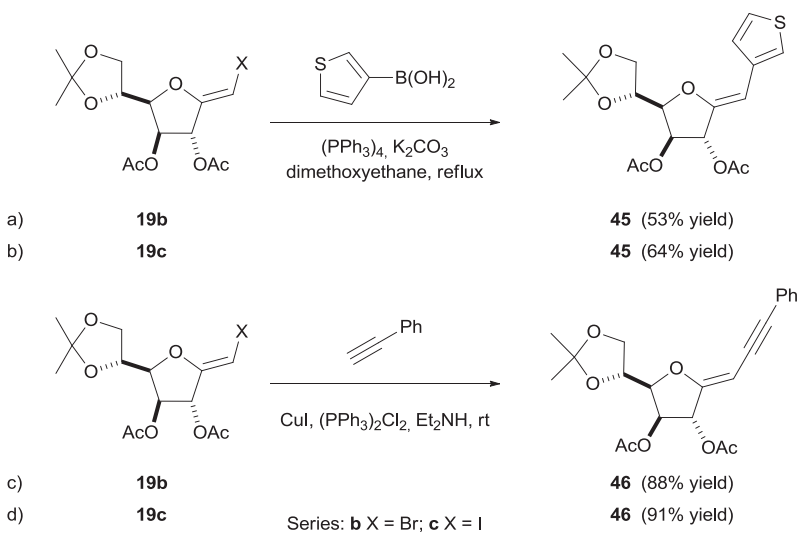
4.1 Cross-coupling reactions of 1'-halo-exo-glycals

We had previously identified the Suzuki-Miyaura[44] and the Sonogashira[45,46] cross-coupling reactions as synthetically useful transformations, and we had shown that they could be performed in furanose 1'-halo-*exo*-glycals, *e.g.* **19b,c**, leading to substituted *exo*-glycals, *e.g.* **45**, **46**, respectively[47,48] (Scheme 14).

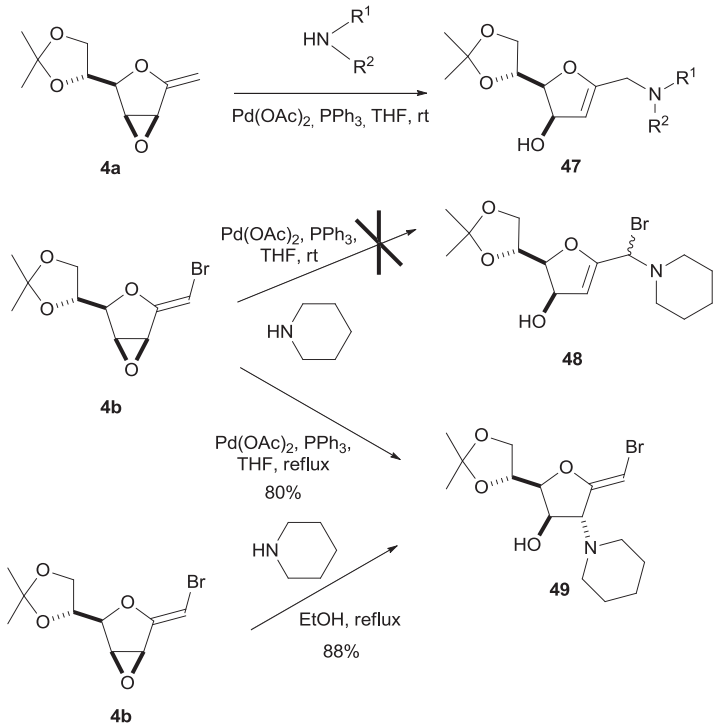
These reactions took place with retention of the configuration at the halide, and led to *exo*-glycals in moderate to good yields. Alkenyl iodides led to higher yields of *exo*-glycals than alkenyl bromides in both Suzuki-Miyaura (Scheme 14a,b) and Sonogashira couplings (Scheme 14c,d).

4.2 Attempted palladium-catalyzed allylic amination of 1'halo-alkenyl derivatives 4b

Furanose *exo*-glycal **4a**, had revealed itself as a valuable substrate, upon treatment with Pd(0), in the preparation of a variety of highly functionalized *C*-1 glycals.



Scheme 14 Suzuki-Miyaura and Sonogashira cross-coupling reactions of 1'halo-exo-glycals **19b,c**.



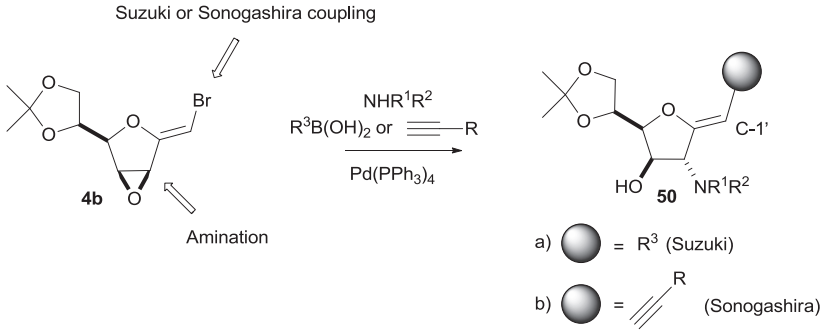
Scheme 15 Reaction of compounds **4a** and **4b** with amines.

In order to compare its behavior, i.e. **4a**→**47** (Scheme 15), in palladium catalyzed aminations *versus* that of 1'halo derivative **4b**, we carried out the reaction of **4b** with piperidine, under experimental conditions similar to those successfully applied to **4a** [$Pd(OAc)_2$, PPh_3, THF, rt (Scheme 15)]. Under these reaction conditions, formation of the expected product **48** was not observed and the starting material was recovered unreacted. It was found, however, that refluxing of the above mentioned solution provided amino-alcohol **49**, in 80% yield (Scheme 15). The fact that the stereochemistry at C-1' was maintained in allylic amine **49**, suggested that no π-allyl palladium intermediate had been involved in the transformation. This was indeed confirmed when a refluxing ethanolic solution containing piperidine and **4b** yielded exclusively allylic amine **49**, in 88% yield (Scheme 15).

These observations seemed to indicate that the reaction taking place was an uncatalyzed nucleophilic opening of the epoxide, and that *the presence of bromine in the vinyl oxirane deactivates the system towards allylic ionization with palladium*.[49–51]

5 Three components assembly of functionalized furanose derivatives 4b,c

From the above mentioned studies, we were aware that: *i)* uncatalyzed amination could be carried out in vinyl oxiranes **4a** **4b** and **4c**, leading to 2-deoxy-2-amino furanose derivatives, *e.g.* **4a–c**→**22a–c** (Scheme 8), and *ii)* that, whereas palladium-catalyzed amination could be performed in deshalo-*exo*-glycal **4a**, it was not effective in alkenyl bromide **4b** (Scheme 15).



Scheme 16 Potential three-components assemblies of 1'halo-epoxy-*exo* glycal **4b**.

With these data in hand, we envisaged that the transformations outlined in Scheme 16 could be feasible, since uncatalyzed nucleophilic opening of the oxirane ring could be followed by cross-coupling reaction of the alkenyl halide, or vice versa, cross-coupling on the alkenyl bromide could be followed by (uncatalyzed) amination.

5.1 One-pot multicomponent amination/Suzuki cross-coupling of 4b

We found that the treatment of **4b** with $Pd(PPh_3)_4$, NaOH, in THF:H_2O, at 60 °C, allowed its multicomponent assembly with boronic acids and amines to give access to a library of compounds with three sites for molecular diversity.

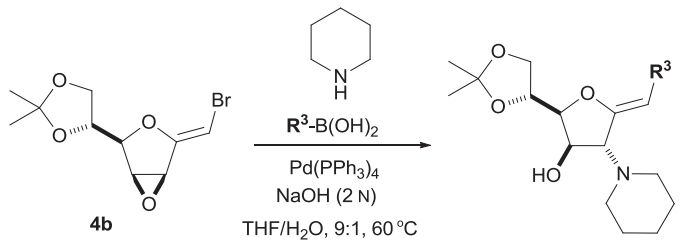
In all instances where a comparison was possible, the one-pot transformation was more efficient than the alternative stepwise synthesis. This transformation could be applied to a variety of aryl and heteroaryl boronic acids (Table 1).

Alkyl boranes have also been used in B-alkyl Suzuki cross-coupling reactions.[52] So, the methodology could also be applied to the coupling of **4b** with alkyl boranes and secondary amines, as shown in Table 2. The reaction of vinyl bromide **4b** with diethyl(3-pyridyl)borane and different cyclic amines led to derivatives **55**–**57**, in good yields.

Unlike the success achieved with the three component reactions involving Suzuki cross-couplings (Tables 1 and 2), the Sonogashira variant did not lead to the desired products. For instance, instead of the desired amino alkynes **50b** (Scheme 16), the reaction of compound **4b** with phenyl acetylene, or trimethylsilylacetylene [alkyne (1.1 eq.), $Pd(PPh_3)_4$ (5%), CuI (10%), Et_2NH (4 mL/mmol), rt] produced a mixture of 2-deoxy enynes **58** [1.2:1, unassigned (*Z*:*E*) mixture] and **59** [2:1, unassigned (*Z*:*E*) mixture], respectively (Table 3, entries *i*, *ii*). Likewise, when dodec-1-yne was employed, a ≈1:1 mixture of enynes **60** (37%) was obtained, although this time accompanied by 2-deoxy-2-amino derivative **61** (17%) (Table 3, entry *iii*). These results seemed to indicate that the Sonogashira cross-coupling was taking place first and it was followed by the formation of a π-allyl palladium complex responsible for the formation of compounds **58**–**60**, as geometric enyne mixtures.[†] Furthermore, the addition of hydride

[†]Will no π-allyl palladium complex be involved, compound **30** would have appeared as a single isomer, owing to the retention of configuration associated with the Sonogashira cross-coupling reaction of alkenyl halides.

Table 1 One-pot multicomponent assembly of 1'-alkenyl *exo*-glycal **4b** with piperidine and boronic acids leading to a library of functionalized allylic amines.



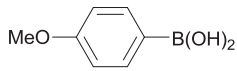
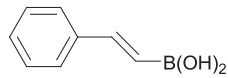
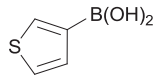
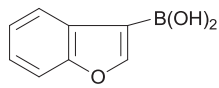
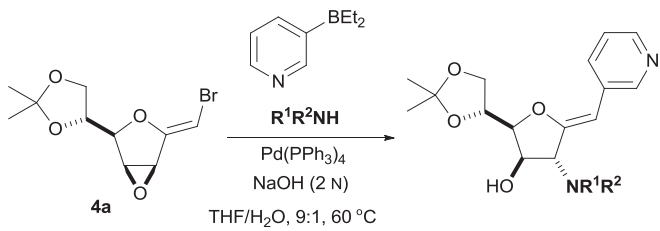
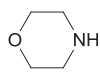
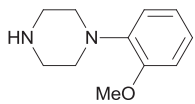
entry	R^3-$B(OH)_2$	product	yield (%)
i		**51**	63
ii		**52**	60
iii		**53**	72
iv		**54**	75

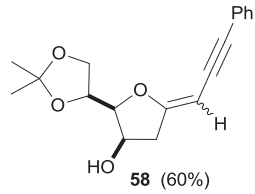
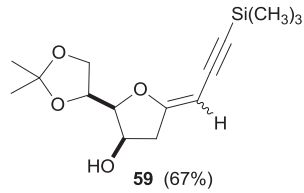
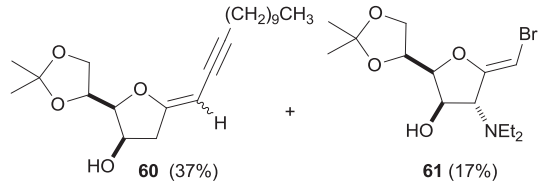
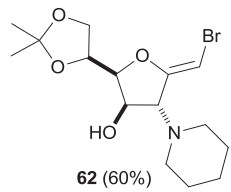
Table 2 One-pot multicomponent assembly of furanose derivatives **55**–**57**, by reaction of alkenyl bromide **4a**, with diethyl(3-pyridyl)borane and amines.



entry	R^1R^2NH	product	yield (%)
i		**55**	78
ii		**56**	71
iii		**57**	74

from a palladium-hydride species to the π-allyl palladium could be responsible for the formation of the 2-deoxy, rather than 2-deoxy-2-amino derivatives.[53]

Table 3 Sonogashira cross-coupling reaction of epoxy bromo-*exo*-glycal **4b**, with phenyl acetylene, trimethylsilylacetylene, and dodec-1-yne, in the presence of amines.

Entry	Reaction conditions	Products (yield)
i	Ph $Pd(PPh_3)_4$, CuI Et_2NH, rt	 **58** (60%)
ii	$Si(CH_3)_3$ $Pd(PPh_3)_4$, CuI Et_2NH, rt	 **59** (67%)
iii	$(CH_2)_9CH_3$ $Pd(PPh_3)_4$, CuI Et_2NH, rt	 **60** (37%) + **61** (17%)
iv	Ph $Pd(PPh_3)_4$, CuI piperidine, 60 °C	 **62** (60%)

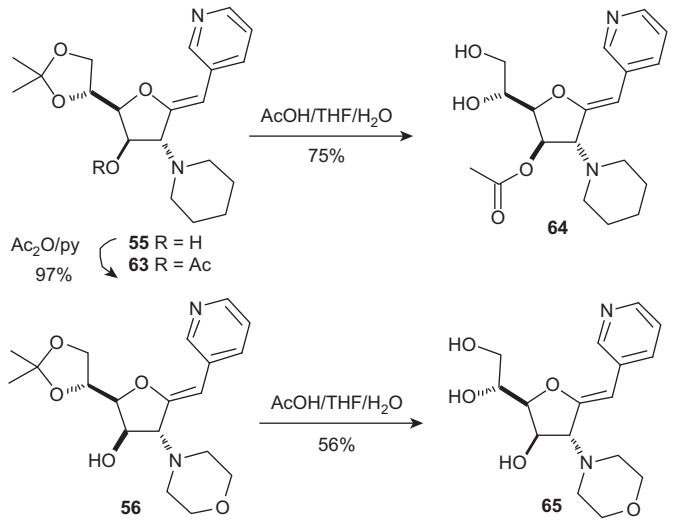
5.2 Attempted one-pot multicomponent amination/Sonogashira cross-coupling[48b]

The presence of an unreacted bromine atom in amino alcohol **61** (Table 3, entry *iii*) was unexpected. Therefore, we carried out several attemps to conduct Sonogashira couplings with **61** and phenylacetylene that were not successful. When the reaction was carried out in the presence of piperidine at 60 °C (Table 3, entry *iv*), only amino alcohol **62** could be isolated.

6 Generation of furanosidic libraries possessing three – or more – sites for molecular diversity

The furanose derivatives **51**–**54** and **55**–**57**, (Tables 1 and 2, respectively) resulting from oxirane opening followed by palladium catalyzed Suzuki cross-coupling, incorporate already two potential pharmacophores. In addition, the presence of the 3-OH group and the 5,6-*O*-isopropylidene acetal makes possible further derivatization of these derivatives.

In order to illustrate this point, amino alcohols **55** and **56** can be transformed into **63**, **64** and **65** (Scheme 17). Thus, laying the bases for the synthesis of additional furanose-based libraries.



Scheme 17 Furanose derivatives **64** and **65**, with more than three points of diversity.

Acetylation of **55** led to acetate **63** with three points of diversity. Interestingly, the enol ether double bond in **63** displayed an enhanced stability towards acid, and consequently the 5,6-*O*-isopropylidene acetal group could be removed without tampering with the anomeric functionality, leading to diol **64**.

These compounds already possess more than three sites for structural diversity, and the 5,6-diol could serve as a handle for subsequent incorporation of molecular diversity.

Likewise, treatment of amino alcohol **56** with AcOH:THF:H_2O (4:2:1) under reflux, yielded triol **65**.

7 Conclusions

Highly functionalized *exo*-glycal **4a** is a useful synthetic intermediate in the generation of molecular diversity in a chemoselective manner. In this report, we have shown that it could be successfully used in the preparation of furanose-based libraries with three or more sites for molecular diversity.

Compound **4a** contains an allylic epoxide moiety that can undergo chemoselective reactions either at the double bond, at the oxirane ring, or at the vinyl epoxide system.

Halogenation of the double bond gives rise to synthetically useful 1'-haloepoxy-*exo*-glycals (**4b,c**). Hydrogenation of **4a** led to the rapid formation of furan derivatives, which could be circumvented by performing the hydrogenation on *exo*-glycals devoid of the oxirane ring. The hydrogenation, then, takes place in a stereocontrolled manner, probably due to the presence of the bulky β-oriented 5,6-*O*-isopropylidene ring.

The oxirane moiety undergoes ring opening with nitrogen nucleophiles or hydroxide in a completely regioselective manner, leading to allylic amino alcohols and allylic diols, respectively.

Nucleophilic ring opening of the oxirane ring in 1'-halo-epoxy-*exo*-glycals leaves the vinyl halide functionality unreacted, so that it can be engaged in Suzuki cross-coupling reactions with boronic acids, thus allowing the generation of a library of furanose-based allylic amines. These amines still bear a OH-3, and a 5,6-isopropylidene acetal ready for further transformations. Acylation can then be performed at O-3, to incorporate a third substituent. Remarkably, these derivatives undergo acid-mediated chemoselective 5,6-*O*-isopropylidene ring opening without aromatization to the corresponding furan derivatives.[21] Thus diols, *e.g.* **64**, and triols, *e.g.* **65**, ready for further derivatization can be readily prepared.

The rich chemistry of these derivatives still merits further scrutiny, and in our opinion further transformation can still be uncovered allowing access to templates with up to five sites of molecular diversity.

Abbreviations

DMF	Dimethylformamide
h	Hour
IDCT	Iodonium dicollidinium triflate
NOE	Nuclear overhauser effect
rt	Room temperature
TBAT	Tetra-*n*-butylammonium tribromide
THF	Tetrahydrofuran

Acknowledgements

This research has been supported with funds from the *Ministerio de Ciencia e Innovación* grants CTQ-2006-C03 and CTQ2009-10343, and *Comunidad de Madrid* grant S2009/PPQ-1752. Generous financial support from Janssen-Cilag S.A. is also acknowledged. Additional support came as predoctoral scholarships to Ana Pedregosa (Janssen-Cilag S.A) and Aitor Barrio [*Consejo Superior de Investigaciones Científicas* (*CSIC*)] who were intensely involved in the development of this research.

References

1 *(a)* W. Meutermans, G. T. Le, and B. Becker, *Chem. Med. Chem.*, 2006, **1**, 1164. *(b)* G. T. Le, G. Abbenante, B. Becker, M. Grathwohl, J. Halliday, G. Tometzki, J. Zuegg and W. Meutermans, *Drug Discovery Today*, 2003, **8**, 701.

2 *(a)* L. Cipolla and F. Peri, *Mini-Rev. Med. Chem.*, 2011, **11**, 39. *(b)* L. Cipolla, A. C. Araujo, D. Bini, L. Gabrielli, L. Russo and N. Shaikh, *Expert Opin. Drug Discov.*, 2010, **5**, 723. *(c)* G. Abbenante, B. Becker, S. Blanc, C. Clark, G. Condie, G. Fraser, M. Grathwohl, J. Halliday, S. Henderson, A. Lam, L. Liu, M. Mann, C. Muldoon, A. Pearson, R. Premraj, T. Ramsdale, T. Rossetti, K. Schafer, G. L. Thanh, G. Tometzki, F. Vari, G. Verquin, J. Waanders, M. West, N. Wimmer, A. Yau, J. Zuegg and W. Meutermans, *J. Med. Chem.*, 2010, **53**, 5576. *(d)* L. Cipolla, F. Peri, B. La Ferla, C. Redaelli and F. Nicotra, *Curr. Org. Synth.*, 2005, **2**, 153.

3 B. Becker, G. C. Condie, G. T. Le and W. Meutermans, *Mini-Rev. Med. Chem.*, 2006, **6**, 1299.

4 *(a)* R. Hirschmann, K. C. Nicolaou, S. Pietranico, J. Salvino, E. M. Leahy, P. A. Sprengeler, G. Furst, A. B. Smith III, C. D. Strader, M. A. Cascieri, M. R. Candelore, C. Donaldson, W. Vale and L. Maechler, *J. Am. Chem. Soc*., 1992, **114**, 9217. *(b)* R. Hirschmann, K. C. Nicolaou, S. Pietranico, E. M. Leahy, J. Salvino, B. Arison, M. A. Cichy, P. G. Spoors, W. C. Shakespeare, P. A. Sprengeler, P. Hamley, A. B. Smith III T. Reisine, K. Raynor, L. Maechler, C. Donaldson, W. Vale, R. M. Freidinger, M. A. Cascieri and C. D. De Strader, *J. Am. Chem. Soc*., 1993, **115**, 12550. *(c)* R. Hirschmann, W. Q. Yao, M. A. Cascieri, C. D. Strader, L. Maechler, M. A. Cichy-Knight, J. Hynes, Jr, R. D. van Rijn, P. A. Sprengeler and A. B. Smith III, *J. Med. Chem*., 1996, **39**, 2441. *(d)* R. F. Hirschmann, K. C. Nicolaou, A. R. Angeles, J. S. Chen, and A. B. Smith III, *Acc. Chem. Res*., 2009, **42**, 1511.

5 J. Boer, D. Gottschling, A. Schuster, B. Holzmann and H. Kessler, *Angew. Chem. Int. Ed*., 2001, **40**, 3870.

6 S. Hanessian, N. Moitessier and S. Wilmouth, *Tetrahedron*, 2000, **56**, 7643.

7 For reviews on carbohydrate-derived structures as templates, see: *(a)* Y. M. Chabre and R. Roy, in: *The Sugar Code. Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, Germany, 2009, p. 53. *(b)* P. V. Murphy and T. Velasco-Torrijos, in: *Glycoscience – Chemistry and Chemical Biology*, 2nd ed. (Eds.: B. Fraser-Reid, K. Tatsuta and J. Thiem), Springer-Verlag, Berlin-Heidelberg, 2008; p. 997. *(c)* P. V. Murphy, *Eur. J. Org. Chem*., 2007, 4177. *(d)* I. Velter, B. La Ferla and F. Nicotra, *J. Carbohydr. Chem*., 2006, **25**, 97. *(e)* P. V. Murphy and J. L. Dunne, *Curr. Org. Synth*., 2006, **3**, 403. *(f)* T. Kanemitsu and O. Kanie, *Trends Glycosci. Glycotechnol*., 1999, **11**, 267. *(g)* M. J. Sofia and D. J. Silva, *Curr. Opin. Drug Discovery Dev*., 1999, **2**, 365.

8 Pyranose-based scaffolds: *(a)* N. Elmouelhi, U. Aich, V. D. P. Paruchuri, M. A. Meledeo, C. T. Campbell, J. J. Wang, R. Srinivas, H. S. Khanna and K. J. Yarema, *J. Med. Chem*., 2009, **52**, 2515. (b) C. T. Campbell, U. Aich, C. A. Weier, J. J. Wang, S. S. Choi, M. M. Wen, K. Maisel, S.-G. Sampathkumar and K. J. Yarema, *J. Med. Chem*., 2008, **51**, 8135. *(c)* Y. Tamaruya, M. Suzuki, G. Kamura, M. Kanai, K. Hama, K. Shimizu, J. Aoki, H. Arai and M. Shibasaki, *Angew. Chem. Int. Ed*., 2004, **43**, 2834. *(d)* N. Moitessier, S. Dufour, F. Chretien, J. P. Thiery, B. Maigret and Y. Chapleur, *Bioorg. Med. Chem*., 2001, **9**, 511. *(e)* P. V. Murphy, J. L. O'Brien, L. J. Gorey-Feret and A. B. Smith III, *Bioorg. Med. Chem. Lett*., 2002, **12**, 1763. *(f)* L. Cipolla, C. Redaelli, F. Granucci, G. Zampella, A. Zaza, R. Chisci and F. Nicotra, *Carbohydr. Res*., 2010, **345**, 1291. *(g)* C. Rosenbohm, D. Van den Berghe, A. Vlietinck and J. Wengel, *Tetrahedron*, 2001, **57**, 6277.

9 Furanose-based scaffolds: *(a)* C. Papageorgiou, R. Haltiner, C. Bruns and T. J. Petcher, *Bioorg. Med. Chem. Lett*., 1992, **2**, 135. *(b)* T. K. Chakraborty, S. Ghosh, M. H. V. R. Rao, A. C. Kunwar, H. Chob and A. K. Ghosh, *Tetrahedron Lett*., 2000, **41**, 10121. *(c)* E. B. Krueger, T. P. Hopkins, M. T. Keaney, M. A. Walters and A. M. Boldi, *J. Comb. Chem*., 2002, **4**, 229. (e) A. Novoa, N. Pellegrini-Moïse, D. Bechet, M. Barberi-Heyob and Y. Chapleur, *Bioorg. Med. Chem*., 2010, **18**, 3285. (f) A. A. Edwards, O. Ichihara, S. Murfin, R. Wilkes, M. Whittaker, D. J. Watkin and G. W. J. Fleet, J. *Comb. Chem*., 2004, **6**, 230.

10 S. Castoldi, M. Cravini, F. Micheli, E. Piga, G. Russo, P. Seneci and L. Lay, *Eur. J. Org. Chem*., 2004, 2853.

11 Sugar-based bicyclic scaffolds: *(a)* F. Peri, R. Bassetti, E. Caneva, L. De Gioia, B. La Ferla, M. Presta, E. Tanghetti and F. Nicotra, *J. Chem. Soc., Perkin Trans*. I, 2002, **5**, 638. *(b)* A. C. Araujo, F. Nicotra, B. Costa, G. Giagnoni and L. Cipolla, *Carbohydr. Res*., 2008, **343**, 1840. (c) A. C. Araujo, F. Nicotra, C.

Airoldi, B. Costa, G. Giagnoni, P. Fumagalli and L. Cipolla, *Eur. J. Org. Chem.*, 2008, 635. *(d)* G. Cervi, F. Peri, C. Battistini, C. Gennari and F. Nicotra, *Bioorg. Med. Chem.*, 2006, **14**, 3349. *(e)* A. Cordeiro, E. Quesada, M.-C. Bonache, S. Velazquez, M.-J. Camarasa and A. San-Felix, *J. Org. Chem.*, 2006, **71**, 7224. *(f)* G. Capozzi, S. Giannini, S. Menichetti, C. Nativi, A. Giolitti, R. Patacchini, E. Perrotta, M. Altamurac and C. A. Maggic, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 2263. *(g)* J. Jimenez-Barbero, E. Dragoni, C. Venturi, F. Nannucci, A. Arda, M. Fontanella, S. Andre, F. J. Cañada, H.-J. Gabius and C. Nativi, *Chem. Eur. J.*, 2009, **15**, 10423. (h) S. Mari, F. J. Cañada, J. Jimenez-Barbero, A. Bernardi, G. Marcou, I. Motto, I. Velter, F. Nicotra and B. La Ferla, *Eur. J. Org. Chem.*, 2006, 2925.

12 *(a)* H. Hünger, J. Ohnsmann and H. Kunz, *Angew. Chem. Int. Ed.*, 2004, **43**, 1104. *(b)* T. Opatz, C. Kallus, T. Wunberg, W. Schmidt, S. Henke and H. Kunz, *Eur. J. Org. Chem.*, 2003, 1527. *(c)* T. Wunberg, C. Kallus, T. Opatz, S. Henke, W. Schmidt and H. Kunz, *Angew. Chem. Int. Ed.*, 1998, **37**, 2503.

13 *(a)* R. Jain, M. Kamau, C. Wang, R. Ippolito, H. Wang, R. Dulina, J. Anderson, D. Gange and M. J. Sofia, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 2185. *(b)* M. J. Sofia, R. Hunter, T. Y. Chan, A. Vaughan, R. Dulina, H. Wang and D. Gange, *J. Org. Chem.*, 1998, **63**, 2802.

14 N. Moitessier, C. Henry, N. Aubert and Y. Chapleur, *Tetrahedron Lett.*, 2005, **46**, 6191.

15 *(a)* A. M. Gómez, A. Pedregosa, M. Casillas, C. Uriel and J. C. López, *Eur. J. Org. Chem.*, 2009, 3579. *(b)* A. M. Gómez, M. Casillas, S. Valverde and J. C. López, *J. Chem. Soc., Chem. Commun.*, 1996, 2357.

16 A. Basha, R. Henry, M. A. McLaughlin, J. D. Ratajczyk and S. J. Wittenberger, *J. Org. Chem.*, 1994, **59**, 6103.

17 J. Gelas and D. Horton, *Carbohydr. Res.*, 1978, **67**, 371.

18 R. E. Ireland, S. Thaisrivongs, N. Vanier and C. S. Wilcox, *J. Org. Chem.*, 1980, **45**, 48.

19 *(a)* A. M. Gómez, A. Pedregosa, C. Uriel, S. Valverde and J. C. López, *Eur. J. Org. Chem.*, 2010, 5619. *(b)* A. M. Gómez, A. Pedregosa, S. Valverde and J. C. López, *Chem. Commun.*, 2002, 2022.

20 R. U. Lemieux and B. Fraser-Reid, *Can. J. Chem.*, 1965, **43**, 1460.

21 *(a)* P. Crotti, V. Di Busolo, L. Favero, F. Macchia and M. Pineschi, *Eur. J. Org. Chem.*, 1998, 1675. *(b)* P. Herdewijn, K. Ruf and W. Pfleiderer, *Helv. Chim. Acta* 1991, **74**, 7. *(c)* B. Fraser-Reid, L. Magdzinski, B. Molino and D. R. Mootoo, *J. Org. Chem.*, 1987, **52**, 4495. *(d)* B. Fraser-Reid, L. Magdzinski and B. Molino, *J. Am. Chem. Soc.*, 1984, **106**, 731.

22 J.-P. Praly, G.-R. Chen, J. Gola and G. Hetzer, *Eur. J. Org. Chem.*, 2000, 2831.

23 *(a)* G. H. Veeneman, S. H. van Leeuwen, H. Zuurmond and J. H. van Boom, *J. Carbohydr. Chem.*, 1990, **9**, 783. (b) D. Noort, G. H. Veeneman, G.-J. Boons, G. A. van der Marel, G. J. Mulder and J. H. van Boom, *Synlett*, 1990, 205.

24 G. Bellucci, C. Chiappe and G. Lo Moro, *J. Org. Chem.*, 1997, **62**, 3176.

25 *(a)* M. Lakhrissi, G. Carchon, T. Schlama, C. Mioskowski and Y. Chapleur *Tetrahedron Lett.*, 1998, **39**, 6453. (b) W.-B. Yang, Y.-Y. Yang, Y.-F. Gu, S.-H. Wang, C.-C. Chang and C.-H. Lin, *J. Org. Chem.*, 2002, **67**, 3773.

26 A. M. Gómez, A. Pedregosa, S. Valverde and J. C. López, *Tetrahedron Lett.*, 2003, **44**, 6111.

27 *(a)* X.-P. He, X.-W. Wang, X.-P. Jin, H. Zhou, X.-X. Shi, G.-R. Chen and Y.-T. Long, *J. Am. Chem. Soc.*, 2011, **133**, 3649. *(b)* S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony and B. G. Davis, *Nature*, 2007, **446**, 1105. *(c)* M. A. Fischbach, H. Lin, D. R. Liu and C. T. Walsh, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 571.

28 A. M. Gómez, A. Pedregosa, C. Uriel, S. Valverde and J. C. López, *Eur. J. Org. Chem.*, 2010, 5619.
29 A. M. Gómez, M. Casillas, B. Rodríguez, S. Valverde and J. C. López, *ARKIVOC*, 2010, **iii**, 288.
30 For reactions of *exo*-glycals with iodonium ions, see: a) L. Lay, F. Nicotra, L. Panza and G. Russo, *Synlett*, 1995, 167. b) A. Haudrechy and P. Sinaÿ, *Carbohydr. Res.*, 1991, **216**, 375.
31 Ring opening of alkyl 2,3-epoxy furanosides, see: *(a)* D. Hou and T. L. Lowary, *Org. Lett.*, 2007, **9**, 4487. *(b)* Y. Bai and T. L. Lowary, *J. Org. Chem.*, 2006, **71**, 9658. *(c)* R. R. Gadikota, C. S. Callam, T. Wagner, B. Del Fraino and T. L. Lowary, *J. Am. Chem. Soc.*, 2003, **125**, 4155. *(d)* C. S. Callam, R. R. Gadikota, D. M. Krein and T. L. Lowary, *J. Am. Chem. Soc.*, 2003, **125**, 13112. *(e)* C. S. Callam, R. R. Gadikota and T. L. Lowary, *Synlett*, 2003, 1271.
32 Ring opening of furanose-derived allylic epoxides, see: *(a)* H. Sajiki, H. Takasu and K. Hirota, *Heterocycles* 2006, **67**, 399. *(b)* H. Takasu, Y. Tsuji, H. Sajiki and K. Hirota, *Tetrahedron*, 2005, **61**, 11027. *(c)* H. Takasu, Y. Tsuji, H. Sajiki and K. Hirota, *Tetrahedron*, 2005, **61**, 8499. *(d)* K. Hirota, H. Takasu, Y. Tsuji and H. Sajiki, *Chem. Commun.*, 1999, 1827.
33 V. Nesterenko, K. S. Putt and P. J. Hergenrother, *J. Am. Chem. Soc.*, 2003, **125**, 14672.
34 *(a)* J. Tsuji, H. Kataoka and Y. Kobayashi, *Tetrahedron Lett.*, 1981, **22**, 2575. *(b)* B. M. Trost and G. A. Molander, *J. Am. Chem. Soc.*, 1981, **103**, 5969.
35 B. M. Trost and D. L. Van Kranken, *Chem. Rev.*, 1996, **96**, 395.
36 A. M. Gómez, A. Barrio, A. Pedregosa, S. Valverde and J. C. López, *Tetrahedron Lett.*, 2003, **44**, 8433.
37 A. M. Gómez, A. Pedregosa, A. Barrio, S. Valverde and J. C. López, *Eur. J. Org. Chem.*, 2009, 4627.
38 These reactions include treatment with *chromium*: O. Fujiwara, K. Takai and K. Utimoto, *J. Org. Chem.*, 1990, **55**, 1705; *tin*: Y. Masuyama, J. Nakata and Y. Kurusu, *J. Chem. Soc., Perkin Trans. 1*, 1991, 2598; *samarium*: J. M. Aurrecoechea and E. Iztueta, *Tetrahedron Lett.*, 1995, **36**, 7129; and *indium*: S. Araki, K. Kameda, J. Tanaka, T. Hirashita, H. Yamamura and M. Kawai, *J. Org. Chem.*, 2001, **66**, 7919; S. Araki, T. Kamei, T. Hirashita, H. Yamamura and M. Kawai, *Org. Lett.*, 2000, **2**, 847; T. Tabuchi, J. Inanaga and M. Yamaguchi, *Tetrahedron Lett.*, 1986, **27**, 1195.
39 Review: Y. Tamaru, *Eur. J. Org. Chem.*, 2005, 2647, and references cited therein.
40 A. M. Gomez and J. C. Lopez, *Specialist Periodical Reports*, *Carbohydrate Chemistry*, 2009, **35**, 290. DOI:10.1039/B901578G.
41 *(a)* A. M. Echavarren, D. R. Tueting and J. K. Stille, *J. Am. Chem. Soc.*, 1988, **110**, 4039. *(b)* D. R. Tueting, A. M. Echavarren and J. K. Stille, *Tetrahedron*, 1989, **45**, 979.
42 *(a)* A. M. Castaño, M. Ruano and A. M. Echavarren, *Tetrahedron Lett.*, 1996, **37**, 6591. *(b)* A. M. Castaño, M. Méndez, M. Ruano and A. M. Echavarren, *J. Org. Chem.*, 2001, **66**, 589.
43 *(a)* K. Nozaki, K. Oshima and K. Utimoto, *J. Am. Chem. Soc.*, 1987, **109**, 2547. *(b)* K. Nozaki, K. Oshima and K. Utimoto, *Tetrahedron*, 1989, **45**, 923.
44 *(a)* N. Miyaura and A. Suzuki, *Chem. Rev.*, 1995, **95**, 2457. *(b)* A. Suzuki, *J. Organomet. Chem.*, 1999, **576**, 147.
45 K. Sonogashira, *J. Organomet. Chem.*, 2002, **653**, 46, and references cited therein.
46 Review: R. Chinchilla and C. Nájera, *Chem. Rev.*, 2007, **107**, 874.

47 Suzuki-Miyaura cross-coupling: A. M. Gómez, G. O. Danelon, A. Pedregosa, S. Valverde and J. C. López, *Chem. Commun*., 2002, 2024.
48 Sonogashira cross-coupling: *(a)* A. M. Gómez, A. Pedregosa, A. Barrio, S. Valverde and J. C. López, *Tetrahedron Lett*., 2004, **45**, 6307. *(b)* A. M. Gómez, A. Barrio, A. Pedregosa, C. Uriel, S. Valverde and J. C. López, *Eur. J. Org. Chem*., 2010, 2910.
49 For studies on allylic ionization versus oxidative addition into vinyl C-Br bonds, see: *(a)* E. Comer, M. G. Organ and S. J. Hynes, *J. Am. Chem. Soc*., 2004, **126**, 16087. *(b)* M. G. Organ, E. A. Arvanitis, C. E. Dixon and J. T. Cooper, *J. Am. Chem. Soc*., 2002, **124**, 1288.
50 For studies on 1-acetoxy-2-bromo-2-alkenes: *(a)* G. C. Nwokogu, *J. Org. Chem*., 1985, **50**, 3900. *(b)* G. C. Nwokogu, *Tetrahedron Lett*., 1984, **25**, 3263. *(c)* B. M. Trost and J. D. Oslob, *J. Am. Chem. Soc*., 1999, **121**, 3057.
51 For studies on α,β-epoxy vinyl triflates: K. Yamamoto and C. H. Heathcock, *Org. Lett*., 2000, **2**, 1709.
52 S. R. Chemler, D. Trauner and S. J. Danishefsky, *Angew. Chem. Int. Ed*., 2001, **40**, 4544.
53 Some related examples have been described: R. Takeuchi and K. Tanabe, *Angew. Chem. Int. Ed*., 2000, **39**, 1975.

Synthesis of carbohydrate-based artificial siderophores and their biological applications

Marta M. Andrade* and Amélia P. Rauter
DOI: 10.1039/9781849734769-00398

During the last decade a series of synthetic siderophores based on monosaccharides has emerged. Such compounds bear ligands of the catecholate and hydroxamate (normal and retro) family and were built around methyl α-D-glucopyranoside, methyl α-D-galactopyranoside, methyl α-D-ribopyranoside and methyl α-D-xylopyranoside. Studies in Gram-negative bacteria and mycobacteria have proven that they can act as effective artificial siderophores.

1 Introduction

Iron is an essential element for microorganisms, plants and animals due to its exclusive chemical properties, such as the ability to coordinate and activate oxygen and the possession of ideal redox chemistry for involvement in electron transport and metabolic processes. Although it is one of the most abundant elements on the planet, iron(III) solubility at neutral pH is around 10^{-10}M, a value that is four orders of magnitude lower than the concentration required by microorganisms and therefore it cannot be utilized by them.[1,2] Consequently, microorganisms, fungi and plants have evolved strategies to scavenge and absorb iron from soil, fresh and marine water, and living organisms (plants and animals), by producing high-affinity iron-binding compounds known as siderophores, capable of chelating iron by allowing its assimilation through cell surface receptors. Siderophores are low-molecular-weight compounds (500–1500 dalton) which can dissolve these ions as soluble Fe^{3+} octahedral complexes with high stability constants ($\log \beta > 30$) that can be taken up by active transport mechanisms (Table 1).[1,3–5] Up to date, there are almost 500 compounds identified as siderophores.[1]

Siderophore research stimulates a multidisciplinary effort in order to facilitate detailed understanding of the essential role of iron in microbial virulence as well as the potential for the development of effective antimicrobial agents.[6]

These compounds can be generally classified according to their iron binding moieties as catechol-based (enterobactin, agrobactin), hydroxamic acid-based (ferrichrome, ferrioxamine B), alpha-hydroxycarboxylic acid-based (vibrioferrin) and mixed-ligand siderophores (aerobactin) (Fig. 1). Their architecture includes linear, tripodal, endocyclic and exocyclic

Faculdade de Ciências da Universidade de Lisboa, Carbohydrate Chemistry Group, Centre of Chemistry and Biochemistry/Department of Chemistry and Biochemistry, Edifício C8, Piso 5, Campo Grande, 1749-016 Lisboa, Portugal. E-mail: marta.m.s.andrade@gmail.com

© The Royal Society of Chemistry 2012

Table 1 Relevant thermodynamic parameters for selected siderophore and synthetic siderophore complexes.[1]

Siderophore	Log β[a]	pFe[b]
Enterobactin	49	35.5
Pyoverdin	30.8	27
Ferrichrome A	32.0	25.2
Ferrioxamine E	32.5	27.7
Ferrioxamine B	30.6	26.6
Aerobactin	22.5	23.3
Alcaligin	64.6	23.0
Rhodotorulic acid	62.3	21.8
Acetohydroxamic acid	28.29	12.5
N-Methylacetohydroxamic acid	29.4	16.2
L-Lysinehydroxamic acid	16.1	7.1
Tiron	46.3	20.94

[a] Overall stability constant at 25 °C for complexation of $Fe(H_2O)_6^{3+}$ by the fully deprotonated ligand; represents β_{110} for the hexadentate ligands, β_{230} for the tetradentate ligands and β_{130} for bidentate ligands.
[b] $-\log$[uncomplexed $Fe^{3+}aq$] at pH 7.4, 1 μM total Fe^{3+} and 10 μM total ligand.

structures. The most common natural siderophores are enterobactin, ferrichrome, coprogen, among others.[5,7–9]

The siderophore excreted by *Streptomyces pilosus*, deferrioxiamine B (**3**, Fig. 1 and Fig. 2), is commercialized as a mesylated ester named Desferal®, and it was for many years the single drug available in the market for iron removal in iron overload diseases. However, its retention time in the body is quite short; its oral activity is very poor, the treatment is expensive and has several collateral effects, illustrating the importance in developing new analogs of the drug. Deferiprone was introduced in the late 80s; it is orally active but does not remove enough iron to keep patients in a negative iron balance (Fig. 2). In Nov 2005, Deferasirox (Fig. 2) was approved by the FDA; it is also orally active, has a long-half-life time and some minor side effects.[10] Also a wide range of catecholates have appeared as possible drugs for iron removal.[11,12]

Several studies of artificial siderophores have been reported during the last two decades.[3,5,13–20] These compounds attracted growing interest as clinically useful iron chelators for the treatment of iron overload diseases and for iron transport-mediated antibiotics or antimicrobial compounds;[5] they are also useful antimalarial agents, acting by causing iron deprivation of malaria parasites.[21] The most common moieties used in artificial siderophores are catecholate and hydroxamate groups. The stereochemistry of these molecules is such that the coordination sphere of iron(III) is completely occupied by the oxygen-containing ligands, resulting in octahedral complexes (Fig. 3).[18]

The development of artificial siderophores that can act as functional model compounds for native siderophores is an essential tool for the investigation of siderophore mechanisms of recognition and metabolism. Molecular recognition and signalling are at the heart of biological processes, by governing intercellular communication and information transfer. The recognition phenomena of siderophore mediated microbial iron(III) uptake involves two processes: recognition of iron(III) by microbial

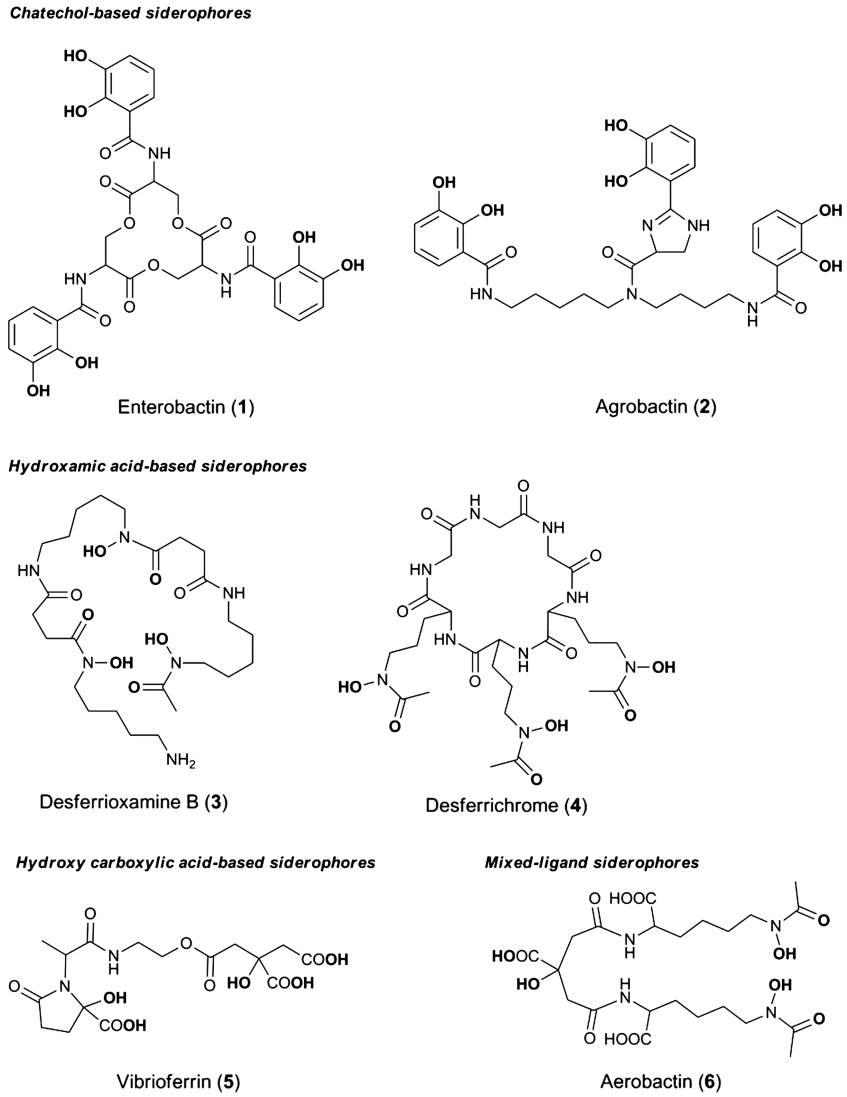
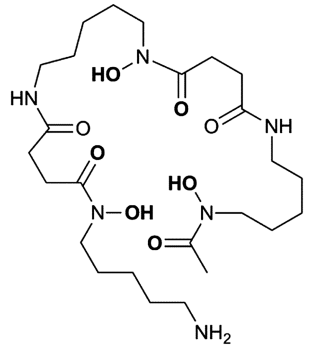


Fig. 1 Examples of siderophores classified according to their iron binding moieties (all shown without iron). The atoms in bold represent the binding sites.

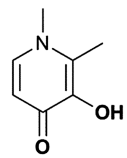
siderophores as iron(III) carriers and recognition and binding of siderophore-iron(III) complexes by microbial membrane receptors. Consequently, many efforts are nowadays devoted to the design and synthesis of less toxic analogs and, in this context, the use of siderophores with carbohydrate skeleton is a promising subject to be exploited.

2 Artificial siderophores - background

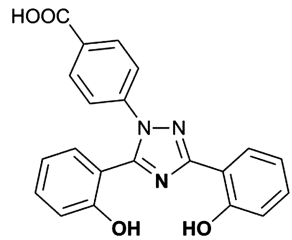
Exocyclic and tripodal siderophore mimics have been successfully built around tertiary amine,[17,19,22] benzene,[23,24] tetrahedral carbon,[25–28] lysine derivatives[20,29] or cyclitols,[30–32] maintaining the chelating capacities but sacrificing its water solubility and the receptor-recognition capability of the



Desferrioxamine B

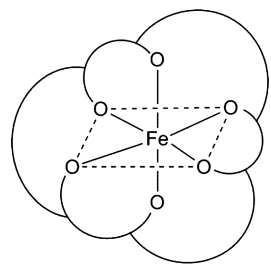


Deferiprone



Deferasirox

Fig. 2 Drugs used nowadays for the treatment of iron overload diseases (binding atoms in bold).



Hexadentate ligand
1:1 complex

Fig. 3 Schematic representation of chelate ring formation for Fe(III)/siderophore complexes.[18]

backbone. Thermodynamically, these synthetic analogues have Fe(III) chelating capabilities that are comparable if not greater than those of their natural counterparts. Inomata *et al*. have also demonstrated that the bulky terminal groups and intramolecular hydrogen bonds on artificial siderophores play important roles in recognition and permeation.[28]

Raymond *et al*.[33] reported the total synthesis of naturally occurring siderophores and Neilands and Miller[34,35] have synthesized various artificial siderophores with binding constants greater than 10^{20} M^{-1}; however, larger binding constants have been obtained by Vögtle and coworkers.[36] More recently, analogues of siderophores with cyclic monosaccharide backbone

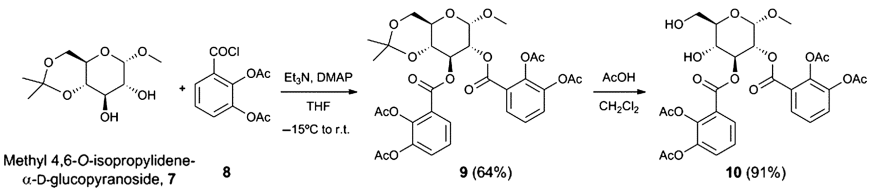
were synthesized and characterized, showing Fe(III) chelating capabilities that are comparable, if not greater than those of their natural counterparts; biologically, they are highly promising synthetic analogues, as *in vivo* studies show significant siderophore activity indicating favorable cell receptor recognition and cellular uptake.[13–16,37,38] The design of siderophore mimics requires some important features to be considered, which include the nature of donor groups, denticity, chelate ring size, architecture, hydrophobicity/hydrophilicity and stability. The availability of different stereoisomeric forms of saccharides and their high functionality render them particularly attractive for the conception of siderophore model compounds of varying structure and polarity and the construction of an optimal octahedral binding cavity for the Fe(III) ion. The chiral saccharide backbone can provide favorable recognition and accessibility to cells, high hydrolytic stability and increased hydrophilicity. The presence of oxygen atoms in and around the saccharide ring gives the cyclic backbone the ability to act as a hydrogen-bond-acceptor during the cell receptor-recognition process. Such compounds and their physicochemical properties are of interest as they can be used as penetration vectors for antibiotics or other biological applications involving Fe(III) uptake and metabolism. These results suggest that the carbohydrate scaffold is a promising biologically viable backbone for siderophore analogues.

3 Synthesis of Carbohydrate-based artificial siderophores

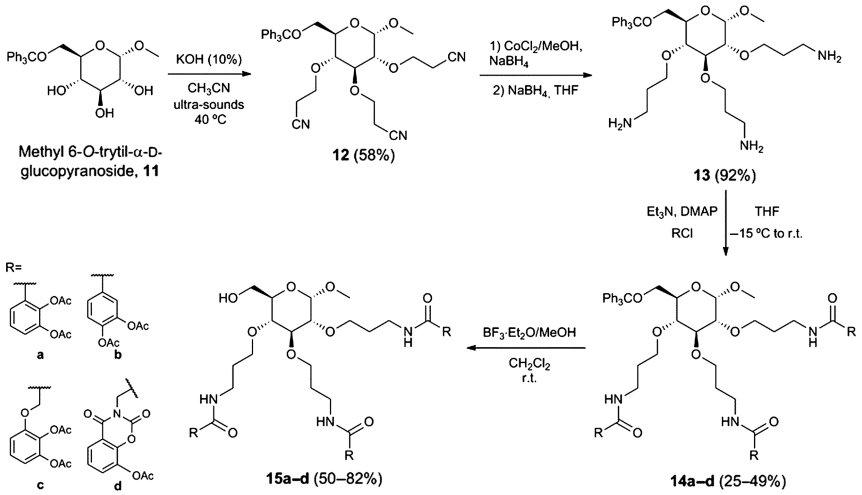
3.1 Catecholate derivatives

Catechol moieties show high affinity for iron(III) as they interact strongly with tripositive metal cations, resulting in high electron density of both oxygen atoms. The binding of cations by catechol has marked pH sensitivity, and the complexes forming at pH 7.0 each bear a net charge and consequently are unlikely to permeate membranes by simple diffusion, thus tending to be trapped in intracellular compartments. A disadvantage for catechol-based ligands is their susceptibility towards oxidation.[18]

The most relevant natural catecholate siderophore is enterobactin (**1**, Fig. 1), which is produced by *E. coli* and other enteric bacteria under iron limitation. Studies indicate that the chiral trilactone backbone and the spacer chain length are to some extent responsible for the high stability of the Fe(III)-enterobactin complex and are critical in the receptor mediated transport and iron release mechanism. In 2001, Heggemann *et al*. have shown that carbohydrates can substitute advantageously the trilactone structure of enterobactin.[14,16] In the last decade, the synthesis of carbohydrate-based artificial siderophores was focused only on monosaccharide scaffolds. The group synthesized several glucose derivatives with two or three 2,3-di(acetoxy)benzoyl ligands at secondary carbon positions of the glucopyranose ring and also with three 2,3-di(acetoxy)phenoxyacetyl and 8-acetoxy-2-oxo-3,4-dihydro-2*H*-benzo[1,3]oxazin-3-yl)acetyl ligands at the same positions. The synthesis of biscatecholate derivatives started from methyl 4,6-*O*-isopropylidene-α-D-glucopyranoside (**7**), which has two free hydroxyl groups ready to react with 2,3-di(acetoxy)benzoyl chloride (**8**) in basic conditions, using THF as solvent. Posterior removal of isopropylidene protecting group afforded biscatecholate **10** with overall yield of 58% (Scheme 1).



Scheme 1 Synthesis of bis-catecholate derivatives with di(acetoxy)benzoyl ligands at carbon positions 2 and 3.[16]



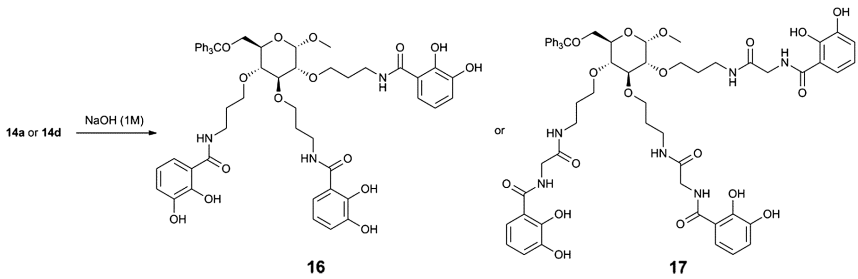
Scheme 2 Synthesis of triscatecholate derivatives with different ligands at carbon positions 2, 3 and 4.[16]

The approach to obtain trischatecolate siderophores started from 6-*O*-trityl protected methyl α-D-glucopyranoside derivative (**11**), which has three free hydroxyl groups at carbon positions 2, 3 and 4. A spacer was now introduced between the sugar and the ligand *via* a Michael type cyanoethylation with acrylonitrile, followed by reduction of the nitrile with $CoCl_2/NaBH_4$ to afford ligand **13** (Scheme 2). Different chatechol-based chelating groups were then introduced in the molecule, *e.g.* by reaction with the acyl chlorides **14a–d**. Final removal of the trityl group with boron trifluoride etherate afforded the triscatecholate siderophores **15a–d**. Compounds **14a** and **14d** were also deacetylated with NaOH to afford **16** and **17** with free catechol-OH groups (Scheme 3).

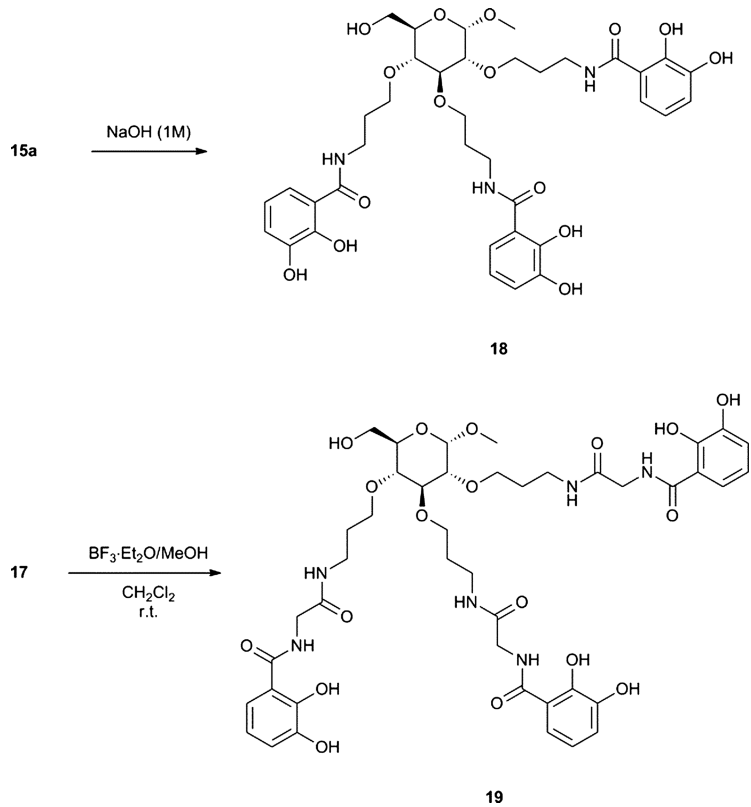
Also Dhungana *et al.*[14] have reported free-catecholate derivatives, *e.g.* starting from **15a** and **17** (Scheme 4).

Another series of triscatecholates based on methyl α-D-glucopyranoside and methyl α-D-galactopyranoside (Schemes 5 and 6) was also reported using synthetic pathways similar to those previously described.[15]

New dihydroxybenzylidenehydrazino derivatives based on methyl α-D-glucopyranoside by replacing the aminopropyl groups by hydrazinocarbonylethyl functional groups have been reported as a new variation of the carbohydrate siderophores (Scheme 7).



Scheme 3 Synthesis of a triscatecholate derivative with free catechol-OH groups.[16]

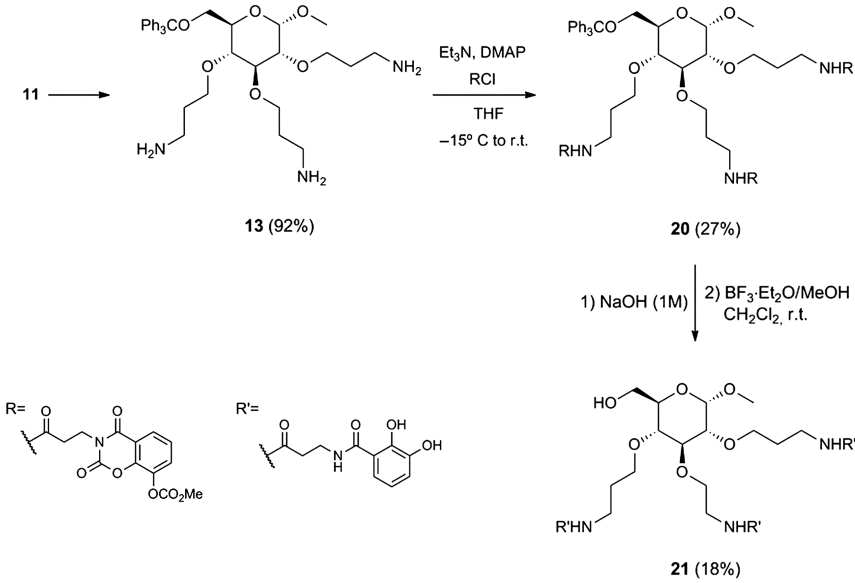


Scheme 4 Triscatecholate derivatives with free catechol-OH groups.[14]

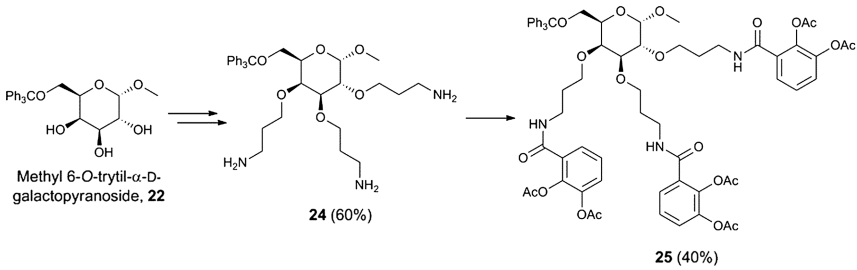
3.2 Retro-hydroxamate and hydroxamate derivatives

Hydroxamate ligands possess lower affinity for iron than catecholates but they have the advantage of forming neutral tris-complexes with iron(III), which are in principle capable of permeating membranes by nonfacilitated diffusion. Many hydroxamates are metabolically labile and are only poorly absorbed *via* the oral route.[18]

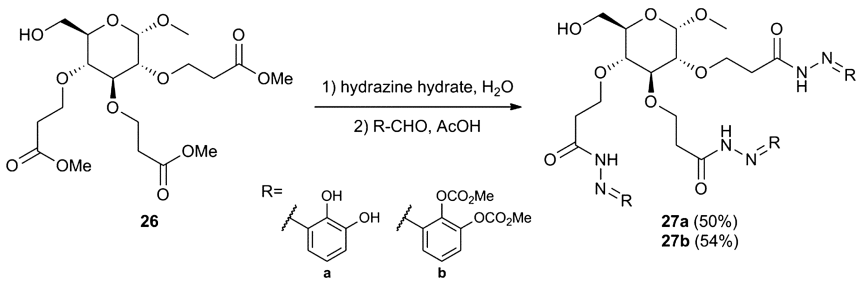
Desferrioxamine B (**3**, Fig. 1) is a hydroxamate-type siderephore used in medicine to treat potentially fatal iron overload resulting from regular blood transfusion of patients with thalassaemia, a genetic blood disease



Scheme 5 Synthesis of new triscatecholate derivatives with methyl α-D-glucopyranoside backbone.[15]



Scheme 6 Synthesis of a triscatecholate derivative with methyl α-D-galactopyranoside backbone.[15]



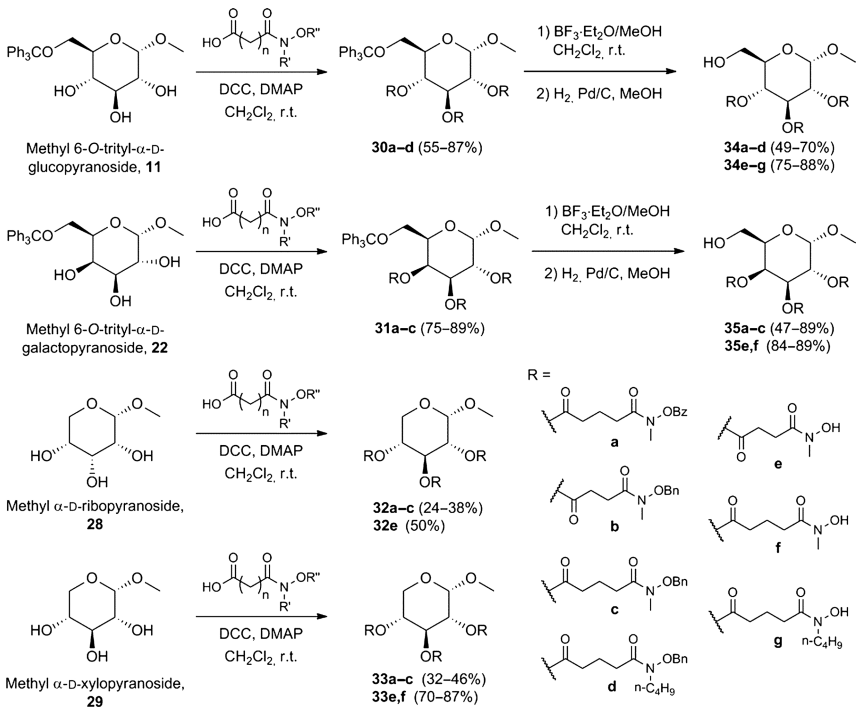
Scheme 7 Synthesis of dihydroxybenzylidenehydrazinocatecholates based on methyl α-D-glucopyranoside.[15]

characterised by the decreased production of normal haemoglobin and resulting in anaemia. Desferichrome (**4**, Fig. 1) is another example of a natural tris-hydroxamate siderophore with an exocyclic structure; its iron (III) complex, called ferrichrome was one of the first siderophores

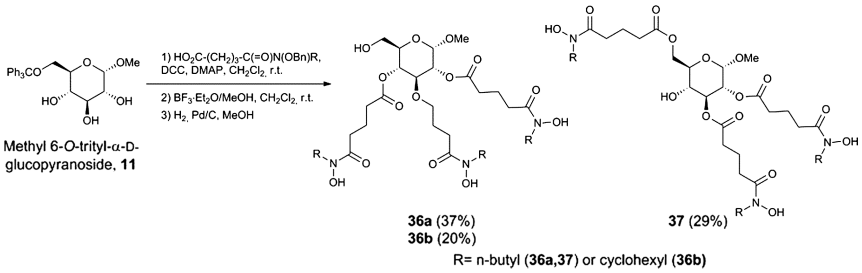
discovered in nature from the smut fungus, *Ustilago sphaerogena*, and is also produced by fungi of the genera *Aspergillus*, *Ustilago* and *Penicillium* among others.[39] Retro-hydroxamate siderophores are analogues in which the position of the hydroxamate nitrogen and carbon are interchanged relative to their position in the natural siderophore and have found application as antimalarial agents.[40–43]

The first synthetic analogues of hydroxamate siderophores based on methyl α-D-glucopyranoside, methyl α-D-galactopyranoside, methyl α-D-ribopyranoside and methyl α-D-xylopyranoside were in fact retro-hydroxamates and were reported by Heggemann *et al.*[15] The two hexoses were first protected with a trityl group at C-6, so that all four monosaccharides studied have only three free OH groups (OH-2, 3 and 4). The corresponding retro-trishydroxamates with different spacer groups were prepared by Steglich esterification with *N*-benzoyloxy-*N*-methylglutaric acid (**30–33a**), *N*-benzyloxy-*N*-methylsuccinic acid (**30–33b**), *N*-benzyloxy-*N*-methylglutaric acid (**30–33c**) and *N*-benzyloxy-*N*-butylglutaric acid (**30d**), using dicyclohexyl-carbodiimide (DCC) and *N,N*-dimethylaminopyridine (DMAP) in dichloromethane. Removal of trityl protecting groups was accomplished for *gluco* and *galacto* derivatives with boron trifluoride etherate, and debenzylation by catalytic hydrogenation afforded compounds **32–35e,f** and **34g** with free hydroxamate groups (Scheme 8).

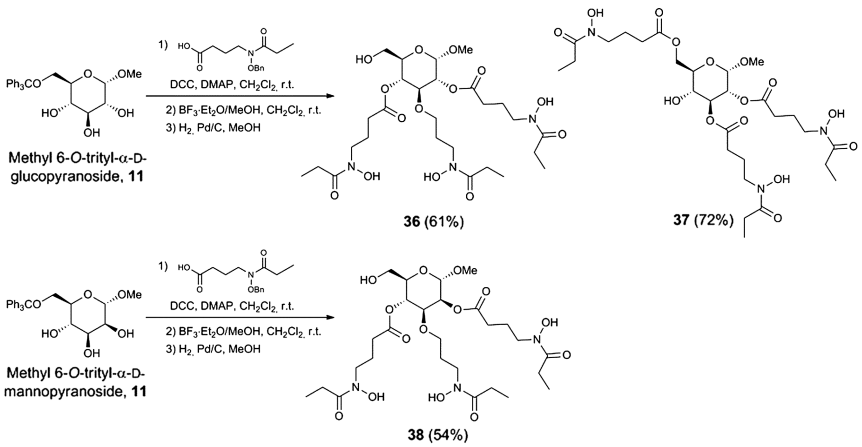
In 2008, Gebhardt *et al.*[38] synthesized new glucose-based retro-hydroxamates, with different *N*-alkyl groups by using the same methods described



Scheme 8 Synthesis of retro-hydroxamate derivatives with hexoses and pentoses.[15]



Scheme 9 Synthesis of glucose-based retro-hydroxamates with long *N*-alkyl chains.[38]



Scheme 10 Synthesis of normal-type hydroxamates based on glucose and mannose.[13,38]

before. Surprisingly, the authors observed the unequivocal formation of a 2,3,6-substituted glucopyranoside (**37**) as by-product, probably due to a Lewis acid catalyzed migration (Scheme 9).

More recently, Dhungana[13] and Gebhardt[38] reported the first synthetic analogues of desferrichrome with normal-hydroxamate geometry based on methyl α-D-glucopyranoside and methyl α-D-mannopyranoside with three asymmetric hydroxamic acid moieties at carbon positions 2, 3 and 4 (**38, 40**), or 2, 3 and 6 (**39**) (Scheme 10).

4 Study of siderophore activity

In Gram-negative bacteria the Fe(III)-siderophore complex is first captured at the cell surface by specific protein receptors, followed by transport to the periplasmic space. For *E. coli*, the most common bacterial strain used in the laboratory, siderophore-iron complexes are recognized by outer membrane receptors which are produced by microbes under iron stressed conditions. Different strains of bacteria produce different types of siderophore receptors. For example, *E. coli* expresses various outer membrane proteins including the ferrichrome receptor (FhuA) and enterobactin receptor (FepA) while *Pseudomonas putida* produces pseudobactin 358 receptor (PupA).[5] The siderophore activity of all siderophore analogs was examined

by growth promotion assays with various Gram-negative bacteria and mycobacteria, which have been well defined in their ability to transport and utilize natural siderophores (siderophore indicator strains).

Gram-negative indicator strains (Iron related marker) included *Salmonella typhimurium* enb-7 (enterobactin-) and TA 2700 (enterobactin-, FEB C-), *E. coli* AB 2847 (aromates biosynthesis-), *E. coli* IR 112 (aroB-, tonB-), *Klebsiella pneumoniae* KN 4401(enterobactin-, aerobactin biosynthesis-), *Morganella morganii* SBK 3 (wild type),*Yersinia enterocolitica* H 5030 (yersiniabactin-) and *Pseudomonas aeruginosa* PAO 6609 (pyoverdin-).[16]

Compounds **9**, **10**, **14a–d**, **15a–d**, **16** and **17** did not stimulated the growth of the indicator strains *P. aeruginosa* PAO 6609, *Klebsiella pneumoniae* KN 4401 and *E. coli* IR 112 missing the protein tonB. With this result the authors suggested that the presence of tonB is necessary for siderophore activity of these compounds. With other bacterial strains, different results were obtained. Biscatechol derivatives **9** and **10** displayed only a low growth promotion of the indicator strains *S. typhimurium* enb-7 and *E. coli* AB 2847. Triscatecholate **15a** promoted very strongly the growth of indicator strains *S. typhimurium* enb-7, TA 2700 and *E. coli* AB 2847 but not that of the other indicator strains; derivative **15b** showed only a minor activity against the strains *P. aeruginosa* ATCC 27853, SG 137 and *E. coli* ATCC 25922. However, compound **15c** promoted *Yersinia enterocolitica* H 5030. Trityl substituted derivatives (**14a–d, 16** and **17**) were of lower activity.[16]

Studies with Gram-negative strains of the antibacterial screening, grown under iron limitation, with *P. aeruginosa* ATCC 27853, SG 137, NCTC 10662 and *E. coli* ATCC 25922 were also accomplished. These experiments showed that compounds **15a–d** and **17** were active in growth promotion of the strains, while compounds **14a,b** and **16** were less active and no activity was found for biscatecholates **9** and **10**.[16]

The previous compounds were also studied for growth promotion of mycobacteria, with wild type strain of *M. smegmatis* SG 987 and mc2155, mutants of *M. smegmatis* M10 (exochelin-) and M24 (mycobactin-), a wild type of *M. phlei* 239 (exochelin-) and the mutant B1 of *M. smegmatis* mc2155 generated by gene replacement (blocked in exochelin biosynthesis).[16]

Here again, bischatecolates were inactive and the most active substances were **17** (the derivative with elongated spacer) and **15a.** The remaining compounds showed diverse results.[16]

Finally, the compounds were tested for their iron complexing capacity using the chromazurol-S test (CAS-assay) and **15a**, **15b**, **15d** and **17** were active.[16]

The complete set of results for siderophore activity of compounds **9**, **10**, **14a–d**, **15a–d**, **16** and **17** can be found in Tables 2 and 3.

Free trishydroxamates **32e**, **33d**, **33f**, **34d**, **34e**, **34g** and **35f** were active on all test strains using wild type strains of Gram-negative bacteria under iron limitation and the enterobactin mutant *S. typhimurium* enb 7. Galactose derivative **35f**, with longer spacer group was more active than **35e**; however, the glucose derivative with shorter spacer (**34d**) was more active than the one with longer spacer (**34e**). These results can indicate that the siderophore

Table 2 Growth promotion tests of the synthesized siderophore analogues on Gram-negative bacteria. Diameter of growth zone (mm), substance application 5 μg.[16]

Compound no.	9	10	14a	14b	14c	14d	15a	15b	15c	15d	16	17	control
Indicator strain													
S. typhimurium:enb-7	10	17	15	0	0	18	32	0	0	0	0	18	28[a]
S. typhimurium TA 2700	0	0	10	n.d.	n.d.	n.d.	33	n.d.	n.d.	n.d.	n.d.	n.d.	30[b]
E. coli AB 2847	0	10	0	0	0	0	25	0	0	0	0	0	28[c]
M. morganii SBK 3	0	0	0	0	0	0	10	0	0	24	0	10	22[d]
Y. enterocolitica H5030	0	0	0	0	10	0	0	0	16	0	10	0	18[c]
Strains of antibacterial screening													
P. aeruginosa ATCC 27853	0	0	0	0	0	0	n.d.	0	15	14	0	12	35[e]
P. aeruginosa SG 137	0	0	15	16	15	11	22	16	16	14	0	15	30[e]
P. aeruginosa NCTC 10662	0	0	15	0	0	14	25	15	22	17	0	15	30[e]
P. aeruginosa ATCC 9027	0	0	0	0	15	0	25	0	17	18	0	15	35[e]
E. coli ATCC25922	0	0	12	0	15	15	35	12	22	21	11	19	38[f]

[a] ferrioxamin G.
[b] enterobactin.
[c] ferrichrome.
[d] 2,3-dihydroxybenzylidene-1,3,5-trimethylalinine.
[e] desferal.
[f] ferricrocin (2 μg).

Table 3 Growth promotion tests of the synthesized siderophore analogues on strains of mycobacteria. Diameter of growth zone (mm), substance application 5 μg.[16]

Compound no.	9	10	14a	14b	14c	14d	15a	15b	15c	15d	16	17	Mycobactin (2 μg)
Indicator strain													
SG 987	0	0	9	0	13	10	23	0	10	10	14	28	17
M10	0	0	16	13	14	13	21	0	0	0	15	20	17
M77	0	0	0	0	0	0	0	10	0	0	0	0	17
mc^2 155	0	0	15	0	16	12	25	13	14	14	15	30	19
M24	0	0	0	0	0	0	26	0	0	0	0	8	19
B1	0	0	10	10	13	11	20	17	15	15	13	25	16

activity results from a joint influence of the length of the spacer groups and of the individual sugar backbone component.[15] The complete set of results can be observed in Table 4.

The results obtained on mycobacteria showed that compounds with free OH groups have siderophore activity as well as effective iron complexation, while **20** did not complexe with iron. Compound **21** was also active on strain mc2155 and B1. The derivatives **27a** (with 2,3-OH groups) and **27b** (with methoxycarbonyloxy groups) exhibited preferred activity on *P. aeruginosa* strains. Compound **34g** bearing a lipophilic butyl component was active on the wild type strain mc2155 and on the mutant B1. Iron complexes of the free hydroxamates **30e**, **33d**, **33f**, **34d**, **34e**, **34g**, **35e** and **35f**, as well as the iron complexes of all triscatecholates were effective growth promoters of the wild type strain mc2155 and the single mycobactin (M24) or exochelin (B1) mutant (Table 5).[15]

Table 4 Growth promotion tests of Gram-negative bacteria, diameter of growth zone (mm), substance application 10 µmol on 6 mm paper discs.[15]

Compound no.	20	21	25	27a	27b	30a	31a	32a	32e	33a	33d	33f	34a	34d	34e	34g	35a	35e	35f	control
Indicator strain																				
P. aeruginosa ATCC 27853	14	25	20	25	16	0	0	0	24	0	17	20	0	18	11	22	0	0	21	35[a]
P. aeruginosa SG 137	12	21	23	23	18	0	0	0	24	0	25	24	0	20	20	20	0	19	25	30[a]
P. aeruginosa NCTC 10662	11	28	25	25	17	0	0	0	24	0	24	29	0	25	19	20	11	14	28	31[a]
P. aeruginosa ATCC 9027	10	n.d.	n.d.	22	0	0	0	0	22	0	20	26	0	19	14	20	0	15	26	34[a]
P. aeruginosa K799/WT	15	23	23	20	15	0	0	0	18	0	25	28	0	20	11	22	0	10	25	35[a]
E. coli ATCC25922	0	27	26	15	0	0	0	0	26	0	23	27	13	22	16	23	0	17	19	33[b]
S. typhimurium enb-7	0	20	26	14	0	0	0	0	17	0	17	27	0	12	10	19	0	0	28	32[a]

[a] desferal G.
[b] ferricrocin.

Table 5 Growth promotion of mycobacteria, diameter of growth zone (mm), substance application 10 µmol (mycobactin 2 µg) on 6 mm paper discs (with and without iron complexation).[15]

Compound no.	20	21	25	27a	27b	30a	31a	32a	32e	33a	33d	33f	34a	34d	34e	34g	35a	35e	35f	A	B
Indicator strain																					
mc^2 155	10	24	22	0	0	0	0	0	0	8	0	0	0	0	0	25	9	0	0	-	-
mc^2 155 + Fe	-	30	27	29	26	-	-	-	33	-	30	34	-	18	24	28	-	22	36	30	14
M24	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	-	-
M24 + Fe	-	18	16	20	21	-	-	-	32	-	24	31	-	22	27	19	-	15	33	28	16
B1	0	21	22	0	0	0	0	0	0	0	0	18	7	0	0	26	0	0	0	-	-
B1 + Fe	-	32	23	26	21	-	-	-	30	-	25	33	-	21	27	25	-	21	36	32	15
B3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	-
B3 + Fe	-	14	0	0	0	-	-	-	19	-	18	15	-	15	12	13	-	13	28	25	15
U3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	-
U3 + Fe	-	13	0	0	0	-	-	-	24	-	20	17	-	16	12	12	-	0	26	22	16

A: ferrichrome; B: mycobactin.

Table 6 Growth promotion of *E. coli* wild type strains and mutants. Diameter of growth zone (mm), substance application 10 µmol.[15]

Compound no.	16	17	21	25	33f	35f	2,3-DHBA[a]	Ferricrocin
Indicator strain								
AB 2847	32	25	25	29	24	23	30	30
BR 158	0	0	0	0	0	0	0	0
H 1443	32	26	25	31	25	23	32	30
H 18765	0	0	0	0	26	24	0	30
H 873	27	25	24	27	25	21	30	31
H 1877	25	25	21	25	27	22	27	32
H a875	25	22	22	25	27	23	30	33

[a] 2,3—DHBA = 2,3-dihydroxybenzoic acid.

The recognition and uptake specificity of the siderophores by use of *E. coli* K12 mutants with alterations in the TonB protein was studied for relevant compounds using ferricrocin as control (Table 6).

Siderophore activity of compounds **36-40** (Table 7) was tested using *M. smegmatis* wild type strains and mutants in siderophore biosynthesis and transport.[38] The authors observed stronger growth promotion activity for iron complexed siderophores compared to their iron free analogs. The hydroxamates **38–40** have shown nearly zero activity, but when complexed with iron strong growth promotion activity was detected. In contrast, retro-hydroxamates **36a**, **36b** and **37** promoted growth of the mycobacteria also in both ferri/desferri forms. These results suggest that the growth promotion activity might depend on lipophilicity rather than on siderophore receptor mediated uptake, since the higher lipophilic character of the retro-hydroxamates facilitates diffusion into the hydrophobic mycobacterial cell envelope.[38] It was observed that glucose-based ferri-hydroxamates **38** and **39** have improved growth promoting activity for both of the mutant strains mc2155-M24-B3 and mc2155-M24-U3, which are completely blocked in their species specific exochelin/mycobactin iron uptake system. Mannose-based ferri-hydroxamate **40** and glucose-based ferri-retrohydroxamates **36a**, **36b** and **37** were less effective for exochelin/mycobactin biosynthesis mutant mc2155-M24-B3 and inactive for the mycobactin biosynthesis/exochelin uptake mutant mc2155-M24-U3 (Table 7).[38]

5 Conclusions and perspectives

During the last decade a series of synthetic analogs of bacterial siderophores built around a carbohydrate skeleton were reported in the literature. These new carbohydrate-based artificial siderophores were focused only on monosaccharide scaffolds. The results obtained indicate that such compounds are highly promising synthetic analogs, as *in vivo* studies showed significant siderophore activity, indicating favorable cell receptor recognition and cellular uptake. In this context, the use of siderophores with carbohydrate skeleton merits to be further exploited. The use of disaccharide platforms (such as sucrose) with several possible hydroxyl groups to explore as attachment point for the ligands would allow to study the importance of

Table 7 Growth promotion of mycobacteria with *Mycobacterium smegmatis* strains.[38]

Compound no.	A[a]	B[a]	C[b]	36a	36a + Fe	36b	36b + Fe	37	37 + Fe	38	38 + Fe	39	39 + Fe	40	40 + Fe	D	E
Indicator strain																	
SG 987	+	+	+	++++	++++	++++	+++	++++	++++	+	++++	+	+++	+	+++	++	++
SG 987–M10	−	+	+	++	+++	+++	+++	++	+++	−	+++	−	++++	−	+++	++	++
mc^2 155	+	+	+	++++	++++	++++	+++	++++	++++	+	++++	−	++++	+	++++	++	++
mc^2 155–M24	+	−	+	−	++++	−	++	−	+++	−	++++	−	++++	−	+++	++	++
mc^2 155–B1	−	+	+	+++	++++	++++	++	++++	++++	+	++++	−	++++	+	++++	++	+
mc^2 155–M24–B3	−	−	+	−	++	−	+	−	++	−	++	−	+++	−	+	++	−
mc^2 155–M24–U3	+	−	−	−	+	−	+	−	+	−	++	−	+++	−	−	++	−

[a] indicates that the strain biosinthesizes the indicated siderophore and has the potential for ligand exchange, yes(+) or no(−).
[b] uptake pf exochelin by the indicated strains, yes(+) or no(−).
[c] diameter of orange halo in mm −, +, ++, +++ indicates the presence and increasing intensity of growth promoting activity.

the position of the ligands in the saccharide backbone and also to evaluate siderophores with different hydrophilicity/lipophilicity. Other promising applications in this area may arise by attaching a fluorescent probe with a linker to the artificial siderophore, which will not interfere with iron chelation and subsequent receptor recognition, and that would make possible to develop new iron-sensitive fluorescent sensors for biological signaling and biomedical or environmental applications.

References

1 H. Boukhalfa and A. L. Crumbliss, *BioMetals*, 2002, **15**, 325–339.
2 V. Braun and H. Killmann, *Trends Biochem. Sci.*, 1999, **24**, 104–109.
3 K. Matsumoto, T. Ozawa, K. Jitsukawa and H. Masuda, *Inorg. Chem.*, 2004, **43**, 8538–8546.
4 B. F. Matzanke, in *Encyclopedia of Inorganic Chemistry*, John Wiley & Sons, Ltd, 2006.
5 J. M. Roosenberg, Y. M. Lin, Y. Lu and M. J. Miller, *Curr. Med. Chem.*, 2000, **7**, 159–197.
6 G. M. Rodriguez, R. Gardner, N. Kaur and O. Phanstiel, *BioMetals*, 2008, **21**, 93–103.
7 *Iron transport in bacteria*, ASM Press, Washington D.C., 2004.
8 A. M. Albrecht-Gary and A. L. Crumbliss, *Metal ions in biological systems*, New York, 1988.
9 A. L. Crumbliss, *Handbook of microbial iron chelates*, CRL Press, Boca Ranton, FL, 1992.
10 E. J. Neufeld, *Blood*, 2006, **107**, 3436–3441.
11 R. J. Bergeron, C. Z. Liu, J. S. McManis, M. X. B. Xia, S. E. Algee and J. Wiegand, *J. Med. Chem.*, 1994, **37**, 1411–1417.
12 I. Turcot, A. Stintzi, J. D. Xu and K. N. Raymond, *J. Biol. Inorg. Chem.*, 2000, **5**, 634–641.
13 S. Dhungana, J. M. Harrington, P. Geblhardt, U. Mollmann and A. L. Crumbliss, *Inorg. Chem.*, 2007, **46**, 8362–8371.
14 S. Dhungana, S. Heggemann, L. Heinisch, U. Mollmann, H. Boukhalfa and A. L. Crumbliss, *Inorg. Chem.*, 2001, **40**, 7079–7086.
15 S. Heggemann, U. Mollmann, P. Gebhardt and L. Heinisch, *BioMetals*, 2003, **16**, 539–551.
16 S. Heggemann, M. Schnabelrauch, D. Klemm, U. Mollmann, R. Reissbrodt and L. Heinisch, *BioMetals*, 2001, **14**, 1–11.
17 L. Heinisch, P. Gebhardt, R. Heidersbach, R. Reissbrodt and U. Möllmann, *BioMetals*, 2002, **15**, 133–144.
18 Z. D. Liu and R. C. Hider, *Coord. Chem. Rev.*, 2002, **232**, 151–171.
19 K. Matsumoto, N. Suzuki, T. Ozawa, K. Jitsukawa and H. Masuda, *Eur. J. Inorg. Chem.*, 2001, 2481–2484.
20 S. Wittmann, L. Heinisch, I. Scherlitz-Hofmann, T. Stoiber, D. Ankel-Fuchs and U. Möllmann, *BioMetals*, 2004, **17**, 53–64.
21 A. Tsafack, J. Libman, A. Shanzer and Z. I. Cabantchik, *Antimicrob. Agents Ch.*, 1996, **40**, 2160–2166.
22 C. Y. Ng, S. J. Rodgers and K. N. Raymond, *Inorg. Chem.*, 1989, **28**, 2062–2066.
23 T. Masuda, S. Asari, S. Tamagaki and T. Nagasaki, *Tetrahedron Lett.*, 2000, **41**, 2411–2414.
24 I. Yoshida, I. Murase, R. J. Motekaitis and A. E. Martell, *Can. J. Chem.*, 1983, **61**, 2740–2744.

25 M. M. Meijler, R. Arad-Yellin, Z. I. Cabantchik and A. Shanzer, *J. Am. Chem. Soc.*, 2002, **124**, 12666–12667.
26 H. Weizman, O. Ardon, B. Mester, J. Libman, O. Dwir, Y. Hadar, Y. Chen and A. Shanzer, *J. Am. Chem. Soc.*, 1996, **118**, 12368–12375.
27 H. Ouchetto, M. Dias, R. Mornet, E. Lesuisse and J. M. Camadro, *Bioorg. Med. Chem.*, 2005, **13**, 1799–1803.
28 T. Inomata, H. Eguchi, K. Matsumoto, Y. Funahashi, T. Ozawa and H. Masuda, *Biosens. Bioelectron.*, 2007, **23**, 751–755.
29 Y. M. Lin and M. J. Miller, *J. Org. Chem.*, 1999, **64**, 7451–7458.
30 A. W. Coleman, C. C. Ling and M. Miocque, *Angew. Chem.-Int. Edit. Engl.*, 1992, **31**, 1381–1383.
31 T. Masuda, J. Hayashi and S. Tamagaki, *J. Chem. Soc.-Perkin Trans.*, 2000, **2**, 161–167.
32 B. Tse and Y. Kishi, *J. Am. Chem. Soc.*, 1993, **115**, 7892–7893.
33 K. N. Raymond and C. J. Carrano, *Acc. Chem. Res.*, 1979, **12**, 183–190.
34 B. H. Lee, M. J. Miller, C. A. Prody and J. B. Neilands, *J. Med. Chem.*, 1985, **28**, 317–323.
35 B. H. Lee, M. J. Miller, C. A. Prody and J. B. Neilands, *J. Med. Chem.*, 1985, **28**, 323–327.
36 W. Kiggen and F. Vögtle, *Angew. Chem., Int. Ed.*, 1984, **23**, 714–715.
37 S. Dhungana, S. Heggemann, P. Gebhardt, U. Mollmann and A. L. Crumbliss, *Inorg. Chem.*, 2003, **42**, 42–50.
38 P. Gebhardt, A. L. Crumbliss, M. J. Miller and U. Mollmann, *BioMetals*, 2008, **21**, 41–51.
39 A. Butler and R. M. Theisen, *Coordin. Chem. Rev.*, 2010, **254**, 288–296.
40 J. Golenser, A. Tsafack, Y. Amichai, J. Libman, A. Shanzer and Z. Cabantchik, *Antimicrob. Agents Chemother.*, 1995, **39**, 61–65.
41 K. P. Holland, H. L. Elford, V. Bracchi, C. G. Annis, S. M. Schuster and D. Chakrabarti, *Antimicrob. Agents Chemother.*, 1998, **42**, 2456–2458.
42 M. J. Miller, A. J. Walz, H. Zhu, C. Wu, G. Moraski, U. Möllmann, E. M. Tristani, A. L. Crumbliss, M. T. Ferdig, L. Checkley, R. L. Edwards and H. I. Boshoff, *J. Am. Chem. Soc.*, 2011, **133**, 2076–2079.
43 P. K. Ojha and K. Roy, *Mol. Simul.*, 2010, **36**, 939–952.

Smart biomaterials: the contribution of glycoscience

Laura Cipolla,* Laura Russo, Francesca Taraballi, Cristina Lupo, Davide Bini, Luca Gabrielli, Alice Capitoli and Francesco Nicotra
DOI: 10.1039/9781849734769-00416

Examples of material functionalisation (''biodecoration'') with signalling and relevant glycidic scaffolds will be outlined. Recent research concerning the development of smart biomaterials for Tissue Engineering (TE) applications will be considered.

1 Introduction: biomaterials and tissue engineering

Humankind's use of materials to repair the body dates to antiquity, when natural materials such as wood were used in an attempt to structurally replace injured tissues by diseases or trauma. In the beginning of the twentieth century, synthetic polymers, ceramics and metal alloys were introduced in place of natural materials; they offer better performance, increased functionality and more reproducibility than their naturally derived counterparts. These advances led to a pronounced increase in the range of use and the efficacy of biomaterials, as a result of which millions of lives have been saved or improved by devices such as vascular stents, dental restoratives, artificial hips and contact lenses. The increasing importance of biomaterials in our society over the past decades can be tracked in a number of ways, including the growth of tissue engineering applications both as an academic discipline and as an important industry.[1]

Sales of regenerative biomaterials already exceed US$240 million per annum[2] and the wider markets that tissue engineering taps into are colossal: costs related to organ replacement account for 8% of global healthcare spending, and by 2040 as much as 25% of the US GDP is expected to be related to healthcare.[3] Nevertheless, if the short history of industrial tissue engineering has taught us anything, it is that the provision of effective products is not in itself sufficient to ensure commercial success. Early TE efforts were plagued by product issues related to scale-up, shelf-life, quality control and distribution, and suffered from inappropriate business models and withdrawal of private finance in the early 2000s.[4] Since then the field has matured, evidenced by the return of large-scale investment and the first regenerative medicine companies becoming profitable. Alongside these positive developments, progress in biomaterials design and engineering are converging to enable a new generation of instructive materials to emerge as candidates for regenerative medicine.[5]

Initially, the main goal of scientists was the design of materials largely performing mechanical functions, being preferably 'inert' without interaction with the biology of the host organism in order to prevent rejection.[6,7]

Dept. of Biotechnology and Biosciences University of Milano-Bicocca, P.za della Scienza 2, 20126 Milano, Italy. E-mail: laura.cipolla@unimib.it

© The Royal Society of Chemistry 2012

Early research and fortuitous accidents linking materials chemistry to biological response provided a rational basis for developing new biomaterials.[1] The molecular biology revolution of the 1970s and advances in genomics, proteomics and more recently in glycomics significantly affected the ways in which biomaterials are nowadays designed and used.

It is well known that a series of interactions occur between the surface of biomaterials and the biological environment after they have been implanted into the human body. Therefore, the biomaterials surface plays an extremely relevant role in the response of artificial medical devices to the biological environment.

The efficacy of artificial implants is determined mainly by their surface characteristics such as surface morphology, microstructure, composition, and functionalization. These properties alter adsorption of proteins which mediate the adhesion of desirable and undesirable cells.[8] Research on biomaterials surface has become one of the hottest topics in biomaterials and biomedical engineering. The highest score for a keyword in the World Biomaterial Congress in 2008 (WBC 2008) is Biomaterials Surface, indicating that the current research focus of the biomaterials community is aimed at understanding the fundamental processes at the interface between implant surfaces and surrounding living tissues.[9]

To this end, scientists are creating new materials including those with improved biocompatibility, stealth properties, cell responsiveness (**smart materials**), specificity and other critical properties. Modern biomaterials science is characterized by a growing need to integrate biomaterials design with new insights emerging from studies of cell–matrix interactions, cellular signalling processes, developmental and systems biology.[10]

Since cell contact with the biomaterial surface significantly influences cell behaviour and performance, trends in biomaterial designs lean towards bioactive materials that can modulate and control cell behaviour.

1.1 Tissue engineering and the biomimetic approach

A variety of new materials are being synthesized from man-made building blocks, and being used to create devices for specific medical applications.[5,11] Materials composed of naturally occurring (biologically derived) building blocks, including extracellular matrix (ECM) components, are being studied for applications such as direct tissue replacement and tissue engineering. The ECM, a complex composite of proteins, glycoproteins and proteoglycans, provides an important model for biomaterials design (Fig. 1).[12]

ECM-derived macromolecules (for example, collagen) have been used for many years in biomaterials applications[5] and it is now possible to create artificial analogues of ECM proteins using recombinant DNA technology.[13] Tissue engineering is an advanced interdisciplinary field that encompasses the design of biomaterials for *in vivo* tissue regeneration where mimicking the natural ECM is often pursued.

The natural ECM supports organ and tissue structure and function, and also regulates basic cellular events like proliferation, growth, migration, differentiation, and survival.[14] These functions are controlled through tissue-specific constituents, such as collagens, laminins, fibronectin or elastins, as well as functional molecules like growth factors or matricellular proteins,

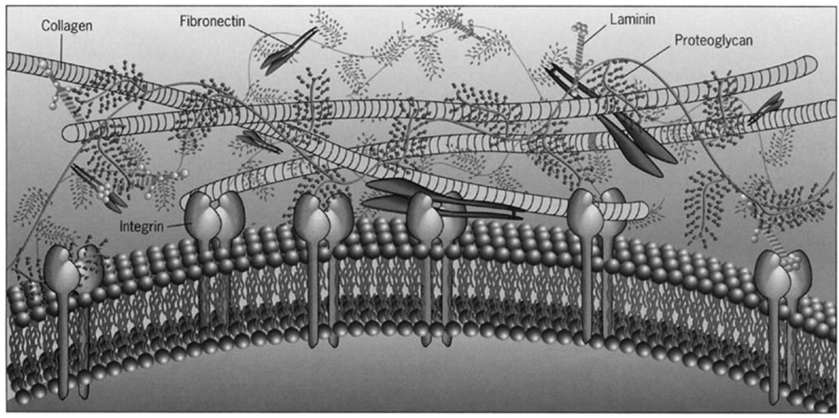


Fig. 1 ECM-cell interaction.

among others. Novel biomaterials should allow for the gradual endogenous remodeling of native tissue leading to the replacement of implant material, manufactured to replace a missing biological structure, with fully functional ECM and cells that existed at the implant site prior to damage.

The outer membrane of a typical cell is covered by specific carbohydrate structures and a forest of at least six different receptor systems that can be activated by interactions with adjacent cells, ligands in the surrounding ECM, and secreted signalling molecules. Hundreds of different proteins play a role in the composite stimulation of cell receptors, which in turn determine a plethora of responses, including cell migration in the early embryo, coordinated organogenesis, and wound repair throughout adult life.[15–19] Collectively, these extrinsic factors make up a highly defined and specialized cell microenvironment, which is essential for correct tissue development and continued function.

The ECM takes a variety of forms and composition in different tissues and at different stages of development in the same tissue.[20] Diversity arises through combinations of specific molecular interactions between numerous isoforms, ratios, and geometrical arrangements of collagens, elastins, proteoglycans, and adhesion proteins such as fibronectins and laminins. This creates an environment that is replete with informational cues. In addition to this, a wealth of molecular mechanisms modulates the dissemination of the biological information.[21]

The critical point during *in vitro* tissue engineering is to at least partially recreate conditions that mimic the natural ECM environment for particular cell types in order to support their function and proliferation. Since cell contact with the biomaterial surface is a key point, in recent years, biomaterials design has focused on the exposition and incorporation of signalling molecules into scaffold materials rather than using them in a diffusive or soluble way (Fig. 2).[22]

Among the most studied molecules are multifunctional proteins like growth factors[23–26] or cytokines,[27] while there are also reports on use of small molecules like neurotransmitters (Fig. 3).[14]

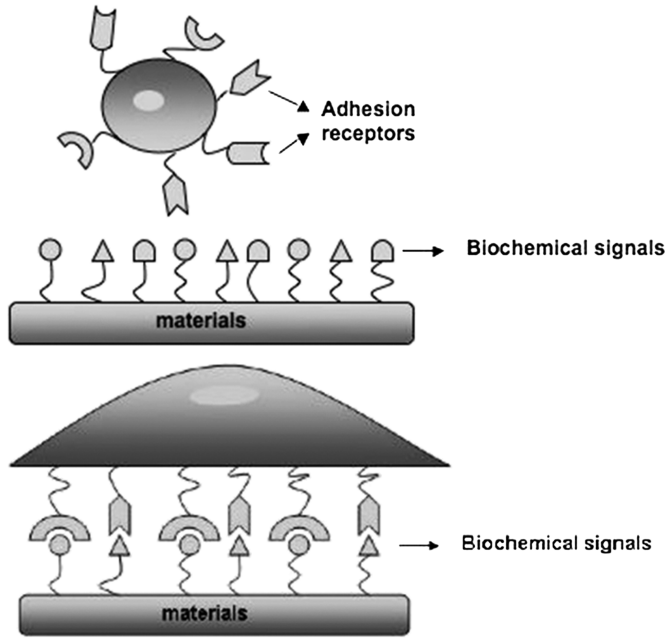


Fig. 2 Surface modifcation in biomaterial design.

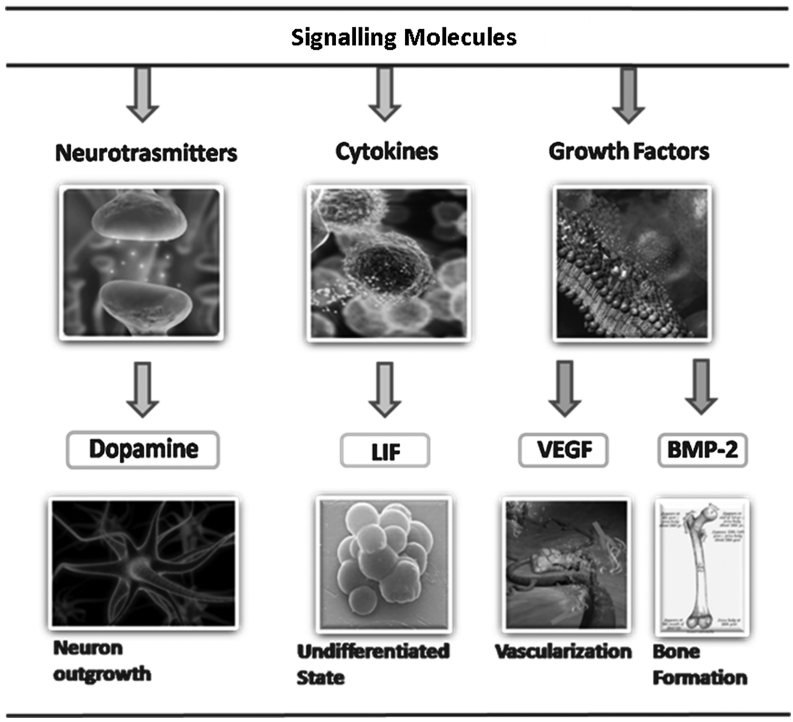


Fig. 3 Examples of signalling molecules modulating cell behaviour in bioactive materials for tissue engineering.

1.2 On the chemical nature of materials

It is sensible to consider today's optimal meaning of 'biomaterial' from two different perspectives, the first being concerned with the evolution of materials science and the wide range of materials options that have opened up during the last decade or so, and the second being the evolution of health care technologies.

The classical view of a material has been "a substance of which things are made." Materials scientists were taught that there were three primary types of material, metallic, being based on the metallic bond, ceramic, based on ionic bonds and polymeric, based on covalent bonds. In addition there were hybrids, which could either be entirely synthetic, usually referred to as composites which typically would be combinations of ceramics and polymers, or the natural equivalents of these composites, including bone, wood, and ivory. Obviously each of these categories contained many subcategories. The metallic materials included pure metals and alloys, ceramics included glasses, glass–ceramics and carbons, the polymers included thermosets, thermoplastics, elastomers and textiles. As biomaterials science emerged, the conventional view of materials, as being tangible pieces of substances from which useful objects were made, prevailed. The stems of hip replacements were made of metals, artificial arteries were made of textiles, dentures and intraocular lenses were made of acrylic polymers. However, these boundaries between material classes have now been eroded; those substances derived from clear, chemically defined primary interatomic and intermolecular bonds are being replaced by those of greater structural complexity that arise from quite different concepts, including those of nanotechnology and self assembly processes inspired by nature. Indeed it is one of the fundamental tenants that is driving nanoscience and nanotechnology that is at the heart of the revolution in materials science (or materials chemistry as it is so often called now), and that is the replacement of top down manufacturing by bottom up synthesis.[28]

Both biologically derived and synthetic materials have been extensively explored in regenerative medicine and tissue engineering. In general, materials from natural sources (*e.g.*, purified protein components such as collagens from animal tissues) are advantageous because of their inherent properties of biological recognition, including presentation of receptor-binding ligands and susceptibility to cell-triggered proteolytic degradation and remodeling. Despite these advantages, many issues have spurred the development of synthetic biomaterials as cellular substrates, including complexities associated with purification, immunogenicity and pathogen transmission. Although some of these limitations can be overcome by recombinant protein expression technologies,[29] greater control over materials properties and tissue responses could be achieved if synthetic analogs were available.[30]

1.3 Functionalization strategies

Different biomedical devices and applications require different properties and functions of materials. Therefore methods to modify materials able to meet the needs of different biomedical systems are extremely variable.

The surfaces of nanostructured materials can be modified and functionalized with different reagents using different methods, including physical,

chemical, or biological. Two things are often achieved by surface modification of materials: 1) enhanced stability and biocompatibility of nanostructured materials in aqueous media and 2) new material functions and properties.[31]

The surface modification of biomaterials with bioactive molecules is a simple way to make smart materials. To date, different strategies have been used for the introduction of biomimetic elements into synthetic materials:

a) Physical adsorption (van der Waals, electrostatic, affinity, adsorbed and cross linked).

b) Physical entrapment attachment (barrier system, hydrogel, dispersed matrix system).

c) Covalent surface immobilization, taking advantage of different natural or unnatural functional groups present both on the biomolecules and on the material surfaces (chemoselective ligation, *via* amino functionalities, heterobifunctional linkers, *etc*.).

The major methods of immobilizing a bioactive compound to a polymeric surface are adsorption *via* electrostatic interactions, ligand–receptor pairing (as in the couple biotin–avidin), and covalent attachment (Fig. 4).

Physical methods, such as molecular coating (or adsorption), surface entrapment, and physical treating with plasma, ozone, or UV have emerged as leading strategies for surface modifications of materials. Through physical modifications, functional molecules and entities, varying charges, or active chemical groups can be introduced onto materials, leading to the functionalization and bioactivation of materials. Physical modifications have obvious advantages, including ease of handling and mild interactions with biomolecules through little or no damage to their bioactivities. These methods, however, also exhibit certain limitations. For example, physical interactions are often formed between the substrates and coatings using these strategies. These physical interactions are usually weak compared to chemical bonds. Therefore, functional molecules and entities may detach from the surfaces especially when certain serum compounds compete for

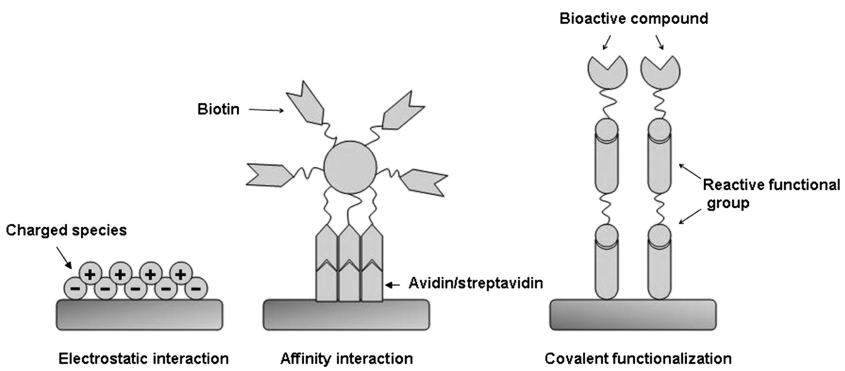


Fig. 4 Different strategies for the introduction of biomimetic elements into synthetic materials.

active binding sites in physiological conditions; detachment is worsened by a competitive displacement of theses functional molecules and entities. Non-covalent adsorption is however sometimes desirable, for example in drug delivery applications.[32]

Covalent immobilizations offer several advantages by providing the most stable bond between the compound and the functionalized polymer surface. A covalent immobilization can be used to extend the half-life of a biomolecule, prevent its metabolism, or allow continued bioactivity of indwelling devices.[33]

Bioactive molecules (growth factors, ECM proteins, *etc.*) that are free in solution, as opposed to those immobilized to the matrix, may induce significantly different biological responses. Growth factors are routinely added to cultures *in vitro*, and have been incorporated and released from polymeric systems with retention of bioactivity, as shown for neurotrophins, BMPs, and VEGF.[34–36] *In vivo*, these soluble factors can be released from the delivery site, and the relevant parameter is the duration over which therapeutic concentrations can be maintained. Alternatively, bioactive molecules can be linked covalently to the scaffold, eventually in a reversible way or exploiting a degradable linking tether.

These scaffolds have been termed "cell-responsive"[37] due to release of the factor upon cellular demand. Once released, these soluble factors can bind their receptors and initiate the signalling cascade.

Alternatively, immobilized biomolecules can ligate their receptors directly from the material surface; however, this type of interaction may not exactly replicate the signalling of soluble factors, as growth factor internalization can stimulate signalling pathways different from those activated at the surface. For example, neuronal growth factor NGF induces neurite outgrowth by signalling at the plasma membrane, yet promotes neuron survival when internalized.[38] Surface immobilization has been successfully used to attach several factors such as EGF,[39] BMP-7,[40] BMP-2,[41] VEGF,[42] NGF,[43] and NT-3[44] to a variety of natural and synthetic biomaterials. Signalling by these immobilized or locally released bioactive ligands may be more potent than signalling by soluble versions added directly to culture media.[45] These studies also demonstrate that the immobilization strategy must consider protein structure and active region topology, when designing suitable delivery systems in order to maximize bioactivity. Ultimately, some factors may be best delivered in a continuous manner, while others benefit from direct attachment to the biomaterial surface.[46]

Different methods have been developed for covalent functionalization of biomolecules to diverse biomaterials. For covalent functionalization to an inert solid polymer, first the surface must be modified to provide reactive groups (-OH, -NH_2, COOH, -SH) for the subsequent functionalization step. When the material does not contain reactive groups, they can be generated by chemical and physical modification. With this goal, a wide number of surface modification techniques have been developed, including plasma,[47] ionic radiation graft polymerization, photochemical grafting, chemical modification and chemical derivatization.

In this approach, however, the biomolecules will be linked to the material surface with a stochastic orientation; this implies that not all the molecules

will be exposed efficiently for receptor recognition. Moreover, the chemical nature of the material can deeply influence the conformation of the molecule,[48,49] that can result in an inactive conformation.

For example peptides can react *via* the N-terminus with different groups on polymers (Fig. 5). This is usually done by reacting an activated carboxylic acid group with the nucleophilic N-terminus of peptides. The carboxylic group can be activated with different peptide coupling reagent, *e.g.* 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, also referred to as water soluble carbodiimide, WSC), dicyclohexylcarbodiimide (DCC) or carbonyl diimidazole (CDI).

More recent approaches capitalize on chemoselective strategies (Fig. 6), where selected pairs of functional groups are used to form stable bonds without the need of an activating agent and without interfering with other functionalities usually encountered in biomolecules. Chemoselective ligation proceeds usually under mild conditions and results in good yields,[50] and allows the linkage of the biomolecule with a controlled spatial orientation.

A biomolecule may also be attached by these coupling methods *via* a spacer group, in order to give better access to the target receptor. One useful and biocompatible spacer is PEG that has been differently functionalized at two extremities. Metal or ceramic surface may also be silanized, exploiting functionalized triethoxysilanes.[51]

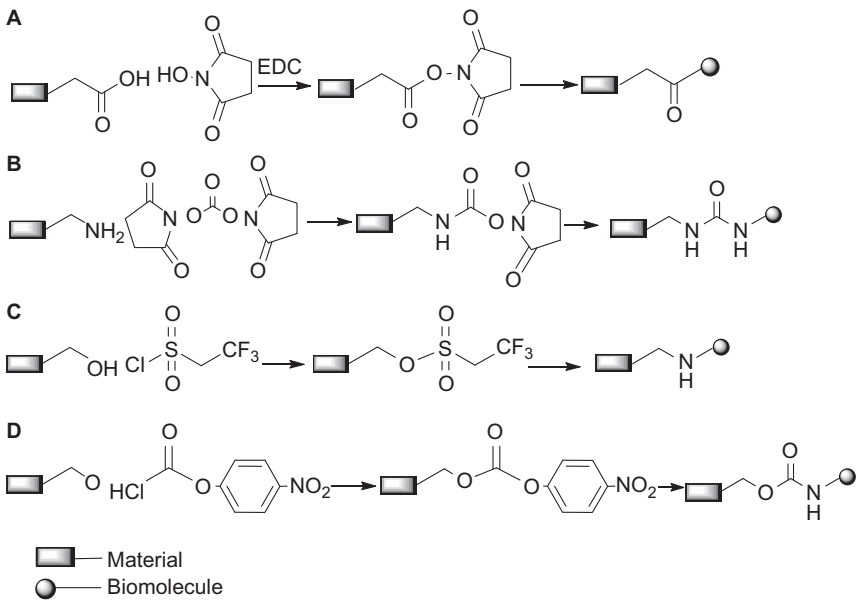


Fig. 5 Coupling methods to different groups on materials.

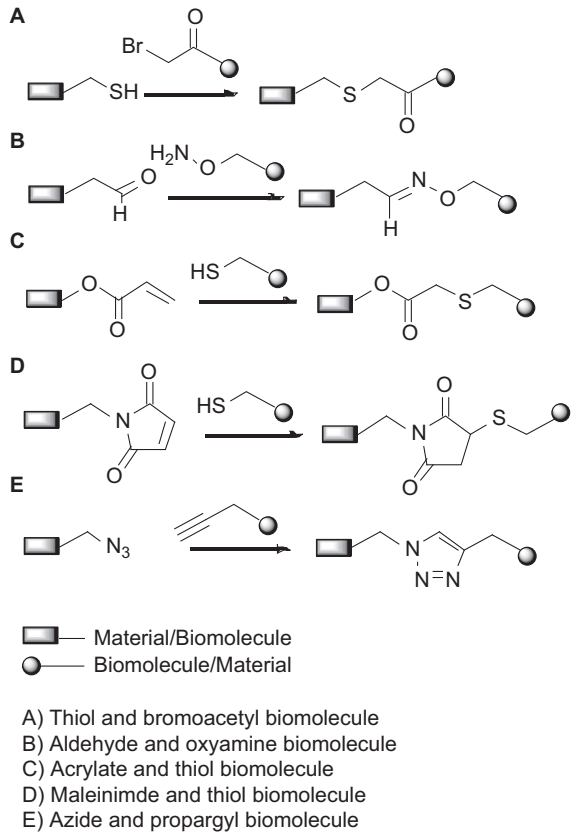


Fig. 6 Examples of chemoselective strategies.

2 Glycoscience for biomaterials design

The glycome, or complement of sugars in the cells, varies among cell type but also within cell type growing to differences such as stage of differentiation and malignant transformation.[52] The traditional view of carbohydrate polymers as energy source (starch and glycogen) and structural materials (cellulose, collagen, and proteoglycans) has expanded. Today, carbohydrates are known to have a wide variety of biological functions. For example, the sulfated polysaccharide, heparin, plays an essential role in blood coagulation.[53] Another related polysaccharide, hyaluronan, which acts as a lubricant in joints, has been used to protect the corneal endothelium during ophthalmologic surgery.[54] In addition to hyaluronan's lubricating and cushioning properties, polysaccharide and glycoprotein participate in a number of recognizing processes. Synthetic carbohydrate based polymers are increasingly being explored as biodegradable, biocompatible and biorenewable materials for use as water absorbent, chromatographic supports and medical devices. In addition, also membrane proteins, *i.e.* lectins, bind specific carbohydrates. This binding is extremely specific,[55] and may present an attractive target for rational design of smart biomaterials. Exploiting lectin–sugar interactions increases design flexibility

by offering the possibility of grafting active sugars on biomaterial surfaces. One successful exploitation of lectin binding is the creation of galactose-grafted scaffolds to support hepatocyte (liver cell) growth (see next sections).[56,57] It is notable that the galactose receptor normally functions to bind soluble sugars, not to mediate adhesion. Thus, it is clear that some molecules not naturally involved with adhesion may be pressed into that service through creative engineering. This is not a universal principle, however, and care must be taken in interpreting results. Most molecules that bind to a cell surface component (either specifically or non-specifically) will allow some degree of adhesion.[52]

Hajime Sato *et al*. exploited the anionic polysaccharide heparin to extend the scope of micropatterned carbohydrate displays by applying electrostatic interaction. They also used micropatterned carbohydrate displays on silicon substrates by a combination of photolithography and self-assembly.[58] These techniques have interesting potential applications in tissue engineering. Top-down photolithography and bottom-up molecular self-assembly are the two main approaches to biomacromolecular micropatterning. Photolithography has been widely used in electronic devices and most practically applied for precise micropatterning of proteins for cell cultivation.[59]

Heparin is a glycosaminoglycan polysaccharide with a variable number of sulfate groups and its binding to basic fibroblast growth factor (bFGF) is known to modulate cell proliferation and differentiation. The interaction of heparin with bFGF induces the dimerization of bFGF, which is required for binding of bFGF to FGF receptor (FGF-R) on cell surfaces,[60,61] as illustrated below (Fig. 7).

The self-assembled monolayer (SAM) of 3-aminopropyltrimethoxysilane (APS) as ammonium-terminated template was prepared on silicon and glass substrates, and micropatterned by photolithography. Self-assembling of heparin onto the APS-SAM through electrostatic interaction and then molecular recognition of bFGF to heparin and of fibroblast cells to bFGF has been achieved.[62] Adsorption of heparin was not observed on a silicon substrate without APS-SAM. Heparin adsorbed on APS-SAM remained on the substrate after rinsing with PBS buffer. The electrostatic interaction was essential for anchoring heparin on APS-SAM. The time-course of heparin adsorption to APS-SAM does not fit with a simple Langmuir model, suggesting that, in addition to electrostatic interaction, some other interaction

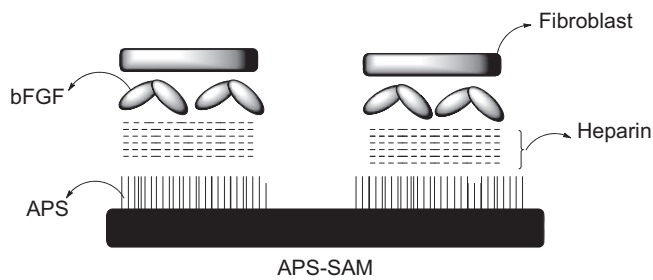


Fig. 7 A three-step method for monosaccharide grafting on PET fibers.

such as hydrogen bonding may contribute a little to the adsorption of heparin to APS-SAM.[63,64]

The above carbohydrate display using heparin has an attractive potential for tissue engineering, since it allows a facile and effective strategy for biomacromolecular patterning with significant bioactive molecules.

In the following sections examples of material functionalization (''biodecoration'') with signalling and relevant glycidic scaffolds will be outlined. Only recent research concerning the development of smart biomaterials for Tissue Engineering (TE) applications will be considered here, while glycoarrays and glycated nanoparticles for theragnostic application will not be included. The sections will be organized based on the chemistry used for covalent attachment of the carbohydrates to the material surfaces.

2.1 Plasma grafting toward "glyco-decoration"

Surface modification finds numerous applications in the preparation of biomaterials, for example, for the control of interfacial interactions between materials and biological molecules.[65–67] Plasma grafting is used in order to modify the surface features,[68] by the introduction of functional groups that can alter the physico-chemical properties of the surface itself (*i.e.* hydrophilicity) and at the same time allow covalent functionalization of the material with polymers,[69] peptides[70] or carbohydrates.[71] Moreover, this method is particularly advantageous when the material surface lacks in functional groups. Extremely high levels of surface functionality can be obtained by this approach. Examples successfully devised in the past include amine,[72] anhydride,[73] epoxide,[74] carboxylic acid,[75] cyano,[76] halide,[77] hydroxyl,[78] furfuryl,[79] and perfluoroalkyl[80] functionalized surfaces.

This method brings many advantages in changing the hydrophobicity or introducing functional groups homogeneously with a defined density over the inert material surfaces.[81] Effectively, any surface that relies on specific chemistry for its performance can, in principle, be generated by the aforementioned pulsed plasmachemical methodology. Then, friction properties and adhesion qualities of the treated material can be improved by many attempts of plasma treatments such as He, Ar or N_2.[82,83] Plasma treatment has also the peculiarity of avoiding the use of hazardous solvents without changing the morphology or the bulk properties of the materials treated.[84] Moreover, this method of surface modification can be used in combination with other well known chemical techniques in order to introduce biomolecules in general and carbohydrates in particular in a chemoselective approach. Two different approaches of the plasma grafting have been used with the final aim of introducing carbohydrates on the surface of different polymers. Bech and coworkers[85] described an interesting procedure to use the plasma grafting method in order to functionalize poly(ethylene terephtalate) (PET) fiber surfaces with monosaccharides. PET is often used in biomedical material as cardiovascular implants, artificial blood vessels, and in other applications due to its excellent mechanical properties and fair biocompatibility.[86–88] However, for biomedical applications, the polymer surfaces are desired to be highly hydrophilic and able to prevent protein adsorption.[89] The modification of polymer fibers to create new bioactive materials is beneficial for a wide range of applications. For example, the

current worldwide crisis in microbial infection and health care of wounds suggests that more research is necessary to understand practical and effective ways of creating safe bioactive textiles. The future development of biomedical and protective textiles with selective properties that benefit the consumer will be based on applying scientific and clinical advances in wound healing, antimicrobial, and enzyme-based fabrics.[90] In particular, PET fibers are particularly suited for the design of blood filters due to their ability to exhibit the highest leukocyte adhesion by the hydrophobic interaction between PET and integrins.[91] Moreover, PET fibers are often used in the biomedical field, as implants, vascular grafts for synthetic vessel replacement, or as cardiac heart valves;[92] among all synthetic fibers, PET fibers challenge cotton as the most common textile fiber, due to numerous advantages, including the cheapness and availability of the raw materials, the possibility to recycle PET fibers, the melt spinning process used for the production of PET fibers, which is clean and economical, the relatively high temperature resistance.

The plasma treatment alone gives surfaces that are relatively unstable (rapid loss of surface hydrophilicity), so one method is to treat the surface of PET with appropriate molecules to maintain the surface with a high level of hydrophilicity and functionality and to achieve a compatible interface with biological materials. A three steps strategy including two different argon plasma treatments was successfully used for PET biodecoration with monosaccharides (Fig. 8). During the first step (pre-activation), the plasma treatment generated radical species on PET surface that, once exposed to air, formed peroxide and hydroperoxide reactive groups. In the second step, the pre-activated fibers were dipped in different monosaccharide solutions, specifically, allyl α-D-galactopyranoside (AG) and 2-methacryloxyethyl glucoside (GEMA). A final plasma treatment allowed the polymerization of the monosaccharides present on the fiber surface. This method allows the control of the grafting density by varying the monosaccharide concentration and the mass of fiber exposed in the plasma

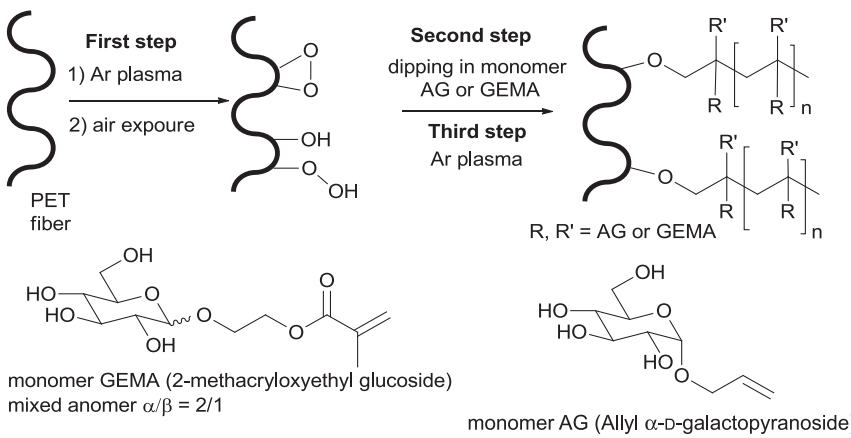


Fig. 8 Plasma grafting of glucose units on PET fibers.

reactor. More interesting results might be obtained with the grafting of oligosaccharides, which can be selectively recognized by target proteins.

A different strategy has been used by our group in a procedure, combining plasma treatment with click-chemistry technique (Fig. 9). In order to apply Huisgen cycloaddition to material modification, suitable methods for azido groups introduction are needed. Several strategies have been proposed for the preparation of azido-functionalized polymers; among them we can mention a general strategy consisting in the (co)polymerization of halide functionalized monomers followed by nucleophilic substitution of the halide groups by sodium azide (NaN_3).[93–95] Alternatively, the azido group can be introduced *via* epoxide opening by NaN_3 on the co-polymerized glycidyl-containing monomers (*i.e.* glycidyl methacrylate).[96] A second approach consists in the (co)polymerization of azido-prefunctionalized monomers. However, both approaches suffer from a strong limitation: the necessity of using only particular functionalized monomers (halide-, epoxide-, azido-functionalized), thus restricting the possibility of tailoring the chemical and physico-mechanical properties of the material, which could result, in this way, unsuitable for the desired application. Plasma treatment was used to introduce amino-groups on polypropylene (PP) membrane surfaces (preactivation). A diazo-transfer reaction transformed the amino groups into azido groups by triflyl azide solution. Finally

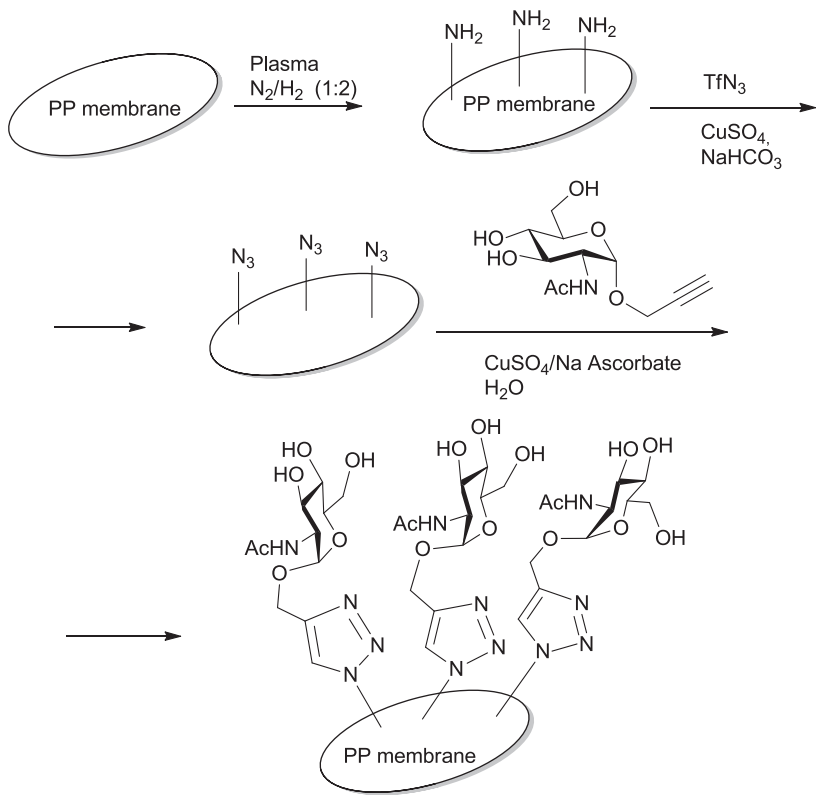


Fig. 9 Diazo-transfer methodology on plasma treated PP surfaces.

azido groups were used to introduce by click reaction a propargyl 2-acetamido-2-deoxy-α-D-glucopyranoside.

The proposed method is a proof of concepts that opens the way to the modification of material surfaces lacking reactive functional groups; it appears an ideal platform for subsequent 'click' reactions, that allows the decoration of the material with any compound, including oligosaccharides, presenting carbon-carbon triple bonds. The method can be of wide applicability, allowing the functionalization of materials of different nature with a wide variety of (bio)active compounds.

2.2 Material functionalization with glycidic structures *via* amide bond formation

A different approach capitalize on the amide bond formation between carboxyl and amino groups suitably located in a complementary position on the material and on the carbohydrate moieties, respectively.

An example of monosaccharide linkage to materials *via* amidation was proposed by our research group on hydroxyapatite.[97] Among various biocompatible materials, synthetic hydroxyapatite (HA) and carbonate hydroxyapatite (CHA) are widely used in many biomedical applications.[98] HA is the natural mineral ingredient of bones, tooth and calcified tissues in vertebrate. Synthetic HA and CHA are used for human implant coatings possessing a proven osteoconductivity and its geometry may deeply influence specific tissue responses, as it happens for vascular in-growth.[99] Nonetheless, researchers try to upgrade these inorganic materials by the addition of biological cues, activating peculiar specific cell-biomaterial interactions, thereby triggering additional, and possibly more specific biological responses. Unfortunately HA possess a paucity of reactive functional groups on its surface; therefore covalent linkage of biomolecules on this material is still a challenging task, despite the observation that the combination of inorganic materials, such as CHA, and organic signalling molecules is very promising in tissue engineering.[100,101] To date, main efforts on apatite covalent functionalization focused on the biodecoration of this inorganic material with whole proteins[102] or short peptide epitopes; the most commonly used peptide for surface modification is RGD,[103–105] the signalling domain derived from fibronectin and laminin.[106] Less exploited as cellular ligands for material functionalization are carbohydrate epitopes that might allow HA-cell interactions. In this context, the possibility to decorate hydroxyapatite with carbohydrates can be of relevant interest. Hence, hydroxyapatite granules (400–600 μm) were treated with aminopropyltriethoxysilane (APTES) in order to introduce free amino groups on HA surface (Fig. 10). The free amino functionalities were then condensed with a carboxyl group introduced onto a C-glycosyl scaffold bearing a carboxylic group. The method highlighted the possibility to link carbohydrates and their analogs to inorganic materials such as HA.

Amide bond formation was also used for the functionalization of chitosan as the scaffold with lactobionic acid.[107] Among natural biomaterials chitosan is a binary heteropolysaccharide consisting of (1→4)-linked 2-acetamido-2-deoxy-β-D-glucopyranose and 2-amino-2-deoxy-β-D-glucopyranose residues,[108,109] which has been widely applied in clinic and

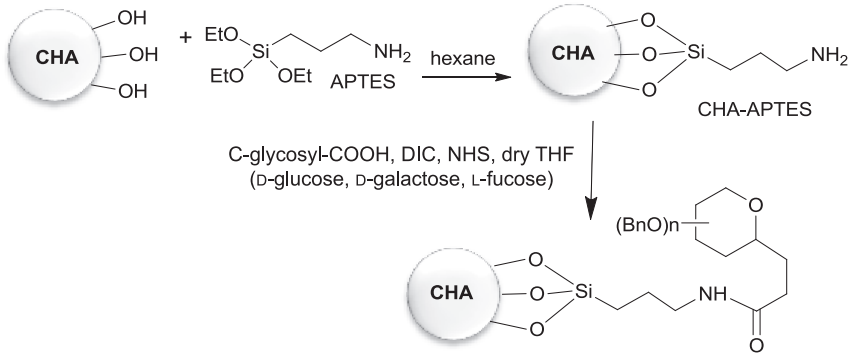


Fig. 10 HA functionalization *via* amide bond formation.

exhibited essential advantage in biodegeneration, biocompatibility and non-immunoreaction to human body when compared to synthetic polymers. The aim of this study was to develop a novel natural nanofibrous scaffold with surface-bound galactose ligands to enhance the bioactivity and mechanical stability of primary hepatocytes in culture for liver regeneration.[110] Primary hepatocyte culture plays an important role in the clinical treatment of hepatic failure patients[111] to avoid natural human immune response and provide sufficient replacement of the synthetic and metabolic functions of the liver, which has been applied in many parts of bioengineering field such as bioartificial liver-assist device (BLA)[112] and tissue engineering for liver regeneration.[113] It is well known that the glycidic part of glycolipids and glycoproteins strongly interacts with lectins present at cell surfaces:[114] In particular, β-galactose residues are recognized by asialoglycoprotein receptors on the surfaces of hepatocytes[115] and promote hepatocyte attachment.[56,116–120] Thus, many works were concerned with developing new scaffolds modified with galactose ligands (sse also next section) to enhance hepatocyte adhesion, maintenance of liver-specific functions and mechanical stability.[121–123]

The needed galactose unit was linked to chitosan through the modification of lactobionic acid (LA), as represented in Fig. 11; the coupling reaction was performed by an activation step using EDC and NHS.[124] The functionalization of chitosan with galactose improved the adhesion of hepatocytes on chitosan membrane underlining the bioactive role carried by the added carbohydrate. In another recent work[125] the same bioactive role of galactose has been exploited in order to differentiate mouse embryonic stem cells towards hepatocytes. Even in this brilliant work galactose units have been covalently linked to a *p*-vinylbenzylamine forming *N*-*p*-vinylbenzyl-O-β-D-galactopyranosyl-(1→4)-D-gluconamide by amidation reaction.[126]

Otherwise, the amide reaction can be used to bind carbohydrates to polysiloxanes (PDMS) in order to obtain "crew-cut" micelles proposed by many authors.[127,128] PDMS could be useful for cell immobilization, medical tubing, column materials, *etc*. Most of these research works focused on the polymers with C–C main chains bearing saccharide residues, with relatively few chemists paying attention to silicon backbone polymers. Compared

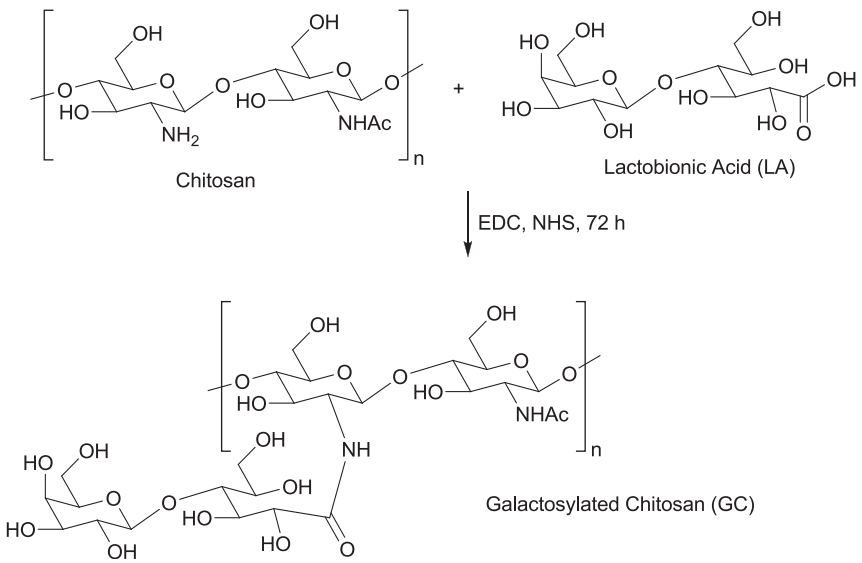


Fig. 11 Lactobionic grafting to chitosan *via* amide bond formation.

with the carbon chain polymers, polysiloxanes have remarkable properties such as flexibility, low cohesive energy, and biocompatibility.

2.3 Surface grafting by 2,4,6-trichloro-1,3,5-triazine (CY)

Gotoh and coworkers[129–131] extensively studied a biomaterial based on a glycoconjugate lactose-CY-silk fibroin (SF) with the aim to expose galactose units on the material for hepatocyte cultures. Silk fibroin (SF) is a natural fibrous protein created by the *Bombyx mori* silkworm, mainly consisting of the Gly–Ala–Gly–Ala–Gly–Ser sequence having its native molecular weight of 370 kDa.[132] SF fiber has been used as a textile material for a long time. Recently, it has been found that SF aqueous solution can be easily converted to gels,[133] sponge,[134] powder,[135] and membranes,[136] which have unique physicochemical properties and biocompatibilities. On the basis of these observations, SF has been reassessed as biotechnological and biomedical materials, in particular many current studies on SF explore SF-based scaffolds for tissue engineering.[137] Therefore, it is expected that the β-galactose non-reducing end of lactose moieties linked to SF may act as an efficient ligand for hepatocyte attachment to SF scaffolds.

The activated lactose-CY (Fig. 12) was synthesized by the reaction of the anomeric hydroxyl group with a chlorine atom of CY, then a second chlorine atom present on the coupling spacer reacts with the phenolic hydroxyl group of tyrosines residue or with the ε-amino group of lysine residues of the protein sequence. The reaction with the CY allows the correct exposure of the β-galactose moiety of the disaccharide on the surface of SF. The covalent linkage of lactose into SF was confirmed by physical methods and lectins recognition. The positive effect of galactose exposure on SF was observed by evaluation of the attachment of hepatocytes using the Lac–CY–SF conjugate-coated dishes, as preliminary tissue enginnering model.

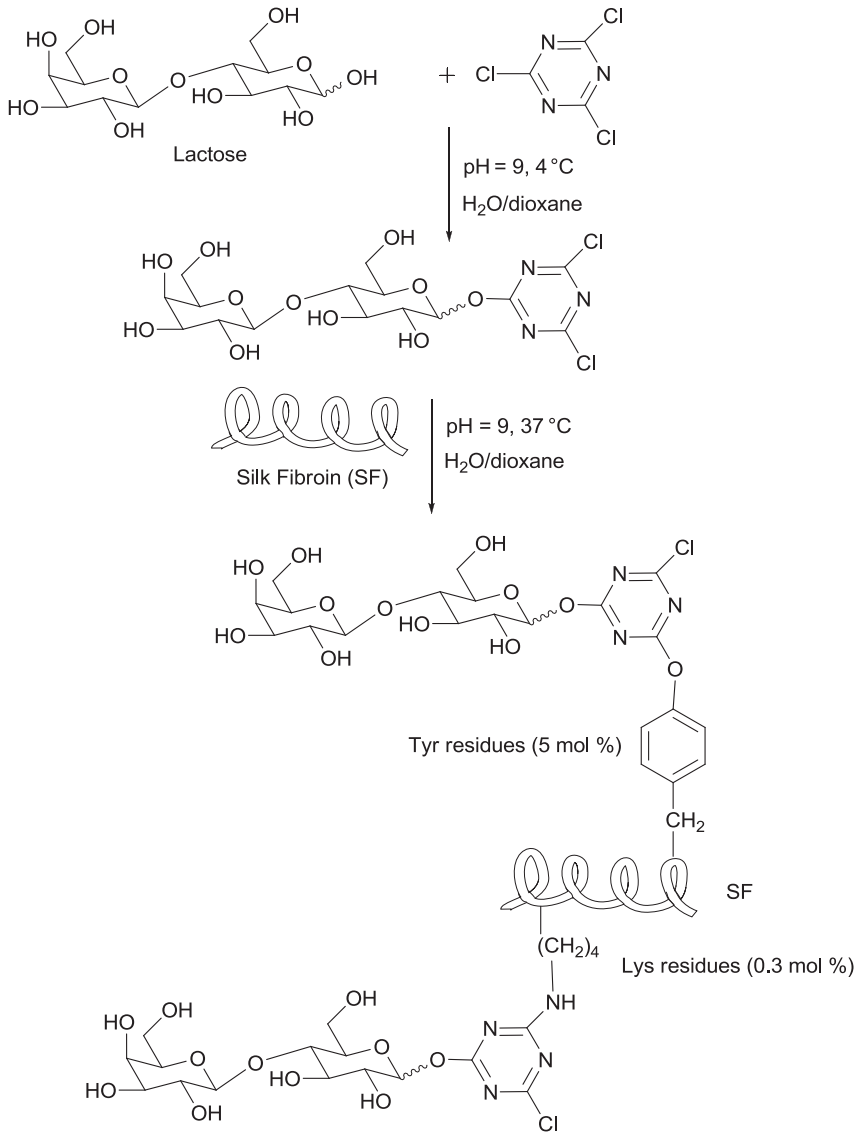


Fig. 12 Lactose conjugation to SF for the design of hepatocyte-specific biomaterials.

The same chemical methodology was used by Acharya *et al.*,[138] in order to remodel the properties of SF toward the adhesion of a different kind of cell, such as fibroblasts with encouraging results.

2.4 Material grafting by click chemistry

A high impact on the glyco-conjugation methodologies has been obtained by the introduction of the ''click-chemistry'' approach.[139] Although click reactions could be classified into four different categories, the copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) reaction represents the ground idea of a click reaction.[140] Nevertheless, the necessity of a more ''bio-safe'' environment has acted as a strong driving force for expansion of

the repertory of practical click reactions.[141] Nowadays, reactions such as the copper(I)-free azide–alkyne cycloaddition with more reactive alkynes,[142] the radical addition between thiols and alkenes or alkynes,[143,144] the Michael addition of thiols with maleimide, the nucleophilic substitution of the *para*-fluoro substituent of pentafluorophenyl groups, and the regular and inverse electron demand Diels–Alder reactions become to be considered as click reactions. In addition, click chemistry is becoming explored also by material scientists both for the processing of bulk materials[145,146] and for the functionalization of material surfaces.[147,148]

In a different approach, Buriak and co-workers[149] functionalized a stainless steel (SS) surface with galactose and *N*-acetylglucosamine moieties by a thiol-ene click reaction. Stainless steel has excellent physical properties such as flexibility and strength, together with appropriate chemical properties that are fundamental requirements for implanted biomedical devices. However, SS prostheses and devices are in constant contact with the aggressive body fluid, and often fail and finally fracture due to corrosion. To improve the biocompatibility and functionality of unmodified stainless steel implants, researchers have devoted substantial attention toward controlling and modifing their surface properties.[150,151]

The process proposed by Buriak and co-workers is illustrated in Fig. 13. A thin (sub-15 nm) and uniform silica coatings were first deposited by atomic layer deposition (ALD) on stainless steel in order to generate a

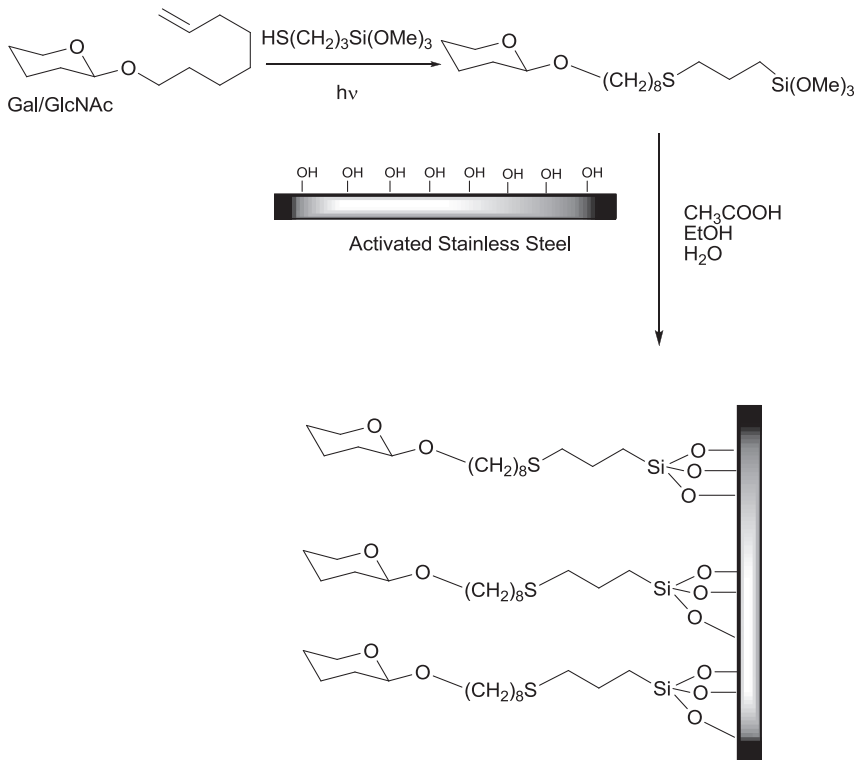


Fig. 13 Methodology for "glycodecoration" of stainless steel surface.

hydroxy-rich interlayer. These ALD coatings act as a platform on which alkoxysilane chemistry can take place. After that, the trialkoxysilane derivatives of the monosaccharides GlcNAc and Gal prepared *via* thiol-ene coupling were conjugated to the surface (Fig. 13). The carbohydrate moieties bound to the stainless steel surface were then detected *via* a complementary enzyme-linked lectin assay (ELLA). A key point of the process is the introduction of a high density of hydroxyl groups on the stainless steel surface in the first step of the functionalization strategy, in order to allow efficient formation of controllable organic monolayers through alkoxysilane coupling reactions.

''Photo-click' reactions were also developed, affording several additional advantages, such as spatio-temporal reaction control *via* focused UV-light irradiation. An example of this class of reaction is the UV light-promoted grafting method of a molecule containing a photoreactive aryl azide group coated on a polymer surface.[152,153] In the case of organic azides, upon photolysis with ultraviolet light, a nitrene radical is generated. The nitrene radical formed is extremely reactive and can undergo a multitude of reactions, for example, insertion into C-H, N-H, and O-H bonds, addition to olefins, proton abstraction reactions to give the corresponding amine, and in the case of aryl azides a number of ring expansion reactions have been observed. Although the precise mechanism of surface modification is not fully understood, this method has been demonstrated to be very effective for a wide range of surfaces. This approach has been used recently to covalently link carbohydrates to PET fibers by Renaudie *et al.*[154] for example by incorporating a sialyl-Lewis X saccharide as it is known to interact specifically with L-selectin.[155] Thus a new method for grafting of a new carbohydrate UV-reactive molecule, an azidophenyl lactamine (AzPhLac), Fig. 14, was proposed.

The phenylazido group on lactose by reductive amination of lactose with an aminated phenyl azide molecule, affording the β-D-galactopyranosyl-(1→4)-1-*N*-[2-(4-azidopnenylamino)ethylamino]-1-deoxy-D-glucitol (AzPhLac); the PET fibers were dipped (dip-coating method) in the AzPhLac aqueous solution, then dried and finally irradiated by UV light for the chemical fixation of the lactamine moiety *via* conversion of the phenylazido group to the highly reactive phenylnitrene. The complete surface modification is still not fully understood; the proposed mechanism by the authors includes also the formation of secondary amines. In addition, the highly reactive phenylnitrene can form covalent bonds not only with the hydrocarbon of the PET surface but also with linkages of AzPhLac molecules in its immediate surroundings, leading to the formation of a polymeric multilayer where the nitrene inserts onto the -OH groups. The method still has to be improved, since PET surface masking was observed during the UV-irradiation step.

The group of Bertozzi[156] proposed a brilliant example of Copper (I)-catalyzed azide-alkyne cycloaddition (CuAAC) for the chemoselective ligation of azide-functionalized pyrene and glycan moieties to the alkyne-functionalized focal point and chain ends of a dendritic scaffold in order to ''glycodecorate'' single-walled carbon nanotubes (SWNTs) by ultrasonication in water. The structural, mechanical, electrical, and optical properties

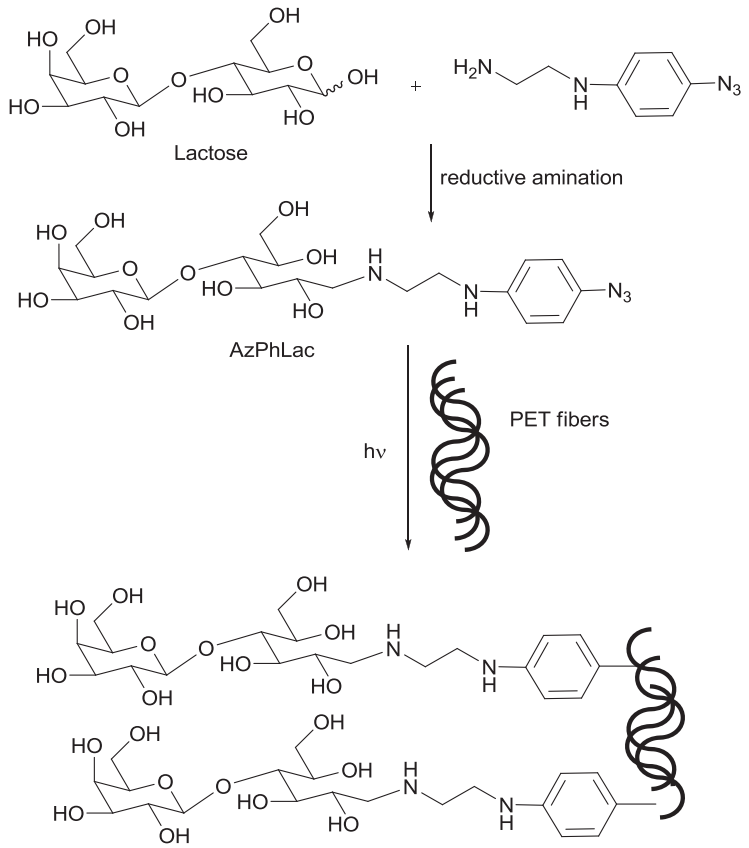


Fig. 14 Grafting by UV-irradiation of azidated lactose to PET fibers.

of single-walled carbon nanotubes (SWNTs) have stimulated considerable interest in their biological applications,[157,158] including biosensing, imaging, intracellular delivery, and cancer cell targeting. However, use of SWNTs in living systems requires strategies to diminish their cytotoxicity.[159] Thus, surface modifications that lower the toxicity of SWNTs while simultaneously enabling specific biological recognition are highly desirable.[160,161] A promising approach is to coat SWNTs with synthetic glycopolymers able to mimic the glycoproteins found on cell surfaces.

The proposed approach involves first the synthesis of dendritic structure ensuring the mimicry of glycoproteins and a suitable moiety for the anchoring to the nanotube surface. Thus, a pyrene tail was conjugated to the focal point of the dendrimer by CuAAC (Fig. 15). The resulting second generation dendrimer was further coupled with pent-4-ynoic anhydride to introduce additional alkyne groups. This third generation dendrimer was then reacted with a 2-azidoethyl mono- or disaccharide using CuAAC once again.

SWNTs coating with glycodendrimers resulted in a bioactive surface that also mitigate their cytotoxicity. The synthetic method proposed by Bertozzi can be readily adapted to ligands for other receptor interactions, in view of novel applications such as biosensors for carbohydrate-binding proteins and delivery agents that target specific cell-surface receptors.

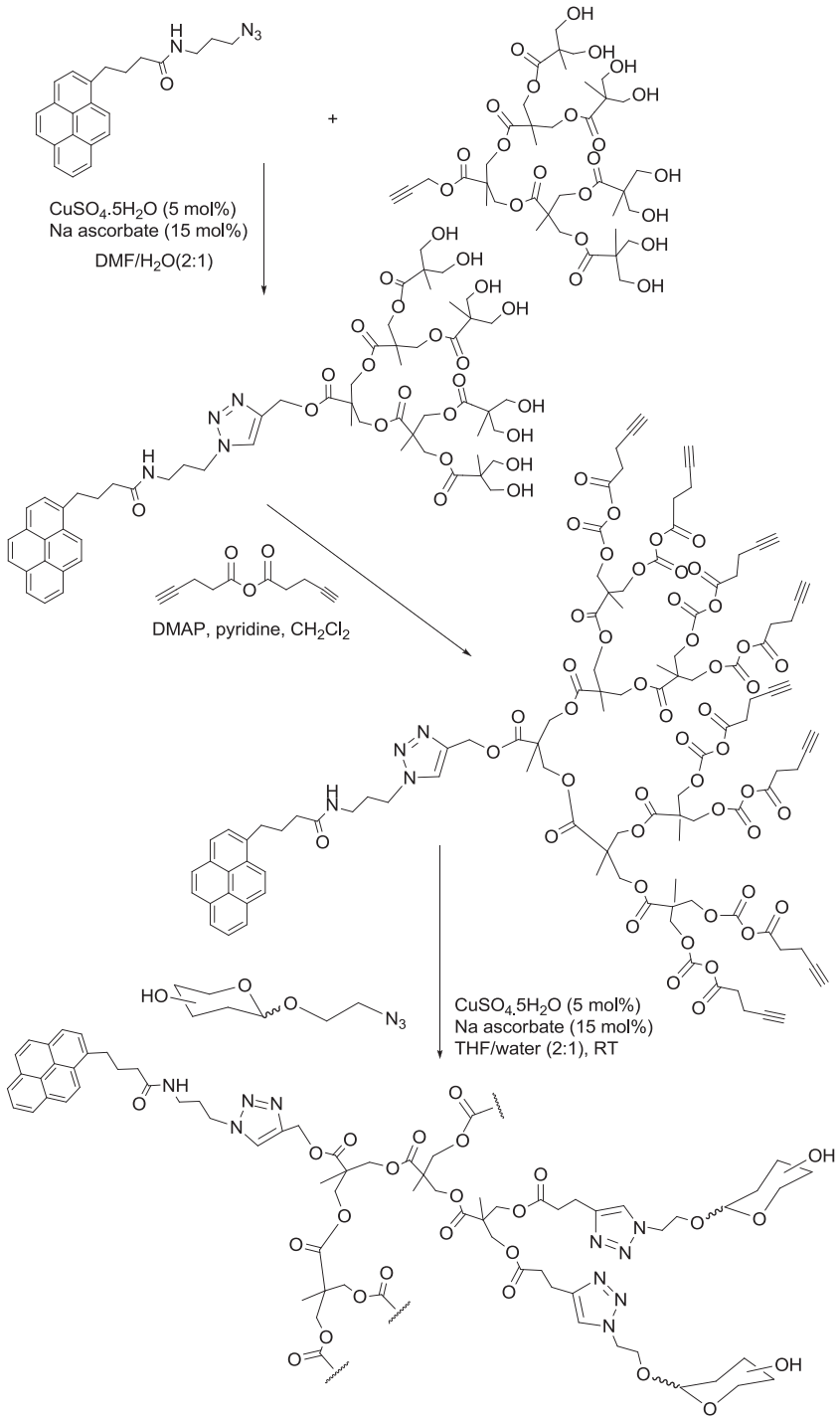


Fig. 15 Generation of glycodendrimers for carbon nanotubes functionalization.

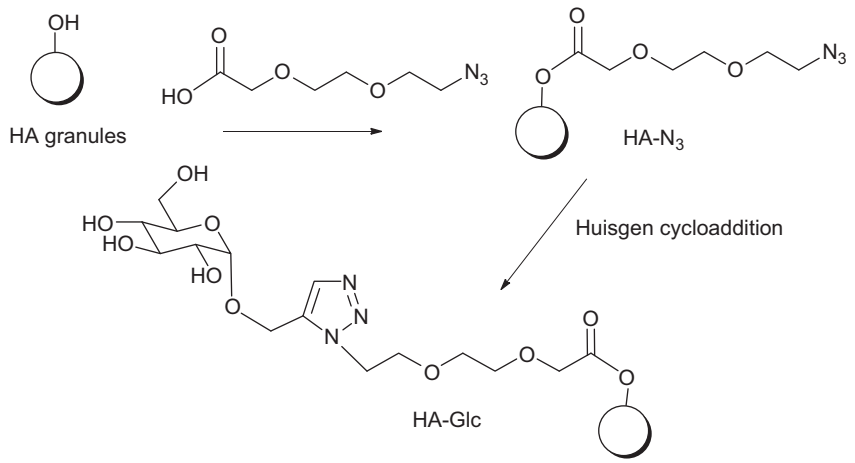


Fig. 16 HA functionalization by Huisgen cycloaddition.

The Huisgen-type cycloaddition was also proposed for the glycodecoration of hydroxyapatite (HA); in this example, a minimal spacer of difunctional PEG was used (Fig. 16), possessing an azido group at one term for the Huisgen cycloaddition, and a carboxyl group on the other side for the covalent bonding to the hydroxyl groups present on the HA surface.[162] The Huisgen cycloaddition was performed onto HA-N_3 with propargyl α-D-glucopyranoside. The "glycosylated" hydroxypatite was characterized by its ability to interact with glucose recognizing lectins.

Xue-Long Sun and co-workers[163] demonstrated the applicability of sequential Diels-Alder and azide-alkyne [3 + 2] cycloaddition reaction for the immobilization of carbohydrates onto a glass-slide surface. A polyethyleneglycol (PEG) linker carrying both an alkyne and a cyclopentadiene terminal groups was synthesized for the sequential click-reactions. The commercially available *N*-(ε-maleimidocaproyl) (EMC)-functionalized glass slide was used as the starting material bearing the dienophile on the material surface (Fig. 17). Alternatively, other groups have developed useful new bioorthogonal cycloaddition with different partner pairs, including azide-oxanorbornadiene,[164] nitrone cyclooctyne[165] and an extraordinarily rapid tetrazine–strained alkene reaction,[166] but none of these has been applied to carbohydrate functionalization of material surfaces.

Crucial to the selection of Diels-Alder cycloaddition as a preferential strategy for surface immobilization of biomolecules are the following key issues:

a) water is the best solvent for biomolecules;
b) water has an extraordinary rate-accelerating effect on the reaction process;
c) the reaction occurs efficiently at room temperature;
d) the diene and dienophiles of choice (cyclopentadiene and EMC) are independently stable and easy to introduce.

The first reaction takes place between the difunctional PEG and the maleimido group of the glass surface, affording the corresponding

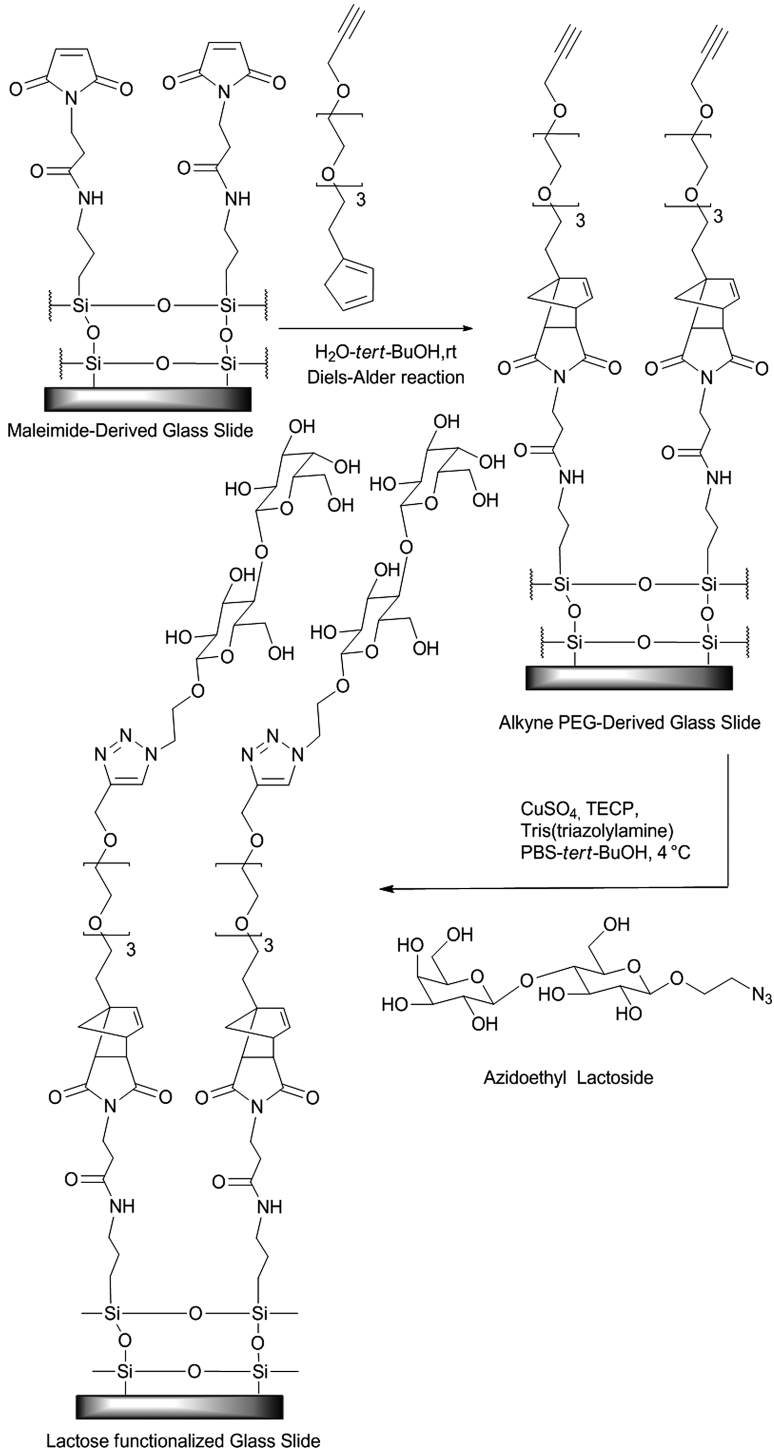


Fig. 17 Double click reaction for the functionalization of glass slides.

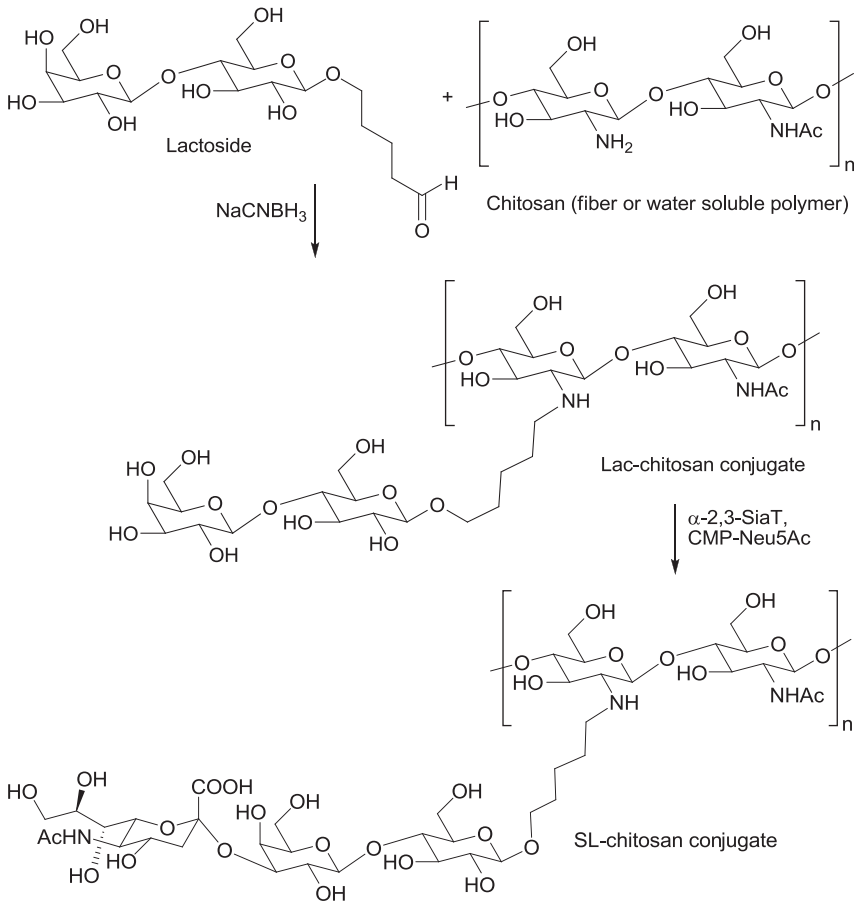


Fig. 18 SL conjugation method to chitosan-based materials.

unsaturated six-membered ring (Fig. 17, step 1). This reaction occurs at room temperature without any influence by solvent and pH. In order to bio-functionalize the PEG surface with carbohydrates, a second step involving a Huisgen type cycloaddition with a suitable azidolactoside (Fig. 17, step 2) was performed. The process efficacy of the coupled click reactions was determined with FITC-galectin and with the use of azidated biotin as the alkyne partner in the second step.

2.5 Miscellaneous grafting methods

A very intersting and up-to-date example is given by the functionalization with carbohydrates of chitosan-based material proposed as anti-viral agents. In fact, the coupling of glycoligands to material surfaces may generate appealing multivalent systems that could be used as diagnostic or therapy tools (theragnostics).

Influence virus start the infection of host cells by its surface hemagglutinin (HA) protein that binds to its glycoligands, such as sialyllactose (SL).

The group of Cheng and co-workers[167] presented the construction of two SL-incorporated chitosan-based materials, either as a water-soluble polymer or as a functional fiber, and demonstrated their abilities for viral adhesion inhibition and decontamination.

The syntheses were accomplished by grafting a lactoside bearing an aldehyde-functionalized aglycone to the amino groups of chitosan or chitosan fiber followed by the enzymatic sialylation with sialyltransferase (Fig. 18). The obtained water-soluble SL – chitosan conjugate was shown to bind hemagglutinin with high affinity and inhibited effectively the viral attachment to host erythrocytes. Moreover, the SL functionalized chitosan fiber was able to efficiently remove the virus from an aqueous medium. The results collectively demonstrate that these potential new materials may function as the virus adsorbents for prevention and control of influenza. In addition, this example represents an appealing approach for presenting a protein ligand on a chitosan backbone, which is a versatile molecular platform for biofunctionalization and, thereby, can be used for not only antiviral design, but also extensive medical development such as diagnosis, drug delivery and biomaterial for tissue engineering applications.

Conclusions

Few examples of materials biofunctionalized with carbohydrates have appeared recently in the literature with the aim of ameliorating materials features, in terms of bioactivity, biocompatibility, and biostability. Despite the increasing development of glycomics and the methods at our disposal for oligossacharide synthesis, application of glycoscience for the design of biomaterials for tissue engineering is still scarce. We think that in the next few years a great contribution by glycochemists will be given to this research field, that may experience a great explosion as observed for glycosylated nanoparticles and glycoarrays.

References

1 N. Huebsch and D. J. Mooney, *Nature*, 2009, **462**, 426.
2 M. J. Lysaght, A. Jaklenec and E. Deweerd, *Tissue Eng Part A*, 2008, **14**, 305.
3 US Department of Health and Human Services. *2020: A New Vision - A Future for Regenerative Medicine*, 2006, http://www.hhs.gov/reference/newfuture.shtml
4 Bouchie, *Nature Biotechnol*, 2002, **20**, 1178.
5 E. S. Place, N. D. Evans and M. M. Stevens, *Nature Mat*, 2009, **8**, 457.
6 B. D. Ratner and S. J. Bryant, *Annu Rev Biomed Eng*, 2004, **6**, 41.
7 J. M. Anderson, A. Rodriguez and D. T. Chang, *Semin Immunol*, 2008, **20**, 86.
8 G. Balasundaram and T. J. Webster, *J Mater Chem*, 2006, **16**, 3737.
9 S. C. G. Leeuwenburgh, J. A. Jansen, J. Malda, A. D. Wouter, J. Rouwkema, C. A. Van Blitterswijk, C. J. Kirkpatrick and D. F. Williams, *Biomaterials*, 2008, **29**, 3047.
10 R. Langer and D. Tirrell, *Nature*, 2004, **428**, 487.
11 A. Hubbel, J. Patterson and M. Martino, *Materials Today*, 2010, **13**, 14.
12 U. Mayer and R. Timpl, *Nidogen: A versatile binding protein of basement membranes. In Extracellular Matrix Assembly and Structure*, 1994, (P. D. Yurchenco, D. E. Birk and R. P. Mecham eds.) Academic Press, San Diego, pp. 318–416.

13 J. C. M. van Hest and D. A. Tirrell, *Chem Commun*, 2001, **19**, 1897.
14 W. P. Daley, S. B. Peters and M. Larsen, *J Cell Sci*, 2008, **121**, 255.
15 N. Zagris, 2001. *Micron: Extracellular matrix in development of the early embryo* (Oxford, England, 1993), 32, pp. 427–438.
16 D. Gullberg and P. Ekblom, *Int J Dev Biol*, 1995, **76**, 845.
17 J. E. Scott, *J Ant*, 1995, **187**, 259.
18 M. Aumailley and B. Gayraud, *J Mol Med*, 1998, **76**, 253.
19 E. I. Wallner, Q. Yang, D. R. Peterson, J. Wada and Y. S. Kanwar, *Am J Physiol*, 1998, **275**, 467.
20 D. J. Behonick and Z. Werb, *Mech Dev*, 2003, **120**, 1327.
21 M. M. Stevens and J. H. George, *Science*, 2005, **310**, 1135.
22 L. Bacáková, E. Filova, F. Rypacek, V. Svorcik and V. Stary, *Physiol Res*, 2004, **53**, 35.
23 J. Gao, Y. M. Kim, H. Coe, B. Zern, B. Sheppard and Y. Wang, *PNAS*, 2006, **103**, 16681.
24 V. H. Fan, A. Au, K. Tanama, R. Littrell, L. B. Richardson, J. W. Wright, A. Wells and L. G. Griffith, *Stem Cells*, 2007, **25**, 1241.
25 A. Sahni, L. A. Sporn and C. W. Francis, *J Biol Chem*, 1999, **274**, 14936.
26 N. Gomez, Y. Lu, S. Chen and C. E. Schmidt, *Biomaterials*, 2007, **28**, 271.
27 K. Alberti, R. E. Davey, K. Onish, S. George, K. Salchert, F. P. Seib, M. Bornhauser, T. Pompe, A. Nagy, C. Werner and P. W. Zandstra, *Nat Methods*, 2008, **5**, 645.
28 D. F. Williams, *Biomaterials*, 2009, **30**, 5897.
29 C. Yang, P. J. Hillas, J. A. Baez, M. Nokelainen, J. Balan, J. Tang, R. Spiro and J. W. Porarek, *BioDrugs*, 2004, **18**, 103.
30 M. P. Lutolf and J. A. Hubbell, *Nature Biotechnol.*, 2005, **23**, 47.
31 T. Xu, N. Zhang, H. L. Nichols, D. Shi and X. Wen, *Mater Sci Eng C*, 2007, **27**, 578.
32 K. Langer, C. Coester, C. Weber, H. von Briesen and J. Kreuter, *Eur J Pharm Biopharm*, 2000, **49**, 303.
33 J. M. Goddard and J. H. Hotchkiss, *Prog Pol Sci*, 2007, **32**, 698.
34 K. J. Whittlesey and L. D. Shea, *Exp Neurol*, 2004, **109**, 1.
35 C. A. Simmons, E. Alsberg, S. Hsiong, W. J. Kim and D. J. Mooney, *Bone*, 2004, **35**, 562.
36 H. C. Kroese-Deutman, P. Q. Ruhe, P. H. Spauwen and J. A. Jansen, *Biomaterials*, 2005, **26**, 1131.
37 M. P. Lutolf, G. P. Raeber, A. H. Zisch, N. Tirelli and J. A. Hubbell, *Adv Mater*, 2003, **20**, 5671.
38 Y. Zhang, D. B. Moheban, B. R. Conway, A. Bhattacharyya and R. A. Segal, *J Neurosci*, 2000, **20**, 5671.
39 P. R. Kuhl and L. G. Griffith-Cima, *Nat Med*, 1996, **2**, 1022.
40 K. Kirkwood, B. Rheude, Y. J. Kim, K. White and K. C. Dee, *J Oral Implantol*, 2003, **29**, 57.
41 V. Karageorgiou, L. Meinel, S. Hofmann, A. Malhotra, V. Volloch and D. Kaplan, *J Biomed Mater Res A*, **71**, 528.
42 A. H. Zisch, M. P. Lutolf, M. Ehrbar, G. P. Raeber, S. C. Rizzi, N. Davies, H. Schmökel, D. Bezuidenhout, V. Djonov, P. Zilla and J. A. Hubbel, *FASEB J*, 2003, **17**, 2260.
43 A. C. Lee, V. M. Yu, J. B. Lowe, M. J. Brenner, D. A. Hunter, S. E. Mackinnon and S. E. Sakiyama-Elbert, *Exp Neurol*, 2003, **184**, 295.
44 S. J. Taylor, J. W. McDonald and S. E. Sakiyama-Elbert, *J Control Relase*, 2004, **98**, 281.

45 D. Seliktar, A. H. Zisch, M. P. Lutolf, J. L. Wrana and J. A. Hubbell, *J Biomed Mater Res A*, 2004, **68**, 704.
46 I. D. Dinbergs, L. Brown and E. R. Edelman, *J Biol Chem*, 1996, **271**, 29822.
47 L. Russo, S. Zanini, C. Riccardi, F. Nicotra and L. Cipolla, *Mat Today*, 2011, **14**, 164.
48 G. Goobes, P. S. Stayton and G. P. Drobny, *Progr. Nuclear Magn. Res. Spectr.*, 2007, **50**, 71.
49 G. Goobes, R. Goobes, W. J. Shaw, J. M. Gibson, J. R. Long, V. Raghunathan, O. Schueler-Furman, J. M. Popham, D. Baker, C. T. Campbell, P. S. Stayton and G. P. Drobny, *Magn. Reson. Chem.*, 2007, **45**, S32.
50 U. Hersel, C. Dahmen and H. Kessler, *Biomaterials*, 2003, **24**, 4385.
51 P. Dubruel, E. Vanderleyden, M. Bergadà, I. De Paepe, H. Chen, S. Kuypers, J. Luyten, J. Schrooten, L. Van Hoorebeke and E. Schacht, *Surface Science*, 2006, **600**, 2562.
52 M. D. Mager, V. LaPointe and M. M. Stevens, *Nature Chem*, 2011, **3**, 582.
53 R. J. Linhardt, T. Toida, *Carbohydrates in Drug Design*, 1997, (Witczak, Z. J., Nieforth, K. A., eds.), Marcel Dekker: New York, pp. 277–341.
54 E. A. Balazs, T. C. Laurent, *The Chemistry*, *Biology and Medical applications of hyaluronan and its derivatives*, 1998, (Laurent, T. C., ed.), London, Portland Press, pp. 325–336.
55 W. I. Weis and K. Drickamer, *Annu Rev Biochem*, 1996, **65**, 441.
56 J. A. Oka and P. H. Weigel, *J Cell Biol*, 1986, **103**, 1055.
57 L. G. Griffith and S. Lopina, *Biomaterials*, 1998, **19**, 979.
58 Y. Miura, H. Sato, T. Ikeda, H. Sugimura, O. Takai and K. Kobayashi, *Biomacromolecule*, 2004, **5**, 1708.
59 C. Love, J. L. Ronan, G. M. Grotenbreg, A. G. van der Veen and H. L. Ploegh, *Nat Biotechnol*, 2006, **24**, 703.
60 Y. Umezewa and H. Aoki, *Anal Chem*, 2004, **76**, 320.
61 U. Lindahl, K. Lidholt, D. Spillmann and L. Kjellèn, *Thrombosis Res*, 1994, **75**, 1.
62 H. Sato, Y. Miura, N. Saito, K. Kobayashi and O. Takai, *Surface Science*, 2007, **601**, 3871.
63 A. B. Foster, *Ann Rev Biochem*, 1961, **30**, 45.
64 H. Verli and J. A. Guimaraes, *Carbohydr Res*, 2004, **339**, 281.
65 K. S. Siow, L. Britcher, S. Kumar and K. J. Griesser, *Plasma Process Polym*, 2006, **3**, 392.
66 F. Poncin-Epaillard, G. J. Legeay, *Biomater Sci Polym Ed*, 2003, **14**, 1005.
67 J. Wang, C. J. Pan, N. Huang, H. Sun, P. Yang, Y. X. Leng, J. Y. Chen, G. J. Wan and P. K. Chu, *Surface and Coatings Technology*, 2005, **196**, 307.
68 K. Kato, E. Uchida, Y. Uyama and Y. Ikada, *Prog Polym Sci*, 2003, **28**, 209.
69 C. Oehr, M. Müller, B. Elkin, D. Hegemann and U. Vohrer, *Surf. Coat. Technol.*, 1999, **3**, 25.
70 J. Wang, C. J. Pan, N. Huang, H. Sun, P. Yang, Y. X. Leng, J. Y. Chen, G. J. Wan and P. K. Chu, *Surface and Coatings Technology*, 2005, **196**, 307.
71 S. Hu, X. Ren, M. Bachman, C. E. Sims, G. P. Li and N. Allbritton, *Anal Chem*, 2002, **74**, 4117.
72 G. Oye, V. Roucoules, L. J. Oates, A. M. Cameron, N. R. Cameron, P. G. Steel, B. G. Davis, D. M. Coe, R. A. Cox and J. P. S. J. Badyal, *Phys Chem B*, 2003, **107**, 3496.
73 M. E. Ryan, A. M. Hynes and J. P. S. Badyal, *Chem Mater*, 1996, **8**, 37.
74 C. Tarducci, E. Kinmond, S. Brewer, C. Willis and J. P. S. Badyal, *Chem Mater*, 2000, **12**, 1884.

75 S. J. Hutton, J. M. Crowther and J. P. S. Badyal, *Chem Mater*, 2000, **12**, 2282.
76 C. Tarducci, W. C. E. Schofield, S. Brewer, C. Willis and J. P. S. Badyal, *Chem Mater*, 2001, **13**, 1800.
77 D. O. H. Teare, D. C. Barwick, W. C. E. Schofield, R. P. Garrod, L. J. Ward and J. P. S. Badyal, *Langmuir*, 2005, **21**, 11425.
78 C. Tarducci, W. C. E. Schofield, S. A. Brewer, C. Willis and J. P. S. Badyal, *Chem Mater*, 2002, **14**, 2541.
79 C. Tarducci, S. A. Brewer, C. Willis and J. P. S. Badyal, *Chem Commun*, 2005, **3**, 406.
80 S. R. Coulson, I. S. Woodward, S. A. Brewer, C. Willis and J. P. S. Badyal, *Langmuir*, 2000, **16**, 6287.
81 R. Q. Kou, Z. K. Xu, H. T. Deng, Z. M. Liu, P. Seta and Y. Y. Xu, *Langmuir*, 2003, **19**, 6869.
82 D. Hegemann, H. Brunner and C. Oehr, *Nuclear Instruments and Methods in Physics Research Section B*: *Beam Interactions with Materials and Atoms*, 2003, **208**, 281.
83 J. López García, A. Asadinezhad, J. Pacherník, M. Lehocký, I. Junkar, P. Humpolíček, P. Sáha and P. Valášek, *Molecules*, 2010, **15**, 2845.
84 K. S. Siow, L. Britcher, S. Kumar and H. J. Griesser, *Plasma Process Polym*, 2006, **3**, 392.
85 L. Bech, B. Lepoittevin, A. Achhab, E. Lepleux, L. Teulè-Gay, C. Boisse-Laporte and P. Roger, *Langmuir*, 2007, **23**, 10348.
86 G. Legeay, F. Poncin-Epaillard and C. R. Arciola, *Int J Artif Organs*, 2006, **29**, 453.
87 K. Sugiyama, K. Kato, M. Kido, K. Shiraishi, K. Ohga, K. Okada and O. Matsuo, *Macromol Chem Phys*, 1998, **199**, 1201.
88 M. W. Huh, I. K. Kang, D. H. Lee, W. S. Kim, D. H. Lee, L. S. Park, K. E. Min and K. H. J. Seo, *J Appl Polym Sci*, 2001, **81**, 2769.
89 H. T. Deng, Z. K. Xu, Z. W. Dai, J. Wu and P. Seta, *Enzyme Microbiol Technol*, 2005, **36**, 996.
90 G. Sun, *Bioactive Fibres and Polymers*, 2001, (J. V. Edwards and T. L. Vigo, eds.), 792, American Chemical Society, Washington, D.C., pp. 243–252.
91 G. Reese, *In Modern Polyesters: Chemistry and Technology of Polyesters and Copolyesters*, 2003, (J. Scheirs, T. E. Long, eds.), Wiley, Chichester, pp. 401–433.
92 L. Barbe, B. Boval, M.-P. Wautier and J.-L. Wautier, *Transfusion*, 2000, **40**, 1250.
93 R. T. Chen, B. W. Muir, G. K. Such, A. Postma, R. A. Evans, S. M. Pereira, K. M. McLean and F. Caruso, *Langmuir*, 2010, **26**, 3388.
94 J. F. Lutz, H. G. Börner and K. Weichenhan, *Macromol Rapid Commun*, 2005, **26**, 514.
95 R. S. Lee and Y. T. Huang, *J Polym Sci Part A: Polym Chem*, 2008, **46**, 4320.
96 N. V. Tsarevsky, S. A. Bencherif and K. Matyjaszewski, *Macromolecules*, 2007, **40**, 4439.
97 M. Sandri, A. Natalello, D. Bini, L. Gabrielli, L. Cipolla and F. Nicotra, *Synlett*, 2011, **13**, 1845.
98 S. V. Dorozhkin, *Biomaterials*, 2010, **31**, 1465.
99 V. Karageorgiou and D. Kaplan, *Biomaterials*, 2005, **26**, 5474.
100 L. L. Hench and J. M. Polak, *Science*, 2002, **295**, 1014.
101 H. S. Mansur and H. S. Posta, *Chem Eng J*, 2007, **137**, 72.
102 K. Zurlinden, M. Laub and H. P. Jennissen, *Materialwiss Werkst*, 2005, **36**, 820.

103 G. Balasundaram, M. Sato and T. J. Webster, *Biomaterials*, 2006, **27**, 2798.
104 M. Durrieu, S. Pallu, F. Guillemot, R. Bareille, J. Amédée and C. H. Baquey, *J Mater Sci-Mater M*, 2004, **15**, 779.
105 M. Nelson, M. Balasundaram and T. J. Webster, *Int J Nanomed*, 2006, **1**, 339.
106 U. Hersel, C. Dahmen and H. Kessler, *Biomaterials*, 2003, **24**, 4385.
107 Z. Q. Feng, X. Chu, N. P. Huang, T. Wang, Y. Wang, X. Shi, Y. Ding and Z. Z. Gu, *Biomaterials*, 2009, **30**, 2753.
108 B. T. Houseman and M. Mrksich, *Chem Biol*, 2002, **9**, 443.
109 B. T. Houseman, E. S. Gawalt and M. Mrksich, *Langmuir*, 2003, **19**, 1522.
110 H. Lis and N. Sharon, *Chem Rev*, 1998, **98**, 637.
111 J. K. Park and D. H. Lee, *J Biosci Bioeng*, 2005, **99**, 311.
112 A. J. Strain and J. M. Neuberger, *Science*, 2002, **295**, 1005.
113 I. K. Kang, J. S. Moon, H. M. Jeon, W. Meng, Y. I. Kim, Y. J. Hwang and S. Kim, *J Mater Sci Mater Med*, 2005, **16**, 533.
114 A. Varki, *Glycobiology*, 1993, **3**, 97.
115 G. Ashwell and J. Harford, *Ann Rev Biochem*, 1982, **51**, 531.
116 P. H. Weigel, R. L. Schnaar, M. S. Kuhlenschmidt, E. Schmell, R. T. Lee, Y. C. Lee and S. Roseman, *J Biol Chem*, 1979, **254**, 10830.
117 A. Kobayashi, T. Akaike, K. Kobayashi and H. Sumitomo, *Chem Rapid Commun*, 1986, **7**, 645.
118 A. Kobayashi, M. Goto, K. Kobayashi and T. Akaike, *J Biomater Sci Polym*, 1994, **6**, 325.
119 J. Yang, M. Goto, H. Ise, C.-S. Cho and T. Akaike, *Biomaterials*, 2002, **23**, 471.
120 J. J. Yoon, Y. S. Nam, J. H. Kim and T. G. Park, *Biotechnol Bioeng*, 2002, **78**, 1.
121 K. A. Peterlinz, R. M. Georgiadis, T. M. Herne and M. J. J. Tarlov, *Am Chem Soc*, 1997, **119**, 3401.
122 M. Schena, D. Shalon, R. W. Davis and P. O. Brown, *Science*, 1995, **270**, 467.
123 R. Georgiadis, K. P. Peterlinz and A. W. Peterson, *J Am Chem Soc*, 2000, **122**, 3166.
124 I. K. Park, J. Yang, H. J. Jeong, H. S. Bom, I. Harada, T. Akaike, S. I. Kim and C. S. Cho, *Biomaterials*, 2003, **24**, 2331.
125 Q. Meng, A. Haque, B. Hexiq and T. Akaike, *Biomaterials*, 2012, **33**, 1414.
126 A. Kobayashi, K. Kobayashi and T. Akaike, *J Biomater Sci Polym Edn*, 1992, **3**, 499.
127 A. Halperin, M. Tirrell and T. P. Lodge, *Adv Polym Sci*, 1992, **100**, 31.
128 L. Zhang and A. Elisenberg, *Science*, 1995, **268**, 1728.
129 Y. Gotoh, M. Tsukada, S.-I. Aiba and N. Minoura, *Int J Biol Macromol*, 1996, **18**, 19.
130 Y. Gotoh, N. Minoura and T. Miyashita, *Colloid Polym Sci*, 2002, **280**, 562.
131 Y. Gotoh, S. Niimi, T. Hayakawa and T. Miyashita, *Biomaterials*, 2004, **25**, 1131.
132 C. M. Mello, K. Senecal, B. Yeung, P. Vouros, D. Kaplan, *Silk polymers: materials science and biotechnology*, 1994, (D. Kaplan, W. W. Adams, B. Farmer and C. Viney eds), American Chemical Society, Washington DC, pp. 67–79.
133 T. Hanawa, A. Watanabe, T. Tsuchiya, R. Ikoma, M. Hidaka and M. Sugihara, M, *Chem Pharm Bull*, 1995, **43**, 284.
134 M. Li, S. Lu, Z. Wu, H. Yan, J. Mo and L. Wang, *J Appl Polym Sci*, 2001, **79**, 2185.
135 D. Akiyama, Y. Kitahara, L. Xuan, M. Murakami, M. Arai and K. Hirabayashi, *J Soc Fiber Sci Tec*, 1991, **47**, 339.

136 N. Minoura, M. Tsukada and M. Nagura, *Biomaterials*, 1990, **11**, 430.
137 S. Sofia, M. B. McCarthy, G. Gronowicz and D. L. Kaplan, *J Biomed Mater Res*, 2001, **54**, 139.
138 C. Acharaya, B. Hinz and S. C. Kundu, *Biomaterials*, 2008, **29**, 4665.
139 C. Remzi Becer, R. Hoogenboom and U. S. Schubert, *Angew Chem Int Ed*, 2009, **48**, 4900.
140 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew Chem Int Ed*, 2001, **40**, 2004.
141 E. M. Sletten and C. R. Bertozzi, *Angew Chem Int Ed*, 2009, **48**, 6974.
142 M. F. Debets, C. W. J. van der Doelen, F. P. J. T. Rutjes and F. L. van Delft, *ChemBioChem*, 2010, **11**, 1168.
143 A. Dondoni, *Angew Chem Int Ed*, 2008, **47**, 8995.
144 R. Hoogenboom, *Angew Chem Int Ed*, 2010, **49**, 3415.
145 M. Heggli, N. Tirelli, A. Zisch and A. Hubbell, *Bioconjugate Chem*, 2003, **14**, 967.
146 D. Fournier, R. Hoogenboom and U. S. Schubert, *Chem Soc Rev*, 2007, **36**, 1369.
147 F. L. van Delft, J. C. M. van Hest, R. J. M. Nolte, A. E. Rowan and F. P. J. T. Rutjes, *QSAR, Comb Sci*, 2007, **26**, 125.
148 F. Santoyo-Gonzalez and F. Hernandez-Mateo, *Chem Soc Rev*, 2009, **38**, 3449.
149 A. M. Slaney, V. A. Wright, P. J. Meloncelli, K. D. Harris, L. J. West, T. L. Lowary and J. M. Buriak, *ACS App Mat Inter*, 2011, **3**, 1601.
150 G. Sydow-Plum and M. Tabrizian, *Mater Sci Technol*, 2008, **24**, 1127.
151 P. Thevenot, W. Hu and L. Tang, *Curr Top Med Chem*, 2008, **8**, 270.
152 M. A. Harmer, *Langmuir*, 1991, **7**, 2010.
153 T. Sugawara and T. Matsuda, *Macromolecules*, 1994, **27**, 7809.
154 L. Renaudie, C. L. Narvor, E. Lepleux and P. Roger, *Biomacromolecules*, 2007, **8**, 679.
155 J. E. Puksas and Y. Chen, *Biomacromolecules*, 2004, **5**, 1141.
156 P. Wu, X. Chen, N. Hu, U. C. Tam, O. Blixt, A. Zetti and C. R. Bertozzi, *Angew Chem Int Ed*, 2008, **47**, 5022.
157 D. Tasis, N. Tagmatarchis, A. Bianco and M. Prato, *Chem Rev*, 2006, **106**, 1105.
158 R. S. Kane and A. D. Stroock, *Biotechnol Prog*, 2007, **23**, 316.
159 V. L. Colvin, *Nat Biotechnol*, 2003, **21**, 1166.
160 X. Chen, G. S. Lee, A. Zettl and C. R. Bertozzi, *Angew Chem Int Ed*, 2004, **43**, 6111.
161 X. Chen, U. C. Tam, J. L. Czlapinski, G. S. Lee, D. Rabuka, A. Zettl and C. R. Bertozzi, *J Am Chem Soc*, 2006, **128**, 6292.
162 L. Russo, E. Landi, A. Tampieri, A. Natalello, S. M. Doglia, L. Gabrielli, L. Cipolla and F. Nicotra, *Carbohydrate Research*, 2011, **346**, 1564.
163 X. L. Sun, *Bioconjugate Chem*, 2006, **17**, 52.
164 S. S. van Berkel, A. T. Dirks, M. F. Debets, F. L. van Delft, J. J. Cornelissen, R. J. Nolte and F. P. Rutjes, *Chembiochem*, 2007, **8**, 1504.
165 C. S. McKay, J. Moran and J. P. Pezacki, *Chem Commun*, 2010, **46**, 931.
166 M. L. Blackman, M. Royzen and J. M. Fox, *J Am Chem Soc*, 2008, **130**, 13518.
167 X. Li, P. Wu, G. F. Gao and S. Cheng, *Biomacromolecules*, 2011, **12**, 3962.