

Modern NMR Approaches to the Structure Elucidation of Natural Products

Volume 2: Data Acquisition and Applications to Compound Classes

Modern NMR Approaches to the Structure Elucidation of Natural Products

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by

Antony J. Williams

ChemConnector, Inc., USA

Email: tony27587@gmail.com

Gary E. Martin

Merck & Co., Inc., USA

Email: gary.martin2@merck.com

and

David Rovnyak

Bucknell University, USA

Email: drovnyak@bucknell.edu



Print ISBN: 978-1-84973-393-9
Two-volume set print ISBN: 978-1-84973-459-2
PDF eISBN: 978-1-84973-468-4
EPUB eISBN: 978-1-78801-072-6

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

© The Royal Society of Chemistry 2017

All rights reserved

Apart from fair dealing for the purposes of research 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 the copyright owner, or in the case of reproduction in accordance with the terms of 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.

The RSC is not responsible for individual opinions expressed in this work.

The authors have sought to locate owners of all reproduced material not in their own possession and trust that no copyrights have been inadvertently infringed.

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

Registered Charity Number 207890

Visit our website at www.rsc.org/books

Foreword

The complex process of the structure elucidation of natural products continues to be one of the most intensely practiced areas of scientific activity. Whether the chemical entity originates as a result of chemical (total) synthesis or represents a purified or semi-purified isolate from natural sources, a *correct* structural verification, including precise stereochemical analysis, remains of critical scientific concern. The effective use of contemporary and evolving nuclear magnetic resonance (NMR) methodologies represents the central tool for the complete establishment of chemical structure. Volume 2 of the series *Modern NMR Approaches to the Structure Elucidation of Natural Products* continues in the same spirit as the treatise that was begun in Volume 1 on contemporary NMR techniques applied to the natural product structural elucidation process.

Volume 2 is composed of 14 chapters and, like Volume 1, is divided into two parts. Following a brief introductory section, which includes a discussion of the overall organization and philosophy of Volume 2, Part 1 of Volume 2 consists of six chapters. These include a detailed account of the respective NMR topics superimposed on the overall theme of these chapters, which emphasizes the successful experimental implementation of the NMR experiments being discussed. Chapter 1 is a concisely written contemporary overview of the nuclear Overhauser effect (nOe) from both a theoretical and an experimental perspective, including descriptions of the two-dimensional NOESY and ROESY experiments. These are traditionally employed for establishing spatial relationships within the proton spin systems *via* through-space dipolar interactions. In addition, there is an introduction to the implementation of a new technique—NMR analysis of molecular flexibility in solution or NAMFIS—and its roll in natural product structure elucidation. The experiments discussed in Chapter 1 and their role in elucidating the correct structure and stereochemistry of natural products cannot be overstated.

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

Chapter 2 consists of a discussion of the determination of both the relative (*J*-based) and absolute stereochemistry of natural products and the NMR experiments used for obtaining this important structural information, including the use of the Mosher ester methodology. Chapter 2 also includes a brief treatment of a cutting-edge advance afforded by a newly developed technique for applying residual chemical shift anisotropy (RCSA) for stereochemical assignment. Going forward, it is quite likely that RCSA measurements will be regularly combined with nOe/rOe (Chapter 1) and residual dipolar coupling (RDC) measurements (Chapter 4). Chapter 3 describes the applications of ^{15}N -NMR spectroscopy in natural product structure elucidation, includes a discussion of the basic experimental parameters associated with ^{15}N -NMR data acquisition and details the experimental implementation of ^{15}N -NMR in the determination of natural product structures. The importance of being able to incorporate results derived from ^{15}N -NMR into the structural elucidation process greatly increases the prospects of arriving at a correct structure. Chapter 4 introduces the theory and applications of RDC, which may be viewed as an extension of or a complimentary technique to the nOe determination (Chapter 1), and its role in natural product structural elucidation is also discussed. Also included are details regarding the experimental implementation of RDC, including sample preparation, data interpretation and a section on the caveats associated with RDC. This will be an important NMR technique in application for small molecule natural products. The use of RDC is also discussed in Chapters 2 and 8. Chapter 5 is an overview of contemporary high-resolution mass spectrometry and its role in natural product structural elucidation. This chapter highlights the fact that, while NMR techniques tend to dominate much of the natural product structure elucidation process, when taken together with the requisite structural information derived from mass spectrometry, the combination provides a powerful one-two punch for producing correct structure determination. It is for this reason that such a chapter is included here. The contents of Chapter 6 focus on recent and future developments in NMR methodologies, with numerous examples illustrating experimental results in terms of structure. Furthermore, this chapter describes completely new NMR experiments, which may ultimately have a significant impact on the structural elucidation process. This chapter is well written and serves to illustrate cutting-edge NMR techniques that will evolve into the mainstream of the structural elucidation process.

Part 2 of Volume 2 consists of eight chapters. The recurring theme here that is superimposed on the material in Part 2 is the application of combined NMR techniques, which is illustrated with appropriate examples. This constitutes a powerful way of introducing useful strategies for achieving success in the structure elucidation process to the spectroscopist or natural product chemist. The chapters contain useful examples and specific case studies for some of the more important natural product classes. Chapter 7 is a well-written chapter and discusses the practical applications of NMR experiments with examples of structure elucidation drawn from the terpene class of natural products (monoterpenes, sesquiterpenes and higher

terpenes). While these structures may not appear to be very complex on the surface, they nevertheless pose their own set of problems and challenges. Chapter 8 describes NMR experiments that are applied to the structure determination of steroids, one of the most studied (by NMR) classes of natural products. Some specific experiments with examples are discussed, including structural details involving ^{19}F -NMR. Chapters 9 and 10 detail NMR experiments that pertain to the structure determination of alkaloids. The strategies employed for structure elucidation are approached in a systematic manner and are illustrated with good examples. ^{15}N -NMR is of considerable utility for providing unambiguous structures of nitrogen-containing natural products. Chapters 11 and 12 provide additional case studies of selected molecules derived from marine natural products and antimicrobial areas, respectively. Because of the structural complexity of many of these molecules, different experimental strategies may be required for full structure elucidation. Each structure may require a unique set of experiments and, in this regard, one size does not fit all! Chapter 13 contains examples of the saponins, which are complex triterpene glycosides. These examples contain considerable structural complexity in that, since they are glycosides, it is necessary to determine the structure of both the aglycone component and also to determine the structural nature of the carbohydrate component of the saponin. Each of these two structural components may require different and unique strategies in order to assess their respective structures and, correspondingly, different NMR experiments for the determination of each structural component within the same molecule may be necessary. This is especially true since the carbohydrate component is usually a complex polysaccharide, typically containing more than one sugar moiety with differing branch points for the sugar residues. Chapter 14 is a summary perspective on both volumes and discusses the future directions of natural product structure elucidation activity.

The collective contents of the two-volume series *Modern NMR Approaches to the Structure Elucidation of Natural Products* present very useful and practical information for the practicing NMR spectroscopist and natural product chemist to implement in their own laboratories. All of the chapters are well written by respected experts in the field. The considerable volume of experimental detail contained in the chapters, together with structure elucidation strategies, when implemented, is certain to contribute to an overall improvement in the speed and accuracy of natural product structures being elucidated.

David C. Lankin

For the most important women in my life—Eirlys, Rae and AnnMarie.
Thank you for your love and support—Antony J. Williams

Gary E. Martin dedicates this volume to his granddaughter, Margaux Electa
Simone Martin, in the hope that she finds science as fascinating as
her grandfather has.

With gratitude to my parents, James and Virginia—David Rovnyak

Contents

Part 1

Chapter 1	Application of the Nuclear Overhauser Effect to the Structural Elucidation of Natural Products	3
	<i>R. R. Gil and A. Navarro-Vázquez</i>	
1.1	The Historical Origin of the Term “Nuclear Overhauser Effect”	3
1.2	The Theory Behind the NOE	6
1.3	Multispin Systems	15
1.4	The Kinetic NOE	16
1.5	Transient NOE	17
1.6	ROE and ROESY	18
1.7	Scalar Couplings—Zero-quantum Filters	21
1.8	NOE Experiments for the Structural Elucidation of Dimeric Compounds	24
1.9	Fully Quantitative NOE in Small Molecules	26
1.10	NOE Deconvolution—NAMFIS	30
1.11	Conclusions	35
	References	36
Chapter 2	Assigning Molecular Configuration by Nuclear Magnetic Resonance	39
	<i>Kirk R. Gustafson, Susanna T. S. Chan and Yizhou Liu</i>	
2.1	Assigning Relative Configuration	39
2.1.1	Coupling Constants and NOEs	40

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

2.1.2	Acetonides	43
2.1.3	<i>J</i> -Based Analysis—The Murata Method	44
2.1.4	Universal Database	51
2.1.5	Computational Methods	55
2.1.6	Residual Chemical Shift Anisotropy	56
2.2	Assigning Absolute Configuration	60
2.2.1	Chiral Anisotropic Reagents	60
2.3	Conclusions	66
	References	68

Chapter 3 Nuclear Magnetic Resonance Experiments Applicable to the Elucidation and Characterization of Nitrogenous Natural Products: ^1H - ^{15}N Heteronuclear Shift Correlation Methods

71

Josep Saurí, Antony J. Williams and Gary E. Martin

3.1	^1H - ^{15}N Direct and Long-range Heteronuclear Shift Correlation	71
3.2	^{15}N Chemical Shift Referencing	72
3.3	The Range of ^{15}N Chemical Shifts	73
3.4	^{15}N Pulse Widths	74
3.5	^{15}N Chemical Shift Prediction	76
3.5.1	Validating ^{15}N Chemical Shift Prediction	80
3.5.2	Setting F_1 Spectral Windows	82
3.5.3	Structure Verification using a ^{15}N Content Database	83
3.5.4	Enhancing ^{15}N Chemical Shift Prediction with a “User-trained” Database	83
3.6	Computer-assisted Structure Elucidation—The Impact of ^{15}N Data	84
3.7	$^1J_{\text{NH}}$ and $^nJ_{\text{NH}}$ Coupling Constants	86
3.8	^1H - ^{15}N Results Obtained Using Various Experiments	86
3.8.1	^1H - ^{15}N Experiments Employing an Initial $^1J_{\text{NH}}$ Magnetization Transfer	87
3.8.2	^1H - ^{15}N Experiments Employing an Initial $^nJ_{\text{NH}}$ Magnetization Transfer	91
3.8.3	^1H - ^{15}N HSQMBC	95
3.8.4	^1H - ^{15}N HSQMBC-TOCSY: A Hyphenated Long-range ^1H - ^{15}N Experiment	98
3.8.5	Triple-resonance ^1H - ^{15}N Experiments	107
3.9	Conclusions	113
	References	113

Chapter 4	Application of Residual Dipolar Couplings to the Structural Analysis of Natural Products	117
	<i>R. R. Gil and A. Navarro-Vázquez</i>	
4.1	Introduction	117
4.2	Alignment Media	130
4.3	The Nature of the Alignment Process	142
4.4	How RDCs are Measured	143
4.5	Data Analysis—How the RDCs are used in the Structural Analysis of Small Molecules	148
4.5.1	Possible Scenarios	148
4.5.2	How to Manage Experimental Uncertainties in RDC Analysis	153
4.5.3	The Conformational Problem	156
4.5.4	How to Handle Symmetrical Rotors	158
4.5.5	Software	159
4.6	RDCs in the Structural Elucidation of Natural Products	159
4.6.1	Dimeric Products	166
4.7	Conclusions	172
	References	172
Chapter 5	Applications of High-resolving Power, High-accuracy Mass Spectrometry for the Structural Elucidation of Natural Products	177
	<i>Gary Kruppa and Wendy Zhong</i>	
5.1	Introduction	177
5.1.1	Mass Spectrometry for the Structural Elucidation of Unknowns	177
5.1.2	Brief Description of FT-ICR	185
5.2	Examples from Literature	190
5.2.1	Ultra-high Resolving Power to Separate the Isotopic Fine Structures	190
5.2.2	Study of Biosynthesis Processes by Measuring Stable Isotopic-labeled Precursors using FT-ICR MS	194
5.2.3	Metabolite Profiling of Triterpene Saponins by Combined Accurate Mass Measurement and MS ⁿ Experiments using LC FT-ICR MS	195
5.3	Conclusions	197
	References	197

Chapter 6	Current Pulse Sequence Developments in Small-molecule Nuclear Magnetic Resonance Spectroscopy	199
	<i>Teodor Parella</i>	
6.1	Introduction	199
6.2	Fast NMR	202
6.3	Pure Shift NMR	206
6.3.1	The ZS Experiment	209
6.3.2	Sensitivity, Strong Coupling Effects, and Spectral Quality	210
6.3.3	ZS Applications	216
6.3.4	Implementing Homodecoupling in 1D and 2D NMR Experiments	218
6.4	Perfect NMR	227
6.4.1	Perfect Spin-echo	228
6.4.2	Using ZQFs	233
6.5	Ultra-long-range Correlation NMR	234
6.5.1	Long-range HSQMBC	236
6.5.2	HSQMBC-TOCSY	237
6.5.3	ADEQUATE	240
6.6	Future Perspectives	243
	Acknowledgements	244
	References	244

Part 2

Chapter 7	Terpenes: Mono-, Sesqui-, and Higher Terpenes	253
	<i>William F. Reynolds and Raul G. Enriquez</i>	
7.1	Introduction	253
7.2	Types of Terpenes and Related Compounds	254
7.3	Basic Approach to Terpene Structure Elucidation by NMR Spectroscopy	257
7.4	Avoiding Determining the Wrong Structure	261
7.5	Specialized Techniques that are Useful for Different Types of Terpenes	267
7.6	Conclusions	272
	Acknowledgements	272
	References	273

<i>Contents</i>	xiii
Chapter 8 Nuclear Magnetic Resonance of Steroids	275
<i>Martin Jaeger and Ruud L. E. G. Aspers</i>	
8.1 Prologue	275
8.2 An Introduction to Steroids	275
8.3 Modern and “Rare” NMR Methods in the Steroid Field	284
8.3.1 Recent General NMR Developments	284
8.3.2 Covariance NMR and Steroids	284
8.3.3 The HSQC-TOCSY Experiment	287
8.3.4 ^{13}C -Detected Experiments	287
8.3.5 1D and High-resolution 1D ^1H Methods	288
8.3.6 ^{19}F NMR	290
8.3.7 Residual Dipolar Couplings	294
8.4 Considerations	298
8.5 Conclusions	307
Acknowledgements	308
References	308
Chapter 9 Nuclear Magnetic Resonance Experiments Applicable to the Elucidation and Characterization of Alkaloid Structures Part I: Direct ^1H–^{13}C Heteronuclear Shift Correlation and Establishing Contiguous Protonated Carbon Spin Systems	315
<i>Josep Saurí and Gary E. Martin</i>	
9.1 Introduction	315
9.2 “First-tier” NMR Methods for Alkaloid Structure Characterization	316
9.2.1 Sample Preparation	317
9.2.2 Probe Selection	317
9.2.3 Parameter Choices	318
9.3 Acquiring NMR Spectra	319
9.4 1D Reference Spectra	319
9.5 Fundamental or “First-tier” 2D NMR Spectra	322
9.5.1 Conventional HSQC	322
9.5.2 Multiplicity-edited HSQC	324
9.5.3 Incorporating BIRD-based Homonuclear Decoupling in the HSQC Experiment: PS-HSQC	324
9.5.4 Summary of One-bond Heteronuclear Correlation Methods	327

9.6	Intentionally Folding HSQC Spectra in Limited-sample Situations	329
9.7	Non-uniform Sampling	332
9.8	Long-range Heteronuclear Shift Correlation	335
9.8.1	HMBC	335
9.8.2	LR-HSQMBC	338
9.9	Defining the Proton–Proton Connectivity Network	342
9.9.1	COSY and TOCSY Spectra	342
9.9.2	HSQC-TOCSY Spectra	346
9.9.3	HSQMBC-COSY and HSQMBC-TOCSY Spectra	349
9.9.4	Strategy for Employing Hyphenated 2D Experiments	354
	References	354

Chapter 10 Nuclear Magnetic Resonance Experiments Applicable to the Elucidation and Characterization of Alkaloid Structures Part II: Advanced Techniques for the Identification of Adjacent Carbons Using H2BC, 1,1-ADEQUATE, and Variants 358

Josep Sauri and Gary E. Martin

10.1	Identification of Adjacent Carbons Using H2BC and 1,1-ADEQUATE Data	358
10.1.1	Identification of Adjacent Protonated Carbons Using the H2BC Experiment	359
10.1.2	Identifying Adjacent Carbons Using 1,1- and 1,1-HD-ADEQUATE	361
10.2	1, <i>n</i> -ADEQUATE and Advanced Variants	371
10.2.1	1, <i>n</i> -ADEQUATE	373
10.2.2	1, <i>n</i> -HD-ADEQUATE	374
10.2.3	Edited 1, <i>n</i> -HD-ADEQUATE	378
10.3	Examples of the Application of Advanced NMR Methods in Structure Elucidation Studies	380
10.3.1	Staurosporine—Is the Crew’s Rule Obsolete?	383
10.3.2	Cryptospirolepine—Resolution of a Long-standing Structural Ambiguity Using 1,1- and 1, <i>n</i> -HD-ADEQUATE Spectra	386
10.3.3	Eudistidine-C	397
10.4	Conclusions	399
	References	400

Chapter 11 Nuclear Magnetic Resonance Case Studies in Marine Natural Products **403**

Angela A. Salim, Andrew M. Piggott and Robert J. Capon

11.1	Introduction	403
11.1.1	Case Studies	404
11.1.2	Snapshots	405
11.2	Marine Natural Product Case Studies	405
11.2.1	Trachycladindoles	405
11.2.2	Franklinolides	406
11.2.3	Bistelletazines	409
11.2.4	Phorbasins	412
11.2.5	Ircinialactams	417
11.2.6	Fascioquinols	420
11.3	Marine Natural Product Snapshots	426
	References	438

Chapter 12 Nuclear Magnetic Resonance Case Studies in Microbial Natural Products **440**

Andrew M. Piggott, Angela A. Salim and Robert J. Capon

12.1	Introduction	440
12.1.1	Case Studies	441
12.1.2	Snapshots	441
12.2	Microbial Natural Product Case Studies	442
12.2.1	Aspergillazines	442
12.2.2	Nocardioazines	444
12.2.3	Heronamides	447
12.2.4	Kibdelones	452
12.2.5	Reveromycins	457
12.3	Microbial Natural Product Snapshots	463
12.3.1	Macrocyclic Lactones	463
12.3.2	Cyclic Peptides	463
12.3.3	Caged Structures	463
12.3.4	Miscellaneous	475
12.4	Conclusion	483
	References	483

Chapter 13 Nuclear Magnetic Resonance in Saponin Structure Elucidation **486**

Antonio Jorge Ribeiro da Silva, Ricardo Moreira Borges and Vitor Soares

13.1	Introduction	486
------	--------------	-----

13.2	General Considerations	487
13.3	Structure and Relative Stereochemistry of the Aglycone	489
13.4	Identity of Component Monosaccharides	493
13.5	Linkage Position and Sequence of the Oligosaccharide Chain	497
13.6	Concluding Remarks	499
13.6.1	Experimental	500
	References	500
Chapter 14	Increasing the Adoption of Advanced Techniques for the Structure Elucidation of Natural Products	502
	<i>Antony J. Williams, Gary E. Martin and David Rovnyak</i>	
	References	506
	Subject Index	508

Part 1

CHAPTER 1

Application of the Nuclear Overhauser Effect to the Structural Elucidation of Natural Products

R. R. GIL^{*a} AND A. NAVARRO-VÁZQUEZ^{*b}

^a Carnegie Mellon University, Department of Chemistry, 4400 Fifth Avenue, Pittsburgh, PA 15213-3890, USA; ^b Institute of Organic Chemistry and Institute for Biological Interfaces Karlsruhe, Institute of Technology (KIT), Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany

*Email: rgil@andrew.cmu.edu; armando.deus@gmail.com

1.1 The Historical Origin of the Term “Nuclear Overhauser Effect”

Albert W. Overhauser, while holding a postdoctoral position at the Department of Physics of the University of Illinois at Urbana (IL), theoretically proposed in 1953 that if the electron spin resonance of the conduction electrons in metals is saturated, the nuclei of the metal will be polarized to the same degree they would be if their gyromagnetic ratio were that of the electron spin.¹ The nuclear spins would exhibit an enhanced polarization by a factor of $\cong 1000$. Such a significant degree of polarization of the metal nuclei predicted by Overhauser was soon after experimentally demonstrated by Carver and Slichter in a sample of metallic lithium.² This is the concept that led to a new field of study known as dynamic nuclear polarization

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

(DNP),^{3,4} in which the electron spin polarization is transferred to nuclei in order to significantly improve their sensitivity. Although not the topic of this chapter, it is important to highlight that DNP is slowly becoming a very powerful tool in the structure analysis of molecules in high-resolution, solid-state nuclear magnetic resonance (NMR) spectroscopy,^{5,6} and led Professor Overhauser to receive the 1994 National Medal of Science.

Note, however, that the nuclear Overhauser effect (nOe), which is the nuclear–nuclear analog of the original electronic Overhauser effect, was first observed by I. Solomon in 1955.⁷ Following the studies on nuclear spin relaxation previously published by Bloembergen, Purcell, and Pound on a single-spin system, Solomon published a seminal article entitled “Relaxation Processes in a System of Two Spins”. He experimentally observed that the longitudinal magnetization of a dipolar-coupled pair of “unlike” spins (different resonance frequencies) does not show simple exponential decay as in the case of a single-spin system. The experiments were done in an anhydrous sample of hydrofluoric acid. Upon saturation (steady-state Overhauser effect) or inversion (transient Overhauser effect) of the ^1H resonance, a maximum increase of $\sim 30\%$ of the ^{19}F signal was observed. The same effect was observed on the proton signal when the ^{19}F resonance was excited in a similar way. Although not explicitly stated in Solomon’s paper, this was the first reported experimental observation of the Overhauser effect between nuclei, and the experiment was later called the “nuclear Overhauser effect” and gave origin to the nOe or NOE acronym. This acronym can also be found in the literature standing for “nuclear Overhauser enhancement” due to the fact that the experiment produces enhancements of the NMR signals. We particularly prefer to use the term “effect”. Solomon also set the theoretical basis of the NOE. He describes in his paper a set of modified Bloch equations that gives the correct equation of motion of the macroscopic magnetic moments for a system of two spins under the influence of dipole–dipole interaction. These equations are known today as the Solomon equations; the diagrams representing the Zeeman levels for a two-spin system are correspondingly known as the Solomon diagrams. Later in the chapter, we will come back to these equations and diagrams.

Intramolecular as well as intermolecular proton–proton NOEs were first reported by Reinhold Kaiser in 1963⁸ and 1965, respectively.⁹ In the former article, Kaiser reported how the Overhauser effect can be used to determine the relative sign of J coupling constants, and to assign spectral lines to transitions between energy levels in *trans*-crotonaldehyde and *m*-dinitrobenzene. In the second article, Kaiser reported the intermolecular NOE in the liquid state between the protons of chloroform and cyclohexane.

Anet and Bourn introduced the first application of the NOE to the configurational and conformational analysis of small organic molecules in a seminal paper published in 1965.¹⁰ In this ground-breaking paper, they observed NOE enhancements when irradiating a sample of 3-methyl-but-2-enoic

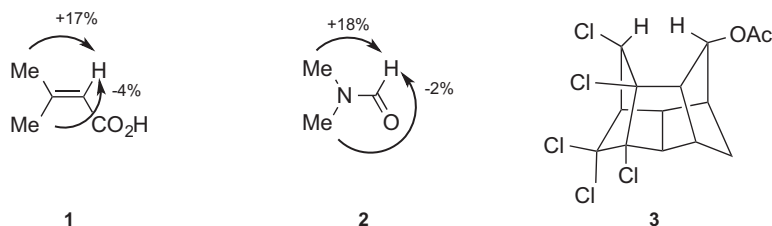


Figure 1.1 First reported intramolecular NOEs by Anet and Bourn.¹⁰

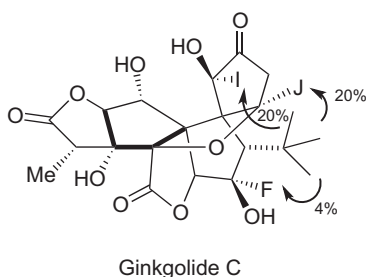


Figure 1.2 First intramolecular NOEs in a natural product. This is the ginkgolide C reported in 1967 by Nakanishi and colleagues.¹¹

acid (1) and dimethylformamide (2), as well as in the half-cage acetate 3 (Figure 1.1). Prior to this article, the assignment of proton resonances in high-resolution spectra was performed on the basis of chemical shifts and J couplings. Shortly thereafter, Nakanishi applied intramolecular NOE analysis to the structural determination of the ginkgolides, a series of polyhydroxylated terpenoids. Irradiation of the ginkgolide C *t*-butyl group resulted in clear intensity enhancements that allowed the determination of the configuration of the related stereogenic centers (Figure 1.2).¹¹ In the following year, the application of the NOE to taxane derivatives was also reported from the group of Nakanishi.¹² With the addition of the NOE to the arsenal of analytical tools, it quickly became evident to chemists that NMR would become one of the most powerful tools for the structural analysis of molecules of any size. The application of the NOE in structural biology is outside the scope of this chapter, and most of the theoretical and practical aspects of the NOE in organic molecules have been covered in the monograph of Neuhaus and Williamson,¹³ as well as other reviews.^{14,15} It is the purpose of this chapter to illustrate the use of not-so-common techniques, mainly involving quantitative aspects of the NOE for the determination of configuration and conformation in organic molecules. It will be assumed in our discussions that concepts such as chemical shift, J coupling, and NMR pulses, as well as the fundamentals of the T_1 and T_2 relaxation processes, are known by the reader.

1.2 The Theory Behind the NOE

In spite of the impressive technological advances that involve very high magnetic fields, cryoprobes, and digital consoles, among others, compared to other spectroscopic techniques, NMR is a horribly insensitive technique due to the very low energy difference between the spin energy states. For protons, the energy difference between the α ($m = +1/2$) and β ($m = -1/2$) states at 500 MHz (magnetic field of 11.75 Tesla) is merely 3.32×10^{-25} Joules, leading to a Boltzmann population distribution of $N_\alpha/N_\beta = 1.000080$ (*i.e.* only 80 out of a million spins in excess at the α lower-energy state). That is why any ingenious experiments that help to significantly increase this population difference are always very welcome, and the Overhauser effect is also used for this purpose, apart from being a powerful technique in structural analysis.

Another consequence of such a low resonance energy is the very weak interaction of the nuclear spins with their surrounding medium, leading sometimes to considerably longer relaxation times. As stated by Abragam, the probability W per unit time of the nuclear spin-1/2 transition from the β into the α state through spontaneous emission of a photon appears to be a negligible phenomenon. For a ^1H spin resonating at 500 MHz, W would be $\approx 10^{-25} \text{ s}^{-1}$. In addition, Abragam also showed that the transition probability for an induced emission or absorption under the effects of a radiation field involving photons would also be very small ($W \approx 10^{-10} \text{ s}^{-1}$), concluding that the coupling with the radiation field is “hopelessly inadequate” as a viable relaxation mechanism. Hence, what would be a viable relaxation mechanism for nuclear spins?

In fact, since the only property of spin-1/2 nuclei that depends on orientation is their magnetic moment, transitions between nuclear spins can only be induced by magnetic fields. Let us first analyze the effect of a radio-frequency (rf) pulse on the macroscopic magnetization (M_z) of a sample of spin-1/2 exposed to a magnetic field B_0 . At thermal equilibrium, the Boltzmann distribution will produce a population excess in the α state for nuclei with a positive gyromagnetic ratio, as is the case for protons. The sum of the projections of the individual spins' magnetic moments onto the z -axis will build the macroscopic magnetization M_z . As we know from basic NMR concepts, no net magnetization is observed in the x, y -plane due to the random nature of the phase of the individual magnetic moments precessing around the axis of the applied magnetic field B_0 . Since M_z is a macroscopic property, like temperature, it can be rotated at our whim around the rf field B_1 applied on the x -axis. Rotation at our whim means that we can take M_z from $+z$ to $+y$ (90° rotation). We can stop here, or we can continue to $-z$ (180° rotation), or we can even continue to a total rotation of 360° and bring M_z back to $+z$ as it was before we turned the rf on. If we pay attention to what happens to the populations as we rotate M_z , we will see that at a 90° or 270° rotation, the populations are equalized ($N_\beta = N_\alpha$) since there is no longer net magnetization in the z -axis; at 180° , the populations are inverted ($N_\beta > N_\alpha$); and after a total rotation of 360° , the populations are back in thermal equilibrium. This simple example experimentally demonstrates that a

fluctuating magnetic field can stimulate the absorption of energy not only by the spins, but also by their relaxation. It is important to keep in mind that these processes only take place if the magnetic field oscillates at the same frequency as the spins.

In a real sample, the spins are not alone. They are surrounded by other spins from the same molecule, by solvent molecules, by impurities, and even by dissolved oxygen gas. These surroundings are known as the “lattice”, a term originating from the early studies of relaxation in solids, where the surroundings were a real solid lattice, and we can say that the spins are coupled and in equilibrium with this lattice. If the spins are driven out of equilibrium by excitation with rf power, once the rf power is turned off, they return to equilibrium by interacting with the lattice. As a result of the Brownian motions in the lattice, molecules randomly reorient in liquid solution, producing fluctuating magnetic fields. These oscillating magnetic fields are not coherent along the whole volume of the sample. Each spin in the sample is affected in a different way as a result of the random nature of the reorientation. However, the frequency of these fluctuating fields will depend on the correlation time (τ_c) of the molecules. The correlation time is defined as the time it takes a molecule to reorient 1 radian. The inverse of the correlation time ($1/\tau_c$) represents the rotational molecular frequency (ω) in radians per second (rad s^{-1}). The lattice is composed of a continuum of energy levels with a distribution of frequencies described by a function known as the spectral density, $J(\omega)$. This spectral density function depends on the molecular correlation time, which in turn depends on the sample temperature and viscosity. Spins return to equilibrium *via* a dynamic interaction with microscopic magnetic fields that randomly turn on and off along the volume of the sample with different intensities and directions. The projections of these magnetic fields along the z -axis (parallel to B_0) will add or subtract to B_0 , introducing acceleration or deceleration of the spins’ precession angular speeds that lead to random changes in their phases. This contribution does not involve energy exchange with the lattice (adiabatic), since no change in the population distribution takes place. This process only affects the loss of transverse magnetization (M_x and M_y) by dephasing the spins contributing to the coherences created in the x, y -plane after a rf pulse is applied.

On the other hand, the components perpendicular to B_0 will in fact produce an effect that is equivalent to that produced by the rf field B_1 by inducing transitions between spin energy levels (non-adiabatic) in both directions ($\alpha \rightarrow \beta$ and $\beta \rightarrow \alpha$), leading to longitudinal relaxation. By longitudinal we refer to changes in the intensity of the magnetization in the z -axis (M_z). These random perpendicular fields also contribute to transverse relaxation. It is important to highlight that these transitions will only happen if these random fields oscillate at the corresponding nuclei Larmor frequencies. During the longitudinal relaxation process, the populations will change until they reach a Boltzmann distribution as predicted by the temperature of the lattice (*i.e.* the populations will continue to change until the

spins reach thermal equilibrium with the lattice). Note that it is assumed that the lattice is always in thermal equilibrium with itself and that its specific heat is infinite.

It is experimentally found that the loss of transverse magnetization as well as the return of the populations follow first-order kinetics (eqn (1.1) and (1.2)). Two first-order rate constants are involved in the relaxation process. These are the longitudinal relaxation or spin–lattice rate constant R_1 , governing the recovery of the equilibrium magnetization M_0 and the transverse, or spin–spin relaxation rate constant R_2 , which dictates the disappearance of observable transverse magnetization. These rate constants will have the associated relaxation times $T_1 = 1/R_1$ and $T_2 = 1/R_2$:

$$\frac{dM_z}{dt} = -R_1[M_z(t) - M_0] \Rightarrow M_z(t) - M_0 = (M_z^{\text{initial}} - M_z^0) e^{-t/T_1} \quad (1.1)$$

$$\frac{dM_{xy}(t)}{dt} = -R_2 M_{xy}(t) \Rightarrow M_{xy}(t) = M_{xy}^{\text{initial}} e^{-t/T_2} \quad (1.2)$$

Why does spin relaxation follow an exponential decay law? In principle, other temporal dependencies—for instance, a Gaussian decay $e^{-\frac{t^2}{T_2}}$ —could be imaginable. Following an exponential dependence implies that the time evolution of magnetization depends only on its particular value at a given time t , as can be seen on the time derivative $\frac{dM_{xy}}{dt} = -\frac{M_{xy}^{\text{initial}}}{T_2} e^{-\frac{t}{T_2}}$. However, other possible forms for magnetization evolution as for instance a Gaussian-like dependence will not fulfill this condition and will therefore imply a memory of the system. We want to thank Stan Sykora for making us aware of this issue.

If a strong rf is applied, as in the spin-lock conditions applied in the ROE experiments, the transverse magnetization is not exactly governed by the T_2 time, but is instead governed by the $T_{1\rho}$ relaxation time.

As noted above, all magnetic field oscillations that can cause a transition between energy levels can be a source of relaxation. Besides the external field, all active nuclei in a molecule, or the surrounding solvent, unpaired electrons in the molecule or external agents (trace metal ions, oxygen, *etc.*) are capable of promoting relaxation since the Brownian motion of molecules and solvents changes the relative orientation of the molecular spins (note, however, that spins themselves are locked in the direction of the strong external field). We list here the sources for spin relaxation experienced by molecules in solution:

- Dipolar relaxation: the source of the NOE and the most effective source of relaxation in solution.
- Chemical shift anisotropy (CSA) relaxation. The chemical shielding is a tensorial orientation-dependent property. Therefore, the effective field

B_{eff} experienced by a particular nucleus depends on molecular re-orientation. For instance, the carbon nuclei in benzene are more shielded when the C_6 axis is collinear with the external field than in the case of a perpendicular orientation. Just as with dipolar coupling, the anisotropy effects vanish in isotropic conditions, but they are still a source of relaxation. However, CSA relaxation is only relevant for nuclei with large chemical shielding tensor anisotropies and at high fields, and it is therefore not so relevant to ^1H NOE experiments.

- Scalar coupling relaxation: contrary to the case of dipolar relaxation, J -coupling values are not affected by rotational motions, but only internal motions, so they are consequently a much poorer source of relaxation than the dipolar coupling. In the case of weak coupling, only the z -component of the local field is affected, producing only transverse relaxation. This is, however, not true in the case of strong coupling, and scalar relaxation may contribute to negative enhancements.
- Paramagnetic relaxation: similar in nature to dipolar relaxation, relaxation caused by interaction with magnetic moments of unpaired electrons is extremely effective due to the high gyromagnetic ratio of the electron being nearly three orders of magnitude larger than that of the proton ($\gamma_e/\gamma_H = -657.42$). Even small traces of paramagnetic agents may cause very effective relaxation of the nuclear spins.

The dipolar interaction between spin-active nuclei is governed by the distance separating the nuclei and the orientation of the internuclear vector with respect to the strong external field:

$$D_{IS} = -\frac{3\mu_0\gamma_I\gamma_S\hbar}{8\pi^2 R_{IS}^3} \left(\cos^2 \theta - \frac{1}{3} \right) \quad (1.3)$$

This direct interaction may have a magnitude of thousands of Hz, as can be observed in solid-state samples. However, under isotropic molecular tumbling conditions, as is commonly the case in liquid-state samples, the equally probable orientations of the internuclear vector in all directions will cause the angular term in eqn (1.3) to vanish (see Chapter 4 on residual dipolar couplings [RDCs]). However, although the D_{IJ} does not show as an observable splitting of spectral lines, the magnetic field oscillation caused by reorientation of the internuclear vector may still cause transitions between spin states. This dipole–dipole relaxation acting on systems driven out of thermal equilibrium is the source of the NOE.

The $n\text{Oe } \eta_{IS}$ can be defined as the fractional enhancement of the NMR signal integral of nucleus I when the magnetization of nucleus S is moved out of equilibrium either by saturation or inversion; if spin I is close in space to S , both nuclei will be involved in a significant dipole–dipole interaction. The perturbation on spin S will be transmitted to spin I , therefore modifying the intensity, measured as the value of the corresponding integral, of its NMR signal, with I^I and I_0^I being the integral values of the NMR signal for

nucleus I under perturbed and unperturbed conditions, respectively. If the intensity increases with the perturbation, then the NOE is defined as positive, and is conversely defined as negative in the case of a decrease in intensity:

$$\eta_{I\{S\}} = 100 \times \frac{I^I - I_0^I}{I_0^I} \quad (1.4)$$

For an NOE enhancement to be observed, the internuclear vector (r_{IS}) has to be reoriented in solution at a rate within the range of the resonance frequencies of the spins I and S , which can be bonded or non-bonded, and may belong to the same or to a different molecule. Note that this reorientation may be caused both by global molecular tumbling and intramolecular conformational changes.

In isotropic solutions, the dipolar coupling is manifested not as an observable splitting of the transition lines, but rather as NOE due to the dipole–dipole relaxation. The fact that the dipolar coupling (D) between two nuclei vanishes due to isotropic molecular tumbling in solution is the reason why the NOE is observed. If the molecule does not tumble (solid state), then the dipolar couplings have a finite value that depends on the internuclear distance and on the angle of the internuclear vector with respect to the axis of the external magnetic field, but the NOE disappears. It is the oscillating nature of the dipole–dipole interaction that produces an efficient cross-relaxation process between the interacting nuclei, giving rise to the NOE. In normal standard liquid NMR conditions (isotropic tumbling), both parameters (the NOE and the dipolar coupling, D) cannot be observed simultaneously. However, if the sample is exposed to an anisotropic environment such as liquid crystals, liquid crystalline solutions, or stretched polymer gels, fractions of the dipolar coupling (known as RDC, see Chapter 4) and NOE are simultaneously observable in solution. If the degree of alignment in the anisotropic media is weak (~ 0.01 – 0.10%), then 99.90–99.99% of the molecules are still tumbling isotropically, thereby producing very homogenous NMR spectra with enough molecular tumbling to display a NOE.

It is beyond the scope of this chapter to treat the theoretical aspects of the NOE in detail. However, without being redundant, it is germane to briefly consider a few key equations, as well as several other basic concepts, in order to better understand the Overhauser effect and provide the reader who has an interest in natural product structure elucidation with an appropriate level of understanding to use these experiments in their research.

The energy level diagram for a system of two dipolar-coupled nuclei I and S is shown in Figure 1.3. We may distinguish in this scheme three types of transitions between energy levels: first, single-quantum transitions W_{SQ} , where either the spin I or spin S state is changed; second, zero-quantum or flip-flop transitions $\alpha\beta \leftrightarrow \beta\alpha$; and finally, the double-quantum $\alpha\alpha \leftrightarrow \beta\beta$ W_{DQ} transitions. Whereas rf irradiation may cause only single-quantum transitions, giving rise to observable spectral lines, relaxation may take place through any of the three mechanisms.

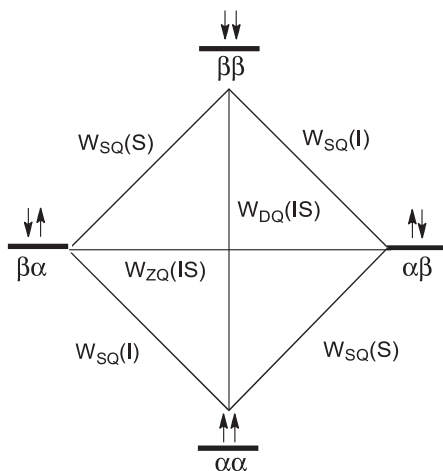


Figure 1.3 Energy levels for a system of two coupled spins.

Consider now what happens through a thermally equilibrated system upon selective irradiation of the spin S (Figure 1.4). Irradiation will cause equalization of the α and β states of spin S through the single-quantum transition processes. If relaxation is not taken into account, we would see how the signal from the S spin disappears while the total intensity of the I signal, defined by its integral, remains the same. However, relaxation-induced non-radiative double-quantum and zero-quantum transition processes are operative. If double-quantum transitions are much more efficient than the zero-quantum transition process ($W_{DQ} \gg W_{ZQ}$), the population of the $\alpha\alpha$ state will grow at the expense of the $\beta\beta$ state in an attempt to restore the Boltzmann relationship ($2\Delta N$). Therefore, the differences between the energy levels for single-quantum transitions of spin I will grow by a factor Δ , increasing the intensity of the I signal. Conversely, if the zero-quantum transition is more efficient than the double-quantum transition, then the population will flow from the $\alpha\beta$ to the $\beta\alpha$ state. This will therefore decrease the population difference by a Δ factor and the integrated intensity of signal I would decrease proportionally. Thus, we will have positive NOE (intensity increases) when double-quantum relaxation predominates over the zero-quantum process, but a negative NOE if the situation is reversed.

Inclusion of all processes in Figure 1.4 in the Solomon equations leads to eqn (1.5). The term σ_{IS} represents the difference between the double-quantum and zero-quantum relaxation rates and it will govern the sign of the NOE. ρ_{IS} is the total dipolar longitudinal relaxation rate constant experienced by spin I , due to the presence of spin S . Note that according to this expression, if the magnetogyric ratios of the coupled nuclei have the same sign, as in homonuclear experiments or ^1H – ^{13}C experiments, then the intensity of the signal is increased (positive NOE), but is reversed if the sign of the nuclei is different. This is an important issue in broadband



Note, however, that spin I can relax through other mechanisms such as CSA or scalar couplings or, in a very effective way, through nucleus–electron relaxation due to the presence of paramagnetic substances (*e.g.* oxygen) in the sample. All of these additional relaxation sources may cause decreases in the NOE enhancement or even cause it to vanish. These additional relaxation pathways can be included in eqn (1.6) by replacing the purely dipolar relaxation rate ρ_{IS} by an effective rate ρ_{IS}^* , which includes all of the additional relaxation effects on spin I :

$$\rho_{IS}^* = \rho_{IS} + \rho_{I\{\text{ext}\}} \quad (1.6)$$

The W_{DQ} , W_{SQ} , and W_{ZQ} rate constants correspond to transitions with very different associated frequencies. If we consider a pair of proton nuclei in a

500 MHz spectrum separated by 1 ppm, the W_{DQ} transition ($\omega_I + \omega_S$) will be of 500 500 MHz + 500 000 MHz = 1 000 500 MHz, whereas it will be only 500 Hz for the zero-quantum transition W_{ZQ} ($\omega_I - \omega_S$). Therefore, they will be induced by fields oscillating at very different frequencies (*i.e.* by molecular motions of very different speeds), and these frequencies will depend on the particular dynamics of the internuclear vectors in the molecule. As a first approximation, we may consider our molecule as a solid body where all of the internuclear vectors are fixed with respect to a rigidly attached molecular frame. The random nature of the molecular motion does not allow the quantification of a rotational speed of this solid body, but we may still define an average rotational speed in terms of the so-called rotational correlation function, τ_c . If we consider a vector $\mathbf{r}$ connecting two points of the body, it is clear that the coordinates of this vector, with respect to the external laboratory frame, will change during the period of time τ . The average speed at which this vector changes its orientation in space can be defined through a time correlation function $g(\tau)$:

$$g(\tau) = \overline{\vec{r}(t)\vec{r}(t+\tau)} \quad (1.7)$$

This rotational correlation function takes the exponential form given in eqn (1.8) where τ_c is the rotational correlation time, defined as the time constant for the exponential decay of the correlation function:

$$g(\tau) = e^{-\frac{\tau}{\tau_c}} \quad (1.8)$$

In our solid body approximations, τ_c will be the same for all of the internuclear vectors considered; note, however, that the presence of internal motion movements may cause this approximation to break down, and very different correlation times may be observed for different parts of the molecule. As a very general rule of thumb, the rotational correlation time for small molecules in organic solvents is given by the expression:

$$\tau_c = 10^{-12} \text{ MW} \quad (1.9)$$

where MW is the molecular weight of the given molecule in mass atomic units. Note, however, that τ_c is proportional to the viscosity of the solvent and decreases with increasing temperature.

The long τ_c of a slowly tumbling molecule will be associated with oscillating fields of low frequency, and a short τ_c will conversely be associated with high-frequency oscillating fields. The available power (*i.e.* how much energy can be retrieved by the lattice at a given frequency) can be quantified by just applying a Fourier transform to the rotational correlation function. This Fourier transform of the exponential correlation function is a Lorentzian function, known as the spectral density function $J(\omega)$:

$$J(\omega) = \frac{\tau_c}{1 + \omega^2 \tau_c^2} \quad (1.10)$$

When τ_c shortens, the spectral density function becomes broader and contributions from higher frequencies become important, as shown in Figure 1.5. The rates for the single-, zero-, and double-quantum processes can be cast now in terms of the spectral density function. For a dipolar-coupled spin pair, the rate constants W are given by the following expressions:

$$W_{\text{DQ}} \propto \frac{\gamma_I^2 \gamma_S^2}{r^6} \frac{12\tau_c}{1 + \tau_c^2(\omega_I + \omega_S)^2} \quad (1.11)$$

$$W_{\text{ZQ}} \propto \frac{\gamma_I^2 \gamma_S^2}{r^6} \frac{2\tau_c}{1 + \tau_c^2(\omega_I - \omega_S)^2} \quad (1.12)$$

$$W_{\text{SQ}} \propto \frac{\gamma_I^2 \gamma_S^2}{r^6} \frac{3\tau_c}{1 + \tau_c^2\omega_I^2} \quad (1.13)$$

Inserting these equations into eqn (1.5), the theoretical stationary NOE for a homonuclear system is given by:

$$\eta_{\text{max}} = \frac{5 + \omega^2\tau_c^2 - 4\omega^4\tau_c^4}{10 + 23\omega^2 + 23\omega^2\tau_c^2 + 4\omega^4\tau_c^4} \quad (1.14)$$

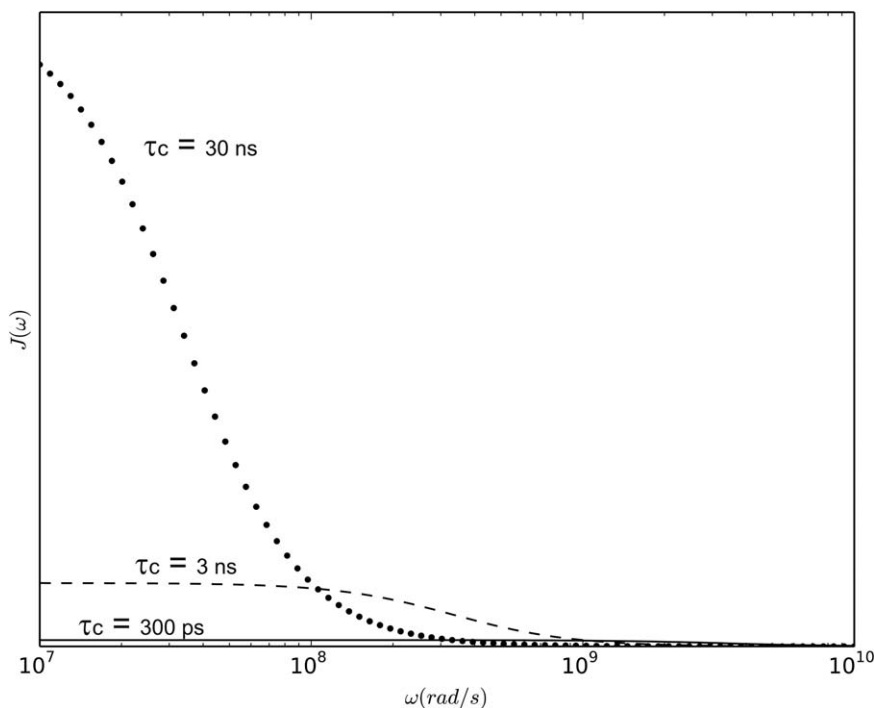


Figure 1.5 Spectral density functions for different rotational correlation times τ_c .

Plotting eqn (1.14) as a function of τ_c indicates that the NOE will be positive at short correlation times, a typical condition of small organic molecules. Around the cross-point, $\omega\tau_c \approx 1.12$ and the NOE falls rapidly until it reaches a maximum negative -100% enhancement. At higher fields, the null condition will be reached at lower τ_c values (Figure 1.6). Therefore, it could be the case that NOE experiments perform better at lower fields, as some complex organic molecules may enter the negative region in high-field instruments.

1.3 Multispin Systems

Things become more complicated, but also more interesting, in the case of a system with more than two coupled dipolar spins. In general, the presence of other X spins close to the observed nucleus I acts as an additional source of leakage, causing a reduction of the observed NOE. For a multispin system, the observed enhancement of I upon S irradiation is given by:

$$f_I\{S\} = \eta_{\max} \frac{r_{IS}^{-6} - \sum_X f_X\{S\} r_{IX}^{-6}}{r_{IS}^{-6} + \sum_X r_{IX}^{-6}} \quad (1.15)$$

Looking at the numerator of eqn (1.15), one can see that the enhancement of spin I depends on the enhancements of the surrounding spins X suffer upon S irradiation. If we just irradiate spin A (Figure 1.7), the perturbation

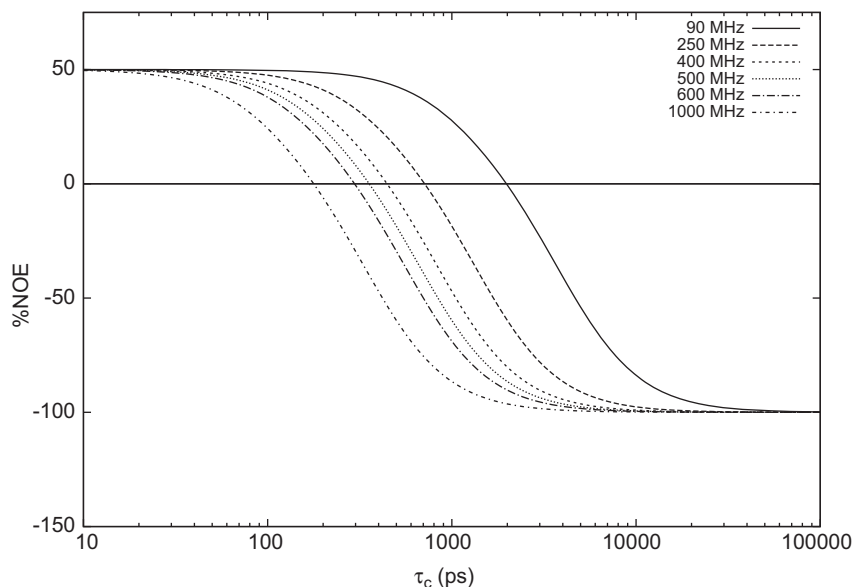


Figure 1.6 NOE intensity *versus* τ_c at different field strengths.

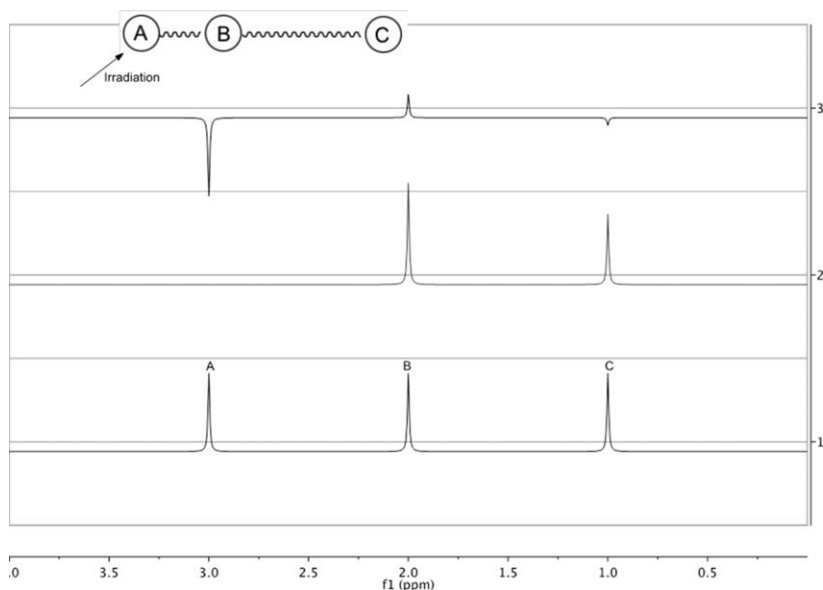


Figure 1.7 Multispin NOE. Bottom: standard ^1H spectrum; middle: ^1H spectrum after irradiation of nucleus A; top: difference spectrum.

created in the populations of B will be transmitted to the population of C, disturbing its signal intensity. In the region of a positive NOE, when double-quantum cross-relaxation predominates, this disturbance can be observed as a negative NOE. Note, however, that in the region of negative NOE, all observed NOEs, either direct or indirect, are of negative sign.

In systems with many spins, the indirect effects may complicate the quantification of the values and a full treatment of the multispin Solomon equations could be needed. However, note that although in a stationary state experiment, the final enhancements are independent of the time, the indirect NOEs need much longer times to reach the stationary state. Thus, the use of non-stationary kinetic NOE experiments will allow easier quantification of the NOE.

1.4 The Kinetic NOE

We will see in this section that quantitative analysis of the NOE can be facilitated by the use of kinetic NOE experiments rather than stationary-state experiments. The most simple way to perform kinetic NOE measurements is to use the truncated NOE experiment (TOE), where the selective irradiation times are kept below the saturation values.¹⁶

For a multi-system, the Solomon equation takes the form:

$$\frac{dI_z}{dt} = -R_I(I_z - I_z^0) - \sigma_{IS}(S_z - S_z^0) - \sum \sigma_{IX}(X_z - X_z^0) \quad (1.16)$$

When spin S is selectively irradiated, at the initial time $t = 0$, the equation reduces to:

$$\left. \frac{dI_z}{dt} \right|_{t=0} = \sigma_{IS} S_z^0 \quad (1.17)$$

Therefore, the initial increase of the I_z population depends linearly only on the cross-relaxation with spin S . This is the basis of the initial rate approximation. However, at longer times, the relaxation of spin I with spins other than S becomes more important. Since for a rigid molecule we can define a global correlation time for the different spins, the NOE enhancements will be directly proportional to the r_{IS}^{-6} distances. Hence, if a valid NOE within the initial rate approximation is measured between two probe nuclei with a well-known distance, it is possible to determine an unknown I - S distance:

$$\begin{aligned} f_I\{S\} &= C r_{IS}^{-6} \tau_m, \\ f_X\{Y\} &= C r_{XY}^{-6} \tau_m, \end{aligned} \quad (1.18)$$

where C is a proportionality constant depending on the Larmor frequency and on the rotational correlation time. Therefore:

$$r_{IS} = r_{XY} = \left(\frac{f_X\{Y\}}{f_I\{S\}} \right)^6 \quad (1.19)$$

The initial rate approximation is the basis of most of the NOE-based conformational analysis that is used due to its simplicity. Note that, similar to the stationary NOE, the TOE is a difference experiment, and therefore has associated with it all of the practical problems inherent to difference spectroscopy.

Most of the modern applications of NOE are based on the use of transient NOE techniques. In the 1D transient NOE or 1D-NOESY experiments, the spin S is driven out of equilibrium by selective inversion. After that, a mixing time period τ_m is allowed to elapse before the application of the observation pulse and acquisition.

1.5 Transient NOE

Since the system is not saturated in the NOESY experiment, the magnetization enhancement on spin I depends now on the relaxation rate of the inverted spin S , denoted as R_S . Near the extreme narrowing limit, this rate is competitive with the cross-relaxation term and the maximum positive enhancement that can be obtained in a transient experiment is only 38.5%. In the $\omega\tau_c \gg 1$ region, enhancement it is still around 100% as cross-relaxation is the most efficient process. Transient NOE measurements can be performed in a 2D fashion in the extremely popular NOESY experiments. In NOESY,

the inversion pulse is split into two 90° pulses and a t_1 chemical shift evolution is inserted between them (Figure 1.8). After the second 90° pulse, z -magnetization is created for all spins, depending on their respective chemical shifts. During the mixing time τ_m , cross-relaxation effects will give rise to NOE enhancements. Note, however, that several artifacts may arise in this experiment. If a scalar coupling exists between the spins, then zero-, single-, and double-quantum coherences evolve during the experiment. Also, axial peaks may arise due to spin relaxation during the t_1 period. These artifacts can be removed using an eight-step phase cycle. Since cross-peaks may have either negative or positive signs with respect to the diagonal, the NOESY spectra should be acquired in phase-sensitive mode using TPPI, States, or States-TPPI protocols.

The advent of pulsed field gradients (PFGs)¹⁷ allowed purging of undesired coherences in a reduced number of scans. By inserting a weak gradient during the whole mixing time τ_m , the number of scans can be reduced to just two (Figure 1.9). Alternatively, a 180° pulse can be flanked by two strong PFGs to suppress the evolution of coherences during the PFG period, with the purpose of the 180° pulse being to refocus the evolution of coherences taking place during the gradient period. Echo-antiecho phase-sensitive NOESY spectra can be recorded by the insertion of a PFG during the evolution period.¹⁸

1.6 ROE and ROESY

When molecules reach a motion regime in which the resonance frequency (ω_0) is in the order of the inverse of their molecular correlation time ($\omega_0\tau_c \cong 1.12$), the NOE enhancement is null and/or very weak. This is a situation that is observed for molecules in a range of molecular weights, reported by some authors to be 1000–2000 or 1000–3000, but a recent review article by Breton and Reynolds¹⁹ mentioned that it can start from molecular weights of ~ 700 . The molecular weight at which the NOE enhancement in small molecules starts to decrease depends on many factors, such as molecular shape (affecting its tumbling rate), solvent viscosity, temperature, and spectrometer frequency. However, it is a situation that has to be judged by the chemist experimentally. Many natural compounds fall into this

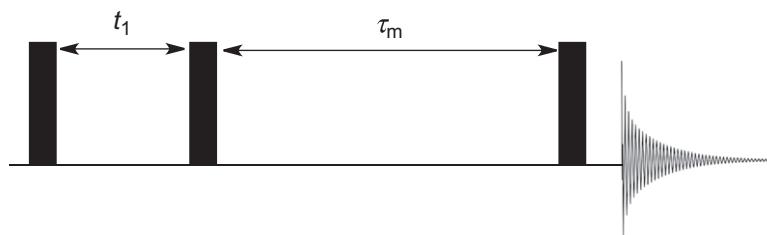


Figure 1.8 2D phase-sensitive NOESY pulse sequence.

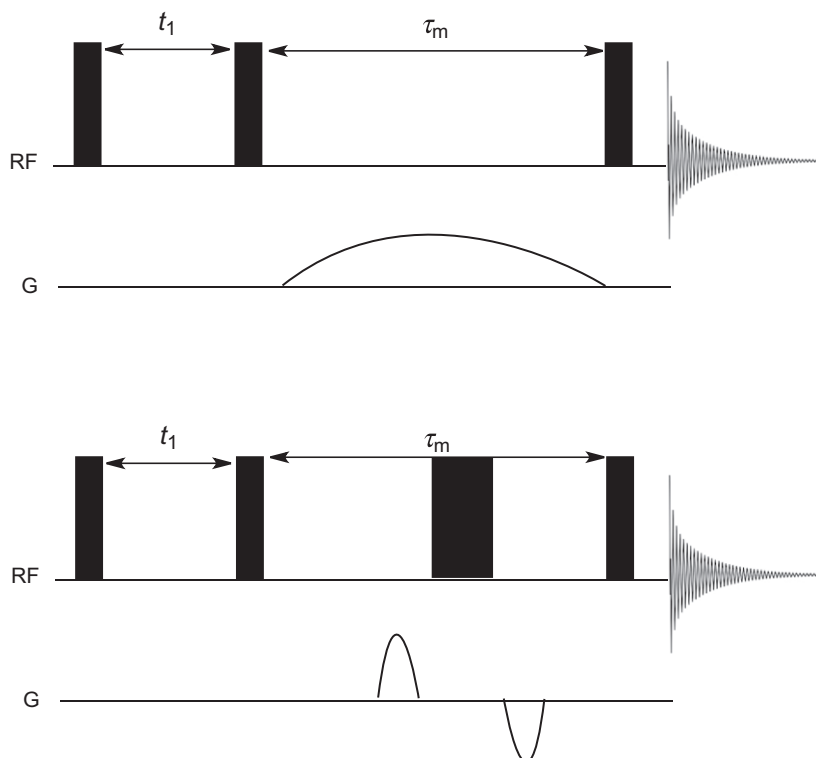


Figure 1.9 2D gradient-selected NOESY pulse sequences.

category, particularly glycosides with several sugar units, peptides, macrocyclic natural compounds, dimers of diterpenes, triterpenes, steroids, or large alkaloids.

Since this undesired phenomenon occurs at $\omega_o\tau_c \cong 1.12$, one of the experimental alternatives is to manipulate $\omega_o\tau_c$ to move it away from the null NOE enhancement point. Around the null point, for a given molecule (constant value of τ_c), the NOE enhancement (η) shows a sharp change as a function of the spectrometer frequency (ω_o). For example, a null NOE at 500 MHz can be turned into an enhancement of +0.10 at 400 MHz and +0.22 at 300 MHz, or -0.10 at 600 MHz and -0.18 at 700 MHz. However, not every laboratory around the world has the luxury of having many NMR spectrometers to manipulate the NOE enhancement as a function of the spectrometer frequency. Nevertheless, for a given spectrometer, it is also possible to manipulate the molecular tumbling rate or correlation time (τ_c), which strongly depends on the solvent viscosity. Since temperature affects viscosity, changes in temperature and/or in the type of solvent will also help to move the experimental conditions away for the null NOE point. These were the only options available to chemists until 1984, when Bothner-By and colleagues introduced a rotating frame NOE experiment that was published

with the name CAMELSPIN (Cross-relaxation Appropriate for Mini-molecules EmuLLated by SPIN-locking) and was later renamed to rotating frame Overhauser enhancement spectroscopy (ROESY).²⁰ In the rotating frame NOE experiment, the cross-relaxation process occurs under the effect of a spin-lock, and the NOE enhancements (also referred to as ROE) are always positive, regardless of the tumbling rate of the molecules. There is not such an NOE null point at $\omega_0\tau_c \cong 1.12$. In addition, direct NOE enhancements have opposite signs to those from spin diffusion, saturation transfer, and exchange for all molecules (small and large).

Similarly to most NMR experiments, the rotating frame NOE can be performed as a selective 1D (selective ROE) or as a 2D (ROESY) experiment. In Figure 1.10, the basic ROESY pulse sequence is shown. Although the experiment does not show an NOE null point as a function of the molecular tumbling rate, leading to a significant advantage over traditional NOE, the ROE experiment presents a series of experimental complications. Since the experiment involves a spin-lock that is not as strong as the one used by TOCSY, in-phase TOCSY transfer peaks are observed and they are not distinguishable from real ROE peaks. In addition, these TOCSY transfers also lead to TOCSY-ROE transfers of magnetization, which give origin to “false” cross-correlation peaks for protons that are far from each other.²¹ The spin-lock is not uniform over the whole proton chemical shift range, leading to non-uniform enhancement and making the experiment not recommendable for quantitative purposes. Finally, anti-phase COSY-like peaks between *J*-coupled spins are also observed, but their intensity is very weak. Further modifications to minimize these unwanted experimental complications were reported. Hwang and Shaka introduced a variant of the experiment that minimizes the TOCSY contributions.²² Its name is transverse ROESY (T-ROESY or Tr-ROESY), and the mixing period consists of a train of 180° with alternate phase ($180^\circ_{-x}, 180^\circ_{+x}$). The pulses are given with no time between them. We strongly recommend this version of ROESY, which significantly reduces the unwanted TOCSY and false ROE peaks. The only shortcoming of this experiment is that the intensities of the cross-peaks are smaller than in the original ROESY experiment. From a conceptual standpoint, NOE and ROE experiments are based on the same physical phenomenon and provide the same structural information. Due to the fact that ROE experiments present so many experimental complications, it is

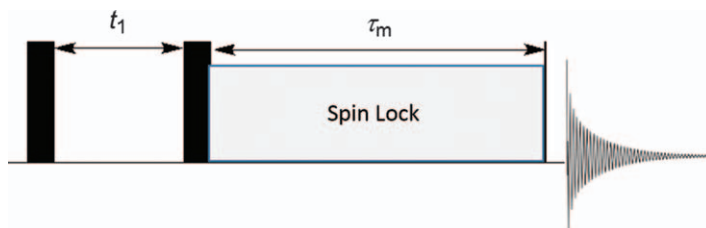


Figure 1.10 2D-ROESY basic pulse sequence.

important to be aware of these interferences when the experiment is used. A series a potential cross-peaks not related to the purpose of the experiment (no ROE peaks) may be observed and can mislead the users regarding their structural problem. Therefore, we strongly recommend not using the ROE experiments if there is no need for them. We fully agree with the following statement from Timothy Claridge, which clearly summarizes the problem:²³

“In view of the various complicating factors associated with the ROESY experiment, it is perhaps prudent to avoid using the technique whenever possible, and instead select a steady-state or conventional transient experiment as first choice.”

We have performed a full-text search of the entire online collection of the *Journal of Natural Products* using the keyword “roesy” and we found 1156 articles. Surprisingly, more than 90% of these articles report the use of ROESY for molecules that do not really need it! And those that use it do not even mention or justify the reasons why. It seems that many natural products chemists use ROESY by tradition without being aware of the differences with regards to a NOESY experiment. If signal overlapping or sensitivity does not become a critical issue, the use of relatively low-field spectrometers (~300–400 MHz) may be considered in order to avoid the use of the ROESY experiment. As a final note, ROE builds twice as fast as NOE, hence shorter mixing time values are recommended (300–400 ms).

1.7 Scalar Couplings—Zero-quantum Filters

If scalar coupling is present between protons, the corresponding term will evolve during the mixing time as a sum of zero- and double-quantum coherences, which will be manifested as antiphase peaks in the 1D- or 2D-NOESY/ROESY spectrum. Although the double-quantum component may be filtered by phase cycling or application of a PFG during the mixing time, this is not possible in the case of the zero-quantum component, as it has the same coherence order as the observed longitudinal magnetization. This zero-quantum component will evolve during the mixing time as the difference Ω_{ZQ} between the offsets of the scalar coupled nuclei, giving a peak of intensity I_{ZQ} :

$$\Omega_{\text{ZQ}} = \Omega_{\text{I}} - \Omega_{\text{S}} \quad (1.20)$$

$$I_{\text{ZQ}} = \cos(\Omega_{\text{ZQ}}\tau) \quad (1.21)$$

Although the total integral of these antiphase peaks will be zero, they may obscure the presence of a true NOE correlation, making the observation of weak correlations difficult and the quantification of the integrals inaccurate. A simple procedure for minimization of the zero-quantum peaks in NOESY experiments is to introduce a small random variation of the mixing time, or either to randomly move the placement of a π pulse in the middle

of the mixing time, between each t_1 increment. The magnitude of the true NOE peak will be minimally affected, especially at long mixing times, whereas the phase of the ZQ peak will differ from experiment to experiment and most of the signal will cancel out. Note that the degree of variation $\Delta\tau_m$ should cover at least one cycle of the zero-quantum coherence, which makes the elimination of ZQ peaks between protons that are very close in chemical shift more difficult to accomplish (see Figure 1.11). However, this procedure introduces some degree of t_1 noise into the spectrum since the ZQ peak is spread out along the F_1 dimension after Fourier transformation. These effects can be clearly shown in the NOESY 2D spectrum of a synthetic isoquinoline derivative, recorded with a 500 ms mixing time delay (Figure 1.12), which clearly shows zero-quantum correlations between vicinal and germinal protons. The introduction of a 30 ms random variation between either scans or experiments effectively removes the zero-quantum peaks.

A 3D-based approach was proposed to suppress ZQC coherence without introducing additional t_1 noise by smearing out the ZQC peak over the third dimension. Clean NOESY spectra are obtained with this methodology, although at the cost of somewhat reduced sensitivity.²⁴ Based on previous

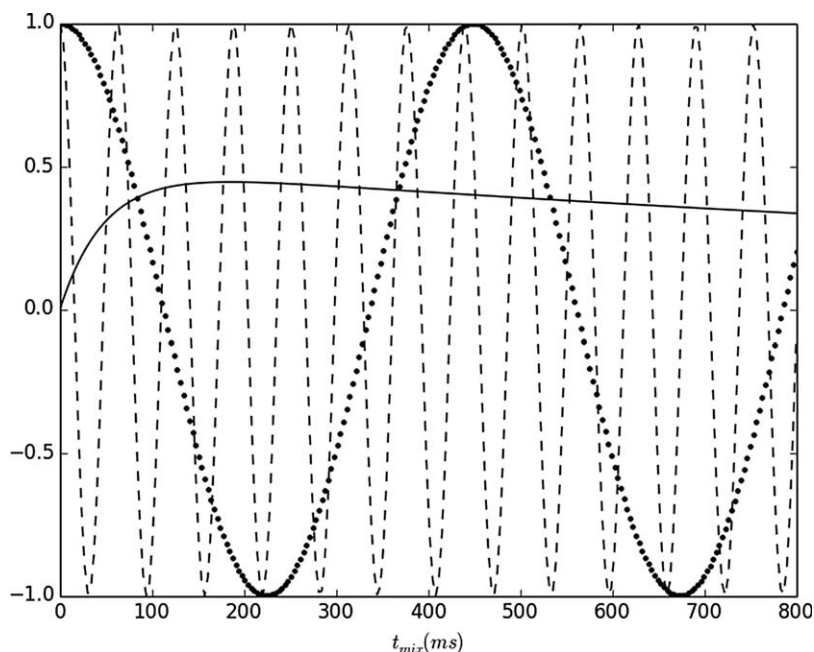


Figure 1.11 Time evolution of zero-quantum coherence *versus* NOE buildup during mixing time. The NOE buildup (continuous line) curve is simulated for a $t_1 = 1$ s value and a cross-relaxation rate $\sigma = 10$ Hz. The offset differences between the two scalarly coupled protons were set to 100 Hz (dashed line) and 14 Hz (dotted line).

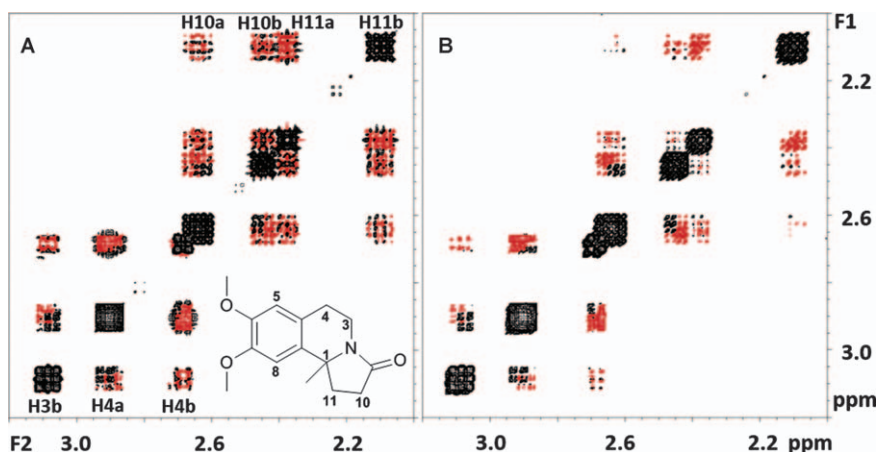


Figure 1.12 500 ms mixing time 2D-NOESY (A) and the same spectra recorded with a 30 ms random variation delay ZQC filter (B).

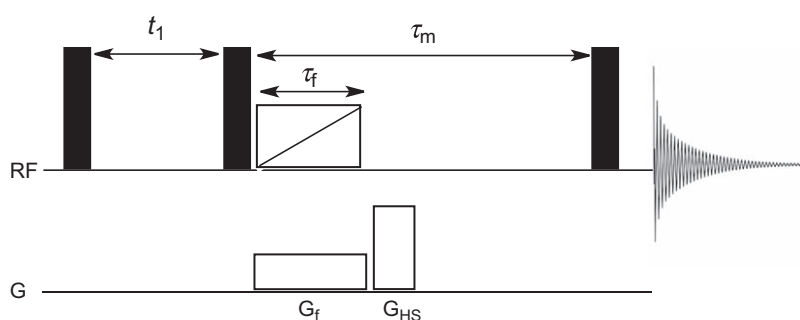


Figure 1.13 Pulse sequence for the Thrippleton-Keeler ZQC filter.

work, Howe presented a ZQ filter based on spin-echo evolution inside a z-filter element. A 180° pulse is placed inside the z-filter block and moved through it in t_1 increments. As ZQ is refocused at $2T-t_1$, the ZQ will evolve during a t_1 time. Typically, 12 increments were needed for a clean elimination of the ZQ peak.²⁵

In 2003, Thrippleton and Keeler²⁶ proposed a ZQC filter based on generating a phase difference between the ZQC magnetizations for different positions of the NMR tube by simultaneous application of a PFG on the z-axis and a swept-frequency 180° pulse (Figure 1.13).

If τ_f is the duration of the gradient/swept-frequency pair, a particular point on the tube will experience the 180° pulse at a time at $x\tau_f$ where $0 < x < 1$. Thus, ZQC is refocused over a time $2\tau_f$, but still evolves for the remaining $(1-x)\tau_f$ time. As x is a function of the position in the tube, the ZQ contribution is effectively dephased over the sample, leading to ZQ

signal attenuation in a single scan. The attenuation A suffered for a I_{ZA} peak between I and J spins is:

$$A = \frac{\sin \Omega_{ZQ} \tau_f}{\Omega_{ZQ} \tau_f} \quad (1.22)$$

Typically, τ_f is in the range of ~ 30 – 50 ms. Although this can become problematic for studies where very short mixing times are needed, it is perfectly acceptable for standard small-molecule NOESY experiments. Note, however, that for coupled spins with closely similar chemical shifts, longer τ_f delays may be necessary. As an alternative, Keeler and Shaka later proposed the use of not one, but a cascade of z -filters, which leads to improved signal suppression.²⁷

1.8 NOE Experiments for the Structural Elucidation of Dimeric Compounds

Dimeric natural products are frequently encountered in nature. Greer and coworkers analyzed a large database of literature entries and estimated that nearly 7% of the papers reported full bilaterally symmetric dimeric structures.^{28,29} NOE experiments in dimeric compounds are difficult to interpret. First, NOEs between equivalent protons are in principle hidden, and second, ambiguities between intra- and inter-monomer correlations can also be present. A way to circumvent the first problem is to exploit the asymmetry found in the ^{13}C isotopomers, as only one of the two equivalent carbons attached to the proton nuclei of interest will be a ^{13}C nucleus, while the other will be attached to ^{12}C . Thus, the ^{13}C filtering of the NOESY experiment, through an HSQC or HMQC step, may allow for differentiation of the two nuclei.³⁰ Editing can be done either prior to the NOE block, as in the case of a HSQC-NOESY,³¹ or after, as in the NOESY-HSQC or NOESY-HMQC experiments.³² If there is an NOE between chemically equivalent protons, the use of F_2 - or F_1 -coupled HSQC experiments allows the observation of a diagonal peak on the ^{12}C isotopomer between the two halves of the doublet split by $^1J_{\text{CH}}$ (Figure 1.14).

In the 800 ms HSQC-NOESY spectrum of carbazol (Figure 1.15), the NOESY peak between the equivalent protons 4 and 4', as well as the

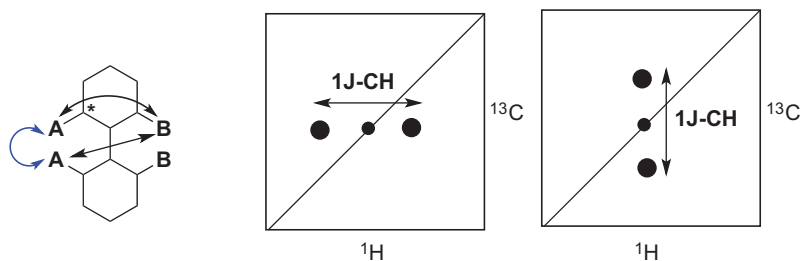


Figure 1.14 Carbon-edited NOESY experiments.

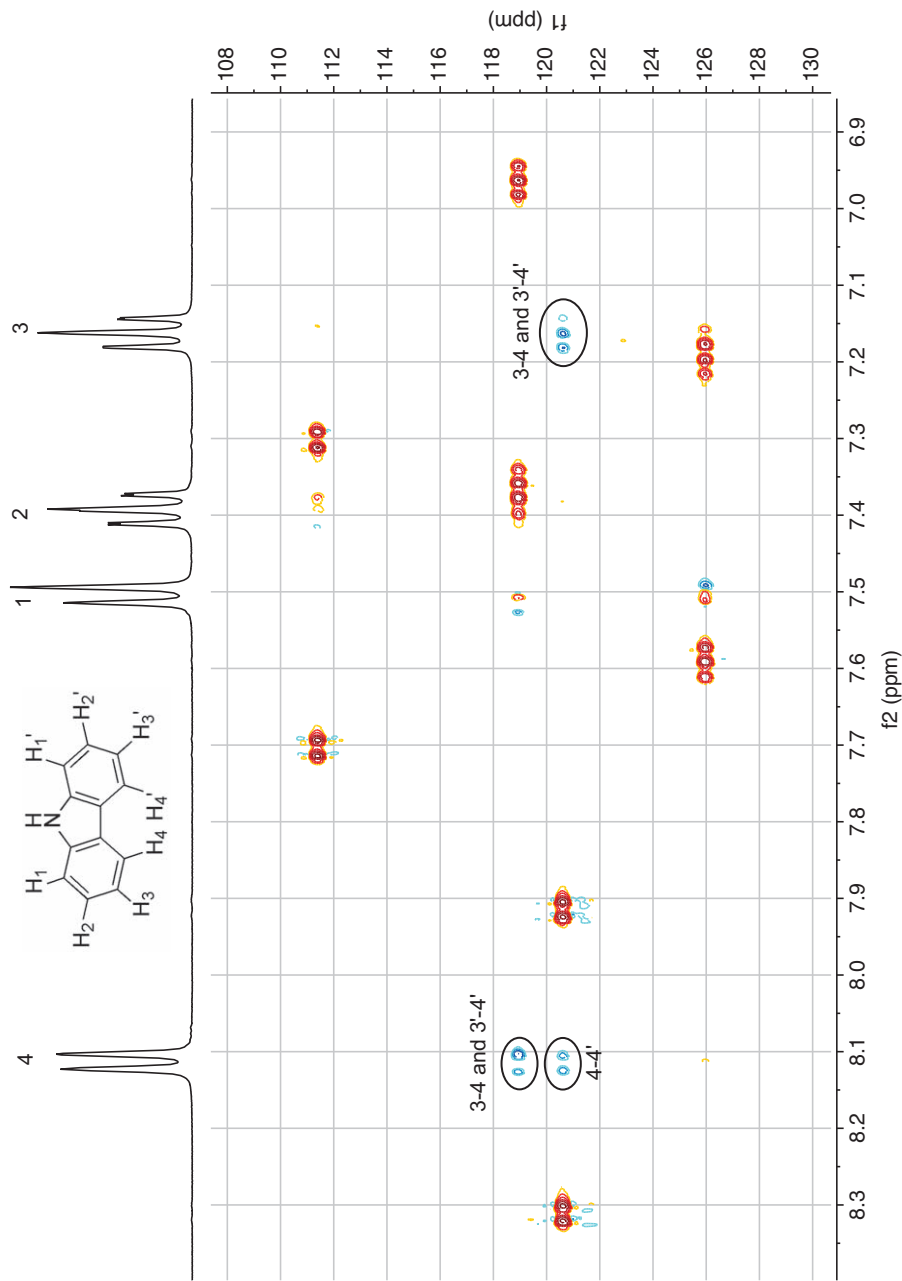


Figure 1.15 HSQC-NOESY spectrum of carbazole in DMSO-*d*₆.

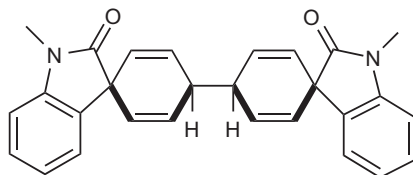


Figure 1.16 HSQC-NOESY-relative NOE intensities in a dimeric amide.

cross-peak between 3 and 4, and 3' and 4' can be seen. Note that the rest of the cross-peaks between close protons are not observed in this spectrum, likely due to different correlation times about the molecular axes.

The intensity of the cross-peaks is, in principle, the same if the carbon editing is done either before (HSQC-NOESY) or after (NOESY-HSQC) the NOESY block. However, experimental intensities may differ due to relaxation effects. Gschwind *et al.*³³ analyzed the effect of diffusion in the gradient-selected HSQC-NOESY and NOESY-HSQC experiments. In the first experiment, the mixing time is placed in the middle of the gradient refocusing, whereas in HSQC-NOESY, refocusing of gradients is separated by short delays. Hence, diffusion may severely decrease signal intensity in the HSQC-NOESY experiment, especially if very long mixing times are employed. Theoretically estimated attenuation factors range from 0.8 at a 0.5 s mixing time to 0.25 for a 3 s mixing time.

With respect to the practical application of these techniques, it should be noted that these are rather insensitive experiments, typically requiring samples in a 0.1–1 M concentration range using standard 5 mm probes, and therefore are only helpful for synthetic compounds where typically larger amounts of sample are available. However, the large boost in sensitivity provided by 3 or 1.7 mm cryoprobes may certainly make this experiment feasible for natural compound elucidation. As an example of this, Buevich *et al.* applied HSQC-NOESY experiments, along with *J*-coupling analysis of ^{13}C satellites, to the determination of the conformation in a dimeric benzamide.³⁴ The relative intensity of the cross-peak between the equivalent methine protons (Figure 1.16) showed a preference for a gauche rather than an anti-conformation of the dimer, in agreement with a *J*-coupling of ~ 4 Hz between these protons observed through the ^{13}C satellites. Generation of artificial HSQC-NOESY data sets, obtained through covariance processing of individual HSQC and NOESY spectra,³⁵ can be useful in structural elucidation, but since the generated spectra does not contain any information present in the parent spectra, they will not show cross-peaks between equivalent protons.

1.9 Fully Quantitative NOE in Small Molecules

Traditionally, organic chemists have employed NOE in a qualitative way where the observation of an NOE correlation simply informs that the two nuclei of interest must be close enough to give an appreciable dipolar

interaction, and the quantitative or semi-quantitative use of NOE experiments is relatively scarce.

Very recently, the Butts group has tried to bring NOE in small molecules to a fully quantitative level (FQ-NOE).³⁶ Their proposal involves the use of techniques that are already known for ameliorating the quantification of NOEs in biopolymers. The proposed methodology involves several procedures that can be summarized in two principal points:

1. Use of 1D-NOESY DPGFSE experiment rather than 2D-NOESY experiments, as they provide more reliable integrals.
2. Calibration of integral areas against irradiated peak areas. This procedure extends the linear behavior of the NOE. After acquisition of the experiments, the integral of the irradiated proton is set to an arbitrary fixed number and all correlation integrals are further related to this integral for each experiment.

Interproton distances are then computed according to:

$$\frac{\eta_{IS}}{\eta_{\text{ref}}} = \frac{r_{IS}^{-6}}{r_{\text{ref}}^{-6}} \quad (1.23)$$

The r_{ref} value is then taken from an internuclear pair that is minimally affected by conformational mobility as a methylene pair or two ortho aromatic protons. Thus, once this reference distance is determined, either from tables or a molecular modeling structure, all interproton distances can be obtained if the corresponding cross-relaxation peaks are observed. A very complete set of 1D-NOESY data was obtained for strychnine, where the average absolute error between interproton distances was below 5% as compared to high-quality DFT and X-ray structures.

However, the use of 2D-NOESY experiments degrades the quality of the fit, especially if short d_1 relaxation delays are used; relaxation delays up to ~ 5 s are necessary in order to improve the quality of the data. However, long delays between transients greatly increases the necessary acquisition time without achieving the accuracy of the 1D procedure. Other sources of error in the 2D-NOESY experiments arise from t_1 noise, which makes the extraction of data very dependent on processing procedures. One could, in principle, argue that many of the sources of t_1 noise (*e.g.* temperature variation) are still present in the 1D experiment in the form of spectral noise; however, the short acquisition time of the 1D experiment lowers the impact of error sources.

Butts and coworkers also observed that degassing the sample did not improve the accuracy of the NOE-determined distances. This is due to a very similar impact of paramagnetic relaxation on the intensity of all observed correlations. Even the presence of oxygen shortens the t_1 relaxation time, allowing the safe use of short d_1 delays in the 1D-NOESY experiments.

The accuracy in the experimentally determined interproton distances allowed the authors to experimentally prove the existence of a minor second

conformation of strychnine. They observed a much larger deviation for the protons $H_{11\beta}$ and $H_{23\beta}$ as the NOE-determined distance, 3.49 Å, compared poorly to DFT and X-ray distances 4.10 Å and 4.12 Å, respectively (Figure 1.17). Further, molecular modeling studies suggested that this effect can be explained by the contribution of a second conformation of strychnine, arising from pseudo-rotation of the C_{22} – C_{23} bond, with a low population that should lie 9.4 kJ mol^{−1} over the basal form according to *ab initio* studies. In this higher-energy conformation, the $H_{11\beta}$ – $H_{23\beta}$ distance is shortened to 2.11 Å according to a computation that, according to Boltzmann-computed populations, corresponds to an average distance of 3.60 Å, very close to the 3.49 Å of the NOESY-determined distance. This neatly shows how the highly non-linear nature of the NOE can be exploited for unraveling low-population conformations that can be very difficult to detect using more linear techniques such as scalar coupling analysis or even RDCs.

The use of quantitative NOESY or ROESY may extend the applicability of NOE experiments to natural products with many stereogenic quaternary centers where no obvious NOE or 3J analysis is possible. By making use of FQ-NOE and ^{13}C DFT chemical shift predictions,^{37,38} Butts and Bifulco determined the relative configuration of conicasterol F,³⁹ a new compound isolated from a *Theonella* marine sponge. 2D-NMR determined the structure to be either the *cis*-diepoxide or the *trans* form (Figure 1.18).

A series of 1D- and 2D-ROESY 300 ms mixing time experiments were conducted and 18 internuclear proton distances determined. The distance between the geminal vinylic protons was taken as a reference. 3D structures were generated at DFT level for the two possible configurations of conicasterol F. The fit between the structures and the data set was expressed in terms of the maximum absolute error (MAE; eqn (1.24)).

$$\text{MAE} = \frac{1}{n} \sum_1^n 100 \times \frac{|r_{\text{ROE}}^i - r_{\text{calc}}^i|}{r_{\text{calc}}^i} \quad (1.24)$$

The differences between the ROE-determined and the DFT-computed proton–proton lengths were determined. MAEs of 2.6 and 5.9 were obtained for the *trans* and *cis* configurations of conicasterol F, respectively. This result was in agreement with ^{13}C chemical shift DFT GIAO/MPWPW91/6-31G* predictions, which furnished MAE deviations of 0.8 and 3.7 for the *trans* and *cis* forms, respectively.

FQ-NOE and quantum mechanical ^{13}C chemical shift predictions were also applied to the determination of the relative configuration in plakilactones G (4) and H (5) (Figure 1.19).⁴⁰ Upon an exhaustive force-field conformational search, structures were refined at the DFT level and ^{13}C shifts computed at the DFT-GIAO level. Boltzmann populations were computed according to the relative energies. The computed Boltzmann populations were then employed to get averaged $\langle r^{-6} \rangle$ distances and chemical

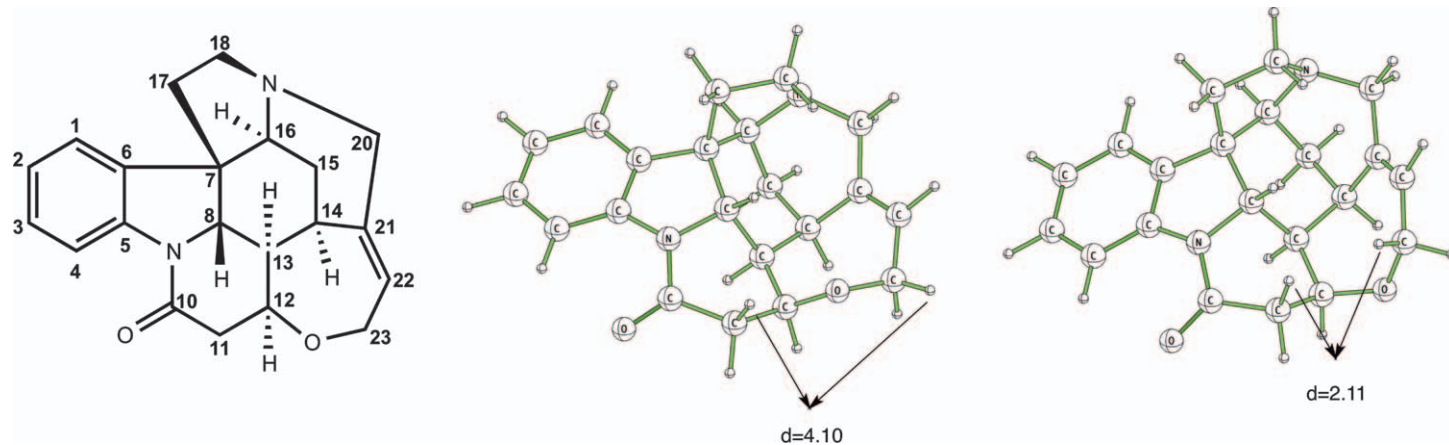


Figure 1.17 Structure of strychnine and DFT geometries for the two lowest-energy conformations. Distances in Å for the $H_{11\beta}$ – $H_{23\beta}$ pair are shown in the figure.

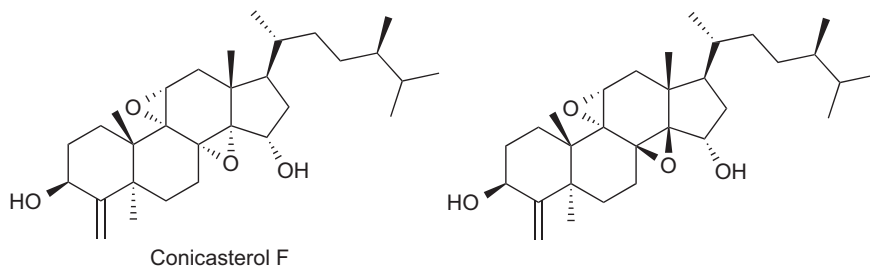


Figure 1.18 FQ-NOE determined configuration of conicasterol F and the rejected diastereoisomer.

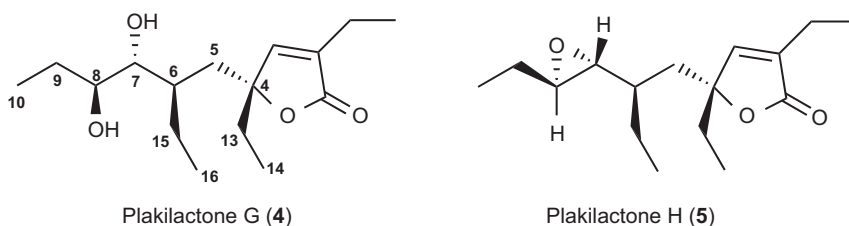


Figure 1.19 Configuration of plakilactones G and H as determined by combined FQ-NOE and DFT chemical shift prediction analysis.

shifts. Both methodologies agreed in the configurational assignment for plakilactone G (5). The determination of configuration in the diol 4 was more cumbersome. As a first step, 4 was determined to possess a 7,8-*erythro* relative configuration by chemical derivatization to the 7,8-*O*-isopropilidene compound and chemical shift analysis of the acetonide methyl groups.⁴¹ Whereas ambiguities arose when determining the relative configuration of the two remaining centers when using the FQ-NOE or chemical shift ¹³C prediction methodologies, the combination of these two methods provided results with higher confidence. Chemical derivatization of the diol moiety in 4 following the Riguera double derivatization method⁴² afforded the absolute configuration of the molecule.

1.10 NOE Deconvolution—NAMFIS

The analysis of NOE experiments in the presence of molecular flexibility is a challenging problem since the very nature of the dipolar relaxation relies on the relative movement of nuclear spins. The analysis of molecular flexibility depends strongly on the ratio between the conformational exchange rates and the global molecular tumbling correlation times, τ_c . In general, NOEs can be only linearly averaged when the conformational exchange rates are lower than NOE buildup rates. In this case, the observed NOE enhancement

is a simple linear combination of the enhancements for every single contribution:

$$\eta_{\text{obs}} = \sum_1^N w_i \eta_i \quad (1.25)$$

where η_i are the individual enhancements and w_i are the corresponding normalized populations. However, only very slow conformational processes such as peptide bond rotations lie in this regime.

Most of the interesting conformational movement types involve rotation around one or more single bonds. These processes generally have low activation barriers and correspondingly high conformational exchange rates. In these cases, conformational exchange is faster than NOE buildup, but still slower than global molecular tumbling. This involves the averaging of auto- and cross-relaxation rates, which are therefore given by:

$$\langle \rho_I \rangle = \sum_i^N w_i \rho_I \quad (1.26)$$

$$\langle \sigma_{IJ} \rangle = \sum_i^N w_i \sigma_{IJ} \quad (1.27)$$

and result in a $\langle r^{-6} \rangle$ averaging of the internuclear distances over the different conformations.

Finally, some molecular conformational movements with very low barriers, such as methyl group rotation, take place at a rate that is even higher than overall tumbling and, in this case, the correlation function is a complex function of spherical harmonics and internuclear distances. Analytical expressions for NOE enhancement between a single proton and a rotating methyl group,⁴³ as well as for intra-methyl relaxation,⁴⁴ are available and have been implemented in the FIRM program.⁴⁵

From this point forward, we will focus on the case of medium-rate $\langle r^{-6} \rangle$ averaged processes where relaxation contributions from different conformations need to be taken into account and NOE enhancements explained (deconvoluted) in terms of a discrete conformational distribution. In this respect, different proposals can be found in the NMR literature, such as MEDUSA,⁴⁶ CPA,⁴⁷ or PEPFLEX-II.⁴⁸

Forster proposed a full-relaxation matrix deconvolution approach.⁴⁹ Assuming a jump-model where the conformational averaging can be described by a combination of discrete geometries, the averaged relaxation matrix is given by:

$$\Gamma_{ab}^{av} = \sum_{i=1}^N w_i \Gamma_{ab}^i \quad (1.28)$$

where w_i are the weights of each population. In this procedure, the effects of spin diffusion are taken into account through the use of relaxation matrix

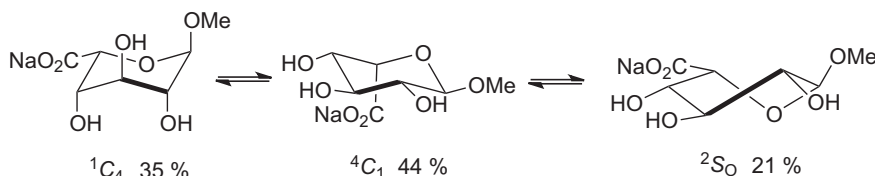


Figure 1.20 Full relaxation matrix NOE deconvoluted populations for *O*-methyl- α -L-iduronate sodium salt.

computations. Populations are computed by applying a non-linear fitting scheme to the experimentally determined NOESY volumes, similar to the Levenberg–Marquardt procedure.^{50,51} The authors applied the developed algorithms to the conformational analysis of the *O*-methyl- α -L-iduronate sodium salt using data from a 3 s mixing time NOESY experiment. Three static conformations (Figure 1.20) consisting of two chair forms, 1C_4 and 4C_1 , and one skew form, 2S_0 , were employed in the analysis. The computed populations were in good agreement with those obtained from *J*-coupling analysis of full MD trajectories using the Haasnoot–Altona equation.⁵² Note, however, that this procedure has rarely been further used for quantitative analysis, and the general viability of this approach still needs to be confirmed, although very recently new software packages have appeared that make use of full relaxation matrix approaches.⁵³

The most widely used approach in conformational analysis of small organic molecules has been the NMR analysis of molecular flexibility in solution (NAMFIS) methodology developed from Cicero, Barbato, and Bazzo.⁵⁴ As a proof of concept, they analyzed the conformation of a cyclic peptide in which previous conformational clustering was done using the MEDUSA procedure. The NAMFIS approach is based on several steps. First, a conformational ensemble should be built using an appropriate force-field and a stochastic conformational search procedure. Since force-field-computed relative energies may have a significant error, care should be taken in order not to lose any structure that is significantly represented in the experimental ensemble. A recommended procedure would be to keep all conformations below $\Delta E \sim 3 \text{ kcal mol}^{-1}$ ($\sim 1\%$ Boltzmann population at 298 K). However, depending on the force-field chosen and the functionality present in the molecule, the necessary energy threshold can be increased to safer limits in the range of 10–12 kcal mol^{-1} .

The key of the NAMFIS method lies in summing up all of the available NMR restraints in a single penalty function P as shown in eqn (1.29). These restraints include not only NOE-estimated distances, but also ${}^3J_{\text{HH}}$ scalar couplings. The penalty function is then built as a sum of square differences:

$$P = \frac{1}{2} \sum_{i=1}^m \left[\frac{d_{\text{calc}}^i - d_{\text{expt}}^i}{d_{\text{err}}^i} \right]^2 + \frac{1}{2} \sum_{i=1}^n \left[\frac{J_{\text{calc}}^i - J_{\text{expt}}^i}{J_{\text{err}}^i} \right]^2 \quad (1.29)$$

The d^i -calculated distances correspond to r^{-6} averages of the internuclear distances of individual geometries as obtained from molecular modeling; in other terms:

$$d^{av} = \left[\sum_{i=1}^N w_i r_i^{-6} \right]^{1/6} \quad (1.30)$$

The ${}^3J_{\text{calc}}$ values are obtained using the simple three-term Karplus equation:

$${}^3J_{\text{calc}} = A \cos^2 \theta + B \cos \theta + C \quad (1.31)$$

or the more complex Haasnoot–Altona equation,⁵² which takes into account the electronegativity of the substituents, as well as their relative orientations. Many molecular modeling packages implement routines for the computation of these couplings from a given 3D structure.

It should be noted that the merit function in NAMFIS weights every NOE and 3J contribution according to their associated error. The use of NOE errors shown in Table 1.1 and a 1.5 Hz general error for ${}^3J_{\text{HH}}$ couplings, computed with the Haasnoot–Altona equation, was recommended.⁵⁵ Once eqn (1.29) is set up, the populations are determined by a linearly constrained non-linear least-squares optimization procedure.

NAMFIS is intended to overcome a general problem with modern computational methods: although they can provide sufficiently accurate geometries, prediction of relative energies in solution can only be considered as a semi-quantitative procedure. Therefore, in the original NAMFIS implementation, computed relative energies are used only to define the threshold for conformer selection.

NAMFIS has been used profusely for the determination of conformational space in tubulin ligands of a macrocyclic nature, such as taxol and its derivatives, discodermolide, laulidamide, and epothilone A (Figure 1.21).

NAMFIS analysis can be performed using scalar couplings other than ${}^3J_{\text{HH}}$. Monteagudo and coworkers⁵⁶ analyzed the conformation of a doxorubicin disaccharide (Figure 1.22) by combination of NOE-derived distances and three-bond proton–carbon ${}^3J_{\text{CH}}$ couplings. The proton–carbon couplings were calculated by using a tailored Karplus relationship.⁵⁷

Table 1.1 Recommended error thresholds for the NAMFIS NOE deconvolution protocol.

Error (Å)	NOE distances (Å)
± 0.1	$x < 2.5$
± 0.2	$2.5 \leq x < 3.0$
± 0.3	$3.0 \leq x < 3.5$
± 0.4	$3.5 \leq x < 6.0$

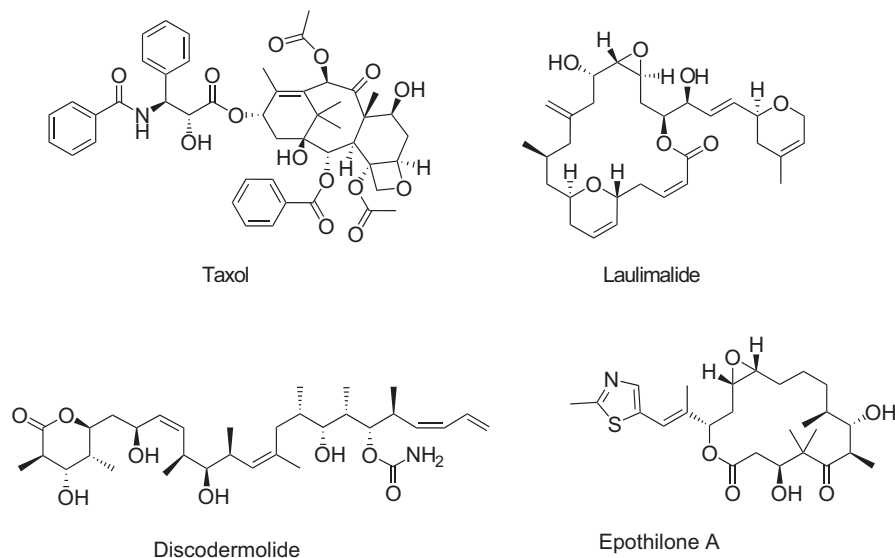


Figure 1.21 Tubulin ligands analyzed with the NAMFIS protocol.

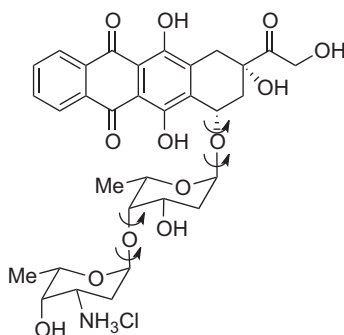


Figure 1.22 Doxorubicin disaccharide.

A graphical frontend for the NAMFIS program has been made available in the Janocchio software package.⁵⁸ Another NAMFIS-like implementation is the DISCON program.⁵⁹ As an additional advantage over the original NAMFIS implementation, DISCON employs a clustering algorithm, based on experimental NOE and scalar coupling data, prior to the NMR data deconvolution step, in this way avoiding the presence of many nearly degenerate solutions of the population determination problem. DISCON was successfully employed in the conformational analysis of (+)-spogistatin-1 (Figure 1.23)⁶⁰ and its analogs⁶¹ using a combination of long-range NOEs, retrieved from a 600 ms NOESY experiment, and $^3J_{\text{HH}}$ couplings.

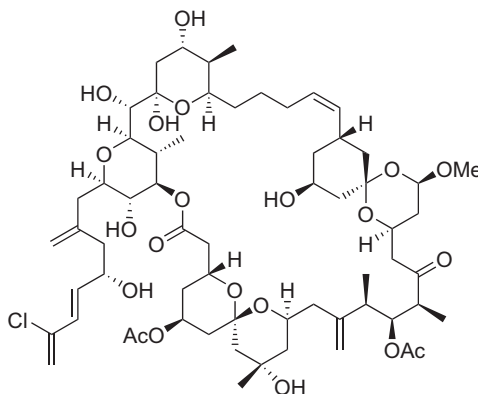


Figure 1.23 (+)-Spoginstantin-1.

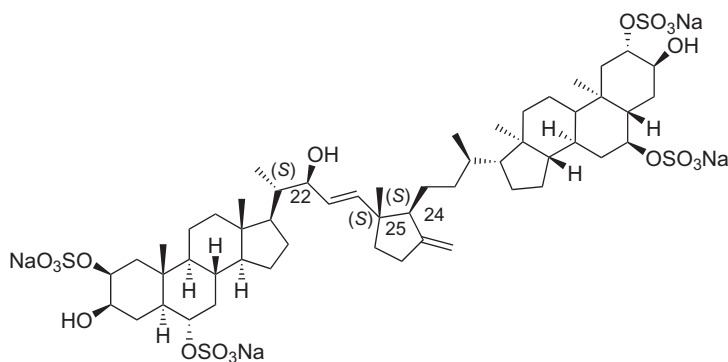


Figure 1.24 Fibrosterol sulfate A.

Griesinger and coworkers employed NOESY-determined distances, in combination with RDCs, for the complete determination of the configuration at the C20,C25,C24' positions of fibrosterol sulfate A, a polysulfated steroid isolated from the sponge *Lissodendoryx fibrosa*.⁶² The authors employed a RDC-based approach where, previous to the RDC fitting step, the conformational space of the molecule was determined by using NOESY-based distances, as well as $^3J_{\text{CH}}$ and $^3J_{\text{HH}}$ couplings, as time-averaged restraints during a MD run.^{63,64} RDC fitting provided strong evidence for the C20,C25,C24'-(SSS) configuration shown in Figure 1.24.

1.11 Conclusions

The NOE family of experiments is still nowadays essential for the configurational/conformational analysis of natural products. The use of PFG-based pulse sequences provides cleaner spectra in a reduced number of scans and

they should be the option of choice. ^{13}C -filtered experiments, such as HSQC-NOESY, may be crucial for the analysis of dimeric structures. By using carefully tailored experimental conditions, the NOE can be brought to a nearly quantitative level, allowing the determination of internuclear distances with high accuracy. In combination with other observables such as scalar couplings, NOE-obtained distances can be employed in NAMFIS and related conformational deconvolution techniques to ascertain the conformational state in solution of highly flexible molecules.

References

1. A. W. Overhauser, *Phys. Rev.*, 1953, **92**, 411.
2. T. Carver and C. Slichter, *Phys. Rev.*, 1956, **102**, 975.
3. A. Abragam and M. Goldman, *Nuclear Magnetism: Order and Disorder*, Oxford University Press, 1982.
4. U. L. Günther, *Top. Curr. Chem.*, 2013, **335**, 23–70.
5. R. A. Wind, M. J. Duijvestijn, C. van der Lugt, A. Manenschijn and J. Vriend, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1985, **17**, 33.
6. A. J. Rossini, A. Zagdoun, M. Lelli, A. Lesage, C. Copéret and L. Emsley, *Acc. Chem. Res.*, 2013, **46**, 1942.
7. I. Solomon, *Phys. Rev.*, 1955, **99**, 559.
8. R. Kaiser, *J. Chem. Phys.*, 1963, **39**, 2435.
9. R. Kaiser, *J. Chem. Phys.*, 1965, **42**, 1838.
10. F. Anet and A. Bourn, *J. Am. Chem. Soc.*, 1965, **87**, 5250.
11. M. Woods, I. Miura, Y. Nakadaira, A. Terahara, M. Maruyama and K. Nakanishi, *Tetrahedron Lett.*, 1967, **8**, 321.
12. M. C. Woods, H-C. Chiang, Y. Nakadaira and K. Nakanishi, *J. Am. Chem. Soc.*, 1968, **90**, 522.
13. D. Neuhaus and M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, 2nd edn, 2000.
14. D. Neuhaus, Nuclear Overhauser Effect, *eMagRes*, 2007, DOI: 10.1002/9780470034590.emrstm0350.pub2. John Wiley & Sons.
15. B. Vögeli, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2014, **78**, 1.
16. R. Wagner and K. Wüthrich, *J. Magn. Reson.*, 1979, **33**, 675.
17. G. Bodenhausen, H. Kogler and R. R. Ernst, *J. Magn. Reson.*, 1984, **58**, 370.
18. T. Parella, F. Sánchez-Ferrando and A. Virgili, *J. Magn. Reson.*, 1997, **125**, 145.
19. R. C. Breton and W. F. Reynolds, *Nat. Prod. Rep.*, 2013, **30**, 501.
20. A. A. Bothner-By, R. L. Stephens, J. Lee, C. D. Warren and R. W. Jeanloz, *J. Am. Chem. Soc.*, 1984, **106**, 811.
21. D. Neuhaus and J. Keeler, *J. Magn. Reson.*, 1986, **68**, 568.
22. T. L. Hwang and A. J. Shaka, *J. Magn. Reson., Ser. B*, 1993, **102**, 155.
23. T. D. W. Claridge, Correlations through space: The nuclear Overhauser effect, *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier, 2009, ch. 8.

24. H. Wang, G. D. Glick and E. R. P. Zuiderweg, *J. Magn. Reson., Ser. A*, 1993, **102**, 116.
25. P. W. A. Howe, *J. Magn. Reson.*, 2006, **179**, 217.
26. M. J. Thrippleton and J. Keeler, *Angew. Chem., Int. Ed.*, 2003, **42**, 3938.
27. K. E. Cano, M. J. Thrippleton, J. Keeler and A. J. Shaka, *J. Magn. Reson.*, 2004, **167**, 291.
28. T. Voloshchuck, N. S. Farina, O. R. Wauchope, M. Kiprowska, P. Haberfield and A. Greer, *J. Nat. Prod.*, 2004, **67**, 1141.
29. G. Lian and B. Yu, *Chem. Biodiversity*, 2010, **7**, 2660.
30. J. Kawabata, E. Fukushima, M. Hara and J. Mizutani, *Org. Magn. Reson.*, 1992, **30**, 6.
31. R. Wagner and S. Berger, *Magn. Reson. Chem.*, 1997, **35**, 199.
32. R. Hoffman, R. Shenhar, I. Willner, H. E. Bronstein, L. T. Scott, A. Rajca and M. Rabinovitz, *Magn. Reson. Chem.*, 2000, **38**, 311.
33. R. M. Gschwind, X. Xie and P. R. Rajamohanan, *Magn. Reson. Chem.*, 2004, **42**, 308.
34. A. V. Buevich, T.-M. Chan, C. H. Wang, A. T. McPhail and A. K. Ganguly, *Magn. Reson. Chem.*, 2005, **43**, 187.
35. K. A. Blinov, A. J. Williams, B. D. Hilton, P. A. Irish and G. E. Martin, *Magn. Reson. Chem.*, 2007, **45**, 544.
36. C. P. Butts, C. R. Jones, E. C. Towers, J. L. Flynn, L. Appleby and N. J. Barron, *Org. Biomol. Chem.*, 2011, **9**, 177.
37. A. Petrovic, A. Navarro-Vázquez and J. Alonso-Gómez, *Curr. Org. Chem.*, 2010, **14**, 1612.
38. M. W. Lodewyk, M. R. Siebert and D. J. Tantillo, *Chem. Rev.*, 2012, **112**, 1839.
39. M. G. Chini, C. R. Jones, A. Zampella, M. V. D'Auria, B. Renga, S. Fiorucci, C. P. Butts and G. Bifulco, *J. Org. Chem.*, 2012, **77**, 1489.
40. S. Di Micco, A. Zampella, M. V. D'Auria, C. Festa, S. De Marino, R. Riccio, C. P. Butts and G. Bifulco, *Beilstein J. Org. Chem.*, 2013, **9**, 2940.
41. M. Lombardo, S. Morganti and C. Trombini, *J. Org. Chem.*, 2003, **68**, 997.
42. F. Freire, J. M. Seco, E. Quiñoá and R. Riguera, *J. Org. Chem.*, 2005, **70**, 3778.
43. J. Tropp, *J. Chem. Phys.*, 1980, **72**, 6035.
44. D. E. Woessner, *J. Chem. Phys.*, 1962, **36**, 1.
45. S. P. Edmondson, *J. Magn. Reson. (1969)*, 1992, **98**, 283.
46. R. Brüschweiler, M. Blackledge and R. R. Ernst, *J. Biomol. NMR*, 1991, **1**, 3.
47. C. Landis and V. S. Allured, *J. Am. Chem. Soc.*, 1991, **113**, 9493.
48. J. Wang, R. S. Hodges and B. D. Sykes, *J. Am. Chem. Soc.*, 1995, **117**, 8627.
49. M. Forster and B. Mulloy, *J. Comput. Chem.*, 1994, **15**, 155.
50. K. Levenberg, *Q. Appl. Math.*, 1944, **2**, 164.
51. D. Marquardt, *J. Soc. Ind. Appl. Math.*, 1963, **11**, 431.
52. C. A. G. Haasnoot, F. A. A. M. De Leeuw and C. Altona, *Tetrahedron*, 1980, **36**, 2783.
53. C. D. Blundell, M. J. Packer and A. Almond, *Bioorg. Med. Chem.*, 2013, **21**, 4976.

54. D. O. Cicero, G. Barbato and R. Bazzo, *J. Am. Chem. Soc.*, 1995, **117**, 1027.
55. N. Nevins, D. Cicero and J. P. Snyder, *J. Org. Chem.*, 1999, **64**, 3979.
56. E. Monteagudo, A. Madami, F. Animati, P. Lombardi and F. Arcamone, *Carbohydr. Res.*, 1997, **300**, 11.
57. C. Morat and F. R. Taravel, *Tetrahedron Lett.*, 1990, **31**, 1413.
58. D. A. Evans, M. J. Bodkin, S. R. Baker and G. J. Sharman, *Magn. Reson. Chem.*, 2007, **45**, 595.
59. O. Atasoylu and A. B. Smith III, DISCON 3.0. <http://onurlab.com/software.html>.
60. O. Atasoylu, G. Furst, C. Risatti and A. B. Smith III, *Org. Lett.*, 2010, **12**, 1788.
61. A. B. Smith III, C. A. Risatti, O. Atasoylu, C. S. Bennett, K. Tendyke and Q. Xu, *Org. Lett.*, 2010, **12**, 1792.
62. H. Sun, U. M. Reinscheid, E. L. Whitson, E. J. d'Auvergne, C. M. Ireland, A. Navarro-Vazquez and C. Griesinger, *J. Am. Chem. Soc.*, 2011, **133**, 14629.
63. A. E. Torda, R. M. Scheek and W. F. van Gunsteren, *Chem. Phys. Lett.*, 1989, **157**, 289.
64. A. E. Torda, R. M. Brunne, T. Huber, H. Kessler and W. F. van Gunsteren, *J. Biomol. NMR*, 1993, **3**, 55.

CHAPTER 2

Assigning Molecular Configuration by Nuclear Magnetic Resonance

KIRK R. GUSTAFSON,^{*a} SUSANNA T. S. CHAN^a AND
YIZHOU LIU^b

^a Molecular Targets Laboratory, Building 562, Room 201, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA;

^b Merck & Co., Inc., NMR Structure Group, Process and Development Research, Mailstop RY-800 D-133, 126 E. Lincoln Ave., Rahway, NJ 07065, USA

*Email: gustafki@mail.nih.gov

2.1 Assigning Relative Configuration

Most natural products are chiral molecules, so the ability to fully assign the relative and absolute configurations of these secondary metabolites is fundamental to natural product structure elucidation. Knowledge of their 3D structure is also key to understanding how these compounds can interact with their biological targets and is essential if they are going to be produced *via* total synthesis. Nuclear magnetic resonance (NMR) is one of the primary spectroanalytical techniques employed to define both the relative and absolute configurations of non-crystalline organic molecules, and recent technological advances allow complete NMR characterization of low-microgram quantities of material.^{1,2} The most common approaches for assigning relative configurations are based on the angular dependence of

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

homonuclear and heteronuclear scalar coupling constants (J), distance constraints from nuclear Overhauser effects (NOEs), empirical NMR database comparisons, density functional theory (DFT) computation-based methods, residual dipolar couplings (RDCs),^{3,4} and residual chemical shift anisotropy (RCSA)^{5–8} measurements. None of these techniques have universal application to the vast array of molecular scaffolds and functional groups found in natural products. The development of a successful strategy to assign the relative configuration of a complex metabolite often requires an integrated application of multiple approaches. New natural products provide unique structure elucidation challenges, and this is particularly true with regard to defining the configuration of stereogenic centers. An extensive review of NMR-based techniques for determining relative stereochemistry was provided by Bifulco and coworkers in 2007,⁹ and a more recent review by Molinski and Morinaka included numerous examples highlighting the application of these techniques to structurally diverse compounds.¹⁰

2.1.1 Coupling Constants and NOEs

The size of vicinal proton–proton couplings ($^3J_{\text{HH}}$) is predictable and can provide valuable information about the relative spatial orientation of the two protons. The Karplus relationship shows that values of $^3J_{\text{HH}}$ are dependent on the dihedral angle (ϕ) between vicinal protons, with maximal couplings for 180° (anti) and 0° (eclipsed) angles, and minimal coupling (near zero) for protons that are oriented 90° from each other.¹¹ Approximate coupling values for unstrained hydrocarbon systems are illustrated in Figure 2.1. The coupling values predicted by the Karplus curve can be impacted by electronegative substituents such as oxygen, as well as by ring constraints, but it has generally proven reliable and is the basis for conformational analysis of many cyclic systems.

For many conformationally rigid ring systems, analysis of vicinal ^1H – ^1H coupling constants and NOE interactions can reveal the relative

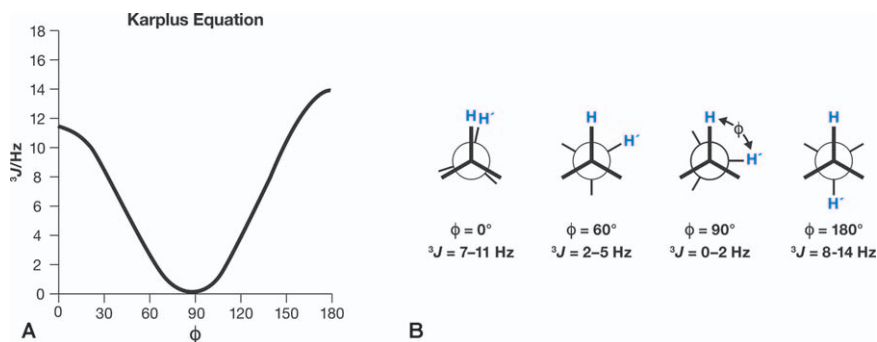


Figure 2.1 (A) The Karplus equation shows that the sizes of vicinal ^1H – ^1H coupling constants are a function of the dihedral angle (ϕ) between the two protons. (B) Newman projections and coupling values of vicinal protons that are eclipsed ($\phi = 0^\circ$), staggered by 60° , offset by 90° , and in an *anti* ($\phi = 180^\circ$) orientation.

configurations of their stereogenic centers. This is especially true with six-membered rings, which are some of the most common structural motifs found in natural products. Six-membered alkyl rings such as cyclohexane generally adopt a low-energy chair conformation and the relationship of vicinal protons can often be inferred from characteristic coupling values. In a chair conformation, 1,2-diaxial protons have large couplings ($J_{aa} = 9\text{--}12\text{ Hz}$), while axial-equatorial and equatorial-equatorial couplings are much smaller (J_{ae} and $J_{ee} = 3\text{--}4\text{ Hz}$) (Figure 2.2A). Identifying axial-axial couplings in these systems is straightforward, but the near identity of J_{ae} and J_{ee} means that couplings to equatorial protons do not provide useful information for defining the configuration of neighboring protons. In six-membered cyclic compounds with electronegative substituents such as pyranose sugars, J values are reduced, with typical J_{aa} values between 7 and 10 Hz. When adjacent protons are both axial, analysis of their coupling constants is sufficient to assign their relative orientation, but when only small vicinal couplings are observed, diagnostic NOE correlations can be useful. A NOE between protons in a 1,3 relationship is indicative of whether they are both axial. NOEs between vicinal protons can also be informative. An equatorial proton can correlate with adjacent axial and equatorial protons, while

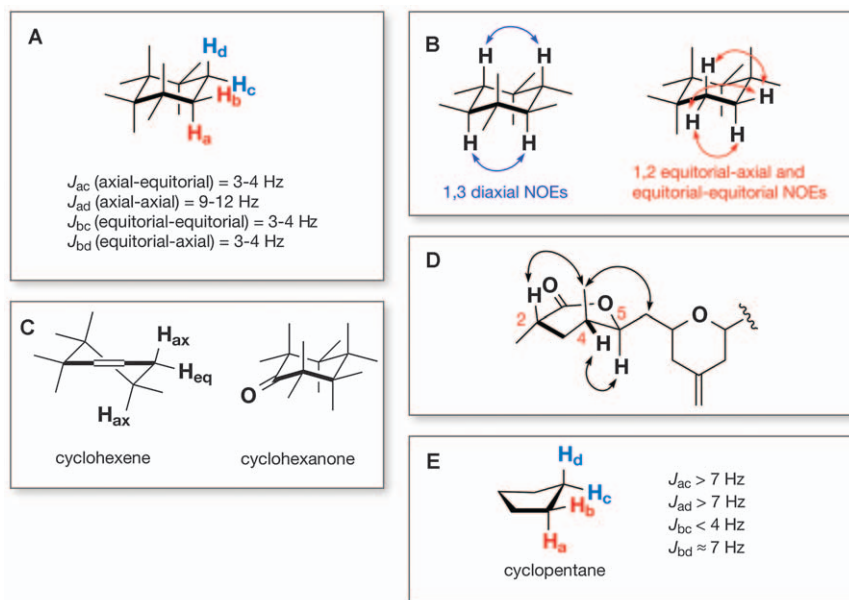


Figure 2.2 (A) Chair conformation of cyclohexane with typical ranges for vicinal $^1\text{H}\text{--}^1\text{H}$ coupling constants. (B) Characteristic NOEs in cyclohexanes. (C) Half-chair conformation of cyclohexene and chair conformation of cyclohexanone. (D) Assignment of the relative configuration of C-2, C-4, and C-5 in the δ -lactone derivative of enigmazole A.¹³ The δ -lactone adopts a half-chair conformation and key NOEs established that H-2, the C-4 methyl group, and the C-5 methylene substituent were all on the same face of the ring. (E) Envelope conformation of cyclopentane with typical vicinal $^1\text{H}\text{--}^1\text{H}$ couplings.

an axial proton only has NOE interactions with adjacent equatorial protons (Figure 2.2B). If 1,4 NOEs are observed in a compound, then the cyclohexane ring exists in a boat conformation and the use of vicinal couplings to define the stereochemistry of the ring is precluded. The characteristic size of the $^3J_{\text{HH}}$ values and the patterns of NOE correlations described above for cyclohexane rings are also observed with cyclohexanones in the chair conformation and for cyclohexenes that adopt the favorable half-chair conformation (Figure 2.2C). For other six-membered ring systems that occur in natural products, $^3J_{\text{HH}}$ and NOEs can be diagnostic if prior studies have carefully defined the properties and conformation that the ring adopts. For example, a previous comprehensive study of δ -lactone rings by Brandänge *et al.*¹² allowed for configurational assignment based on NOE interactions of a key δ -lactone derivative in the structural elucidation of enigmazole A (Figure 2.2D).¹³

Conformational analysis of cyclopentanes is more complicated because the energy difference between the various envelope and twist forms tends to be small, so numerous conformations may be possible for a given structure. For conformationally flexible systems, the measured spin–spin coupling constants are a weighted average from all of the major conformers. Typical vicinal coupling values for a cyclopentane ring in the envelope conformation are illustrated in Figure 2.2E. However, interconversion between different envelope and twist forms of a five-membered ring, along with the effects that different substituents have on $^3J_{\text{HH}}$ values in these systems, can cause wide deviations from these typical values. A recent approach for assigning the relative configuration of five-membered oxolane rings was described by Napolitano *et al.* that utilizes a combination of $^3J_{\text{HH}}$ and $^{2,3}J_{\text{CH}}$ coupling values.¹⁴ The authors proposed that this methodology can be applied to a wide range of carbocyclic and heterocyclic five-membered ring systems, so it could have wide application in natural product studies.

Cyclobutane rings are generally even flatter than cyclopentanes; however, the steric effects of ring substituents or a *trans* ring fusion can cause strong puckering of the four-membered ring. For relatively flat cyclobutanes, *cis* coupling (6–10 Hz) is generally larger than *trans* (2–7 Hz), but when the ring is sufficiently puckered, then *trans* coupling can become larger than *cis*. Due to this type of conformational ambiguity and the effects that different substituents have on $^3J_{\text{HH}}$ values in these ring systems, the size of vicinal couplings alone is generally unreliable for making stereochemical assignments for cyclopentanes and cyclobutanes. For four- and five-membered rings, the NOEs between adjacent *cis* and *trans* protons are often similar in magnitude, but the observation of NOE correlations between 1,3-*cis* protons can help define the relative configuration of the ring at these positions. The dihedral angles in cyclopropanes and three-membered ring heterocycles are rigidly fixed by the ring system geometry. For cyclopropanes, $^3J_{\text{cis}}$ (7–10 Hz) is always larger than $^3J_{\text{trans}}$ (2–6 Hz), so the magnitude of these vicinal coupling constants can be used for configurational assignments. The same relative relationship is true for three-membered heterocyclic rings, but the range of observed couplings is smaller due to electronegative effects of the heteroatom. For epoxides, $^3J_{\text{cis}} \approx 4\text{--}5$ Hz and $^3J_{\text{trans}} \approx 2\text{--}3$ Hz, while

aziridines have couplings of $^3J_{\text{cis}} \approx 6\text{--}7$ Hz and $^3J_{\text{trans}} \approx 3\text{--}4$ Hz. Additional information on the NMR characteristics of a variety of three- to six-membered heterocyclic compounds has been described by Pinto *et al.*¹⁵

2.1.2 Acetonides

The relative configuration of 1,3-diols can be assigned by a ^{13}C -acetonide method originally described by the Rychnovsky laboratory.^{16,17} Formation of an acetonide derivative by reacting a 1,3-diol with 2,2-dimethoxypropane provides a six-membered dioxane ring with a *gem*-dimethyl group. Characteristic ^{13}C chemical shifts of these two acetal methyls are indicative of the relative stereochemistry of the diol. The 1,3-dioxane ring for acetonides derived from *syn*-1,3-diols exists as a well-defined chair conformation with the bulky alkyl substituents oriented in equatorial positions (Figure 2.3). In this acetonide conformation, the ^{13}C chemical shift of the axial methyl group is approximately 19 ppm, while the chemical shift of the equatorial methyl is approximately 30 ppm. In contrast, *anti*-1,3-diol acetonides adopt a twist-boat conformation in which the methyl groups are magnetically equivalent and they both resonate at approximately 25 ppm. The acetal carbons also provide diagnostic chemical shifts of $\delta_{\text{C}} \approx 98.1$ and 100.6 for the *syn* (chair) and *anti* (twist-boat) diol acetonides, respectively.¹⁸ These characteristic ^{13}C chemical shifts of the acetal methyl groups and acetal carbons in 1,3-diol

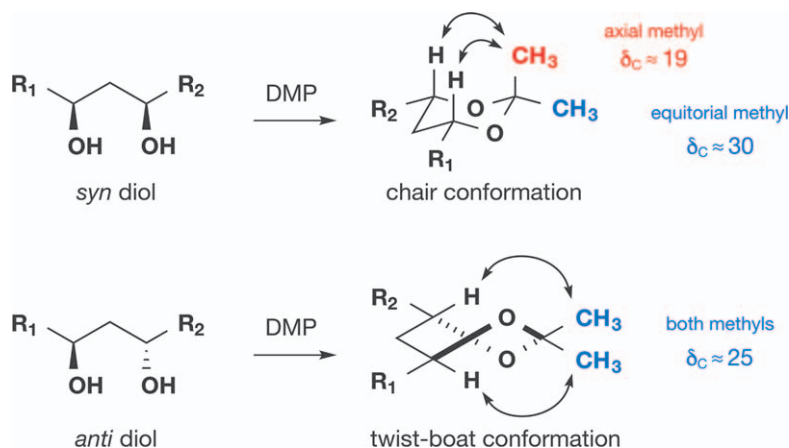


Figure 2.3 Acetonide formation by treating 1,3-diols with 2,2-dimethoxypropane (DMP) provides a six-membered dioxane ring. The acetonide generated from a *syn* diol adopts a chair conformation where the ^{13}C chemical shift of the axial methyl group ($\delta_{\text{C}} \approx 19$) is characteristically lower than the equatorial methyl ($\delta_{\text{C}} \approx 30$) and both oxymethine protons have NOE interactions (black arrows) with the axial methyl protons and not the equatorial methyl. The acetonide from an *anti* diol adopts a twist-boat conformation in which both methyls are magnetically equivalent with similar chemical shifts ($\delta_{\text{C}} \approx 25$). Each oxymethine proton has a NOE interaction with only one methyl group, which is distinct from the methyl that the other oxymethine has a NOE with.

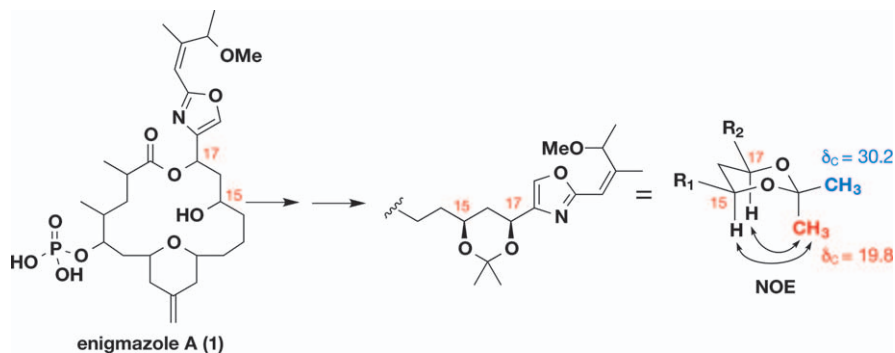


Figure 2.4 Assignment of the relative configurations at C-15 and C-17 of enigmazole A (1).¹³ Hydrolysis of the macrolide ring in 1 gave a linear product with hydroxyl groups at C-15 and C-17. Treatment of this diol with 2,2-dimethoxypropane provided a six-membered 1,3-dioxane ring that was in a chair conformation. One acetonide methyl had a chemical shift that was indicative of an equatorial orientation ($\delta_{\text{C}} = 30.2$) and showed no NOEs with the oxymethine protons, while the other methyl had a chemical shift that was indicative of an axial orientation ($\delta_{\text{C}} = 19.8$) and it showed NOEs with both oxymethine protons. This established the relative configurations at C-15 and C-17 as *syn*.

acetonides uniquely identify the relative configuration of the diol starting material as *syn* or *anti*. This method can also be applied to larger alternating (1,3,5...)polyol chains by preparing two different polyacetonide derivatives that are frameshifted relative to each other.

Diagnostic NOEs between the dioxane ring oxymethine protons and the acetonide methyl groups can also be used to assign the configuration of the diol starting material. In the chair conformation of the *syn* diol acetonide, both oxymethine protons have NOEs with the axial methyl due to their 1,3-diaxial orientation with this group, and they have no NOEs with the equatorial methyl. In the twist-boat conformation of the *anti*-diol acetonide, one oxymethine proton has a NOE with one acetonide methyl, while the other oxymethine proton will correlate with the other methyl group (Figure 2.3). The ability to utilize these characteristic NOEs for configurational assignments can be helpful for mass-limited samples, as these measurements generally require a smaller sample size and less instrument time than is needed to get appropriate ^{13}C NMR data. Application of the acetonide method for assigning the configuration of a 1,3-diol is illustrated in Figure 2.4.

2.1.3 J-Based Analysis—The Murata Method

The relative configuration of adjacent chiral centers in acyclic compounds is defined as *threo* (*syn*) when the highest-priority substituent at each center projects on the same side of the plane that is defined by a zigzag projection of the main carbon chain. When the substituents project on opposite sides of the plane, the relative configuration is *erythro* (*anti*). For adjacent methine carbons, there are three possible staggered rotamers for a *threo*

configuration and three rotamers for an *erythro* configuration, as illustrated by the Newman projections in Figure 2.5A. In these projections, the methine protons have an *anti* orientation when they are opposite from each other and their dihedral angle is approximately 180° . Protons with adjacent orientations have a dihedral angle of approximately 60° and are defined as *gauche*. Following the Karplus relationship, scalar coupling between two *anti* protons is characteristically large (8–12 Hz), while coupling between *gauche* protons is smaller (2–5 Hz). If the compound of interest adopts one main conformer (>85% of the total), then the size of the coupling constant between the adjacent methine protons is indicative of their relative orientations. In a similar manner, the magnitude of vicinal ^1H – ^{13}C couplings ($^3J_{\text{CH}}$) is also dihedral angle dependent, and the size of this long-range heteronuclear coupling is characteristic of an *anti* (large, 6–8 Hz) or *gauche* (small, 1–3 Hz) orientation between the proton and vicinal carbon (Figure 2.5B).

A ground-breaking method first described by Murata and coworkers utilizes these characteristic vicinal ^1H – ^1H and ^1H – ^{13}C couplings, along with $^2J_{\text{CH}}$ values and NOE or ROE interactions to establish the relative configuration of adjacent stereogenic methines.¹⁹ The application of this *J*-based method in the structural elucidation of natural products involves conceptually breaking the structure into a series of fragments and then assigning the relative configurations of the stereogenic centers within these fragments. Appropriate homonuclear ($^3J_{\text{HH}}$) and heteronuclear ($^{2,3}J_{\text{CH}}$) coupling constants, as well as NOE interactions, are determined for each fragment, and these are then compared to predicted values. As illustrated in Figure 2.5C, *threo* and *erythro* diastereomers each have three different staggered rotamers and the theoretical distribution of small and large scalar coupling values can be used to distinguish which rotamer is present in a given system. The four possible *gauche* orientations all provide a unique pattern of small and large couplings, and thus the configuration and rotamer can be identified in this manner. The two *anti* rotamers generate the same coupling pattern, which is distinct from all of the *gauche* rotamers, but this analysis alone cannot distinguish between the *threo* and *erythro* diastereomers. In this case, characteristic NOEs can often be used to make the assignment. When the chiral methine protons in the *threo* diastereomer are *anti*, then the C_1 and C_4 groups are *gauche* and NOE interactions between their respective protons (H_1 and H_4) should be possible (Figure 2.5C). In the *erythro* diastereomer, C_1 and C_4 are *anti*, so no NOE is measured between these groups, but NOEs between C_1 –Y (H_1 and H_Y) and C_4 –X (H_4 and H_X) may be observed if X and Y are protonated. Diagnostic NOEs can be used to assign the configuration of the two different *anti* rotamers, and they can help verify the identity of *gauche* rotamers assigned solely on the basis of homonuclear and heteronuclear coupling constant analysis.

When X or Y is an electronegative oxygen substituent (OH or OR), the geminal C–H couplings ($^2J_{\text{CH}}$) can also be employed in stereochemical assignments. The $^2J_{\text{CH}}$ values are generally negative and their absolute magnitude does not show a C–H dihedral angle dependence, because the proton and carbon do not constitute a dihedral angle. However, the

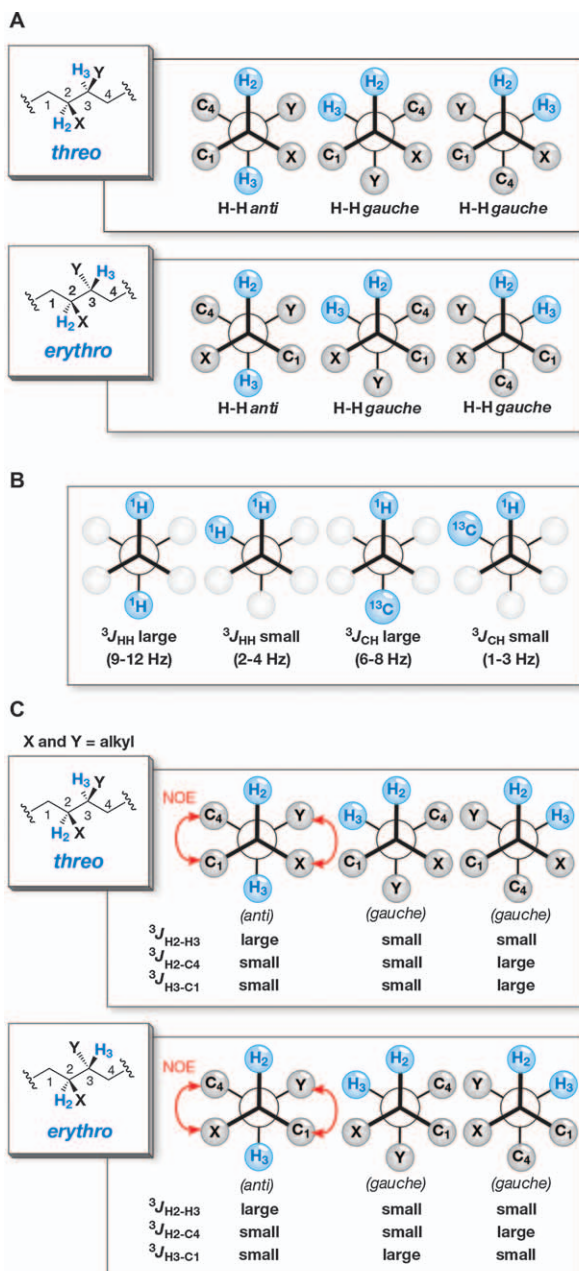


Figure 2.5 (A) The six staggered rotamers of 2,3-substituted *threo* and *erythro* diastereomers. (B) Vicinal (3J) ^1H - ^1H and ^1H - ^{13}C coupling constants for *anti* and *gauche* orientations. J values adapted from Matsumori *et al.*¹⁹ (C) Relative $^3J_{\text{HH}}$ and $^3J_{\text{CH}}$ coupling values for the rotamers of 2,3-substituted methine diastereomers when X and Y are alkyl groups. Adapted and used with permission from: *Handbook of Marine Natural Products*, vol. 1, Springer, London, 2012, pp. 547–570. Copyright 2012, Springer.

magnitude of $^2J_{CH}$ is related to the dihedral angle between the proton and the oxygen substituent on the carbon (Figure 2.6A). When the proton and the oxygenated group on the geminal carbon are *anti*, the absolute magnitude of

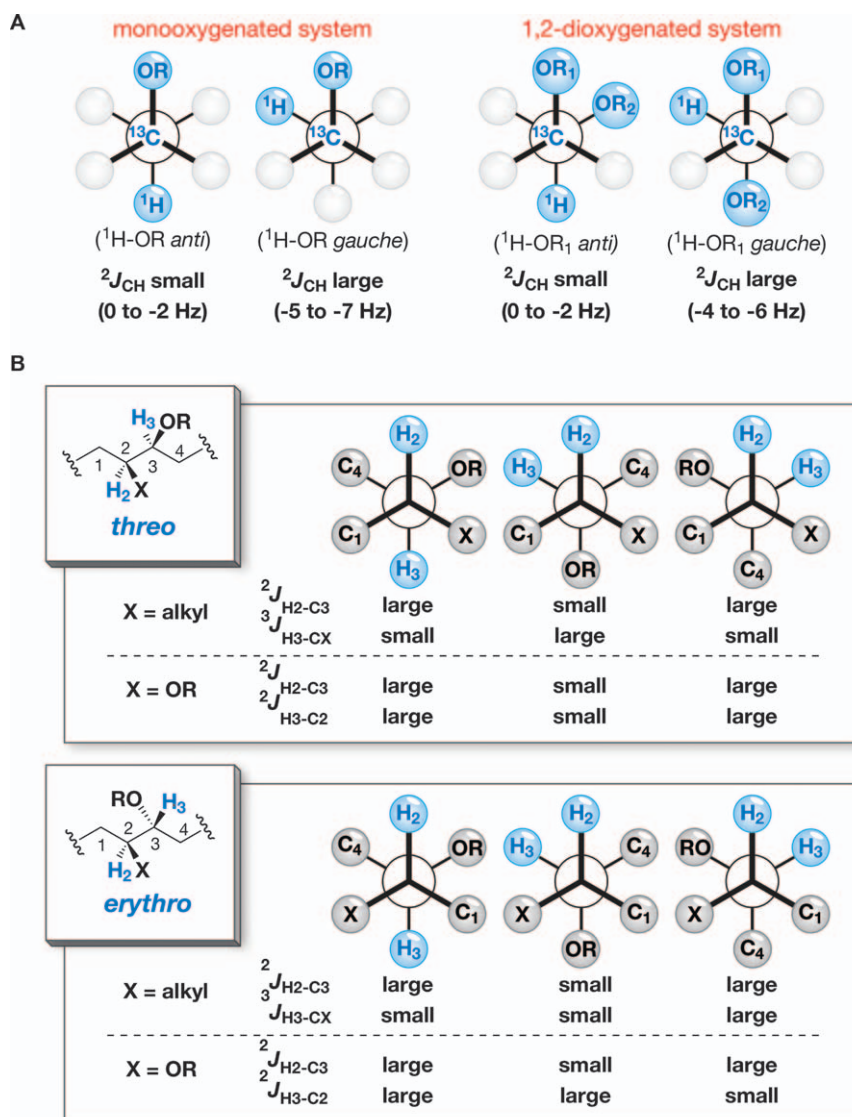
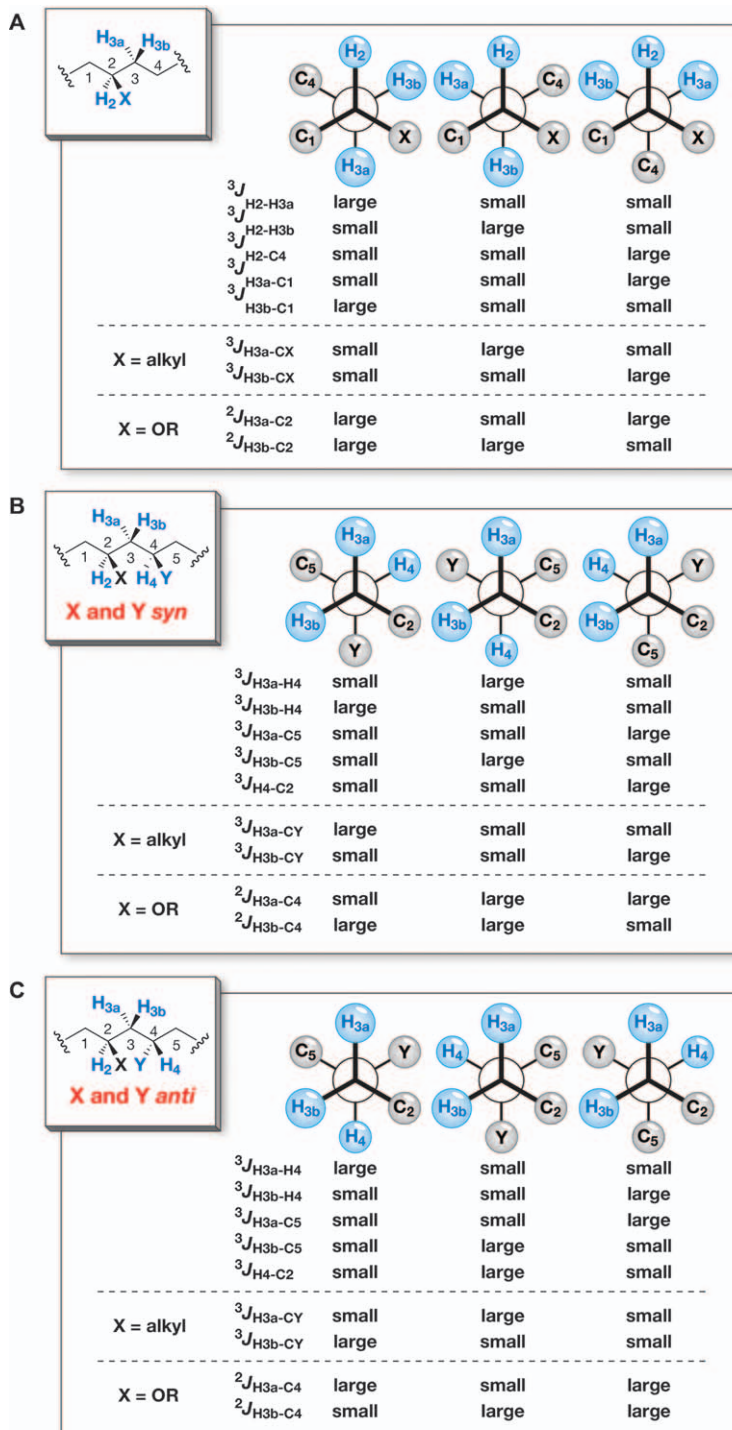


Figure 2.6 (A) Dihedral angle dependence in both monooxygenated and 2,3-dioxygenated systems for the $^2J_{CH}$ coupling between a proton and an adjacent coupled carbon that bears an electronegative oxygen substituent, where R=H or alkyl. (B) $^2J_{CH}$ coupling patterns for the six rotamers of monooxygenated 2,3-methine diastereomers. Adapted and used with permission from: *Handbook of Marine Natural Products*, vol. 1, Springer, London, 2012, pp. 547–570. Copyright 2012, Springer.



$^2J_{\text{CH}}$ is small (0 to -2 Hz). When the proton and oxygenated group are *gauche*, then the magnitude of $^2J_{\text{CH}}$ is larger (-5 to -7 Hz). A similar pattern of *anti* (small) and *gauche* (large) $^2J_{\text{CH}}$ values is also observed with chiral 1,2-dioxygenated systems (Figure 2.6B). The distribution of large and small $^3J_{\text{HH}}$ and $^3J_{\text{CH}}$ couplings in these oxygenated systems is the same as those for alkyl-substituted 1,2-methines, but the magnitude of these three-bond vicinal couplings may be reduced. To assign the *threo* or *erythro* configuration of 1,2-monooxygenated or 1,2-dioxygenated systems, the pattern of the $^2J_{\text{CH}}$ couplings is used in conjunction with NOE measurements and 3J mononuclear and heteronuclear couplings. In spite of the additional insights provided by $^2J_{\text{CH}}$ data, J couplings alone are unable to differentiate between *threo* and *erythro* diastereomers when the methine protons are *anti* in these oxygenated systems. Assignment of the relative configuration in this situation requires the observation of diagnostic NOE interactions.

The J -based configurational method can also be employed in structural fragments with 1,3-stereogenic methines that are separated by a methylene group. To successfully employ this technique, the diastereotopic methylene protons must be sufficiently resolved to allow their stereospecific assignment relative to the adjacent methines. If accurate measurement of the $^3J_{\text{HH}}$ and $^{2,3}J_{\text{CH}}$ coupling constants for these protons is possible, then their *threo* or *erythro* orientations relative to the first methine can be determined from the unique pattern of J values for the six possible rotamers (Figure 2.7). Appropriate NOEs may be necessary to make or confirm these assignments. Once the stereospecific relationship between the individual methylene protons and the first methine is established, their orientations relative to the second methine can be defined in a similar manner using their J values. By analyzing the 1,3-methine system as two overlapping two-carbon fragments, it is possible to interrelate the orientation of each diastereotopic methylene proton with each methine proton and thus correlate the relative configuration of the two methines.

It is possible to extend this J -based approach of conceptually breaking systems into a series of overlapping two-carbon fragments for methine pairs that are separated by more than one methylene group, but lack of dispersion of the methylene protons generally prevents its application beyond 1,4-methine pairs. Kobayashi and coworkers made extensive use of J -based

Figure 2.7 Configurational analysis of a 2,4-methine system when $\text{X} = \text{alkyl}$ or $-\text{OR}$ and $\text{Y} = \text{alkyl}$. (A) The *erythro* or *threo* orientations, relative to H_2 , of the diastereotopic methylene protons H_{3a} and H_{3b} can be assigned based on the pattern of their respective $^1\text{H}-^1\text{H}$ and $^1\text{H}-^{13}\text{C}$ J values. (B & C) Then, the configuration of H_4 relative to H_{3a} and H_{3b} can be established in a similar manner. This process correlates the stereochemical orientation of both H_2 and H_4 with the two diastereotopic methylene protons, and thus the relative configurations at C_2 and C_4 can be assigned as *syn* or *anti*. Adapted and used with permission from: *Handbook of Marine Natural Products*, vol. 1, Springer, London, 2012, pp. 547–570. Copyright 2012, Springer.

methodologies in their assignment of the stereochemistry of the theonezolidides, a series of complex marine macrolides with 23 chiral centers.²⁰ They were able to assign the relative configurations of adjacent methines, 1,3-methines, and 1,4-methine pairs based on *J* values. The application of this technique to two adjacent stereogenic methines is illustrated in Figure 2.8 for a portion of the macrolide ring in theonezolidide B. Other recent examples in which *J*-based techniques were successfully applied include structural studies of the amino acid-derived alkaloid thelepamide,²¹ callyspongiolide, a 14-membered sponge macrolide,²² the linear sugar portion of the alkaloid hunanamyacin A,²³ the mandelalide family of macrolide glycosides,²⁴ and the dinoflagellate toxin karlotoxin-2,²⁵ and they were broadly employed in the structural elucidation of gargantulide A, a complex marine macrolide with 50 stereogenic methine carbons.²⁶ *J*-based

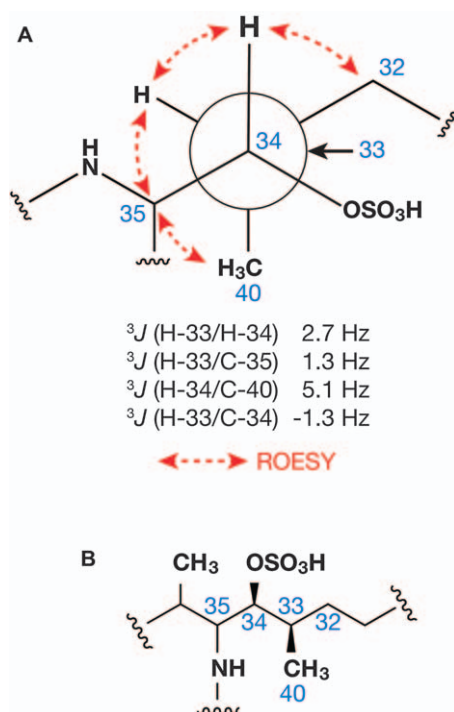


Figure 2.8 (A) Newman projections of the C-34 to C-33 portion of theonezolidide B. H-33 had a small vicinal coupling to H-34 and a small 3J coupling to C-35, thus it was *gauche* to both of these substituents. H-33 had a small 2J coupling to C-34, indicating it was *anti* to the electronegative $-\text{OSO}_3\text{H}$ group. H-34 had a large 3J coupling to the C-40 methyl, indicating an *anti* orientation for these two substituents. These data established that the methyl substituent on C-33 and the sulfate group on C-34 were *syn*. ROESY correlations between H-32 and H-34, H-34 and H-33, H-33 and H-35, and between H-35 and H₃-40 confirmed the *syn* configuration. (B) Structural fragment of theonezolidide B with a *syn* relative configuration at C-33 and C-34.

configurational analysis has proved to be a very powerful tool with wide applications in organic structural studies. It can provide stereochemical details for many conformationally flexible systems that are not amenable to study by other means and it is now routinely utilized in many contemporary natural product studies.

2.1.4 Universal Database

Kishi's group at Harvard advanced the concept and validity of applying a universal NMR database approach in order to assign the relative configurations of complex polyol structures. It is most applicable for polyketide-type natural products composed of polyhydroxylated and methylated linear chains or large flexible macrocycles. Compounds are conceptually broken down into a series of separate structural "stereoclusters" that ideally are separated from other functional groups by two or more bridging atoms, such as methylene carbons. The structural and spectroscopic properties of these stereoclusters are not impacted by steric and stereoelectronic effects from the rest of the molecule, but they are directly related to the specific stereochemical arrangement of substituents within the stereocluster. Thus, the ^1H and ^{13}C NMR chemical shift patterns of a stereocluster are dependent on the configuration of the substituents on the carbon chain and they are unique from the patterns of the other diastereomers. A variety of different polyol-containing structural motifs (Figure 2.9) were stereospecifically synthesized

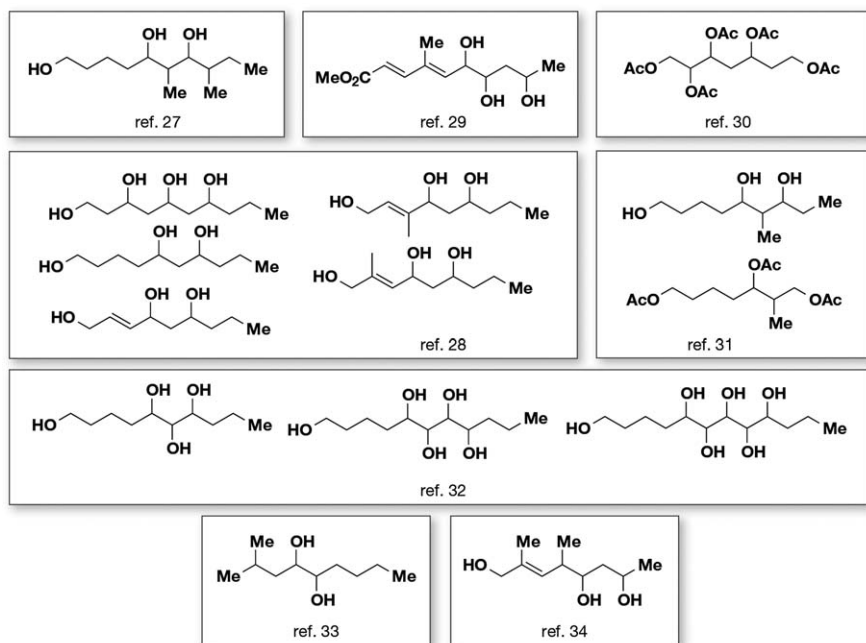
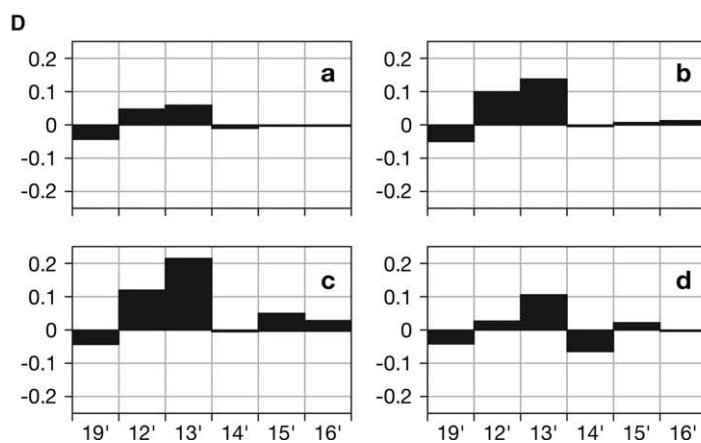
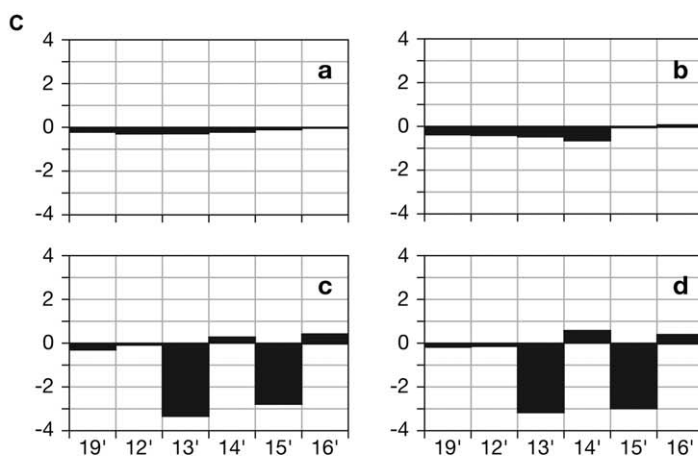
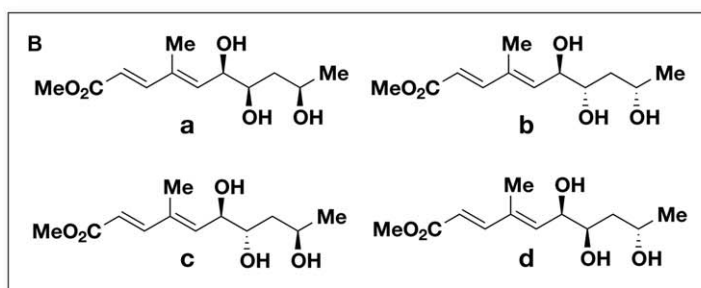
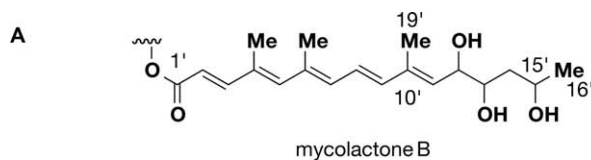


Figure 2.9 Structures of the synthetic polyols in Kishi's universal NMR database.



to generate all of the diastereomers, and these were characterized by ^1H and ^{13}C NMR in order to provide the database.^{27–34}

If a conformationally flexible natural product contains a structural component that is represented in one of the stereocluster libraries, comparison of its NMR data with the synthetic diastereomers of that motif can identify the best match. Comparisons are often done by generating a histogram of the difference ($\Delta\delta$) in the measured chemical shifts in the natural product with the chemical shifts of each associated diastereomer. Close correspondence between the chemical shifts of the natural product and one of the diastereomers results in a histogram with little deviation. When the $\Delta\delta$ values between the natural product and one of the diastereomers are very small, the resulting histogram has little up or down variation, and this indicates that the natural product and the corresponding synthetic compound share the same relative configuration. Comparisons of ^{13}C NMR data are most commonly employed when making these database determinations, but ^1H NMR comparisons can often provide complementary evidence to resolve or reinforce tentative assignments. Application of this approach with the terminal fatty acid portion of mycolactone B²⁹ is illustrated in Figure 2.10.

A straightforward and widely used application of the universal database is assigning the relative configuration of polyols based on the ^{13}C NMR chemical shifts of the carbinol carbons. A pattern of characteristic chemical shifts for alternating polyols was noted by Kishi and colleagues that correlated with the relative stereochemistry of the hydroxyl groups.²⁸ For 1,3,5-triol moieties in conformationally flexible molecules, the chemical shift of the central carbon atom (C-3) is dependent on the 1,3-configuration and the 3,5-configuration, but is largely independent of other structural motifs in a molecule. When the C-1/C-3 and C-3/C-5 configurations are both *anti*, the C-3 carbon falls in the range 66.0 ± 0.5 ppm (CD_3OD). When the configurations are *anti/syn* or *syn/anti*, the chemical shift range is 68.0 ± 0.5 ppm, while C-3 in *syn/syn* configurations resonates at 70.0 ± 0.5 ppm. A similar configuration-dependent distribution of chemical shifts is observed when

Figure 2.10 (A) The universal database was used to assign the relative configurations of C-12', C-13', and C-15' in the terminal fatty acid portion of mycolactone B.²⁹ (B) Synthetic diastereomers **a–d** were compared with the ^1H and ^{13}C NMR data recorded for mycolactone B. (C) Histograms showing the difference in the ^{13}C chemical shifts between select carbons in mycolactone B and the corresponding carbons in diastereomers **a–d**. The x-axis depicts the carbon number and the y-axis is $\Delta\delta$, defined as $\delta_{\text{a–d}} - \delta_{\text{mycolactone B}}$ in ppm. Diastereomers **a** and **b** had very similar carbon chemical shifts with mycolactone B, while **c** and **d** had significant differences, thus C-13'/C-15' are *syn*. (D) Similar comparison between mycolactone B and diastereomers **a–d**, but using proton chemical shift data. Diastereomer **a** had the greatest similarity, so the C-12'/C-13'/C-15' relative configuration could be assigned as *syn/syn*.

^{13}C NMR spectra are acquired in $\text{DMSO}-d_6$, but the chemical shift ranges are slightly different.²⁸ Thus, *anti/anti* and *syn/syn* systems can be directly assigned from their characteristic ^{13}C chemical shift values. While *syn/anti* and *anti/syn* systems can be distinguished from all *anti* and all *syn*, they cannot be differentiated from each other based solely on ^{13}C data. The same approach used with 1,3,5-polyols can be applied to larger alternating polyol moieties by conceptually breaking the larger structure into a series of overlapping 1,3,5-polyol segments. The relative configuration of a series of alternating polyols in the cryptocaryol class of metabolites was assigned in this manner (Figure 2.11),³⁵ and later confirmed by total synthesis.³⁶ Other polyol systems have also been amenable to this type of approach. Isolated 1,3-diols can be assigned from their ^{13}C NMR data since the chemical shift of *syn* diols (≈ 71.5 ppm) is approximately 2 ppm downfield from *anti* diols (≈ 69.2 ppm).²⁸ Databases were also created for adjacent 1,2,3-triol, 1,2,3,4-tetraol, and 1,2,3,4,5-pentaol systems that can be used when these moieties are separated from other structural motifs in a molecule by at least two bridging (CH_2) carbons.³² For extended 1,5-polyols, an empirical stereochemical relationship based on ^1H NMR data was developed by Miyata and Matsunaga, since the central (C-3) methylene protons have degenerate chemical shifts in 1,5-*anti* diols and nonequivalent chemical shifts in 1,5-*syn* diols.³⁷

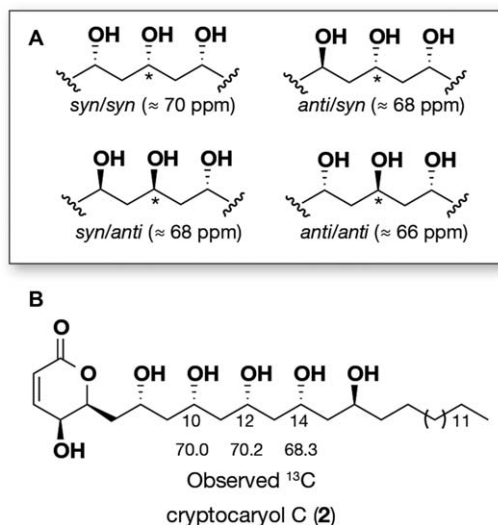


Figure 2.11 (A) Alternating polyols with characteristic ^{13}C NMR chemical shifts (CD_3OD) for the central carbon. (B) Assignment of the relative configurations of C-8, C-10, C-12, C-14, and C-16 in cryptocaryol C (2).³⁵ The chemical shifts of C-10 and C-12 indicated C-8/C-10/C-12/C-14 were *syn/syn/syn*. The chemical shift of C-14 was appropriate for a *syn/anti* configuration; therefore, C-14/C-16 were *anti*.

Kishi has extended the universal database approach to the assignment of absolute configurations of alcohols by acquiring the ^{13}C NMR spectra of the natural product and appropriate diastereomeric reference compounds in a chiral NMR solvent such as (*R*)- and (*S*)-*N*, α -dimethylbenzylamine.^{38,39} Empirical rules were proposed to assign the absolute stereochemistry of saturated secondary alcohols, aryl alcohols, and saturated tertiary alcohols. In a related effort, his group has also explored the use of chiral lanthanide shift reagents to predict the absolute stereochemistry of secondary and tertiary alcohols.^{40,41}

2.1.5 Computational Methods

The increased speed and accuracy of NMR chemical shift prediction by quantum mechanical DFT-based computational methods has made this approach a valuable tool in organic structural analysis, including applications in relative stereochemical assignments.^{9,42} In natural products studies, comparison of the experimentally measured ^{13}C NMR chemical shifts with those calculated for the different potential diastereomers in the molecule can help to define the correct stereostructure. NMR shielding tensors and magnetic susceptibilities can be calculated for each theoretical structure using the DFT gauge-independent atomic orbital (GIAO)⁴³ method, once the structure has been geometry optimized at the quantum mechanical level (*e.g.* using the B3LYP functional/6-31G[d] basis set)⁴⁴ to a population of low-energy conformers. ^{13}C NMR chemical shift calculations are made for each of the conformers and the Boltzmann-averaged chemical shift can then be calculated for each stereoisomer. The calculated values are compared with the corresponding chemical shifts in the natural product and the best fit between the experimental and theoretical datasets is used to assign the relative configuration. A quantitative probability analysis protocol between the experimental and computational chemical shifts, termed DP4, has recently been described by the Goodman group.⁴⁵ DP4 probability analysis of individual chemical shift differences was superior to probabilities based on mean absolute error parameters and correlation coefficients. It is especially well suited for natural products studies where one set of experimental data has to be assigned to one of many possible diastereomers, and the protocol for DP4 calculations is available online (<http://www-jmg.ch.cam.ac.uk/tools/nmr/DP4/>). Recent examples applying these ^{13}C NMR computational methods to elucidate the relative stereochemistry of natural products include studies of thelepamide,²¹ okadaic acid,⁴⁶ and guibourtinidol,⁴⁷ while a more comprehensive coverage of prior applications is provided in the review by Bifulco and colleagues.⁴² Stereochemical analyses using computationally derived ^1H chemical shifts, ^1H - ^1H coupling constants, and ^1H - ^{13}C heteronuclear coupling constants are also possible,⁹ but as yet these techniques have not found broad application in most natural products laboratories.

2.1.6 Residual Chemical Shift Anisotropy

NMR analysis of molecules weakly aligned in anisotropic media in order to assign stereochemical configuration is an area of growing research focus. Partial molecular alignment in cross-linked polymeric gels that swell in deuterated solvents reintroduces measurable anisotropic NMR phenomena such as dipolar couplings and chemical shift anisotropy for molecules in solution. RDCs are treated in Chapter 4 of this volume and have also been the subject of several other chapters.^{3,4} RDCs are measured as changes in nuclear spin–spin couplings upon alignment of a molecule in an anisotropic medium. The same technique used for RDC measurement also makes another residual anisotropic interaction, namely RCSA, observable, at least in principle. RCSA arises from the anisotropic nature of the chemical shielding tensor⁴⁸ and accordingly can provide information on the relative orientations of different such tensors within a molecule. However, the use of RCSA was largely limited by difficulties inherent in its measurement. Indeed, until very recently, the application of RCSA to small molecules was the subject of a scant few reports in the literature,^{6–8} in stark contrast to that of RDC. The major hurdle with RCSA measurement is that changes in isotropic chemical shifts usually take place due to chemical environmental changes manifested by most available alignment techniques; these changes coincide with frequency changes due to RCSA and consequently make extraction of RCSA data equivocal. However, a very recent report by Griesinger, Liu, and coworkers⁵ described two approaches to the acquisition of RCSAs for small-molecule configuration analysis that bypass the aforementioned impediments. The first method makes use of a two-stage NMR tube composed of two sections with different inner diameters (IDs): 4.2 and 3.2 mm (or any ID smaller than 4.2 mm). The wide and narrow sections are used for chemical shift measurements under “weak” and “strong” alignment conditions, respectively, while changes in isotropic chemical shifts under these conditions are avoided. This method is related to a previously published rubber-based gel-stretching apparatus designed for RCSA measurement,^{49–51} but with the advantages of being more stable and solvent compatible due to the use of glass material to support stretching. The second method utilizes a gel-compression device, which offers better sensitivity than the stretching device, but introduces undesired isotropic chemical shift changes into RCSA measurement. However, it was shown that such changes can be eliminated by a highly robust, post-acquisition correction method. It is also worth noting that the variable-angle spinning method has also been used for RCSA measurement in the past, but this approach requires special hardware that is not routinely available.^{7,51,52}

RCSA and RDC have virtually identical realms of application, such as small-molecule constitution and configuration analysis, but they usually have complementary information content. In small molecules, the most readily measurable RDC, namely the one-bond ^1H – ^{13}C RDC, is often limited by availability, while the long-range ^1H – ^{13}C RDC is much more challenging

to measure.^{53,54} Including ^{13}C RCSA data can greatly increase the diversity of orientation sampling, because the ^{13}C chemical shielding tensors generally have principal axes that are not parallel with ^1H – ^{13}C bonds. Therefore, combining RDC and RCSA data will provide more robust and reliable structural diagnosis. Figure 2.12 demonstrates the utility of RCSA in differentiating 13 potential diastereomers of strychnine whose structures were predicted by DFT calculations. The *Q*-factor indicates the level of agreement between experimental values and values predicted from a structural model; a smaller *Q*-factor indicates better agreement. A low *Q*-factor of 0.05 clearly differentiates the correct diastereomer “RSSRRS” from the other candidates, with the second lowest *Q*-factor of 0.10 being for “SSSRSS”. The same result is also provided by RDC data, with the lowest *Q*-factor of 0.07 associated with the correct stereochemistry. Moreover, the variation of RCSA *Q*-factors with different diastereomers also follows a similar trend to that of RDC *Q*-factors. Clearly, having two orthogonal types of data giving the same diagnosis greatly enhances the reliability of the analysis. In Figure 2.12, analysis was performed separately for RCSA and RDC with consistent results. Alternatively, because RCSA and RDC arise from the same molecular alignment process, they can also be combined into one analysis imposing a single uniform alignment tensor during fitting. Figure 2.13 displays two representative correlation plots from a combined analysis, showing pair-wise comparisons of experimental data and data predicted from the correct diastereomer “RSSRRS” (Figure 2.13A) and a typical incorrect diastereomer “RSRSSH” (Figure 2.13B). The correlation plots are good at revealing outliers that arise from structural changes due to incorrect stereochemistry, and can facilitate more rational result interpretation.

RCSA also has certain advantages over RDC. For example, ^1H – ^{13}C RDCs cannot be applied to non-protonated carbons unless significantly less sensitive long-range RDC measurements are performed.^{53,54} In contrast, RCSAs are equally applicable to both protonated and non-protonated carbon resonances. Hence, for proton-deficient compounds, or compounds with quaternary carbons of unknown stereochemistry, information from RCSA can be unique. For example, although RDCs were unable to fully define the configuration of the alkaloid retrorsine since one of the chiral centers, C-11, is a non-protonated carbon, it was instead possible to successfully assign the configuration of retrorsine using RCSAs.⁵ Another advantage of RCSA is its experimental simplicity, since chemical shift measurement is much simpler than coupling constant measurement. This often translates into higher sensitivity with challenging samples. For example, the INEPT transfer used in all modern ^1H – ^{13}C coupling measurements becomes very inefficient when the molecule is strongly aligned, due to transfer loss from ^1H – ^1H RDCs and a high dynamic range of ^1H – ^{13}C RDCs. In contrast, the simple 1D ^{13}C – $\{^1\text{H}\}$ experiment is not complicated by these issues. It was estimated that a few tens of micrograms of material could be sufficient on a 1.7 mm cryoprobe for RCSA measurement.⁵

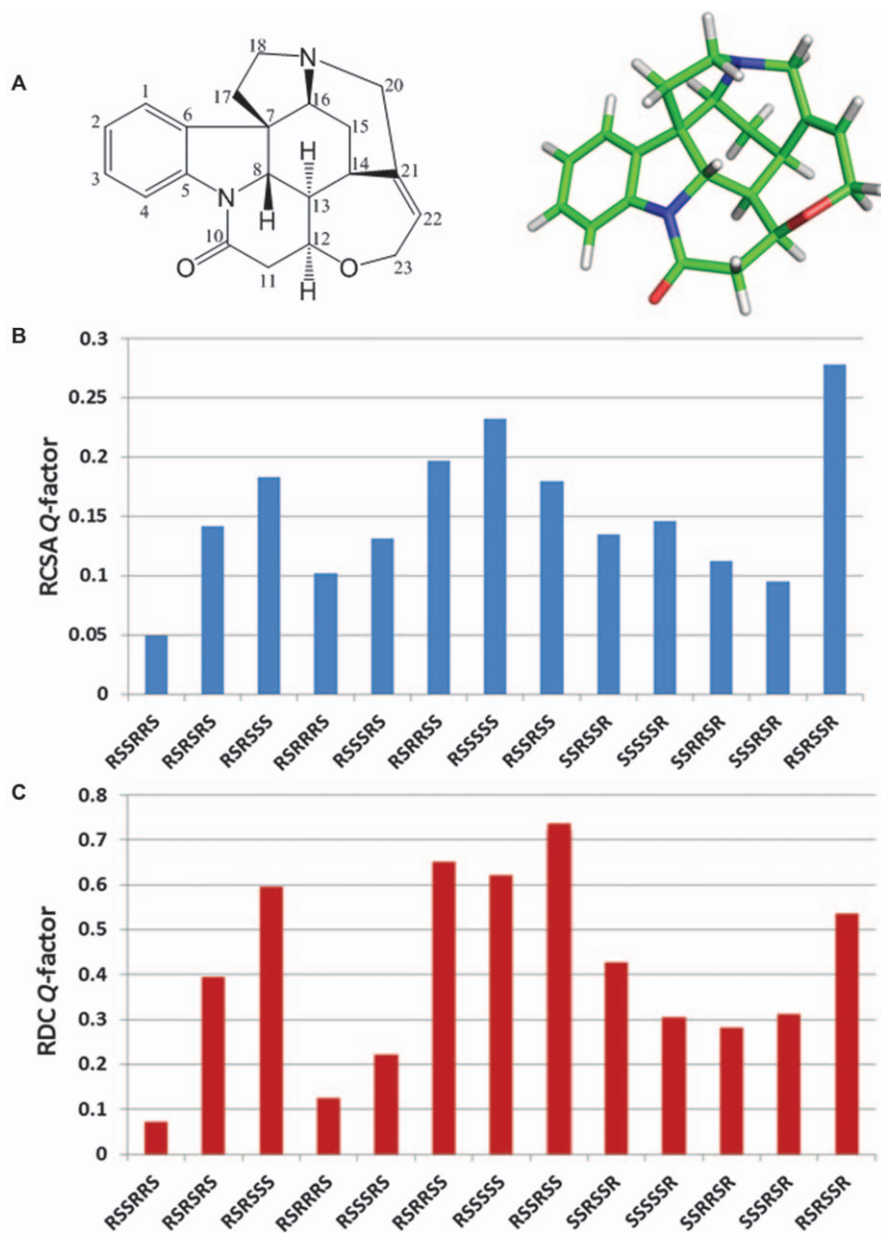


Figure 2.12 Stereochemistry analysis on strychnine by RCSA and RDC. (A) Structure of strychnine “R5SRR5”. (B) RCSA Q-factors of different low-DFT energy diastereomers. (C) RDC Q-factors of different low-DFT energy diastereomers.

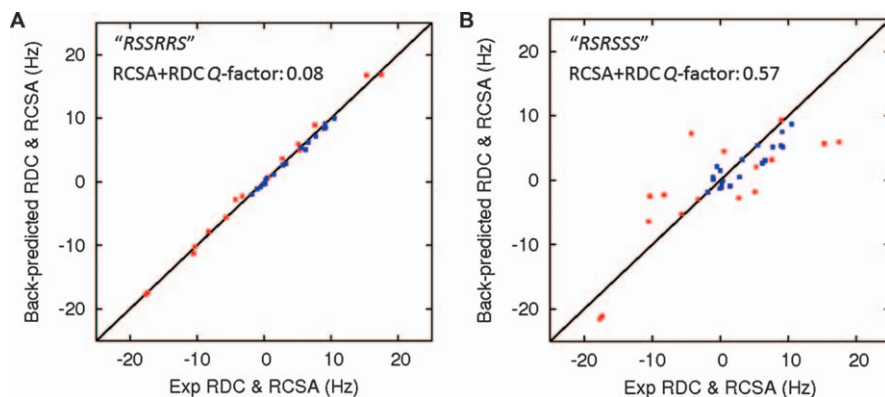


Figure 2.13 “Experimental *versus* back-predicted” correlation plots from an analysis combining RCSA and RDC data from strychnine. RCSA and RDC data are plotted in blue and red, respectively. Results from the correct stereochemistry “RSSRRS” and a representative incorrect stereochemistry “RSRSSS” are displayed in (A) and (B), respectively.

However, RCSA has its own limitations. The sp^2 carbons usually have large chemical shielding anisotropies, and therefore tend to have large RCSA values as well. In contrast, RCSAs from sp^3 carbons are generally small, which makes them more prone to measurement error. For compounds mainly composed of sp^3 carbons, stronger alignment is preferred in order to increase the RCSA data span, and in combination with RDC measurements, this should lead to more robust results. An issue also arises for compounds containing significant numbers of both sp^2 and sp^3 carbons, as the conventional RCSA Q -factor may not be diagnostic regarding the stereochemistry of the sp^3 carbons due to the Q -factor being dominated by data from sp^2 carbons. To address this problem, an improved Q -factor, namely Q_{CSA} , was recently introduced. This takes into consideration the varying CSA sizes of differently hybridized carbons and equalizes the contributions from different types of carbons in the quality-factor calculation.⁵ Finally, application of RCSA usually requires *a priori* knowledge of chemical shielding tensors for a molecule. For carbon atoms, this information can be obtained from GIAO-DFT calculations with accuracy that is sufficient for RCSA-based structural analysis, based on previous publications.^{5–7} However, for heavier nuclear spins, such as ^{15}N and ^{19}F , the accuracy of theoretical chemical shift prediction degrades significantly,⁵⁵ and accurate prediction is even more challenging for elements after the second row, such as ^{31}P ,⁵⁶ which potentially limits the usage of RCSA for these atoms unless information on the chemical shielding tensor is available from experimental sources.

In summary, recent progress in RCSA measurement and data analysis has offered a new route to the structural characterization of small molecules that are challenging for RDC and other more conventional NMR methods. Challenges still exist in application of RCSA to heavier atoms, but the

situation should improve with future developments in quantum chemistry methods.

2.2 Assigning Absolute Configuration

2.2.1 Chiral Anisotropic Reagents

The ability to assign the absolute configuration of chiral carbons in organic compounds is essential for establishing their 3D structure. A variety of techniques including X-ray diffraction analysis, electronic or vibrational circular dichroism, and total synthesis have been employed in this regard. However, there are no universally applicable methodologies that can be effectively applied to all, or even a majority of chiral compounds. For many non-crystalline compounds, defining the absolute configuration of the stereogenic centers they contain can present significant structural elucidation challenges. This is especially true when working with new, structurally complex molecular scaffolds often found in natural products. NMR-based methodologies to make absolute configurational assignments have generally utilized chiral anisotropic reagents to generate specific diastereomeric derivatives that can be used for spectroscopic comparisons. These approaches are useful for characterizing the functional groups that are frequently found in natural products such as secondary alcohols, α -substituted primary amines, and α -substituted carboxylic acids.

The use of chiral anisotropic reagents involves derivatization of a specific functional group associated with the stereogenic center in question with a chiral auxiliary. The chiral reagents contain benzene or other aryl rings that can perturb the local magnetic field experienced by nearby protons due to the induction of an aromatic ring current. Ring currents result from the circulation of aromatic π -electrons in the applied magnetic field of an NMR spectrometer and they induce a secondary magnetic field that can either oppose or reinforce the spectrometer magnetic field. The sign and magnitude of the induced secondary magnetic field experienced by neighboring protons in the molecule are dependent on the spatial orientation of the proton relative to the aromatic ring and the distance from the ring. Thus, ring currents associated with chiral anisotropic reagents generate spatially oriented shielding (+) and deshielding (−) effects on nearby protons that result in characteristic changes in the proton chemical shifts.

One of the most widely used chiral anisotropic moieties is α -methoxy- α -trifluoromethylphenylacetic acid (MTPA), which is commercially available as an acid chloride reagent. Separate treatment of a chiral secondary alcohol with the (*R*)- and (*S*)-enantiomers of MTPA chloride provides (*S*)- and (*R*)-MTPA ester derivatives of the chiral alcohol, respectively. It should be noted that the MTPA ester products have the opposite absolute configuration from the acid chloride reagents due to the functional group priorities and IUPAC rules for nomenclature. The theoretical basis and application of this methodology was pioneered by the Mosher laboratory in 1973 when an empirical

correlation was observed between the absolute configuration of a chiral alcohol and the ^1H NMR chemical shifts of the resulting diastereomeric ester products.^{57,58} This method was further developed and refined using high-field NMR to assign virtually all of the proton resonances of the MTPA esters, and in current applications is commonly referred to as the advanced Mosher's method.⁵⁹ A model was developed to explain the characteristic chemical shift differences between the (*S*)- and (*R*)-MTPA esters and to utilize this difference to assign the configuration of chiral alcohols. It was theorized that the MTPA ester derivatives have a dominant conformation in which the carbonyl and trifluoromethyl groups of MTPA are in the same plane as the original methine carbinol proton (Figure 2.14). The plane defined by these three substituents is referred to as the MTPA plane and this conformational hypothesis was subsequently confirmed by X-ray analysis of various MTPA ester derivatives. The net result is that substituents projecting on the same side of the plane as the MTPA benzyl ring experience a shielding effect from its ring current and their proton resonances are shifted upfield. As illustrated in Figure 2.14, the R_2 substituent of the (*S*)-MTPA ester is co-facial with the benzyl ring, so its protons will resonate at a higher field (reduced ppm) relative to the R_2 protons of the (*R*)-MTPA ester derivative. In a similar manner, R_1 is shielded by the benzyl ring in the (*R*)-MTPA ester and its protons are upfield of the R_1 protons of the (*S*)-MTPA derivative.

The absolute configuration of a chiral secondary alcohol is assigned by preparing both the (*S*)- and (*R*)-MTPA ester derivatives and measuring the precise chemical shifts of all the R_1 and R_2 protons in close proximity to the stereogenic center for each ester. Each proton chemical shift in the (*R*)-MTPA ester (δ_{R}) is then subtracted from the chemical shift of the corresponding proton in the (*S*)-MTPA ester (δ_{S}) and the difference in chemical shifts ($\Delta\delta$) is defined as $\Delta\delta^{\text{S-R}} = \delta^{\text{S}} - \delta^{\text{R}}$. If the MTPA ester adopts a conformation that is consistent with the MTPA plane, then there should be a systematic distribution of positive $\Delta\delta$ values on one of the R substituents and negative $\Delta\delta$

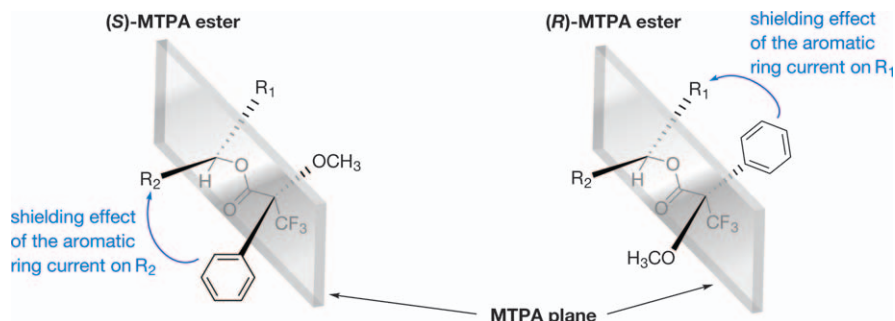


Figure 2.14 The ideal conformation of the MTPA ester in which the original carbinol methine proton, the ester carbonyl, and the trifluoromethyl group all lie in the same plane (the “MTPA plane”). In this conformation, R_2 in the (*S*)-MTPA ester experiences a shielding effect from the phenyl ring current, while R_1 is shielded in the (*R*)-MTPA ester.

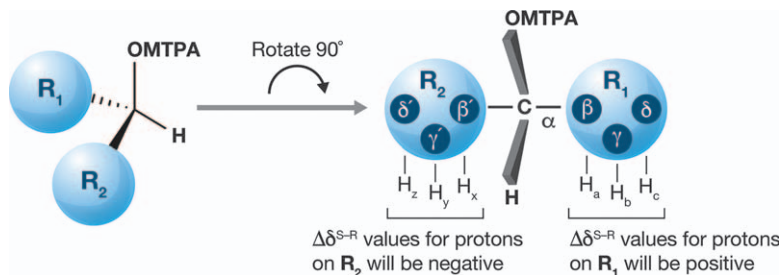


Figure 2.15 When the MTPA ester is viewed as illustrated, protons on R_1 have $\Delta\delta^{S-R}$ values >0 and those on R_2 have $\Delta\delta^{S-R}$ values <0 . If there is a consistent pattern of positive and negative $\Delta\delta^{S-R}$ values about the MTPA plane that clearly delineates R_1 and R_2 , then the absolute configuration of the stereogenic carbon can be assigned.

values on the other substituent. With the $-\text{OMTPA}$ group and carbonyl proton oriented as illustrated in Figure 2.15, then the $\Delta\delta^{S-R}$ values for protons on substituent R_1 will be positive and those on R_2 will be negative. Therefore, when the $\Delta\delta^{S-R}$ values are consistently positive for one substituent and negative for the other, it is possible to define R_1 and R_2 , and thus assign the absolute configuration of the chiral alcohol carbon.

The anisotropic shielding effect of the aryl ring on a neighboring proton is inversely proportional to the distance between the proton and the $-\text{OMTPA}$ group. Thus, the absolute magnitude of $\Delta\delta^{S-R}$ values is dependent on both the distance between a proton and the $-\text{OMTPA}$ group, and their spatial orientation. Protons with $\Delta\delta^{S-R} = 0$ are either too far from the MTPA aryl ring to be impacted by its ring current or they lie on the MTPA plane. In the ideal MTPA ester conformation, the carbonyl proton aligns with the MTPA plane and its $\Delta\delta^{S-R}$ value is zero. However, at times, this proton may deviate from the MTPA plane and/or it can experience anisotropic effects from the MTPA ester carbonyl group that result in $\Delta\delta^{S-R}$ values that are non-zero. As such, $\Delta\delta^{S-R}$ values for the carbonyl proton are unreliable and should not be used when making a configurational assignment. For some MTPA ester derivatives, the positive and negative $\Delta\delta^{S-R}$ values are not systematically distributed on the two R substituents relative to the MTPA plane. This results when steric factors cause the MTPA ester to adopt a conformation that deviates significantly from the ideal. When both positive and negative $\Delta\delta^{S-R}$ values are observed for a series of protons that the model predicts should all be on the same side of the MTPA plane, then the advanced Mosher's method cannot be used to assign the configuration. In some instances, it may only be possible to determine $\Delta\delta^{S-R}$ values for protons on one of the R substituents. This situation can result when one substituent has no protons in close proximity to the $-\text{OMTPA}$ group or when signal overlap in the ^1H NMR spectra prevents accurate chemical shift measurements. When $\Delta\delta^{S-R}$ values are irregularly distributed relative to the MTPA plane, or when they cannot be determined for both R substituents, then this technique is unreliable and

cannot be used to unambiguously assign the absolute configuration of the secondary alcohol in question.

Application of the advanced Mosher's method with a natural product involves separate treatment of the chiral secondary alcohol with (*R*)- and (*S*)-MTPA chloride to provide the (*S*)- and (*R*)-MTPA ester derivatives, respectively. The reaction is generally run in a mild base such as pyridine or triethylamine, with 4-dimethylaminopyridine used as an acylation catalyst. The assignment and accurate chemical shift measurement of all of the relevant proton signals can normally be accomplished by analysis of ^1H , COSY, and HSQC NMR datasets. Accurate comparison of the chemical shifts of the two diastereomeric MTPA esters requires that both NMR samples are made up at the same concentration and precisely referenced. Caution should be exercised when a compound under study contains two or more chiral alcohols that can be derivatized with MTPA chloride. When two MTPA esters are present in the same molecule, the local magnetic field that a proton experiences can potentially be impacted by the ring currents of both MTPA aryl groups. If the sign of the induced magnetic fields that a proton experiences are the same, then they complement each other and interpreting the $\Delta\delta^{\text{S-R}}$ values is straightforward, as illustrated in Figure 2.16. However, if the induced magnetic fields have opposite signs in a proton's environment, then they will counteract each other and the net effect on the proton chemical shift is difficult to predict. This can lead to a non-uniform distribution of positive and negative $\Delta\delta^{\text{S-R}}$ values relative to the two MTPA planes. If two chiral alcohols are spatially far removed in a compound, then the resulting MTPA esters may be sufficiently separated that neighboring protons only experience a measureable effect of one of the esters. However, when two MTPA esters are close enough that they both impact neighboring proton chemical shifts, then the Mosher's method may be unsuitable for absolute configuration assignments. Riguera's group developed a method to define the absolute configuration of 1,2-, 1,3-, 1,4-, and 1,5-diols using α -methoxyphenylacetic acid (MPA) as the chiral auxiliary to produce bis-(*R*)- and bis-(*S*)-MPA esters. The configuration of the diol is correlated with the $\Delta\delta$ values of protons on the R_1 and R_2 substituents and with the $\Delta\delta$ of the methine protons attached to the stereogenic carbons in question.⁶⁰

Chiral anisotropic reagents can also be employed to define stereogenic methine carbons that are substituted with a primary amine group.⁶¹ When this functional group is present in a natural product, the (*S*)- and (*R*)-MTPA amides are prepared and analyzed in an analogous manner to the esters. When the -NHMTPA amide is viewed as illustrated in Figure 2.17A, the R_1 ($\Delta\delta^{\text{S-R}} > 0$) and R_2 ($\Delta\delta^{\text{S-R}} < 0$) groups of the methine are assigned based on their $\Delta\delta$ values, and this defines the absolute configuration of the methine carbon. Chiral methines that are α to a carboxylic acid can be elucidated with reagents such as phenylglycine methyl ester (PGME). Condensation of the carboxylic acid with this chiral auxiliary is usually accomplished in the presence of a peptide coupling reagent (PyBOP[®]), a racemization inhibitor (HOBt), and a base such as triethylamine. Both the (*S*)- and (*R*)-PGME amides

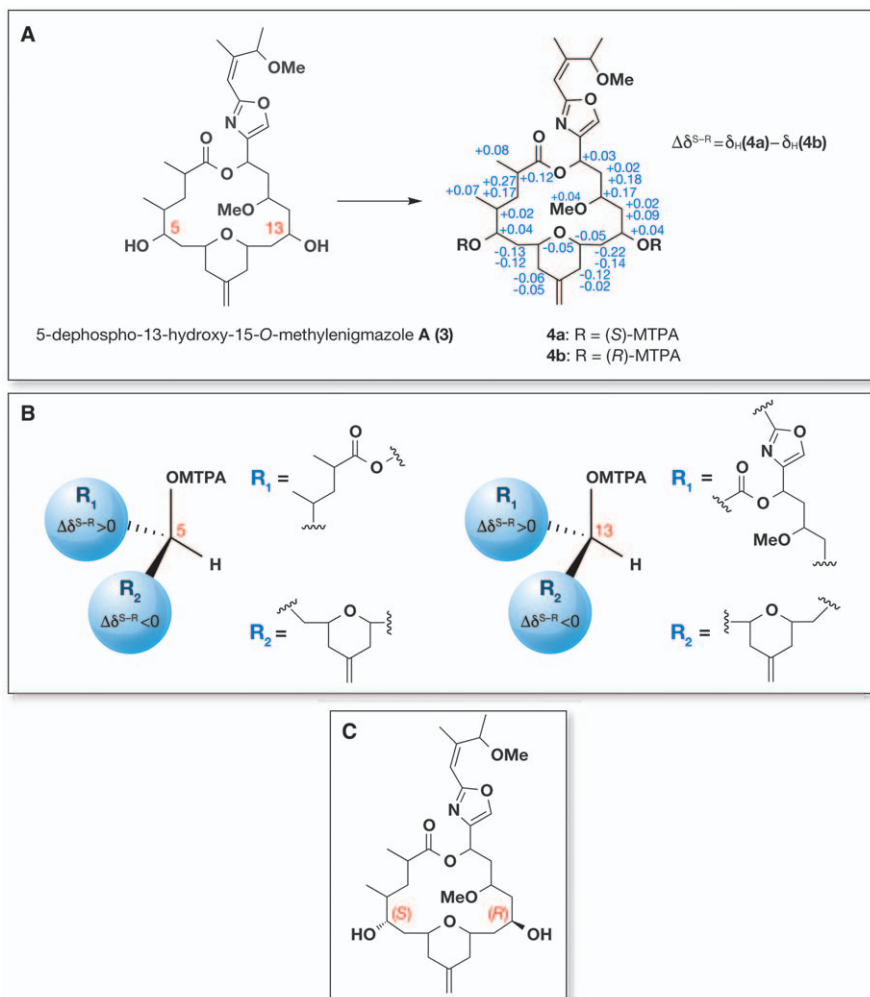


Figure 2.16 Assignment of the absolute configurations of C-5 and C-13 of 5-dephospho-13-hydroxy-15-O-methylenigmazole **A** (**3**).¹³ (A) Treatment of diol **2** with (*R*)- or (*S*)-MTPA chloride provided the *bis*-(*S*)- and *bis*-(*R*)-MTPA esters, respectively. The $\Delta\delta$ values were calculated for protons around the macrocyclic ring and they showed a systematic distribution of positive values (northern hemisphere) and negative values (southern hemisphere). (B) When the MTPA esters were modeled as shown for C-5 and for C-13, the R_1 and R_2 substructures at each stereogenic center could be identified. This established the absolute configuration at C-5 as *S* and at C-13 as *R*. (C) The structure of 5-dephospho-13-hydroxy-15-O-methylenigmazole **A** (**2**) with *S*-configuration at C-5 and *R*-configuration at C-13.

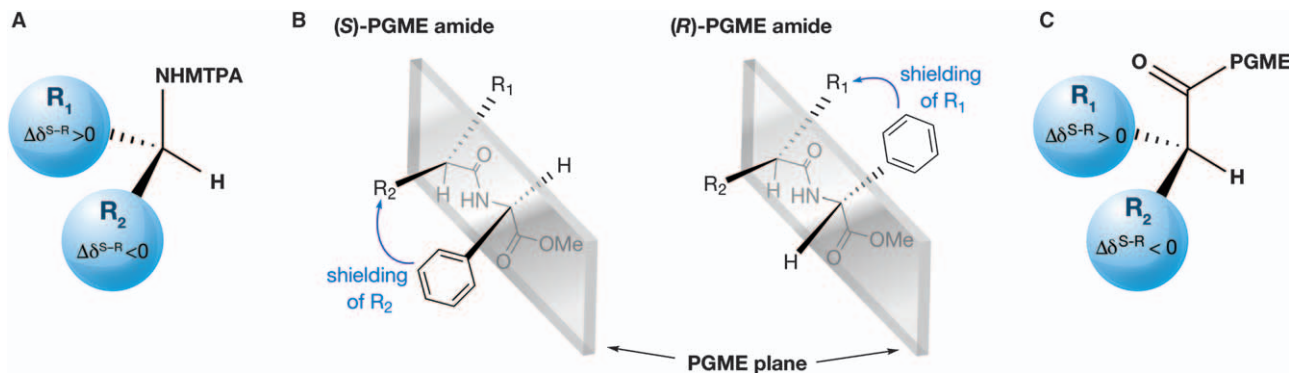


Figure 2.17 (A) When the MTPA amide is viewed as shown, $\Delta\delta^{S-R}$ values for R_1 are positive and are negative for R_2 . (B) The PGME amides adopt a conformation in which the α -methine proton, amide carbonyl and ester carbonyl define the PGME plane. R_2 is shielded by the phenyl ring in the (S)-PGME amide and R_1 is shielded in the (R)-PGME amide. (C) Characteristic distribution of $\Delta\delta^{S-R}$ values in a PGME amide used to assign the absolute configuration of the α -methine carbon. Adapted and used with permission from: *Handbook of Marine Natural Products*, vol. 1, Springer, London, 2012, pp. 547–570. Copyright 2012, Springer.

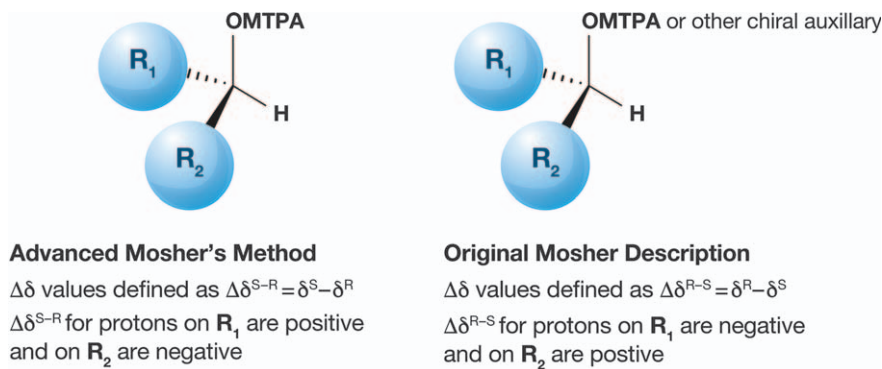


Figure 2.18 How $\Delta\delta$ is calculated in the advanced Mosher's method⁵⁹ differs from the original Mosher description,^{57,58} and this impacts which substituent, R_1 or R_2 , has positive values and which has negative.

are generated from (*S*)- and (*R*)-PGME, respectively, and the amide products adopt a conformation that defines a PGME plane (Figure 2.17B). The $\Delta\delta^{S-R}$ values for the neighboring protons on the diastereomers are calculated, and the R_1 and R_2 substituents are determined as shown in Figure 2.17C. A variety of other chiral reagents, including phenylglycine dimethylamide, have also been used to assign the configuration of carboxylic acids with a chiral α -methine.⁶¹

While MTPA is the most widely used chiral auxiliary reagent in natural product structure elucidation studies, there are a variety of other chiral anisotropic reagents one could use. The theory behind their application is similar to that for MTPA, and in many cases they have been reported to produce larger proton chemical shift differences in the resulting ester diastereomers.⁶² MPA has been extensively studied and utilized by the Riguera laboratory and they have advocated its use over MTPA.^{63,64} An interesting feature of this reagent is that the $\Delta\delta$ values for the diastereomeric MPA esters increase when the ^1H NMR spectra are acquired at lower temperatures. It should be noted that the chemical literature detailing the use of chiral anisotropic reagents can be confusing due to differences in the way $\Delta\delta$ is calculated. The advanced Mosher's method defined $\Delta\delta$ as equal to $\delta^S - \delta^R$ ($\Delta\delta^{S-R}$), and this convention has been adopted in many natural product studies. However, the original description of the technique by Mosher employed $\Delta\delta = \delta^R - \delta^S$ ($\Delta\delta^{R-S}$), and this definition has been widely utilized by Riguera and others. These two contrasting definitions of $\Delta\delta$ still lead to the same stereochemical assignments because the model used to assign the R_1 and R_2 substituents is dependent on which definition of $\Delta\delta$ is employed (Figure 2.18).

2.3 Conclusions

In this chapter, we have subdivided the important topic of assigning molecular configuration into two broad categories: assigning relative

configuration and assigning absolute configuration. The former is more extensively covered since it lends itself more readily to NMR methods and is probably more familiar to many readers. The latter category is treated more briefly and reflects the narrower focus of NMR-based work in that area, as well as the relative difficulty of assigning absolute configuration.

Analysis of proton–proton coupling constants and NOE (or ROE) measurements are fundamental techniques in natural products structural elucidation studies, and the Karplus relationship provides a straightforward approach to assigning relative configuration in many instances. Rychnovsky and coworkers recognized that the configuration of 1,3-diols could be assigned by a ^{13}C acetone method that involves derivatizing the 1,3-diol with 2,2-dimethoxypropane to form a six-membered 1,3-dioxane ring. The characteristic ^{13}C shifts of the resulting *gem*-dimethyl groups form the basis for the configuration assignment. Murata developed what has become known as *J*-based configuration analysis, frequently referred to by the acronym JBCA. The relative magnitudes of heteronuclear ^1H – ^{13}C coupling constants follow well-defined patterns and the method has been used to assign the configurations of some extremely complicated molecules, including gargantulide A, a complex marine macrolide with a remarkable 50 stereogenic methine carbons. Work in Kishi's laboratories at Harvard developed and validated a universal NMR database approach to assigning the relative configurations of complex polyol structures. While the method is limited to polyols, given the abundance of highly oxygenated natural products, there is a wide array of structural problems that can be addressed by this approach. Computing power and speed have increased dramatically over the past decade, making it feasible to routinely utilize quantum mechanical DFT to address a diverse assortment of chemical structure problems. Applications range from chemical shift and coupling constant predictions for energy-minimized structures to calculations of the chemical shielding tensor used in RCSA methods for assigning configurations to molecules in a partially anisotropic environment. Although RCSAs have been known of for many years, it is only recently that the difficulties inherent to their measurement have been overcome, thus making this an exciting area of investigation for assigning molecular configuration. RCSAs are experimentally simpler to measure than RDCs and they can be readily applied to non-protonated stereogenic carbons. While RDCs can be applied to quaternary carbons, this requires measurement of $^nD_{\text{CH}}$, which is a relatively low-sensitivity measurement. It remains to be seen how researchers will apply RCSA data, but the range of possibilities is broad. The last section of this chapter is devoted to the assignment of absolute configuration by NMR methods that utilize chiral anisotropic reagents to derivatize specific functional groups associated with the stereogenic center of a molecule. For most complex natural products, application of various NMR methods is required to assign the configuration of all of the stereogenic centers. This chapter is not intended to be a rigorous treatise on methods for stereochemical analysis, but rather is intended to provide the interested reader with a pathway into

the extensive literature associated with this area of NMR investigation that can be explored in more depth as necessary to solve a given problem.

References

1. T. F. Molinski, *Nat. Prod. Rep.*, 2010, **27**, 321.
2. T. F. Molinski, *Curr. Opin. Biotechnol.*, 2010, **21**, 819.
3. G. Kummerlöwe and B. Luy, in *Annual Reports on NMR Spectroscopy*, ed. G. A. Webb, Academic Press, 2009, vol. 68, p. 193.
4. R. R. Gil, C. Griesinger, A. Navarro-Vázquez and H. Sun, in *Structure Elucidation in Organic Chemistry – The Search for the Right Tools*, eds. M.-M. Cid and J. Bravo, Wiley-VCH, New York, 2015, p. 279.
5. N. Nath, M. Schmidt, R. R. Gil, R. T. Williamson, G. E. Martin, A. Navarro-Vázquez, C. Griesinger and Y. Liu, *J. Am. Chem. Soc.*, 2016, **138**, 9548.
6. F. Hallwass, M. Schmidt, H. Sun, A. Mazur, G. Kummerlow, B. Luy, A. Navarro-Vázquez, C. Griesinger and U. M. Reinscheid, *Angew. Chem., Int. Ed.*, 2011, **50**, 9487.
7. G. Kummerlowe, S. L. Grage, C. M. Thiele, I. Kuprov, A. S. Ulrich and B. Luy, *J. Magn. Reson.*, 2011, **209**, 19.
8. Y. Z. Liu and J. H. Prestegard, *J. Biomol. NMR*, 2010, **47**, 249.
9. G. Bifulco, P. Dambruoso, L. Gomez-Paloma and R. Riccio, *Chem. Rev.*, 2007, **107**, 3744.
10. T. F. Molinski and B. I. Morinaka, *Tetrahedron*, 2012, **68**, 9307.
11. M. Karplus, *J. Am. Chem. Soc.*, 1963, **85**, 2870.
12. S. Brandänge, M. Färnäck, H. Leijonmarck and A. Sundin, *J. Am. Chem. Soc.*, 2003, **125**, 11942.
13. N. Oku, K. Takada, R. W. Fuller, J. A. Wilson, M. L. Peach, L. K. Pannell, J. B. McMahon and K. R. Gustafson, *J. Am. Chem. Soc.*, 2010, **132**, 10278.
14. J. G. Napolitano, J. A. Gavín, C. García, M. Norte, J. J. Fernández and A. H. Daranas, *Chem. – Eur. J.*, 2011, **17**, 6338.
15. D. C. G. A. Pinto, C. M. M. Santos and A. M. S. Silva, in *Recent Research Developments in Heterocyclic Chemistry*, ed. M. V. D. Pinho e Melo, Research Signpost, Kerala, India, 2007, 397–475.
16. S. D. Rychnovsky and D. J. Skaltzky, *Tetrahedron Lett.*, 1990, **31**, 945.
17. S. D. Rychnovsky, B. Rogers and G. Yang, *J. Org. Chem.*, 1993, **58**, 3511.
18. D. A. Evans, D. L. Rieger and J. R. Gage, *Tetrahedron Lett.*, 1990, **31**, 7099.
19. N. Matsumori, D. Kaneno, M. Murata, H. Nakamura and K. Tachibana, *J. Org. Chem.*, 1999, **64**, 866.
20. K. Nozawa, M. Tsuda, N. Tanaka, T. Kubota, E. Fukushi, J. Kawabata and J. Kobayashi, *Tetrahedron Lett.*, 2013, **54**, 783.
21. J. Rodríguez, R. M. Nieto, M. Blanco, F. A. Valeriote, C. Jiménez and P. Crews, *Org. Lett.*, 2014, **16**, 464.
22. C.-D. Pham, R. Hartmann, P. Böhrer, B. Stork, S. Wesselborg, W. Lin, D. Lai and P. Proksch, *Org. Lett.*, 2014, **16**, 266.

23. Y. Hu, K. Wang and J. B. MacMillan, *Org. Lett.*, 2013, **15**, 390.
24. J. Sikorska, A. M. Hau, C. Anklin, S. Parker-Nance, M. T. Davies-Coleman, J. E. Ishmael and K. L. McPhail, *J. Org. Chem.*, 2012, **77**, 6066.
25. J. Peng, A. R. Place, W. Yoshida, C. Anklin and M. T. Hamann, *J. Am. Chem. Soc.*, 2010, **132**, 3277.
26. J.-R. Rho, G. Subramaniam, H. Choi, E.-H. Kim, S. P. Ng, K. Yoganathan, S. Ng, A. D. Buss, M. S. Butler and W. H. Gerwick, *Org. Lett.*, 2015, **17**, 1377.
27. Y. Kobayashi, J. Lee, K. Tezuka and Y. Kishi, *Org. Lett.*, 1999, **1**, 2177.
28. Y. Kobayashi, C.-H. Tan and Y. Kishi, *Helv. Chim. Acta*, 2000, **83**, 2562.
29. S. Fidanze, F. Song, M. Szlosek-Pinaud, P. L. C. Small and Y. Kishi, *J. Am. Chem. Soc.*, 2001, **123**, 10117.
30. Y. Kobayashi, C.-H. Tan and Y. Kishi, *Angew. Chem., Int. Ed.*, 2000, **39**, 4279.
31. Y. Kobayashi, C.-H. Tan and Y. Kishi, *J. Am. Chem. Soc.*, 2001, **123**, 2076.
32. S. Higashibayashi, W. Czechtizky, Y. Kobayashi and Y. Kishi, *J. Am. Chem. Soc.*, 2003, **125**, 14379.
33. S. Higashibayashi and Y. Kishi, *Tetrahedron*, 2004, **60**, 11977.
34. Y. Kishi, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 6703.
35. T. Grkovic, J. S. Blees, N. H. Colburn, T. Schmid, C. L. Thomas, C. J. Henrich, J. B. McMahon and K. R. Gustafson, *J. Nat. Prod.*, 2011, **74**, 1015.
36. Y. Wang and G. A. O'Doherty, *J. Am. Chem. Soc.*, 2013, **135**, 9334.
37. Y. Miyata and S. Matsunaga, *Tetrahedron Lett.*, 2008, **49**, 6334.
38. Y. Kobayashi, N. Hayashi, C.-H. Tan and Y. Kishi, *Org. Lett.*, 2001, **3**, 2245.
39. Y. Kobayashi, N. Hayashi and Y. Kishi, *Tetrahedron Lett.*, 2004, **44**, 7489.
40. I. Ghosh, H. Zeng and Y. Kishi, *Org. Lett.*, 2004, **6**, 4715.
41. I. Gosh, Y. Kishi, H. Tomoda and S. Omura, *Org. Lett.*, 2004, **6**, 4719.
42. S. Di Micco, M. Giovanna Chini, R. Riccio and G. Bifulco, *Eur. J. Org. Chem.*, 2010, **8**, 1411.
43. K. Wolinski, J. F. Hinto and P. Pauly, *J. Am. Chem. Soc.*, 1990, **112**, 8251.
44. K. Kim and K. D. Jordan, *J. Chem. Phys.*, 1994, **98**, 10089.
45. S. G. Smith and J. M. Goodman, *J. Am. Chem. Soc.*, 2010, **132**, 12946.
46. H. J. Domínguez, G. D. Crespín, A.-J. Santiago-Benítez, J. A. Gavín, M. Norte, J. J. Fernández and A. H. Daranas, *Mar. Drugs*, 2014, **12**, 176.
47. K. D. Klika, I. Ricarte, M. T. S. Trevisan, M. G. de Vasconcelos Silva and R. W. Owen, *Tetrahedron: Asymmetry*, 2015, **26**, 247.
48. H. Saitō, I. Ando and A. Ramamoorthy, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **57**, 181.
49. P. W. Kuchel, B. E. Chapman, N. Mueller, W. A. Bubbb, D. J. Philp and A. M. Torres, *J. Magn. Reson.*, 2006, **180**, 256.
50. G. Kummerlowe, E. F. McCord, S. F. Cheatham, S. Niss, R. W. Schnell and B. Luy, *Chem. - Eur. J.*, 2010, **16**, 7087.
51. J. Courtieu, J. P. Bayle and B. M. Fung, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1994, **26**, 141.

52. A. Grishaev, L. S. Yao, J. F. Ying, A. Pardi and A. Bax, *J. Am. Chem. Soc.*, 2009, **131**, 9490.
53. N. Nath, E. J. d'Auvergne and C. Griesinger, *Angew. Chem., Int. Ed.*, 2015, **54**, 12706.
54. P. Trigo-Mouriño, A. Navarro-Vázquez, J. Ying, R. R. Gil and A. Bax, *Angew. Chem., Int. Ed.*, 2011, **50**, 1.
55. S. Patchkovskii, J. Autschbach and T. Ziegler, *J. Chem. Phys.*, 2001, **115**, 26.
56. S. Patchkovskii and T. Ziegler, *J. Phys. Chem.*, 2002, **106**, 1088.
57. J. A. Dale and H. S. Mosher, *J. Am. Chem. Soc.*, 1973, **95**, 512.
58. G. R. Sullivan, J. A. Dale and H. S. Mosher, *J. Org. Chem.*, 1973, **38**, 2143.
59. I. Ohtani, T. Kusumi, Y. Kashman and H. Kakisawa, *J. Am. Chem. Soc.*, 1991, **113**, 4092.
60. F. Freire, J. M. Seco, E. Quiñoá and R. Riguera, *J. Org. Chem.*, 2005, **70**, 3778.
61. J. M. Seco, E. Quiñoá and R. Riguera, *Tetrahedron: Asymmetry*, 2001, **12**, 2915.
62. J. M. Seco, S. Latypov, E. Quiñoá and R. Riguera, *Tetrahedron Lett.*, 1994, **35**, 2921.
63. J. M. Seco, E. Quiñoá and R. Riguera, *Chem. Rev.*, 2004, **104**, 17.
64. J. M. Seco, E. Quiñoá and R. Riguera, *Chem. Rev.*, 2012, **112**, 4603.

CHAPTER 3

Nuclear Magnetic Resonance Experiments Applicable to the Elucidation and Characterization of Nitrogenous Natural Products: ^1H - ^{15}N Heteronuclear Shift Correlation Methods

JOSEP SAURÍ,^a ANTONY J. WILLIAMS^b AND GARY E. MARTIN^{*a}

^a Merck & Co., Inc., Process & Analytical Chemistry, NMR Structure Elucidation, Rahway, NJ 07065, USA; ^b ChemConnector Inc., Wake Forest, NC 27587, USA

*Email: gary.martin2@merck.com

3.1 ^1H - ^{15}N Direct and Long-range Heteronuclear Shift Correlation

While direct observation of ^{15}N is notoriously challenging, the same is not true for the acquisition of ^1H - ^{15}N heteronuclear shift correlation data using HSQC, heteronuclear multiple bond correlation (HMBC), IMPEACH-MBC,¹ constant time inverse-detection gradient accordion (CIGAR)-HMBC,²

Modern NMR Approaches to the Structure Elucidation of Natural Products:

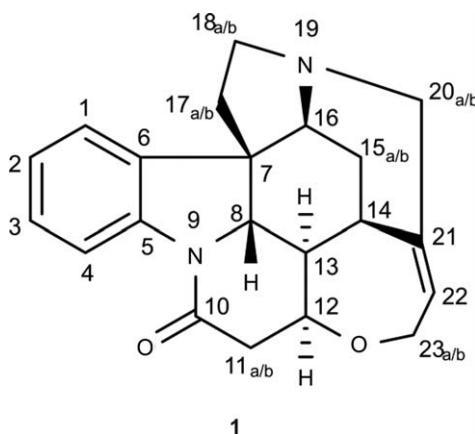
Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

LR-HSQMBC,³ and related experiments. Indeed, one of the current authors has demonstrated the feasibility of acquiring ^1H - ^{15}N HMBC data for a sample of <50 μg of strychnine (**1**) over a weekend using a 1.7 mm MicroCryoProbeTM at a proton observation frequency of 600 MHz.⁴ Parella and coworkers⁵ have also demonstrated the simultaneous acquisition of ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra. The application of long-range ^1H - ^{15}N heteronuclear shift correlation in alkaloid chemistry was the subject of a comprehensive review in 2015,⁶ as well as in a 2007 review.⁷ Although this chapter highlights the application of ^{15}N to alkaloids and peptides as model compounds, the experiments discussed in the chapter are directly applicable to any category of nitrogen-containing natural products. As an additional resource, the interested reader is referred to the series of reviews dealing with long-range ^1H - ^{15}N heteronuclear shift correlation that cover the period from approximately 1995 to the present.⁶⁻¹² There is also a wealth of ^{15}N chemical shift information contained in a number of chapters and monographs devoted to ^{15}N nuclear magnetic resonance (NMR) spectroscopy that have appeared.¹³⁻²²



3.2 ^{15}N Chemical Shift Referencing

Chemical shift referencing of ^{15}N spectra presents a curious conundrum. There are presently two conflicting chemical shift referencing schemes in usage and investigators should be aware of both. In addition, when ^{15}N chemical shift data are reported, it is important to note the chemical shift reference employed. Older ^{15}N data in the literature have generally, but not always, been referenced to nitromethane, which was assigned a chemical shift of 0 ppm. As with most chemical shift referencing conventions, resonances downfield of nitromethane were assigned a positive (+) chemical shift, while those upfield were assigned a negative (−) chemical shift. Since nitromethane resonates quite far downfield of most nitrogens of interest to natural products chemists, virtually all natural product nitrogen resonances

Table 3.1 ^{15}N chemical shift referencing schemes.^a

	^{15}N chemical shift reference				
	Liq. NH_3	CH_3NO_2	$^{15}\text{NH}_4\text{NO}_3$	$\text{NH}_4^{15}\text{NO}_3$	$^{15}\text{NH}_4\text{Cl}$
Liq. NH_3	0.0	379.5	−19.9	383.5	−26.6
CH_3NO_2	379.5	0.0	359.6	4.0	352.9
$^{15}\text{NH}_4\text{NO}_3$	19.9	359.6	0.0	363.6	−6.7
$\text{NH}_4^{15}\text{NO}_3$	383.4	−4.0	363.6	0.0	356.9
$^{15}\text{NH}_4\text{Cl}$	26.6	352.9	6.7	356.9	0.0

^aThe chemical shift referencing schemes shown in this table are the ones that have been most commonly reported, although others have been reported, including nitric acid, formamide, and potassium nitrate.

would have negative chemical shifts. A number of years ago, the nitromethane chemical shift referencing sign convention was reversed as a result of an IUPAC committee, and nitrogen resonances upfield of nitromethane are now assigned positive (+) chemical shifts. Hence, for nitromethane-referenced ^{15}N chemical shifts, it is important to know when the data were reported and which sign convention was in usage.

Relative to nitromethane, liquid ammonia, NH_3 , resonates quite far upfield at +379.5 ppm. In recent years, chemical shift referencing to liquid ammonia has become prevalent, with an explosion in the volume of protein ^{15}N NMR data being reported. Following the chemical shift convention associated with the use of tetramethylsilane (SiMe_4), which is used for ^1H and ^{13}C chemical shift referencing, protein ^{15}N NMR data are usually assigned relative to liquid ammonia, with the NH_3 resonance assigned a chemical shift of 0 ppm. On this scale, the authors have seen the ^{15}N chemical shift of nitromethane variously reported as +379.5, +380.2, and +381.7 ppm. We will note here that all of the ^{15}N data reported in this chapter have been referenced relative to liquid ammonia.

If the referencing scheme described above was not sufficiently confusing, several other ^{15}N chemical shift referencing conventions have also been reported over the years. The more prevalent referencing schemes of which the authors are aware are summarized in Table 3.1, with conversion factors that allow readers to convert between them.

3.3 The Range of ^{15}N Chemical Shifts

The range across which ^{15}N chemical shifts are observed is quite broad, encompassing approximately 900 ppm. For most commonly encountered nitrogen-containing natural products, however, the practical range of ^{15}N chemical shifts is somewhat more compact, yet still broad in terms of excitation requirements, particularly since ^{15}N is a low γ nuclide. Aliphatic ^{15}N resonances can range upfield to ~20 ppm. The downfield limit for most sp^2 nitrogen resonances is ~350 ppm, with a few exceptions. The exceptions are species such as methoximes ($\text{C}=\text{N}-\text{OCH}_3$), which can resonate as far

downfield as ~ 400 – 410 ppm. The sp nitrogens of nitriles generally resonate in the range of ~ 350 – 370 ppm.

3.4 ^{15}N Pulse Widths

While the range of typically encountered ^{15}N chemical shifts for natural products is approximately 300 ppm, at an observation frequency of ~ 60 MHz on a spectrometer operating at 600 MHz for ^1H observation, it can still be challenging with many older NMR probes, and some cryogenic NMR probes, to deliver pulses short enough to effectively excite this bandwidth. Generally, newer generations of gradient-inverse and cryogenic NMR probes will be capable of shorter ^{15}N 90° pulses than were available with older hardware. In a 2010 study, one of the current authors reported results that could be obtained with a new design for Bruker TXI 1.7 mm Micro-CryoProbesTM.²³ The excitation bandwidths that could be obtained with various 90° pulse lengths are shown in Figure 3.1. Using a sample of ^{15}N -enriched urea, with a $25\ \mu\text{s}$ 90° ^{15}N pulse, it was possible to excite a bandwidth of ± 185 ppm relative to the transmitter with 95% peak intensity.

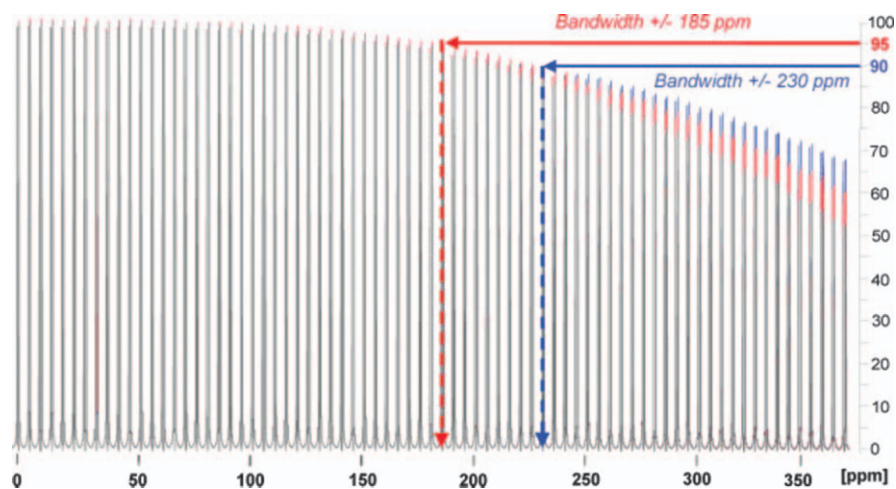


Figure 3.1 ^{15}N excitation profiles achieved with varied 90° ^{15}N pulse widths of $25\ \mu\text{s}$ (blue), $22\ \mu\text{s}$ (red), and $20\ \mu\text{s}$ (black).²³ A sample of ^{15}N -enriched urea was employed and the spectra are 1 transient each in order to eliminate any relaxation effects. Pulses were moved off resonance in successive 5 ppm steps. When a $25\ \mu\text{s}$ pulse was employed, 95% intensity was achieved at ± 185 ppm from the transmitter; 90% intensity was achieved ± 230 ppm relative to the transmitter. A ^{15}N spectral width of 600 ppm, which will include the nitrogen resonances from virtually any natural product, was achieved with $\sim 70\%$ intensity using a $22\ \mu\text{s}$ 90° ^{15}N pulse. (Reproduced from G. E. Martin, B. D. Hilton, D. Moskau, N. Freytag, K. Kessler and K. Colson, *Magn. Reson. Chem.*, 2010, **48**, 935–937 with permission from John Wiley and Sons. Copyright © 2010 John Wiley & Sons, Ltd.)

Peak intensity of 90% was achieved with a bandwidth of ± 230 ppm. With a $22\ \mu\text{s}$ 90° pulse, an excitation bandwidth of 600 ppm was achieved with 70% intensity, which will encompass virtually any ^{15}N resonance ever likely to be encountered in a natural product.

The impact and importance of short ^{15}N 90° pulses is illustrated in Figure 3.2. Using methyl orange (2) as a test compound, a pair of ^{15}N HMBC

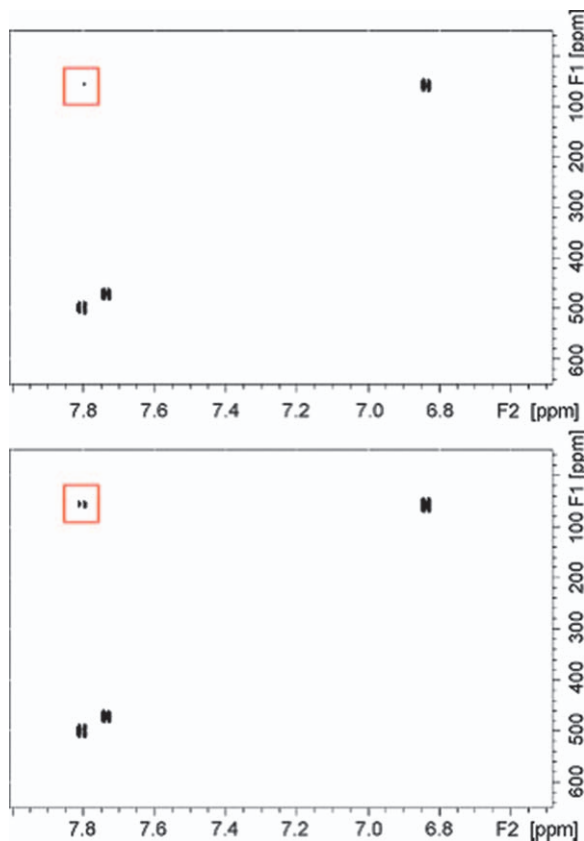
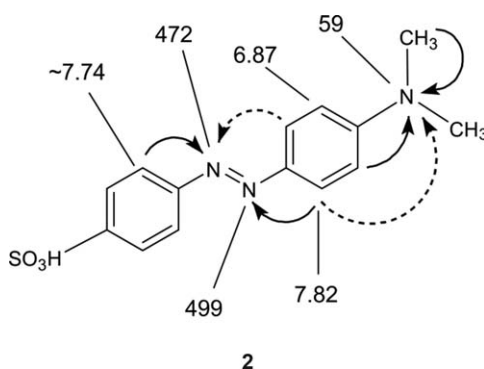


Figure 3.2 ^1H - ^{15}N HMBC spectra of methyl orange (2) acquired at 600 MHz using a 1.7 mm TXI triple-resonance MicroCryoProbeTM. The spectrum shown in the top panel was acquired with the ^{15}N 90° pulse set to $35\ \mu\text{s}$. Data shown in the bottom panel were acquired with the ^{15}N 90° pulse set to $25\ \mu\text{s}$. Both spectra employed adiabatic 180° pulses, and both spectra were plotted with identical threshold levels. The weak correlations boxed in red correspond to the $^4J_{\text{NH}}$ correlation from the proton resonating at 7.82 ppm with the N,N -dimethyl aliphatic nitrogen resonating at 59 ppm. The beneficial impact of shortening the duration of the 90° pulse in the spectrum shown in the bottom panel is obvious.²³ (Reproduced from G. E. Martin, B. D. Hilton, D. Moskau, N. Freytag, K. Kessler and K. Colson, *Magn. Reson. Chem.*, 2010, **48**, 935–937 with permission from John Wiley and Sons. Copyright © 2010 John Wiley & Sons, Ltd.)

spectra were recorded with the ^{15}N 90° pulse set to first 35 μs and then 25 μs . A ^{15}N 90° pulse width in the range of 35–40 μs is, in the authors' experience, typical of older triple-resonance cryoprobes manufactured during the period from about 2000 to 2005. Probes from that era were designed predominantly for triple-resonance NMR experiments on peptides and proteins, where the range of amide ^{15}N resonance chemical shifts is relatively narrow. In contrast, methyl orange (2) was chosen as a model compound that might be considered an “extreme” example since the aliphatic *N,N*-dimethyl nitrogen resonates at 59 ppm, while one of the diazo nitrogens resonates at 499 ppm. These chemical shifts encompass a chemical shift range that is broader than that likely to be encountered in virtually any natural product.



Ideally, the duration of the ^{15}N 90° pulse should be kept as short as possible, with the probe hardware being used to acquire the data. For older probes with longer ^{15}N 90° pulses (*e.g.* >40 μs), it may be advisable to acquire a pair of long-range ^1H – ^{15}N spectra if the molecule being characterized is expected to have ^{15}N resonances with significantly different ^{15}N shifts.

3.5 ^{15}N Chemical Shift Prediction

The process of identifying an unknown can be accelerated using a database of chemical structures and associated spectral parameters, specifically experimental chemical shifts and coupling constants. This is especially true when the data are utilized to produce NMR prediction algorithms, and these are now available for a number of nuclei, most commonly ^{13}C and ^1H , but also ^{31}P and ^{19}F . Electronic content databases of ^{15}N data are available from a number of sources.^{24–26} The authors have direct experience with the ACD/NNMR Predictor program (ACD/NNMR Predictor, Advanced Chemistry Development, Inc., Toronto, ON, Canada) as technical manager for the software for a number of years (AJW) and as a laboratory user (GEM).

The latest form of the ACD/NNMR version 2015 content database contains >10 000 chemical structures associated with >23 700 ^{15}N chemical shifts and >5000 coupling constants.²⁷ Data continue to be extracted from the literature on an ongoing basis and are checked according to a number of stringent criteria. An individual compound record includes the chemical structure, original literature reference, one or more ^{15}N chemical shifts, and, where available, associated coupling constants. The latter data are, unfortunately, only infrequently measured for ^{15}N . These data can be searched by different parameters including structure, substructure, and chemical shift, and this database provides the foundation data for the derivation of prediction algorithms for the ^{15}N nucleus.

NMR prediction offers the possibility of structure verification based on chemical shifts and has been recommended to help optimize experimental acquisition parameters.^{6,7,10,12} With ACD/NNMR Predictor, details of the calculation are provided, as well as the possibility of creating a user database for both reference purposes and to facilitate and increase the accuracy of future predictions.

The application of NNMR Predictor in order to calculate chemical shifts for staurosporine (**3**) is reviewed below (as shown as a comparison of experimental *versus* predicted shifts in Table 3.2), as this provides an opportunity to examine the interface elements of the software (see Figure 3.3). In the results window, the predicted values are listed in the “Value” column, which includes an estimated error. When a number of representative fragments are utilized for the prediction, a calculation error is included. When there are no fragments that can be used, then a value is estimated using a lookup table.

The ability to review the details of how a predicted spectrum is generated is a very useful aspect of the software program. If a particular structure is included in the database, then it will be retrieved and displayed. However, when a prediction is performed for a chemical that is absent from the database, then the different structures used to produce the predicted spectrum are listed in a “Calculation Protocol” window. This allows the environments of the nitrogen atoms for which the chemical shifts are calculated to be compared with the environments and chemical shifts of the related structures used for prediction. The analysis of the spectrum-structural information presented in the calculation protocol therefore allows the chemist to validate the reliability of the NMR prediction.

Table 3.2 Experimental and predicted ^{15}N chemical shifts for the alkaloid staurosporine (**3**).

	Experimental	Predicted
N-13	137.5	136.0
N-6	111.1	117.4
N-12	145.5	143.0
NHCH ₃	20.3	67.8

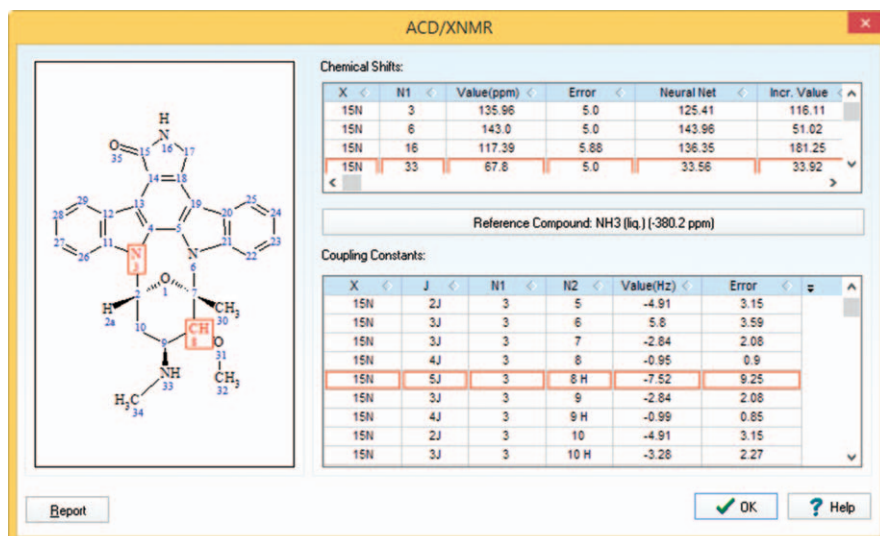
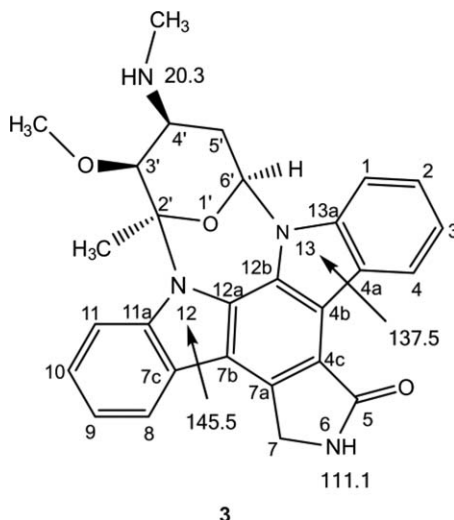


Figure 3.3 The results window associated with the prediction of the ^{15}N shifts for staurosporine (3).²⁸ Hovering over a particular nitrogen atom highlights the associated shift in the table. Estimated coupling constants are also produced for comparison.



For example, for the chemical shift at 111.1 ppm in staurosporine,²⁸ shown above (and discussed in further detail later in this chapter), the calculation protocol is as shown in Figure 3.4. This shows that 16 hits were used covering a range from ~107 to 134 ppm with a calculated value of 117.4 ± 5.9 ppm. The histogram indicates the distribution of hits with chemical shifts assigned to the nitrogen nuclei used for prediction. In the

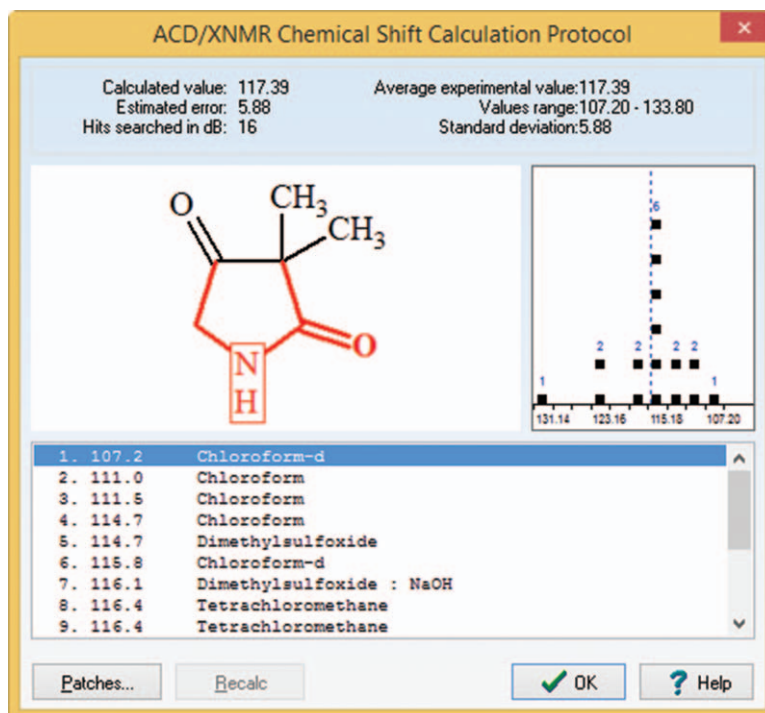


Figure 3.4 The calculation protocol for the amide NH in the structure of staurosporine (3).²⁸ Because of the paucity of available data for ^{15}N chemical shifts, it is generally the situation that data obtained in a variety of solvents are used.

left panel of Figure 3.4, an example structure from the set of hits is shown. The fragment centered on the nitrogen atom of interest is colored red.

The calculation protocol associated with a particular atom can occasionally display an error message, indicating an issue with the prediction. Examples can include when there is no structure found containing a nitrogen-centered fragment suitable for chemical shift calculation for a given atom, a limitation generally resulting from the small size of the database, but more specifically the lack of structural diversity in the training dataset. In such situations, a shift calculation is performed using an “incremental approach”, whereby an alternative algorithm is used to give a “best-guess” estimate of the shifts when there are no similar structures or structural fragments available to inform the normal algorithm. It should be noted that there are additional *alternative* prediction algorithm results also listed in the table under the columns of neural net and incremental value, and these values are closer to the measured experimental value in this case. The neural networks function as a “black box”; they do not provide any access to the details associated with how the chemical shift for a given nitrogen atom is calculated. However, neural networks are significantly

faster for ^{15}N NMR chemical shift calculation, which makes them more effective in the prediction of assignments for large files of chemical structures. A rational combination of different approaches—an incremental approach, an approach based on a fragment library, and neural network methods—allows the user to examine the predicted *versus* the experimental shifts in more detail.

The ^{15}N database is small relative to the available databases of ^1H and ^{13}C spectra: 23 702 *versus* 246 431 and 227 404 structures, respectively.²⁷ While it is possible to perform “solvent-dependent predictions” for both ^1H and ^{13}C nuclei, this is not feasible for ^{15}N , as the dataset is too small. Data from a variety of solvents are therefore utilized to provide a prediction, but ^{15}N shifts can be significantly dependent on the solvent, and this can lead to rather a large error in a predicted value. Solvents, for example, would influence tautomeric forms, and this can influence chemical shifts as detailed below for the donnazole compounds (*vide infra*). Since there are generally only a few nitrogen atoms in any molecule being predicted, solvent-dependent estimates are rarely an issue for the user to get shift values for setting up an experiment or assisting in the validation of a structural hypothesis.

3.5.1 Validating ^{15}N Chemical Shift Prediction

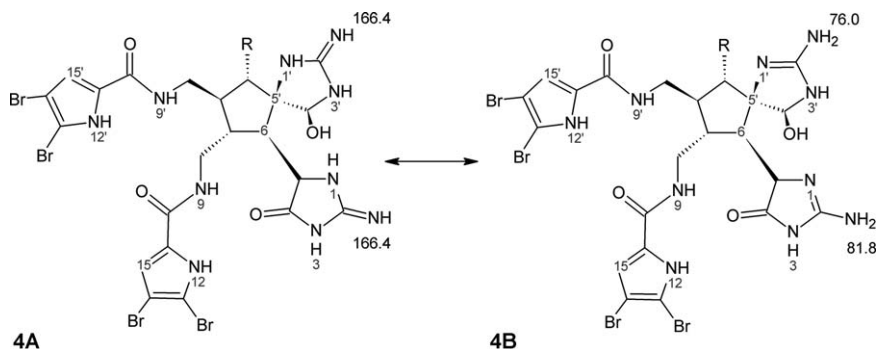
To validate the performance of the ^{15}N chemical shift prediction algorithms, a statistical analysis has previously been performed using an earlier version of the software product (ACD/NNMR v8.08).⁶ Using a classical leave-one-out approach, the ^{15}N shifts associated with ~8300 individual chemical structures were predicted. The resulting analysis gave a correlation coefficient of $R^2 = 0.97$ over 21 244 points. The standard error of about ± 15 ppm is quite reasonable and affords a basis for setting the F_1 windows to acquire long-range ^1H – ^{15}N data. Other comparisons have been performed by the authors using smaller subsets of data. When a particular review or book chapter is written, we commonly perform an analysis of the specific set of chemicals discussed within that manuscript for the purposes of self-consistency and to relate the prediction algorithms to the compounds discussed in detail.

In an earlier work,⁶ the ACD/NNMR 2012 predictor was applied to a selection of 45 nitrogen-containing structures with experimentally measured ^{15}N NMR chemical shifts described in the review. Of the compounds chosen, 11 were already present in the NNMR database, and these were removed from the set so that all compounds being compared were absent from the ACD/NNMR database—a total of 34 molecules. A total of 117 chemical shifts were predicted and a plot of the values against experimental shifts indicated some obvious outliers. A total of 13 predicted chemical shifts with a difference of >50 ppm from the measured shifts were identified, and some suggestions were made for the discrepancies observed.

For example, a comparison of predicted *versus* experimental shifts for donnazole A and B (**4A** and **4B**, respectively)²⁹ indicated that the structure representation shown is incorrect.



Other examples of large differences could be explained in terms of folding of the data, the simple inclusion of a decimal point, and transcription errors. However, lack of direct access to the data does not allow confirmation of any of these hypotheses, but in our estimation, the reported data are suspicious. While we would never suggest that predictions should be treated as more important than experimental data, should large differences be observed, we would suggest that careful analysis should be undertaken to confirm the observations. Certainly, predictions will fail for particularly exotic structures where there are no representative molecular fragments in the database or where the connectivities simply introduce a

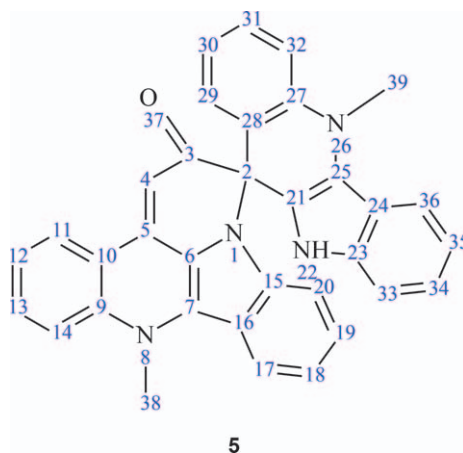


Scheme 3.1 Impact of tautomerization of the exo C=NH group (**4A**) *versus* C-NH₂ (**4B**) on the calculated ¹⁵N chemical shifts of donnazole forms A and B.

Table 3.3 Predicted chemical shifts for cryptospirolepine (5) as produced by ACD/NNMR and DFT calculations at the b3lyp/6-311+G(d,p)/-belyp/6-31G(d) level of theory.³²

Position	Experimental ¹⁵ N shifts	Calculated ¹⁵ N chemical shifts			
		HOSE	Neural net	Incr.	DFT
1	163.7	169.6	147.5	194.9	130.1
8	135.1	114.7	101.6	93.4	89.7
22	144.0	124.2	116.4	54.8	95.7
26	95.0	131.4	114.4	128.4	61.4

molecular architecture producing dramatically different shifts from predictions. One such example is the complex, spiroonacyclic alkaloid cryptospirolepine (5).^{30–32} The predicted chemical shifts obtained using each of the methods available in ACD/NNMR Predictor, including those produced using density functional theory (DFT) calculations, provide predictions outside of the expected range of ± 15 ppm cited earlier, and are dramatically different for N-26. However, based on the application of the newly reported 1,1- and 1,*n*-HD-ADEQUATE experiments, combined with ¹H–¹⁵N HMBC and 2 Hz-optimized LR-HSQMBC experiments (see Section 3.8.4.4 later in this chapter), there is now complete confidence in the accuracy of the revised structure of cryptospirolepine³² (Table 3.3).



3.5.2 Setting F₁ Spectral Windows

As mentioned earlier, ¹⁵N shift prediction algorithms can provide a useful starting point for estimating shifts in order to set up spectral acquisition parameters, which is especially useful for less experienced investigators with little knowledge of the ¹⁵N nuclide. It is very prudent to consider shift

calculations for plausible structural motifs prior to setting F_1 spectral windows for the acquisition of long-range ^1H - ^{15}N correlation spectra. This is especially true when working on a spectrometer equipped with a probe that does not have short ^{15}N 90° pulses to excite a broad spectral window.

As mentioned earlier, we suggest that calculating the expected ^{15}N chemical shifts of a potential molecular structure and then adding a buffer of ± 15 – 20 ppm to the range of potential ^{15}N chemical shifts is a reasonable way to establish a starting F_1 spectral window in most cases. The ideal, however, is improved spectrometer and probe efficiency to irradiate the entire ^{15}N spectral width, in which case such calculations would become irrelevant; great advances continue in this area.

3.5.3 Structure Verification using a ^{15}N Content Database

As discussed above, the databases underlying the prediction tools are built on training data extracted from the literature. In many cases, only one representative dataset for a particular compound may exist, but it is also possible that multisolvent data are available. An existing database of ^{15}N data allows us to search these data and, armed with substructural moieties assembled from ^{13}C and ^1H NMR data, and with ^{15}N chemical shifts in hand, it is possible to perform a number of flexible searches to identify potential classes of compounds. Various options can be set in the search interface (see Figure 3.5). These include the looseness factor, the minimum number of query shifts, and the hit quality index (HQI).

The “looseness factor” is the deviation allowed around each chemical shift during the search. The “minimum number of query shifts” to match is the number that should be found during a search. For example, if, in the “Enter Query Shifts” field, a list of shifts is entered and the minimum to match is 2, then the program will find all of the records containing at least two chemical shifts from this list. The option to sort the results by the HQI based on minimal distance, orders the results such that the best matches are listed first (*i.e.* with the highest HQI). The ability to intersect search lists allows substructure searches to be performed followed by shift searching, to prune result sets, to attempt structure identification. This process will become much more valuable as the dataset increases in size.

3.5.4 Enhancing ^{15}N Chemical Shift Prediction with a “User-trained” Database

Novel compounds, classes, or structural fragments that may be under investigation in a particular laboratory may not yet be described in the literature. Consequently, they will not be represented in the predictor training databases. It would clearly be of value for a scientist investigating new compounds to capture and catalog their own structures and assignments

Figure 3.5 The chemical shift search interface for querying the database.

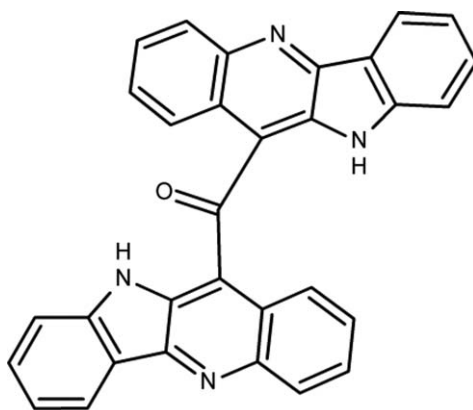
and to use these data to improve the prediction algorithms. All ACD/NNMR predictors allow the chemist to create *user databases* of structures and assignments. As new chemical shifts are assigned to atoms, the program can use these data to improve the predictions. This user database training can have a dramatic impact on the ability of an organization to predict NMR spectra for a diverse array of compounds containing structural moieties that have been represented at a fairly minimal level of one to two structures in the training database.

3.6 Computer-assisted Structure Elucidation—The Impact of ^{15}N Data

There have been a number of evaluations of the ability of computer-assisted structure elucidation (CASE) programs (see Volume 1, Chapter 9 of this series for an overview of CASE) to generate the correct structure with and without the availability of long-range ^1H - ^{15}N heteronuclear 2D NMR data.^{33–35}

Even incomplete long-range ^1H - ^{15}N heteronuclear chemical shift data can have a very beneficial impact on the performance of CASE programs.

An example is found in the application of the Structure Elucidator CASE program to the dimeric indoloquinoline alkaloid cryptomisine (6).



6

Without any ^1H - ^{15}N chemical shift correlation data, the program ran for 210 hours, generating >75 million structures, of which >22 000 remained after filtration and removal of duplicates. In contrast, when only simple ^1H - ^{15}N HMQC direct correlation data were added to the input data, only five structures were generated in ~ 1 minute, with the correct structure among them.³⁶

Cryptospirolepine (5) is a far more complex example. When an attempt was made to solve the structure using CASE methods with the 1993 data (HSQC and ^1H - ^{13}C HMBC), the Structure Elucidator program ran for >420 hours without ever generating the correct structure. It is unlikely that augmenting the input data with ^1H - ^{15}N HMBC or other long-range proton-nitrogen data as the sole additional input would have had much, if any, impact on the output. Including 60 Hz 1,1-HD-ADEQUATE (see Volume 2, Chapter 10) and IDR-HSQC-TOCSY data in the input file allowed the program to generate the correct structure in a 140 second calculation among the 6816 structures generated. Further augmenting the input data with 2 Hz-optimized ^1H - ^{15}N LR-HSQMBC reduced the calculation time to 40 seconds and the number of structures generated by the program to 522. Without the 1,1-HD-ADEQUATE data included in the input file, the program was still unable to generate the correct structure, even with the ^{15}N data incorporated in the input file³⁷ (Table 3.4)

These results, and those with cryptomisine (6), clearly underscore the value of including ^1H - ^{15}N heteronuclear chemical shift correlation data in the investigation of unknown nitrogen-containing structures. The impact of these data on multiple CASE program applications in terms of the reduction of the numbers of structural possibilities generated as output, very dramatically illustrates the benefit of having these data available.^{33,36-40}

Table 3.4 Summary of the results obtained by incorporating various combinations of the available 2D NMR as inputs for the ACD Structure Elucidator CASE program. Generation times and the numbers of structures generated are shown in the columns to the right. The top row of the table shows the results with the NMR data available in 1993 (COSY, HMQC, and 8 Hz HMBC). The program failed to generate the revised structure by the time the calculation was terminated at 420 hours.³⁷

1,1-HD-ADEQUATE	^1H - ^{13}C HMBC	IDR-HSQC-TOCSY	^1H - ^{15}N LR-HSQMBC	GENERATION	
60 Hz	8 Hz	15 ms	2 Hz	Time (s)	# structures
	+			>420 hours	>10 400
+	+	+		140	6816
+	+	+	+	40	522

3.7 $^1J_{\text{NH}}$ and $^nJ_{\text{NH}}$ Coupling Constants

Unlike ^1H - ^{13}C heteronuclear coupling constants that are fairly regular and predictable, ^1H - ^{15}N heteronuclear coupling constants tend to be more variable. Whereas $^3J_{\text{CH}}$ coupling constants tend to be larger than $^2J_{\text{CH}}$ and $^4J_{\text{CH}}$ coupling constants in most cases, rendering $^3J_{\text{CH}}$ correlations most prevalent in ^1H - ^{13}C HMBC spectra, the same is not true of long-range ^1H - ^{15}N heteronuclear shift correlation spectra. In many instances, $^2J_{\text{NH}}$ correlations are the most prevalent responses in ^1H - ^{15}N HMBC⁴¹ spectra when the experiment is optimized in the vicinity of ~ 5 Hz.

In part, the variability of ^1H - ^{15}N heteronuclear couplings can be attributed to the orientation of the nitrogen lone pair of electrons. Differences in the sizes of coupling constants can be observed as a function of whether a C-H bond vector is synclinal or anticlinal relative to the orientation of the nitrogen lone pair. We addressed some of these questions in early work published during the development of ^1H - ^{15}N HMBC, and the reader is referred there⁴² or to some of the primary literature references on ^{15}N NMR.^{13–22}

3.8 ^1H - ^{15}N Results Obtained Using Various Experiments

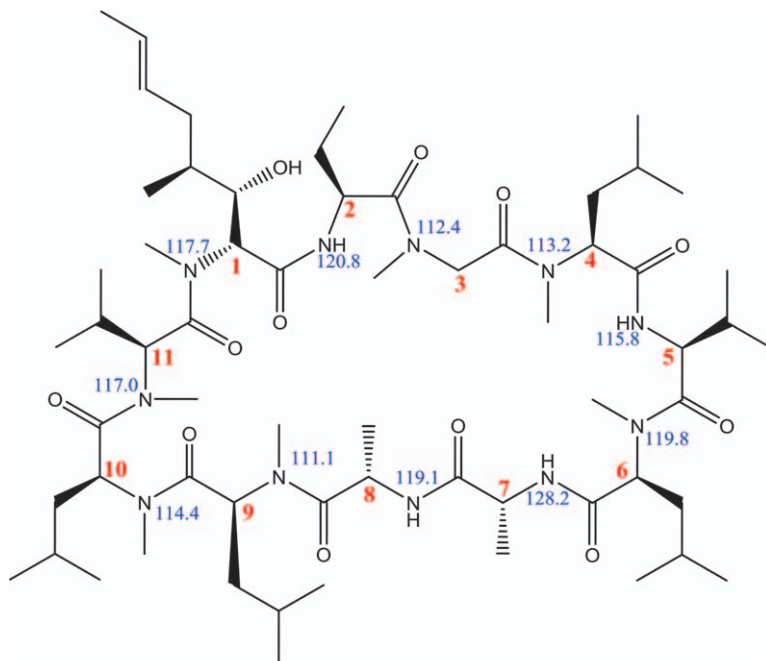
Thus far in this chapter, we have considered many of the attributes of ^1H - ^{15}N heteronuclear chemical shift correlation, and it is now germane to consider some of the experiments that can be used to acquire these data. Although strychnine (**1**) has generally been employed as the model compound for illustrating experiment performance, strychnine unfortunately does not have a protonated nitrogen. Hence, we will employ the cyclic peptide cyclosporine (**7**) as a model compound for direct ^1H - ^{15}N chemical shift correlation experiments. Both strychnine (**1**) and cyclosporine (**7**) will be utilized to illustrate the results that can be obtained with long-range ^1H - ^{15}N correlation experiments. Experiments will be sub-grouped into those utilizing an initial $^1J_{\text{NH}}$ magnetization transfer (see Section 3.8.1) and those employing an initial $^nJ_{\text{NH}}$ magnetization transfer (see Section 3.8.2).

3.8.1 ^1H - ^{15}N Experiments Employing an Initial $^1J_{\text{NH}}$ Magnetization Transfer

Logical subdivision of experiments applicable to ^1H - ^{15}N heteronuclear shift correlation can be based on the initial magnetization transfer. Hence, we will begin by discussing those experiments involving an initial $^1J_{\text{NH}}$ transfer step as opposed to those involving an initial $^nJ_{\text{NH}}$ transfer. The former group includes the HSQC and HSQC-TOCSY experiments, while the latter group (see Section 3.8.2) includes experiments such as HMBC, HSQMBC, LR-HSQMBC, and HSQMBC-TOCSY. Finally, there are also several experiments that can be considered as triple-resonance experiments (see Section 3.8.3). These experiments include the simultaneous acquisition of ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra and the HCNMBC experiment that utilizes an initial $^1J_{\text{CH}}$ transfer followed by a natural abundance $^1J_{\text{CN}}$ or $^nJ_{\text{CN}}$ transfer. While the former two groups of experiments can be performed on a two-channel spectrometer with a conventional inverse-detection probe, the latter group requires a three-channel spectrometer and a triple-resonance probe.

3.8.1.1 ^1H - ^{15}N HSQC of Cyclosporine

Cyclosporine (7), like strychnine, has been a useful model compound for the development of NMR experiments that are applicable to peptides. The structure is shown below with ^{15}N resonance assignments, and for purposes of ^{15}N NMR experiments, the protonated amide nitrogens in the structure make cyclosporine a useful model compound for this chapter.



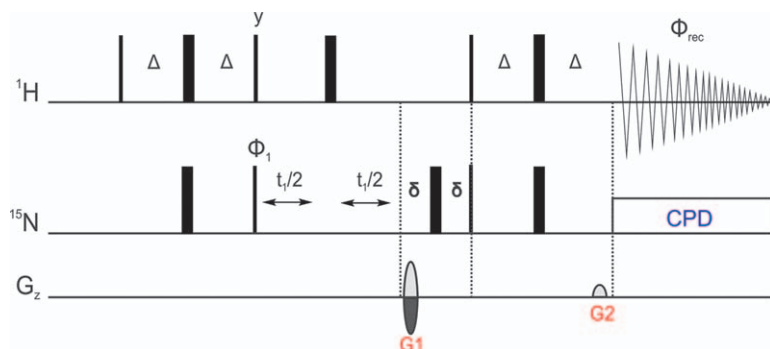


Figure 3.6 The pulse sequence for the ^1H - ^{15}N HSQC experiment is shown. The delays for the $^1J_{\text{NH}}$ coupling constant, Δ , can be optimized for 95 Hz. The gradient ratio needs to be adjusted to reflect the gyromagnetic ratios of ^1H and ^{15}N , but for most users with access to a modern spectrometer, this requirement is transparent to the user.

The ^1H - ^{15}N HSQC pulse sequence is shown in Figure 3.6. Unlike $^1J_{\text{CH}}$ coupling constants that are highly variable, ranging from ~ 125 Hz to as large as ~ 250 Hz, the $^1J_{\text{NH}}$ coupling constants have a relatively narrow range of ~ 90 – 100 Hz. Consequently, the $^1J_{\text{NH}}$ coupling constant can be uniformly optimized at 95 Hz.

Although the total ^{15}N chemical shift range is quite broad, the range of protonated ^{15}N shifts is much narrower, ranging from approximately 25 to 150 ppm, which makes ^1H - ^{15}N HSQC data relatively simple to acquire. Furthermore, since most molecules will have a limited number of nitrogen resonances, there is no need for high digital resolution in the F_1 frequency domain. Typically, the current authors have found it acceptable to digitize the F_1 domain of proton–nitrogen chemical shift correlation spectra of natural products more coarsely, with a data point every ~ 2 ppm for ^1H - ^{15}N HSQC spectra.

The pulse sequence for the ^1H - ^{15}N HSQC spectrum is shown in Figure 3.6. For the relatively narrow range of $^1J_{\text{NH}}$ coupling constants, a compromise optimization of the Δ delay in the pulse sequence for ~ 95 Hz works for virtually any protonated nitrogen-containing natural product one is likely to encounter.

The ^1H - ^{15}N HSQC spectrum of cyclosporine (7) is shown in Figure 3.7. While the range of ^{15}N chemical shifts for protonated nitrogens is relatively narrow, the range of amide ^{15}NH chemical shifts is narrower still. Practically speaking, most protonated amide ^{15}N resonances are observed in the relatively narrow range from ~ 80 to 135 ppm (downfield of liq. $\text{NH}_3 = 0$ ppm). The spectrum shown in Figure 3.7 was acquired with broadband ^{15}N decoupling. It is also possible to acquire ^1H - ^{15}N HSQC data with multiplicity editing, if it were necessary to differentiate NH from NH_2 resonances.

3.8.1.2 ^1H - ^{15}N HSQC-TOCSY of Cyclosporine

Another experiment that is quite feasible with peptide natural products but has received very little attention is the ^1H - ^{15}N HSQC-TOCSY experiment. An

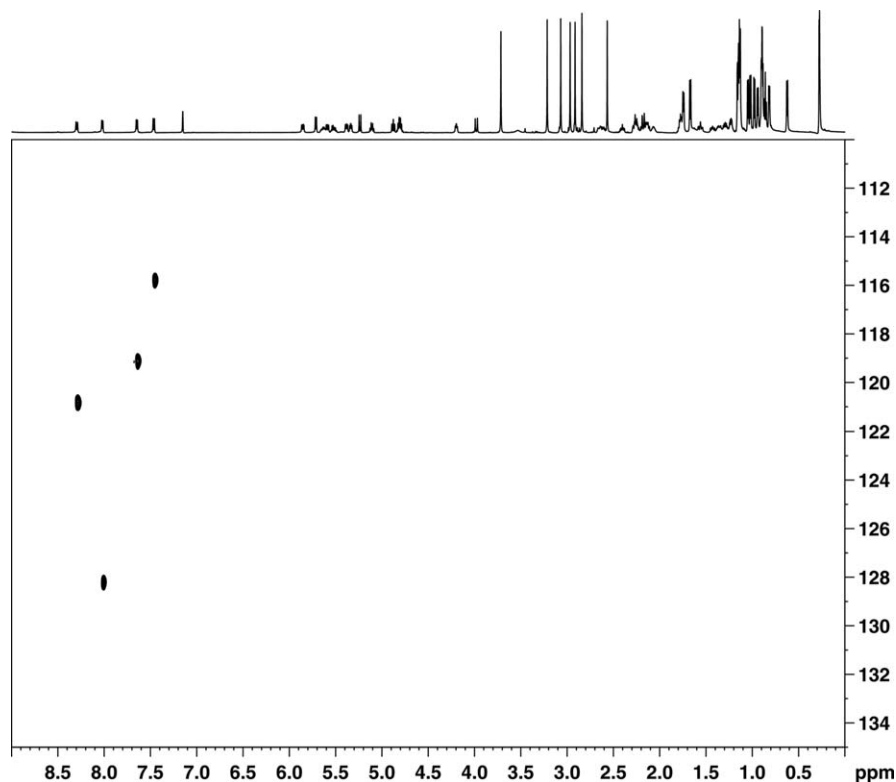


Figure 3.7 ^1H - ^{15}N HSQC spectrum of a 1.5 mg sample of cyclosporine (**7**) dissolved in 35 μL of d_6 -benzene using a 600 MHz 1.7 mm gradient triple-resonance MicroCryoProbeTM. The data were acquired as 1K points in F_2 with eight transients for each of the 96 increments used to digitize the 40 ppm F_1 spectral window, giving an acquisition time of 20 minutes. The data were processed using a $\pi/2$ phase-shifted sine bell to afford the final $2\text{K} \times 1\text{K}$ point spectrum shown.

early application of the ^1H - ^{15}N PEP-HSQC-TOCSY (PEP = preservation of equivalent pathways) experiment to the linear peptide bombesin (**8**) was reported by Williamson, Marquez, and Gerwick.⁴³ The contemporary ^1H - ^{15}N HSQC-TOCSY pulse sequence is shown in Figure 3.8 and the corresponding ^1H - ^{15}N HSQC-TOCSY spectrum of cyclosporine is shown in Figure 3.9. Following the “labeling” of the NH resonance with the ^{15}N chemical shift during the evolution period, proton-proton coherence is propagated by the subsequent TOCSY transfer in a fashion analogous to the ^1H - ^{13}C IDR-HSQC-TOCSY experiment discussed in the Chapter 9 of this volume. Hence, the proton spin systems of the amino acids with 2° NH resonances are sorted in the F_1 frequency domain by the amide ^{15}N chemical shift. Practically speaking, ^1H - ^{15}N HSQC-TOCSY, when the experiment is performed on a spectrometer equipped with a cryogenic NMR probe, is a relatively high-sensitivity experiment, despite the fact that it is very infrequently employed. The correlations shown for cyclosporine (**7**) in

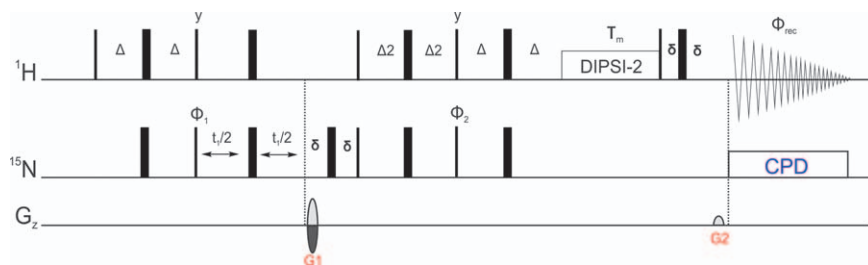


Figure 3.8 Pulse sequence for ^1H - ^{15}N HSQC-TOCSY. The experiment is essentially identical to the ^1H - ^{13}C experiment (see Figure 9.31 in Chapter 9) with the exception of the X-nucleus and the gradient ratio, which must be adjusted for the gyromagnetic ratio (γ_{N}).

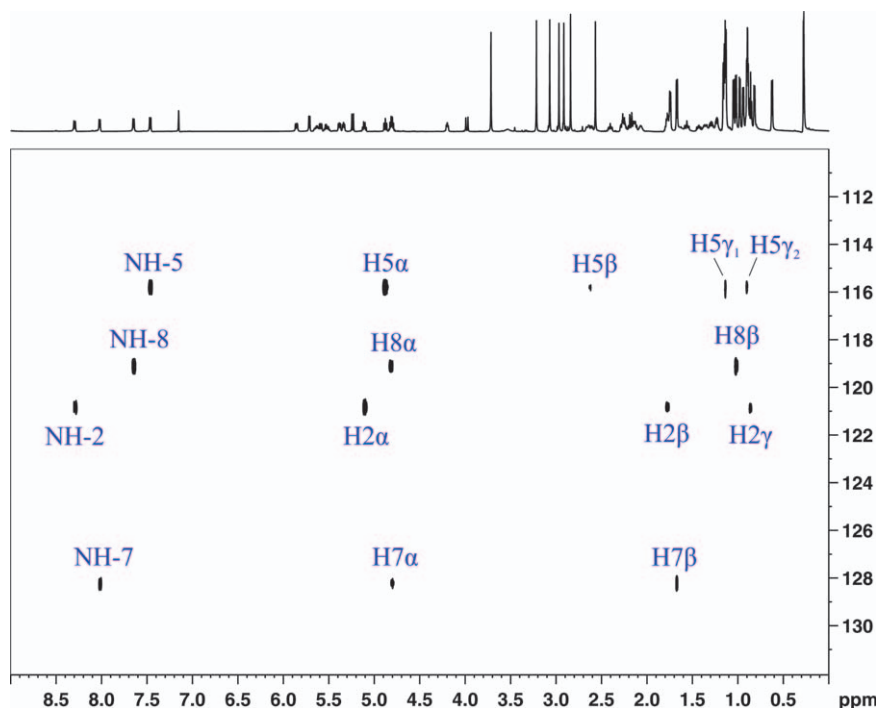
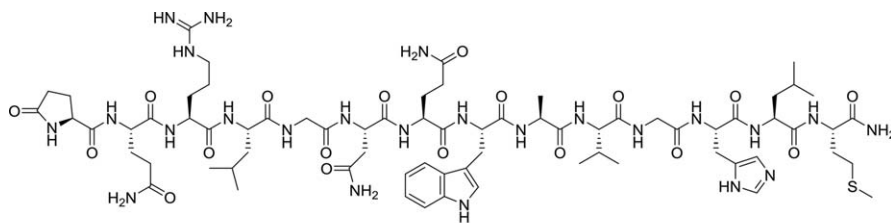


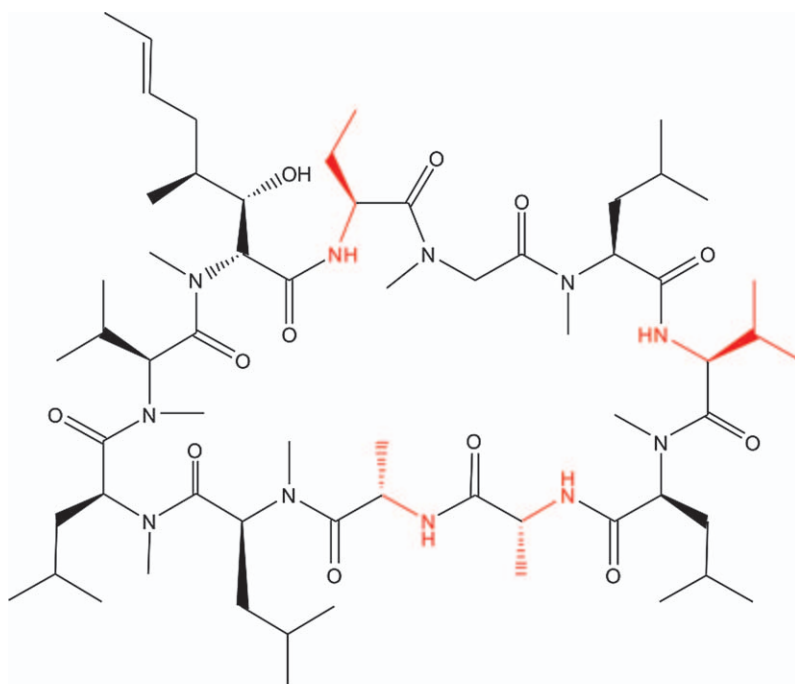
Figure 3.9 ^1H - ^{15}N HSQC-TOCSY spectrum of the cyclic peptide cyclosporine (7). The direct responses can be identified by referring to Figure 3.8, although since all of the nitrogen resonances of interest are amides, the NH resonances will obviously be those furthest downfield at each F_1 frequency. The spectrum was acquired as 2K points in F_2 with 16 transients for each of the 96 increments of the evolution period and a TOCSY mixing time of 40 ms. The data were processed using a $\pi/2$ phase-shifted sine bell to afford the $2\text{K} \times 1\text{K}$ spectrum shown.

Figure 3.9, translate to the identification of the segments of the structure designated by red bonds on structure 9. Since there are only four protonated ^{15}N resonances, the structural information provided by the HSQC-TOCSY

spectrum is more limited than would be the case for many peptide natural products that do not have as many *N*-methyl resonances.



8



9

3.8.2 ^1H - ^{15}N Experiments Employing an Initial $^nJ_{\text{NH}}$ Magnetization Transfer

The previous section of this chapter considered experiments that employ an initial $^1J_{\text{NH}}$ magnetization transfer. This section delves into the assortment of experiments that utilize an initial $^nJ_{\text{NH}}$ magnetization transfer. Long-range ^1H - ^{15}N heteronuclear shift correlation experiments are inherently more challenging due to the lack of uniformity of $^nJ_{\text{NH}}$ coupling constants. As might be expected, research in this area began with the HMBC experiment and progressed from there to the development of experiments intended to deal the variability of $^nJ_{\text{NH}}$ coupling constants such as

IMPEACH-MBC¹ and CIGAR-HMBC,² the much more recently developed LR-HSQMBC³ experiment, and others.

3.8.2.1 ^1H - ^{15}N HMBC

Applications of the ^1H - ^{15}N HMBC experiment date back further than any other ^1H - ^{15}N applications to natural products. The first experimental results were communicated during two scientific meetings in 1993. Non-gradient results, obtained with several indoloquinoline alkaloids, were reported in a poster presentation at a meeting of the American Society of Natural Products by Martin and coworkers in July, 1993.⁴⁴ The first results using gradient methods were reported by Koshino and coworkers in November of that year.⁴⁵ A gap of nearly 2 years followed before the next published reports from the laboratory of one of the current authors (GEM).⁴⁶ The period from 1995 to 2000 saw the publication of >50 reports, which were surveyed in the first review of the topic by Martin and Hadden in 2000.⁸ Subsequent reviews and chapters followed that have been cited earlier in this chapter.

The pulse sequence for the ^1H - ^{15}N HMBC experiment is shown in Figure 3.10. The sequence is shown without a low-pass J -filter. Although it is certainly possible to filter out the direct responses, in the experience of the authors, it is seldom worth doing so. The direct responses will be ~ 95 Hz doublets, providing the equivalent of a ^1H - ^{15}N HSQC spectrum while the HMBC data are being acquired. ^1H - ^{15}N spectra can, of course, be recorded using the same pulse sequence as that for ^1H - ^{13}C long-range correlation, which does have a low-pass J -filter. To retain the direct correlations, one only has to set the $^1J_{\text{NH}}$ delay to a large value (*e.g.* 300 Hz), which will allow the direct correlation doublets to be observed, albeit with slightly attenuated intensity. The only other change that needs to be made to use the ^1H - ^{13}C experiment is to adjust the gradient ratio accordingly, which is generally transparent to the user and controlled by the spectrometer software, when ^{15}N is selected as the X-nucleus.

Optimization of the $^nJ_{\text{NH}}$ delay for ^1H - ^{15}N HMBC experiments is less straightforward than the corresponding ^1H - ^{13}C experiments due to the variability of the $^nJ_{\text{NH}}$ delay, as noted above. Various initial optimizations have been suggested in the published literature, and a reasonable starting point is probably in the range of 5–8 Hz. Although most of the published reports in the literature have utilized the ^1H - ^{15}N HMBC experiment, the variability of the $^nJ_{\text{NH}}$ coupling constant makes it attractive to employ a long-range experiment that utilizes accordion optimization of the long-range delay (see Section 3.8.4.3).

3.8.2.1.1 ^1H - ^{15}N HMBC of Cyclosporine. Cyclosporine (7) has four NH resonances that provide a convenient illustration of the retention of direct correlation NH doublets in a ^1H - ^{15}N HMBC spectrum, acquired using a pulse sequence without a J -filter (see Figure 3.10). The ^1H - ^{15}N direct correlations are readily identified by their characteristic ~ 90 – 95 Hz doublet in F_2 , and there is essentially no reason to exclude the $^1J_{\text{NH}}$ correlations through the use of a J -filter since these correlations provide the equivalent of a ^1H - ^{15}N HSQC spectrum. Correlations observed in the ^1H - ^{15}N HMBC spectrum of cyclosporine (7), shown

in Figure 3.11, are summarized by correlation arrows as shown in **10**. Weak correlations are designated by dashed arrows. Note that all of the correlations observed are either *via* $^2J_{\text{NH}}$ or $^3J_{\text{NH}}$ in the 5 Hz-optimized spectrum shown.

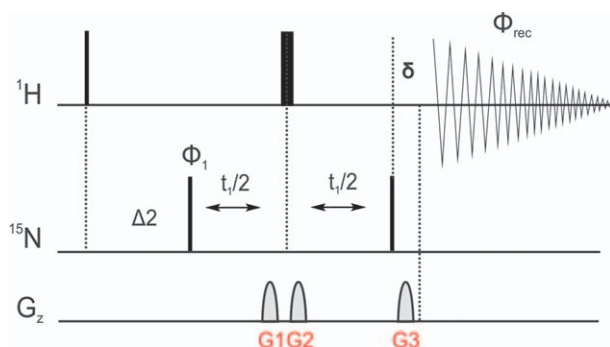
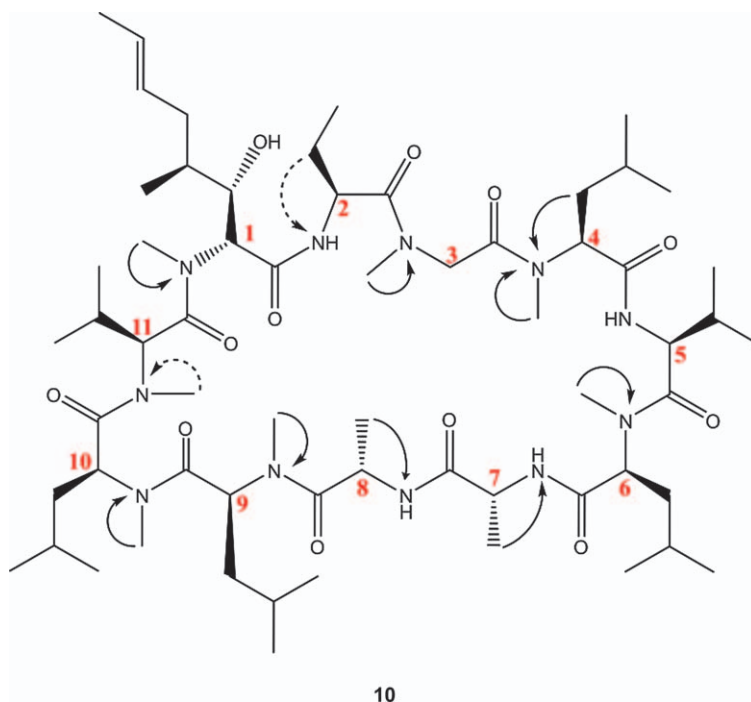


Figure 3.10 Pulse sequence for ^1H - ^{15}N HMBC without a low-pass J -filter to remove direct ^1H - ^{15}N correlations. Typical optimization of the delay, Δ , for the long-range $^nJ_{\text{NH}}$ correlation is usually in the range of 3–8 Hz. As a general guideline, the number of transients accumulated/ t_1 increment when optimized at 3 Hz should probably be two- to four-times that of an experiment optimized at 8 Hz. Because ^{15}N is a low γ nuclide and the range of ^{15}N chemical shifts can be relatively wide, particularly if there are heteroaromatic nitrogens in a molecule, it is preferable to acquire the data using a probe with the shortest ^{15}N 90° pulse possible (see Figure 3.2). Obviously, for sensitivity reasons, acquiring these data using either a LN_2 or helium-cooled cryoprobe (see Volume 1, Chapter 3) is preferable to acquiring the data with a room temperature probe.

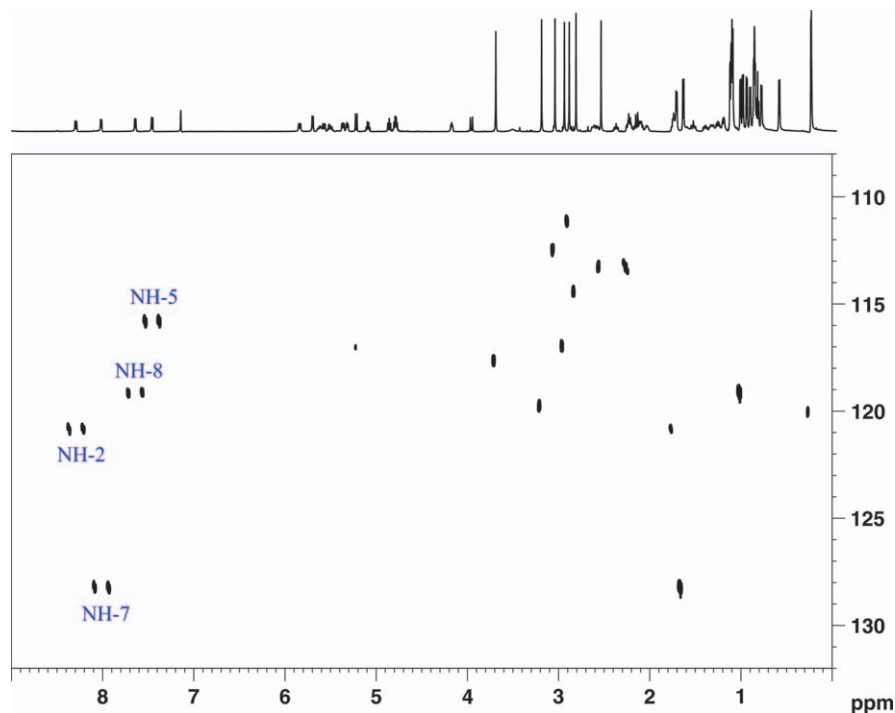


Figure 3.11 ^1H - ^{15}N HMBC spectrum of cyclosporine (7). The spectrum was optimized for 5 Hz and the data were acquired without a low-pass f -filter. The data were acquired as 4K points in F_2 and the F_1 frequency domain was digitized with 96 increments. Twelve transients were acquired for each of the t_1 increments. Data were processed using $\pi/2$ phase-shifted sine bell apodization followed by zero-filling to afford the $4\text{K} \times 1\text{K}$ spectrum shown. The acquisition time was 29 minutes. The balance of the correlations observed were predominantly from the N -methyl groups, as shown by 10.

3.8.2.1.2 ^1H - ^{15}N HMBC of Strychnine. Strychnine (1) was among the first molecules examined by ^1H - ^{15}N HMBC.⁴⁷ The experiments were performed using the standard HMBC experiment employed for ^1H - ^{13}C experiments with the gradient ratios changed to accommodate the gyromagnetic ratio of ^{15}N . The 8 Hz-optimized ^1H - ^{15}N HMBC spectrum is shown in Figure 3.12.

One inherent disadvantage to the HMBC experiment when small, very-long-range correlations are sought is the fact that the data are antiphase following the evolution period and are presented as magnitude mode spectra. Hence, small, weak antiphase correlations have a tendency to self-cancel, thus obliterating them. Such is not the case with other long-range ^1H - ^{15}N experiments discussed later in this chapter. As will be noted from the summary of the DFT-calculated ^1H - ^{15}N coupling constants for strychnine, the majority of the coupling constants are <3 Hz, which makes the availability of experiments such as LR-HSQMBC and HSQMBC-TOCSY particularly valuable for probing ^1H - ^{15}N heteronuclear coupling pathways.

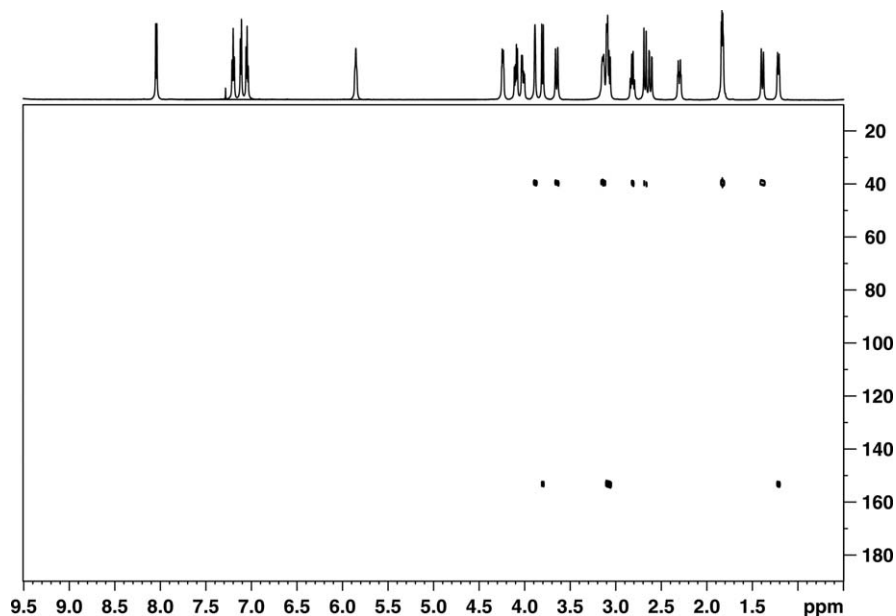


Figure 3.12 ^1H - ^{15}N HMBC spectrum of strychnine (**1**) optimized for 8 Hz. The data were acquired as 4K points with 16 transients for each one of the 96 t_1 increments, using a 4 mg sample of strychnine dissolved in 40 μL of deuteriochloroform and a 600 MHz 1.7 mm triple-resonance Micro-CryoProbeTM over 37 minutes.

Correlations are observed in the 8 Hz-optimized spectrum to the tertiary aliphatic N19 resonance and to the N9 amide nitrogen. All of the correlations observed are either *via* $^2J_{\text{NH}}$ or $^3J_{\text{NH}}$. In the experience of the current authors, $^4J_{\text{NH}}$ correlations are generally not observed in ^1H - ^{15}N HMBC spectra optimized in the range of 5–8 Hz.

In the 8 Hz-optimized ^1H - ^{15}N HMBC spectrum of strychnine (**1**) shown in Figure 3.12, it should also be noted that there are no correlations observed with the N9 amide nitrogen resonance from any of the aromatic protons, including the $^3J_{\text{H4N9}}$ correlation. DFT calculations performed on strychnine estimated the $^3J_{\text{H4N9}}$ and $^4J_{\text{H3N9}}$ coupling constants at -0.77 and -0.6 Hz, respectively. In contrast, in the 3 Hz-optimized ^1H - ^{15}N HMBC spectrum of strychnine shown in Figure 3.13, correlations are observed with N9 from both the H3 and H4 aromatic protons.

3.8.3 ^1H - ^{15}N HSQMBC

One of the inherent problems in acquiring ^1H - ^{15}N HMBC data is the anti-phase nature of the observed correlations, as well as the need for a magnitude calculation of the data. This is because the pulse sequence allows for the evolution of multiple-quantum coherences, which introduce severe

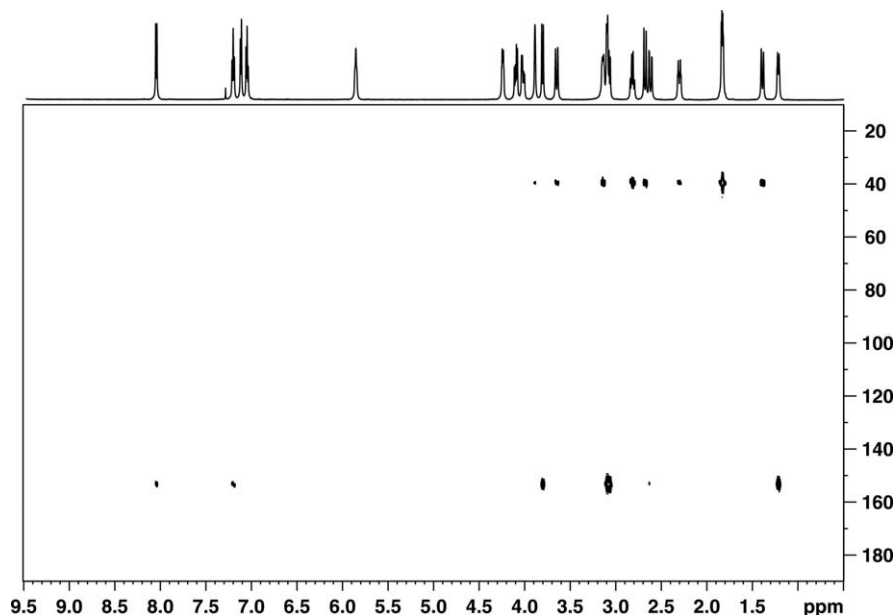


Figure 3.13 ^1H - ^{15}N HMBC spectrum of strychnine (**1**) optimized for 3 Hz with 64 transients accumulated for each t_1 increment, giving an acquisition time of 2 hours and 40 minutes. The correlations observed in the 8 Hz-optimized spectrum shown in Figure 3.12 persist and additional correlations are observed. In particular, correlations from the aromatic H3 and H4 resonances with the indole N9 resonance are observed, which would link that part of the molecule to the amide nitrogen, if this were an unknown structure being elucidated.

distortions into the final lineshapes of the multiplets in both F_2 and F_1 dimensions. In addition, J_{HH} evolves during the variable t_1 evolution period. Owing to these reasons, the HSQMBC experiment was developed and reported⁴⁸ in the late 1990s. The pulse sequence is basically a long-range optimized HSQC sequence and is based on an $^nJ_{\text{XH}}$ transfer mechanism *via* single-quantum coherences. In order to maximize response intensity, the experiment initially lacked a reverse INEPT element, hence magnetization was detected antiphase with respect to the active $^nJ_{\text{XH}}$ coupling constant. However, this can lead to partial cancellation of weak long-range correlation cross-peaks, which, in the worst case, may lead to missing correlations. A refocused HSQMBC was later reported⁴⁹ so that the antiphase term was transferred back to in-phase magnetization by applying a reverse-INEPT, analogous to the HSQC pulse sequence. Moreover, under these conditions, heteronuclear decoupling can be applied during acquisition, with a consequent sensitivity improvement. The pulse sequence for ^1H - ^{15}N HSQMBC is shown in Figure 3.14. An 8 Hz ^1H - ^{15}N HSQMBC of strychnine (**1**) is shown in Figure 3.15.

In HMBC and HSQMBC experiments, the signal intensity of the responses strongly depends on the congruence between the long-range delay

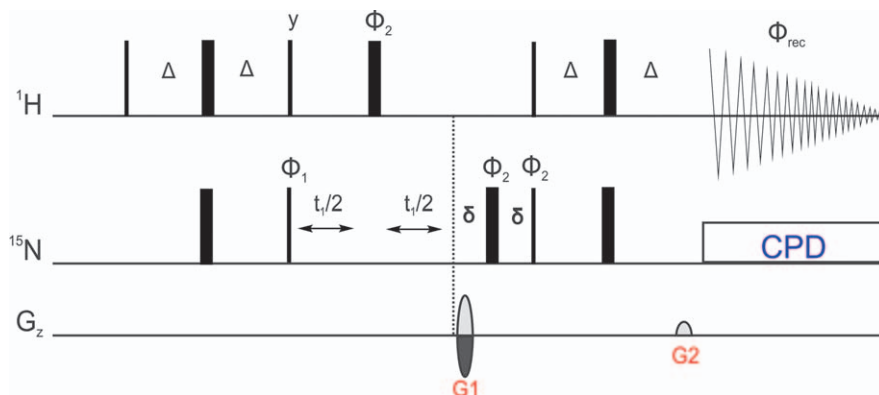


Figure 3.14 Pulse sequence for ^1H - ^{15}N HSQMBC. The delay, Δ , in the two INEPT transfers is optimized as a function of $1/4 \ ^nJ_{\text{NH}}$. Phases are cycled as $\Phi_1 = 0, 2$; $\Phi_2 = 0, 0, 2, 2$; and $\Phi_{\text{rec}} = 0, 2, 2, 0$.

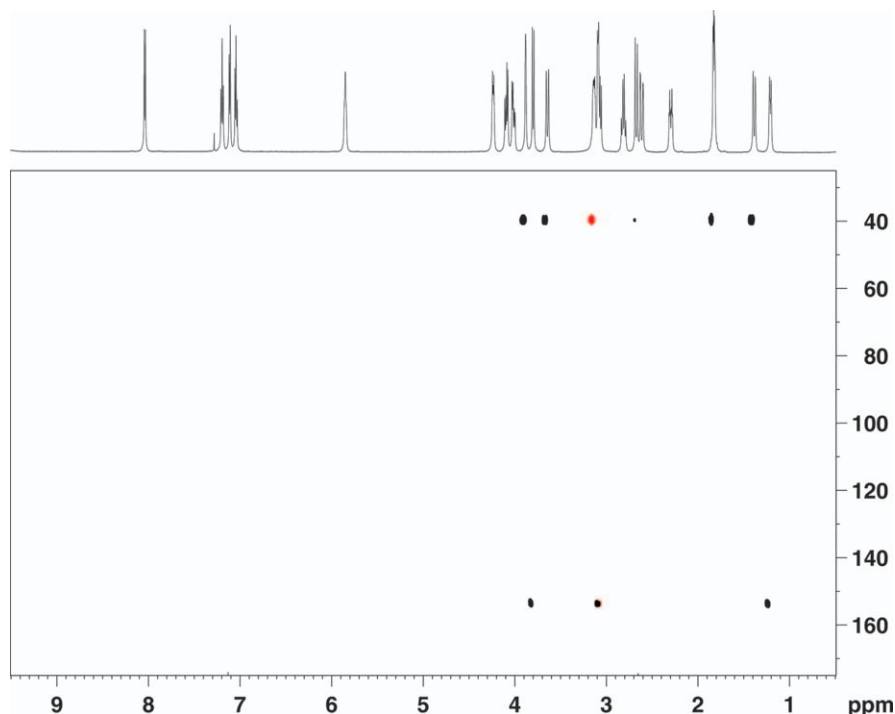


Figure 3.15 ^1H - ^{15}N HSQMBC spectrum of strychnine (1). The experiment was optimized for 8 Hz. The data were acquired as 3K points in F_2 using 48 transients for each one of the 80 increments to digitize the F_1 frequency domain. Data were zero-filled to afford the $4\text{K} \times 1\text{K}$ spectrum shown. The acquisition time was 1 hour and 36 minutes.

optimization and the actual coupling constant, by means of sine function $\sin(\pi^n J_{\text{XH}} \Delta)$. Therefore, it is rather common to have a scenario where correlations with a small $^n J_{\text{XH}}$ appear either as extremely weak responses or not

at all, which can lead to misinterpretation of the data or, in the worst case, misassignment of a structure.

3.8.4 ^1H - ^{15}N HSQMBC-TOCSY: A Hyphenated Long-range ^1H - ^{15}N Experiment

The HSQMBC-TOCSY⁵⁰ experiment has recently been reported and provides an elegant way to restore some of the potentially weak and/or missing correlations. The experiment is based on the extension of the HSQMBC pulse sequence by the addition of a TOCSY module.⁵⁹ In this way, the initial $^nJ_{\text{XH}}$ information, which is codified in the first part of the pulse sequence (HSQMBC step), is further transferred along a remote proton spin system through the TOCSY transfer, thus extending the long-range connectivity to the farthest protons of a given spin system irrespective of the real $^nJ_{\text{XH}}$ value between the heteronucleus and the remote proton(s). The magnetization transfer pathway is illustrated schematically in Figure 3.16. The pulse sequence, shown in Figure 3.17, is a slight modification of the HSQMBC experiment. Following the HSQMBC pulse sequence element, built from an initial INEPT, the nitrogen evolution period, and finally, a reverse INEPT element, a TOCSY module is appended to propagate the information to a given proton spin system.

The intensity of the relayed responses depends first on the initial $^nJ_{\text{NH}}$ magnetization transfer and second on the existence of an effective TOCSY pathway. For example, the signal intensity of the relayed $^4J_{\text{N-H3}}$ will only depend on the initial $^3J_{\text{NH2}}$ (black arrow) + J_{H2H3} (red arrow). The HSQMBC step is usually optimized to 8 Hz so that strong correlations are obtained, whereas the TOCSY mixing time is usually optimized for some value between 20 and 60 ms. Owing to this dependence of the transfer mechanisms, very-long-range correlations can be obtained even if the actual long-range heteronuclear coupling constant is close to zero. These correlations can be useful in confirming ^{15}N assignments when the proton that is long-range coupled to ^{15}N (e.g. $^1\text{H}_3$ in Figure 3.16) is overlapped with other proton resonances.

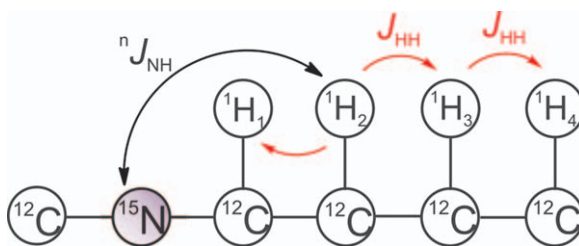


Figure 3.16 Schematic representation of the magnetization transfer pathway in the HSQMBC-TOCSY experiment. The initial magnetization transfer in the HSQMBC step is represented by the black double-headed arrow between $^1\text{H}_2$ and the ^{15}N resonance. Once $^1\text{H}_2$ is labeled with the ^{15}N chemical shift information, the subsequent TOCSY transfers are denoted by the red arrows.

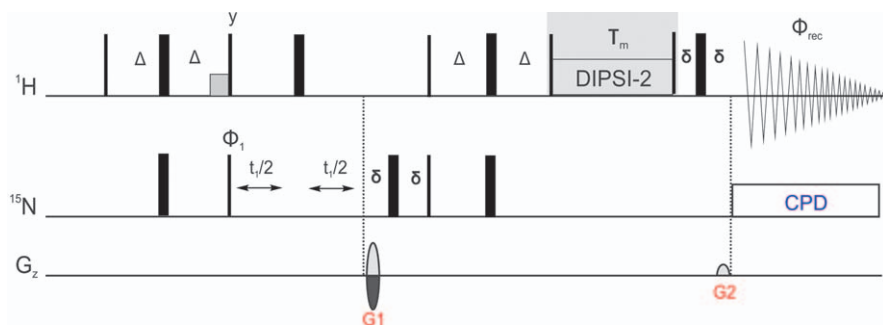


Figure 3.17 Pulse sequence for ^1H - ^{15}N HSQMBC-TOCSY. The experiment is essentially identical to the ^1H - ^{13}C experiment (see Figure 9.39 in Chapter 9) with the exception of the X-nucleus and the gradient ratio, which must be adjusted for the gyromagnetic ratio of ^{15}N (γ_{N}). The delay, Δ , in the two INEPT transfers is optimized as a function of $1/4 \ ^nJ_{\text{NH}}$. Phases are cycled as $\Phi_1 = 0, 2$ and $\Phi_{\text{rec}} = 0, 2$. The mixing time for the TOCSY transfer is usually set to the 20–60 ms range.

3.8.4.1 ^1H - ^{15}N HSQMBC-TOCSY of Cyclosporine

The HSQMBC-TOCSY spectrum of cyclosporine (7) is shown in Figure 3.18. The spectrum was acquired with the initial HSQMBC step optimized for 3 Hz and the subsequent TOCSY transfer optimized for 40 ms. Unlike the ^1H - ^{15}N HSQC-TOCSY spectrum of cyclosporine (7) shown in Figure 3.9, which contains only correlations from the four protonated nitrogens in the structure, in contrast, correlations can potentially be observed from all of the cyclosporine nitrogens in the HSQMBC-TOCSY experiment.

3.8.4.2 ^1H - ^{15}N HSQMBC-TOCSY of Strychnine

An example of the ^1H - ^{15}N HSQMBC-TOCSY experiment using strychnine (1) is shown in Figure 3.19. The experiment was optimized to 8 Hz with a 60 ms mixing time. When compared to the standard 8 Hz-optimized ^1H - ^{15}N HMBC shown in Figure 3.13, five additional correlations are obtained for nitrogen N19, while another two new correlations appear for nitrogen N9. One particularly useful correlation is that between N19 and the H22 vinyl proton established *via* the TOCSY transfer.

As will be discussed later in this chapter, the LR-HSQMBC³ experiment (see Section 3.8.4.4) represents an alternative tool to gain access to either weak or very long-range correlations. Unlike the HSQMBC-TOCSY experiment, the LR-HSQMBC transfer mechanism is based on a direct $^nJ_{\text{NH}}$ pathway, where the delay optimization is typically set to 2–4 Hz in order to observe the desired weak responses. Thus, in LR-HSQMBC, no proton spin systems are required for good performance of the experiment, and therefore its usage is especially recommended for proton-deficient compounds. LR-HSQMBC is a much more sensitive technique than, for example,

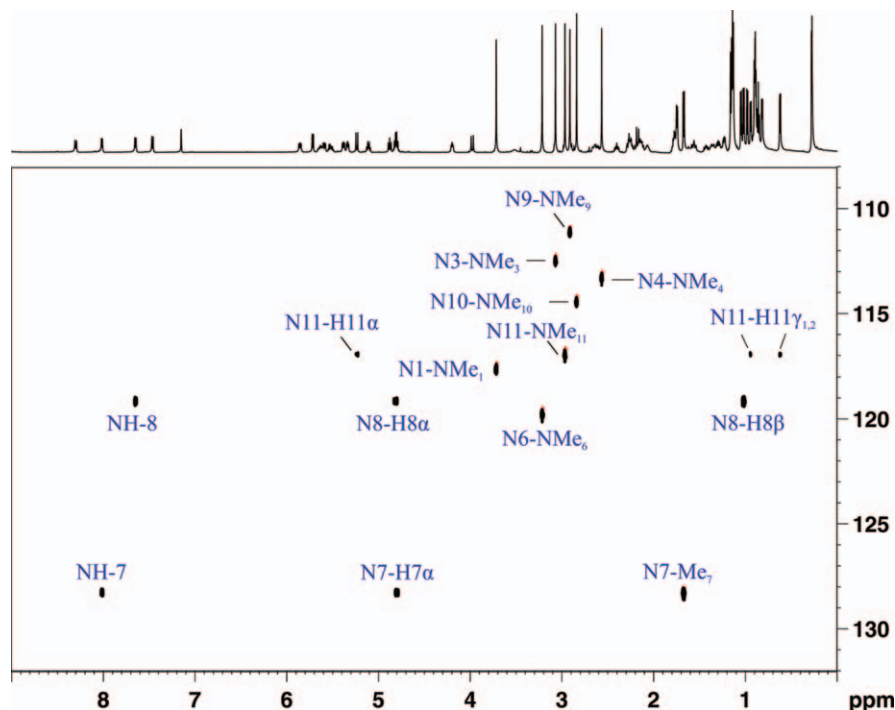


Figure 3.18 HSQMBC-TOCSY spectrum of cyclosporine (7) acquired with the initial long-range ^1H - ^{15}N transfer of the HSQMBC segment of the experiment optimized for 3 Hz, with the subsequent TOCSY transfer optimized for 40 ms. Correlations in the spectrum are labeled using the numbering scheme shown on 7.

optimizing the HMBC experiment in the range of 2–3 Hz. However, if a chemist is looking for additional correlations in a proton-rich compound due to missing responses in conventional HMBC or HSQMBC spectra, then HSQMBC-TOCSY might be a better option in terms of general sensitivity as compared to LR-HSQMBC. This difference can be understood by considering the length of the pulse sequences of both experiments. While the LR-HSQMBC can last up to ~500 ms (when optimized for 2 Hz), an 8 Hz-optimized HSQMBC-TOCSY will usually last about 150–190 ms (depending on the TOCSY mixing time being used). Hence, losses due to T_2 relaxation can be more problematic in the LR-HSQMBC experiment. Figure 3.20 shows 1D slices from nitrogen N19 of strychnine extracted from the LR-HSQMBC (2 Hz) and the HSQMBC-TOCSY (8 Hz + 60 ms mixing time). Clearly, not only are more correlations obtained in the HSQMBC-TOCSY experiment, but also they are obtained with a better overall sensitivity.

As mentioned above, the HSQMBC-TOCSY experiment is usually optimized to 8 Hz to establish strong correlations that will be subsequently propagated by the TOCSY transfer. It is interesting to note, though, that for the nitrogen N9 shown in Figure 3.21, the signal intensity is rather similar to the

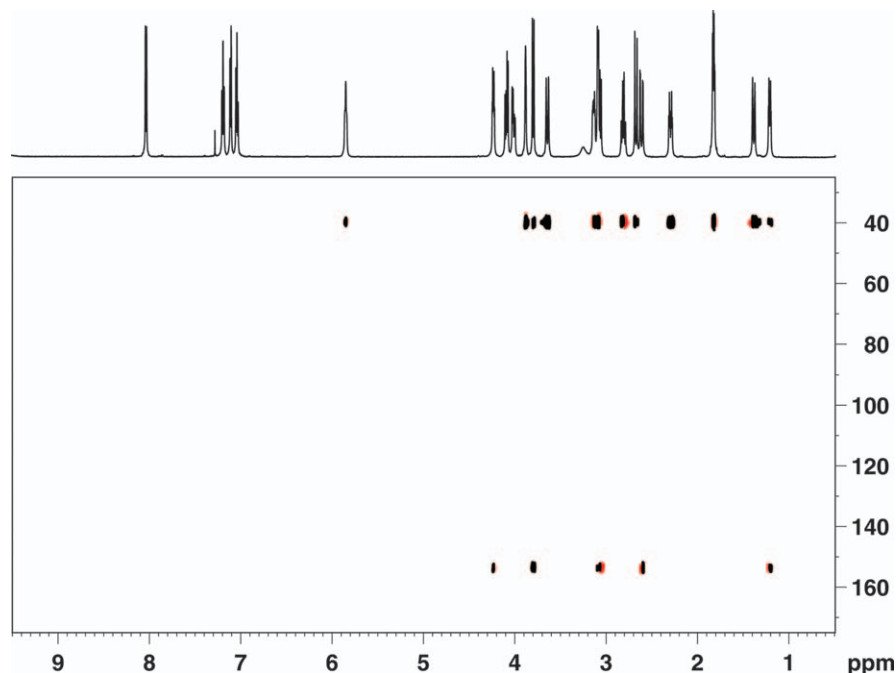


Figure 3.19 ^1H - ^{15}N HSQMBC-TOCSY spectrum of strychnine (**1**). The spectrum was optimized for 8 Hz plus 60 ms for the TOCSY. The data were acquired as 2K points in F_2 using 64 transients for each one of the 80 increments to digitize the F_1 frequency domain. Data were zero-filled to afford the $4\text{K} \times 1\text{K}$ spectrum shown. The acquisition time was 2 hours and 40 minutes.

2 Hz-optimized LR-HSQMBC (or even somewhat better in the latter). This can be explained by the fact that the heteronuclear coupling constants to nitrogen N9 are smaller than those associated with N19 (see Table 3.5). Thus, an 8 Hz HSQMBC-TOCSY is somewhat non-optimal for nitrogen N9. An alternative would be to acquire the experiment using a more appropriate delay (*e.g.* 5 Hz), but in that case, the duration of the pulse sequence would become correspondingly longer and could lead to losses of response intensity due to more extensive relaxation effects. The most viable alternative is probably employing a 2 Hz-optimized LR-HSQMBC experiment (see Section 3.8.4.4).

3.8.4.3 ^1H - ^{15}N CIGAR-HMBC

Static optimization of the HMBC experiment for a single fixed value of the long-range $^nJ_{\text{NH}}$ coupling constant is problematic in ^1H - ^{15}N heteronuclear shift correlation experiments, due to the range of $^nJ_{\text{NH}}$ coupling constants. An alternative to static optimization is provided by the concept of accordion optimization.⁵¹ Quite simply, when the duration of the evolution period is short, the duration of the long-range delay is optimized for a small coupling constant, which corresponds to a relatively long delay. In successive

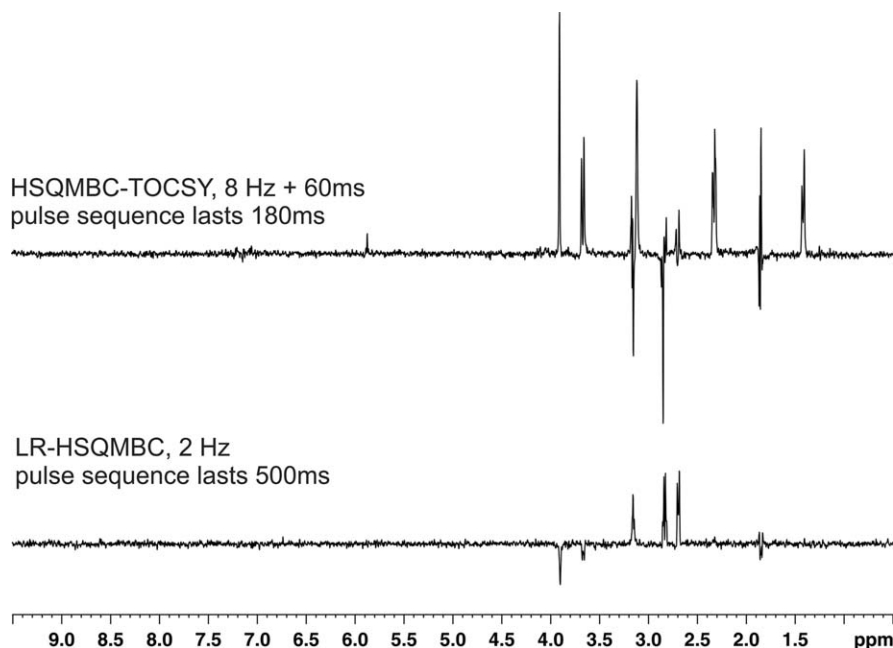


Figure 3.20 1D slices from nitrogen N19 of strychnine extracted from the ^1H - ^{15}N HSQMB-TOCSY spectrum (top) and the 2 Hz LR-HSQMB (bottom; see Section 3.8.4.4). The data were acquired and processed in exactly the same way as 2K points in F_2 and the F_1 frequency domain was digitized with 80 increments, with 64 transients. These data were zero-filled to afford the $4\text{K} \times 1\text{K}$ spectrum shown, and the vertical scaling of the two plots is identical.

experiments, as the duration of the evolution period grows longer from incrementation, the duration of the long-range delay is optimized for progressively larger heteronuclear couplings. In this fashion, as the evolution time lengthens while the F_1 frequency domain is digitized, the duration of the long-range delay is correspondingly decremented, thereby keeping the overall duration of the experiment as short as possible. The first such experiment was the ACCORD-HMBC experiment reported by Wagner and Berger.⁵² It was obvious in a study by one of the current authors (GEM) that the width of responses in the F_1 frequency domain was a function of the number of increments used to digitize F_1 .⁵³ A series of reports followed that addressed the F_1 dispersion of responses inherent to the ACCORD-HMBC experiment. The first modification was the development of the IMPEACH-MBC experiment reported in 1999.⁵⁴ The first application of accordion optimization for the acquisition of ^1H - ^{15}N data was the report by Martin and Hadden¹ in 2000 when they compared the results obtained using both the ACCORD-HMBC and IMPEACH-MBC experiments. A clear advantage over conventional ^1H - ^{15}N HMBC was noted for the data acquired using ^1H - ^{15}N IMPEACH-MBC. The development of the IMPEACH-MBC experiment was followed by the CIGAR-HMBC experiment in 2000.⁵⁵ Subsequent work by

HSQMBC-TOCSY, 8 Hz + 60ms
pulse sequence lasts 180ms



LR-HSQMBC, 2 Hz
pulse sequence lasts 500ms

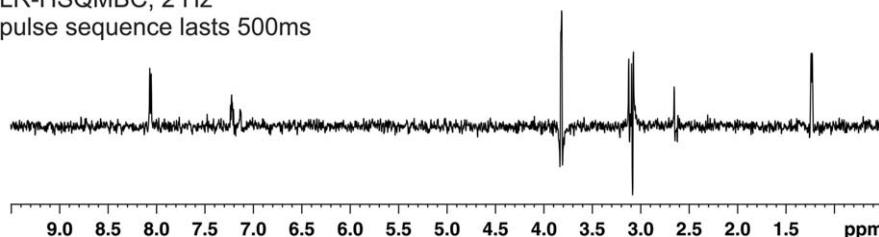


Figure 3.21 1D slices from nitrogen N9 of strychnine extracted from ^1H - ^{15}N HSQMBC-TOCSY spectrum (top) and the 2 Hz LR-HSQMBC (bottom). The data were acquired and processed in exactly the same way as 2K points in F_2 and the F_1 frequency domain was digitized with 80 increments, with 64 transients. That data were zero-filled to afford the $4\text{K} \times 1\text{K}$ spectrum shown.

Table 3.5 Calculated ^1H - ^{15}N heteronuclear coupling constants for strychnine (**1**). The coupling constants were calculated using DFT methods at the b3lyp/6-311 + G(d,p)//belyp/6-31G(d) level of theory. A selection of these calculated ^1H - ^{15}N coupling constants for strychnine was published previously,³ and the complete list is presented here.

Proton	N9	N19
1	-0.31	0.02
2	-0.09	0.00
3	-0.77	0.01
4	-0.60	-0.01
8	-1.62	0.00
11a	-2.84	0.02
11b	0.07	0.02
12	0.19	0.03
13	-2.19	-0.06
14	-0.17	0.38
15a	-0.02	-3.62
15b	-0.09	-0.47
16	-0.16	-7.18
17a	0.09	-0.09
17b	0.03	-4.92
18a	-0.02	-7.65
18b	0.09	-0.09
20a	0.01	0.35
20b	-0.04	-1.43
22	0.01	0.35
23a	0.00	-0.22
23b	0.02	-0.01

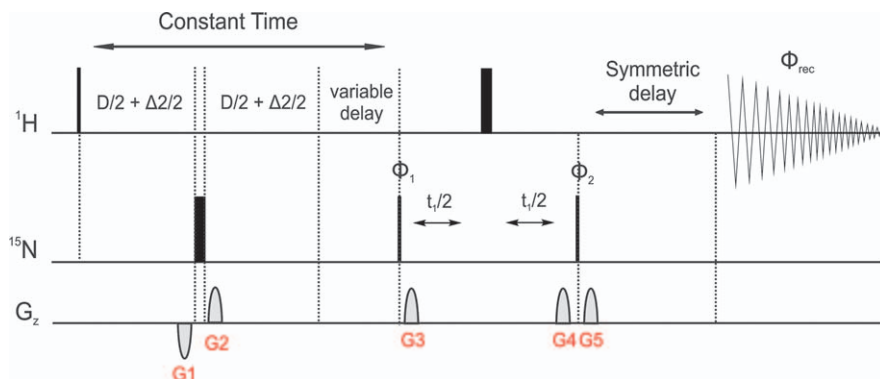


Figure 3.22 Pulse sequence for the accordion-optimized ^1H - ^{15}N CIGAR-HMBC experiment.⁵⁵ The experiment utilizes the concept of the constant time variable delay introduced in the IMPEACH-MBC experiment.⁵⁴ The 180° pulse is successively repositioned as the evolution time is incremented. The variable delay, v_d , in the constant time variable delay is decremented to go from optimization for the smallest optimization value (longest delay interval; e.g. 3 Hz) in the first increment of the evolution time to the largest optimization (shortest delay interval; e.g. 8 Hz). In parallel, the delays flanking the 180° pulse are successively incremented, keeping the overall duration of the time variable delay constant and equal to the symmetrical delay at the end of the pulse sequence.

Cheatham and coworkers⁵⁶ and also by Russell in an Agilent application note⁵⁷ made the same observation regarding the advantages of accordion optimization over conventional static optimization for ^1H - ^{15}N long-range heteronuclear shift correlation experiments.

The CIGAR-HMBC pulse sequence is shown in Figure 3.22. The experiment is intended to reintroduce, in a controlled fashion, whatever degree of F_1 modulation or “twist” that an investigator wishes, as a means of authenticating weak correlations and discriminating them from fortuitous noise. Legitimate correlations will be modulated in F_1 , whereas noise will not. The degree of modulation introduced is controlled by a parameter, the J -scale. The modulation is clearly illustrated in the expansion of the correlations for the N19 resonance of strychnine (1) shown in the inset in Figure 3.23.

3.8.4.4 ^1H - ^{15}N LR-HSQMBC

The LR-HSQMBC experiment is discussed in Volume 2, Chapter 9, to which the interested reader is referred for a more complete discussion of the experiment.³ The experiment was initially developed to probe molecular structures for very-long-range heteronuclear correlations (e.g. $>^4J_{\text{CH}}$), the observation of which can be problematic with experiments such as HMBC, where magnetization is antiphase at the end of the pulse sequence. HMBC data are magnitude calculated for presentation, which can result in the cancellation of very weak correlations. The LR-HSQMBC experiment, in contrast, employs a reverse-INEPT step to refocus magnetization after

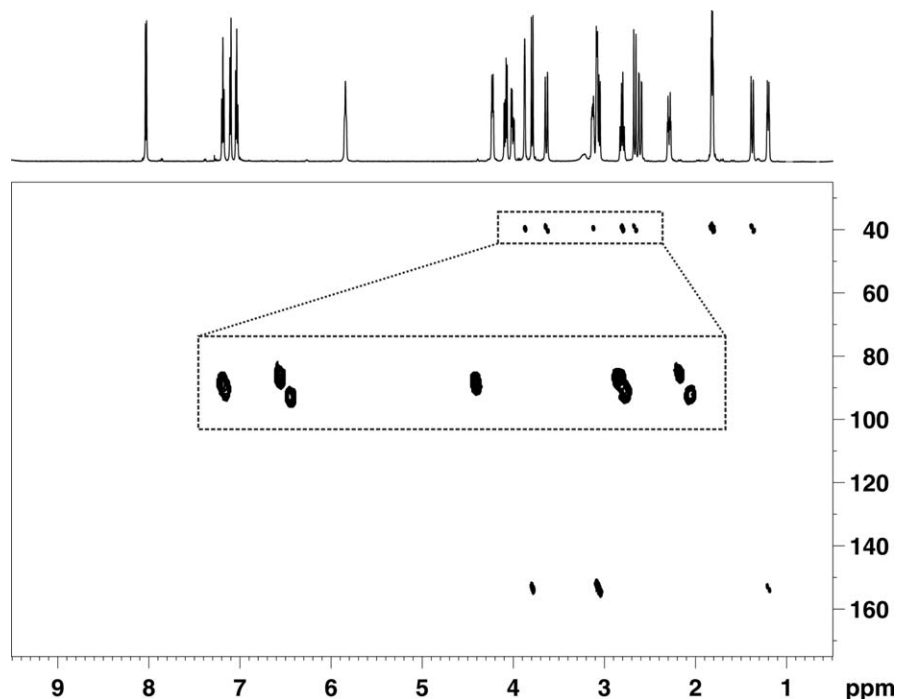


Figure 3.23 ^1H - ^{15}N CIGAR-HMBC spectrum of strychnine (**1**). The data were acquired with the scaling parameter J -scale = 6 to reintroduce a degree of modulation in F_1 for response authentication. The data were acquired as 4K points in F_2 using 16 transients for each one of the 128 increments, for a duration of 1 hour and 20 minutes.

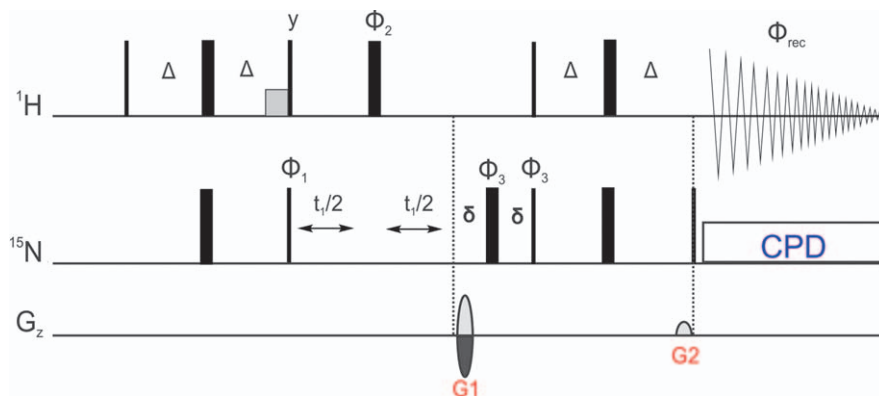


Figure 3.24 Pulse sequence for ^1H - ^{15}N LR-HSQMBC. The experiment employs a reverse-INEPT step and a clean in-phase module composed of the G2 gradient and a 90° pulse to remove any undesirable antiphase contributions to the in-phase magnetization prior to detection.

evolution, allowing broadband X-nucleus decoupling during acquisition. The pulse sequence is shown in Figure 3.24.

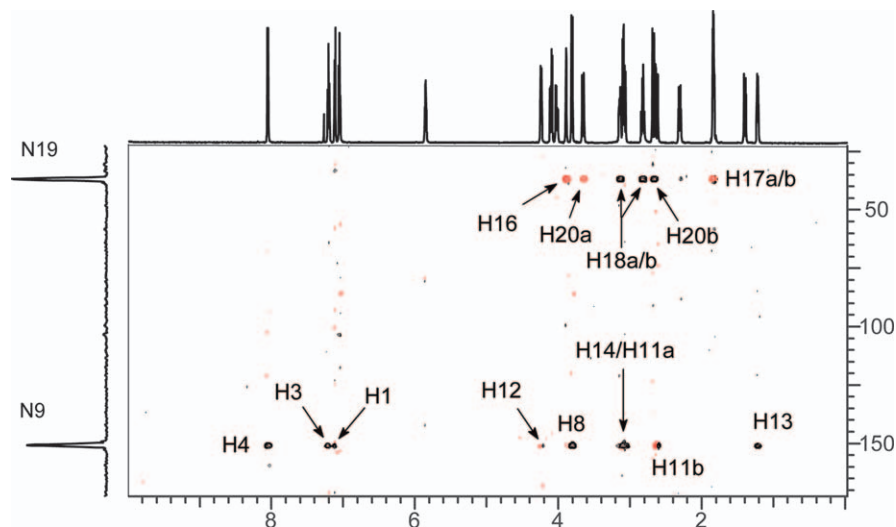
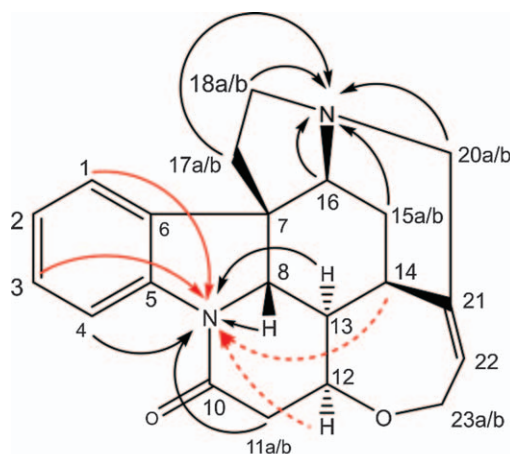


Figure 3.25 ^1H - ^{15}N LR-HSQMBC spectrum of strychnine optimized for 2 Hz. While $^2J_{\text{NH}}$ and $^3J_{\text{NH}}$ correlations typify ^1H - ^{15}N HMBC spectra, which are typically optimized in the range of 5–8 Hz, the refocused nature of the LR-HSQMBC pulse sequence facilitates optimization at 2 Hz, which frequently leads to the observation of $^4J_{\text{NH}}$ correlations as shown in 11.

The 2 Hz-optimized ^1H - ^{15}N LR-HSQMBC spectrum of strychnine (**1**) is shown in Figure 3.25. Correlations observed in the 2 Hz ^1H - ^{15}N LR-HSQMBC spectrum of strychnine are summarized in 11. Typically observed $^2J_{\text{NH}}$ and $^3J_{\text{NH}}$ correlations are designated by black arrows. Correlations *via* $^4J_{\text{NH}}$ are designated by red arrows. Weak correlations are denoted by dashed arrows. Several of the $^4J_{\text{NH}}$ correlations observed in the spectrum shown in Figure 3.25 had never been previously reported for strychnine.³



3.8.5 Triple-resonance ^1H - ^{15}N Experiments

To date, there are only two ^1H - ^{15}N triple-resonance NMR experiments that have been considered. The first can be described as a time-shared experiment that was reported by Parella and colleagues⁵ in 2007 that provides for the simultaneous acquisition of ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra. The second experiment is the HCNMBC experiment reported late in 2014 by Cheatham and coworkers⁵⁸ that establishes ^{13}C - ^{15}N correlations at natural abundance *via* either $^1J_{\text{CN}}$ or $^nJ_{\text{CN}}$ in a manner analogous to the magnetization transfer pathways of the 1,1- and 1,*n*-ADEQUATE experiments (see Volume 2, Chapter 10). From a sensitivity standpoint, the HCNMBC experiment is still lower in sensitivity than the corresponding ADEQUATE experiments, since it operates as a function of the natural abundance of both ^{13}C (1.13%) and ^{15}N (0.37%).

While the former of the two experiments described in this section has obvious applicability in natural products structure characterization, it remains to be seen if the latter experiment will be employed in the characterization of natural products. To date, the only applications of the HCNMBC experiment have been the application to nicotine reported by Cheatham and coworkers and the data for strychnine (**1**) that appear below.

3.8.5.1 Simultaneous Acquisition of ^1H - ^{13}C and ^1H - ^{15}N HMBC Spectra

Time-shared acquisition of ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra can be accomplished using the triple-resonance CN-HMBC experiment developed in 2007 by Parella and coworkers.⁵ The pulse sequence for the experiment is shown in Figure 3.26. To discriminate correlations originating from ^{13}C and ^{15}N , two datasets are recorded with a minimum two-step phase cycle: 1) $\varphi_1 = \varphi_2 = (\text{x}, -\text{x})$; and 2) $\varphi_1 = (-\text{x}, \text{x})$, $\varphi_2 = (\text{x}, -\text{x})$, with the receiver phase always set as x , $-\text{x}$. Following data acquisition, the two datasets are added/subtracted to afford separate ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra.

Employing the pulse sequence shown in Figure 3.26, the time-shared CN-HMBC spectra presented in Figure 3.27 were obtained. Figures 3.27A and 3.27B show the individually stored spectra. Addition of the two spectra affords the ^1H - ^{13}C HMBC spectrum shown in Figure 3.27C. Subtraction of the two spectra provides the ^1H - ^{15}N HMBC spectrum that is shown in Figure 3.27D.

When the experiment was initially reported, Parella and coworkers⁵ also employed strychnine (**1**) as a model compound to test the experiment. In their initial study, the results from the CN-HMBC experiment were compared with separately acquired ^1H - ^{13}C HMBC and ^1H - ^{15}N HMBC spectra. The C10 carbonyl and N9 amide resonances were used for the basis of the performance comparison, which is presented in Figure 3.28. When the CN-HMBC data were acquired using the same experimental time as the

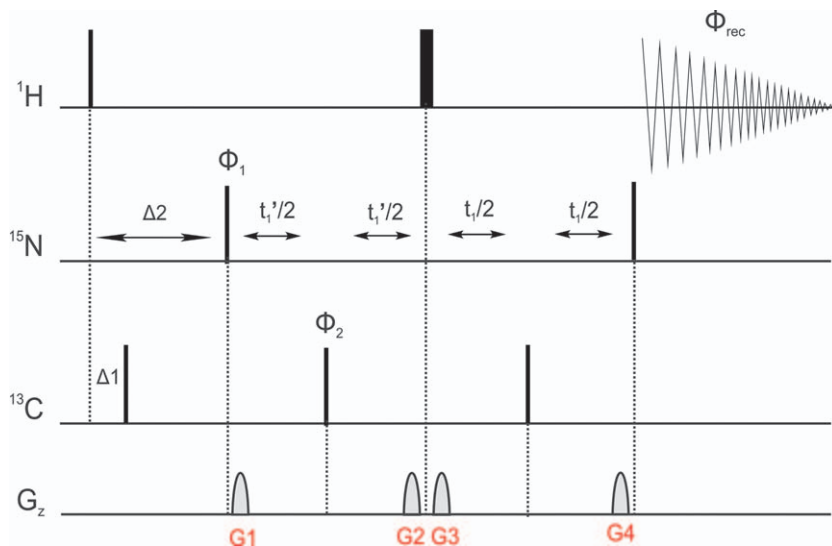


Figure 3.26 Pulse sequence for the time-shared CN-HMBC experiment.⁵ The experiment is triple resonance and provides for the simultaneous acquisition of ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra. Two datasets are recorded with a minimum two-phase cycle: (1) $\phi_1 = \phi_2 = (x, -x)$; and (2) $\phi_1 = (-x, x)$, $\phi_2 = (x, -x)$, with the receiver phase $\phi_{\text{rec}} = (x, -x)$.⁵

individual HMBC experiments (Figures 3.28A and 3.28B, respectively), the s/n ratio for the C10 carbonyl and N9 amide resonances was comparable. However, when the CN-HMBC experiment was performed for the aggregate time expended to acquire the two separate HMBC spectra, better s/n was obtained, as shown in Figure 3.28C.

3.8.5.2 ^{13}C - ^{15}N HCNMBC— ^{13}C - ^{15}N Correlation at Natural Abundance

One other experiment that is potentially applicable to alkaloid structure characterization is the newly reported HCNMBC triple-resonance experiment, first described in late 2014 for the establishment of ^{13}C - ^{15}N heteronuclear correlations at natural abundance.^{58,59} To date, the experiment has been applied predominantly to small, nitrogen heterocycles. The pulse sequence for the experiment is shown in Figure 3.29. A schematic representation of the magnetization transfer pathways in the HCNMBC experiment is shown in Figure 3.30 and, finally, the HCNMBC spectrum of strychnine (**1**) optimized for $J_{\text{CN}} = 6$ Hz is shown in Figure 3.31. Obviously, performing the experiment requires a gradient triple-resonance probe and a three-channel NMR spectrometer console. The acquisition of these data has never been reported using a conventional room temperature probe because of the low statistical probability of both ^{13}C and ^{15}N in the same molecule.

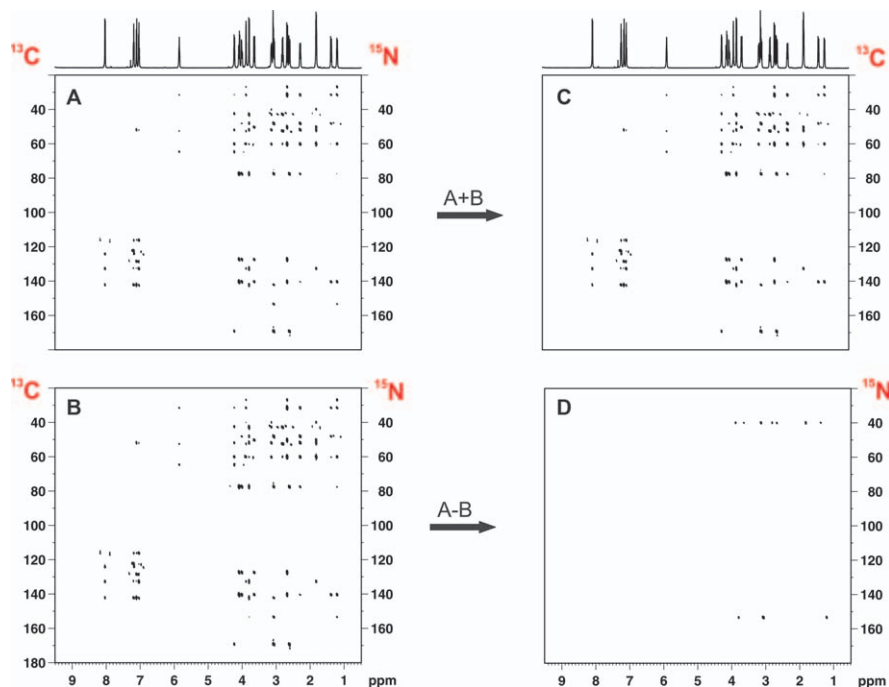


Figure 3.27 2D time-shared CNHMBc spectra of strychnine (**1**) acquired with relative phases $\phi_1 = (-x, x)$ in (A), and with a relative phase $\phi_1 = (x, -x)$ in (B). Linear combinations of (A) and (B) followed by conventional processing provides two separate spectra with the nuclei edited: (C) ^{13}C -HMBC spectra and (D) ^{15}N -HMBC spectra with enhanced sensitivity.

The initial magnetization transfer in the HCNMBc experiment is *via* $^1J_{\text{CH}}$, although presumably (albeit probably at a severe sensitivity penalty) one could also begin with an initial $^nJ_{\text{CH}}$ transfer. These transfers are designated by black arrows in the schematic shown in Figure 3.30. The subsequent magnetization transfer can be either *via* $^1J_{\text{CN}}$, designated by a green arrow, or *via* $^nJ_{\text{CN}}$, designated by a red arrow. As in the ADEQUATE experiment (see Volume 2, Chapter 10), HCNMBc is an out-and-back experiment and the return magnetization transfer path is identical to the outbound path.

The HCNMBc spectrum of strychnine (**1**) is shown in Figure 3.31 and the correlations in the spectrum are summarized by **12**. Correlations *via* $^1J_{\text{CN}}$ coupling pathways are denoted by red bonds. Dashed bonds or arrows denote weak correlations. Long-range correlations *via* $^nJ_{\text{CN}}$ are designated by red arrows.

It is far too early in the development of the HCNMBc experiment to even consider speculating on its potential utility for natural product structure elucidation applications. One inherent difficulty in using the experiment

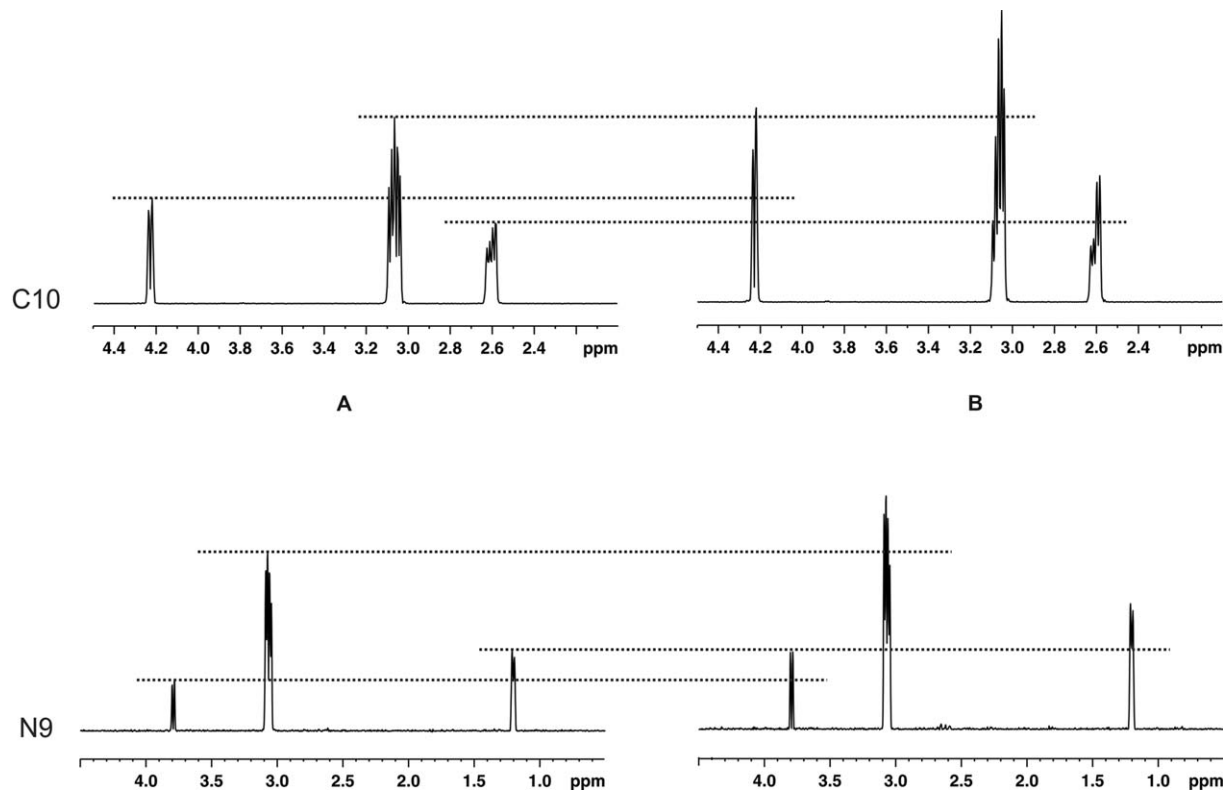


Figure 3.28 Slices taken at the ^{13}C chemical shift of the C10 carbonyl resonance and the N9 amide ^{15}N resonance of strychnine (**1**).⁵ (A) Slices from the ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra of strychnine acquired separately. (B) Individual slices extracted from the CN-HMBC spectra acquired with the same number of scans/ t_1 increment as the individual HMBC spectra. (C) CN-HMBC spectra acquired with the same total measuring time as expended for both spectra in (A). (Reproduced with permission from M. Pérez-Trujillo, P. Nolis, and T. Parella, *Org. Lett.*, **9**, 29–32 (2007). Copyright (2007) American Chemical Society.)

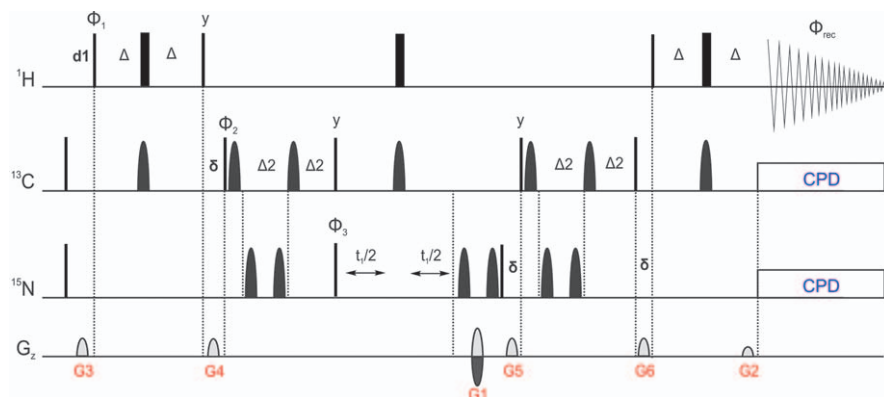


Figure 3.29 Pulse sequence for the HCNMBC experiment described by Cheatham and coworkers.^{58,59} The delays were optimized as follows: $\Delta = 1/4 (J_{\text{CH}})$, being $^1J_{\text{CH}}$ set to 145 Hz; $\Delta 2 = 1/4 (J_{\text{CN}})$, being J_{CN} set 6 Hz. Phases were cycled as follows: $\phi_1 = 0, 0, 0, 0, 2, 2, 2, 2, 2, 2$; $\phi_2 = 0, 0, 0, 2, 2$; $\phi_3 = 0, 2$; $\phi_{\text{rec}} = 0, 2, 2, 0, 2, 0, 0, 2$. Gradients were optimized as follows: $G1 : G2 = 80 : 8.1$; $G3 = 31$; $G4 = 19$; $G5 = 11$; $G6 = 7$.

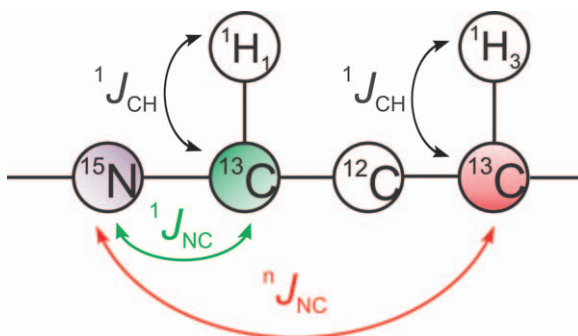


Figure 3.30 Schematic representation of the magnetization transfer pathways in the HCNMBC experiment (see Figures 3.29 and 3.31). The initial magnetization transfer is via $^1J_{\text{CH}}$ as designated by the black arrows. Analogous to the ADEQUATE experiments (see Volume 2, Chapter 10), the HCNMBC experiment is an out-and-back experiment. Following the initial magnetization transfer, magnetization is transferred either via $^1J_{\text{NC}}$ (designated by the green arrow) or via $^nJ_{\text{NC}}$ (designated by the red arrow). Finally, magnetization is transferred back to the starting proton via the same pathway for detection. Correlations observed in the HCNMBC experiment for strychnine optimized for $J_{\text{NC}} = 6$ Hz are shown in Figure 3.31 and are illustrated in 12.

is determining the optimization of the delay for the J_{CN} coupling. DFT calculations were used to set the optimization of the J_{CN} delay and were used to record the spectrum shown in Figure 3.31.

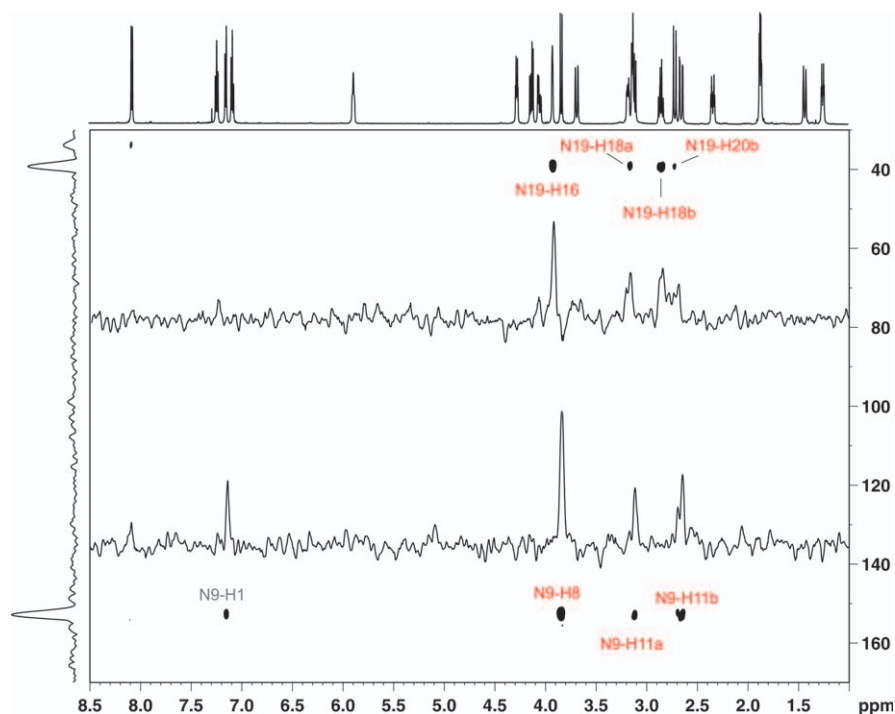
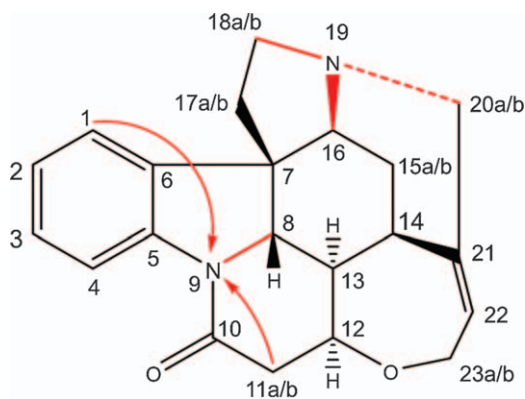


Figure 3.31 HCNMBC spectrum of strychnine (**1**) acquired with the delay for the J_{CN} heteronuclear coupling optimized for 6 Hz. The data were acquired using a 4 mg sample of strychnine dissolved in 35 μL of deuteriochloroform in a 1.7 mm NMR tube. The data were acquired using a 600 MHz spectrometer equipped with a $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ gradient triple-resonance Micro-CryoProbeTM over 15 hours. Slices extracted from the spectrum at the N9 and N19 F_1 chemical shifts are superimposed on the spectrum to show the signal-to-noise obtained in the experiment. DFT calculation of the J_{CN} coupling constants was used in the optimization of the experiment.



3.9 Conclusions

Since the first reported applications of ^1H - ^{15}N HMBC in poster presentations in 1993, followed by the first published reports in 1996, the utilization of ^1H - ^{15}N long-range heteronuclear chemical shift correlation has become an active area of NMR research. There have been multiple reviews and chapters detailing the applications of these experiments in structure elucidation problems, originating from both the natural product and synthetic research communities. There is now a reasonably robust arsenal of experiments that can be utilized to probe long-range ^1H - ^{15}N correlation pathways of complex natural products. These experiments, in combination with cryogenic NMR probe technology, especially the 1.7 mm Micro-CryoProbe™, bring these experiments within reach even for samples in the low-microgram range, especially for the ^1H - ^{15}N HSQC experiment.

Given the slow adoption rate of new experimental NMR methods by the natural products chemistry community (see Volume 1, Chapter 1,), it may be some time before some of the newer experiments, particularly the triple-resonance HCNMBC experiment for the natural abundance ^{13}C - ^{15}N correlation, begin to see application in natural product structure elucidation studies. However, the value of ^1H - ^{15}N heteronuclear shift correlation data is firmly established and, as has been shown in a number of studies, also has a very significant impact on CASE program performance (see Section 3.6). It will be quite interesting to see what new experiments are developed in the years to come in order to probe long-range ^1H - ^{15}N heteronuclear correlations.

References

1. G. E. Martin and C. E. Hadden, *Magn. Reson. Chem.*, 2000, **38**, 251–256.
2. M. Kline and S. Cheatham, *Magn. Reson. Chem.*, 2003, **41**, 307–314.
3. R. T. Williamson, A. V. Buevich and G. E. Martin, *Tetrahedron. Lett.*, 2015, **55**, 3365–3366.
4. B. D. Hilton and G. E. Martin, *J. Nat. Prod.*, 2010, **73**, 1465–1470.
5. M. Pérez-Trujillo, P. Nolis and T. Parella, *Org. Lett.*, 2007, **9**, 29–32.
6. G. E. Martin and A. J. Williams, Applications of ^1H - ^{15}N Long-range Heteronuclear Shift Correlation and ^{15}N NMR in Alkaloid Chemistry, in *Annual Reports in NMR Spectroscopy*, ed. G. A. Webb, Elsevier, New York, 2015, vol. 84, pp. 1–76.
7. G. E. Martin, M. Solntseva and A. J. Williams, Applications of ^{15}N NMR in Alkaloid Chemistry, in *Modern Alkaloids*, ed. E. Fattorusso and O. Talialatela-Scafati, Wiley-VCH, New York, 2007, pp. 411–476.
8. G. E. Martin and C. E. Hadden, *J. Nat. Prod.*, 2000, **63**, 543–585.
9. R. Marek and A. Lyčka, *Curr. Org. Chem.*, 2002, **6**, 35–66.
10. G. E. Martin and A. J. Williams, Long-Range ^1H - ^{15}N Heteronuclear Shift Correlation, in *Annual Reports in NMR Spectroscopy*, ed. G. A. Webb, Elsevier, New York, 2005, vol. 56, pp. 1–119.

11. R. Marek, A. Lyčka, E. Kolehmainen, E. Sievänen and J. Toušek, *Curr. Org. Chem.*, 2007, **11**, 1154–1205.
12. G. E. Martin and A. J. Williams, Utilization of Long-Range ^1H - ^{15}N 2D NMR Spectroscopy for Chemical Structure Elucidation and Confirmation, in *Encyclopedia of Magnetic Resonance*, ed. R. K. Harris and R. A. Wasylishen, Wiley, New York, 2010 on-line: DOI: 10.1002/9780470034590.emrstm1083.
13. G. C. Levy and R. L. Lichter, *Nitrogen-15 NMR*, Wiley, New York, 1979.
14. M. Witanowski and R. A. Web, *Nitrogen NMR*, Plenum Press, London, 1973.
15. G. J. Martin, M. L. Martin, and J.-P. Gouesnard, ^{15}N NMR Spectroscopy, Springer-Verlag, Berlin, 1981.
16. S. Berger, S. Braun, and H.-O. Kalinowski, *NMR Spectroscopy of the Non-Metallic Elements*, Wiley, New York, 1997, pp. 111–318.
17. E. F. Moonney and P. H. Wilson, Nitrogen Magnetic Resonance Spectroscopy, in *Annual Reports in NMR Spectroscopy*, ed. E. F. Mooney, Academic Press, New York, 1968, vol. 2, pp. 125–152.
18. M. Witanowski, L. Stefaniak and G. A. Webb, Nitrogen NMR, in *Annual Reports in NMR Spectroscopy*, ed. E. F. Mooney, Academic Press, New York, 1972, vol. 5a, pp. 395–457.
19. M. Witanowski, L. Stefaniak and G. A. Webb, Nitrogen NMR, in *Annual Reports in NMR Spectroscopy*, ed. E. F. Mooney, Academic Press, New York, 1977, vol. 7, pp. 118–239.
20. M. Witanowski, L. Stefaniak and G. A. Webb, Nitrogen NMR, in *Annual Reports in NMR Spectroscopy*, ed. G. A. Webb, Academic Press, New York, 1981, vol. 11b.
21. M. Witanowski, L. Stefaniak and G. A. Webb, Nitrogen NMR, in *Annual Reports in NMR Spectroscopy*, ed. E. F. Mooney, Academic Press, New York, 1986, vol. 18.
22. M. Witanowski, L. Stefaniak and G. A. Webb, Nitrogen NMR, in *Annual Reports in NMR Spectroscopy*, ed. E. F. Mooney, Academic Press, New York, 1993, vol. 25.
23. G. E. Martin, B. D. Hilton, D. Moskau, N. Freytag, K. Kessler and K. Colson, *Magn. Reson. Chem.*, 2010, **48**, 935–937.
24. NMRShiftDB, Accessed on November 10th 2015: <http://nmrshiftdb.nmr.uni-koeln.de/>.
25. Bio-Rad NMR Spectral Databases, Accessed on November 10th 2015: <http://www.bio-rad.com/en-us/product/nmr-spectral-databases>.
26. ACD/NNMR Predictor, Accessed on November 10th 2015: http://www.acdlabs.com/products/adh/nmr/nmr_pred/.
27. ACD/Labs NMR Predictors What's New 2015: Accessed on November 10th 2015. http://www.acdlabs.com/products/adh/nmr/nmr_pred/whatsnew.php.
28. M. M. Senior, R. T. Williamson and G. E. Martin, *J. Nat. Prod.*, 2013, **76**, 2088–2093.
29. J. Muñoz, C. Moriou, J.-F. Gallard, P. D. Marie and A. Al-Mourabit, *Tetrahedron Lett.*, 2012, **53**, 5828–5832.

30. A. N. Tackie, G. L. Boye, M. H. M. Sharaf, P. L. Schiff Jr., R. C. Crouch, T. D. Spitzer, R. L. Johnson, J. Dunn, D. Minick and G. E. Martin, *J. Nat. Prod.*, 1993, **56**, 653–670.
31. G. E. Martin, C. E. Hadden, D. J. Russell, B. D. Kaluzny, J. E. Guido, W. K. Duholke, B. A. Stiemsma, T. J. Thamann, R. C. Crouch, K. Blinov, M. Elyashberg, E. R. Martirosian, S. G. Molodotsov, A. J. Williams and P. L. Schiff Jr., *J. Heterocycl. Chem.*, 2002, **39**, 1241–1250.
32. J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff Jr., T. Parella, R. T. Williamson and G. E. Martin, *Angew. Chem., Int. Ed.*, 2015, **54**, 10160–10164.
33. M. Köck, J. Junker and T. Lindel, *Org. Lett.*, 1999, **1**, 2041–2044.
34. J.-M. Nuzillard, J. D. Cormolly, C. Delavde, B. Richard, M. Ches-Hanrot and L. Le Men-Olivier, *Tetrahedron*, 1999, **55**, 11511–11518.
35. A. Grube and M. Köck, *J. Nat. Prod.*, 2006, **69**, 1212–1214.
36. G. E. Martin, C. E. Hadden, D. J. Russell, B. D. Kaluzny, G. E. Guido, W. K. Duholke, B. A. Stiemsma, T. J. Thamann, R. C. Crouch, K. Blinov, M. E. Elyashberg, E. R. Martirosian, S. G. Molodtsov, A. J. Williams and P. L. Schiff Jr., *J. Heterocycl. Chem.*, 2002, **39**, 1241–1250.
37. J. Saurí, K. A. Blinov, A. V. Buevich, M. H. M. Sharaf, P. L. Schiff Jr., T. Parella, R. T. Williamson and G. E. Martin, *J. Nat. Prod.*, manuscript in preparation (2015).
38. M. E. Elyashberg, A. J. Williams, G. E. Martin, Computer-Assisted Structure Elucidation, in *Progress in NMR Spectroscopy*, ed. J. Feeney and L. Sutcliffe, Pergamon, London, 2008, vol. 53, pp. 1–104.
39. M. E. Elyashberg, A. J. Williams and K. A. Blinov, *Contemporary Computer-Assisted Approaches to Molecular Structure Elucidation*, RSC Publishing, London, 2012, xxii + 482.
40. M. E. Elyashberg and A. J. Williams, *Computer-Based Structure Elucidation from Spectral Data*, Springer, Berlin, 2015, vol. xvi, p. 447.
41. A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, 1986, **108**, 2093–2094.
42. R. C. Crouch and G. E. Martin, *J. Heterocycl. Chem.*, 1995, **32**, 1665–1669.
43. R. T. Williamson, B. L. Marquez and W. H. Gerwick, *Tetrahedron Lett.*, 1999, **55**, 2881–2888.
44. G. E. Martin, R. C. Crouch, M. H. M. Sharaf, and P. L. Schiff Jr., “ ^1H - ^{15}N Direct and Long-Range Heteronuclear Shift Correlation Techniques – Potential Applications,” 34th Annual Meeting of the American Society of Pharmacognosy, San Diego, CA, July 1993, poster #101.
45. J. Uzawa, H. Utumi, H. Koshino, T. Hinomoto and K. Anzai, 32nd NMR Conference, Tokyo, Japan, November, 1993, pp. 147–150.
46. R. C. Crouch and G. E. Martin, *J. Heterocycl. Chem.*, 1995, **32**, 1665–1669.
47. G. E. Martin, R. C. Crouch and C. W. Andrews, *J. Heterocycl. Chem.*, 1995, **32**, 1759–1766.
48. R. Marek, L. Kralik and V. Sklenar, *Tetrahedron Lett.*, 1997, **38**, 665–668(b) R. T. Williamson, B. L. Marquez, W. H. Gerwick and K. E. Kover, *Magn. Reson. Chem.*, 2000, **38**, 265.
49. S. Gil, J. F. Espinosa and T. Parella, *J. Magn. Reson.*, 2010, **207**, 312–321.

50. J. Saurí, N. Marcó, R. T. Williamson, G. E. Martin and T. Parella, *J. Magn. Reson.*, 2015, **258**, 25–32.
51. G. Bodenhausen and R. R. Ernst, *J. Am. Chem. Soc.*, 1982, **104**, 1304–1309.
52. R. Wagner and S. Berger, *Magn. Reson. Chem.*, 1998, **36**, S44–S46.
53. G. E. Martin, C. E. Hadden, R. C. Crouch and V. V. Krishnamurthy, *Magn. Reson. Chem.*, 1999, **37**, 517–518.
54. C. E. Hadden, G. E. Martin and V. V. Krishnamurthy, *J. Magn. Reson.*, 1999, **140**, 274–280.
55. C. E. Hadden, G. E. Martin and V. V. Krishnamurthy, *Magn. Reson. Chem.*, 2000, **38**, 143–147.
56. M. Kline and S. Cheatham, *Magn. Reson. Chem.*, 2003, **41**, 307–314.
57. D. J. Russell, ^1H - ^{15}N CIGAR – Optimizing for Long-Range Nitrogen Detection – Application Note: https://www.agilent.com/cs/library/applications/5990-9062en_hi.pdf.
58. S. Cheatham, P. Gierth, W. Bermel and Ě. Kupče, *J. Magn. Reson.*, 2014, **247**, 38–41.
59. S. Cheatham, M. Kline and Ě. Kupče, *Magn. Reson. Chem.*, 2015, **53**, 363–368.

CHAPTER 4

Application of Residual Dipolar Couplings to the Structural Analysis of Natural Products

R. R. GIL^{*a} AND A. NAVARRO-VÁZQUEZ^{*b}

^a Carnegie Mellon University, Department of Chemistry, 4400 Fifth Avenue, Pittsburgh, PA 15213-3890, USA; ^b Institute of Organic Chemistry and Institute for Biological Interfaces Karlsruhe, Institute of Technology (KIT), Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany

*Email: rgil@andrew.cmu.edu; armando.deus@gmail.com

4.1 Introduction

Nuclear magnetic resonance (NMR) is unquestionably one of the most powerful spectroscopic techniques for analyzing the structural and dynamic properties of molecules in general. The atomic constitution as well as the 3D structure of small molecules can be determined using a combination of three well-established standard NMR parameters observed in the liquid state: (i) chemical shift; (ii) nuclear Overhauser enhancement (or effect; NOE); and (iii) scalar coupling constants (J ; homonuclear and heteronuclear).¹ Chemical shifts are used to determine the identities of the atoms as well as their environment within the molecule. Proton–proton distances can be obtained using through-space spin–spin interactions observed *via* NOE experiments.² NOE has been extensively used in the configurational and conformational analysis of organic molecules since the pioneering work by Anet and Bourn dating from 1965.³ Proton–proton and proton–carbon torsional angles can be determined *via* the analysis of three-bond scalar

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

coupling constants ($^3J_{\text{HH}}$ and $^3J_{\text{HC}}$), which strongly depend on the dihedral angle between the three bonds in consideration, as established by the Karplus equation.^{4,5} An empirical generalization of the Karplus equation that accounts for the electronegativity of the substituents, developed by Altona and coworkers,⁶ is currently in wide use for obtaining dihedral angle information in small molecules. In addition, one-bond (1J) and long-range (2J and 3J) heteronuclear scalar couplings (*e.g.* ^1H – ^{13}C) are routinely used to gather additional connectivity information between protons and carbons. However, although extremely powerful, conventional NMR experiments cannot always provide a unique structural solution for the chemist, since in particular cases the technique presents limitations. One of its major shortcomings is the lack of ability to distinguish symmetry. Symmetric molecules will show fewer resonances than expected and could be half, a quarter, *etc.*, depending on the type of molecular symmetry. In symmetrical molecules, the J -coupling between atoms with the same chemical environment will vanish, preventing its use for structural purposes. This situation is only valid in conventional NMR, when we only observe correlations between protons attached to ^{12}C atoms, such as ^1H – ^{12}C – X – ^{12}C – ^1H , where X can be any molecular linker or nothing. These molecular species represent 99% of the sample. However, it is still possible to see NOEs between protons with identical chemical environments by looking at the species ^1H – ^{13}C – X – ^{12}C – ^1H , where the chemical shift degeneracy between protons is broken by the $^1J_{\text{CH}}$ coupling between the ^1H and the ^{13}C atoms. This is possible by observing the cross-correlation peak between the proton attached to the ^{12}C atom with the ^{13}C atom in a ^{13}C -coupled HSQC-NOESY experiment.^{7–9} This experiment was not extensively used due to its poor sensitivity, since the NOE is already a fraction (<10%) of the low molar concentration of the isotopomeric species in the sample (1%), providing a signal that represents <0.1% of the proton signal in the spectrum. In addition, this limitation due to symmetry extends to molecular entities presenting enantiomorphous conformations (*e.g.* mirror image enantiomers), which cannot be distinguished by NMR unless chiral auxiliary agents are added to the solution. It is important for the reader to bear in mind that NMR parameters (chemical shift, scalar, and dipolar couplings) are insensitive to inversion or reflection of the atomic coordinates, making it impossible to determine absolute molecular configuration without the aid of chiral auxiliary agents.

Another structural limitation is that 3J -coupling constants cannot distinguish positive from negative dihedral angles due to the scalar coupling dependence with the cosine of this angle. Finally, the $1/r^6$ dependence of NOE sets a limitation on the distance information between a pair of protons. For macromolecules, where the rotational diffusion correlation time τ_c is long, it is possible to accurately derive distances between protons that are no more than ~ 5 Å apart, and the accuracy of the NOE-derived distance is affected by spin diffusion for a pair of protons that are further apart.^{10,11} Only in perdeuterated protein samples was it possible to obtain accurate NOE-derived distances for pairs of methyl groups located 12 Å apart.¹² These

considerations limit the application of NOE studies to the local environment. The scenario is even worse for small molecules since they tumble in a fast motion regime (extreme narrowing limit), where their very short correlation times, τ_c , leads to a maximum NOE enhancement of only +50% compared to -100% for macromolecules.² This fact limits the applications of NOE in small molecules to much shorter distances between protons; the determination of the relative configuration between remotely located stereocenters becomes problematic. In addition, different stereoisomeric compounds can sometimes equally yield the same set of 3J and NOE values, leading to ambiguous solutions to the structural problem. Consequently, despite being extremely powerful NMR “tools”, NOE and 3J -coupling constants provide structural information of local character, and these two NMR parameters cannot be used to establish correlations between nuclei that are remotely located one from another. If the chain of short-range interactions is “interrupted” by, for example, a magnetically inactive linker, there is no way to determine the relative configuration of the two disconnected molecular fragments by conventional NMR.

For example, if we want to determine the relative orientation of the hydroxyl groups at positions C-3 and C-17 in compound **1** (Figure 4.1), it is possible to do so *via* a relay of NOE and 3J -coupling correlations from H-3 all the way up to H-17. A relay of short-range correlations (local character) clearly permits, in this case, the determination of the relative configuration of two stereocenters remotely located from each other. However, if we

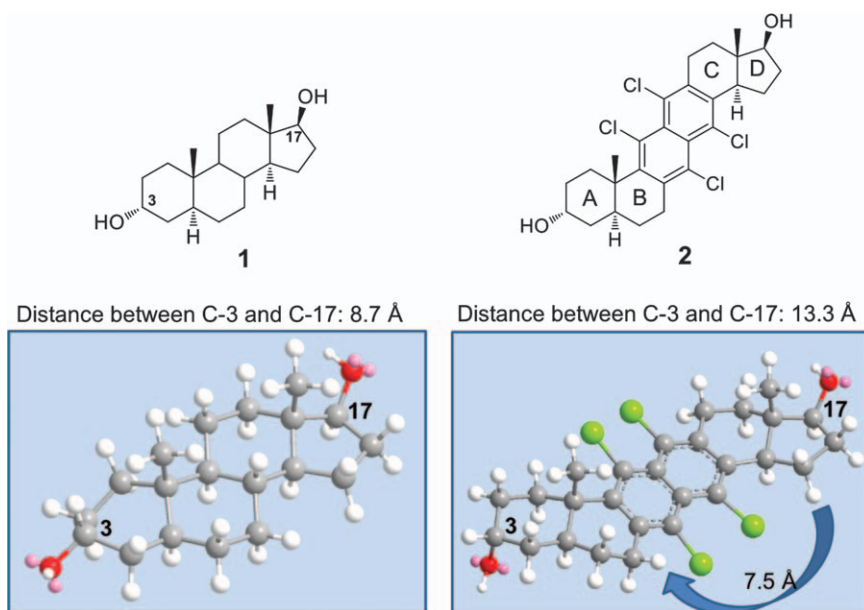


Figure 4.1 The presence of a magnetically inactive linker breaks the chain of NOE and 3J correlations.

interrupt the transfer of NOE and 3J information between rings B and C of the steroidal skeleton, by inserting a hypothetical non-protonated linker as shown in compound 2, the relay of short-range correlations is broken at that point and there is no longer a way to determine the relative orientation of the two OH groups by means of conventional NMR experiments. The distance between point A and B in compound 2 is 13.3 Å and the closest proton–proton interaction is located 7.5 Å away, as shown in the 3D structure of 2 in Figure 4.1. A distance of 7.5 Å is beyond the limit to observe proton–proton NOEs in small molecules.

This situation is not uncommon and the frequency of occurrence is difficult to assess, since unsolved structures will be “waiting” in the freezer of the laboratory for a better methodology that permits their identification. Single-crystal X-ray diffraction analysis is, without question, the most powerful tool to establish both the relative and absolute configuration of a given molecule. However, suitable single crystals for X-ray analysis may not always be available. A recent article has reported a new X-ray analysis method on the nanogram to microgram scale using porous complexes rather than a single crystal for the structural analysis of the compound. Although a very ingenious idea, its applicability to the analysis of complex small molecules is yet to be demonstrated.¹³

With the type of structural problem just used as an example, we need an NMR parameter of non-local character that allows us to determine the relative orientation of bonds regardless of the distance between them. As shown in the equation in Figure 4.2, the direct spin–spin coupling D_{ij} (through space) between two nuclei (bonded or non-bonded) is the NMR parameter that meets these requirements.[†] The dipolar coupling depends on the distance (r_{ij}) between the coupled pair of nuclei as well as on the angle (θ) between the internuclear vector ($\vec{r}_{ij}$) and the applied magnetic field (B_0).¹⁴ Notice that the magnetic field represents in the equation a global axis of reference because the spatial orientation of a given internuclear vector with respect to it is encoded in the value of the dipolar coupling D_{ij} . It is evident from the equation in Figure 4.2 that in the hypothetical case of having a single molecule adopting a given orientation with respect to the magnetic field, it will show a unique combination of dipolar couplings as a result of the intrinsic spatial arrangement of its atoms. This intrinsic spatial arrangement of the atoms in a molecule, by definition, represents the molecular configuration and its preferred conformation. Dipolar couplings are therefore an NMR phenomenon that encodes very powerful structural information of a non-local character. It does not matter how far two spins are

[†]The expression of the dipolar coupling we employ here follows that given by Kramer *et al.*¹⁴ and corresponds to a definition of the weak dipolar coupling spin Hamiltonian as:

$$H_D = 2\pi D \hat{I}_z \hat{S}_z$$

and therefore the contribution of the dipolar coupling to the total splitting T is given by $T = J + D$. If the Hamiltonian is defined as $H_D = 2\pi D' 2I_z \hat{S}_z$, the total splitting T is given by $T = J + 2D'$.

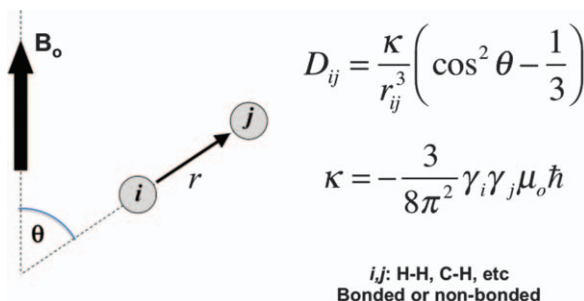


Figure 4.2 Dipolar coupling (D_{ij}) for any bonded or non-bonded pair spins i, j in a static molecule (e.g. solid state) as a function of the distance (r_{ij}) and the angle (θ) of the internuclear vector ($\vec{r}_{ij}$) with the applied magnetic field (B_o), where μ_o is the vacuum permeability constant, γ is the gyromagnetic constant, and $\hbar$ is the Planck constant divided by 2π .

from each other, the dipolar coupling will always encode their relative orientation with respect to the applied magnetic field (B_o), and hence the relative orientation between them.

The application of dipolar coupling in order to determine the relative orientation of pairs of spins (e.g. bonds) explained in this way sounds very straightforward and easy. However, in a real solid powder sample, we do not have only a single molecule, and the molecules in the powder adopt all possible orientations, leading to a broad distribution of D_{ij} values. In addition, to make it even more complex, the dipolar coupling can show maximum values in the order of kilohertz. For example, the maximum value for the term $\cos^2 \theta$ is 1 for $\theta = 0$ or π . Hence, the maximum value of D_{ij} can be derived as:

$$D_{ij}^{\max} = \frac{\kappa}{r_{ij}^3} \left(1 - \frac{1}{3} \right) = \frac{2}{3} \frac{\kappa}{r_{ij}^3} \quad (4.1)$$

For the spin pairs $^1\text{H}-^1\text{H}$, $^{13}\text{C}-^1\text{H}$, and $^{15}\text{N}-^1\text{H}$, the values for κ (in kHz $\text{\AA}^3$) are -360.3 , -90.6 , and 36.5 , respectively. For example, an aliphatic C-H single bond of length (r) $1.10 \text{ \AA}$ will show a maximum D_{ij} value of -45.38 kHz . Since κ is positive for $^1\text{H}-^1\text{H}$ and $^{13}\text{C}-^1\text{H}$, their D_{ij} values will be negative if the internuclear vector is parallel to the magnetic field ($\theta = 0$ or π) and positive if it is perpendicular to the magnetic field ($\theta = \pi/2$ or $3\pi/2$). From the equation in Figure 4.2, it can also be derived that the minimum D_{ij} will be:

$$D_{ij}^{\min} = \frac{\kappa}{r_{ij}^3} \left(0 - \frac{1}{3} \right) = -\frac{1}{3} \frac{\kappa}{r_{ij}^3} \Rightarrow \frac{D_{ij}^{\max}}{D_{ij}^{\min}} = -2 \quad (4.2)$$

In the case of the C-H bond, the D_{ij} minimum will be $-(-45.38/2) = +22.69 \text{ kHz}$. All of these complications make the extraction of dipolar couplings from solid-state NMR spectra an almost impossible task.

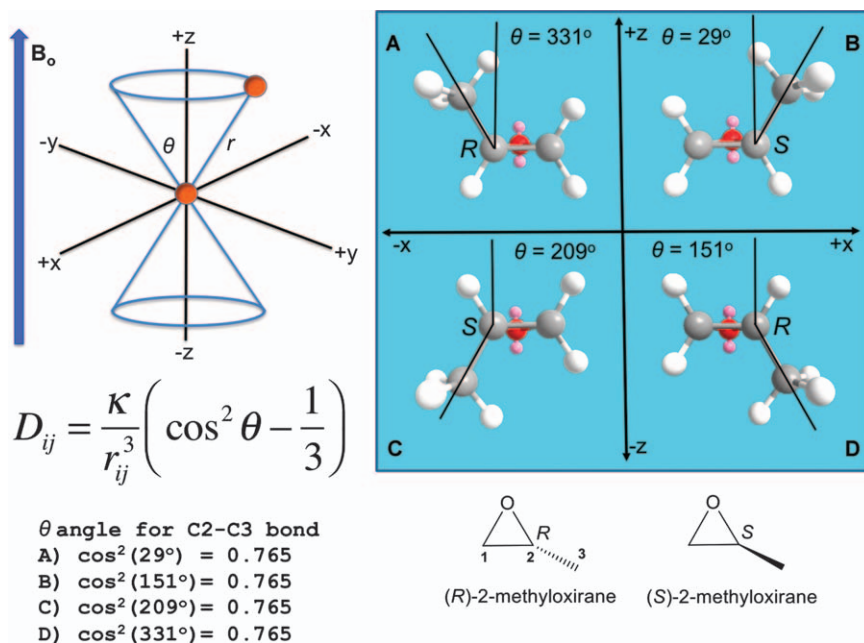


Figure 4.3 The dipolar coupling is insensitive to inversion or plane reflection symmetry operations.

Before moving on to liquid-state NMR spectroscopy, another important aspect of dipolar coupling (D_{ij}) needs to be highlighted. Due to its dependence on the $\cos^2$ of the angle θ , in a 3D representation of an internuclear vector, as we can see in Figure 4.3, the two solid angles of the θ value, reflected in the x, y plane, satisfy the same D_{ij} value. In other words, D_{ij} is insensitive to a reflection and/or inversion of the 3D molecular coordinate. This intrinsic property of the dipolar coupling makes it insensitive to symmetry, hence preventing its application to the assignment of absolute configuration. Two enantiomorphous conformations (mirror images), such as the example of the enantiomers *R* and *S* of the rigid small molecule 2-methyloxirane shown in Figure 4.3, clearly illustrate this limitation of dipolar couplings. The figure shows four structures that are to each other either a reflection on the x, y (A, C and B, D) and y, z (A, B and C, D) planes, as well as inversions through the origin of the system coordinates (A, D and B, C). In each case, the angle θ for the carbon-carbon bond C2-C3, which represents the relative orientation of the methyl group with respect to the magnetic field (B_0), adopts values of 331° , 29° , 209° , and 151° for A, B, C, and D, respectively. As shown in the lower left corner of Figure 4.3, the value of $\cos^2 \theta$ for these four orientations is the same (0.765), leading to the same D_{ij} value for the internuclear vector C2-C3 in each case.

In isotropic solution (composed of solute and solvent only), small molecules tumble at correlation times (τ_c) that are much shorter than the reciprocal of the dipolar couplings, and during these motions, the internuclear vectors adopt all possible orientations with respect to the magnetic field (B_0) with equal probability. As a result, the dipolar coupling averages out, and it is not observed in conventional solution NMR experiments. However, these time-dependent dipole–dipole interactions between spins are responsible for inducing spin relaxation as well as generating NOEs between nuclei that are close enough in space. Differently from the solid state, in solution, the term $\cos^2 \theta$, as well as the dipolar coupling D_{ij} , are time dependent. D_{ij} is observed as the time average value $\bar{D}_{ij}$.

$$\bar{D}_{ij} = \frac{\kappa}{r_{ij}^3} \left(\overline{\cos^2 \theta} - \frac{1}{3} \right) \quad (4.3)$$

For one-bond couplings, the distance r_{ij} between nuclei will commonly change in the timescale of stretching vibrations and therefore it can be substituted by its vibrationally averaged value (eqn (4.4)). Note that this may no longer be the case for long-range coupling where conformational changes may dramatically change the distance between nuclei.

$$\bar{D}_{ij} = \frac{\kappa}{r_{ij}^3} \left(\overline{\cos^2 \theta} - \frac{1}{3} \right) \quad (4.4)$$

The term $\overline{\cos^2 \theta}$ can be represented in terms of the probability matrix (or probability tensor) $\hat{P}$, whose matrix components are the time average of the external field components with respect to a fixed molecular frame (*i.e.* $\hat{P}$ represents the probability of the external field pointing in a particular direction of space). It is nicely and clearly demonstrated by Glaser and coworkers¹⁴ that:

$$\overline{\cos^2 \theta} = \vec{r}^T \hat{P} \vec{r} = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (4.5)$$

where x , y , and z are the 3D coordinates of the unit vector r_{ij} in an arbitrary molecular frame. For simplicity, we can say that $\hat{P}$ represents in a certain way the distribution probability of the internuclear vector ($\vec{r}$) with respect to the magnetic field B_0 . The matrix $\hat{P}$ is real, symmetric, and has a trace of 1 ($P_{xx} + P_{yy} + P_{zz} = 1$).¹⁴ Under isotropic conditions, the off-diagonal terms of $\hat{P}$ are zero, and each of the principal components of P_{xx} , P_{yy} , and P_{zz} adopt a value of 1/3, meaning that there is an equal probability distribution of $\vec{r}$ in all of the spatial directions. Simply restated, as the molecule tumbles in solution, a third of the time B_0 will be aligned along the x -axis of the molecular frame, another third of the time B_0 will be aligned along the y -axis,

and the final third of the time B_0 will be aligned along the z-axis of the arbitrary molecular frame:

$$\hat{P} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \quad (4.6)$$

If we insert the expansion of eqn (4.5) in eqn (4.4) using the isotropic value of $\hat{P}$ as shown in eqn (4.6), it is clearly shown that the dipolar coupling value vanishes since subtracting a third of the isotropic tensor $\hat{P}$ results in a null value, as shown below:

$$\bar{D}_{ij} = \frac{\kappa}{r_{ij}^3} \left(\vec{r}^T \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \vec{r} - \frac{1}{3} \right) = 0 \quad (4.7)$$

However, if we force the molecules to adopt a minor degree of alignment (or orientation) by exposing them to an anisotropic medium, a measurable fraction of the maximum value of dipolar coupling D_{ij} (0.01–0.1%) can be observed in the NMR spectrum. This fraction of the dipolar coupling is called the residual dipolar coupling (RDC). RDCs maintain the same angular information as the original dipolar coupling and, as stated above, this information is of non-local character.

The NMR literature uses the tensor $\hat{A}$,¹⁴ known as the alignment tensor, instead of $\hat{P}$, which is defined as:

$$\hat{A} = \hat{P} - \frac{1}{3} \hat{1} \quad (4.8)$$

where $\hat{1}$ is a unitary diagonal tensor, or the Saupe matrix $\hat{S} = \frac{3}{2} \hat{A}$. Hence, the matrix $\hat{S}$ corresponds to an angular part $\frac{3 \cos^2 \theta - 1}{2}$.

Since $\overline{\cos^2 \theta} = \vec{r}^T \hat{P} \vec{r}$ (eqn (4.3)), it is easy to show that $\overline{\cos^2 \theta} - 1/3 = \vec{r}^T \hat{A} \vec{r}$,¹⁴ hence the dipolar coupling D_{ij} can be expressed in a general equation in tensorial form as follows:

$$D_{ij} = \frac{\kappa}{r_{ij}^3} \vec{r}_{ij}^T \hat{A} \vec{r}_{ij} \quad (4.9)$$

The use of the alignment tensor instead of the probability makes more physical sense since for isotropic molecular tumbling $\hat{A} = 0$ (no alignment). This can be easily shown if we expand eqn (4.8) using the isotropic value for the probability tensor $\hat{P}$ as follows:

$$\hat{A} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 0 \quad (4.10)$$

clearly showing that D_{ij} vanishes in isotropic solution. The beauty of eqn (4.9) is that all of the possible motion regimes of a molecule can be encoded in the tensor $\hat{A}$, from solid state (no motion) to isotropic solution, going through all of the possible alignment (or orientational) probabilities that a molecule can adopt in an anisotropic medium.

If the molecule stops moving (solid state), by definition there is a maximum probability (1) of the magnetic field B_0 being aligned along the z -axis of the molecular frame, while no B_0 is present along the x - and y -axes ($P_{xx} = P_{yy} = 0$; $P_{zz} = 1$), hence the alignment tensor will have the form:

$$\hat{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix} \quad (4.11)$$

For example, a unit internuclear vector aligned along the z -axis (parallel to B_0) as shown in Figure 4.4A can be represented in the molecular frame by the following matrices:

$$\vec{r} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \quad \vec{r}^T = (0 \quad 0 \quad 1) \quad (4.12)$$

If we plug eqn (4.11) and (4.12) into eqn (4.9) to calculate the D_{ij} value of an internuclear vector parallel to the magnetic field B_0 , we have:

$$D_{ij} = \frac{\kappa}{r_{ij}^3} (0 \quad 0 \quad 1) \begin{pmatrix} -1/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{2}{3} \frac{\kappa}{r_{ij}^3} \quad (4.13)$$

The result is identical to eqn (4.1), which represents the maximum value for the dipolar coupling when the internuclear vector is parallel to B_0 .

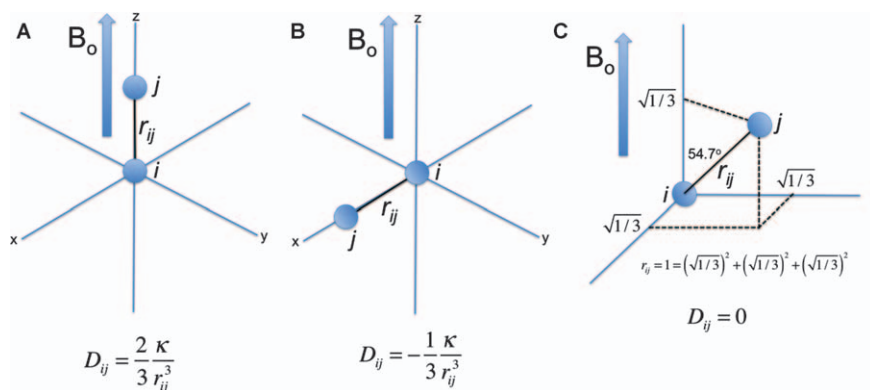


Figure 4.4 Dipolar coupling for an internuclear vector lying in parallel, perpendicular and magic angle orientations with respect to the external magnetic field.

An internuclear vector perpendicular to B_0 can be represented by a unit vector along either the x - or the y -axes; the dipolar coupling value for a vector in the x -axis (Figure 4.4B) will be:

$$D_{ij} = \frac{\kappa}{r_{ij}^3} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = -\frac{1}{3} \frac{\kappa}{r_{ij}^3} \quad (4.14)$$

Hence, the result is identical to eqn (4.2). Since κ has a negative sign for positive γ values, the D_{ij} value for a vector perpendicular to B_0 will be positive and of half the value of a vector parallel to B_0 , as already derived in eqn (4.2). Finally, consider what happens if the internuclear vector is located equidistant from each axis (Figure 4.4C). In this case, the coordinate values in x , y , and z will be the same and the angle between the internuclear vector and each axis will be 54.7° . The column vector representing this situation will be:

$$\vec{r}_{ij} = \begin{pmatrix} \sqrt{1/3} \\ \sqrt{1/3} \\ \sqrt{1/3} \end{pmatrix} \quad (4.15)$$

and its dipolar coupling value can be calculated as follows:

$$D_{ij} = \frac{\kappa}{r_{ij}^3} \begin{pmatrix} \sqrt{1/3} & \sqrt{1/3} & \sqrt{1/3} \end{pmatrix} \begin{pmatrix} -1/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix} \begin{pmatrix} \sqrt{1/3} \\ \sqrt{1/3} \\ \sqrt{1/3} \end{pmatrix} = 0 \quad (4.16)$$

Hence, the dipolar coupling in solid-state NMR vanishes at an angle of 54.7° between B_0 and the internuclear vector. This angle is known as the “magic angle” because it makes D_{ij} “disappear”. In solid-state NMR, if the sample is spun at high speed and the axis of sample rotation is set at 54.7° from B_0 , the broadening due to the unwanted dipolar couplings from all of the possible orientations of the molecules in the solid powder is “magically” removed from the spectrum. This technique is known as magic angle spinning.

The alignment tensor is diagonal when represented in the laboratory frame, a frame where the magnetic field B_0 is aligned along the z -axis. In the cases shown above, the internuclear vector was represented in the same coordinate system, the laboratory frame. However, if the molecule is represented by x , y , z coordinates belonging to a different frame (an arbitrary molecular frame), the relative orientation between the laboratory frame and the molecular frame is encoded in the non-diagonal terms of the alignment tensor $\hat{A}$. In practice, when the molecule is exposed to an alignment medium (anisotropic medium), the orientation of the molecule is not known *a priori* and the alignment tensor needs to be determined experimentally. In fact,

this is the *key step* in the use of RDCs for structural analysis. Once the alignment tensor is known, the relative orientation of any internuclear vector in the molecule can be determined provided that experimental RDC values are available for that particular internuclear vector. In practice, the most commonly used internuclear vector in small organic molecules is the C–H bond, since its RDC ($^1D_{\text{CH}}$) is relatively easy to measure, as we will discuss later. Long-range C–H RDCs ($^2D_{\text{CH}}$, $^3D_{\text{CH}}$) can be also measured, but it is more challenging to measure them with the desired accuracy since they are one order of magnitude smaller than the $^1D_{\text{CH}}$. In addition, H–H internuclear vectors are also used, with the H–H geminal RDCs ($^2D_{\text{HH}}$) being a popular option. The use of one-bond carbon–carbon RDCs ($^1D_{\text{CC}}$) has also been reported,^{15,16} but their application is limited by the very low sensitivity of their measurement. Access to cryogenic NMR probes (see Chapters 3 and 4 in the first volume of this book) and the use of inverse-detected J -modulated 1,1-ADEQUATE experiments^{17,18} will likely allow more facile measurement of $^1D_{\text{CC}}$ in the near future.

Before moving on to the application of RDCs in the structural analysis of small organic molecules with a special emphasis on natural products, it is necessary to describe some properties of the probability ($\hat{P}$) and alignment ($\hat{A}$) tensors in order to better understand the application of this methodology.

As mentioned above, the alignment tensor is not known *a priori* and has to be determined experimentally. We already mentioned that the probability tensor:

$$\hat{P} = \begin{pmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{pmatrix} \quad (4.17)$$

is symmetrical since $P_{xy} = P_{yx}$, $P_{xz} = P_{zx}$, and $P_{yz} = P_{zy}$, and that its trace is 1 ($P_{xx} + P_{yy} + P_{zz} = 1$). However, the alignment tensor:

$$\hat{A} = \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} \quad (4.18)$$

which is also symmetrical, is traceless:

$$A_{xx} + A_{yy} + A_{zz} = 0 \Rightarrow A_{xx} + A_{yy} = -A_{zz} \quad (4.19)$$

For both tensors, due to their symmetry and the fact that one of the diagonal terms can be calculated from the other two, the number of unknown variables is reduced from nine to five. Hence, in order to determine the alignment tensor of any molecule exposed to an anisotropic environment, we need at least five equations (eqn (4.9)) with at least five independent D_{ij} values. By independent we mean D_{ij} values from non-parallel internuclear vectors with the condition that no more than three vectors lie in the plane.¹⁹ It is important to note that parallel vectors will have the same angle θ with

respect to the magnetic field B_0 , regardless of the orientation of the molecule, and will provide the same angular information. For example, menthol does not have enough independent C–H vectors to determine the alignment tensor. As shown in Figure 4.5, where the different types of CH bonds in menthol (including the fast-rotating CH_3) are color coded, the same color means that those bonds are parallel. Only four sets of linearly independent one-bond CH RDC values can be measured, marked in green, pink, red, and yellow. Hence, its alignment tensor cannot be calculated from these experimental data. Although there are a total of 13 possible RDC values to be measured, the resulting matrix for solving the equations will be ill conditioned.

Another important point is that the structural information encoded in the set of more than five independent RDC values is only limited to the relative orientation of their corresponding internuclear vectors, but the information about the relative positions of these vectors in the molecule is not available. In other words, the distance between the internuclear vectors is not encoded in the RDCs. Hence, *ab initio* determination of a 3D structure is not possible solely from RDCs, regardless of the number of D_{ij} values available. However, this is not a problem since the 2D structure of small molecules can be determined using the molecular formula obtained from high-resolution mass spectrometry in combination with a series of multinuclear and multi-dimensional NMR experiments (*i.e.* COSY, TOCSY, NOESY, HSQC, HMBC, HSQC-TOCSY, ADEQUATE, INADEQUATE, *etc.*). At this point, the reader is perhaps wondering how RDCs are used to determine the constitution, the relative configuration, and/or the preferred conformation of a small molecule. In 1999, Prestegard and coworkers²⁰ proposed the use of singular value decomposition (SVD) to solve a set of D_{ij} equations and hence calculate the alignment tensor. The SVD calculations have proven to be robust and

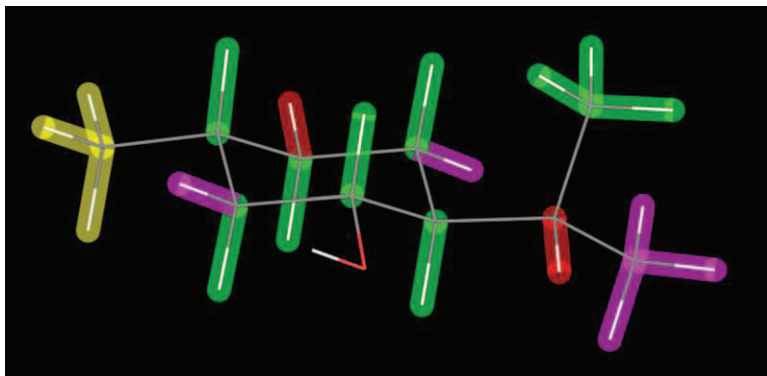


Figure 4.5 3D structure of menthol showing the parallel CH bonds in the same colors. Note that for menthol it is possible to measure only four independent one-bond ^1H – ^{13}C RDCs (D_{CH}) marked in green, pink, red, and yellow.

efficient, even when a very limited set of experimental RDCs is available. Popular programs for RDC data analysis that implement SVD are PALES²¹ or MSpin,²² as will be discussed later. When solving the SVD problem, the dipolar couplings D_{ij} are transformed into their reduced counterparts $d_{ij} = D_{ij} \frac{r_{ij}^3}{\kappa} = \overline{\cos^2 \theta} - \frac{1}{3}$ in order to eliminate dependence on the internuclear distance or the magnetogyric ratios. Therefore, all measured couplings will have the same weight in the SVD treatment, regardless of their nature.

$$d_{ij} = (x \ y \ z) \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{xy} & A_{yy} & A_{yz} \\ A_{xz} & A_{yz} & A_{zz} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (2xy \ 2xz \ y^2 - x^2 \ 2yz \ z^2 - x^2) \begin{pmatrix} A_{xy} \\ A_{xz} \\ A_{yy} \\ A_{yz} \\ A_{zz} \end{pmatrix} \quad (4.20)$$

For a set of N dipolar couplings, we may then set up a set of linear equations:

$$\begin{pmatrix} d_{ij}^1 \\ \vdots \\ d_{ij}^N \end{pmatrix} = \begin{pmatrix} 2x_1y_1 & 2x_1y_1 & y_1^2 - x_1^2 & 2y_1z_1 & z_1^2 - x_1^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2x_Ny_N & 2x_Ny_N & y_N^2 - x_N^2 & 2y_Nz_N & z_N^2 - x_N^2 \end{pmatrix} \begin{pmatrix} A_{xx} \\ A_{xz} \\ A_{yy} \\ A_{yz} \\ A_{zz} \end{pmatrix} \quad (4.21)$$

or, in an abbreviated matrix notation:

$$\hat{d} = \hat{M} \hat{T} \quad (4.22)$$

If five RDCs are available, the linear system can be solved by any suitable method. Note, however, that if some of the internuclear vectors are parallel, they will furnish nearly identical rows in the director cosine matrix and the problem will be ill conditioned (*i.e.* a very small change in the experimental values will cause a large change in the determined tensor). The condition of the problem can be quantified by the condition number κ , which, for a symmetric diagonalizable matrix, is the ratio between the absolute value of the largest and smallest eigenvalues.

In general, more than five RDCs will be available and therefore the system is over-determined and has no solution. However, the best solution in a least-squares sense (*i.e.* the tensor that would furnish the best fitting between experimental and calculated values) can be calculated using the SVD.

The $N \times 5$ matrix $\hat{M}$ can be decomposed as $\hat{M} = \hat{U} \hat{\Sigma} \hat{V}^T$ where $\hat{U}_{N \times N}$ and $\hat{V}_{N \times N}$ contain the left and right singular vectors and $\hat{\Sigma}$ is a $N \times 5$ diagonal matrix that contains the singular values of $\hat{M}$. Thus, the best solution will be given by:

$$\hat{T} = \hat{V} \hat{\Sigma}^{-1} \hat{U}^T \hat{d} \quad (4.23)$$

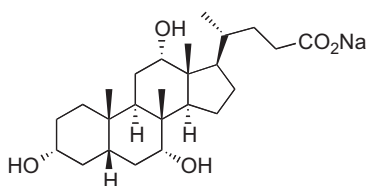
A condition number will be given by the ratio between the largest and smallest singular values $\kappa = \frac{\sigma_{\text{high}}}{\sigma_{\text{low}}}$. As a general rule of thumb, if values of the condition number are above 100, then the problem is considered badly conditioned. Values of a condition number between 30 and 100 may be acceptable, but the acceptability depends on how well the tensor is defined in the region where the stereogenic centers are located.

Note that the least-squares problem can be also solved using optimization algorithms such as Powell or Levenberg–Marquardt.²³ Although computationally less efficient, they may be of use if additional restraints on the solution must be imposed.

4.2 Alignment Media

Before showing a series of examples on how RDCs are measured and used in the structural analysis of natural and synthetic small molecules, the history of aligning small molecules, as well as the current status of alignment media, will be discussed.

The first compound to be aligned in order to perform an NMR experiment was a small molecule. In 1963, Saupe and Englert²⁴ oriented benzene in the nematic phase of *p*-azoxyanisole. The alignment was so strong that H–H RDCs of up to 2000 Hz were observed with line widths ranging from 5 to 15 Hz. The six protons of benzene are chemically equivalent and its isotropic ¹H NMR spectrum shows a single resonance, but unlike the case of *J*-coupling (or indirect coupling), the dipolar coupling does not vanish under chemical equivalence conditions. Two chemically equivalent nuclei with the possibility of exhibiting a dipolar coupling between them will show a signal splitting equal to 3/2-times the dipolar coupling constant $\left(\frac{3}{2}D_{ij}\right)$ (3D if the alternative form of the Hamiltonian is used; see above).²⁵ This is not a common experimental situation, but the reader should be aware of it at the time of extracting the RDC value from the signal splitting. Unfortunately, nematic liquid crystal phases are not of practical application to the structural analysis of small molecules due to the strong degree of alignment induced by them. To successfully analyze small molecules, it is desirable to have a weak degree of alignment with dipolar couplings in the order of 0.01–0.1% of their maximum values in solid state. Still, that first experiment by Saupe and Englert allowed the development of the theory of alignment, as well as the empirical determination of the shielding anisotropy of the benzene ring. Unfortunately, practical applications for the structural elucidation of non-trivial organic compounds had to wait nearly 40 years. In 2003, Mangoni *et al.*²⁶ and Yan *et al.*²⁷ almost simultaneously reported the first application of RDCs to the configurational analysis of small molecules. Mangoni *et al.* aligned sodium cholate (3)²⁶ in a strained polyacrylamide cross-linked gel and determined the configuration of the three hydroxylated stereogenic centers.



3

In the meantime, since the first report by Tjandra and Bax in 1997,²⁸ RDCs have been extensively used in studies of the dynamics and structures of proteins and nucleic acids, and a vast number of alignment media compatible with water were developed. However, this topic is beyond the scope of this chapter, but interested readers can consult many excellent review articles on the subject.^{11,29–33} The major impediment to aligning small organic molecules for structural analysis was initially the lack of alignment media compatible with organic solvents. After being successfully used by Courtieu and coworkers for the NMR analysis of enantiomeric mixtures,³⁴ three research groups introduced in 2003 (almost simultaneously) the use of CDCl_3 lyotropic liquid crystalline (LLC) solution of poly- γ -benzyl-L-glutamate (PBLG) to align small organic molecules for the determination of relative configuration.^{15,35,36} The large magnetic susceptibility anisotropy of these mesophases causes them to spontaneously align in the presence of an intense external magnetic field. This alignment is then partially transmitted to the solvent and the molecules in solution. Other homopolypeptides have been further developed as orienting media in CDCl_3 , such as poly- γ -benzyl-D-glutamate (PBDG), poly- γ -ethyl-L-glutamate (PELG), and poly- ϵ -carbobenzyl-L-lysine (PCBLL).³⁷ PBLG also forms LLC phases in DMF, as well as in mixtures with up to 50% of $\text{DMSO}-d_6$ in CDCl_3 .³⁸ More recently, LLC phases in CDCl_3 based on polyguanidines and polyacetylenes have been reported.^{39,40} All of these LLC phases form helical secondary structures with macrohomochirality. In addition to acting as orienting media, they also show enantiodiscriminatory power. The L and D versions of PBLG and PBDG form secondary structures of opposite helicity, affording enantiomeric chiral aligning phases.⁴¹ The LLC phases have a limited lower degree of alignment, and this lower limit is still high compared to the strained aligning gels (SAGs) that will be discussed below, leading to broader signals. Efforts were reported by Thiele's group to improve the resolution of the spectra by using high-molecular-weight PBLG/PBDG.⁴² The use of these aligning media is, however, limited by the fact that they are very expensive and sample preparation can be tedious. Other helical lyotropic phases have recently been reported. Polyguanidines show an even stronger degree of alignment, but the recently reported polyacetylene LLC phases have shown an excellent degree of alignment, with $^1D_{\text{CH}}$ values within a range of -10 to $+15$ Hz.^{40,43} In addition, polyacetylenes show NMR spectra with fairly narrow signals and pronounced enantiodiscriminatory capabilities. Also, helical phases based

on polyisocyanides have recently been reported⁴⁴ to provide a fair degree of alignment. These developments open new and promising avenues in lieu of the determination of absolute configuration by RDCs, but no significant progress has yet been made in achieving this goal. There are two publications that created false expectations by claiming the assignment of absolute configuration by RDCs.^{45,46} These two reports led a group of experts in the field to write an article that explains the reasons why it is not yet possible to determine absolute configuration even when using chiral alignment media such as the LLC chiral phases described above.⁴⁷ We have already shown above that the D_{ij} value provided by the equation shown in Figure 4.2 is insensitive to either an inversion or a reflection of the molecular 3D geometry. Two enantiomorphous (mirror image) rigid molecules such as the *R* and *S* pair of 2-methyloxirane (see Figure 4.2) can be described by two sets of x, y, z 3D coordinates with opposite signs. In other words, every atom coordinate (x, y, z) is replaced by the coordinate $(-x, -y, -z)$, and the coordinate vectors for *R* and *S* isomers are related according to the following equation:

$$\vec{r}_{Sij} = -\vec{r}_{Rij} \quad (4.24)$$

The *R* and *S* pair of 2-methyloxirane will show the same alignment tensor in a non-chiral orienting media, and the following condition holds:

$$\hat{A}_R = \hat{A}_S \quad (4.25)$$

Hence, since both enantiomers show the same alignment tensor in a non-chiral alignment medium, replacing r_{ij} with $-r_{ij}$ in eqn (4.9) (tensorial form) does not change the value of D_{ij} . In other words, a set of RDCs equally satisfies two enantiomorphous structures (mirror images) and, consequently, only the relative configuration of the molecule can be determined by using RDCs. In contrast, if these two structures are oriented in a chiral, non-racemic alignment medium, due to the fact that both enantiomers may interact differently with the chiral phase (forming diastereomorphous associations), the following condition holds:

$$\hat{A}_R \neq \hat{A}_S \quad (4.26)$$

This means that each enantiomer will show a different set of D_{ij} values. Though each enantiomer aligns differently in the chiral, non-racemic orienting medium, there is no way of experimentally knowing which set corresponds with which enantiomer. Both sets will equally satisfy r_{ij} and $-r_{ij}$ as a solution for eqn (4.9). Consequently, there is no way to determine the absolute configuration of the molecule. However, if the nature of the interaction between the chiral phase and each enantiomer could be predicted, then the determination of absolute configuration is possible. Theoretical prediction of the alignment tensor under chiral, non-racemic conditions has not yet been achieved, and this is an area of research with significant room for exploration. When the alignment tensor of enantiomers in chiral phases can be accurately predicted in the same way that circular dichroism or

vibrational circular dichroism spectra are now predicted using density functional theory (DFT) calculations, the determination of absolute configuration by RDCs will be very straightforward. As stated in the recent critical article,⁴⁷ the two articles that have claimed the assignment of absolute configuration “are either technically unconvincing⁴⁵ or contradict the principles of physics.”⁴⁶ Many chiral, non-racemic aligning phases have been developed thus far, but we will have to wait for the development of computational methods to predict the alignment tensors of enantiomorphous conformations.

Thus far, only PBLG, PELG, and PCBL are commercially available. It is expected that high-molecular-weight PBLG and PBDG will become commercially available in the near future.⁴⁸ An aligning helical β -peptide phase has also been reported with a very good degree of alignment.⁴⁹ In addition, some of the self-orienting media employed for biomolecule alignment have been used for small-molecule elucidation, such as for cetylpyridinium halogenides,⁵⁰ which were employed for alignment of protonated azacycles,⁵¹ or disodium cromoglycate,^{52,53} which was employed in the alignment of cyclodextrins.⁵⁴ Pentaethylene glycol monodecyl ether has also been reported as an alignment medium for DMSO/water mixtures.⁵⁵

Another group of more widely used alignment media are the SAGs. The SAG methodology was originally proposed as part of the development of orienting media for biomolecules in water, independently and simultaneously, by Tycko *et al.*⁵⁶ and Grzesiek and coworkers.⁵⁷ The anisotropy was mechanically generated either by compressing or stretching the gels, based on a method of analyzing the elasticity of elastomers published by Deloche and Samulski in 1981.⁵⁸ Based on this concept, Luy *et al.*⁵⁹ proposed in 2004 an ingenious idea consisting of swelling a polystyrene gel stick 10 mm long and 3 mm diameter inside a 5 mm NMR tube with CDCl_3 . The gel swells in all directions until the cylinder touches the wall of the NMR tube. Once the gel touched the wall of the NMR tube, it continues swelling, but only in the axial direction of the tube. This radially constricted swelling is the equivalent of mechanically stretching the gel in the vertical direction, hence creating axial anisotropy, as shown in Figure 4.6.

The orienting properties of this gel were demonstrated by aligning strychnine and menthol. Luy *et al.* have also shown that the degree of alignment increases as the cross-link density of the gel increases, with divinylbenzene as the cross-linking agent in this case.⁶⁰ It soon became evident that any swollen gel would show the same orienting properties, and further developments were aimed at using them with other organic solvents. Most of the development of aligning gels was done by the Luy group,^{61–69} including a stretching apparatus that is compatible with organic solvents,⁷⁰ similar to the one developed by Kuchel *et al.*⁷¹ for rapid and reversible stretching of gelatin (in water) in order to discriminate enantiomers using RDCs and ^2H quadrupolar splitting. The Griesinger group has developed a polyacrylamide gel (PH gel)⁷² that is compatible with water and $\text{DMSO-}d_6$.

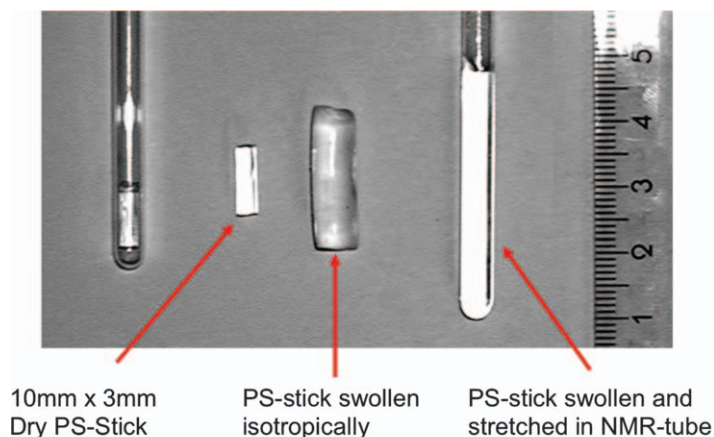


Figure 4.6 A dry polystyrene (PS) stick (3 mm OD, 10 mm long) swells inside a 5 mm NMR tube and is mechanically stretched in the vertical direction. Picture reproduced with permission from Luy *et al.*⁵⁹ Copyright © 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Table 4.1 lists the gel/solvent combinations that are available thus far to orient small molecules. For the traditional method proposed by Luy *et al.* in 2004, the sample can be dissolved in the solvent that will be used to swell the gel, and the sample then penetrates inside the gel as the gel swells, or once the gel is swollen and stabilized, the sample can be allowed to slowly diffuse into the gel. From a technical standpoint, this is a lengthy procedure, and this is the principal bottleneck in the duration of the analysis. Gels can take several days to swell and stabilize. In the case of poly(methylmethacrylate) (PMMA) gels swollen in CDCl_3 , this procedure can take 20–30 days, and the sample takes 2–3 days to diffuse inside the gel.⁷³

Apart from the dipolar coupling, the quadrupolar coupling (Q) and the chemical shift (δ) are anisotropic NMR parameters that also depend on the orientation of molecules with respect to the magnetic field. In isotropic solution, the chemical shift is observed as the average value of the three principal components of the chemical shift tensor, a topic that will not be discussed here. However, in orienting media, the observed chemical shift is a result of a weighted average determined by the molecular alignment tensor. The difference between this weighted average and the isotropic chemical shift value is known as residual chemical shift anisotropy (RCSA). RCSA is not very significant in weakly aligned molecules, but it should not surprise us if we see it in NMR spectra collected in anisotropic conditions. Unlike RDCs, RCSA is strongly dependent on the electronic structure of the molecule. Its measurement is very challenging and the application of it to structure problems in small molecules has not been very successful thus far.⁷⁷ The other anisotropic NMR parameter, the quadrupolar coupling (Q), also averages to zero due to isotropic molecular tumbling, but it is not always

Table 4.1 Self-aligned media.

	Solvent	Chiral	Remarks
Poly- γ -benzyl-L-glutamate (PBLG) ³⁷	CDCl ₃ , CD ₂ Cl ₂ , DMF, THF, dioxane	Yes	Commercial
Poly- γ -ethyl-L-glutamate (PELG) ³⁷	CDCl ₃ , CD ₂ Cl ₂	Yes	Commercial. Lower degree of alignment than PELG
Poly- ϵ -carbobenzoyloxy-L-lysine (PCBLL) ³⁷	CDCl ₃ , CD ₂ Cl ₂	Yes	
Polyguanidines ³⁹	CDCl ₃	Yes	
Helical β -peptides ⁴⁹	CDCl ₃	Yes	
Polyacetylenes ⁴⁰	Water	Yes	
Polyisocyanides ⁷⁴	CDCl ₃	Yes	
Sodium cromoglycate ⁵²	Water	No	Commercial
Cetylpyridinium chloride ⁵⁰	Water	No	Prepared from commercial sources
Pentaethylene glycol monodecyl ether (C12E5) ⁷⁵	DMSO/water mixture		Prepared from commercial sources
Polyacrylamide ^{56,57}	Water	No	
Dimethylacrylamide copolymers (PH.PDMAA) ⁷²	Water/DMSO	No	
PMMA ^{73,76}	CDCl ₃ , CDCl ₂		
Polyacrylonitrile (PAN) ^{65,68}	DMSO	No	
Polyvinylacetate (PVAc) ⁶²	CDCl ₃ , CD ₂ Cl ₂ , CD ₃ OD, CD ₃ CN, DMSO, DMF, C ₆ D ₆ , acetone, THF, AcOEt, dioxane	No	
Polystyrene (PS) ⁵⁹	CDCl ₃ , CD ₂ Cl ₂ , THF, C ₆ D ₆ , dioxane	No	
Polydimethylsilylone (PDMS) ⁶¹	CDCl ₃ , CD ₂ Cl ₂ , THF, C ₆ D ₆ , dioxane, <i>n</i> -hexane	No	Clean spectral window

necessarily present in anisotropic conditions. Quadrupolar nuclei have a spin quantum number greater than 1/2 and present a nuclear electric quadrupole moment (eQ). The eQ is a result of a non-spherically symmetric charge distribution in the nucleus. The shape of quadrupolar nuclei can be either prolate (like an American football) or oblate (flattened on the poles like the shape of our planet, Earth). As we all know, nuclei with spin-1/2 (¹H, ¹³C, ³¹P, ¹⁹F, and ¹⁵N) split into only two energy levels under the influence of a magnetic field B_0 and, regardless of the orientation of the molecules, their dipole moments always stay aligned (or locked) in the direction of B_0 . This is in fact the key for nuclei with spin-1/2 to have

excellent relaxation properties and to show sharp NMR signals, as well as to show signal splitting due to J -coupling with other nuclei. However, the eQ property of the quadrupolar nucleus locks to the electric field gradient (EFG) property of the molecule and follows the EFG as the molecules change direction during molecular tumbling in solution. In highly symmetric molecules, the EFG vanishes and the quadrupolar nuclei behave similarly to nuclei with spin-1/2. This is a phenomenon that is commonly observed in molecules containing ^{14}N , where sharp ^{14}N NMR lines and the coupling of ^{14}N to other nuclei is observed only if the ^{14}N -containing molecules are symmetrical and show zero EFG. Quadrupolar nuclei split into more than two energy levels in the presence of B_0 based on their spin quantum number. For example, deuterium (^2H) is quadrupolar, has spin-1, and splits into three energy levels (+1, 0, -1). The two energy transitions are degenerate and, in principle, a single NMR signal should be observed for ^2H . However, in non-symmetric molecular environments, the interaction of the eQ with the EFG breaks the degeneracy of the energy transition, leading to the observation of a signal splitting, known as the quadrupolar interaction or quadrupolar splitting (Q -splitting). The value of the Q -splitting depends on the size of the EFG, the size of the quadrupole moment, and the relative orientation of the EFG with respect to the magnetic field B_0 . As with dipolar coupling, the Q -splitting averages out in isotropic solution due to isotropic molecular tumbling, but its effect can be observed by solid-state NMR. This physical property of deuterium makes it a tool of great diagnostic value for probing the degree of anisotropy generated by an orienting medium. Deuterated solvents are routinely used in NMR spectroscopy. A residual quadrupolar interaction or quadrupolar splitting will be observed in the ^2H NMR signal of the deuterated solvent when it is exposed to the anisotropic environment generated by the orienting medium. The size of the ^2H Q -splitting of the deuterated solvent will depend on the factors discussed above. At this point, it is evident that for the same degree of anisotropy (strength of alignment), different solvents will show different sizes of ^2H Q -splitting. Probably the most sensitive solvent is CDCl_3 , since it is an asymmetric molecule with a strong EFG. Acetone- d_6 and DMSO- d_6 have both been shown to be very poor solvents for probing the degree of anisotropy based on ^2H Q -splitting. Gels will swell differently in different solvents; hence, in order to compare the degree of anisotropy, it is a good idea to add CDCl_3 to the desired solvent and check its ^2H Q -splitting. A nice comparison using this approach was done by Luy and coworkers⁶¹ when the orienting properties of stretched poly(dimethylsiloxane) gels for apolar (hexane and benzene) and weakly polar organic solvents (CD_2Cl_2 , CDCl_3 , and d_8 -THF) were reported. It is also very important to monitor the evolution of the ^2H Q -splitting as the gel swells in the desired solvent until the ^2H signal reaches stable and homogenous splitting. Figure 4.7 shows the ^2H Q -splitting of CDCl_3 as a PMMA gel stick (3 mm $\times$ 10 mm) swells in a 5 mm NMR tube.

If the gel is made by polymerization of the desired monomer in the presence of the cross-linker and a radical initiator, it is highly likely that the

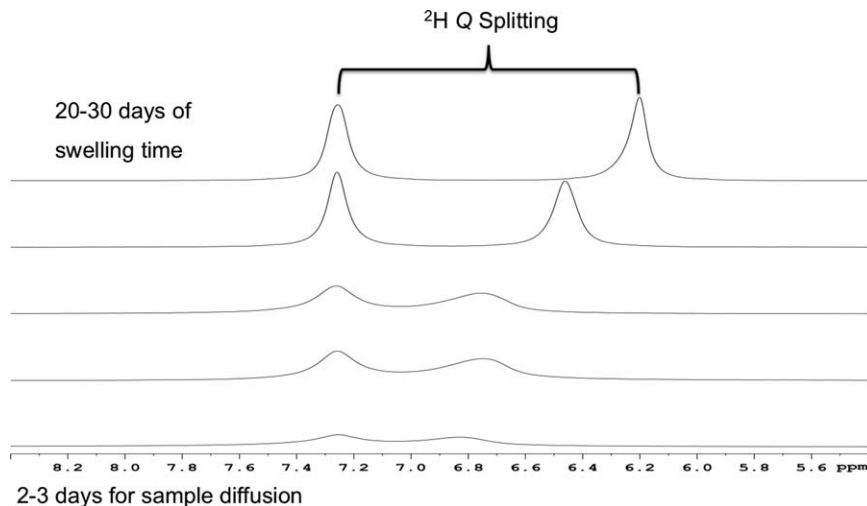


Figure 4.7 Evolution of the ^2H Q-splitting of the CDCl_3 signal in the ^2H NMR spectrum of a PMMA gel swollen in a 5 mm NMR tube. The top trace shows homogenous Q-splitting, indicating that the gel is anisotropic and ready to use.

polymer stick will contain residual amounts of free monomer. The monomer has to be washed out by dialysis. For PMMA gels, dialysis is not possible because the gels crack during the drying process and become unusable. In this case, if the signals of the monomer do not interfere with the signals of the compound to be analyzed and the contamination of the sample with residual monomer is not an issue, then the gels can be used as they are. Polyacrylamide-based gels have the advantage of being very easy to wash. The presence of residual monomer is not an issue in polymer gels prepared by monomer-free polymer chains cross-linked by radiation.⁶¹ Thus far, several polymer gels/solvents covering almost the whole range of polarity have been developed. However, none of these gels are yet commercially available. Although making them is not difficult for a chemist, the lack of commercial sources certainly discourages widespread use of RDCs for structure elucidation.

Our group has developed the use of PMMA gels as alignment media in CDCl_3 , and they have been successfully used for the structural analysis of natural products, as discussed below. However, the main drawbacks of using them with the traditional swelling approach proposed by Luy *et al.*⁵⁹ in 2004 is the extremely long swelling time and the unavoidable presence of residual monomer. In order to overcome these problems and to take advantage of the great flexibility shown by PMMA gels swollen in CDCl_3 , a new swelling procedure was developed by the Gil group.⁷⁶ Instead of using a 3 mm $\times$ 10 mm polymer stick, narrower gel sticks were prepared in 3 mm NMR tubes, giving origin to gel rods of 2 mm in diameter. They were cut into stick lengths of 25 mm and put in a regular 5 mm NMR tube. Enough CDCl_3

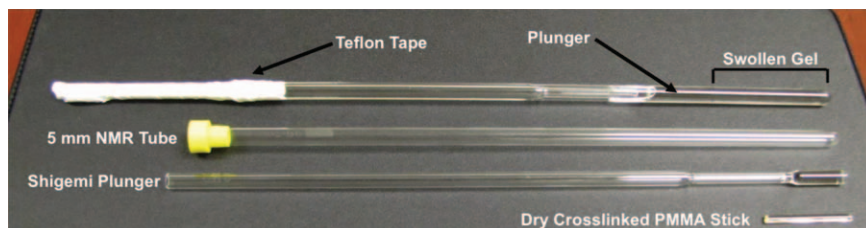


Figure 4.8 Experimental setup to swell a 3 mm PMMA gel stick using a regular NMR tube, Shigemi plunger, and Teflon tape.

Picture reproduced with permission from Gayathri *et al.*⁷⁶ Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

was added to cover the gel and a Shigemi plunger was inserted into the NMR tube to restrict the vertical growth of the gel as it swelled in the CDCl_3 . The Shigemi plunger was held in position by wrapping Teflon tape around the upper part of the tube, as seen in Figure 4.8.

Only radial swelling was allowed until the gel touched the walls of the tube. In this case, radial anisotropy is created, and it is perpendicular compared to the anisotropy generated by stretching using the Luy approach (Figure 4.9A and 4.9B). PMMA gels with a cross-link density of ethylene glycol diacrylate of around 0.3 molar% showed a CDCl_3 ^2H Q -splitting of ~ 50 Hz. If the plunger is removed, then the gel relaxes to dimensions of 40 mm long and 4 mm wide, and the ^2H Q -splitting is lost (Figure 4.9D).

The use of the Shigemi plunger is critical during the swelling process since the dry polymer sticks are brittle and fast vertical swelling produces radial cracks and, in certain cases, the gel breaks into two or more pieces. This method has shown significant advantages over the traditional approach. The PMMA gel shows incredible flexibility and it can be compressed and relaxed at our whim. If the gel is compressed back to the original position, the original ^2H Q -splitting is recovered. The compression is reversible, so the degree of alignment is “tunable”. It behaves as a sponge and residual monomer (Figure 4.10A, marked with asterisks) and can be easily removed by gently pumping the gel with fresh solvent using the plunger. It takes 8–10 washes with about 100 μL (Figure 4.10B) monomer-free gel. The same way in which the monomer is washed out of the polymer can be used to quickly infuse the sample inside the gel. The swelling process takes a few hours compared to 2–3 days using the traditional method, and after the monomer was washed out of the gel, the compound can be diffused inside the gel in a matter of minutes. As seen in Figure 4.10C–4.10F, four different compounds were loaded, analyzed, and washed out from the same gel.

With all of these advantages in hand, we see no reason to use the traditional method with PMMA gels any longer. As of today, the PMMA gels can only be used with CDCl_3 and CD_2Cl_2 , hence limiting the analysis to samples that are soluble in these two solvents. The procedure was further perfected by the development of a compression device (see Figure 4.11) that allows

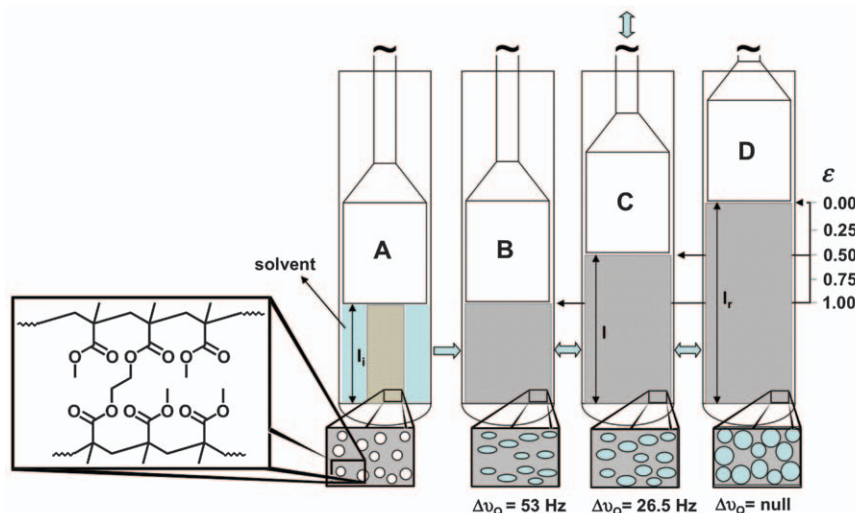


Figure 4.9 Representation (not to scale) of the reversible gel compression/relaxation experimental setup. (A) A dry cross-linked PMMA stick radially swells in CDCl_3 under the restriction of a Shigemi plunger and the gel “cavities”, adopting an oblate character (B) that conveys anisotropy to the gel, showing a ^2H quadrupolar splitting $\Delta\nu_Q$ of 53 Hz. (C) As the plunger is released, this oblate character decreases gradually to give an isotropic gel (D) for which the $\Delta\nu_Q$ collapses to zero. The process is linearly reversible from (D) to (B) as a function of the compression factor ε .

Figure reproduced with permission from Gayathri *et al.*⁷⁶ Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

precise control of the degree of alignment. The device is currently commercially available from New Era Enterprises, Inc.⁷⁸

Figure 4.12 shows evolution plots of the ^2H Q -splitting as a function of the compression in the device from fully relaxed (0 Hz) to maximum compression (30 Hz for this gel) at 180° turns of the knob on top of the threaded plunger. At the beginning and end of the compression process, the response is not linear (top plot) and it is in the center (bottom plot).

A stretching device for gels swollen in organic compounds was developed by the Luy group in collaboration with DuPont.⁷⁰ It is a modification of Kuchel's⁷¹ stretching device (which is only for water), but using all-organic, solvent-resistant flexible tubing made out of the perfluorinated elastomer Kalrez[®] 8002. A picture of the device is shown in Figure 4.13.

The swollen gel with the sample is inserted into the flexible tube before assembling the device. The rubber tubing is inserted into an open-ended tube. The flexible tubing is held by a Teflon or brass plug at the bottom and it is stretched at the top. The stretching position is locked with a fixation device. The major advantage of using stretching and/or compression devices is the ability to tune the degree of alignment.

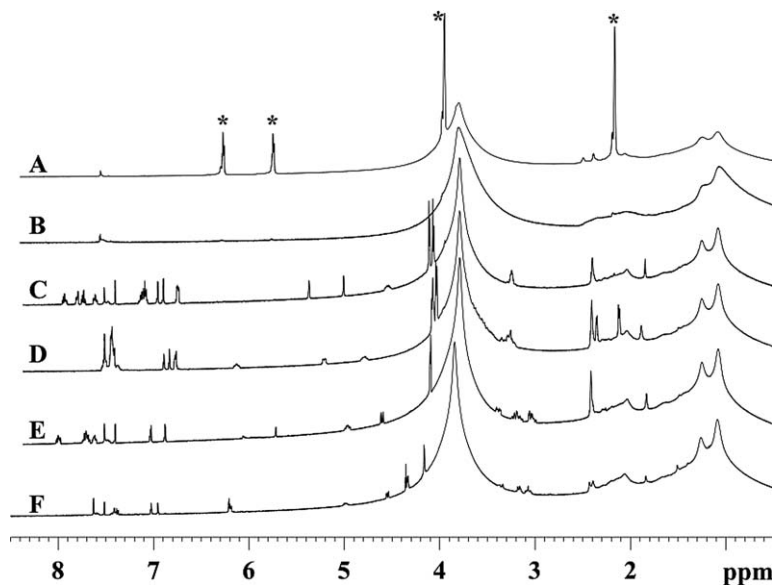


Figure 4.10 ^1H NMR spectra of the same PMMA gel in the same NMR tube right after the first swelling (A) with monomers marked by asterisks, and after five washes each with 200 μL CDCl_3 (B; no monomer) and suctioning in and squeezing out of four consecutive samples (C–F).⁷⁶

Figure reproduced with permission from Gayathri *et al.*⁷⁶ Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

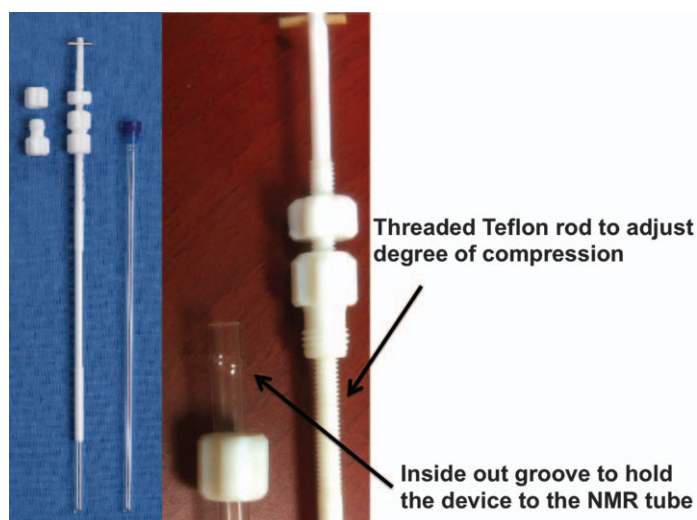


Figure 4.11 Compression device for precise control of the degree of alignment in compressed PMMA gels.

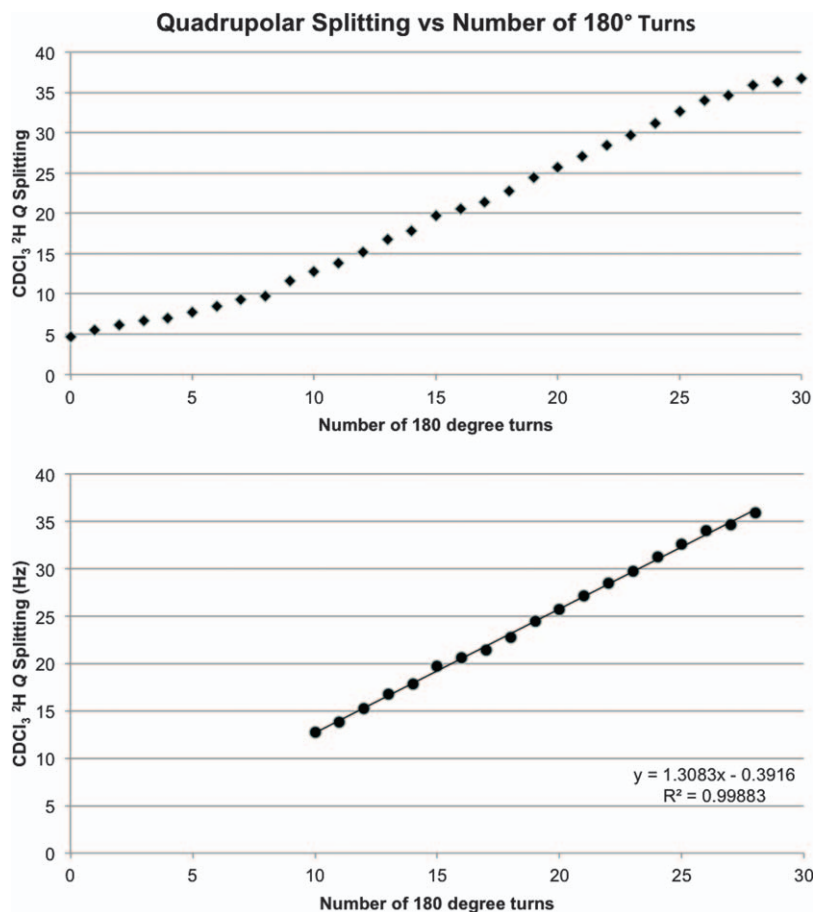


Figure 4.12 ²H Q-splitting of the ²H CDCl₃ signal as a function of the number of 180° turns of the knob on the compression device. Between turns 10 and 28, the response is linear, with $R^2 = 0.9988$.

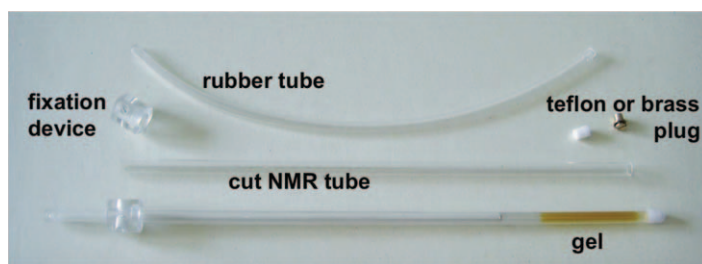


Figure 4.13 Stretching device made out of the perfluorinated elastomer Kalrez[®] 8002 for regulating the alignment of different swollen gels. Figure reproduced with permission from Kummerlöwe *et al.*⁷⁰ Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

4.3 The Nature of the Alignment Process

We mentioned above that when molecules tumble isotropically in solution, the probability over time of any molecule's internuclear vector pointing in any direction in space is the same, leading to a probability tensor with principal components $P_{xx}=P_{yy}=P_{zz}=1/3$ and an alignment tensor of null value. If we want to observe residual values for any of the anisotropic NMR parameters, such as dipolar couplings, chemical shift anisotropy, or quadrupolar couplings, we need to manipulate those probabilities and move them away from their isotropic value of $1/3$. To achieve this goal, we need to impose order on the molecular motion of the molecules in solution. In the macroscopic world, we are continuously imposing order on moving objects (people, animals, vehicles, *etc.*) by setting oriented obstacles in their path of motion. For example, grids of streets and mandatory direction of circulation impose order on vehicles' direction of motion. Walls and building prevent people from circulating in any desired direction. Returning to the microscopic world, if we want to manipulate the values of the probability tensor, we need to impose organized restrictions (obstacles) in the path of the motion of the small molecules, but we do not want them to bind to those obstacles. The orientation mechanism has to be purely steric. When molecules are in solution in the NMR tube, they can freely rotate and diffuse in any direction, as shown in the upper center part of Figure 4.14. The wall of the NMR tube in the microscopic world needs to be very far away from the molecules in order to have significant impact on the average orientation.

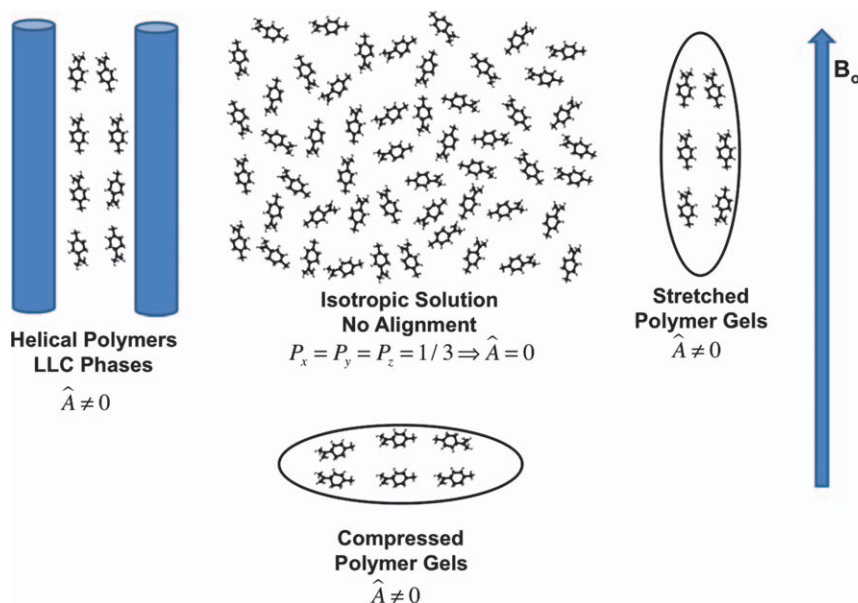


Figure 4.14 Different alignment scenarios depending on the type of alignment medium.

To generate microscopic order, we need oriented microscopic columns, as is the case for helical polymer-forming LLC phases (Figure 4.14, left), or we can even create cavities with asymmetric shapes that are all oriented in the same direction, as is the case for stretched polymer gels (Figure 4.14, right) or compressed polymer gels (Figure 4.14, bottom). These well-organized and oriented microscopic obstacles induce a repopulation of states at the expense of others by manipulating the orientational probabilities of the molecules. Please note that the pictures in Figure 4.14 are not to scale. The degree of alignment will strongly depend on the space between the obstacles.

Since we desire a weak alignment, the molecules have to spend 99.90–99.99% of the time moving in an isotropic space. The major problem with the LLC solutions, and even more so with nematic liquid crystals, is that the molecules spend less time in the isotropic motion states; probably 80–85% of the time in the case of nematic liquid crystals. The degree of alignment will also depend on the overall shape of the molecules. Globular molecules will align poorly, but flatter molecules will show good alignment properties. Also note that, in general, the orientational preferences of molecules aligned in LLC phases should be quite similar, while the orientational preferences of molecules between stretched and compressed gels should be rotated 90° with respect to B_0 (see Figure 4.14).

The differences in orientation between stretched and compressed polymer gels were reported as a part of the development of the compressing method using PMMA gels. As shown in Figure 4.15, the pseudoaxial protons at the seven-membered ring are, on average, almost parallel to B_0 in the compressed gel (negative RDCs), while they are perpendicular in the stretched gel. Figure 4.15 gives a representation of the corresponding probability tensors in the same molecular frame in both situations.⁷⁶

We have mentioned above that there is not yet a computational method for accurately predicting the alignment tensor for small molecules. In the case of large molecules, the program PALES²¹ has been successful in predicting the alignment tensor of proteins using walls and cylinders as model aligning phases. The problem with small molecules is that, unlike large molecules, they can also interact with the fine structure of the wall (grooves, pockets, cavities, *etc.*) of the orienting medium, making the prediction problem much more complicated. This is another area of research with considerable possibilities for further developments in both chiral and non-chiral alignment media.

4.4 How RDCs are Measured

Once we have the target small molecule oriented in the desired medium in terms of polarity and degree of alignment, the next step is to actually measure the desired RDCs. The signal splitting observed in isotropic solutions is due to indirect spin–spin couplings known also as scalar coupling or J -coupling. When the dipolar coupling is observed for a pair of nuclei ij

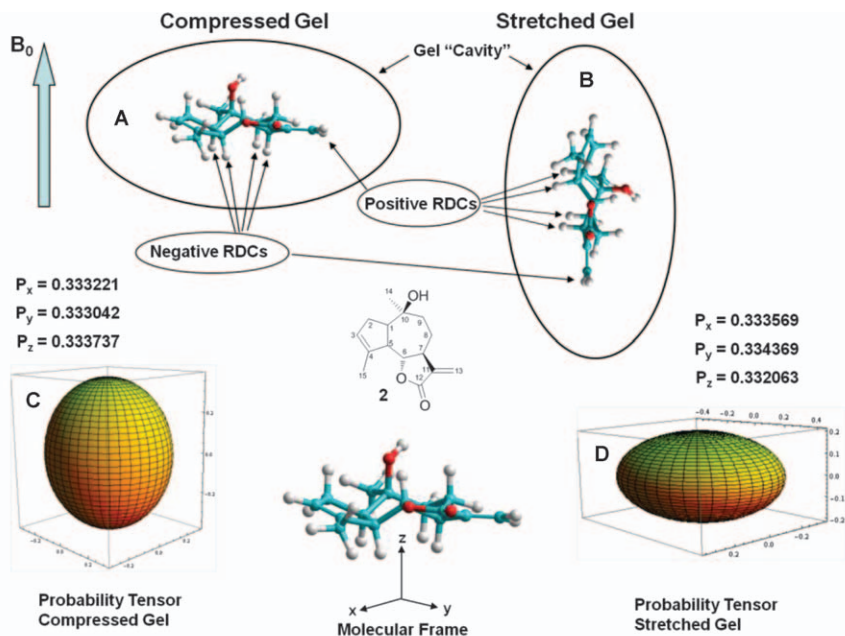


Figure 4.15 Comparison of the alignments generated by a compressed gel versus a stretched gel.

Figure reproduced with permission from Gayathri *et al.*⁷⁶ Copyright © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

already showing J -coupling, since the Hamiltonians of their J -coupling and their dipolar coupling commute, instead of observing a doublet of doublets of J_{ij} and D_{ij} , the signal splitting is equal to the sum $J_{ij} + D_{ij}$. Hence, in anisotropic conditions, we observe NMR signals with a total splitting $T_{ij} = J_{ij} + D_{ij}$. At this point, it is evident that in order to extract RDCs (D_{ij}), we need to collect two sets of NMR experiments: one in isotropic conditions that allows us to measure the J_{ij} values; and the other in anisotropic conditions to measure the total splitting T_{ij} and extract the RDCs from the difference $D_{ij} = T_{ij} - J_{ij}$. Special care must be taken with the signs since J and D can be either positive or negative. As a rule of thumb, $^3J_{HH}$, $^1J_{CH}$, and $^3J_{CH}$ are positive, while $^2J_{HH}$ and $^2J_{CH}$ are negative, except for some exceptions. If the signal for a pair of nuclei shows a null J_{ij} value but a measurable D_{ij} value, then its sign cannot be experimentally determined. In those cases, both signs need to be used in the calculations and we need to see which one fits better. Any experiment to measure J -couplings is suitable for measuring RDCs.

Nearly all of the work published thus far on the application of RDCs to the structural analysis of small molecules, mainly natural products, involved the use of only one-bond 1H - ^{13}C RDCs ($^1D_{CH}$) at natural abundance. The simplest experiment is just a ^{13}C proton-coupled spectrum (see Figure 4.16),

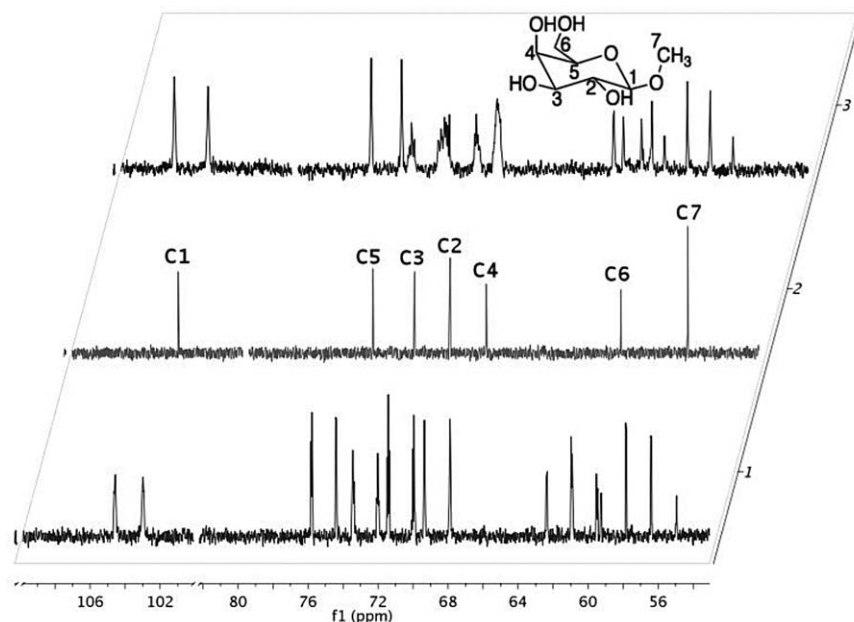


Figure 4.16 ^{13}C methyl- β -D-galactopyranoside spectra. Top: gated decoupled for an aligned sample in disodium cromoglycate. Middle: CPD ^{13}C non-aligned sample. Bottom: gated decoupled non-aligned sample.

usually using gated decoupling in order to benefit from NOE during the relaxation time.

The experiment benefits from good digital resolution, which allows accurate extraction of splitting, but suffers from the lack of sensitivity of a ^{13}C -detected experiment. In addition, overlapping can be a serious problem when molecular complexity increases. This last problem can be alleviated through the use of HETCOR-type experiments, but at the cost of additional experimental time. Due to the aforementioned problem with ^{13}C -detected experiments, inverse-detected HSQC-type experiments are currently preferred for D_{CH} determination. These $^1D_{\text{CH}}$ values are relatively easy to measure, either from F_2 or F_1 cross-sections, through ^1H -coupled HSQC spectra. The values of the $^1J_{\text{CH}}$ coupling constants in small organic molecules are in the range of 120–250 Hz. With $^1J_{\text{CH}}$ couplings being large and well resolved, an additional $^1D_{\text{CH}}$ contribution to the splitting upon inducing weak alignment is easily quantified from the change in the splitting. In order to measure $^1D_{\text{CH}}$ couplings from the F_2 cross-sections of proton-coupled HSQC experiments, a simple modification of the HSQC pulse program is necessary; ^{13}C decoupling during the acquisition simply needs to be turned off (Figure 4.17).

The main advantage of this type of experiment is that the couplings of each CH bond in methylene groups can be measured separately. The key to the collection of clean data, apart from having a homogenous and

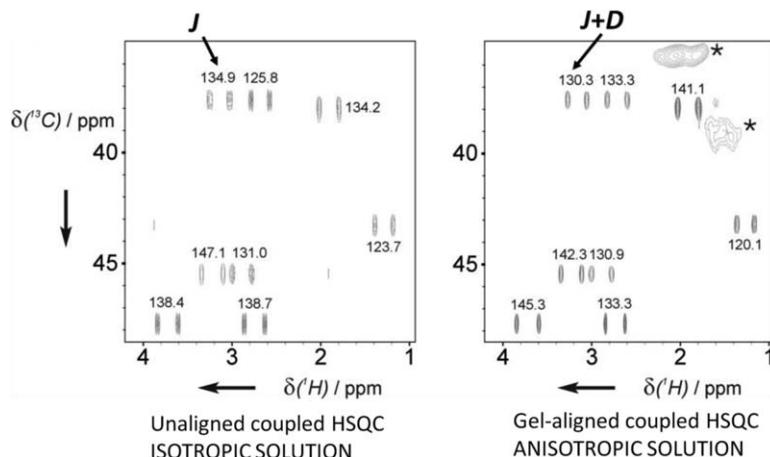


Figure 4.17 F_2 proton-coupled HSQC experiment in isotropic (left) and anisotropic (right) conditions. In isotropic conditions, the splitting of the signals in F_2 corresponds to the $^1J_{CH}$ value, while in anisotropic conditions it corresponds to $^1J_{CH} + ^1D_{CH} = ^1T_{CH}$. Figure kindly provided by Prof. Burkhard Luy.

well-shimmed gel sample, is to work in weak alignment conditions. This is extremely important in order to avoid peak distortion artifacts in the 1H , ^{13}C doublets due to a mismatching problem with the INEPT delay value. In the HSQC experiments, an average value of 145 Hz for $^1J_{CH}$ is used in order to set up the value of the delay for INEPT transfer. Large RDC values significantly move the signals, splitting $^1T_{CH}$ away from 145 Hz, giving rise to the presence of antiphase components superimposed on the in-phase component of each leg of the doublet. When antiphase distortions are present, it is not possible to accurately measure the $^1T_{CH}$ values. There are two solutions to the problem: either using very weak alignment media or using the clean in-phase (CLIP)-HSQC experiment proposed by Luy and coworkers. It is important to note that using weak alignment media do not totally remove these artifacts but alleviate them.⁷⁹ CLIP experiments allow for the collection of proton-coupled HSQC spectra with pure absorptive multiplets. In most of the articles reporting the use of RDCs for the analysis of small molecules, there is a preference for using F_2 proton-coupled HSQC. The major reason for this is that the experiment is fast, the J_{ij} and T_{ij} splittings are measured from the F_2 cross-sections with good resolution, and the splittings for non-equivalent protons of methylene groups can be measured separately. On the other hand, the major shortcoming of using these experiments is the signal broadening introduced by additional proton–proton RDCs. This is not a significant problem when very small molecules are being analyzed, but for more complex structures containing several methylene groups (*e.g.* steroids, triterpenes, complex alkaloids, *etc.*), the broadening introduced by proton–proton RDCs reaches unacceptable levels and the extraction of accurate $^1D_{CH}$

couplings is no longer possible. To avoid this problem, our group has adopted the use of F_1 proton-coupled HSQC experiments to extract RDCs from the F_1 cross-sections. Excellent results were obtained with the J -scaled BIRD HSQC experiments proposed by Luy.^{80,81} The J -scaled BIRD experiment incorporates a scaling factor κ for the J and T values that helps to minimize the impact of the low resolution in F_1 on the experimental error (see Error Treatment below, Section 4.5.2). The experiment is standard in the Bruker pulse program library. Due to the artifacts introduced by the INEPT transfer when dealing with ^1H - ^{13}C couplings, only the outer lines of the methylene triplets (or doublet of doublets) are observed in the F_1 cross-sections of F_1 -coupled HSQCs. As a result, the only available information is the sum of the couplings of the methylene carbon with its two protons ($^1J_{\text{CHa}} + ^1J_{\text{CHb}}$ in isotropic solution and $^1T_{\text{CHa}} + ^1T_{\text{CHb}}$ in anisotropic conditions). For methyl groups, the outer quartet components are three-times more intense than the two inner components, and the average $^1J_{\text{CH}_3}$ or $^1T_{\text{CH}_3}$ values can be extracted as a third of the distance between these two outer components. Although we can obtain only one RDC per CH_2 group (the sum of the couplings), this is not a problem since the sums can be easily treated in the computational calculations of the alignment tensor, as we will see later in this chapter. Figure 4.18 shows an example of F_1 proton-coupled J -scaled BIRD HSQC spectra collected in isotropic and anisotropic media.

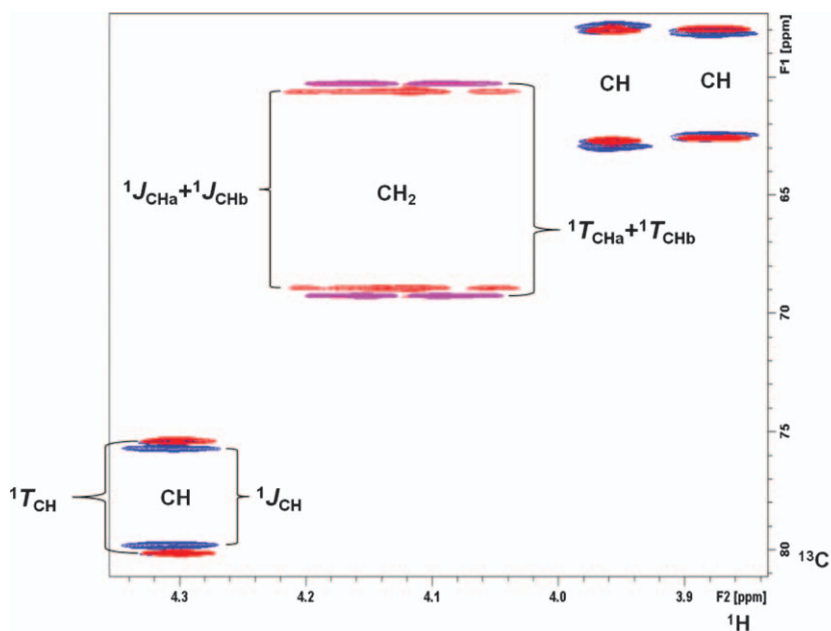


Figure 4.18 Overlapping of F_1 proton-coupled J -scaled BIRD HSQC experiments in isotropic and anisotropic conditions. For CH groups, blue = isotropic and red = anisotropic; for CH_2 groups, red = isotropic and pink = anisotropic. The experiment has a J -scaling factor $\kappa = 4$.

The studies using only one-bond CH RDCs have shown that $^1D_{\text{CH}}$ couplings can suffice to solve structural problems in small rigid or semi-rigid molecules. However, there are many situations in which it is desirable or essential to have more independent RDCs, such as: (a) molecules in which the individual fragments containing the stereocenters have too few CH bonds, or where the CH bonds are close to being parallel to one another, as is the case for menthol, which was shown in Figure 4.5; or (b) rigid or semi-rigid molecules with multiple unknown chiral centers, in which different configurations and/or conformations fit equally well to the experimental set of $^1D_{\text{CH}}$ data. Such cases are actually quite common and the measurement of long-range proton–carbon RDCs ($^2D_{\text{CH}}$ and $^3D_{\text{CH}}$) is readily feasible and can solve such problems. Long-range proton–carbon RDCs are about an order of magnitude smaller than the corresponding $^1D_{\text{CH}}$ values, and their measurement requires much higher precision. We have developed a selective J -scaled HSQC experiment in order to accurately measure long-range J_{CH} and T_{CH} values from the F_1 cross-sections. The pulse program also includes a scaling factor $\kappa = 20$, and the splitting in F_1 for the small value of long-range CH couplings is magnified by a factor of 20, allowing an accurate measurement of these small couplings.⁸² The other popular option, although it has not been extensively used as yet, is the measurement of $^2D_{\text{HH}}$ using the P.E. COSY,⁸³ E. COSY,⁸⁴ and P.E. HSQC experiments.⁸⁵

Finally, despite the very low sensitivity of the experiment, one-bond ^{13}C , ^{13}C RDCs ($^1D_{\text{CC}}$) have been measured using the INADEQUATE experiment.^{15,16} In addition, accurate measurement of these $^1D_{\text{CC}}$ values is also challenging since they are in the order of the long-range CH RDCs ($^2D_{\text{CH}}$ and $^3D_{\text{CH}}$). With the development of ultra-high magnetic fields and highly sensitive cryoprobes, we expect that these RDCs could be easily measured in the future using proton-detected experiments such as the J -modulated ADEQUATE.^{17,18} We have noted above that a set of RDCs encodes information about the relative orientation of the internuclear vectors, but it does not encode information about either their connectivities or their relative distances in space. However, imagine having access to the carbon–carbon connectivity network from INADEQUATE and/or J -modulated ADEQUATE spectra in combination with the relative orientation of each carbon–carbon bond from carbon–carbon RDCs. This would be like having crystallographic information in solution of the carbon skeleton of the molecule.

4.5 Data Analysis—How the RDCs are used in the Structural Analysis of Small Molecules

4.5.1 Possible Scenarios

Once the sample under study is aligned and the RDCs are collected, what do we do next? Different possible scenarios may be encountered in this situation. We mentioned above that in order to calculate the alignment tensor, we need at least five independent RDCs. Calculating the alignment tensor

using five independent RDCs for the sake of calculating an alignment tensor does not really help much in the process of the determination of the constitution, the configuration, and/or the preferred conformation of the molecule. We have also mentioned that a set of RDCs only encodes the relative orientation of the internuclear vectors, but it does not encode information about the distances between them. However, in practice, at least one or a few planar (2D) structures can be proposed from the combined analysis of mass spectrometry data and from a series of multidimensional and multinuclear NMR experiments. This information can even be given to structure elucidation programs in order to get a set of proposed 2D structures. Many structure elucidators rank the proposed structures by comparing predicted with experimental ^{13}C NMR data. 3D structures with the correct configuration can be determined in many cases using NOE and 3J -coupling constants analysis, but it is not always possible to determine the configuration using these two experiments. Hence, in order to successfully use RDCs in the structural analysis of small molecules, we have to have access to a complete set of all possible solution structures. Sometimes, the problem may be as simple as discriminating between two possible diastereomers. Keep in mind that we cannot use RDCs to discriminate between two enantiomers. In other cases, we may need to discriminate the correct structure from a set of $2^n/2$ diastereomers, where n is the number of stereocenters. The factor of two is due to the fact that the set of 2^n structures contains the pairs of enantiomers. In practice, we may face different scenarios: (1) a set of rigid configurations; (2) a set of structures that are mostly rigid, but containing small fragment with moderate flexibility; (3) a set of flexible structures, but the overall shape of the different conformations is very similar; and (4) a set of flexible structures in which the different conformations are very different in shape. Scenarios 1 and 2 are very easy to resolve using RDCs. Scenario 3 must be treated with care and needs to be carefully analyzed on a case-by-case basis, while scenario 4 is still a major problem and there is not yet a generally accepted and reliable method of applying RDCs in these cases.

4.5.1.1 Scenario 1

If we want to use RDCs to discriminate the correct structure from a set of rigid diastereomers, the procedure is very simple and straightforward. The group from which the correct structure has to be discriminated can be a complete set of $2^n/2$ diastereomers or just a reduced group of diastereomers that cannot be discriminated using other NMR experiments such as NOE or 3J -coupling constants analysis. The first step is to generate accurate 3D structures. Due to its accuracy, it is preferred that the computer-generated structures are energy minimized using DFT calculations. However, other energy minimization methods such as molecular mechanics or semi-empirical methods can be used, but with less accuracy.

Although it is not possible to calculate a structure solely from RDCs, it is possible to find which structures better satisfy the experimental set of RDCs.

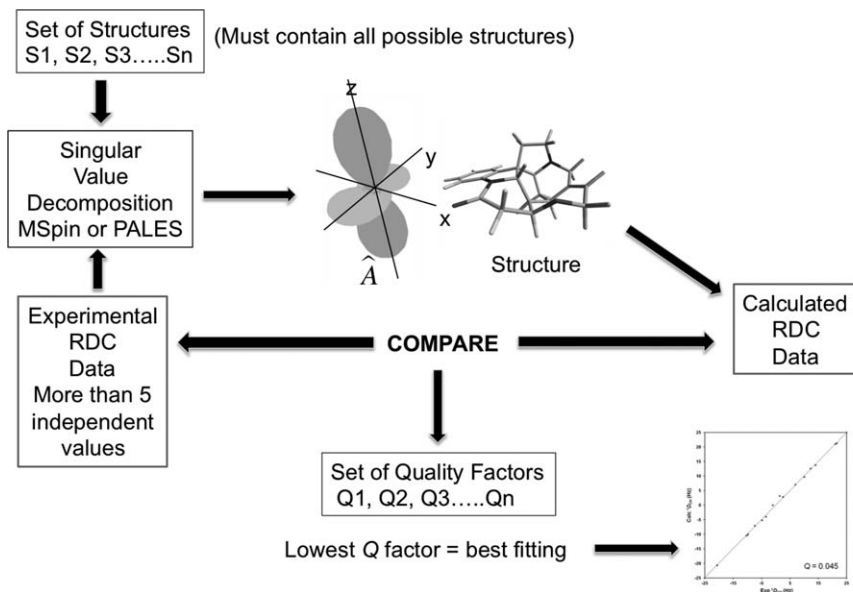


Figure 4.19 SVD fitting of the RDC data to a complete set of possible structures.

This is done by performing a SVD fitting of the experimental data to each of the possible structures. The rationale behind the fitting procedure is presented in Figure 4.19 and is described as follows. Using the experimental RDC data consisting of more than five RDC values—the more the better—and the coordinates for each internuclear vector extracted from a computer-generated structure, an alignment tensor is calculated.

The SVD method will always generate an alignment tensor from the best fitting of the experimental RDCs (D_{exp}) to the proposed structure. Using the calculated alignment tensor, the RDC value for each internuclear vector will be back calculated (D_{calc}). SVD is the most commonly used method, but similar results are obtained if the alignment tensor is determined by minimizing the difference between the observed (D_{exp}) and the back-calculated (D_{calc}) RDCs as a function of the matrix elements of the tensor A using a least-squares method. The quality of the fitting (D_{exp} vs. D_{calc}) is commonly expressed in terms of the Cornilescu quality factor Q :⁸⁶

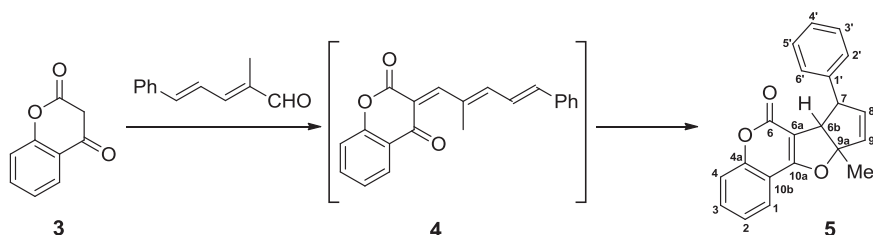
$$Q = \sqrt{\frac{\sum_{ij} (D_{ij}^{\text{exp}} - D_{ij}^{\text{calc}})^2}{\sum_{ij} (D_{ij}^{\text{exp}})^2}} \quad (4.27)$$

For the correct structure, Q should ideally be zero, when the back-calculated (D_{calc}) RDCs are identical to the experimentally measured RDCs (D_{exp}), but there is always noise in the measured RDCs, as well as in the structural model, due to geometric inaccuracies in the method employed to generate

the structure, and vibrational corrections may not be taken into account. Hence, the quality factor Q is normally larger than zero, but the lowest Q corresponds to the correct structure provided that all possible structures were part of the set that was used for the fitting. It is very important to note that the more similar the geometry of the two structures, the lower the difference in Q factor. However, if the RDCs are measured with good accuracy, the lowest Q corresponds to the correct structure. The impact of errors on structure discrimination will be discussed later. In practice, the procedure is very straightforward for end users. The structure files and a separate file with the RDCs values have to be provided to the software in the format required by them (PALES²¹ or MSpin²²) and the output file will provide the Q factor for each structure, as well as information about the alignment tensor, a list of calculated versus experimental values, the condition of the matrix, and the global degree of order (GDO). If fact, it is not necessary to be an expert in order to perform this analysis. An example of scenario 1 is the determination of the relative configuration of the cyclopenta[b]benzofurane-type derivative (5); a product of the reaction of the 1,3-dicarbonyl compound (3) with the $\alpha,\beta,\gamma,\delta$ -unsaturated aldehyde (4), as shown in Scheme 4.1.

Compound 5 is a conformationally rigid structure and has three chiral centers (C6b, C7, and C9a), leading to four possible diastereomers (5a₁, 5a₂, 5a₃, and 5a₄), as shown in Figure 4.20.

Compound 5 was aligned in a 3 mm PMMA gel stick swollen in CDCl₃, and only $^1D_{CH}$ RDCs were collected at different degrees of compression using the method described above. $^1D_{CH}$ values were collected for C1, C2, C3, C4, C6b, C7, C8, C9, C9a (CH₃), C2', C3', and C4' from the cross-sections of an F₂ proton-coupled HSQC spectrum. In this particular case, since RDCs were collected at four different degrees of compression, SVD fitting of RDCs to each structure was performed using the slopes of each $^1D_{CH}$ as a function of the 2H quadrupolar splitting of CDCl₃ at each degree of compression. The SVD fitting selected in one-shot the configuration 5a₄ with a $Q = 0.133$, while configurations 5a₁, 5a₂, and 5a₃ showed Q values of 0.851, 0.473, and 0.373, respectively. For the SVD fitting, the RDC average values for the CH₃ and phenyl rotors were used, since the feature is implanted in MSpin. The structural discriminating power of RDCs is clearly demonstrated here without the need for using either NOE or 3J -coupling constants analyses, though NOE and DFT ^{13}C chemical shift prediction using the DP4 method



Scheme 4.1

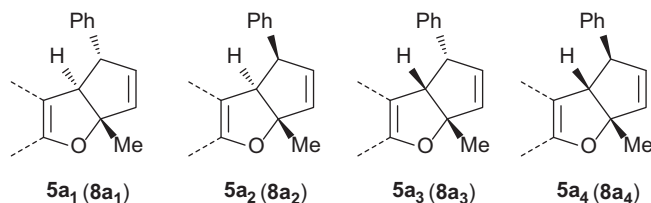
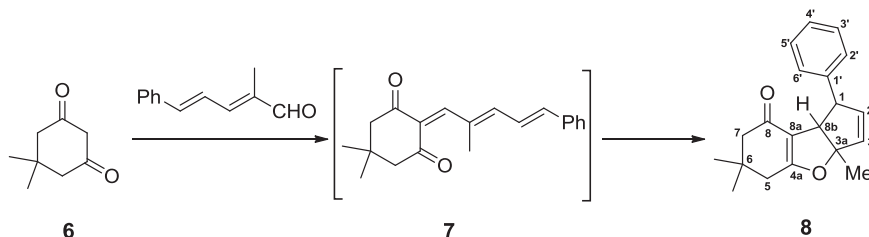


Figure 4.20 Relative configurations for compound 5.



Scheme 4.2

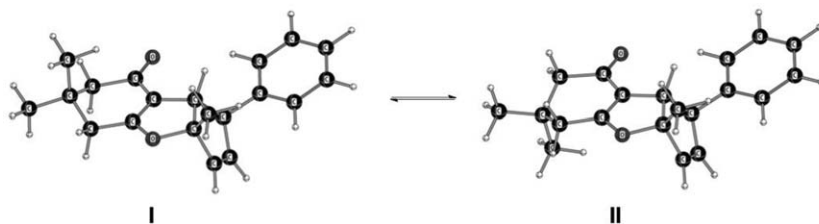
developed by Smith and Goodman⁸⁷ confirmed the same configuration. A similar approach, as will be discussed later, was used to discriminate the correct configuration of the natural sesquiterpene lactones ludartin and 10-*epi*-8-deoxycumambrin B.

4.5.1.2 Scenario 2

In the same work, the reaction of the 1,3-dicarbonyl (**6**) with the $\alpha,\beta,\gamma,\delta$ -unsaturated aldehyde (**7**) gave compound **8** (see Scheme 4.2), which is not as rigid as **5** and presents moderate flexibility at carbons C5, C6, and C7.

DFT calculations have shown that compound **8** has two conformations in fast exchange, as shown in Scheme 4.3.

We have shown above that SVD analysis straightforwardly selects the correct conformation in cases of rigid molecules, but for less conformationally rigid molecules, the analysis is slightly more complicated. However, from a visual inspection of the 3D models shown in Scheme 4.3, the geometry of the molecule is almost identical at all of its atoms, except for C6 and the two methyl groups attached to this carbon atom, which are involved in a flip-flop conformational exchange. Here, we can assume that this conformational exchange does not significantly disturb the overall shape of the molecule and, as a gross approximation, the alignment of each conformation is described by the same alignment tensor. Hence, a single alignment tensor approximation can be applied with confidence in this case. However, it is necessary to define a common molecular frame of reference for both conformations in order to perform the SVD calculations. In this case, we just superimposed the atomic coordinates of the heavy atoms of the two cyclopentyl rings, and the SVD fitting of the RDC data to each



Scheme 4.3

conformational pair for each configuration is weighted by contribution of the Boltzmann populations computed from the DFT free energies. The populations for the SVD calculation are given to the software as additional input parameters. RDCs for compound **8** were collected in the same way as for compound **5**, and the SVD calculations revealed the same configuration for compound **8**. The conclusion from these scenarios is that as long as the flexible portion of the molecule does not significantly disturb the overall molecular shape, the single-tensor approximation can be used with confidence. In some cases, it is not even necessary to include the RDCs values from the flexible part if they are not critical to determining the configuration of the molecule.

Scenarios 3 and 4 will be discussed in more detail in the section below that is specifically dedicated to the conformation problem.

4.5.2 How to Manage Experimental Uncertainties in RDC Analysis

Several sources of error may arise in the determination of the RDC data and the fitting procedure. The first and, in many cases, principal source of error is the determination of the splitting between lines either in F_1 or F_2 proton-coupled experiments due to the uncertainty in the peak position. As stated by Kontaxis *et al.*,⁸⁸ the error σ associated with the measurement of a coupling splitting T is equal to:

$$\sigma = \frac{LW}{SN \times \kappa} \quad (4.28)$$

where SN is the signal-to-noise ratio for the double peaks, κ is the J -scaling factor, and LW is the signal linewidth of the broader signal, and hence those signals in anisotropic conditions. Therefore, it is important to acquire the spectrum with a good signal-to-noise ratio. Note that errors in the alignment conditions are much larger due to the physical inhomogeneity of the gel, the difficulty in properly shimming the samples, and the presence of additional dipolar proton-proton or long-range proton-carbon couplings. In some cases, these problems can be alleviated by modification of the pulse sequences; for example, through the introduction of BIRD pulses to get rid of the small long-range couplings. Another way to reduce the error associated

with the splitting measurement when using F_1 -type experiments is the introduction of a J -evolution module in the sequence with a J -scaling factor κ that multiplies the total splitting T . This technique is very valuable for the measurement of long-range couplings where κ factors of up to 20 have been used,⁸² as well as for minimizing the error of one-bond proton-carbon splitting in F_1 without requiring the collection of large numbers of increments in the second frequency domain in order to improve resolution.

Even in the best aligning conditions and with the experimental setup, the determination of the error in the couplings may be extremely complicated in the case of a strongly coupled system. Consider for instance an ABX system such as a methylene group in a ^{13}C -gated decoupled experiment. Equations for the line pattern of the X part of dipolar-coupled ABX systems have been given by Emsley and colleagues.⁸⁹ Strong coupling between the AB protons dramatically affects the X part of the spectrum, which prevents direct extraction of individual D_{AC} and D_{BC} couplings. Note, however, that the separation between the outer lines remains constant with a value equal to $J_{\text{AC}} + J_{\text{BC}} + D_{\text{AC}} + D_{\text{BC}}$ (see Figure 4.21). Thus, a safer procedure in case of strong coupling is to report couplings for methylene groups as sum or half-sum values.

The presence of strong coupling between protons may greatly affect the separation between lines either in F_1 - or F_2 -coupled HSQC experiments. As in the case of ^{13}C -gated decoupled experiments, $^1\text{D}_{\text{CH}}$ methylene groups can be more safely reported not as individual C-H splittings, but as the sum or half-sum of the individual splittings (*i.e.* separation between outer lines of the lines manifold in the X part of the spectrum).

Additional source of errors may arise from the structural models used in the fitting. In general, geometries are obtained through different molecular modeling techniques. In our experience, any modern DFT methodology provides good geometries for RDC analysis. Note that the use of static geometries will neglect vibrational corrections to the RDCs; however, this factor will be mainly absorbed in the total degree of order of the obtained tensor, although it may cause some errors in flexible groups such as vinyl or phenyl groups. The major problem arises from a deficient treatment of long-amplitude conformational changes, as we will see in the following section.

How do we take into account all of these sources of error in the treatment of the data? In many cases, Monte Carlo bootstrapping techniques are employed. In a typical bootstrapping calculation, the error is estimated either for all of the measured data or particular errors are associated with different RDCs; for instance, one may assume a larger error for long-range data or assign a very large error to a particularly poorly resolved splitting. In this procedure, N series of synthetic data are built, starting from the originally reported one, by generating a random distribution of data around each datum. Typically, this distribution takes a Gaussian form and its standard deviation is the standard error associated with each measurement. Then, alignment tensor fitting is performed for each of the N series and properties such as the Q factor values are averaged over the set of calculations.

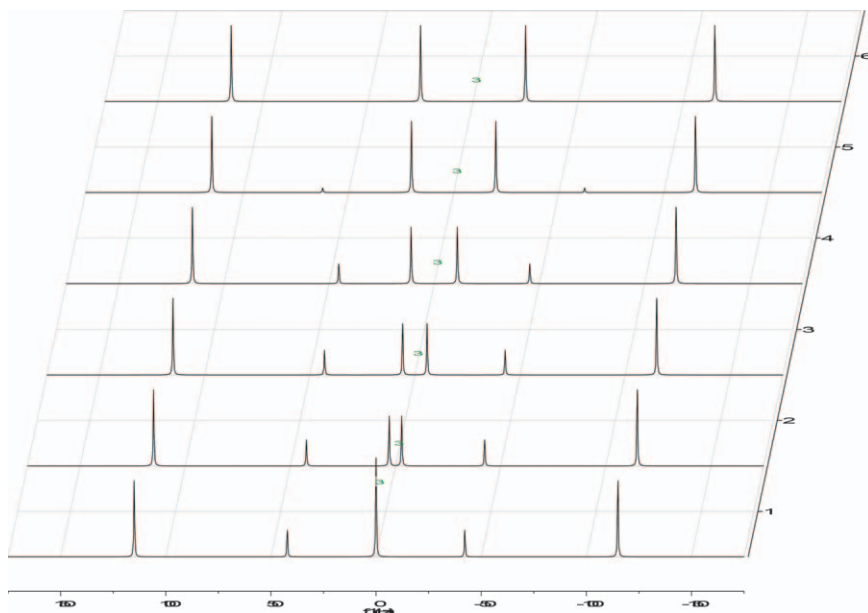


Figure 4.21 ABX simulated spectra. $J_{AB} = -14$ Hz, $D_{AB} = +40$ Hz, $J_{AC} = +180$ Hz, $J_{BC} = +120$ Hz, $D_{AC} = -40$ Hz, $D_{BC} = -30$ Hz. From bottom to top: $\delta_{AB} = 5, 10, 20, 50, 500$ Hz. Whereas the splitting between the outer lines remains constant with a separation $J_{AC} + J_{BC} + D_{AC} + D_{BC} = 230$ Hz, strong AB coupling prevents accurate extraction of individual D_{AC} and D_{BC} values. Spectra simulated with the MestReNova 7.1 software.

Thus, if a conservative error is given for all measurements and the distribution of quality factors is far enough (typically 2–3 standard deviations) for the best scoring structure, a statistical confidence criterion can be established to ensure that the correct structure has been found.

A simpler procedure to ascertain the impact of errors in the RDC fitting is to include the estimated errors in the expression of the quality factor, as in the following equation:⁹⁰

$$Q = \sqrt{\frac{\sum (D_i^{\text{calc}} - D_i^{\text{exp}})^2 / (\sigma_i^{\text{exp}})^2}{\sum (D_i^{\text{exp}} / \sigma_i^{\text{exp}})^2}} \quad (4.29)$$

Alternatively, the χ^2 factor or its inverse n/χ^2 , where n is the number of RDC data, can be used in a similar manner. Note that, in contrast to Q or χ^2 estimators, n/χ^2 larger estimators indicate a better fit.⁹¹ If $n/\chi^2 \geq 1$, the back-computed values fulfill the experimental restraints within the experimental uncertainty.

$$\chi^2 = \sum \frac{(D_i^{\text{exp}} - D_i^{\text{calc}})^2}{\sigma_i^2} \quad (4.30)$$

4.5.3 The Conformational Problem

The relationship between molecular conformation and the alignment mechanism has been studied since the beginning of research in the field. *A priori*, the aligning mechanism is not independent of the molecular conformation and therefore we cannot, in principle, derive a single alignment tensor for all of the conformations in the molecule. We can visualize the conformation dependence of the molecular alignment by considering a situation where a molecule interchanges between extended and closed conformations inside an aligning medium. While in the extended conformation, some of the possible orientations are forbidden due to the steric interaction with the walls. In contrast, this problem does not occur in the closed conformation, which will show a lower degree of order and therefore will furnish dipolar couplings of much smaller absolute values.

Most of the early studies on aligned media NMR of flexible molecules dealt with rotation around single bonds in very simple organic compounds. Two main approaches were based on the maximum entropy (ME) principle⁹² and the additive potential (AP) method.⁹³ The AP method highlights the important shortcoming that *a priori* knowledge of the potential surface is required. A similar constraint is not implicit in the ME methodology, but the lack of constraint correspondingly results in potential surfaces that are too flat. The APME hybrid model of Maliniak and colleagues^{94,95} claimed to solve these problems. Their approach allows the combination of RDC data with other observables as NOE and ³J-couplings. However, these methods cannot be applied in a straightforward manner to more complex organic compounds.

We will consider in the following a conformational jump model where the molecule lies in well-defined potential energy wells. Each particular conformation will have associated an own alignment tensor. Following the tensorial notation given above, the conformationally averaged RDC between nuclei *I* and *J* will be given by:

$$D_{IJ} = \kappa \sum \frac{p_i}{R_i^3} \sum \vec{r}^T \hat{A}_i \vec{r} \quad (4.31)$$

where *R* is the *IJ* internuclear distance in each conformation, $\hat{A}_i$ is the alignment tensor for the *i*th conformation, and *p_i* is the normalized weight of the *i*th conformation. Note that, for a two-conformation problem, determination of each alignment tensor will require us to fit a total of 11 unknowns (five tensors elements, multiplied by 2 + 1, population degrees of freedom). Clearly, a multitensor fit would not be practical in general.⁹⁰

The problem is even worse as the populations and the total degree of order are not independent quantities. Thus, in eqn (4.31), if we divide a given

population p_i by 2, we can just multiply the correspondent tensor $\hat{A}_i$ by the same quantity and the equality will still hold. A way to circumvent this problem is to assume that the total degree of order, which we can express in terms of the GDO, is the same for all conformations,⁹⁰ or we can make a fair estimation of the populations through the use of other experimental observables (J -couplings, NOE, and chemical shifts) or computational approaches. However, in many cases, things can be much easier if the conformational movement involves only a small part of the molecule and does not imply a large change in overall molecular shape or charge distribution. In this case, one can assume, as an approximation, that all of the considered conformations share a common alignment tensor (*i.e.* the alignment in the medium does not depend on the conformational state of the molecule). In this case, eqn (4.31) can be approximated to:

$$\langle D_{ij} \rangle = \kappa \sum \frac{p_i}{R_i^3} \sum \vec{r}^T \hat{A}_{ST} \vec{r} \quad (4.32)$$

where $\hat{A}_{ST}$ is the common single-alignment tensor for all the conformations. Note that if the distance R_i does not appreciably change between the different conformations, as is the case for the $^1D_{CH}$ RDCs, the equation can be rearranged to:

$$\langle D_{ij} \rangle = \frac{\kappa}{R_i^3} \sum p_i \sum \vec{r}^T \hat{A}_{ST} \vec{r} \quad (4.33)$$

One question arises at this point: if a common tensor is shared by all conformations, how does one define the relative orientation between them? A naive approach would be to just superimpose all of those atoms that are not involved in the conformational process. Another more “physical” approach has been suggested by Thiele *et al.*⁹⁰ based on an earlier suggestion by Burnell and De Lange,⁹⁶ and is to use that system of coordinates that throws out the rotational and translational degrees of freedom from the internal ones. This can be accomplished through the Eckart transform.⁹⁷ Note, however, that mathematically the Eckart transform is equivalent to the least-squares superimposition of atom coordinates, but in mass-weighted coordinates. Thus, for molecules with no particularly asymmetric mass distributions, the Eckart transform should give nearly equivalent results to just plain least-squares atom superimposition. On the other hand, the Eckart transform is defined for infinitesimal displacements and is therefore restricted to conformational movements of a local nature. Most of the practical applications in flexible molecules have made use of the “single-tensor” approach where a common alignment tensor is shared by all conformations (or at least the most populated ones). The application of “molecular dynamics protocols”⁹⁸ where the structure is refined against available data implies the use of the single-tensor approximation. Multitensor approaches are very scarce due to the difficulty in obtaining a

sufficiently large number of independent RDCs in small molecules, although the use of long-range data may somewhat alleviate this situation. When the single-tensor approach is applicable, it is possible to optimize the populations of molecules with a discrete conformational space by using a combination of Levenberg–Marquardt optimization of the population degrees of freedom with SVD optimization of tensor variables. A grid search starting from different initial populations is employed in order to avoid being trapped in local minima. Applying this technique, we quantified the populations of the three possible rotamers of phenylmenthol by recording RDCs from the ^{13}C -gated decoupled spectra collected in a PBLG/ CHCl_3 solution.⁹⁹

4.5.4 How to Handle Symmetrical Rotors

In most occasions, the molecules to be analyzed will contain symmetrical rotors as methyl or phenyl groups. The fast rotation of these groups causes averaging of the RDC values, which must be taken into account when fitting RDCs to 3D structures. For a methyl group, the $^1D_{\text{CH}}$ RDC is given by:

$$\langle ^1D_{\text{CH}_3} \rangle = \frac{3 \cos^2 \theta - 1}{2} D_{\parallel} = -\frac{1}{3} D_{\parallel} \quad (4.34)$$

Where D_{P} is the RDC for a virtual CH vector pointing in the direction of the rotation axis and with the same bond length. For a perfect tetrahedral geometry, the $^1D_{\text{CH}_3}$ value is just $-1/3$ of that virtual RDC. Many of the RDC analysis programs cannot deal directly with the averaged methyl data and they are transformed into corresponding C– CH_3 couplings.

$$D_{\text{CC}} = -\frac{1}{3} \frac{\gamma_{\text{C}} r_{\text{CC}}^3}{\gamma_{\text{H}} r_{\text{CH}}^3} = -0.22 D_{\parallel} \quad (4.35)$$

These back-converted D_{CC} couplings can then be input into programs such as XPLOR or others.

In a phenyl group, fast rotation will average the couplings of all ortho and meta C–H vectors. Considering a two-fold rotation model, the averaged CH for ortho and meta C–H vectors is given by:

$$\langle ^1D_{\text{CH}} \rangle_{\text{ortho}} = \langle ^1D_{\text{CH}} \rangle_{\text{meta}} = \frac{1}{4} D_{\parallel} + \frac{3}{4} D_{\perp} \quad (4.36)$$

where $D_{\parallel}$ represents virtual CH vectors pointing in the direction of the rotation axis and $D_{\perp}$ is the coupling of a vector that is orthogonal to the axis, but lying on the plane of the ring. Note that this equation only holds for the case of a symmetrical two-state jump model.

A modified setup of the SVD equations in order to explicitly account for averaging of the free rotors has been proposed.⁹⁹ In this scheme, RDCs are input as their actual measured values and the software applies the required average of data by summing up the rows of the cosine director matrix $\mathbf{M}$ for

the vectors to be averaged. This methodology has been also applied for methylene groups where only the sum of their corresponding splittings could be measured with confidence.¹⁰⁰

4.5.5 Software

Several programs have been developed and employed for RDC fitting in some of the groups working in the field. Here, we present a brief survey of the available software and their main characteristics.

PALES^{21,101} was initially developed in the Ad Bax group and is currently maintained in the group of Markus Zweckstetter in Göttingen. PALES is mainly oriented toward the prediction of sterical alignment, also including electrostatic corrections. The program runs in command line mode, but a graphical interface is also available. PALES accepts molecular structures in PDB models and can perform SVD fitting if RDCs and molecular 3D structures are provided.

MSpin¹⁰² was developed in the group of one of the current authors. MSpin is a commercially distributed program²² that is strongly oriented for small-molecule analysis. It is a graphically driven program with several options for group averaging and conformational analysis using the single-tensor approximation. The program also handles RCSA data by reading chemical shift tensor information from *ab initio* computational files.

Alternatively, in restrained MD simulations, the RDC SANI extension or XPLOR can be used.^{98,103} Note that these modules require *a priori* evaluation of the alignment tensor. Multitensor fitting has been performed by the group of Navarro-Vázquez by making use of Open Babel-based Python scripts. These are available upon request from the author.

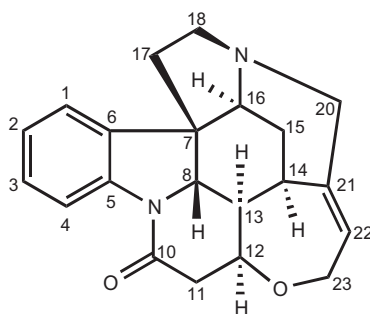
4.6 RDCs in the Structural Elucidation of Natural Products

The maturity of the technique is demonstrated by the determination of the relative configuration of several complex natural compounds. During the initial development of the RDC methodology, different natural products with well-characterized structures were used as a proof of concept in order to demonstrate the power of this technology in the determination of the configurations and preferred conformations of small molecules. As mentioned above, sodium cholate (3) was the first small molecule, as well as the first small natural compound, to be used to prove the concept from a structural standpoint.

Griesinger and coworkers¹⁵ measured $^1D_{CH}$, $^2D_{CH}$, $^3D_{CH}$, and $^1D_{CC}$ couplings in a sample of monoterpene menthol aligned in PBLG/ $CDCl_3$ using F_2 proton-coupled HSQC for $^1D_{CH}$, a modified HMBC for $^{2,3}D_{CH}$, and INADEQUATE for $^1D_{CC}$. In this work, the limitation on calculating the alignment tensor imposed by the presence of too many parallel CH bonds in

menthol was lifted by the use of ^{13}C – ^{13}C and long-range ^{13}C – ^1H RDCs. This is one of the first examples showing the application of RDCs to the structural analysis of water-insoluble molecules. SVD fitting of RDC data to the structure of menthol using the PALES program showed excellent agreement between the experimental and back-calculated RDCs. In addition, the assignment of methylene diastereotopic protons by permutation of their assignment in the SVD calculations in order to obtain the lowest Q factor was also demonstrated.

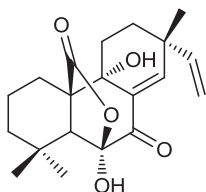
Almost simultaneously, the indole alkaloid strychnine (**9**) aligned in PBLG/ CDCl_3 was used by Thiele and Berger to probe the assignment of diastereotopic protons in methylene groups. SVD analysis of the RDC data for this work was also done with PALES.^{36,104}



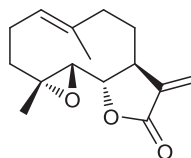
Strychnine (**9**)

The monoterpene norcamphor and the pimarane diterpene sphaeropsidin A (**10**) were used by Luy and coworkers to show the aligning properties of stretched poly(vinyl acetate) (PVC) gels in polar organic solvents.⁶² An excellent correspondence between the experimental and back-calculated $^1D_{\text{CH}}$ for sphaeropsidin A (**10**) was observed for a sample aligned in PAVC/DMSO.

The conformational analysis by RDCs of the sesquiterpene lactone parthenolide (**11**) was reported by Luy and coworkers as part of the work in which the Kalrez stretching apparatus was developed.⁷⁰



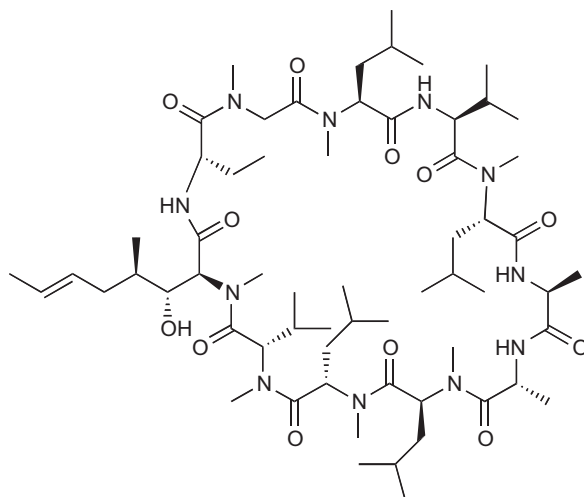
Sphaeropsidin A (**10**)



Parthenolide (**11**)

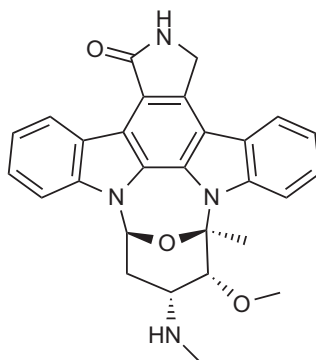
Refinement of the structure of cyclosporine A (**12**) using RDCs was performed with a sample aligned in stretched polydimethylsiloxane (PDMS) swollen in CDCl_3 . Only $^1D_{\text{CH}}$ values were measured using F_2 proton-coupled

HSQC, and they were fitted to existing crystal and NOE-derived structures using PALES.⁶³



Cyclosporine A (12)

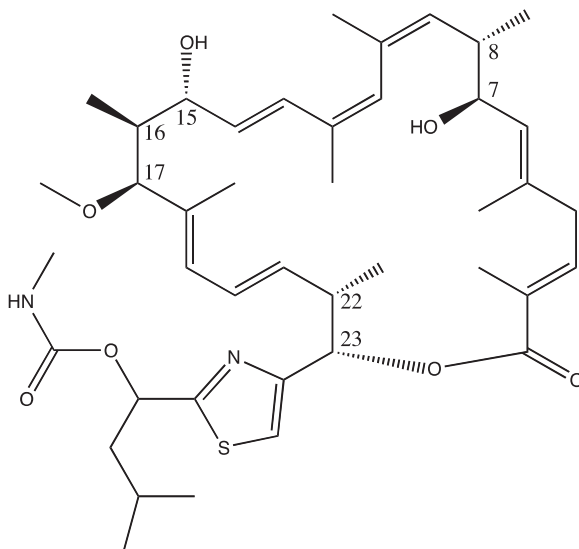
Staurosporine (13) was aligned in a stretched perdeuterated polystyrene gel swollen in CDCl_3 .⁶⁶ Only $^1D_{\text{CH}}$ values were measured using from F2 proton-coupled HSQC experiments. Configuration and conformation discrimination was performed on this compound by fitting 13 RDC values to different potential structures using PALES. The selection of the correct structure was straightforward, but the main goal of this study was to demonstrate that perdeuteration of the polymer could yield NMR spectra that were free of polymer background signal, permitting the measurement of RDCs at very low concentrations.



Staurosporine (13)

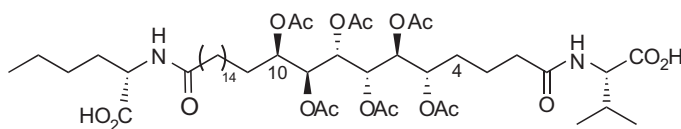
In 2008, Carlomagno and coworkers used a combined NOE/RDC MD-restrained approach to obtain the relative configuration of the macrolide archazolid A (14).¹⁰³ This molecule has seven independent stereogenic

centers, yielding a total of 64 possible relative configurations. All of these possible configurations were subjected to NOE- 3J -restrained molecular dynamics in XPLOR. The low-energy conformations obtained were scored against the RDC values by SVD fitting. Although the MD-restrained procedure gave similar energies for the *SS-RSS-RR* and *SS-RSR-RR* [C7, C8]–[C15, C16, C17]–[C22, C23], the *Q* factors were systematically lower for this last configuration, in agreement with the previous results from total synthesis.



Archazolid A (**14**)

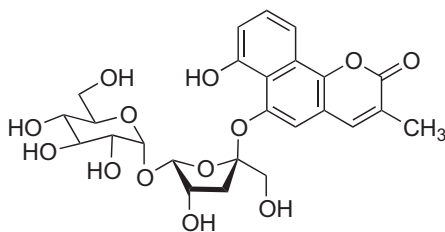
In the same year, Griesinger and coworkers employed RDCs to establish the configuration of several stereocenters of saggitamide A (**15**),¹⁰⁵ a peracetylated natural compound extracted from a tropical tunicate. Combined $^3J_{\text{HH}}$, $^2J_{\text{CH}}$, and $^3J_{\text{CH}}$ scalar coupling analysis resulted in four distinct configurations. On the basis of scalar couplings, the authors determined the populations arising from rotation of the C6 and C7 bonds. Saggitamide A was aligned in two different media—PH.PDMAA and poly(acrylonitrile) (PAN) gels—and $^1D_{\text{CH}}$ RDCs from C4 to C10 centers were recorded from F_2 -coupled HSQC spectra. Much lower *Q* factors were obtained for the depicted configuration, in agreement the one that had been previously proposed by Kishi.



Saggitamide A (**15**)

Griesinger and coworkers also established the configuration of the hexofuranose ring of sucro-neolambertellin.⁹⁸ Peracetylation of the structure

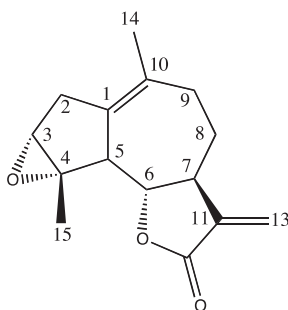
was required to reduce signal overlap. $^1D_{CH}$ RDCs were obtained by aligning the peracetylated derivative in PH gels ($DMSO-d_6$) and recording F_2 -coupled HSQC spectra. The RDCs obtained were analyzed using the floating chirality protocol in XPLOR NIH. Whereas NOE violations could not distinguish a single configuration, cross-validation with the RDC data clearly signaled *RSSR* as the best fitting configuration.



Sucro-neolambertellin (**16**)

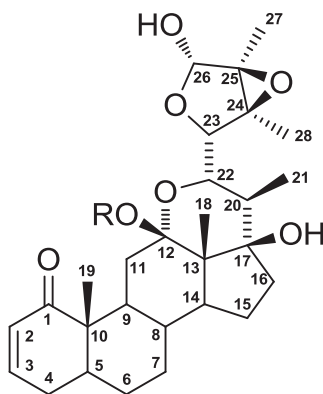
The first application of the PMMA gels as an alignment medium in $CDCl_3$ was demonstrated with the structural analysis of ludartin (**17**), a bioactive natural product (sequiterpene lactone) with gastric cytoprotective effects¹⁰⁶ and antiaromatase enzyme activity.¹⁰⁷ In this application, the structure of ludartin was used as a proof of concept of the structural discriminatory power of RDCs. Though a very rigid molecule, the determination of the configuration of the epoxide ring at carbons C3 and C4 was very challenging. When ludartin was first isolated in 1972,¹⁰⁸ the orientation of the epoxide ring was proposed as α , based on the chemical shift of H6 and H15. However, in 1989, it was shown that this criterion was not sufficient for the unambiguous determination of its configuration.¹⁰⁹ From a geometrical standpoint, the inversion of the configuration of the epoxide ring in this compound does not significantly change the orientation of proton H2, nor of the methyl group H15. In both isomers, α and β , the distance between the methyl group H14 and H2 is the same, leading to similar NOE values. The orientation of the methyl group H15 is not different enough to differentiate the α and β isomers based on the chemical shift, and H15 and H5 are at a distance that will exhibit a NOE between them in both isomers. In 1989, the structure was confirmed based on a series of chemical modifications that generated both isomers and hence allowed the determination of the correct configuration of the epoxide ring based on the effect produced by the epoxide oxygen on the chemical shift of H5 and H6. These chemical modifications were time consuming, and the HPLC separation of the products was difficult because of their similarity in chromatographic behavior. It took nearly a month for one of the authors during his graduate studies to achieve this goal. Nearly 19 years later, the same compound was aligned in a 5 mm PMMA stick swollen in $CDCl_3$. Ten $^1D_{CH}$ values were measured from the cross-sections of an F_2 proton-coupled HSQC collected in isotropic and anisotropic conditions. SVD fitting of these data to the structures of both isomers generated by AM1 semiempirical calculations showed a *Q* factor of

0.048 for the α -epoxide isomer (17) and of 0.221 for the β -epoxide isomer.⁷³ Since this analysis was done using the traditional swelling method developed by Luy *et al.*,⁵⁹ the bottleneck was the long swelling time (20–30 days). However, this is no longer a problem due to the development of the reversible compression/relaxation methods using the 3 mm PMMA sticks described above.⁷⁶



Ludartin (17)

Another successful application of the 5 mm PMMA stick swollen in CDCl_3 using the traditional method is the structural analysis of the steroidal lactone jaborosalactol 24 (18).¹¹⁰

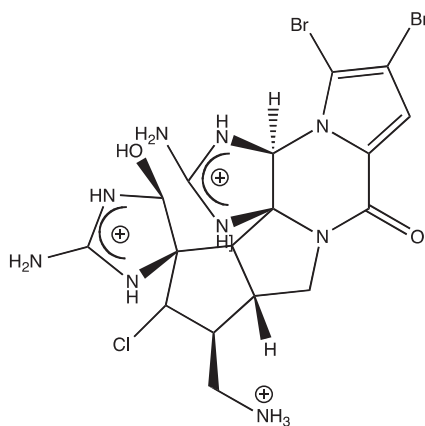


Jaborosalactol 24 (18)

The configuration at C23, C24, C25, and C26 of this compound isolated from *Jaborosa parviflora* was unknown. Attempts to determine this configuration only using a combination of NOE and $^3J_{\text{HH}}$ coupling constants analysis on all of the possible configuration/conformations generated by DFT calculations led to three possible solution structures. This ambiguity can be lifted by introducing RDCs into the analysis. This is a typical case of Scenario 2 discussed above. The main steroidal skeleton is very rigid and its configuration can be easily determined by conventional NMR experiments (NOE and $^3J_{\text{HH}}$ analysis), and the rotation of the side chain around the

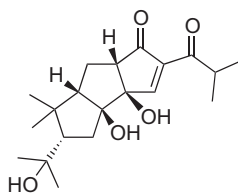
C22–C23 bond does not produce a significant impact on the alignment tensor. Hence, fitting of the lowest energy conformation for each of the eight possible configurations to a total of 15 $^1D_{CH}$ constants straightforwardly selected 23*S*, 24*S*, 25*S*, 26*S* as the correct absolute configuration relative to the well-known absolute configuration of the steroidal skeleton.

In a combined experimental and theoretical approach, the absolute configuration of the alkaloid dibromopalau'amine (**19**) was determined by Griesinger and coworkers.¹¹¹ The compound was aligned in stretched PAN gel swollen in DMSO-*d*₆. The structure is very rigid and the RDC analysis was very straightforward.



Dibromopalau'amine (**19**)

Luy and coworkers determined the relative configuration of tricyclohumol (**20**),⁹¹ a transformation product of beer. Since the absolute configuration at C4 and C5 centers could be derived from those of parent compounds in the biosynthetic transformation, three independent stereocenters still remained to be determined. The molecule was allowed to diffuse in a per-deuterated PAN gel in DMSO-*d*₆, which was then stretched using the device described earlier in this chapter.



Tricyclohumol (**20**)

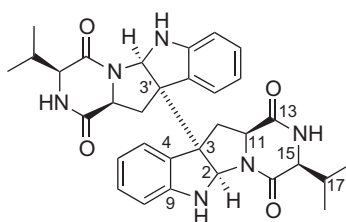
$^1D_{CH}$ RDCs were obtained from a CLIP-HSQC experiment, while proton–proton couplings ($^2D_{HH}$) were also obtained from a P.E. HSQC experiment. Since the fused rings contained three prochiral centers, assignment of the relative configuration must be performed on each configuration for every

combination of assignment of prochiral groups, resulting in a total of 64 possible combinations (permutations). RDCs were able to clearly discern the correct configuration of the molecule. Besides, in combination with the measured $^3J_{\text{HH}}$ couplings, the configuration of the prochiral groups could also be distinguished.

4.6.1 Dimeric Products

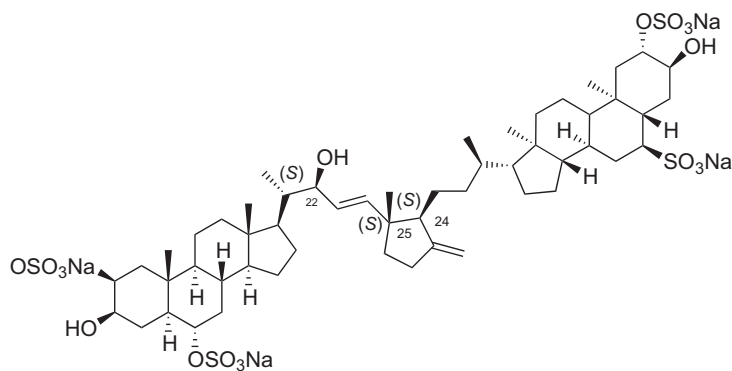
The power of RDCs in solving stereochemical problems in non-contiguous spin systems can be exploited in the determination of the relative configuration of dimeric molecules. As exposed in the NOE chapter of this monograph, traditional structural elucidation of dimers is especially difficult due to the fact that inter- and intra-monomer correlations cannot be easily distinguished. Illustrating the usefulness of RDCs for the elucidation of dimers, Navarro-Vázquez *et al.* determined the relative configuration of a synthetic analog (**21**) of asperdinin.¹¹²

Analog **21** was diffused in a PDMS gel and 10 $^1D_{\text{CH}}$ RDCs were measured by using the F_2 -coupled CLIP-HSQC sequence. Since the bispyrrolidine fusion was already known to be *cis*, four trial configurations were possible for this analog. MMFF94 force-field conformational search and NOE experiments indicated a strong preference for a *syn* conformation around the bond connecting the two dimeric units. On the other hand, the observation of strong optical activity indicated that the subunits have to possess the same configuration rather than mirror image ones, indicating a formal C_2 symmetry of the dimer. For a C_2 symmetric molecule like **21**, the alignment tensor has only three independent variables. This restriction was enforced by inputting the RDC values for each of the subunits inside the SVD module of the MSpin program. Therefore, the authors separately fitted each member of the conformational ensembles obtained for each configuration making use of the SVD procedure. RDC fitting indicated a 11*R*,15*R* configuration of the monomer units, as well as an *anti* conformation in accordance with the NOE and molecular mechanics data. *A priori*, a minor contribution of the higher energy (–)-synclinal form cannot be completely ruled out on the basis of existing data. In a second stage of their investigation, the authors simultaneously fitted both conformations to the RDC data by making use of a multiple-tensor approach. Since both conformations have C_2 symmetry, there are only three degrees of freedom for each tensor. Note, however, that, as stated before, populations and total degree of order are not independent, and one of the two must be fixed when the data are fitted. Since the two conformations in this case are similar in shape, the alignment tensors were constrained to have the same degree of order, therefore making a total of six unknowns in the fitting procedure, five alignment tensor components ($3 + 3 - 1$), and one degree of freedom for the populations. This multitensor approach afforded a 70:30% ratio of the two conformations, which is in good agreement with high-level DFT computations and chemical shift predictions.



Asperdimin analog (21)

Fibrosterol sulfate A (22), a polysulfated steroid isolated from the sponge *Lissodendoryx* (*Acanthadoryx*) *fibrosa*,¹¹³ is a heterodimeric molecule composed of two steroidal subunits and a linker containing a *exo*-methylenic cyclopentyl unit. The relative and absolute configurations of the steroid rings were determined on the basis of *J*-couplings, chemical shifts comparison with alkaloids of known configuration, and biosynthetic configurations. However, the relative configurations of the stereogenic centers in the linker were unknown.



Fibrosterol sulfate A (22)

Application of RDCs to this dimeric molecule is a formidable problem since a very large conformational space should *a priori* arise from the eight rotatable bonds connecting the rigid steroid units,¹¹⁴ making the use of the single-tensor approximation unreliable. Still, Griesinger and coworkers elected to determine the configuration of the unknown centers by employing a strategy combining NOE, *J*-coupling analysis, and RDCs.

NOE/*J*-coupling force-field-restrained dynamics was performed on the eight different possible configurations. Conformational clustering showed up to four different conformations with different degrees of contact between the steroid units, from an open form to a closed conformation with face-to-face contact of the steroid subunits. Populations were determined by simply counting the number of MD frames in each cluster. The authors allowed 5 mg of fibrosterol sulfate A to diffuse into a PAN polymer stick swollen in a

5 mm tube. Forty-four $^1D_{CH}$ RDCs were obtained with the help of F_2 -coupled CLIP-HSQC experiments. Good digital resolution (23.6 Hz) was necessary in the indirect dimension in order to clearly discern the carbon resonances.

Given the large number of RDCs obtained, the authors applied the multitensor approach. Since populations were obtained in the course of the MD simulations, they can be fixed, allowing quantification of the degree of order for each conformation.

In order to fit the tensors, the authors expanded the SVD equations:

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \dots \\ d_{n-1} \\ d_n \end{pmatrix} = \hat{M} \begin{pmatrix} A_{xy}^1 \\ A_{xz}^1 \\ A_{yy}^1 \\ A_{yz}^1 \\ A_{zz}^1 \\ \dots \\ A_{xy}^n \\ A_{xz}^n \\ A_{yy}^n \\ A_{yz}^n \\ A_{zz}^n \end{pmatrix} \quad (4.37)$$

where $d_i = D_i \frac{R^3}{K}$ is the reduced RDC, $A_{\alpha\beta}^i$ is the component of the i th alignment tensor, and $\hat{M}$ is the model matrix composed of the direction cosines of the internuclear vectors:

$$\hat{M} = \begin{pmatrix} p_1 a_{1,1} & p_1 b_{1,1} & p_1 c_{1,1} & p_1 d_{1,1} & p_1 e_{1,1} & \dots & p_n a_{n,1} & p_n b_{n,1} & p_n c_{n,1} & p_n d_{n,1} & p_n e_{n,1} \\ p_1 a_{1,2} & p_1 b_{1,2} & p_1 c_{1,2} & p_1 d_{1,2} & p_1 e_{1,2} & \dots & p_n a_{n,2} & p_n b_{n,2} & p_n c_{n,2} & p_n d_{n,2} & p_n e_{n,2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p_1 a_{1,m} & p_1 b_{1,m} & p_1 c_{1,m} & p_1 d_{1,m} & p_1 e_{1,m} & \dots & p_n a_{n,m} & p_n b_{n,m} & p_n c_{n,m} & p_n d_{n,m} & p_n e_{n,m} \end{pmatrix}$$

$$a = 2 \cos \phi_x \cos \phi_y, b = 2 \cos \phi_x \cos \phi_z, c = (\cos^2 \phi_y - \cos^2 \phi_x),$$

$$d = 2 \cos \phi_x \cos \phi_z, e = (\cos^2 \phi_z - \cos^2 \phi_x) \quad (4.38)$$

The authors codified the algorithm in Python and the 44 $^1D_{CH}$ RDCs were fitted to each of the four representative cluster conformations using the MD populations for each configuration. The fitting was clearly better for the 22S,

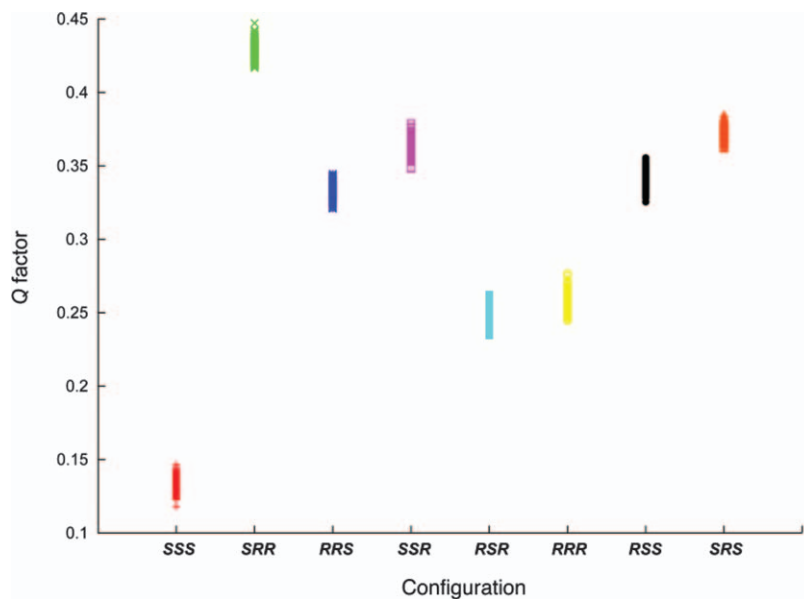


Figure 4.22 Multiple alignment tensor fit, showing the error bars of the Q factors of eight possible configurations of fibrosterol A sulfate (**22**) obtained by bootstrapping error analysis using a standard deviation value of 0.5 Hz.

24*S*, 25*S* configuration for **22** and statically sound, as determined by bootstrapping calculations (Figure 4.22).

A general problem with many natural products is that they are isolated in very small amounts. Alignment media, even working under optimal experimental conditions, introduces significant line broadening that leads to a considerable loss in signal-to-noise. Hence, such experiments require a minimum amount of sample that is normally higher than the amount available. In order to overcome this problem, Griesinger and coworkers have developed a PH gel with a diameter that is suitable for use in a 1.7 mm MicroCryoProbe™. The gels are polymerized in a 2 mm chamber and, after washing and rinsing, they fit neatly in 1.7 mm NMR tubes and can be re-swollen in only milligrams of DMSO- d_6 . The authors noted an important sharpening of the NMR lines as compared to gels prepared in 5 mm tubes, which they attributed not only to better shimming of the smaller-diameter tube, but also to a higher homogeneity of the gel inside the thinner tube. After excluding some of the possible relative configurations of vatiparol (**23**), due to geometric conformation restraints for each possible configuration, the remaining conformation/configuration combinations were optimized with the CFF force-field and refined at the DFT B3LYP/6-31G* level of theory. The authors determined the configuration of the stereogenic centers of vatiparol (**23**), excluding from the SVD fitting those $^1D_{CH}$ values from the rotatable phenyl rings (Figure 4.23).

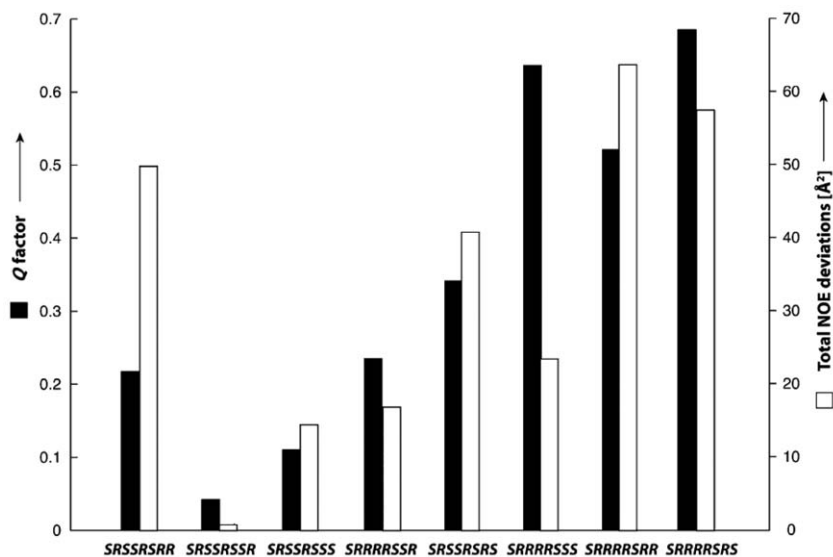
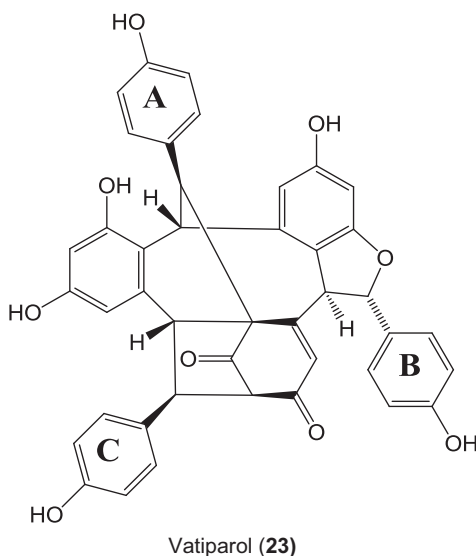


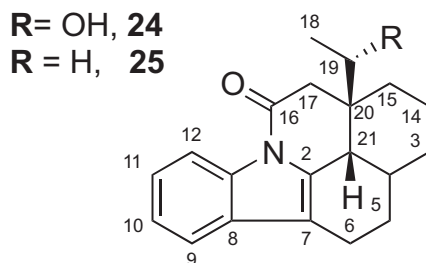
Figure 4.23 NOE deviations and RDC fitting in vatiparol (23).



Important NOE deviations were observed for those proton pairs involving the rotatable phenyl rings B and C. A conformational search showed two possible torsional states for each ring amounting to a total of four conformations. Inclusion of the RDC data from the B and C rings enabled the obtaining of populations for each conformation. This torsional information was then employed in the computation of ECD and ORD curves, which allowed the determination of the absolute configuration of vatiparol (23), since proper knowledge of the conformational preference of the phenyl

chromophores was essential to accurately describe the chiroptical properties. Note, therefore, that although RDCs alone cannot not be used to determine absolute configuration, the structural information they provide is instrumental when chiroptical techniques are used for this purpose.

Another example of the interplay between RDCs and chiroptical spectroscopies is the determination of the absolute configuration of a new *Vinca* alkaloid derivative. Structure **24** is similar to the known *Vinca* alkaloid eburnamonine (**25**) (or vincamine), save for the presence of a hydroxyl group in the ethyl chain. Natural eburnamonine is known to exist in both levo- and dextro-rotatory enantiomeric forms, and therefore the absolute configuration of compound **24** cannot be assigned on the basis of structural similarity.



Gil and coworkers¹¹⁵ recorded $^1D_{CH}$ spectra by aligning a sample of **24** in a compressed PMMA gel using the aforementioned compressing device. The J -scaled F_1 -HSQC experiment was used for recording $^1D_{CH}$. DFT-based conformational analysis of **24** showed that the hydroxyethyl sidechain of eburnamonine can adopt three different conformations of (–)-synclinal, (+)-synclinal, and antiperiplanar for the C15–C20–C19–C18 dihedral angle (Figure 4.24). The relative populations of each conformation for the four possible configurations were estimated at the DFT level of theory and RDCs were averaged according to a single-tensor scheme superimposing the heavy atoms of the fixed eburnan skeleton. From the RDCs, the Q factor was much lower for the 19*S*, 20*R*, 21*S* form ($Q = 0.085$) than for any of the other forms. The result was in complete agreement with DP4, a Bayesian estimator of the probability of a given structure in terms of DFT computed versus predicted chemical shifts. The very good agreement between the computed DFT ECD curve and the experimental one (data not shown) verified that the absolute configuration was indeed 19*S*, 20*R*, 21*S* and not its enantiomer.

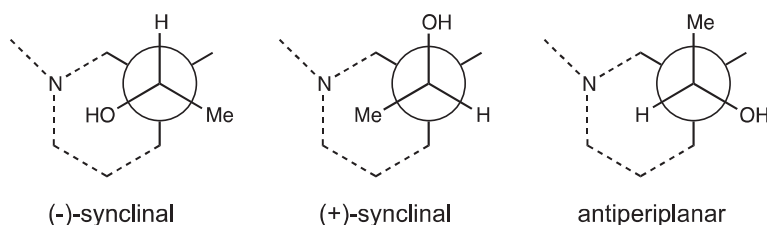


Figure 4.24 Three different conformations adopted by alkaloid **24**.

4.7 Conclusions

RDCs are nowadays an excellent tool for the configurational/conformational analysis of natural products. A large variety of weak-alignment media allows measurement of RDCs in standard NMR solvents. One-bond proton-carbon $^1D_{CH}$ couplings can be easily measured through HSQC experiments and later analyzed with the help of least-squares SVD or other fitting procedures. The structural information then obtained might not only solve configurational problems, but can also provide additional conformational information, which in turn may be crucial for the determination of the absolute configuration *via* chiroptical techniques.

References

1. T. D. W. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier, Amsterdam, Boston, 2nd edn, 2009.
2. D. Neuhaus and M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, 2nd edn, 2000.
3. F. A. L. Anet and A. J. R. Bourn, *J. Am. Chem. Soc.*, 1965, **87**, 5250–5251.
4. M. Karplus, *J. Chem. Phys.*, 1959, **30**, 11–15.
5. J. L. Marshall, *Methods in Stereochemical Analysis, Vol. 2: Carbon-Carbon and Carbon-Proton NMR Couplings: Applications to Organic Stereochemistry and Conformational Analysis*, Deerfield Beach, Verlag Chemie International, 1983.
6. C. A. G. Haasnoot, F. A. A. M. Deleeuw and C. Altona, *Tetrahedron*, 1980, **36**, 2783–2792.
7. R. C. Crouch, R. B. McFadyen, S. M. Daluge and G. E. Martin, *Magn. Reson. Chem.*, 1990, **28**, 792–796.
8. R. M. Gschwind, X. Xie and P. R. Rajamohanan, *Magn. Reson. Chem.*, 2004, **42**, 308–312.
9. R. Wagner and S. Berger, *Magn. Reson. Chem.*, 1997, **35**, 199–202.
10. A. Bax, G. Kontaxis and N. Tjandra, *Nuclear Magnetic Resonance of Biological Macromolecules, Part B*, 2001, vol. 339, pp. 127–174.
11. A. Bax, *Protein Sci.*, 2003, **12**, 1–16.
12. R. Sounier, L. Blanchard, Z. R. Wu and J. Boisbouvier, *J. Am. Chem. Soc.*, 2007, **129**, 472–473.
13. Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen and M. Fujita, *Nature*, 2013, **495**, 461–466.
14. F. Kramer, M. V. Deshmukh, H. Kessler and S. J. Glaser, *Concepts Magn. Reson., Part A*, 2004, **21A**, 10–21.
15. L. Verdier, P. Sakhaii, M. Zweckstetter and C. Griesinger, *J. Magn. Reson.*, 2003, **163**, 353–359.
16. M. Sarfati, P. Lesot, D. Merlet and J. Courtieu, *Chem. Commun.*, 2000, 2069–2081.
17. R. T. Williamson, A. V. Buevich and G. E. Martin, *Org. Lett.*, 2012, **14**, 5098–5101.

18. C. M. Thiele and W. Bermel, *Magn. Reson. Chem.*, 2007, **45**, 889–894.
19. G. Kummerlöwe, B. Luy and A. W. Graham, *Annu. Rep. NMR Spectrosc.*, 2009, **68**, 193–232.
20. J. A. Losonczi, M. Andrec, M. W. F. Fischer and J. H. Prestegard, *J. Magn. Reson.*, 1999, **138**, 334–342.
21. M. Zweckstetter, *Nat. Protoc.*, 2008, **3**, 679–690, 679.
22. *MSpin. MESTRELAB RESEARCH SL, Santiago de Compostela, SPAIN.*
<http://www.mestrelab.com>.
23. K. Madsen, H. B. Nielsen and O. Tingleff, *Methods for Non-Linear Least Squares Problems*, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Denmark, 2nd edn, 2004.
24. A. Saupe and G. Englert, *Phys. Rev. Lett.*, 1963, **11**, 462–464.
25. M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, John Wiley & Sons, Chichester, New York, 2001.
26. A. Mangoni, V. Esposito and A. Randazzo, *Chem. Commun.*, 2003, 154–155.
27. J. L. Yan, A. D. Kline, H. P. Mo, M. J. Shapiro and E. R. Zartler, *J. Org. Chem.*, 2003, **68**, 1786.
28. N. Tjandra and A. Bax, *Science*, 1997, **278**, 1111–1114.
29. A. Bax, G. Kontaxis and N. Tjandra, *Methods Enzymol.*, 2001, **339**, 127–174.
30. A. Bax and A. Grishaev, *Curr. Opin. Struct. Biol.*, 2005, **15**, 563–570.
31. J. H. Prestegard, *Nat. Struct. Biol.*, 1998, **5**(Suppl), 517–522.
32. J. R. Tolman, H. M. Al-Hashimi, L. E. Kay and J. H. Prestegard, *J. Am. Chem. Soc.*, 2001, **123**, 1416–1424.
33. J. H. Prestegard, C. M. Bougault and A. I. Kishore, *Chem. Rev.*, 2004, **104**, 3519–3540.
34. J. P. Bayle, J. Courtieu, E. Gabetty, A. Loewenstein and J. M. Pechine, *New J. Chem.*, 1992, **16**, 837–838.
35. C. Aroulanda, V. Boucard, F. Guibe, J. Courtieu and D. Merlet, *Chem. – Eur. J.*, 2003, **9**, 4536–4539.
36. C. M. Thiele and S. Berger, *Org. Lett.*, 2003, **5**, 705–708.
37. C. Aroulanda, M. Sarfati, J. Courtieu and P. Lesot, *Enantiomer*, 2001, **6**, 281–287.
38. A. Marx, B. Boettcher and C. M. Thiele, *Chem. – Eur. J.*, 2010, **16**, 1656–1663.
39. L. Arnold, A. Marx, C. M. Thiele and M. Reggelin, *Chem. – Eur. J.*, 2010, **16**, 10342–10346.
40. N.-C. Meyer, A. Krupp, V. Schmidts, C. M. Thiele and M. Reggelin, *Angew. Chem., Int. Ed.*, 2012, **51**, 8334–8338.
41. A. Marx, V. Schmidts and C. M. Thiele, *Magn. Reson. Chem.*, 2009, **47**, 734–740.
42. A. Marx and C. Thiele, *Chem. – Eur. J.*, 2009, **15**, 254–260.
43. A. Krupp and M. Reggelin, *Magn. Reson. Chem.*, 2012, **50**, S45–S52.
44. M. Dama and S. Berger, *Org. Lett.*, 2012, **14**, 241–243.
45. V. M. Marathias, G. J. Tawa, I. Goljer and A. C. Bach, II, *Chirality*, 2007, **19**, 741–750.

46. N. Nath and N. Suryaprakash, *J. Phys. Chem. B*, 2011, **115**, 6868–6875.
47. R. Berger, J. Courtieu, R. R. Gil, C. Griesinger, M. Koeck, P. Lesot, B. Luy, D. Merlet, A. Navarro-Vázquez, M. Reggelin, U. M. Reinscheid, C. M. Thiele and M. Zweckstetter, *Angew. Chem., Int. Ed.*, 2012, **51**, 8388–8391.
48. C. M. Thiele, *personal communication*.
49. C. M. Thiele, W. C. Pomerantz, N. L. Abbott and S. H. Gellman, *Chem. Commun.*, 2011, **47**, 502–504.
50. R. S. Prosser, J. A. Losonczi and I. V. Shiyanovskaya, *J. Am. Chem. Soc.*, 1998, **120**, 11010–11011.
51. P. Trigo-Mouriño, R. Santamaría-Fernandez, V. M. Sánchez-Pedregal and A. Navarro-Vázquez, *J. Org. Chem.*, 2010, **75**, 3101–3104.
52. L. J. Yu and A. Saupe, *Mol. Cryst. Liq. Cryst.*, 1982, **80**, 129–134.
53. E. Troche-Pesqueira, M.-M. Cid and A. Navarro-Vázquez, *Org. Biomol. Chem.*, 2014, **12**, 1957–1965.
54. J.-M. Pechine, A. Meddour and J. Courtieu, *Chem. Commun.*, 2002, 1734–1735.
55. V. V. Klochkov, A. V. Klochkov, C. M. Thiele and S. Berger, *J. Magn. Reson.*, 2006, **179**, 58–63.
56. R. Tycko, F. J. Blanco and Y. Ishii, *J. Am. Chem. Soc.*, 2000, **122**, 9340–9341.
57. H.-J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield and S. Grzesiek, *J. Biomol. NMR*, 2000, **18**, 303–309.
58. B. Deloche and E. T. Samulski, *Macromolecules*, 1981, **14**, 575–581.
59. B. Luy, K. Kobzar and H. Kessler, *Angew. Chem., Int. Ed.*, 2004, **43**, 1092–1094.
60. B. Luy, K. Kobzar, S. Knör, J. Furrer, D. Heckmann and H. Kessler, *J. Am. Chem. Soc.*, 2005, **127**, 6459–6465.
61. J. C. Freudenberger, P. Spiteller, R. Bauer, H. Kessler and B. Luy, *J. Am. Chem. Soc.*, 2004, **126**, 14690–14691.
62. J. C. Freudenberger, S. Knör, K. Kobzar, D. Heckmann, T. Paululat, H. Kessler and B. Luy, *Angew. Chem., Int. Ed.*, 2005, **44**, 423–426.
63. J. Klages, C. Neubauer, M. Coles, H. Kessler and B. Luy, *ChemBioChem*, 2005, **6**, 1672–1678.
64. K. Kobzar, H. Kessler and B. Luy, *Angew. Chem., Int. Ed.*, 2005, **44**, 3145–3147.
65. G. Kummerlöwe, J. Auernheimer, A. Lendlein and B. Luy, *J. Am. Chem. Soc.*, 2007, **129**, 6080–6081.
66. G. Kummerlöwe, S. Knör, A. O. Frank, T. Paululat, H. Kessler and B. Luy, *Chem. Commun.*, 2008, 5722–5724.
67. G. Kummerlöwe, M. U. Kiran and B. Luy, *Chem. – Eur. J.*, 2009, **15**, 12192–12195.
68. G. Kummerlöwe, M. Behl, A. Lendlein and B. Luy, *Chem. Commun.*, 2010, **46**, 8273–8275.
69. S. S. D. Büchler, G. Kummerlöwe and B. Luy, *Int. J. Artif. Organs*, 2011, **34**, 134–138.

70. G. Kummerlöwe, E. F. McCord, S. F. Cheatham, S. Niss, R. W. Schnell and B. Luy, *Chem. – Eur. J.*, 2010, **16**, 7087–7089.
71. P. W. Kuchel, B. E. Chapman, N. Mueller, W. A. Bubbs, D. J. Philp and A. M. Torres, *J. Magn. Reson.*, 2006, **180**, 256–265.
72. P. Haberz, J. Farjon and C. Griesinger, *Angew. Chem., Int. Ed.*, 2005, **44**, 427–429.
73. R. R. Gil, C. Gayathri, N. V. Tsarevsky and K. Matyjaszewski, *J. Org. Chem.*, 2008, **73**, 840–848.
74. S. Mayer and R. Zentel, *Prog. Polym. Sci.*, 2001, **26**, 1973–2013.
75. T. N. Pham, S. L. Hinchley, D. W. H. Rankin, T. Liptaj and D. Uhrin, *J. Am. Chem. Soc.*, 2004, **126**, 13100–13110.
76. C. Gayathri, N. V. Tsarevsky and R. R. Gil, *Chem. – Eur. J.*, 2010, **16**, 3622–3626.
77. F. Hallwass, M. Schmidt, H. Sun, A. Mazur, G. Kummerlöwe, B. Luy, A. Navarro-Vázquez, C. Griesinger and U. M. Reinscheid, *Angew. Chem., Int. Ed.*, 2011, **50**, 9487–9490.
78. *New Era Enterprises, Inc. Vineland, NJ, USA*. <http://www.newera-spectro.com>.
79. A. Enthart, J. C. Freudenberger, J. Furrer, H. Kessler and B. Luy, *J. Magn. Reson.*, 2008, **192**, 314–322.
80. J. Furrer, M. John, H. Kessler and B. Luy, *J. Biomol. NMR*, 2007, **37**, 231–243.
81. C. M. Thiele and W. Bermel, *J. Magn. Reson.*, 2012, **216**, 134–143.
82. P. Trigo-Mouriño, A. Navarro-Vázquez, J. Ying, R. R. Gil and A. Bax, *Angew. Chem., Int. Ed.*, 2011, **50**, 7576–7580.
83. L. Mueller, *J. Magn. Reson.*, 1987, **72**, 191–196.
84. C. Griesinger, O. W. Soerensen and R. R. Ernst, *J. Am. Chem. Soc.*, 1985, **107**, 6394–6396.
85. P. Tzvetkova, S. Simova and B. Luy, *J. Magn. Reson.*, 2007, **186**, 193–200.
86. G. Cornilescu, J. L. Marquardt, M. Ottiger and A. Bax, *J. Am. Chem. Soc.*, 1998, **120**, 6836–6837.
87. S. G. Smith and J. M. Goodman, *J. Am. Chem. Soc.*, 2010, **132**, 12946–12959.
88. G. Kontaxis, G. M. Clore and A. Bax, *J. Magn. Reson.*, 2000, **143**, 184–196.
89. M. Edgar, J. W. Emsley and M. I. C. Furby, *J. Magn. Reson.*, 1997, **128**, 105–113.
90. C. M. Thiele, V. Schmidts, B. Böttcher, I. Louzao, R. Berger, A. Maliniak and B. Stevansson, *Angew. Chem., Int. Ed.*, 2009, **48**, 6708–6712.
91. D. Intelmann, G. Kummerlöwe, G. Haseleu, N. Desmer, K. Schulze, R. Fröhlich, O. Frank, B. Luy and T. Hofmann, *Chem. – Eur. J.*, 2009, **15**, 13047–13058.
92. D. Catalano, L. Di Bari, C. A. Veracini, G. N. Shilstone and C. Zannoni, *J. Chem. Phys.*, 1991, **94**, 3928–3935.
93. J. W. Emsley, G. R. Luckhurst and C. P. Stockley, *Proc. R. Soc. London, Ser. A*, 1982, **381**, 117–138.

94. B. Stevansson, C. Landersjö, G. Widmalm and A. Maliniak, *J. Am. Chem. Soc.*, 2002, **124**, 5946–5947.
95. B. Stevansson, D. Sandstrom and A. Maliniak, *J. Chem. Phys.*, 2003, **119**, 2738–2746.
96. E. E. Burnell and C. A. De Lange, *J. Magn. Reson.*, 1980, **39**, 461–480.
97. C. Eckart, *Phys. Rev.*, 1935, **47**, 552–558.
98. A. Schuetz, T. Murakami, N. Takada, J. Junker, M. Hashimoto and C. Griesinger, *Angew. Chem., Int. Ed.*, 2008, **47**, 2032–2034.
99. V. M. Sánchez-Pedregal, R. Santamaria-Fernández and A. Navarro-Vázquez, *Org. Lett.*, 2009, **11**, 1471–1474.
100. P. Trigo-Mouriño, R. Santamaria-Fernández, V. M. Sánchez-Pedregal and A. Navarro-Vázquez, *J. Org. Chem.*, 2010, **75**, 3101–3104.
101. M. Zweckstetter and A. Bax, *J. Am. Chem. Soc.*, 2000, **122**, 3791–3792.
102. A. Navarro-Vázquez, *Magn. Reson. Chem.*, 2012, **50**, S73–S79.
103. C. Farès, J. Hassfeld, D. Menche and T. Carlomagno, *Angew. Chem., Int. Ed.*, 2008, **47**, 3722–3726.
104. C. M. Thiele, *J. Org. Chem.*, 2004, **69**, 7403–7413.
105. A. Schuetz, J. Junker, A. Leonov, O. F. Lange, T. F. Molinski and C. Griesinger, *J. Am. Chem. Soc.*, 2007, **129**, 15114–15115.
106. O. S. Giordano, E. Guerreiro, M. J. Pestchanker, J. Guzman, D. Pastor and T. Guardia, *J. Nat. Prod.*, 1990, **53**, 803–809.
107. J. G. Blanco, R. R. Gil, C. I. Alvarez, L. C. Patrino, S. Genti-Raimondi and A. Flury, *FEBS Lett.*, 1997, **409**, 396–400.
108. T. A. Geissman and T. S. Griffin, *Phytochemistry*, 1972, **11**, 833–835.
109. V. E. Sosa, J. C. Oberti, R. R. Gil, E. A. Rúveda, V. L. Goedken, A. B. Gutierrez and W. Herz, *Phytochemistry*, 1989, **28**, 1925–1929.
110. M. E. García, S. Pagola, A. Navarro-Vázquez, D. D. Phillips, C. Gayathri, H. Krakauer, P. W. Stephens, V. E. Nicotra and R. R. Gil, *Angew. Chem., Int. Ed.*, 2009, **48**, 5670–5674.
111. U. M. Reinscheid, M. Köck, C. Cychon, V. Schmidts, C. M. Thiele and C. Griesinger, *Eur. J. Org. Chem.*, 2010, 6900–6903.
112. C. Pérez-Balado, H. Sun, C. Griesinger, A. R. de Lera and A. Navarro-Vázquez, *Chem. – Eur. J.*, 2011, **17**, 11983–11986.
113. E. L. Whitson, T. S. Bugni, P. S. Chockalingam, G. P. Concepcion, X. Feng, G. Jin, M. K. Harper, G. C. Mangalindan, L. A. McDonald and C. M. Ireland, *J. Org. Chem.*, 2009, **74**, 5902–5908.
114. H. Sun, U. M. Reinscheid, E. L. Whitson, E. J. d'Auvergne, C. M. Ireland, A. Navarro-Vázquez and C. Griesinger, *J. Am. Chem. Soc.*, 2011, **133**, 14629–14636.
115. P. Trigo-Mouriño, R. Sifuentes, A. Navarro-Vázquez, C. Gayathri, H. Maruenda and R. R. Gil, *Nat. Prod. Commun.*, 2012, **7**, 735–738.

CHAPTER 5

Applications of High-resolving Power, High-accuracy Mass Spectrometry for the Structural Elucidation of Natural Products

GARY KRUPPA*^a AND WENDY ZHONG^b

^a Life Science Mass Spec, Bruker Daltonics, 40 Manning Road, Billerica, MA 01821, USA; ^b Analytical Research & Development, Merck Research Laboratories 126 E. Lincoln Ave., Rahway, NJ 07065, USA

*Email: Gary@GaryKruppa.com

5.1 Introduction

5.1.1 Mass Spectrometry for the Structural Elucidation of Unknowns

5.1.1.1 *Natural Products Complexity makes Direct Structural Elucidation with NMR Only Quite Difficult*

Mass spectrometry (MS) has played an increasingly important role in the analysis of biological molecules since the development of electrospray ionization (ESI)¹ and matrix-assisted laser desorption ionization.² These techniques have revolutionized the field, allowing the MS analysis of intact large biomolecules such as proteins, glycans, oligonucleotides, and lipids.

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

Natural products are another class of biomolecules for which data can easily be obtained by ESI-MS.³ Tandem MS (MS/MS) is capable of revealing the structural information of intact molecules by using a number of techniques to energize and fragment the molecules followed by a second stage of MS to analyze the fragments.⁴ Biomolecules such as proteins, glycans, and oligonucleotides are made up from a limited set of building blocks (*e.g.* 20 amino acids for protein or four nucleotides for DNA), connected by known and uniform chemistry. The MS/MS fragmentation patterns of such molecules are predictable; consequently, their MS/MS spectra can be easily or even automatically analyzed, giving the sequence information of the building blocks.⁵ Natural products, however, draw on a much wider range of building blocks and enzymatic chemistry for their biosynthesis.⁶ Natural products may contain the “standard” set of 20 amino acids, as well as less common amino acids, unusually modified amino acids, nucleic acids, lipid moieties, carbohydrate moieties, less common heteroatoms, and many other naturally occurring building blocks assembled by enzymes with connectivity that is not readily predictable. The complete elucidation of natural product structures involves the use of multi-dimensional nuclear magnetic resonance (NMR) to completely solve all aspects of the structure, including stereochemistry. However, complete determination of these complex structures by NMR alone is time consuming and can be error prone. For example, some commonly occurring heteroatoms such as sulfur are difficult to detect by NMR.

The molecular weight of the natural product, obtained using MS, can be used to validate the complete structural elucidation by NMR and to speed up the process of interpreting and assigning multi-dimensional NMR data. The way in which the mass spectrometric data are used depends on their accuracy and the availability of MS/MS data. When low-resolution mass spectrometers such as single- or triple-quadrupole instruments are used, the molecular mass and mass of MS/MS fragments will usually be accurate only to within a few tenths of a Dalton. Data obtained from such instruments can still be used as a good check of NMR data, as the calculated mass of the proposed structure from NMR must match the measured molecular weight from the low-resolution mass spectrum. However, there are limits to this, as, for example, one cannot distinguish the mass difference between one sulfur and two oxygen atoms using such mass spectrometric data, and there are even more commonly occurring isobaric substitutions that cannot be distinguished using low-resolution, low-mass accuracy mass spectrometers.

High mass accuracy is defined as mass accuracy better than 5 ppm, and can often be obtained from mass spectrometers having high resolving power, which is generally defined to be $m/\Delta m > 10\,000$ (where m is the measured mass and Δm is the full peak width at half maximum of the peak height). If data from high-mass accuracy, high-resolving power mass spectrometers are available, then as described in Section 5.1.1.2 below, it is generally possible to assign the complete elemental composition of a natural product. This eliminates some of the ambiguities described above, and is

very valuable in combination with multi-dimensional NMR data in order to solve the structures of natural products. For larger natural products, as we will discuss with examples in the later sections of this chapter, the use of ultra-high-resolving power mass spectrometers, capable of achieving $m/\Delta m > 100\,000$, can be very useful for unambiguously determining the presence and number of many heteroatoms. The remainder of this chapter will focus on the application of such high-mass accuracy, ultra-high-resolving power mass spectrometers to the elucidation of natural products structures.

5.1.1.2 Use of Accurate Mass together with NMR

5.1.1.2.1 How Accurate Mass gives an Elemental Composition. When high-resolution, high-mass accuracy MS is available, the first step in the structure identification of an unknown natural product is usually to use such instruments to determine the molecular formula. High-resolution ($m/\Delta m > 10\,000$ and mass error < 5 ppm) spectrometers, such as time of flight (TOF), magnetic sector, Orbitrap,⁷ and Fourier transform ion cyclotron resonance (FT-ICR),⁸ provide a mass accuracy of 1–5 ppm depending on the instruments. Table 5.1 shows a comparison of the mass accuracies and resolutions provided by three different mass spectrometers, and it is clear that FT-ICR MS offers the best performance. FT-ICR instruments can achieve < 1 ppm mass error routinely with $> 100\,000$ mass resolution. After obtaining the accurate molecular weight information, computer software will generate molecular formulae based on specified input constraints. There are many commercially and publicly available programs for determining the molecular formulae of unknowns from accurate mass data, and they share many common features. All such software allows the user to input constraints in order to specify what elements are possibly present in the natural product and what the mass measurement accuracy of the instrument is, and most programs consider many optional constraints such as the minimum or maximum number of double bond equivalents in the structure. Without a reason to suspect otherwise, it is often assumed that natural products contain only the atoms C, H, N, O, and S, and, less commonly, Cl, Br, and others must be included. The known exact masses of each element are not integers; for example, C is 12.0000, H is 1.0078, N is 14.0031, O is 15.9949, and S is 31.97207. The software

Table 5.1 Mass resolution and mass accuracy provided by different mass spectrometers.

MS	Mass resolution	Mass accuracy
TOF	$< 80\,000$	< 5 ppm Internal calibration
Orbitrap	$< 250\,000$	< 5 ppm External calibration
FT-ICR	Up to $10\,000\,000$	< 1 ppm External calibration

can generate possible combinations of the specified elements and adds up to the measured molecular mass within the specified error bar. The rounded nearest integer mass of the most abundant isotope of each element or molecule is known as the nominal mass, and the difference between the exact mass and the nominal mass is known as the mass defect. For example, hydrogen has a nominal mass of 1 u and a positive mass defect of 0.0078 u; carbon, by definition of the atomic mass scale, has a nominal mass of 12 u and a zero mass defect; oxygen has a nominal mass of 16 u and a negative mass defect of -0.0051 u. The software, besides using the known accurate masses of the elements, also generally applies constraints based on the known valency of each element in order to eliminate chemically improbable results (e.g. C_3H_{500}). The more possible elements that are included in the input for the unknown compounds, the higher the number of the possible formulae the software will generate. Also, for a given mass measurement accuracy, the number of possible combinations for a given set of elemental constraints increases dramatically with molecular weight. On the other hand, the lower the mass error (with high mass accuracy), the lower the number of the possible formulae the computer will generate. For example, for a mass at m/z 500.0000 with possible elements C, H, N, and O, the computer can generate nine elemental compositions with a mass error of 5 ppm. On the other hand, if the mass error is lowered to 1 ppm, the possible formulae will be cut down to only one. For an unknown at m/z 1000.0000, the possible number of elemental combinations of C, H, N, and O is seven within 1 ppm. For such large molecules, it is sometimes impossible to determine an unambiguous molecular formula based on the mass measurement alone, and so the use of MS/MS data and isotopic profile information (both discussed in the following sections) is necessary for such large natural products. Nevertheless, regardless of the molecular weight, it is quite clear that higher mass accuracy provides more reliable molecular formulae.

Determination of the molecular formula provides NMR spectroscopists with excellent supporting data for spectral interpretation. Since natural products are often not fully purified by LC prior to acquiring NMR data, knowing the exact number of each element in the molecule of interest will help determine which peaks in the NMR spectrum belong to the molecule of interest and which are from other molecules or contaminants in the analyte. In addition, some elements such as S and O cannot be directly observed in the NMR spectrum, but they can be easily determined based on the accurate mass measurement.

5.1.1.3 Use of Accurate Mass in MS/MS

Multistage MS (MS^n) is a powerful technique for unknown natural product structure identification. The superscript n refers to the number of mass filters or stages of mass selection that are used in the experiment. The most common type is known as MS/MS, where $n=2$, and it involves the use of two

mass filters. In a MS/MS experiment, ions are passed through the first mass filter, commonly a quadrupole mass filter, which is set to transmit only the ion of interest, known as the parent or precursor ion. Upon exiting the first mass filter, the parent ion is then accelerated and fragmented by energetic collisions with a gas such as He or Ar in the collision cell of the mass spectrometer. The product ions are then extracted from the collision cell and detected after passing through the second mass filter, which is scanned to measure the full set of fragment ions. Other than energetic collisions, many alternative ion activation techniques are available, such as infrared multiphoton dissociation (IRMPD), surface induced dissociation, electron capture dissociation (ECD), electron transfer dissociation (ETD), electronic excitation dissociation, and electron ionization dissociation.^{9–17} These different activation methods break the molecules in many different ways, resulting in extensive structural information, and some are usually preferred for certain classes of molecules (*e.g.* ECD and ETD work best with multiply charged peptide and protein precursor ions). The fragments from a MS/MS experiment may be further fragmented and analyzed by adding a second collision cell and third mass filter to perform a MS³ experiment. However, the preferred instruments for MS^{*n*} (*n* > 2) are trapping instruments, such as quadrupole RF ion traps, Orbitrap, and FT-ICR mass spectrometers. In these types of mass spectrometers, the mass filtering step is actually a mass selection step, performed by ejecting the ions that are not of interest. Then, the selected precursor ions are excited by RF so that they undergo energetic collisions with gas that is either present or added to the trap to produce fragments within the same trap. The mass selection step can be repeated to detect one of the fragment ions, and it can in turn be further fragmented for a MS³ experiment and so on. After the final stage of fragmentation, the fragments are then mass analyzed by the normal mass measurement scan of the instruments. Thus, unlike instruments that have multiple mass filters, the MS^{*n*} experiments performed on ion trapping instruments are separated in time rather than space. In this way, MS^{*n*} can easily be performed in a single trap analyzer, with *n* > 4 having been shown to be useful for the analysis of natural products, as will be shown in examples later in this chapter.

Quadrupole and ion trap instruments only provide nominal masses for the fragment ions (accurate only to a few tenths of a Dalton). Using such nominal mass information, one is less likely to obtain the molecular formulae of fragment ions. On the other hand, quadrupole-TOF (Q-TOF), Orbitrap, and FT-ICR instruments provide high mass accuracy for all fragment ions. Based on the high mass accuracy of the fragment ions, molecular formulae for each individual fragment ion can be readily determined, which contain unique and valuable substructural information regarding the precursor ion. High mass accuracy increases the confidence in the assignment of the elemental composition of each fragment ion. Furthermore, MS^{*n*} experiments provide the relative connectivity of the fragments. The information obtained from these experiments can be used to propose substructure moieties. This is particularly useful for larger natural products. As noted

above, accurate mass measurement alone may not yield a single, unambiguous molecular formula. In such cases, accurate fragment masses help to determine the unambiguous formulae for these large natural products, which must be subsets of the full molecular formula and, by putting together all of the constraints from the fragments ions, the molecular formula can be unambiguously determined. These substructures also assist in the NMR data interpretation.

5.1.1.4 Advantages of Ultra-high Resolution and Mass Accuracy

Isotope patterns in mass spectra can also help in elemental composition assignments. Isotope patterns in mass spectra arise from the fact that most elements in nature are found not as single atomic masses, but have isotopes resulting from different numbers of neutrons in the nucleus. For example, carbon in nature has 98.9% ^{12}C (exact mass 12.000000) and 1.1% ^{13}C (exact mass 13.003503). Thus, the mass spectrum of buckminsterfullerene, or C_{60} , is not a single peak at nominal mass 720 u, but a distribution of peaks as shown in Figure 5.1. The relative intensities of the isotopic peaks can be calculated by the binomial method,¹⁸ but as the molecule becomes larger and the number of different atoms considered increases, the complexity of this calculation explodes. A sophisticated Fourier transform method for the calculation of theoretical isotope patterns has been described,¹⁹ and this has been implemented in some commercial software packages for the determination of the molecular formulae of mass spectral peaks. If the mass

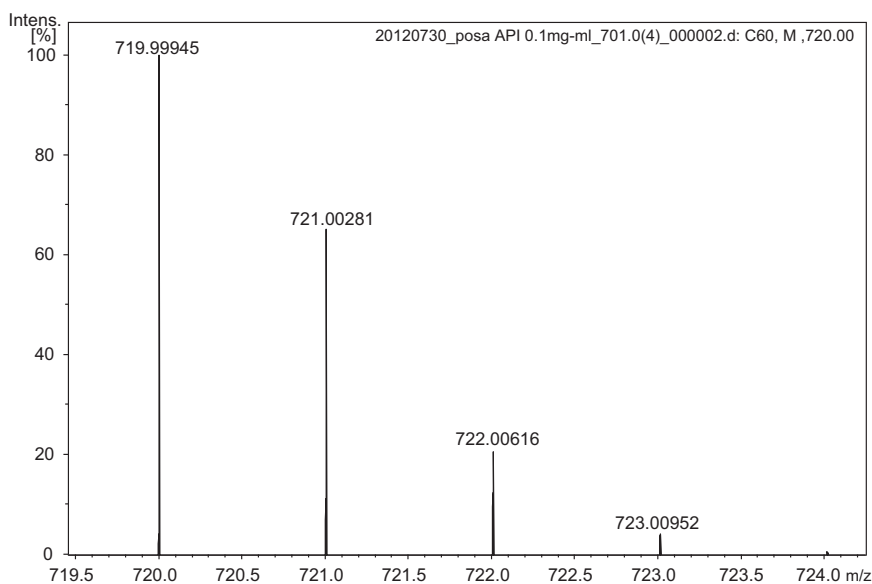


Figure 5.1 Calculated C_{60} isotopic patterns at a resolving power of 600 000.

Table 5.2 The mass calculated based on the combination of lightest isotopes for each element in the molecular formula is referred to as the monoisotopic mass, and the corresponding peak in the mass spectrum is known as the monoisotopic peak.

<i>m/z</i>	<i>I</i> (%)
719.99945	100
721.00281	64.9
722.00616	20.7
723.00952	4.3
724.01287	0.7

spectrometer is capable of measuring the relative isotopic abundances with good fidelity, the isotopic peak intensities measured experimentally can be compared to formulae predicted based on the accurate mass measurement, and can be used as an additional method to rank the results, as shown for the C₆₀ example in Table 5.2. The mass calculated based on the combination of the lightest isotopes for each element in the molecular formula is referred to as the monoisotopic mass, and the corresponding peak in the mass spectrum is known as the monoisotopic peak. For C₆₀, the monoisotopic mass is 719.99945 u, which is 60-times the mass of ¹²C (12.000000 u by definition of the atomic mass scale) minus the mass of an electron, appearing at nominal mass 720 u. The peak appearing at nominal mass 721 u has the composition ¹²C₅₉¹³C₁, and has a theoretical relative intensity of 64.89% based on the relative natural abundance of ¹³C of 1.109%. It should be noted that the relative abundance given here is an average of the relative abundance of ¹³C found in many naturally occurring substances. A commonly used shorthand nomenclature is to refer to the monoisotopic peak as A, the second peak in the isotopic pattern is referred to as A + 1, the third as A + 2, *etc.* For molecules containing heteroatoms, the A + 1 and A + 2 peaks will contain contributions from the isotopes of these atoms. Some elements commonly occurring in natural products, like chlorine and bromine, have very distinct isotope distributions: chlorine has two isotopes, with ³⁷Cl present at an abundance of about a third that of ³⁵Cl; bromine has two isotopes at nominal mass ⁷⁹Br and ⁸¹Br with about equal abundance. In the mass spectrum, the high relative abundance of the A + 2 peak, caused by the heavier isotopes of chlorine and bromine, can often be visually identified. Such relative abundance information is usually reflected in the ranking of match quality between the theoretical and the experimental isotopic distribution measured in the mass spectrum.

5.1.1.5 Isotopic Fine Structure Obtained via Ultra-high-resolution MS

In all except for the mass spectrometers with the highest resolving power, the contributions of different elements to the A + 1, A + 2, *etc.*, peaks are not

resolved and are simply added to the intensity of the isotopic peak. For example, if the natural product contains carbon, chlorine, and sulfur, the intensity of the peak at the $A + 2$ position will be the sum of contributions from ^{34}S , ^{37}Cl , and $^{13}\text{C}_2$ on a regular mass spectrometer. However, ^{34}S , ^{37}Cl , and ^{13}C have different mass defects, and if the resolution of the mass spectrometer is sufficiently high, the individual $A + 2$ peaks from these isotopes can be resolved, allowing for the direct counting of the number of these heteroatoms present in the molecule. The pattern that appears when the contributions from various elements to the $A + 2$ or higher isotopic peaks is resolved is termed the isotopic fine structure. The unambiguous determination of the number of the sulfur atoms in a compound shown in Figure 5.2 is one example of the application of the isotopic fine structure obtained from ultra-high-resolution, high-accuracy MS. Figure 5.2 shows the theoretical simulation of a compound with the formula $\text{C}_{36}\text{H}_{62}\text{N}_5\text{O}_7\text{S}_1$ at m/z 708 calculated at different resolving powers that can be obtained using three different types of mass spectrometers. The insets show expansions of the $A + 2$ isotopic peak. ^{34}S can only be separated from $^{13}\text{C}_2$ in a FT-ICR mass spectrometer with a resolving power of 250 000. The number of sulfur atoms

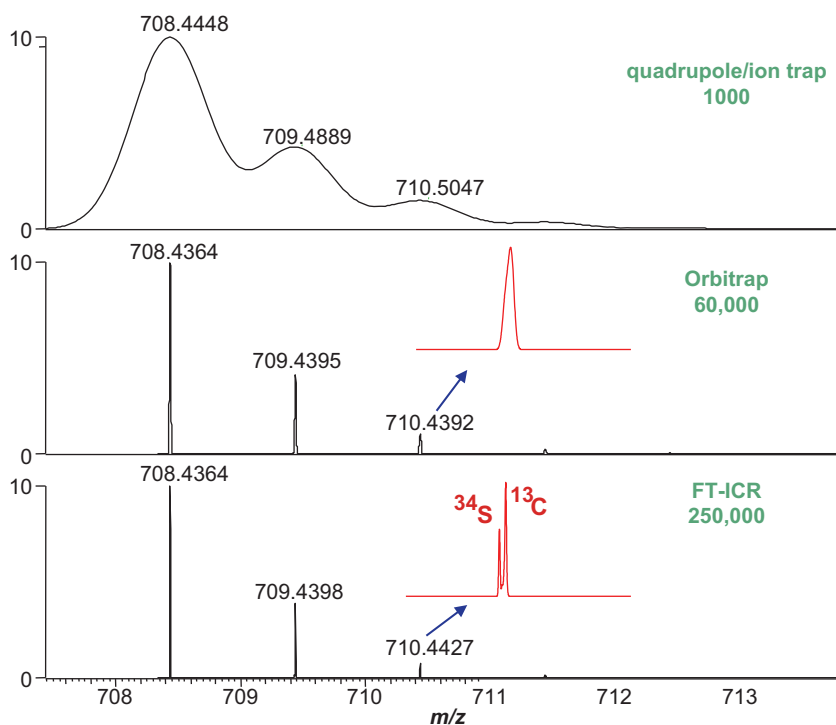


Figure 5.2 The theoretical simulation of the isotope distribution of a compound with the formula $\text{C}_{36}\text{H}_{62}\text{N}_5\text{O}_7\text{S}_1$ at m/z 708. Different mass resolving powers can be obtained *via* three different mass spectrometers.

in the molecule can then be directly determined since the height of the ^{34}S peak relative to the monoisotopic peak is equal to the number of sulfurs multiplied by 2.4%. Determination of the presence of sulfur atoms is critical in the structural identification of unknown natural products. First of all, it significantly narrows down the search of elemental compositions; secondly, sulfur atoms cannot be observed by NMR. Without MS data, sulfur atoms are often missed during NMR data interpretation. FT-ICR is capable of providing mass resolutions of up to 10 000 000 depending on the m/z ratio and the magnetic field. The most challenging application in small-molecule and natural product analysis using high-resolution FT-ICR is to separate ^{34}S from ^{37}Cl , the mass difference of which is 1.25 mDa. To achieve a baseline separation between them, theoretically a resolving power of 640 000 is required for an ion at m/z 400. With this resolving power, separation can also be achieved between ^{18}O and ^{13}C in the $A + 2$ isotopic peak, as well as ^{15}N from ^{13}C in the $A + 1$ isotopic peak.

5.1.2 Brief Description of FT-ICR

FT-ICR MS is the only technique that has sufficient resolving power to routinely give isotopic fine structure for molecules in the molecular weight range of natural products. At the same time, it also confidently measures the relative abundances of the isotopic peak distribution. FT-ICR instruments also offer many MS/MS options for the energetic activation and fragmentation of precursor ions. Furthermore, FT-ICR is an ion trapping instrument and is capable of performing MS^n experiments. These characteristics make FTMS ideally suited for natural product structural analysis. However, the attributes of a FT-ICR instrument do not mean that the information from other types of mass spectrometers is not useful for natural product structure elucidation. Modern Orbitrap and Q-TOF instruments also offer relatively high resolving powers and mass accuracies. In terms of measuring relative abundances, Q-TOF instruments measure the relative isotopic abundances with good fidelity as ions are measured by the electron multiplier. The examples and discussions presented later in this chapter apply to all high-performance mass spectrometers, except where the discussion involves the use of isotopic fine structure measurements, which can generally only be achieved using FT-ICR instruments.

A full discussion of the theory and performance of FT-ICR instruments is beyond the scope of this chapter; however, books and review articles are available for the interested reader.⁸ For the purposes of this discussion, an outline of the basic principles and performance characteristics of FT-ICR that are important for the measurement of natural products will be presented. The FT-ICR mass analyzer is typically a cylindrically shaped cell capped by two trapping plates, with the cylinder walls divided into four quadrants. Opposing pairs of plates in the four quadrants are known as the excite plates and detect plates (Figure 5.3).

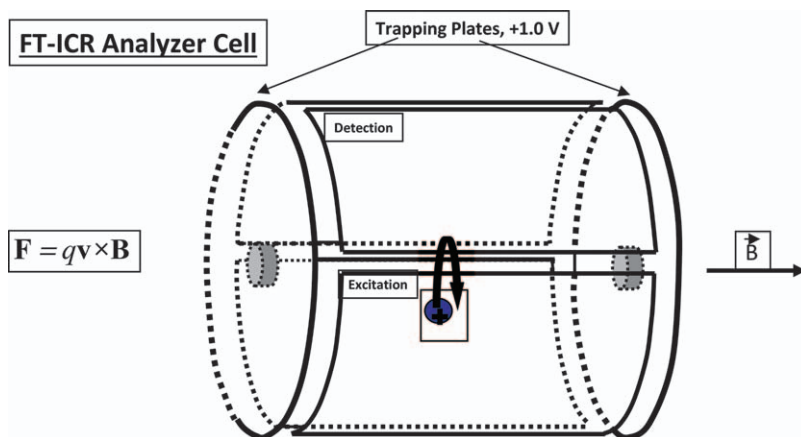


Figure 5.3 The FT-ICR analyzer cell and the motion of ions in the presence of the magnetic field.

The cylindrical mass analyzer cell is located within a vacuum chamber that sits inside the homogenous magnetic field region of a superconducting solenoid magnet. The basic equation of ion motion in FT-ICR is:

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \quad (5.1)$$

where $\mathbf{F}$ is the force on the ion, $\mathbf{v}$ is the ion velocity, and $\mathbf{B}$ is the applied magnetic field that lies along the axis of the cylindrical analyzer cell (Figure 5.3). q is the charge on the ion, and variables in bold font in eqn (5.1) are vector variables. The ion motion equation cross-product is zero when the velocity vector of the ion is parallel to the magnetic field. In a physical sense, this means that the ions are unconstrained by the magnetic field, as they move along the axial axis of the cylindrical mass analyzer cell. For this reason, a “trapping” voltage, typically about +1.0 V, is applied to the trapping plates (a positive trapping voltage for positive ions or a negative trapping voltage for negative ions) in order to trap ions in the analyzer cell.

For velocity components of ions that are not parallel to the magnetic field, the ions experience a force perpendicular to both the magnetic field and the velocity vector of the ions, which causes them to move in a circular path around the magnetic field axis and is known as cyclotron motion (Figure 5.3). The radius of the cyclotron motion depends on the kinetic energy of the ions. Ions are typically injected into the mass analyzer cell, with low energies resulting in small initial cyclotron radii of 0.01–0.1 mm. The frequency, f , of the cyclotron motion is:

$$f = qB/2\pi m \quad (5.2)$$

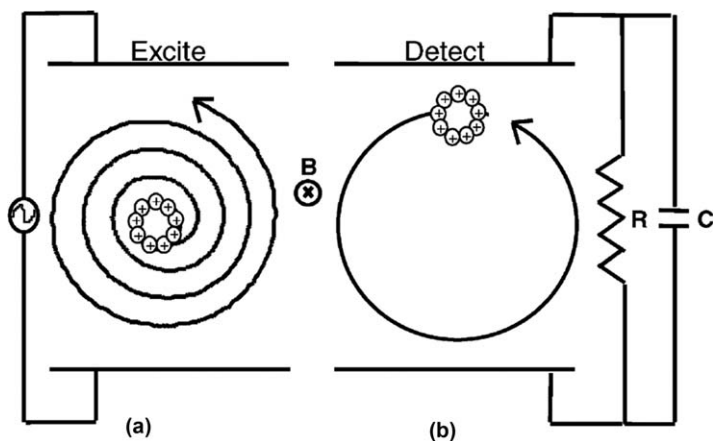


Figure 5.4 (a) Excitation and (b) detection of the ion cyclotron motion.

where B is the magnetic field strength in Tesla, q is the charge on the ion, and m is the mass of the ion.[†] Thus, the mass-to-charge ratio ($m/z = m/q$) can be determined from the frequency of the ion cyclotron motion (f) based on eqn (5.2). The frequency of the cyclotron motion or cyclotron frequency (f) is measured by excitation of the trapped ions. This is achieved through application of a RF field to the excite plates of the analyzer cell that is in resonance with the cyclotron frequency of the ions. By absorbing energy from this RF field, all ions of a given mass-to-charge ratio coherently move together to a larger radius, as shown in Figure 5.4A. The ion cloud induces an image current, which can be measured and amplified as shown in Figure 5.4B. The ion clouds continue to move in a coherent packet at this radius until they are eventually knocked out of phase with each other, which is usually caused by collisions with neutral gases or by space charge effects. Thus, the image current decays with time where the decay time directly depends on the gas pressure in the mass analyzer cell. To achieve high resolution and mass accuracy, ultra-high vacuum (UHV) is required for the analyzer cell of the FT-ICR, as this permits slower signal decay.

A typical peptide mass spectrum is shown in Figure 5.5. The excitation and detection of ions of more than one mass-to-charge ratio leads to a time domain signal (*i.e.* intensity as a function of time) that consists of a superposition of sinusoidal waveforms for each mass-to-charge ratio present within the analyzer cell (Figures 5.5A through 5.5C), analogous to the FID signal obtained in an NMR experiment. Fourier transformation of the

[†]Eqn (5.2) describes the ideal situation (*i.e.* with no applied electric field); however, this is an approximation, as there are electric fields from both the trapping plates and other ions within the analyzer cell, and so calibration equations have been developed in order to correct these electric field effects.

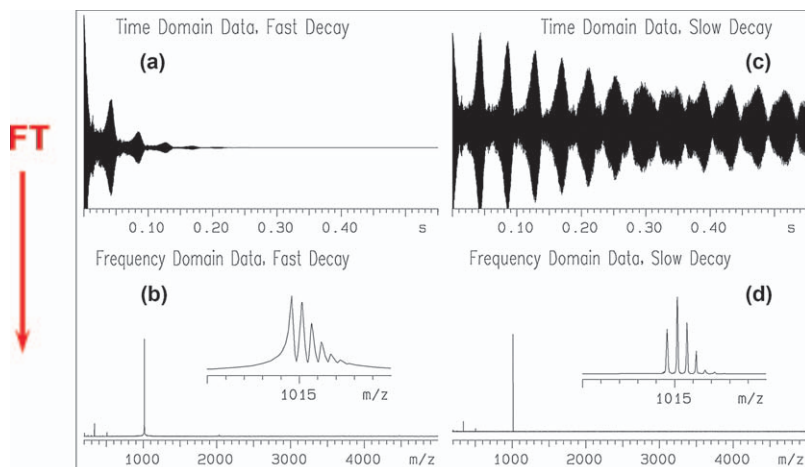


Figure 5.5 Fourier transform and resolution.

time domain signal yields a spectrum that shows the intensity as a function of frequency, also known as the frequency domain spectrum, and as described above, the mass-to-charge ratio ($m/z = m/q$) can then be derived using eqn (5.2) (Figures 5.5B and 5.5D). The high gas pressure in the analyzer cell leads not only to a more rapid decay of the image current signal (Figure 5.5A), but also to poor resolution and peak shape in the mass spectrum (Figure 5.5B). In UHV conditions, the transient signal is much longer (Figure 5.5C) and ideal peak shapes and resolutions can be achieved (Figure 5.5D), offering low- or even sub-ppm mass measurement accuracy.

A typical commercial FT-ICR MS system is pictured in Figure 5.6, with a schematic cutaway showing the regions of differential pumping required to achieve UHV conditions in the mass analyzer cell (Figure 5.6B). Ions are produced at atmospheric pressure by ESI in the source region, then transmitted to the high-vacuum region by RF ion guides through several regions of differential pumping, and are detected in the analyzer with the pressure at 10^{-10} mbar. The ion of interest can also be isolated by a quadrupole in the source region; the quadrupole can also be used for ion transmission, for performing MS/MS experiments in the collision cell, or for other purposes. FT-ICR MS relies on high-field superconducting magnets to achieve its high performance characteristics, and the ability to achieve high resolution and thus isotopically resolved spectra increases with magnetic field strength; therefore, higher-field magnets are preferred. Commercial systems are now available with field strengths of up to 15.0 T, though 7.0 T, 9.4 T, and 12.0 T field strength systems are more common. Other important characteristics of FT-ICR MS include versatility in performing MS/MS experiments and the availability of highly accurate

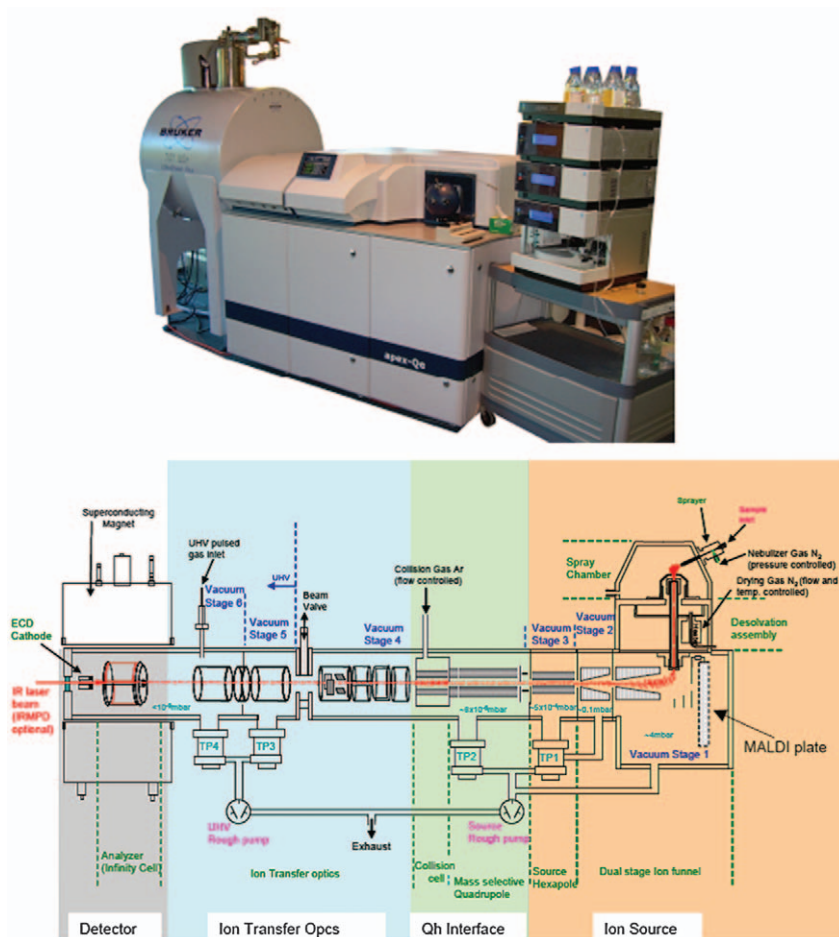


Figure 5.6 A commercial FT-ICR MS system with a nanoflow LC system in front (above), and (below) a cutaway view of the differential vacuum stages that are required to transfer ions formed at atmospheric pressure from the ESI source to the analyzer cell of the FT-ICR MS at 10^{-10} mbar. With permission from Bruker Daltonics, Inc.

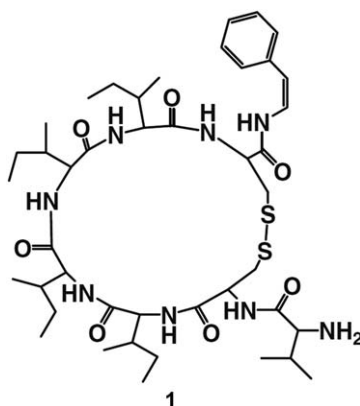
mass measurements. These last two features can be of great benefit in determining the molecular formulae and obtaining the structural information of natural products. Using well-established methods, a single natural product or several adducts can be isolated in the analyzer cell of a FT-ICR mass spectrometer and energized to induce fragmentation using a laser beam or collisions with neutral gas. The fragments can then be measured with high mass accuracy and resolution. The suite of tools offered by FT-ICR MS makes it an ideal instrument for work on natural products, which will be illustrated with examples later in this chapter.

5.2 Examples from Literature

5.2.1 Ultra-high Resolving Power to Separate the Isotopic Fine Structures

5.2.1.1 Determine Number of S Atoms to Reduce the Molecular Formulae Hit for Microcionamide A

Microcionamide A (**1**) was one of the peptides isolated from a number of marine sponge *Clathria abietina* showing high levels of cytotoxicity.²⁰ FT-ICR MS was used to aid the structure determination of these new biologically active peptides. Figure 5.7 shows the high-resolution mass spectrum of microcionamide A at m/z 875.48753, which was acquired at a resolving power of 82 000.²¹ Expansion of the $A + 2$ isotope showed a split of ^{34}S from $^{13}\text{C}_2$ with a mass difference of 10.9 mDa. The relative natural abundance of ^{34}S is $\sim 4\%$ relative to ^{32}S . The number of sulfur atoms present in this peptide can be inferred from the accurate measurement of relative abundance of the ^{34}S peak relative to that of the monoisotopic peak. FT-ICR MS separated the ^{34}S peak from the $^{13}\text{C}_2$ peak, allowing the determination of the number of the sulfur atoms by comparing the relative ratio of the ^{34}S peak to the monoisotopic peak. As shown in Figure 5.7, the measured relative abundance of ^{34}S in the $A + 2$ isotope (m/z 877.48243) was $\sim 8\%$ relative to the monoisotopic peak (which contains ^{12}C and ^{32}S , among other isotopes) at m/z 875.48753. The number of sulfur atoms in the microcionamide A was thus determined to be two. Knowing the absolute number of sulfur atoms significantly reduces the number of molecular formulae within the known experimental mass accuracy. As a result, the molecular formula for **1** shown in Figure 5.7 was quickly determined to be $\text{C}_{43}\text{H}_{70}\text{N}_8\text{O}_7\text{S}_2$ with a mass error of -0.63 mDa between the experimental and theoretical data.



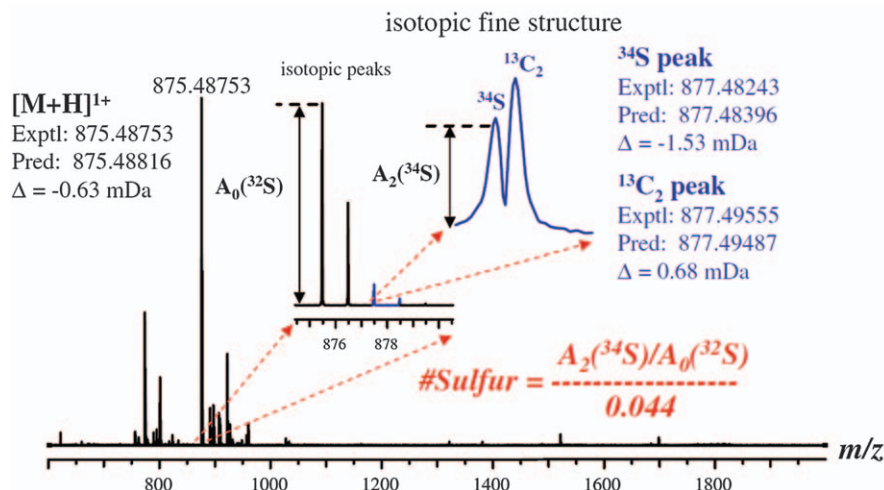
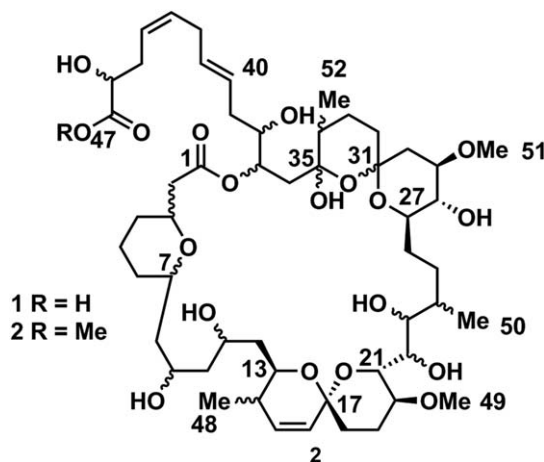


Figure 5.7 The mass spectrum of the protonated molecule of microcionamide A (1) at m/z 875.²¹

(X. Feng and M. M. Siegal, FTICR-MS applications for the structure determination of natural products, *Analytical and Bioanalytical Chemistry*, 2007, **389**, 1341–1363, © Springer-Verlag 2007, with kind permission from Springer Science and Business Media.)

5.2.1.2 Combining High-resolution, Highly Accurate MS and MS^n to Solve a Complicated Natural Product Structure—Methylspirastrellolide A

Methylspirastrellolide A (2) is a novel antimitotic macrolide isolated from the Caribbean marine sponge *Spirastrella coccinea*. This novel compound



was initially characterized using NMR by Andersen and coworkers as shown in structure **2b**.²² This structure was obtained based primarily on NMR data and HRCIMS data. The accurate mass data of the methyl ester **2b** generated via a HRCIMS instrument was at m/z 1027.5805 ($[M + H]^+$), corresponding to an empirical formula of $C_{53}H_{86}O_{19}$.

The high-resolution FT-ICR MS measurement later obtained different molecular weight information, which is inconsistent with the previously published structure.^{21,23} As shown in Figure 5.8, the accurate mass of the sodiated molecule is at m/z 1049.52054 ($[M + Na]^+$) and the accurate mass of the deprotonated molecular ion was also observed at m/z 1025.52476. The accurate mass data were inconsistent with the molecular formula $C_{53}H_{86}O_{19}$. As illustrated in Figure 5.8, the $A + 2$ isotopic peak had a split, which was due to the separation of ^{37}Cl from $^{13}\text{C}_2$ at the resolving power of 100 000. The data suggest that one chlorine atom was present in the compound, indicating a new molecular formula, $C_{53}H_{83}O_{17}\text{Cl}$, with an error of -0.47 mDa. It is worth noting that unlike small molecules, where it is generally easy to pick out the characteristic $^{35}\text{Cl}/^{37}\text{Cl}$ isotopic pattern, in this large molecule, the $A + 2$ isotope is significantly lower in abundance than that of the $A + 1$ isotope. The smaller relative abundance of the $A + 2$ isotope is due to the

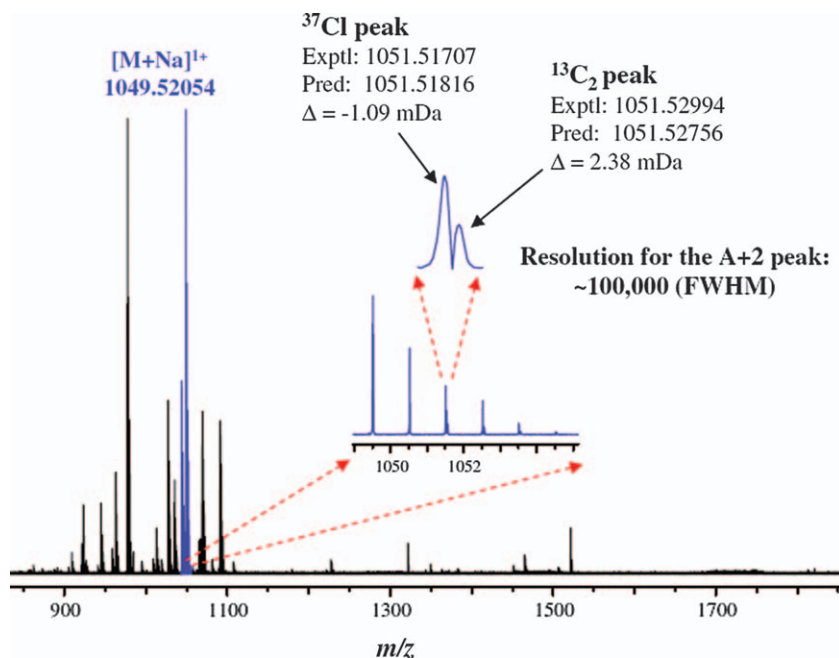
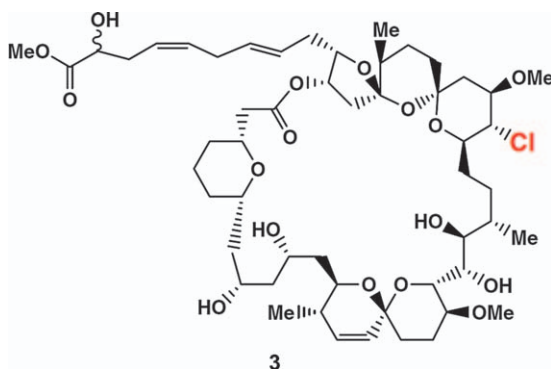


Figure 5.8 The molecular structure and positive ion mode nano-ESI FT-ICR mass spectrum of methylspirastrellolide A (**3**).²¹

(X. Feng and M. M. Siegal, FTICR-MS applications for the structure determination of natural products, *Analytical and Bioanalytical Chemistry*, 2007, **389**, 1341–1363, © Springer-Verlag 2007, with kind permission from Springer Science and Business Media.)

relatively large number of carbon atoms. In this particular case, there are 53 carbon atoms in the molecule, which contributes 58.3% of the ^{13}C in total (natural abundance of one ^{13}C is 1.1%, which constitutes a majority of the intensity of the $A + 1$ isotopic peak). In terms of the $A + 2$ isotope, the major contributions come from ^{37}Cl and $^{13}\text{C}_2$. The natural abundance of one ^{37}Cl is 32%, whereas one $^{13}\text{C}_2$ peak is 0.12% relative to the intensity of the mono-isotopic peak. With 53 carbons, the $A + 2$ isotope is the sum of three major contributors: 32% from ^{37}Cl , 15.04% from $^{13}\text{C}_2$, and 3.4% from ^{18}O . The mass difference between ^{37}Cl and $^{13}\text{C}_2$ is 10 mDa, while the mass difference between $^{13}\text{C}_2$ and ^{18}O is 2.5 mDa. The resolving power in this example was not able to separate $^{13}\text{C}_2$ from ^{18}O . Therefore, the doublet observed in the inset in Figure 5.8 consists of one peak from ^{37}Cl and the other peak is from the $^{13}\text{C}_2$ and ^{18}O peaks combined. Nevertheless, separating ^{37}Cl from $^{13}\text{C}_2$ made it possible to determine the correct molecular formula for this complicated natural product spirastrellolide A, which led to the determination of the correct structure as shown in structure 3.

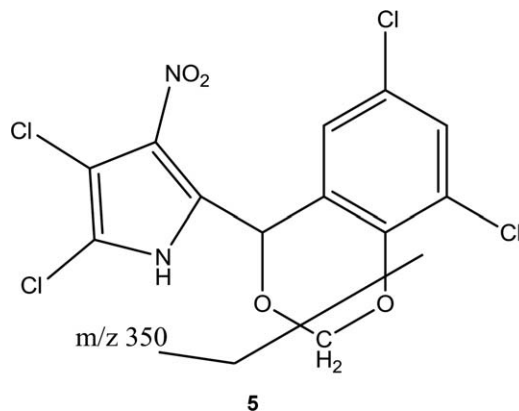


In order to confirm the accuracy of the new structure, MS/MS experiments on the sodiated molecule at m/z 1049.52063 were conducted. An interesting fragment ion at m/z 663.37110 was observed, which corresponds to an empirical formula of $\text{C}_{34}\text{H}_{56}\text{O}_{11}\text{Na}^+$ with an error of -0.38 mDa, while another fragment ion at m/z 386.14953 was observed with an empirical formula of $\text{C}_{19}\text{H}_{27}\text{O}_6\text{Cl}$ with an error of -0.09 mDa. The sum of the elemental composition of these two fragment ions gives the correct elemental composition of the precursor ion for the molecular formula of $\text{C}_{53}\text{H}_{83}\text{O}_{17}\text{Cl}$, which provided additional confirmation for structure 3. In addition, a MS/MS experiment was also conducted on its protonated form at m/z 1027.53971. The major fragment ion at m/z 977.50294 was the result of the loss $\text{CH}_3\text{OH} + \text{H}_2\text{O}$ from the molecule ion. A MS^3 experiment of the fragment ion at m/z 977.50294 was then performed, which producing fragment ions with a series of water losses (up to nine water molecules were observed).²¹ Many of these fragment ions were detected as doublets under ultra-high resolution, which would have been detected as singlet peaks under low resolution. The exact mass of the doublet at m/z 923 demonstrated that the lower mass of the doublet peak of m/z 923 was the result of the loss of $3\text{H}_2\text{O}$ with the exact

mass of 54.0317, while the higher mass of m/z 923 was generated *via* the loss of $\text{HCl} + \text{H}_2\text{O}$ with the exact mass of 53.9872 from the MS/MS product ion (m/z 977). The mass difference between $3\text{H}_2\text{O}$ and $\text{HCl} + \text{H}_2\text{O}$ was 0.0445 Da, which requires a resolving power of 21 000 to separate these two peaks. These data provided additional evidence for the presence of a Cl atom in the molecule.

5.2.2 Study of Biosynthesis Processes by Measuring Stable Isotopic-labeled Precursors using FT-ICR MS

Another application for ultra-high-resolution MS is in the study of biosynthetic processes by measuring stable isotopic-labeled compounds. One example shown here is a pyrrole-containing antibiotic dioxapyrrolomycin (**5**), which is active against Gram-positive bacteria.²⁴ The biosynthetic mechanism for the nitration of **5** was studied by feeding a series of ^{13}C and ^{15}N stable isotope-labeled precursors. The method of choice for the monitoring and the determination of both the extent and position of enrichment of the ^{13}C - and ^{15}N -labeled **5** was high-resolution FT-ICR MS.



Negative ion MS/MS data of unlabeled deprotonated dioxapyrrolomycin (**5**) at m/z 380 and mono- ^{13}C -labeled deprotonated dioxapyrrolomycin (**5**) at m/z 381^{21,24} show the same fragment ion at m/z 351 for both compounds, which was the result of the loss of $\text{H}_2\text{C}=\text{O}$ and $\text{H}_2^{13}\text{C}=\text{O}$, respectively. The data indicate that the carbon in between two oxygens contains the entire ^{13}C enrichment. The extent of ^{13}C labeling for **5** therefore could be calculated as shown by Charan *et al.*^{21,24}

Determination of the ^{15}N -labeled dioxapyrrolomycin is much more challenging. The theoretical m/z for the deprotonated mono- ^{13}C -labeled dioxapyrrolomycin $\text{C}_{11}^{13}\text{H}_5\text{Cl}_4\text{N}_2\text{O}_4$ $[\text{M} - \text{H}]^-$ is 381.90426, while that for the deprotonated mono- ^{15}N -labeled dioxapyrrolomycin $\text{C}_{12}\text{H}_5\text{Cl}_4\text{N}^{15}\text{NO}_4$ $[\text{M} - \text{H}]^-$ is 381.89793. The mass difference between the two is 6.9 mDa and a resolving power above 100 000 is required in order to separate them. As shown in Figure 5.9, ^{15}N and ^{13}C isotopes were separated by FT-ICR MS at a resolving power of 100 000. The high mass accuracy confirmed the ^{15}N

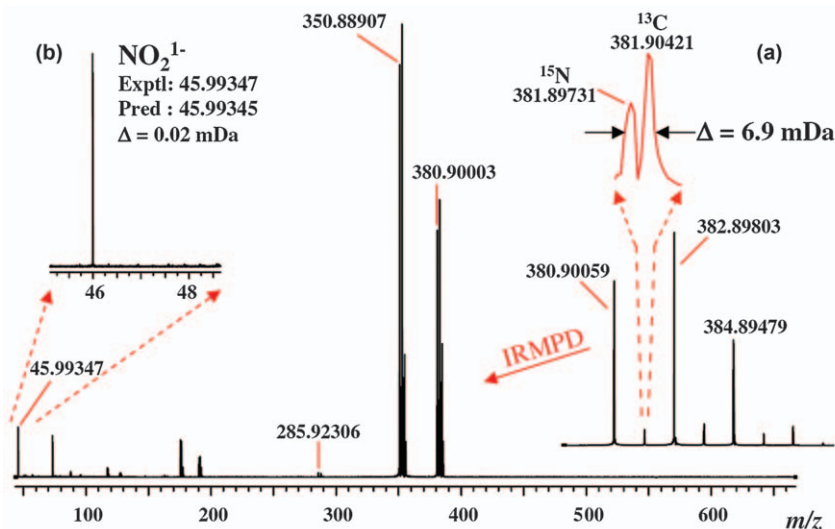


Figure 5.9 Negative mode ESI FT-ICR IRMPD MS/MS spectrum for the ^{15}N -enriched dioxapyrrolomycin (5).²⁴ (Reprinted with permission from R.D. Charan, G. Schlingmann, V. S. Bernan, X. D. Feng, and G. T. Carter, *J. Nat. Prod.* 2006, **69**(1), 29. Copyright (2006) American Chemical Society.)

isotopic enrichment with high confidence. This could easily be mis-assigned as the naturally occurring ^{13}C species using a low-resolution spectrum, which would only show a single peak. In addition, IRMPD MS/MS experiments were further conducted in order to determine the specificity of the N labeling. Among the product ions, the presence of the NO_2^- species (m/z 45.99347) with a strong signal-to-noise ratio (Figure 5.9B) and the absence of the $^{15}\text{NO}_2^-$ species strongly indicated that the nitro group was not ^{15}N labeled. The average relative abundance for the ^{15}N ($A+1$) peak was 9.81% of the monoisotopic ion peak at m/z 381, and therefore the extent of the ^{15}N labeling could be estimated.^{21,24}

5.2.3 Metabolite Profiling of Triterpene Saponins by Combined Accurate Mass Measurement and MSⁿ Experiments using LC FT-ICR MS

As discussed in Section 51.1.2.1, with better mass accuracy, it becomes easier to assign molecular formulae. Here, we present an example of using LC/MS accurate mass measurement and LC/MS/MS experiments to quickly determine the molecular formulae and tentative structures of 79 different triterpene saponins from the hairy roots of the model *Medicago truncatula*.²⁵ Triterpenes are one of the largest classes of plant natural products. Due to their vast variety of structures and bioactivities, it is very important to identify their metabolite structures. Figure 5.10 shows the deprotonated molecule of an unknown saponin ($\text{RT} = 23.19$ minutes) at m/z 945.50956.

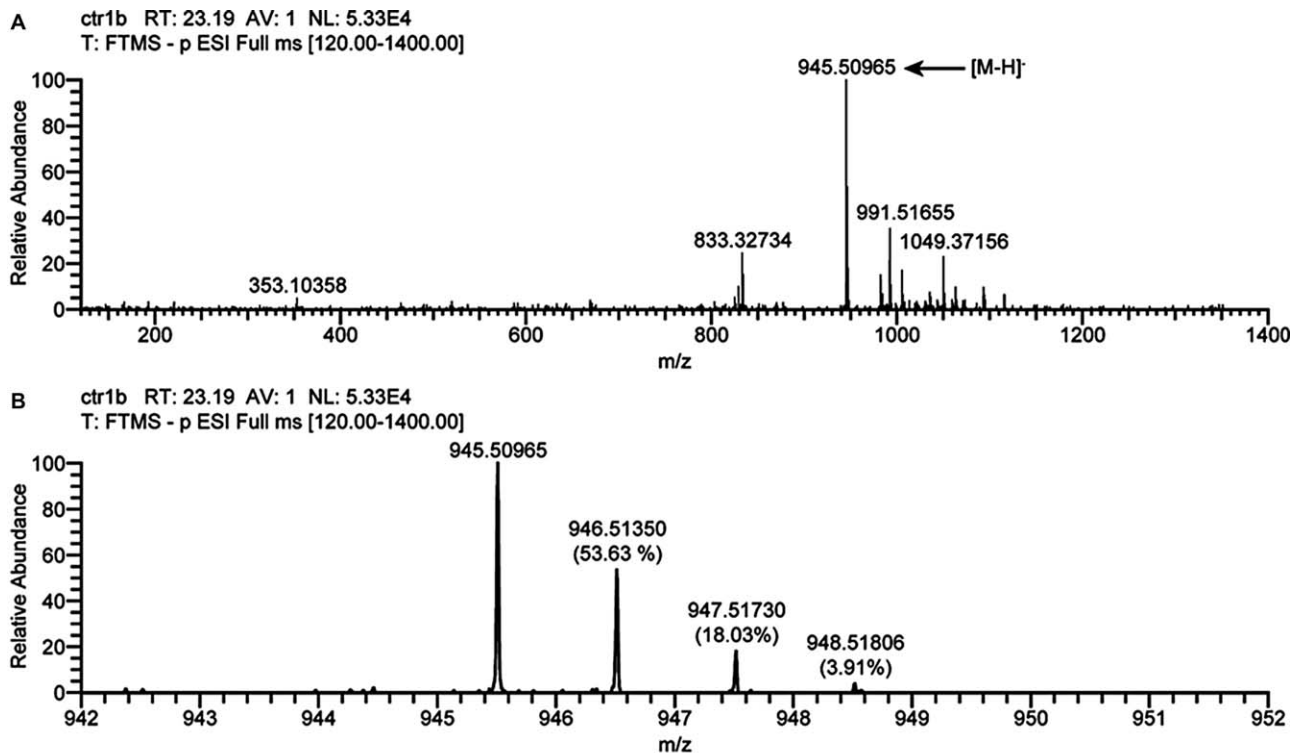


Figure 5.10 (A) High-resolution, accurate MS spectrum of the unknown at RT = 23.19 min. (B) The expansion of the isotope cluster at m/z 945, showing that the protonated molecule is accompanied by additional isotopic ions. The relative intensities of the isotopic ions are indicated in brackets.²⁵
(Reprinted with permission from J. Pollier, K. Morreel, D. Geelen, A. Goossens, *J. Nat. Prod.* 2011, 74(6), 1462. Copyright (2011) American Chemical Society.)

Three possible molecular formulae were generated *via* the computer software within a 5 ppm mass error window: $C_{65}H_{69}O_6$ (-0.33 ppm), $C_{40}H_{81}O_{24}$ (-2.83 ppm), and $C_{47}H_{77}O_{19}$ (3.38 ppm). Assignment of the correct formula from these three possibilities was achieved by evaluating the relative intensity between the $A + 1$ isotope peak and the monoisotopic peak. As discussed in Section 5.1.1.4, carbon in nature contains 98.89% ^{12}C and 1.11% ^{13}C . Similarly, hydrogen and oxygen in nature contain 0.015% 2H and 0.038% ^{17}O , respectively. For instance, if a compound contains 65 carbon atoms, 69 hydrogen atoms and 6 oxygen atoms, the relative intensity between the $A + 1$ peak and the monoisotopic peak is calculated as 73.4% (from the sum of the contributions of ^{13}C , 2H , and ^{17}O). Accordingly, the theoretical relative abundances of the $A + 1$ isotopic peaks for $C_{65}H_{69}O_6$, $C_{40}H_{81}O_{24}$, and $C_{47}H_{77}O_{19}$ are 73.41%, 46.53%, and 54.05%, respectively.²⁵ As shown in Figure 5.10B, the experimental relative abundance of the $A + 1$ isotope was 53.63%, which suggests that the correct molecular formula for the unknown at $RT = 23.19$ minutes is $C_{47}H_{77}O_{19}$ ($\Delta = 3.38$ ppm). In the same manner, the molecular formulae for all of the other unknown saponins were readily determined.²⁵ Low-mass accuracy MS/MS measurements were then used to assign molecular structures that fit the assigned molecular formulae, as described in the ref. 25.

5.3 Conclusions

FT-ICR MS can provide extremely useful information for the structural elucidation of nature products. Ultra-high mass accuracy allows unambiguous assignment of the molecular formulae of small molecules. High-resolution MS also provides additional unique information regarding the isotopic fine structure, such as the presence of sulfur and chlorine, and yields the exact number of atoms for that particular element. With the combination of multistage MS/MS, it is also possible to obtain unique molecular formulae for large molecules by determining the molecular formulae of the fragment ions of the large molecules.

References

1. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, *Science*, 1989, **246**(4926), 64.
2. M. Karas, D. Bachmann and F. Hillenkamp, *Anal. Chem.*, 1985, **57**(14), 2935.
3. F. E. Koehn and G. T. Carter, *Nat. Rev. Drug Discovery*, 2005, **4**(3), 206.
4. F. W. McLafferty, *Acc. Chem. Res.*, 1980, **13**(2), 33.
5. J. K. Eng, A. L. McCormack and J. R. Yates, *J. Am. Soc. Mass Spectrom.*, 1994, **5**(11), 976.
6. D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2007, **70**(3), 461.
7. A. Makarov, *Anal. Chem.*, 2000, **72**(6), 1156.

8. A. G. Marshall, C. L. Hendrickson and G. S. Jackson, *Mass Spectrom. Rev.*, 1998, **17**(1), 1.
9. J. Laskin and J. H. Futrell, *Mass Spectrom. Rev.*, 2005, **24**(2), 135.
10. R. A. Zubarev, *Mass Spectrom. Rev.*, 2003, **22**(1), 57.
11. S. A. McLuckey and J. L. Stephenson, *Mass Spectrom. Rev.*, 1998, **17**(6), 369.
12. X. Yu, Y. Huang, C. Lin and C. E. Costello, *Anal. Chem.*, 2012, **84**(17), 7487.
13. M. D. A. Mabud, M. J. Dekrey and R. G. Cooks, *Int. J. Mass Spectrom.*, 1985, **67**(3), 285.
14. R. G. Cooks, T. Ast and A. Mabud, *Int. J. Mass Spectrom.*, 1990, **100**, 209.
15. T. Ast, M. A. Mabud and R. G. Cooks, *Int. J. Mass Spectrom.*, 1988, **82**(1–2), 131.
16. G. Tsaprailis, H. Nair, A. Somogyi, V. H. Wysocki, W. Q. Zhong, J. H. Futrell, S. G. Summerfield and S. J. Gaskell, *J. Am. Chem. Soc.*, 1999, **121**(22), 5142.
17. W. Q. Zhong, E. N. Nikolaev, J. H. Futrell and V. H. Wysocki, *Anal. Chem.*, 1997, **69**(13), 2496.
18. J. A. Yergey, *Int. J. Mass Spectrom. Ion Processes*, 1983, **52**(2–3), 337.
19. A. L. Rockwood, S. L. Vanorden and R. D. Smith, *Anal. Chem.*, 1995, **67**(15), 2699.
20. R. A. Davis, G. C. Mangalindan, Z. P. Bojo, R. R. Antemano, N. O. Rodriguez, G. P. Concepcion, S. C. Samson, D. de Guzman, L. J. Cruz, D. Tasdemir, M. K. Harper, X. Feng, G. T. Carter and C. M. Ireland, *J. Org. Chem.*, 2004, **69**(12), 4170.
21. X. Feng and M. M. Siegel, *Anal. Bioanal. Chem.*, 2007, **389**(5), 1341.
22. D. E. Williams, M. Roberge, R. Van Soest and R. J. Andersen, *J. Am. Chem. Soc.*, 2003, **125**(18), 5296.
23. D. E. Williams, M. Lapawa, X. Feng, T. Tarling, M. Roberge and R. J. Andersen, *Org. Lett.*, 2004, **6**(15), 2607.
24. R. D. Charan, G. Schlingmann, V. S. Bernan, X. D. Feng and G. T. Carter, *J. Nat. Prod.*, 2006, **69**(1), 29.
25. J. Pollier, K. Morreel, D. Geelen and A. Goossens, *J. Nat. Prod.*, 2011, **74**(6), 1462.

CHAPTER 6

Current Pulse Sequence Developments in Small-molecule Nuclear Magnetic Resonance Spectroscopy

TEODOR PARELLA

Servei de Ressonància Magnètica Nuclear, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Barcelona), Catalonia, Spain
Email: teodor.parella@uab.cat

6.1 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a powerful and a very exciting analytical tool for solving many structural and dynamics chemical issues of small molecules in solution.^{1–4} NMR has been and is persistently updated and improved by the incorporation of new developments in instrumentation, software, and methodology. The most expensive advances in NMR instrumentation are mainly focused on the availability of higher magnetic fields that offer better sensitivity and larger spectral dispersion.⁵ Other developments involve improved electronics, computer technology, and NMR accessories, with a particular emphasis on the design of cryogenically cooled probes that provide sensitivity gains of up to a theoretical factor of four,^{6,7} capillary and microprobes that allow working at the sub-milligram level using very reduced nanoliter- to microliter-size active volumes,^{6–9} the use of restricted-volume NMR tubes to increase mass

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

sensitivity, or the use of NMR tubes of smaller diameter to reduce thermal noise at the expense of fill factor. In addition, the flexibility to update basic spectrometer configurations with other NMR capabilities, such as HR-MAS and solid-state techniques, the use of combined hyphenated analytical techniques (LC/MS-NMR),¹⁰ or the current compatibility with sensitivity-enhanced dynamic nuclear polarization transfer techniques¹¹ allows the successful application of NMR to a broad range of samples and experimental conditions. On the other hand, advances in NMR software provide user-friendly interfaces for automating data acquisition and data processing, which greatly facilitates an easy set-up and performance, even for non-specialized NMR users. Important novel approaches are focused on automated NMR data analysis using computer-assisted structure elucidation (CASE) programs^{12,13} and on the analysis of complex mixtures containing multiple components, with special interest in metabonomic studies.

Another essential point in order to understand the huge potential and versatility of NMR spectroscopy is its tremendous capacity to manipulate spin systems in an incredible number of different ways using dedicated pulse schemes. In modern spectrometers, a complete library containing hundreds of pulse sequences is available for a routine use. In terms of NMR methodology, a basic strategy for the structure elucidation and characterization of small molecules is well established. It consists of the concerted analysis of a suite of basic 1D and 2D NMR experiments that provide complementary chemical information. Thus, after the analysis of the regular 1D ^1H and 1D $^{13}\text{C}\{^1\text{H}\}$ spectra, high-sensitivity COSY and/or TOCSY experiments quickly establish through-bond J_{HH} connectivities and facilitate ^1H chemical shift assignments. Heteronuclear correlation HSQC and HMBC experiments provide direct ($^1J_{\text{CH}}$) and long-range heteronuclear ($^nJ_{\text{CH}}$; $n > 1$) connectivities, respectively, which help trace out the carbon skeleton. NOESY/ROESY experiments afford through-space ^1H - ^1H correlations in order to obtain 3D molecular views and to determine relative configurations and some dynamic aspects such as chemical exchange or conformational behavior. A considerable number of additional NMR experiments are also available to obtain diverse and complementary information. For instance, diffusion/DOSY experiments provide helpful insights into molecular sizes, and its application has proved to be highly useful for monitoring and characterizing intermolecular interactions, for exploring reaction mechanisms by studying intermediate compounds with short lifespans and/or with uncertain stabilities, for characterizing major and minor compounds from crude reactions and complex mixtures without the need for further purification, or for characterizing degradation products, among others. Another interesting aspect for NMR users is the quantitative determination of NMR parameters (like homonuclear and heteronuclear coupling constants, relaxation times, or diffusion coefficients) with simplicity and accuracy, without the use of sophisticated post-processing protocols. On the other hand, recent interest has been focused on the determination of residual dipolar coupling (RDC) constants of small molecules dissolved in weakly

aligned anisotropic media to obtain long-distance structural information in the form of relative orientational chemical bond restraints.^{14–17} Important advances have been made in sample preparation, the design of appropriate solvent-dependent anisotropic media, and the development of novel NMR methods for the accurate and precise measurement of experimental RDCs.

After three decades of fruitful developments, the field of small-molecule NMR has achieved an optimal level in terms of the number and the robustness of available NMR experiments. Current advances in pulse sequences are focused on some general objectives (Figure 6.1): (i) the improvement of signal sensitivity (signal-to-noise ratio; SNR) in order to work with micrograms/nanograms of a sample; (ii) economizing spectrometer time by collecting NMR data much faster; (iii) a better performance that offers general applicability in a wide range of experimental conditions and samples, allowing for full automation without previous or with minimum set-up and minimizing user intervention in both data acquisition and data processing steps; (iv) methods to achieve a more efficient resolution/editing/dispersion between NMR signals; (v) approaches to get a more accurate and easy determination of NMR parameters; and (vi) the design of novel pulse schemes that afford new, additional, multiple, complementary, or specific NMR information compared with the existing NMR experiments.

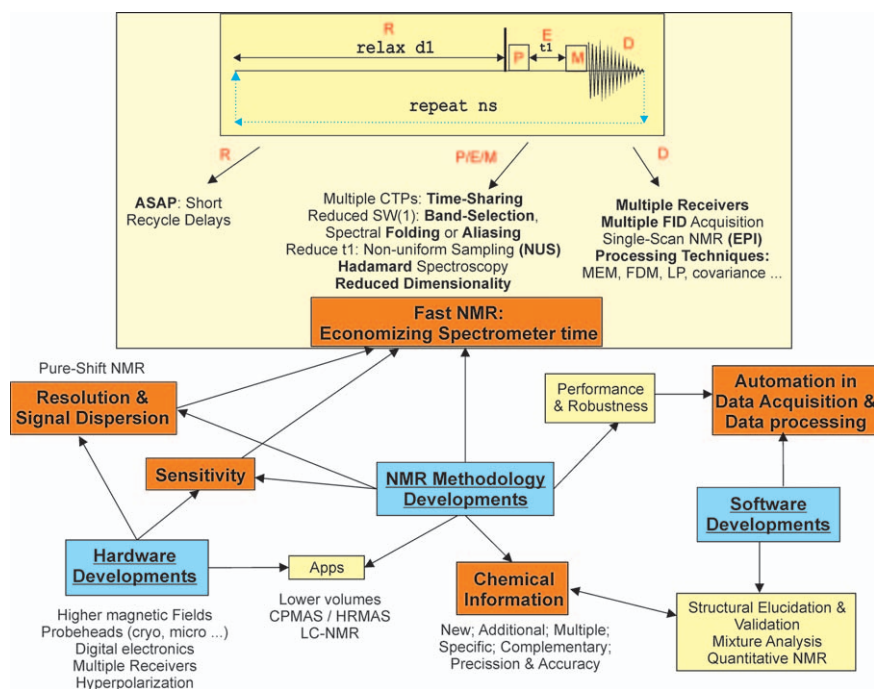


Figure 6.1 General scheme summarizing the main areas involving pulse sequence design, hardware and software developments in small-molecule NMR spectroscopy.

In this chapter, a review of the latest developments in the design and application of novel NMR experiments for the structural characterization of small molecules in solution is presented. Specific emphasis is placed on new concepts, NMR pulse sequence elements, and novel pulse schemes that have been developed to perfect existing experiments or to afford new NMR solutions. First, several “fast NMR” techniques are overviewed. Then, the very old but exciting idea about how to obtain broadband ^1H -homodecoupled NMR spectra is introduced by describing some modern implementations. The incorporation of this “pure shift NMR” concept in existing experiments is described and illustrated with some practical applications. The advent of ultra-high-resolution NMR spectroscopy based on the combination of several resolution-enhanced techniques in a single NMR experiment is also evaluated as a promising tool for analyzing very complex mixtures and highly congested spectra. Additionally, the design of “perfect NMR” experiments with the aim of achieving high-quality spectra with non-distorted cross-peaks that are amenable to a more accurate analysis is also outlined. The description of the use of robust and, at the same time, simple NMR building blocks, such as perfect-echo elements or a zero-quantum filter (ZQF), is exemplified for the accurate and straightforward extraction of NMR parameters. Finally, a compilation of novel long-range correlation experiments that may become excellent complements to conventional HMBC/HSQMBC experiments and their use in CASE programs will be presented. It is recommended that the reader refers to a review covering developments up until November 2012 for identifying and characterizing natural products as a complement to this chapter.²

6.2 Fast NMR

The proper optimization of spectrometer times is a major issue in any NMR laboratory. Beyond the proper choice and correct set-up of the NMR experiments to be executed, there is a minimum experimental time that can not be avoided. A major breakthrough in NMR pulse sequence design was the incorporation of pulsed-field magnetic gradients (PFGs) as key elements for the efficient selection of specific coherence transfer pathways and for purging undesired contributions.¹⁸ In practice, this development changed the concept of routine NMR, allowing high-quality, noise-free NMR spectra to be collected under less demanding experimental conditions and using shorter acquisition times. In recent years, several new strategies have been proposed to accelerate data acquisition. Techniques that fall into this category include: (i) the reduction of the usually long pre-scan delay using acceleration by sharing adjacent polarization (ASAP);^{19–22} (ii) the acquisition of multiple datasets in a single NMR experiment (single-scan NMR,^{23–26} time-sharing [TS]-NMR,²⁷ multiple-FID acquisition,²⁸ the use of multiple receivers,^{29,30} *etc.*); (iii) the reduction of the number of variable time increments required to monitor indirect dimensions in multidimensional NMR (use of spectral folding or aliasing,^{31–33} non-uniform sampling [NUS]

techniques,^{34–40} region-selective excitation, Hadamard spectroscopy, reduced dimensionality experiments, *etc.*); (iv) the use of enhanced NMR instrumentation to improve sensitivity and performance (cryoprobes, high magnetic fields, and hyperpolarized techniques); and (v) the complementary use of advanced processing mathematical tools such as forward linear prediction, maximum entropy reconstruction, filter diagonalization methods,⁴¹ or generalized indirect covariance co-processing.^{42,43}

In absolute terms, the initial recovery delay represents a large percentage of the overall experimental time of any pulse sequence. Thus, the use of a reduced recovery delay in between scans is a simple approach to shortening measurement times. In order to avoid important sensitivity losses by signal saturation, the ASAP technique¹⁹ has been combined with optimized Ernst angle-type excitation in heteronuclear experiments.^{20–22} For instance, a modified HSQC pulse scheme (ASAP-HSQC)²⁰ that includes a short homonuclear isotropic mixing period of approximately 40 ms as a recovery delay to transfer polarization from nearby ^1H – ^{12}C spins to ^1H – ^{13}C spins, an optimized INEPT transfer, and a modified gradient scheme that retains the polarization reservoir for all of the passive spins not directly bound to the NMR-active heteronucleus has been proposed to speed up the acquisition of HSQC spectra. Thus, an ASAP-HSQC spectrum of menthol was quickly recorded in 28 seconds using a 512×128 data point matrix and one scan per t_1 increment and 16 dummy scans. A related clean in-phase (CLIP)-ASAP-HSQC was subsequently proposed for the fast determination of $^1J_{\text{CH}}$ along the F_2 dimension, including RDCs of partially aligned molecules, in about 25 seconds.²² These experiments can be further improved by using NUS and are also of potential interest for the rapid screening of natural products.⁴⁴

Ultrafast (UF) or single-scan NMR is another exciting concept initially developed by Frydman and coworkers.^{23–26} It is based on the equivalent acquisition of different t_1 increments of a 2D experiment simultaneously in different parts of the NMR tube. The method uses a one-shot spatial encoding strategy, where slice selection along the z -axis is performed by the conjuncted application of z -axis PFGs and frequency-modulated pulses. In this way, a 2D experiment can be recorded in a single scan or, in other words, in less than 1 second. The uses and applications of the UF NMR methods for monitoring fast organic reactions and dynamic processes has been recently reviewed,⁴⁵ although the techniques do present some practical drawbacks: (i) they have very low sensitivity due to the slice selection process and J -modulation effects; (ii) they have low resolution and limited spectral widths (SWs) along the indirect dimension due to the limitation of current gradient capabilities; and (iii) they are difficult to set up and automate for general and routine use. Alternatively, hybrid UF techniques can be used to improve the performance of multidimensional UF NMR while remaining faster than conventional acquisitions.⁴⁶ Much better results in terms of SNR can be achieved with the multi-scan single-shot approach or with interleaved acquisitions. Recently, the covariance method and a pattern recognition

algorithm have been applied in order to improve resolution in the spatially encoded dimension of homonuclear UF TOCSY experiments.⁴⁷

The development of NMR methods to obtain additional and/or multiple forms of information in a given spectrometer period and/or from a single NMR experiment is another topic of current interest. TS methods rely on the simultaneous evolution of multiple frequencies in the same period of a NMR pulse sequence and they require a triple-resonance spectrometer configuration.²⁷ Thus, ^1H , ^{13}C and ^1H , ^{15}N correlation spectra (TS-HSQC,⁴⁸ TS-HSQC-TOCSY,⁴⁹ TS-HSQMBC,⁴⁹ or TS-HMBC⁵⁰) have been simultaneously recorded in nitrogen-containing molecules. This approach has shown potential in the simultaneous measurement of one-bond and/or long-range J_{CH} and J_{NH} coupling constants from F_2 -coupled versions of these TS-HSQC, TS-HSQMBC, or TS-HSQC-TOCSY experiments, and of both the magnitude and the sign of a complete set of multiple heteronuclear coupling constants (J_{CH} , J_{NH} , J_{CF} , J_{NF} , and J_{FH}) from a TS-selHSQMBC-TOCSY experiment on fluorine organic compounds.⁵¹

The concept of sequential multiple-FID acquisitions within the same scan also allows the simultaneous one-shot collection of multiple NMR spectra. This approach requires the presence of free evolution delays in different parts of the pulse sequence to accommodate different data acquisition periods. The major inconvenience of this approach is associated with the duration of each FID, which must sometimes be short. For instance, multiple-FID acquisition and TS concepts have been combined in a single ^1H , ^{13}C , and ^1H , ^{15}N HMBC/HMBC-COSY experiment.²⁸ Another example involves the simultaneous acquisition of a reference and a T_2 -filtered ^{19}F experiment to afford a two-fold reduction in the measurement time of 1D ^{19}F spectra in screening studies.⁵² Similar ideas have been applied to NMR systems incorporating multiple receivers, allowing the simultaneous acquisition of several NMR spectra in a single shot, referred to as parallel acquisition NMR spectroscopy.^{29,30} These experiments can require additional hardware (standard “triple-resonance” or broadband probes equipped with a separate receiver for each nuclear species under investigation) that is not available in most conventional spectrometer configurations.⁵³ The original publication described the simultaneous recording of proton-detected ^1H - ^1H COSY or TOCSY and carbon-detected CH correlations.²⁹ On the other hand, 2D ^{13}C - ^{13}C INADEQUATE, single- or multiple-bond correlation spectra, and a standard 1D $^{13}\text{C}\{^1\text{H}\}$ spectrum, all at the same time, are simultaneously collected in the PANACEA experiment.⁵⁴ The method requires a cryogenically cooled probe with optimized ^{13}C detection and it can be extended to ^{15}N using three parallel receivers and triple-resonance capabilities. A modified HR-PANACEA affords additional HMBC data⁵⁵ and its application to ^{29}Si NMR has also been reported.⁵⁶ Triple-resonance ^1H -X- ^{19}F experiments for the parallel acquisition of ^1H -X and ^{19}F -X (X = ^{13}C or ^{15}N) HSQC and HMBC correlation experiments, using simultaneous ^1H and ^{19}F detection, has been described.⁵⁷ Parallel multiple-receiver acquisition has been combined with UF NMR in the so-called parallel UF NMR spectroscopy.⁵⁸ The potential of

this combination has been demonstrated by the parallel acquisition of 2D ^1H - ^1H and ^1H -X ($\text{X} = ^{19}\text{F}$ or ^{31}P) correlation spectra within a single transient. On the other hand, G-matrix Fourier transformation NMR spectroscopy that uses the reduced dimensionality approach⁵⁹ and multiple receivers has been applied to simultaneously acquiring 3D HSQC-TOCSY, 2D ^1H - ^1H TOCSY, and 2D ^{13}C - ^1H HETCOR spectra, offering time savings of an order of magnitude and potential application to the analysis of mixtures and metabolic studies.⁶⁰

The experimental time and the resolution of a multidimensional NMR experiment are proportional to the extent of data sampling in the indirect dimension. Thus, several approaches have been used to significantly shorten measurement times by reducing the number of recorded t_1 increments or proportionally increasing the digital resolution along the indirect F_1 dimension for a given experimental time. For instance, it is now possible to execute NUS^{34–38} in modern spectrometers in a very straightforward manner. In practice, 10–50% of NUS is now feasible in most 2D NMR experiments of small molecules, offering time savings or resolution enhancements of a factor between 2 and 10. Recently, it has been shown that with a careful choice of the sampling schedule and the reconstruction algorithm, 3% of NUS (a practical factor of 32) can be applied in order to increase substantially the resolution of the ^{13}C dimension 2D HSQC spectra of complex mixtures of small molecules, with no increase in experimental time.⁶¹ In this same work, TOCSY multiplet patterns have also been highly resolved along the indirect dimension in a reasonable experimental time. As an example, Figure 6.2 shows the four-fold F_1 resolution enhancement achieved in a routine HSQMBC spectrum of strychnine, whereas excellent results of up to a 10-fold enhancement have also been reported in a related 10% NUS HSQMBC-TOCSY experiment.⁶² However, as the overall sensitivity of the experiment

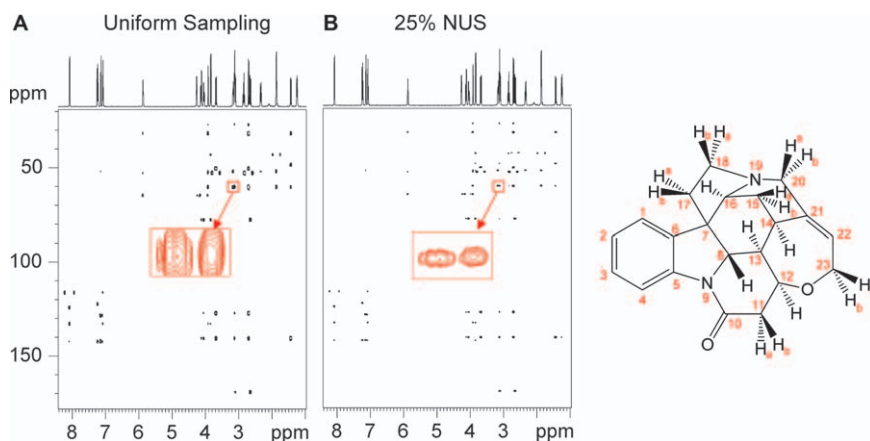


Figure 6.2 Experimental enhancement on F_1 digital resolution achieved after using 25% NUS in HSQMBC spectra of strychnine. Both (A) and (B) spectra have been acquired in the same experimental time.

decreases, the extent to which NUS can be applied simultaneously is also limited because spectral quality can be easily degraded.

Another simple approach for significantly increasing digital resolution along the indirect dimension of multidimensional NMR experiments is the use of spectral folding or spectral aliasing.^{31–33} The major benefits of this approach can be achieved in heteronuclear experiments by simple reduction of the indirect ^{13}C spectral width ($\text{SW}[^{13}\text{C}]$), but at an initial cost of ambiguities in frequencies if no additional information is used. For instance, reducing $\text{SW}(^{13}\text{C})$ from the conventional 160 ppm to only 2 ppm in experiments like HSQC or HMBC drastically increases the digital resolution by about two orders of magnitude. As an advantage, ^{13}C chemical shift information is not completely lost in aliased experiments, and it can be retrieved using reference datasets, such as conventional 1D ^{13}C ,⁶³ a complementary low-resolution 2D experiment that can be acquired faster, or the collection of two aliased datasets using slightly different SWs.⁶⁴ Recently, spectral reconstruction algorithms based on modified HSQC pulse sequences have been reported to afford conventional 2D representations that correspond to those obtained with conventional F_1 SWs. In these aliased HSQC experiments, the missing chemical shift information is encoded in the form of chemical shift multiplets (DENA and AMNA experiments)^{65,66} or phase properties (PHANA experiment).⁶⁷ Additionally, spectral aliasing has been combined with other resolution-enhanced techniques, including NUS or broadband homodecoupling, into the same F_1 dimension⁶⁸ or combined with pure shift techniques along the direct F_2 dimension (see Section 6.3.4).^{69–71}

6.3 Pure Shift NMR

The suppression of the typical J -coupling pattern observed in NMR signals represents a powerful tool for the fast and successful analysis and interpretation of simplified NMR data. For instance, chemists have always been thankful for the simple representation of the regular 1D $^{13}\text{C}\{^1\text{H}\}$ spectrum, where all resonances appear as a fully heterodecoupled singlets without the J_{CH} splittings. Broadband heterodecoupling is easily performed by irradiating ^1H frequencies while ^{13}C nuclei are detected. For many years, the tremendous impact and complementarity of broadband ^1H homodecoupling has been recognized, but a general experimental solution for the acquisition of these data has not been developed. The resulting simplified singlet ^1H signals would afford better spectral resolution and signal dispersion by several orders of magnitude, greatly minimizing the perpetual problem of resonance overlap. Figure 6.3 shows the theoretical enhanced signal resolution that could be achieved by efficient ^1H homodecoupling as a function of the magnetic field strength. Even in a 1.0 GHz spectrometer, two complex and overlapped signals separated by only 0.01 ppm would not be resolved. Theoretically, these signals could only appear resolved in a conventional 1D spectrum recorded in hypothetical 2.0 or 3.0 GHz spectrometers. In contrast, they could be clearly distinguished in a broadband ^1H -homodecoupled

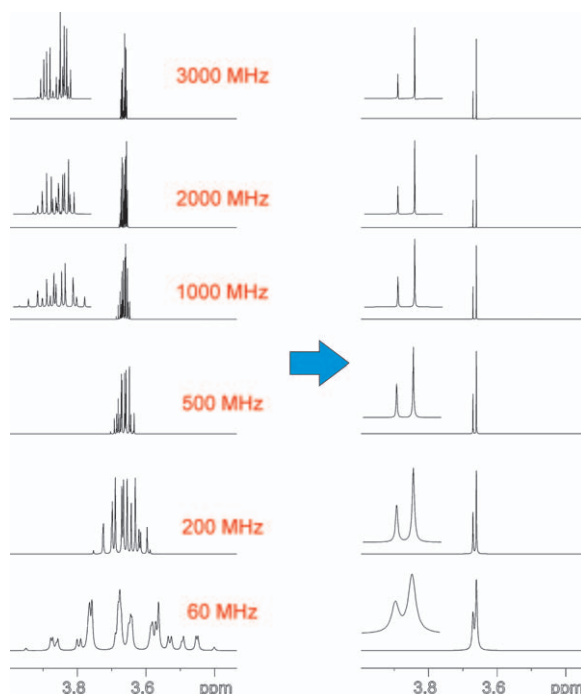


Figure 6.3 Theoretical simulations of an AB spin system (separated by 0.01 ppm and with linewidths of 1 Hz) belonging to a more complex system as a function of the magnetic field strength. Left: conventional 1D ^1H spectra; right: pure chemical shift ^1H spectra.

spectrum recorded at very low magnetic fields, even in a low-cost benchtop NMR system working at 60 MHz. The availability of pure chemical shift spectra could allow an accurate determination of chemical shifts and provide a clear distinction between small chemical shift differences ($\Delta\delta$) and a much easier and simpler analysis, including a precise identification of the number of peaks resonating in overcrowded regions or the easy monitoring of signal intensities in array experiments, among others.

A number of interesting questions about ^1H homodecoupling would quickly arise from non-specialist NMR users. Are the resulting homodecoupled spectra obtained directly from the execution of a specific pulse sequence or are additional and sophisticated post-processing and spectral reconstruction methods required? Which NMR experiments/pulse sequences can achieve this? In terms of the pulse sequences, are they robust experiments that can be routinely applied to any type of molecule and NMR sample? What are the requirements in terms of hardware and software? Is the set-up easy and ready to be used in an automated mode? What about the sensitivity or spectral quality when we are referring to linewidths, the presence of artifacts, or undesired strong J -coupling effects? Can these NMR spectra be quantified? Can the pure shift NMR concept be implemented in

other NMR experiments? Where can we apply these methodologies? Classical and modern NMR approaches to achieving broadband ^1H homodecoupled NMR spectra have been recently reviewed^{72–75} and, in this chapter, efforts have been focused on describing the current NMR methods developed from the original Zangger–Sterk (ZS) experiment.⁷⁶

To understand the fundamentals of the ZS experiment, some brief comments about key spin-echo building blocks are recommended. Probably the easiest way to obtain a broadband homodecoupled ^1H spectrum is the familiar 2D J -resolved experiment (Figure 6.4A) that separates chemical shift and J_{HH} information orthogonally in a 2D “map”. The internal 1D projection along the detected F_2 dimension after a 45° tilting process affords a pure shift ^1H NMR spectrum, but the phase-twist nature of the lineshapes requires absolute value mode calculation, and the presence of unavoidable strong coupling effects usually provides degraded resolution and distorted intensities. Several attempts have been made to achieve pure absorption data, but always with a significant and thus far unavoidable sensitivity penalty, as recently shown for a pure shift yielded by chirp excitation (PSY-CHE)-based 2D J version.^{77,78} Similar results can be also obtained from the 45° projection of the diagonal-peak multiplets of an anti z-COSY spectrum.⁷⁹ A second approach is the band-selective homodecoupling (BASHD) method (Figure 6.4B), where the inversion block consists of a sequential pair of non-selective 180° -selective $180^\circ(\text{A})$ pulses.^{80,81} In the resulting region-selective experiment, only nuclei A experiencing the selective pulse are detected with complete F_1 -homodecoupling from the rest of the nuclei. A simple slice-selective variant of the BASHD method would afford the entire F_1 -homodecoupled NMR spectrum (Figure 6.4C). In this experiment, a spatial encoding element (consisting of the simultaneous application of a selective 180° pulse and a weak encoding gradient) excites each individual

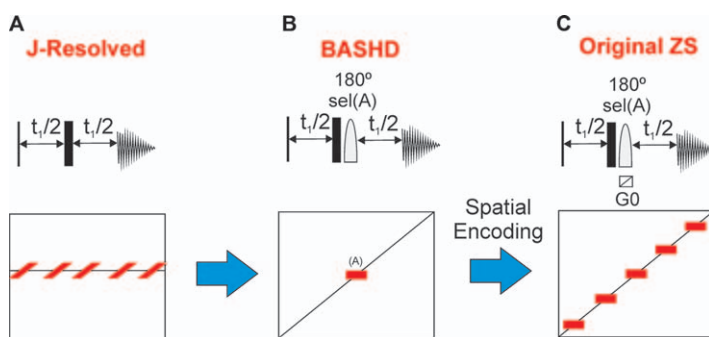


Figure 6.4 Spin-echo NMR building blocks and spectral patterns obtained with different 2D approaches: (A) J -resolved map where only J evolves in the indirect F_1 dimension; (B) frequency-selective version to achieve broadband F_1 homodecoupling in a selected region; and (C) slice-selective version to achieve complete broadband F_1 homodecoupling chemical shift correlation maps by introducing spatial encoding during the selective pulse.

signal in a separate position of the NMR tube. In a conventional probehead configuration, PFGs are applied from the z -axis, thus the F_1 dimension of the resulting 2D spectrum would correspond to the z -position along the NMR tube. The main disadvantage of slice selection is the pronounced loss of sensitivity due to the spatial encoding process, because each signal only contributes from a particular z -slice of the NMR tube. Better SNR can be achieved from a recent PSYCHE-TOCSY experiment (see Section 6.3.4)⁸² or by using the constant-time (CT) approach,⁸³ as reported for the CT-COSY, CT-TOCSY, and CT-NOESY experiments.^{84–86}

6.3.1 The ZS Experiment

The recent interest in broadband homodecoupling is based on the original idea reported by Zangger and Sterk in 1997,⁷⁶ and later revisited by Morris's group.⁸⁷ The method requires a 2D acquisition mode and a data reconstruction protocol to generate a 1D homodecoupled FID (Figure 6.5A). The 2D measurement is repeated by incrementing the variable t_1 period in steps of $1/SW$, where the indirect SW is large compared to the widths of the proton multiplets. Typically, SW is set to 40–100 Hz, and 16–32 t_1 increments give reasonable results in 5–10 minutes with several milligrams of sample (Figure 6.5B). The FID reconstruction is done by concatenating the first ~ 10 ms from the FID of each sequential increment and the final homodecoupled ^1H spectrum is obtained after conventional Fourier transformation. This procedure is possible because the temporal evolution of scalar J -coupling is much slower than that of the chemical shift. However, a certain amount of evolution of the scalar J -coupling does occur for the data block length of 10 ms, introducing some spectral artifacts such as undesired sidebands and some extra line broadening. Two additional factors limit the level of homonuclear decoupling. The first is the bandwidth of the selective 180° pulse used, since decoupling requires that it affects only one coupling partner. The second factor is strong coupling, which leads to “extra” responses, mostly at frequencies between the two coupled chemical shifts. There is a trade-off between sensitivity and the minimum chemical shift difference decoupled, since a narrower bandwidth corresponds to a thinner sample slice.

Different strategies have been proposed to improve the ZS experiment: (i) by improving the SNR per time unit; (ii) by obtaining direct homodecoupled data without need of the 2D reconstruction method; (iii) by using alternative selective inversion elements; and (iv) by improving spectral quality and performance, which means better lineshapes, minimization/suppression of undesired sidebands, and attenuation of strong coupling effects. Figure 6.6 shows several homodecoupling NMR building blocks based on a combined 180° (hard)-selective inversion element. All of these elements selectively invert the selected spins (active spins) while leaving all their J -coupled partners unaffected (passive spins). The distinction between active and passive protons can be accomplished by three different approaches: (i) protons located in different parts of the NMR tube using spatial encoding methods

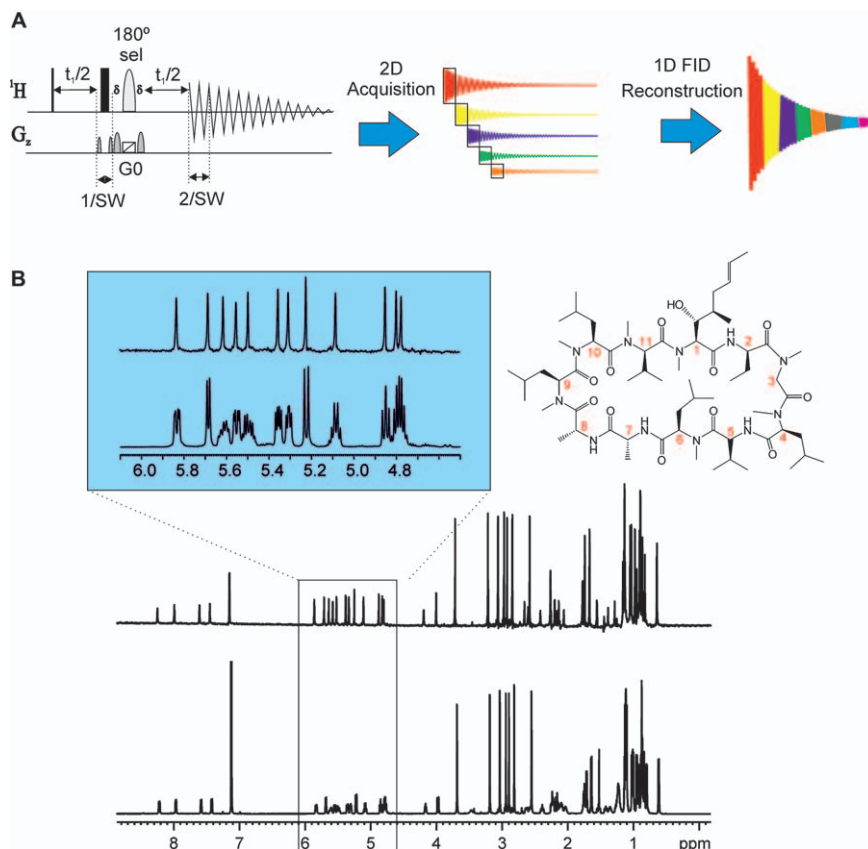


Figure 6.5 (A) Pulse scheme of the ZS experiment and processing protocol that is followed in order to reconstruct a 1D homodecoupled FID. (B) 600 MHz (bottom) conventional and (top) broadband homodecoupled ${}^1\text{H}$ NMR spectra of the peptide cyclosporine. Sixteen data chunks of 10 ms were co-added using a 60 ms Rsnob $180^\circ {}^1\text{H}$ pulse and an encoding gradient of 0.75 G cm^{-1} , giving a total experimental time of 5 minutes.

(Figure 6.6A); (ii) protons located in different parts of the NMR spectrum using frequency-selective pulses (Figure 6.6B); and (iii) protons of different chemical natures, for instance, using isotope-edited BIRD clusters to differentiate ${}^1\text{H}$ - ${}^{13}\text{C}$ and ${}^1\text{H}$ - ${}^{12}\text{C}$ in heteronuclear experiments (Figure 6.6C).

6.3.2 Sensitivity, Strong Coupling Effects, and Spectral Quality

As a major drawback, the original ZS experiment provides very low levels of SNR per time unit due to two main factors: (i) the requirement for a time-consuming 2D acquisition mode to obtain a 1D homodecoupled FID; and (ii) the sensitivity loss associated with the slice-selection process that is also

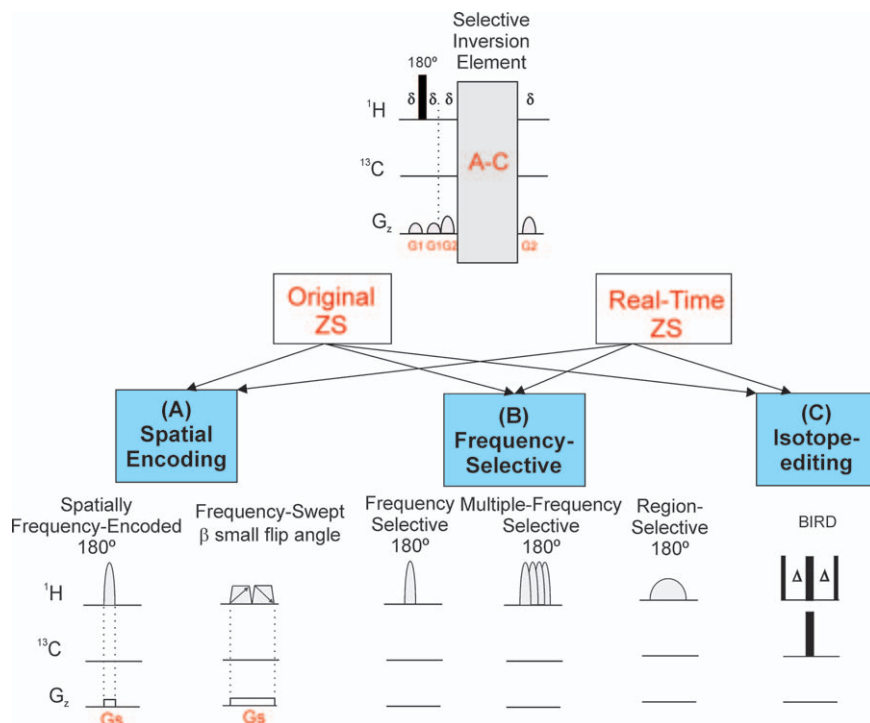


Figure 6.6 Basic NMR building blocks to achieve broadband homonuclear decoupling.

directly correlated to the unwanted effects of strong coupling and to the bandwidth of the chosen selective pulse. The slice-selection process involved in spatially encoded NMR experiments can be easily controlled by fine adjustment of the three basic equations shown in eqn (6.1) through (6.3), and visualized as shown in Figure 6.7. Spatial frequency encoding is performed by simultaneous application of a frequency-selective 180° pulse and a spatial-encoding gradient, G_s , both with the same duration. Experimentally, the range of sampled frequencies (SW_G) is defined by the strength of gradient G_s according to:

$$SW_G = \gamma \times L \times G_s \quad (6.1)$$

where γ is the gyromagnetic ratio of the spatially encoded nucleus and L is the active volume coil length (Figure 6.7A). Furthermore, the carrier frequency (Ω) and the selective pulse bandwidth ($\Delta\omega$), which is inversely proportional to the pulse length, determine the z-position of each nuclear spin (Figure 6.7B):

$$z = \Omega / (\gamma \times G_s) \quad (6.2)$$

and the slice thickness (Δz) (Figure 6.7C):

$$\Delta z = \Delta\omega / (\gamma \times G_s) \quad (6.3)$$

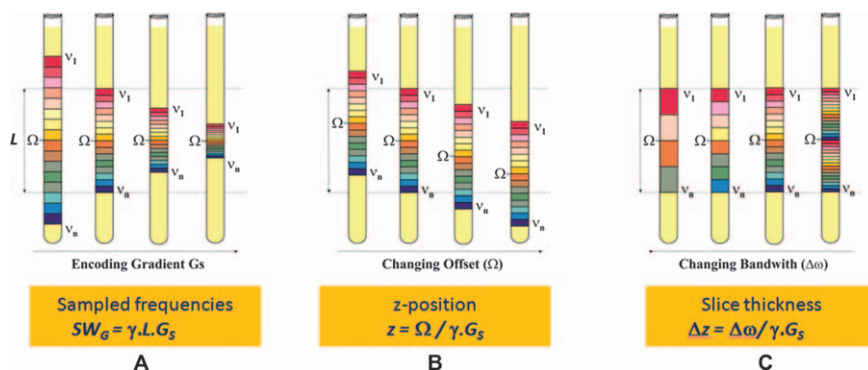


Figure 6.7 Schematic representation of the slice selection procedure as a function of the experimental NMR parameters. The NMR spectrum is displayed as discrete colored slices along the z-axis. (A) The range of sampled frequencies (SW_G) is determined as a function of the gradient strength (G_s) and the active volume coil length (L). (B) The z-position of each individual resonance is defined from the carrier frequency (Ω). (C) The slice thickness (Δz) and therefore the number of slices is determined by the selective pulse bandwidth ($\Delta\omega$).

The overall SNR of a slice-selective experiment will depend on both the strength of the encoding gradient and on the selectivity of the 180° pulse. For instance, a typical 20 ms Gaussian-shaped 180° ^1H pulse (effective bandwidth of 60.7 Hz) applied simultaneously with a gradient of 0.743 G cm^{-1} (this should be equivalent to 1.39% of the maximum gradient strength of a gradient unit delivering 53.5 G cm^{-1}) splits the sample height ($L = 18 \text{ mm}$) into approximately 94 slices along the z-axis, defining a slice thickness of about 0.019 cm and covering a SW_G of 5694 Hz (9.47 ppm in a 600 MHz spectrometer). Thus, under these general conditions, the single-slice selection procedure would only afford about 1% of the sensitivity of the conventional ^1H spectrum. This low SNR could be improved using shorter and correspondingly less selective pulses (Gaussian-, Rsnob-, and Reburp-shaped pulses are usually applied) and/or less intense encoding gradients, but always at the expense of increasing the probability of accidental excitation of two mutually J -coupled spins into the same slice. Obviously, any spatial frequency-encoding NMR technique, including the ZS experiment, is best performed in higher fields as demonstrated using an oligomeric saccharide sample.⁸⁸

Sequential multi-slice excitation⁸⁹ can be exploited for cases in which nuclear spins in different parts of the NMR tube are exclusively excited during subsequent transients. This is experimentally performed by changing the offset frequency while the previously used spins have time to relax towards equilibrium before being excited again. This approach allows a faster repetition rate (the recycle delay can be reduced down to 100 ms) and significantly shorter overall acquisition times. It has been shown that SNR gains of a factor of two- to three-fold can be achieved using eight different

excitations. On the other hand, a moderate sensitivity gain of an average factor of two-fold can also be achieved using the ASAP technique as a fast pulsing method.⁹⁰ Another alternative affording a more than one order of magnitude improvement in sensitivity is based on simultaneous multi-slice excitation that is performed by applying multiple-frequency modulated pulses^{91,92} or with related multiple spatial encoding using adiabatic chirped pulses.^{82,93} The initial proposal was based on the careful and optimized setting of multiple offsets in order to avoid the excitation of mutually J -coupled protons within the same slice, which would result in distorted multiplets due to J_{HH} evolution.⁹¹ An alternative non-equidistant modulation ZS scheme recommends the use of 50 non-equidistant modulations of 120 ms Rsnob pulses without previous knowledge of frequencies.⁹² An improved method, referred to as PSYCHE,⁹³ uses a pair of low flip ($\beta = 15\text{--}20^\circ$) angle swept-frequency pulses simultaneously applied to a very weak encoding gradient as a selective inversion element (Figure 6.8A). The distinction between passive and active spins in PSYCHE is purely statistical, the former being a proportion $\sin 2\beta$ of the whole and the latter $\cos 2\beta$, rather than being determined by spatial position within the sample (ZS method) or by coupling to a dilute heteronucleus (BIRD). Figure 6.8B through 6.8D compare the performance of the original ZS and the PSYCHE experiments in the aliphatic overcrowded region of the steroid estradiol. The sensitivity of the PSYCHE experiment using 15 ms chirped pulses is similar to the ZS experiment with a 10 ms Rsnob pulse (Figure 6.8B vs. 6.8D). However, imperfect homodecoupling is clearly observed in the latter. In order to achieve a thinner slice thickness and perfect homodecoupling, a more selective 40 ms Rsnob pulse is required in the ZS experiment, but with a sensitivity loss of a factor of four (Figure 6.8C vs. 6.8D). In practice, spatial encoding only affords 0.1–2% of the maximum ^1H sensitivity (depending of the bandwidth of the selective pulse), whereas PSYCHE can reach to 1–10%. Interestingly, the strongly coupled AB spin system resonating at around 3.2 ppm (separated by 30 Hz in a 600 MHz spectrometer) in estradiol is not perfectly homodecoupled in any case. An improved PSYCHE pulse scheme using compensated adiabatic pulses instead of the hard 180° pulse affords better results in terms of spectral quality, but still retains the undesired strong coupling effects and with a minimal penalty in sensitivity.⁷⁸ The PSYCHE approach is very robust in terms of set-up and automation and it can be used to improve significantly the low sensitivity of existing multi-dimensional ZS-based NMR experiments.

A real-time version of the ZS experiment has been proposed for the simpler and faster acquisition of homodecoupled 1D spectra.⁹⁴ Data collection requires neither 2D data acquisition nor additional post-acquisition reconstruction methods. In this pulse scheme, the acquisition period consists of n concatenated loops that include a pair of hard/selective 180° ^1H pulses (each one flanked by the G3 and G4 gradients) applied at intervals of 2τ periods (τ is set to $AQ/2n$) (Figure 6.9A). As in the original ZS experiment, the method also involves a significant cost in sensitivity due to slice selection

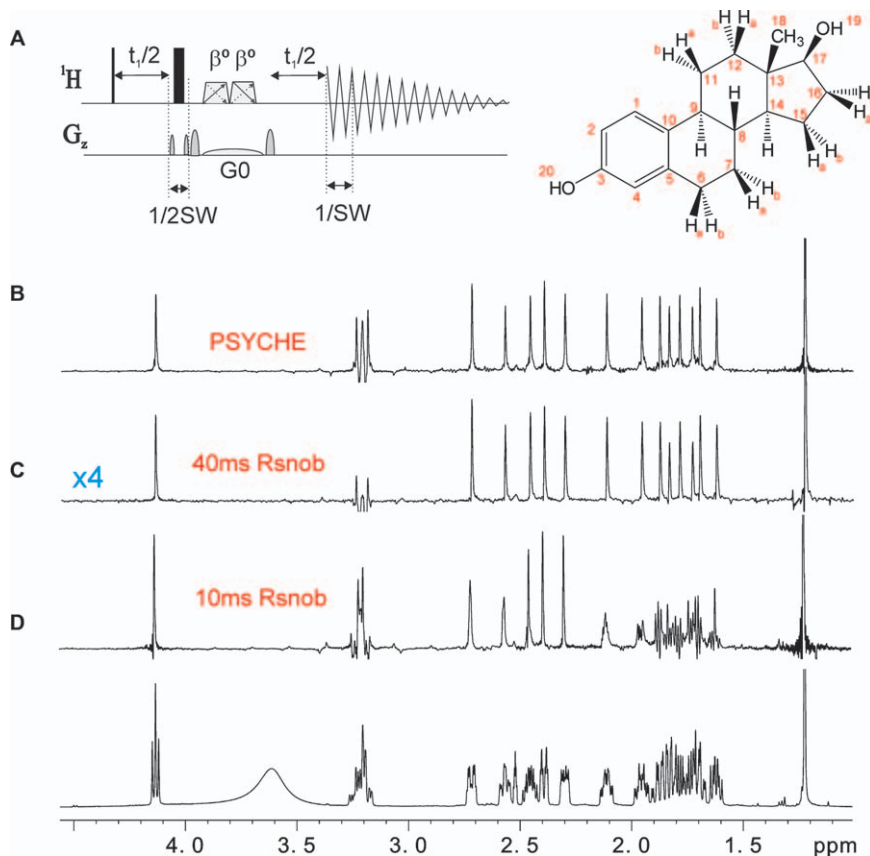


Figure 6.8 (A) Pulse scheme of the PSYCHE experiment. (B–D) Comparison between the PSYCHE (using two 15 ms chirped $[\beta]$ 15° pulses; (B) and ZS spectra using 10 ms Rsnob (C) and 40 ms Rsnob (D), respectively, of estradiol. All spectra have been acquired with a 0.75 G cm^{-1} encoding gradient.

(compare Figure 6.9B and 6.9C) and, in favorable cases, the SNR can be improved by a factor of two to three by rapidly repeating the experiment sequentially on adjacent, non-perturbed regions of the sample⁹⁵ or by simultaneous multiple slice selection.⁹¹ However, there are two major inconveniences associated with the real-time ZS approach. First, the length of the selective pulse applied during the interrupted FID is significantly limited and must be kept below 10–15 ms in order to avoid severe loss of signal due to transverse relaxation. This limitation offers very poor slice selectivity, a problem that is soon observed when nearby coupled protons need to be homodecoupled. Second, FID interruption generates poor spectral quality due to the presence of wider lineshapes for shorter FID chunks and important sidebands at multiples of n/τ from the main signal, with intensities decreasing rapidly with increasing n and for long FID chunks. The variation of the FID chunk lengths between individual transients significantly

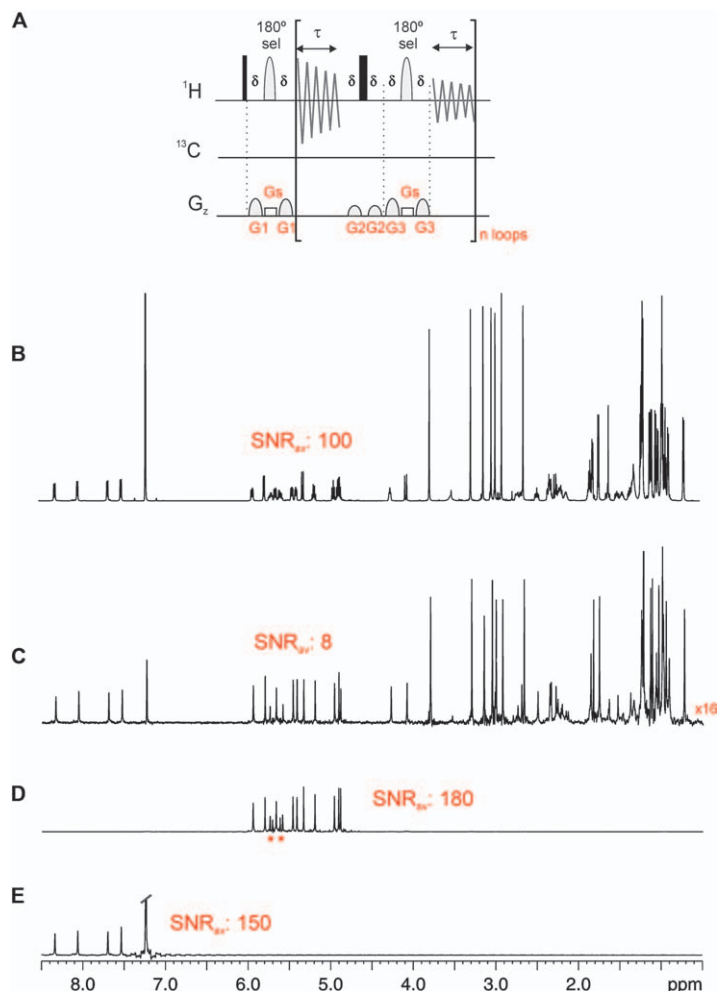


Figure 6.9 (A) Pulse sequence of the real-time ZS experiment. Broadband homo-decoupling during detection is achieved by applying a pair of hard/selective 180° ^1H pulses (represented as solid and shaded shapes) at the middle of $2\tau = \text{AQ}/n$ periods, where AQ is the acquisition time and n is the number of concatenated loops. (B) Conventional ^1H spectrum of cyclosporine normalized with a $\text{SNR} = 100$. (C) Real-time ZS spectrum acquired with a 5 ms REBURP-shaped pulse and a weak 0.7 G cm^{-1} encoding gradient. (D & E) 1D HOBS spectra of the H_α and NH regions acquired as in (C), but without the encoding gradient G_s . (Reproduced from ref. 96 with permission from John Wiley and Sons. Copyright © 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

suppresses decoupling sidebands and thus can allow, for instance, the analysis and clear distinction of signals belonging to minor components in compound mixtures.⁹⁵ It has been reported that $\sim 25\%$ of the total ^1H sensitivity can be recovered for a $8 \mu\text{g}$ azithromycin aglycone sample dissolved

in a 3 mm chloroform-matched Shigemi NMR tube on a 700 MHz spectrometer equipped with a 5 mm cryoprobe (experimental time: about 2 minutes). In practice, the original ZS and PSYCHE experiments offer better general applicability and more optimal linewidths than real-time ZS.

A very simple modification of the real-time ZS pulse scheme has been proposed for the rapid acquisition of full-sensitivity, homodecoupled band-selective (HOBS) NMR spectra.^{96–98} The major novelty in this approach is that the spatial encoding gradients applied simultaneously with the selective pulses are omitted (see Figure 6.9A), avoiding sensitivity losses due to slice selection. The HOBS method retains all of the advantages/disadvantages of the real-time ZS experiment, and excellent results are achieved using short 180° REBURP semi-selective pulses (~5–10 ms). Figure 6.9D and 6.9E show the individual single-scan 1D HOBS spectra that are rapidly obtained after selection of the H_α and NH regions, respectively, of the peptide cyclosporine. It is worth emphasizing that HOBS spectra present better SNRs than the conventional ¹H spectrum due to signal collapsing.

6.3.3 ZS Applications

The different ZS versions have been applied in a wide range of topics, including: (i) the analysis of complex mixtures and the accurate determination of small $\Delta\delta$, with a particular interest in the differentiation of enantiomers under the presence of a chiral solvating agent (CSA);^{70,99,100} (ii) the distinction and unambiguous assignment of diastereoisomers⁹⁸ or the characterization of similar molecules having nearly identical NMR spectra;⁷¹ (iii) the measurement of homo- and hetero-nuclear coupling constants from simplified multiplet structures;^{101–104} (iv) the direct analysis of mono-exponential intensity decays in overlapped signals, including the measurement of T₁ and T₂ relaxation times^{105,106} and diffusion data;^{87,107–109} (v) the implementation in practically any type of homonuclear and heteronuclear 2D experiments; and (vi) the further combination with other resolution-enhanced techniques (spectral aliasing, NUS, and other homodecoupling schemes) and spectral reconstruction methods.^{69–71}

For instance, both ZS and HOBS techniques have been used to rapidly detect small $\Delta\delta$ in enantiodifferentiation studies.^{70,99,100} Figure 6.10 shows the sensitive discrimination of several ¹H resonances belonging to a racemic mixture of a small molecule in the presence of Pirkle alcohol as a CSA. A straightforward comparison between the conventional (Figure 6.10A) and the fully homodecoupled multiplets (Figure 6.10B) makes it evident that a more facile and reliable determination of $\Delta\delta$ and R/S molar ratios is feasible considering the highly dispersed singlet signals, independent of the original multiplet complexity. This clear identification can even be improved by using homodecoupled HSQC spectra (see Section 6.3.4).

The original ZS method has been applied to the direct measurement of heteronuclear coupling constants of high-abundant nuclei (*J*_{FH} and *J*_{PH}) from simplified ¹H-homodecoupled 1D ¹H multiplets.^{110,111} In a similar way,

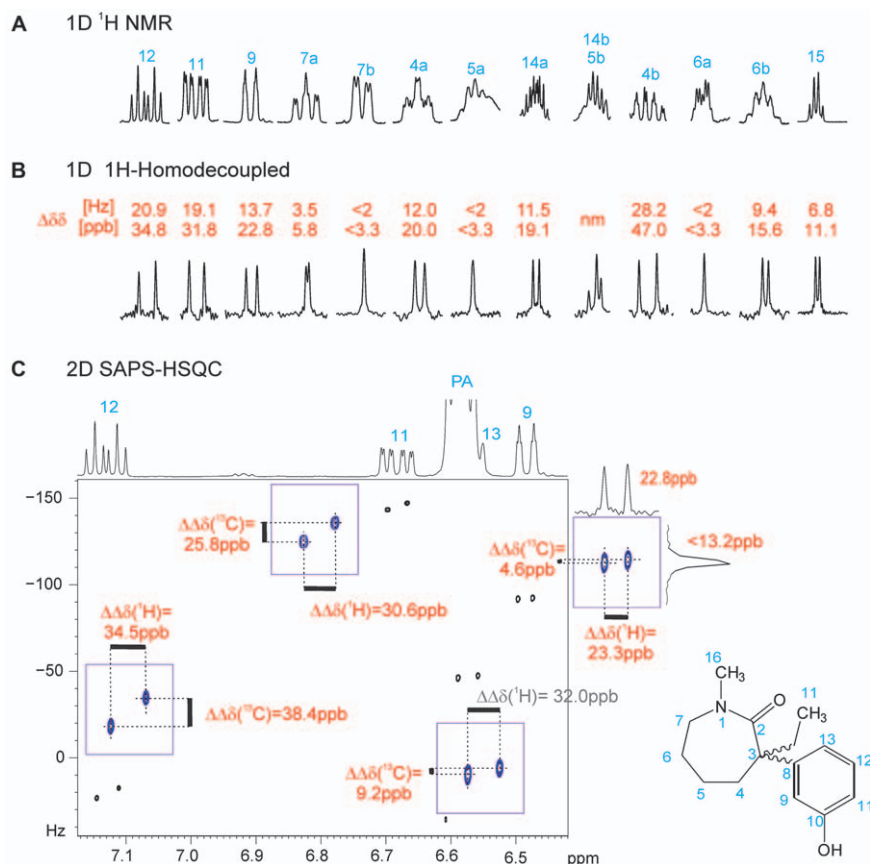


Figure 6.10 Pure shift NMR experiments in enantiodifferentiation studies. (A) Conventional ^1H multiplets; (B) homodecoupled ^1H multiplets obtained from a ZS experiment; (C) expansion of the SAPS-HSQC spectrum of a racemic mixture of a small organic molecule in the presence of Pirkle alcohol as a CSA agent.

(Reproduced from ref. 70 with permission from the Royal Society of Chemistry.)

modified ZS schemes have been proposed for the selective measurement of specific J_{HH} coupling constants by reintroducing some of them in homodecoupled 1D spectra,^{101,102,112} where scalar coupling to only one selected signal is visible, whereas all other couplings are removed from the spectrum. An interesting real-time J -upscaled 1D NMR method has been proposed in order to visualize unresolved scalar couplings.¹¹³ This technique, which works by introducing additional scalar coupling evolution delays within the recording of the FID, not only stretches the recorded coupling patterns, but also actually enhances the resolution of multiplets by reducing signal broadening due to magnetic field inhomogeneities during the interrupted data acquisition. Enlarging scalar couplings also enables their

determination in situations where the spectral resolution is limited, such as in the acquisition dimension of heteronuclear broadband decoupled HSQC spectra. In an opposite sense, a related method has been proposed to achieve real-time homonuclear J -downscaling, where all J -values are uniformly reduced by an arbitrary scaling factor. In the resulting 1D spectra, signal overlap is reduced, while scalar coupling information remains available.¹¹⁴ Homodecoupled versions of the G-SERF experiment have been reported for the accurate measurement of selective J_{HH} coupling constants from a 2D J -resolved format.^{103,104}

6.3.4 Implementing Homodecoupling in 1D and 2D NMR Experiments

An important feature of the described ZS pulse schemes is their easy implementation as a general building block in regular 1D and 2D homonuclear and heteronuclear NMR experiments, retaining the same selectivity and sensitivity features discussed previously for each ZS version. Novel homodecoupled experiments can be classified as a function of the homodecoupling element (BIRD, spatial encoding, or HOBS) and of the acquisition mode (2D *vs.* real-time) (Table 6.1). Broadband homodecoupling along the acquisition dimension of a 1D or a 2D experiment can be implemented by replacing the conventional data acquisition of the original ZS or real-time ZS building blocks. Homodecoupled versions of the DOSY, TOCSY, NOESY, and ROESY experiments based on the original ZS experiment suffer both low spatially encoded sensitivity and 3D acquisition modes, and therefore long measurement times are required. Real-time ZS versions use conventional 2D data acquisition and processing, but low sensitivity is again the main drawback. Alternatively, HOBS versions offer full sensitivity but only have a guaranteed successful application in specific spin systems and have been limited to selective regions containing non- J -coupled protons. As an example, Figure 6.11A shows how real-time HOBS is implemented in a 2D homonuclear pulse sequence. The comparative analysis between the regular TOCSY and the HOBS-TOCSY spectra, both acquired under the same experimental conditions and time, reveals a much better signal resolution and sensitivity in the direct dimension of the HOBS version without affecting spectral quality and performance (Figure 6.11B).

Different approaches have been reported to enhance the concept of ultra-high-resolution by obtaining ^1H homodecoupled signals along both F_1 and F_2 dimensions in 2D NMR spectra. CT versions of the homonuclear COSY, TOCSY, and NOESY experiments have been combined with indirect covariance processing to produce full homodecoupled 2D correlation maps. These versions have been demonstrated to be able to detect and characterize impurities in synthetic and natural products samples.^{85,86} The classical BASHD method achieves equivalent broadband homodecoupling along the F_1 dimension in a reduced area of the spectrum using band selection,^{80,81,115} and a related

Table 6.1 Summary of reported broadband ^1H -homodecoupled 1D and 2D NMR experiments.

NMR experiment		Homodecoupling		Acquisition mode		Original references	Related references
		BIRD	Spatial encoding	Frequency selection	Pseudo-2D/3D	Real-time	
1D	^1H NMR	✓			✓		111,117
			✓		✓		70,76,87,90,91,93,110,118
		✓				✓	119,120
			✓			✓	121
				✓		✓	96–98,100
				✓		✓	100
			✓			✓	101,102
2D	SelTOCSY Quick-SERF Inversion recovery CPMG-PROJECT			✓		✓	105,106
				✓		✓	105
	COSY <i>J</i> -resolved G-SERF			✓		✓	122
				✓		✓	92,115
			✓		✓		103,104
	TOCSY		✓		✓		82,123,124
			✓		✓		121
				✓		✓	96
	DOSY		✓		✓		87,107,108
			✓			✓	109
	NOESY		✓		✓		85
	ROESY			✓		✓	97,106
			✓		✓		125
				✓		✓	125
	HSQC/HSQCed		✓		✓		126
		✓			✓		127–132
		✓				✓	70,133
	HSQC-TOCSY HSQMBC			✓		✓	96,97
				✓		✓	71
				✓		✓	51,71,134,135
	ADEQUATE	✓	✓		✓		132,136 137,138

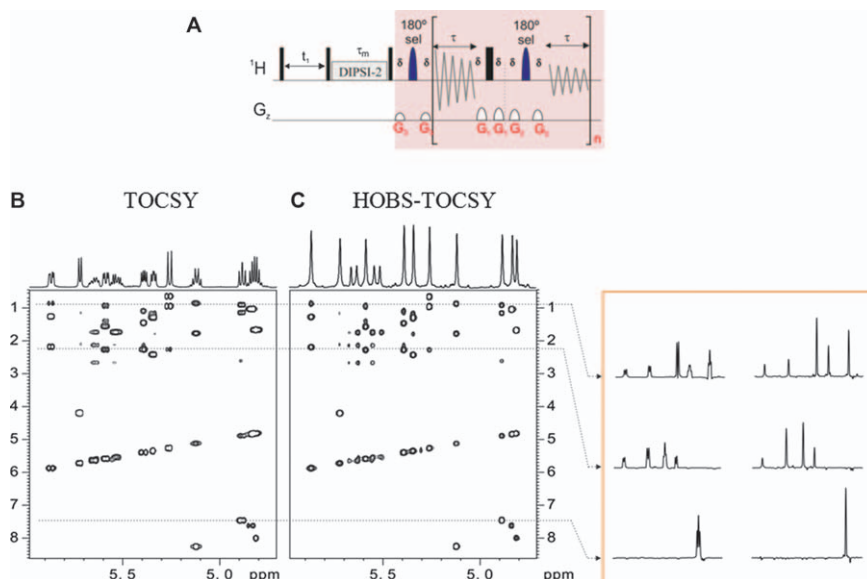


Figure 6.11 (A) Pulse scheme of the HOBS-TOCSY experiment. (B & C) Comparison between the (B) TOCSY and (C) H_x region-selective regular HOBS-TOCSY spectra of cyclosporine (mixing time of 60 ms). (C) 1D slices taken at two different frequencies to compare the relative sensitivities and resolution levels. Four scans were collected for each 128 t_1 increments of 2K complex points, giving an experimental time of 13 minutes for each 2D spectrum. Homodecoupling was achieved using 50 loops and $\tau = 4.3$ ms (AQ = 170 ms). (Reproduced from ref. 96 with permission from John Wiley and Sons. Copyright © 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

PSYCHE-TOCSY pulse scheme (Figure 6.12A) supplies information on the entire spectrum.⁸² This latter method only has about 10% of the overall sensitivity and requires high digitization in the F_1 dimension because indirect covariance is used to generate pure shift data in both dimensions (Figure 6.12C vs. 6.12D). NUS can be used to reduce the overall acquisition time. Alternatively, a band-selective F_1 -homodecoupled PSYCHE-TOCSY experiment has demonstrated its usefulness as an alternative to HOBS for obtaining BASHD in highly congested areas with strong coupling effects, as is traditionally found in the 3.5–4.5 ppm area in carbohydrates, by minimizing the need for a large number of t_1 increments that increases the total experimental time.¹¹⁶

Figure 6.13 shows the straightforward implementation of the HOBS technique in the standard inversion recovery (IR) experiment for measuring T_1 relaxation times in overlapped ${}^1\text{H}$ resonances using conventional mono-exponential curve-fitting methods, without need for additional data analysis based on deconvolution or line-fitting techniques. The proposed HOBS-IR experiments¹⁰⁵ can use the same automated data acquisition, processing,

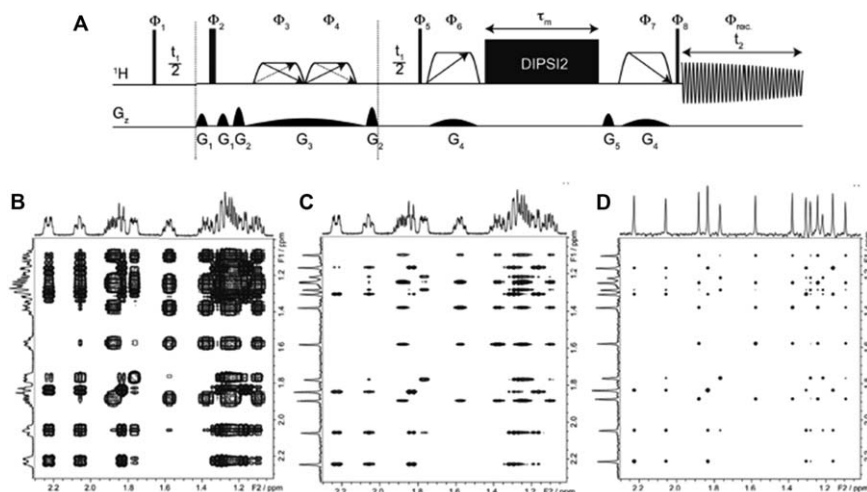


Figure 6.12 (A) Pulse scheme of the PSYCHE-TOCSY experiment. (B) Regular 2D TOCSY spectrum of estradiol. (C) F_1 -homodecoupled TOCSY spectrum acquired with the PSYCHE-TOCSY scheme of (A). (D) Pure shift TOCSY spectrum after applying indirect covariance. (Reproduced with permission from ref. 82.)

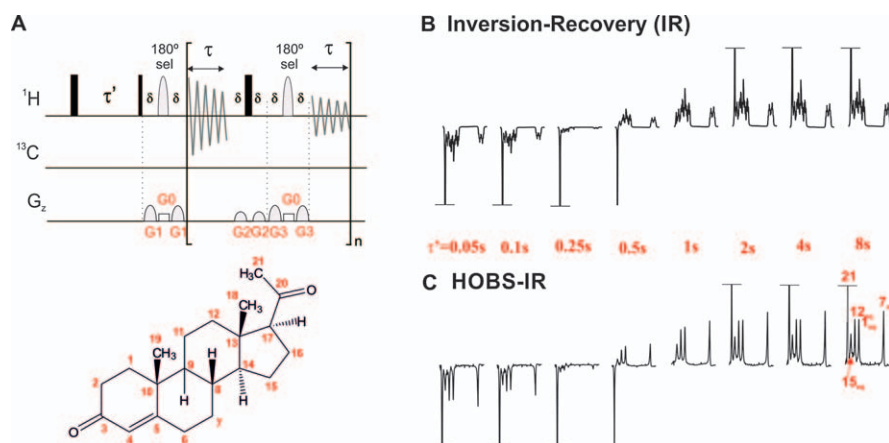


Figure 6.13 (A) Pulse scheme of the HOBS-IR experiment used to measure T_1 relaxation times in overlapped proton signals. (B & C) Intensity decays obtained from (B) the conventional IR and (C) the HOBS-IR experiments. All spectra were collected under the same experimental conditions and plotted on the same absolute vertical scale. Homodecoupling was achieved using a 5 ms Reburp 180° pulse, $\tau = 8.9$ ms, $AQ = 569$ ms and $n = 32$.

(Reproduced from L. Castañar *et al.*, *J. Measurement of T_1/T_2 relaxation times in overlapped regions from homodecoupled ^1H singlet signals*, *Magn. Reson.*, **244**, 30–35, Copyright (2014) with permission from Elsevier.¹⁰⁵)

and fitting analysis subroutines as the original IR experiment. The same strategy has also been exploited for T_2 measurements in overlapped signals using the CPMG-based PROJECT and HOBS-PROJECT experiments (see Section 6.4.1) and has potential utility in reaction monitoring or reaction kinetics.

Broadband homodecoupling using real-time BIRD-based or HOBS approaches has been successfully introduced in some 2D heteronuclear NMR experiments, without introducing additional penalties in sensitivity.^{70,131,133} These approaches have been incorporated in different HSQC schemes, namely HSQC, sensitivity-enhanced HSQC-PEP and multiplicity-edited HSQC versions. In real-time BIRD-based homodecoupled experiments, $180^\circ(^1\text{H})$ -BIRD inversion modules are used for homodecoupling and heteronuclear decoupling during the τ acquisition periods are also applied in order to obtain fully decoupled ^1H signals. As an exception, diastereotopic protons belonging to CH_2 spin systems appear as doublets because $^2J_{\text{HH}}$ is also inverted during the BIRD isotope filter cluster and therefore is not decoupled (Figure 6.14A). In general, such HSQC methods provide an increase in both resolution and sensitivity, so long as T_2 is not limiting. Linewidths are similar to those in the conventional spectrum, although signal losses from imperfect pulses, mismatch between τ and $^1J_{\text{CH}}$, and transverse relaxation should, in principle, lead to broader lines in the pure shift spectrum. Any FID discontinuities in the decoupled signal lead to small F_2 sidebands that can be minimized using variable data chunks.⁹⁵ A comparative study between different versions of the pure shift HSQC experiment has been made, evaluating signal intensity and efficient water suppression in ^1H - ^{15}N HSQC on labeled proteins.¹³¹ The original ZS strategy has also been implemented in F_2 -heterodecoupled (RESET-HSQC)¹²⁹ and F_2 -heterocoupled versions (pure-shift CLIP-HSQC) in order to measure $^1J_{\text{CH}}$ in isotropic and anisotropic media,^{127,130} but a time-consuming 3D acquisition mode is required. The major advantages are the simplified doublet signals and the better monitoring of strong coupling effects, although $^2J_{\text{HH}}$ splitting cannot be removed. A modified pure shift HSQC version incorporating a perfect BIRD inversion element allows the complete decoupling of geminal $^2J_{\text{HH}}$ interactions.¹²⁸

The traditional heteronuclear proton-detected experiments (HSQC, HSQC-TOCSY, and HSQMBC) can profit from additional resolution-enhanced techniques to offer ultra-enhanced resolution spectra: (i) spectral aliasing can be incorporated to improve resolution along the F_1 dimension by one or two orders of magnitude without increasing the total experimental time by using a reduced ^{13}C SW; (ii) resolution can be further improved using NUS along the indirect dimension; (iii) broadband homodecoupling along the detected dimension can be implemented using some of the pure shift methods described previously; and (iv) additional data processing tools can be used, such as zero-filling or linear prediction (Figure 6.15).

Spectral aliased pure shift (SAPS)-HSQC spectra have been shown to be very effective for the detection and accurate, fast quantification of small

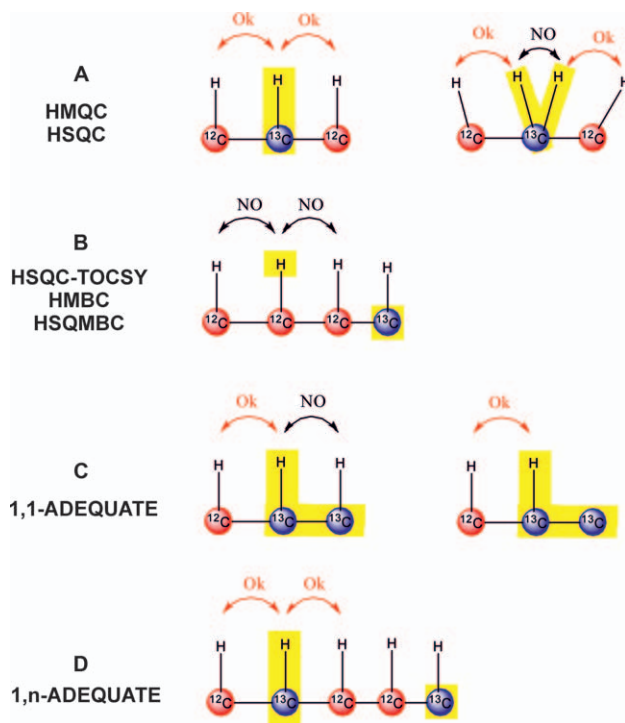


Figure 6.14 Different levels of BIRD-based broadband homodecoupling achieved in inverse proton-detected NMR experiments as a function of the involved spin system. (A) In HMQC/HSQC experiments, the detected $^1\text{H}\{^{13}\text{C}\}$ signals are efficiently homodecoupled from the rest of $^1\text{H}\{^{12}\text{C}\}$. As an exception, geminal $^2J_{\text{HH}}$ splittings in diastereotopic methylene systems are not eliminated from the multiplet and their signals appear as doublets. (B) Homodecoupling does not work in HSQC-TOCSY and HMBC/HSQMBC experiments because the detected $^1\text{H}\{^{12}\text{C}\}$ signals are mostly coupled to indistinguishable $^1\text{H}\{^{13}\text{C}\}$ signals. (C) Partial homodecoupling is achieved in 1,1-ADEQUATE experiments because the vicinal J_{HH} between the detected $^1\text{H}\{^{13}\text{C}\}/^1\text{H}\{^{13}\text{C}\}$ spin systems cannot be eliminated. (D) In 1,n-ADEQUATE, the selected $^1\text{H}\{^{13}\text{C}\}/^1\text{H}\{^{13}\text{C}\}$ pairs are sufficiently distant and therefore surrounded by $^1\text{H}\{^{12}\text{C}\}$, which can be efficiently homodecoupled.

$\Delta\delta(^1\text{H})$ and $\Delta\delta(^{13}\text{C})$, with interesting applications to enantiodifferentiation studies,⁷⁰ to the identification and assignment of similar molecules containing near-identical ^1H and ^{13}C NMR spectra (Figure 6.15),⁷¹ and to the measurement of the sign and magnitude of small coupling constants from highly resolved E.COSY cross-peaks.⁶⁹ For instance, SAPS-HSQC has been used to simultaneously discriminate ^1H and ^{13}C signals separated by less than 1 Hz in a 600 MHz spectrometer (~ 1 ppb for ^1H and ~ 3 ppb for ^{13}C) in short measurement times, even for NMR signals with no apparent splitting in the 1D ^1H ZS or the conventional $^{13}\text{C}\{^1\text{H}\}$ spectra (Figure 6.10).

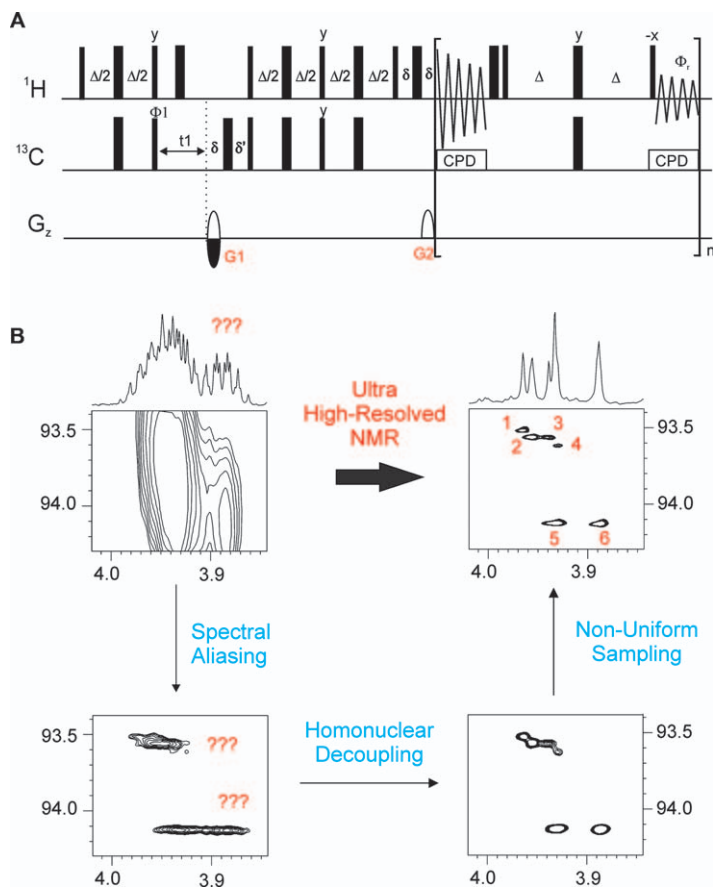
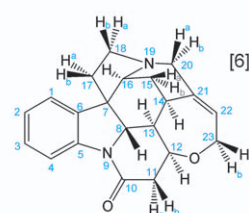


Figure 6.15 (A) Pulse scheme of the real-time BIRD-based homodecoupled HSQC-PEP experiment. (B) Strategy to achieve ultra-high-resolved HSQC spectra. Improvements of up to two orders of magnitude in digital resolution along the F_1 dimension can be achieved by spectral aliasing and NUS, whereas signal resolution in F_2 is strongly enhanced with broadband homodecoupling. (Reproduced from ref. 71 with permission from John Wiley and Sons. © 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

Broadband BIRD-based homodecoupling is not feasible in HSQC-TOCSY and HMBC/HSQMBC experiments because there is no way to distinguish between neighboring ^1H - ^{13}C signals (Figure 6.14B). However, the full-sensitivity HOBS-selHSQMBC experiment has been proposed to afford pure shift long-range heteronuclear correlation sub-spectra with considerable enhancement in terms of both resolution and sensitivity.^{51,71,134,135} In an improved version, even C/CH_2 and odd CH/CH_3 carbon multiplicity information can be directly distinguished from the relative positive/negative phase of cross-peaks in a modified multiplicity-edited HOBS-selHSQMBC experiment (Figure 6.16B).¹³⁵ This is a good example demonstrating how



(Reproduced from ref. 135 with permission from John Wiley and Sons.
Copyright © 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

homodecoupling offers a completely new approach to the simultaneous and precise determination of $^nJ_{\text{CH}}$ from the pure in-phase doublet nature of cross-peaks originating along the F_2 dimension (see pulse scheme in Figure 6.16A and spectrum in Figure 6.16C). Optionally, broadband heteronuclear ^{13}C decoupling can be applied during the FID acquisition periods (τ) in order to obtain a real pure shift selHSQMBC spectrum along both dimensions (Figure 6.16D). The method can be further extended by a TOCSY propagation step, and their utility has been demonstrated for the structure elucidation of the alkaloid staurosporine and the complex, dimeric *Strychnos* alkaloid sungucine,¹³⁹ as well as for the simultaneous determination of the magnitude and sign of up to three different heteronuclear coupling constants from each individual cross-peak detected in molecules containing highly abundant passive nuclei ($Z = ^{19}\text{F}$ or ^{31}P).⁵¹ Whereas J_{HZ} and J_{XZ} coupling constants are measured from E.COSY multiplet patterns, J_{XH} is independently extracted from the complementary IPAP pattern generated

along the detected F_2 dimension. The broadband homodecoupled CPMG-HSQMBC method has also been proposed for the precise and direct measurement of long-range heteronuclear coupling constants.^{132,136} In contrast to selHSQMBC, these methods provide complete 2D spectra with simplified and pure AP doublet cross-peaks. Unfortunately, major inconveniences are the need for a 3D acquisition mode and the significant penalty in sensitivity due to the use of spatial encoding as a homodecoupling method, requiring long acquisition times and highly concentrated samples. Applications have been demonstrated for the extraction of $^nJ_{\text{SeH}}$ and $^nJ_{\text{CH}}$ values in carbohydrates. The concerted use of HOBS-HSQC, HOBS-selHSQMBC, and HOBS-HSQC-TOCSY is able to differentiate and assign complex mixtures containing diastereoisomers with near identical ^1H and ^{13}C NMR spectra.⁷¹

Partial BIRD-based homodecoupling has also been implemented in 1,1- and 1,*n*-ADEQUATE experiments (Figure 6.17A).¹³⁷ In the 1,1-ADEQUATE experiment, correlations between protonated and adjacent quaternary carbons (^{13}CH - ^{13}C) will be decoupled, whereas correlations between adjacent protonated carbons (^{13}CH - ^{13}CH) will not be decoupled from each other, analogous to the anisochronous protons of a geminal methylene in a pure shift HSQC spectrum (Figure 6.14C). In the 1,*n*-ADEQUATE experiment, long-range correlations corresponding to $^2J_{\text{CC}}$ (^{13}CH - ^{12}C - ^{13}CH spin systems) and $^3J_{\text{CC}}$ (^{13}CH - ^{12}C - ^{12}C - ^{13}C) will be decoupled, except for the small homonuclear proton-proton coupling between the remote proton(s) on ^{13}C , which generally will not be observed given the typical resolution in this type of experiment (Figure 6.14D). The homodecoupled 1,1- and 1,*n*-ADEQUATE (HD-ADEQUATE) methods have been used to resolve long-standing structural questions associated with the complex alkaloid cytospirolepine (Figure 6.17B).¹³⁷ Characterization of several degradants of the original, sealed NMR sample a decade after the initial report called the validity of the originally proposed structure into question. Optimal results were achieved with 700 μg of sample dissolved in 35 μl of deuterated solvent using a 1.7 mm MicroCryoProbeTM. Sensitivity increases in the range typically realized by applying broadband homonuclear decoupling (40–60%) could reduce data acquisition times to up to half or less, enhancing the utility and viability of these techniques in a wide variety of challenging and heretofore inaccessible structure elucidation problems. The partially homodecoupled 1,1-HD-ADEQUATE experiment has also been used to characterize the carbon skeleton of spiroketals that cannot be achieved using HMBC data.¹⁴⁰ Related homodecoupled *J*-modulated HD-ADEQUATE experiments have been reported for the measurement of $^1J_{\text{CC}}$ and $^nJ_{\text{CC}}$ coupling constants.¹³⁸ By combining homodecoupling and 50% NUS, significant gains in F_1 resolution can be realized while economizing spectrometer time by a factor of four. Smaller $^1J_{\text{CC}}$ couplings ($< \sim 30$ Hz) that were not resolved with conventional data sampling can be visualized with 50% NUS data acquisition. On the other hand, the $^nJ_{\text{CC}}$ -modulated ADEQUATE data allowed the successful visualization and quantification of $^nJ_{\text{CC}}$ coupling constants that were as small as 2.3 Hz.

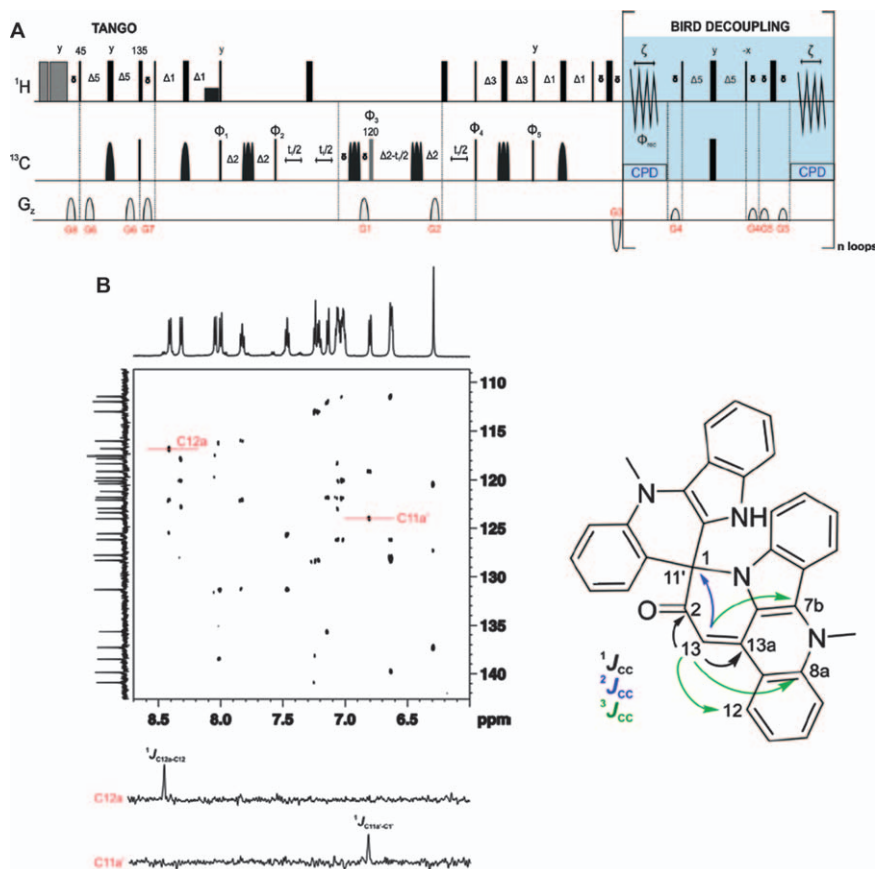


Figure 6.17 (A) Pulse sequence of the homodecoupled ADEQUATE experiment. (B) Experimental $1,n$ -HD-ADEQUATE spectrum performed on a 700 μg sample of crytospirepine dissolved in 35 μL of d_6 -DMSO. (Reproduced from ref. 137 with permission from John Wiley and Sons. © 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

6.4 Perfect NMR

Another interesting area in NMR spectroscopy is the development of robust pulse sequences that provide “perfect” NMR spectra, where signals exhibit undistorted phase properties, allowing the extraction of NMR parameters in a more straightforward way and without tedious post-processing. Two concepts that have been incorporated in modern NMR experiments are revised in this section. First, a “perfect echo” element is used as a general NMR building block to remove/minimize the undesired J_{HH} evolution effects during spin-echo periods. Second, an adiabatic ZQF is employed to eliminate unwanted dispersive AP contributions and to afford pure in-phase signals that are more amenable to a direct and simple analysis.

6.4.1 Perfect Spin-echo

The basic spin-echo element (Figure 6.18A.1) is commonly included in different forms into a great variety of homonuclear and heteronuclear NMR building blocks. For a homonuclear two-spin H_1-H_2 system with a mutual J_{H1-H2} coupling, chemical shift (δ) effects are refocused, whereas J_{HH} evolves during the entire spin-echo according to the following expression:

$$-H_{1y} - H_{2y} \xrightarrow{\frac{\Delta}{2} - 180^\circ(1H, x) - \frac{\Delta}{2}} -H_{1y}c + 2H_{1x}H_{2z}s - H_{2y}c + 2H_{2x}H_{1z}s \quad (6.4)$$

where c is $\cos(\pi J_{H1H2}\Delta)$ and s is $\sin(\pi J_{H1H2}\Delta)$.

The evolution of J_{HH} defined by eqn (6.4) generates a mix of IP and AP contributions that produces phase distortions of the observed signal.

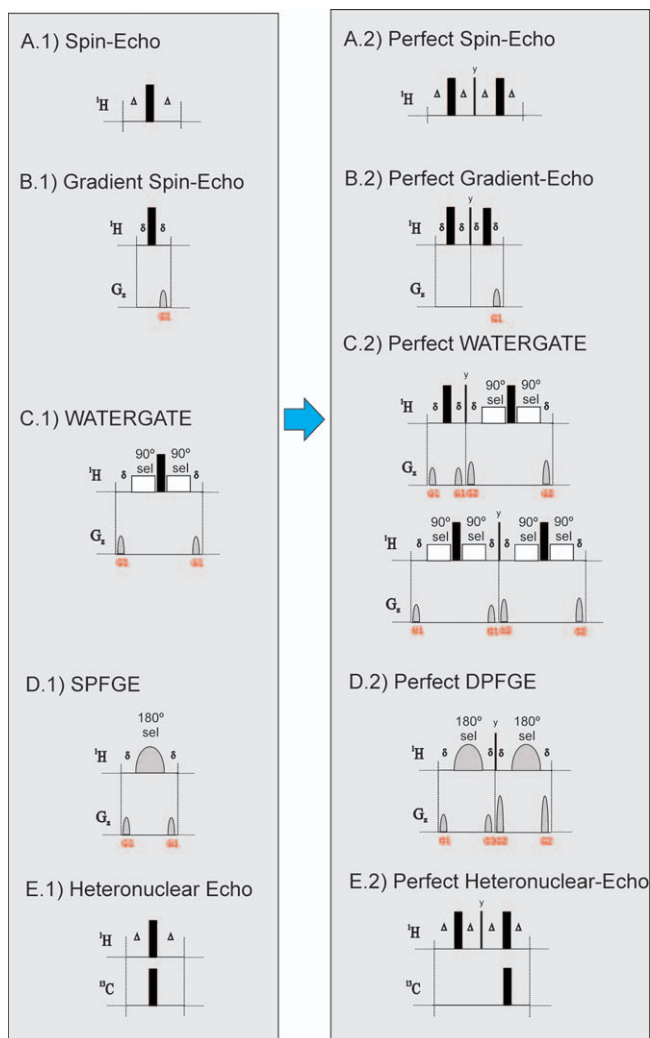


Figure 6.18 NMR building blocks based on a spin-echo or a perfect spin-echo element.

An alternative J -compensated spin-echo building block, referred to as perfect echo (Figure 6.18A.2), simultaneously refocuses both J_{HH} and δ effects by using two sequential spin-echoes separated by a $90^\circ(y)$ pulse:^{141,142}

$$\begin{aligned}
 -H_{1y} - H_{2y} &\xrightarrow{\frac{\Delta}{2} - 180^\circ(1H, x) - \frac{\Delta}{2}} -H_{1y}c + 2H_{1x}H_{2z}s - H_{2y}c + 2H_{2x}H_{1z}s \\
 &\xrightarrow{90^\circ(1H, y)} -H_{1y}c - 2H_{1z}H_{2x}s - H_{2y}c - 2H_{2z}H_{1x}s \\
 &\xrightarrow{\frac{\Delta}{2} - 180^\circ(1H, x) - \frac{\Delta}{2}} -H_{1y}c^2 + 2H_{1x}H_{2z}cs - 2H_{1z}H_{2x}sc - H_{2y}s^2 \\
 &\quad - H_{2y}c^2 + 2H_{2x}H_{1z}cs - 2H_{2z}H_{1x}sc - H_{1y}s^2 \\
 &= -H_{1y} - H_{2y}
 \end{aligned} \tag{6.5}$$

A number of different perfect echo NMR building blocks (Figure 6.18) have been successfully implemented in a series of NMR applications in order to solve some traditional issues: (i) the accurate determination of T_2 relaxation times from undistorted multiplets in “perfect echo CPMG” experiments¹⁴² and their pure shift versions;¹⁰⁵ (ii) during gradient spin-echo elements typically found in diffusion^{143,144} or heteronuclear experiments;¹⁴⁵ (iii) the suppression of J_{HH} evolution during solvent-suppressed WATERGATE periods¹⁴⁶ or selective excitation double-PFG echo (DPFGE) schemes; (iv) the suppression of chemical exchange effects in DOSY experiments;¹⁴⁷ (v) the improvement of heteronuclear transfers by means of the “perfect echo INEPT” element,¹⁴⁸ as demonstrated in a perfect HSQC experiment;¹⁴⁹ (vi) as a method for homodecoupling during the indirect t_1 period in homonuclear (preventing the generation of ZQ artifacts in “perfect NOESY” experiments¹⁵⁰) and heteronuclear experiments (“perfect-HMQC” experiments¹⁵¹); and (vii) to refocus the $^2J_{\text{HH}}$ evolution during BIRD clusters by using a perfect BIRD element, as reported for the measurement of $^1D_{\text{CH}}$ in diastereotopic CH_2 spin systems from fully homodecoupled HSQC spectra.¹²⁸

As an example, Figure 6.19 illustrates the benefits of combining the concepts of a perfect echo and pure shift NMR into the same experiment. Improved T_2 measurements can be obtained from a novel perfect CPMG sequence, referred to as periodic refocusing of J -evolution by coherence transfer (PROJECT) (Figure 6.19A vs. 6.19B).¹⁴² Clearly, the AP contributions observed in the original CPMG sequence are completely removed using the perfect version, allowing much better and more accurate measurement of T_2 values by fitting the experimental data to a clean exponential decay. The IP nature of the PROJECT cross-peaks is suitable for the implementation of broadband homodecoupling. Thus, a HOBS-PROJECT experiment has been proposed to measure accurate T_2 relaxation times in overlapped signals (Figure 6.19C).¹⁰⁵ The resulting singlet signals are easily amenable to conventional mono-exponential curve fitting analysis, without the need for additional data analysis techniques such as deconvolution or line-fitting techniques. The aforementioned PROJECT methods could also be applied as

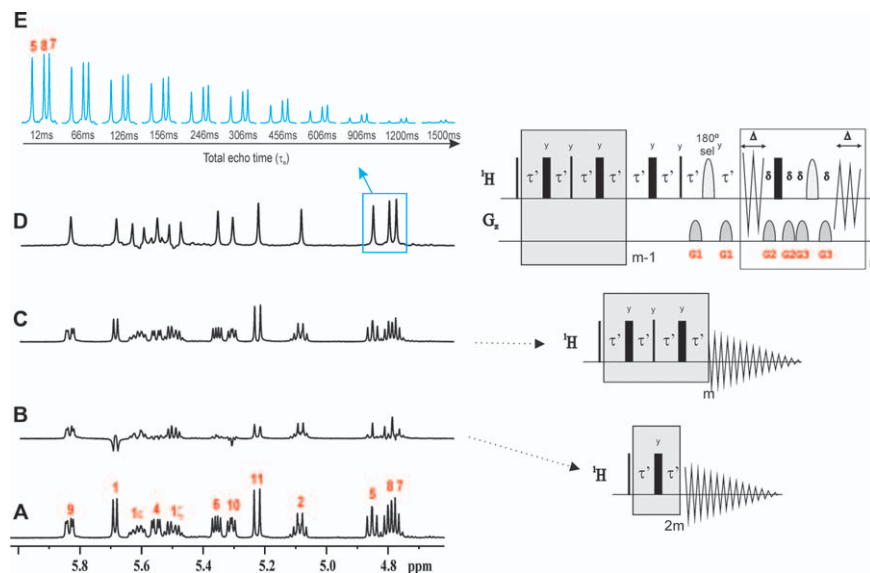


Figure 6.19 Comparison of 1D (A) conventional ^1H , (B) CPMG, (C) PROJECT, and (D) HOBS-PROJECT spectra of cyclosporine acquired with a total echo time of $\tau_e = 156$ ms ($m = 26$ and $\tau' = 1.5$ ms). All spectra were collected under the same experimental conditions as described in Figure 6.2 and are plotted at the same absolute vertical scale. (E) Signal T_2 decays for the H5, H8, and H7 protons in the HOBS-PROJECT experiment. (Reproduced from L. Castañar *et al.*, Measurement of T_1/T_2 relaxation times in overlapped regions from homodecoupled ^1H singlet signals, *J. Magn. Reson.*, **244**, 30–35, Copyright (2014) with permission from Elsevier.¹⁰⁵)

a T_2 filter to suppress signals from high-molecular-weight molecules as found in metabonomic studies or STD NMR experiments.

The perfect gradient echo (Figure 6.18B.2) is another simple building block to remove both δ and J effects during the application of PFGs. The first example was reported to remove the large J -coupling peak distortions in translational diffusion studies using the PFG spin-echo (PGSE) scheme with long PFGs.¹⁴³ Another implementation has been demonstrated in the pure in-phase (PIP)-HSQMBC experiment for measuring accurate $^nJ_{\text{CH}}$ from pure IP cross-peaks in long-range heteronuclear correlation experiments (see Section 6.4.2). Perfect echo versions of the WATERGATE and double-WATERGATE or excitation sculpting schemes (PE-WATERGATE and PE-ES-WATERGATE, respectively) (Figure 6.18C.2) have been proposed to improve J -modulation lineshape distortions on multiplets, and to achieve cleaner and more efficient solvent suppression under highly frequency-selective conditions requiring long-duration pulses without affecting the original excitation profiles (Figure 6.20).¹⁴⁶ These improved elements have also been incorporated in solvent-suppressed PGSE experiments,¹⁴⁴ allowing narrow suppression bandwidths as well as the measurement of diffusion for peaks

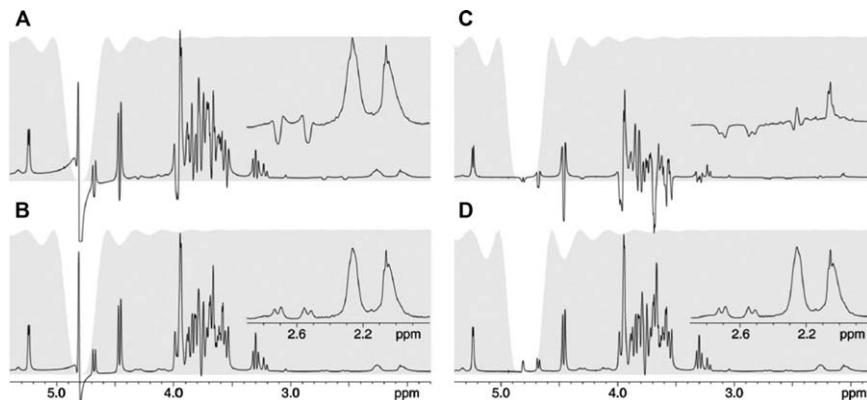


Figure 6.20 Spectra of semi-skimmed (1.7% fat) milk recorded using: (A) conventional WATERGATE; (B) PE-WATERGATE; (C) WATERGATE with excitation sculpting; and (D) PE-ES-WATERGATE, all using 10.5 ms rectangular 90° selective pulses. Calculated excitation profiles are overlaid. (Reproduced from ref. 146 with permission from the Royal Society of Chemistry.)

close to the solvent peak. Both perfect echo WATERGATE sequences perform admirably, and which is the better option depends on the priority one puts on the quality of the solvent suppression or signal loss due to T_2 weighting. Similarly, single-PFG echoes and DPFGEs are the most widely accepted methods for IP selective excitation, both based on the use of a central selective 180° ^1H pulse (Figure 6.18D.1). However, the excitation of a wide region containing mutually J -coupled resonances generates important multiplet distortions, which are more evident in the double-echo version. A perfect echo DPFGE element (Figure 6.18D.2) can be designed for the selective IP excitation of coupled protons.

The heteronuclear version of the perfect echo has also been proposed to optimize heteronuclear transfers (Figure 6.18E.2). The major inconvenience is that the duration of this element is doubled because whereas J_{HH} is perfectly refocused, J_{CH} only evolves during one of the two echoes.¹⁴⁸ An obvious implementation is the perfect INEPT element, which can be incorporated into many heteronuclear NMR pulse sequences. For instance, it has been shown experimentally and by simulation that the J_{HH} interferences present in conventional HSQC experiments can be suppressed in a perfect HSQC pulse scheme that uses perfect echo INEPTs in both defocusing and refocusing heteronuclear transfer periods.¹⁴⁹ Figure 6.21 shows the enhanced results obtained in a simulated perfect CLIP-HSQC experiment (Figure 6.21A) to extract $^1J_{\text{CH}}$ from undistorted cross-peaks along the F_2 dimension. These distortions can be particularly pronounced when large J_{HH} values are present, as found in studies involving molecules in weakly aligned anisotropic media. Experimental results have been provided to determine experimental $^1D_{\text{CH}}$ and $^2D_{\text{HH}}$ RDCs in CH_2 spin systems.

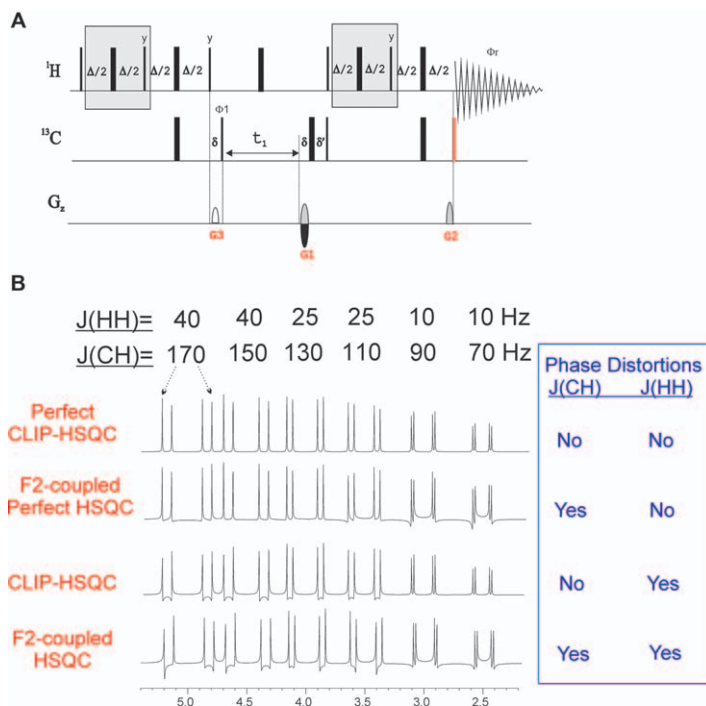


Figure 6.21 (A) Perfect HSQC pulse sequence; (B) simulated spectra showing the phase peak distortion effects in several 140 Hz-optimized F_2 -heterocoupled HSQC experiments: conventional HSQC; CLIP-HSQC; perfect HSQC; and perfect CLIP-HSQC. Six protons have been simulated with different J_{HH} and J_{CH} values, as shown in the upper part. (Reproduced from ref. 149 with permission from John Wiley and Sons. Copyright © 2014 John Wiley & Sons, Ltd.)

Perfect echoes have been proposed to perform broadband homodecoupling in the F_1 dimension of homonuclear and heteronuclear 2D experiments. In conventional NOESY, homonuclear J_{HH} modulation during t_1 creates antiphase coherences that lead to unwanted ZQ and other higher-order coherences. A broadband homodecoupled t_1 period can prevent the creation of AP coherences for shorter periods, whereas chemical shift encoding is still effective. This perfect NOESY experiment¹⁵⁰ will raise the sensitivity of nuclear Overhauser effect cross-peaks for protons involved in J -coupling. It has been shown that further enhancements are achieved using an additional ZQF. On the other hand, broadband homonuclear decoupling has been achieved in the indirect t_1 evolution period of the HMQC scheme by incorporating blocks of perfect echo.¹⁵¹ This perfect HMQC showed better resolution and sensitivity than conventional HMQC spectra, and the F_1 linewidths were very close to those of HSQC for CH_2 groups.

6.4.2 Using ZQFs

The simultaneous application of a swept-frequency pulse and a PFG enables *z*-magnetization to be distinguished from ZQ coherence (ZQC). The implementation of a ZQF efficiently suppresses unwanted AP contributions that usually distort the multiplet pattern of cross-peaks, thereby hampering their analysis.¹⁵² This powerful NMR building block, which enables the elimination of the strong interference caused by ZQC as originally demonstrated for NOESY¹⁵⁰ and TOCSY¹⁵³ experiments, can be readily implemented in many NMR pulse schemes.¹⁵⁴ For instance, the chirp-based *z*-filter was found to be excellent at suppressing AP magnetization while leaving the IP magnetization basically intact in spin-echo and stimulated echo diffusion NMR experiments.¹⁵⁵ *J*-resolved-type experiments incorporating an adiabatic ZQF element at the end of the sequence for the selection of IP magnetization and quadrature detection in the indirect dimension have been reported. The phase-sensitive recording scheme enables the distinction of strong coupling artefacts from symmetry relations, and the full multiplet obtained by the experiment can be reduced by a simple pattern recognition algorithm.¹⁵⁶ Related examples are also found in the nemoDIAG⁹² and G-SERF¹⁰³ experiments.

A ZQF has been implemented in heteronuclear correlation experiments, as reported in PIP-HSQC,¹⁴⁵ PIP-HSQC-TOCSY,¹⁵⁷ and PIP-HSQMBC¹⁴⁵ experiments (Figure 6.22). The clean PIP pattern of cross-peaks is suitable for the direct extraction of coupling constants in resolved signals, as shown in the expanded area of Figure 6.22D. Alternatively, the measurement can be performed using a peak-fitting process from a reference signal, applying the IPAP technique in non-resolved multiplets or using heterodecoupled multiplets as a reference for a fitting process in cases where the additional splitting due to the active $^nJ_{\text{CH}}$ cannot be directly extracted from the non-decoupled multiplet.

A selective HSQMBC-TOCSY pulse scheme incorporating a ZQF has been proposed for the accurate determination of small $^nJ_{\text{CH}}$ coupling values and also, very importantly, their sign (Figure 6.23).¹⁵⁸ The general applicability and high versatility of selHSQMBC-TOCSY solve the major problems that are typically encountered in the original HSQC-TOCSY and HSQMBC experiments while retaining the most important advantages: (i) both protonated and non-protonated carbons can be observed; (ii) the magnitude and the sign of $^nJ_{\text{CH}}$ are directly determined from the analysis of the relative displacement of α/β multiplets, avoiding the need for complex and time-consuming analysis and individualized fitting procedures; (iii) the measurement is made in the detected dimension and, consequently, large numbers of t_1 increments are not required; (iv) pure IP multiplets are obtained that allow the easy and accurate extraction of $^nJ_{\text{CH}}$, even for broad, unresolved, or highly complex multiplets; (v) coupling constants smaller than the linewidth can be determined; and (vi) the combination of different forms of selective proton excitation and TOCSY editing offer excellent versatility and access to non-overlapped resonances, making measurements of coupling constants possible, even in congested areas.

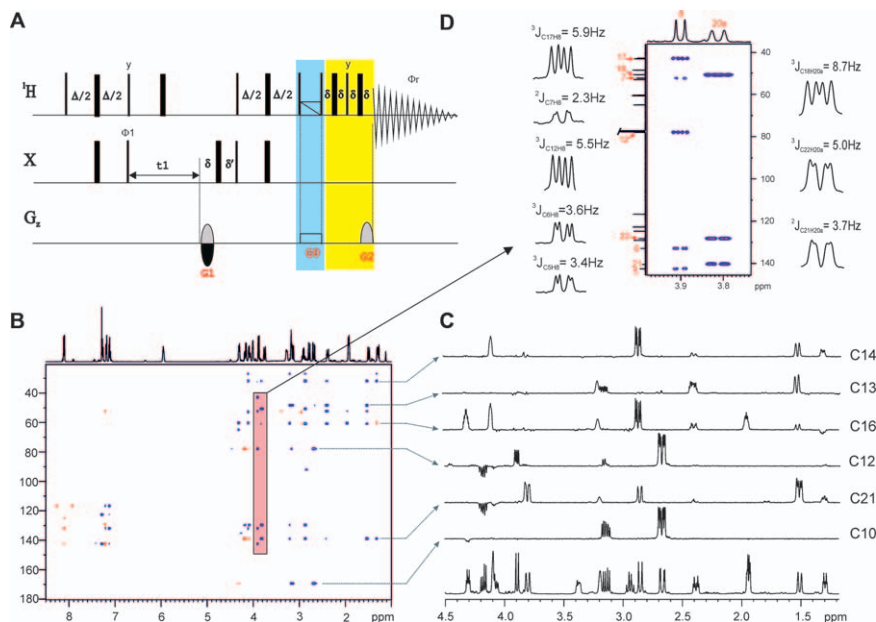


Figure 6.22 (A) Pulse sequence designed to obtain PIP cross-peaks in heteronuclear correlation experiments. The interpulse delay, Δ , is set to $1/(2 \times {}^1J_{\text{CH}})$ or $1/(2 \times {}^nJ_{\text{CH}})$ in PIP-HSQC or PIP-HSQMBC experiments, respectively. The ZQF incorporates a chirped adiabatic 180° ^1H pulse applied simultaneously with a purging G_0 gradient. Broadband heteronuclear decoupling during proton acquisition is optional. (B) 8 Hz-optimized PIP-HSQMBC spectrum of strychnine. (C) 1D row slices taken at different ^{13}C frequencies showing in-phase multiplet patterns for all observed cross-peaks. (D) Expanded area showing how the magnitude of ${}^nJ_{\text{CH}}$ can be easily determined from direct analysis of undistorted and resolved IP peaks. (Reproduced from ref. 145 with permission from John Wiley and Sons. Copyright © 2014 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

6.5 Ultra-long-range Correlation NMR

2D HSQC/HMQC, HSQC-TOCSY, and HMBC/HSQMBC correlation experiments are the most often used heteronuclear NMR experiments for the structure elucidation of natural products (Figure 6.24). All of these pulse schemes use an out-and-back magnetization transfer mechanism, where the starting point and the detected nucleus is always ^1H due to its optimal relaxation and for sensitivity reasons. HSQC/HMQC experiments provides information about directly attached C–H chemical bonds. Alternatively, the multiplicity-edited version of the HSQC experiment affords direct connectivities and carbon multiplicity information in a single and sensitive NMR experiment, significantly economizing spectrometer time. HMBC/HSQMBC experiments provide complementary long-range CH correlation, usually two-, three-, and, only in particular conditions, four-bond separated

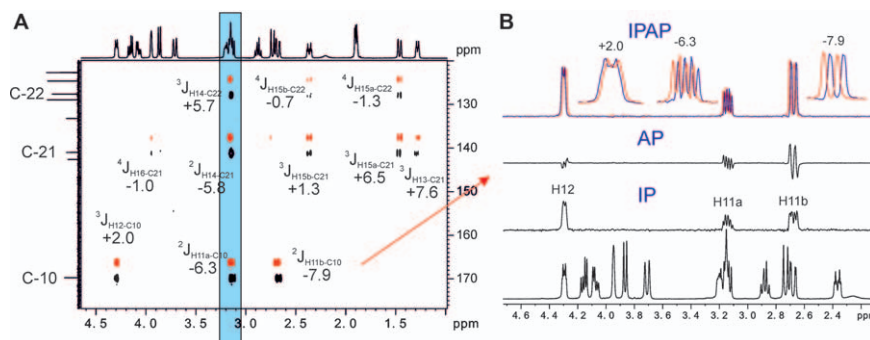


Figure 6.23 Selective HSQMC-TOCSY experiment for the extraction of $^nJ_{CH}$ using the IPAP methodology. 2D IP and AP datasets are separately recorded and further combined ($IP \pm AP$) to provide two different α (in red)/ β (in black) 2D sub-spectra, as shown in (A). The analysis of 1D rows by determining the relative displacement between α/β cross-peaks provides the sign and the magnitude of $^nJ_{CH}$, as shown in (B). (Reproduced from ref. 158 with permission from John Wiley and Sons. Copyright © 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

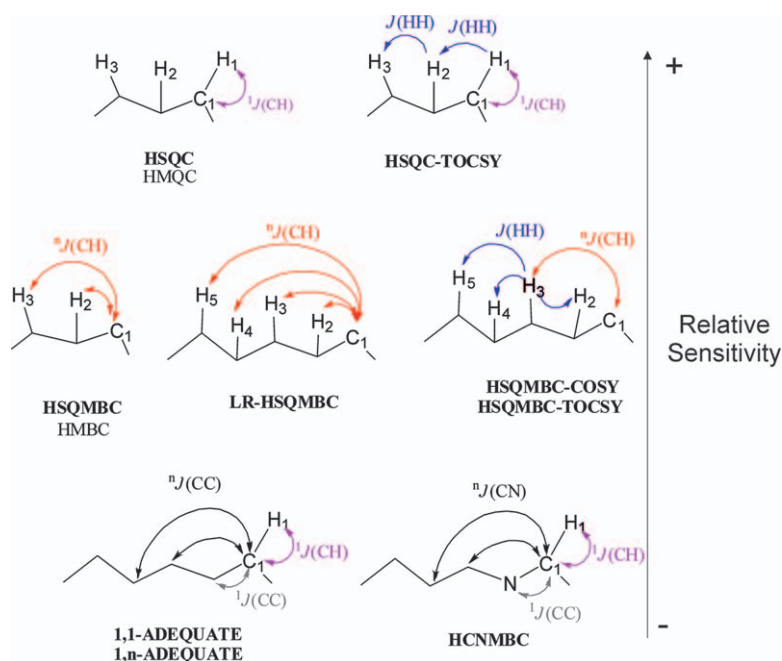


Figure 6.24 Structural fragments derived from direct and long-range heteronuclear shift correlation NMR experiments.

heteronuclear connectivities. HSQC-TOCSY is a hyphenated pulse scheme consisting of a sequential dual-step transfer *via* $^1J_{CH} + J_{HH}$, being another example of those experiments containing multiple pieces of information in a

single spectrum. The experiment simultaneously provides ^1H and ^{13}C correlations of complete spin networks. All of these experiments are routinely used for qualitative analysis and for chemical shift assignment purposes. Modified versions of all these proton-detected experiments are also available for the quantitative measurement of $^1J_{\text{CH}}$ and $^nJ_{\text{CH}}$,¹⁵⁹ and they can all be easily tuned to other heteronuclei, such as ^{19}F , ^{31}P , ^{15}N , ^{29}Si , *etc.*

The detection of long-range heteronuclear correlations corresponding to $^nJ_{\text{CH}}$ values smaller than 1–2 Hz is a challenge in the structural analysis of small molecules and natural products. Recently, several proposals have been made for overcoming the typical $^{2,3}J_{\text{CH}}$ limitation of HMBC by extending the visualization of long-range correlation data to four-, five-, and even six-bond $^nJ_{\text{CH}}$ couplings. These techniques should prove to be effective complements to HMBC spectra for probing challenging, proton-deficient chemical structures, where there is an increasing need to establish $^4J_{\text{CH}}$ and sometimes still longer-range connectivities. A simple strategy should be to optimize a HMBC experiment for a small long-range coupling in the range of 2–4 Hz. However, this optimization choice comes at the expense of considerable loss in sensitivity due to antiphase cancellation of responses when the data are magnitude calculated for presentation and to relaxation due to the longer duration of the delays necessitated by the targeted 2–4 Hz $^nJ_{\text{CH}}$ correlations.

6.5.1 Long-range HSQMBC

A long-range HSQMBC experiment (LR-HSQMBC)¹⁶⁰ has been reported as a complement to the standard 8 Hz-optimized HMBC. It affords comparably more “very-long-range” $^nJ_{\text{CH}}$ correlations (where $n = 4\text{--}6$), with higher sensitivity than can be realized with either HMBC or D-HMBC experiments optimized in the range of 2–4 Hz. As a main feature, this novel HSQMBC pulse scheme uses a refocusing period to generate in-phase correlations with respect to $^nJ_{\text{CH}}$ and therefore minimizes the AP cancellation of responses with very small coupling constants due to the low resolution that is usually attainable in the direct F_2 dimension. In addition, this refocusing allows the application of broadband ^{13}C (or ^{15}N) heteronuclear decoupling during the acquisition period, partially recovering some of the SNR that is lost as a result of relaxation processes encountered during the long duration of the interpulse delays. As an additional feature, HSQMBC data are phase sensitive rather than magnitude calculated as in the HMBC experiment. Table 6.2 summarizes the number of correlations observed in the 2 Hz-optimized LR-HSQMBC spectrum of strychnine (Figure 6.25A) and compares them to other long-range heteronuclear shift correlation experiments. The LR-HSQMBC experiment provided a total of 68 very-long-range correlations that reached further than three bonds ($>^3J_{\text{CH}}$). There were 14 more very-long-range responses compared with the 2 Hz HMBC, and 24 more very-long-range responses than were observed in the D-HMBC data. It should also be noted that the LR-HSQMBC also revealed the largest total number of $^nJ_{\text{CH}}$ correlations out of all experiments tested, with a total of 160 $^nJ_{\text{CH}}$ correlations. In

Table 6.2 Comparison of the number and the nature of long-range heteronuclear responses observed in different long-range correlation NMR experiments performed on the alkaloid strychnine.

	2J	3J	4J	5J	6J	Total $>^3J$	Total
2 Hz LR-HSQMBC ^a	33	59	55	11	2	68	160
2 Hz D-HMBC ^a	29	43	34	8	2	44	116
2 Hz HMBC ^a	34	54	43	10	1	54	142
8 Hz HMBC ^a	34	53	36	4	1	41	129
HSQC-TOCSY ^b	23	15	11	2	0	13	51
8 Hz HSQMBC-Refoc ^b	36	58	40	1	0	41	135
8 Hz HSQMBC-COSY ^b	38	58	58	14	2	74	170
8 Hz HSQMBC-TOCSY ^b	38	58	65	18	3	86	182

^aTaken from Williamson *et al.*¹⁶⁰^bTaken from Sauri *et al.*⁶² TOCSY mixing time of 60 ms.

summary, the 2 Hz LR-HSQMBC experiment affords 24.1% more $^4J_{\text{CH}}$ correlations, a 10% increase in $^5J_{\text{CH}}$ correlations, and a 100% increase in the number of $^6J_{\text{CH}}$ correlations relative to the next most efficient 2 Hz-optimized HMBC experiment. It was also shown that DFT calculations can give reliable estimations of quantitative longer-range couplings. In a comparative study, the LR-HSQMBC experiment also promised to extend the range of HMBC to provide data that are similar to those afforded by 1,*n*-ADEQUATE, even in sample-limited situations and under much better sensitivity conditions. Complementary 2 Hz optimized ^1H - ^{15}N LR-HSQMBC data collected on strychnine showed $^4J_{\text{NH}}$ correlations that had not been previously reported in any HMBC counterpart.¹⁶¹

The capability of the LR-HSQMBC experiment has been challenged with a 400 μg sample of the proton-deficient xanthone antibiotic cervinomycin A₂; DFT calculations were employed to determine whether the very-long-range correlations observed were reasonable on the basis of the calculated long-range couplings (Figure 6.26). Table 6.2 shows the impact of LR-HSQMBC data as a supplement to HMBC data in CASE programs.¹⁶² Similar conclusions were drawn from more extensive NMR data for the alkaloids staurosporine and crytospirepine, including 1,1- and 1,*n*-ADEQUATE data, demonstrating that LR-HSQMBC data provide a very significant improvement in CASE program calculation times and a sizeable reduction in the numbers of structures generated in the output file.

6.5.2 HSQMBC-TOCSY

Analog 2D HSQMBC-COSY and HSQMBC-TOCSY pulse schemes⁶² based on a sequential dual transfer step $^nJ_{\text{CH}} + J_{\text{HH}}$, have also been evaluated as complementary NMR methods to HMBC/HSQMBC and LR-HSQMBC. The incorporation of an additional J_{HH} transfer step at the end of a refocused HSQMBC scheme (Figure 6.27A) can favor the sensitive observation of traditionally missing or very weak correlations and, in addition, facilitates the

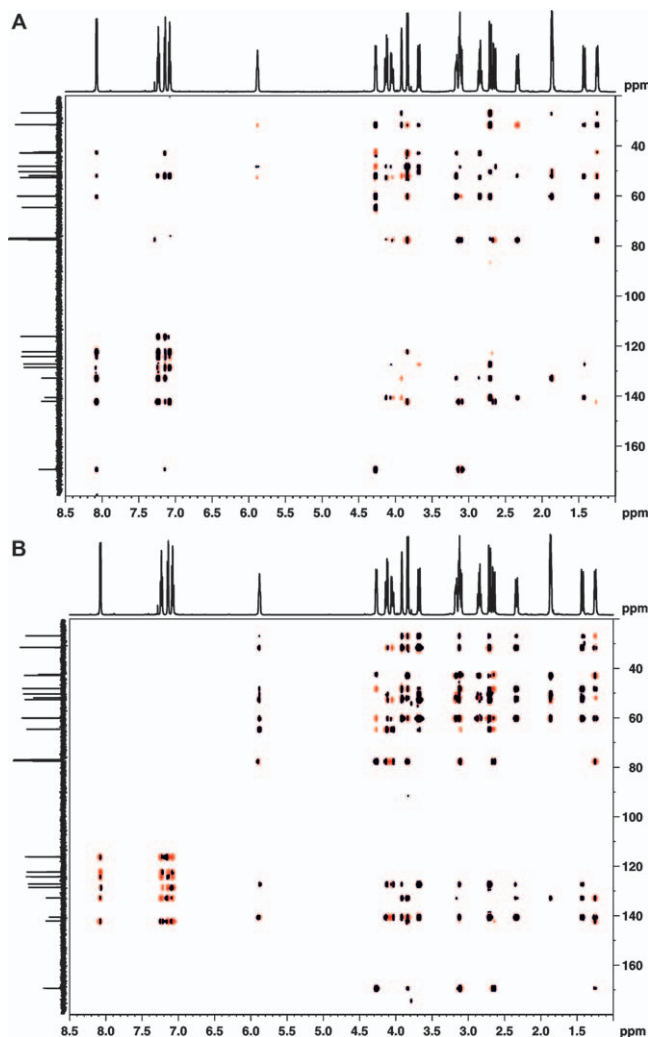


Figure 6.25 (A) 2 Hz-optimized LR-HSQMBC and (B) 8 Hz-optimized HSQMBC-TOCSY spectra of strychnine in CDCl_3 . For a confidence comparison, both experiments were sequentially recorded in a 1.7 mm MicroCryo-ProbeTM using eight scans per each of the 128 t_1 increments. The overall acquisition time for each experiment was about 30 minutes.

detection of a significant number of still-longer-range correlations with both protonated and non-protonated carbons under optimum sensitivity conditions (Figure 6.27B). Taking as a basis the theoretical J_{CH} values of strychnine, the following small coupling correlations were expected: five correlations were expected with $^2J_{\text{CH}} < 1$ Hz; 12 $^2J_{\text{CH}}$ correlations in the range of 1–2 Hz; 11 with $^3J_{\text{CH}} < 1$ Hz; and eight $^3J_{\text{CH}}$ correlations in the range of 1–2 Hz. This means that 42.5% of $^2J_{\text{CH}}$ and 28% of $^3J_{\text{CH}}$ connectivities

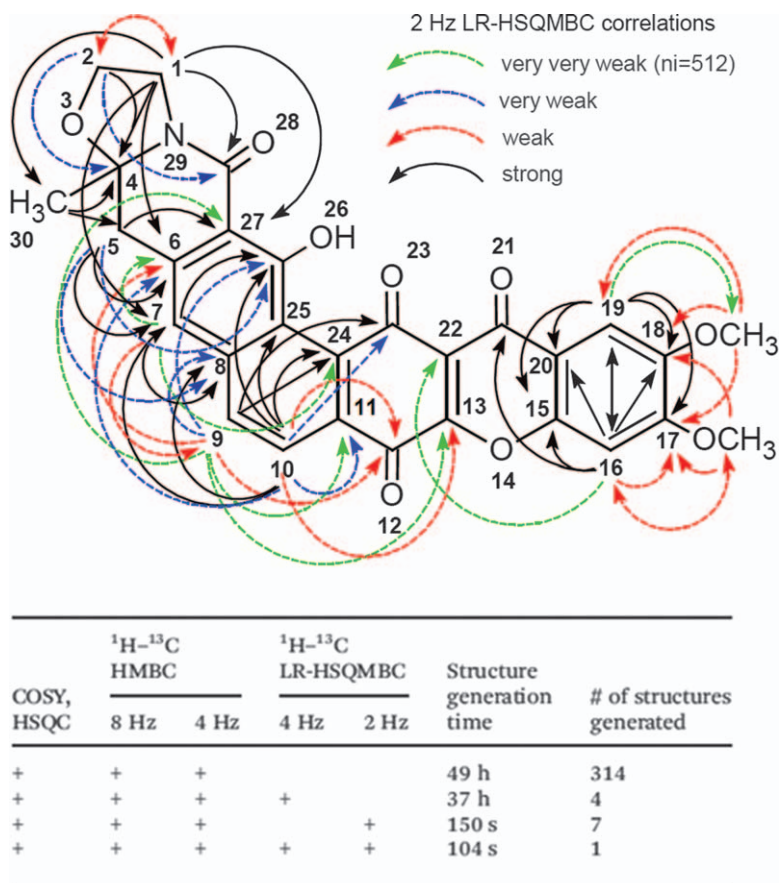


Figure 6.26 Correlations observed in a 2 Hz LR-HSQMBC spectrum of cervinomycin A₂ and a summary of Structure Elucidator™ results with various 2D NMR used as a input. (Reproduced from ref. 162 with permission from the Royal Society of Chemistry.)

present couplings smaller than 2 Hz and, *a priori*, they would exhibit very small intensities or would even be absent in a conventional 8 Hz-optimized HSQMBC experiment. The results obtained with the HSQMBC-TOCSY experiment exceeded the performance of all other experiments. A total of 182 correlations are observed, 86 of them being of $>^3J_{CH}$. This represents a 14% increase over the total number of correlations and a 26% increase in the number of correlations over $^3J_{CH}$ observed in a 2 Hz reference LR-HSQMBC spectrum. The intensity of relayed cross-peaks does not depend of its corresponding $^nJ_{CH}$ value and therefore HSQMBC-TOCSY is highly useful for detecting correlations with $^nJ_{CH} < 1$ Hz in optimum sensitivity conditions. In addition, the number of connectivities could be increased provided that the use of a longer TOCSY mixing time allows the J_{HH} propagation

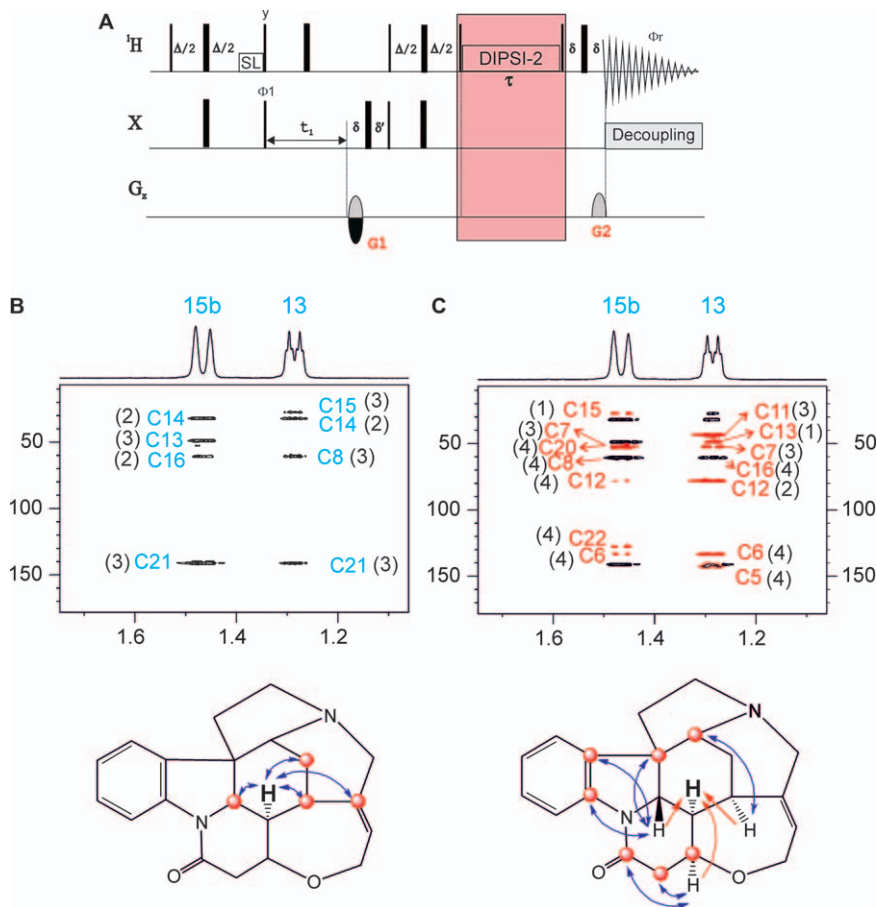


Figure 6.27 (A) Pulse sequence of the HSQMBC-TOCSY experiment. (B & C) Expanded area including the H13 and H15b protons of the 8 Hz-optimized HSQMBC and HSQMBC-TOCSY spectra of strychnine, respectively. (Reproduced from J. Sauri *et al.*, Extending long-range heteronuclear NMR connectivities by HSQMBC-COSY and HSQMBC-TOCSY experiments, *J. Magn. Reson.*, 258, 25–32, Copyright (2015) with permission from Elsevier.⁶²)

without serious sensitivity losses due to relaxation. The analog ^1H – ^{15}N HSQMBC-TOCSY spectrum on strychnine was also reported in the same study.

6.5.3 ADEQUATE

Despite its low sensitivity due to the need for the J_{CC} in molecules at natural abundance, ADEQUATE experiments has proved to be indispensable for tracing out the structure skeleton in highly challenging molecules, especially in those parts of the molecular structure with a low density of protons.^{163–167}

Of the several versions initially proposed, 1,1- and 1,*n*-ADEQUATE have more practical utility. 1,1-ADEQUATE provides a conventional 2D ^1H - ^{13}C correlation map displaying only two-bond H(C)C connectivities through a sequential $^1J_{\text{CH}} + ^1J_{\text{CC}}$ out-and-back transfer mechanism. This experiment is essential to differentiating $^2J_{\text{CH}}$ from $^3J_{\text{CH}}$ connectivities in conventional HMBC or HSQMBC experiments. On the other hand, three-bond H(CC)C and four-bond H(CCC)C connectivities are detected in a less sensitive 1,*n*-ADEQUATE experiment using equivalent $^1J_{\text{CH}} + ^nJ_{\text{CC}}$ transfers. Assuming that $^3J_{\text{CC}}$ usually predominates in 1,*n*-ADEQUATE, $^4J_{\text{CH}}$ connectivities can be roughly estimated by comparing HMBC/HSQMBC and ADEQUATE datasets, as reported for the complex, highly proton-deficient alkaloid staurosporine.^{168,169} Of interest, $^1J_{\text{CC}}$ and $^nJ_{\text{CC}}$ connectivities are both detected in a 1,*n*-ADEQUATE experiment, and a asymmetric modification of the original pulse sequence, referred to as an inverted $^1J_{\text{CC}}$ 1,*n*-ADEQUATE experiment,¹⁷⁰ allows the editing of these correlations in the form of opposite signal phases. It has been reported that spectra that are well digitized in F_1 can be acquired overnight with a 4.5 mg sample of strychnine in a 1.7 mm MicroCryoProbe™. A further improvement involves a dual optimization of the experiment with judiciously chosen values for the $^1J_{\text{CC}}$ and $^nJ_{\text{CC}}$ delays, summing and normalizing them in a single experiment. In this way, it is possible to achieve broadband inversion of the $^1J_{\text{CC}}$ correlations across the range of $^1J_{\text{CC}} = 29\text{--}82$ Hz in a single NMR experiment (Figure 6.28).¹⁷¹ Advantages of the dual-optimized approach are the generally better averaged SNR for observed cross-peaks, the improved detection of weaker long-range correlations, and no knowledge of the size of J_{CC} being required for acquisition parameter set-up. Recently, BIRD-based homo-decoupled versions of 1,1- and 1,*n*-ADEQUATE¹³⁷ and their *J*-modulated counterparts, optionally combined with NUS, have been proposed in order to improve their performance in terms of sensitivity and measurement time (see Section 6.3.4).

Several attempts have been made at using covariance processing to mathematically combine pairs of NMR spectra that share a common frequency domain in order to facilitate the extraction of correlation information. Examples have been reported of mixing ADEQUATE with HSQC or HMBC data into the same spectral representation without the need for newly designed pulse sequences or new and time-consuming data acquisitions. Combinations between HSQC and 1,1-ADEQUATE,^{172,173} 1,*n*-ADEQUATE,^{173,174} and inverted $^1J_{\text{CC}}$ 1,*n*-ADEQUATE spectra¹⁷⁵ have been reported to afford carbon-carbon correlation spectra that allow the extraction of direct $^1J_{\text{CC}}$, long-range ($^nJ_{\text{CC}}$, where $n \geq 2$), and $^1J_{\text{CC}}$ -edited long-range correlation data, respectively. Covariance processing of HMBC and 1,1-ADEQUATE spectra has also recently been reported,^{176,177} allowing convenient, high-sensitivity access to $^nJ_{\text{CC}}$ correlation data, and these processes are equivalent to the much lower-sensitivity *n*,1-ADEQUATE experiment. Furthermore, HMBC-1,1-ADEQUATE correlations are observed in the F_1 frequency domain at the intrinsic chemical shift of the ^{13}C resonance in

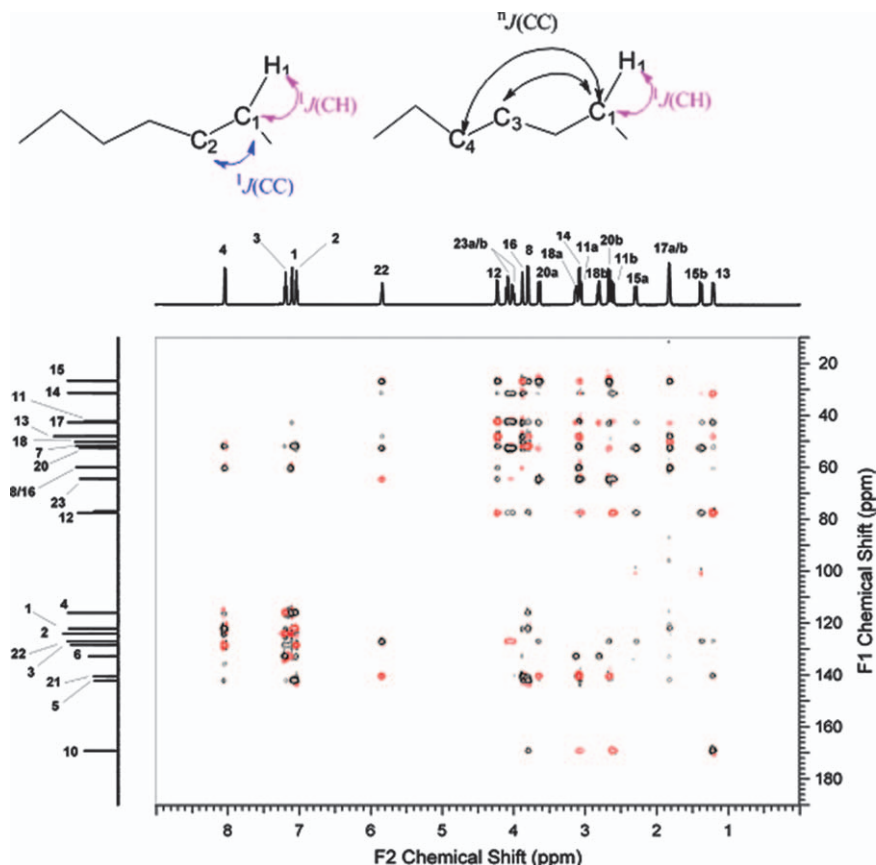


Figure 6.28 Dual-optimized inverted $^1J_{CC}$ 1,*n*-ADEQUATE spectrum of strychnine. The opposite phase between cross-peaks allows us to distinguish $^1J_{CC}$ (red) from $^nJ_{CC}$ (black) connectivities. The data were acquired as 3072×200 real points, with 128 transients accumulated for each set of optima. The experiment was dual optimized for $^1J_{CC} = 57$ Hz and $^nJ_{CC} = 9.5$ Hz for the first optimization and $^1J_{CC} = 64$ Hz and $^nJ_{CC} = 8$ Hz for the second optimization. The acquisition time for the spectrum was 48 hours and 42 minutes. (Reproduced from M. Reibarkh *et al.*, Broadband inversion of $^1J_{CC}$ responses in 1,*n*-ADEQUATE spectra, *J. Magn. Reson.*, **236**, 126–133, Copyright (2013), with permission from Elsevier.¹⁷¹)

question, rather than at the double-quantum frequency of the pair of correlated carbons, as visualized by the *n*,1- and *m*,*n*-ADEQUATE experiments, greatly simplifying data interpretation. Additionally, HMBC and 1,*n*-ADEQUATE¹⁷⁸ spectra were combined to generate long-range carbon–carbon correlation data that were equivalent to the very-low-sensitivity *m*,*n*-ADEQUATE experiment. These HMBC–1,*n*-ADEQUATE spectra afford a potential means of establishing the structures of severely proton-deficient molecules.

In analogy with ADEQUATE experiments, a triple-resonance HCNMBC experiment has been reported for multiple-bond H-(C)-N correlation experiments using a sequential magnetization transfer *via* $^1J_{\text{CH}}$ and one- or multiple-bond $^nJ_{\text{CN}}$ coupling constants (typically $n = 1-3$) at the natural isotopic abundance.^{179,180} The pulse sequence reported provides high-quality spectra, even for sub-milligram samples. For instance, a 10 μg sample of glycine in D_2O is sufficient to obtain the HCNMBC spectrum in ~ 12 hours using a 1.7 mm cryoprobe. These preliminary results indicate that the pulse sequence has great potential in the structural determination of nitrogen heterocycles. The experiment produces data that are highly complementary to ^1H - ^{15}N HMBC-type correlations in that they can provide ^{15}N chemical shift data for nitrogen that may not show up in the HMBC and, in addition, they can also be used to directly measure J_{CN} coupling constants.

6.6 Future Perspectives

In the coming years, a consolidation period of the different concepts described in this chapter is expected, with the advent of new developments that will indicate the possibilities and also practical limitations of each technique in the structure elucidation and mixture analysis of small molecules. It will be interesting to see whether such improvements will positively affect detection limits, speed up measurement times, or allow widespread analytical use in automation mode.

Certainly, the current major challenge in pure shift NMR spectroscopy is to improve sensitivity. The most sensitive spatial encoding methods only recover 5–10% of the signal, and although the HOBS methodology offers full sensitivity, it is limited only to particular spin systems (*e.g.* NH or H_α regions in peptides and proteins). Novel and general strategies are needed in order to achieve efficient broadband homodecoupling for all signals in the entire NMR spectrum and to reinforce spectral quality, namely the improvement of strong coupling effects (particularly in diastereotopic CH_2 spin systems), the improvement of linewidths, and the minimization of the presence of unwanted artefacts and sidebands.

We will also have to be alert to the new advances in NUS schedules and spectral reconstruction methods, which will play fundamental roles in the reduction of overall measurement times in high-dimensional NMR experiments. It is expected that the combination of multiple resolution-enhanced techniques in a single NMR experiment will simultaneously reinforce the concepts of signal resolution and digital resolution in so-called ultra-high-resolution NMR spectroscopy. For instance, it will be interesting to evaluate the potential use of pure shift HSQC combined with other resolution-enhanced techniques in the automated analysis of complex mixtures (*e.g.* in metabolomic studies^{61,181–183} or for quantitative NMR).

A key remaining challenge in automated structure elucidation protocols is the unambiguous distinction between $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ in HMBC experiments and between $^2J_{\text{CC}}$ and $^3J_{\text{CC}}$ in ADEQUATE experiments. An important step is

the incorporation of these unambiguous data into CASE programs in order to help reduce calculation times and generate a minor number of solutions. The design of sign-sensitive determination of the experiments complemented with theoretical DFT calculations can be of interest.

Acknowledgements

Financial support for this research provided by the MINECO of Spain (project CTQ2015-64436-P) is gratefully acknowledged.

References

1. T. D. W. Claridge, *High-resolution NMR Techniques in Organic Chemistry*, 1999.
2. R. C. Breton and W. F. Reynolds, *Nat. Prod. Rep.*, 2013, **30**, 501.
3. T. Carlomagno, *Nat. Prod. Rep.*, 2012, **29**, 536.
4. M. Halabalaki, K. Vougianniopoulou, E. Mikros and A. L. Skaltsounis, *Curr. Opin. Biotechnol.*, 2014, **25**, 1–7.
5. R. Teodorescu, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 26.
6. C. Anklin, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 38.
7. K. L. Colson, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 58.
8. T. F. Molinski, *Nat. Prod. Rep.*, 2010, **27**, 321.
9. A. Krunić and J. Orjala, *Magn. Reson. Chem.*, 2015, **53**, 1043.
10. M. Spraul, U. Braumann, M. Godejohann, C. Daolio and L.-H. Tseng, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 71.
11. P. Giraudeau, Y. Shrot and L. Frydman, *J. Am. Chem. Soc.*, 2009, **131**, 13902.
12. M. E. Elyashberg, A. J. Williams and G. E. Martin, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2008, **53**, 1.
13. M. E. Elyashberg, A. J. Williams and K. A. Blinov, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 187.
14. H. M. Ge, H. Sun, N. Jiang, Y. H. Qin, H. Dou, T. Yan, Y. Y. Hou, C. Griesinger and R. X. Tan, *Chemistry*, 2012, **18**, 5213.
15. P. Trigo-Mouriño, R. Sifuentes, A. Navarro-Vázquez, C. Gayathri, H. Maruenda and R. R. Gil, *Nat. Prod. Commun.*, 2012, **7**, 735.
16. P. Trigo-Mouriño, A. Navarro-Vázquez, J. Ying, R. R. Gil and A. Bax, *Angew. Chem., Int. Ed.*, 2011, **50**, 7576.
17. R. R. Gil, *Angew. Chem., Int. Ed.*, 2011, **50**, 7222.
18. T. Parella, *Magn. Reson. Chem.*, 1998, **36**, 467.
19. R. Freeman and E. Kupče, *Magn. Reson. Chem.*, 2007, **45**, 2.

20. D. Schulze-Sünninghausen, J. Becker and B. Luy, *J. Am. Chem. Soc.*, 2014, **136**, 1242.
21. J. Furrer, *Chem. Commun.*, 2010, **46**, 3396.
22. J. Becker and B. Luy, *Magn. Reson. Chem.*, 2015, **53**, 878.
23. L. Frydman, T. Scherf and A. Lupulescu, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 15858.
24. L. Frydman, A. Lupulescu and T. Scherf, *J. Am. Chem. Soc.*, 2003, **125**, 9204.
25. P. Pelupessy, *J. Am. Chem. Soc.*, 2003, **125**, 12345.
26. A. Tal and L. Frydman, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **57**, 241.
27. T. Parella and P. Nolis, *Concepts Magn. Reson., Part A*, 2010, **36**, 1.
28. P. Nolis, M. Pérez-Trujillo and T. Parella, *Angew. Chem., Int. Ed.*, 2007, **46**, 7495.
29. E. Kupče, R. Freeman and B. K. John, *J. Am. Chem. Soc.*, 2006, **128**, 9606.
30. E. Kupče, *Top. Curr. Chem.*, 2013, **335**, 71.
31. G. B. B. Njock, D. E. Pegnyem, T. A. Bartholomeusz, P. Christen, B. Vitorge, J.-M. Nuzillard, R. Shivapurkar, M. Foroozandeh and D. Jeannerat, *Chimia*, 2010, **64**, 235.
32. W. F. Reynolds and R. G. Enriquez, *Magn. Reson. Chem.*, 2003, **41**, 927.
33. D. Jeannerat, *J. Magn. Reson.*, 2007, **186**, 112.
34. K. Kazimierczuk and V. Y. Orekhov, *Angew. Chem., Int. Ed.*, 2011, **50**, 5556.
35. V. Y. Orekhov and V. A. Jaravine, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **59**, 271.
36. S. G. Hyberts, H. Arthanari and G. Wagner, *Top. Curr. Chem.*, 2012, **316**, 125.
37. D. Rovnyak, M. Sarcone and Z. Jiang, *Magn. Reson. Chem.*, 2011, **49**, 483.
38. M. Mobli and J. C. Hoch, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2014, **83**, 21.
39. K. Kazimierczuk and V. Orekhov, *Magn. Reson. Chem.*, 2015, **53**, 921.
40. M. R. Palmer, R. A. Gupta, M. E. Richard, C. L. Suiter, T. Polenova, J. C. Hoch and D. Rovnyak, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 93.
41. K. Blinov and A. J. Williams, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 259.
42. F. Zhang and R. Brüschweiler, *J. Am. Chem. Soc.*, 2004, **126**, 13180.
43. D. A. Snyder and R. Brüschweiler, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 244.
44. W. Reynolds, R. Breton and D. Burns, *Planta Med.*, 2014, **80**, IL26.
45. A. Herrera, E. Fernández-Valle, R. Martínez-Álvarez, D. Molero-Vílchez, Z. D. Pardo-Botero and E. Sáez-Barajas, *Magn. Reson. Chem.*, 2015, **53**, 952.
46. S. Akoka and P. Giraudeau, *Magn. Reson. Chem.*, 2015, **53**, 986.

47. W. Qiu, Y. Chen, Z. Wei, J. Yang, Y. Lin and Z. Chen, *Magn. Reson. Chem.*, 2015, **53**, 945.
48. M. Pérez-Trujillo, P. Nolis, W. Bermel and T. Parella, *Magn. Reson. Chem.*, 2007, **45**, 325.
49. P. Nolis, M. Pérez-Trujillo and T. Parella, *Magn. Reson. Chem.*, 2006, **44**, 1031.
50. M. Pérez-Trujillo, P. Nolis and T. Parella, *Org. Lett.*, 2007, **9**, 29.
51. J. Saurí, P. Nolis and T. Parella, *Magn. Reson. Chem.*, 2015, **53**, 427.
52. C. Dalvit, A. D. Gossert, J. Coutant and M. Piotto, *Magn. Reson. Chem.*, 2011, **49**, 199.
53. R. Freeman and Ě. Kupče, *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*, 2015, **1**, 119.
54. Ě. Kupče and R. Freeman, *J. Am. Chem. Soc.*, 2008, **130**, 10788.
55. R. Freeman and Ě. Kupče, *Magn. Reson. Chem.*, 2010, **48**, 333.
56. Ě. Kupče and B. Wrackmeyer, *Appl. Organomet. Chem.*, 2010, **24**, 837.
57. Ě. Kupče, S. Cheatham and R. Freeman, *Magn. Reson. Chem.*, 2007, **45**, 378.
58. Ě. Kupče and R. Freeman, *J. Magn. Reson.*, 2011, **213**, 1.
59. S. Kim and T. Szyperski, *J. Am. Chem. Soc.*, 2003, **125**, 1385.
60. S. M. Pudakalakatti, A. Dubey, G. Jaipuria, U. Shubhashree, S. K. Adiga, D. Moskau and H. S. Atreya, *J. Biomol. NMR*, 2014, **58**, 165.
61. A. Le Guennec, J.-N. Dumez, P. Giraudeau and S. Caldarelli, *Magn. Reson. Chem.*, 2015, **53**, 913.
62. J. Saurí, N. Marcó, R. T. Williamson, G. E. Martin and T. Parella, *J. Magn. Reson.*, 2015, **258**, 25.
63. D. Jeannerat, *Magn. Reson. Chem.*, 2003, **41**, 3.
64. B. Vitorge, S. Bieri, M. Humam, P. Christen, K. Hostettmann, O. Muñoz, S. Loss and D. Jeannerat, *Chem. Commun.*, 2009, 950.
65. M. Foroozandeh and D. Jeannerat, *Chemphyschem*, 2010, **11**, 2503.
66. M. Foroozandeh and D. Jeannerat, *Magn. Reson. Chem.*, 2015, **53**, 894.
67. K. Ramírez-Gualito and D. Jeannerat, *Magn. Reson. Chem.*, 2015, **53**, 901.
68. A. Cotte, M. Foroozandeh and D. Jeannerat, *Chimia*, 2012, **66**, 764.
69. N. Marcó, A. Fredi and T. Parella, *Chem. Commun.*, 2015, **51**, 3262.
70. M. Pérez-Trujillo, L. Castañar, E. Monteagudo, L. T. Kuhn, P. Nolis, A. Virgili, R. T. Williamson and T. Parella, *Chem. Commun.*, 2014, **50**, 10214.
71. L. Castañar, R. Roldán, P. Clapés, A. Virgili and T. Parella, *Chem. – Eur. J.*, 2015, **21**, 7682.
72. K. Zangger, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2015, **86–87**, 1.
73. N. H. Meyer and K. Zangger, *ChemPhysChem*, 2014, **15**, 49.
74. L. Castañar and T. Parella, *Magn. Reson. Chem.*, 2015, **53**, 399.
75. R. W. Adams, *eMagRes*, 2014, **3**, 295.
76. K. Zangger and H. Sterk, *J. Magn. Reson.*, 1997, **489**, 486.
77. A. J. Pell and J. Keeler, *J. Magn. Reson.*, 2007, **189**, 293.

78. M. Foroozandeh, R. W. Adams, P. Kiraly, M. Nilsson and G. A. Morris, *Chem. Commun.*, 2015, **51**, 15410.
79. A. J. Pell, R. A. E. Edden and J. Keeler, *Magn. Reson. Chem.*, 2007, **45**, 296.
80. R. Brüschweiler, C. Griesinger, O. W. Sørensen and R. R. Ernst, *J. Magn. Reson.*, 1988, **78**, 178.
81. V. V. Krishnamurthy, *Magn. Reson. Chem.*, 1997, **35**, 9.
82. M. Foroozandeh, R. W. Adams, M. Nilsson and G. A. Morris, *J. Am. Chem. Soc.*, 2014, **136**, 11867.
83. A. Bax, A. Mehlkopf and J. Smidt, *J. Magn. Reson.*, 1979, **35**, 167.
84. Y. Xia, G. Legge, K. Y. Jun, Y. Qi, H. Lee and X. Gao, *Magn. Reson. Chem.*, 2005, **43**, 372.
85. J. A. Aguilar, A. A. Colbourne, J. Cassani, M. Nilsson and G. A. Morris, *Angew. Chem., Int. Ed.*, 2012, **51**, 6460.
86. J. A. Aguilar, J. Cassani, M. Delbianco, R. W. Adams, M. Nilsson and G. A. Morris, *Chem. – Eur. J.*, 2015, **21**, 6623.
87. J. A. Aguilar, S. Faulkner, M. Nilsson and G. A. Morris, *Angew. Chem., Int. Ed.*, 2010, **49**, 3901.
88. D. Pitoux, Z. Hu, B. Plainchont, D. Merlet, J. Farjon, D. Bonnañfé and N. Giraud, *Magn. Reson. Chem.*, 2015, **53**, 836.
89. P. Sakhaei, B. Haase, W. Bermel, R. Kerssebaum, G. E. Wagner and K. Zangger, *J. Magn. Reson.*, 2013, **233**, 92.
90. N. Suryaprakash, *Chem. Commun.*, 2014, **50**, 8550.
91. L. Castañar, P. Nolis, A. Virgili and T. Parella, *Chem. – Eur. J.*, 2013, **19**, 15472.
92. A. Cotte and D. Jeannerat, *Angew. Chem., Int. Ed.*, 2015, **54**, 6016.
93. M. Foroozandeh, R. W. Adams, N. J. Meharry, D. Jeannerat, M. Nilsson and G. A. Morris, *Angew. Chem., Int. Ed.*, 2014, **53**, 6990.
94. N. H. Meyer and K. Zangger, *Angew. Chem., Int. Ed.*, 2013, **52**, 7143.
95. J. Mauhart, S. Glanzer, P. Sakhaei, W. Bermel and K. Zangger, *J. Magn. Reson.*, 2015, **259**, 207.
96. L. Castañar, P. Nolis, A. Virgili and T. Parella, *Chem. – Eur. J.*, 2013, **19**, 17283.
97. J. Ying, J. Roche and A. Bax, *J. Magn. Reson.*, 2014, **241**, 97.
98. R. W. Adams, L. Byrne, P. Király, M. Foroozandeh, L. Paudel, M. Nilsson, J. Clayden and G. A. Morris, *Chem. Commun.*, 2014, **50**, 2512.
99. M. Pérez-Trujillo, T. Parella and L. T. Kuhn, *Anal. Chim. Acta*, 2015, **876**, 63.
100. L. Castañar, M. Pérez-Trujillo, P. Nolis, E. Monteagudo, A. Virgili and T. Parella, *Chemphyschem*, 2014, **15**, 854–857.
101. N. Gubensäk, W. M. F. Fabian and K. Zangger, *Chem. Commun.*, 2014, **50**, 12254.
102. N. Lokesh, S. R. Chaudhari and N. Suryaprakash, *Chem. Commun.*, 2014, **50**, 15597.
103. D. Pitoux, B. Plainchont, D. Merlet, Z. Hu, D. Bonnañfé, J. Farjon and N. Giraud, *Chem. – Eur. J.*, 2015, **21**, 9044.

104. J. E. Herbert Puchet, D. Putoux, C. M. Grison, S. Robin, D. Merlet, D. J. Aitken, N. Giraud and J. Farjon, *Chem. Commun.*, 2015, **51**, 7939.
105. L. Castañar, P. Nolis, A. Virgili and T. Parella, *J. Magn. Reson.*, 2014, **244**, 30.
106. J. M. McKenna and J. A. Parkinson, *Magn. Reson. Chem.*, 2014, **53**, 249.
107. M. Nilsson and G. A. Morris, *Chem. Commun.*, 2007, 933.
108. S. Islam, J. A. Aguilar, M. W. Powner, M. Nilsson, G. A. Morris and J. D. Sutherland, *Chem. – Eur. J.*, 2013, **19**, 4586.
109. S. Glanzer and K. Zangger, *Chem. – Eur. J.*, 2014, **20**, 11171.
110. S. R. Chaudhari and N. Suryaprakash, *RSC Adv.*, 2014, **4**, 15018.
111. J. A. Aguilar, G. A. Morris and A. M. Kenwright, *RSC Adv.*, 2014, **4**, 8278.
112. S. R. Chaudhari and N. Suryaprakash, *Chemphyschem*, 2015, **16**, 1079.
113. S. Glanzer and K. Zangger, *J. Am. Chem. Soc.*, 2015, **137**, 5163.
114. S. Glanzer and K. Zangger, *Chemphyschem*, 2015, **16**, 3313.
115. A. Cotte, M. Foroozandeh and D. Jeannerat, *Chim. Int. J. Chem.*, 2012, **66**, 764.
116. V. M. R. Kakita, S. P. B. Vemulapalli and J. Bharatam, *Magn. Reson. Chem.*, 2016, **54**, 308.
117. J. A. Aguilar, M. Nilsson and G. A. Morris, *Angew. Chem., Int. Ed.*, 2011, **50**, 9716.
118. P. Sakhaei, B. Haase, W. Bermel, R. Kerssebaum, G. E. Wagner and K. Zangger, *J. Magn. Reson.*, 2013, **233**, 92.
119. A. Lupulescu, G. L. Olsen and L. Frydman, *J. Magn. Reson.*, 2012, **218**, 141.
120. K. J. Donovan and L. Frydman, *Angew. Chem., Int. Ed.*, 2014, **54**, 594.
121. N. H. Meyer and K. Zangger, *Angew. Chem., Int. Ed.*, 2013, **52**, 7143.
122. K. Rachineni, V. M. R. Kakita, S. Dayaka, S. P. B. Vemulapalli and J. Bharatam, *Anal. Chem.*, 2015, **87**, 7258.
123. G. A. Morris, J. A. Aguilar, R. Evans, S. Haiber and M. Nilsson, *J. Am. Chem. Soc.*, 2010, **132**, 12770.
124. J. J. Koivisto, *Chem. Commun.*, 2013, **49**, 96.
125. V. M. R. Kakita and J. Bharatam, *Magn. Reson. Chem.*, 2014, **52**, 389.
126. N. H. Meyer and K. Zangger, *Chem. Commun.*, 2014, **50**, 1488.
127. T. Reinsperger and B. Luy, *J. Magn. Reson.*, 2014, **239**, 110.
128. L. Kaltschnee, A. Kolmer, I. Timari, V. Schmidts, R. W. Adams, M. Nilsson, K. E. Köver, G. A. Morris and C. M. Thiele, *Chem. Commun.*, 2014, **50**, 2512.
129. P. Sakhaei, B. Haase and W. Bermel, *J. Magn. Reson.*, 2009, **199**, 192.
130. I. Timári, L. Kaltschnee, A. Kolmer, R. W. Adams, M. Nilsson, C. M. Thiele, G. A. Morris and K. E. Kövér, *J. Magn. Reson.*, 2014, **239**, 130.
131. P. Kiraly, R. W. Adams, L. Paudel, M. Foroozandeh, J. A. Aguilar, I. Timári, M. J. Cliff, M. Nilsson, P. Sándor, G. Batta, J. P. Waltho, K. E. Kövér and G. A. Morris, *J. Biomol. NMR*, 2015, **62**, 43.
132. I. Timári, T. Z. Illyés, R. W. Adams, M. Nilsson, L. Szilágyi, G. A. Morris and K. E. Kövér, *Chem. – Eur. J.*, 2015, **21**, 3472.

133. L. Paudel, R. W. Adams, P. Király, J. A. Aguilar, M. Foroozandeh, M. J. Cliff, M. Nilsson, P. Sándor, J. P. Waltho and G. A. Morris, *Angew. Chem., Int. Ed.*, 2013, **52**, 11616.
134. L. Castañar, J. Saurí, P. Nolis, A. Virgili and T. Parella, *J. Magn. Reson.*, 2014, **238**, 63.
135. J. Saurí, E. Sistaré, R. T. Williamson, G. E. Martin and T. Parella, *J. Magn. Reson.*, 2015, **252**, 170.
136. I. Timári, L. Szilágyi and K. E. Kövér, *Chem. – Eur. J.*, 2015, **21**, 13939.
137. J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff, T. Parella, R. T. Williamson and G. E. Martin, *Angew. Chem. Int. Ed.*, 2015, **54**, 10160.
138. J. Saurí, T. Parella, R. T. Williamson and G. E. Martin, *Magn. Reson. Chem.*, 2015, DOI: 10.1002/mrc.4322.
139. J. Saurí, M. Frédérick, A. T. Tchinda, T. Parella, R. T. Williamson and G. E. Martin, *J. Nat. Prod.*, 2015, **25**, 2236.
140. C. Lorenc, J. Saurí, A. Moser, A. V. Buevich, A. J. Williams, R. T. Williamson, G. E. Martin and M. W. Peczu, *ChemistryOpen*, 2015, **4**, 577.
141. K. Takegoshi, K. Ogura and K. Hikichi, *J. Magn. Reson.*, 1989, **84**, 611.
142. J. A. Aguilar, M. Nilsson, G. Bodenhausen and G. A. Morris, *Chem. Commun.*, 2012, **48**, 811.
143. A. M. Torres, G. Zheng and W. S. Price, *Magn. Reson. Chem.*, 2010, **48**, 129.
144. D. Sinnaeve, *J. Magn. Reson.*, 2014, **245**, 24.
145. L. Castañar, J. Saurí, R. T. Williamson, A. Virgili and T. Parella, *Angew. Chem. Int. Ed.*, 2014, **53**, 8379.
146. R. W. Adams, C. M. Holroyd, J. A. Aguilar, M. Nilsson and G. A. Morris, *Chem. Commun.*, 2013, **49**, 358.
147. J. A. Aguilar, R. W. Adams, M. Nilsson and G. A. Morris, *J. Magn. Reson.*, 2014, **238**, 16.
148. B. Baishya and C. L. Khetrpal, *J. Magn. Reson.*, 2014, **242**, 143.
149. L. Castañar, E. Sistaré, A. Virgili, R. T. Williamson and T. Parella, *Magn. Reson. Chem.*, 2014, **53**, 115.
150. B. Baishya and A. Verma, *J. Magn. Reson.*, 2015, **252**, 41.
151. B. Baishya, C. L. Khetrpal and K. K. Dey, *J. Magn. Reson.*, 2013, **234**, 67.
152. K. E. Cano, M. J. Thrippleton, J. Keeler and A. J. Shaka, *J. Magn. Reson.*, 2004, **167**, 291.
153. P. W. A. Howe, *J. Magn. Reson.*, 2006, **179**, 217.
154. M. J. Thrippleton and J. Keeler, *Angew. Chem., Int. Ed.*, 2003, **42**, 3938.
155. A. M. Torres, R. Dela Cruz and W. S. Price, *J. Magn. Reson.*, 2008, **193**, 311.
156. B. Luy, *J. Magn. Reson.*, 2009, **201**, 18.
157. K. Kobzar and B. Luy, *J. Magn. Reson.*, 2007, **186**, 131.
158. J. Saurí, J. F. Espinosa and T. Parella, *Angew. Chem., Int. Ed.*, 2012, **51**, 3919.
159. T. Parella and J. F. Espinosa, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2013, **73**, 17.

160. R. T. Williamson, A. V. Buevich, G. E. Martin and T. Parella, *J. Org. Chem.*, 2014, **79**, 3887.
161. R. T. Williamson, A. V. Buevich and G. E. Martin, *Tetrahedron Lett.*, 2014, **55**, 3365.
162. K. A. Blinov, A. V. Buevich, R. T. Williamson and G. E. Martin, *Org. Biomol. Chem.*, 2014, **12**, 9505.
163. B. Reif, M. Köck, R. Kerssebaum, H. Kang, W. Fenical and C. Griesinger, *J. Magn. Reson., Ser. A*, 1996, **118**, 282.
164. M. Köck, B. Reif, W. Fenical and C. Griesinger, *Tetrahedron Lett.*, 1996, **37**, 363.
165. M. Köck, R. Kerssebaum and W. Bermel, *Magn. Reson. Chem.*, 2003, **41**, 65.
166. G. E. Martin, *Annu. Rep. NMR Spectrosc.*, 2011, **74**, 215.
167. G. E. Martin, M. Reibarkh, A. V. Buevich, K. A. Blinov and R. T. Williamson, *eMagRes*, 2007, **3**, 215–234.
168. A. V. Buevich, R. T. Williamson and G. E. Martin, *J. Nat. Prod.*, 2014, **77**, 1942.
169. M. M. Senior, R. T. Williamson and G. E. Martin, *J. Nat. Prod.*, 2013, **76**, 2088.
170. G. E. Martin, R. T. Williamson, P. G. Dormer and W. Bermel, *Magn. Reson. Chem.*, 2012, **50**, 563.
171. M. Reibarkh, R. T. Williamson, G. E. Martin and W. Bermel, *J. Magn. Reson.*, 2013, **236**, 126.
172. G. E. Martin, B. D. Hilton, M. R. Willcott and K. A. Blinov, *Magn. Reson. Chem.*, 2011, **49**, 350.
173. G. E. Martin, B. D. Hilton and K. A. Blinov, *J. Nat. Prod.*, 2011, **74**, 2400.
174. G. E. Martin, B. D. Hilton, M. Robert Willcott and K. A. Blinov, *Magn. Reson. Chem.*, 2011, **49**, 641.
175. G. E. Martin, K. A. Blinov, M. Reibarkh and R. T. Williamson, *Magn. Reson. Chem.*, 2012, **50**, 722.
176. S. F. Cheatham, M. Kline, R. R. Sasaki, K. A. Blinov, M. E. Elyashberg and S. G. Molodtsov, *Magn. Reson. Chem.*, 2010, **48**, 571.
177. G. E. Martin, R. T. Williamson, K. A. Blinov, C. G. Anklin and W. Bermel, *Magn. Reson. Chem.*, 2012, **50**, 691.
178. G. E. Martin, K. A. Blinov and R. T. Williamson, *Magn. Reson. Chem.*, 2013, **51**, 299.
179. S. Cheatham, P. Gierth, W. Bermel and Ě. Kupče, *J. Magn. Reson.*, 2014, **247**, 38.
180. S. Cheatham, M. Kline and Ě. Kupče, *Magn. Reson. Chem.*, 2015, **53**, 363.
181. T. Öman, M.-B. Tessem, T. F. Bathen, H. Bertilsson, A. Angelsen, M. Hedenström and T. Andreassen, *BMC Bioinform.*, 2014, **15**, 413.
182. K. Bingol, D.-W. Li, L. Bruschweiler-Li, O. A. Cabrera, T. Megraw, F. Zhang and R. Bruschweiler, *ACS Chem. Biol.*, 2015, **10**, 452.
183. K. Bingol, L. Bruschweiler-Li, D.-W. Li and R. Bruschweiler, *Anal. Chem.*, 2014, **86**, 5494.

Part 2

CHAPTER 7

Terpenes: Mono-, Sesqui-, and Higher Terpenes

WILLIAM F. REYNOLDS*^a AND RAUL G. ENRIQUEZ^b

^a Department of Chemistry, University of Toronto, Toronto ON, M5S 3H6, Canada; ^b Instituto de Quimica, Universidad Nacional Autonoma de Mexico, Mexico D.F. 04100, Mexico

*Email: wreynold@chem.utoronto.ca

7.1 Introduction

Terpenes are naturally occurring organic compounds containing some multiple of five carbons.¹ They are composed of two or more isoprene units (1), most often attached in a head-to-tail fashion (where the head is defined as the branched end of the isoprene unit). The biogenetic isoprene rule postulates that terpenes containing one or more rings can be regarded as biosynthesized from the corresponding acyclic terpene *via* cyclization reactions.² This assumption provides a convenient method for checking the probability of any proposed terpene structure. Terpenes contain from two up to twelve or more isoprene units, but the most common have three (sesquiterpenes), four (diterpenes), or six (triterpenes) isoprene units. The development of 2D nuclear magnetic resonance (NMR) methods³ has greatly aided the determination of terpene structures and there has been a dramatic increase in the number of new terpenes reported in the literature over the past 25 years. For example, review articles in 2010 showed that 452 sesquiterpene structures had been reported in the previous year,⁴ and 978 triterpene structures had been reported in the past 2 years,⁵ while the total

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

number of known diterpene structures of one structural class (labdane diterpenes) was close to 7000.⁶

This rapid proliferation of known structures poses two distinct problems for a natural product chemist using NMR. First, in view of the large number of known terpenes, it is critical to have methods to quickly identify whether an isolated terpene is a new or known compound. This *dereplication* approach is described in detail in Volume 1, Chapter 8. Second, it is becoming increasingly rare to isolate an unknown compound at significant yield, and thus one is often faced with the task of determining both the skeletal structure and stereochemistry of a compound while working with as little as 1 mg or less of sample. This is often further complicated by serious overlap of peaks in the proton and/or carbon spectra in many terpenes. The development of cryogenically cooled probes and particularly of cold microprobes (see Volume 1, Chapters 3 and 4) has significantly lowered the amount of sample that is required to get a set of adequate spectra for structure elucidation in reasonable time.^{7,8} However, most natural product laboratories do not have access to cryogenically cooled microprobes.

In this review, we first list the various types of terpenes and the classes of compounds derived from terpenes, and provide representative examples of each type. Second, we discuss the basic approaches for dereplicating known structures and determining the structures of unknown compounds. Next, we discuss precautions one can take to avoid the embarrassment of reporting an incorrect structure, a problem that still occurs surprisingly often.⁹ This includes a discussion of the pitfalls in interpreting 2D spectra, including how much concern there should be about missing expected correlation peaks and the observation of unexpected peaks. We also point out the risks of misinterpretation associated with deceptively simple spectra. Finally, we describe several specialized approaches that are helpful in overcoming problems in structure determination, particularly in the case of severe spectral crowding. The examples chosen to illustrate the various points will draw heavily on our own research in this area, involving the investigation of several hundred terpenes and related compounds over the past 30 years.

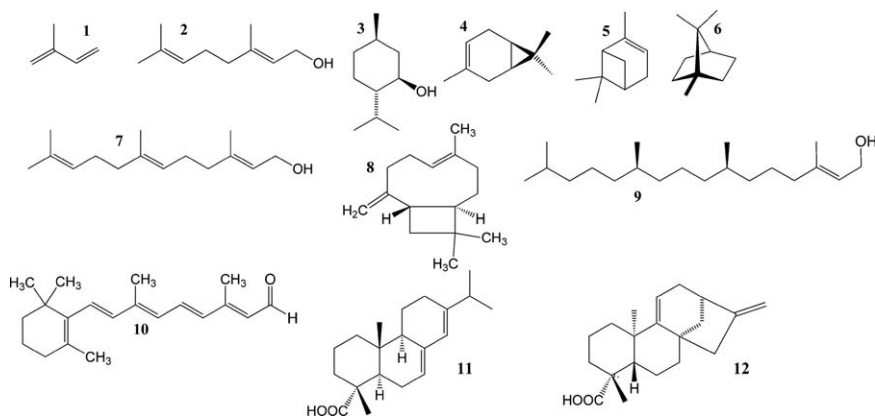
7.2 Types of Terpenes and Related Compounds¹⁰

Monoterpenes (often just called terpenes) are composed of two isoprene units attached in a head-to-tail fashion. Many of these have pleasant odors and are often referred to as essential oils. Acyclic and monocyclic terpenes are the most common. Geraniol (rose oil) (2) is a representative acyclic terpene while menthol (3) is typical of monocyclic terpenes. Possible bicyclic terpene ring structures include [4.1.0], [3.1.1], and [2.2.1] bicycloheptane structures, with Δ^3 -carene (4) α -pinene (5), and bornane (6) being respective examples of the three types of structures.

Sesquiterpenes are 15-carbon compounds composed of three head-to-tail isoprene units. They can be considered to be derived from farnesol (7) sometimes as acyclic compounds, but usually by cyclization to form a very

wide variety of monocyclic, bicyclic, and tricyclic compounds. As a result of this structural diversity, sesquiterpenes are significantly more challenging structural problems than monoterpenes. One example of a bicyclic sesquiterpene is β -caryophyllene (**8**), which is isolated from cloves.

Diterpenes have four isoprene units in a head-to-tail arrangement. Phytol (**9**) can be regarded as the parent acyclic diterpene. Many monocyclic diterpenes have important biological functions (*e.g.* retinal, **10**, which is important in vision, and vitamin A, in which the aldehyde of retinal is reduced to a primary alcohol). While bicyclic diterpenes are known, it is our impression that the majority of diterpenes from plant sources have tricyclic or tetracyclic structures. Two examples of tricyclic and tetracyclic diterpenes are, respectively, abietic acid (**11**) and kauradienoic acid (**12**). Many other diterpenes share the A/B ring structure shown in structures **11** and **12**, but with differences in the upper right corner of the molecular structure. Figure 7.1 illustrates that the skeletal structures of compounds **3**, **6**, **8**, and **12** are all consistent with the biogenetic terpene rule since their structures, respectively, correspond to two, two, three, and four isoprene units joined head-to-tail, followed by cyclization.



Sesterpenes have five isoprene units bonded head-to-tail (*i.e.* a total of 25 carbons) and are relatively uncommon in nature. The parent sesterpene, geranylarnesol, can be regarded as a combination of geraniol (**2**) and farnesol (**7**).

The parent triterpene can be regarded as squalene (**13**). It consists of two farnesol units joined tail-to-tail. There are many different triterpenes, with tetracyclic and pentacyclic triterpenes being particularly common. An example of the former is lanosterol (**14**), and of the latter is oleanic acid (**15**). Steroids (which are discussed in detail in Volume 2, Chapter 9) are biogenetically derived from triterpenes. The parent steroid can be regarded as cholesterol (**16**), which is derived by cleavage of three methyl groups from lanosterol.

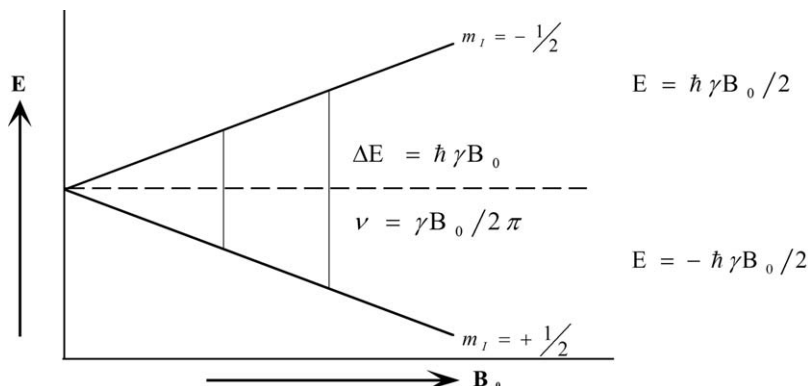


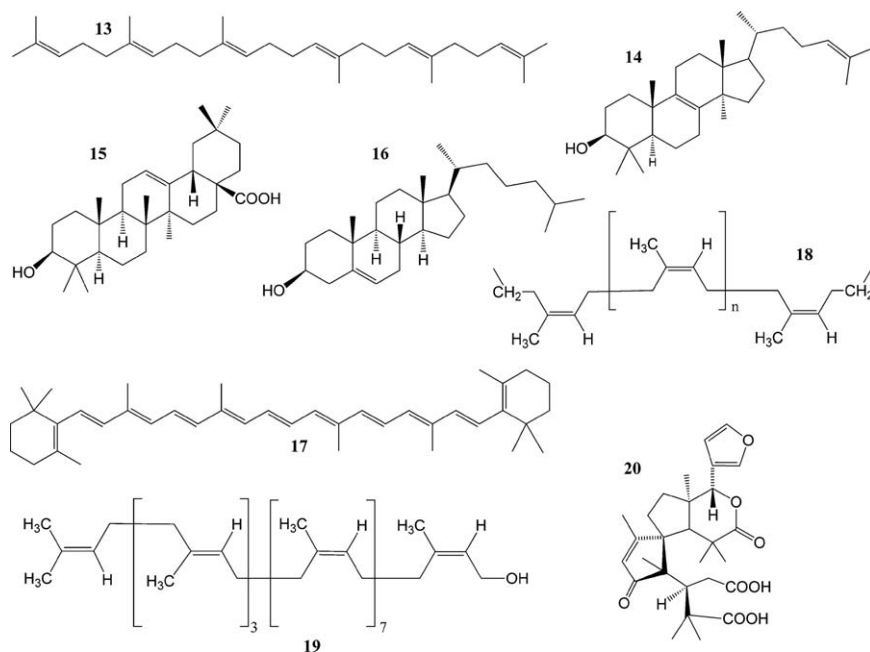
Figure 7.1 Illustration that the carbon skeletal structures of (A) menthol (3), (B) bornane (6), (C) β -caryophyllene (8), and (D) kauradienoic acid (12) are consistent with the biogenetic isoprene rule. The heavy lines correspond to isoprene units, the light lines to bonds in the corresponding acyclic terpene, and the dashed lines to bonds formed *via* cyclization.

Tetraterpenes have 40 carbons and have skeletons derived from two phytol units attached tail-to-tail. Many common tetraterpenes have one or two terminal six-membered rings with a linear central segment composed of conjugated double bonds. Thus, they are often strongly colored. A typical example is β -carotene (17). The primary spectral problem is in the assignment of double bond carbons and their stereochemistries (usually *E*).

Polyterpenes consist of a linear chain of 10 or more isoprene units. The most important polyterpene is natural rubber, which consists of a large number of polymerized *Z*-isoprenoid units (18). Other than rubber, the most common natural polyterpenes are a family of compounds called polyprenols. They commonly have a head isoprene unit, followed by three *E*-isoprene units and then an even number of *Z*-isoprene units with a terminal hydroxyl group (e.g. polyprenol-12; 19). Polyprenols show particularly severe spectral crowding.

Saponins are terpenes (most commonly triterpenes) with attached monosaccharide or polysaccharide units. The latter are attached to the terpene unit (called the aglycone) *via* linkage of the anomeric hydroxyl of a sugar unit to either a hydroxyl group (ether linkage) or a carboxylic acid group (ester linkage) in the aglycone unit. Many of these contain aglycones that are derivatives of oleanic acid (15). Most of the structural/spectral assignment problems of saponins are associated with the polysaccharide components, including identification of individual monosaccharide units and determination of the linkage sites between different units.

Tetranortriterpenoids are compounds where there has been substantial rearrangement and degradation of a triterpene skeleton. A typical example is guyanin (20), which we believe was the first natural product with a totally unprecedented skeleton that was elucidated by 2D NMR methods.¹¹ The main problem with investigating tetranortriterpenoids is the great variety of possible structures.



7.3 Basic Approach to Terpene Structure Elucidation by NMR Spectroscopy

A good discussion of the use of NMR for organic structure elucidation and a discussion of the different pulse sequences that can be used for this purpose is provided by Lambert and Mazzola,¹² while Reynolds and Enriquez¹³ provide a review that discusses the relative merits of different pulse sequences and processing strategies. For that reason, specific pulse sequences will not be described in detail here or cited. Rather, we discuss how they are applied to structure determination and how the data they provide are interpreted.

Unless one is so sample limited that it is impossible to easily acquire a ^{13}C spectrum, the logical initial approach is to acquire a ^1H and a ^{13}C spectrum, followed by either a DEPT spectrum or an edited HSQC spectrum. The latter provides the same information as a DEPT spectrum in less time, while providing the additional bonus of correlating ^1H and ^{13}C signals.¹³ In extremely sample-limited cases, we suggest acquiring a proton spectrum plus an edited HSQC spectrum, preferably with an inverse detection probe to take advantage of its higher proton sensitivity. In either case, but particularly in the latter situation, it is also important to obtain a high-resolution mass spectrum of the parent ion of the compound in order to determine the empirical formula. These data can be then entered into one of the existing spectral databases¹⁴ to determine whether one is dealing with a known

compound. However, the problem is that these databases do not include all known compounds, which is hardly surprising in view of the vast numbers of these compounds. Also, many compounds discovered in the past do not have completely reported ^{13}C spectra. Thus, we encourage investigators working in the natural product area to determine and assign spectra of known compounds where data are lacking. Unfortunately, unless one has completed assignments of a significant number of structurally related known compounds, it is now difficult to persuade journal editors to publish these types of experimental data papers. Thus, there has been a clear need for a procedure where reliable data of this kind can be made widely available. Fortunately, the ChempSpider website,¹⁵ which was recently established by the Royal Society of Chemistry, nicely fills this requirement. This same website also provides other spectroscopic and physical data for known compounds that can be helpful in the dereplication process.

Following dereplication, if the structure of the compound turns out to be unknown or uncertain, the next logical step is to acquire a set of 2D spectra. The first step is to determine the skeletal structure of the compound. The basic approach, which has been applied in numerous cases, is to acquire a ^1H - ^1H COSY spectrum and one-bond and n -bond ($n = 2, 3$) ^1H - ^{13}C correlation spectra.¹⁶ The first two spectra assign molecular fragments involving connected protonated carbons, while the third ties the skeletal structure together (see Figure 7.2). The one-bond correlation spectrum is best obtained with the HSQC sequence since the alternative HMQC spectrum has inferior ^{13}C resolution (see Reynolds and Enriquez¹³ for a discussion of the relative merits of different alternative sequences for structure determination), while the n -bond spectrum is usually obtained with the HMBC sequence. Both sequences normally run with gradient selection to suppress ^1H signals bonded to ^{12}C . In many cases, the correlation data obtained from these spectra will be sufficient to determine the skeletal structure of an unknown compound. An example showing how the correlation data can be used to determine the skeletal structure of a terpene is shown in Figure 7.3 for kauradienoic acid, the first compound that we investigated using this approach.¹⁶ As can be seen, the molecule is assembled in a manner similar to putting together the pieces of a jigsaw puzzle. The safest way to interpret correlation data and convert it to the most probable structure is discussed in the following section, while additional techniques, which can be used when the data obtained in this way are insufficient to determine the structure, are discussed in Section 7.5.

Unless one is dealing with a planar structure, an investigator is next faced with the often difficult task of determining the 3D skeletal structure of the molecule plus the stereochemistry of any substituent groups. Usually, a combination of approaches must be used. The two most commonly employed are the use of vicinal proton-proton coupling constants and the use of nuclear Overhauser effect (nOe) data, the latter often being obtained with 2D NOESY or ROESY pulse sequences. The two-bond and three-bond ^1H - ^{13}C couplings can also provide helpful stereochemical information, although

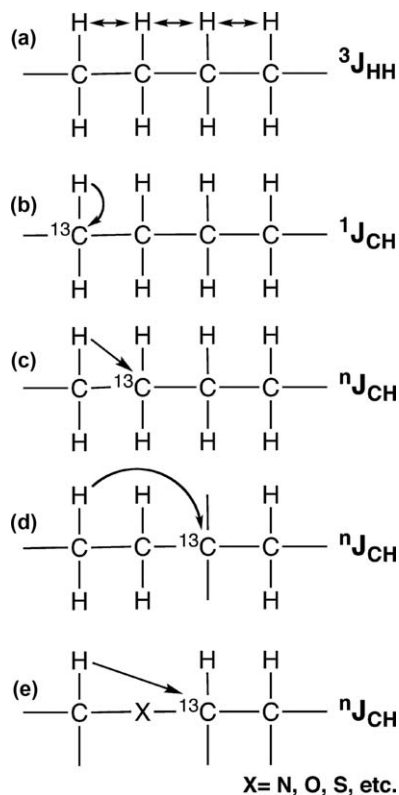


Figure 7.2 Illustration of the use of 2D NMR correlation data for determining skeletal structures of organic compounds. (A) The use of COSY spectra to establish ^1H - ^1H connectivities in a fragment of protonated carbons. (B) The use of one-bond ^{13}C - ^1H correlation spectra to assign the attached carbons in the same fragment. (C) The use of two-bond ^{13}C - ^1H correlations to determine the attachment of fragments of protonated carbons to quaternary carbons. (D) The use of three-bond ^{13}C - ^1H connectivities to provide the same information as in (C). As discussed in the text, one problem is distinguishing whether the observed correlations are due to two-bond or three-bond connectivities. (E) The use of three-bond ^{13}C - ^1H connectivities to determine the connection between two protonated carbon fragments separated by a heteroatom.

their accurate determination is a non-trivial task.¹⁷ One helpful characteristic of terpenes is that monoterpenes often contain a cyclohexane ring in a chair conformation, while diterpenes and triterpenes often have two or more fused cyclohexane rings, again in chair conformations. However, one must always be aware that the presence of substituents may force a cyclohexane ring of a diterpene or triterpene into a non-chair conformation. For that reason, it is useful to combine the NMR data with calculations of energetically preferred conformations. While this is still most commonly done using empirical molecular mechanics programs,¹⁸ the rapid increase in

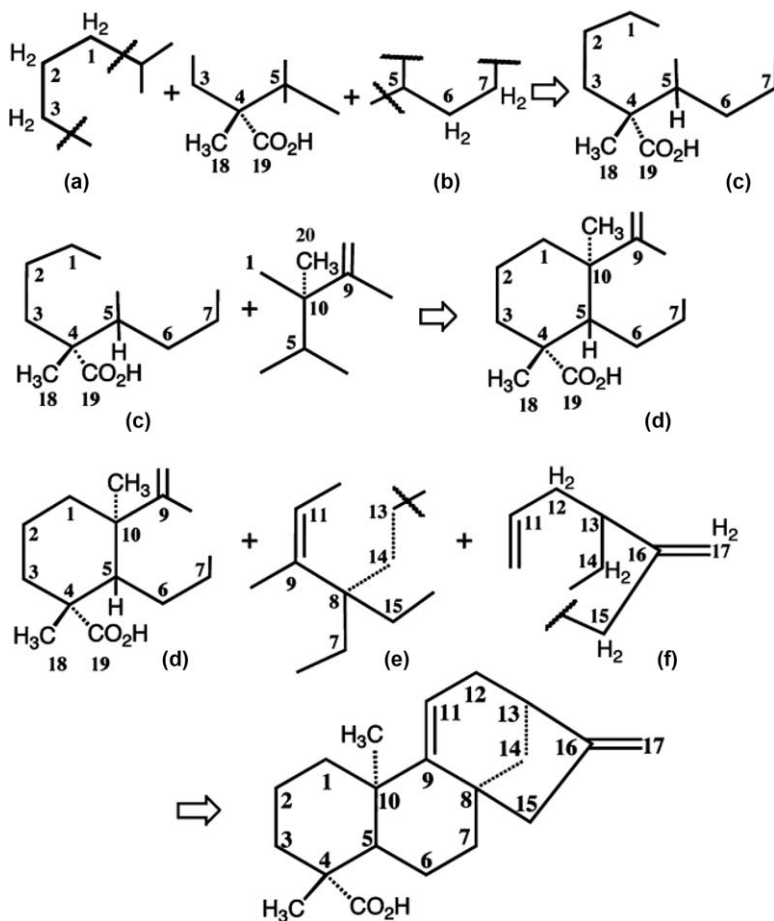


Figure 7.3 Illustration of the use of connectivity information of the kind outlined in Figure 7.2 to determine the skeletal structure of kauradienoic acid (12). Protonated carbon fragments **a** and **b** are connected *via* quaternary carbon C(4) by long-range cross-peaks between the C(18) methyl singlet and C(3), C(4), and C(5). The resultant part structure **c** is further elaborated by long-range correlations from the C(20) methyl singlet, which ties C(1) and C(5) *via* C(10), yielding part structure **d**. Long-range correlations of several protons to quaternary carbons C(8) and C(9) elucidate part structure **e**, while COSY and one-bond ^{13}C - ^1H correlations provide the last fragment **f**. Noting carbons common to these fragments, part structures **d**, **e**, and **f** are then combined to yield the skeletal structure of **12**.

computer speed and memory makes quantum mechanical calculations a viable option.¹⁹

An alternative approach, which is applicable to the specific case of secondary hydroxyl substituents, is to derivatize the hydroxyl group with a chiral reagent such as Mosher's reagent.²⁰ While the rules for predicting

stereochemistry from derivatization shifts are empirical, they also appear to be highly reliable. Finally, one can partially orient the molecule in one of a number of orienting media and use the resultant residual dipolar ^1H - ^{13}C couplings (usually determined by comparing the couplings measured from coupled HSQC spectra in orienting and isotropic media, and described in detail in Volume 2, Chapter 4) as input to predict the 3D structure of the molecule.²¹ However, this is again a non-trivial procedure and it may sometimes be difficult to recover the compound from the orienting medium for further study.

7.4 Avoiding Determining the Wrong Structure

In spite of the availability of increasingly sophisticated spectroscopic methods for organic structure determination, there still has been a surprisingly large number of incorrect structures reported.⁹ While many of these have involved errors in stereochemistry, others have involved errors in skeletal structures. There are a number of precautions that one can take to at least minimize this risk. First, one must be cautious in assigning a carbon (or proton) to a particular structural feature solely on the basis of its chemical shift. For example, Figure 7.4 illustrates two different compounds, each of which has a non-protonated carbon in a region (near δ 80 ppm) that is often associated with a hydroxyl-bearing quaternary carbon. However, this is not true in either case, but rather is due to a combination of factors that move the chemical shifts for the carbons from their expected regions to this one.

A second, general precaution is to avoid trying to force the data to fit a preconceived notion of a structure (*e.g.* by assuming that a noise spike or other small artifact is an expected correlation). Instead, one should systematically examine the correlation data to find structure(s) that are entirely

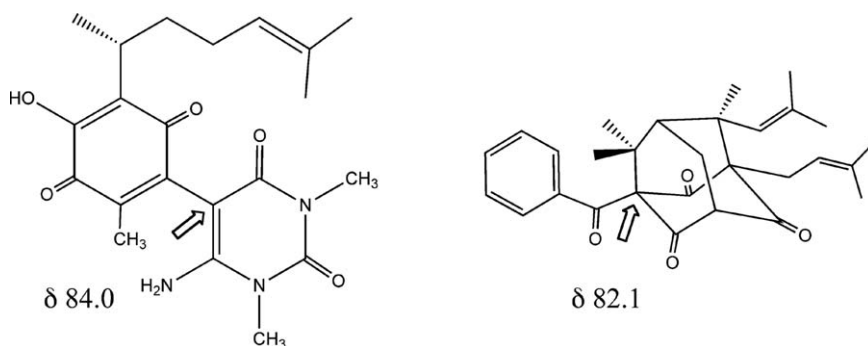


Figure 7.4 Example of two molecules, each with a carbon near δ 80 ppm. While one might assume that these carbons are hydroxyl-bearing quaternary carbons, this is not true in either case, illustrating the need for caution in making assignments solely based on chemical shifts. The compound on the left is from our own unpublished research, while the compound on the right was supplied by Professor Helen Jacobs.

consistent with the data. An understandable trap to avoid here is to assume that the first structure that you find that fits the data must be the correct one. Since it is often difficult to determine whether HMBC correlations to quaternary carbons in particular are two-bond or three-bond correlations (or even unexpectedly large four-bond correlations), there may be more than one structure that is consistent with the available data. Here, the best approach is to use one of the available computer programs that determine all possible structures that are consistent with the available correlation data.²² These programs will also predict ^{13}C and ^1H chemical shifts for the different structures, and this will often allow one to determine the most probable structure. However, one should be aware that these predictions are based on datasets for known compounds with assigned spectral data. If one has an unusual compound with structural features that are not well represented in the database, the predictions may not be accurate enough to enable a reliable choice to be made between different structures. An alternative approach is to use one of the increasingly sophisticated *ab initio* quantum mechanical programs to predict the shifts for different structures.¹⁹ As computer speed increases, this is becoming an increasingly viable approach, and we suspect that a hybrid approach combining empirical data with quantum mechanical calculations will eventually prove to be the most reliable method.

However, even with these sophisticated computer-based methods, there is still an important role for the spectroscopist in evaluating the data, one that can be a symbiotic relationship with the software tools. Here there are two key points to consider. First, how important is it if one or more expected correlation peaks is not observed, particularly in HMBC spectra? Here it is important to have knowledge of the expected sizes of two-bond and three-bond ^{13}C – ^1H couplings.²³ If the coupling has a near-zero value, no peak may be observed, even with excellent signal/noise. Alternatively, if the coupling is small compared to the value used to optimize the delay, the expected correlation peak may not be observed if the signal/noise is marginal. A common example of the former situation is the two-bond ^1H – ^{13}C correlation for olefinic or aromatic groups, while an example of the latter problem is the three-bond correlation for an axial proton in a cyclohexane-like chair. The latter problem is particularly likely to occur if the proton has a complex multiplet structure, although this problem can be minimized by the appropriate choice of acquisition and processing parameters.²⁴ Thus, under these circumstances, the absence of an expected correlation is not a reason to reject a structure. However, if the coupling that should give rise to a correlation is comparable in magnitude to the value chosen to calculate the delay (typically 8 Hz), this should be a very serious concern.

The alternative situation is when a strong correlation is observed that would not be expected for a particular structure. Unless there is good evidence that this is due to some type of instrumental artifact, this is usually a reason to reject a proposed structure. However, in the case of a HMBC spectrum, one should first check to make sure the peak is not actually one

component of an incompletely canceled one-bond ^{13}C - ^1H doublet. Even though HMBC sequences usually contain some form of “ J filter” to suppress one-bond peaks, these will still often appear, particularly for strong proton signals such as methyl signals. These occur at the chemical shift of the carbon in F_1 and at $\nu_{\text{H}} \pm J_{\text{CH}/2}$ in F_2 . If another proton has a chemical shift that is the same as the frequency of one of the doublet peaks, this could be mistaken as a correlation between the former proton and the carbon causing the doublet. The best way to avoid this mistake is to visually inspect the contour plot for the spectrum, watching for a doublet centered on the chemical shift of a proton with a splitting whose magnitude is consistent with the expected value for $^1J_{\text{CH}}$. It should also be at the F_1 frequency of the corresponding attached carbon. On the other hand, a singlet at the chemical shifts of a directly bonded CH pair is actually a two-bond or three-bond peak between two CH groups that are equivalent due to symmetry. This is a useful clue that a molecule may have some type of symmetric structure, perhaps a dimeric structure.

A common source of error is severe overlap of two or more (usually proton) signals. Under these circumstances, the correlation data could be misinterpreted, leading to a wrong structure. One recent example, which generated considerable controversy, is hexacyclinol, where the structure was originally incorrectly assigned as structure **a**²⁵ in Figure 7.5 rather than structure **b** in the same figure, which was later shown to be correct.²⁶ The original mistake was very likely due to misinterpretation of the HMBC data in a severely crowded ^1H spectral region. This was possibly also due to an attempt to fit the correlation data to a preconceived notion of the structure, as mentioned at the beginning of this section. Again, this is an area where the use of a computer program to determine all possible structures that fit the data would enable the scientist to recognize the existence of alternatives, alerting him or her to possible problems.²² However, since overlap provides a potential risk of error, it is important, whenever possible, to take steps to minimize this risk. This is discussed in Section 7.5. Finally, one should always check whether a proposed terpene skeletal structure is consistent with

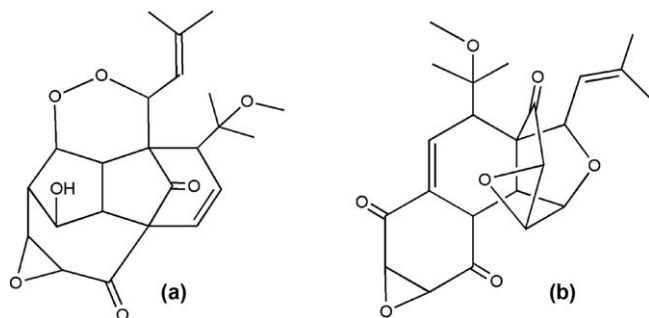


Figure 7.5 Illustration of two different structures of the same compound that have been reported in the literature. The incorrect structure **a** is from Schlegel *et al.*,²⁵ while the correct structure **b** is from Porco *et al.*²⁶

the biogenetic isoprene rule (see Figure 7.1). If not, one should carefully re-examine their analysis of the data for possible errors. While exceptions to this rule are known, they are relatively rare.

A related problem associated with spectral overlap is the appearance of deceptively simple spectra. While everyone using NMR is trained at an early stage that protons that are equivalent by symmetry will not show coupling to each other, one must also be aware of the possibility of accidental near or complete overlap of coupled protons in the absence of symmetry. Even with the availability of higher-field spectrometers, it is our experience that this can still be surprisingly common and easily misleading. An example is given in Figure 7.6, which shows the deceptive 500 MHz ^1H spectrum of a steroid containing a $[-\text{CH}(\text{OH})\text{CH}_2\text{OH}]$ sidechain that would normally be expected to give an ABC spectrum (since the methylene protons should be diastereotopic). Instead, it appears as a sharp singlet, integrating to three protons, and with a chemical shift consistent with an OCH_3 group. However, an edited HSQC spectrum showed that the apparent singlet was actually due to total overlap of the CH and CH_2 protons. A similar problem observed in another steroid was the appearance of a singlet for a side chain CH_3 group, which should normally be expected to be a doublet due to coupling with an adjacent CH group.²⁷ However, an explanation was provided by a HMBC spectrum, which showed that the methyl protons were coupled to a methine carbon whose proton had an identical chemical shift to the methyl protons.

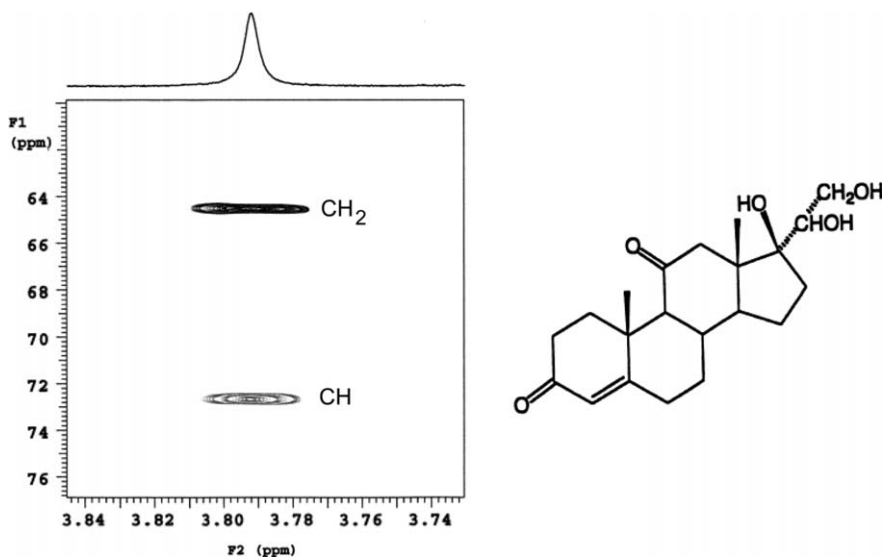


Figure 7.6 Illustration of the accidental equivalence of the three protons of the $-\text{CH}(\text{OH})-\text{CH}_2\text{OH}$ fragment of the steroid illustrated in the figure. The appearance of a sharp three-proton singlet at δ 3.79 ppm could easily be mistaken for an OCH_3 group. However, the edited HSQC spectrum shown in the figure confirmed that it was due to overlapping CH and CH_2 protons. The compound was supplied by Professor Paul Reese.

Accidental overlap can also lead to incorrect stereochemical assignments, particularly if one is relying on vicinal proton–proton couplings for this purpose. If two protons are strongly coupled to one another, they will appear to be equally coupled to a third proton, even when the couplings are actually different (“virtual coupling”).²⁸ This is illustrated in Figure 7.7 for the X resonances of a series of simulated ABX spectra as the AB chemical shift difference approaches zero. Even though the actual AX and BX couplings used for the simulation are 10 Hz and 2 Hz (consistent with the X proton being axial in a cyclohexane ring), they both appear as 6 Hz splittings in the AA' limit. This could be mistaken as indicating an equatorial proton.

Solvent-induced chemical shifts (see next section) can sometimes be used to remove accidental degeneracy.

Another potential problem occurs when a molecule undergoes some type of dynamic process, most often an exchange between different conformations or

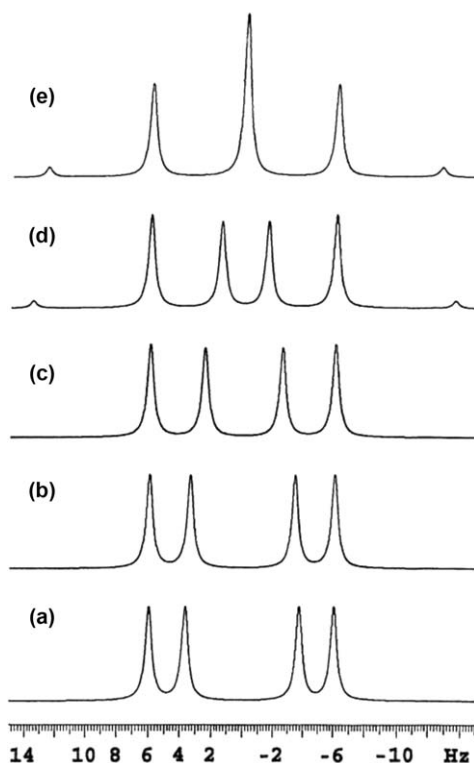
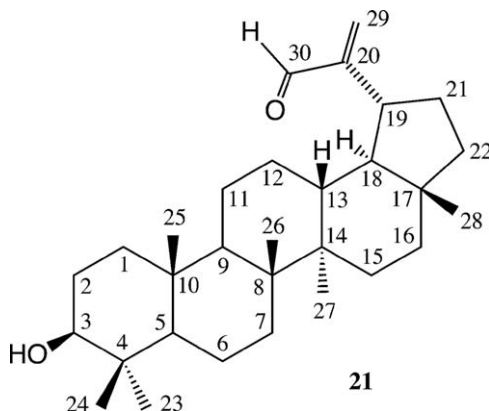


Figure 7.7 Simulated X resonances for a series of ABX spectra, each with $J_{AB} = -12.0$ Hz, $J_{AX} = +10.0$ Hz, and $J_{BX} = +2.0$ Hz. The AB chemical shift differences are (A) +30 Hz (B) +20 Hz (C) +10 Hz (D) +5 Hz, and (E) 0 Hz. The splittings in the doublet of doublet X spectra are (A) 9.7, 2.3 Hz (B) 9.4, 2.6 Hz (C) 8.5, 3.5 Hz (D) 7.5, 4.5 Hz, and (E) 6.0, 6.0 Hz, illustrating the increased discrepancy between splittings and couplings as the AB chemical shift difference decreases.

different tautomers. If the exchange rate is rapid on the so-called NMR time scale (a term we dislike because there are many different NMR time scales), one will observe what appears to be a normal spectrum, and may not be aware of the dynamic process. In the very slow exchange limit, one will observe two (or more) sets of spectra. While this complicates the assignment process, a NOESY or ROESY spectrum often shows off-diagonal EXSY peaks (which are of opposite phase to nOe peaks) between pairs of exchanging signals. These can be very helpful in making assignments.¹³ However, if the exchange rate is in the intermediate range, the appearance of the ^{13}C spectrum in particular can be quite misleading. If the chemical shift difference between a pair of exchanging carbons is relatively large, then the resultant ^{13}C peak may be so broad that it can easily be lost in the noise and thus missed. An unusual feature of this problem is that if sensitivity is not a limiting factor, there is actually an advantage in acquiring a ^{13}C spectrum on a lower-field spectrometer, since the smaller frequency differences will make it easier to approach the fast exchange limit for different carbons at a reasonable temperature.

An example of this type of problem was provided by a derivative of the common triterpene, lupeol, in which the methyl group of the propene side chain had been converted to an aldehyde (**21**). The initial ^{13}C spectrum appeared to show 24 or possibly 25 carbons, which suggested that the compound might be a sesterpene or a steroid.²⁹ However, the appearance of the ^1H spectrum, including the presence of six methyl singlets, seemed more consistent with a triterpene or a tetranortriterpenoid. By reacquiring the ^{13}C spectrum at 35 °C using a more sensitive 3 mm probe, we were able to observe five additional broadened signals (from 10 to 40 Hz wide), while the presence of a sixth very broad (~100 Hz) signal was confirmed by observing an HSQC correlation to a sharp ^1H peak. Low-temperature (−40 °C) spectra indicated the presence of two conformations of the side chain group that were slowly exchanging by hindered rotation about the C(19)–C(20) bond. Consistent with this conclusion, the broadened peaks in the room temperature ^{13}C spectrum were all due to carbons that were spatially close to the side chain.²⁹



7.5 Specialized Techniques that are Useful for Different Types of Terpenes

While the basic approach outlined in Section 7.2 will work well in many cases, this is not always true. Severe spectral crowding of ^1H and/or ^{13}C spectra is often the biggest problem. The most obvious solution is to rerun spectra on a higher-field spectrometer. However, this may either not be available or may still be insufficient to overcome the overlap problem. However, it is our experience that both the resolution and sensitivity of 2D spectra (especially HSQC and HMBC) can be significantly improved by correct choices of acquisition parameters.^{13,24} In particular, the attempt to save time by limiting the number of acquisition points and time-incremented spectra may actually waste time by providing spectra with insufficient digital resolution. In addition, the well-established post-acquisition technique of forward linear prediction³⁰ can significantly improve F_1 resolution at no cost in experiment time.^{13,31} Contrary to widely expressed views that only two-fold linear prediction should be used, we have found that four-fold linear prediction can usually be reliably carried out, even with marginal signal/noise,³¹ and, in one case, even 16-fold prediction.³²

Alternatively, provided that overlap only occurs in specific ^1H spectral regions, this can sometimes be overcome with the old but still valuable approach of aromatic solvent-induced shifts (ASIS).³³ An approach that we have found useful is what could be called “benzene titration”. For example, if one has originally obtained a set of different NMR spectra in an organic solvent such as CDCl_3 , one then adds a small increment of C_6D_6 to the sample and then repeats the ^1H spectrum (note that since C_6D_6 is significantly less dense than CDCl_3 , it is necessary to ensure that the solution is well mixed). Hopefully, the overlapped protons will be differently affected by the added C_6D_6 and thus will no longer overlap. Alternatively, if the resolution is improved but there is still partial overlap, then one or more increments of C_6D_6 can be added. Since ASIS of ^{13}C signals are generally relatively smaller than those of ^1H signals, having HSQC spectra for the original and titrated samples will usually allow unambiguous ^1H assignments. An example illustrating the value of this approach is shown in Figure 7.8. The addition of a small amount of C_6D_6 is sufficient to completely reveal otherwise overlapped ^1H signals, allowing stereochemistry assignments.

Another old but still useful technique is to use ^{13}C satellites in proton spectra in order to determine vicinal coupling constants between protons that are equivalent either by symmetry or accidentally.³⁴ A later modification incorporated a BIRD pulse in order to eliminate the peak due to protons bonded to ^{12}C .³⁵ An alternative approach that is particularly useful in the case of complex spectra is to use a high-resolution ^{13}C -coupled spectrum for this purpose.^{36,37} This has the additional advantage of providing the magnitude of the one-bond ^{13}C - ^1H couplings, which provides useful information, particularly concerning the stereochemistry of anomeric protons in monosaccharide units. An example of this technique is illustrated in

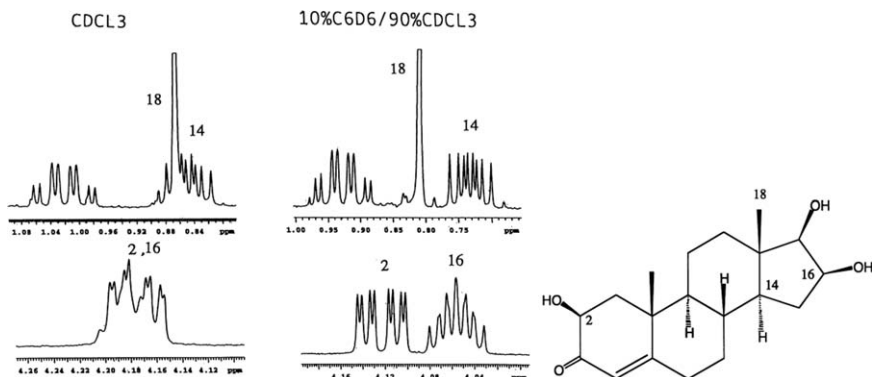


Figure 7.8 Illustration of the use of “benzene titration” to resolve overlapping peaks in the spectrum of the steroid illustrated at the figure. The sample was supplied by Professor Paul Reese.

Figure 7.9. Here, four strongly overlapped protons in a monosaccharide unit can be resolved (along with two almost overlapped carbon peaks), and the magnitudes of vicinal ^1H – ^1H coupling constants are sufficiently well determined (~ 8 – 9 Hz) to identify the monosaccharide as glucose.

There are also assignment techniques that are particularly suitable for specific classes of terpenes. For example, triterpenes typically contain up to eight methyl singlets. These not only produce particularly strong two-bond and three-bond ^1H – ^{13}C correlations, but also provide a network of connectivities that effectively constitutes a “molecular zipper”, tying much of the skeletal structure together (*e.g.* see Figure 7.10). This allowed the authors a number of years ago to provide the first unambiguous, complete ^1H and ^{13}C assignments for several triterpenes,³⁸ as well as to choose the correct one of three alternative structures that had been proposed for the same triterpene.³⁹ In the latter case, benzene titration was needed to resolve overlapping CH_3 proton singlets.

^1H -detected ^1H – ^{13}C shift-correlated sequences (HSQC, HMQC, HMBC, *etc.*) have generally replaced the older corresponding ^{13}C -detected sequences (HETCOR, COLOC, FLOCK, *etc.*).¹² However, one should remember that, although the latter sequences are significantly less sensitive, they can also give better resolution in the case of severe spectral crowding. The reason for this is that it is easier to obtain good resolution in a reasonable time for the detection axis F_2 than for the time-incremented axis F_1 . Since the ^{13}C spectral window is much larger than the ^1H spectral window, one can achieve a high resolution along both axes with ^{13}C -detected sequences. However, it is difficult with ^1H -detected sequences to obtain very high ^{13}C resolution, although the use of extensive forward linear prediction can partially overcome the problem. A good example of the advantage of ^{13}C detection was provided by our investigation of polyphenol-12 (19), a compound containing a sequence of 12 isoprene units.⁴⁰ With the use of HETCOR and FLOCK, it was possible to make complete assignments of peaks as close as 0.01–0.02 ppm

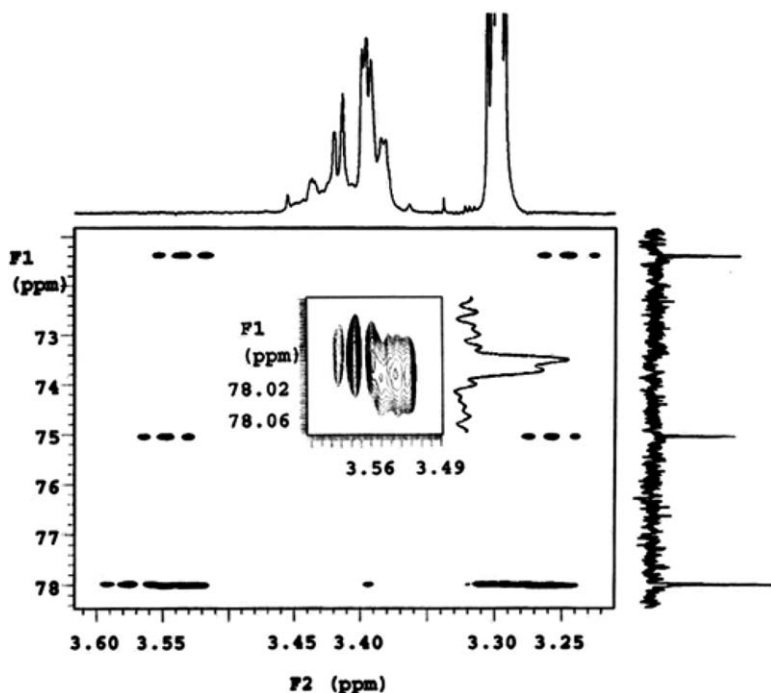


Figure 7.9 Illustration of the use of a coupled high-resolution HSQC spectrum to determine the ^1H - ^1H multiplet patterns of the glucose unit of a saponin, in spite of almost complete overlap of protons H-2 to H-5. The observation of ~ 8 –9 Hz triplets for H-2 to H-4 confirms the monosaccharide as β -D-glucose. The ^1H spectrum is illustrated along the top (with the multiplet at δ 3.30 ppm being due to CD_2HOD from the deuterated solvent). The inset spectrum is an expansion of the high-frequency multiplets from the two protons near δ 78 ppm, illustrating that the ^{13}C resolution in the HSQC spectrum is comparable with that in the 1D ^{13}C spectrum. The sample was provided by Dr Russel Ramsewak.

apart on both axes, something that was impossible with HSQC and HMBC (e.g. see Figure 7.11). With the availability of ^{13}C -optimized cryogenically cooled probes, it will now be possible to obtain these ^{13}C -detected sequences with significantly less sample, and we believe that there should be a resurrection of the use of these older sequences, particularly in cases of severe spectral crowding.

Another class of terpenes where spectral crowding is often a serious problem is saponins, particularly those with four or more attached monosaccharides. Here, selective 1D-TOCSY experiments can be very helpful. Provided that at least one ^1H signal from each monosaccharide unit is resolved, one can apply a selective excitation to that signal and, by arraying the TOCSY mixing time, assign the entire sequence of protons in each monosaccharide unit. However, this approach is hindered if one or more of the protons is in a pseudo-equatorial orientation, since the

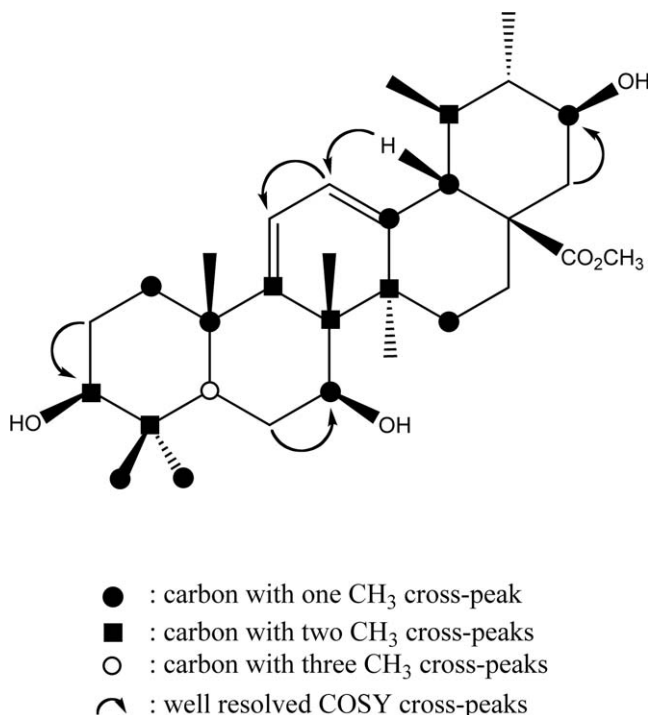


Figure 7.10 Illustration of how CH₃ HMBC cross-peaks tie together the structure of a triterpene and, along with COSY peaks from resolved proton multiplets, allow complete assignment of the skeletal structure and ¹H and ¹³C spectra of the triterpene.

(Reproduced with permission from W. F. Reynolds and R. G. Enriquez, *J. Nat. Prod.*, 2002, 65, 221. Copyright (2002) American Chemical Society.¹³)

resultant small ¹H–¹H couplings act as bottlenecks to magnetization transfer and thus require longer mixing times. If there is a partial overlap of peaks from two different monosaccharide units that one wants to excite, then benzene titration may again provide sufficient resolution. In addition, one can set the selective pulse regions to be alternatively centered at the two ends of the overlapped region and hope that there is sufficient differentiation to allow assignment. Finally, since the ¹³C spectrum may be better resolved than the proton spectrum, a HSQC-TOCSY experiment could be considered as an alternative. This is a very-low-sensitivity experiment and would be much more time consuming, particularly if one wanted to repeat it with different TOCSY mixing times. However, it has been demonstrated that applying covariance processing techniques in order to generate HSQC-TOCSY spectra can significantly reduce the time requirements for this experiment.⁴¹

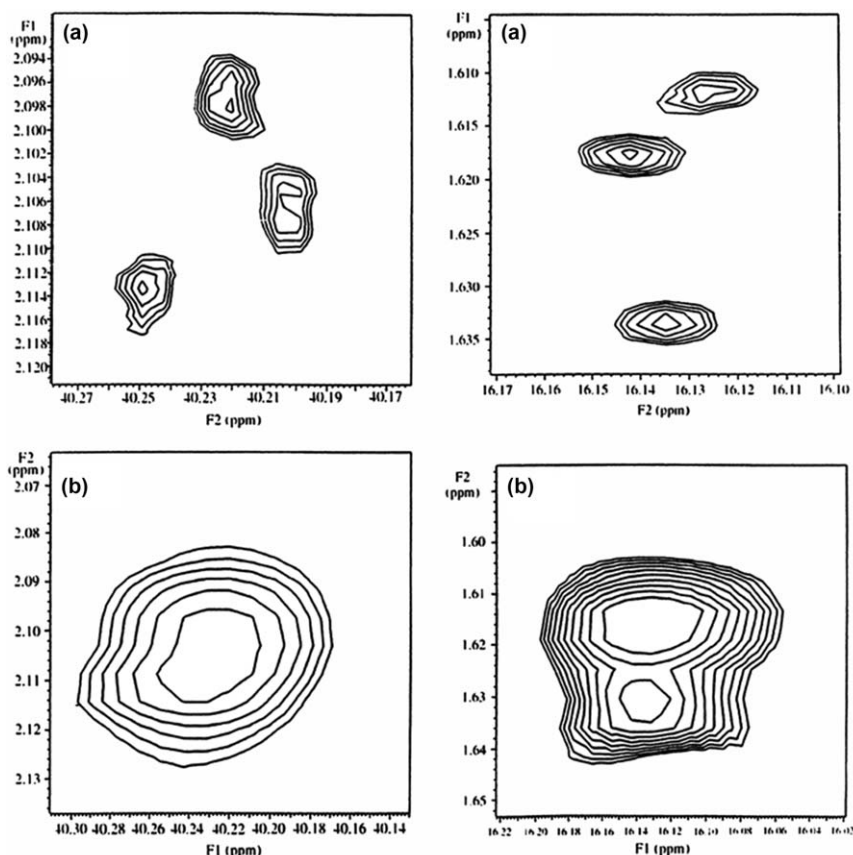


Figure 7.11 Illustration of the superior resolution of a ^{13}C -detected HETCOR spectrum compared to a ^1H -detected HSQC spectrum, in each case allowing clear resolution of three ^{13}C - ^1H peaks, which are unresolved in the HSQC spectrum. The compound is polyprenol-12 (**19**) and the spectra are taken from Reynolds *et al.*⁴⁰

With permission of the publisher © Canadian Science Publishing or its licensors.

A number of alternative long-range ^{13}C - ^1H sequences exist that have been designed to distinguish two-bond and three-bond connectivities involving protonated carbons (*e.g.* XCORFE, 2J , 3J -HMBC, and H2BC).⁴² However, as mentioned in Section 7.4, one major hindrance in interpreting HMBC spectra has, until recently, been the lack of any unambiguous method of distinguishing two-bond and three-bond connectivities between protons and quaternary carbons in particular. Fortunately, it has recently been shown that 1,1-ADEQUATE spectra can be used for this purpose and that this information can dramatically simplify the task of computer-aided structure elucidation^{3,43,44} (described in Volume 2, Chapter 11). While this is an

intrinsically very-low-sensitivity experiment, ^{13}C -optimized cryogenically cooled probes can significantly reduce sample requirements.⁴⁴ Also, it has recently been shown that covariance processing of a 1,1-ADEQUATE spectrum with a HSQC spectrum can significantly improve the signal/noise of the 1,1-ADEQUATE spectrum, and thus further reduce sample or time requirements.⁴⁵ This occurs because the signal/noise is determined by the more sensitive HSQC spectrum used in the processing, although the peaks in the original 1,1-ADEQUATE spectrum must still be sufficiently higher than the noise level in order to guarantee reliable results. Under favorable circumstances and using a 1.7 mm micro-cold probe on a 600 MHz instrument, sample requirements can be reduced to <500 μg with a weekend-long experiment. However, as noted in Section 7.1, few natural product laboratories have probes of this type available to them at present.

Finally, the ultimate approach for determining a molecular skeleton is to use the INADEQUATE sequence, which establishes carbon-carbon connectivities *via* ^{13}C - ^{13}C coupling (a 0.01% probability in unlabeled compounds). While the optimum conditions for carrying out this experiment have recently been reported,⁴⁶ this remains the least sensitive of the available 2D structure determination experiments. Again, this problem has been partially alleviated by the development of carbon-optimized cold probes. However, as noted in the introduction, it is increasingly common that unknown compounds are often isolated in very small amounts, meaning that INADEQUATE is still impossible with many new natural products. Hopefully, continuing improvements, including symmetry-based methods for extracting correlation peaks from noisy data,⁴⁷⁻⁴⁹ non-linear sampling,⁴⁹ and the use of selective pulses in combination with Hadamard processing methods,⁵⁰ will further lower the sample requirements for this sequence and extend its range of application.

7.6 Conclusions

Terpenes pose challenging structural problems for a natural product chemist due to the difficulties of distinguishing new compounds from the very large number of known compounds, the wide range of possible skeletal structures and stereochemistries, and the problems of severe spectral overlap. We hope that we have demonstrated how modern NMR methods (particularly 2D methods) can aid in structure determination, as well as minimize the risk of determining an incorrect structure.

Acknowledgements

We thank Professor Helen Jacobs, Professor Paul Reese (both of the Department of the West Indies, Mona Campus, Kingston, Jamaica), and Dr Russel Ramsewak (Department of Chemistry, University of the West Indies, St. Augustine Campus, St Augustine, Trinidad) for supplying compounds that were used in preparing four of the figures.

References

1. O. Wallach, Nobel Lecture, 1910, in *Nobel Lectures, Chemistry*, 1901–1921, Elsevier, Amsterdam, 1966.
2. L. Ruzicka, *Proc. Chem. Soc.*, 1959, 341.
3. W. P. Aue, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.*, 1976, **64**, 2229.
4. B. M. Fraga, *Nat. Prod. Rep.*, 2010, **27**, 1681.
5. J. D. Connolly and R. D. Hill, *Nat. Prod. Rep.*, 2010, **27**, 79.
6. R. J. Peters, *Nat. Prod. Rep.*, 2010, **27**, 1521.
7. D. S. Dalisay, E. W. Rogers, A. S. Edison and T. F. Molinski, *J. Nat. Prod.*, 2009, **72**, 732.
8. B. D. Hilton and G. E. Martin, *J. Nat. Prod.*, 2010, **73**, 1465.
9. K. C. Nicolau and S. A. Snyder, *J. Nat. Prod.*, 2005, **44**, 1012.
10. *Chemistry of Terpenes and Terpenoids*, ed. A. A. Newman, Academic Press, New York, 1981.
11. H. Jacobs, F. Ramdayal, W. F. Reynolds and S. McLean, *Tetrahedron Lett.*, 1986, **27**, 1453.
12. J. B. Lambert and E. P. Mazzola, *Nuclear Magnetic Resonance Spectroscopy: An Introduction to Principles, Applications and Experimental Methods*, Pearson Education Inc., Upper Saddle River, NJ, 2004.
13. W. F. Reynolds and R. G. Enriquez, *J. Nat. Prod.*, 2002, **65**, 221.
14. See www.nmrwiki.org/wiki/index.php?title=Databases for a listing of data bases.
15. www.chemspider.com.
16. W. F. Reynolds, R. G. Enriquez, L. I. Escobar and X. Lozoya, *Can. J. Chem.*, 1984, **62**, 2421.
17. B. L. Marquez, W. H. Gerwick and R. T. Williamson, *Magn. Reson. Chem.*, 2001, **39**, 499.
18. e.g. M. A. Munoz and P. Joseph-Nathan, *Magn. Reson. Chem.*, 2010, **48**, 458.
19. e.g. S. G. Smith and J. M. Goodman, *J. Am. Chem. Soc.*, 2010, **132**, 12946.
20. J. A. Dale and H. S. Mosher, *J. Am. Chem. Soc.*, 1973, **95**, 512.
21. C. M. Thiele, *Eur. J. Org. Chem.*, 2008, **34**, 5673.
22. e.g. A. J. Williams, M. E. Elyashberg, K. A. Blinov, D. C. Lankin, G. E. Martin, W. F. Reynolds, A. Porco, C. A. Singleton and S. Su, *J. Nat. Prod.*, 2007, **71**, 581.
23. P. E. Hansen, in *Progress in Nuclear Magnetic Resonance Spectroscopy*, ed. J. W. Emsley, J. Feeney and L. H. Sutcliffe, Pergamon Press, Oxford, 1981, vol. 14.
24. T. Burrow, R. G. Enriquez and W. F. Reynolds, *Magn. Reson. Chem.*, 2009, **47**, 2086.
25. B. Schlegel, A. Härtl, H.-M. Dahse, F. A. Gollmick, U. Gräfe, H. Dörfelt and B. Kappes, *J. Antibiot.*, 2002, **55**, 814.
26. J. A. Porco, Jr., S. Su, X. Lei, S. Bardhan and S. D. Rychnovsky, *Angew. Chem., Int. Ed.*, 2006, **45**, 5790.
27. H. T. D'Armas, B. S. Mootoo and W. F. Reynolds, *J. Nat. Prod.*, 2000, **63**, 1669.

28. J. I. Musher and E. J. Corey, *Tetrahedron*, 1962, **18**, 791.
29. D. Burns, W. F. Reynolds, G. Buchanan, P. B. Reese and R. G. Enriquez, *Magn. Reson. Chem.*, 2000, **38**, 488.
30. H. Barkhuisjen, H. deBeer, W. M. M. J. Bovee and D. van Ormondt, *J. Magn. Reson.*, 1985, **61**, 465.
31. W. F. Reynolds, M. Yu, R. G. Enriquez and I. Leon, *Magn. Reson. Chem.*, 1997, **35**, 505.
32. W. F. Reynolds and R. G. Enriquez, *Magn. Reson. Chem.*, 2003, **41**, 927.
33. P. Laszlo, in *Progress in Nuclear Magnetic Resonance Spectroscopy*, ed. J. W. Emsley, J. Feeney and L. H. Sutcliffe, Pergamon Press, Oxford, 1967, vol. 3.
34. J. J. Turner and N. Sheppard, *Proc. R. Soc.*, 1959, **A252**, 506.
35. R. T. Williamson, J. R. Carney and W. H. Gerwick, *J. Nat. Prod.*, 2000, **63**, 876.
36. W. F. Reynolds, S. McLean, L.-L. Tay, M. Yu, R. G. Enriquez, D. M. Estwick and K. O. Pascoe, *Magn. Reson. Chem.*, 1997, **35**, 455.
37. S. Simova, *Magn. Reson. Chem.*, 1998, **36**, 505.
38. W. F. Reynolds, S. McLean, J. Poplawski, R. G. Enriquez and L. I. Escobar, *Tetrahedron*, 1986, **42**, 3419.
39. S. McLean, M. Perpich-Dumont, W. F. Reynolds, H. Jacobs and S. S. Lachmansing, *Can. J. Chem.*, 1987, **65**, 2519.
40. W. F. Reynolds, S. McLean, H. Jacobs and W. W. Harding, *Can. J. Chem.*, 1999, **77**, 1922.
41. K. A. Blinov, N. I. Larin, M. P. Kvasha, A. Moser, A. J. Williams and G. E. Martin, *Magn. Reson. Chem.*, 2005, **43**, 999.
42. See W. Schoefberger, J. Schlagnitweit and N. Müller in *Annual Reports on NMR Spectroscopy*, ed. G. Webb, Elsevier, London, 2011, vol. 72, p. 1 for a comprehensive survey of all HMBC variants including ^2J , ^3J -HMBC and H2BC.
43. S. W. Meyer and M. Köck, *J. Nat. Prod.*, 2008, **71**, 1524.
44. S. F. Cheatham, M. Kline, R. R. Sasaki, K. A. Blinov, M. E. Elyashberg and S. G. Molodtsov, *Magn. Reson. Chem.*, 2010, **48**, 571.
45. (a) G. E. Martin, B. D. Hilton and K. A. Blinov, *Magn. Reson. Chem.*, 2011, **49**, 248; (b) G. E. Martin, B. D. Hilton, M. R. Willcott, III and K. A. Blinov, *Magn. Reson. Chem.*, 2011, **49**, 350.
46. A. D. Bain, D. W. Hughes, C. K. Anand, Z. Nie and V. J. Robertson, *Magn. Reson. Chem.*, 2010, **48**, 630.
47. R. Dunkel, C. Mayne, R. Pugmire and D. Grant, *J. Magn. Reson.*, 1990, **90**, 290.
48. T. Nakazawa, H. Sengtschmid and R. Freeman, *J. Magn. Reson.*, 1996, **120A**, 269.
49. C. K. Anand, A. D. Bain and S. C. Watson, *J. Magn. Reson.*, 2011, **210**, 146.
50. E. Kupce and R. Freeman, *J. Magn. Reson.*, 2010, **206**, 147.

CHAPTER 8

Nuclear Magnetic Resonance of Steroids

MARTIN JAEGER^{*a} AND RUUD L. E. G. ASPERS^b

^aNiederrhein University of Applied Sciences, Department of Chemistry, Instrumental Analytical Chemistry and ILOC Institute for Coatings and Surface Chemistry, Frankenring 20, D-47798 Krefeld, Germany; ^bRadboud University, Institute for Molecules and Materials, Biophysical Chemistry, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

*Email: martin.jaeger@hs-niederrhein.de

8.1 Prologue

When invited to write a chapter on the nuclear magnetic resonance (NMR) of steroids, one immediately thinks of the classic treatises by Kirk *et al.*,^{1–3} succeeded by Kasal *et al.*,⁴ and the textbook by Croasmun and Carlson.⁵ In the general steroid field, the masterpiece by Fieser and Fieser⁶ is instantly memorable, as are the autobiographical essays of Djerassi.^{7,8} Our intention was therefore not to rewrite the art of our predecessors, but rather to hope that our readers will allow themselves to be intrigued by our selection and compilation of contemporary NMR topics that have not been assembled previously.

8.2 An Introduction to Steroids

Chemical compounds of the steroid class possess a noticeable common feature. Their skeleton consists of four fused rings, most often one five-membered and three six-membered carbocycles. The unsubstituted

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

carbohydrate scaffold with a total of 17 carbon atoms has been baptized gonane. Although this structure does not actually exist in nature, it forms the base of the chemical class of the steroids. Steroids that naturally occur as derivatives from squalene or isoprenes as precursors usually possess methyl groups at the A/B and C/D ring junction (*cf.* Figure 8.1) that can be modified or eliminated during metabolism or other enzymatic reactions. Examples include aldosterone or estradiol. The number of carbon atoms, which is augmented by the substituents on the gonane scaffold, determines the membership of the compound in the steroid subclass. Cholestanes comprise 27 carbon atoms; hence, cholesterol is a member of that series. The cholanes and pregnanes have 24 and 21 carbon atoms, respectively. Examples are cholic acid and progesterone, respectively. Androstanes and estranes complete the sequence of steroid classes; testosterone (**1**) and estradiol are probably the best-known representatives of these classes, possessing 19 and 18 carbon atoms, respectively. In parallel to this chemical classification, a historical nomenclature exists, which is based on the origin of the compound or its main natural source: progestagens, androgens, estrogens, glucocorticoids, and mineralocorticoids. Taxonomy and biological function provide a third classification. Steroids are found in animals, plants, and fungi. The latter are a source of ergosterols and vitamin D. In plants, phytosterols are present, among them ergosterol, sitosterol, and diosgenin. As late as 1979, brassinosteroids were discovered in plants; their B ring is constituted as a seven-membered ring.

Animal steroids are divided into insect and vertebrate steroids. Among insect steroids, ecdysteroids are the most prominent class, with ecdysterone itself as an example. Steroids in vertebrates usually serve as hormones. These are sex steroids, including androgens, estrogens, and progestogens,[†] which play various roles in sexual differentiation and reproduction. Further, there are the corticosteroids that regulate metabolism and immune functions, blood volume, and excretion. Anabolic steroids influence muscle and bone synthesis. Finally, the parent compound of all vertebrate steroids is cholesterol (**2**), which is in itself a modulator of cell membrane fluidity and whose plaques may cause atherosclerosis.

An overview of the biochemical pathways of steroidal compounds is illustrated in the famous Boehringer scheme,⁹ and a concise, modernized scheme has recently been published on the Internet.¹⁰ A few representative steroids are collected in Figure 8.1. The structures should guide the reader through this text and at the same time exemplify the traditional classes of steroids.

For an introduction on the history of steroids, we would like to direct the reader's attention to a collection of essays and narrative reviews that appeared in two special issues of *Steroids* in August and December 1992,^{11–25} followed by two contributions in August 1996.^{26,27}

[†]Progestagens, progestogens, and progestins are synonymously and non-consistently used for exogenous or synthetic gestagen hormones.

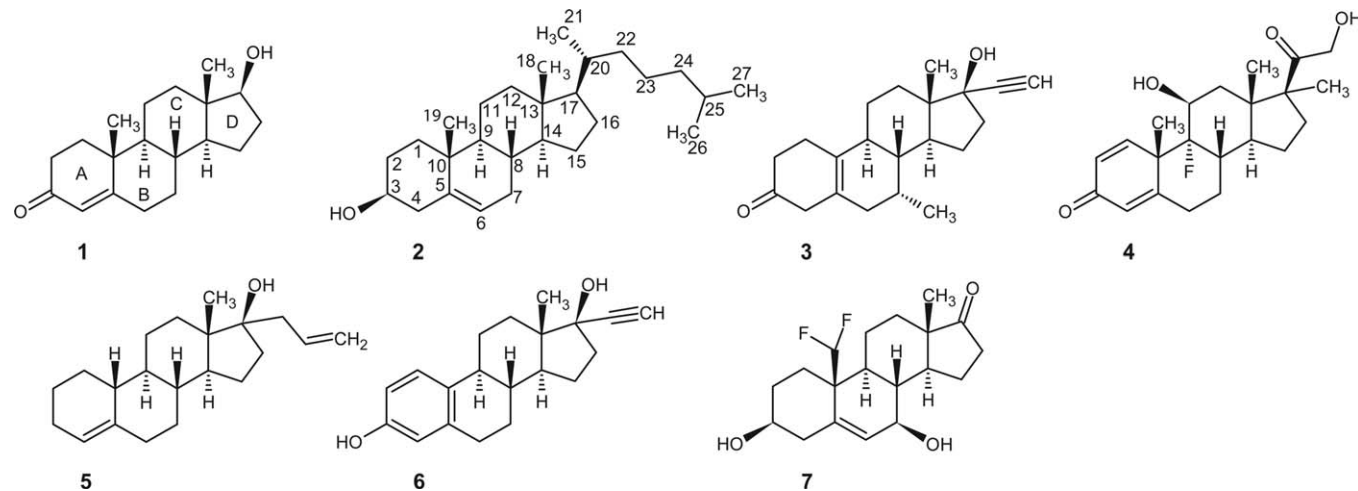


Figure 8.1 Selected steroidal structures used in this chapter: testosterone (1), cholesterol (2), tibolone (3), dexamethasone (4), allylestrenol (5), 17 α -ethinylestradiol (6), and 19,19-difluoro-3 β ,7 β -dihydroxy-androst-5-en-17-one (7).

In the history of steroids, the 1930s mark a turning point in their molecular perception. Not only were the first steroid chemical constitutions fully and correctly elucidated and published in 1932, but it was also demonstrated that spectroscopy would greatly contribute to the future identification of steroids. Still, chemical reactions or degradations were the predominant means of structure identification, where the elucidation of cholanolic acid,^{28,29} testosterone (1),³⁰ estrone,^{31,32} progesterone,^{33,34} and cortisone³⁵ were early milestones of the pre-spectroscopic era. It was not until the late 1950s that NMR spectroscopy started to slowly grow into the primary analytical tool that it is today for the elucidation of chemical constitution and stereochemistry, not only for steroids, but also for all kinds of chemical and biological compounds. In 1958, Shoolery and Rogers published a remarkable collection of NMR spectra of approximately 50 steroids, recorded at 40 MHz.³⁶ The interested and especially younger readers are referred to this article for an impression of what crowded NMR spectra look like at 40 MHz. To the same purpose, a 40 MHz 1D ¹H spectrum of 4,4-dimethyl- $\Delta^{5,10}$ -3-ceto-estrenol is presented in Figure 8.2.³⁷

The analysis of steroid NMR spectra in the late 1950s and early 1960s was limited to a habitus or pattern description supported by a copy of the spectrum; due to the “hump” or “envelope” of the methylene and methine resonances, individual chemical shifts could only be determined for special absorptions. In that respect, NMR at that time was comparable to infrared (IR) spectroscopy. Shoolery and Rogers³⁶ thus suggested using the hump in a manner similar to the IR fingerprint region, but stated that the chemical substitution of a single proton produced a marked and characteristic effect on the spectrum. In their discussion of the applicability of the NMR method to the elucidation of unknown steroids, they assured that “as a means of determining the number of methyl groups, and to some extent their position, the NMR method may very well be unrivaled. Although based on empirical correlation with established known structures...” Nevertheless, the distinction of axial and equatorial protons was already mentioned. Using 60 MHz spectrometers in 1959, stereochemical aspects of sapogenins were investigated.³⁸ The investigations on steroids using NMR continued, relying mainly on the methyl resonances and certain prominent signals. In the early 1960s, extensive reviews on steroid shift data already comprising 160 known and 100 new steroids were published by Zuercher.^{39,40} From the analysis of the chemical shifts of the 18- and 19-methyl resonances, a system to calculate their chemical shifts depending on the substitution of the steroid skeleton was devised. A value was attributed to a substituent taking into account its chemical constitution and its relative position towards the nucleus whose chemical shift was to be calculated. For multiple substitutions, these increments could be added.

It was not until the advent of FT-NMR around 1970 that NMR could be elevated to the next level of important contributions to the steroid field. This

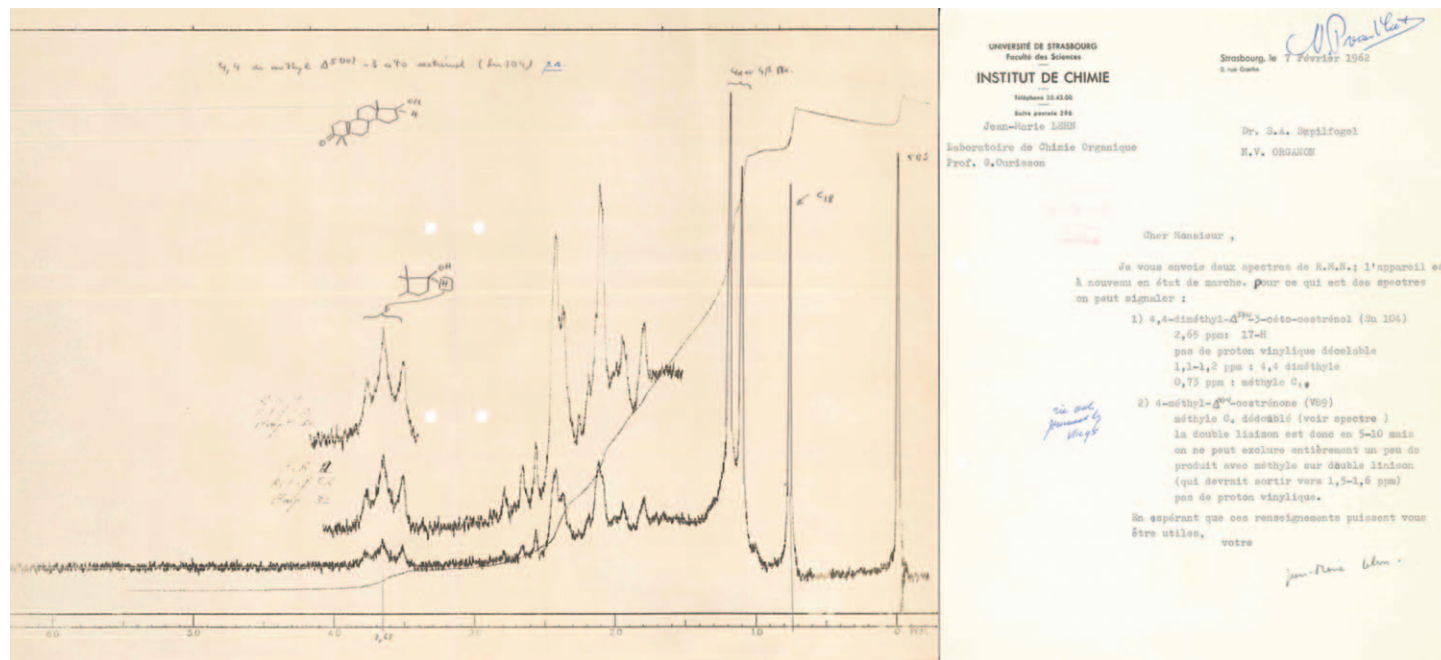


Figure 8.2 1D ^1H spectrum of 4,4-dimethyl- $\Delta^{5,10}$ -3-ceto-estrenol acquired at 40 MHz and the corresponding interpretation sheet. The spectrum was recorded for N. V. Organon at the University of Strasbourg and interpreted by the later Nobel Laureate J.-M. Lehn. The original spectrum was provided by M. B. Groen. (Reproduced from Jaeger and Aspers, *Steroids and NMR*, *Annu. Rep. NMR Spectrosc.*, 77, 115–285 Copyright (2012) with permission from Elsevier.³⁷⁾

achievement was not so much attained by the ^1H resolution increasing along with the magnetic field from 100 to 200 MHz, but by the utilization of ^{13}C spectroscopy performed with proton broadband decoupling, where the apparently higher resolution is due to the larger spectral dispersion and the absence of multiplet resonances. The increase in resolution depending on the field strength is illustrated for tibolone (3) in Figure 8.3.

From that period on, the molecular constitutional investigations could be carried out in a straightforward manner by means of ^{13}C NMR, first creating and then relying on the wealth of compiled reference data, *cf.* the compilation of 30 ^{13}C spectra,⁴¹ the assignment of ketosteroids,⁴² the distinction of 42 phytosterols,⁴³ and the differentiation of C-24 isomeric sterols.⁴⁴ The assignment of the carbon chemical shifts was supported by selective decoupling techniques, chemical reactions, lanthanide shift reagents, and selective and partial deuteration.⁴⁵

In 1977, Blunt and Stothers published a huge compilation with more than 400 steroids, and also revising the assignment of earlier works.⁴⁶ The result of all of the reference compilations was the establishment of a chemical shift increment system for the analysis and prediction of steroid ^{13}C chemical shifts. The system was continuously extended, since the interpretation of chemical shifts remained a prime means of structure confirmation. The work by Blunt and Stothers also summarized the methods available at that time: shielding data interpretation, off-resonance decoupling, selective proton decoupling, spectral comparison, isotopic labeling, the use of lanthanide shift reagents, and spin-lattice relaxation time measurements.

Yet the issue of the complete elucidation of stereochemistry had remained unsolved. With field strength and thus resolution further increasing, and the exploitation of nuclear Overhauser enhancement (NOE) spectroscopy,⁴⁷ the assessment of the full 3D structure became possible from 1980 onward. As the first example, the conformation of 17 α -acetoxy-6 α -methylpregn-4-ene-3,20-dione was fully analyzed by Farrant *et al.*⁴⁸ using 2D NMR and NOE spectroscopy. While the 1D techniques continued to be used, 2D techniques became more and more common after the advent of 400 MHz spectrometers. The first full proton assignment was described by Hall and Sanders for 1-dehydrotestosterone in 1980 and for 11 β -hydroxyprogesterone in 1981.^{47,49} A culmination of the description of 2D NMR applied to steroids occurred in 1985 when several research groups used homo- and hetero-nuclear correlation techniques such as COSY-90, COSY-45, heteronuclear detected C-H correlation experiments, 2D *J*-resolved experiments, and ^{13}C - ^{13}C INADEQUATE experiments.⁵⁰⁻⁵² For the first time, the considerable importance of NMR was overtly stated for norethisterone: "Two-dimensional NMR appeared to be superior to mass and IR spectroscopy in identifying the isomers."⁵³ By then, the principal experiments that were considered essential for a full structure analysis according to today's standards were developed.

In the years that followed, instrumentation evolved further, providing ever-increasing field strength, gradient technology,^{54,55} small-volume,^{56,57}

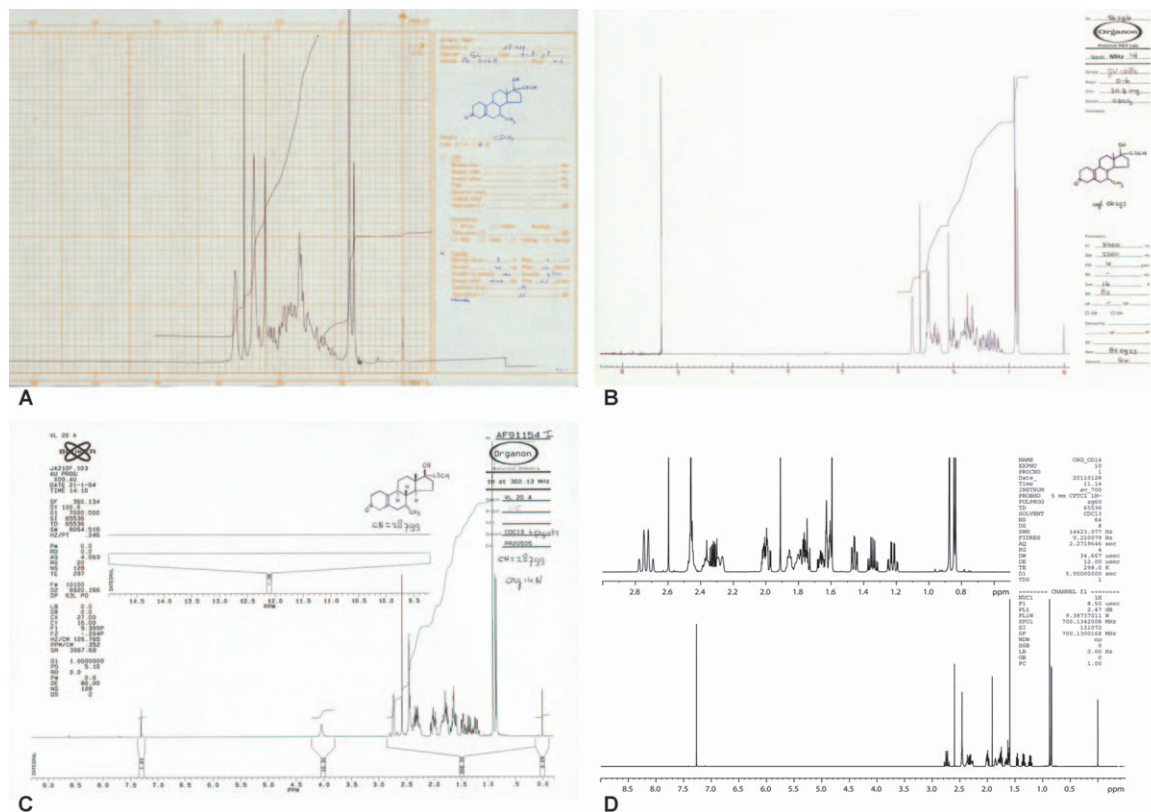


Figure 8.3 1D ^1H NMR spectra of tibolone (3) in CDCl_3 at different field strengths: (A) 90 MHz, (B) 200 MHz, (C) 360 MHz, and (D) 700 MHz. Historically, all spectra are recorded with different spectral resolutions and processed with different window functions. At 400 MHz, mathematical resolution enhancement using a Gaussian window function was employed, whereas no window function at all was applied at 700 MHz. (Reproduced from Jaeger and Aspers, *Steroids and NMR*, *Annu. Rep. NMR Spectrosc.*, 77, 115–285 Copyright (2012) with permission from Elsevier.³⁷)

and cryogenically cooled probes.^{58,59} These state-of-the-art devices are now widely and routinely available. Experiments were adapted to the new technologies, taking advantage of inverse detection⁶⁰ and gradient-enhanced experiment versions.⁵⁵ Several strategies were available to approach stereochemistry. The original Karplus equation relating 3J -couplings to the dihedral angles of the coupled nuclei dating from 1959⁶¹ was refined in 1980 by Haasnoot *et al.*⁶² On the other hand, NOE spectroscopy provided a complementary tool based on dipolar and thus through-space mechanisms. Its use for steroid stereochemical definition was noted above. In 1987, the conformation of the A ring in solution was studied and the result compared to the solid-state structure obtained from X-ray crystallography.^{63,64} The dialectic of solution versus solid-state structure in conformational space remains—especially when structure–activity relations are discussed—a point of attention that remains an issue. For a thorough description of historical methodologies, the reader is referred to the book chapter by Croasmun and Carlson⁶⁵ contained within the book by the same authors.⁵ As structure elucidation was not a routine task until the mid-1990s, strategies for steroid structure elucidation are discussed based on the experiments available at that time.^{2,65–67}

Despite the availability of techniques that allowed full chemical shift assignment in the mid-1980s, it was as late as 1990 that the full assignments of one of the best-known steroids, testosterone, and its epimer 17 α -methyltestosterone were published.^{68,69} Remarkably, the publications contained interchanged proton assignments for H16 α and H16 β . The assignment of Kirk *et al.* turned out to be the correct one,⁶⁹ yet the supporting experiments were not finally published until 2011.⁷⁰

By the 1980s, the interest of the pharmaceutical industry in steroids had declined and the golden age of steroids had faded, despite the fact that steroidal compounds are still very widespread in contraception and inflammation (*cf.* Table 8.1).

In addition, there are 15 more steroidal drugs among the drugs ranked 100–200 according to prescriptions in the USA in 2010.

Due to their molecular properties, availability, the wealth of data, and their favorable NMR features, steroids continue to hold a valuable place in the development of NMR methodology, such as NOE, inverse experiments, shaped pulses, 3D experiments, residual dipolar couplings (RDCs), and pure shift or covariance methods. As a concluding remark to this section, it should be mentioned that six Nobel Prizes were received for work related to or based on steroids. The chemistry Nobel Laureates were Heinrich Wieland in 1927, Adolf Windaus in 1928, Adolf Butenandt and Leopold Ruzicka in 1939, and Derek Barton and Odd Hassel in 1969. For medicine and physiology, the Nobel Prize was awarded to Philip S. Hench, Edward C. Kendall, and Tadeus Reichstein in 1950 and to Feodor Lynen and Konrad E. Bloch in 1964.

Table 8.1 Marketed drugs that contain steroidal compounds as sold on the US market in 2010 and according to worldwide sales as of 2009. Source: <http://njardarson.lab.arizona.edu/content/top-pharmaceuticals-poster>.

Active ingredient	Drug name	Ranking wrt prescription in the USA (prescriptions in million units)	Ranking wrt worldwide sales 2009 in million USD	Company	Indication
Fluticasone and salmeterol	Advair, Seretide	9 16.6	4 8094	GSK	Asthma
Mometasone	Nasonex	18 8.8	68 1517	Merck & Co.	Nasal
Conjugated estrogens	Premarin	24 7.7	138 833	Pfizer	Menopausal symptoms
Norethindrone acetate and ethynyl estradiol	Loestrin 24 Fe	31 6.7	—	Warner Chilcott	Anticonception
Fluticasone propionate	Flurent HFA, Flixotide	44 5.1	83 1330	GSK	Asthma
Etonogestrel and ethynyl estradiol	Nuvaring	45 4.8	—	Merck & Co.	Anticonception
Drospirenone and ethynyl estradiol	Yaz	52 4.6	128 885	Bayer	Anticonception
Norgestimate and ethynyl estradiol	Ortho Tri-Cyclen Lo	52 4.2	—	Ortho-McNeil	Anticonception
Budesonide and formoterol	Symbicort	59 3.6	41 2353	AstraZeneca	Asthma
Dutasteride	Avodart	63 3.5	133 865	GSK	BPH Products
Estradiol	Vivelle-Dot	75 2.8	—	Novartis	Estrogen replacement
Conjugated estrogens and medroxyprogesterone	Prempro	77 2.7	—	Pfizer	Menopausal symptoms
Desogestrel and ethynyl estradiol	Apri	87 2.4	—	Teva	Anticonception
Ciprofloxacin and dexamethasone	Ciprodex Otic	94 2.2	—	Alcon	Anti-infection and anti-inflammation
Testosterone	Androgel	103 2.0	166 708	Solvay	Menopausal symptoms (?)
Finasteride	Propecia, Proscar	138 1.3	177 663	Merck & Co.	Benign prostatic hypertrophy
Drospirenone and ethynyl estradiol	Yasmin	177 0.8	151 766	Bayer	Anticonception
Butenoside	Pulmicort	196 0.7	70 1505	AstraZeneca	Asthma

8.3 Modern and “Rare” NMR Methods in the Steroid Field

8.3.1 Recent General NMR Developments

The following section shall point out the most recent methodological studies on steroids and also remind us of some examples of experiments that have not often been applied to steroids. Historically, there was vibrant cross-fertilization between steroid research and NMR spectroscopy. Although quite a few steroidal compounds are still successfully marketed drugs (*cf.* Table 8.1), the liaison between steroids and modern NMR developments has cooled in recent years. Nevertheless, newly developed experiments and technologies are implemented in laboratories that study steroids and are applied for structure elucidation purposes. However, the results are not consistently made public. For those spectroscopists who are just entering the field of small molecules, there is a variety of textbooks and reviews to choose suitable experiments from.^{71–76} The experiments described in these references also work well for steroids.

In search of consistently improving the already advanced modern standard set of experiments, sophisticated hardware and software extensions add to the spectroscopist's toolbox. Special attention was paid to fast NMR spectroscopic methods.^{77–80} These methods either constitute experimental techniques where several experiments are combined into one sequence,^{81,82} time-shared experiments, or parallel acquisition and detection.^{83–87} Other fast methods make use of non-linear or statistical sampling of data points in multidimensional experiments^{88,89} or rely on computational techniques, such as covariance processing.^{90–95}

8.3.2 Covariance NMR and Steroids

Covariance NMR is of special interest in those cases where there is only limited experimental time since the spectral computation involved in covariance processing does not require further or extended experiments, but mathematically improves the visualization of information by enhancing spectral resolution in sparsely sampled spectral dimensions,^{92,93} or can link and transform heterogeneous datasets.^{90,96} The power of covariance spectroscopy in the form of unsymmetrical indirect covariance (UIC) processing allows access to low-sensitivity or time-consuming correlation experiments (*e.g.* HSQC-NOESY or ^{13}C – ^{13}C COSY).^{90,91,96} The higher sensitivity of UIC hyphenated spectra (*e.g.* HSQC-TOCSY) as compared to their fully experimental counterparts stems from the component spectra, HSQC and TOCSY in this case, which are more sensitive than their combined pulse sequence.^{90,97} For steroids, the revelation of long-range couplings resulting from the enhanced sensitivity of the UIC method⁹⁸ is of special interest in the field of steroid analysis.

Aspers *et al.*⁷⁰ studied four steroids in order to investigate the effect of spectral congestion on the usefulness of UIC processing of spectral data. The signal density and resulting spectral congestion increased from dexamethasone (4) through testosterone and allylestrenol (5) to tibolone (6). Experimental and UIC-computed hyphenated spectra, such as HSQC-TOCSY and HSQC-NOESY, were compared. Depending on the signal crowding, the computed spectra proved to be superior with regard to sensitivity and fast accessibility. In the case of the heavily congested spectra of tibolone, artifacts from residual noise and spectral overlap hampered the distinction between true correlations and artifacts. The usefulness for structure elucidation was thus considered less important. The recording of 3D HSQC-NOESY data was not described as advantageous for steroids, since the necessary information could be derived from 2D spectra or UIC 2D-HSQC-NOESY. Covariance transformation of hyphenated experiments resulted in $\text{CH}_m\text{-CH}_n$ COSY and $\text{CH}_m\text{-CH}_n$ NOESY spectra. Although applicable for structure verification in the cases of low and moderate spectral density, they were regarded as inferior to a $^{13}\text{C}\text{-}^{13}\text{C}$ INADEQUATE experiment with respect to readily available structure information. While the HSQC-TOCSY and HSQC-NOESY experiments were found to be very useful in general, as they were quickly derived with sufficient quality from data recorded in screening mode, the use of $\text{CH}_m\text{-CH}_n$ COSY and NOESY was recommended only for special purposes. The effects of spectral noise and insufficient experimental parameter optimization on the computed spectra were demonstrated as well.

The NOESY, HSQC, and UIC HSQC-NOESY spectra of tibolone are shown in Figure 8.4. Analyses of a correct C15–H16 correlation and of an artifact between C1–H4 will highlight the benefits and concerns in performing UIC. Legitimate correlations, such as C15–H16, and artifacts, such as C1–H4, are labeled in the spectra of Figure 8.4. The NOE cross-peak H15–H16 was correlated to C15–H15 in the HSQC spectrum during covariance processing. After this transformation, the peak C15–H16 was obtained in the HSQC-NOESY spectrum. The artifact C1–H4 observed in the UIC spectrum could analogously be traced back to the NOE signals H1–H2 and H2–H4: the signals of H1 and H2 overlapped in the proton dimension of the HSQC and therefore interfered with the covariance calculation and caused the false correlation C1–H4. In the UIC HSQC-NOESY, direct H–C correlations can be distinguished from H–C NOE correlations by their opposite phases, indicated as different colors within the spectra in Figure 8.4.

On the basis of the spectra shown in Figure 8.4, the effects of t_1 noise, the range of NOE intensities, and the use of non-optimized experimental parameters were discussed for structure elucidation and structure verification. Since NOESY spectra were found to suffer from these effects more than COSY or TOCSY data, the optimization of experimental parameters was considered to be a prerequisite for further covariance processing.⁷⁰

The combinations that result from using phase-sensitive and magnitude-mode spectra as component spectra for UIC transformations were compiled and analyzed with respect to the information that could be obtained. The use

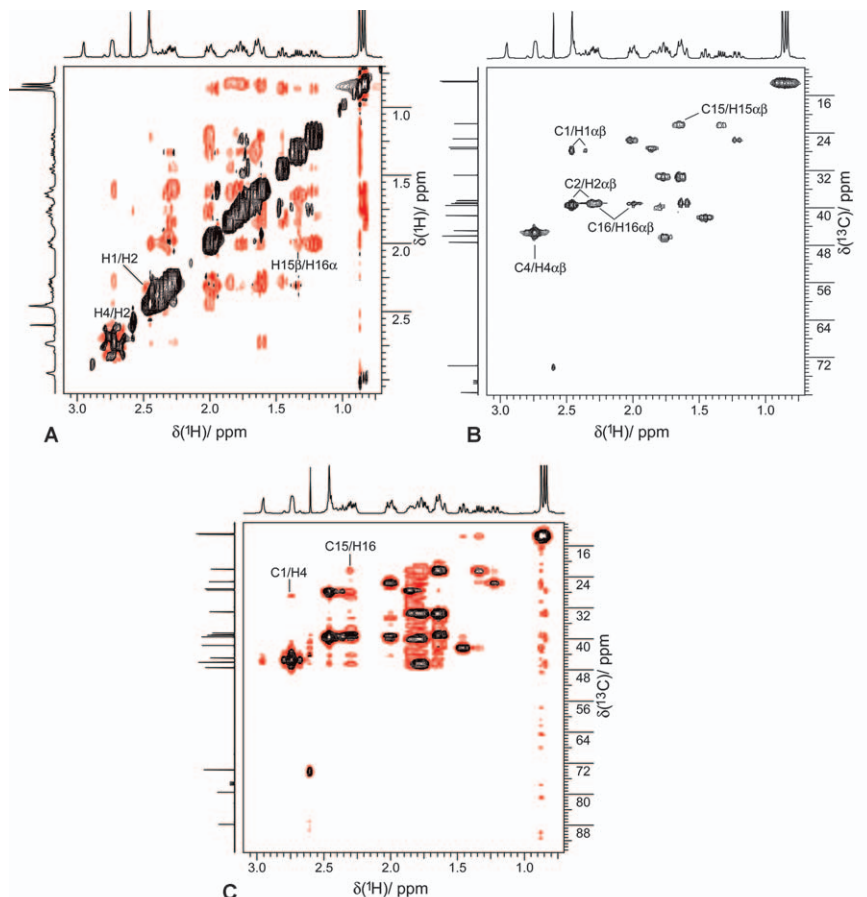


Figure 8.4 UIC calculated H-C HSQC-NOESY spectrum (C) from H-H NOESY (A) and H-C HSQC (B) of tibolone. The signals leading to the true signal C15-H16 and the artifact C1-H4 in the UIC HSQC-NOESY are indicated in the NOESY and HSQC spectra. The t_1 noise in the UIC spectrum is a remainder from the NOESY spectra acquired with generic parameters without sample optimization. (Reprinted from Aspers *et al.*⁷⁰ with permission from John Wiley and Sons. Copyright © 2011 John Wiley & Sons, Ltd.)

of covariance-computed spectra should be attractive to routine laboratories since it is fast to apply. If routine usage of covariance processing is considered, we suggest that the covariance spectrum should determine whether a phase-sensitive or magnitude experiment should be employed, since many spectral combinations maintain the phase information.⁷⁰

In the field of steroid analysis, the challenge is not necessarily the development of new experiments to allow for the solution of a specific problem, but rather to choose the most suitable experiment and avoid falling into a trap of interpreting data in a vicious circle. Sitosterol serves as a

good example.⁷² The C7 methylene and C8 methine signals overlap in a 1D ^{13}C experiment, leading to ambiguity in interpreting $^{13}\text{C}\{^1\text{H}\}$, ^{13}C DEPT, and APT spectra. This problem is of course not encountered when utilizing edited H-C HSQC spectra.

In the following section, we will examine studies in which techniques were applied that we believe were not exploited to their full potential.

8.3.3 The HSQC-TOCSY Experiment

With modern NMR instrumentation, the somewhat low sensitivity of the HSQC-TOCSY experiment may be overcome such that experiments can be recorded on a few milligrams of material within a few hours,^{70,99} and the utilization of this experiment in natural product structure elucidation will hopefully continue to grow.¹⁰⁰ Interesting phase-sensitive variants of HSQC-TOCSY have been described that allow for the distinction of direct C-H correlations from relayed ones^{99,101} or the discrimination of C-H multiplicities.¹⁰² The use of the latter was demonstrated on a dimeric cholestene derivative where all resonances could be assigned on the basis of the phase-sensitive HSQC-TOCSY experiment. Although the multiplicity information can, in principle, be derived from 1D techniques such as DEPT or APT, the edited HSQC-TOCSY will be helpful for decreasing signal density in strongly crowded spectra or spectral regions, as is the case for androstane (6), tibolone (3), and others. The HSQC-TOCSY experiment with inverted-direct responses or correlations, IDR-HSQC-TOCSY, was shown to enhance resolution as well.¹⁰¹ Its great strength lies in the separation of direct responses from relayed ones and the help with interpretation therein. Nevertheless, this experiment has not been described for use with steroids. The successful covariance construction of IDR-HSQC-TOCSY spectra was reported.⁹⁰ A method to enhance the spectral dispersion by spreading the resonances into three dimensions was termed 3D MAXY-HMQC for maximum-quantum correlation spectroscopy.¹⁰³ The separation of the CH_n resonances of dexamethasone was successfully demonstrated. The experiment was shown to provide essentially the same sensitivity as other 3D HMQC or HSQC experiments.

8.3.4 ^{13}C -Detected Experiments

A very promising tool for the structure elucidation of small molecules, including steroids, is the ^{13}C -sensitive cryoprobe that allows ^{13}C direct detection and, especially, ^{13}C - ^{13}C INADEQUATE experiments on only a few milligrams of sample in less than a day.^{78,79} Also, other heteronuclear direct-detection experiments benefit greatly from these hardware developments. As an example, a new update of the edited 1D ^{13}C NMR experiment was recently given in the form of the DEPTQ experiment,^{104,105} and its improved sensitivity was demonstrated using cholesterol as an example molecule. It may be predicted that the ^{13}C - ^{13}C INADEQUATE experiment will be considered

more often in future analyses. Various versions and practical aspects of the ^{13}C - ^{13}C INADEQUATE experiment were reviewed in 2002.¹⁰⁶ More recently, INADEQUATE was investigated with regard to problems, artifacts, and solutions thereof.¹⁰⁷ Among other non-steroidal compounds, 17α -ethinyl-estradiol (**6**) was studied. Due to the very high signal-to-noise ratio obtained, not only direct ^{13}C - ^{13}C correlations, but also long-range correlations of up to three bonds could be visualized (*cf.* Figure 8.5). This visualization was achieved by varying the interpulse delay within the spin-echo part of the pulse sequence from values of 0.005 s to 0.05 s, corresponding to a $^1J_{\text{CC}}$ equal to 50 Hz and an $^nJ_{\text{CC}}$ equal to 5 Hz, respectively. While most of the direct correlations were observed at a delay τ equal to $1/(4J_{\text{CC}})$ with $J = 50$ Hz, the missing ethinyl correlations for C19-C20 and C19-C17 could be detected at $\tau = 3/(4J_{\text{CC}})$, since the magnitude of the couplings are $^1J_{\text{C19-C20}} = 163$ Hz and $^1J_{\text{C19-C17}} = 72.1$ Hz.

The long-range correlations $^nJ_{\text{CC}}$ were observed at $\tau = 0.05$ s; they can be recognized from the dashed lines in Figure 8.5. The spectra were obtained from overnight experiments on 10 to 25 mg of sample.

8.3.5 1D and High-resolution 1D ^1H Methods

In many routine laboratories, the use of maximally generic experiment sets sometimes prevents specific experiments from being considered. A 2D experiment is often believed to provide all of the information from which the necessary parts can be extracted as needed. Occasionally, the application of selective experiments may lead to a more straightforward, problem-oriented solution and also to an answer that is more reliable with respect to data quality. In 1997, the accuracy of 1D pulsed-field gradient (PFG) NOE experiments, PFG-NOE, was analyzed, and NOE enhancements of 0.02% could be reliably measured.¹⁰⁸ Based on the high-quality data thus obtained, the initial linear part of the NOE build-up curve could be determined, from which intermolecular distances could be derived. The steroid 11β -hydroxyprogesterone served as an example. In another study, selective excitation in combination with PFGs was also used for the chemical shift assignment of stigmasterol.¹⁰⁹ The complete proton assignment succeeded, solely based on the 1D selective COSY, relayed COSY, TOCSY, and NOESY. As an advantage of the 1D methods, the high resolution allowing the direct determination of coupling constants was emphasized.

Together with NOE data, scalar coupling constant analysis remains a central pillar of conformational structure elucidation. In the field of steroids, the reduction of spectral overlap is hence a driver of the search for experiments that are suitable for the determination of coupling constants. Proton-proton and one-bond proton-carbon splittings of chenodeoxycholic acid were therefore determined from phase-sensitive H-C HMQC spectra.¹¹⁰ More recent experiments for the measurement of $^nJ_{\text{CH}}$ coupling constants have also been compiled in the literature.¹¹¹

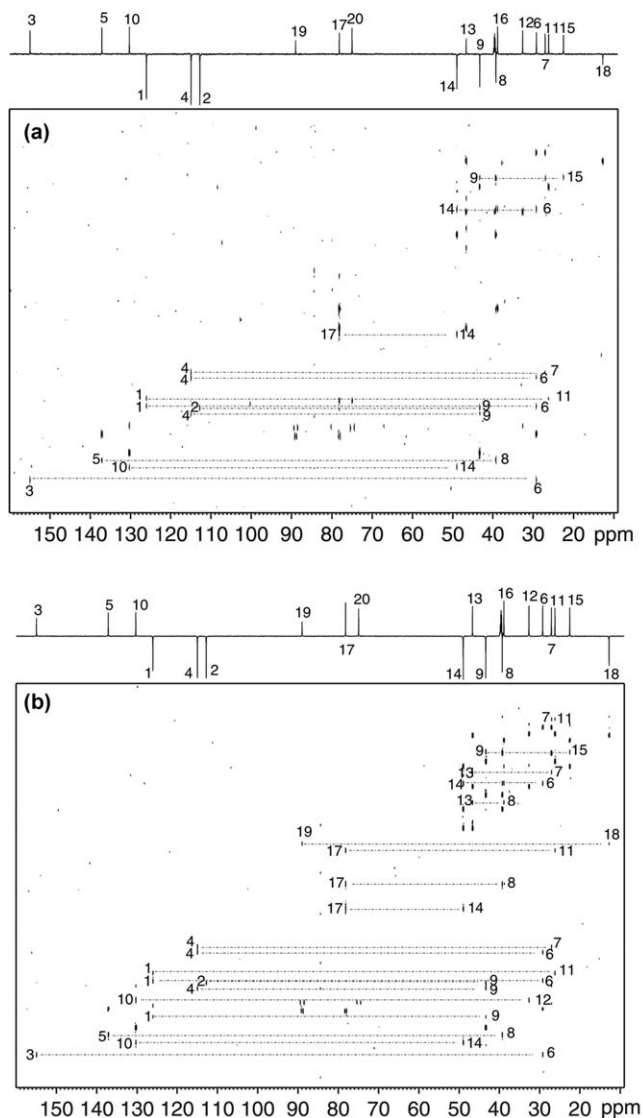


Figure 8.5 Comparison of the ^{13}C – ^{13}C INADEQUATE spectra of 17 α -ethynylestradiol (6): (A) the standard experiment recorded with $\tau = 0.005$ s from $\tau = 1/(4J_{CC})$ where $J_{CC} = 50$ Hz and (B) the long-range version recorded with $\tau = 0.05$ s from $\tau = 1/(4^n J_{CC})$ where $^n J_{CC} = 5$ Hz. The dashed lines indicate correlations arising from two- and three-bond C–C coupling interactions. (Reprinted from Bain *et al.*¹⁰⁷ with permission from John Wiley and Sons. Copyright © 2010 John Wiley & Sons, Ltd.)

Despite their value in performing structure elucidation, coupling constants also create complexity in spectra in terms of making the identification and interpretation of signals a challenging task, especially in crowded

spectra such as those of many steroids. While heteronuclear decoupling was successfully applied early within ^{13}C NMR to reduce signal complexity, homonuclear decoupling was used as a selective decoupling technique prior to the advent of 2D techniques. Decoupling over the whole spectral range remained impossible.

In the study of steroids, three attempts resulted in the equivalent of a full-scale $^1\text{H}\{^1\text{H}\}$ NMR spectrum, also referred to as a pure shift spectrum. In 1999, Guenneau *et al.* reported on the application of a *J*-resolved experiment for the determination of relaxation times.¹¹² The method proved unsuccessful for transverse relaxation times due to pulse imperfections and non-resolved long-range *J*-couplings. In contrast, longitudinal and rotating frame relaxation times could be determined provided an appropriate preparation period preceded the 2D spin-echo sequence of the *J*-resolved experiment. Lorentzian lines were obtained in the power spectra, which were found to be suitable for the estimation of ^1H relaxation parameters. Using this method, the relaxation times of androstene were determined. The authors also concluded that fully decoupled spectra could be obtained from spin-echo experiments, although in a non-quantitative manner. Due to the 2D character of the experiments, they proved to be rather time consuming for a pure shift spectrum.

In 2007, Pell and Keeler applied 2D *J*-spectroscopy in order to observe homonuclear decoupled spectra.¹¹³ They introduced the Zangger-Sterk pulse sequence¹¹⁴ into the *J*-resolved experiment, with the result that absorption mode lineshapes were obtained. Although the spectral sensitivity was found to be low, the integral of the absorption lines reflected the number of protons causing the resonance. The resolving power of this experiment was demonstrated for dehydroisoandrosterone and a series of other compounds. The regular and the decoupled ^1H NMR spectra of the steroidal compound are presented in Figure 8.6.

In 2011, Kolkman *et al.*¹¹⁵ developed a method based on reconstruction. Selectively decoupled spectral regions were summed up to yield the full-range spectrum. The experiment may be considered to be a development of decoupling using spatial encoding through gradient pulses.¹¹⁶ The improvements consist of the full utilization of the sample size as compared to the original sample slicing, fast pulse repetition, and a sensitivity increase while retaining quantitative information. The experiment was applied to blood plasma samples and testosterone.

Fully decoupled or pure shift ^1H NMR experiments are likely to assume a role in the identification of compounds in mixtures, particularly when the compounds themselves provide complex or crowded spectra. Two important features might be considered essential for a future role: a reasonable experiment time and the signals being quantitative.

8.3.6 ^{19}F NMR

Fluorine NMR spectroscopy has frequently been used for steroid structure analysis, although it does not seem to be as widely exploited as it could be.¹¹⁷

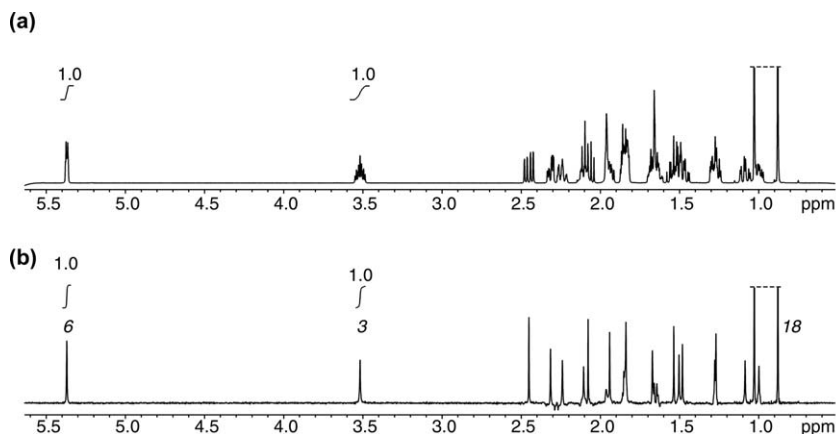


Figure 8.6 The 1D ^1H spectrum (A) and projection of the absorption-mode J -spectrum (B) of dehydroisoandrosterone. In (A) is shown the conventional spectrum of the full chemical shift range. The J -spectrum was acquired with 44 scans per time increment. The spectral widths in ω_1 and ω_2 are 100 and 2561 Hz, respectively. The acquisition time in t_2 was 1.6 s; 200 t_1 increments were recorded, giving a maximum value of t_1 of 2.0 s. (Reproduced from Pell and Keeler, Two-dimensional J -spectra with absorption-mode lineshapes, *J. Magn. Reson.*, **189**, 293–299, Copyright (2007) with permission from Elsevier.¹¹³)

Fluorine NMR profits from the favorable properties of the ^{19}F nucleus (*i.e.* its receptivity being 83% of that of ^1H) and from the interest in the medicinal chemistry community^{118,119} (*i.e.* in improving biological activity and physicochemical properties such as activity enhancement as exhibited by 9α -fluorocortisol,^{16,120} alterations of the metabolic stability of a compound, changing basicity and lipophilicity, and increasing binding affinity). Fluorine may be introduced into a compound during the pharmaceutical lead optimization process to protect the drug against oxidation or dealkylation by cytochrome P450 enzymes.¹²¹ Through changes in basicity, the bioavailability can also be adjusted. The influence of a fluorine substituent on stereochemistry and electron distribution may exercise positive effects on target–ligand binding. Further, recent applications of fluorine as a unique probe for metabolite investigations were described that highlighted the advantage of the absence of interfering signals, since naturally occurring compounds usually lack fluorine substituents.^{122–124}

A 3D F–H heteronuclear TOCSY filtered/edited experiment was used to characterize the individual steroids from a mixture of four dexamethasone (4) derivatives without prior separation.¹²⁵ The corresponding spectral cube is presented in Figure 8.7. Pulse sequences were developed for F–H TOCSY edited H–H TOCSY and F–H TOCSY edited H–C HSQC experiments, which provide sufficient signals for compound identification. Hu *et al.*¹²⁵ nevertheless suggested that the experiment should be combined with LC–MS methods in order to increase confidence. As a drawback, the authors

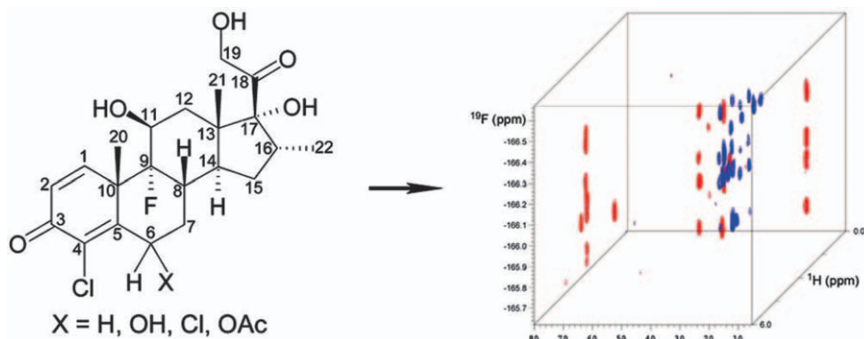


Figure 8.7 3D ^{19}F - ^1H heteronuclear TOCSY-edited, multiplicity-edited ^1H - ^{13}C HSQC. Positive and negative contours are plotted in red and blue, respectively. The spectrum was recorded at 25.0 °C on a 600 MHz spectrometer equipped with a 3 mm ^1H - ^{19}F { ^1H - ^{19}F / ^{13}C } PFG triple-resonance probe and z-gradient accessory. (Reprinted with permission from H. Hu, P. Kulanthaivel and K. Krishnamurthy, *J. Org. Chem.*, 2007, 72, 6259–6262. Copyright (2007) American Chemical Society.¹²⁵)

acknowledge that in cases of a single fluorine substituent, which is often the case with steroids, all detectable magnetization stems from a single fluorine nucleus, thereby resulting in a relatively low sensitivity of the experiment. However, the lack of fluorine in biological matrices may render this experiment attractive for metabolic studies.

In the second recent investigation, the use of a series of fluorine correlation spectra such as F-H and H-F HOESY and F-C HMQC were described for steroid configuration assignments.¹¹⁷ The F-C HMQC for 19,19-difluoro-3 β ,7 β -dihydroxy-androst-5-en-17-one (7) is illustrated as an example in Figure 8.8. Due to the complexity and signal overlap in the ^1H spectra of the fluorinated steroids under consideration, fluorine detection provided a superior and elegant manner to unequivocally assign diastereotopic fluorine nuclei and to determine the configuration of the carbon atoms C4 and C5, to which a cyclopropyl substituent was attached. In order to determine the configuration of the latter difluoro-cyclopropylandrostande derivative, F-H HOESY was applied, since the configuration could not be unequivocally assigned from the synthetic route or from molecular modeling approaches. After the assignment of the ^1H signals to H4, H6, and H10, it could easily be deduced from the HOESY experiment that the fluorine *trans* to H4 was in close proximity to both H10 β and H3 β , and the *cis*-fluorine was in proximity to H4 α and H6 β . Therefore, the cyclopropyl moiety was found situated above the plane and H4 was oriented below the plane. The other stereocenters were defined by the naturally occurring steroid scaffold and the absolute configuration at C4 and C5 is *S* and *R*, respectively. The H-F or F-H HOESY experiments were thus reported as sensitive and simple experiments for use in structure elucidation, if modern NMR instrumentation is available.¹¹⁷ In an extension of that study, full exploitation of the F-C HMQC experiment

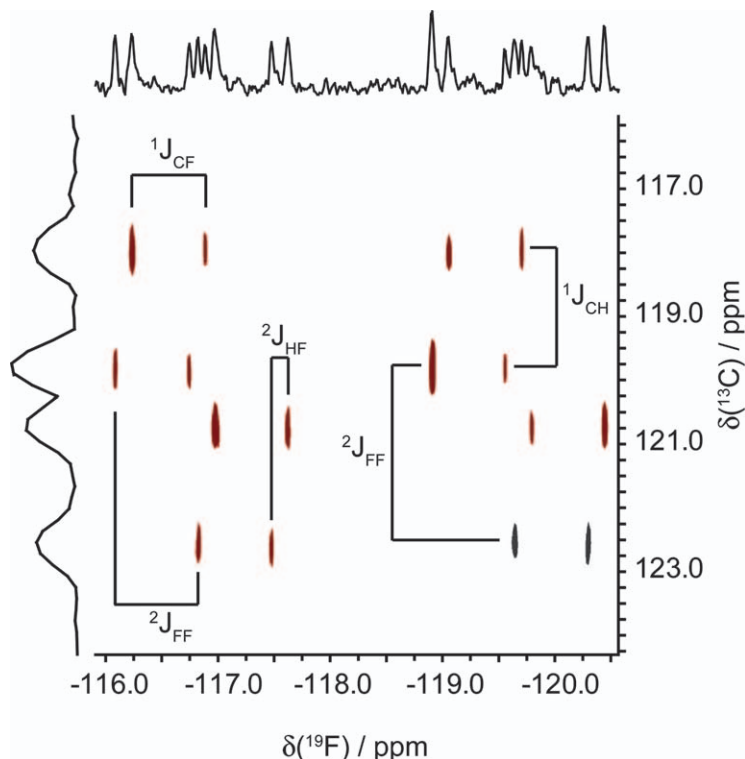


Figure 8.8 F–C HMQC of 19,19-difluoro-3 β ,7 β -dihydroxy-androst-5-en-17-one (**7**) in DMSO- d_6 with couplings assigned.
(Reprinted from Ampt *et al.*¹¹⁷ with permission from John Wiley and Sons. Copyright © 2011 John Wiley & Sons, Ltd.)

towards the determination of the magnitudes and signs of the J_{HF} , J_{CF} , J_{CH} , and J_{FF} couplings was demonstrated (*cf.* Figure 8.9), which may be of potential interest in structure analysis using RDCs.

Another aspect of that investigation was devoted to the determination of homo- and hetero-nuclear coupling constants and their signs.¹²⁶ In the steroidal compound **7**, all fluorine atoms were found to be magnetically and chemically inequivalent, and their F–F coupling constant was 280 Hz. Figure 8.9D shows the expanded region of the fully coupled ^{19}F – ^{13}C HMQC spectrum of compound **7**. As is easily recognized from the figure, an E.COSY-like pattern is displayed. The study further elaborated on the full analysis of the F1–C19 correlation signal. From the theoretically possible 32 cross-peaks, only a reduced multiplet with eight correlations was observed due to E.COSY-type cancellations. A simplified view of how the cross-peaks cancel out is presented in Figure 8.9. The approach to spectral interpretation was based on spin states. Cross-peaks within a given multiplet were thus observed when the passive spin had the same spin state in both dimensions ($\alpha\alpha$ or $\beta\beta$). The sign of a coupling was derived from the slope or tilt of the

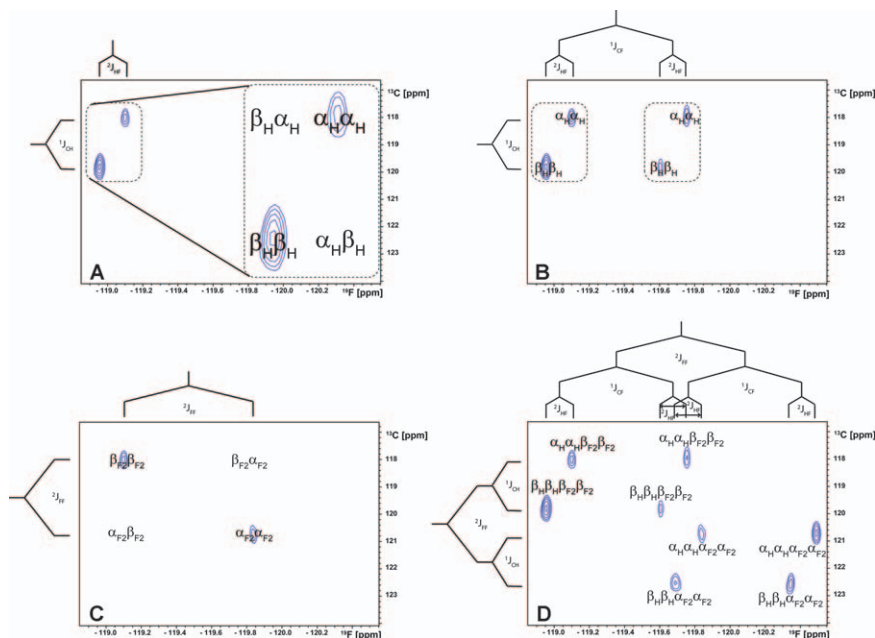


Figure 8.9 Expansion of the correlation between C19 (120.2 ppm) and F1 (−119.6 ppm) of the fully coupled HMQC of 19,19-difluoro-3β,7β-dihydroxy-androst-5-en-17-one (7) with the coupling tree displayed on the axes. (A) Effect of passive spin H (6.22 ppm) on this correlation, where α and β represent the spin state of passive spin H. (B) Effect of the CF coupling in the fluorine dimension. (C) Effect of the homonuclear F1F2 coupling on this correlation, where α and β represent the spin state of F2. (D) Resulting spectrum.

(Reproduced from Ampt *et al.*, Determination of size and sign of heteronuclear coupling constants from 2DF-C correlation spectra, *J. Magn. Reson.*, **215**, 27–33, Copyright (2012) with permission from Elsevier.¹²⁶)

cross-peak pattern. The sizes of the coupling constants $^1J_{CF}$, $^1J_{CH}$, $^2J_{HF}$, and $^2J_{FF}$ were determined from the splittings and the displacements of the peaks.

8.3.7 Residual Dipolar Couplings

RDCs have established themselves as a third pillar of stereochemical structure elucidation¹²⁷ and, with ever-improving experimental ease of handling,¹²⁸ their use will become even more widespread.¹²⁹ Due to the dipolar interaction mechanism, the size of a RDC depends on the length and the angle of the magnetic field and the internuclear vector of the two nuclei that are coupled to each other. The RDC contribution of the observed coupling is superimposed onto the scalar coupling and can be derived as the difference between the scalar coupling and the observed splitting.¹³⁰ To this purpose, the molecule is oriented in a weakly aligning medium, such as bicelles, phages, and polymers,^{131–133} depending on the solubility of the

compound and the desired strength of alignment.¹³⁰ The splitting (*e.g.* H–H, C–H, or N–H) is obtained from suitable *J*-resolved spectra. The scalar coupling may be measured from the same type of spectra recorded in the isotropic phase. The orientation of the molecule with respect to the magnetic field is computed by rotating a structural model (*e.g.* from X-ray diffraction data or molecular modeling) with respect to a magnetic field axis such that the best agreement between calculated and observed RDCs is obtained.^{134,135} Structure elucidation is achieved *via* comparison of a number of models with regard to their matching of calculated and observed RDCs. It is possible to deduce the conformation,^{136,137} the relative configuration of a given compound,^{138–141} and, if a known stereocenter is available as reference, the absolute configuration.^{142,143} Recently, the use of RDCs for molecular constitution elucidation was described.¹⁴⁴ RDCs have great potential for structure elucidation, especially when NOEs are not detectable due to the lack of suitable nuclei in proximity, or if coupling constants cannot be reliably analyzed according to the Karplus or Haasnoot equations. In the case of many natural products, the molecular structures can be proton deficient, making the measurement of NOEs difficult, but steroids are rarely hindered by this issue.

In 2002, RDCs were applied to the configurational analysis of sodium cholate.¹⁴⁵ The extracted values were compared to models of sodium cholate and several of its epimers. The analysis revealed that indeed only the correct epimer yielded a match, thereby allowing the unambiguous assignment of the stereocenters under study.

Kummerloewe *et al.*¹⁴⁶ described how RDCs can be used to distinguish the diastereomers of both 5 α -cholestan-3-one and cholesterol by cross-fitting. Cross-fitting implies the comparison of RDCs computed from structural models of the diastereomers versus their experimentally determined counterparts and vice versa. The method included the transfer of structural information, the alignment tensors, and cross-fitting of the anisotropic parameters of steroids, which have similar overall structures, and therefore allowed for the use of reduced subsets of RDCs. It was also shown that RDCs could be transferred between steroidal compounds in flexible moieties that are sufficiently similar in structure and dynamic behavior. An example of spectra recorded in an alignment medium and isotropic phase is given in Figure 8.10 showing 5 α -cholestan-3-one. For the spectra in the anisotropic phase, 16 mg of the steroid was dissolved in a mixture of polydimethylsiloxane (PDMS) and chloroform and the resulting gel was stretched using a mechanical device. The corresponding isotropic coupling constants were determined in chloroform.

Within a study of testosterone and 17 α -epitestosterone,^{147,148} cross-fitting was performed for the distinction of both epimers. Comparison of experimental and calculated data presented good agreement between both isomers as can be seen in Figure 8.11. On inspection of the quality factors within Table 8.2, it can be seen that only RDCs calculated for the correct epimer match the corresponding experimental data. Scalar coupling

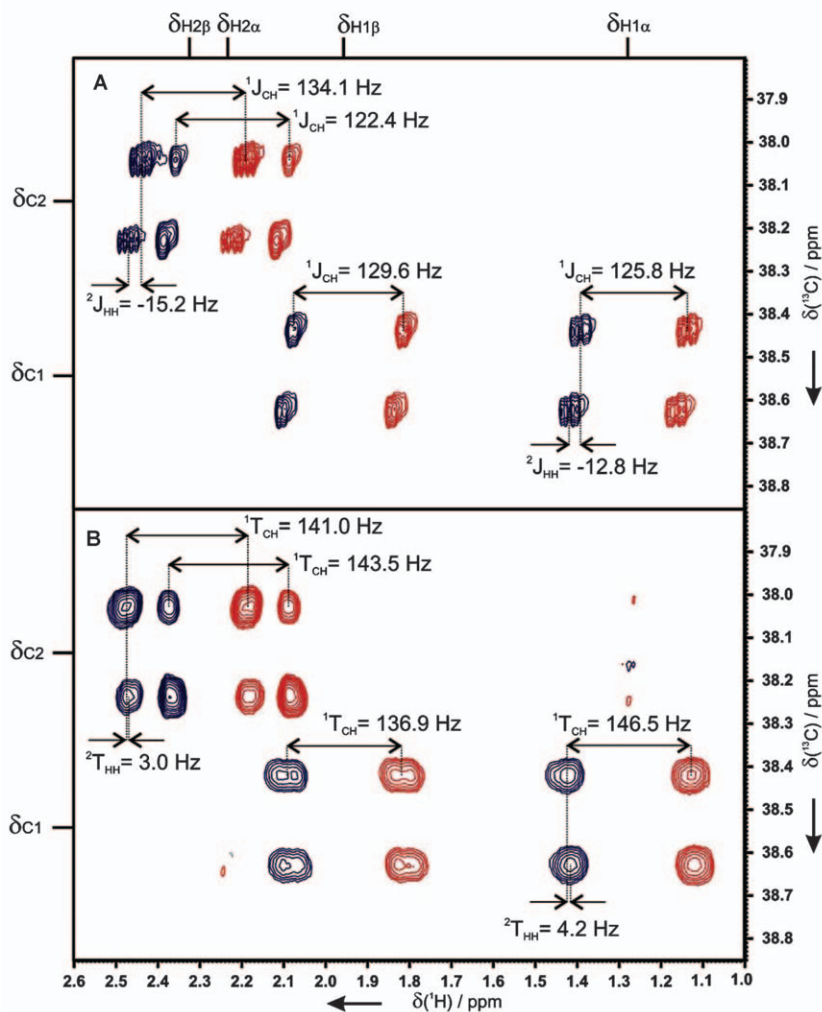


Figure 8.10 Details of P.E.-HSQC spectra acquired on 5 α -cholestan-3-one in a chloroform solution (A) and in a stretched PDMS/CDCl₃ gel (B), showing the multiplets of the methylene groups at C1 and C2. One-bond C–H couplings and geminal H–H couplings are assigned in the spectra. Note that the sign information of the homonuclear couplings is given by the tilt of the multiplet and that this tilt changes upon alignment in stretched PDMS gel as the negative $^2J_{HH}$ couplings (A) are compensated by the larger positive D_{HH} couplings of both methylene groups in the aligned spectrum (B). (Reprinted with permission from Kummerloewe *et al.*¹⁴⁶ as published in *Open Spectrosc. J.*, 2010.)

constants were determined in an isotropic solution of chloroform. The corresponding RDCs were derived from aligned samples using poly- γ -benzyl-L-glutamate and chloroform. In contrast to the weaker alignment caused by

Table 8.2 Cross-fitting of RDC data obtained for testosterone and epitestosterone with the number of RDCs used for fitting being 23 for testosterone 18 for epitestosterone. The abbreviations RMS, n/χ^2 , and r are used following the common statistical conventions.

RDC input	Model testosterone	Model epitestosterone
Testosterone	RMS = 3.289 $n/\chi^2 = 0.350$ $r = 0.993$	RMS = 15.871 $n/\chi^2 = 0.020$ $r = 0.820$
Epitestosterone	RMS = 15.101 $n/\chi^2 = 0.016$ $r = 0.740$	RMS = 3.649 $n/\chi^2 = 0.242$ $r = 0.984$

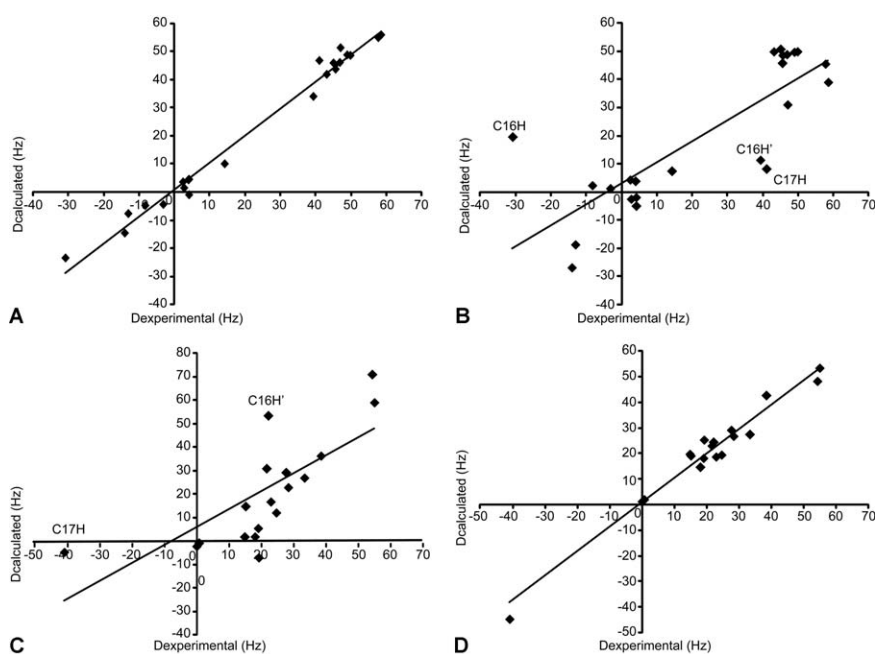


Figure 8.11 Scatterplots of experimental versus calculated RDC values, $D_{\text{experimental}}$ vs. $D_{\text{calculated}}$: testosterone RDCs against a molecular model of testosterone (A) and of epitestosterone (B); epitestosterone RDCs against a model of testosterone (C) and of *epi*-testosterone (D).^{147,148}

PDMS gels, polyglutamates generally induce stronger alignment, which in turn leads to larger RDCs. As a consequence, higher-order effects often hamper the interpretation of proton-detected spectra. The recording of ^{13}C -detected spectra provides a suitable remedy. In the study described, 1D ^{13}C - and carbon-detected J -resolved experiments were used. From Figure 8.11, it is easy to observe that the largest deviations between experimental and calculated RDCs occurred for the stereocenter at C17 and the adjacent C16, indicating the specificity of the method. This observation is in

line with the finding of Kummerloewe *et al.* that a reduced dataset can be used for similar compounds. It is remarkable that even for stereocenters in the rather flexible D ring of steroids, RDCs allowed not only the distinction of two epimers when both epimers were studied, but also the absolute configuration assignment. The quality of the fit is superior for the correct stereoisomer, such that the configurations of the steroids can be elucidated on the basis of a single epimer. A stereomodel obtained from molecular modeling proved sufficient for configuration investigations within this study. Both conclusions are in agreement with the earlier findings for sodium cholate.¹⁴⁵

8.4 Considerations

The structure elucidation of steroids by NMR using the set of experiments^{72,74,76} available today is mostly a routine task of moderate complexity, while in the 1980s, assignment strategies were on the basis of the then new 2D techniques and were heavily discussed for steroids.^{45,65,149,150} The variety of currently available experiments provides an ample choice for the most suitable route to assign all signals and thus perform a full structure elucidation. The task has been compared to solving an over-determined matrix.¹⁵¹ Straightforward signal assignment is usually hampered only in the case of strong signal crowding or overlap as for tibolone (**3**)⁷⁰ or in cases of limited availability of experimental time, which may lead to parameter choices resulting in insufficient resolution in 2D datasets or a reduced set of spectral data. As a second cause of more difficult spectral interpretation, the availability of only submicrogram amounts of sample may be considered (e.g. metabolites or natural product isolates).

Contemporary structure elucidation needs to take two aspects into account, provided the steroid skeleton itself does not deviate totally from its natural chemical constitution and configuration. This condition very often holds true, although a few exceptions are known, such as the completely inverted *ent*-19-nortestosterone,¹⁵² unnaturally configured 13-episteroids,^{152,153} or the brassinosteroids that were discovered as late as 1979.^{154,155} The first question is: what spectra or experiments are needed to assign all signals (*i.e.* mostly from ¹H and ¹³C and, more rarely, ¹⁹F) and to establish the connectivity and describe the stereochemistry with respect to conformation and configuration? In principle, this question refers to *de novo* structure elucidation. Secondly, what spectra or experiments constitute the minimal set in terms of type and recording time needed to test a structural hypothesis? This aspect refers to the verification or rebuttal of a structure proposal, be it from chemical synthesis or purification or lead identification after chemical library screening.

From the previous sections, it has been concluded that the principal aspects of steroid structure to be characterized are the signal assignment of ¹H and ¹³C in order to describe skeleton and substituents, establish atom-to-atom connectivity, identify α and β positions of substituents including

hydrogen isotopes, and recognize conformational preferences or flexibility. When a structure proposal is to be verified, these attributes may be reduced to a certain part or moiety of the steroidal compound. A somewhat different task is the identification of a steroid in total based on a reduced dataset.

Within the structure elucidation protocol, the determination of the molecular formula should be most conveniently left to high-resolution mass spectrometry from a few nano- or micro-grams. It should not be forgotten that the number of double bonds can be derived from MS data as well.¹⁵⁶ The identification of heteroatoms by proton and carbon NMR is an indirect method and might be supported by directly detected fluorine or phosphorus NMR spectra, but requires more work when compared to extracting this information from accurate mass data. After determining the molecular formula, the acquisition of NMR spectra usually starts with a 1D ^1H spectrum, while a 1D ^{13}C spectrum is often omitted for sensitivity reasons. Yet the authors consider a composite-pulse decoupled 1D ^{13}C spectrum very valuable due to its superior resolution and its independence of coupling constants, which might lead to unexpected or missing responses (*e.g.* in alkynyl or fluorinated steroids). When modern cryogenic probe technology allowing high-sensitivity carbon direct detection is available, 1D ^{13}C spectra can be recorded within an hour on a sample of several hundred micrograms (see Volume 1, Chapter 3). As a next step, the easiest link between ^1H and ^{13}C spectra may be achieved *via* an ^1H - ^{13}C HSQC experiment. In our experience, gradient multiplicity-edited HSQC provides excellent results with good sensitivity and delivers the ^1H - ^{13}C correlations as well as all carbon multiplicities when compared to the 1D ^{13}C spectrum. At this point, signal assignment can start if a structure hypothesis and expert knowledge are available. This knowledge might stem from a human expert, from electronic or paper databases, or from chemical shift predictions of any kind. Rules of thumb as given in Table 8.3 provide good starting points.

Prior to recording steroid spectra, the operator should consider spending time on parameter optimization. A few considerations will greatly enhance spectral quality as compared to standard set parameters within sample series of diverse compounds. In the proton dimension of 2D experiments,

Table 8.3 ^{13}C chemical shifts of methyl, methylene, and methine moieties and H-H coupling constants of steroids.

Fragment	$\delta(^{13}\text{C})/\text{ppm}$	Fragment with H-H connectivity	$^nJ_{\text{HH}}/\text{Hz}$
Primary alkyl carbons	12–24	–CH ₂ – (geminal)	–14 to –12 Hz
Secondary alkyl carbons	20–21	–CH–CH– (vicinal)	
Tertiary alkyl carbons	35–57	Axial–axial	10–14 Hz
Quaternary carbons	65–91	Axial–equatorial	3.5–5 Hz
Olefinic carbons	119–172	Equatorial–equatorial	2.5–4 Hz
Carbonyl carbons	177–220	Five-membered ring	Depending on conformation
Carbons with fluorine	88–102		

the size of the smallest coupling constant desired usually determines the necessary resolution. For example, hydrocortisones display long-range W-couplings of 0.8 Hz and M-couplings of 0.2 Hz due to the rings' trans-arrangement. These 4J W-pathways exist from Me18 to H12 β and H12 α , and from Me19 to H1 α . Similarly, 5J pathways are often observed from Me19 to H2 β . For the carbon dimension, overlap or similar shifts need to be taken into account. Spectral congestion in the carbon domain is, however, of lesser concern due to the reduction of signal density by the second dimension. Nevertheless, if, for example, methyl resonances exhibit severe overlap, the ^{13}C - ^1H HETCOR equipped with a BIRD filter has been shown to provide the highest resolution for CH resonances (0.001 ppm in each frequency dimension) and is proven in the specific case of steroids.^{157,158} For a variety of small molecules, the resolution of a H-H COSY in magnitude mode was suggested to be 4 Hz per point,⁷² since the maximum intensity is obtained if the resolution equals J , but can *in praxi* be only 20% the size of the splitting. For steroids, the presence of multiple long-range couplings demands higher resolution, and this is easily achieved with modern spectrometers.

With regard to stereochemistry, it might be safely assumed that equatorial protons usually resonate at lower fields than axial protons. For the concordance of α and β to axial and equatorial protons, a model proves helpful. Nevertheless, there are several ways to establish direct proton-carbon bonding and multiplicity. The connectivity of the CH_n fragments can be visualized *via* the proton network through-bond or through-space, *via* the carbon network, or *via* a combination of these methods. The first method is easily addressed by an H-H COSY-type experiment. When an H-H TOCSY experiment is used instead of COSY, as might be the case in routine sample queues, some thought should be given to the number of signals expected for steroids. The total number of correlations in a TOCSY spectrum of a steroid is usually high, especially for longer mixing times. Because of the significant number of long-range H-H correlations, a TOCSY experiment with mixing times above 25 ms may already result in a high number of correlation peaks.⁷⁰ Hence, correlation times of 18–20 ms are usually sufficient. The common analysis of the proton pathway often follows the generic connectivity diagrams described for norethisterone and 5α -androstane.^{4,151} For the assignment, a connectivity and through-space proximity matrix proved useful. Examples of such correlation matrices are presented for dexamethasone and tibolone in Figure 8.12.

Alternatively, the proton network can be mapped by NOESY spectroscopy. Since the ring juncture stereochemistry can vary with the type of steroid (*e.g.* 5α versus 5β , or 14α versus 14β , nor- or *epi*-steroid), standard correlation diagrams often refer to the 5α -androstane as the generic skeleton. However, the NOE parameters have to be chosen carefully in order not to excite spin diffusion and create transferred NOEs or give rise to COSY or TOCSY artifacts for NOESY or ROESY spectra, which may cause contradictions, especially during *ab initio* structure elucidation.

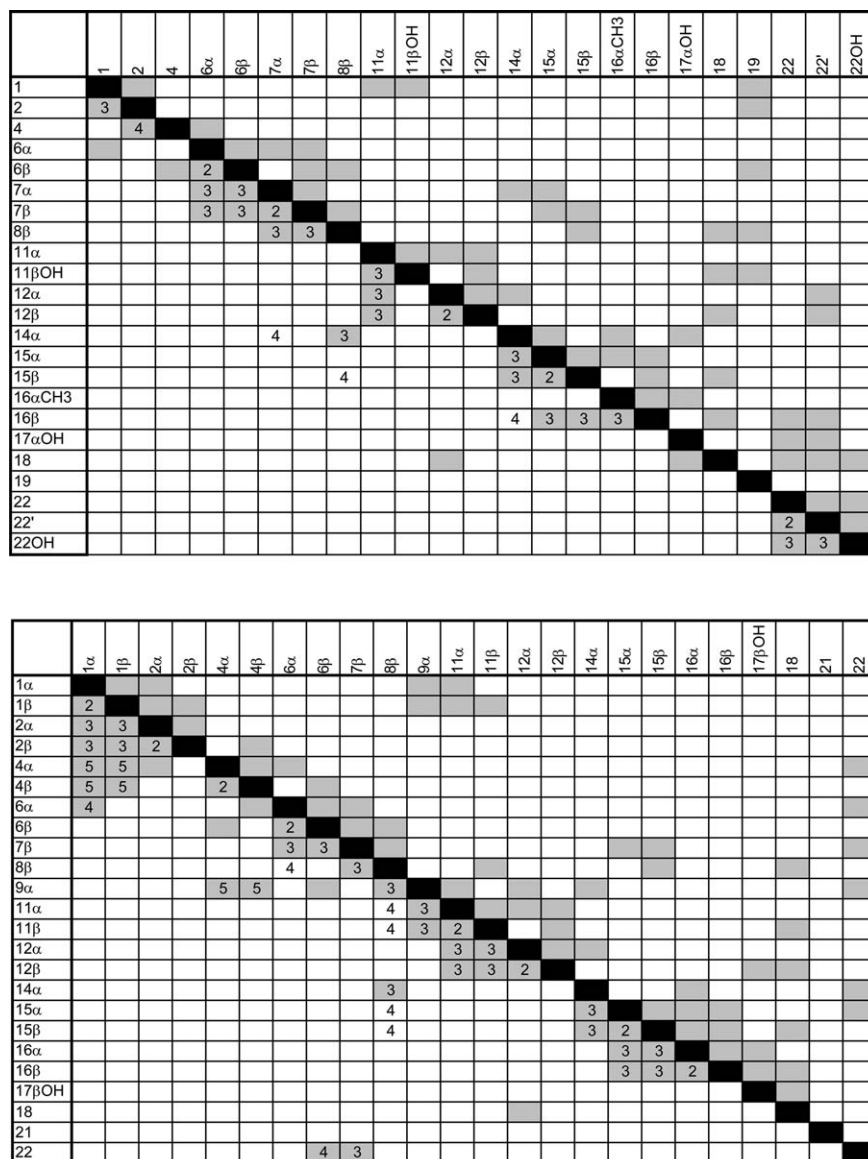


Figure 8.12 COSY, TOCSY connectivity, and NOESY correlation matrices for dexamethasone (4) (top) and tibolone (3) (bottom). The NOESY correlations are marked above the diagonal, COSY and TOCSY correlations below. Numbers indicate the type of J -coupling. Mixing times were set to 18 ms for TOCSY and 500 ms for NOESY experiments. (Reproduced from Jaeger and Aspers, *Steroids and NMR*, *Annu. Rep. NMR Spectrosc.*, 77, 115–285 Copyright (2012) with permission from Elsevier.³⁷)

A pure carbon network experiment is of course ^{13}C – ^{13}C INADEQUATE. In recent example studies, by using cryoprobe technology at 600 MHz, high-quality spectra were obtained on 10–20 mg samples within approximately 24 hours of recording time.^{37,70,107} The value of INADEQUATE information is irrefutable and will not be further discussed. An example is given in Figure 8.13 for tibolone (3).

Among the mixed nuclei pathways, the inverse-detected (*i.e.* ^1H -detected) experiments have proven their advantage over their heteronuclear-detected variants, although the latter may have the advantage of higher resolution.⁷²

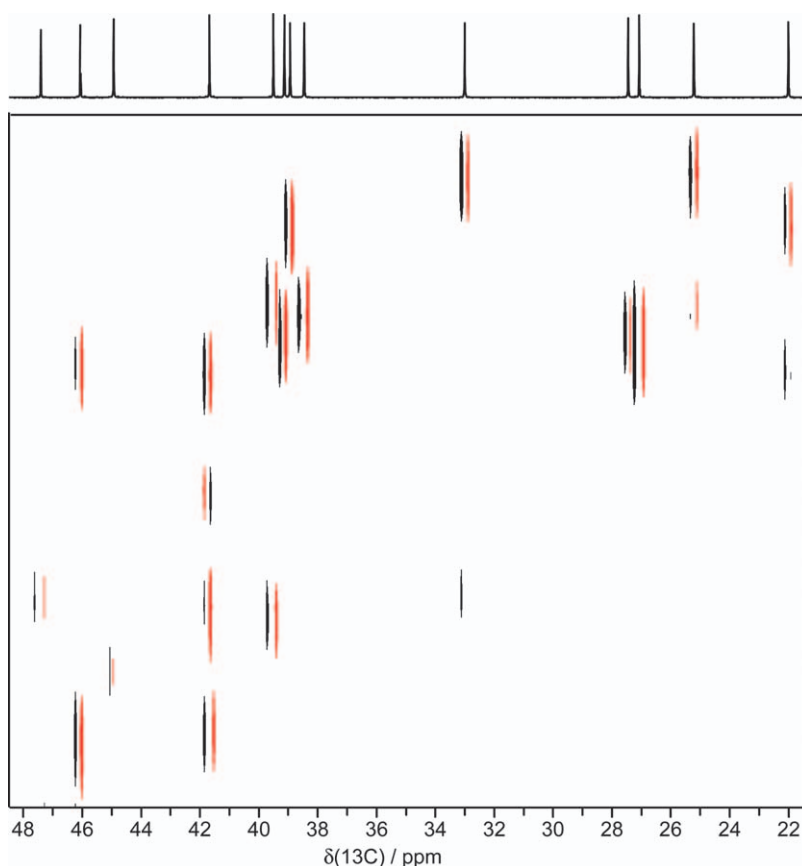


Figure 8.13 ^{13}C – ^{13}C INADEQUATE of 250 mg tibolone (3) in 700 μl CDCl_3 recorded on a Bruker DPX 400 FT-NMR spectrometer operating at a 400 MHz ^1H frequency and equipped with a PABBO 5 mm room-temperature broadband probe. The delay was set to 50 Hz for J_{CC} . The experiment time amounted to 23 hours for 256 accumulations on 64 increments in the indirect dimension and a relaxation delay of 5 seconds. (Reproduced from Jaeger and Aspers, *Steroids and NMR, Annu. Rep. NMR Spectrosc.*, 77, 115–285 Copyright (2012) with permission from Elsevier.³⁷)

For steroids, a lack of resolution may be compensated for by extracting the necessary information from another experiment. Still, the heteronuclear-detected experiments may experience a revival if dynamic nuclear polarization techniques^{159,160} become more widespread, since the sensitivity gain should be larger on the time scale of carbon detection than of proton detection due to the longer carbon relaxation times.

Among the inverse-detected long-range correlation experiments, H-C HMBC is the most commonly used despite its disadvantages with regard to distinguishing $^2J_{CH}$ and $^3J_{CH}$ couplings, a situation that can be found for unsaturated steroids (e.g. A ring correlations in estradiol). More recently, H-C H2BC spectra were proposed as a remedy.¹⁶¹ Spectral editing was subsequently included in order to reduce overlap (e.g. in the spectra of prednisolone).¹⁶² Only a few more reports applying H2BC to steroids have been published, among them the analysis of saponins,¹⁶³ sapogenins,¹⁶⁴ and bile acids.^{165,166} For tibolone, the assignment strategy using HSQC and H2BC is illustrated in Figure 8.14.

In the discussion of spectra that are suitable for computer-assisted structure elucidation systems (see Volume 1, Chapter 9), 1,1-ADEQUATE spectra were reinvestigated as a remedy to the distinction problem for two and three bonds.^{167,168}

As can be seen from Figure 8.15A, only $^nJ_{CH}$ correlations were observed for tibolone. Again, the most recent cryoprobe technology allows the recording of a 1,1-ADEQUATE spectrum within a few hours on less than 10 mg of sample. A common feature of HMBC and ADEQUATE experiments is that reasonable results can be obtained for steroids without the necessity of individual optimization of parameters, such as delays corresponding to C-H couplings. Both H2BC and 1,1-ADEQUATE are expected to have the greatest potential for significant contributions to problem-oriented as well as automated structure analysis.

After the molecular constitution is established, the stereochemistry has to be addressed. H-H NOESY or H-H ROESY experiments are presently the means of choice. In the case of fluorinated steroids, H-F or F-H HOESY experiments have been applied successfully¹¹⁷ for the fast assignment of diastereotopic fluorines and to determine the configuration of two carbon atoms bearing a cyclopropyl substituent. In terms of H-H Overhauser spectroscopy, the NOESY experiment usually works fine for steroids in chloroform solution at 400 and 600 MHz based on the current authors' experience. For steroids with large substituents, such as cholestane and ergostane, field strengths of 400 MHz proved preferable. Although the theoretical advantages of ROESY are well known, this type of experiment has rarely been used in the current authors' laboratories. Based on our experience, the gradient version of H-H NOESY, in most cases of routine application, gave reliable results with a better signal-to-noise ratio than the ROESY experiment.

In analogy to the COSY pathway, the stereochemical assignment and, as a consequence, the stereochemical characterization starts from a distinct

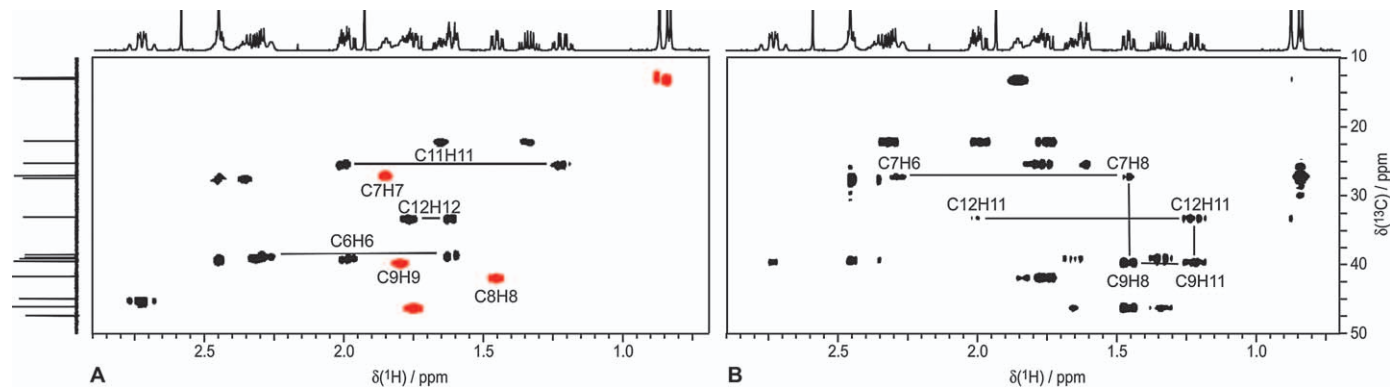


Figure 8.14 H-C correlation spectra of 2.5 mg of tibolone (3) in 0.65 ml CDCl_3 recorded in a 5 mm quartz tube at 298 K at a 600 MHz proton frequency using an H,C-sensitive cryogenic probe. Each experiment was acquired within 43 minutes. (A) Phase-sensitive H-C HSQC with four accumulations on 256 increments in the indirect dimension using a relaxation delay of 1.5 seconds. (B) H-C H2BC with four accumulations on 256 increments in the indirect dimension using a 1.5 second relaxation delay. For the J_{CH} filter, a range from 125 to 165 Hz was chosen, 18 ms was given for the constant time evolution for J_{HH} . (Reproduced from Jaeger and Aspers, *Steroids and NMR, Annu. Rep. NMR Spectrosc.*, 77, 115–285 Copyright (2012) with permission from Elsevier.³⁷)

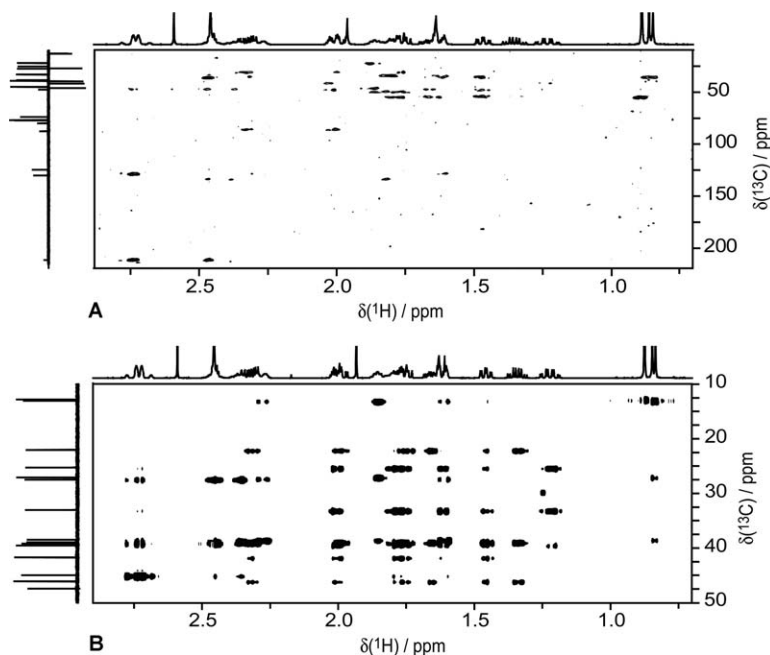


Figure 8.15 H-C correlation spectra of 11.50 mg of tibolone (**3**) in 0.7 ml CDCl_3 recorded in a 5 mm quartz tube at 298 K at a 500 MHz proton frequency using an H_C -sensitive cryogenic probe. (A) 1,1-ADEQUATE spectrum with 64 accumulations for 128 increments in the indirect dimension. The relaxation delay was set to 2 seconds and the delays for coupling were chosen for $J_{\text{CH}} = 135$ Hz and $J_{\text{CC}} = 43$ Hz. The resulting experimental time was 5 hours. (B) H-C HSQC-TOCSY with eight accumulations for 256 increments in the indirect dimension. The relaxation delay was set to 2 seconds, and for the mixing time 60 ms was chosen. The experiment time amounted to 80 minutes. (Reproduced from Jaeger and Aspers, *Steroids and NMR*, *Annu. Rep. NMR Spectrosc.*, 77, 115–285 Copyright (2012) with permission from Elsevier.³⁷)

proton signal such as H8, H9, H14, CH_3 -18, CH_3 -19, or a proton geminal to a hydroxyl substituent of a known configuration. As the reader may expect, this method is very robust. Only signal overlap or insufficient resolution can obstruct the straightforward assignment. The protons at C16 in testosterone represent illustrative examples.^{68–70} When increasing the spectral resolution does not resolve the signal overlap, then other experiments, such as hyphenated pulse sequences, need to be performed. This combination of pulse sequences such as HSQC and NOESY or HSQC and TOCSY allows for the introduction of additional resolution *via* heteronuclear editing, even if the spectra are recorded in 2D variants.^{74,169} The application of a HSQC-NOESY

in 2D and 3D modes to the assignment of H16 α and H16 β in testosterone was demonstrated by Aspers *et al.*⁷⁰

Since NOE data are easily generated with the NMR spectrometers that have been produced within the last two decades, detailed analyses of coupling constants for qualitative assignment purposes seem to have become somewhat redundant in the steroid field, partially as a result of the rigidity of the steroid skeleton. This is in contrast to the 1980s, when *J*-resolved spectroscopy was an important tool for structure elucidation. In conclusion, today's standard set of experiments shall prove sufficient for structure analysis of steroids.

What experiments should now be considered to be the spectroscopic minimum? A high-resolution mass spectrum and a 1D ^1H NMR are both indisputable. Both H–H COSY and H–C HSQC require shorter experiment times than 1D ^{13}C , but present only a selection of signals, whereas H–H COSY reflects the connectivity network of protons over the steroid scaffold, and HSQC provides the necessary information over the carbon atoms without the quaternary carbons. In principle, COSY information may be read out of the 1D ^1H , but this is difficult for steroids since strong signal crowding is observed in the proton-only domain. Clearly, ^{13}C – ^{13}C INADEQUATE, 1,1-ADEQUATE, H–C HMBC, and any 3D spectroscopy are too time consuming or otherwise too insensitive to be considered as fast, routine applications in a high-throughput mode. As a lot of the structural information contained within the HSQC data is contained in the chemical shifts, expert knowledge either in the form of a human expert or of computerized data needs to be combined with the spectroscopic data that have been recorded in order to achieve structure confirmation. An experiment worth considering would be H–C HSQC-TOCSY. Proton connectivity information and carbon shift information is combined in this hyphenated experiment. The TOCSY part should be recorded with short mixing times (*e.g.* around 20 ms) to avoid signal crowding due to long-range couplings transferring magnetization too widely over the steroid scaffold. A HSQC-TOCSY is shown for tibolone in Figure 8.15B. The spectrum was recorded with a mixing time of 60 ms. While the correlations appear well separated, almost all signals are related to each other and consequently do not provide much help for interpretation when compared to spectra recorded with shorter mixing times. In our experience, recording times of 1.5–4.5 hours are needed on a few milligrams of a steroid using a first-generation 5 mm proton-selective cold probe.⁷⁰ Yet indirect covariance transformation processing techniques⁹⁵ (see Volume 1, Chapter 10) can generate this type of spectrum within a few seconds from COSY or TOCSY and HSQC spectra.⁹⁰ It has been found that, due to the higher sensitivity of the individual experiments as compared to the hyphenated sequence, the signal-to-noise ratio of the computed spectrum is superior to the experimental one.^{70,90,98} A comparison of experimental and UIC-computed HSQC-TOCSY spectra of allylestrenol (5) is presented in Figure 8.16. It therefore seems worth trying to base the structure confirmation of steroids on a H–C HSQC-TOCSY dataset.

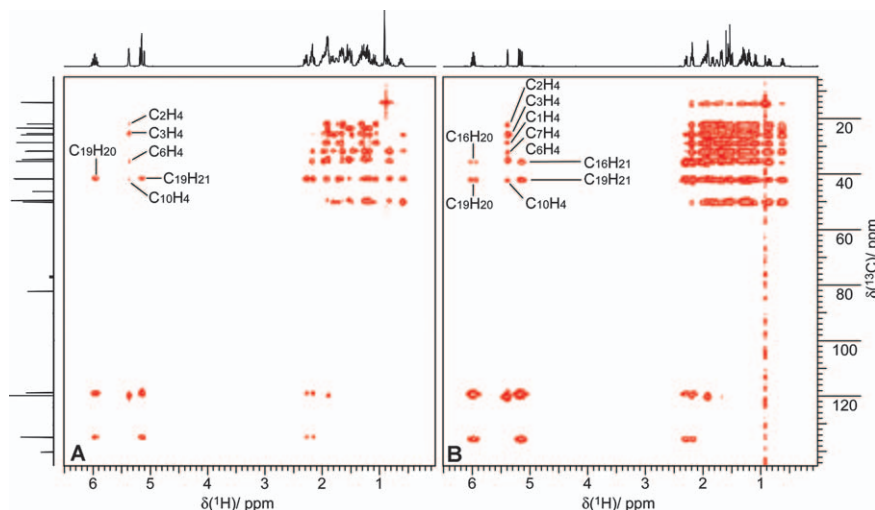


Figure 8.16 H-C HSQC-TOCSY spectra of allylestrenol (**5**) in CDCl_3 . (A) Experimental spectrum; (B) spectrum computed from H-H TOCSY and H-C HSQC by unsymmetrical indirect covariance transformation. (Adapted from Aspers *et al.* with permission from John Wiley and Sons. Copyright © 2011 John Wiley & Sons, Ltd.⁷⁰)

8.5 Conclusions

Philosophically, one might contemplate whether the presence of a correlation in a spectrum that is due to a through-bond coupling is considered a direct proof of a chemical bond.¹⁷⁰ Perhaps the correlation that originates from a mathematical treatment of spectra should be considered as a proof of that bond, too? A structural problem arising from a synthetic modification may be answered by other methods more specifically, such as COSY, NOESY, or HMBC for the introduction of a double bond, for a methylation in the α or β position, or an acetylation of a hydroxyl group, respectively. As was illustrated above, gradient-selective 1D experiments are powerful tools for addressing this specific type of issue since they provide the high resolution of 1D NMR spectra. For the recognition of the aromatization of the A ring, a simple 1D ^1H might be sufficient. In metabolite investigations of steroids, COSY and TOCSY experiments play the most prominent role in identifying the position and configuration of the hydroxyl group, whereas for metabolite elucidation of small molecules in general, HSQC has probably become more important today.

In conclusion, as with any other small molecule, steroids may be investigated using general strategies regarding how to verify structural hypotheses in the quickest and most generic manner. It is up to the spectroscopist's imagination and experience to decide which experiment will yield the most straightforward access to the structural information that is sought.

Acknowledgements

The authors are gratefully indebted to their former colleagues from N. V. Organon, Schering-Plough and Merck & Co. for many fruitful discussions. They are especially thankful to L. Fielding, P. J. A. v. Tilborg, F. M. Kaspersen, G. N. Wagenaars, M. Honing, M. A. Lunenburg, C. A. A. v. Boeckel, F. T. L. Brands, and G. E. Martin. Their heartfelt gratitude is due to Marinus B. Groen for sharing his experience and providing the 40 MHz spectrum of 4,4-dimethyl- $\Delta^{5,10}$ -3-ceto-estrenol signed by J.-M. Lehn.

References

1. D. N. Kirk, *Nat. Prod. Rep.*, 1986, **3**, 505–513.
2. D. N. Kirk, *Nat. Prod. Rep.*, 1989, **6**, 393–404.
3. D. N. Kirk, S. J. Gaskell and B. A. Marples, *Steroid Anal.*, 1995, 25–113.
4. A. Kasal, M. Budesinsky and W. J. Griffiths, Spectroscopic Methods of Steroid Analysis, in *Steroid Analysis*, ed. H. L. J. Makin and D. B. Gower, Springer, Netherlands, Dordrecht, Heidelberg, London, New York, 2010, pp. 27–161.
5. W. R. Croasmun and R. M. K. Carlson, *Two-Dimensional NMR Spectroscopy. Applications for Chemists and Biochemists*, VCH, Weinheim, 1st edn, 1987 and 2nd expanded and updated ed. VCH Publishers, Weinheim, 1994.
6. L. F. Fieser and M. Fieser, *Steroids*, Reinhold Publishing Corporation, New York, 1959.
7. C. Djerassi, *Steroids Made It Possible (Profiles, Pathways, and Dreams: Autobiographies of Eminent Chemists)*, American Chemical Society, Washington, DC, 1990.
8. C. Djerassi, *The Pill, Pygmy Chimps, and Degas' Horse*, Basic Books, New York, 1992.
9. *Biochemical Pathways (Poster)* Boehringer Mannheim.
10. *Human Steroidogenesis* Wikipedia <http://en.wikipedia.org/wiki/Steroid>.
11. S. Bernstein, *Steroids*, 1992, **57**, 392–402.
12. A. J. Birch, *Steroids*, 1992, **57**, 363–377.
13. K. Bloch, *Steroids*, 1992, **57**, 378–383.
14. F. B. Colton, *Steroids*, 1992, **57**, 624–630.
15. C. Djerassi, *Steroids*, 1992, **57**, 631–641.
16. J. Fried, *Steroids*, 1992, **57**, 384–391.
17. H. Herzog and E. P. Oliveto, *Steroids*, 1992, **57**, 617–623.
18. R. Hirschmann, *Steroids*, 1992, **57**, 579–592.
19. J. A. Hogg, *Steroids*, 1992, **57**, 593–616.
20. E. R. H. Jones, *Steroids*, 1992, **57**, 357–362.
21. P. A. Lehmann, *Steroids*, 1992, **57**, 403–408.
22. E. Mappus, M. Renaud, M. Rolland de Ravel, C. Grenot and C. Y. Cuilleron, *Steroids*, 1992, **57**, 122–134.
23. K. Nakanishi, *Steroids*, 1992, **57**, 649–657.

24. G. Rosenkranz, *Steroids*, 1992, **57**, 409–418.
25. A. Zaffaroni, *Steroids*, 1992, **57**, 642–648.
26. K. Heusler and J. Kalvoda, *Steroids*, 1996, **61**, 492–503.
27. S. A. Szpilfogel and F. J. Zeelen, *Steroids*, 1996, **61**, 483–491.
28. H. Wieland and E. Dane, *Z. Physiol. Chem.*, 1933, **216**, 91–104.
29. H. Wieland, E. Dane and C. Martius, *Z. Physiol. Chem.*, 1933, **215**, 15–24.
30. A. Butenandt, *Nature*, 1932, **130**, 238.
31. A. Butenandt, H. A. Weidlich and H. Thompson, *Ber. Dtsch. Chem. Ges. (A and B Series)*, 1933, **66**, 601–604.
32. A. Butenandt and U. Westphal, *Z. Physiol. Chem.*, 1934, **223**, 147–168.
33. K. H. Slotta, H. Ruschig and E. Fels, *Ber. Dtsch. Chem. Ges. (A and B Series)*, 1934, **67**, 1624–1626.
34. O. Wintersteiner and W. M. Allen, *J. Biol. Chem.*, 1934, **107**, 337–350.
35. T. Reichstein, *Helv. Chim. Acta*, 1936, **19**, 29–63.
36. J. N. Shoolery and M. T. Rogers, *J. Am. Chem. Soc.*, 1958, **80**, 5121–5135.
37. M. Jaeger and R. L. E. G. Aspers, *Annu. Rep. NMR Spectrosc.*, 2012, **77**, 115–258.
38. W. E. Rosen, J. B. Ziegler, A. C. Schabica and J. N. Shoolery, *J. Am. Chem. Soc.*, 1959, **81**, 1687–1689.
39. R. F. Zuercher, *Helv. Chim. Acta*, 1961, **44**, 1380–1395.
40. R. F. Zuercher, *Helv. Chim. Acta*, 1963, **46**, 2054–2088.
41. H. J. Reich, M. Jautelat, M. T. Messe, F. J. Weigert and J. D. Roberts, *J. Am. Chem. Soc.*, 1969, **91**, 7445–7454.
42. H. Eggert and C. Djerassi, *J. Org. Chem.*, 1973, **38**, 3788–3792.
43. I. Rubinstein, L. J. Goad, A. D. H. Clague and L. J. Mulheirn, *Phytochemistry*, 1976, **15**, 195–200.
44. M. J. Thompson, S. R. Dutky, G. W. Patterson and E. L. Golden, *Phytochemistry*, 1972, **11**, 1781–1790.
45. P. K. Agrawal, D. C. Jain, R. K. Gupta and R. S. Thakur, *Phytochemistry*, 1985, **24**, 2479–2496.
46. J. W. Blunt and J. B. Stothers, *Org. Magn. Reson.*, 1977, **9**, 439–464.
47. L. D. Hall and J. K. M. Sanders, *J. Chem. Soc., Chem. Commun.*, 1980, 368–370.
48. R. D. Farrant, D. N. Kirk, J. D. Merish and J. K. M. Sanders, *J. Steroid Biochem.*, 1983, **19**, 181–187.
49. L. D. Hall and J. K. M. Sanders, *J. Org. Chem.*, 1981, **46**, 1132–1138.
50. H. J. Schneider, U. Buchheit, N. Becker, G. Schmidt and U. Siehl, *J. Am. Chem. Soc.*, 1985, **107**, 7027–7039.
51. A. G. J. Sedee, G. M. J. Beijersbergen van Henegouwen, M. E. De Vries and C. Erkelens, *Steroids*, 1985, **45**, 101–118.
52. H. Duddeck, D. Rosenbaum, M. Hani, A. Elgamal and M. B. E. Fayez, *Magn. Reson. Chem.*, 1986, **24**, 999–1003.
53. A. G. J. Sedee, G. M. J. Beijersbergen van Henegouwen, W. Guijt and C. A. G. Haasnoot, *Pharm. Weekbl., Sci. Ed.*, 1985, **7**, 202–207.

54. G. E. Martin, J. E. Guido, R. H. Robins, M. H. M. Sharaf, P. L. Schiff, Jr. and A. N. Tackie, *J. Nat. Prod.*, 1998, **61**, 555–559.
55. T. Parella, *Magn. Reson. Chem.*, 1998, **36**, 467–495.
56. G. E. Martin, *Annu. Rep. NMR Spectrosc.*, 2005, **56**, 1–96.
57. G. E. Martin, *Encycl. Nucl. Magn. Reson.*, 2002, **9**, 98–112.
58. H. Kovacs, D. Moskau and M. Spraul, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2005, **46**, 131–155.
59. G. E. Martin, *Encycl. Nucl. Magn. Reson.*, 2002, **9**, 33–35.
60. A. Bax, *Two-dimensional Nuclear Magnetic Resonance in Liquids*, D. Reidel Publishing Company, Dordrecht, Holland, 1982.
61. M. Karplus, *J. Chem. Phys.*, 1959, **30**, 11–15.
62. C. A. G. Haasnoot, F. A. A. M. de Leeuw and C. Altona, *Tetrahedron*, 1980, **36**, 2783–2792.
63. K. Marat, J. F. Templeton, R. K. Gupta and V. P. S. Kumar, *Magn. Reson. Chem.*, 1987, **25**, 730–733.
64. K. Marat, J. F. Templeton and V. P. S. Kumar, *Magn. Reson. Chem.*, 1987, **25**, 25–30.
65. W. R. Croasmun and R. M. K. Carlson, *Methods Stereochem. Anal.*, 1987, **9**, 387–424.
66. L. L. Smith, *Steroids*, 1964, **4**, 395–414.
67. P. K. Agrawal, D. C. Jain and A. K. Pathak, *Magn. Reson. Chem.*, 1995, **33**, 923–953.
68. K. Hayamizu, T. Ishii, M. Yanagisawa and O. Kamo, *Magn. Reson. Chem.*, 1990, **28**, 250–256.
69. D. N. Kirk, H. C. Toms, C. Douglas, K. A. White, K. E. Smith, S. Latif and R. W. P. Hubbard, *J. Chem. Soc., Perkin Trans. 2*, 1990, 1567–1594.
70. R. L. E. G. Aspers, P. E. T. J. Geutjes, M. Honing and M. Jaeger, *Magn. Reson. Chem.*, 2011, **49**, 425–436.
71. S. Berger and S. Braun, *200 and More NMR Experiments – A Practical Course*, Wiley-VCH, Weinheim, 2004.
72. W. F. Reynolds and R. G. Enriquez, *J. Nat. Prod.*, 2002, **65**, 221–244.
73. T. D. W. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier Ltd., Oxford, 2009.
74. G. E. Martin, C. E. Hadden and D. J. Russell, Solution NMR Spectroscopy, in *Handbook of Spectroscopy*, ed. G. Gauglitz and D. S. Moore, Wiley-VCH, Weinheim, 2014, pp. 209–268.
75. G. E. Martin, D. J. Russell, K. A. Blinov and M. E. Elyashberg, *Ann. Magn. Reson.*, 2003, **1**, 1–31.
76. E. E. Kwan and S. G. Huang, *Eur. J. Org. Chem.*, 2008, 2671–2688.
77. H. S. Atreya, *J. Indian Inst. Sci.*, 2010, **90**, 87–104.
78. W. Bermel, I. Bertini, I. C. Felli and R. Pierattelli, *J. Am. Chem. Soc.*, 2009, **131**, 15339–15345.
79. I. C. Felli and B. Brutscher, *ChemPhysChem*, 2009, **10**, 1356–1368.
80. R. Freeman and E. Kupce, *J. Biomol. NMR*, 2003, **27**, 101–114.
81. E. Kupce and R. Freeman, *Magn. Reson. Chem.*, 2010, **48**, 333–336.
82. E. Kupce and R. Freeman, *J. Magn. Reson.*, 2010, **206**, 147–153.

83. G. Fisher, C. Petucci, E. MacNamara and D. Raftery, *J. Magn. Reson.*, 1999, **138**, 160–163.
84. L. Frydman, T. Scherf and A. Lupulescu, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 15858–15862.
85. E. Kupce, R. Freeman and B. K. John, *J. Am. Chem. Soc.*, 2006, **128**, 9606–9607.
86. T. Parella and P. Nolis, *Concepts Magn. Reson., Part A*, 2010, **36A**, 1–23.
87. A. G. Webb, J. V. Sweedler and D. Raftery, Parallel NMR detection, in *On-Line LC-NMR Related Techniques*, ed. K. Albert, John Wiley & Sons Ltd, Chichester, UK, 2002, pp. 259–279.
88. M. W. Maciejewski, A. S. Stern, G. F. King and J. C. Hoch, *Mod. Magn. Reson.*, 2006, **2**, 1287–1293.
89. M. Mobli, A. S. Stern and J. C. Hoch, *J. Magn. Reson.*, 2006, **182**, 96–105.
90. K. A. Blinov, N. I. Larin, A. J. Williams, K. A. Mills and G. E. Martin, *J. Heterocycl. Chem.*, 2006, **43**, 163–166.
91. K. A. Blinov, A. J. Williams, B. D. Hilton, P. A. Irish and G. E. Martin, *Magn. Reson. Chem.*, 2007, **45**, 544–546.
92. R. Bruschweiler, *J. Chem. Phys.*, 2004, **121**, 409–414.
93. R. Bruschweiler and F. Zhang, *J. Chem. Phys.*, 2004, **120**, 5253–5260.
94. D. A. Snyder and R. Bruschweiler, *J. Phys. Chem. A*, 2009, **113**, 12898–12903.
95. D. A. Snyder and R. Bruschweiler, Multidimensional correlation spectroscopy by covariance NMR, in *Multidimensional Methods for the Solution State*, John Wiley & Sons Ltd., Chichester, UK, 2010, pp. 97–105.
96. F. Zhang and R. Bruschweiler, *J. Am. Chem. Soc.*, 2004, **126**, 13180–13181.
97. G. E. Martin, B. D. Hilton, P. A. Irish, K. A. Blinov and A. J. Williams, *J. Nat. Prod.*, 2007, **70**, 1393–1396.
98. G. E. Martin, B. D. Hilton, K. A. Blinov and A. J. Williams, *J. Heterocycl. Chem.*, 2008, **45**, 1109–1113.
99. B. D. Hilton and G. E. Martin, *J. Nat. Prod.*, 2010, **73**, 1465–1469.
100. G. E. Martin and R. C. Crouch, *Mod. Methods Plant Anal.*, 1994, **15**, 25–89.
101. C. E. Hadden, G. E. Martin, J. K. Luo and R. N. Castle, *J. Heterocycl. Chem.*, 1999, **36**, 533–539.
102. P. Bigler and C. Mueller, *Spectrochim. Acta, Part A*, 1994, **50A**, 297–306.
103. M. Liu, X.-A. Mao, C. Ye, J. K. Nicholson and J. C. Lindon, *J. Magn. Reson.*, 1997, **129**, 67–73.
104. P. Bigler, *Spectrosc. Lett.*, 2008, **41**, 162–165.
105. P. Bigler, R. Kuemmerle and W. Bermel, *Magn. Reson. Chem.*, 2007, **45**, 469–472.
106. J. Buddrus and J. Lambert, *Magn. Reson. Chem.*, 2002, **40**, 3–23.
107. A. D. Bain, D. W. Hughes, C. K. Anand, Z. Nie and V. J. Robertson, *Magn. Reson. Chem.*, 2010, **48**, 630–641.
108. K. Stott, J. Keeler, Q. N. Van and A. J. Shaka, *J. Magn. Reson.*, 1997, **125**, 302–324.

109. P. Forgo and K. E. Kover, *Steroids*, 2004, **69**, 43–50.
110. D. Yang, X. Xu and C. Ye, *Magn. Reson. Chem.*, 1992, **30**, 711–715.
111. B. L. Marquez, W. H. Gerwick and R. T. Williamson, *Magn. Reson. Chem.*, 2001, **39**, 499–530.
112. F. Guenneau, P. Mutzenhardt, D. Grandclaude and D. Canet, *J. Magn. Reson.*, 1999, **140**, 250–258.
113. A. J. Pell and J. Keeler, *J. Magn. Reson.*, 2007, **189**, 293–299.
114. K. Zangger and H. Sterk, *J. Magn. Reson.*, 1997, **124**, 486–489.
115. A. J. Kolkman, J. Draaisma, S. S. Wijmenga and M. Tessari, *Spectral Simplification by Collapsing J-multiplet Structures: An Improved Technique for Dilute Samples*, EUROMAR, Frankfurt, Germany, 2011.
116. J. A. Aguilar, S. Faulkner, M. Nilsson and G. A. Morris, *Angew. Chem., Int. Ed.*, 2010, **49**, 3901–3903.
117. K. A. M. Ampt, R. L. E. G. Aspers, M. Jaeger, P. E. T. J. Geutjes, M. Honing and S. S. Wijmenga, *Magn. Reson. Chem.*, 2011, **49**, 221–230.
118. K. L. Kirk, *J. Fluorine Chem.*, 2006, **127**, 1013–1029.
119. K. L. Kirk, *Curr. Top. Med. Chem.*, 2006, **6**, 1447–1456.
120. J. Fried and E. F. Sabo, *J. Am. Chem. Soc.*, 1957, **79**, 1130–1141.
121. M. Morgenthaler, E. Schweizer, A. Hoffmann-Röder, F. Benini, R. E. Martin, G. Jaeschke, B. Wagner, H. Fischer, S. Bendels, D. Zimmerli, J. Schneider, F. Diederich, M. Kansy and K. Müller, *ChemMedChem*, 2007, **2**, 1100–1115.
122. R. E. London and S. A. Gabel, *J. Am. Chem. Soc.*, 1994, **116**, 2570–2575.
123. L. A. Luck, J. E. Vance, T. M. O'Connell and R. E. London, *J. Biomol. NMR*, 1996, **7**, 261–272.
124. R. Martino, V. Gilard, F. Desmoulin and M. Malet-Martino, *J. Pharm. Biomed. Anal.*, 2005, **38**, 871–891.
125. H. Hu, P. Kulanthaivel and K. Krishnamurthy, *J. Org. Chem.*, 2007, **72**, 6259–6262.
126. K. A. M. Ampt, R. L. E. G. Aspers, P. Dvortsak, R. M. van der Werf, S. S. Wijmenga and M. Jaeger, *J. Magn. Reson.*, 2012, **215**, 27–33.
127. C. M. Thiele, *Eur. J. Org. Chem.*, 2008, 5673–5685.
128. B. Luy, K. Kobzar and H. Kessler, *Angew. Chem., Int. Ed.*, 2004, **43**, 1092–1094.
129. G. Kummerloewe and B. Luy, *TrAC, Trends Anal. Chem.*, 2009, **28**, 483–493.
130. C. M. Thiele and S. Berger, *Org. Lett.*, 2003, **5**, 705–708.
131. G. Kummerloewe and B. Luy, *Annu. Rep. NMR Spectrosc.*, 2009, **68**, 193–230.
132. J. C. Freudenberger, P. Spiteller, R. Bauer, H. Kessler and B. Luy, *J. Am. Chem. Soc.*, 2004, **126**, 14690–14691.
133. M. Zweckstetter and A. Bax, *J. Biomol. NMR*, 2001, **20**, 365–377.
134. M. Zweckstetter, *Nat. Protoc.*, 2008, **3**, 679–690.
135. J. A. Losonczi, M. Andrec, M. W. F. Fischer and J. H. Prestegard, *J. Magn. Reson.*, 1999, **138**, 334–342.
136. B. Stevansson, C. Landersjö, G. Widmalm and A. Maliniak, *J. Am. Chem. Soc.*, 2002, **124**, 5946–5947.

137. C. Landersjö, B. Stevansson, R. Eklund, J. Ostervall, P. Soderman, G. Widmalm and A. Maliniak, *J. Biomol. NMR*, 2006, **35**, 89–101.
138. M. Sarfati, J. Courtieu and P. Lesot, *Chem. Commun.*, 2000, 1113–1114.
139. M. Sarfati, P. Lesot, D. Merlet and J. Courtieu, *Chem. Commun.*, 2000, 2069–2081.
140. C. Aroulanda, V. Boucard, F. Guibe, J. Courtieu and D. Merlet, *Chem. – Eur. J.*, 2003, **9**, 4536–4539.
141. R. R. Gil, C. Gayathri, N. V. Tsarevsky and K. Matyjaszewski, *J. Org. Chem.*, 2008, **73**, 840–848.
142. V. M. Marathias, G. J. Tawa, I. Goljer and A. C. Bach, *Chirality*, 2007, **19**, 741–750; R. Berger, J. Courtieu, R. R. Gil, C. Griesinger, C. Koeck, P. Lesot, B. Luy, D. Merlet, A. Navarro-Vazquez, M. Reggelin, U. M. Reinscheid, C. M. Thiele and M. Zweckstetter, *Angew. Chem., Int. Ed.*, 2012, **51**, 8388–8391.
143. L. Ziani, P. Lesot, A. Meddour and J. Courtieu, *Chem. Commun.*, 2007, 4737–4739.
144. G. Kummerloewe, B. Crone, M. Kretschmer, S. F. Kirsch and B. Luy, *Angew. Chem.*, 2011, **123**, 2693–2696.
145. A. Mangoni, V. Esposito and A. Randazzo, *Chem. Commun.*, 2003, 154–155.
146. G. Kummerloewe, S. Schmitt and B. Luy, *Open Spectrosc. J.*, 2010, **4**, 16–27.
147. P. Wijten, *Structure elucidation and chiral discrimination of small molecules using Residual Dipolar Couplings*, Thesis Wageningen University, 2010.
148. P. Wijten, R. L. E. G. Aspers, M. Wagener and M. Jaeger, in preparation.
149. H. Duddeck, *Nat. Prod. Anal., [Symp.]*, 1998, 105–118.
150. L. D. Hall and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1980, **102**, 5703–5711.
151. A. G. J. Sedee, G. M. J. Beijersbergen van Henegouwen, W. Guijt and C. A. G. Haasnoot, *J. Chem. Soc., Perkin Trans. 2*, 1984, 1755–1759.
152. N. M. Teerhuis, I. A. M. Huisman and M. B. Groen, *Tetrahedron Lett.*, 2001, **42**, 2869–2871.
153. L. Fielding, Y. Diepeveen, D. Fletcher and N. Hamilton, *Magn. Reson. Chem.*, 2001, **39**, 323–328.
154. M. D. Grove, G. F. Spencer, W. K. Rohwedder, N. Mandava, J. F. Worley, J. D. Warthen, G. L. Steffens, J. L. Flippen-Anderson and J. C. Cook, *Nature*, 1979, **281**, 216–217.
155. G. Adam and V. Marquardt, *Phytochemistry*, 1986, **25**, 1787–1799.
156. M. Hesse, H. Meier and B. Zehe, *Spektroskopische Methoden in der Organischen Chemie*, Georg Thieme Verlag, Stuttgart, 1991.
157. W. F. Reynolds, M. Yu and R. G. Enriquez, *Magn. Reson. Chem.*, 1997, **35**, 614–618.
158. W. F. Reynolds, S. McLean, H. Jacobs and W. W. Harding, *Can. J. Chem.*, 1999, **77**, 1922–1930.
159. J. H. Ardenkjaer-Larsen, B. Fridlund, A. Gram, G. Hansson, L. Hansson, M. H. Lerche, R. Servin, M. Thaning and K. Golman, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 10158–10163.

160. C. Song, K.-N. Hu, C.-G. Joo, T. M. Swager and R. G. Griffin, *J. Am. Chem. Soc.*, 2006, **128**, 11385–11390.
161. N. T. Nyberg, J. O. Duus and O. W. Sorensen, *J. Am. Chem. Soc.*, 2005, **127**, 6154–6155.
162. N. T. Nyberg, J. O. Duus and O. W. Sorensen, *Magn. Reson. Chem.*, 2005, **43**, 971–974.
163. A. A. Kicha, N. V. Ivanchina, T. T. T. Huong, A. I. Kalinovsky, P. S. Dmitrenok and P. Q. Long, *Russ. Chem. Bull.*, 2010, **59**, 2133–2136.
164. S. Uhlig, H. Wisloff and D. Petersen, *J. Agric. Food Chem.*, 2007, **55**, 6018–6026.
165. R. Holm, J. C. Madsen, W. Shi, K. L. Larsen, L. W. Staede and P. Westh, *J. Inclusion Phenom. Macrocycl. Chem.*, 2011, **69**, 201–211.
166. C. Schoenbeck, P. Westh, J. C. Madsen, K. L. Larsen, L. W. Staede and R. Holm, *Langmuir*, 2011, **27**, 5832–5841.
167. S. F. Cheatham, M. Kline, R. R. Sasaki, K. A. Blinov, M. E. Elyashberg and S. G. Molodtsov, *Magn. Reson. Chem.*, 2010, **48**, 571–574.
168. G. E. Martin, B. D. Hilton, M. R. Willcott, III and K. A. Blinov, *Magn. Reson. Chem.*, 2011, **49**, 350–357.
169. G. E. Martin and R. C. Crouch, *J. Nat. Prod.*, 1991, **54**, 1–70.
170. M. H. Levitt, *Spin Dynamics. Basics of Nuclear Magnetic Resonance*, John Wiley & Sons Ltd, Chichester, UK, 2001.

CHAPTER 9

Nuclear Magnetic Resonance Experiments Applicable to the Elucidation and Characterization of Alkaloid Structures Part I: Direct ^1H – ^{13}C Heteronuclear Shift Correlation and Establishing Contiguous Protonated Carbon Spin Systems

JOSEP SAURÍ AND GARY E. MARTIN*

Merck & Co., Inc., Process & Analytical Chemistry, NMR Structure Elucidation, Rahway, NJ 07065, USA

*Email: gary.martin2@merck.com

9.1 Introduction

Alkaloids are nitrogen-containing compounds that are generally considered as secondary metabolites of the organism or plant that elaborates them.

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

It seems germane to again cite the definition of alkaloids from Wikipedia that was recently used in a review of ^{15}N nuclear magnetic resonance (NMR) of alkaloids.¹

“The boundary between alkaloids and other nitrogen-containing natural compounds is not clear-cut. Compounds like amino acids, peptides, proteins, nucleotides, nucleic acids, amines, and antibiotics are usually not called alkaloids. Natural compounds containing nitrogen in the exocyclic position (mescaline, serotonin, dopamine, *etc.*) are usually attributed to amines rather than alkaloids. Some authors, however, consider alkaloids a special case of amines.”²

Alkaloids are an exceptionally diverse family of compounds that have been the subject of many volumes, and in that sense, any attempt to review NMR applications on a class-by-class basis, as has been done in recent reviews of the applications of ^{15}N NMR in alkaloids, would be far beyond the scope of this chapter.^{1,3} There have also been several reviews of the applications of 2D NMR methods to alkaloids as a class, although those are now rather dated and do not deal with any of the modern NMR techniques now available.^{4,5} Indeed, there has been a highly focused effort on the development of homodecoupled or what are generally referred to as “pure shift” (PS) NMR methods that are the topic of several recent reviews,^{6–8} as well as other new NMR methods that are certainly applicable in the elucidation or structure characterization of alkaloids.

2D NMR methods that can be applied to alkaloids are nearly as diverse as the range of structures that can be labeled as alkaloids. Fortunately, however, many of the reports of new NMR methods have utilized strychnine as a model compound over the past 20–30 years. In a discussion of NMR methods that are applicable to alkaloid structure determination, we will begin by employing strychnine as well as several other alkaloids as model compounds to illustrate the function of a selection of 2D NMR methods that the current authors have found to have utility for alkaloid structure elucidation. In very general terms, the NMR methods that an investigator would probably consider using as soon as a sample is available and ready for NMR data acquisition might be labeled as “first-tier” methods. The order in which the various experiments from such a group might be performed will undoubtedly vary from one investigator to the next. There is not necessarily any right or wrong order to acquiring these data, and the order in which we will present them in this chapter represents nothing more than the predilection of the authors.

9.2 “First-tier” NMR Methods for Alkaloid Structure Characterization

Irrespective of which laboratory an investigator works in, the first NMR experiment of any type will undoubtedly be a proton reference spectrum.

Proton spectra serve as the basis of evaluating the likely purity of a sample and defining proton chemical shift ranges for subsequent experiments as well as the potential range of F_1 chemical shifts for experiments such as multiplicity-edited heteronuclear single-quantum correlation (ME-HSQC) that will undoubtedly follow.

9.2.1 Sample Preparation

Preparing a sample for NMR data acquisition involves a continuum of choices. Generally, the first choices will be the selection of a deuterated NMR solvent as well as a tube format, which will, in turn, be governed by the instrumentation available in the laboratory where the data will be acquired. Obviously, solubility will be a key criterion in selecting a solvent. While many investigators are inclined to employ deuteriochloroform as a first choice, chloroform can be acidic and protonation can cause broadening of the proton spectra of alkaloids. Deuteriochloroform can be readily neutralized or we have found it useful on some occasions to bubble a bit of ammonia gas up through an NMR sample, provided of course that the diameter of the NMR tube being used will allow this, which 3 and 5 mm tubes do. Another point that should be kept in mind is that NMR solvents such as d_4 -methanol readily exchange NH or OH resonances of whatever molecule is dissolved in the solvent, which is undesirable if it is necessary to acquire a ^1H - ^{15}N HSQC spectrum.

The sample concentration will in many instances be governed by the quantity of material available for NMR investigation. If more generous quantities of material are available, the amount dissolved will be a combination of the experiment(s) to be performed, spectrometer frequency, and the NMR probe in which the data will be acquired. Clearly, smaller tube diameters are advantageous when the sample to be investigated is scarce. In addition, smaller tube diameters (*e.g.* 3 and 1.7 mm) put progressively greater percentages of the actual sample in the rf coil of the NMR probe. Small format NMR probes and probe performance have been reviewed several times and the interested reader is referred to those reports for a more comprehensive treatment of the topic.^{9–15} Higher sample concentrations can also be advantageous for the acquisition of ^{13}C NMR spectra, as well as for the acquisition of more sophisticated NMR data (*e.g.* 1,1- and 1,*n*-ADEQUATE spectra (see Chapter 10, Sections 10.1.2 and 10.2)).^{16,17}

9.2.2 Probe Selection

Probe selection may or may not be a choice that an investigator can make. Generally, the probe used will be a function of the NMR laboratory where the investigator is working. Probe geometry can either be “conventional”, where the X-observe coil is closer to the sample with the ^1H coil further removed and generally used for broadband proton decoupling, or “inverse”, where the proton coil is closer to the sample with the X-coil further removed. Inverse-detection NMR probes are best suited for heteronuclear 2D NMR

experiments where ^1H is the detected nucleus, with these experiments giving much better sensitivity than X-nucleus detection probes, which were the norm prior to the early 1990s.

Another facet of probe choice is that of two- *versus* three-channel probes. The former will generally have a proton coil and an X-coil that can be tuned across a range that, for natural products, would generally range from ^{15}N to ^{13}C . Three-channel probes provide an X-coil that is doubly tuned to both ^{13}C and ^{15}N . Triple-resonance probes offer greater flexibility that, until recently, was more of a factor of convenience for the natural products chemist utilizing NMR. With a triple-resonance probe, it is possible for an investigator to queue a series of 2D NMR experiments that might include both ^1H - ^{13}C and ^1H - ^{15}N 2D NMR experiments overnight or over the span of a weekend. Experiments correlating ^{13}C and ^{15}N in a single experiment were not feasible at natural abundance prior to late 2014.^{18,19} Instead, it was necessary to utilize unsymmetrical indirect covariance processing²⁰ of ^1H - ^{13}C and ^1H - ^{15}N spectra that share a common dimension, the ^1H frequency domain, to calculate a ^{13}C - ^{15}N correlation spectrum.^{21–24} In late 2014, Cheatham and coworkers^{18,19} introduced the HCNMBC experiment that utilizes an initial ^1H - ^{13}C transfer *via* $^1J_{\text{CH}}$ followed by a natural abundance ^{13}C - ^{15}N transfer *via* either $^1J_{\text{CN}}$ or $^nJ_{\text{CN}}$, after which the magnetization is transferred back and detected *via* the starting proton in a fashion analogous to the 1,1- and 1,*n*-ADEQUATE experiments.^{16,17} Since the HCNMBC experiment applies pulses on ^1H , ^{13}C , and ^{15}N during the course of the pulse sequence (see Chapter 3, Section 3.8.5.2), a triple resonance probe is mandatory to perform the experiment.

The other choice alluded to in the previous section of this chapter is that of a room temperature *versus* a cryogenically cooled NMR probe. Cryogenically cooling the rf coils of the NMR probe and the preamplifiers reduces noise, affording correspondingly higher signal/noise (s/n) and as much as four-times the sensitivity of a probe operating at room temperature for helium-cooled cryoprobes. Until fairly recently, cryogenically cooled probes or cryoprobes utilized a closed loop cooling system that employed circulating helium gas as the refrigerant, which gave an operating rf coil temperature in the range of approximately 18–25 K. More recently, Bruker introduced Prodigy™ LN₂ probes that utilize a nitrogen Dewar and the boil-off from the Dewar to cool the rf coils. While LN₂-cooled rf coils do not offer quite the sensitivity gain of helium-cooled cryoprobes, they nevertheless afford 2- to 2.5-times the sensitivity of a conventional room temperature probe and are considerably less expensive and cheaper to operate than helium-cooled cryoprobes. For experiments such as 1,1- or 1,*n*-ADEQUATE^{16,17} or HCNMBC,^{18,19} using either a LN₂- or preferably a helium-cooled cryoprobe is highly advised in order to keep data acquisition times reasonable.

9.2.3 Parameter Choices

Parameter choices vary widely as a function of the experiment being performed. Choices will generally involve ^1H and X (most typically ^{13}C and

^{15}N for natural products) nuclei spectral widths, choices of the F_2 digitization, the number of increments to be performed to digitize the second frequency domain F_1 , whether or not to employ non-uniform sampling (NUS), the number of transients accumulated/ t_1 increment, interpulse delays, and delays based on homo- and/or heteronuclear coupling constants utilized in manipulating the components of magnetization in the experiment. Another choice that must sometimes be made is whether or not it is feasible to “fold” the 2D experiment into the second frequency domain. For example, if a HSQC spectrum is being acquired for a small sample (*e.g.* $<50\ \mu\text{g}$), it may be advantageous to limit the F_1 spectral range to perhaps 10–80 ppm. The region from approximately 80 ppm to the beginning of the aromatic region at perhaps 115–120 ppm is generally devoid of responses and there is generally little to be gained by digitizing this region of the 2D NMR spectrum, although there are a few structural moieties that do exhibit resonances in the region of 80–110 ppm. Since there is no filter in the F_1 dimension, responses below 80 ppm, if that is the lower limit of the F_1 range selected, will simply fold back into the spectrum from the opposite end of the spectrum (*e.g.* from 10 ppm). The location of the folded resonance will be governed by the chemical shift of the resonance in a predictable manner and will be discussed in the context of the HSQC spectrum (see Section 9.6).

9.3 Acquiring NMR Spectra

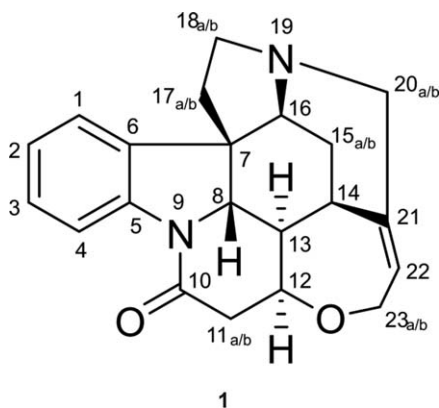
Opinions will vary, as noted in the introduction to this chapter, as to the order in which various NMR spectra should be acquired by an investigator. We will take the route that we happen to prefer, acquiring a proton reference spectra followed by the acquisition of some form of HSQC spectrum. We specifically note here that HSQC is strongly preferred over the acquisition of the older HMQC experiment. Once a proton and HSQC spectrum are in hand, logical decisions can be made on the next series of experiments that it may be advisable to acquire.

9.4 1D Reference Spectra

There is relatively little need to discuss 1D reference spectra. A good proton spectrum is the nominal starting point for any alkaloid structure elucidation study. Generally, better-quality proton spectra can be acquired with less concentrated solutions in whatever deuterated solvent that the alkaloid happens to be soluble in. The majority of alkaloids that the authors have encountered have been soluble in chloroform, methanol, acetonitrile, or dimethyl sulfoxide (see Section 9.2.1). The acquisition of a proton reference spectrum will generally afford an investigator with an idea of whether or not the acquisition of a ^{13}C reference spectrum is going to be a feasible undertaking. When working with a scarce sample (*e.g.* submilligram samples), it may take tens of thousands of transients to record a usable ^{13}C reference

spectrum, which in many cases will not be worth the investment in spectrometer time.

For the purpose of illustration, in this chapter we will first utilize the readily available *Strychnos* alkaloid strychnine (**1**). For investigators wishing to acquire any of the various spectra discussed in this chapter themselves, we would recommend the purchase of a 250 mg bottle of Fluka's PestanalTM analytical reagent-grade strychnine. The purity is extremely high, and it is not necessary to recrystallize the alkaloid prior to using a sample of this material in order to begin acquiring data. In contrast, strychnine that has been sitting around on a laboratory shelf will likely require recrystallization from ethanol to rid the sample of strychnine 19-*N*-oxide that forms over time from simple air oxidation.²⁵ The alkaloid is freely soluble in deuteriochloroform (~4 mg in 35 μ L CDCl₃) and we will utilize that sample in a 1.7 mm NMR tube in conjunction with a 600 MHz three-channel Bruker AVANCE III NMR spectrometer equipped with a 1.7 mm TXI MicroCryoProbeTM in order to acquire the NMR spectra used for illustration in this chapter. A comparable concentration could be used in a 3 or 5 mm tube format.



The 600 MHz ¹H and 150 MHz ¹³C reference spectra of the sample just noted in the previous paragraph are shown in Figure 9.1A and 9.1B, respectively. Opinions differ on the acquisition of carbon data. Some authors prefer to acquire a carbon reference spectrum, but the current authors find it preferable to extract carbon chemical shift data from a direct ¹H-¹³C correlation HSQC spectrum and some variant of a long-range heteronuclear shift correlation spectrum (e.g. HMBC or LR-HSQMBC). The question of carbon resonance multiplicity was for many years addressed using either an APT²⁶ or DEPT²⁷ experiment, but this same information can be gleaned much more quickly from a multiplicity-edited HSQC spectrum on a modern spectrometer.²⁸

Inspection of the proton and carbon reference spectra shown in Figure 9.1 shows a well-dispersed proton spectrum, with aliphatic resonances suggestive of the possibility of oxygen-bearing carbons that is cross-confirmed

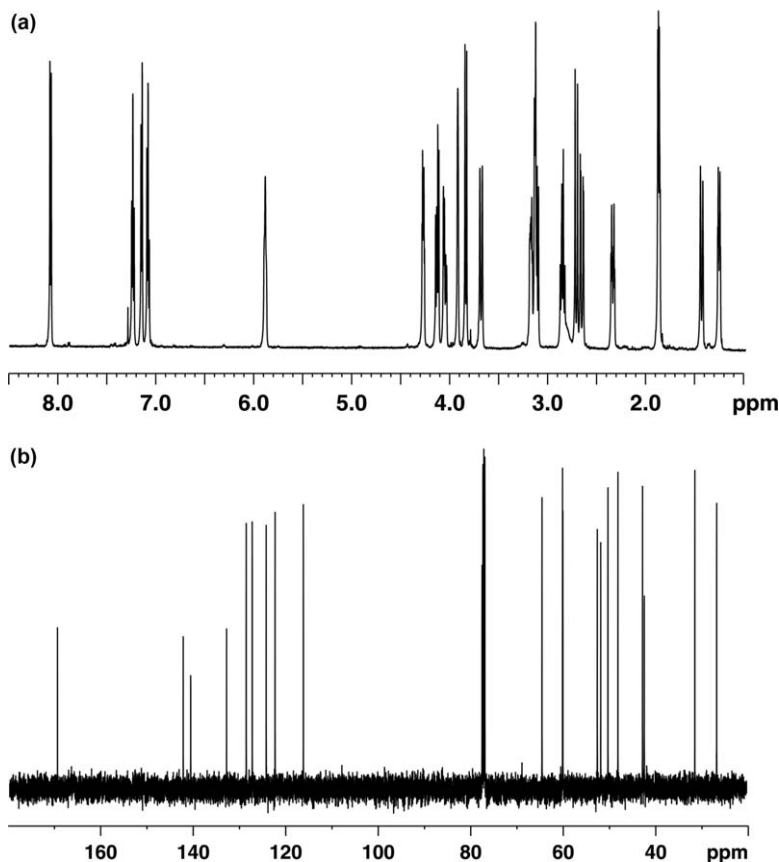


Figure 9.1 (A) ^1H NMR spectrum of an ~ 4 mg sample of strychnine (**1**) dissolved in $35\ \mu\text{L}$ of deuteriochloroform in a 1.7 mm NMR tube recorded at an observation frequency of 600 MHz. The spectrum was recorded in four transients digitized with 64K data points. (B) ^{13}C reference spectrum of **1** recorded at 150 MHz using the same sample in 128 transients digitized with 64K data points. The acquisition of the spectrum required 3 minutes and 50 seconds. The C12 resonance is just downfield of the chloroform multiplet.

by examination of the carbon spectrum. There are no methyl resonances observed in the proton spectrum. There is a vinyl proton at ~ 6.22 ppm that is broad with a “triplet” character, which suggests that the proton in question might reside next to a methylene; there are four well-resolved aromatic protons that are consistent with a contiguous four-spin system, suggesting a 1,2-disubstituted phenyl. The carbon spectrum contains a resonance just downfield of the chloroform $1:1:1$ triplet, consistent with an oxygen-bearing carbon, eight resolved sp^2 carbons and a carbonyl. Short of beginning to interpret multiplet structures in the proton spectrum, there is not a great deal more that can be gleaned from simple inspection of the 1D spectra.

9.5 Fundamental or “First-tier” 2D NMR Spectra

At this juncture, some investigators will have a philosophical difference with the authors of this chapter in terms of what experiment should be done next. Many will choose to acquire a COSY²⁹ spectrum next, based on the premise that it is a quick experiment to do and provides considerable insight into the interpretation of the proton NMR spectrum. Heteronuclear shift correlation, prior to the advent of inverse detection, would have indeed taken considerably longer to acquire than a COSY spectrum. Thus, while 20 years ago COSY might have been the next experiment to logically acquire in the structure elucidation process, this is no longer the case when proton-detected heteronuclear experiments are considered. Indeed, the development and continually increasing access to cryogenic probes makes it possible to acquire an HSQC spectrum on the sample of strychnine (**1**) being used to illustrate this chapter in less than 5 minutes if desired. When PS methods are factored into consideration, for smaller samples than that being used as an example, both resolution and, since homonuclear couplings are largely collapsed (with the exception of germinal methylene proton–proton couplings, since both protons are on ¹³C and hence the bilinear rotational decoupling (BIRD)-based homodecoupling cannot operate), sensitivity are enhanced. Beyond these considerations, there is also the added benefit of multiplicity editing^{30,31} so that CH/CH₃ resonances will exhibit the opposite phase to that of CH₂ resonances. The synergy of these factors provides a strong argument for acquiring a HSQC spectrum immediately after a proton spectrum is in hand.

9.5.1 Conventional HSQC

To illustrate the various ways in which HSQC spectra can be acquired, we will begin with the HSQC spectrum without either multiplicity-editing or BIRD-based homonuclear decoupling that is shown in Figure 9.2.²⁸ The data shown in Figure 9.2 were acquired using the pulse sequence shown in Figure 9.3. The spectrum was acquired using 128 increments to digitize the spectrum in the F₁ frequency domain; the data were linear predicted to 256 points in F₁ and zero-filled during processing to afford the 2K×1K spectrum shown. The data were acquired in 10 minutes. The transmitter was positioned at 100 ppm in the F₁ frequency domain. Note that in the aliphatic region, the F₁ chemical shift range downfield of the transmitter is devoid of responses, as is the region of the spectrum upfield of the transmitter in the vinyl/aromatic region of the spectrum. This empty “real estate” provides several opportunities for the better utilization of spectrometer time and sensitivity that will be addressed later in this chapter.

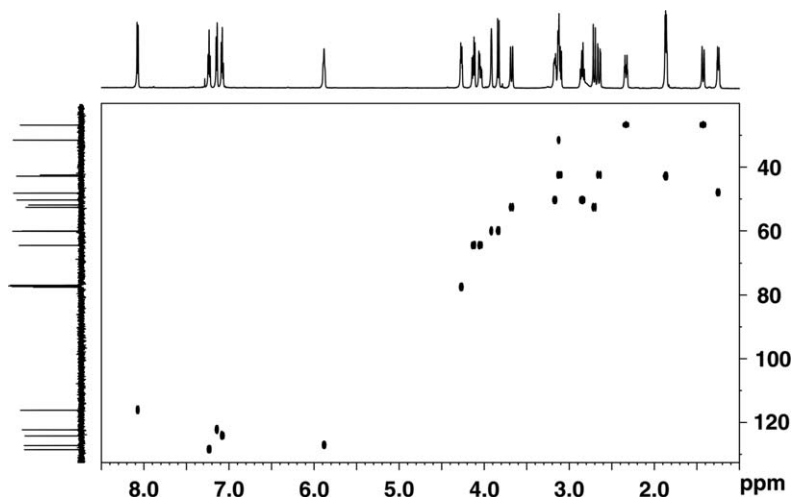


Figure 9.2 HSQC spectrum of a ~ 4 mg sample of strychnine (**1**) dissolved in 35 μL deuteriochloroform in a 1.7 mm NMR tube. The data were acquired as $2\text{K} \times 128$ points in ~ 10 minutes using a Bruker 600 MHz spectrometer equipped with a 1.7 mm TXI MicroCryoProbeTM. Data were processed using linear prediction in F_1 to 256 points followed by zero-filling to afford the $2\text{K} \times 1\text{K}$ spectrum shown.

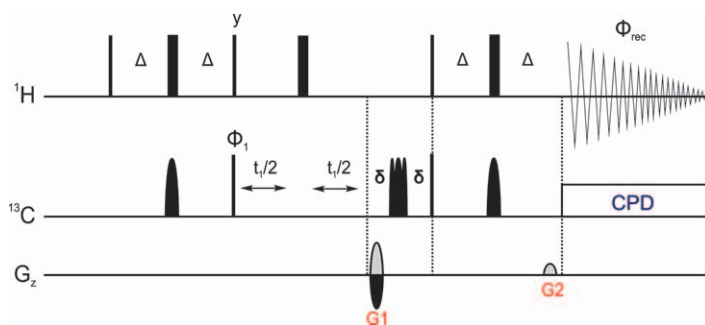


Figure 9.3 Pulse sequence for the basic HSQC experiment used to acquire the spectrum shown in Figure 9.2. Narrow solid lines denote 90° pulses; thick solid black lines denote 180° hard pulses. Shaped pulses are denoted by solid half ellipses; composite shaped pulses are denoted by the triplicate half ellipse shown just after the G1 gradient on the ^{13}C channel. This convention will be followed throughout this chapter. The delay Δ for the heteronuclear couplings should be set to $1/4(J_{\text{CH}})$. As a compromise, this delay is typically optimized for 145 Hz. Phases are x unless otherwise noted: $\Phi_1 = 0, 2$; $\Phi_{\text{rec}} = 0, 2$. Gradient ratios must be adjusted for the heteronucleide pair being investigated.³⁰ Hence, the gradient ratio is set to $\text{G1}:\text{G2} = 80:20.1$ for ^1H - ^{13}C heteronuclear correlation experiments and should be set to $80:8.1$ for ^1H - ^{15}N experiments.

9.5.2 Multiplicity-edited HSQC

The next increase in sophistication of the HSQC experiment is the incorporation of a multiplicity-editing block at the end of the pulse sequence after ^{13}C chemical shift labeling of the proton resonances has been accomplished during the evolution period. The pulse sequence for multiplicity-edited HSQC is shown in Figure 9.4.³¹ Advantage can be taken of the even/odd nature of ^{13}C -coupled ^1H multiplet structure. Hence, CH/CH₃ resonances (even) will end up with a positive phase, while CH₂ resonances (odd) will be inverted. We will utilize the convention of positively phased resonances being plotted in black and negatively phased resonances being plotted in red throughout this chapter. Following the multiplicity-editing step, decoupling is applied during acquisition in the usual fashion. The multiplicity-edited HSQC spectrum is shown in Figure 9.5.

As expected, based on the multiplicity-editing step, the spectrum shown in Figure 9.5 is plotted with positively phased responses (only CH in the case of **1**) shown in black and methylene responses inverted and plotted in red.

A comparison of the aliphatic regions of the spectra shown in Figures 9.2 and 9.5 is presented in Figure 9.6. The multiplicity-editing step in the pulse sequence shown in Figure 9.4 provides carbon resonance multiplicity information that is analogous to what would be garnered from a ^{13}C DEPT spectrum.

9.5.3 Incorporating BIRD-based Homonuclear Decoupling in the HSQC Experiment: PS-HSQC

Considerable effort has been devoted, beginning in 2013, to the development of PS-HSQC methods based on BIRD-based³² homonuclear decoupling, beginning with the work of Morris and coworkers.³³ Covering PS

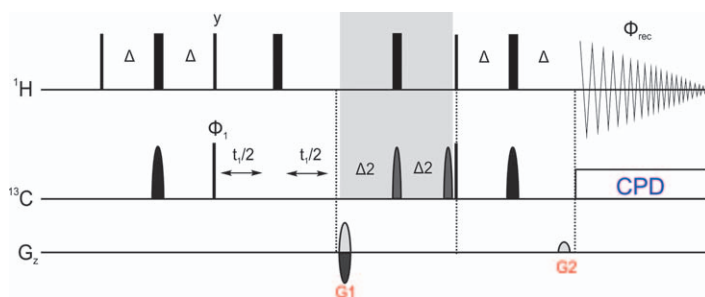


Figure 9.4 Pulse sequence for multiplicity-edited HSQC. Following the evolution period, multiplicity editing is accomplished using the π -sandwich shown in the shaded block. Delays were set: $\Delta = 1/4(^1J_{\text{CH}})$ and $\Delta 2 = 1/2(^1J_{\text{CH}})$. Phases were x unless otherwise noted: $\Phi_1 = 0, 2$; $\Phi_{\text{rec}} = 0, 2$. Again, the gradient ratios should be set: the $G_1 : G_2 = 80 : 20.1$ for ^1H - ^{13}C heteronuclear correlation experiments and $80 : 8.1$ for ^1H - ^{15}N experiments.

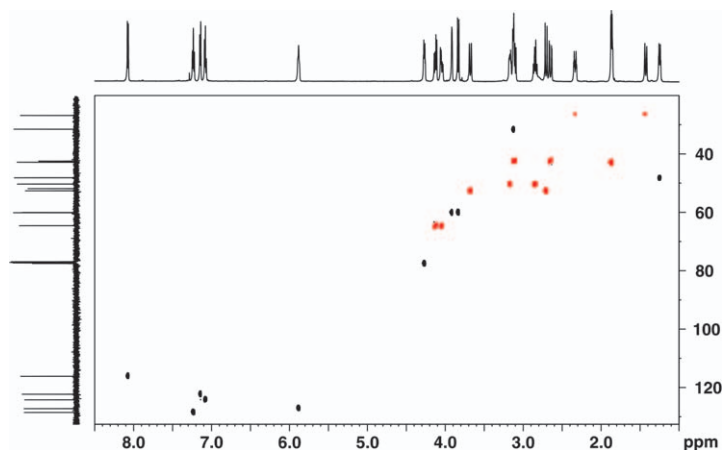


Figure 9.5 Multiplicity-edited HSQC spectrum of strychnine (**1**). Methine/methyl resonances have positive phase and are plotted in black; methylene resonances are inverted and are plotted in red. Anisochronous methylene protons are readily identified in this presentation. The data were acquired as $3\text{K} \times 128$ points (F_2 , F_1 , respectively), accumulating four transients/ t_1 increment. Data acquisition took 11 minutes (overkill). The data were processed using linear prediction in the second frequency domain to 256 points followed by zero-filling to 1K points. Apodization was $\pi/2$ phase-shifted sine bell in both frequency domains to afford the $2\text{K} \times 1\text{K}$ point spectrum shown.

NMR is well beyond the scope of this chapter, but efforts in this area of NMR investigation have also been the subject of several recent reviews, to which the interested reader is directed for further information.^{6–8} The so-called PS-HSQC experiments operate using what can be referred to as “chunked” data acquisition. Quite simply, a block of data points is acquired, a BIRD pulse followed by a hard 180° pulse is applied, and the next “chunk” of data is acquired and the process is repeated n times until the complete FID has been accumulated. It must be recalled that the BIRD pulse operates by inverting only protons on ^{12}C , leaving those protons directly attached to ^{13}C unaffected to be refocused by the hard 180° pulse, providing the desired homonuclear decoupling. In the PS-HSQC experiment, the homonuclear decoupling is afforded by the pulses in the shaded area of the pulse sequence.

The impact of the number of loops through the BIRD-based decoupling element of the pulse sequence shown in Figure 9.7 is addressed in Figure 9.8. The H4/C4 aromatic correlation of strychnine is used as an example. In the first trace of the stack plot, the loop counter, n , was set to 1. It is readily apparent that there was effectively no decoupling, with the H4 resonance remaining a doublet. When $n = 2$, notable collapse of the doublet begins to occur and the resonance has assumed the appearance of a triplet. With successively greater values of loop counter n , the outer limbs of the triplet diminish in intensity until they are only vestigial in the range $n = 6$ –8.

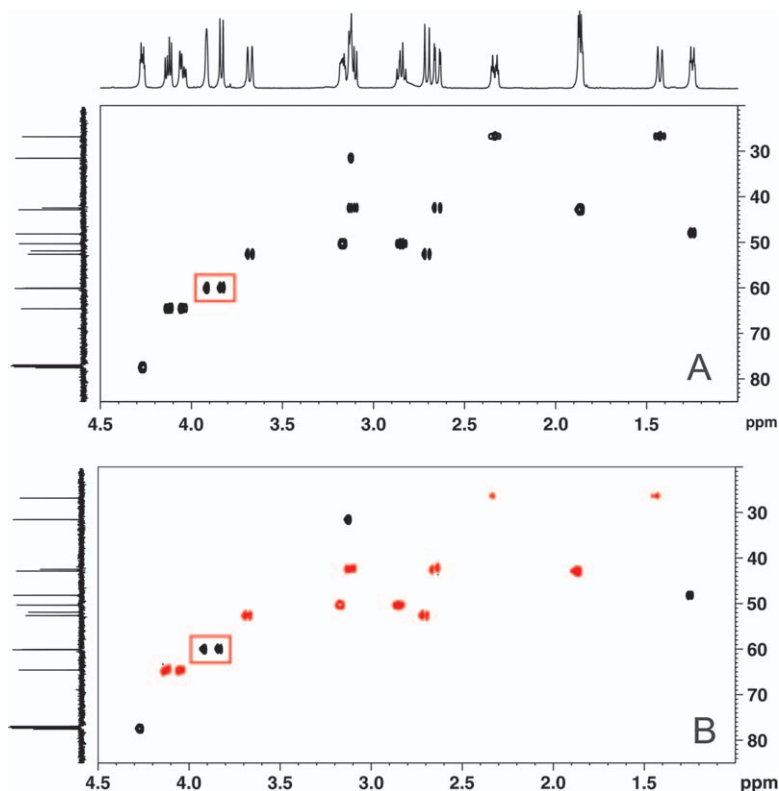


Figure 9.6 (A) Aliphatic region from the unedited HSQC spectrum of strychnine (**1**) shown in Figure 9.2. (B) Corresponding region from the multiplicity-edited HSQC spectrum shown in Figure 9.5. If this were an unknown molecule, the pair of resonances shown in the red-boxed region in (A) might be mistaken for a methylene when, in fact, these are the C8 and C16 CH correlations, with the two carbons having nearly identical carbon chemical shifts. The positive phase of the same resonances in the boxed region of (B) confirms that these are methine resonances. Compare, for example, the anisochronous 23-methylene proton resonances just downfield of the 8/16 resonances in (B). The negative phase confirms that those are methylene protons.

While it might seem attractive to opt for greater values of n , as the number of loops through the pulse sequence element increases, baseline noise and line width both begin to increase, offsetting some of the gain accrued by resorting to BIRD-based decoupling.

It was noted above that the BIRD element only operates to invert those protons bound to ^{13}C , with refocusing of the ^1H - ^{12}C resonances affording broadband homonuclear decoupling. In the case of methine or methyl resonances, the multiplet is collapsed to a singlet as in the case of the H4 resonance of strychnine (**1**) used as an example in Figure 9.8. In contrast, for the case of anisochronous methylene protons, remote protons (*i.e.* $^nJ_{\text{HH}}$

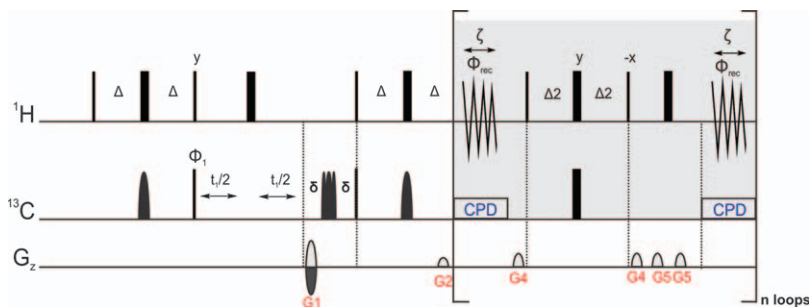


Figure 9.7 Pulse sequence for PS-HSQC. The pulse sequence is analogous to that shown for HSQC in Figure 9.2. Broadband homonuclear decoupling is afforded by the shaded region of the pulse sequence that gives what can be termed “chunked” data acquisition. Delays were set: $\Delta = 1/4(^1J_{\text{CH}})$ and $\Delta 2 = 1/2(^1J_{\text{CH}})$. Phases were x unless otherwise noted: $\Phi_1 = 0$, 2; $\Phi_{\text{rec}} = 0$, 2. The interval $\zeta = aq/2n$. The gradient ratios should be set: the $G1 : G2 = 80 : 20.1$ for ^1H - ^{13}C heteronuclear correlation experiments and $80 : 8.1$ for ^1H - ^{15}N . The other gradients should be set: $G4 = 3$; $G5 = 5$. The loop counter should be set to a value to give complete decoupling as shown in Figure 9.8.

where $n > 2$) will be refocused and hence decoupled, while the geminal coupling $^2J_{\text{HH}}$ will not be refocused, since both protons reside on the same ^{13}C and are hence affected in the same way by the operation of the BIRD element, which is to say that they are not decoupled.

Based on the discussion in the above, it is instructive to compare the results from a multiplicity-edited HSQC with a multiplicity-edited PS-HSQC spectrum. The comparison can be readily drawn for the two segments of the strychnine spectrum shown in Figure 9.9. The H23a/b resonances are anisochronous and both are vicinally coupled to the H22 vinyl proton, accounting for the doubled doublet observed for each of the protons. The H20a resonance is also shown in the spectral segment. The 20-methylene is isolated with no vicinal coupling partners. From the hypothetical structural fragment shown in Figure 9.10B, we would expect the H23a/b proton resonances to collapse to a pair of doublets and the H20a resonance to remain a geminal doublet. That is exactly what is noted when Figure 9.9A and 9.9B are compared.

9.5.4 Summary of One-bond Heteronuclear Correlation Methods

The authors of this chapter firmly believe that following the acquisition of a proton reference spectrum, the most meaningful second experiment that can be acquired, assuming that there is sufficient sample and/or access to appropriate probe technology, is undoubtedly a multiplicity-edited PS-HSQC spectrum. Multiplicity-editing provides carbon resonance multiplicity without having to resort to the acquisition of a ^{13}C reference spectrum,

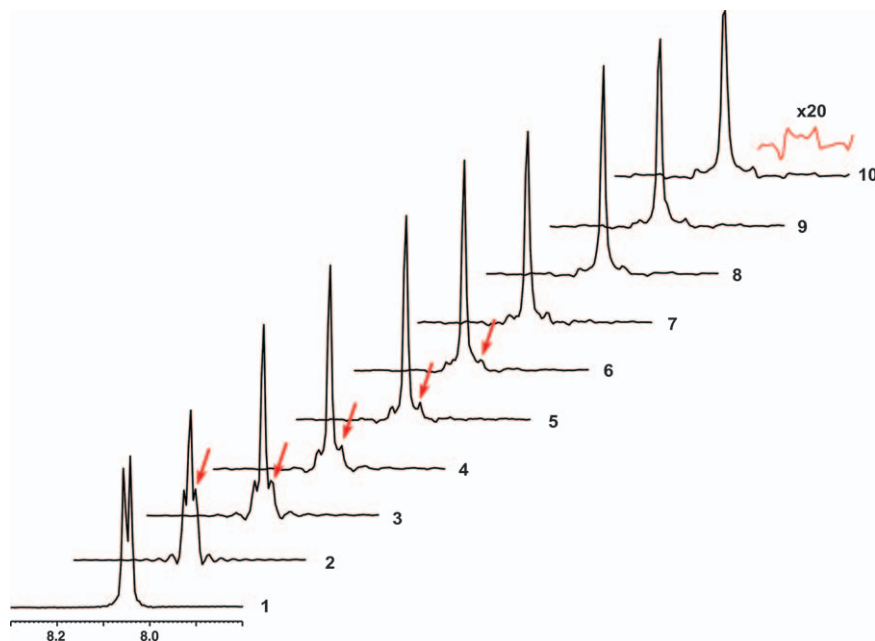


Figure 9.8 Series of traces from PS-HSQC spectra of strychnine (**1**) acquired by successively varying the setting of the loop counter n , which is used to control the number of loops through the BIRD-based decoupler element used during acquisition. The H4/C4 doublet was chosen for this illustration. As will be noted in the first trace ($n=1$), there is effectively no decoupling. With $n=2$, collapse of the multiplet has begun, effectively reducing the resonance to a pseudo-triplet. In successive experiments, the outer limbs of the pseudo-triplet diminish, collapsing into the decoupled resonance at the chemical shift of H4. We have found that a value in the range of $n=6-8$ provides satisfactory decoupling for standard acquisition times ($t_2=150-300$ ms). While it might be tempting to resort to higher values of n , there is a trade-off between the number of decoupling loops, the line width, and baseline noise. When $n=10$, the H4 resonance has visibly broadened, which is costing s/n, and the baseline noise is increasing as shown in the $\times 20$ vertical expansion to the right of the resonance.

which will likely be impossible in many scarce sample situations. Incorporating intentional F_1 aliasing combined with NUS provides another very significant time saving. By combining these techniques, we have shown that a multiplicity-edited PS-HSQC spectrum of a $3\ \mu\text{g}$ sample of a 661 Da metabolite can be acquired in just over 14 hours.³⁴ Similar work using more conventional HSQC methods has been demonstrated by Molinski and coworkers for a number of marine natural products for which samples of only a few micrograms were available.³⁵⁻³⁹ Indeed, in a benchmarking exercise by one of the current authors, it was demonstrated that HSQC data on samples of strychnine as small as 470 ng can be acquired over a long weekend using a 1.7 mm MicroCryoProbeTM.¹⁴

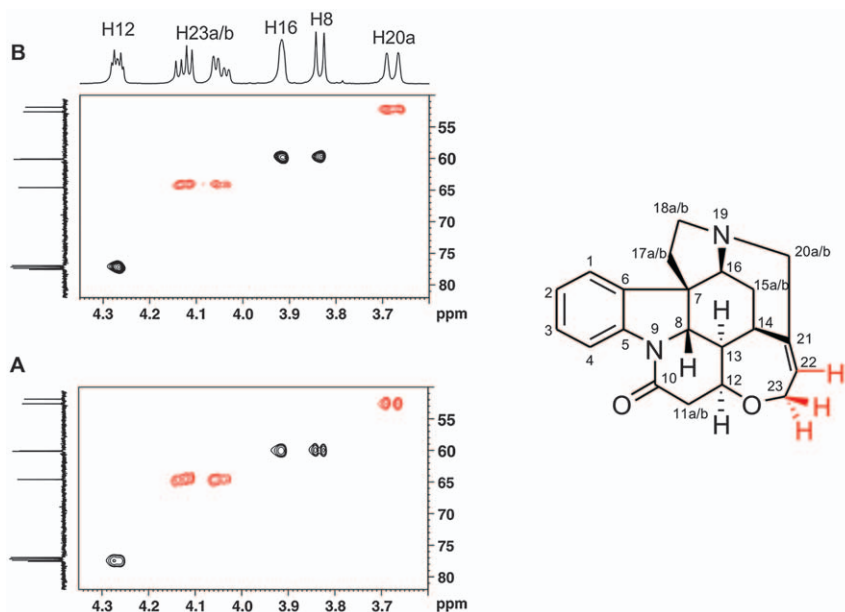


Figure 9.9 Segments of the multiplicity-edited HSQC and multiplicity-edited PS-HSQC spectra of strychnine (**1**). (A) Segment of the aliphatic region of the multiplicity-edited HSQC spectrum shown in Figure 9.5 containing the H12, H23a/b, H16, H8, and H20a resonances. (B) Segment showing the same region from the multiplicity-edited PS-HSQC spectrum recorded using the pulse sequence shown in Figure 9.11. Obviously, in such situations, it behooves an investigator to resort to the smallest-diameter NMR probe. Full homonuclear couplings are retained for the correlations in the spectral segment from the multiplicity-edited HSQC spectrum shown in (A) (see also Figures 9.10 and 9.12B). In contrast, in (B), homonuclear vicinal couplings are collapsed, while geminal $^2J_{\text{HH}}$ couplings are unaffected by the BIRD-based decoupling (see also Figures 9.10 and 9.12C). Hence, the correlations for the H23a/b resonances in (A) are simplified from a pair of doubled doublets to geminal doublets in (B). Since the H20a/b resonances are isolated with no vicinal coupling partners, the correlations remain as unchanged geminal doublets in the multiplicity-edited PS-HSQC spectrum.

9.6 Intentionally Folding HSQC Spectra in Limited-sample Situations

In many instances, highly interesting natural product samples may be available in only severely limited quantities. Obviously, in such situations, it behooves an investigator to resort to the smallest-diameter NMR probe available or to run smaller diameter tubes coaxially in a larger probe (*e.g.* running a 3 mm tube in a 5 mm cryoprobe).^{12–15} Another approach that can be used to the investigator's advantage is intentionally folding any of the various HSQC spectra in the F_1 frequency domain. In the aliphatic region of

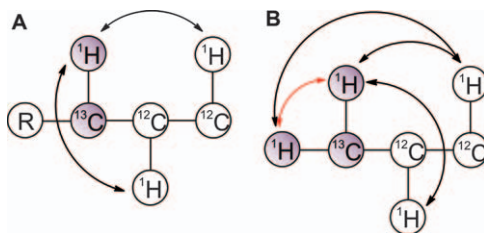


Figure 9.10 (A) A hypothetical structure fragment with a methine proton on ^{13}C and the protons to which it is coupled located on ^{12}C . The BIRD pulse during the acquisition in the PS-HSQC experiment will invert the ^{13}C proton(s), while the hard 180° pulse refocusing the ^1H - ^{12}C proton(s) effectively decouples the methine resonance, which is exactly analogous to the H4 resonance of strychnine (**1**) used as an example in Figure 9.8. These homonuclear couplings are denoted by black arrows. (B) In the case of anisochronous methylene proton resonances, both resonances are on the same ^{13}C resonance and hence the geminal coupling between them (red double-headed arrow) will not be refocused by the BIRD element. As in the example shown in (A), the remote protons on ^{12}C will be refocused by the BIRD element, resulting in the simplification of the methylene resonance to a pair of geminal doublets.

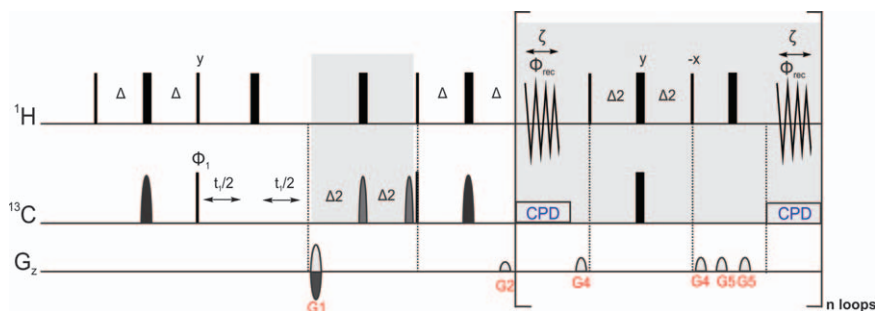


Figure 9.11 Pulse sequence incorporating multiplicity editing into PS-HSQC. Multiplicity editing is afforded by the pulse sequence elements in the first shaded block. BIRD-based homodecoupling is provided by the second shaded block.

most NMR spectra, the region downfield of ~ 80 – 85 ppm in F_1 will be devoid of responses. Likewise, in the aromatic region, there are only very rarely correlations upfield of 100 ppm in F_1 . Consequently, when the full F_1 spectral width is digitized in a HSQC spectrum, there are significant regions in the F_1 dimension that are being digitized for no purpose. In such instances, an investigator has the option of intentionally folding the spectrum in F_1 . There are no F_1 filters, so resonances downfield or upfield of the digitized frequency range will be aliased back into the frequency range that is being digitized.^{40,41} In the case of strychnine (**1**), if the spectral width in F_1 is set to cover from the upfield C15 resonance to just below the C12

resonance (e.g. 10–90 ppm), the C22 vinyl and four protonated aromatic carbons will be aliased and observed in the digitized F_1 frequency range.

Intentionally aliasing or “folding” spectra in F_1 provides a means of obtaining higher digital resolution in F_1 where it may be needed or of obtaining higher s/n ratios by being able to accumulate more transients/ t_1 increment in a given block of spectrometer time. Figure 9.13 shows a multiplicity-edited PS-HSQC spectrum of strychnine (**1**) with the F_1 spectral width intentionally narrowed to the range of 10–90 ppm.

When the spectrum of strychnine (**1**) shown in Figure 9.5 was acquired, the transmitter was located at ~ 80 ppm. As will be noted from that spectrum, the C3 and C22 resonances, which are the furthest downfield, were

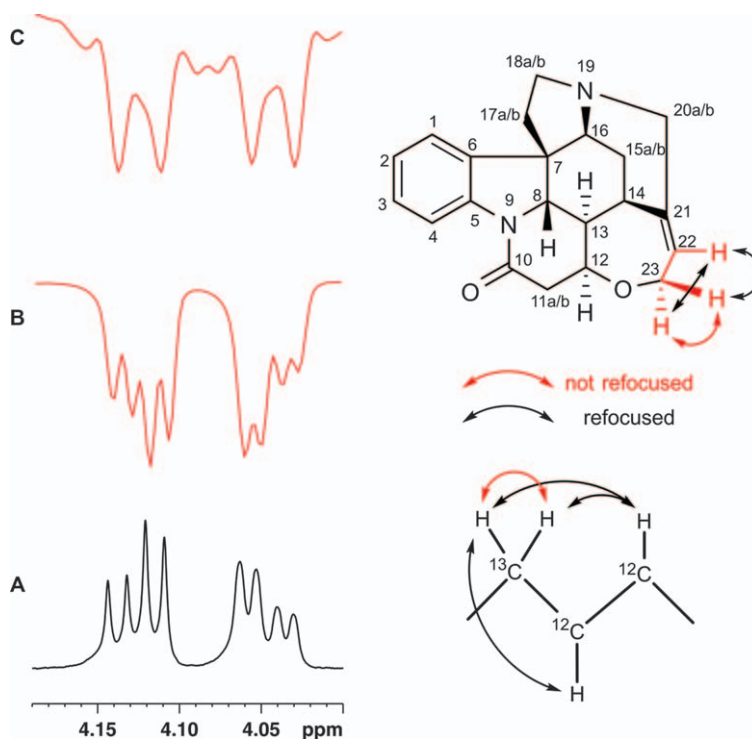


Figure 9.12 Comparison of multiplicity-edited HSQC and multiplicity-edited PS-HSQC results. (A) Proton reference spectrum of the H23a/b resonances of strychnine (**1**). (B) Slice at the F_1 frequency of the H23/C23 correlation from the multiplicity-edited HSQC spectrum of **1** shown in Figure 9.5. (C) Corresponding H23/C23 correlation slice from the multiplicity-edited PS-HSQC spectrum (Figure 9.11). The numbering scheme and structure of **1** are shown in the right panel. As would be expected from the hypothetical structural fragment shown in the bottom-right panel, $^nJ_{\text{HH}}$ couplings where $n > 2$ will be refocused and decoupled. In contrast, geminal $^nJ_{\text{HH}}$ couplings where $n = 2$ will not be refocused and will remain in the spectrum. Hence the H23a/b multiplets will be simplified from doubled doublets to geminal doublets in the multiplicity-edited PS-HSQC spectrum as shown in Figure 9.9.

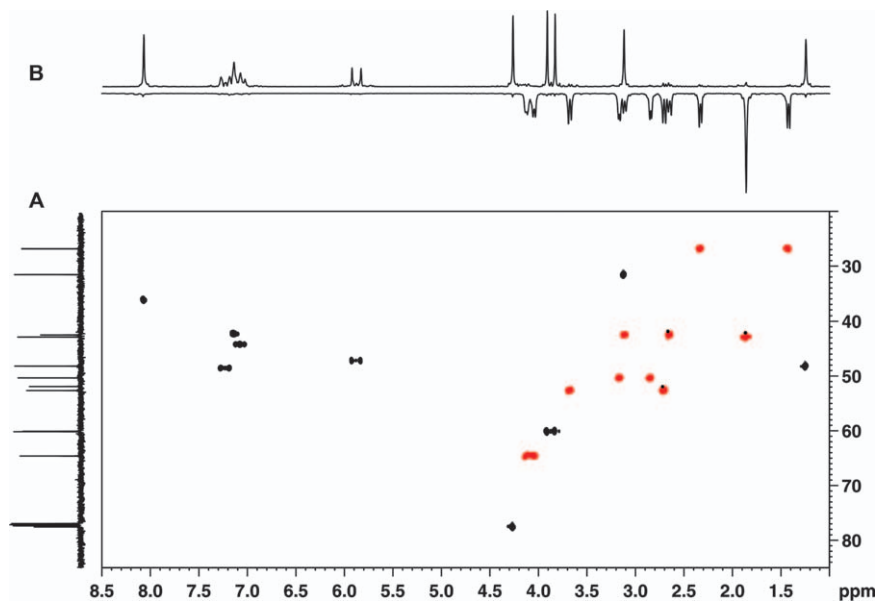


Figure 9.13 Intentionally folded multiplicity-edited PS-HSQC spectrum of strychnine (**1**). (A) Multiplicity-edited PS-HSQC spectrum acquired with an F_1 spectral width set to 10–90 ppm. Aliphatic resonances will be observed at their intrinsic F_1 chemical shifts. Aromatic resonances will be aliased back into the aliphatic region of the spectrum and will appear X ppm downfield of the upfield limit of the spectrum in F_1 , where X corresponds to the difference between the actual chemical shift and the 90 ppm chemical shift range selected. (B) Positive/negative projection of the F_2 dimension. Note, for example, that the C22 and C3 resonances are obvious doublets in F_2 (see Figure 9.14), whereas they should be collapsed to singlets by the BIRD-based decoupling.

both still fully decoupled in F_2 . However, when the transmitter was shifted to 45 ppm as shown in Figure 9.14 and the F_1 spectral width narrowed to 10–90 ppm, the decoupling efficiency of the GARP composite pulse decoupling scheme used to acquire the data shown in Figure 9.13 was insufficient to fully collapse the $^1J_{CH}$ coupling in F_2 , leading to the 53 and 49 Hz doublets observed for C3 and C22, respectively. Changing the decoupling from GARP to Bi-lev again fully collapsed the $^1J_{CH}$ couplings in F_2 as shown in Figure 9.14C. While intentional folding is a useful tool, investigators should be careful to ensure that the efficiency of the X-nucleus decoupling scheme employed has sufficient bandwidth to accommodate the offset of the protonated carbon resonances furthest away from the transmitter (Figure 9.15).

9.7 Non-uniform Sampling

Another approach that can be employed strategically to conserve spectrometer time, to increase s/n for a given block of instrument time, or to

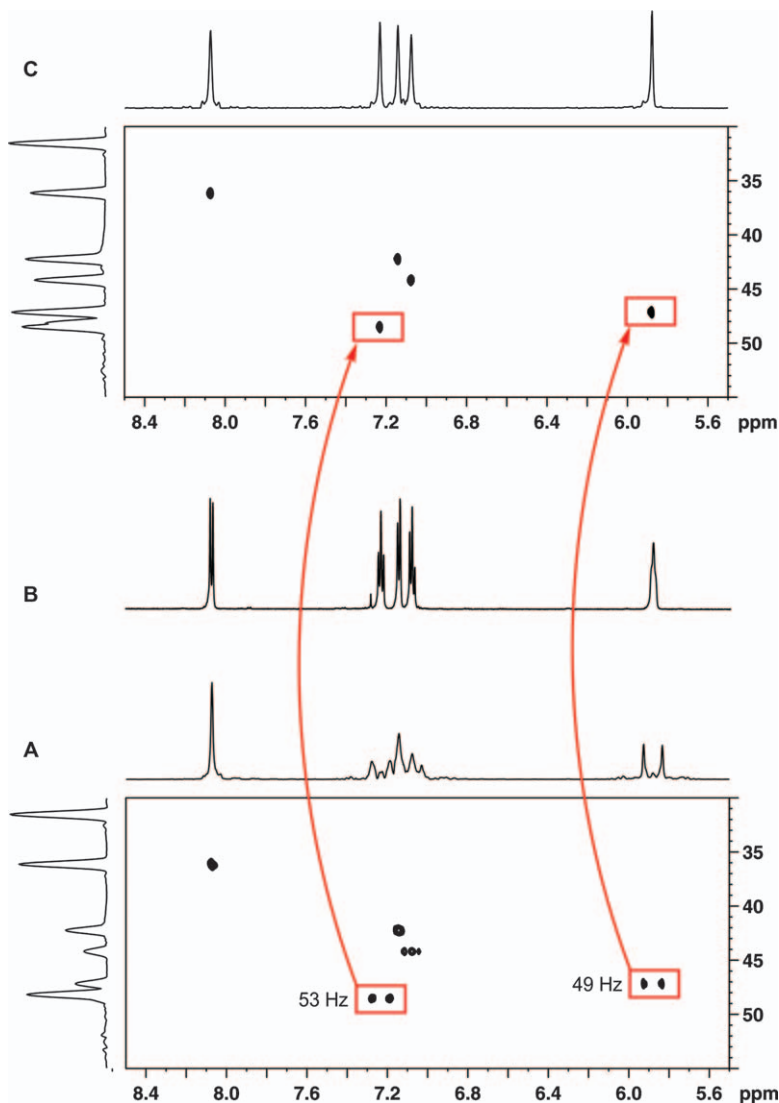


Figure 9.14 (A) Expansion of the aromatic region of the intentionally folded multiplicity-edited PS-HSQC spectrum of **1** shown in Figure 9.13. Note that the C3 and C22 resonances are observed as 53 and 49 Hz doublets in F_2 , respectively. (B) High-resolution proton spectrum of **1**. (C) Repeat of the experiment shown in Figure 9.13 with Bi-level X-nucleus decoupling employed rather than the GARP decoupling used to acquire the data shown in (A). The F_2 projection is shown above both (A) and (C). Note that all resonances are fully decoupled in the projection above (C) as a result of the change in the decoupling scheme.

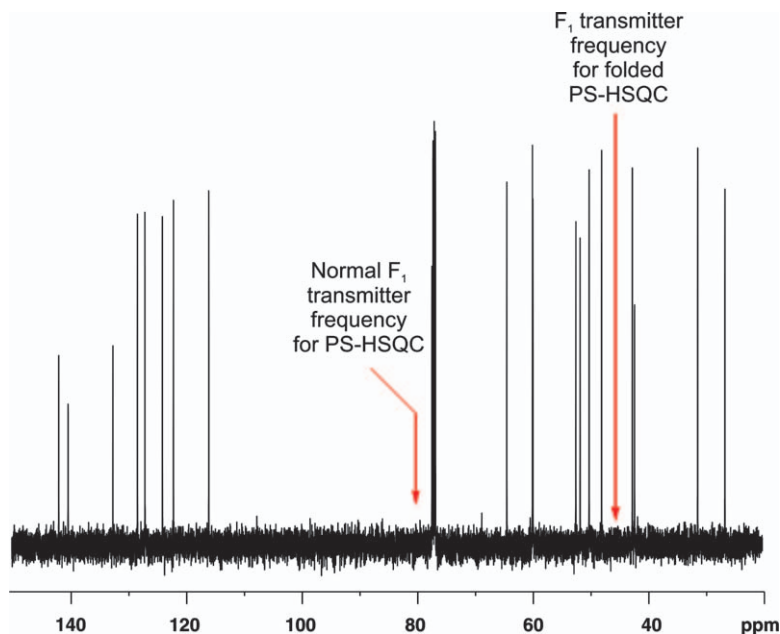


Figure 9.15 ^{13}C reference spectrum of strychnine (**1**) showing the transmitter location for a normal full F_1 spectral width PS-HSQC spectrum *versus* the location of the transmitter for the folded spectrum shown in Figure 9.13A. The C3 and C22 carbons are the methine carbons furthest downfield of the protonated carbons and are not completely decoupled when GARP decoupling was employed to generate the data in Figures 9.13 and 9.14A.

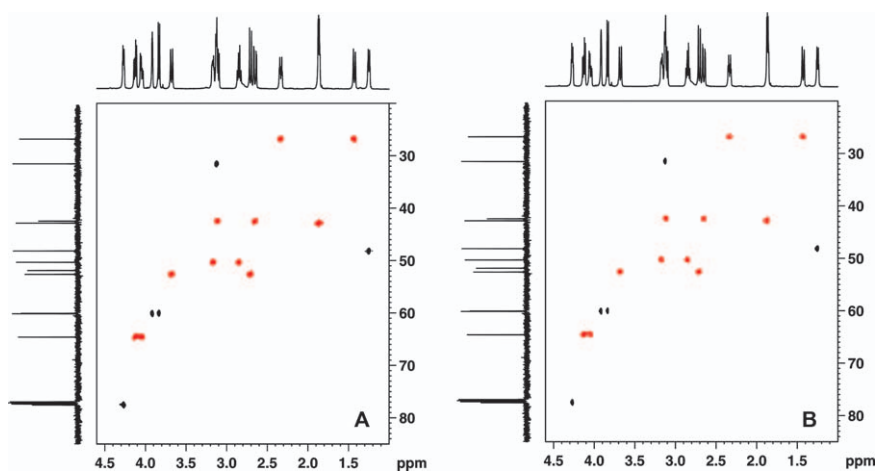


Figure 9.16 Comparison of conventional (linear) F_1 sampling *versus* 20% NUS sampling. Both spectra were acquired as $3\text{K} \times 256$ points. (A) A conventionally sampled spectrum acquired in 25 minutes. (B) A 20% Sparse sampled spectrum acquired in 5 minutes. Processing of both spectra was identical and the plotted spectra are, in practical terms, indistinguishable from one another.

significantly improve F_1 digital resolution is through the use of non-uniform sampling or NUS. A comprehensive discussion of NUS is beyond the scope of this chapter, but the interested reader is referred to the Chapter 6 in Volume 1 of this series by Rovnyak *et al.*⁴² For the purposes of this chapter, we will simply show a comparison of two multiplicity-edited PS-HSQC strychnine spectra, one acquired as $3\text{K} \times 256$ points with conventional, linear sampling in F_1 and a second acquired as $3\text{K} \times 256$ points with 20% sampling in F_1 . The former was acquired in 25 minutes while the latter required only 5 minutes. The more congested aliphatic regions of the two spectra are compared side-by-side in Figure 9.16.

9.8 Long-range Heteronuclear Shift Correlation

There are numerous proton-detected methods now available for long-range heteronuclear shift correlation. These methods have been discussed in a number of reviews and the interested reader is referred to these for a more in-depth treatment of long-range heteronuclear correlation methods than is possible in this chapter.^{43–48} The oldest and quite likely still most widely used long-range heteronuclear shift correlation experiment is the HMBC experiment described in 1986 by Bax and Summers.⁴⁹ Indeed, the seminal report of this experiment may be the most highly cited NMR paper ever written. Several years after the introduction of the HMBC experiment, a decoupled variant, D-HMBC,⁵⁰ was also introduced, although it has never enjoyed the popularity of the HMBC experiment.⁴⁹ During the period from 1998 to 2000, a number of investigations of “accordion”-optimized long-range correlation experiments were reported, beginning with ACCORD-HMBC^{51,52} and followed by further modifications that included IMPEACH-MBC⁵³ and then CIGAR-HMBC.⁵⁴ Readers interest in an in-depth treatment of the various analogs of the HMBC and related experiments are referred to the three comprehensive reviews by Furrer.^{46–48} Accordion optimization has been demonstrated to have an advantage for the acquisition of long-range ^1H - ^{15}N heteronuclear shift correlation spectra, which is considered in Chapter 3, Section 3.8.4.3 due to the variability of $^nJ_{\text{NH}}$ long-range coupling constants. For ^1H - ^{13}C correlation experiments, however, we will focus on the HMBC⁴⁹ and LR-HSQMBC⁵⁵ experiments in the following sections.

9.8.1 HMBC

The pulse sequence for the HMBC experiment, modified to include gradients, is shown in Figure 9.17. The pulse sequence is relatively simple and very robust. The shape of the FID from a HMBC experiment is that of a spin-echo (see Figure 9.18) and it is advisable to digitize in F_2 with a minimum of 4K points since there is no X-nucleus decoupling used in the experiment. The processed data are mixed phase and are typically presented as the magnitude-calculated spectrum, although there are phase-sensitive versions of the experiment.

The 8 Hz-optimized HMBC spectrum of strychnine is shown in Figure 9.19. The data were acquired as $4\text{K} \times 128$ points, with eight transients



Figure 9.17 The HMBC pulse sequence first described in 1986 by Bax and Summers.⁴⁹ The initial 90° ^1H pulse is followed by a delay, Δ , and then a 90° ^{13}C pulse, which serves as the low-pass J -filter to suppress unwanted $^1J_{\text{CH}}$ doublets flanking the proton chemical shift. The delay should be set to $\Delta = 1/2(^1J_{\text{CH}})$; 145 Hz is generally a good compromise value. The delay $\Delta 2$ following the first ^{13}C pulse is the long-range delay, $\Delta 2 = 1/2(^nJ_{\text{CH}})$, and should generally be set in the range of ~ 3 –10 Hz. Phase cycling is simple: $\Phi_1 = 0, 2$; $\Phi_{\text{rec}} = 0, 2$, and as would be expected, the gradient ratio must be adjusted as a function of the heteronuclei, hence $G1 : G2 : G3 = 50 : 30 : 40.1$ for ^1H – ^{13}C and $70 : 30 : 50.1$ for ^1H – ^{15}N . No decoupling is applied during acquisition and it is advisable to acquire the data using at least 4K points.

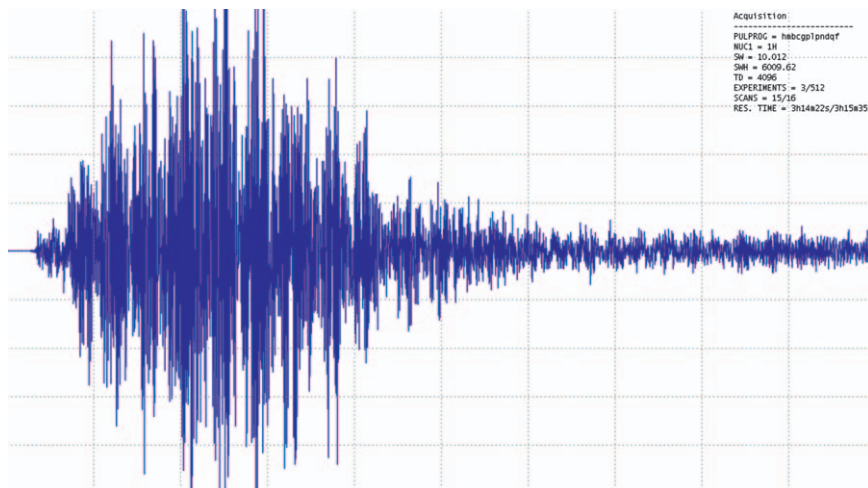


Figure 9.18 FID from an 8 Hz-optimized HMBC spectrum. The data were acquired as 4K points with an acquisition time of 340 ms.

accumulated/ t_1 increment, giving an acquisition time of 25 minutes. Delays were optimized for 145 and 8 Hz for the $^1J_{\text{CH}}$ and $^nJ_{\text{CH}}$ couplings, respectively. Data were linear predicted in F_1 to 512 points. Both frequency domains were subjected to $\pi/2$ phase-shifted sine bell apodization.

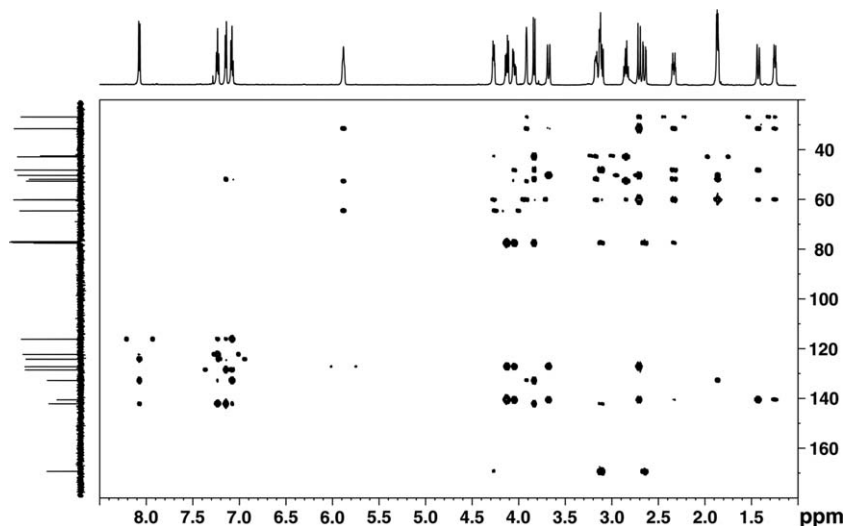
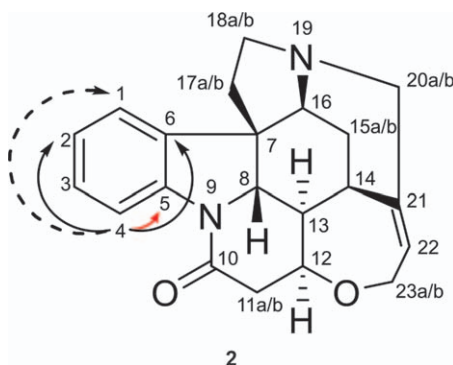


Figure 9.19 HMBC spectrum of strychnine (**1**) optimized for 8 Hz and acquired as $4\text{K} \times 128$ points. The data were linear predicted to 512 points in F_1 , subjected to $\pi/2$ phase-shifted sine bell apodization and zero-filled to afford the $2\text{K} \times 1\text{K}$, magnitude-calculated spectrum shown.

The optimization of the long-range delay in the HMBC experiment based on $^nJ_{\text{CH}}$ does not lead to a spectrum in which all correlations have a particular coupling pathway as is the case in HSQC experiments where all correlations are one-bond *via* $^1J_{\text{CH}}$. Instead, correlations in the HMBC spectrum are indeterminate insofar as the long-range coupling pathway length, although $^3J_{\text{CH}}$ correlations do tend to predominate. Correlations *via* $^2J_{\text{CH}}$ and $^4J_{\text{CH}}$ are not uncommon and still-longer correlations have been observed in HMBC spectra.^{56,57}

To examine in more detail part of the data shown in Figure 9.20, consider the correlations associated with the H4 aromatic proton resonance shown in **2**.



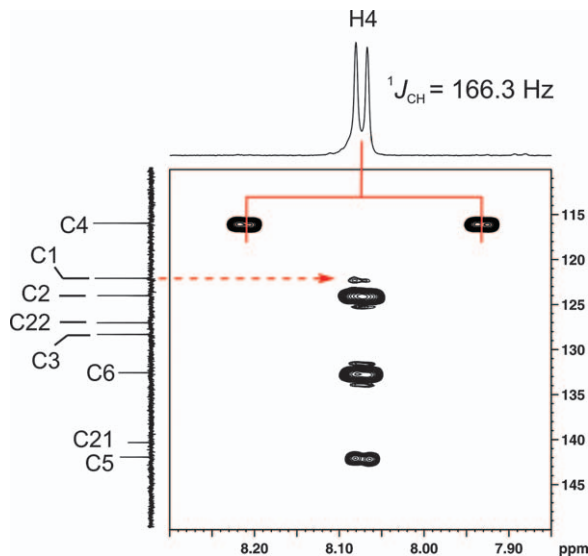


Figure 9.20 Segment of the 8 Hz-optimized HMBC spectrum of strychnine (**1**) showing correlations associated with the H4 resonance. As is readily apparent, the low-pass J -filter was not completely effective at suppressing the $^1J_{\text{CH}}$ 166.3 Hz doublet associated with $^1J_{\text{H4C4}}$. Although such residual one-bond doublets can be problematic in congested regions of the spectrum, they also conveniently provide one-bond coupling constant information. The two strongest long-range correlations are *via* $^3J_{\text{CH}}$ to C2 and C6. A slightly weaker $^2J_{\text{CH}}$ correlation to C5 is observed, as is a weak $^4J_{\text{CH}}$ correlation to C1. These correlations are summarized by **2**.

As would be expected, changing the long-range optimization affords different correlations. Adjusting the $\Delta 2$ delay for a smaller coupling (e.g. 3–4 Hz) can afford longer-range correlations that are not observed when the experiment is optimized in the range of perhaps 8–10 Hz. The differences in the various couplings observed with **2** and 8 Hz optimization are shown in Table 9.1. Quick inspection of Table 9.1 shows that for $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ correlations, the numbers of correlations observed are quite comparable. For $^4J_{\text{CH}}$ and $^5J_{\text{CH}}$ correlations, there are markedly more correlations observed for 2 Hz optimization than for 8 Hz optimization. Both optimizations afford a single $^6J_{\text{CH}}$ correlation.

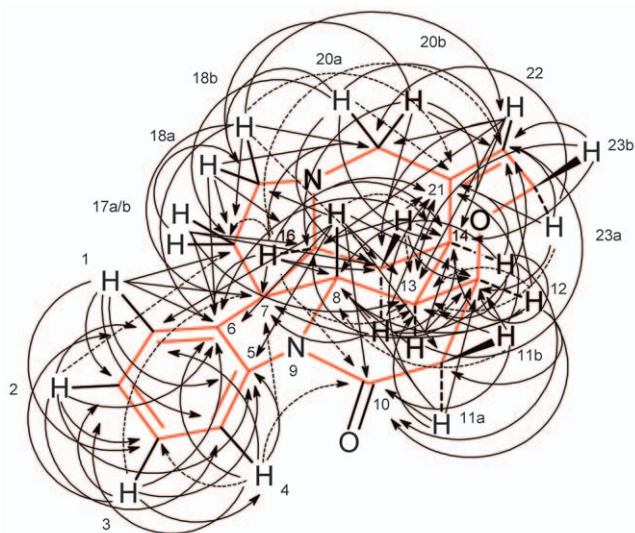
If all of the 129 correlations observed in the 8 Hz-optimized HMBC spectra are superimposed on the structure, there is a bewildering array of correlation information, as shown in Figure 9.21.

9.8.2 LR-HSQMBC

The LR-HSQMBC⁵⁵ experiment has only very recently been reported, in contrast to the very-well-establish HMBC experiment. The experiment is

Table 9.1 Comparison of the various long-range correlations observed in 2 Hz- versus 8 Hz-optimized HMBC spectra.

	$^2J_{\text{CH}}$	$^3J_{\text{CH}}$	$^4J_{\text{CH}}$	$^5J_{\text{CH}}$	$^6J_{\text{CH}}$	Total $>^3J_{\text{CH}}$	Total
2 Hz	34	54	43	10	1	54	142
8 Hz	35	53	36	4	1	41	129

**Figure 9.21** Optimization of a HMBC spectrum of strychnine for 8 Hz yields 129 long-range correlations (see Table 9.1). When all of the observed correlations are superimposed on the structure, the volume of information that can be derived from a single spectrum can be bewildering. However, when the information is sorted into two frequency domains, it is manageable and forms a cornerstone of modern structure elucidation strategies.

derived from the HSQMBC experiment, which is a single quantum-based method to establish long-range heteronuclear correlations. For further details on the HSQMBC experiment, there are several very good reviews.^{58,59} The pulse sequence for the LR-HSQMBC experiment is shown in Figure 9.22.

The LR-HSQMBC experiment begins and ends with INEPT pulse sequence elements. The refocused nature of the experiment provides in-phase magnetization, which avoids antiphase cancellation of very small correlations, which can be problematic when HMBC data are magnitude calculated for presentation. Refocusing of magnetization allows the application of X-nucleus decoupling during acquisition. Thus, while LR-HSQMBC may not equal the sensitivity of a HMBC experiment when both are optimized for 8 Hz, in contrast, LR-HSQMBC does afford significantly better performance

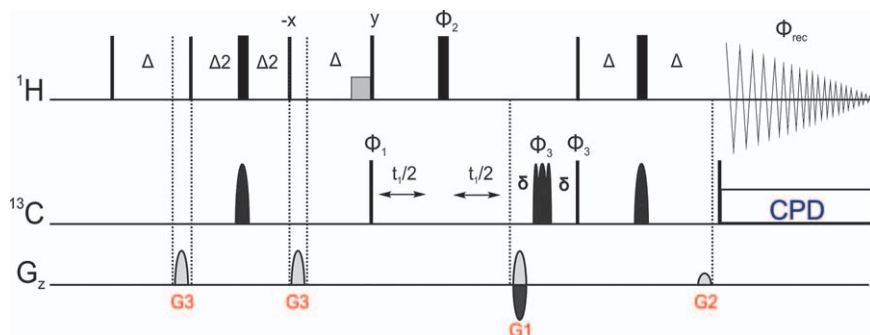


Figure 9.22 Narrow bars correspond to 90° pulses and wide bars correspond to 180° pulses. All ^{13}C pulses were applied as adiabatic CHIRP pulses. The pulse sequence begins with a $^nJ_{\text{CH}}$ transfer of magnetization to ^{13}C ; the subsequent G-BIRD_{R,X} sequence serves to suppress unwanted $^1J_{\text{CH}}$ responses. The delays are optimized as $\Delta = 1/4(^nJ_{\text{CH}})$ and $\Delta 2 = 1/4(^1J_{\text{CH}})$. Phases are cycled as: $\Phi_1 = 0, 2$; $\Phi_2 = 0, 0, 2, 2$; $\Phi_3 = 0, 0, 0, 0, 2, 2, 2, 2$; and $\Phi_{\text{rec}} = 0, 2, 0, 2, 2, 0, 2, 0$. The nuclide-dependent gradient ratios are $G1:G2 = 80:20.1$ for $^1\text{H}-^{13}\text{C}$ and $80:8.1$ for $^1\text{H}-^{15}\text{N}$. $G3 = 17$. A clean in-phase pulse sequence element was included in the original report of the experiment,⁵⁵ but has since been found to be unnecessary.

when both experiments are compared when optimized for 2 Hz. In part, the better performance of LR-HSQMBC when optimized for 2 Hz can be attributed to the collapse of the heteronuclear coupling, which would be expected to partially recover some of the signal lost in the long delays when the experiment is optimized for 2 Hz.

The 2 Hz-optimized LR-HSQMBC spectrum of strychnine is shown in Figure 9.23. The data were acquired using 512 increments in the second frequency domain. To observe small, very-long-range correlations, it is necessary to more heavily digitize in F_1 than would be typical with a HMBC spectrum optimized in the range of perhaps 6–8 Hz. In some cases, such as that of the xanthone antibiotic cervinomycin A_2 , which was also employed as a model compound in the initial report of the LR-HSQMBC experiment, there was still considerable signal after 768 increments of the evolution time, and indeed, some of the very-long-range correlations were not observed until very late in the digitization of F_1 .

When the 2 Hz-optimized LR-HSQMBC, D-HMBC, and HMBC spectra were compared, as is noted in Table 9.2, the best results insofar as very-long-range correlations were obtained with the LR-HSQMBC data, followed by HMBC. It is also interesting to compare the correlations observed in the LR-HSQMBC data that were not observed in the HMBC data. This comparison is shown in Figure 9.24. Correlations shown in red were observed in an 8 Hz-optimized HMBC spectrum, but not in the 2 Hz LR-HSQMBC data. The correlations shown in black were not observed in either 2 or 8 Hz-optimized HMBC spectra.

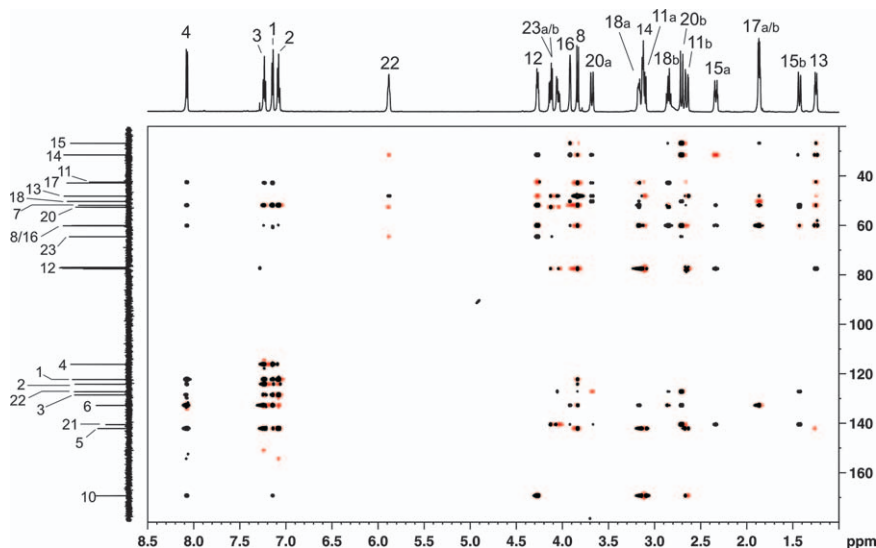


Figure 9.23 LR-HSQMBC spectrum of a 4.5 mg sample of strychnine (**1**). The data were acquired as 4096×512 points, accumulating 32 transients/ t_1 increment, with an acquisition time of 10 hours and 45 minutes. It should be noted that all correlations were observable after 2 hours and 45 minutes of accumulation, although the very-long-range correlations were extremely weak so early in the acquisition.

Table 9.2 Comparison of 2 Hz-optimized LR-HSQMBC, D-HMBC, and HMBC data for strychnine (**1**).

	$^2J_{\text{CH}}$	$^3J_{\text{CH}}$	$^4J_{\text{CH}}$	$^5J_{\text{CH}}$	$^6J_{\text{CH}}$	Total $>^3J_{\text{CH}}$	Total
2 Hz HMBC	34	54	43	10	1	54	142
2 Hz LR-HSQMBC	33	59	55	11	2	68	160
2 Hz D-HMBC	29	43	34	8	2	44	115

Another example of the “reach” of the LR-HSQMBC experiment when optimized for 2 Hz is found in the recent revision of the structure of the *Cryptolepis* alkaloid cryptospirolepine.⁶⁰ The object of acquiring the data was to probe for the $^2J_{\text{CH}}$ correlation between the vinyl proton and the adjacent carbonyl (see inset in Figure 9.25). Correlations were observed from both methyl groups *via* $^6J_{\text{CH}}$ and $^7J_{\text{CH}}$ coupling pathways and there was a hint of the $^2J_{\text{CH}}$ correlation sought in the study. Based on DFT calculations, the $^6J_{\text{CH}}$ and $^7J_{\text{CH}}$ couplings observed had calculated coupling constants of -0.27 and 0.31 Hz, respectively. In contrast, the $^2J_{\text{CH}}$ coupling between the vinyl proton and the adjacent carbonyl was calculated to be only 0.03 Hz. Because of the very weak potential correlation between the vinyl and carbonyl, the unequivocal solution of the structure was based on a newly developed 1,1-HD-ADEQUATE experiment (see Chapter 10, Section 10.1.2.2).

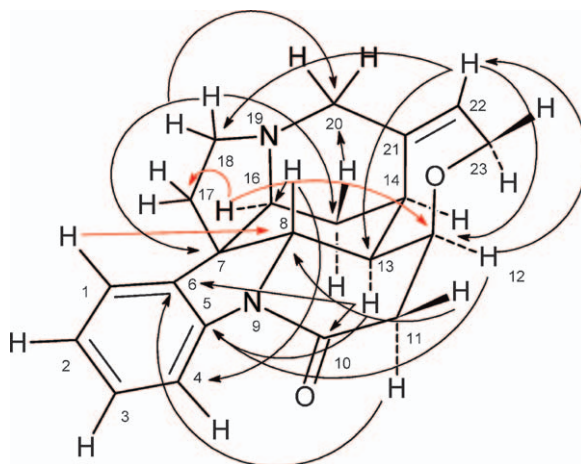


Figure 9.24 Comparison of correlations observed in LR-HSQMBC *versus* HMBC spectra. The black arrows denote correlations observed in the 2 Hz LR-HSQMBC spectrum that were not observed in the 2 or 8 Hz HMBC data. The red correlations were observed in the 8 Hz HMBC data, but not in the 2 Hz LR-HSQMBC spectrum.

9.9 Defining the Proton–Proton Connectivity Network

A key facet of structure characterization involves piecing together the proton–carbon fragments defined in the HSQC spectrum. Classically, COSY has been used to establish proton–proton homonuclear coupling pathways in structure elucidation exercises. We will first consider COSY spectra, followed by first homonuclear TOCSY and HSQC-TOCSY spectra, the latter being used to sort proton–proton homonuclear correlations by ^{13}C chemical shifts in the second dimension. COSY and TOCSY, of course, are much more sensitive experiments than HSQC-TOCSY and should be the first choice for defining the proton homonuclear coupling network. However, when molecules are being studied with complex, overlapped proton resonances that can lead to ambiguities in the structure being elucidated, the less sensitive HSQC-TOCSY spectrum is a powerful tool.

9.9.1 COSY and TOCSY Spectra

For molecules with a well-resolved proton spectrum, COSY provides an effective means of establishing the proton homonuclear coupling network. The simple COSY pulse sequence is shown in Figure 9.26. Strychnine (**1**) is certainly an example of a molecule with a well-resolved proton spectrum, and the COSY spectrum of strychnine is shown in Figure 9.27.

Correlations that can be extracted from the spectrum shown in Figure 9.27 (*e.g.* for the H13 resonance) are shown in the annotated expansion shown in

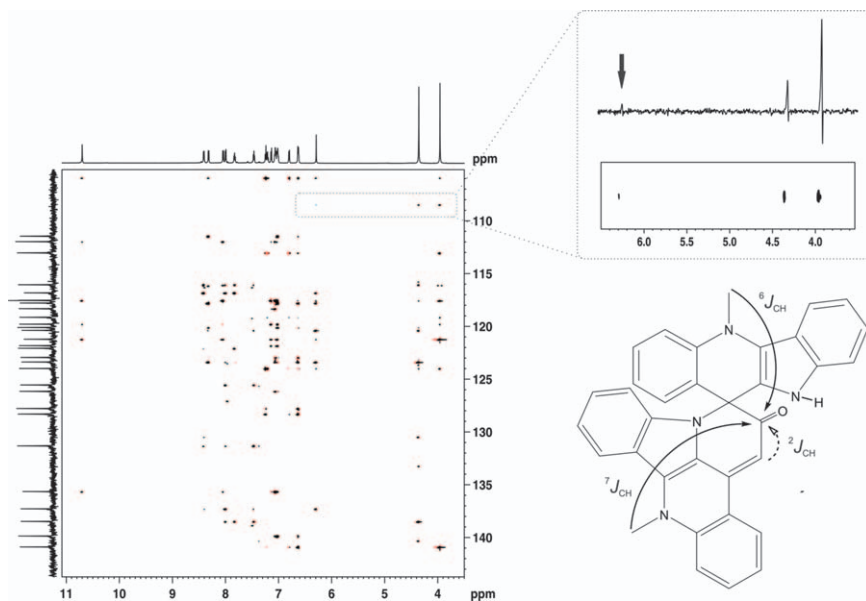


Figure 9.25 LR-HSQMBC spectrum of cryptospirolepine (structure shown in insert) optimized for 2 Hz. The data were acquired as 768 increments on a sample of 700 μg of the alkaloid dissolved in 40 μL DMSO- d_6 using a 600 MHz 1.7 mm MicroCryoProbeTM. The inset shows the intentionally folded correlations to the carbonyl from the two *N*-methyl groups via $^6J_{\text{CH}}$ and $^7J_{\text{CH}}$ coupling pathways and, designated by the arrow, a hint of the $^2J_{\text{CH}}$ correlation sought in the experiment. Based on DFT calculations, the coupling constants for the three correlations were predicted to be -0.27 and 0.31 Hz for the two methyls and a tiny 0.03 Hz for the $^2J_{\text{CH}}$ correlation from the vinyl proton. The weak nature of the $^2J_{\text{CH}}$ correlation mandated an alternative approach to the revision of the structure that utilized the newly developed 1,1-HD-ADEQUATE experiment.⁶⁰

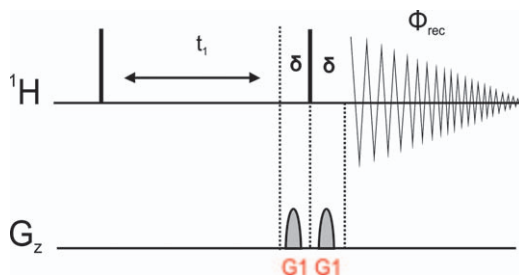


Figure 9.26 The COSY experiment is one of the simplest 2D NMR experiments, involving only a pair of proton 90° pulses flanking the evolution period, t_1 .

Figure 9.28 and by 3. Note that the H13 and H8 protons have a *trans*-diaxial arrangement in the structure, while H13 is *gauche* to H14 and H12. The spatial arrangement of these protons carries with it the expectation of a strong H13–H8 correlation and somewhat weaker correlations for H13–H12 and H13–H14, which is borne out by the intensity of the various resonances in the slice from the COSY spectrum plotted above the high-resolution reference spectrum in Figure 9.28.

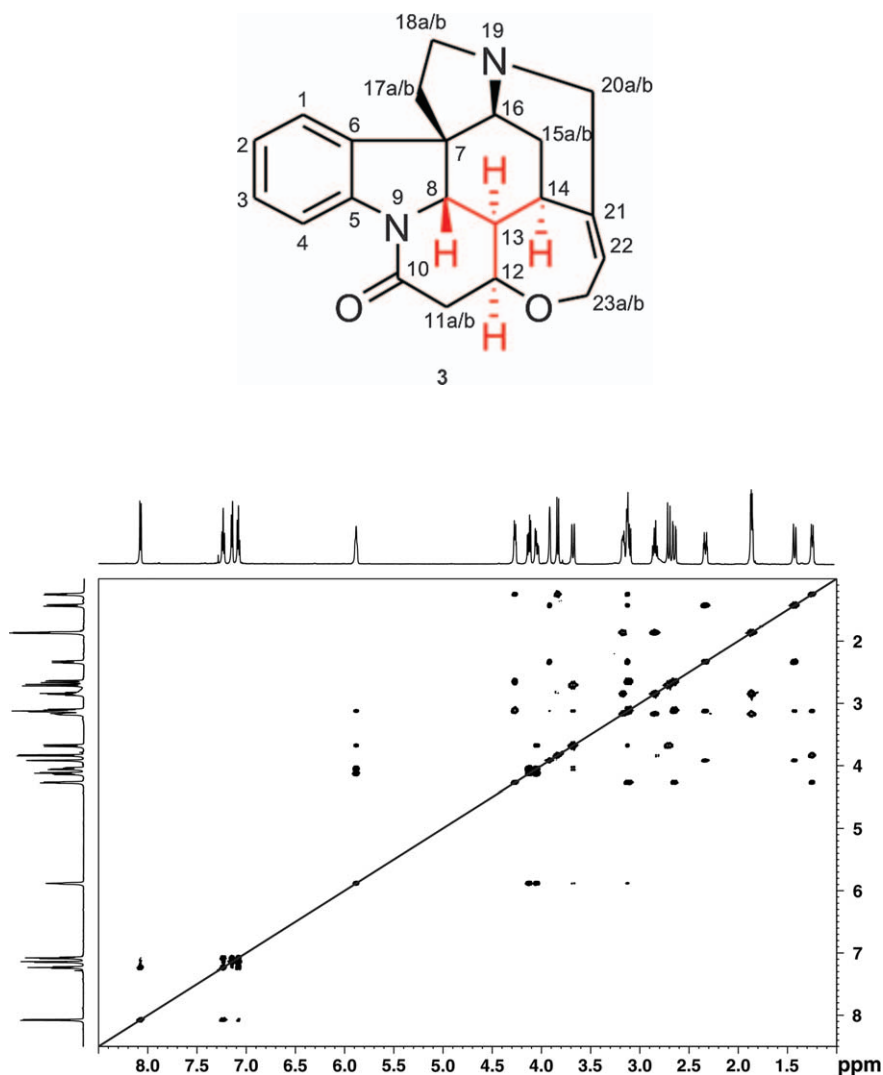


Figure 9.27 COSY spectrum of strychnine (1) acquired as $2K \times 256$ points, linear predicted and zero-filled to afford the $2K \times 2K$ contour plot shown. This familiar spectrum establishes proton–proton homonuclear correlations that are symmetrical about the diagonal of the spectrum.

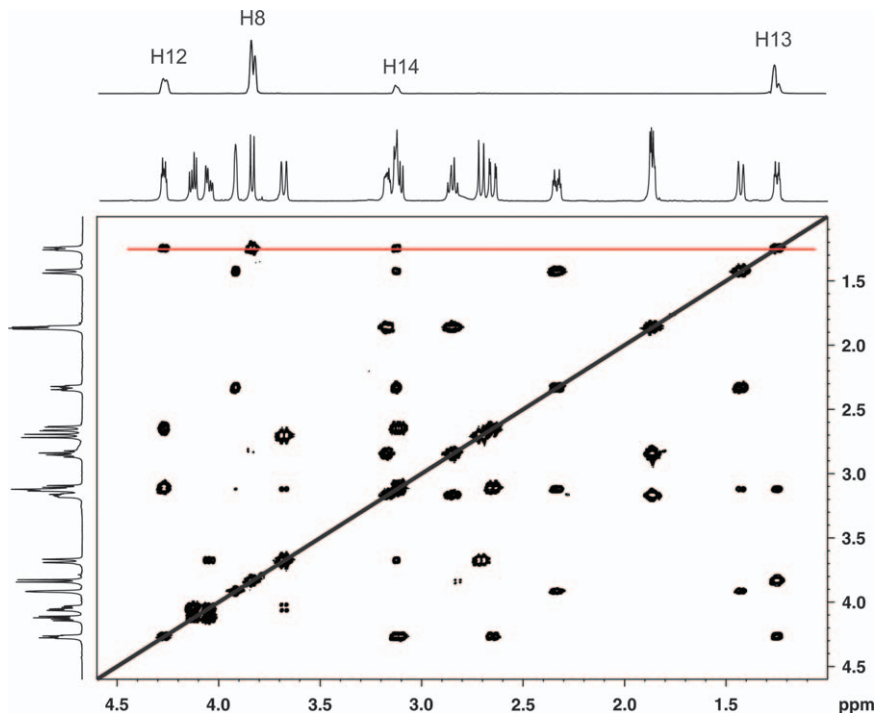


Figure 9.28 Expansion of the aliphatic region of the COSY spectrum of strychnine shown in Figure 9.27. The slice extracted from the spectrum at the F_1 chemical shift of the H13 methine resonance is plotted above the high-resolution proton spectrum. Correlations in the slice are highlighted by 3.

When more congested proton NMR spectra are encountered, tracing out proton-proton coupling networks can be considerably more difficult than this example and can lead to ambiguity when a correlation occurs between a proton and one of several proton resonances that happen to be overlapped in the proton spectrum. In some cases, simply transferring magnetization further from the starting point will provide the means of resolving ambiguities. In the case of strychnine, the H14 proton resonance partially overlaps the H18a and H11a resonances. Homonuclear TOCSY spectra provide the means of connecting H13 to resonances beyond H14, thereby sidestepping the overlap issue associated with H14/H18a/H11a. The homonuclear TOCSY pulse sequence is shown in Figure 9.29.

When interpreting spectra, novice investigators should note that spectra can be interpreted horizontally along a slice taken at a specific F_1 frequency as shown in Figure 9.28 or vertically at a specific F_2 frequency. The latter approach is employed for the correlations observed for the H22 vinyl resonance of strychnine as shown in Figure 9.30. As will be quickly noted, although COSY and TOCSY spectra are related in terms of information content, they will not necessarily afford the same correlations. When

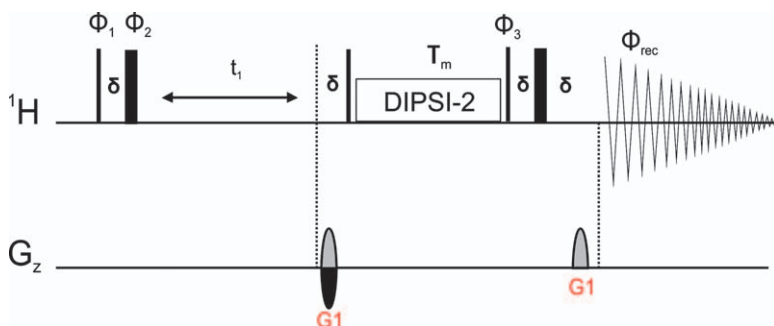


Figure 9.29 Pulse sequence for homonuclear TOCSY. Following the evolution period t_1 , a DIPSI-2 sequence is applied to propagate magnetization in order to further remote proton resonances. The extent to which magnetization can be propagated is a function of the duration of the mixing time. Phases for the experiment are as follows: $\Phi_1 = 0, 2$; $\Phi_2 = 0, 0, 0, 0, 2, 2, 2, 2$; $\Phi_3 = 0, 0, 2, 2$; $\Phi_{\text{rec}} = 0, 2, 2, 0, 2, 0, 0, 2$, and the gradient strength is $G1 = 30$.

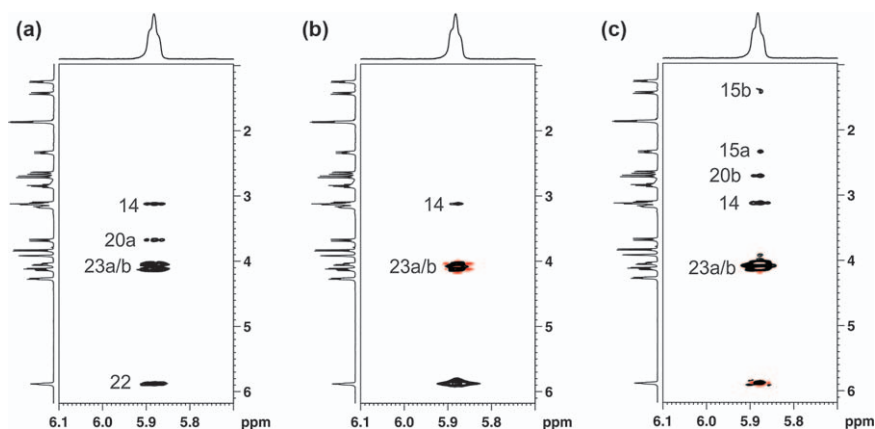


Figure 9.30 Correlations of the H22 vinyl proton of strychnine (**1**). (A) Segment of the COSY spectrum shown in Figure 9.27. (B) Segment of the 20 ms TOCSY spectrum. (C) Segment of the 40 ms TOCSY spectrum.

comparing the TOCSY spectra shown in Figure 9.30B and 9.30C, the 40 ms TOCSY spectrum shown in Figure 9.30C does contain the correlations of the 20 ms spectrum shown in Figure 9.30B, as well as several additional correlations.

9.9.2 HSQC-TOCSY Spectra

A further alternative for establishing the proton homonuclear coupling network of a molecule being investigated is provided by the HSQC-TOCSY experiment, whose pulse sequence is shown in Figure 9.31. The specific

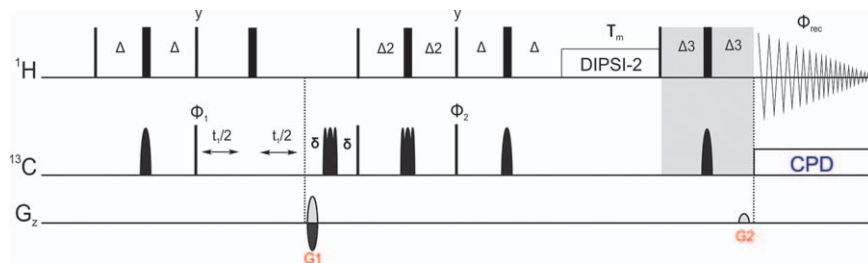
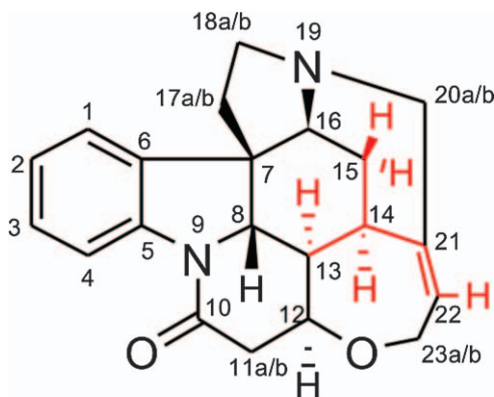


Figure 9.31 Pulse sequence for HSQC-TOCSY with inverted direct responses. Following the labeling of protons with the ^{13}C chemical shift of the directly attached proton, magnetization is refocused and then magnetization is propagated along the homonuclear coupling network in the DIPSI-2 block of the experiment. Finally, the inversion of the direct responses is provided by the shaded block of the pulse sequence. Key parameters and delays in the experiment are defined as follows: the delays in the experiment are defined as: $\Delta = 1/4(^1J_{\text{CH}})$; $\Delta 2 = 1/8(^1J_{\text{CH}})$ for all multiplicities or $1/4(^1J_{\text{CH}})$ for CH only; and $\Delta 3 = 1/2(^1J_{\text{CH}})$. Phases are defined as $\Phi_1 = 0, 2$; $\Phi_2 = 1, 1, 3, 3$; and $\Phi_{\text{rec}} = 0, 2, 2, 0$. Gradient ratios can be set for either ^1H - ^{13}C or ^1H - ^{15}N as follows: $G1 : G2 = 80 : 20.1$ (^1H - ^{13}C) and $80 : 8.1$ (^1H - ^{15}N). Mixing times for the DIPSI-2 pulse envelope can be varied to control the extent of magnetization propagation. Mixing times of 12–20 ms will typically afford a single TOCSY transfer, whereas mixing times in the range of 35–40 ms will transfer magnetization two or more bonds from the starting proton.

sequence shown is that for inverting direct responses, which facilitates the interpretation of the spectrum since the inverted responses are the analog of a non-edited HSQC spectrum.

The 20 ms HSQC-TOCSY spectrum of strychnine (**1**) with inverted direct responses is shown in Figure 9.32. The convention of plotting the negatively phased direct responses in red and the positively phased TOCSY correlations in black will be employed.



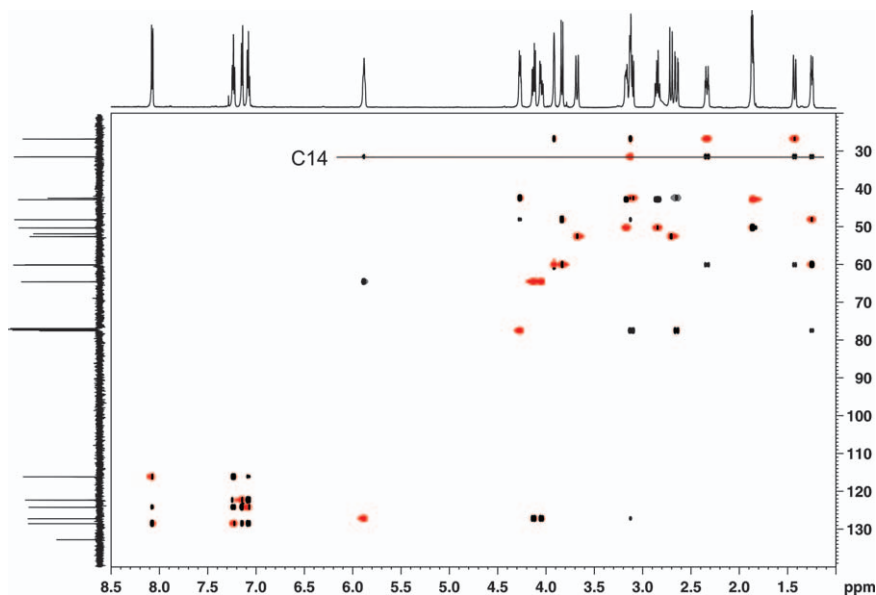
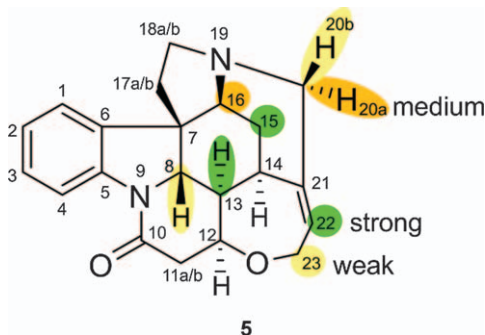


Figure 9.32 HSQC-TOCSY spectrum of strychnine (**1**) recorded with a 20 ms mixing time with the direct responses (plotted in red) inverted. To illustrate the interpretation of these data, we will consider the F_1 slice taken from the spectrum at the ^{13}C chemical shift of the C14 resonance. The C14 F_1 slice superimposed on a high-resolution proton spectrum is shown in Figure 9.33 and correlations are summarized in 4.

As with the homonuclear TOCSY experiment, magnetization can be propagated further from the starting proton by resorting to a longer mixing time. When a mixing time of 40 ms is employed, the weak proton resonances in the boxed region of the C14 slice shown in Figure 9.33 become more prominent, as shown in the segment of the C14 slice from a HSQC-TOCSY spectrum recorded with a 40 ms mixing time presented in Figure 9.34.



If still-longer mixing times were employed, magnetization could be propagated still further from the starting 14-position. It is likely in a 60 ms

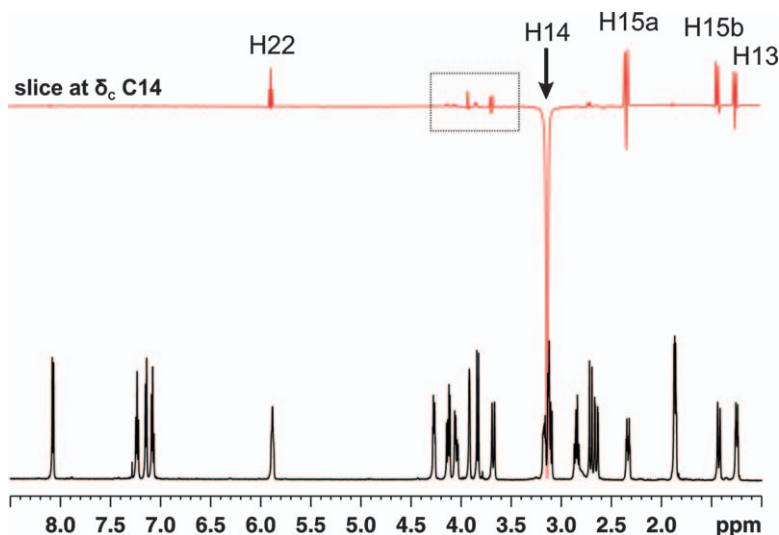


Figure 9.33 The slice extracted at the F_1 chemical shift of the C14 resonance of strychnine is superimposed over the high-resolution proton spectrum. The H14 correlations are inverted, whereas the correlations for the H22, H15a/b, and H13 have positive phases. There are a number of weak correlations shown in the boxed region of the slice. The coupling network defined by the slice is shown by 4.

experiment that correlations to the H12 and H11a/b resonances would begin to be observed; it is less likely that correlations to either the 17a/b or 18a/b resonances would be observed. Responses on 5 are color coded as a function of the strength of the correlation.

9.9.3 HSQMBC-COSY and HSQMBC-TOCSY Spectra

In the same sense that it is possible to append a TOCSY mixing step to the HSQC experiment, a recent report has demonstrated the possibility of doing the same with the HSQMBC experiment.⁶¹ While the HSQC-TOCSY experiment first establishes the direct ^1H - ^{13}C correlation *via* the $^1J_{\text{CH}}$ coupling pathway, in contrast, the HSQMBC-COSY and -TOCSY experiments first establish a long-range correlation *via* $^nJ_{\text{CH}}$ coupling, followed by either a COSY or TOCSY transfer from the remote site. In the case of an initial transfer *via* $^nJ_{\text{CH}}$ to a quaternary carbon, there can, of course be no further propagation. When the initial transfer is to a protonated carbon, then magnetization can be further propagated.

9.9.3.1 HSQMBC-COSY

The pulse sequence schematic for the HSQMBC-COSY experiment is shown in Figure 9.35 and the schematic showing the magnetization transfer pathway is shown in Figure 9.36. The HSQMBC-COSY spectrum of

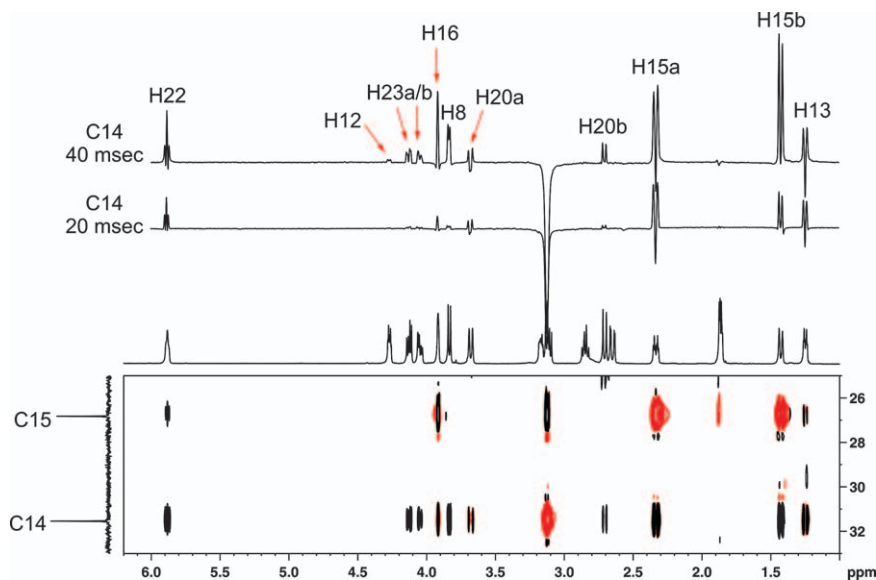


Figure 9.34 Segment of the 40 ms HSQC-TOCSY spectrum showing correlations for the C14 and C15 resonances. The slice from the 20 ms spectrum discussed previously and the 40 ms slice are plotted above the contour plot segment. The correlations that were weak in the corresponding slice from the 20 ms spectrum are now more intense. The correlations from this experiment (mixing time = 20 ms) are illustrated by 5. It should also be noted that, as with the comparison of COSY and 20 and 40 ms TOCSY spectra presented in Figure 9.30, HSQC-TOCSY can also be interpreted vertically using a slice at a specific F_2 frequency.

strychnine optimized for an 8 Hz long-range transfer is shown in Figure 9.37. Insight into the correlations expected in the HSQMBC-COSY experiment can be gained by comparing the correlations for the H22 resonance in an 8 Hz HMBC spectrum (see Figure 9.19), a 20 ms HSQC-TOCSY spectrum, and the 8 Hz HSQMBC-TOCSY spectrum shown in Figure 9.38. All of the long-range correlations observed in the HMBC spectra arise from the long-range heteronuclear coupling pathways. In the IDR-HSQC-TOCSY spectrum shown in Figure 9.38B, the H22/C22 correlation is inverted and plotted in red. In the HSQMBC-COSY spectrum, correlations can occur *via* an initial $^nJ_{CH}$ transfer followed by a homonuclear J_{HH} transfer. Thus, the correlations designated by black arrows in Figure 9.36 arise *via* $^nJ_{CH}$, while the correlation designated by a red arrow arises from the homonuclear J_{HH} transfer from H2 to H3.

9.9.3.2 HSQMBC-TOCSY

The HSQMBC-TOCSY experiment is a homolog of the HSQMBC-COSY experiment in that the COSY block appended to the HSQMBC long-range

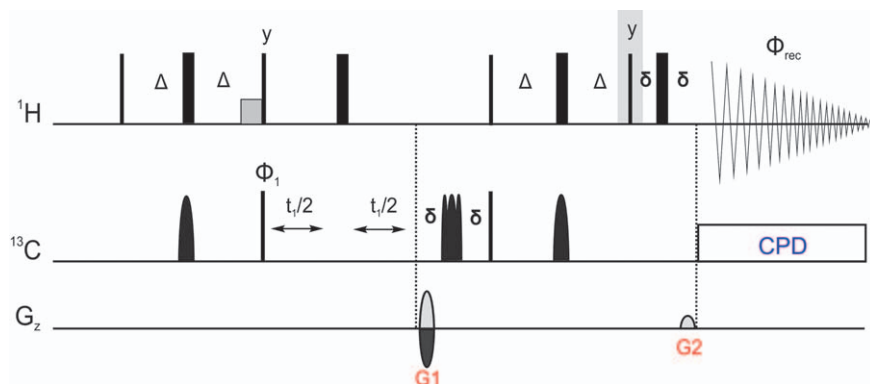


Figure 9.35 Pulse sequence schematic for the HSQMBC-COSY experiment. The delay, Δ , in the two INEPT transfers is optimized as a function of $1/4(J_{\text{CH}})$. After the protons are labeled with the ^{13}C chemical shift during the evolution period, magnetization is refocused and then the COSY transfer step (90° pulse in the grey-shaded box) occurs. Phases are cycled as $\Phi_1 = 0, 2$ and $\Phi_{\text{rec}} = 0, 2$. The gradients are optimized for $G1 : G2 = 80 : 20.1$ and $80 : 8.1$ for ^1H - ^{13}C and ^1H - ^{15}N , respectively.

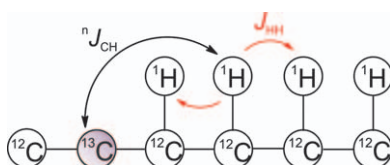


Figure 9.36 Schematic showing the magnetization transfer pathway in the HSQC-COSY experiment. Following the initial $^nJ_{\text{CH}}$ transfer and subsequent refocusing, the magnetization is transferred to adjacent protons *via* J_{HH} .

portion of the experiment is simply replaced by a TOCSY period as shown in Figure 9.39.⁶¹ Following the initial long-range heteronuclear transfer, magnetization is further propagated by a TOCSY mixing period, the duration of which can be varied as in the HSQC-TOCSY experiment described in Section 9.9.2. The schematic showing the magnetization transfer in the HSQMBC-TOCSY experiment is shown in Figure 9.40, and a side-by-side comparison of the HSQMBC-COSY and HSQMBC-TOCSY spectra of strychnine (**1**) is presented in Figure 9.41.

Comparing the experimental results shown in the two spectra in Figure 9.41, it can easily be seen that there are considerably more correlations associated with the HSQMBC-TOCSY spectra in Figure 9.41B than with the HSQMBC-COSY spectrum in Figure 9.41A. For example, the H22 vinyl resonance in Figure 9.41A shows correlations to the C21, C12, C23, and C20 resonances and what appears to be a weak correlation to either the C8 or

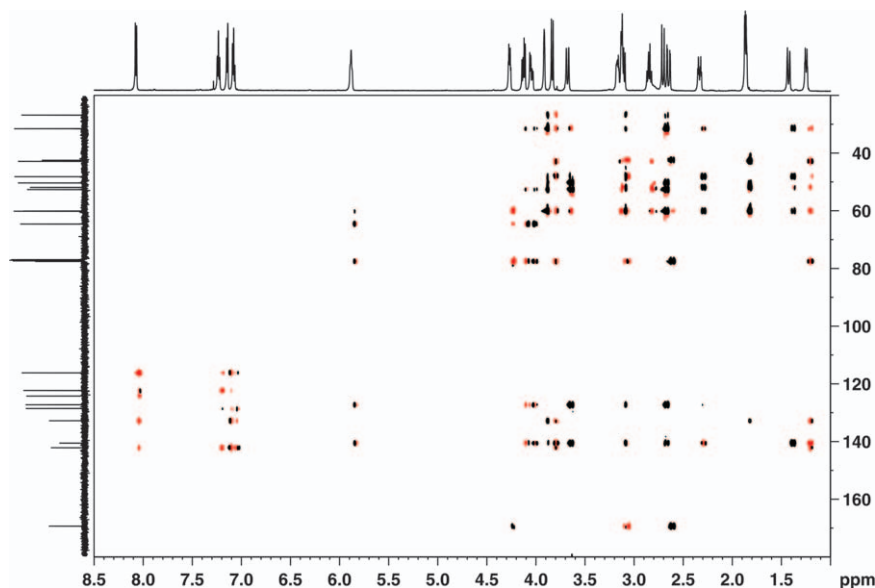


Figure 9.37 8 Hz-optimized HSQMBC-COSY spectrum of strychnine (**1**). The data were acquired as $4\text{K} \times 128$ points in four scans/ t_1 increment, giving an acquisition time of just over 11 minutes. The data were linear predicted and zero-filled to afford the $8\text{K} \times 1\text{K}$ spectrum shown.

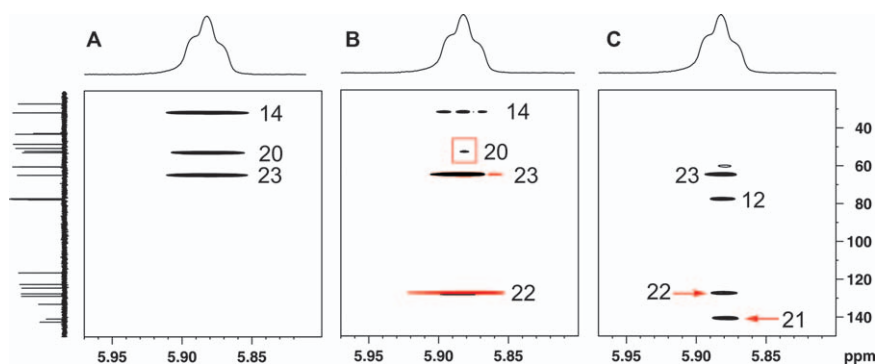


Figure 9.38 Comparison of the correlations of the H22 vinyl resonance of strychnine (**1**). (A) Segment of the 8 Hz-optimized HMBC spectrum. (B) Segment of the 20 ms IDR-HSQC-TOCSY spectrum. (C) Segment of the 8 Hz-optimized HSQMBC-COSY spectrum.

C16 resonance. In contrast, the 20 ms HSQMBC-TOCSY spectrum shown in Figure 9.41B exhibits the same correlations as in the HSQMBC-COSY spectrum, as well as additional correlations from the H22 vinyl proton to the C14 and C13 resonances. The correlation that was weakly observed to either C8 or C16 is also now observed with considerable intensity.

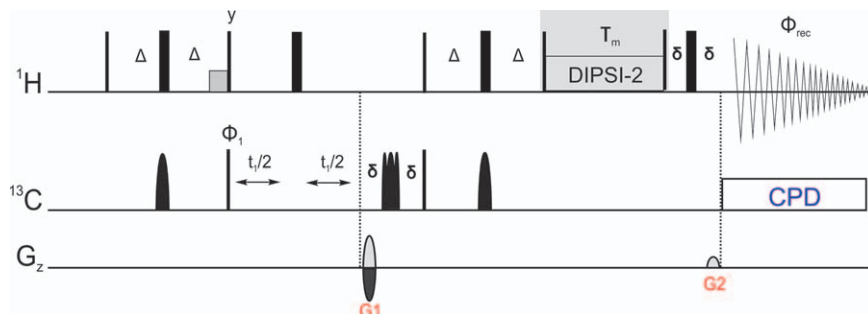


Figure 9.39 Pulse sequence for the HSQMB-COSY experiment.⁶¹ The COSY building block in the HSQMB-COSY experiment (Figure 9.36) is replaced by the DIPSI-2 TOCSY interval in the shaded block. The delay, Δ , in the two INEPT transfers is optimized as a function of $1/4(J_{\text{CH}})$. Phases are cycled as $\Phi_1 = 0, 2$ and $\Phi_{\text{rec}} = 0, 2$. The gradients are optimized for $G1 : G2 = 80 : 20.1$ and $80 : 8.1$ for ^1H - ^{13}C and ^1H - ^{15}N , respectively.

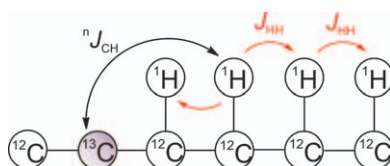


Figure 9.40 Schematic showing the magnetization transfer pathways in the HSQMB-COSY experiment. The long-range heteronuclear transfer via $^nJ_{\text{CH}}$ is denoted by the black arrow. The subsequent homonuclear transfers via J_{HH} are denoted by the red arrows.

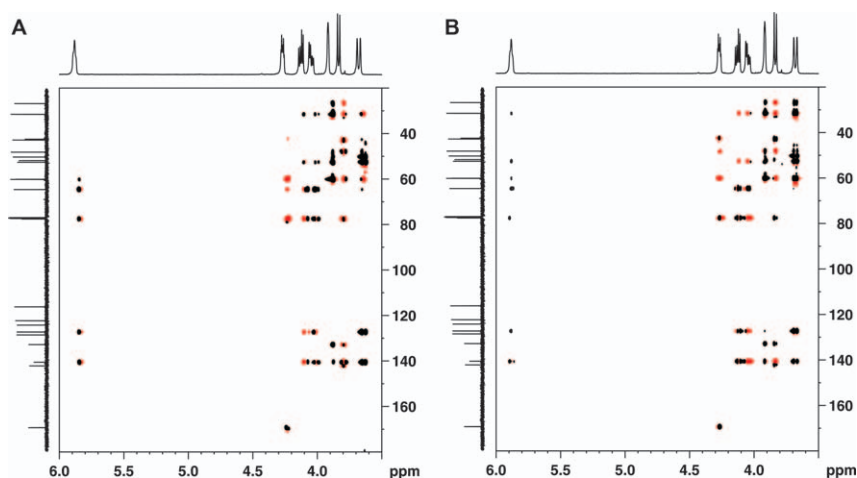


Figure 9.41 Comparison plots of the vinyl and aliphatic portions of the HSQMB-COSY (A) and 20 ms HSQMB-TOCSY (B) spectra of strychnine (1).

9.9.4 Strategy for Employing Hyphenated 2D Experiments

It is difficult to predict the structure elucidation strategies that might involve the newly developed, hyphenated HSQMB-COSY and -TOCSY experiments since there has been a long-standing dearth of applications of the HSQC-TOCSY experiment. In part, the hesitancy to employ HSQC-TOCSY can probably be attributed to the perception that it is much lower in sensitivity than HSQC and COSY. While that is in part true, the experiment is still, in the experience of the current authors, no less sensitive than HMBC. It will be interesting to see how the utilization of the new HSQMB-COSY and -TOCSY experiments is approached by the natural products chemistry community in the coming years.

References

1. G. E. Martin and A. J. Williams, Applications of ^1H - ^{15}N Long-range Heteronuclear Shift Correlation and ^{15}N NMR in Alkaloid Chemistry, in *Annual Reports in NMR Spectroscopy*, ed. G. A. Webb, Elsevier, New York, 2015, vol. 84, pp. 1–76.
2. Wikipedia. <http://www.wikipedia.org>.
3. G. E. Martin, M. Solntseva and A. J. Williams, Applications of ^{15}N NMR in Alkaloid Chemistry, in *Modern Alkaloids*, ed. E. Fattorusso and O. Talialatela-Scafati, Wiley-VCH, New York, 2007, pp. 411–476.
4. G. E. Martin and R. C. Crouch, Inverse-detected 2D NMR Applications in Alkaloid Chemistry, in *Modern Methods in Plant Analysis*, ed. H. F. Linskens and J. F. Jackson, Springer-Verlag, New York, 1994, vol. 15, pp. 25–89.
5. P. Forgo, J. Hohmann, G. Dombi and L. Máthé, Advanced Multi-dimensional NMR Experiments as Tool for Structure Determination of *Amaryllidacea* Alkaloids, in *Poisonous Plants and Related Toxins*, ed. T. Acamovic, S. Steward and T. W. Pennycott, CABI, Wallingford, UK, 2004, pp. 322–328.
6. R. W. Adams, Pure Shift NMR Spectroscopy, *eMagRes*, 2014, 3, 1–15.
7. K. Zangger, Pure shift NMR, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2015, 86–87, 1–20.
8. L. Castañar and T. Parella, *Magn. Reson. Chem.*, 2015, 53, 399–426.
9. D. J. Russell, C. E. Hadden, G. E. Martin, A. A. Givson, A. P. Zens and J. L. Carolan, *J. Nat. Prod.*, 2000, 63, 1047–1049.
10. G. E. Martin, Applications of NMR Microprobes and Methodologies for Spectral Assignments, in *Encyclopedia of Nuclear Magnetic Resonance*, ed. D. M. Grant and R. K. Harris, Wiley, New York, 2002, vol. 9, supplement, pp. 98–112.
11. G. E. Martin, Cryogenic NMR Probes, *Encyclopedia of Nuclear Magnetic Resonance*, ed. D. M. Grant and R. K. Harris, Wiley, New York, 2002, vol. 9, supplement, pp. 33–35.

12. G. E. Martin, Small Volume High Sensitivity NMR Probes, *Annual Reports on NMR Spectroscopy*, ed. G. A. Webb, Elsevier, London, 2005, vol. 56, pp. 1–99.
13. G. E. Martin, Application of Cryogenic NMR Probe Technology for the Identification of Low Level Impurities in Pharmaceuticals, in *Handbook of Modern Magnetic Resonance*, ed. G. A. Webb, Elsevier, London, 2006, vol. II, pp. 1187–1194.
14. B. D. Hilton and G. E. Martin, *J. Nat. Prod.*, 2010, **73**, 1465–1470.
15. G. E. Martin, Small Sample Cryoprobe NMR Applications, *Encyclopedia of NMR*, ed. R. K. Harris and R. A. Wasylishen, Wiley, New York, on-line, 2011, DOI: 1002/9780470034590.emrstm1300.
16. G. E. Martin, Using 1,1- and 1,n-ADEQUATE 2D NMR data in Structure Elucidation Protocols, *Annual Reports on NMR Spectroscopy*, ed. G. A. Webb, Elsevier, London, 2011, vol. 74, pp. 215–291.
17. G. E. Martin, M. Reibarkh, A. V. Buevich, K. A. Blinov and R. T. Williamson, Development of 1,n-ADEQUATE and Modified Variants and Their Application to Structure Elucidation and Spectral Assignments Problems, *eMagRes*, 2014, **3**, 215–234, DOI: 10.1002/9780470034590.emrstm1370.
18. S. Cheatham, P. Gierth, W. Bermel and Ě. Kupče, *J. Magn. Reson.*, 2014, **247**, 38–41.
19. S. Cheatham, M. Kline and Ě. Kupče, *Magn. Reson. Chem.*, 2015, **53**, 363–368.
20. D. A. Snyder and R. Brüschweiler, Multidimensional Correlation Spectroscopy by Covariance NMR, in *Multidimensional NMR Methods for the Solution State*, ed. G. A. Morris and J. W. Emsley, Wiley, New York, 2010, pp. 97–105.
21. G. E. Martin, P. A. Irish, B. D. Hilton, K. A. Blinov and A. J. Williams, *Magn. Reson. Chem.*, 2007, **45**, 624–627.
22. G. E. Martin, P. A. Irish, B. D. Hilton, K. A. Blinov and A. J. Williams, *Magn. Reson. Chem.*, 2007, **45**, 883–888.
23. G. E. Martin, B. D. Hilton, P. A. Irish, K. A. Blinov and A. J. Williams, *J. Heterocycl. Chem.*, 2007, **44**, 1219–1222.
24. G. E. Martin, B. D. Hilton, K. A. Blinov and A. J. Williams, *J. Nat. Prod.*, 2007, **70**, 1966–1970.
25. C. E. Hadden, B. D. Kaluzny, R. H. Robins and G. E. Martin, *Magn. Reson. Chem.*, 1999, **37**, 325–327.
26. S. L. Patt and J. Shoolery, *J. Magn. Reson.*, 1982, **46**, 535–539.
27. D. M. Doddrell, D. T. Pegg and M. R. Bendall, *J. Magn. Reson.*, 1982, **48**, 323–327.
28. V. Rutar and T. C. Wong, *J. Magn. Reson.*, 1983, **53**, 495–499.
29. W. P. Aue, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.*, 1976, **64**, 2229–2246.
30. T. Parella, *Magn. Reson. Chem.*, 1998, **36**, 467–495.
31. T. Parella, J. Belloc, F. Sánchez-Fernando and A. Virgili, *Magn. Reson. Chem.*, 1998, **36**, 715–719.

32. J. R. Garbow, D. P. Weitekamp and A. Pines, *Chem. Phys. Lett.*, 1982, **93**, 504–508.
33. L. Paudel, R. W. Adams, P. Király, J. A. Aguilar, M. Fooroozandeh, M. J. Cliff, M. Nilsson, P. Sándor, J. P. Waltho and G. A. Morris, *Angew Chem., Int. Ed.*, 2013, **52**, 11616–11619.
34. Y. Liu, M. D. Green, R. Marques, T. Pereira, R. Helmy, R. T. Williamson, W. Bermel and G. E. Martin, *Tetrahedron Lett.*, 2014, **55**, 5450–5453.
35. T. Molinski, *Curr. Opin. Drug Discovery Dev.*, 2009, **12**, 197–206.
36. D. S. Salisay, B. I. Morinaka, C. K. Skepper and T. F. Molinski, *J. Am. Chem. Soc.*, 2009, **131**, 7552–7553.
37. D. S. Dalisay, E. W. Rogers, A. S. Edison and T. F. Molinski, *J. Nat. Prod.*, 2009, **72**, 732–738.
38. D. S. Dalisay and T. F. Molinski, *Org. Lett.*, 2009, **11**, 1967–1970.
39. T. F. Molinski, *Curr. Opin. Biotechnol.*, 2010, **21**, 819–826.
40. M. Pérez-Trujillo, L. Castañar, E. Monteagudo, L. T. Kuhn, P. Nolis, A. Virgili, R. T. Williamson and T. Parella, *Chem. Commun.*, 2014, **50**, 10214–10217.
41. N. Marco, A. Fredi and T. Parella, *Chem. Commun.*, 2015, **51**, 3262–3265.
42. M. Palmer, R. Gupta, M. Richards, C. L. Suiter, T. Polenova, J. C. Hoch and D. Rovnyak, Application of Non-Uniform Sampling for Sensitivity Enhancement of Small Molecule Heteronuclear Correlation Spectra, in *Modern NMR Approaches to the Structure Elucidation of Natural Products*, ed. A. J. Williams, G. E. Martin and D. J. Rovnyak, RSC, London, 2015, vol. 1, in press.
43. W. F. Reynolds, *Encycl. NMR Spectrosc.*, 2011, DOI: 10.1002/97804700334590.emrstm1176.
44. W. Schoefberger, J. Schlagnitweit and N. Müller, Recent Developments in Heteronuclear Multiple Bond Correlation Experiments, in *Annual Reports on NMR Spectroscopy*, ed. G. A. Webb, Elsevier, Amsterdam, 2011, vol. 72, pp. 1–60.
45. J. Furrer, Recent Developments in HMBC Studies, in *Annual Reports on NMR Spectroscopy*, ed. G. A. Webb, Elsevier, Amsterdam, 2011, vol. 74, pp. 294–354.
46. J. Furrer, *Concepts Magn. Reson. A*, 2012, **40**, 101–127.
47. J. Furrer, *Concepts Magn. Reson. A*, 2012, **40**, 146–169.
48. J. Furrer, *Concepts Magn. Reson. A*, 2015, **43**, 177–206.
49. A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, 1986, **108**, 2093–2094.
50. K. Furihata and H. Seto, *Tetrahedron Lett.*, 1995, **36**, 2817–2820.
51. R. Wagner and S. Berger, *Magn. Reson. Chem.*, 1998, **36**, S44–S46.
52. G. E. Martin, C. E. Hadden and V. V. Krishnamurthy, *J. Magn. Reson.*, 1999, **140**, 274–280.
53. C. E. Hadden, G. E. Martin and V. V. Krishnamurthy, *Magn. Reson. Chem.*, 2000, **38**, 143–147.
54. V. V. Krishnamurthy, D. J. Russell, C. E. Hadden and G. E. Martin, *J. Magn. Reson.*, 2000, **146**, 232–239.

55. R. T. Williamson, A. V. Buevich, G. E. Martin and T. Parella, *J. Org. Chem.*, 2014, **79**, 3387–3394.
56. R. Araya-Maturana, T. Delgado-Castro, W. Carona and B. E. Weiss-López, *Curr. Org. Chem.*, 2001, **5**, 253–263.
57. R. Araya-Maturana, H. Pessoa-Mahana and B. E. Weiss-López, *Nat. Prod. Commun.*, 2008, **3**, 445–450.
58. R. T. Williamson, B. L. Marquez, W. H. Gerwisk and K. E. Kövér, *Magn. Reson. Chem.*, 2000, **38**, 265–273.
59. L. Castanar and T. Parella, Recent Advances in Small Molecule NMR: Improved HSQC and HMQBC Experiments, in *Annual Reports on NMR Spectroscopy*, ed. G. A. Webb, Elsevier, Amsterdam, 2015, vol. 84, pp. 163–232.
60. J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff Jr., T. Parella, R. T. Williamson and G. E. Martin, *Angew. Chem., Int. Ed.*, 2015, **54**, 10160–10164.
61. J. Saurí, N. Marcó, R. T. Williamson, G. E. Martin and T. Parella, *J. Magn. Reson.*, 2015, **252**, 25–32.

CHAPTER 10

Nuclear Magnetic Resonance Experiments Applicable to the Elucidation and Characterization of Alkaloid Structures Part II: Advanced Techniques for the Identification of Adjacent Carbons Using H2BC, 1,1-ADEQUATE, and Variants

JOSEP SAURÍ AND GARY E. MARTIN*

Merck & Co., Inc., Process & Analytical Chemistry, NMR Structure Elucidation, Rahway, NJ 07065, USA

*Email: gary.martin2@merck.com

10.1 Identification of Adjacent Carbons Using H2BC and 1,1-ADEQUATE Data

In some cases, critical coupling constants may either be too small to be successfully exploited or it may be necessary to unequivocally assign a directly adjacent carbon resonance in order to assign a structure. The H2BC

Modern NMR Approaches to the Structure Elucidation of Natural Products:

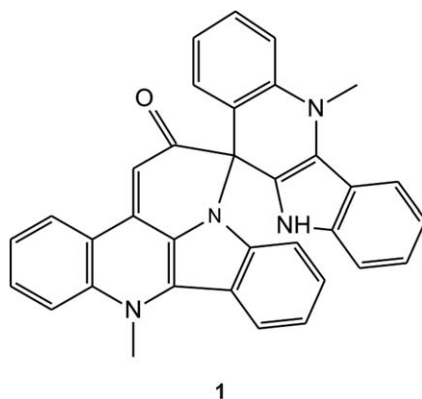
Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

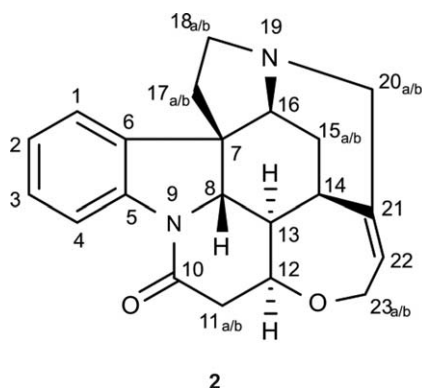
Published by the Royal Society of Chemistry, www.rsc.org

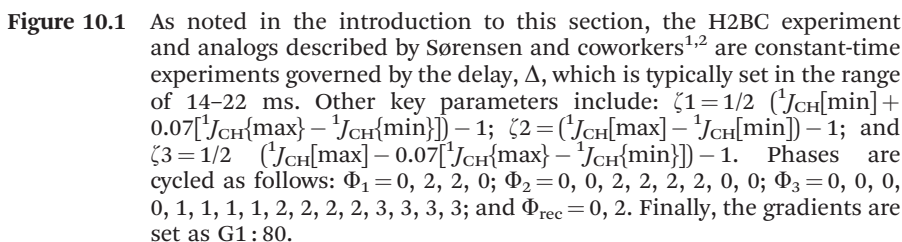
experiment was, in part, developed by Sørensen and coworkers for such instances and, aside from some caveats, works for adjacent protonated carbon resonances.^{1,2} H2BC is a constant-time experiment, a feature that limits the F_1 digital resolution that can be attained, which can limit the application of the experiment in some instances. Indeed, during the recent investigation and revision of the structure of the complex *Cryptolepis* aromatic alkaloid cryptospirolepine (1), an attempt to utilize H2BC was unsuccessful due to the F_1 resolution limitations imposed by the constant-time nature of the pulse sequence.³ Nevertheless, H2BC is certainly a higher-sensitivity choice than ADEQUATE, although the latter works for adjacent quaternary carbons, whereas H2BC cannot provide quaternary carbon assignments.



10.1.1 Identification of Adjacent Protonated Carbons Using the H2BC Experiment

The pulse sequence for the H2BC experiment^{1,2} is shown in Figure 10.1. H2BC is a constant-time experiment, which in some cases limits access to high F_1 digital resolution. The application of the H2BC experiment is also restricted to the establishment of correlations between adjacent protonated carbons.





Surveying the literature has shown that the H2BC experiment has not been widely utilized in natural product structure elucidation studies. In a SciFinder™ search, 37 references to the H2BC experiment were found. Of the references found, 15 referred to carbohydrate applications, three referred to alkaloid applications, and four citations pertained to references to the H2BC experiment in reviews or reports in which the application was not noted.

Two of the citations found addressed the impact of employing the H2BC experiment to unequivocally identify adjacent protonated carbons and the impact of these data on the performance of computer-assisted structure elucidation (CASE) programs.^{4,5} Similar benefits have been ascribed to the inclusion of 1,1-ADEQUATE data in the CASE program data input file (the interested reader is referred to Volume 1, Chapter 9 for further information).^{4,6}

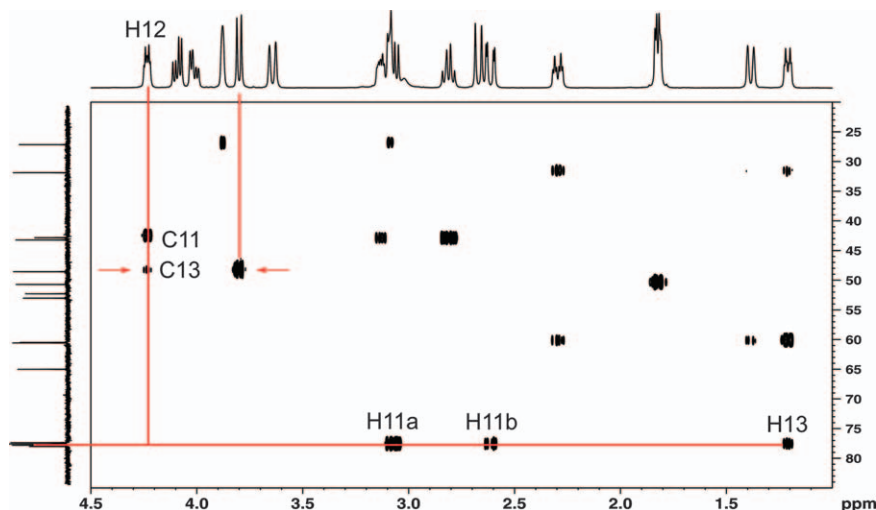


Figure 10.2 The aliphatic portion of the H2BC spectrum of strychnine (**2**) is shown. The spectrum can be interpreted “vertically” along the column defined, for example by the H12 chemical shift, which highlights correlations to the adjacent C11 (strong) and C13 (weak) resonances. Alternatively, the spectrum can be interpreted “horizontally” along the row defined by the C12 chemical shift, which pinpoints the H11a/b anisochronous geminal methylene resonances and the H13 methine resonance. Compare the H12/C12 correlations to that of H8/C8, which is flanked by the C7 quaternary carbon and the H13/C13 methine. A correlation is observed at the F_1 shift of the C13 resonance, but none is observed for C7, since the H2BC experiment cannot identify adjacent quaternary carbon resonances.

10.1.2 Identifying Adjacent Carbons Using 1,1- and 1,1-HD-ADEQUATE

Development of the 1,1-ADEQUATE experiment^{7,8} predated the commercial introduction of cryogenic NMR probes by several years, and the relatively low sensitivity of the experiment compared to HMBC, LR-HSQMBC, and H2BC, coupled with the timing of the introduction of the experiment, has probably contributed to the scant adoption of the experiment for natural product structure elucidation. All of the ADEQUATE-type experiments are out-and-back experiments that begin with the transfer of magnetization from a proton to either the directly bound carbon *via* $^1J_{CH}$ or to a remote carbon *via* $^nJ_{CH}$. The experiments have been the subject of two recent, comprehensive reviews.^{9,10} In practical terms, only two of the experiments, 1,1- and 1,*n*-ADEQUATE, are likely to be useful in the context of natural product structure elucidation. Both of these experiments rely on an initial $^1J_{CH}$ transfer, followed either by a $^1J_{CC}$ or $^nJ_{CC}$ carbon–carbon transfer, with $n=3$ predominating in the latter experiment. Schematically, the magnetization transfer pathways for the 1,1-ADEQUATE experiment is illustrated in Figure 10.3.

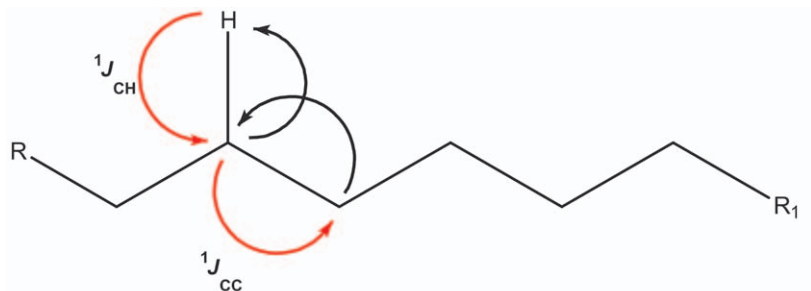


Figure 10.3 Out-and-back magnetization transfer pathway of the 1,1-ADEQUATE experiment. Following the initial $^1J_{\text{CH}}$ transfer to the directly bound carbon, magnetization is transferred to the adjacent ^{13}C via the $^1J_{\text{CC}}$ coupling. The delay for the $^1J_{\text{CH}}$ coupling is typically optimized for ~ 145 Hz. The delay for the homonuclear ^{13}C – ^{13}C transfer is typically optimized in the range of $^1J_{\text{CC}} = 40$ – 60 Hz.

The 1,1-HD-ADEQUATE experiment was only described in the literature in 2015 with two applications, including the initial report.^{3,11} The information content of the 1,1-HD-ADEQUATE is identical to the original 1,1-ADEQUATE experiment described in 1996.⁷ However, since the 1,1-HD-ADEQUATE experiment employs BIRD-based homonuclear decoupling during acquisition, some of the proton multiplets will be collapsed, while others may be simplified, remaining as geminal doublets. From the homodecoupling (HD) applied in the experiment, the sensitivity of the 1,1-HD-ADEQUATE is correspondingly better than the original experiment. When 1,1-HD-ADEQUATE is employed in conjunction with a small-diameter (*e.g.* 1.7 mm) cryoprobe, it becomes feasible to perform the experiment on submilligram samples. Following the description of the 1,1-ADEQUATE spectrum and a discussion of the interpretation of the spectrum, we will consider the 1,1-HD-ADEQUATE experiment.

10.1.2.1 Acquiring and Interpreting 1,1-ADEQUATE Data

The pulse sequence for the experiment is shown in Figure 10.4. Interpretation of 1,1-ADEQUATE spectra is straightforward. Correlations in F_1 correspond to the carbons, either protonated or quaternary, that are directly adjacent to the starting protonated carbon. Again, it is instructive to employ strychnine as a model compound. The full 60 Hz-optimized 1,1-ADEQUATE spectrum is shown in Figure 10.5. To facilitate the interpretation of a 1,1-ADEQUATE spectrum, we will overlay a multiplicity-edited pure shift heteronuclear single-quantum correlation (PS-HSQC) over the aliphatic region of the 1,1-ADEQUATE spectrum as shown in Figure 10.6.

A useful comparison of the H2BC experiment discussed in Section 10.1.1 and the 1,1-ADEQUATE experiment is seen if we contrast the correlations of

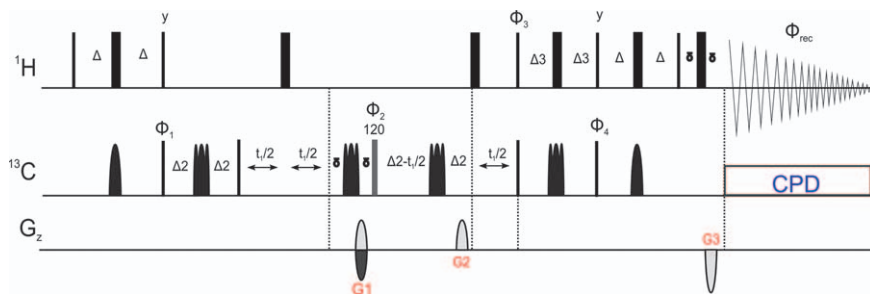


Figure 10.4 Pulse sequence for the 1,1- and 1,*n*-ADEQUATE experiments. Key parameters and delays are as follows: $\Delta = 1/4(^1J_{\text{CH}})$; $\Delta 2 = 1/4(^1J_{\text{CC}})$ for 1,1-ADEQUATE and $1/4(^nJ_{\text{CC}})$ for 1,*n*-ADEQUATE, with the delays optimized for 40–60 and 3–7 Hz, respectively;^{9,10} $\Delta 3 = 1/8(^1J_{\text{CH}})$ for all multiplicities or $1/4(^1J_{\text{CH}})$ for CH only. Phases are cycled as: $\Phi_1 = 0, 2$; $\Phi_1 = 0(\times 8), 2(\times 8)$; $\Phi_3 = 0, 0, 2, 2$; $\Phi_4 = 1, 1, 3, 3$; $\Phi_{\text{rec}} = 0, 2, 2, 0, 2, 0, 0, 2, 2, 0, 2, 2, 0$ —care should be taken to set the number of scans to an integer multiple of 16. Gradients should be set as: $G1 : G2 : G3 = 78.5 : 78.6 : -59$ ($^1\text{H}-^{13}\text{C}$). It is also preferable to cycle any of the ADEQUATE experiments so that the sum of the acquisition time plus the interpulse delay is about 3 seconds.

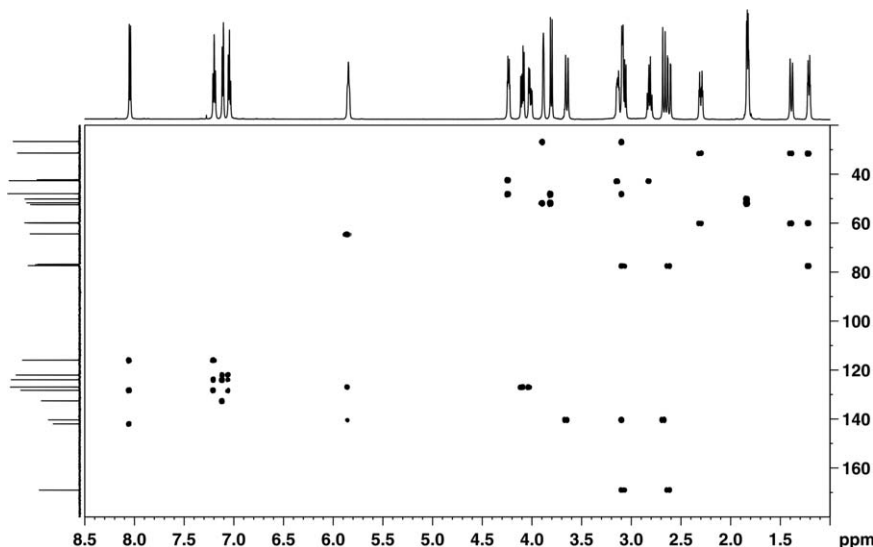


Figure 10.5 The 60 Hz-optimized 1,1-ADEQUATE spectrum of strychnine (2). Correlations in the spectrum can be interpreted as shown in Figure 10.6.

the H12/C12 and H8/C8 resonances. Examining the structure of strychnine (2), the C12 methine is flanked by the 11-methylene and the 13-methine resonances. If we examine the H2BC spectrum of the aliphatic region of strychnine shown in Figure 10.2, consistent with the constitution of

strychnine, correlations are observed in the H2BC spectrum with both the C13 and C11 resonances, albeit weakly with the former. If we compare the aliphatic region of the 1,1-ADEQUATE spectrum shown in Figure 10.6, we observe the same correlations with the C13 and C11 resonances, although both correlations are about equal in intensity. Whereas the $^2J_{\text{CH}}$ correlations can vary considerably, in contrast, $^1J_{\text{CC}}$ correlations tend to be more uniform and can more reliably be utilized to establish the identity of vicinal carbons.¹²

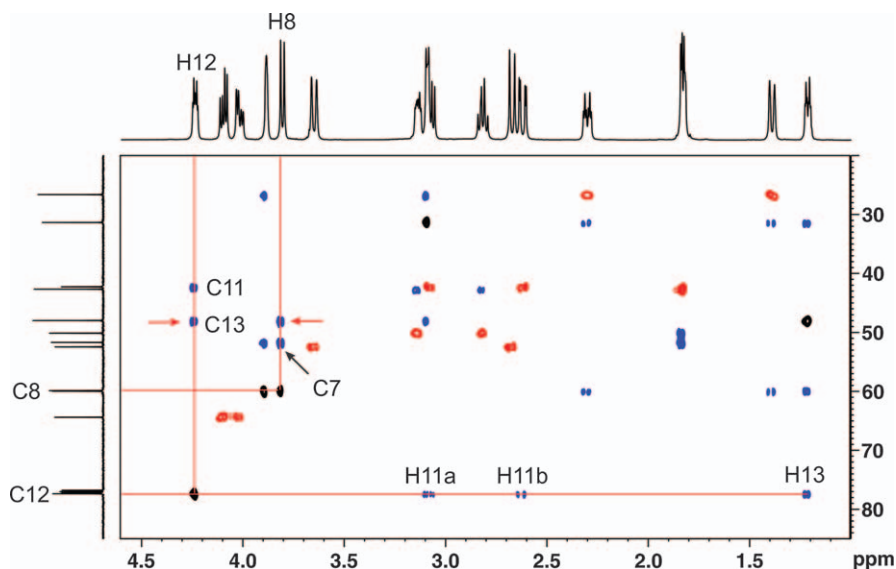


Figure 10.6 Multiplicity-edited PS-HSQC spectrum of strychnine (2) overlaid on the 60 Hz-optimized 1,1-ADEQUATE spectrum. Correlations in the multiplicity-edited PS-HSQC spectrum are plotted using the convention established at the beginning of this chapter: CH/CH₃ resonances have positive intensity and are plotted in black; CH₂ resonances have negative intensity and are plotted in red. 1,1-ADEQUATE correlations are plotted in blue. Comparing this presentation to that of the H2BC spectrum shown in Figure 10.2, the correlations that can be established from the H12/C12 are identical. Using the column defined by the H12 chemical shift, correlations are observed in F₁ corresponding to the C11 and C13 resonances that flank C12. Interpreting along the row defined by the F₁ chemical shift of C12, correlations are observed at the F₂ frequencies of H11a/b and H13. Again referring to Figure 10.2, it was noted that the C7 quaternary carbon and the H13/C13 methine flank H8/C8. In the H2BC spectrum, no correlation is observed at the F₁ shift of C7 since the experiment is incapable of identifying adjacent quaternary carbon resonances. In contrast, in the 1,1-ADEQUATE spectrum, correlations are observed in the column defined by the F₂ chemical shift of H8 at both the C13 and C7 F₁ frequencies. Despite the inherently lower sensitivity of the 1,1-ADEQUATE experiment relative to H2BC, the former has the advantage of allowing the unequivocal identification of adjacent quaternary carbons.

Shifting our consideration to the correlations associated with the 8-methine, we obtain a different picture from comparing the two spectra. Since the H2BC experiment operates only for adjacent protonated carbons, as expected, only a correlation to the 13-methine is observed in Figure 10.2. In contrast, in the aliphatic segment of the 1,1-ADEQUATE spectrum shown in Figure 10.6, strong correlations are observed to both C13 and the C7 quaternary carbon, which is consistent with the nearly uniform $^1J_{CC}$ couplings between C8 and the C13 and C7 carbons. Indeed, the DFT-calculated coupling constants were: $^1J_{C8C7} = 34.1$; $^1J_{C8C13} = 32.8$; $^1J_{C12C11} = 36.5$; and $^1J_{C12C13} = 35.3$ Hz. The measured $^1J_{CC}$ couplings were: $^1J_{C8C7} = 34$; $^1J_{C8C13} = 32$; $^1J_{C12C11} = 37$; and $^1J_{C12C13} = 34$ Hz. In contrast, the DFT-calculated $^2J_{CH}$ coupling constants were: $^2J_{C12H11a} = -0.61$; $^2J_{C12H11b} = -6.74$; and $^2J_{C13H12} = 1.62$ Hz. Based on these couplings, the weaker correlation between H12 and C13 observed in the H2BC spectrum shown in Figure 10.2 is not surprising.

The difference between the H2BC and 1,1-ADEQUATE experiments is essentially a tradeoff between the variability of $^2J_{CH}$ couplings and higher sensitivity for the former *versus* better uniformity of the $^1J_{CC}$ coupling constants exploited in the latter and lower overall sensitivity. Investigators with sufficient samples and access to a spectrometer with an appropriate LN_2 - or helium-cooled cryoprobe would do well to consider acquiring 1,1-ADEQUATE data in those cases where a critical correlation is not observed in a H2BC spectrum or when the nature of the sample is such that the F_1 digital resolution requirements are too high for H2BC to be used because of the constant-time nature of the experiment.³

10.1.2.2 Introducing Partial Homodecoupling—The 1,1-HD-ADEQUATE Experiment

Intensive research focus on the development of pure shift NMR experiments has been the subject of several very recent reviews.^{13–15} Collapsing proton multiplets, when it is possible with the method being employed for HD, translates to correspondingly high signal/noise (s/n) ratios in the experiment being performed, as well as enhanced resolution in the proton spectrum. Indeed, one of the current authors has recently shown the beneficial impact that employing PS-HSQC can have on metabolite characterization.¹⁶ Based on the sensitivity gains obtained for heteronuclear shift correlation experiments, the authors were logically interested in obtaining corresponding improvements in the sensitivity of the much less sensitive ADEQUATE experiments, which prompted the development and recent communication detailing the application of the 1,1- and 1,*n*-HD-ADEQUATE pulse sequences shown in Figure 10.7 that were utilized in the revision of the structure of the alkaloid cryptospiropepine.³

Considerations implicit in the PS-HSQC experiment for anisochronous methylene protons also apply in the HD-ADEQUATE experiments, although

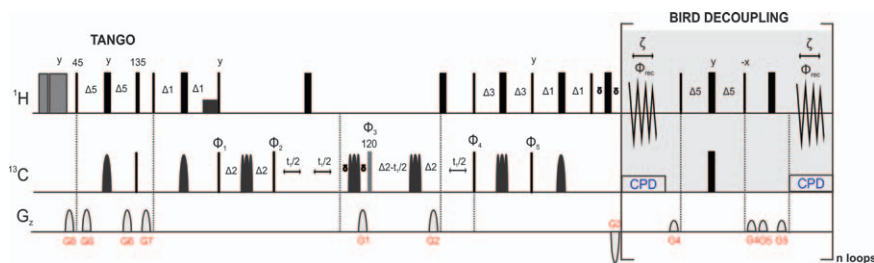
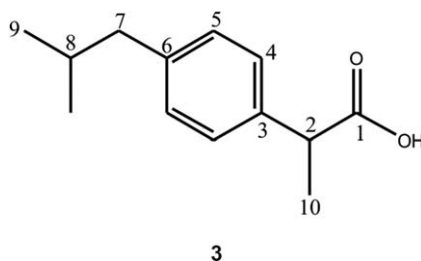


Figure 10.7 Pulse sequence for 1,1-^{7,8} and 1,*n*-HD-ADEQUATE.³ The two experiments utilize the same pulse sequence and differ only in the optimization of the J_{CC} delays; the 1,1-HD-ADEQUATE experiment employs a $^1J_{CC}$ delay typically optimized in the range of 40–60 Hz, while the 1,*n*-HD-ADEQUATE experiment utilizes an $^nJ_{CC}$ delay typically optimized in the range of 3–7 Hz. HD-ADEQUATE employs BIRD-based HD and a “chunked” data acquisition analogous to the PS-HSQC experiment discussed in Chapter 9, Section 9.5.3. Due to the nature of this data acquisition strategy and the increased number of RF pulses, an increase in the level of artifacts is expected. This “noise” will primarily be associated with ¹³C-bound protons and protons bound to an isolated ¹³C atom, while the signals of interest emanate from molecules with ¹³C-coupled atom pairs (¹H¹³C–¹³C or ¹H¹³C–(¹³C)_{*n*}–¹³C where *n* = 2–3).^{9,10} Hence, signals from a ¹H–¹²C or isolated ¹H–¹³C moieties should be suppressed as much as possible to minimize artifacts. In order to achieve this goal, additional “building blocks” have been inserted in the standard ADEQUATE pulse sequence. These include a purge element (2.5 ms[*x*] 5 ms[*y*] at 10 kHz RF field followed by a spoil gradient) prior to the relaxation delay in order to destroy all undesired components of magnetization. We also incorporated a TANGO¹⁷ gradient element to excite and dephase ¹²C-bound protons, while leaving ¹³C-bound protons unaffected. Finally, a 1 ms high-power trim pulse was applied at the end of the initial INEPT transfer step to help destroy any remaining undesired transverse magnetization.

in a slightly different manner. The 1,1-ADEQUATE experiment is designed to establish correlations between directly adjacent carbons, irrespective of whether or not the neighboring carbon is protonated or a quaternary carbon. We will first consider the case of adjacent protonated and non-protonated carbons. Protonated carbons adjacent to a quaternary carbon (e.g. ¹H–¹³C_A–¹³C_BR₃) will experience HD in the 1,1-HD-ADEQUATE experiment. Protons vicinal to ¹³C_A will be decoupled since all but 1% of those protons are of the type ¹H–¹²C. Hence, correlations between adjacent protonated and non-protonated carbons in the 1,1-HD-ADEQUATE experiment will benefit from HD, with a corresponding enhancement in s/n. In the PS-HSQC experiment, the homonuclear coupling between anisochronous geminal methylene protons is not collapsed since both protons reside on the same ¹³C resonance (see Chapter 9, Figure 9.12). In the 1,1-ADEQUATE experiment, protons on adjacent protonated carbons (e.g. those of the type ¹H–¹³C_A–¹H–¹³C_B) will likewise not be decoupled in BIRD-based HD schemes since, again, both protons reside on ¹³C, albeit different carbons in this case.

Before considering the 1,1-HD-ADEQUATE spectrum of strychnine (2) that we have been using as a model compound throughout this chapter, it is germane to examine correlations from ibuprofen (3), which was used for comparative evaluation during the development of the 1,1-HD-ADEQUATE experiment. The cases just described above are illustrated in the two panels in Figure 10.8.



The aliphatic portion of the 1,1-HD-ADEQUATE spectrum of strychnine (2) is shown in Figure 10.9. Once again, exactly as in Figure 10.6, the H8 resonance shows correlations to the C13 methine and C7 quaternary carbon resonances. As expected based on the panels in Figure 10.8, the traces

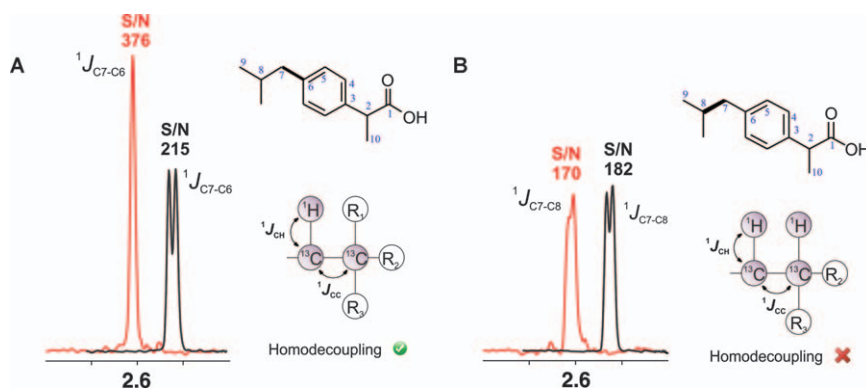


Figure 10.8 (A) Case for adjacent protonated and non-protonated carbons of the type $^1\text{H}-^{13}\text{C}_\text{A}-^{13}\text{C}_\text{B}\text{R}_3$, which are represented by the C7 methine and C6 quaternary aromatic carbons of ibuprofen (3). The H7 doublet in the 1,1-HD-ADEQUATE experiment would be homodecoupled since 99% of the H8 protons reside on ^{12}C . The image in (A) shows the comparison from two identically acquired spectra. The black trace shows the H7 doublet from the conventional 1,1-ADEQUATE experiment with s/n of 215:1 for the correlation. In comparison, the red trace shows the corresponding H7 singlet from the 1,1-HD-ADEQUATE spectrum with s/n of 376:1. The benefit derived from HD is obvious. (B) The correlations from the conventional and HD-ADEQUATE spectra for the correlation between the C7 and C8 methine carbon resonances. Here, the carbons are of the type $^1\text{H}-^{13}\text{C}_\text{A}-^1\text{H}-^{13}\text{C}_\text{B}$, which leads to the expectation that the vicinal coupling between H7 and H8 will not be collapsed, which is indeed the case, although the couplings of the H8 methine to the two methyl resonances are collapsed. The s/n from the two traces are comparable.

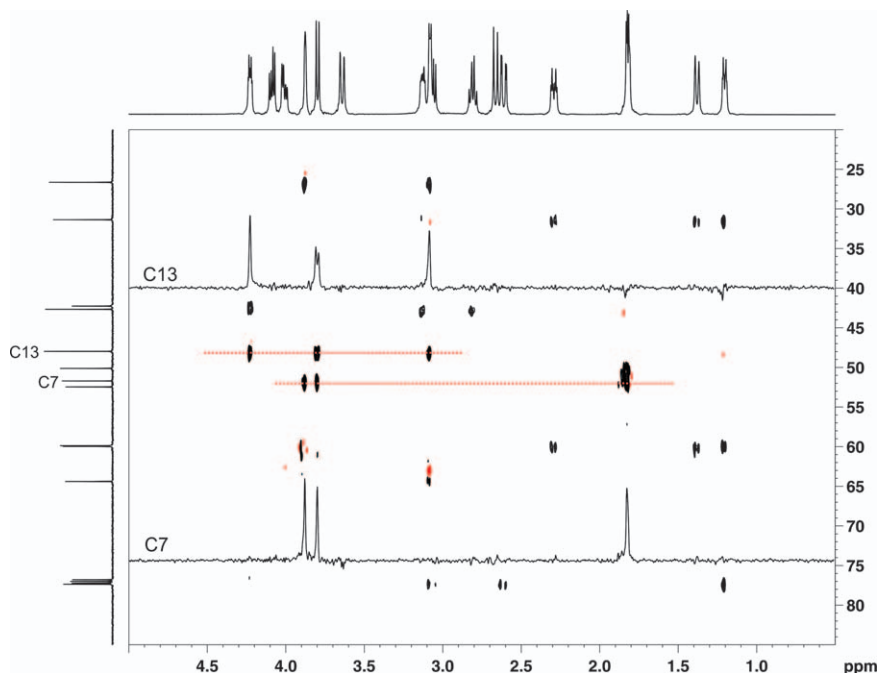
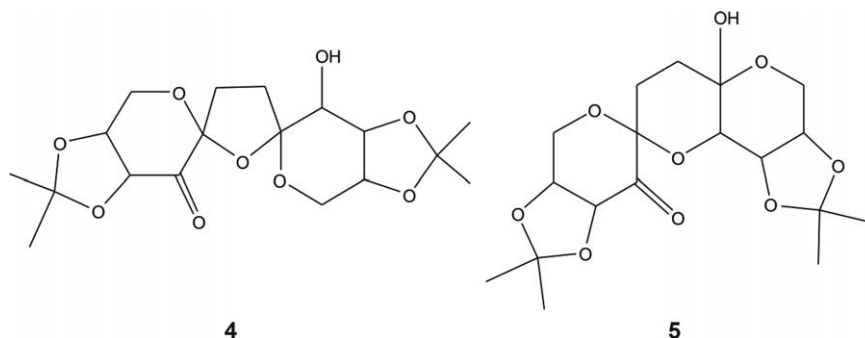


Figure 10.9 Segment of the aliphatic region of the 1,1-HD-ADEQUATE³ spectrum of strychnine (**2**). The experiment was optimized for $^1J_{CC} = 50$ Hz and the data were acquired as $3K \times 160$ points with 64 transients accumulated/ t_1 increment, giving an acquisition time of 9 hours and 52 minutes. The data were linear predicted to 320 points in F_1 and zero-filled to afford the $2K \times 2K$ spectrum shown. Slices extracted at the F_1 chemical shifts of the C7 and C13 resonances are superimposed over the contour plot. As expected based on the results shown in Figure 10.8, the C8–C7 correlation is collapsed to a singlet since C7 is a quaternary carbon. In contrast, the C8–C13 correlation remains a doublet since both protons are directly bound to ^{13}C .

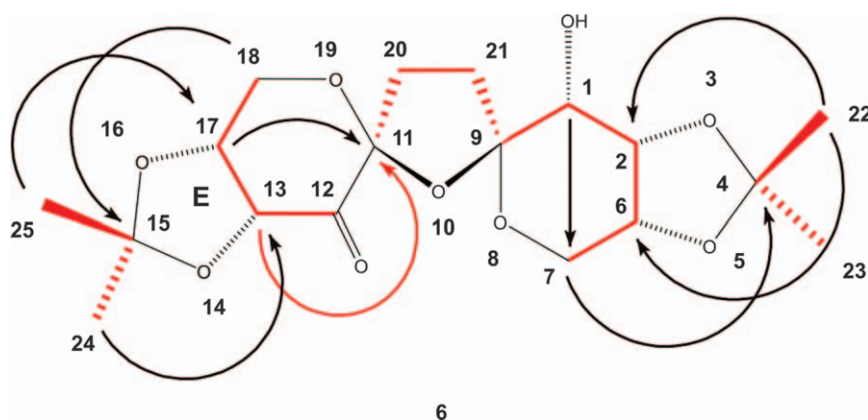
extracted from the spectrum at the F_1 shifts of the C7 and C13 resonances show the H8 methine doublet for the former to be collapsed to a singlet by the HD, while the correlation with the C13 methine carbon remains a doublet as expected.

Preliminary work done with ibuprofen provided a comparison of the performance of the conventional *versus* the 1,1-HD-ADEQUATE experiments. While useful and valid, the comparison was done on a model compound at high concentration. Although comparisons on “real-world” samples are not always available, during the development of the 1,1-HD-ADEQUATE experiment, we did have occasion to perform both the conventional and HD experiment on a sample of a synthetic compound from a collaborative investigation.¹¹ This study was also particularly interesting in that it highlighted a potential drawback of the improved sensitivity embodied in the homodecoupled experiment. The molecule being studied was not an

alkaloid but rather a rearrangement product of a spiroketal triggered by an enol ether epoxidation. Conventional 2D experiments (COSY, HSQC, and HMBC) were incapable of unequivocally defining the structure. Inputting those data into the Structure Elucidator™ CASE program (see Volume 1, Chapter 9) suggested structures 4 and 5 as the most probable structures for the rearrangement product, which prompted the ADEQUATE investigation.



An initial 40 Hz-optimized 1,1-ADEQUATE spectrum was acquired using the 2.5 mg sample of the reaction product available. The s/n ratio of the experiment was less than optimal, which prompted the decision to repeat the data acquisition using the then newly developed 1,1-HD-ADEQUATE experiment. Several comparison traces from the two spectra are shown in Figure 10.10. From the 1,1-HD-ADEQUATE spectrum, the structure of the reaction product was unequivocally assigned as 4. Key correlations from the 1,1-HD-ADEQUATE (bonds shown in red) and 2 Hz-optimized LR-HSQC spectra (black arrows) are shown by 6.



These data also highlight a potential drawback associated with the higher sensitivity of the 1,1-HD-ADEQUATE experiment. Specifically, the correlation from the H13 methine carbon *via* $^2J_{CC}$ to C11 is observed with unmistakable intensity in the 1,1-HD-ADEQUATE spectrum, whereas this correlation might be lost in the noise floor of the conventional 1,1-ADEQUATE

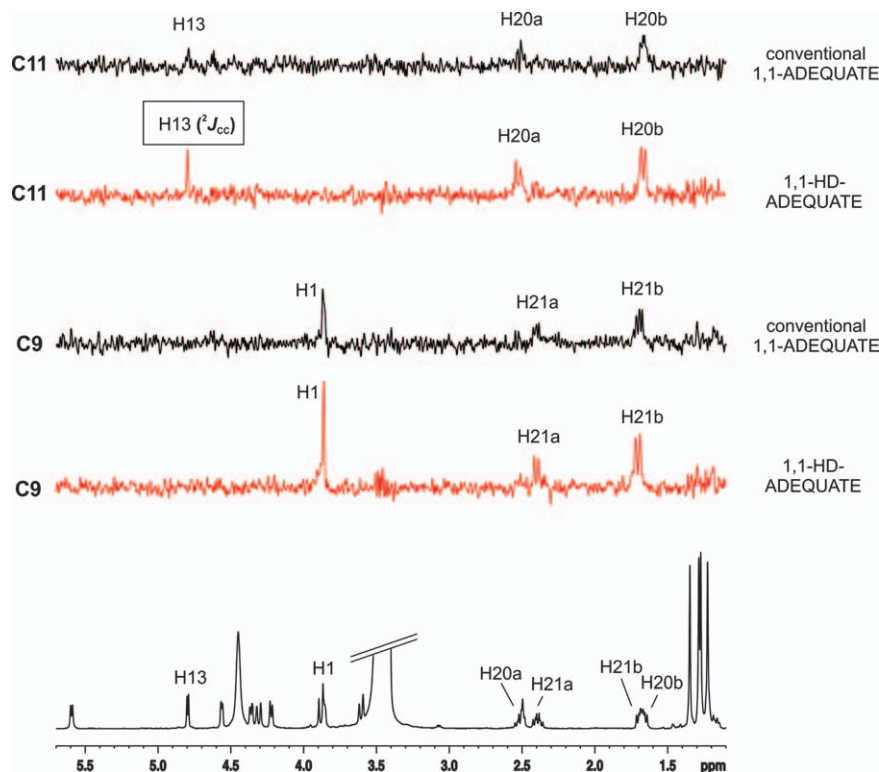


Figure 10.10 Traces for the C9 and C11 spiro carbons extracted from the 40 Hz conventional 1,1-ADEQUATE (black) and 1,1-HD-ADEQUATE (red) spectra of **4** (see also **6**).¹¹ The benefits of even partially collapsing multiplet structures are most evident for the correlations associated with the 11-spiro carbon. It should be noted, however, that the prominent $^2J_{CC}$ correlation of C13–C11 across the C12 carbonyl is also significantly enhanced by the HD provided by the 1,1-HD-ADEQUATE experiment. (Used with permission from: C. Lorenc, J. Sauri, A. Moser, A. V. Buevich, A. J. Williams, R. T. Williamson, G. E. Martin and M. W. Peczu, *ChemistryOpen*, 2015, **4**, 577–580; DOI: 10.1002/open. 201500122;¹¹ Copyright 2015 John Wiley and Sons.)

spectrum. DFT calculations predicted a 12.8 Hz $^2J_{CC}$ coupling between C13 and C11 across the C12 carbonyl.

When the amplitude transfer curves are calculated for various optimizations of the 1,1-HD-ADEQUATE experiment as shown in Figure 10.11, we can see that in a 40 Hz-optimized experiment, a 12.8 Hz $^2J_{CC}$ coupling constant would be expected to give rise to a correlation with 24% of the maximum possible intensity in a 40 Hz-optimized experiment.

To verify the unexpected observation of a $^2J_{CC}$ correlation in a 1,1-ADEQUATE spectrum, the actual carbon–carbon coupling constant was measured using a J -modulated ADEQUATE experiment.^{18,19} The actual $^2J_{CC}$ coupling constant was 11.3 Hz. Hence, while the intensity of the response

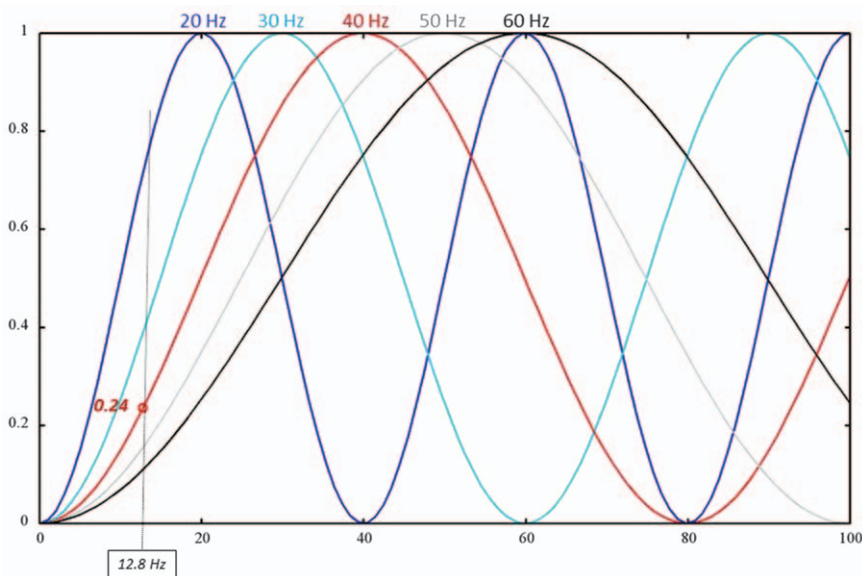


Figure 10.11 Calculated amplitude transfer curves for various optimizations of the 1,1-HD-ADEQUATE experiment. An anticipated 12.8 Hz ${}^2J_{CC}$ correlation would be expected to arise in a 40 Hz-optimized experiment at a correlation with an intensity of 0.24.

observed would be somewhat less than the calculated 0.24, it would still be expected and was sufficiently strong to be observed without question. This observation highlights a newly discovered potential ambiguity associated with 1,1-ADEQUATE data that should be taken into account whenever ${}^2J_{CC}$ correlations across a carbonyl are possible. It remains to be determined if other structural motifs will also give rise to similar detectable ${}^2J_{CC}$ correlations.

10.2 1,*n*-ADEQUATE and Advanced Variants

The 1,*n*-ADEQUATE experiment and variants were reviewed in 2014.¹⁰ The prefixes 1 and *n* correspond to the initial ${}^1J_{CH}$ and subsequent ${}^nJ_{CC}$ magnetization transfers, respectively. Once again, as with the 1,1-ADEQUATE experiment just considered, the 1,*n*-ADEQUATE experiment is an out-and-back experiment whose magnetization transfer pathways are shown schematically in Figure 10.12. While the HMBC²⁰ experiment affords predominantly ${}^3J_{CH}$ correlations, which can be supplemented by the greater “reach” of the LR-HSQMBC experiment,²¹ in contrast, the 1,*n*-ADEQUATE experiment affords predominantly ${}^3J_{CC}$ correlations, which are equivalent to ${}^4J_{CH}$ heteronuclear correlations in the HMBC and LR-HSQMBC experiments. Although the 1,*n*-ADEQUATE experiment will usually represent a “last resort” for reasons of the lower sensitivity when compared to HMBC and LR-HSQMBC, the experiment can, nevertheless, provide a critical

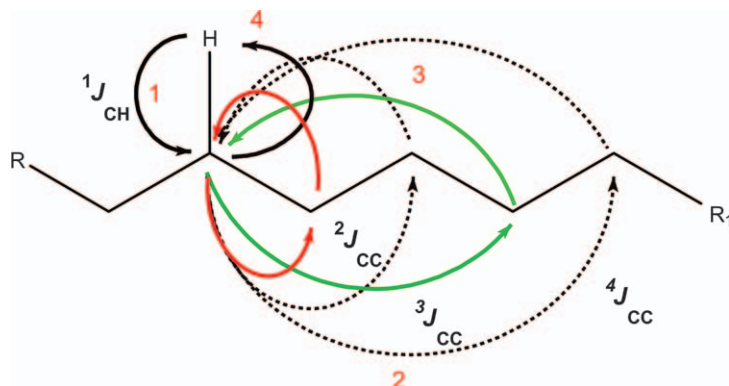


Figure 10.12 Magnetization transfer pathways of the $1,n$ -ADEQUATE experiment. The initial transfer is *via* the large $^1J_{\text{CH}}$ heteronuclear coupling to the directly attached ^{13}C . The subsequent transfer is nominally *via* $^nJ_{\text{CC}}$, which is most commonly with $n=3$ (solid green arrows), which corresponds to a $^4J_{\text{CH}}$ heteronuclear correlation in HMBC and LR-HSQMBC spectra. The $^2J_{\text{CC}}$ and $^4J_{\text{CC}}$ correlations are usually much weaker than the $^3J_{\text{CC}}$ correlations in $1,n$ -ADEQUATE spectra and are denoted by dashed black arrows. Unfortunately, unlike the HMBC experiment, there is no equivalent to the low-pass J -filter, so there is a high likelihood that a significant number of $^1J_{\text{CC}}$ correlations will “leak” into a $1,n$ -ADEQUATE spectrum (red arrows). The 5 Hz-optimized $1,n$ -ADEQUATE spectrum of strychnine (2) is shown in Figure 10.13, as well as a segment of the aliphatic region of the $1,n$ -ADEQUATE spectrum with the corresponding segment of the $1,1$ -ADEQUATE spectrum superimposed and offset in order to illustrate the intrusion of $^1J_{\text{CC}}$ correlations shown in Figure 10.14.

correlation(s) when they are not observed for whatever reason when using the more sensitive experimental alternatives.

$1,1$ - and $1,n$ -ADEQUATE share the same pulse sequence that was shown in Figure 10.4. The two experiments differ only in their optimization of the delays for $^1J_{\text{CC}}$ *versus* $^nJ_{\text{CC}}$ and, of course, their relative sensitivities, with $1,n$ -ADEQUATE being less sensitive than $1,1$ -ADEQUATE in the same sense that HMBC is less sensitive than HSQC. Delays for the $^nJ_{\text{CC}}$ magnetization transfers in a $1,n$ -ADEQUATE experiment are typically optimized in the range of 3–7 Hz¹⁰ as compared to $^1J_{\text{CC}}$ delays in the $1,1$ -ADEQUATE experiment that are typically optimized in the range of 40–60 Hz.⁹

As with correlations in the HMBC experiment, long-range carbon–carbon correlations, $^nJ_{\text{CC}}$, in a $1,n$ -ADEQUATE experiment are of indeterminate length and are typically $n=2$ –4, with $n=3$ predominating. A further caveat associated with the $1,n$ -ADEQUATE experiment pertains to $^1J_{\text{CC}}$ correlations. Unlike the HMBC experiment (see Chapter 9, Figure 9.19), which has a low-pass J -filter specifically designed to suppress direct correlations, there is no equivalent for $^1J_{\text{CC}}$ correlations that can “leak” into a $1,n$ -ADEQUATE spectrum.¹⁰ Hence, with any conventional $1,n$ -ADEQUATE spectrum,

$^1J_{CC}$ correlations must be specifically identified. A H2BC spectrum would provide the means of specifically identifying unwanted $^1J_{CC}$ correlations between pairs of protonated carbons, but a 1,1-ADEQUATE spectrum would be necessary to pinpoint correlations between a protonated and non-protonated carbon.

10.2.1 1,*n*-ADEQUATE

The 5 Hz-optimized 1,*n*-ADEQUATE spectrum of strychnine (**2**) is presented in Figure 10.13. *A priori*, there is no way to differentiate unwanted $^1J_{CC}$ correlations from the desired $^nJ_{CC}$ correlations in the spectrum shown in Figure 10.13. A segment of the aliphatic region of the 1,*n*-ADEQUATE spectrum is presented in Figure 10.14, with the corresponding segment from the 1,1-ADEQUATE spectrum shown in Figure 10.15 superimposed and offset slightly to the left, in order to make correlation overlaps more discernable.

Despite the challenges associated with the interpretation of 1,*n*-ADEQUATE spectra, they have been used several times in structure elucidation studies, including in the characterization of the marine natural alkaloid 5,6-dehydrolamellarin (**7**)⁷ in the initial report of the ADEQUATE experiments and, more recently, in the characterization of several phakellin and isophakellin alkaloids.²²

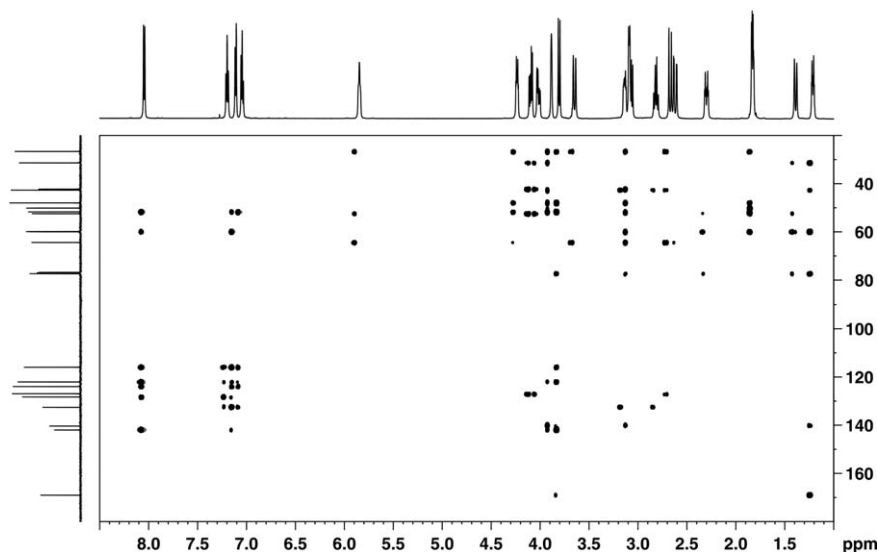


Figure 10.13 1,*n*-ADEQUATE spectrum of strychnine (**2**) optimized for 5 Hz. The data were acquired as 3K×160 points with an acquisition time of 48 hours, although usable data could be obtained in less time. As noted in the text and in the captioning of Figure 10.12, there are $^1J_{CC}$ correlations in this spectrum that need to be identified in some way in order to prevent misinterpretation if this were an unknown structure.

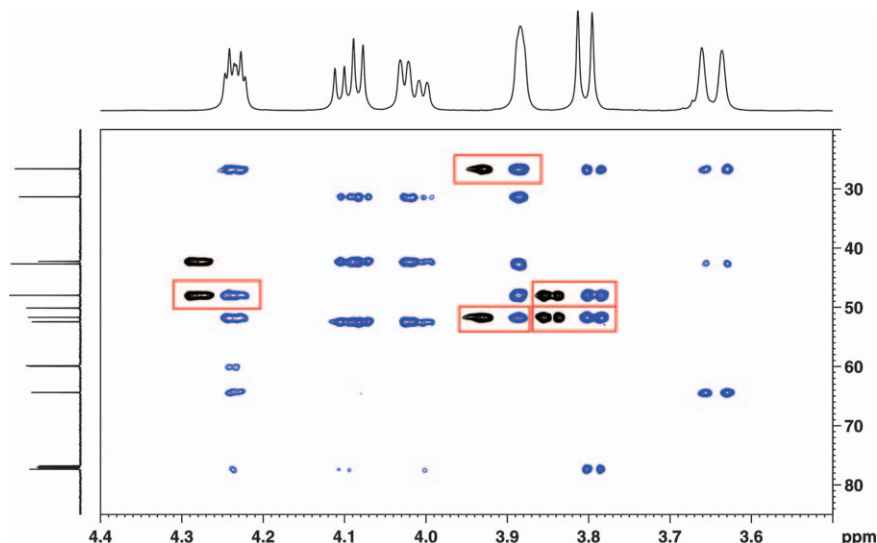


Figure 10.14 Superimposed aliphatic segments of the 5 Hz-optimized $1,n$ -ADEQUATE spectrum shown in Figure 10.13 (blue correlations) and the corresponding segment of the $1,1$ -ADEQUATE spectrum of strychnine (**2**) shown in Figure 10.5. The correlations from the latter are plotted in black and are slightly offset to the left to make them more discernable. Overlapping correlations in the $1,1$ - and $1,n$ -ADEQUATE spectra are boxed in red. Note there is one correlation from the $1,1$ -ADEQUATE spectrum (C12–C11) that does not overlap with a $1,n$ -ADEQUATE correlation.

10.2.2 $1,n$ -HD-ADEQUATE

As with the $1,1$ -ADEQUATE experiment, the low inherent sensitivity of the experiment, which is a direct consequence of the $\sim 1.1\%$ relative abundance of ^{13}C , has had a significant impact on the utilization of the experiment. However, in the same sense that $1,1$ - and $1,n$ -ADEQUATE share the same pulse sequence, differing only in the duration of the delays used for ^{13}C – ^{13}C transfer, the HD pulse sequence shown in Figure 10.7 can also be used to acquire $1,n$ -ADEQUATE data with HD. Again returning to ibuprofen as a model system to compare the results from $1,n$ -ADEQUATE and $1,n$ -HD-ADEQUATE spectra, comparison traces are shown in Figure 10.15. In this example, the $^3J_{\text{CC}}$ correlation of C2–C5 was chosen. It is important to note that both of these carbons are protonated. The black trace shows the $^3J_{\text{CC}}$ correlation of C2–C5, with the H2 resonance observed as a quartet as expected from homonuclear coupling to the adjacent methyl group. The corresponding correlation from the 5 Hz-optimized $1,n$ -HD-ADEQUATE spectrum is shown in red. The multiplet structure has been completely collapsed. The HD beneficially increased the s/n ratio for the correlation from 146:1 in the conventional experiment to 243:1 in the

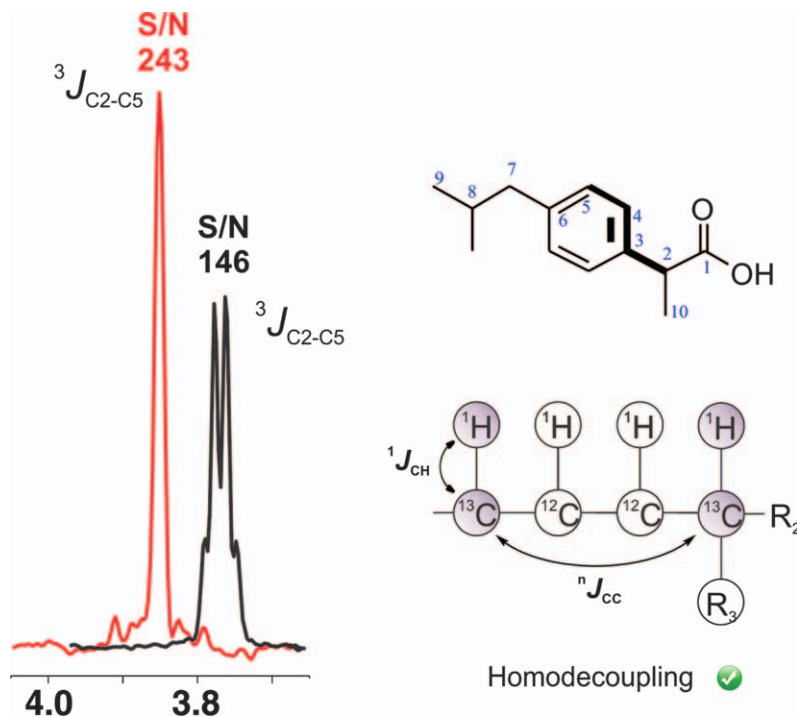
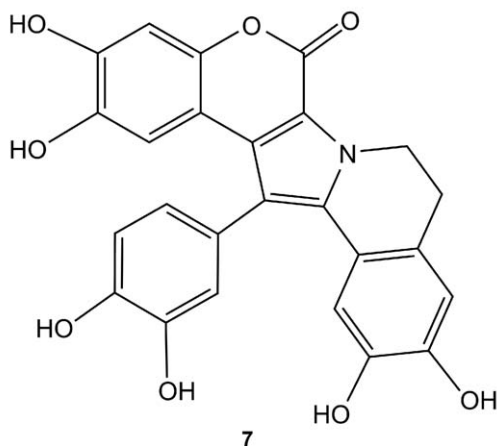


Figure 10.15 Comparison of the C2–C5 $^3J_{CC}$ correlations of ibuprofen (**3**) from the conventional, 5 Hz-optimized $1,n$ -ADEQUATE spectrum (quartet shown in the black trace) and the 5 Hz-optimized $1,n$ -HD-ADEQUATE spectrum (red trace). In addition to the collapse of the multiplet structure, there is also a >66% increase in s/n for the $^3J_{CC}$ correlation between C2 and C5.

$1,n$ -HD-ADEQUATE spectrum. The schematic representation for the $^nJ_{CC}$ coupling pathway is also shown in Figure 10.15. Unlike the $^1J_{CC}$ correlation between two protonated carbons that was not collapsed by the BIRD-based HD scheme, the only correlation involving two protons on ^{13}C is the $^5J_{HH}$ correlation between $^1\text{H}^{13}\text{C}(2)$ and $^1\text{H}^{13}\text{C}(5)$, which will almost always be too small to be resolved in a 2D NMR experiment. In contrast, all of the vicinal proton–proton homonuclear couplings will be of the type $^1\text{H}^{13}\text{C}_A\text{--}^1\text{H}^{12}\text{C}_B$, which will be collapsed by the BIRD-based HD used in the experiment. The important point to be gleaned from this example is that all homonuclear proton–proton couplings in a $1,n$ -HD-ADEQUATE spectrum will be effectively collapsed, with the sole exception of the unresolved $^nJ_{HH}$ coupling where $n = 4\text{--}6$ between the two protons on ^{13}C that are correlated in the $1,n$ -HD-ADEQUATE experiment. Since the $1,n$ -ADEQUATE experiment is substantially lower in sensitivity than $1,1$ -ADEQUATE, the gain in s/n based on the collapse of proton multiplets is significant and highly beneficial. There is also the ancillary benefit derived from the improved resolution in

the proton spectrum, although that will likely be secondary to the improvement in sensitivity in most cases.



The 5 Hz-optimized $1,n$ -HD-ADEQUATE spectrum of strychnine (2) is shown in Figure 10.16. Quite obviously, a $1,n$ -HD-ADEQUATE spectrum will present the same challenges as a conventional $1,n$ -ADEQUATE spectrum insofar as differentiating ${}^nJ_{CC}$ correlations from the ${}^1J_{CC}$ correlations that unavoidably leak into a $1,n$ -ADEQUATE spectrum.

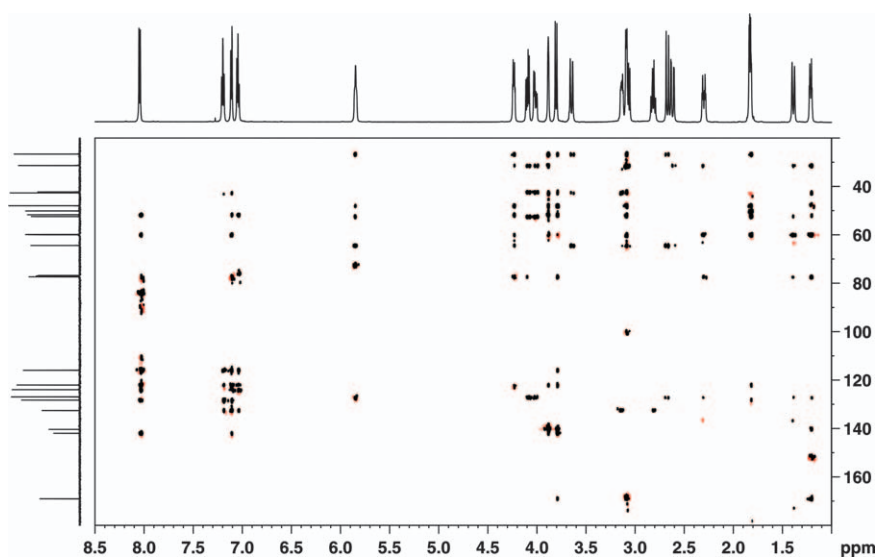


Figure 10.16 $1,n$ -HD-ADEQUATE spectrum of strychnine (2) optimized for 5 Hz. The data were acquired using the pulse sequence shown in Figure 10.7. The expansion of the aliphatic region is shown in Figure 10.17 and an overlay of the slice extracted at the F_1 chemical shift of C15 is presented in Figure 10.18.

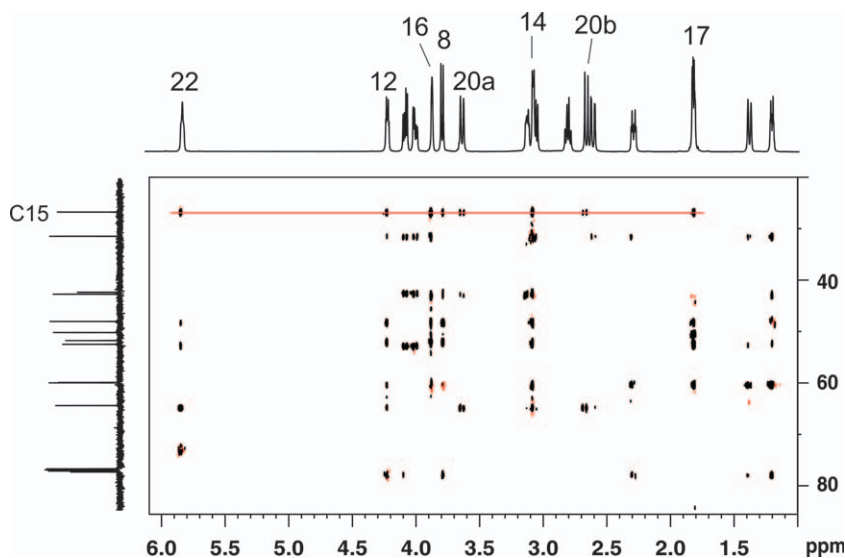


Figure 10.17 Expansion of the aliphatic region of the 5 Hz-optimized $1,n$ -HD-ADEQUATE spectrum of strychnine (2). Correlations for the C15 resonance as labeled and the corresponding slice is compared to the proton spectrum in Figure 10.18.

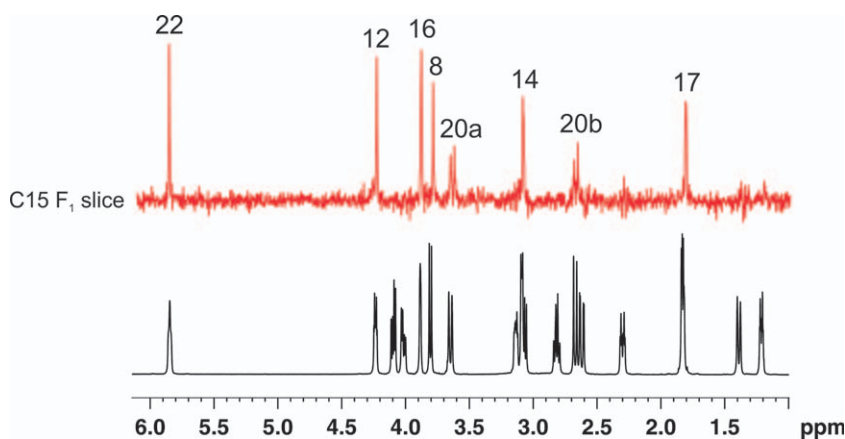


Figure 10.18 Proton reference spectrum (black trace) and the C15 F_1 slice (red trace) extracted from the 5 Hz-optimized $1,n$ -HD-ADEQUATE spectrum of strychnine (2) shown in Figure 10.17. As expected, the H16, H20a, and H20b multiplets were not collapsed, although the couplings of the H16 resonance are not resolved. The other homonuclear J_{HH} couplings are collapsed by the experiment as expected and as shown in the ibuprofen slices in Figure 10.15.

10.2.3 Edited 1,*n*-HD-ADEQUATE

Multiplicity editing of HSQC spectra (see Section 9.5.2) is accomplished by the simple insertion of a π -pulse sandwich in the pulse sequence as shown in Figure 9.4 (see Chapter 9). Multiplicity editing in the HSQC experiment is based on resonance multiplicity and is relatively simple to introduce. In contrast, it would be desirable to “edit” 1,*n*-ADEQUATE spectra on the basis of the size of the homonuclear J_{CC} coupling. We could perhaps label this type of spectral manipulation as “scalar editing”. Differentiating $^1J_{CC}$ from $^nJ_{CC}$ homonuclear coupling cannot, however, be accomplished by anything as simple as a π -sandwich. Indeed, if that approach were tried, everything would be inverted, which would not be especially useful. ^{13}C – ^{13}C homonuclear couplings, however, are significantly different in size, with $^1J_{CC}$ couplings typically in the range of 30–70 Hz (there are exceptions on both ends of the range), while $^nJ_{CC}$ couplings where $n = 2$ –4 are typically in the range of 2–7 Hz, although some $^2J_{CC}$ couplings can be as large as 15 Hz.

If we take advantage of the difference in the oscillatory rates of the $^1J_{CC}$ and $^nJ_{CC}$ correlations and adjust the duration of the delays in the pulse sequence, we can select conditions where $^1J_{CC}$ correlations have negative intensity while $^nJ_{CC}$ correlations have positive intensity, thereby achieving the goal of “scalar editing”.^{23,24} After the initial INEPT transfer from ^1H to ^{13}C , both $^1J_{CC}$ and $^nJ_{CC}$ components of magnetization evolve according to $\sin(J\pi\Delta)$. Since the ADEQUATE experiments are out-and-back experiments, the same amplitude transfer function governs the evolution of magnetization during the delay prior to the reverse INEPT step in the pulse sequence. Hence, the overall amplitude transfer is governed by $\sin^2(J\pi\Delta)$, and both the $^1J_{CC}$ and $^nJ_{CC}$ components of magnetization have positive intensity, precluding differentiation of $^1J_{CC}$ from $^nJ_{CC}$ correlations. However, by making the sequence asymmetric in terms of the delay durations before and after evolution, transfer amplitudes can be obtained that are negative for $^1J_{CC}$ correlations but remain positive for $^nJ_{CC}$ correlations, thereby providing a basis for differentiating the two types of correlations. The amplitude transfer delays were modified as follows: the delay prior to evolution was lengthened as a function of $(2m+1)/(2^1J_{CC})$; and the delay following evolution was shortened according to the function $(2m-1)/(2^1J_{CC})$. The rapid oscillatory nature of $^1J_{CC}$ is shown in Figure 10.19, where response intensity is plotted as a function of $^1J_{CC}$. Even cursory inspection shows that most of the typical range of $^1J_{CC}$ coupling constants is negative. Conversely, $^nJ_{CC}$ correlations from 0 to ~11 Hz have positive intensities, as shown in Figure 10.20, which provides a basis for differentiating $^1J_{CC}$ from $^nJ_{CC}$ correlations in a modified 1,*n*-ADEQUATE experiment.

While the discussion above does provide the means of editing a 1,*n*-ADEQUATE spectrum,²³ there were still positive lobes in the amplitude transfer curve shown in Figure 10.19 that could lead to positive $^1J_{CC}$

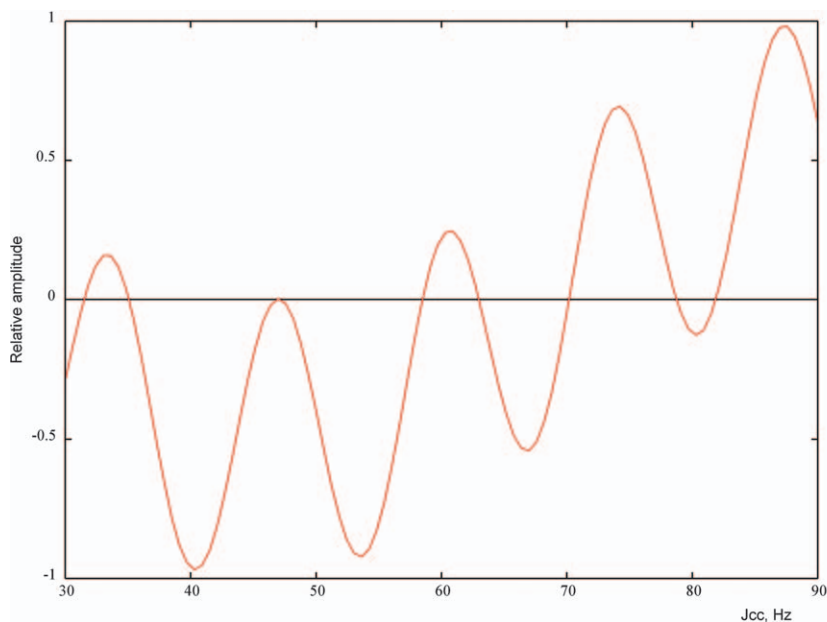


Figure 10.19 Amplitude transfer curve for $^1J_{CC}$ correlations across the range of 30–90 Hz for asymmetrical delays of the type $(2m+1)/(2^1J_{CC})$ and $(2m-1)/(2^1J_{CC})$ flanking the evolution period of a modified ADEQUATE pulse sequence. (Reproduced from: *Magn. Reson. Chem.*, 2012, **50**, 563–568 with permission from John Wiley and Sons. Copyright © 2012 John Wiley & Sons Ltd.)

correlations in an inverted $^1J_{CC}$ 1,*n*-ADEQUATE spectrum. Indeed, that is exactly the case for the C22 correlation of strychnine (**2**), which has a 72 Hz $^1J_{CC}$ coupling to C21, which places that correlation in a region that is expected to exhibit a positive correlation, which is what happened experimentally. A further modification of the pulse sequence was subsequently introduced²⁴ that employed dual optimization to give a much broader inversion of $^1J_{CC}$ correlations, as shown by the amplitude transfer curves in Figure 10.21, while maintaining a positive intensity for $^nJ_{CC}$ correlations. As shown in Figure 10.22, $^nJ_{CC}$ correlations have a positive intensity across the range of 0–15 Hz (blue trace).

To illustrate the function of the dual-optimized inverted $^1J_{CC}$ 1,*n*-ADEQUATE experiment, the spectrum of strychnine (**2**) acquired using optimizations of $^1J_{CC} = 57$ Hz and $^nJ_{CC} = 9.5$ Hz and optima of $^1J_{CC} = 64$ Hz and $^nJ_{CC} = 8$ Hz for the second set of parameters is shown in Figure 10.23. The problematic $^1J_{CC} = 72$ Hz coupling between C22 and C21, which resulted in a positive intensity for this correlation when single optimization of the experiment was done, gave a correctly inverted correlation in the dual-optimized experiment.

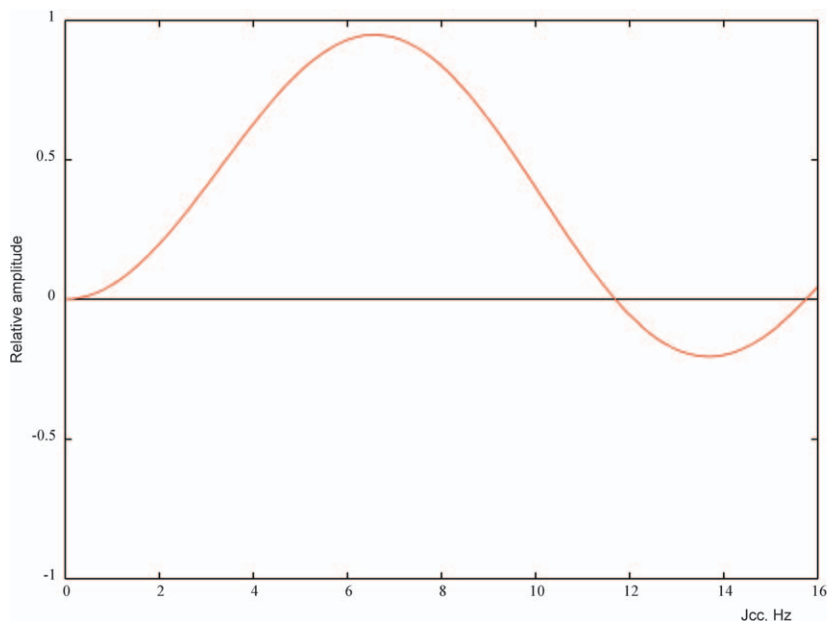


Figure 10.20 Amplitude transfer curve for $^nJ_{CC}$ correlations across the range of 0–16 Hz for asymmetrical delays in a modified ADEQUATE experiment. The $^nJ_{CC}$ correlations have positive intensity across the most commonly encountered range for these correlations.

(Reproduced from: G. E. Martin, W. Bermel, R. T. Williamson and P. G. Dormer, *Magn. Reson. Chem.*, 2012, **50**, 563–568²³ with permission from John Wiley and Sons. Copyright © 2012 John Wiley & Sons Ltd.)

A comparison of the F_1 slices taken at the C21 chemical shift is presented in Figure 10.24. The C21 slice from a single-optimum inverted $^1J_{CC}$ 1, n -ADEQUATE spectrum²³ is shown in Figure 10.24. The correlation for C21–C22 is strongly positive, whereas the C20–C21 correlation is inverted as intended by the experiment. When the data were acquired using the dual-optimized approach, with the optimization values used to generate the amplitude transfer curves shown in Figure 10.21, the C21–C22 correlation was now successfully inverted. It is also worth noting that the intensity of the correlations at the 1H shifts of H8 and H13 were both significantly more intense than they were in the single-optimized experiment.

10.3 Examples of the Application of Advanced NMR Methods in Structure Elucidation Studies

There appears to be an induction period between when techniques are first developed and applied in the laboratories of the originators and the

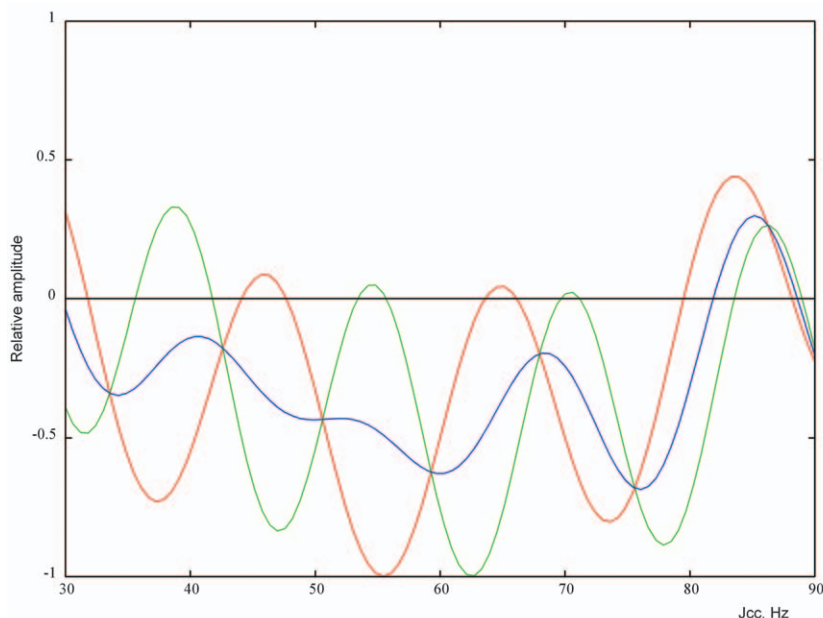


Figure 10.21 Amplitude transfer curves for $^1J_{CC}$ correlations across the range of 30–90 Hz for a dual-optimized inverted $^1J_{CC}$ 1,*n*-ADEQUATE experiment. Optimization of the delays for $^1J_{CC} = 57$ Hz and $^2J_{CC} = 9.5$ Hz gives the result shown by the red curve; optimization for $^1J_{CC} = 64$ Hz and $^2J_{CC} = 8$ Hz gives the result shown by the green curve. When the two curves are summed and normalized, the result is shown by the blue curve. Using this approach, $^1J_{CC}$ correlations across the range of 27–82 Hz will be inverted, which covers most of the range of commonly encountered $^1J_{CC}$ coupling constants. (Reproduced from: M. Reibarkh, R. T. Williamson, G. E. Martin and W. Bermel, Broadband inversion of $^1J_{CC}$ responses in 1,*n*-ADEQUATE spectra, *J. Magn. Reson.*, **236**, 126–133, Copyright (2013)²⁴ with permission from Elsevier.)

more general acceptance of techniques by the NMR community. One very notable example is the length of time it took for HSQC to supplant the older HMQC experiment that affords significantly lower F_1 digital resolution. Indeed, impromptu surveys at 5 year intervals have shown that even after 20 years the HMQC experiment is still widely and perhaps inexplicably utilized by the natural products community. How long it may take some of the experiments described in this chapter to become widely adopted is almost impossible to predict. Nevertheless, there are several examples of applications of these techniques to spectral assignment and structure elucidation studies that are worth noting to close this chapter.

The previous chapter discussed a range of what might be considered first- and second-tier experiments that can be employed in the

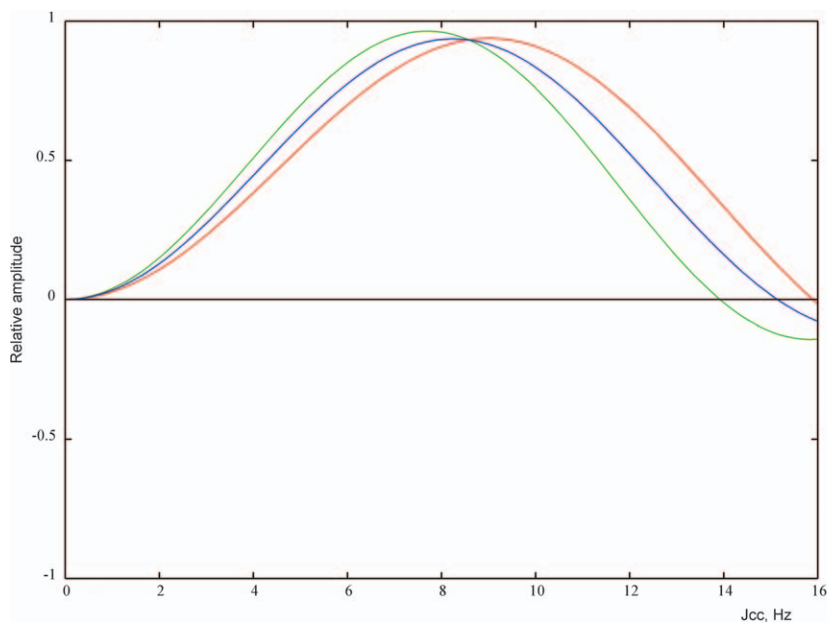


Figure 10.22 Response intensity for ${}^nJ_{CC}$ correlations across the range of 0–16 Hz in a dual-optimized inverted ${}^1J_{CC}$ 1, n -ADEQUATE experiment using the optimizations specified in the caption of Figure 10.21. The ${}^nJ_{CC}$ correlations maintain useful positive intensity across the range commonly encountered for these couplings.
(Reproduced from: M. Reibarkh, R. T. Williamson, G. E. Martin and W. Bermel, Broadband inversion of $1J_{CC}$ responses in 1, n -ADEQUATE spectra, *J. Magn. Reson.*, **236**, 126–133, Copyright (2013)²⁴ with permission from Elsevier.)

characterization of alkaloids, or any other class of natural products for that matter. This chapter has delved into what might be labeled third-tier experiments—the group of experiments that you consider resorting to when nothing else has worked to provide the necessary correlations to deal with a troublesome structure characterization problem. At this point, it is appropriate to examine some recently reported structural work that has applied some of the range of experiments described in this and the previous chapter. We will begin with a study of the alkaloid staurosporine (**8**), followed by a discussion of the revision of the structure of the complex *Cryptolepis* alkaloid cryptospirolepine (**1**). We will conclude the chapter with some work done to define the structure of the novel alkaloid eudistidine-C (**10**). The structure of eudistidine-C was suggested by a calculation using the Structure Elucidator™ CASE program, but the necessary long-range heteronuclear correlations to confirm the structure were not observable in HMBC²⁰ spectra, irrespective of the optimization. The structure was confirmed using the LR-HSQMBC experiment.²¹

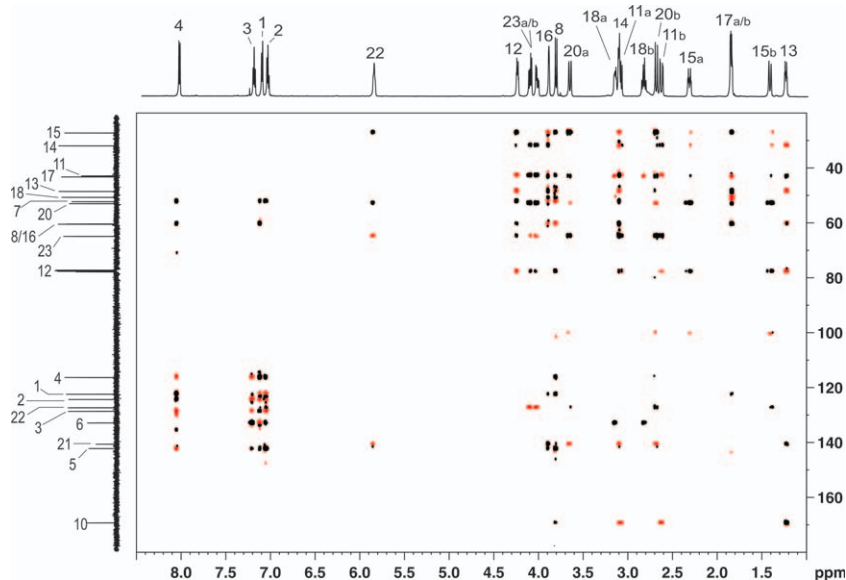


Figure 10.23 Dual-optimized inverted $^1J_{CC}$ 1, n -ADEQUATE spectrum of strychnine (2). These data were acquired with optima of $^1J_{CC}=57$ Hz and $^nJ_{CC}=9.5$ Hz for the first set and optima of $^1J_{CC}=64$ Hz and $^nJ_{CC}=8$ Hz for the second set. Inverted $^1J_{CC}$ correlations are plotted in red, while $^nJ_{CC}$ correlations have a positive intensity and are plotted in black.

10.3.1 Staurosporine—Is the Crew's Rule Obsolete?

Staurosporine (8) is a complex indolocarbazole alkaloid that was first characterized by Meksuriyen and Cordell in 1988.²⁵ We have employed the molecule to highlight the application of the dual-optimized inverted $^1J_{CC}$ 1, n -ADEQUATE experiment²⁶ and as a model compound during the course of a study to contrast the HMBC and the 1,1- and 1, n -ADEQUATE experiments.¹² The two four-spin aromatic systems were sequenced and unequivocally assigned using a series of IDR-HSQC-TOCSY spectra with mixing times ranging from 12 to 48 ms (see Chapter 9, Section 9.9.2).

When the sample of staurosporine was interrogated using the dual-optimized inverted $^1J_{CC}$ 1, n -ADEQUATE spectrum shown in Figure 10.25, correlations were observed for all possible $^1J_{CC}$ correlations. To maintain a sufficiently high level of digital resolution in the aromatic region of the spectrum, the spectrum was intentionally folded in F_1 . The green arrows denote $^nJ_{CC}$ correlations that were extracted for the aliphatic carbon-carbon correlations from the folded data. Black arrows denote $^nJ_{CC}$ correlations; dashed black arrows denote weak $^nJ_{CC}$ correlations. From the ensemble of $^nJ_{CC}$ correlations shown on 8, the connectivity network for the molecule was clearly well determined, including multiple assignments to the C12a and

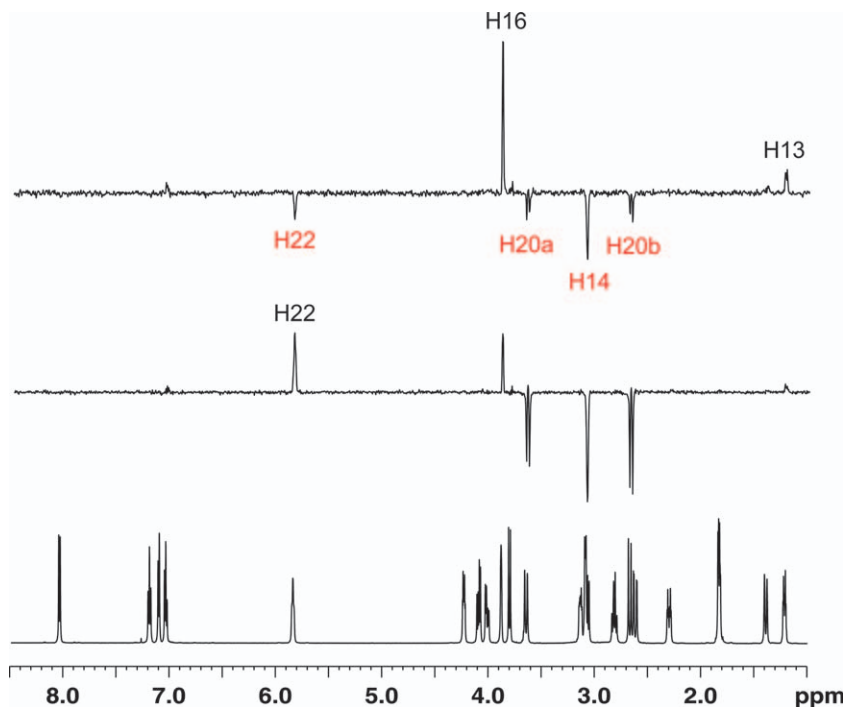


Figure 10.24 Comparison of the F_1 slices at the chemical shift of C21 from the single- and dual-optimized inverted $^1J_{CC}$ 1, n -ADEQUATE spectra.^{23,24} (Bottom) Proton reference spectrum of strychnine (2). (Middle) Slice from the single-optimized spectrum where response intensity is governed by the amplitude transfer curve shown in Figure 10.19.²³ The $^1J_{CC}$ coupling between C21 and C22 is 72 Hz and would be expected to exhibit a strong positive correlation and does. (Top) Slice at the C21 F_1 frequency from the dual-optimized inverted $^1J_{CC}$ 1, n -ADEQUATE spectrum acquired with optimization pairs of 57/9.5 and 64/8 Hz for $^1J_{CC}/^nJ_{CC}$. Response intensity is governed by the amplitude transfer curves shown in Figure 10.21. The expectation that the C21–C22 correlation would be properly inverted in this experiment is met.²⁴ It is also interesting to note that the correlations of H8 and H13 with C21 also have greater intensity in the trace from the dual-optimized spectrum (top panel) than they do in the single-optimized spectrum (middle panel). (Reproduced from: M. Reibarkh, R. T. Williamson, G. E. Martin and W. Bermel, Broadband inversion of $^1J_{CC}$ responses in 1, n -ADEQUATE spectra, *J. Magn. Reson.*, **236**, 126–133, Copyright (2013)²⁴ with permission from Elsevier.)

C12b resonances in the central core of the molecule that could be problematic to observe in HMBC data. While HMBC might be problematic in terms of correlations to the C12a and C12b resonances, it is quite probable that these correlations would be observable in a 2 or 4 Hz-optimized LR-HSQMBC spectrum (see Chapter 9, Section 9.8.2).²¹

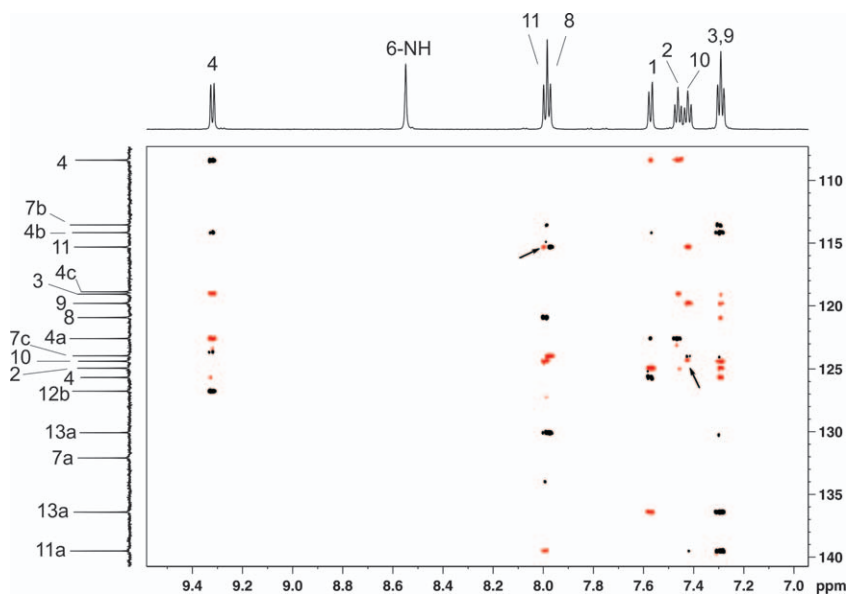
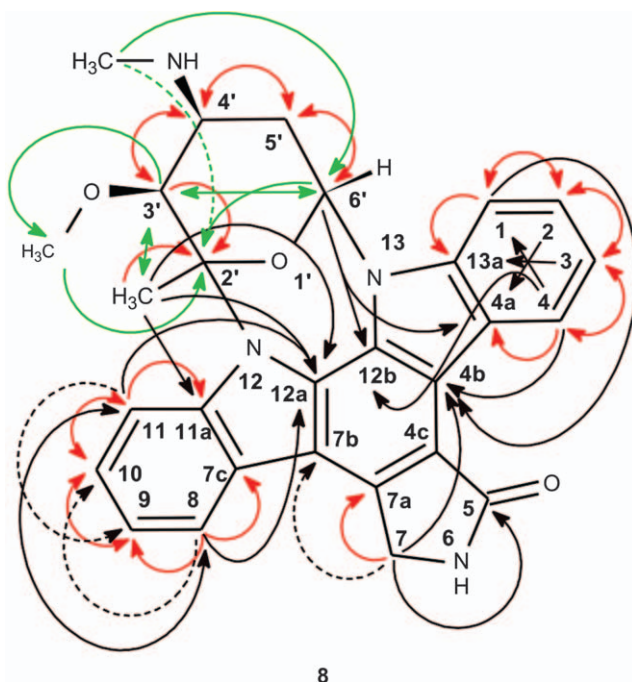


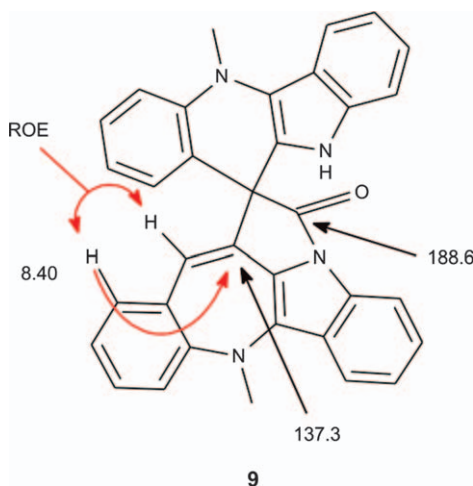
Figure 10.25 Dual-optimized inverted $^1J_{CC}$ 1, n -ADEQUATE spectrum of staurosporine (8). The data were acquired with the following pairs of optima: $^1J_{CC}/nJ_{CC} = 57/9.5$ and $64/8$ Hz. The amplitude transfer curves governing the $^1J_{CC}$ and nJ_{CC} correlations are those shown in Figures 10.21 and 10.22, respectively.

The informal Crew's rule suggests that a ratio of heavy elements (C, N, O, and perhaps S) to hydrogen of <2 can lead to a situation where the molecule can be difficult or even impossible to characterize.^{27,28} While that may well have been true using the conventional suite of what might be considered first-tier NMR experiments (HSQC, COSY or TOCSY, and HMBC), with experiments such as LR-HSQMBC, HSQMBC-TOCSY, and the experiments described in Chapter 9, it is quite likely that a revision of the Crew's rule is in order.^{27,28} Maybe the ratio needs to be set to 1, or perhaps even 0.5?

10.3.2 Cryptospirolepine—Resolution of a Long-standing Structural Ambiguity Using 1,1- and 1,*n*-HD-ADEQUATE Spectra

The structure of the complex, sprioononacyclic *Cryptolepis* alkaloid cryptospirolepine²⁹ (**9**) was reported in 1993. The complexity of this alkaloid was unprecedented from any *Cryptolepis* species and predated the discovery of bis-indoloquinoline alkaloids by several years. The structure was assembled using HMQC and HMBC data, which then represented cutting-edge 2D NMR experiments. The structure elucidation also predated the development of ¹H–¹⁵N HMBC experiments by several years, and CASE programs were then too immature to even consider trying to determine a molecule the size of cryptospirolepine.

There were several key features of the structure that warrant mention. First was a carbonyl resonating at 188.4 ppm that exhibited no HMBC correlations, irrespective of the optimization of the experiment. Based on this observation, it was logical to conclude that this carbon had to be a minimum of four bonds from the nearest proton to account for the lack of HMBC correlations. Second, there was a vinyl singlet resonating at 6.27/93.2 ppm that exhibited a ROE correlation to an aromatic proton resonating at 8.40 ppm. Third was a quaternary carbon resonating at 137.3 ppm that exhibited a correlation to the proton resonating at 8.40 ppm in the HMBC spectrum. In concert, these data led to the proposal of an indolobenzazepine system, which at the time was unprecedented in *Cryptolepis* species, although the alkaloid homocryptolepinone, which had an indolobenzazepine skeleton, was isolated and characterized in 1995.³⁰ The structure proposed on the basis of the available NMR data in 1993 is shown by **9**. There were, however, several features of the structure that were troubling. First, the carbonyl chemical shift, as an amide, was far downfield of the normal range of amide carbonyl chemical shifts at 188.6 ppm. The intensity of the HMBC correlation from the proton resonating at 8.40 ppm to the quaternary carbon at 137.4 ppm seemed too great for a ⁴J_{CH} correlation.



In 2002, the original NMR tube was cut open and chromatographed. Shockingly, there were 26 peaks in the chromatogram (Figure 10.26), with no intact cryptospirolepine remaining.

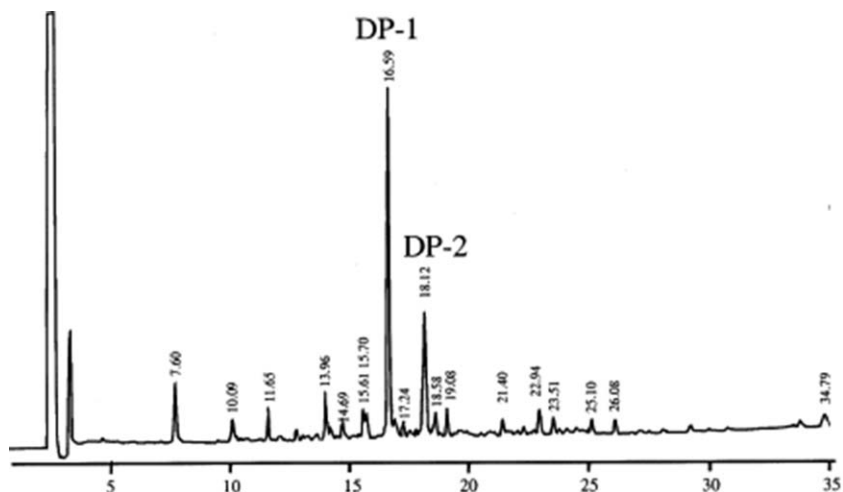
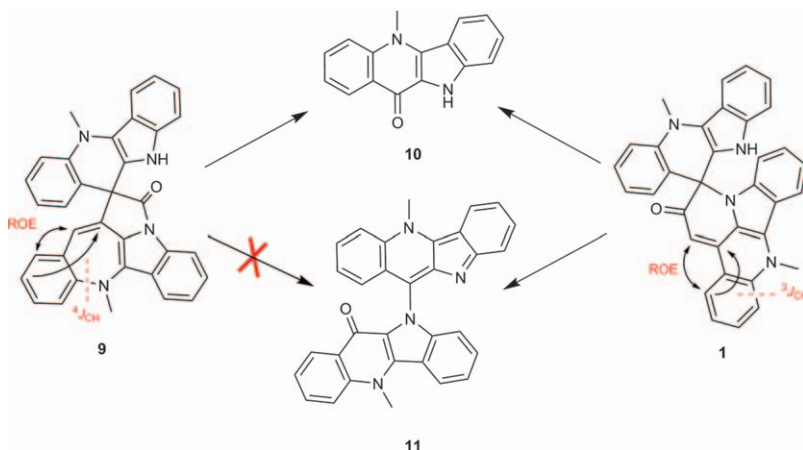


Figure 10.26 Chromatogram of the original cryptospirolepine sample (9) recorded in 2002 when the NMR tube containing the original sample in DMSO- d_6 was cut open. None of the intact alkaloid remained. The two largest components were isolated and characterized.³⁰

(Reproduced from: G. E. Martin, C. E. Hadden, D. J. Russell, B. D. Kaluzny, J. E. Guido, W. K. Duholke, B. A. Stiemsma, T. J. Thamann, R. C. Crouch, K. Blinov, M. Elyashberg, E. R. Martirosian, S. G. Molodtsov, A. J. Williams and P. L. Schiff, Jr., *J. Heterocycl. Chem.*, 2002, **39**, 1241–1250³¹ with permission from John Wiley and Sons. Copyright © 2002 Journal of Heterocyclic Chemistry.)

The two isolated major degradants of cryptospirolepine were unequivocally characterized as shown in Scheme 10.1. Cryptolepinone, the structure shown in the top center of the scheme, could be mechanistically rationalized from cryptospirolepine, whereas the dimeric bis-indoloquinoline shown in the bottom center could not be. Because of the limited sample remaining (<2 mg), work to revise the structure of cryptospirolepine had to be postponed for another 12 years.

In the interim from the 2002 report characterizing two of the degradants of cryptospirolepine,³¹ there were significant advances in our ability to deal with small samples and challenging structure elucidation problems. Small-volume and cryogenic probes are now widely available.^{32–42} A significant number of powerful new NMR methods have been developed, most of which have been discussed earlier in this chapter. There were also significant advances in the capabilities of CASE programs^{43–45} and in computing capabilities that now make it feasible to perform DFT calculations on a molecule the size of cryptospirolepine. These developments essentially set the stage to facilitate the revision of the structure of cryptospirolepine reported in 2015.³



Scheme 10.1 The structure of cryptospirolepine, **9**, proposed in 1993, is shown on the left.²⁹ The structures determined for the two major degradants, **10** and **11**, isolated from the original degraded sample in 2002, are shown in the center.³¹ While the structure of cryptolepinone (**10**) could be rationalized as a degradant of cryptospirolepine, in contrast, the dimeric bis-indoloquinoline structure (**11**) could not be rationalized mechanistically from cryptospirolepine.

(Reproduced from: J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff, Jr., T. Parella, R. T. Williamson and G. E. Martin, *Angew. Chem., Int. Ed.*, 2015, **54**, 10160–10164³ with permission from John Wiley & Sons. Copyright © 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

10.3.2.1 Cryptospirolepine—Attacking the Problem with High-sensitivity Experiments

Reinvestigating the structure of cryptospirolepine has several inherent challenges. First, the solubility of the alkaloid is limited to $\sim 700 \mu\text{g } 35 \mu\text{L}^{-1}$ DMSO- d_6 , the volume of a 1.7 mm MicroCryoProbe™ sample. Given the solubility limitations, the investigation was initially based on high-sensitivity 2D NMR experiments, beginning with PS-HSQC (see Chapter 9, Section 9.5.3). Because of partial overlaps in the 600 MHz ^1H NMR spectrum, the four-spin systems were sequenced using IDR-HSQC-TOCSY (see Chapter 9, Section 9.9.2) to provide unequivocal ^1H and ^{13}C chemical shift assignments for all of the protonated carbon resonances. The 45 ms IDR-HSQC-TOCSY spectrum is shown in Figure 10.27.

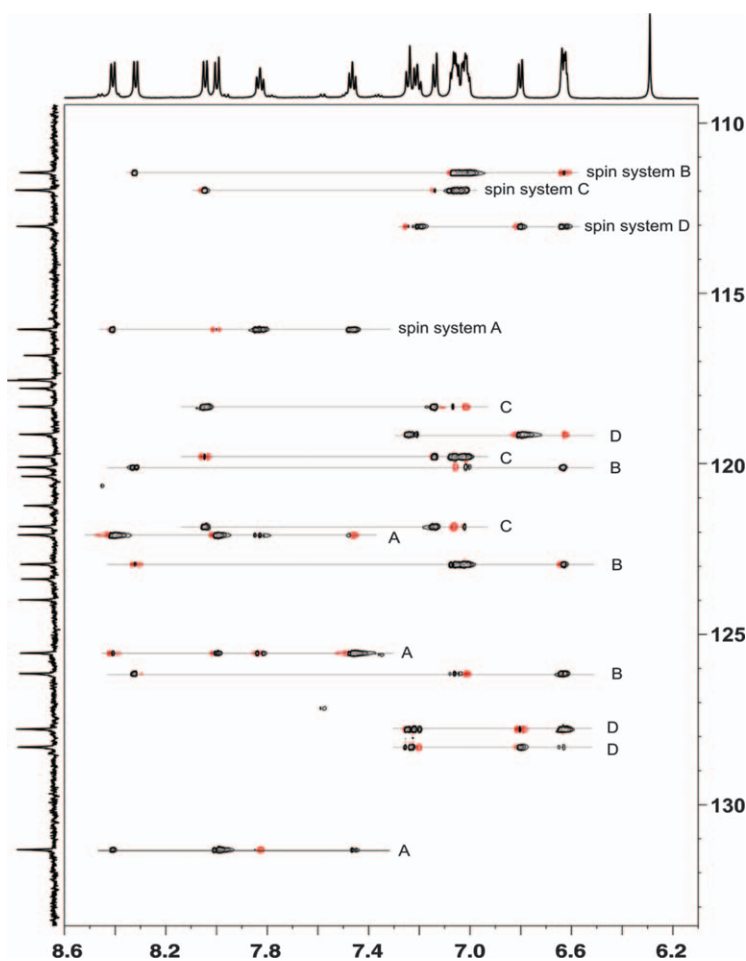


Figure 10.27 IDR-HSQC-TOCSY spectrum of cryptospirolepine (1) recorded with a 45 ms mixing time. The component resonances of the four four-spin systems are labeled A through D.

Next, 8 and 4 Hz-optimized HMBC (see Chapter 9, Section 9.8.1) spectra were acquired. Despite the vastly improved sensitivity of modern cryoprobes relative to early 1990 vintage room temperature gradient-inverse probes, there was still no hint of a correlation with the carbonyl resonance in either of the two HMBC spectra. Despite the inability of the HMBC experiment to visualize correlations with the carbonyl resonance of cryptospirolepine, the same was not true for a 2 Hz-optimized LR-HSQMBC (see Chapter 9, Section 9.8.2) experiment.²¹ To maintain high F_1 digital resolution in the aromatic region of the spectrum, the carbonyl and spiro carbon resonances were intentionally folded in the LR-HSQMBC spectrum that was acquired, which is shown in Figure 10.28. In the inset in Figure 10.28, there are two frank correlations to the carbonyl carbon from the two *N*-methyl resonances and a tantalizing hint of a correlation from the vinyl singlet. Although their presence in the spectrum was gratifying, the former did not afford structurally relevant information. Indeed, when the structure was fully established, it could be determined that the two *N*-methyl correlations were $^6J_{CH}$ and $^7J_{CH}$ correlations, the latter being the longest correlation in a LR-HSQMBC spectrum that we are aware of. The putative correlation from the vinyl proton singlet was rather weak, with a *s/n* ratio of $\sim 2:1$, so it could not be ascertained whether or not this was a legitimate $^2J_{CH}$ correlation from the vinyl proton to the carbonyl or fortuitous noise.

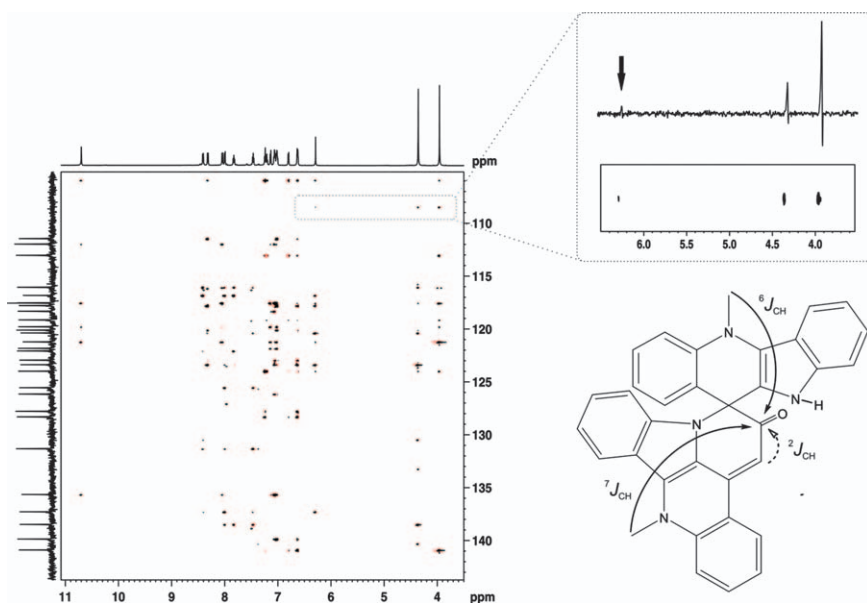


Figure 10.28 ^1H - ^{13}C LR-HSQMBC spectrum of **1** optimized for 2 Hz. Note that the spectral window along the carbon dimension has been reduced down to 40 ppm and therefore the carbonyl resonance is intentionally folded in the F_1 frequency domain appearing at 108.4 ppm rather than the normal 188.4 ppm. An expanded area corresponding to the carbonyl frequency is shown in the expansion. The trace from the spectrum is plotted above the contour plot.

10.3.2.2 ^1H - ^{15}N HMBC and LR-HSQMBC

When the structure of cryptospirolepine was original reported,²⁹ the work done on the molecule preceded the first reported applications of ^1H - ^{15}N HMBC to small molecules by several months.^{46,47} The first ^1H - ^{15}N HMBC spectrum to be directly related to cryptospirolepine was during the characterization of the bis-indoloquinoline degradant of cryptospirolepine shown in Scheme 10.1 that was reported in 2002,³¹ in which three of the four nitrogens were observed in the spectrum (Table 10.1). Since the ^1H - ^{15}N LR-HSQMBC experiment reported in 2014⁴⁸ has considerable utility for observing small (*e.g.* $^4J_{\text{NH}}$) long-range correlations, a 2 Hz-optimized spectrum was recorded and is shown in Figure 10.29.

Although all four nitrogens of cryptospirolepine (**1**) were observed in the 2 Hz-optimized LR-HSQMBC spectrum, they were not particularly diagnostic, nor can comparisons be readily drawn to the bis-indoloquinoline degradant **11**, which was characterized in 2002,³¹ since the conjugation in the top portion of that molecule is considerably different.

10.3.2.3 Application of 1,1- and 1,*n*-HD-ADEQUATE

Since higher-sensitivity experiments failed to provide the means of unequivocally establishing the structure of cryptospirolepine, we turned to the application of the newly developed 1,1- and 1,*n*-HD-ADEQUATE experiments, knowing that it would be difficult to perform them on the small ~ 700 μg sample that could be dissolved in a 1.7 mm MicroCryoProbe™ NMR tube. A segment of the 60 Hz-optimized 1,1-HD-ADEQUATE spectrum of cryptospirolepine is shown in Figure 10.30B. Three correlations were observed from the vinyl carbon. $^1J_{\text{CC}}$ correlations were observed to the carbonyl resonating at 188.6 ppm and to the quaternary carbon, C13a, resonating at 137.3 ppm. These correlations irrefutably position the vinyl carbon between the carbonyl and the 137.3 ppm quaternary carbon. Hence, the [7.5.5] central core of the

Table 10.1 DFT-calculated ^1H - ^{15}N coupling constants for the various ^1H - ^{15}N long-range correlations of cryptospirolepine.³ It is interesting to note that all of the calculated values are <2 Hz, which suggests that observing some of these using, for example, a 3 or 4 Hz-optimized HMBC spectrum would be difficult at best.

	N5'	N8	N10'	N3
H5'-N-Me	0.61*			0.16*
H8'-N-Me		0.46*		
H13		0.30		0.34
H4				0.70
H1'	0.57			
H8'			0.47	
H9'			1.35	
H4'	1.59			
H9		1.60		
H10'-NH			73.23	

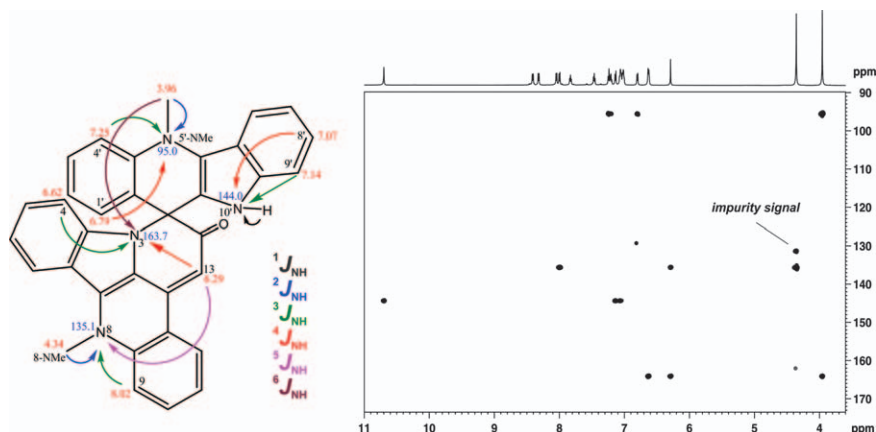


Figure 10.29 ^1H – ^{15}N LR-HSQC spectrum of cryptospirolepine (**1**) optimized for 2 Hz. The rigid skeletal framework of the alkaloid, coupled with multiple long-range correlation pathways, probably accounts for the observation of the $^5J_{\text{NH}}$ and $^6J_{\text{NH}}$ correlations observed in the spectrum.

pentacyclic portion of the structure must be revised to a [6.6.5] system. There was also a third correlation observed that can be attributed to the $^2J_{\text{CC}}$ correlation across the carbonyl to the C1/C11' spiro carbon resonating at 65.9 ppm (folded into this presentation). A similar $^2J_{\text{CC}}$ correlation was noted in the 1,1-HD-ADEQUATE spectrum of **6**.¹¹ A segment of the 5 Hz-optimized 1,*n*-HD-ADEQUATE spectrum of cryptospirolepine is shown in Figure 10.30D. The correlations are predominantly *via* $^3J_{\text{CC}}$ with the exception of the two $^1J_{\text{CC}}$ correlations to C2 and C13a and the $^2J_{\text{CC}}$ correlation to C1/C11'.

10.3.2.4 Using Advanced NMR Data as CASE Program Input

Another aspect of acquiring sophisticated NMR data in order to characterize challenging structures that should not be ignored is the impact of those data on the performance of CASE programs. A study in 2014⁶ showed that the availability of 2 Hz LR-HSQC data as CASE program input could have a profound impact on both the numbers of structures generated and the generation time. It was in that light that we elected to evaluate various subsets of the cryptospirolepine NMR data available as CASE program input for the ACD Structure Elucidator™ program.

We began by restricting the input data (see Table 10.2) to those data that would have been available in 1993 when the structure of cryptospirolepine was first reported.²⁹ It is interesting to note that with the 1993 data, which consisted of HSQC (HMQC data were available in 1993) and 8 Hz-optimized HMBC, the program ran for 3 full weeks (420 hours) without ever generating the correct structure (neither revised nor that reported in 1993) from among the >10 400 structures generated in the output file. The computation was terminated at that point.

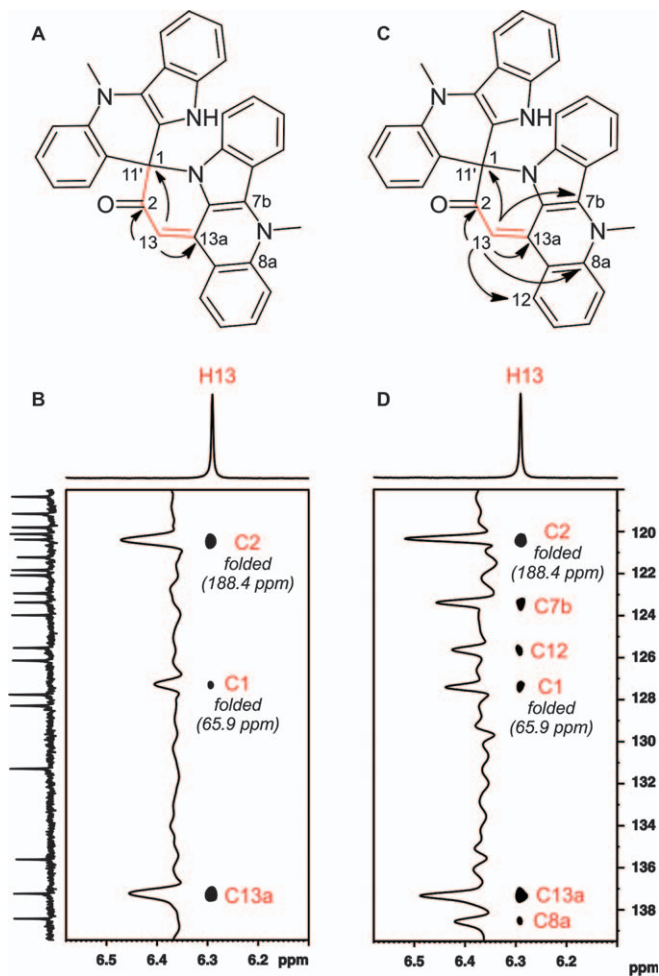


Figure 10.30 (A) Revised structure of cryptospirolepine (1) showing the two $^1J_{CC}$ correlations to the flanking C2 and C13a carbons and the $^2J_{CC}$ correlation to the C1/C11' carbon. These data allow the revision of the structure to that shown above from that contained in the original 1993 report (see also Scheme 10.1). (B) Segment of the 60 Hz-optimized 1,1-HD-ADEQUATE spectrum of cryptospirolepine showing the correlations summarized on the structure in (A). (C) Revised structure of cryptospirolepine showing the $^nJ_{CC}$ correlations observed in the 5 Hz-optimized 1,1-HD-ADEQUATE spectrum segment shown in (D). The full 1,1-HD-ADEQUATE spectrum is shown in Figure 10.31. (D) Segment of the 5 Hz-optimized 1,1-HD-ADEQUATE spectrum. The long-range correlations observed were *via* $^3J_{CC}$ with the exception of the $^2J_{CC}$ correlation across the carbonyl, which is analogous to that observed for 6.¹¹ The full 5 Hz-optimized 1,1-HD-ADEQUATE spectrum is shown in Figure 10.32. (Reproduced from: J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff, Jr., T. Parella, R. T. Williamson and G. E. Martin *Angew. Chem., Int. Ed.*, 2015, 54, 10160–10164³ with permission from John Wiley & Sons. Copyright © 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

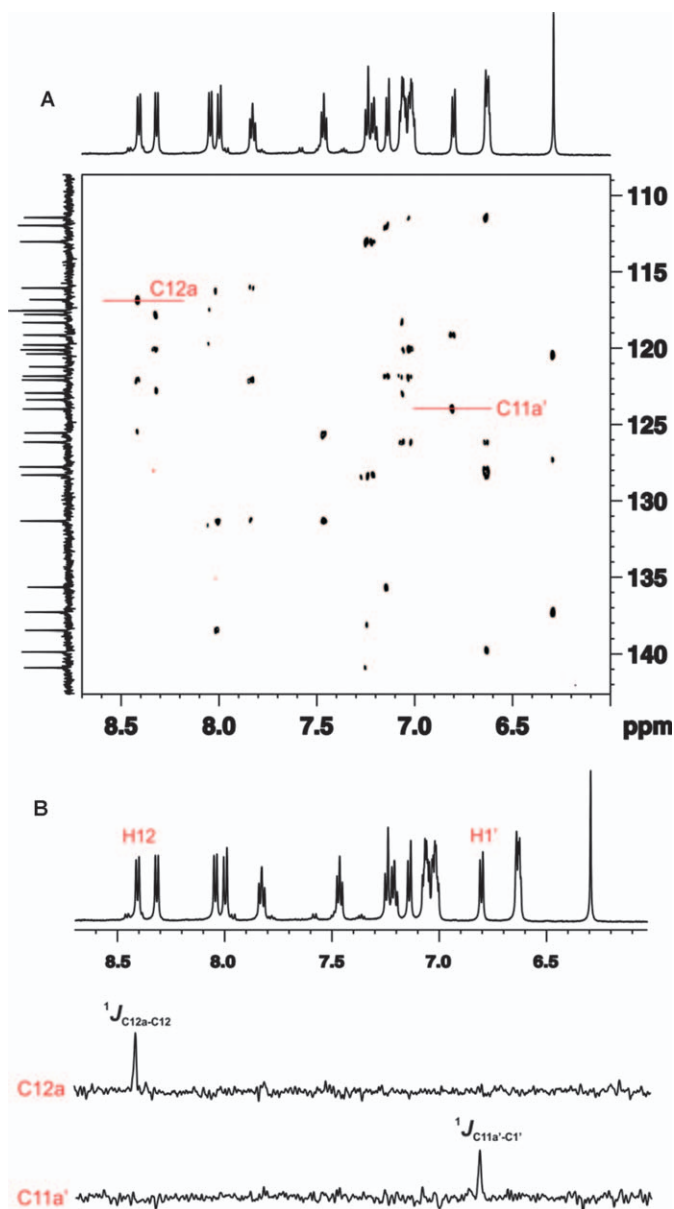


Figure 10.31 (A) 60 Hz 1,1-HD-ADEQUATE spectrum acquired for cryptospirolepine (1). The spiro and carbonyl carbons were intentionally folded in order to improve F_1 digital resolution in the aromatic region of the spectrum. (B) Proton reference spectrum and selected slices showing the HD afforded by the BIRD-based decoupling. (Reproduced from: J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff, Jr., T. Parella, R. T. Williamson and G. E. Martin, *Angew. Chem., Int. Ed.*, 2015, 54, 10160–10164³ with permission from John Wiley & Sons. Copyright © 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

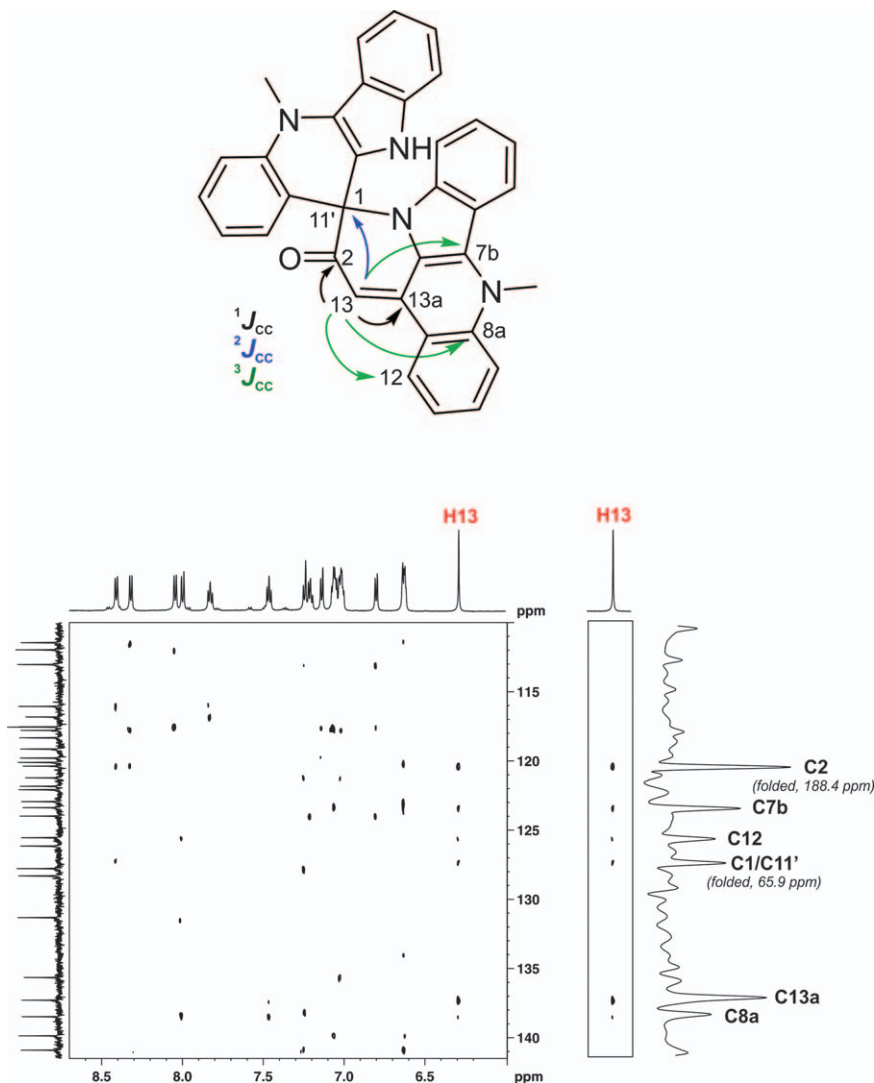


Figure 10.32 Plot of the full 5 Hz-optimized $1,n$ -HD-ADEQUATE spectrum of cryptospirolepine (**1**). The key correlations from the carbonyl carbon, C13, are shown on the structure. The vertical segment plotted to the right shows the correlations observed for the C13 vinyl carbon.

(Reproduced from: J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff, Jr., T. Parella, R. T. Williamson and G. E. Martin, *Angew. Chem., Int. Ed.*, 2015, **54**, 10160–10164³ with permission from John Wiley & Sons. Copyright © 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.)

Table 10.2 Results obtained when various subsets of the NMR data available for cryptospirolepine were utilized as input files for the Structure Elucidator™ CASE program.

1,1-ADEQUATE	1, <i>n</i> -ADEQUATE	¹ H- ¹³ C HMBC		IDR HSQC-TOCSY	¹ H- ¹³ C LR-HSQMBC		¹ H- ¹⁵ N LR-HSQMBC	Generation	
60 Hz	7 Hz	8 Hz	4 Hz	15 ms	2 Hz	4 Hz	2 Hz	Time (s)	Structures (n)
		+						> 420 hours	> 10 400
+		+		+				140	6816
+		+		+			+	40	522
+		+		+	+			142	3360
+		+		+	+		+	45	258
+		+	+	+				18	360
+		+	+	+			+	7	40
+		+	+	+	+		+	7	24
+	+	+	+	+	+	+	+	7	24

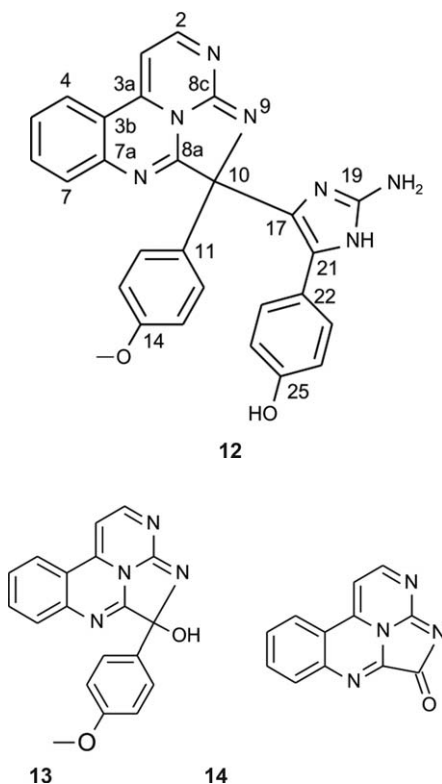
It will be noted from the next row of data, which added only the 1,1-HD-ADEQUATE and 15 ms IDR-HSQC-TOCSY data, that those data had a profound impact on the program's ability to utilize the data. The program ran for only 140 seconds, generating 6816 structures, but the revised structure (9) was in the output file! As successively more complete data packages were added to the input file, the structure generation time decreased still further and the number of structures generated plummeted. When the complete ensemble of the data were used as input, the program ran for only 7 seconds, generating only 24 structures, with the correct (revised) structure of cryptospirolepine, 9, being first in the output file.

10.3.2.5 Computational Considerations of Cryptospirolepine

With the revised structure of cryptospirolepine in hand, there remained several unanswered questions. First was the enigmatic absence of a $^2J_{\text{CH}}$ correlation from the H13 vinyl proton to the C2 carbonyl in the 2 Hz-optimized LR-HSQMBC experiment. From DFT calculations on the revised structure, for whatever reason, the calculated $^2J_{\text{CH}}$ coupling constant was a miniscule 0.03 Hz! Thus, it is probably more surprising that there was any hint of a potential correlation from H13 to C2 in the 2 Hz-optimized LR-HSQMBC than it would have been for no correlation to be observed. There also remained the question of the stereochemistry at the spiro center. CD spectra of cryptospirolepine exhibited no distinct features above the noise, hence it was concluded that the sample was likely racemic. Attempts were also made using chiral HPLC methods to resolve the enantiomers, but those attempts were unsuccessful. There remains the possibility that the alkaloid as produced by the plant is chiral, but underwent racemization during the isolation procedure.

10.3.3 Eudistidine-C

As a final example, we will consider a unique marine alkaloid, eudistidine-C (12). A *Eudisoma* sp. produces several polycyclic marine alkaloids that affect the transcription factor hypoxia-inducing factor-1. Two of the homologs, eudistidine-A (13) and -B (14), were characterized using a conventional suite of 2D NMR experiments, including ^1H - ^{15}N HMBC data.⁴⁹ A third analog was more complex and could not be fully characterized using the standard suite of 2D NMR experiments. Utilizing the data available from the COSY, HSQC, and ^1H - ^{13}C and ^1H - ^{15}N HMBC spectra as well as the empirical formula derived from accurate mass data as input for the Structure Elucidator™ CASE program led to the proposed structure as 12.⁵⁰ While the available NMR data were consistent with the structure proposed by Structure Elucidator™, it was critical to be able to confirm the structural proposal with very-long-range heteronuclear shift correlation data. Hence, a 2 Hz-optimized LR-HSQMBC²¹ spectrum of the alkaloid was acquired at 600 MHz using a 1.7 mm Micro-CryoProbe™, which gave the spectrum shown in Figure 10.33.



Key long-range correlations observed in the 2 Hz-optimized LR-HSQMBC spectrum shown in Figure 10.33 are summarized on 15. The values shown on the correlation arrows are the DFT-calculated $^nJ_{\text{CH}}$ coupling constants.

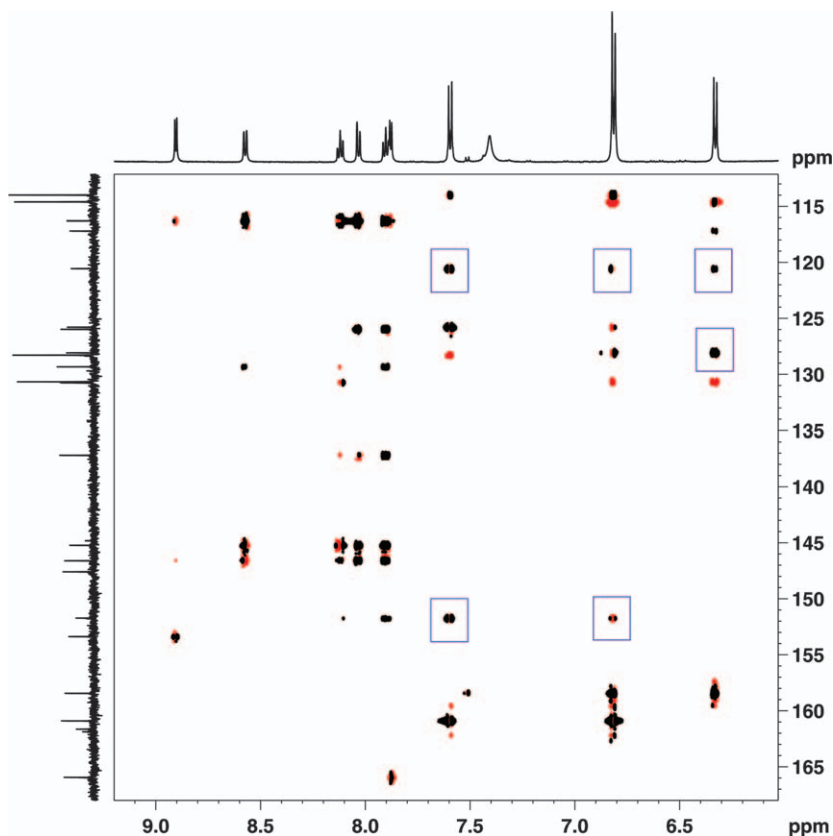
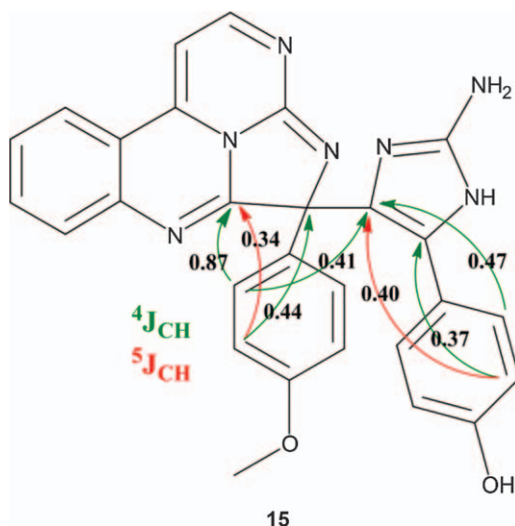


Figure 10.33 LR-HSQMBC spectrum of eudistidine-C (12) optimized for 2 Hz. The data were acquired with 768 increments in the second frequency domain. Key long-range correlations are boxed and correspond to the correlations shown by 15.

Based on the small size of the critical $^4J_{\text{CH}}$ and $^5J_{\text{CH}}$ coupling constants shown on 13, it is not particularly surprising that these correlations were not observed in a 4 Hz-optimized HMBC experiment. With these correlations in hand that support the results of the Structure Elucidator™ calculation, the structure was considered to be firmly established.

10.4 Conclusions

This chapter has presented a number of innovative NMR experiments, including some very recently reported experiments, that can be focused on the structure elucidation of alkaloids or, for that matter, any other class of natural product structures. There will undoubtedly be an induction period that can vary wildly from one laboratory to the next in terms of the rate of adoption of these experiments. As noted in Volume 1, Chapter 1, even 20 years after the introduction of the HSQC experiment, HMQC continues to be in wide usage. Other laboratories seem reticent to utilize multiplicity editing, probably under the mistaken impression that editing is too costly in terms of sensitivity. In contrast, when the time that it takes to acquire, for example, a DEPT experiment is considered, any time penalty that accrues by using multiplicity-edited HSQC rather than an unedited version of the experiment will be inconsequential.

With regard to how long it will take for experiments like LR-HSQMBC²¹ to be adopted, only time will tell. A SciFinder™ search of the literature returns relatively few citations of HSQC-TOCSY applications in natural product structure elucidation. It is likely, however, that the heteronuclear experiments such as LR-HSQMBC, or HSQMBC-COSY/-TOCSY,⁵¹ will be adopted more quickly than any of the homonuclear carbon-carbon-based experiments such as 1,1- or 1,*n*-HD-ADEQUATE. Nevertheless, for adventuresome investigators, there is a growing arsenal of powerful new experimental methods that can be applied to challenging structure elucidation problems. The future is bright!

The three alkaloid examples selected as illustrations in the preceding sections were chosen to illustrate some of the highlights associated with the new NMR experimental methods that are available. As was shown with staurosporine (8),²⁶ the definition of what constitutes a “proton-deficient” molecule is changing. The Crews’ rule^{27,28} may well have to be revised as some of the now available NMR methods come into more routine usage.

Cryptospirolepine (9) represents another challenging alkaloid structure characterization that took more than two decades to resolve. Characterization of two degradants of the original sample nearly 10 years after the original report of the structure suggested that the original structure was likely to be incorrect.³¹ Although performed after the structure was successfully revised, Structure Elucidator™ calculations showed that the current version of the 1993 data was insufficient for the CASE program to successfully define the structure of this complex alkaloid. It was necessary to use

some of the most advanced NMR techniques available to successfully revise the structure of cryptospirolepine, as detailed in the recent communication by the authors.³

The example provided by the structural characterization of eudistidine-C (12) illustrates another approach to contemporary structure elucidation. The available data from more commonly utilized 2D NMR experiments can be used as input for one of the several CASE programs now available, which can in turn be used to calculate structures for consideration by investigators. Once prospective structures for a molecule are in hand, specific advanced NMR experiments can be selected in order to observe key correlations, either via $^nJ_{CH}$, $^1J_{CC}$, or $^nJ_{CC}$ coupling pathways—the former in this example and the latter in the cryptospirolepine example.

References

1. N. T. Nyberg, J. Ø. Duus and O. W. Sørensen, *J. Am. Chem. Soc.*, 2005, **127**, 6154–6155.
2. N. T. Nyberg, J. Ø. Duus and O. W. Sørensen, *Magn. Reson. Chem.*, 2005, **43**, 971–974.
3. J. Saurí, W. Bermel, A. V. Buevich, E. C. Sherer, L. A. Joyce, M. H. M. Sharaf, P. L. Schiff, Jr., T. Parella, R. T. Williamson and G. E. Martin, *Angew. Chem., Int. Ed.*, 2015, **54**, 10160–10164.
4. S. F. Cheatham, M. Kline, R. R. Sasaki, K. A. Blinov, M. E. Elyashberg and S. G. Molodtsov, *Magn. Reson. Chem.*, 2010, **48**, 571–574.
5. J. Hubert, S. Chollet, S. Purson, R. Reynaud, D. Harakat, A. Martinez, J.-M. Nuzillard and J.-H. Renault, *J. Nat. Prod.*, 2015, **78**, 1609–1617.
6. K. A. Blinov, A. V. Buevich, R. T. Williamson and G. E. Martin, *Org. Biomol. Chem.*, 2014, **12**, 9505–9509.
7. M. Köck, B. Reif, W. Fenical and C. Griesinger, *Tetrahedron Lett.*, 1996, **37**, 363–366.
8. B. Reif, M. Köck, R. Kerssebaum, H. Kang, W. Fenical and C. Griesinger, *J. Magn. Reson.*, 1996, **118A**, 282–285.
9. G. E. Martin, Using 1,1- and 1,n-ADEQUATE 2D NMR data in Structure Elucidation Protocols, in *Annual Reports in NMR Spectroscopy*, ed. G. A. Webb, Elsevier, London, 2011, vol. 74, pp. 215–291.
10. G. E. Martin, M. Reibarkh, A. V. Buevich, K. A. Blinov and R. T. Williamson, Development of 1,n-ADEQUATE and Modified Variants and Their Application to Structure Elucidation and Spectral Assignments Problems, *eMagRes*, 2014, **3**, 215–234, DOI: 10.1002/9780470034590.emrstm1370.
11. C. Lorenc, J. Saurí, A. Moser, A. V. Buevich, A. J. Williams, R. T. Williamson, G. E. Martin and M. W. Pecuh, *ChemistryOpen*, 2015, **4**, 577–580, DOI: 10.1002/open. 201500122.
12. A. V. Buevich, R. T. Williamson and G. E. Martin, *J. Nat. Prod.*, 2014, **77**, 1942–1946.

13. R. W. Adams, Pure Shift NMR Spectroscopy, *eMagRes*, 2014, **3**, 1–15.
14. K. Zangger, Pure shift NMR, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2015, **86–87**, 1–20.
15. L. Castañar and T. Parella, *Magn. Reson. Chem.*, 2015, **53**, 399–426.
16. Y. Liu, M. D. Green, R. Marques, T. Pereira, R. Helmy, R. T. Williamson, W. Bermel and G. E. Martin, *Tetrahedron Lett.*, 2014, **55**, 5450–5453.
17. S. Wimperis and R. Freeman, *J. Magn. Reson.*, 1984, **58**, 348–353.
18. K. E. Köver and P. Forgó, *J. Magn. Reson.*, 2004, **166**, 47–52.
19. C. M. Thiele and W. Bermel, *Magn. Reson. Chem.*, 2007, **45**, 889–894.
20. A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, 1986, **108**, 2093–2094.
21. R. T. Williamson, A. V. Buevich, G. E. Martin and T. Parella, *J. Org. Chem.*, 2014, **79**, 3387–3394.
22. W. W. Meyer and M. Köck, *J. Nat. Prod.*, 2007, **71**, 1524–1526.
23. G. E. Martin, W. Bermel, R. T. Williamson and P. G. Dormer, *Magn. Reson. Chem.*, 2012, **50**, 563–568.
24. M. Reibarkh, R. T. Williamson, G. E. Martin and W. Bermel, *J. Magn. Reson.*, 2013, **236**, 126–133.
25. D. Meksuriyen and G. A. Cordell, *J. Nat. Prod.*, 1988, **51**, 884–892.
26. M. M. Senior, R. T. Williamson and G. E. Martin, *J. Nat. Prod.*, 2013, **76**, 2088–2093.
27. P. Falifo and P. J. Crews, *J. Org. Chem.*, 2004, **69**, 9025–9029.
28. T. F. Molinski and B. I. Morinaka, *Tetrahedron*, 2012, **68**, 9307–9343, this citation represents the first actual mention of the “Crews Rule” by name in the published chemical literature.
29. A. N. Tackie, G. L. Boye, M. H. M. Sharaf, P. L. Schiff, Jr., T. D. Spitzer, R. L. Johnson, J. Dunn, D. Minick, R. C. Crouch and G. E. Martin, *J. Nat. Prod.*, 1993, **56**, 653–670.
30. M. H. M. Sharaf, P. L. Schiff, Jr., A. N. Tackie, C. H. Phoebe, A. O. Davis, C. W. Andrews, R. C. Crouch and G. E. Martin, *J. Heterocycl. Chem.*, 1995, **32**, 1631–1636.
31. G. E. Martin, C. E. Hadden, D. J. Russell, B. D. Kaluzny, J. E. Guido, W. K. Duholke, B. A. Stiemsma, T. J. Thamann, R. C. Crouch, K. Blinov, M. Elyashberg, E. R. Martirosian, S. G. Molodtsov, A. J. Williams and P. L. Schiff, Jr., *J. Heterocycl. Chem.*, 2002, **39**, 1241–1250.
32. G. E. Martin, Applications of NMR Microprobes and Methodologies for Spectral Assignments, in *Encyclopedia of Nuclear Magnetic Resonance*, ed. D. M. Grant and R. K. Harris, Wiley, New York, 2002, vol. 9, supplement, pp. 98–112.
33. G. E. Martin, Cryogenic NMR Probes, in *Encyclopedia of Nuclear Magnetic Resonance*, ed. D. M. Grant and R. K. Harris, Wiley, New York, 2002, vol. 9, supplement, pp. 33–35.
34. G. E. Martin, Small Volume High Sensitivity NMR Probes, in *Annual Reports in NMR Spectroscopy*, ed. G. A. Webb, Elsevier, London, 2005, vol. 56, pp. 1–99.
35. G. E. Martin, Application of Cryogenic NMR Probe Technology for the Identification of Low Level Impurities in Pharmaceuticals, in *Handbook*

- of *Modern Magnetic Resonance*, ed. G. A. Webb, Elsevier, London, 2006, vol. II, pp. 1187–1194.
36. B. D. Hilton and G. E. Martin, *J. Nat. Prod.*, 2010, **73**, 1465–1470.
37. T. Molinski, *Curr. Opin. Drug Discovery Dev.*, 2009, **12**, 197–206.
38. D. S. Salisay, B. I. Morinaka, C. K. Skepper and T. F. Molinski, *J. Am. Chem. Soc.*, 2009, **131**, 7552–7553.
39. D. S. Dalisay, E. W. Rogers, A. S. Edison and T. F. Molinski, *J. Nat. Prod.*, 2009, **72**, 732–738.
40. D. S. Dalisay and T. F. Molinski, *Org. Lett.*, 2009, **11**, 1967–1970.
41. T. F. Molinski, *Curr. Opin. Biotechnol.*, 2010, **21**, 819–826.
42. G. E. Martin, Small Sample Cryoprobe NMR Applications, in *Encyclopedia of NMR*, ed. R. K. Harris and R. A. Wasylishen, Wiley, New York, 2011, on-line, DOI: 1002/9780470034590.emrstm1300.
43. M. E. Elyashberg, A. J. Williams, and G. E. Martin, Computer-Assisted Structure Elucidation, in *Progress in NMR Spectroscopy*, ed. J. Feeney and L. Sutcliffe, Pergamon, London, 2008, vol. 53, pp. 1–104.
44. M. Elyashberg, A. Williams and K. Blinov, *Contemporary Computer-Assisted Approaches to Molecular Structure Elucidation*, RSC Publishing, London, 2012, vol. xxii, p. 482.
45. M. E. Elyashberg and A. J. Williams, *Computer-Based Structure Elucidation from Spectral Data*, Springer, Berlin, 2015, vol. xvi, p. 447.
46. G. E. Martin, R. C. Crouch, M. H. M. Sharaf, P. L. Schiff, Jr., Annual Meeting American Society of Pharmacognosy, San Diego, CA, July 18–22, 1993, Abstract P101.
47. J. Uzawa, H. Utumi, H. Koshino, T. Hinomoto, K. Anzai, 32nd NMR Conference, Tokyo, Japan, November 4–6, 1993, pp. 147–150.
48. R. T. Williamson, A. V. Buevich and G. E. Martin, *Tetrahedron Lett.*, 2014, **55**, 3365–3366.
49. S. T. S. Chan, P. R. Patel, T. R. Ransom, C. J. Henrich, T. C. McKee, A. K. L. Goey, K. M. Cook, W. D. Figg, J. B. McMahon, M. J. Schnermann and K. R. Gustafson, *J. Am. Chem. Soc.*, 2015, **137**, 5569–5575.
50. J. Saurí, S. T. S. Chan, A. V. Buevich, K. R. Gustafson, R. T. Williamson, and G. E. Martin, 2016, submitted.
51. J. Saurí, N. Marcó, R. T. Williamson, G. E. Martin and T. Parella, *J. Magn. Reson.*, 2015, **258**, 25–32.

CHAPTER 11

Nuclear Magnetic Resonance Case Studies in Marine Natural Products

ANGELA A. SALIM, ANDREW M. PIGGOTT AND
ROBERT J. CAPON*

Division of Chemistry and Structural Biology, Institute for Molecular
Bioscience, The University of Queensland, St Lucia, Queensland 4072,
Australia

*Email: r.capon@uq.edu.au

11.1 Introduction

From its origins as an esoteric sub-branch of natural products chemistry in the early 1970s, marine natural products chemistry has grown into a vibrant, multidisciplinary field of research, extending well beyond a core capability in organic chemistry to embrace disciplinary expertise across biodiversity (*e.g.* microbiology, zoology, botany, and ecology), biology (*e.g.* biochemistry, pharmacology, molecular, cell, and developmental biology, and genetics), and bioactives (*e.g.* analytical, synthetic, and medicinal chemistry).¹ Notwithstanding this expanded reach, at its most fundamental, marine natural products research is driven by the need to detect, isolate, characterize, and identify structurally diverse metabolites, rich in complex carbon skeletons, heterocycles, functionality, and stereochemistry. Knowledge of marine natural products informs our understanding of biosynthetic pathways and chemical reactivity, while access to these remarkable substances accelerates

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

our ability to probe new biological (disease) targets, inspiring the discovery and development of next-generation drugs and biomolecular tools. That we can successfully deliver on the marine natural products potential is due in no small part to the contribution made by nuclear magnetic resonance (NMR) spectroscopy. In this chapter, we seek to illustrate this process through six case studies, each focusing on a class of natural product unique to the marine environment, and 11 snapshots, highlighting a diversity of NMR spectroscopic signatures.

11.1.1 Case Studies

The isolation and identification of the trachycladindole alkaloids (Figure 11.1) and the franklinolide polyketides (Figure 11.6), as first-in-class exemplars of new bioactive (anticancer) natural products, was challenging on a number of fronts, not the least being their propensity for chemical transformation during handling, storage, and data acquisition (*e.g.* elimination, isomerization, solvolysis, and hydrolysis). The terpenyl-pyrrolizidine bistellettazines (Figure 11.10) and terpenyl-taurine phorbasins (Figure 11.14) were also first-in-class natural products, with NMR data rendered all the more complex by their quasi-symmetric dimeric structures. Structure elucidation of the bistellettazines and phorbasins required analyzing NMR data acquired in multiple solvents, combined with biosynthetic considerations that invoked combinatorial Diels–Alder-like intermolecular cycloadditions to deliver the bistellettazines and cascading intramolecular Michael additions to assemble dimeric phorbasins. Although less structurally complex,

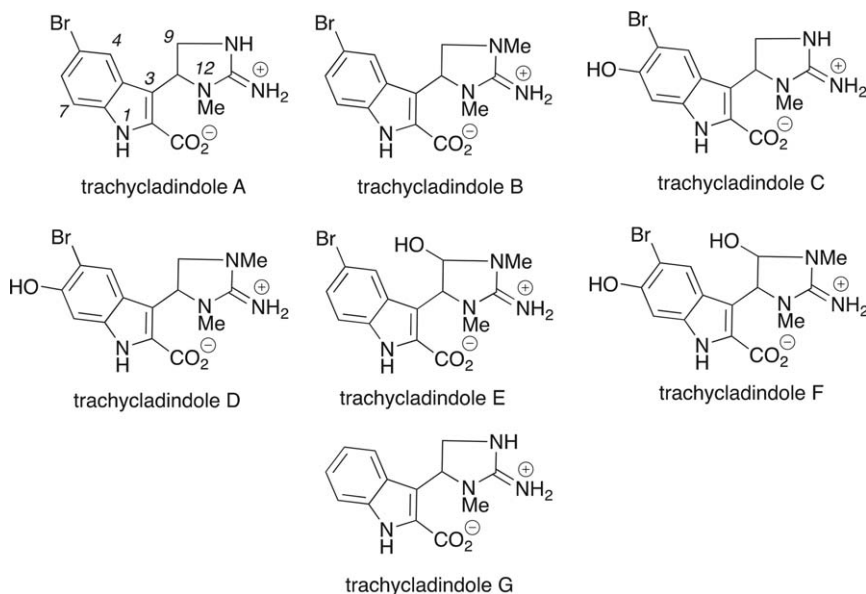


Figure 11.1 Trachycladindoles A–G.

the sesterterpene glycinyllactam ircinialactams (Figure 11.18) and the meroterpene fascioquinols (Figure 11.22) were nevertheless exciting discoveries, in that they exhibited biological properties that are relevant to chronic inflammatory pain and infectious disease. As ircinialactams are scarce minor metabolites recovered from highly complex marine extracts, their structure elucidation benefited from the use of high-performance liquid chromatography–diode array detection–solid-phase extraction–NMR (HPLC–DAD–SPE–NMR) technology. The isolation and identification of acid labile fascioquinols were challenging due to their inherently acidic (phenolic) nature, necessitating considerable care during handling, storage, and data acquisition. Collectively, these case studies provide a glimpse into the role that NMR spectroscopy plays in the structure elucidation of marine natural products.

11.1.2 Snapshots

Eleven examples, covering a wide cross-section of chemical, biosynthetic, biological, structural, and spectroscopic properties, have been selected in order to emphasize the diversity of ^1H NMR signatures presented by marine natural products. Suitable references have been provided for those wishing to delve further into the characterization and structure elucidation of these remarkable compounds and their co-metabolites.

11.2 Marine Natural Product Case Studies

11.2.1 Trachycladindoles

A southern Australian marine sponge, *Trachycladus laevispirulifer*, collected during commercial trawling operations (Orange roughly bycatch) in the Great Australian Bight yielded the cytotoxic agents trachycladindoles A–G as a selection of novel indole-2-carboxylic acids bearing a 2-amino-4,5-dihydroimidazole moiety (Figure 11.1). The trachycladindoles displayed potent and selective cytotoxicity against a panel of human cancer cell lines and their structures were assigned on the basis of detailed spectroscopic analysis.²

As small-molecular-weight, highly substituted alkaloids, isolation and structure elucidation of the trachycladindoles presented a number of significant challenges. In addition to the difficulties of purifying and characterizing a complex mixture of structurally related alkaloids from a rare, irreplaceable deep-sea sponge extract, the pure trachycladindoles exhibited significant chemical instability under non-native conditions. Prolonged (>3 month) storage, even in the dark at -30°C , resulted in the 9-hydroxyl-10-N-methyl trachycladindoles E and F undergoing facile loss of HCHO, leading to trachycladindoles A and C. Furthermore, interpretation of already “information-poor” NMR data was rendered all the more difficult by the presence of a unique internal salt bridge between the 2-carboxylate and the aminoimidazole moieties, as visualized in the energy-minimized molecular model

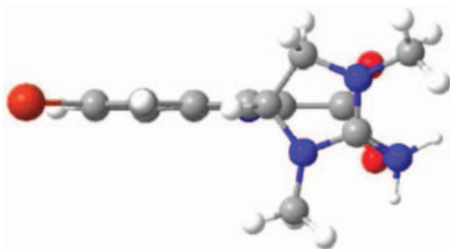


Figure 11.2 Energy-minimized (MM2) model of trachycladindole B.

shown in Figure 11.2. This internal salt bridge, a common feature of all trachycladindoles, has restricted rotation about C-3/C-8, leading to diagnostic deshielding and excessive broadening of ^1H NMR resonances for H-4 and H-8 (and H-9b). As evidenced in Figure 11.3, resonances for H-8 in some trachycladindoles were broadened to the degree of being rendered effectively invisible. Notwithstanding these challenges, the 2D NMR data (Figure 11.4) proved immensely valuable, and were critical to elucidating the structure of this class of unprecedented bioactive marine alkaloid.

A plausible biosynthesis leading to the trachycladindoles (Figure 11.5) has tryptophan undergoing intramolecular CO_2H transfer to C-2 *via* a cyclic acetal intermediate, with concomitant conversion of the amino group to a guanidinoenamine. Subsequent intramolecular stereo-controlled E2 addition/cyclization delivers the 2-amino-4,5-dihydroimidazole heterocycle, with methylation, oxidation, and bromination elaborating trachycladindoles A–D and G. An alternative stereo-controlled epoxidation of the guanidinoenamine, followed by intramolecular $\text{S}_{\text{N}}2$ -mediated addition/cyclization, together with methylation, oxidation, and bromination, elaborates trachycladindoles E and F.

11.2.2 Franklinolides

A marine sponge sample consisting of a massive *Geodia* sp. thinly encrusted with a *Halichondria* sp., collected during deep-water (–105 m) scientific trawling operations in the Great Australian Bight, yielded the cytotoxic franklinolides A–C as unprecedented polyketide phosphodiester (Figure 11.6). The structure–activity relationship (SAR) as assessed by co-metabolite studies demonstrated that the phosphodiester moiety unique to the franklinolides enhanced the cytotoxicity of the parent polyketide by >300-fold.³ Discovery of the franklinolides and their heightened anticancer properties led to speculation that glyceric acid phosphodiester may have broader application as conjugates capable of enhancing the potency and therapeutic utility of other cytotoxic agents (including current clinically approved cancer chemotherapeutics that are rendered less effective due to multi-drug resistance).

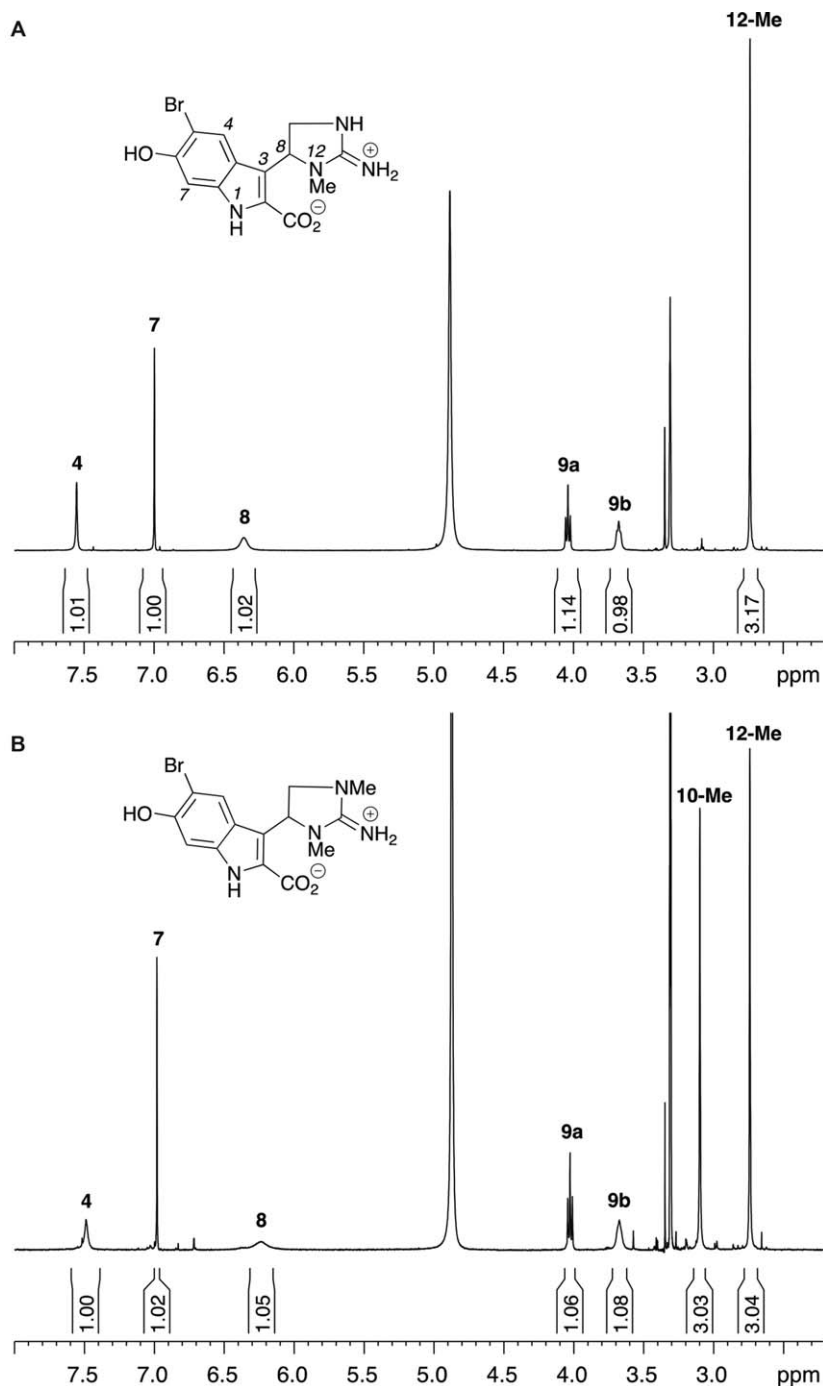


Figure 11.3 ^1H NMR (600 MHz, methanol- d_4) spectra of (A) trachycladindole C and (B) trachycladindole D.

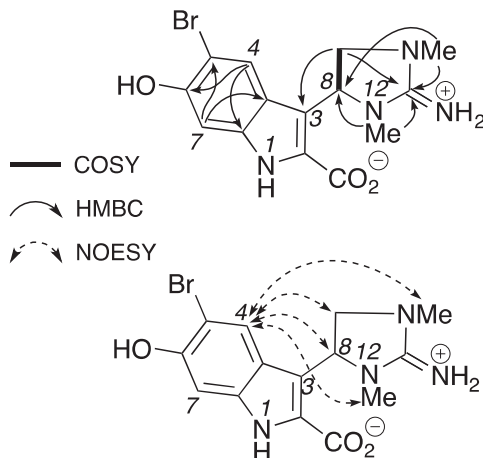


Figure 11.4 Selected NMR (600 MHz, methanol-*d*₄) correlations for trachycladindole D.

Handling, storage, and data acquisition on franklinolides were complicated by a propensity to undergo double-bond isomerization, solvolysis of the glyceric acid, and hydrolysis of the phosphodiester, all exacerbated by the need to employ a reverse-phase gradient HPLC fractionation strategy incorporating trace levels of trifluoroacetic acid (0.01%) as a mobile phase modifier. To avoid isomerization, hydrolysis, or solvolysis, it was essential that all HPLC fractions were buffered to pH 7 prior to concentrating *in vacuo*. Even with pure franklinolides in hand, identification of the phosphodiester moiety proved particularly challenging. ¹H NMR (CDCl₃) data for franklinolide A rapidly broadened, possibly due to acid-mediated isomerization, whereas prolonged storage in methanol-*d*₄ (Figure 11.7) led to *in situ* solvolysis, yielding glyceric acid deuteromethyl esters.

Notwithstanding these challenges, rapid acquisition of NMR data supported by chemical derivatization and degradation was pivotal to the structure elucidation of the franklinolides. For example, analysis of 1D and 2D NMR (methanol-*d*₄) data for franklinolide A established the extended polyketide and smaller glyceric acid substructures (Figures 11.7 and 11.8), while high-resolution electrospray ionization mass spectrometry operating in switchable positive and negative detection modes suggested a phosphodiester linkage. Initial efforts at confirming the phosphodiester by ³¹P NMR (CDCl₃) failed, most likely due to isomerization-related broadening. Fortunately, rapid acquisition of the ³¹P NMR (methanol-*d*₄) spectrum for franklinolide A revealed a diagnostic resonance (δ_P 0.3, dd, *J*_{P-H} 8.9, 8.3 Hz) that compared very well with that obtained from an authentic sample of commercially available (*S*)-2-phosphoglyceric acid disodium salt (δ_P 0.8, d, *J*_{P-H} 9.5 Hz). Supportive of this phosphodiester structure, fractionation of acid-hydrolyzed franklinolide A yielded the known natural product polyketide bitungolide A and (*S*)-3-*O*-methyl-2-phosphoglyceric acid (Figure 11.9),

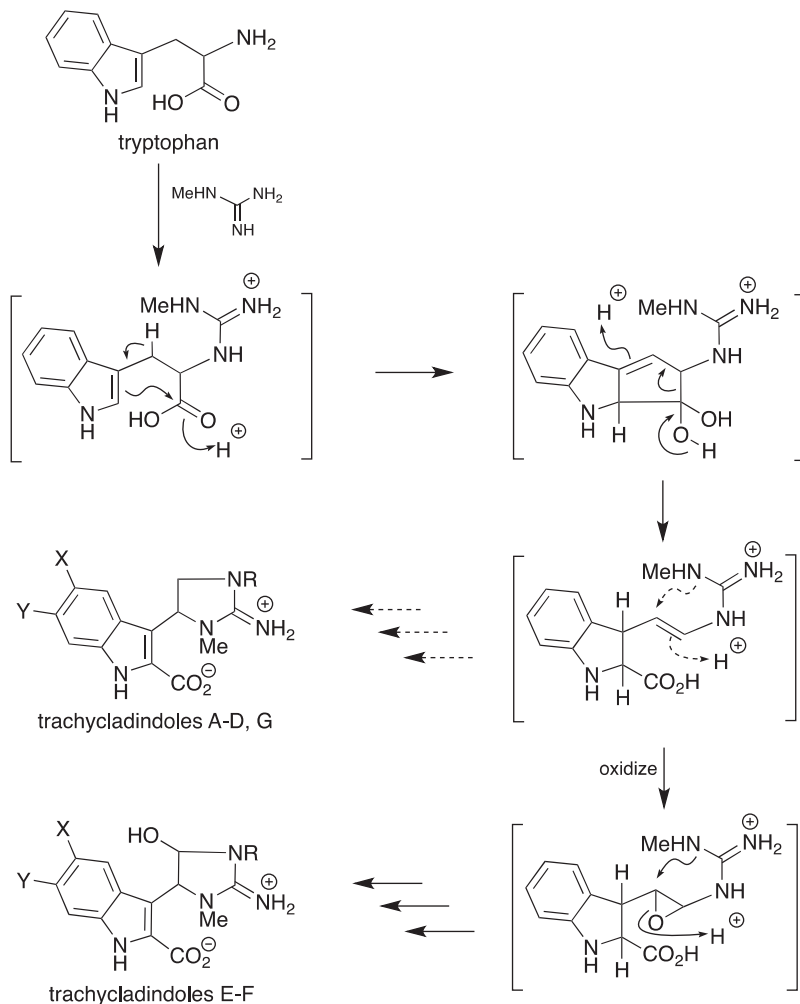


Figure 11.5 Plausible biosynthesis of trachycladindoles.

while careful re-examination and comparison of the ¹H NMR (methanol-*d*₄) data for franklinolide A and bitungolide A revealed broadening of H-9 in the former, consistent with a C-9 to C-2' phosphodiester linkage.

11.2.3 Bistelletazines

A southern Australian marine sponge, *Stelletta* sp., collected during deep-water (−90 m) scientific trawling operations in the Great Australian Bight, yielded the cytotoxic bistelletazines A–C and bistelletazole A (Figure 11.10) as rare examples of terpenyl-pyrrolizidine and terpenyl-imidazole conjugates.⁴ The quasi-symmetric dimeric and stereo complex nature of the

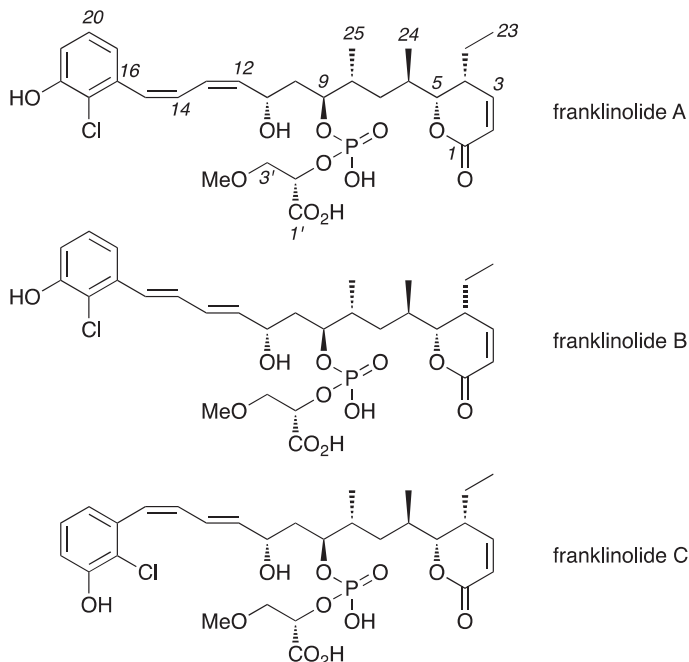


Figure 11.6 Franklinolides A–C.

bistellettazines, coupled with heavily broadened NMR resonances for “near-identical” pyrrolizidine moieties, presented a significant challenge to structure elucidation. Notwithstanding these challenges, structure elucidation was largely achieved through analysis of the 900 MHz NMR data, supported by biosynthetic considerations.

NMR (DMSO- d_6) spectra of bistellettazine A provided an exceptionally rich dataset, with COSY and HMBC correlations clearly defining the planar structure of the terpenyl fragment (Figures 11.11 and 11.12). Relative configurations about the cyclohexenyl residue (C-8 and C-13) and all double bonds were assigned on the basis of 3J -coupling and ROESY correlations, with the cyclohexenyl configuration being fully consistent with the proposed Diels–Alder-mediated biosynthesis (Figure 11.13). By contrast, ^1H NMR resonances attributed to both pyrrolizidine moieties were heavily coupled/broadened (Figures 11.11 and 11.12), requiring that structure assignment rely on hydrolysis of bistellettazine A, followed by detailed NMR analysis of the recovered intact pyrrolizidine amine residue. Although the structures of bistellettazines A–C and bistellettazole A span three different homologous carbon skeletons, a highly convergent biosynthetic pathway was proposed, involving combinatorial Diels–Alder-like dimerization, and amidation of hypothetical C_9 and C_{11} norsesquiterpene carboxylic acids (Figure 11.13).

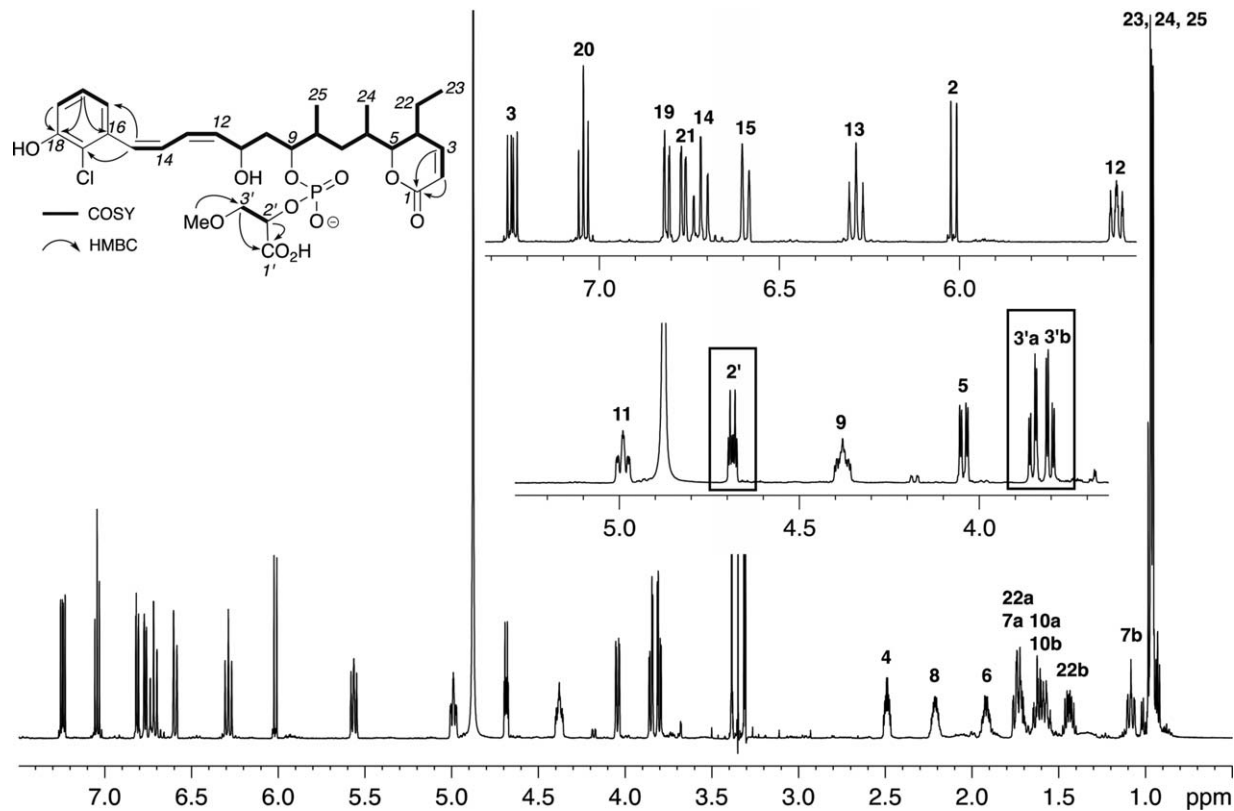


Figure 11.7 ^1H NMR (600 MHz, methanol- d_4) spectrum of franklinolide A and selected 2D NMR correlations (dashed boxes indicated the glyceric acid residue).

11.2.4 Phorbasins

A southern Australian marine sponge, *Phorbas* sp., collected during scientific trawling operations in the Great Australian Bight, yielded phorbasins B–K as a structurally and biosynthetically related library of cytotoxic terpenes and

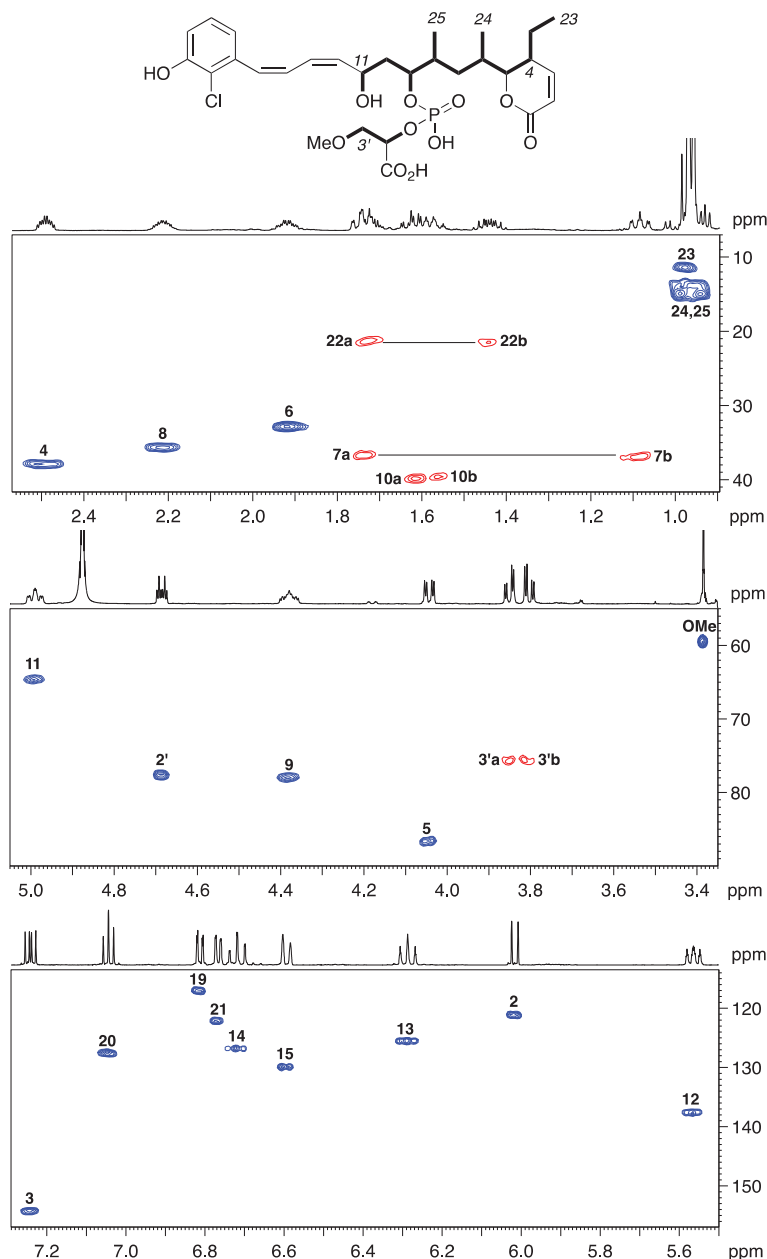


Figure 11.8 ^1H – ^{13}C HSQC NMR (600 MHz, methanol- d_4) spectrum of franklinolide A.

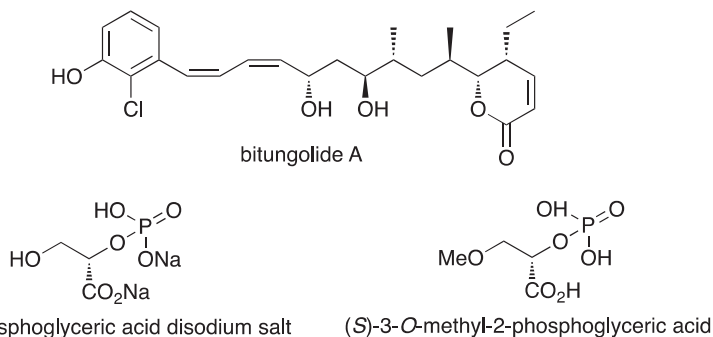


Figure 11.9 Bitungolide A and phosphoglyceric acids.

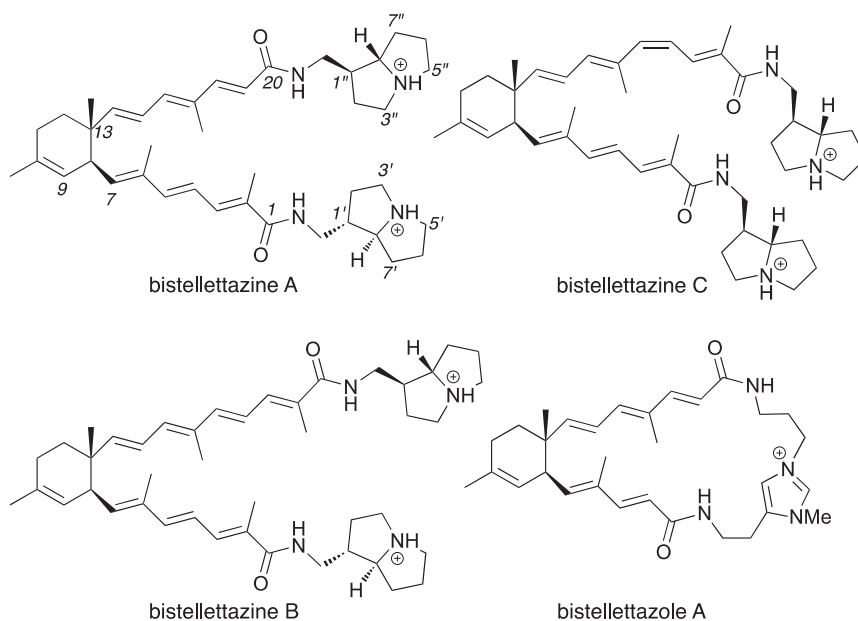


Figure 11.10 Bistellettazines A–C and bistellettazole A.

terpenyl-taurine conjugates (Figure 11.14). Incorporation of the taurinyl residue in phorbassins D–F was particularly noteworthy, as was the unprecedented heterocyclic system evident in the quasi-symmetric dimeric phorbassins E and F.^{5,6} Structure elucidation of the phorbassins was largely achieved by analysis of NMR data, with identification of the unprecedented heterocyclic system embedded in phorbassins E and F being particularly challenging, and requiring the added consideration of a plausible biosynthesis invoking cascading Michael additions.

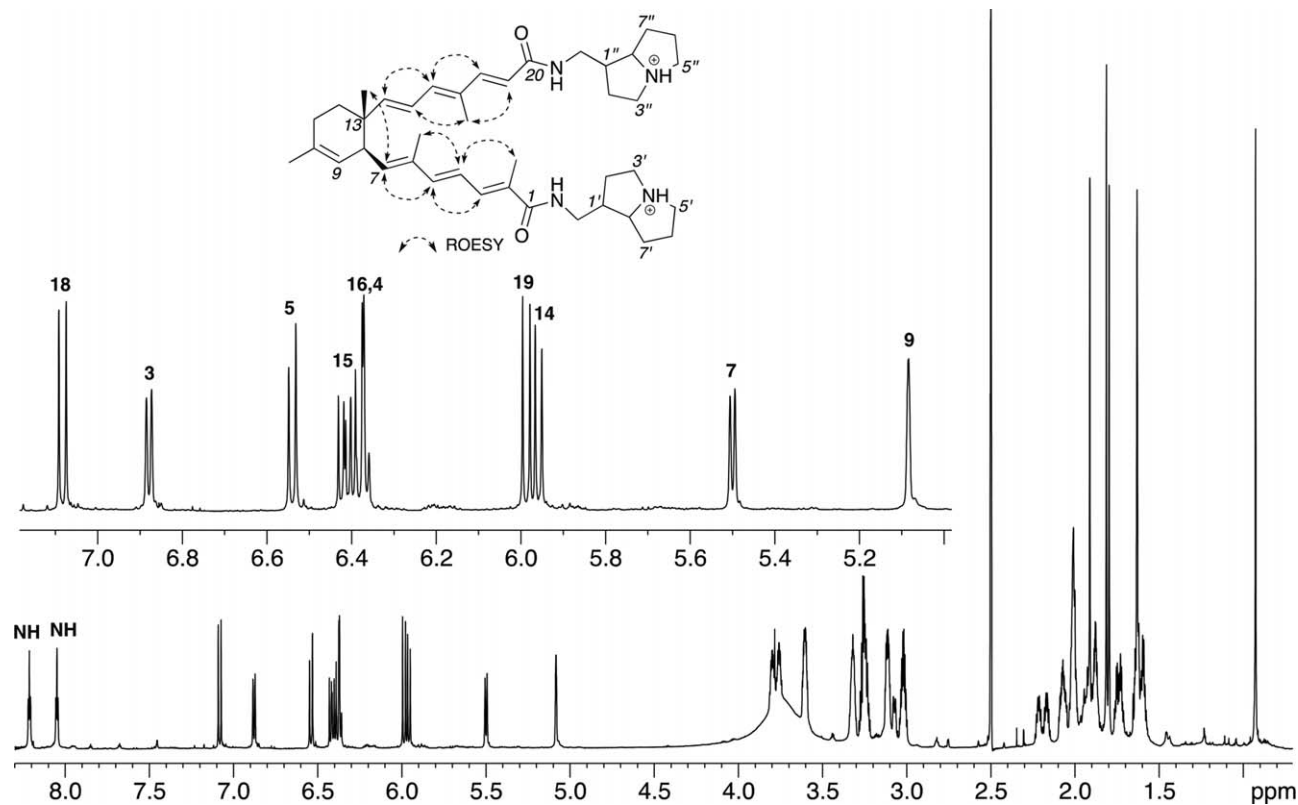


Figure 11.11 ^1H NMR (900 MHz, $\text{DMSO}-d_6$) spectrum of bistelletazine A and selected ROESY correlations.

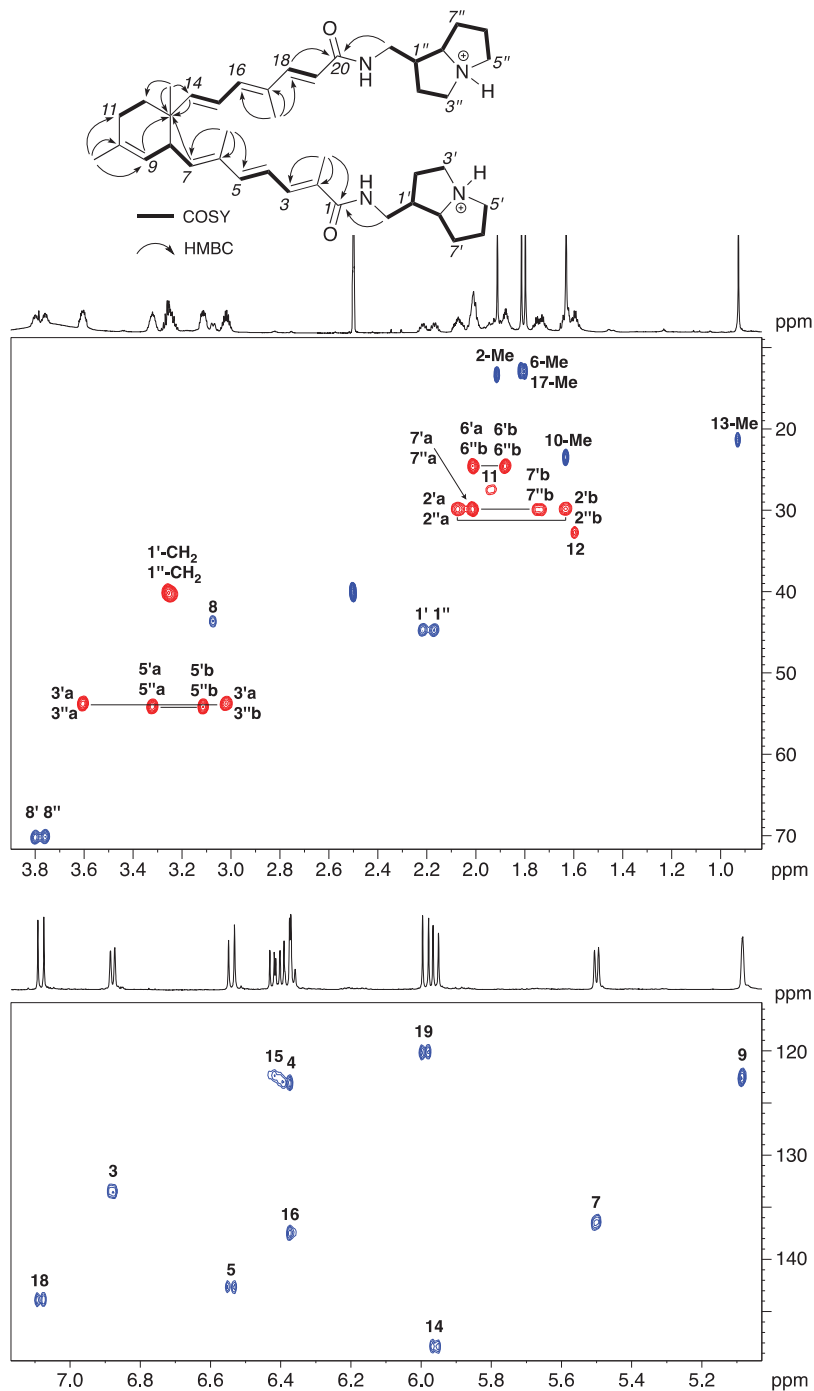
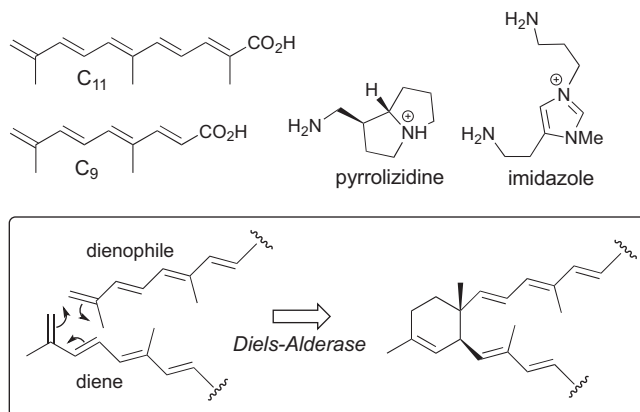


Figure 11.12 ^1H - ^{13}C HSQC NMR (900 MHz, $\text{DMSO}-d_6$) spectrum of bistellettazine A and selected 2D NMR (COSY and HMBC) correlations.



diene C₁₁ + dienophile C₉ + pyrrolizidine = bistellettazine A
 diene C₁₁ + dienophile C₁₁ + pyrrolizidine = bistellettazines B-C
 diene C₁₁ + dienophile C₉ + imidazole = bistellettazole A

Figure 11.13 Plausible Diels–Alder-mediated biosynthesis of bistellettazines A–C and bistellettazole A.

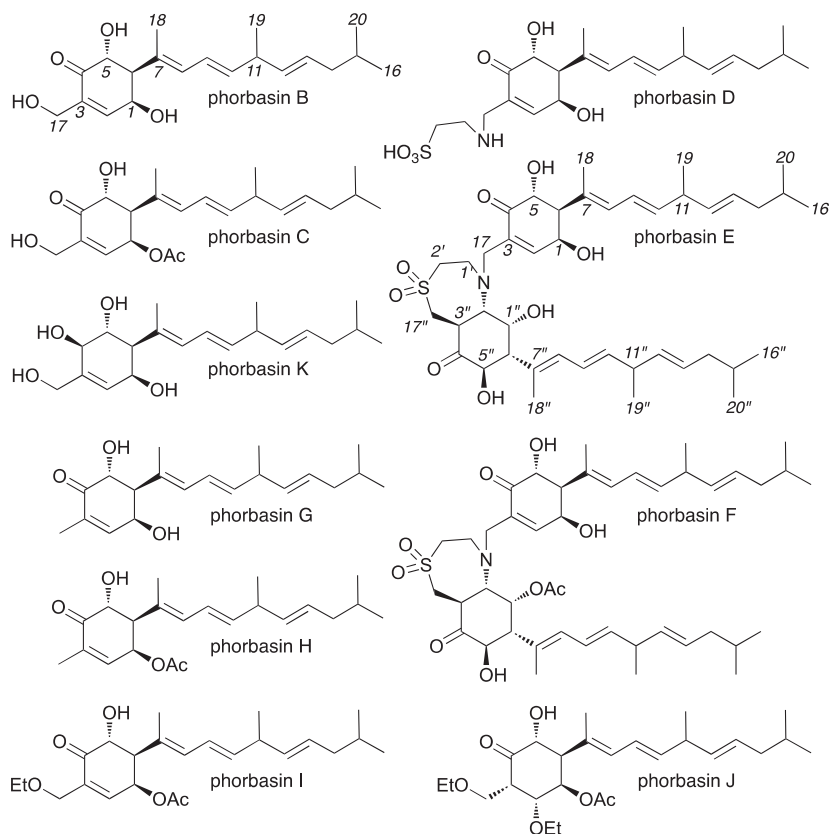


Figure 11.14 Phorbasins B–K.

The quasi-symmetric dimeric phorbasin E incorporates two near-identical diterpenyl residues fused to a taurinyl heterocycle. The NMR (methanol- d_4) data for phorbasin E (Figures 11.15 and 11.16) proved particularly complex, with structure elucidation necessitating the interpretation of data acquired in multiple solvents (*e.g.* methanol- d_4 and pyridine- d_5). For example, whereas the NMR (methanol- d_4) data for phorbasin C was effective at assigning diterpenyl structure fragments with relative configurations, one of which incorporated a taurinyl residue, it failed to provide the diagnostic correlations needed to assemble these fragments. Fortunately, the NMR (pyridine- d_5) data successfully resolved resonances for C-2'' (δ_C 69.3) and C-1 (δ_C 68.7), enabling assignment of a diagnostic HMBC correlation from H-17b to C-2'', which permitted assembly of the structure for phorbasin E (Figure 11.16).

A plausible biosynthesis for the unprecedented heterocyclic system in phorbasin E invoked cascading Michael additions (Figure 11.17), in which a hypotauninyl S_N2 addition adduct of phorbasin B forms a Michael addition adduct with a second equivalent of phorbasin B, thereby assembling the quasi-symmetric dimeric framework. Elimination of H_2O re-establishes the Michael acceptor enone moiety, which undergoes an intramolecular Michael addition from the adjacent sulfinic acid moiety to deliver the heterocycle and phorbasin E.

11.2.5 Ircinialactams

Glycine-gated chloride channel receptors (GlyRs) represent novel molecular targets for the development of drugs to treat a range of indications, including chronic inflammatory pain. Screening of a library of ~2500 southern Australian and Antarctic marine invertebrates and algae for potent and isoform-selective modulators of GlyRs prioritized extracts derived from several sponges of the family Irciniidae (two *Ircinia* spp. and one *Psammocinia* sp.). Bioassay-guided fractionation of these extracts returned an array of biosynthetically related sesterterpene tetronic acids, a structure class unique to marine sponges. Included among isolated natural products were several new minor co-metabolites, the ircinialactams (Figure 11.18), all of which incorporated a very rare glycynyl lactam moiety.⁷

The isolation and structure elucidation of ircinialactams were greatly assisted by the application of HPLC-DAD-SPE-NMR technology, which delivered high-quality data on very small (200 μ g) amounts of sample recovered *in situ* from extremely complex marine extracts. Double-bond geometries were assigned by ROESY correlations and diagnostic ^{13}C NMR chemical shifts of olefinic methyls, while COSY and HMBC correlations readily distinguished regioisomers (*e.g.* 8-hydroxyircinialactam A and 8-hydroxyircinialactam B; Figures 11.19 and 11.20). The SAR as assessed by co-metabolite investigations established that the very minor ircinialactams, featuring terminal tetronic acid and glycynyl lactam moieties, were the GlyR modulators. A plausible biosynthesis (Figure 11.21) has these rare glycynyl

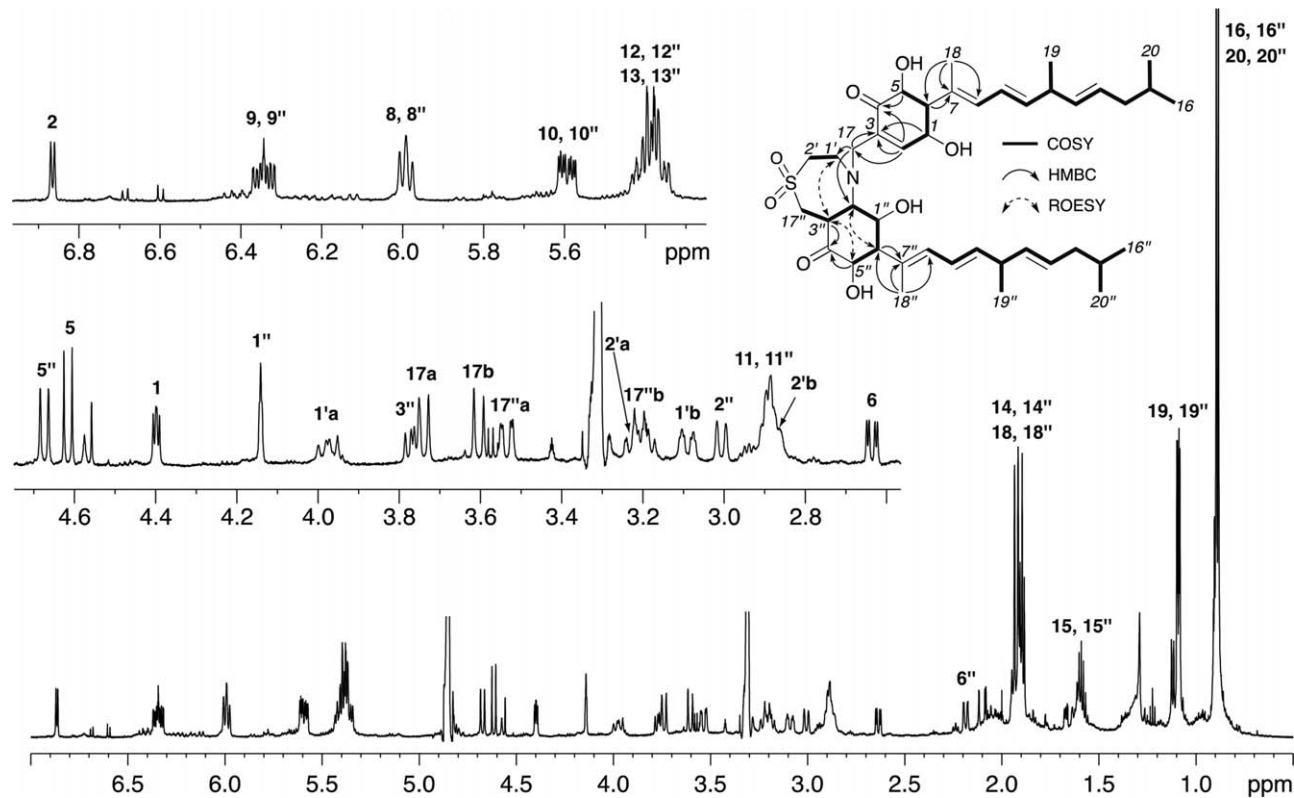


Figure 11.15 ^1H NMR (600 MHz, methanol- d_4) spectrum of phorbacin E and selected 2D NMR correlations.

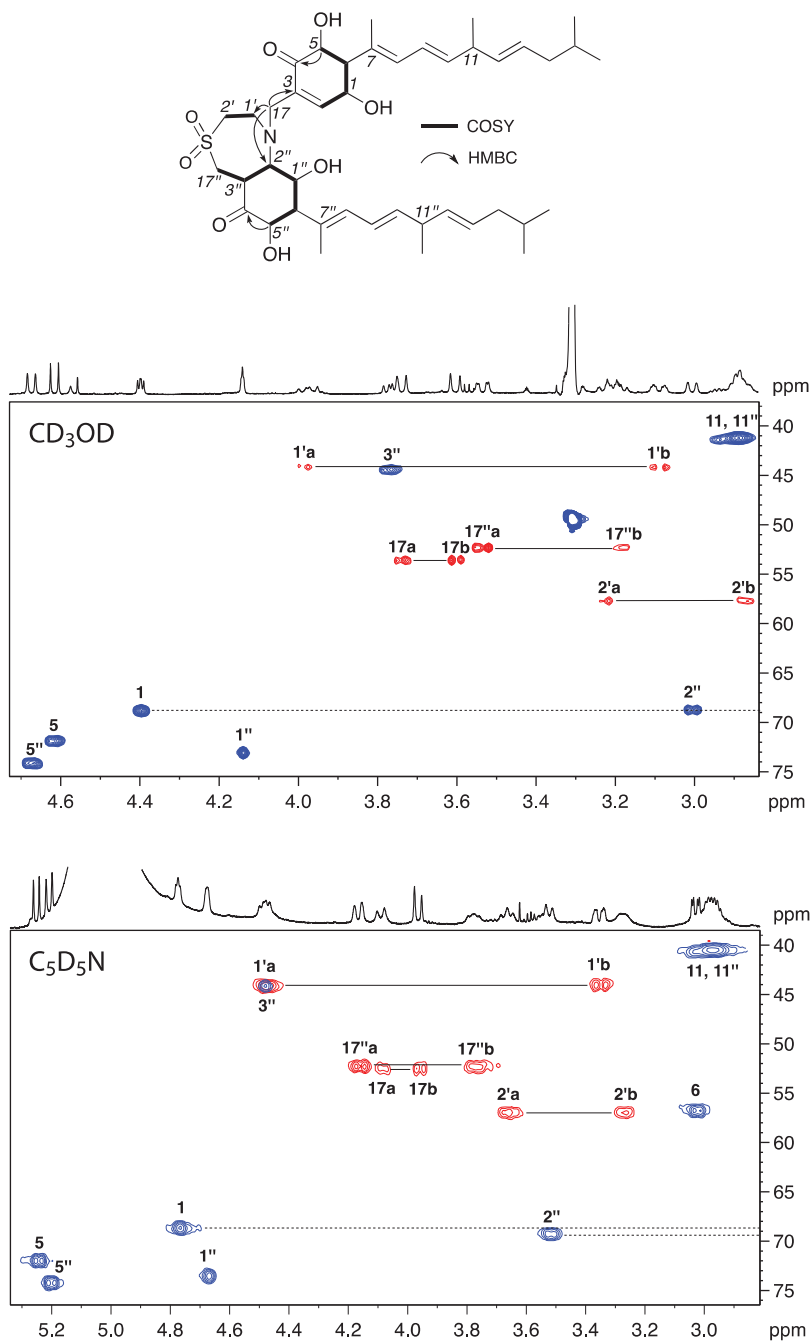


Figure 11.16 ^1H - ^{13}C HSQC NMR (600 MHz) spectra of phorbacin E in (top) methanol- d_4 and (bottom) pyridine- d_5 .

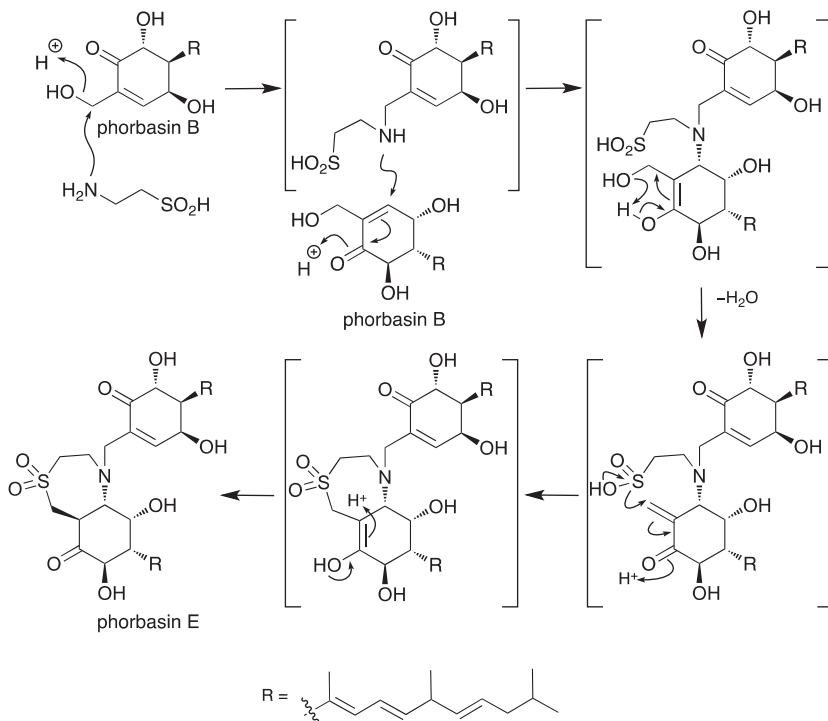


Figure 11.17 Plausible biosynthetic pathway leading to quasi-symmetric dimeric phorbasins.

lactams (e.g. 8-hydroxyircinialactams) derived from oxidative insertion of glycine into far more abundant furan co-metabolites (e.g. [12*E*,20*Z*,18*S*]-8-hydroxyvariabilin). Indeed, this biosynthetic hypothesis is currently the basis behind (unpublished) biomimetic syntheses of ircinialactams and related metabolites.

11.2.6 Fascioquinols

A southern Australian deep-water marine sponge, *Fasciospongia* sp., collected during scientific trawling operations in the Great Australian Bight, yielded a selection of biosynthetically related meroterpenes, including the new fascioquinols A–F (Figure 11.22).⁸ Fascioquinols A and B displayed promising antibacterial activity against the Gram-positive bacteria *Staphylococcus aureus* (IC₅₀ 0.9–2.5 μM) and *Bacillus subtilis* (IC₅₀ 0.3–7.0 μM),⁸ while fascioquinol E exhibited noteworthy inhibitory activity against capsule polysaccharide B, a metal-dependent protein tyrosine phosphatase responsible for the virulence of bacterial pathogens such as *Streptococcus pneumoniae*.⁹

Spectroscopic analysis and structure elucidation of the fascioquinols were largely routine. For example, the isomeric fascioquinols A and E returned very informative NMR data (Figure 11.23), which were readily interpreted,

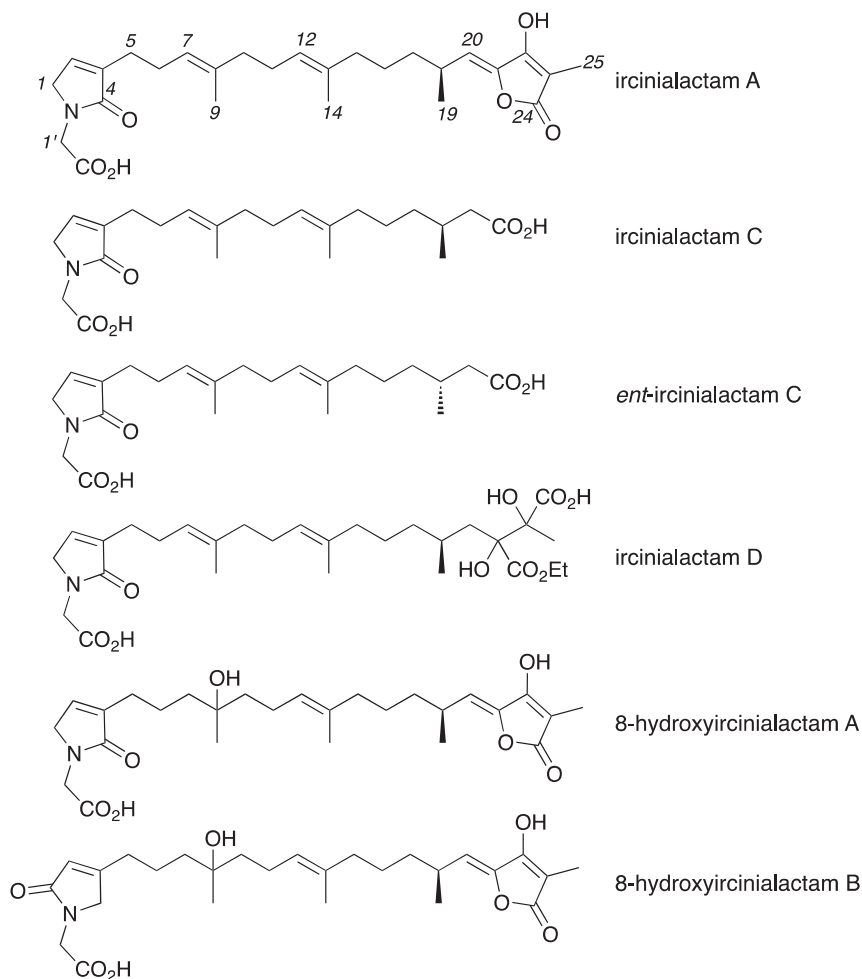


Figure 11.18 Ircinialactams.

assigning sulfate regiochemistry, double-bond geometry, and carbocycle relative configurations. Notwithstanding the ease of analysis of the NMR data, the isolation and structure elucidation of the fascioquinols did present challenges, not the least being that they were intrinsically acidic (phenolic), while at the same time acid labile. Incautious handling of the sulfates fascioquinols A and E resulted in acid-mediated (auto)conversion to the 1,4-hydroquinones fascioquinol B and geranylgeranyl-1,4-hydroquinone, respectively. Likewise, prolonged exposure of fascioquinols A and E to acidic conditions resulted in transformation to fascioquinols C, D, and F, as well as stronglylphorine-22. A plausible biosynthetic pathway linking fascioquinols and co-metabolites (Figure 11.24) includes a number of transformations that could easily be acid rather than enzyme mediated.

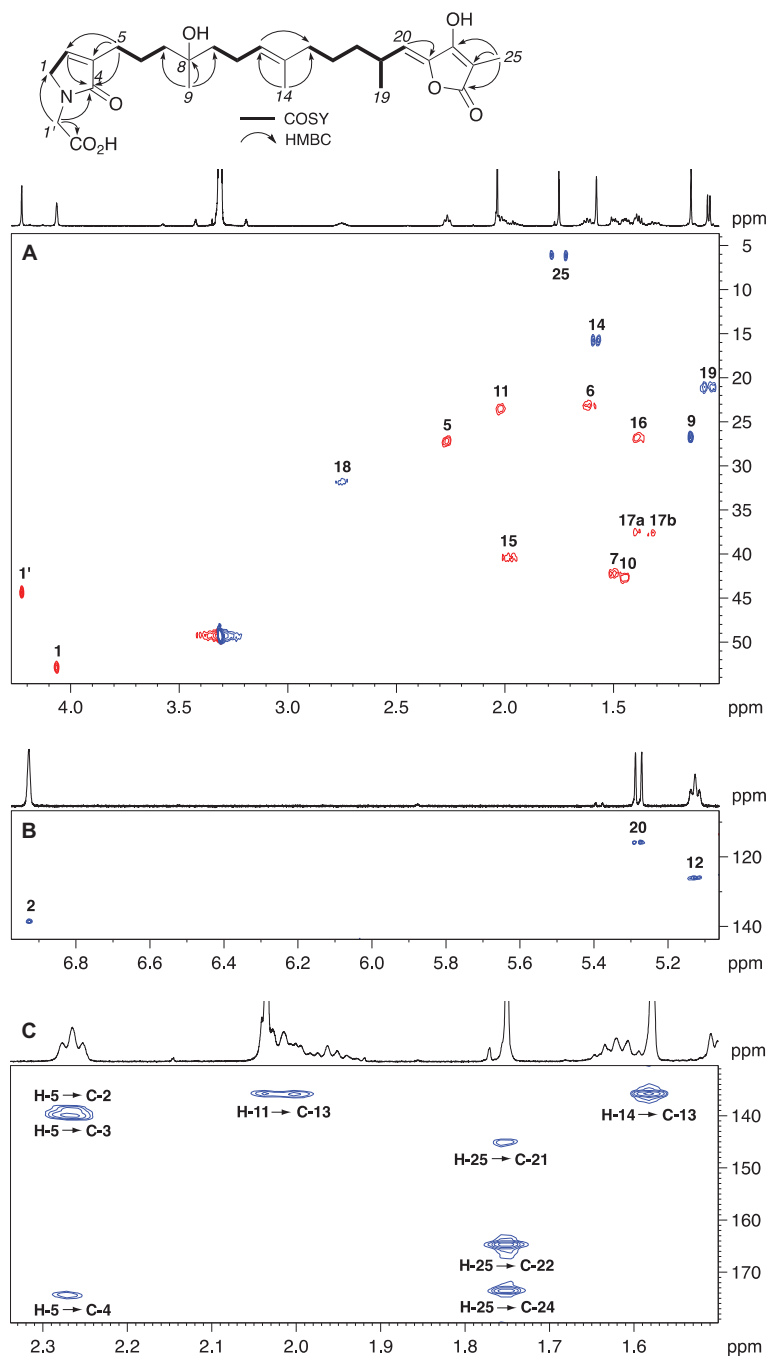


Figure 11.19 (A & B) ^1H - ^{13}C HSQC and (C) ^1H - ^{13}C HMBC NMR (600 MHz, methanol- d_4) spectra of 8-hydroxyircinialactam A and selected 2D correlations.

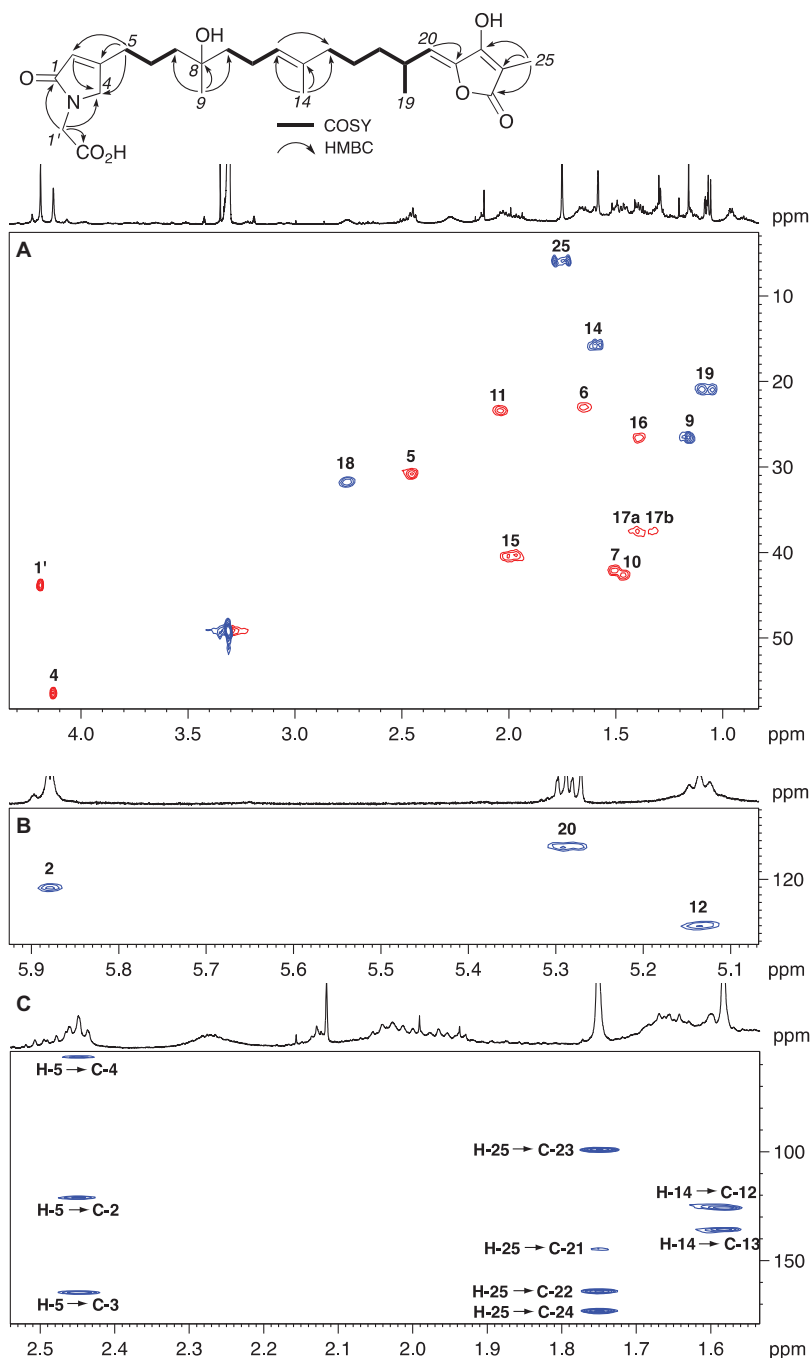


Figure 11.20 (A & B) ^1H - ^{13}C HSQC and (C) ^1H - ^{13}C HMBC NMR (600 MHz, methanol- d_4) spectra of 8-hydroxyircinalactam B and selected 2D correlations.

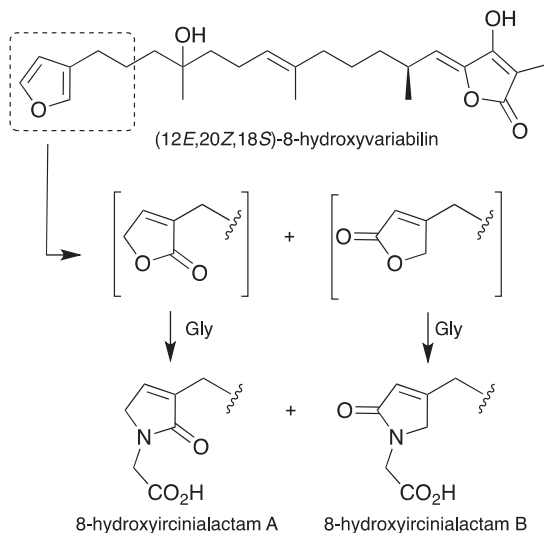


Figure 11.21 Plausible biosynthesis of glycinyllactams from furans.

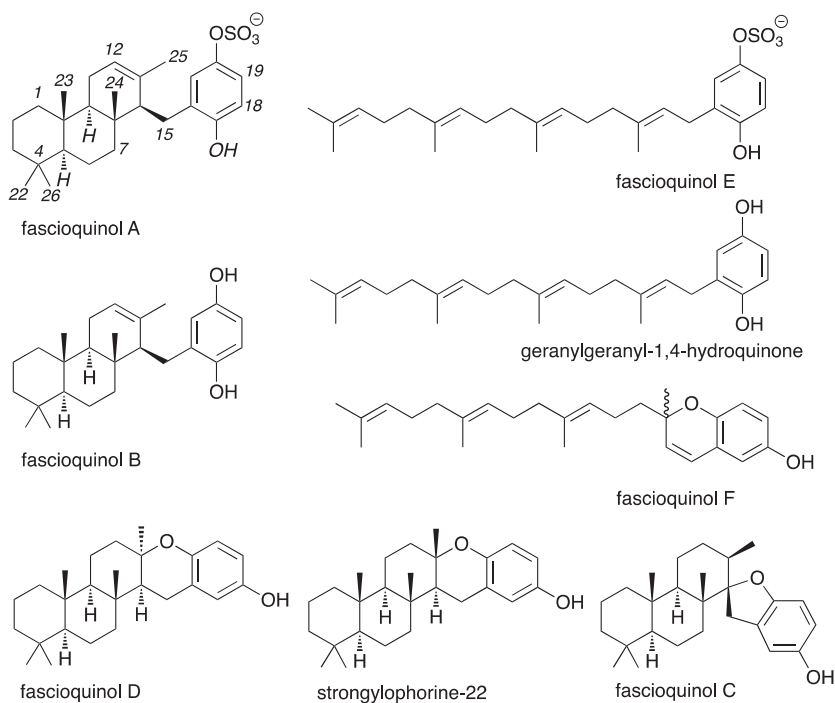


Figure 11.22 Fascioquinols and related meroterpenes.

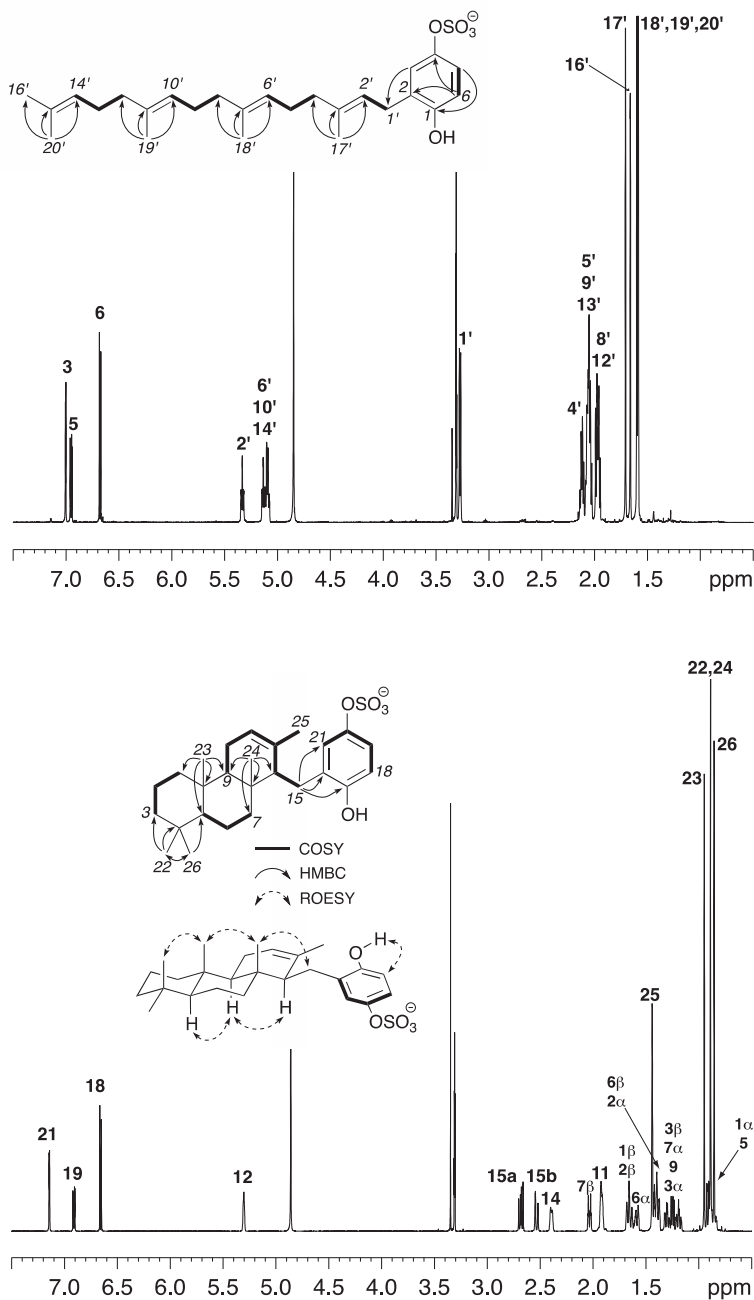


Figure 11.23 ¹H NMR (methanol-*d*₄) spectra and selected 2D NMR correlations for (top) fascioquinol E and (bottom) fascioquinol A.

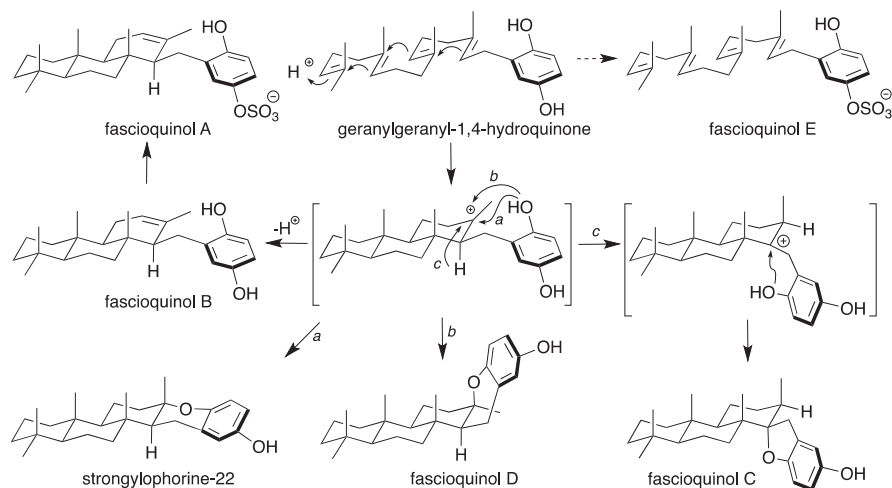


Figure 11.24 Plausible biosynthetic/chemical relationship between fascioquinols.

11.3 Marine Natural Product Snapshots

(±)-Purealin (Figure 11.25), reported in 2012 from an Australian sponge, *Pseudoceratina* sp., and a modest Gram-positive antibiotic, is the first recorded example of a naturally racemic bromotyrosine-derived spiroisoxazole,¹⁰ while massadine-14-*O*-sulfate (Figure 11.26), reported in 2012 from a deep-water Great Australian Bight sponge, *Axinella* sp., is the first reported example of a sulfated bromopyrrolo-2-aminoimidazole.¹¹ Discorhabdin X (Figure 11.27), reported in 2009 from an Australian sponge, *Higginsia* sp., features a rare, fused heterocyclic pyrroloiminoquinone scaffold,¹² while mirabilin I (Figure 11.28), reported in 2010 from an Australian sponge, *Clathria* sp., incorporates a stereochemically challenging fused cyclic scaffold.¹³ Lamellarin S (Figure 11.29), reported in 1996 from an Australian tunicate, *Didemnum* sp., and the first example of this marine alkaloid class to exhibit atropoisomerism,¹⁴ was subsequently identified as an inhibitor of the neurodegenerative disease kinases CK1 and CDK5.¹⁵ Ianthellidone F (Figure 11.30), reported in 2012 from an Australian sponge, *Ianthella* sp., and a first-in-class exemplar of a new family of pyrrolidone alkaloids, is an inhibitor of the neurodegenerative disease protease β -secretase.¹⁶ Ningalin G (Figure 11.31), reported in 2012 from an Australian tunicate, *Didemnum* sp., is an inhibitor of the kinases CK1, CDK5, and GSK3 β ,¹⁷ while dictyodendrin J (Figure 11.32), reported in 2012 from an Australian sponge, *Ianthella* sp., and bearing a unique *seco*-carbon skeleton with an unusual 1,2-diketone functionality, inhibits the protease β -secretase.¹⁸ *N*³,5'-cycloxanthosine (Figure 11.33), reported in 2005 from a deep-water Great Australian Bight sponge, *Erylus* sp., was the first recorded natural occurrence of a cyclonucleoside,¹⁹ whereas 9-(5'-deoxy-5'-thio- β -D-xylofuranosyl)adenine disulfide (Figure 11.34), reported in 2010 from an Australian sponge, *Trachycladus*

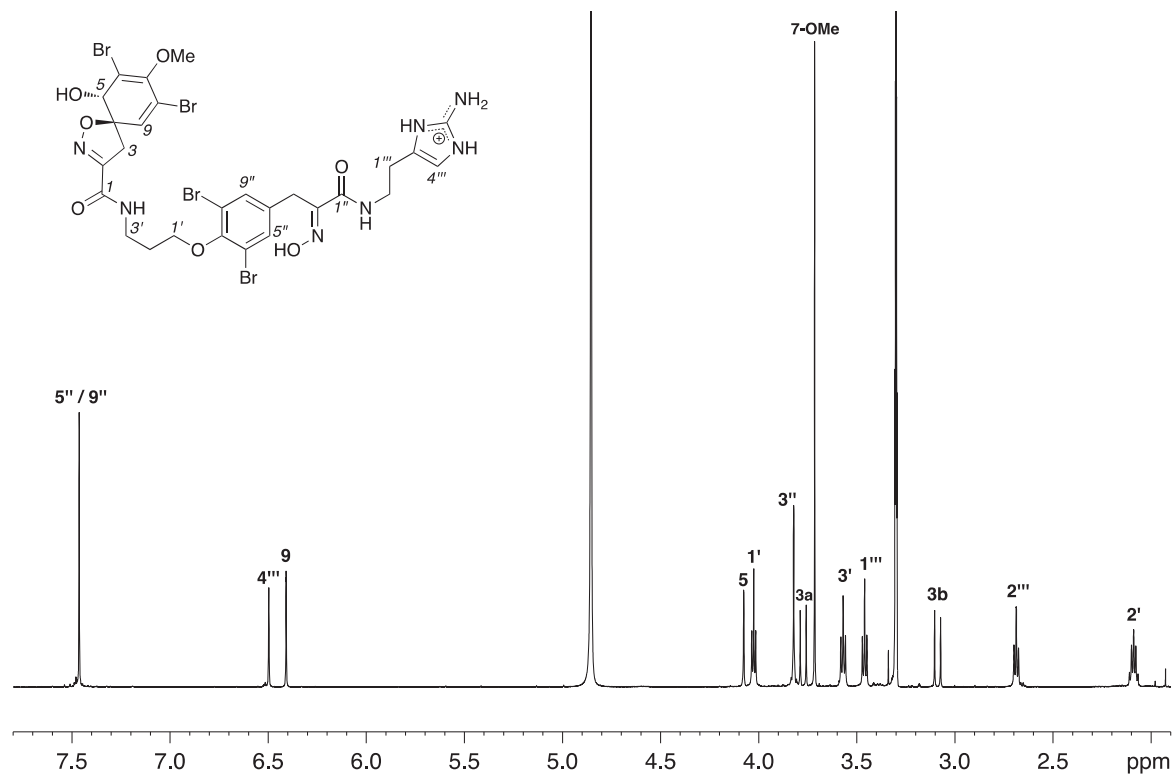


Figure 11.25 ^1H NMR (600 MHz, CD_3OD) spectrum of (±)-purealin (single arbitrary enantiomer displayed).¹⁰

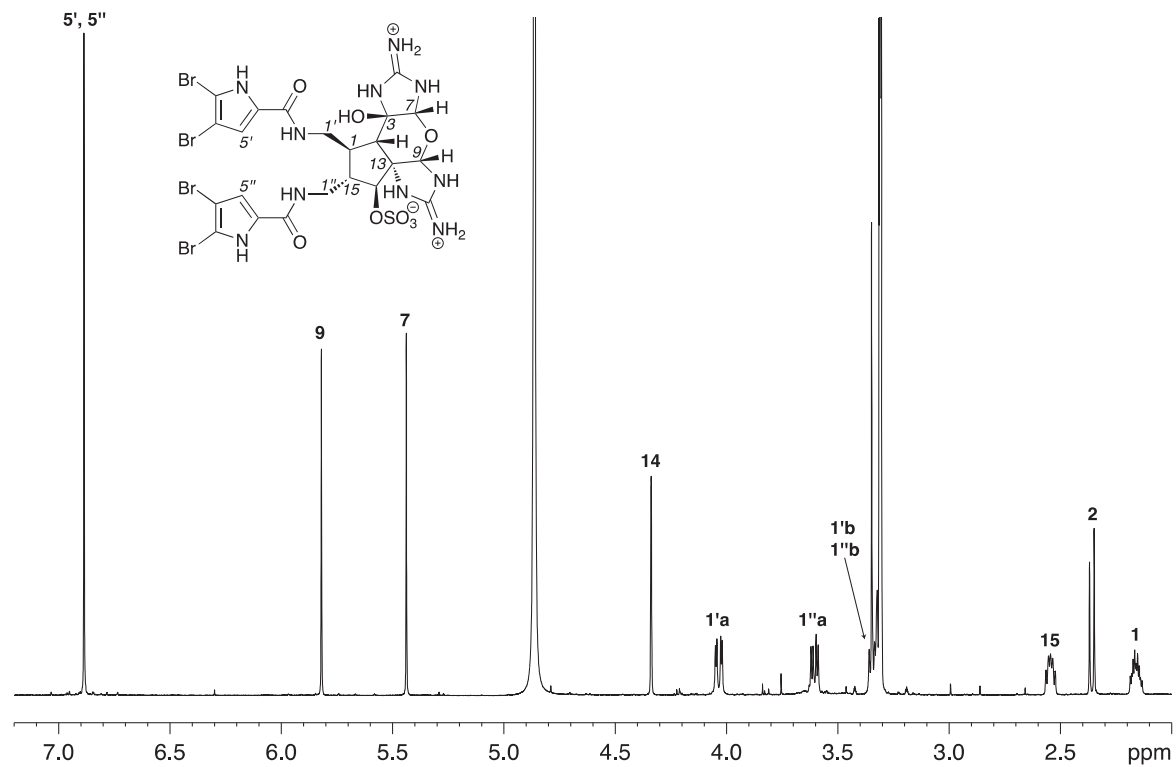


Figure 11.26 ^1H NMR (600 MHz, $\text{methanol-}d_4$) spectrum of massadine-14-O-sulfate.¹¹

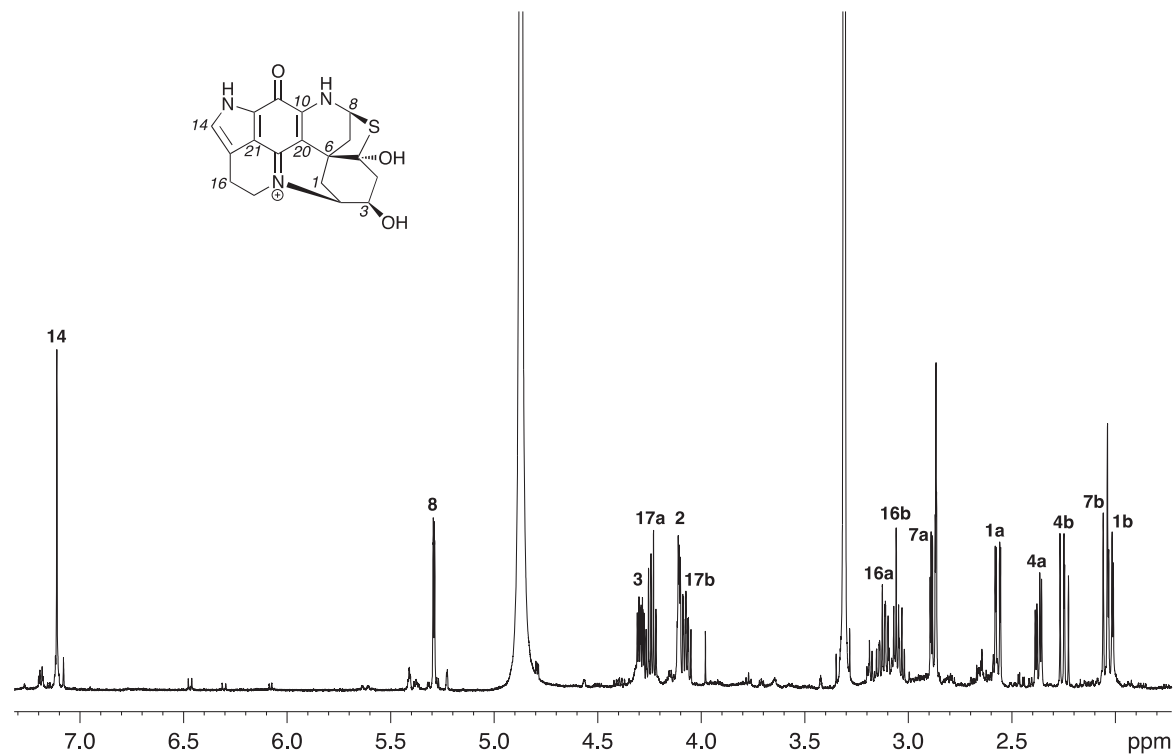


Figure 11.27 ^1H NMR (600 MHz, $\text{methanol-}d_4$) spectrum of discorhabdin X.¹²

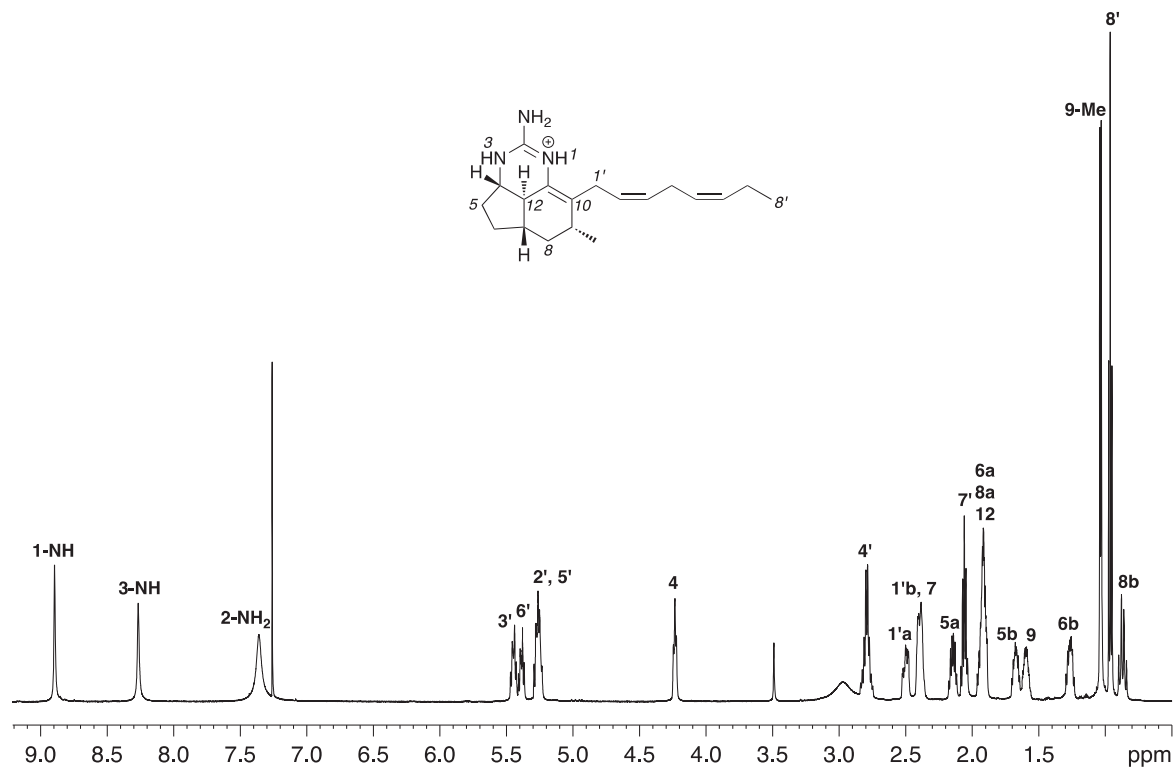


Figure 11.28 ^1H NMR (600 MHz, CDCl_3) spectrum of mirabilin I.¹³

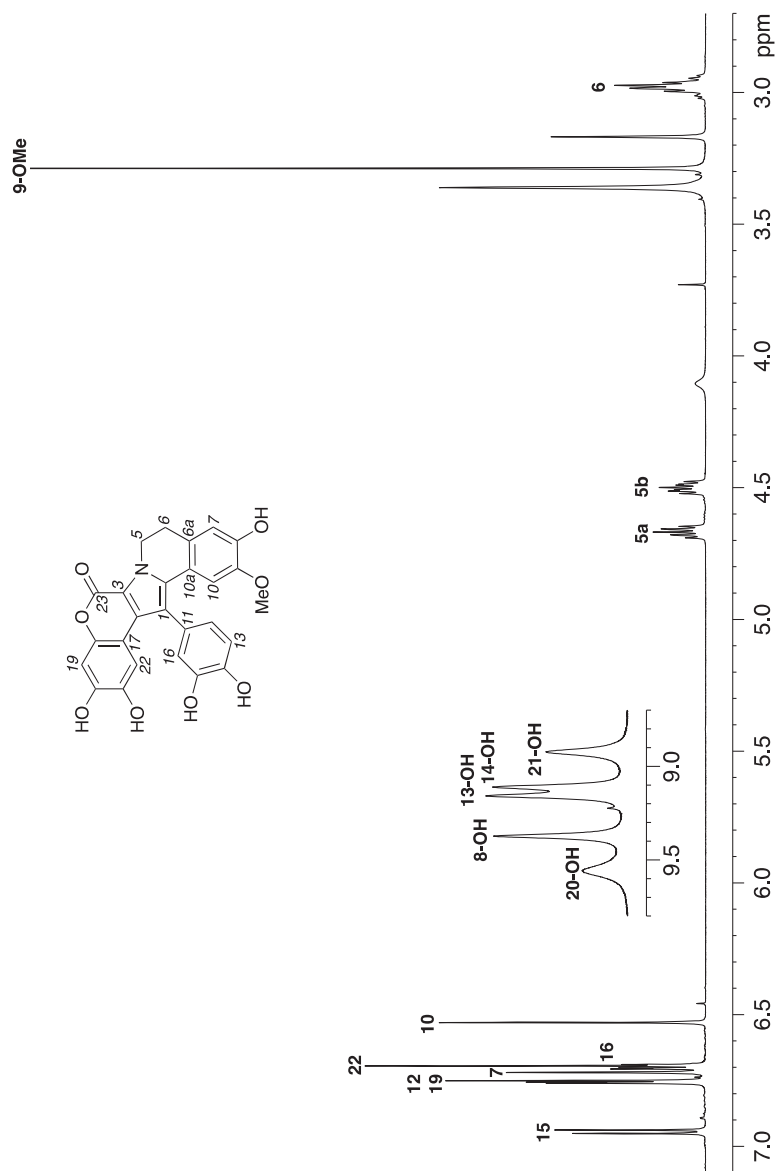


Figure 11.29 ^1H NMR (600 MHz, $\text{DMSO-}d_6$) spectrum of lamellarin S.^{14}

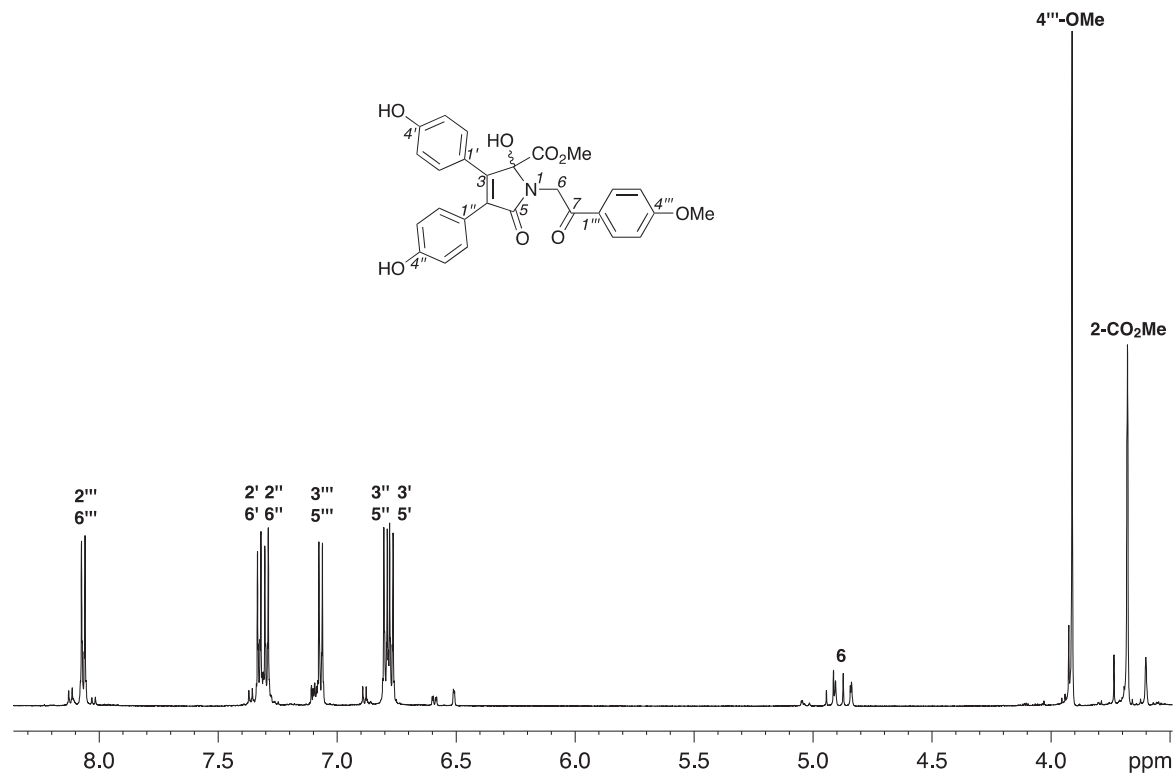


Figure 11.30 ^1H NMR (600 MHz, acetone- d_6) spectrum of ianthellidone F.¹⁶

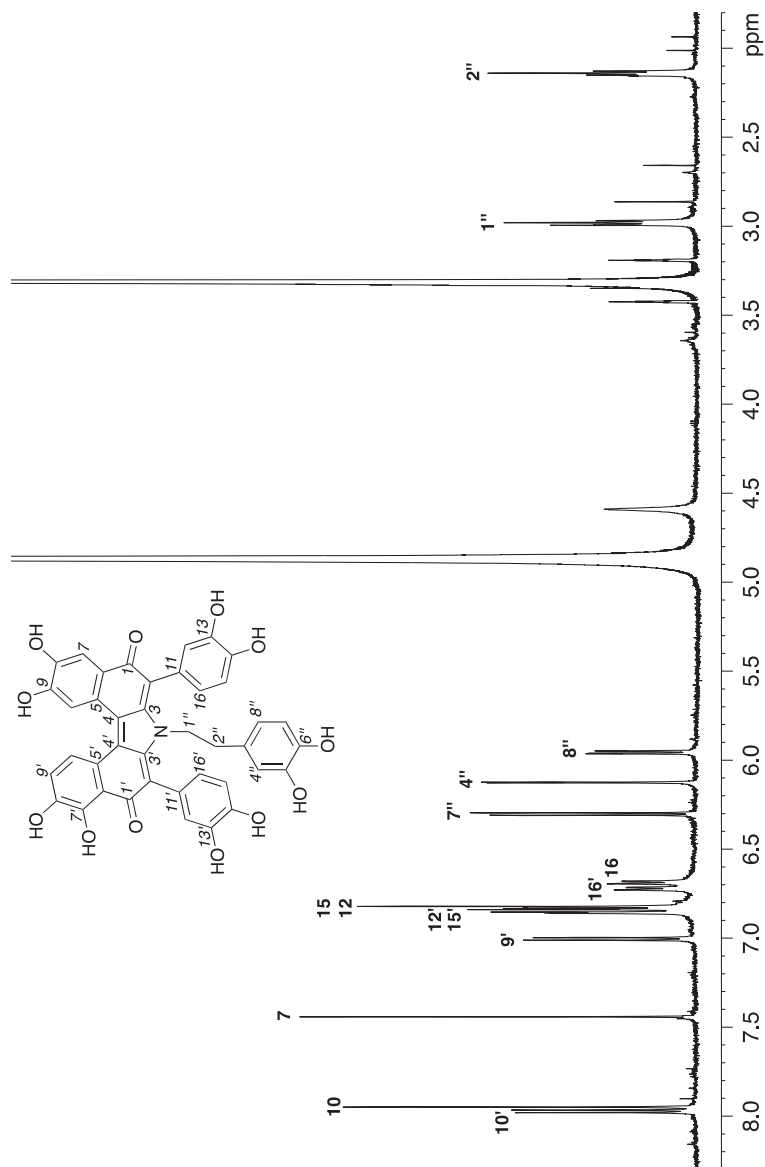


Figure 11.31 ^1H NMR (600 MHz, $\text{methanol-}d_4$) spectrum of ningalin G.¹⁷

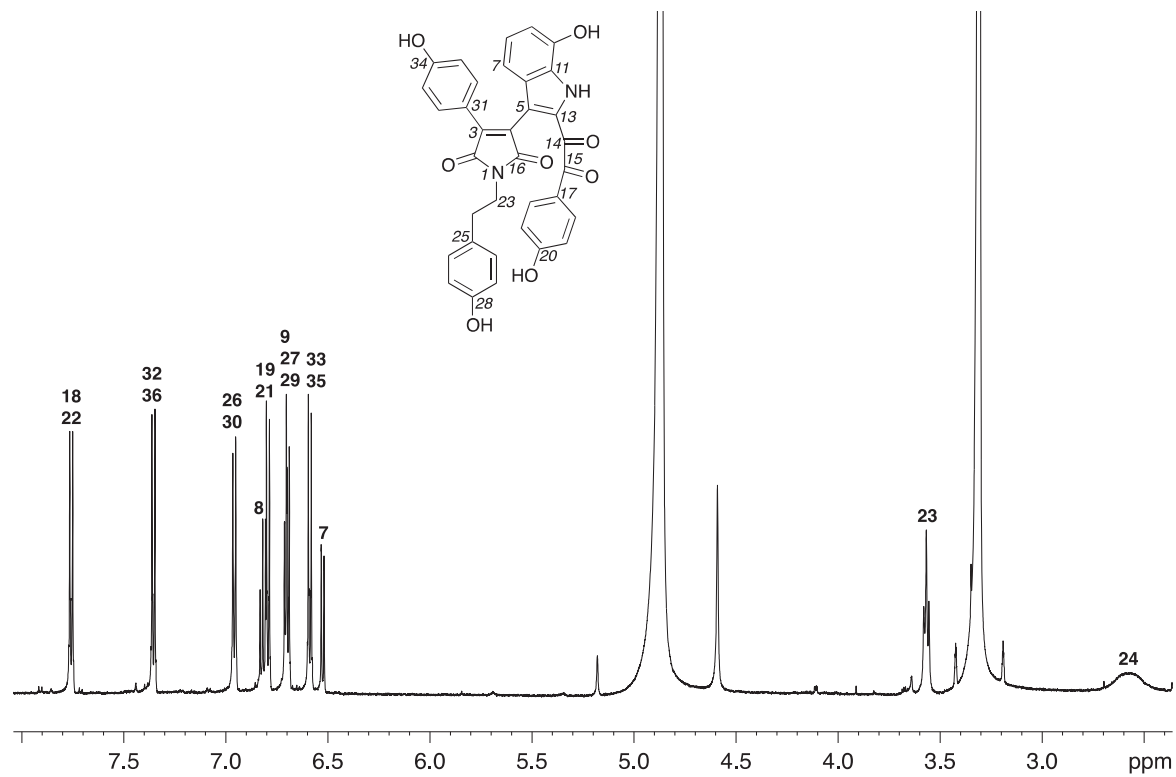


Figure 11.32 ^1H NMR (600 MHz, methanol- d_4) spectrum of dictyodendrin J.¹⁸

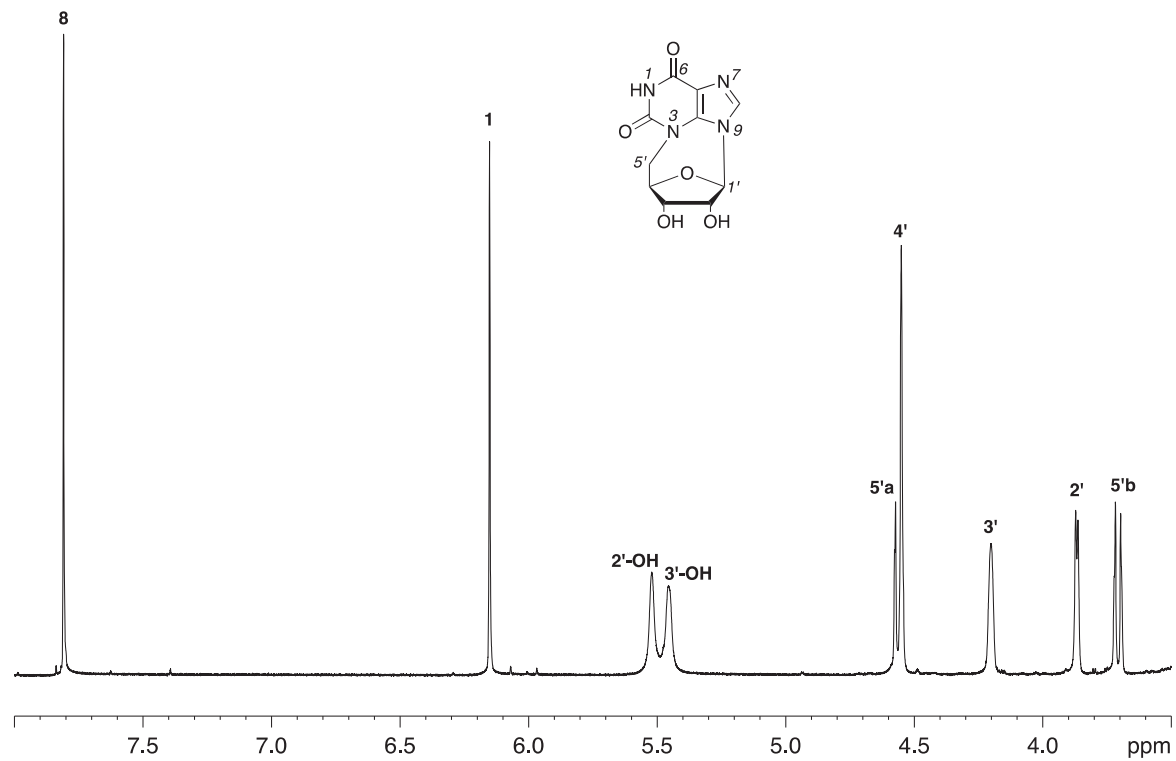


Figure 11.33 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of $N^3,5'$ -cycloxanthosine.¹⁹

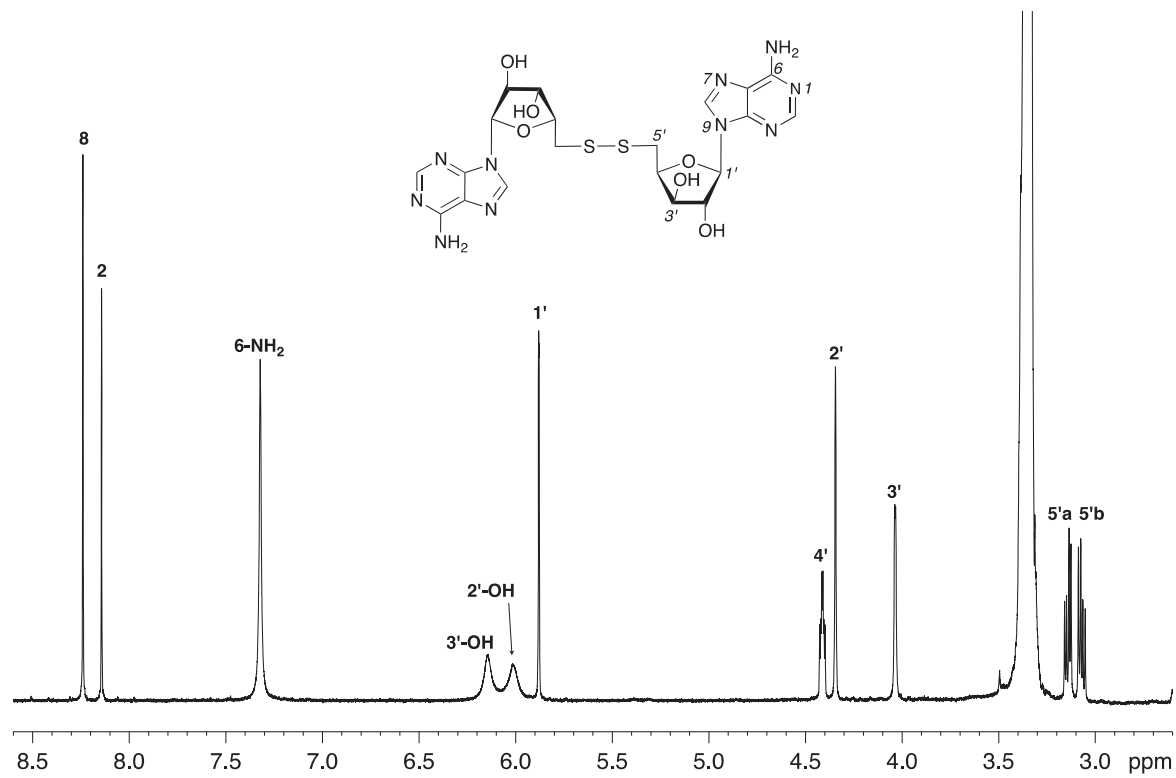


Figure 11.34 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of 9-(5'-deoxy-5'-thio- β -D-xylofuranosyl)adenine disulfide.²⁰

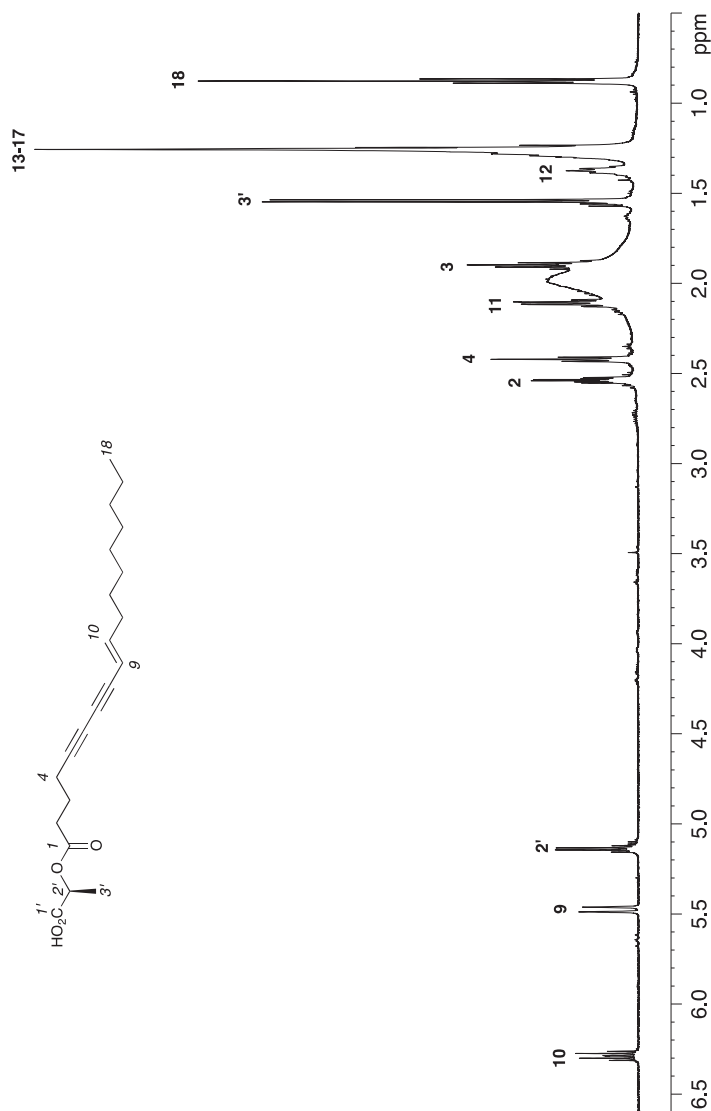


Figure 11.35 ^1H NMR (600 MHz, CDCl_3) spectrum of heterofibrin A2.²¹

laevispirulifer, was the first recorded natural occurrence of a nucleoside disulfide and only the second of a xylo nucleoside.²⁰ Heterofibrin A2 (Figure 11.35), reported in 2010 from an Australian sponge, *Spongia* (*Heterofibrinia*) sp., was the first natural example of a fatty acid lactyl or dilactyl ester,²¹ with the free carboxylic acid (heterofibrin A1) inhibiting lipid droplet formation in fibroblasts and hepatocytes.²²

As demonstrated above, whether the NMR data are bewilderingly complex or disturbingly simple, NMR spectroscopy is well established as the pre-eminent tool for characterizing and assigning molecular structures to marine natural products.

References

1. R. J. Capon, *Aust. J. Chem.*, 2010, **63**, 851.
2. R. J. Capon, C. Peng and C. Dooms, *Org. Biomol. Chem.*, 2008, **6**, 2765.
3. H. Zhang, M. M. Conte and R. J. Capon, *Angew. Chem., Int. Ed.*, 2010, **49**, 9904.
4. M. El-Naggar, A. M. Piggott and R. J. Capon, *Org. Lett.*, 2008, **10**, 4247.
5. H. Zhang and R. J. Capon, *Org. Lett.*, 2008, **10**, 1959.
6. H. Zhang, J. M. Major, R. J. Lewis and R. J. Capon, *Org. Biomol. Chem.*, 2008, **6**, 3811.
7. W. Balansa, R. Islam, F. Fontaine, A. M. Piggott, H. Zhang, T. I. Webb, D. F. Gilbert, J. W. Lynch and R. J. Capon, *Bioorg. Med. Chem.*, 2010, **18**, 2912.
8. H. Zhang, Z. G. Khalil and R. J. Capon, *Tetrahedron*, 2011, **67**, 2591.
9. A. J. Standish, A. A. Salim, H. Zhang, R. J. Capon and R. Morona, *PLoS One*, 2012, **7**, e36312.
10. A. A. Salim, Z. G. Khalil and R. J. Capon, *Tetrahedron*, 2012, **68**, 9802.
11. H. Zhang, Z. Khalil, M. M. Conte, F. Plisson and R. J. Capon, *Tetrahedron Lett.*, 2012, **53**, 3784.
12. M. El-Naggar and R. J. Capon, *J. Nat. Prod.*, 2009, **72**, 460.
13. M. El-Naggar, M. Conte and R. J. Capon, *Org. Biomol. Chem.*, 2010, **8**, 407.
14. S. Urban and R. J. Capon, *Aust. J. Chem.*, 1996, **49**, 711.
15. F. Plisson, X.-C. Huang, H. Zhang, Z. Khalil and R. J. Capon, *Chem. – Asian J.*, 2012, **7**, 1616.
16. H. Zhang, M. M. Conte, X.-C. Huang, Z. Khalil and R. J. Capon, *Org. Biomol. Chem.*, 2012, **10**, 2656.
17. F. Plisson, M. Conte, Z. Khalil, X.-C. Huang, A. M. Piggott and R. J. Capon, *ChemMedChem*, 2012, **7**, 983.
18. H. Zhang, M. M. Conte, Z. Khalil, X.-C. Huang and R. J. Capon, *RSC Adv.*, 2012, **2**, 4209.
19. R. J. Capon and N. S. Trotter, *J. Nat. Prod.*, 2005, **68**, 1689.

20. C. Peng, G. M. K. B. Gunaherath, A. M. Piggott, Z. Khalil, M. Conte and R. J. Capon, *Aust. J. Chem.*, 2010, **63**, 873.
21. A. A. Salim, J. Rae, F. Fontaine, M. M. Conte, Z. Khalil, S. Martin, R. G. Parton and R. J. Capon, *Org. Biomol. Chem.*, 2010, **8**, 3188.
22. J. Rae, F. Fontaine, A. A. Salim, H. P. Lo, R. J. Capon, R. G. Parton and S. Martin, *PLoS One*, 2011, **6**, e22868.

CHAPTER 12

Nuclear Magnetic Resonance Case Studies in Microbial Natural Products

ANDREW M. PIGGOTT, ANGELA A. SALIM AND
ROBERT J. CAPON*

Division of Chemistry and Structural Biology, Institute for Molecular
Bioscience, The University of Queensland, St Lucia, Queensland 4072,
Australia

*Email: r.capon@uq.edu.au

12.1 Introduction

To maintain and advance the quality of life offered by modern healthcare requires an ongoing commitment to the development of new drugs, to improve and replace those that have become less effective, and to bring to the community safer treatments for an ever-wider array of important diseases. Regardless of the specific medical need, the drug discovery pipeline is critically dependent on access to diverse, high-quality molecular libraries that are capable of inspiring the discovery process. A poor choice of chemistry leads to wasted effort and resources, and no new drugs! Historically, the pharmaceutical industry has relied heavily on natural products, particularly microbial metabolites, which represent an extraordinarily diverse, pre-assembled pool of biologically active molecules, programmed by evolution to be potent and selective modulators of key biopolymers, cells, tissues, organs, and animals. Access to nature's intellectual property, derived from the evolutionary

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

equivalent of a billion-year global drug discovery program supported by an unlimited budget and a workforce of trillions, has delivered an array of privileged bioactive scaffolds that informed and inspired modern drug discovery, re-purposing ecological advantage to pharmaceutical benefit. *Actinomycetes* species, for example, have delivered >50% of all antibiotics in use today, including macrolides (e.g. erythromycin), glycopeptides (e.g. vancomycin), aminoglycosides (e.g. streptomycin), anthracyclins (e.g. tetracycline), and polyenes (e.g. amphotericin), as well as important anticancer (e.g. doxorubicin), immunosuppressive (e.g. rapamycin), and anthelmintic (e.g. ivermectins) agents. Fungi have also been significant producers of important drugs, with the genus *Penicillium* yielding arguably the most well-recognized name in modern drugs, the antibiotic penicillins, along with one of the most commercially successful drug classes in history, the antilipidemic statins. The journey to discover, isolate, identify, evaluate, develop, manufacture, and deliver microbial metabolites (and their derivatives) to market sparked and fueled a revolution in global science, commercialization, and healthcare. Notwithstanding their impact on human health, the structure elucidation of microbial metabolites, which are rich in diverse carbon and heterocyclic skeletons and unusual functionality, represents a substantial challenge. This chapter seeks to briefly illustrate the role played by nuclear magnetic resonance (NMR) spectroscopy in meeting this challenge through consideration of five microbial natural product case studies and 11 snapshots.

12.1.1 Case Studies

The heavily modified dipeptide aspergillazines illustrate both the opportunities and challenges encountered when comparing the spectroscopic properties of biosynthetically related co-metabolites. For example, although the known trichodermamide A and the new aspergillazine A appear to differ only by replacement of the 5-oxo moiety in the former with a 5-thio moiety in the latter, the inherent and unexpected chemical properties of this heterocyclic framework dictated a far more complex outcome. Likewise, the recent discovery of the diketopiperazine (DKP) nocardioazines demonstrates that microbial natural products continue to provide new structural, biosynthetic, and biological insights, even in well-studied systems. The polyketide heronamides, kibdelones, and reveromycins provide excellent platforms to demonstrate the role of NMR spectroscopy in assigning complex molecular structures, inclusive of multiple regiochemical, conformational, and stereochemical isomers. All of these case studies emphasize how biosynthetic considerations can be used to probe structural, chemical, and biological properties, supported by well-dispersed and information-dense NMR spectra.

12.1.2 Snapshots

The 11 snapshots are inclusive of a wide selection of structural and biosynthetically diverse microbial metabolites and have been chosen to

illustrate both the complexity and the value of NMR spectroscopy. Suitable references have been provided for those wishing to delve further into the isolation, characterization, and structure elucidation of these remarkable compounds.

12.2 Microbial Natural Product Case Studies

12.2.1 Aspergillazines

Chemical analysis of an Australian strain of the fungus *Aspergillus unilateralis* (MST-F8675), isolated from a soil sample collected near Mount Isa, Queensland, yielded a complex array of highly modified dipeptides, including the new aspergillazines A–E and the known trichodermamides A (*cf.* penicillazine) and B (Figure 12.1).¹ The chemical and spectroscopic properties of these metabolites ensured that the structure elucidation was particularly challenging. Indicative of this challenge, a 2000 isolation of penicillazine from a marine-derived *Penicillium* sp. employed X-ray analysis to incorrectly assign a quinolone-ester structure,² while a 2003 re-isolation from another marine-derived fungus *Trichoderma virens* employed NMR technology to arrive at the correct chromone-amide structure.³ The 2005 investigation of *A. unilateralis* expanded our knowledge of this rare structure class, encountering an unprecedented array of thiophane analogs (*e.g.* aspergillazine A).

When confronted with the need to solve the structure of a complex new natural product, one particularly powerful approach is to compare NMR data with closely related known natural products. This process is greatly accelerated if the known natural products are themselves isolated and characterized as co-metabolites. For example, a comparison of the NMR data for aspergillazine A with those for trichodermamide A (Figure 12.2) established

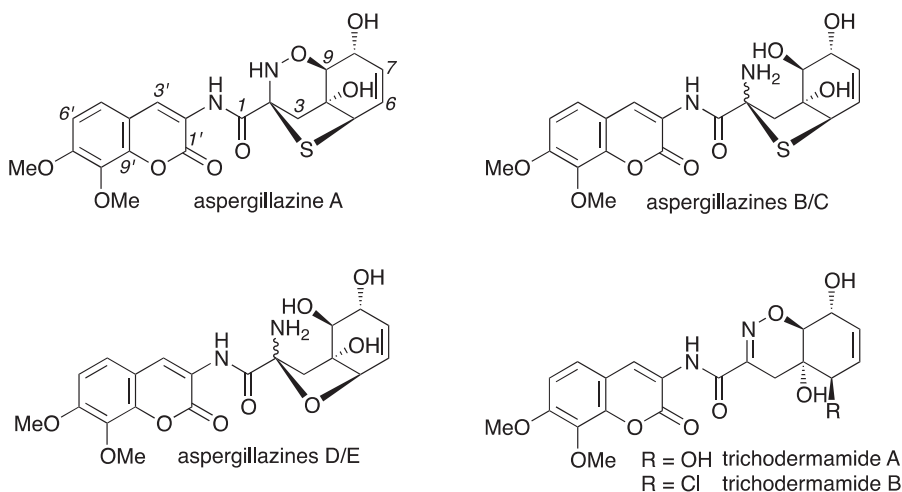


Figure 12.1 Aspergillazines A–E and trichodermamides A–B.

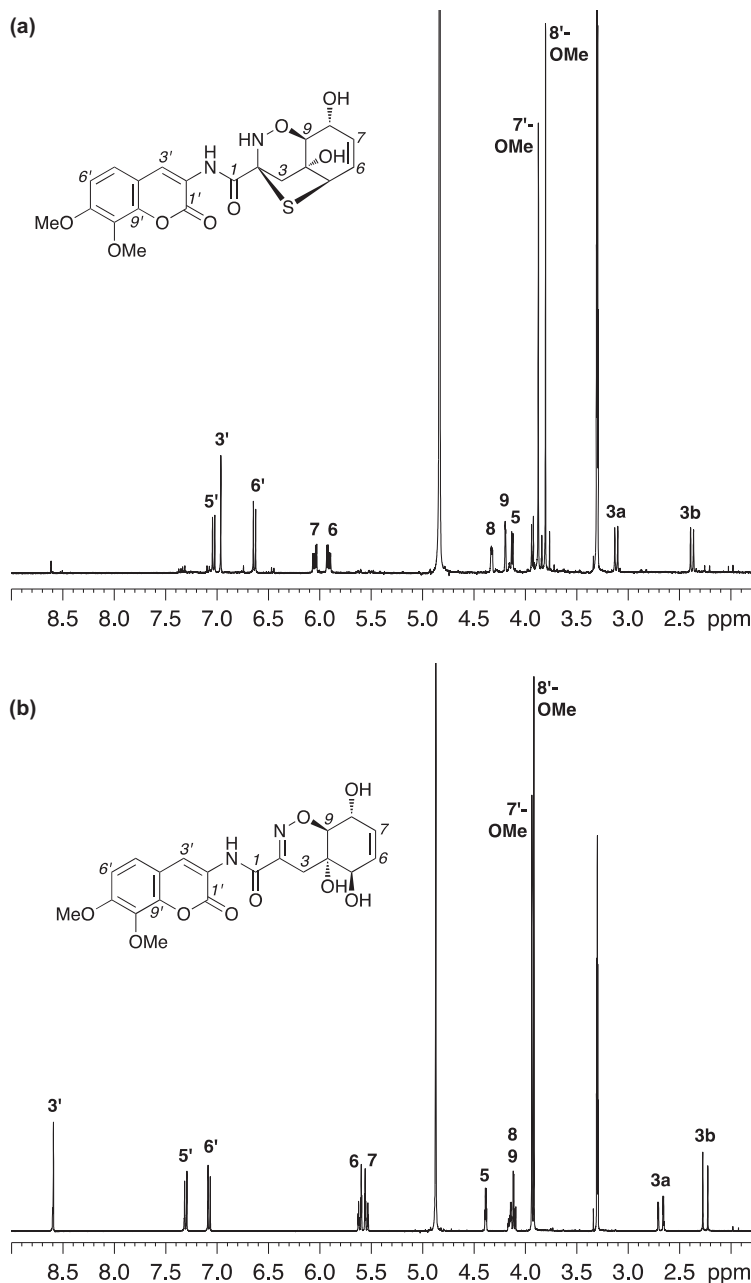


Figure 12.2 ^1H NMR spectra (400 MHz, CD_3OD) of (a) aspergillazine A and (b) trichodermamide A.

a common dimethoxyaminocoumarin N-terminus (C-1' to C-9'), correctly focusing attention on structural differences centered about the C-terminus (C-1 to C-9). More specifically, whereas trichodermamide A possessed a

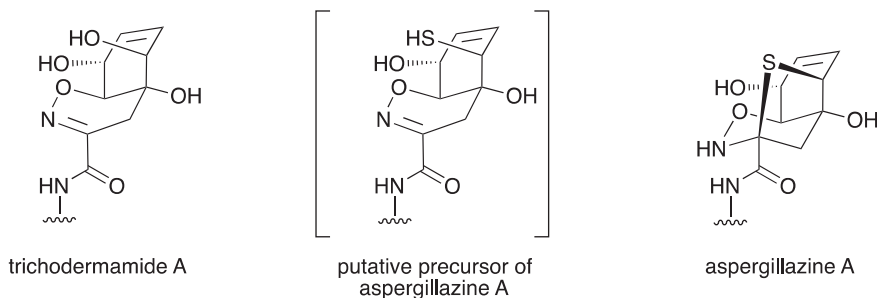


Figure 12.3 Representation of trichodermamide A, aspergillazine A and its putative biosynthetic precursor.

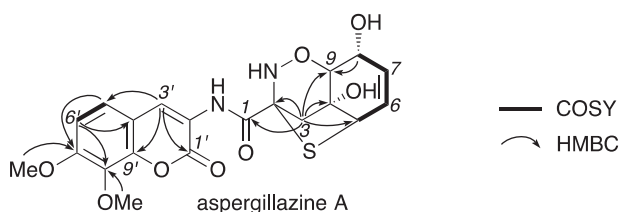


Figure 12.4 Selected 2D NMR correlations for aspergillazine A.

deshielded C-5 oxymethine (δ_{C} 73.5), aspergillazine A possessed a less deshielded C-5 thiomethine (δ_{C} 47.7). Moreover, conversion of the C-2 sp^2 oximino carbon in trichodermamide A (δ_{C} 151.4) to a quaternary sp^3 carbon in aspergillazine A (δ_{C} 77.2) was attributed to intramolecular addition and thiophane ring formation to the 5-thio moiety (Figure 12.3), with a comprehensive analysis of the 1D and 2D NMR data (Figure 12.4) completing the structure assignment. Similar NMR analyses confirmed the structures of aspergillazines B and C (Figure 12.1) as stable (non-interconverting) epimeric ring-opened amino-thiophane analogs of aspergillazine A. The stability of these amino-thiophanes was especially noteworthy, particularly as the epimeric amino-tetrahydrofuran co-metabolites aspergillazines D and E proved to be remarkably unstable, undergoing rapid epimerization immediately after high-performance liquid chromatography (HPLC) separation (Figure 12.5). These observations highlight the impact that even small functional group changes (*e.g.* a hydroxyl *versus* a thiol) can have on highly tuned natural product scaffolds.

12.2.2 Nocardioazines

Chemical analysis of a *Nocardioopsis* sp. (CMB-M0232) recovered from a marine sediment sample (−55 m) collected off South Molle Island near Brisbane, Australia, revealed two new prenylated DKPs, nocardioazines A and B (Figure 12.6), as well as two known DKPs, *cyclo*-(L-Trp-L-Trp) and *cyclo*-(L-Trp-D-Trp).⁴ Detailed analysis of the 1D (Figure 12.7) and 2D (Figure 12.8) NMR data

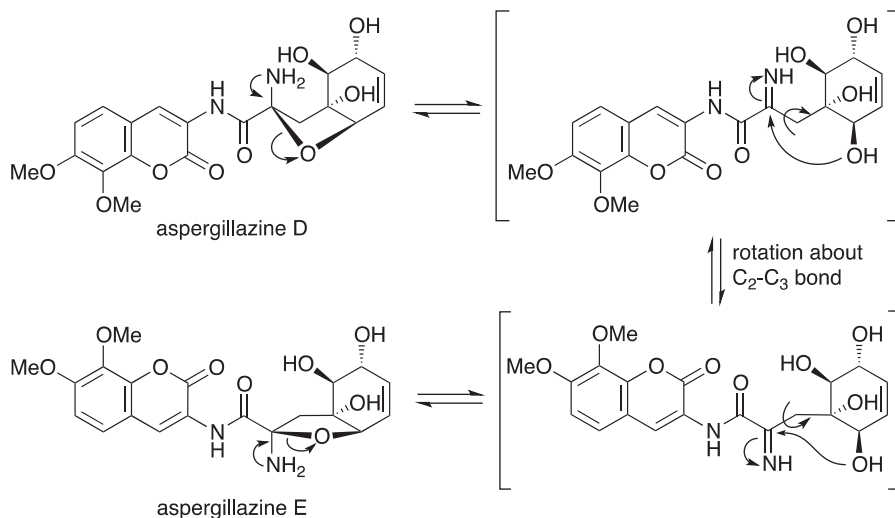


Figure 12.5 Proposed mechanism for C-2 epimer equilibration between aspergillazines D and E.

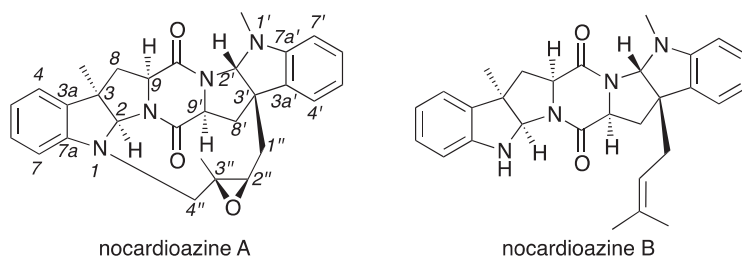


Figure 12.6 Nocardioazines A and B.

for nocardioazine A revealed a DKP scaffold consisting of two tricyclic indoline systems bridged by an unprecedented N-1 to C-3' epoxy-isopentanyl “strap”. Diagnostic ROESY correlations allowed assignment of independent relative configurations for both tricyclic indoline systems and established a *trans*-epoxide configuration. Although overlapping NMR resonances for H-9 and H-9' prevented detection of the critical ROESY correlation needed to fuse the relative configurations about both nocardioazine A indoline systems, the NMR data for the “unstrapped” nocardioazine B were far more accommodating. The nocardioazines are biosynthetically noteworthy as they represent: (i) the first C-prenylated DKPs to be reported from a marine-derived bacterium; (ii) the first C-3 *normal* prenylated DKPs to be reported from any source (Figure 12.9); and (iii) the first example of a new class of bridged (strapped) DKP scaffold. Although the nocardioazines were inactive against a panel of bacteria and human cancer cell lines, nocardioazine A was found to be an inhibitor of the ABC transporter and trans-membrane efflux pump P-glycoprotein (P-gp). P-gp is over-expressed in many multidrug-resistant cancer cell lines and

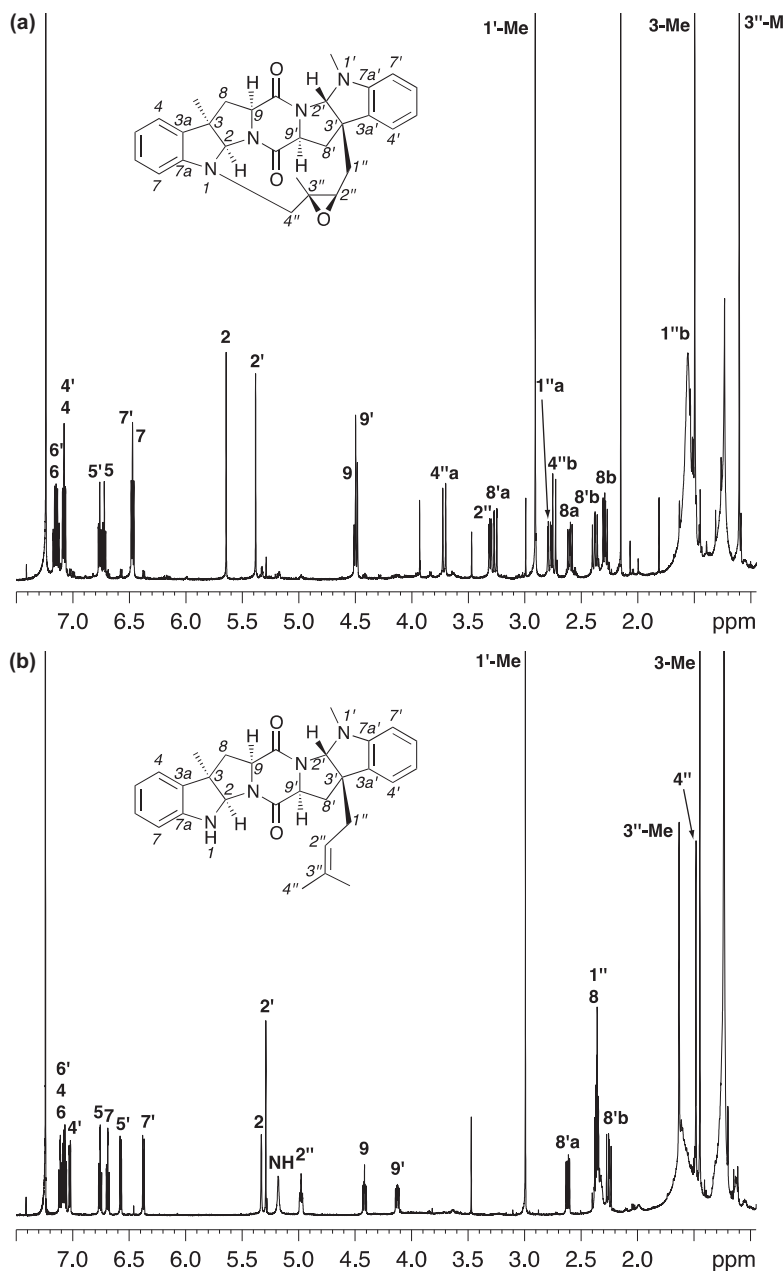


Figure 12.7 ^1H NMR spectra (600 MHz, CDCl_3) of (a) nocardioazaine A and (b) nocardioazaine B.

is an important mechanism by which cancer cell lines acquire resistance to anticancer agents. Inhibitors of P-gp represent a potential strategy to overcome this resistance, and nocardioazaine A represents a promising new lead.

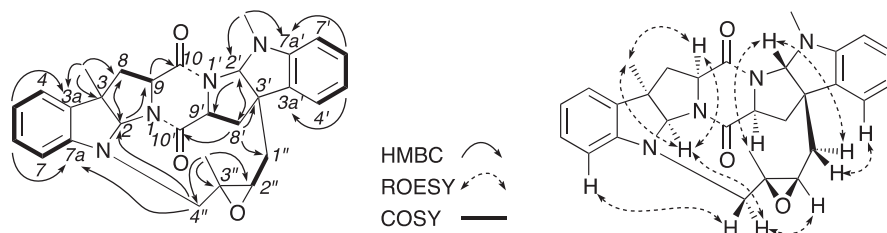


Figure 12.8 Selected 2D NMR correlations for nocardioazaine A.

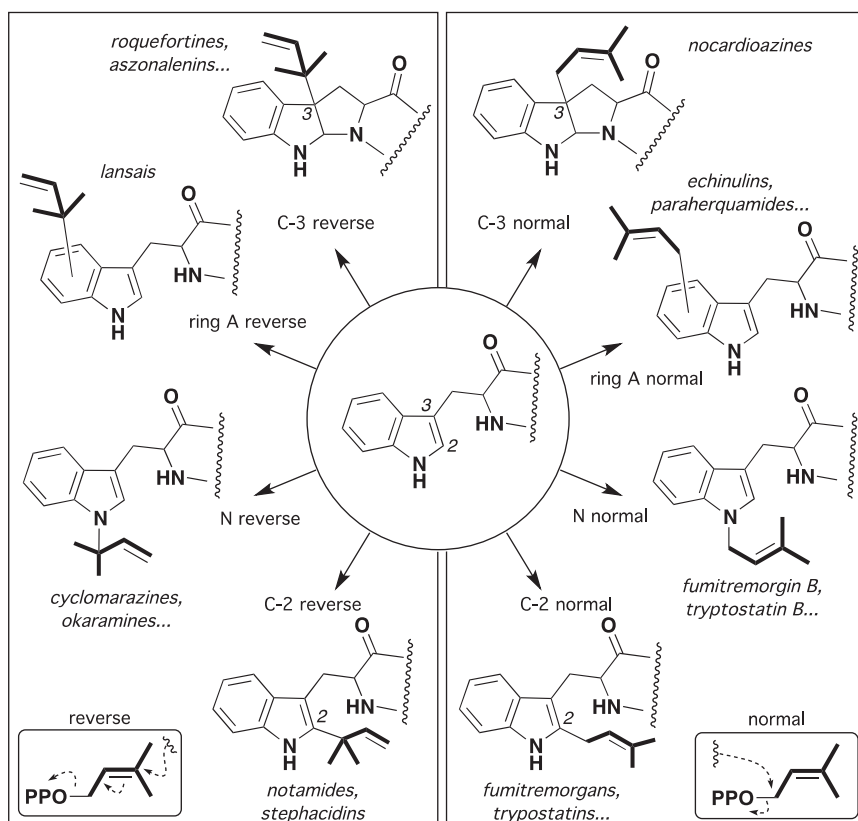


Figure 12.9 Indole prenylation patterns for diketopiperazines incorporating tryptophan.

12.2.3 Heronamides

A *Streptomyces* sp. (CMB-M0406) isolated from shallow-water sediment (−1 m) collected off Heron Island, Australia, yielded heronamides A–C (Figure 12.10) as new examples of a rare class of polyketide 20-membered macrolactams featuring unprecedented carbon/heteroatom skeletons.⁵ Detailed analysis of the 1D (Figure 12.11a) and 2D (Figure 12.13) NMR data

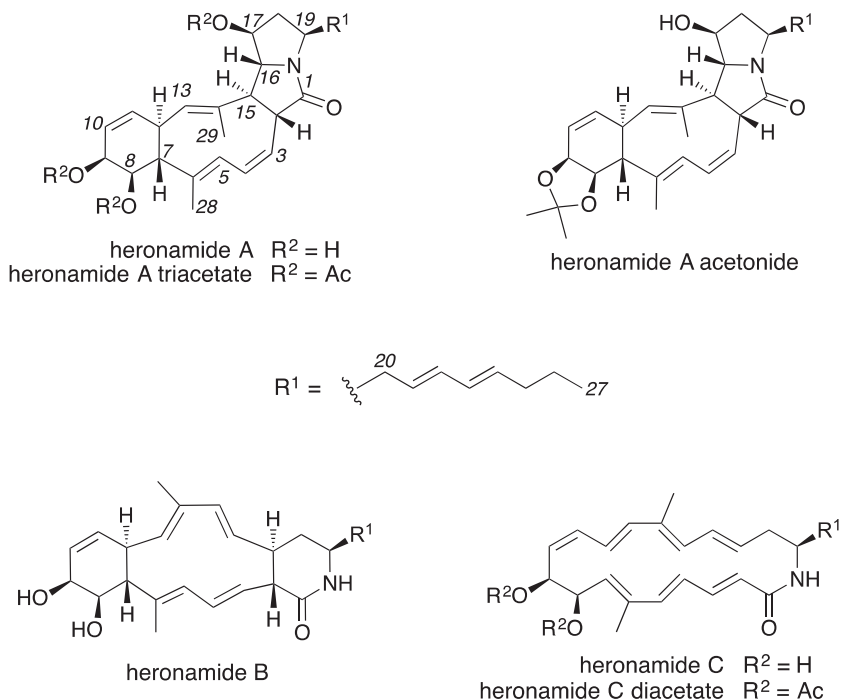


Figure 12.10 Heronamides A–C and synthetic derivatives.

for heronamide A revealed three distinct structure fragments (subunits A–C). The COSY and HMBC data allowed subunits B and C to be unambiguously connected, and established the C-7 to C-12 cyclohexenyl moiety. However, due to overlapping signals, the NMR data did not allow unambiguous identification of the heterocyclic core of heronamide A. To achieve better dispersion of NMR signals, heronamide A acetonide and heronamide A triacetate were prepared. Fortunately, the NMR data for the heterocyclic core of the triacetate were better resolved than in heronamide A (Figure 12.11b), permitting full assignment of the planar structure and relative configuration. Similarly, detailed analysis and comparison of the 1D (Figure 12.12a) and 2D (Figure 12.13) NMR data for heronamide B with those of heronamide A also permitted full assignment of the planar structure and relative configuration. Heronamide C proved difficult to handle and characterize, displaying very poor solubility in a range of common NMR solvents ($CDCl_3$, acetone- d_6 , DMSO- d_6 , and methanol- d_4). While reasonable-quality 1D (Figure 12.12b) and 2D (Figure 12.13) NMR data were obtained for heronamide C in pyridine- d_5 , improved solubility in methanol- d_4 and increased signal dispersion were obtained by preparing the diacetate derivative, allowing unambiguous assignment of the relative stereostructure.

A plausible biosynthetic process linking heronamides A–C (Figure 12.14) has heronapyrrole C as a pivotal polyene precursor, capable of undergoing either a synchronous $6\pi + 6\pi$ tandem electrocyclozation to generate

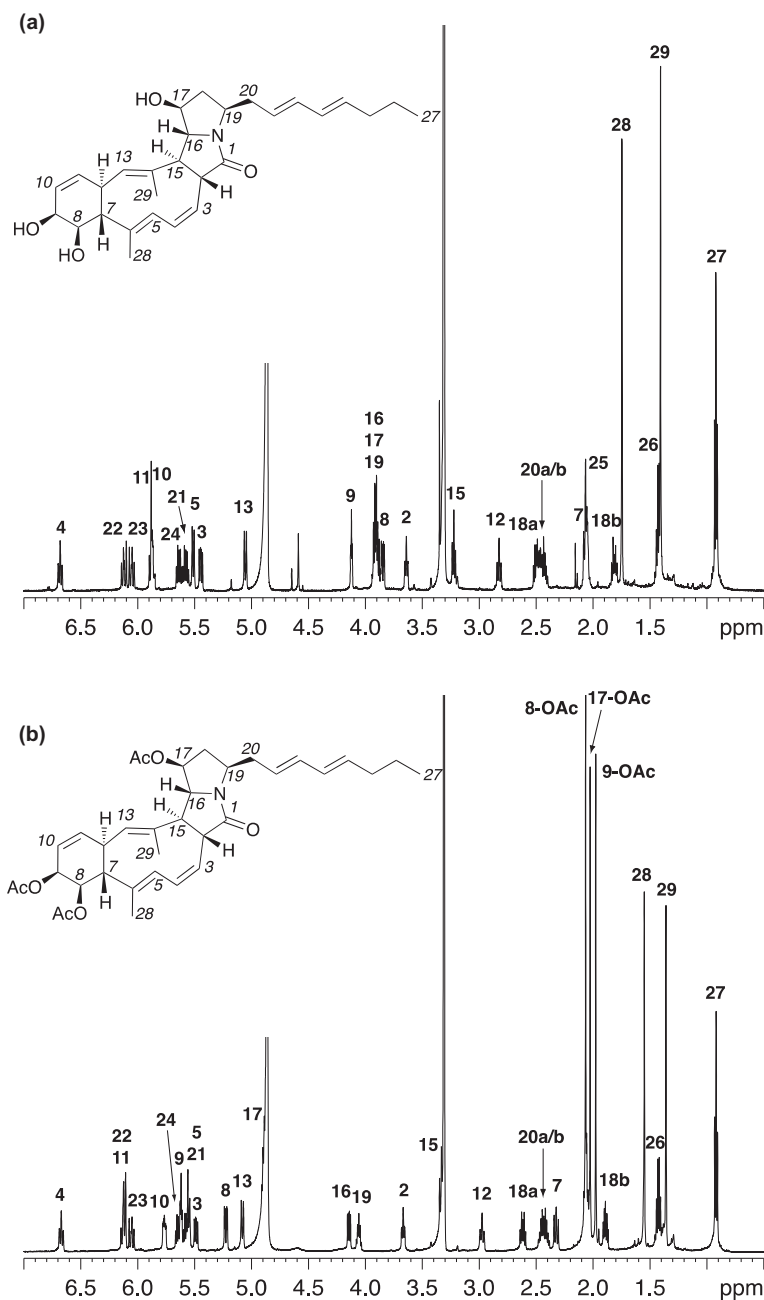


Figure 12.11 ^1H NMR spectra (600 MHz, CD_3OD) of (a) heronamide A and (b) heronamide A triacetate.

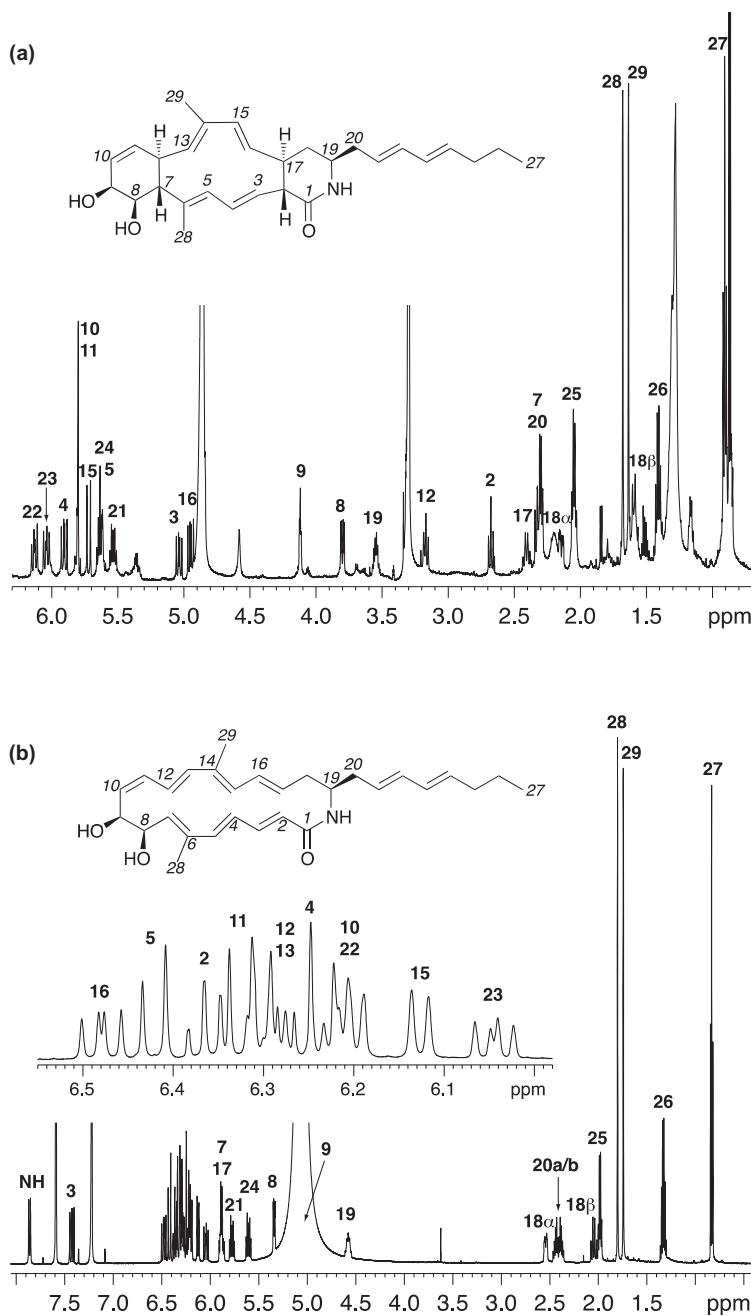


Figure 12.12 ^1H NMR spectra (600 MHz) of (a) heronamide B (CD_3OD) and (b) heronamide C (pyridine- d_5).

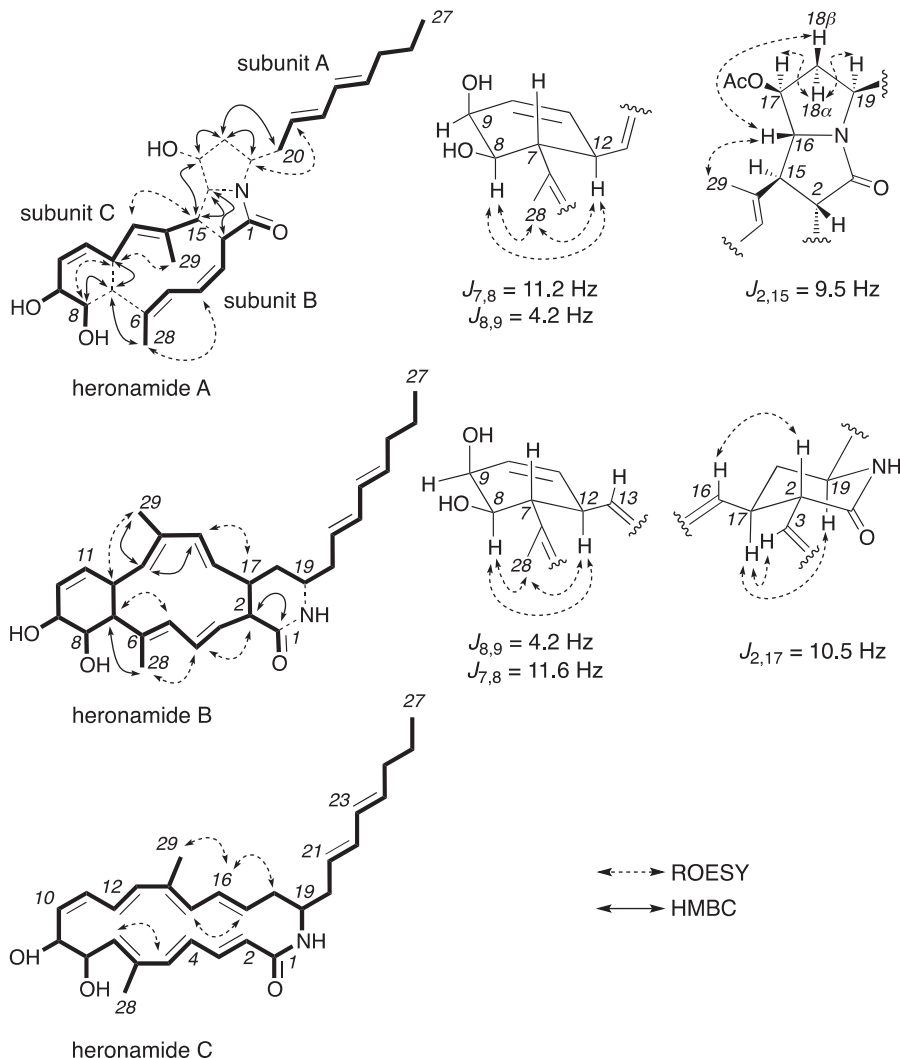


Figure 12.13 Selected 2D NMR correlations for heronamides A–C.

heronamide B, or selective epoxidation and S_N2 -mediated macrolactam-pyrrolidine formation, followed by a synchronous $4\pi + 6\pi$ tandem electrocyclization to generate heronamide A. It is proposed that the tandem electrocyclizations proceed in a conformationally synchronized conrotatory fashion, permitting tentative assignment of the complete relative configurations for heronamides A and B.

To assign absolute configurations to the heronamides, the *S*-MTPA and *R*-MTPA esters of heronamide A acetone were prepared as part of a Mosher analysis. As the available supply of all heronamides and derivatives was limited, the derivatization and subsequent NMR analyses were undertaken

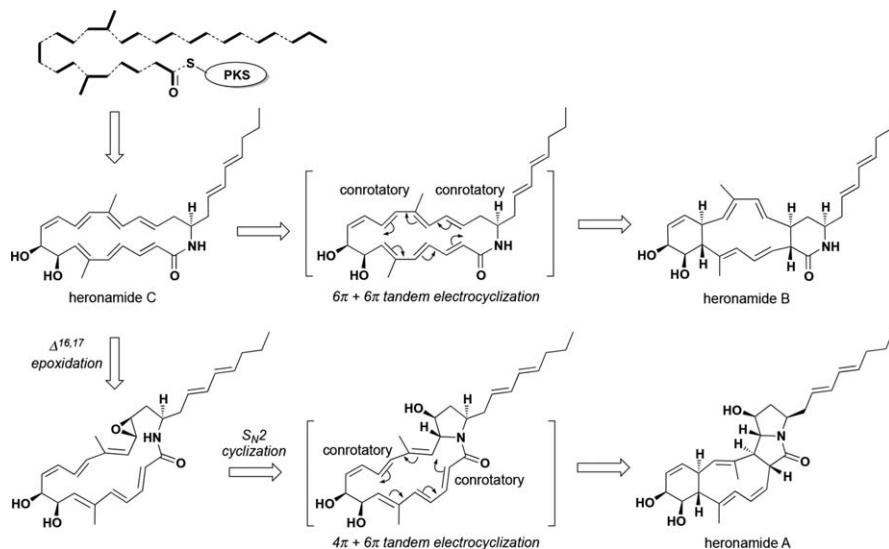


Figure 12.14 Plausible biosynthetic relationship between heronamides A–C.

in situ at an analytical scale (150 μg). The ^1H NMR spectra of the Mosher esters revealed a positive $\Delta\delta^{\text{SR}}$ for H-16 (+42 Hz) and a negative $\Delta\delta^{\text{SR}}$ for H-19 (–30 Hz), confirming a 17*S* configuration and permitting the assignment of absolute configurations to all chiral centers across heronamides A–C.

Although none of the heronamides or derivatives displayed anticancer or antibiotic activity, heronamide C elicited an unexpected and dramatic vacuolization of HeLa cells, which dissipated with no apparent adverse effects on removal of the compound from the cell medium.

12.2.4 Kibdelones

Kibdelones⁶ and isokibdelones⁷ (Figure 12.15) are a novel family of bioactive heterocyclic polyketides produced by a rare soil actinomycete, *Kibdelosporangium* sp. collected from a timber woolshed 15 km north of Port Augusta in South Australia in 1996. The ^1H NMR data (Figure 12.16) suggested a very close structural similarity between the kibdelones and isokibdelones, even though their UV spectra and HPLC retention times differed significantly. Analysis of 2D NMR data (COSY, HMBC, and ROESY) for kibdelone A revealed correlation sequences consistent with a substructure consisting of rings A and C–F (Figure 12.17). However, final assembly of the kibdelone A skeleton presented a challenge as the available data did not unambiguously differentiate between the structure as shown and an isomer with an alternative regiochemistry about the ring A/B junction. This issue was resolved by NMR spectroscopic analysis of the hydroquinone obtained by *in situ* (NMR tube) reduction of kibdelone A with sodium dithionite. The ^1H NMR spectrum of dihydrokibdelone A displayed downfield resonances for two

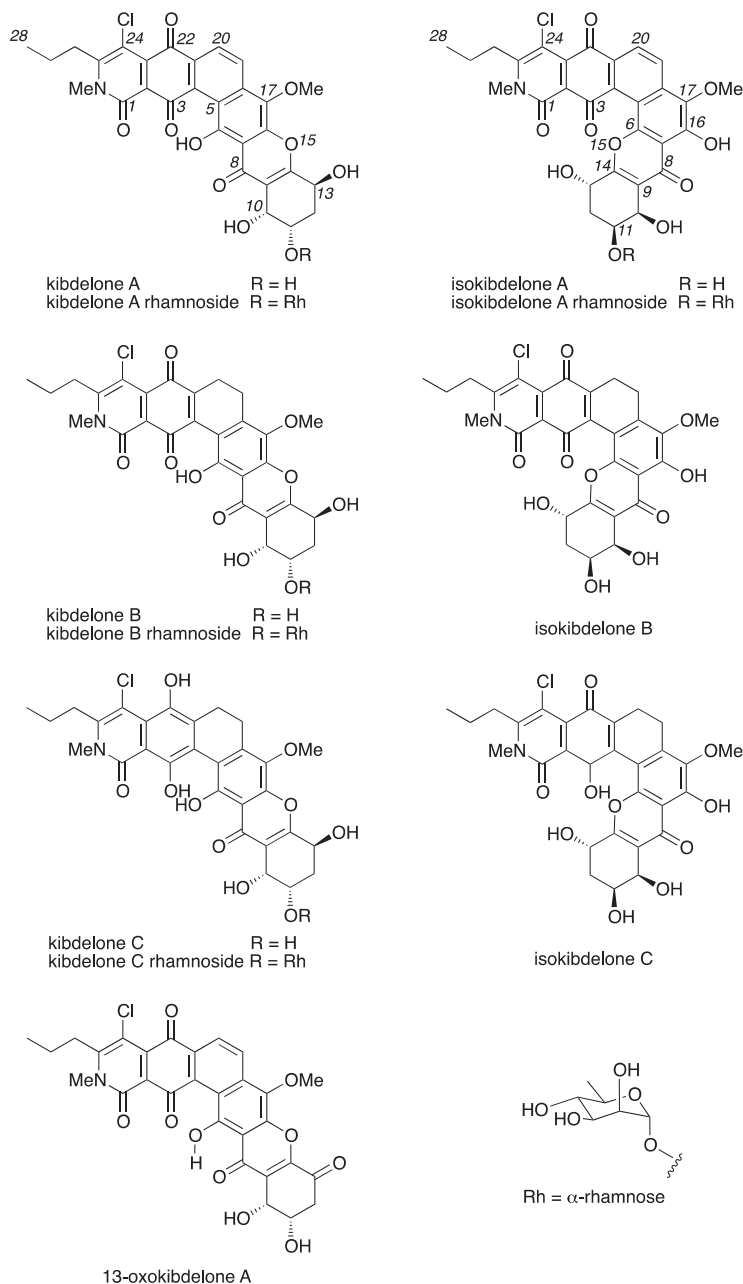


Figure 12.15 Kibdelones and isokibdelones.

chelated phenolic protons, 3-OH and 6-OH, and one non-chelated phenolic proton, 22-OH, as would be expected from the structure as shown. A set of HMBC correlations observed for dihydrokibdelone A further supported the

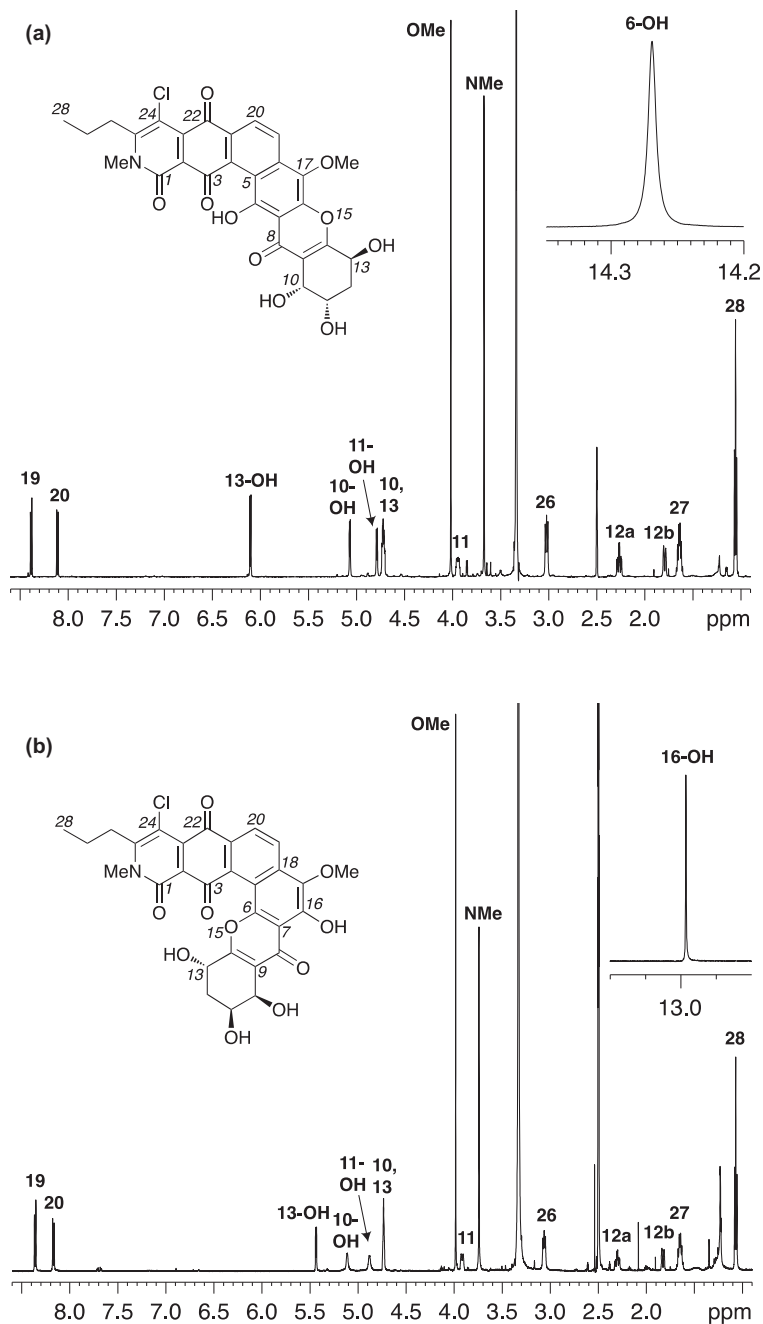


Figure 12.16 ^1H NMR spectra (600 MHz, $\text{DMSO}-d_6$) of (a) kibelone A and (b) isokibelone A.

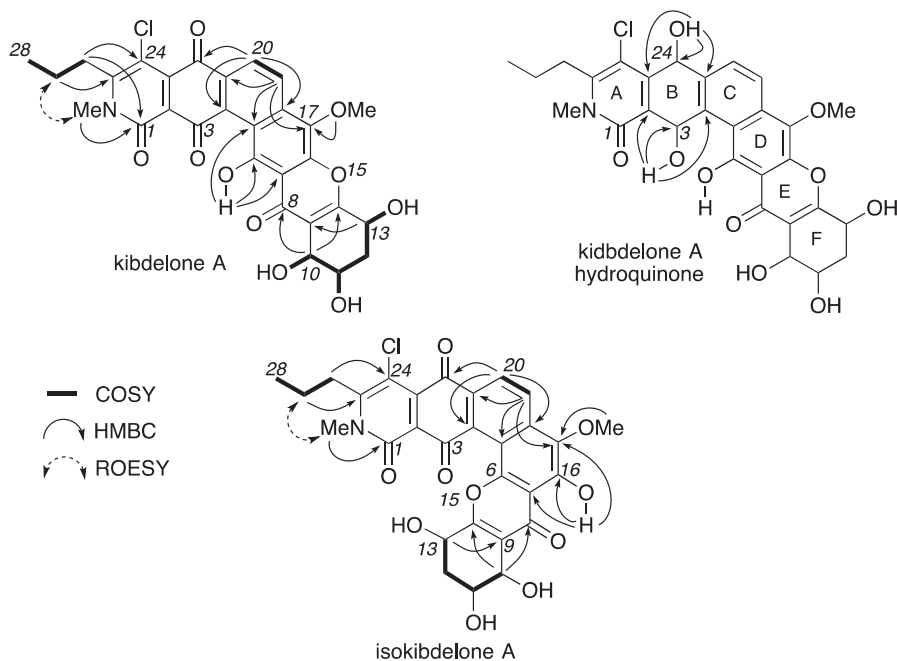


Figure 12.17 Selected 2D NMR correlations for kibelone A, kibelone A hydroquinone and isokibelone A.

connectivity of rings A and B as shown (Figure 12.17). Similarly, analysis of 1D and 2D NMR data for isokibelone A revealed the structure as shown, differing in the position of the ring D/E fusion compared to kibelone A (Figure 12.17). This ring D/E linkage also explained the chelated character of the C-16 hydroxyl in isokibelone A. Comparison of ^{13}C NMR data for kibelone A and isokibelone A confirmed a common regiochemistry for the fusion between rings A/B.

Under mild conditions, kibelones B and C undergo rapid equilibration to kibelones A–C, while kibelone B rhamnoside equilibrates to a mixture of kibelone A–C rhamnosides. This can be observed by the changing in color of MeOH solutions of kibelones B and C, from orange (B) and bright yellow (C) to greenish yellow—a process that was accelerated by heating. A plausible mechanism for this equilibration involves air oxidation, quinone/hydroquinone redox transformations, and a choreographed sequence of keto/enol tautomerizations that aromatize ring C *via* a quinone-methide intermediate. In the proposed mechanism, kibelone C undergoes air oxidation to yield kibelone B, which undergoes an acid-mediated double-keto/enol tautomerization to yield an unstable quinone-methide intermediate, which can in turn undergo a third keto/enol tautomerization to give kibelone A hydroquinone and, upon air oxidation, is converted to kibelone A (Figure 12.18). The same equilibration is observed in the isokibelone series.

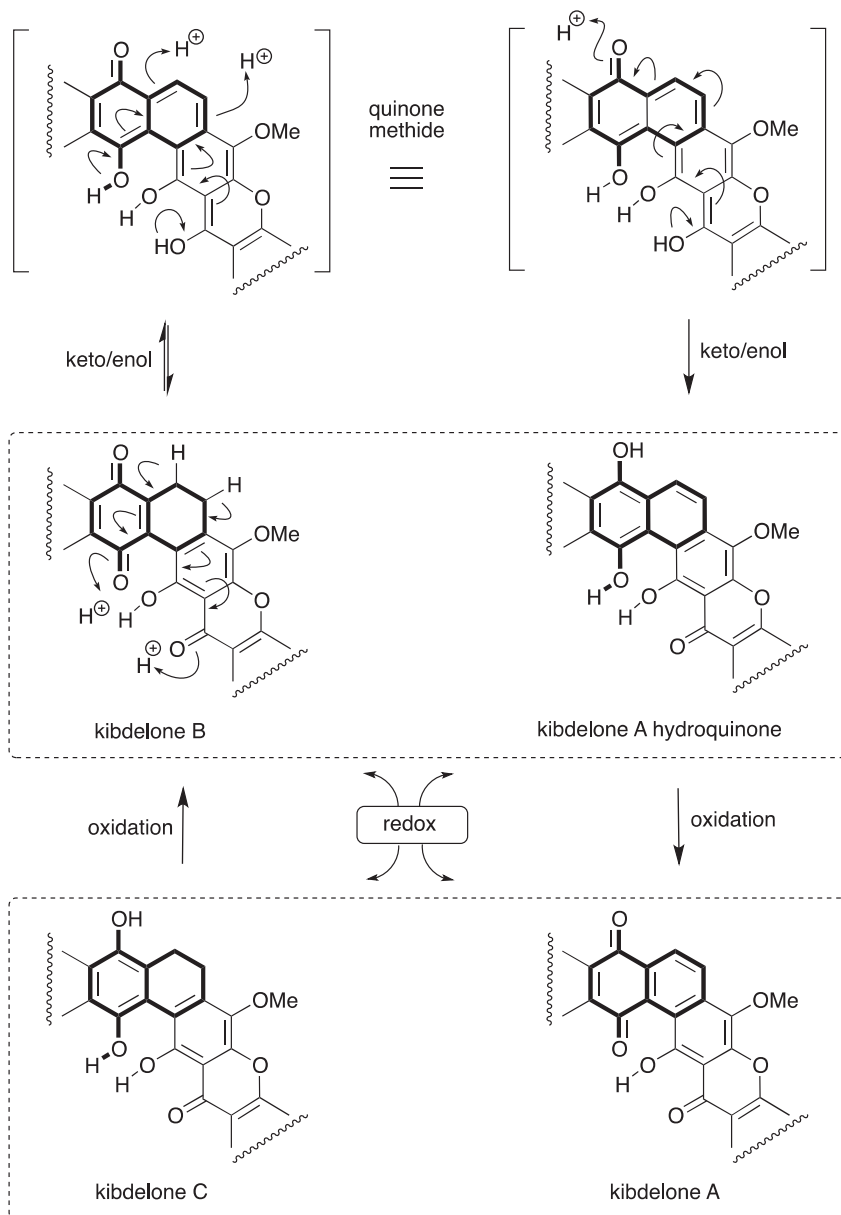


Figure 12.18 Plausible mechanism for the equilibration of kibelones A–C.

Kibelones exhibited potent and selective cytotoxicity against a panel of human tumor cell lines and displayed significant antibacterial and nematocidal activity. Isokibelones, on the other hand, were 10–200-times less cytotoxic than the kibelones, which gave an ideal opportunity to expand structure–activity relationship investigations into this unique family of polyketides.

12.2.5 Reveromycins

The polyketide spiroketal reveromycins A–D (Figure 12.19) were first described from a Japanese soil *Streptomyces* sp. (SN-593) in the early 1990s by researchers from the RIKEN Antibiotics Laboratory.^{8–11} More recently, chemical analysis of two Australian actinomycetes, the marine-derived *Streptomyces* sp. MST-MA568 and the terrestrial *Streptomyces* sp. MST-RA7781, revealed a complex mixture of more than 16 reveromycins.¹² The complexity of the crude extracts challenged and defied traditional isolation techniques, prompting the consideration of a HPLC–diode array detector–solid-phase extraction–NMR (HPLC-DAD-SPE-NMR) approach, involving: (i) 500 µg injections of each fraction through an analytical C₈ HPLC column; (ii) the collection of peaks of interest onto individual SPE cartridges; and (iii) the elution of metabolites from each SPE cartridge into the LC-NMR probe with a fixed aliquot of MeOH-*d*₄. This approach successfully identified the four known reveromycins, as well as 12 new reveromycin analogs (Figure 12.19), spanning homologs of the 6,6-spiroketal (*hemi*-succinate E, the 4'-methyl esters of A, D, and E, the *hemi*-fumarates H–J, and the *hemi*-furanoates K–M) and 5,6-spiroketal (*hemi*-succinates F and G) reveromycin motifs. The advantages of a HPLC-DAD-SPE-NMR strategy include: (i) the use of analytical HPLC to maximize resolution in the fractionation process; (ii) the use of *in situ* robotics to coordinate the handling of low-mass samples, minimizing contamination risk; and (iii) the use of a high-sensitivity, low-volume LC-NMR probe to maximize sample concentration and NMR data quality (see Figures 12.20 and 12.21 for examples).

During the HPLC purification and analysis of the reveromycins, the presence of a minor later-eluting “contaminant” was observed in each of the supposedly pure samples of the *hemi*-succinate reveromycins (A, C, D, and E), but not for any of the other reveromycins. The HPLC-DAD-SPE-NMR analysis established that the contaminants were in equilibrium with the parent reveromycins and were virtually identical in structure. It was hypothesized that the *hemi*-succinate moiety can undergo intramolecular cyclization to establish an equilibrium with a *ketal*-succinyl species (Figure 12.22). By contrast, the co-metabolite *hemi*-succinate 4'-methyl esters, *hemi*-fumarates, and *hemi*-furanoates are incapable of forming a *ketal*-succinyl and therefore do not equilibrate, explaining the lack of “contaminant” peaks during HPLC analysis. The *ketal*-succinyl is likely stabilized by a seven-membered hydrogen-bonded ring linking the *ketal*-succinyl 1'-OH and the 6,6-spiroketal (Figure 12.22), explaining the inability of 5,6-spiroketal reveromycin *hemi*-succinates to achieve a comparable equilibrium. This proposed *ketal*-succinyl equilibrium raises the possibility of a rearrangement pathway directly linking the natural 6,6-spiroketal reveromycin *hemi*-succinates to their 5,6-spiroketal isomers and brings into question the biosynthetic origins of the 5,6-spiroketal reveromycins.

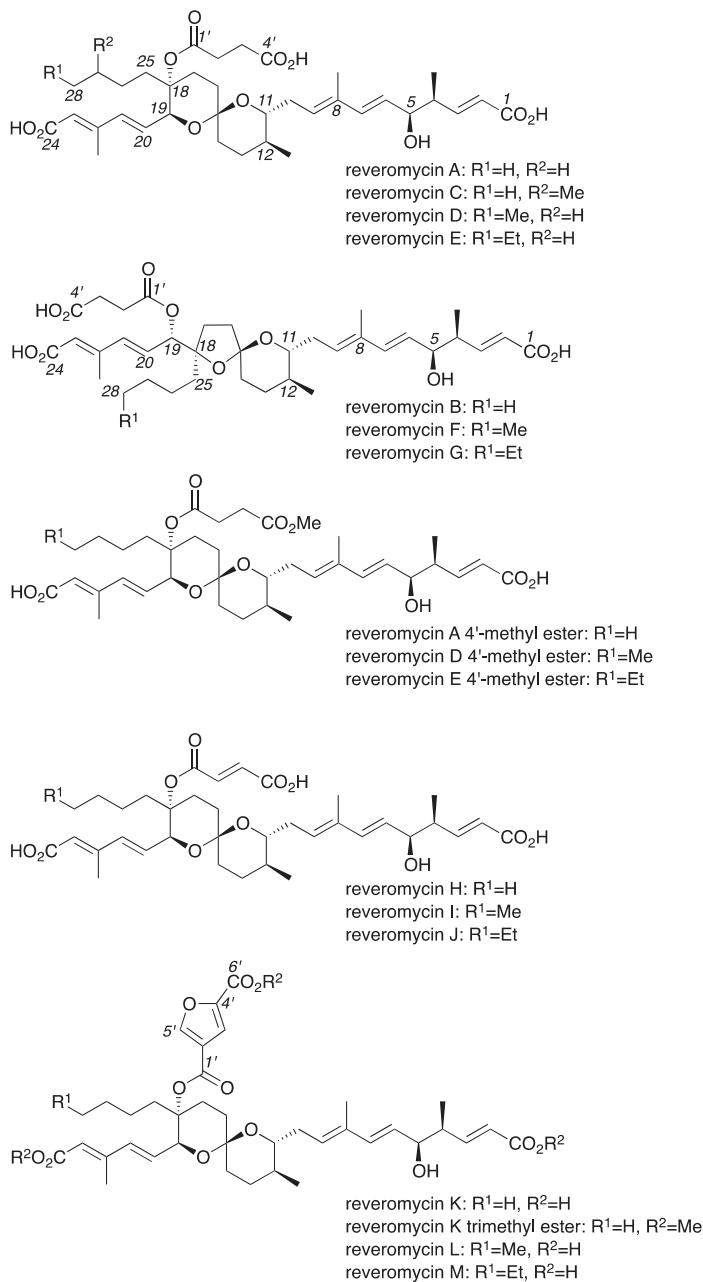


Figure 12.19 Reveromycins.

A plausible reveromycin polyketide biosynthesis (Figure 12.23) involves: (i) a 2-methylsuccinyl-CoA starter residue; (ii) flexibility to incorporate a selection of pre-formed C_{6-8} polyketide residues to account for the range

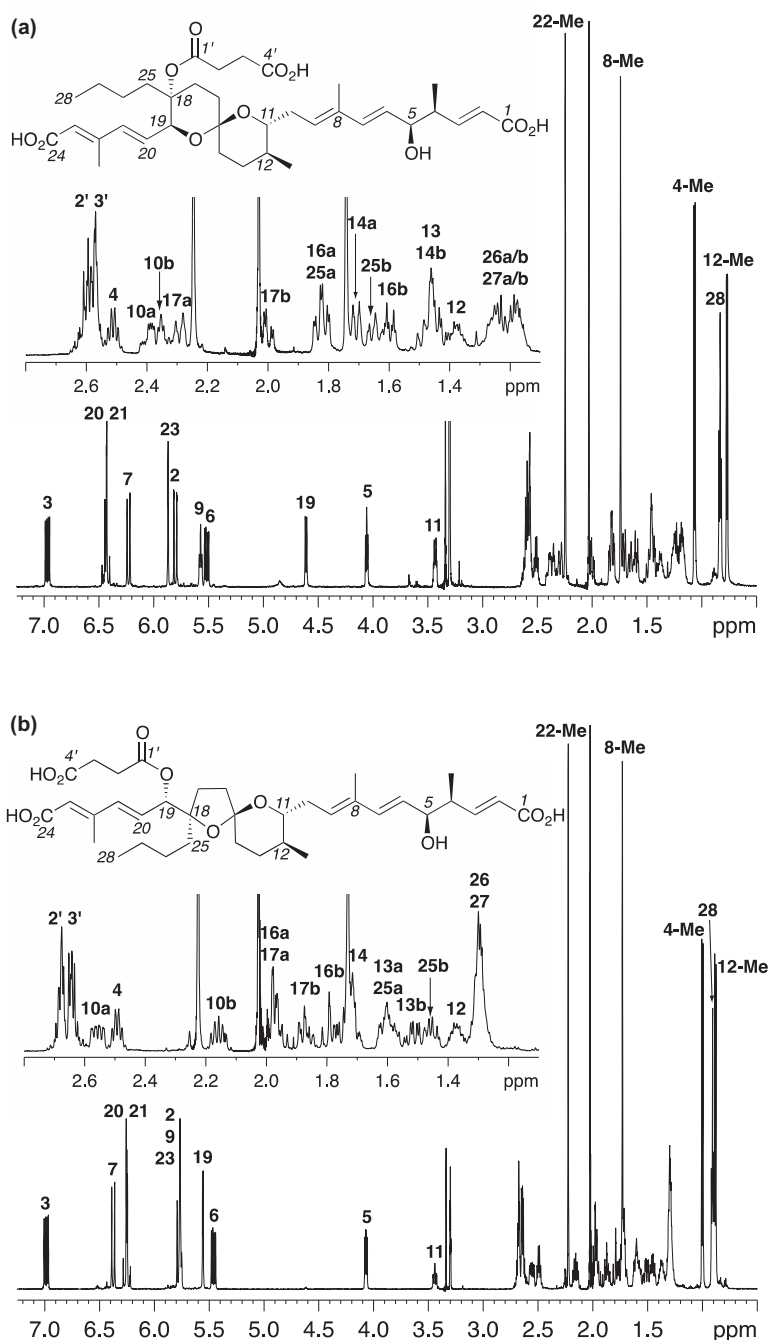


Figure 12.20 ^1H NMR spectra (600 MHz, CD_3OD) of (a) reveromycin A and (b) reveromycin B.

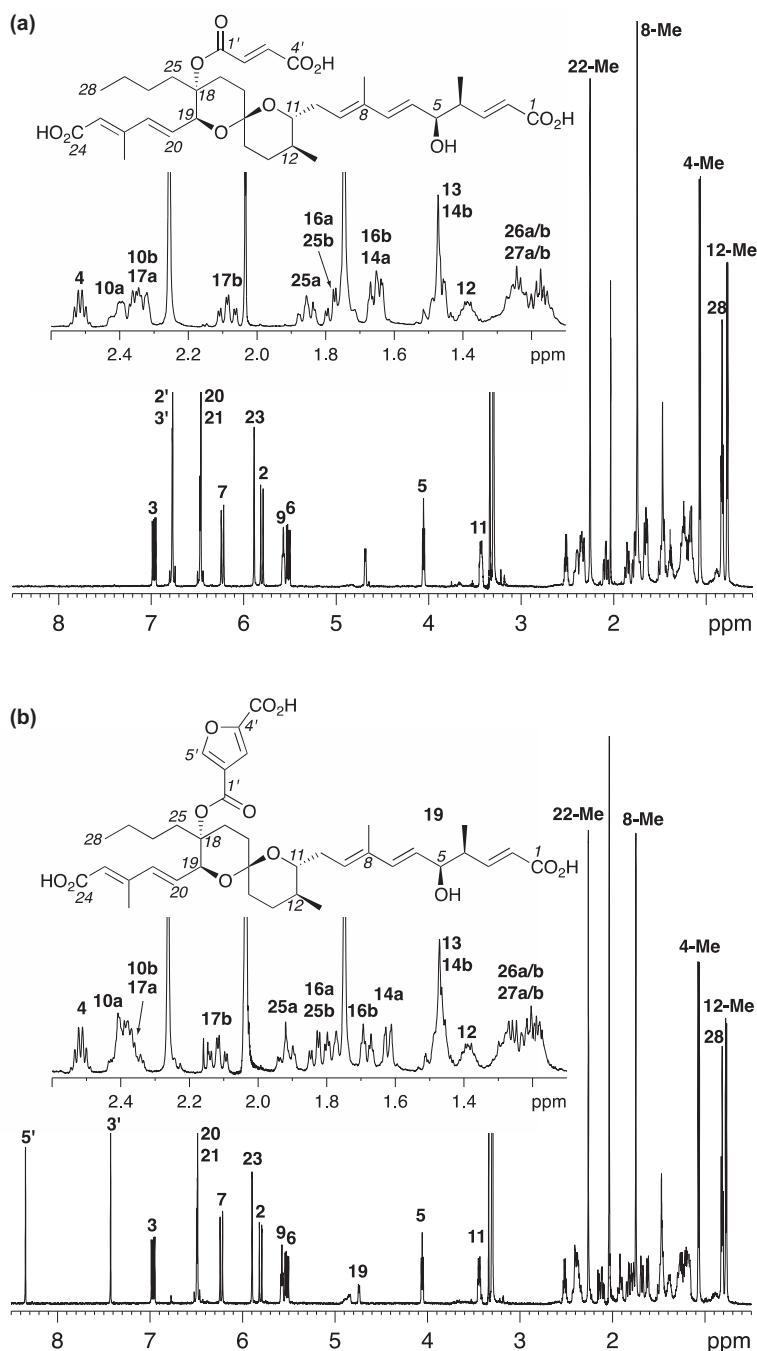


Figure 12.21 ^1H NMR spectra (600 MHz, CD_3OD) of (a) reveromycin H and (b) reveromycin K.

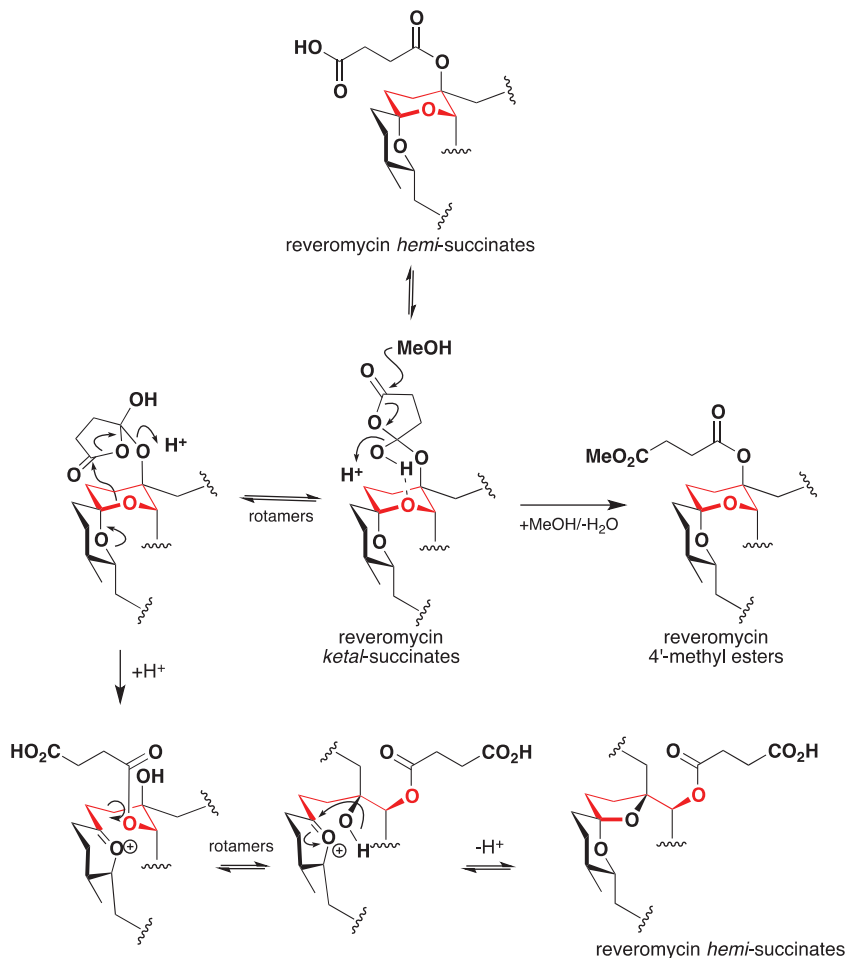


Figure 12.22 Hemi-succinyl–ketal-succinyl equilibrium as a mechanism for the formation of 4'-methyl ester and 5,6-spiroketal artifacts.

of C-18 side chains (*n*-butyl, *n*-pentyl, *n*-hexyl, and *n*-isohexyl); and (iii) coordinated but flexible delivery of a selection of succinate, fumarate, and furanoate moieties to account for the range of ester residues.

The reveromycins displayed modest cytotoxicity, with the 4'-methyl esters being notably cytotoxic to AGS (stomach) and HeLa (cervical) (IC₅₀ 2–10 μM) cell lines, but not HT29 (colon) and HFF-1 (fibroblast) (IC₅₀ > 30 μM) cell lines. Whilst not antibacterial, reveromycins A, C, and D have an *in vitro* MIC against *Candida albicans* that is superior to the commercial antifungal agent fluconazole at pH 3. The 4'-methyl esters displayed a two- to four-fold improvement in antifungal activity against *C. albicans*, *C. krusei*, and *C. parapsilosis* at pH 3, the hemi-furanoates showed a two- to four-fold decrease in antifungal activity, while the 5,6-spiroketals showed no appreciable antifungal activity.



12.3 Microbial Natural Product Snapshots

12.3.1 Macrocyclic Lactones

The 42-membered macrocyclic lactone desertomycin A (Figures 12.24 and 12.25) was first reported in 1958 from an African desert sand isolate of *Streptomyces flavofungini*.¹³ After repeated failed attempts at growing crystals suitable for X-ray diffraction studies, its structure was finally elucidated by spectroscopic analysis in 1986.¹⁴ Desertomycin A exhibits broad-spectrum antibiotic activity against bacteria and fungi and cytotoxicity against leukemia.

The 18-membered macrocyclic lactone fidaxomicin (Figures 12.26 and 12.27) was isolated from a soil isolate of *Dactylosporangium aurantiacum* in 1987 and its structure assigned by spectroscopic analysis.^{15,16} Fidaxomicin exhibits narrow-spectrum activity against Gram-positive anaerobic bacteria and was approved by the Food and Drug Administration in 2011 for the treatment of *Clostridium difficile*-associated diarrhea.

The 26-membered macrocyclic lactone oligomycin A (Figures 12.28 and 12.29) was first isolated in 1954 from a strain of *Streptomyces diastochromogenes* as a complex with two other closely related oligomycin analogs.¹⁷ The pure compound was recovered in 1958,¹⁸ and its structure was elucidated by spectroscopic analysis in the mid-1980s.^{19,20} Oligomycin A exhibits broad-spectrum antifungal activity, but its clinical use is limited due to its toxicity. The compound also inhibits ATPase activity by causing a conformational change in the F_0 portion of the complex, which is transmitted to F_1 and results in impaired substrate binding.²¹

12.3.2 Cyclic Peptides

The 28-membered cyclic depsipeptide telomycin (Figures 12.30 and 12.31) was first reported in 1958 from two strains of *Streptomyces canus*.²² Its structure was partially elucidated by chemical degradation in 1963,²³ and further refined by a combination of Edman degradation, synthesis, and spectroscopic analysis in 1968.²⁴ Telomycin exhibits potent and selective antibiotic activity against Gram-positive bacteria.²⁵

The cyclic hexapeptide desotamide (Figures 12.32 and 12.33) was isolated from a soil *Streptomyces* sp. and its structure assigned by a combination of spectroscopic analysis and chemical degradation and synthesis in 1997.²⁶ With no reported biological activity, desotamide is an example of the legions of known microbial metabolites that, over the last 60+ years, were isolated, identified, and (prematurely) confined to the virtual world that is the scientific literature. Such compounds represent a remarkable resource of untapped biological potential.

12.3.3 Caged Structures

The unprecedented dimeric thioether polyketide abyssomicin J (Figures 12.34 and 12.35), first reported in 2013 from a South China Sea deep-sea

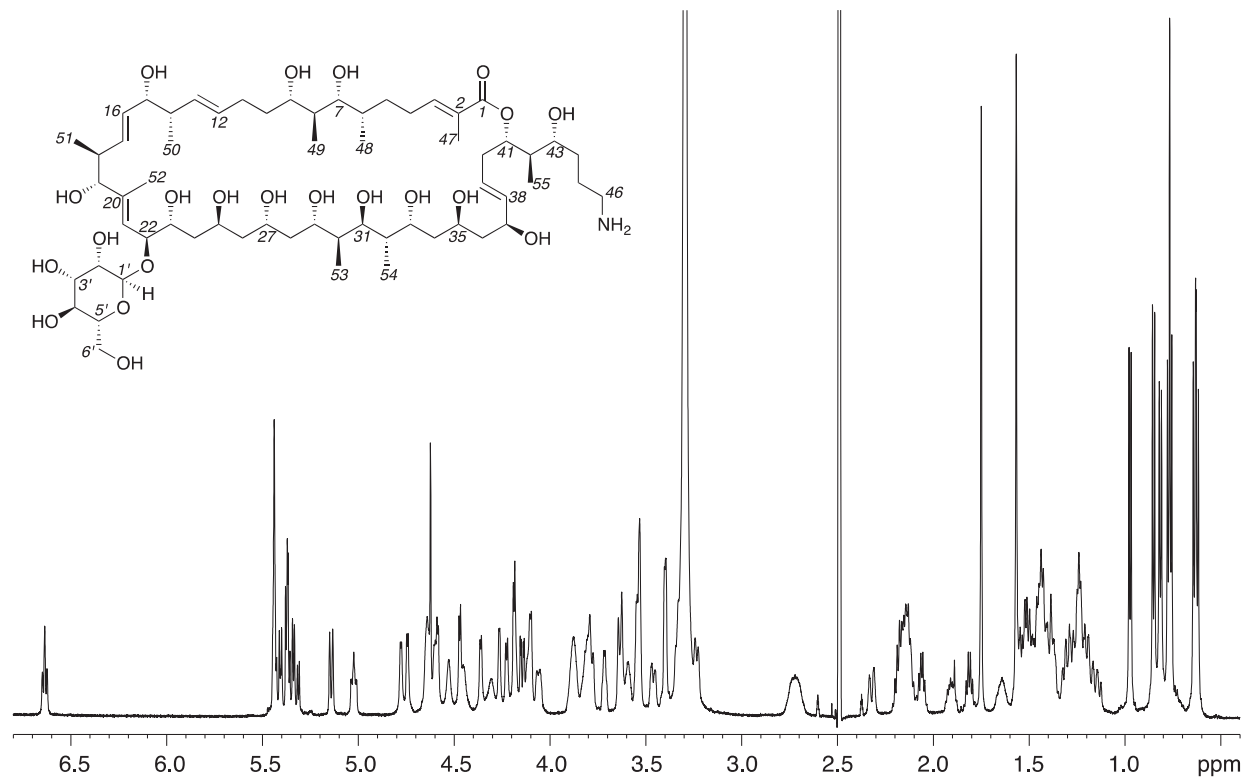


Figure 12.24 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of desertomycin A.

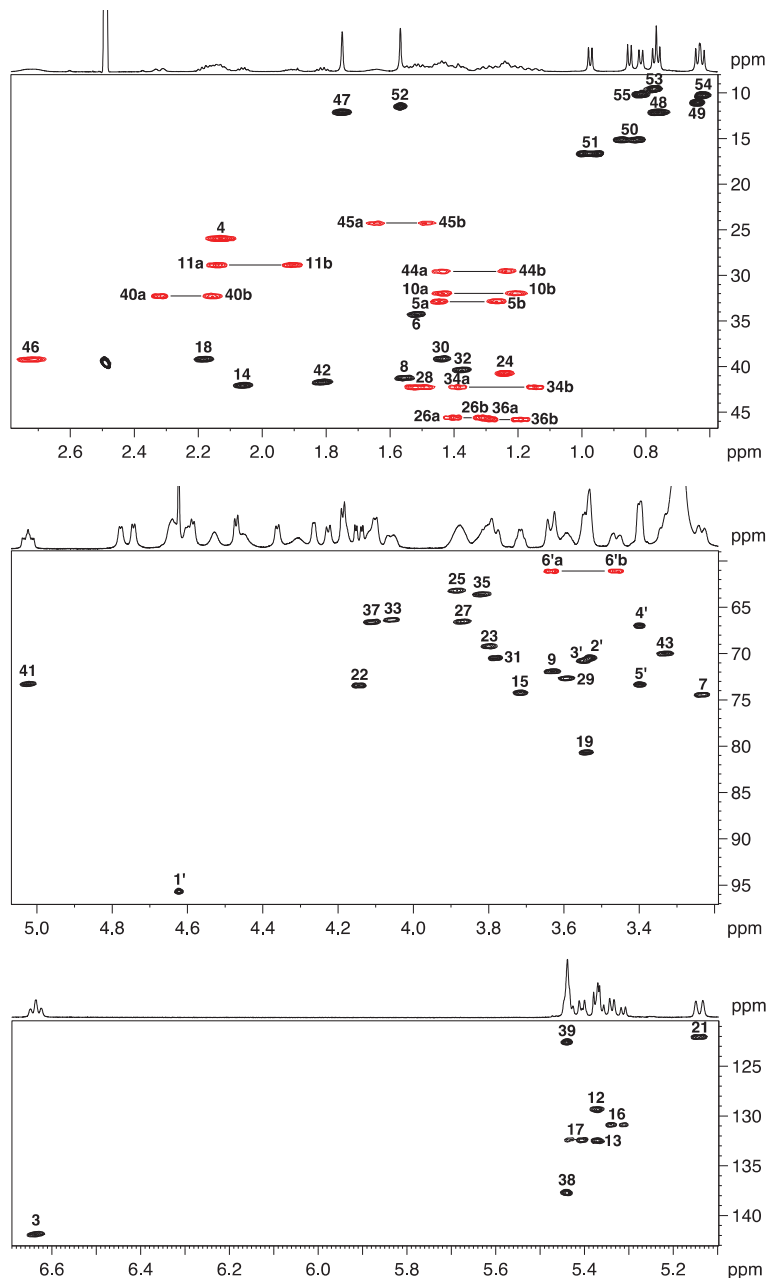


Figure 12.25 HSQC NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of desertomycin A.

Verrucosipora sp.,²⁷ is a unique example of a masked Michael acceptor and antitubercular prodrug. *In situ* enzymatic oxidative activation of abyssomicin J initiates a reverse Michael addition cascade that delivers the highly potent Michael acceptor *atrop*-abyssomicin C as an intracellular antitubercular agent.

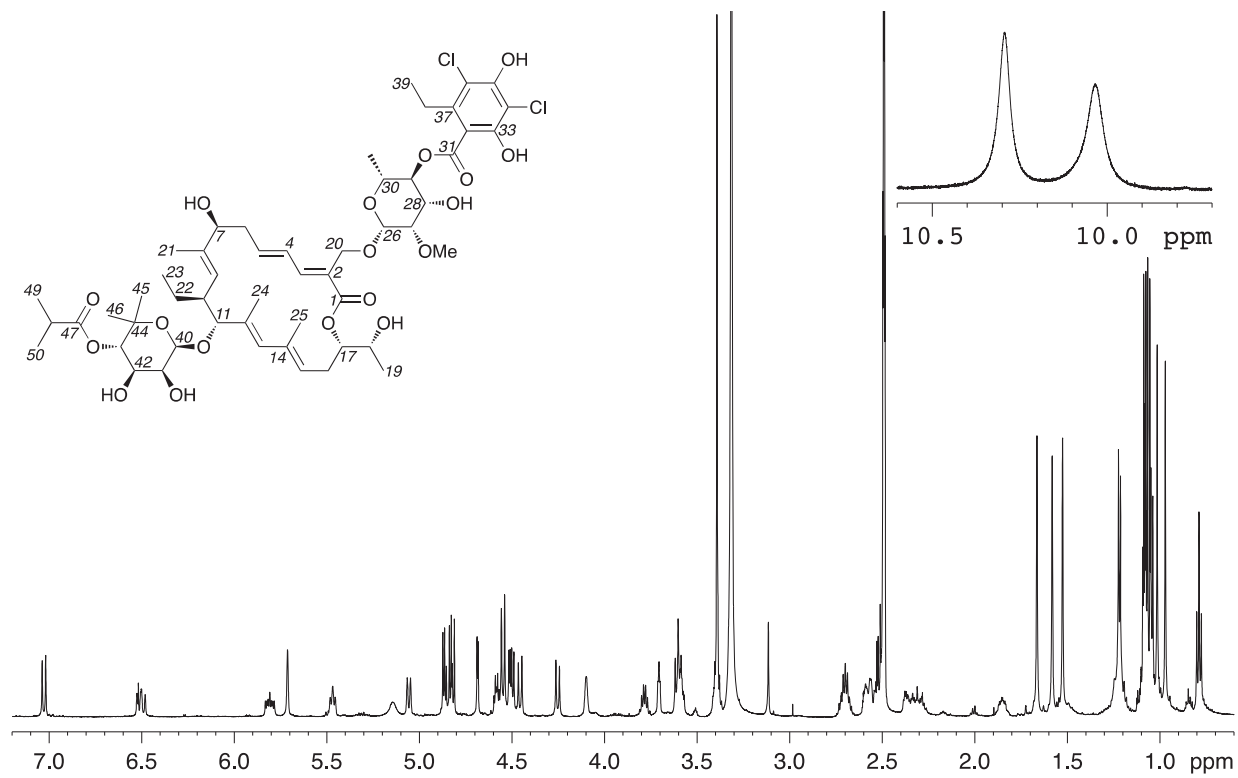


Figure 12.26 ^1H NMR (600 MHz, $\text{DMSO-}d_6$) spectrum of fidaxomicin.

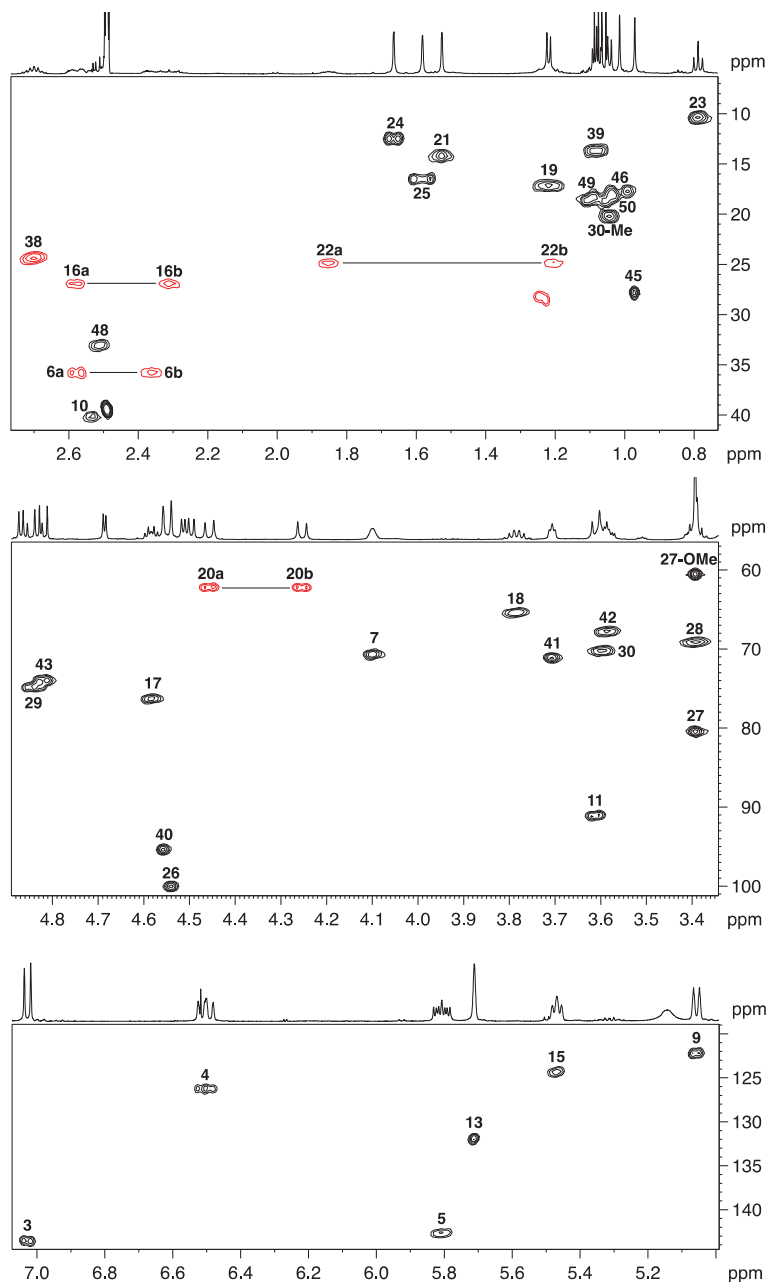


Figure 12.27 HSQC NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of fidaxomicin.

Platensimycin (Figure 12.35) and platencin (Figure 12.36) are unprecedented antibiotics that were first reported in 2006^{28,29} and 2007,³⁰ respectively, from a South African soil isolate of *Streptomyces platensis*. Both

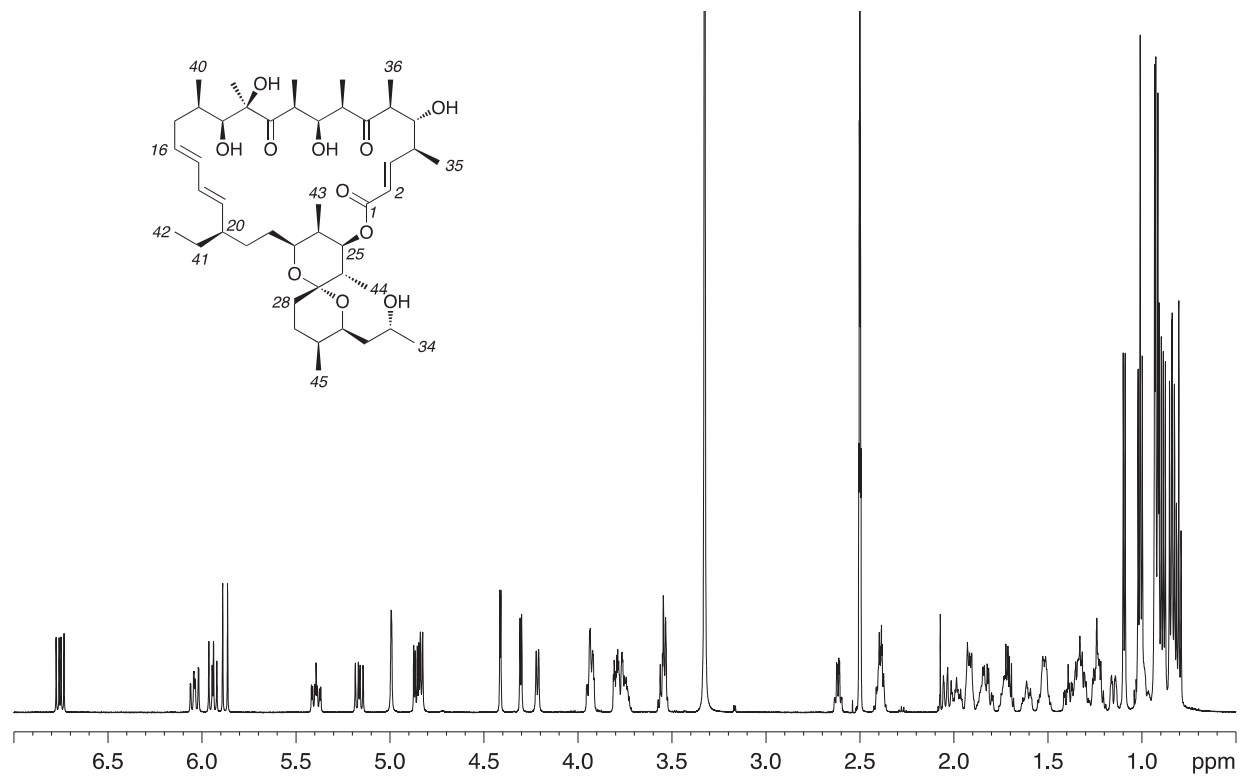


Figure 12.28 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of oligomycin A.

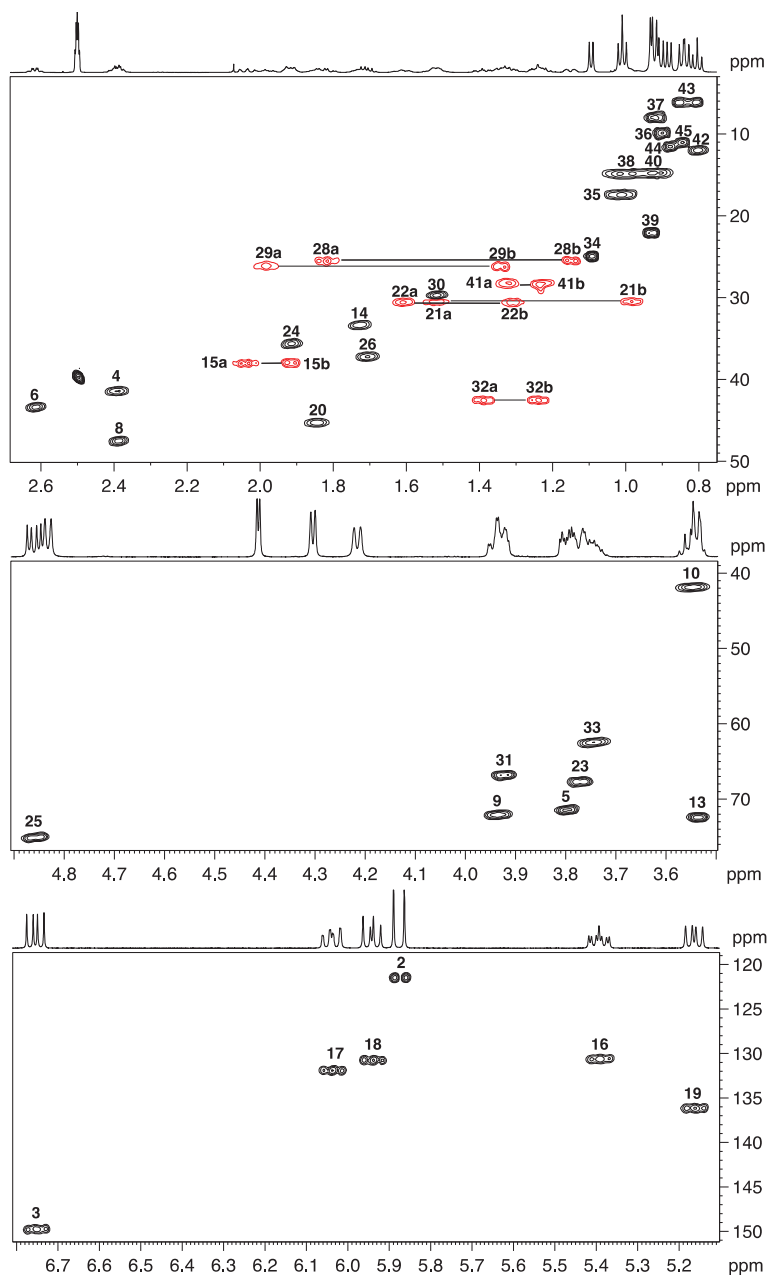


Figure 12.29 HSQC NMR (600 MHz, DMSO-*d*₆) spectrum of oligomycin A.

compounds consist of two distinct structural elements connected *via* an amide bond—a 3-amino-2,4-dihydroxybenzoic acid polar unit biosynthesized from pyruvate and acetate *via* the TCA cycle, and a lipophilic polycyclic

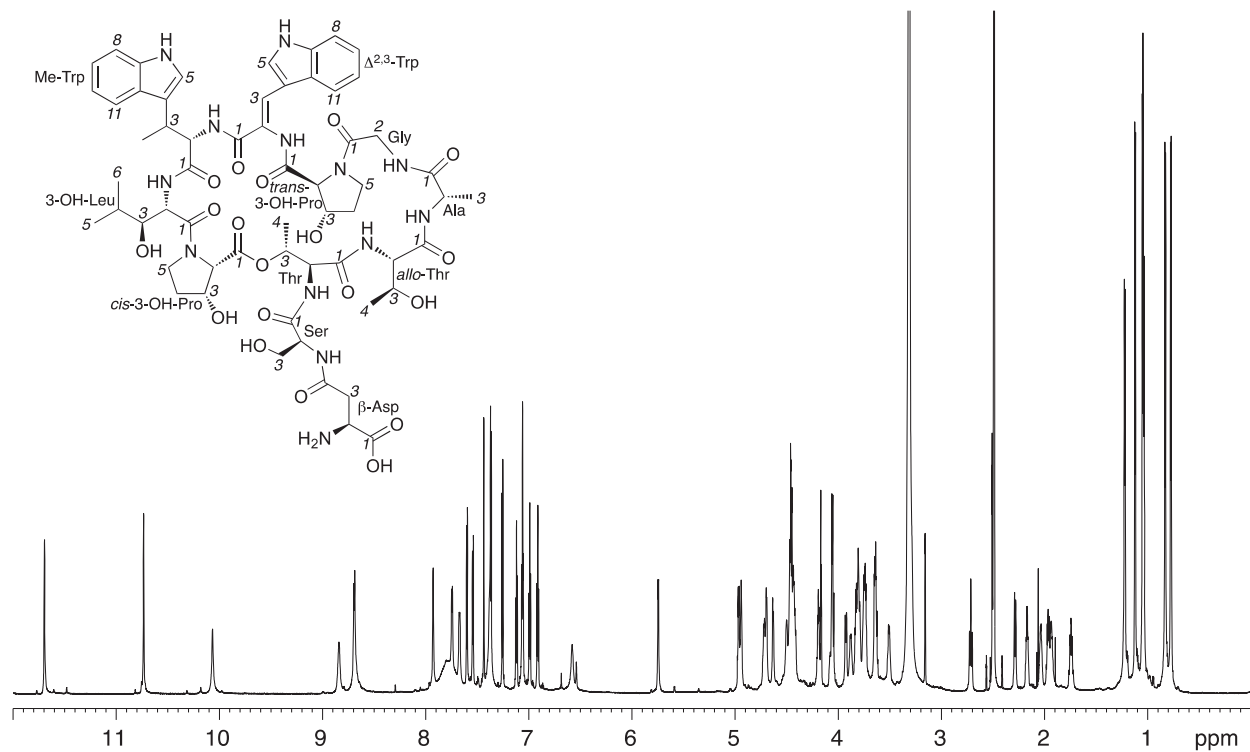


Figure 12.30 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of telomycin.

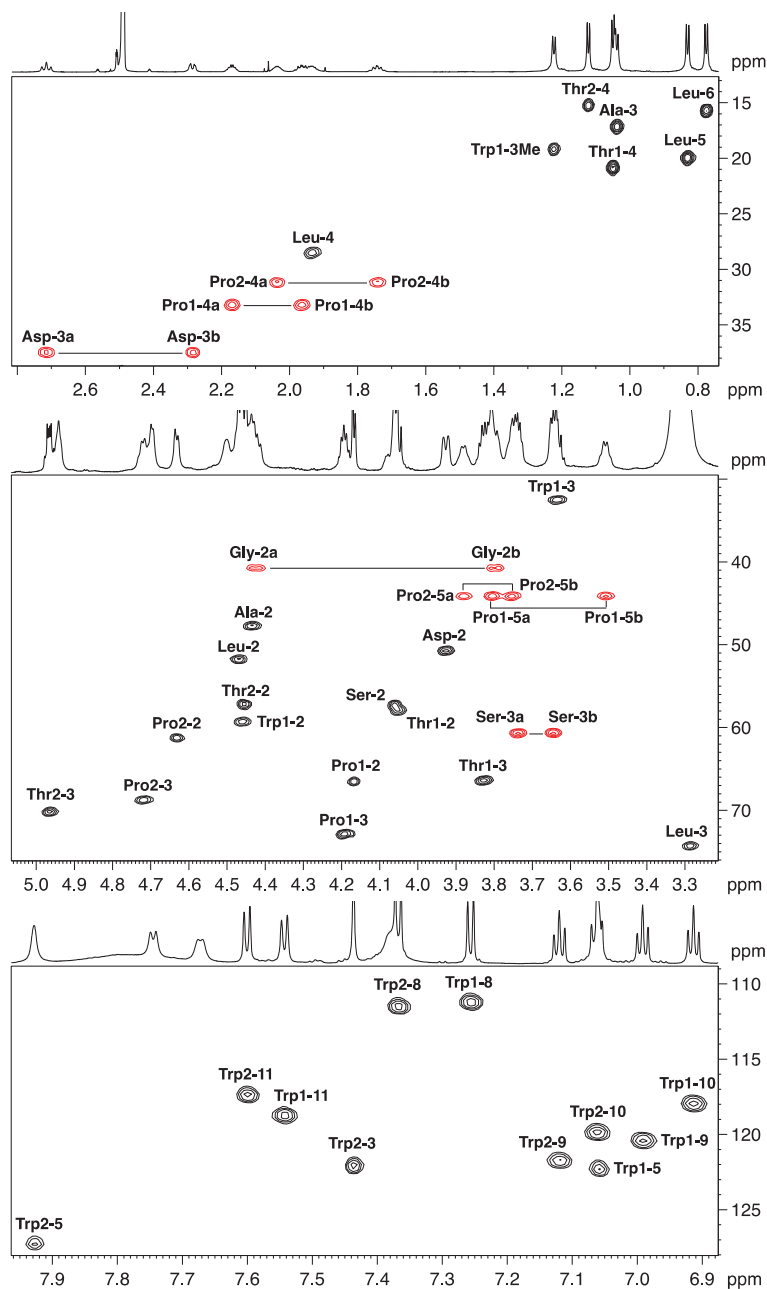


Figure 12.31 HSQC NMR (600 MHz, DMSO- d_6) spectrum of telomycin.

ketolide unit biosynthesized from the non-mevalonate terpenoid pathway. Platensimycin and platencin exhibit broad-spectrum Gram-positive antibacterial activity through the inhibition of fatty acid biosynthesis.³¹

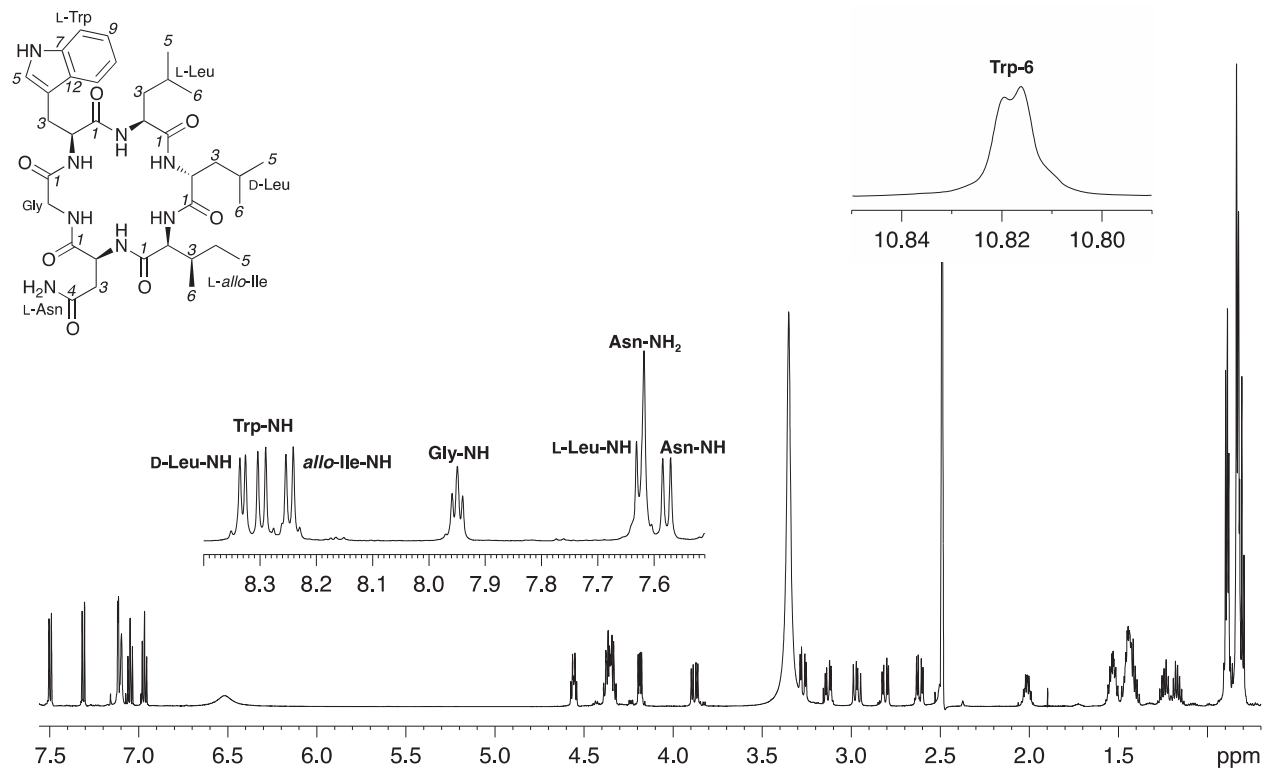


Figure 12.32 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of desotamide.

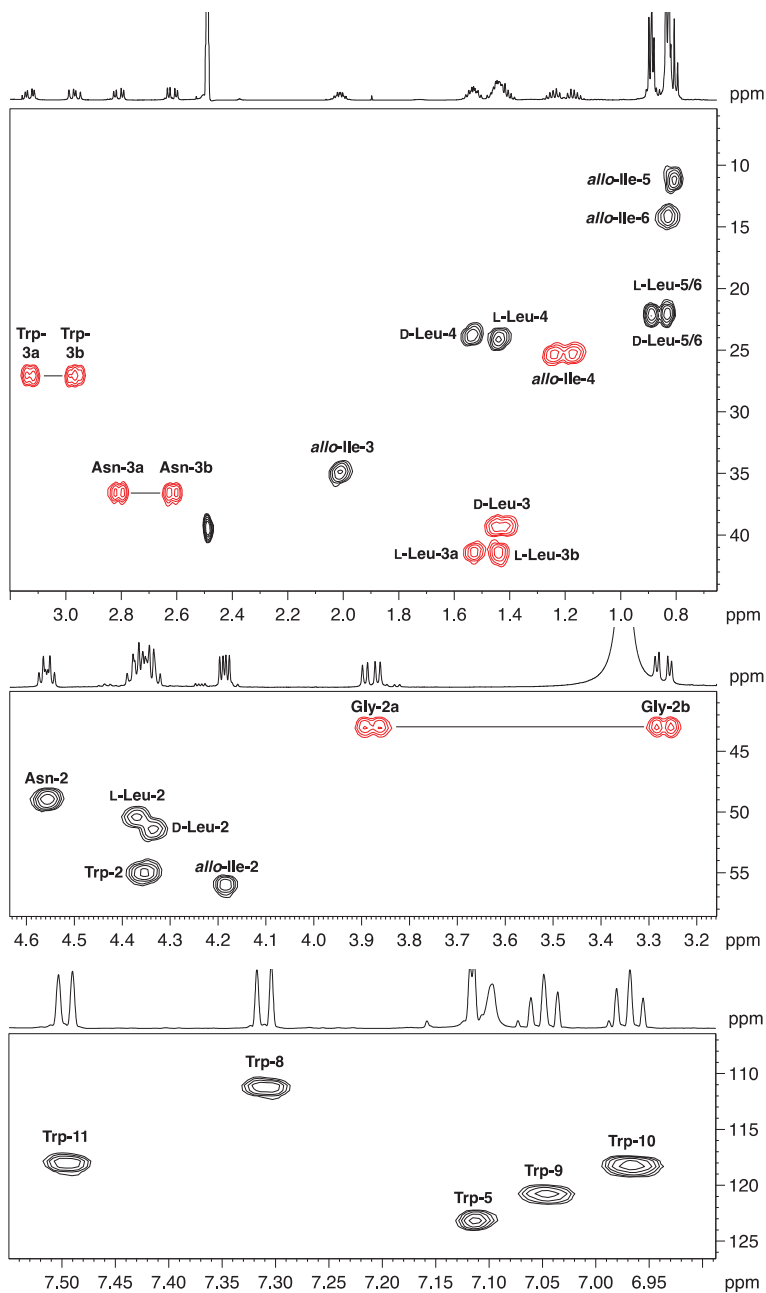


Figure 12.33 HSQC NMR (600 MHz, DMSO- d_6) spectrum of desotamide.

Enterocin (Figure 12.37) and 5-deoxyenterocin (Figure 12.38) are polyketide antibiotics that were first reported in 1976^{32,33} and 1987,³⁴ respectively, from soil isolates of *Streptomyces candidus* and *Streptomyces*

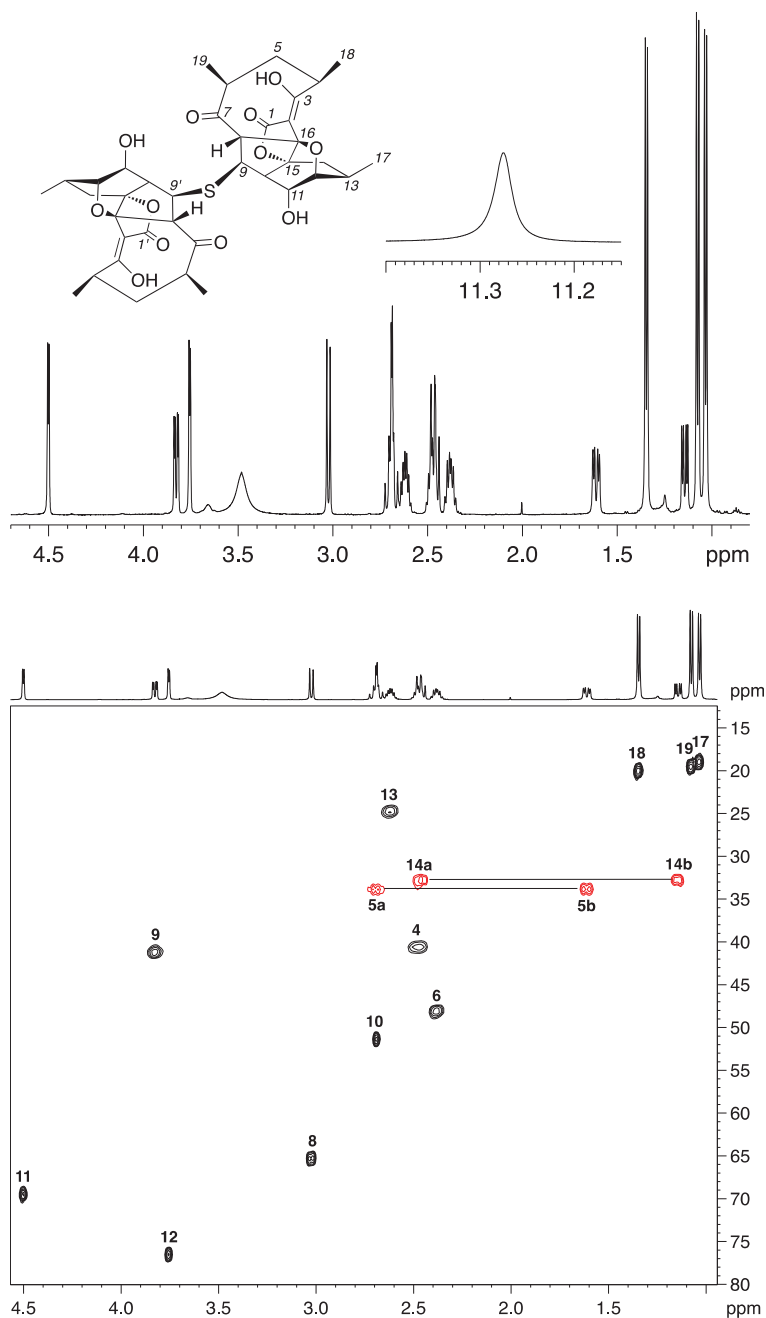


Figure 12.34 ^1H NMR and HSQC (600 MHz, $\text{DMSO}-d_6$) spectra of abyssomicin J.

viridochromogenes. Both compounds are polyketides derived from an uncommon benzoate starter and seven malonate units,³⁵ and both exhibit broad-spectrum bacteriostatic activity with low acute toxicity.

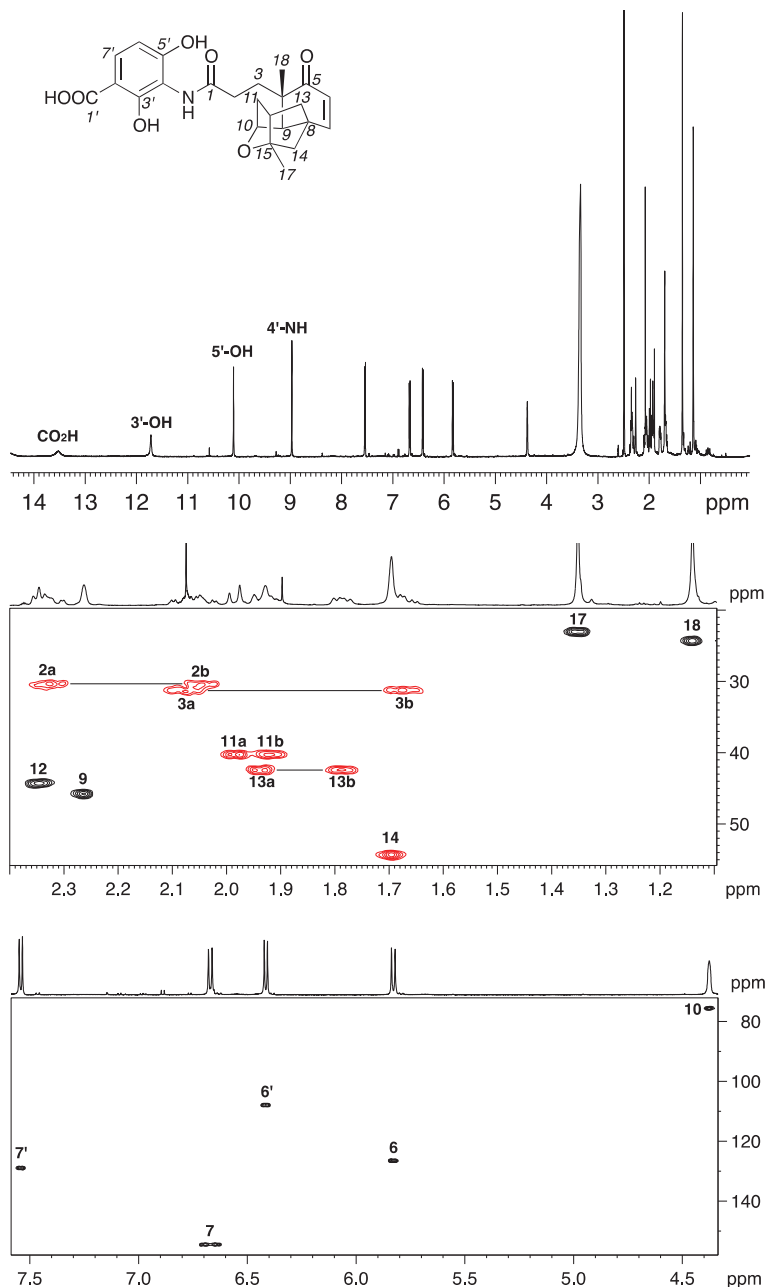


Figure 12.35 ^1H NMR and HSQC (600 MHz, $\text{DMSO}-d_6$) spectra of platensimycin.

12.3.4 Miscellaneous

Phenelfamycin E (Figures 12.39 and 12.40) is an elfamycin-type antibiotic that was first reported in 1988 from soil isolates of *Streptomyces violaceoniger*^{36,37} that exhibits activity against Gram-positive anaerobic microorganisms, including *C. difficile*.³⁸

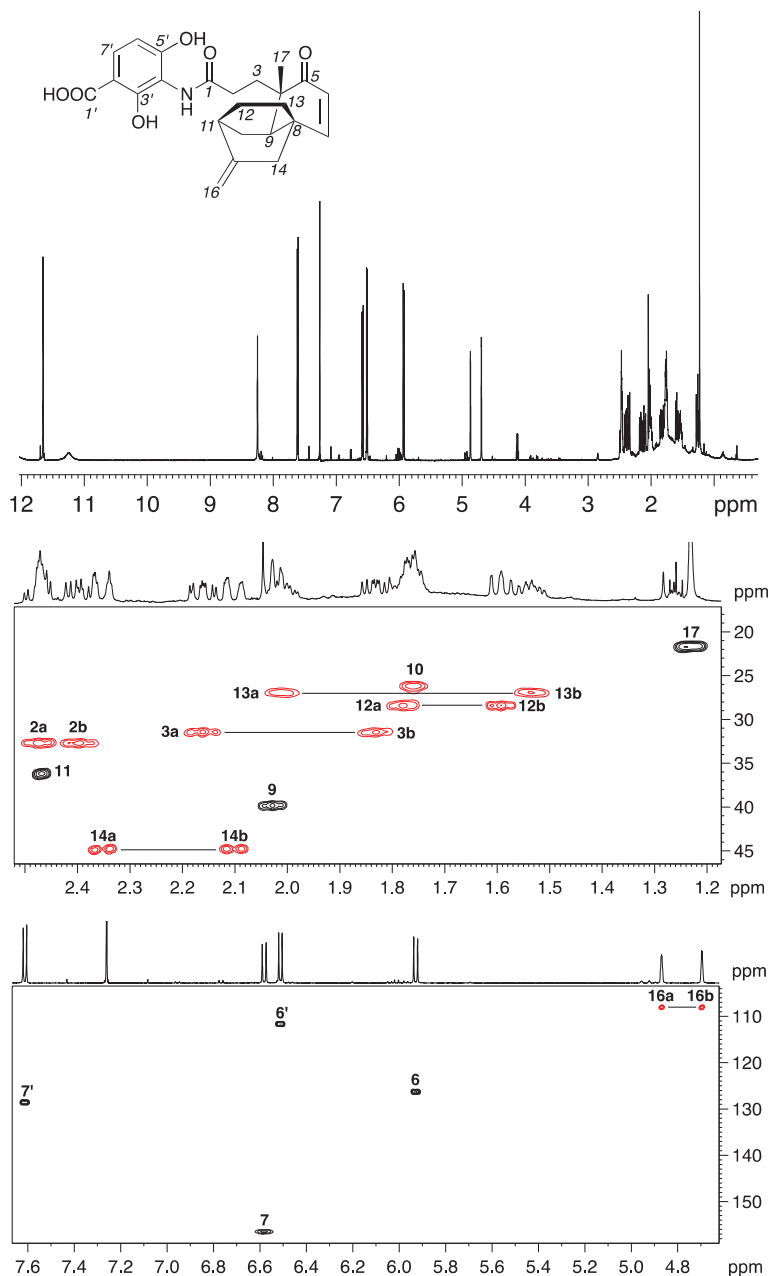


Figure 12.36 ^1H NMR and HSQC (600 MHz, $\text{DMSO}-d_6$) spectra of platensin.

The macrocyclic spiro-tetrone acid chlorothricin (Figures 12.41 and 12.42) that was first reported in 1969 from an isolate of *Streptomyces antibioticus*^{39,40} is active against Gram-positive bacteria through the inhibition of

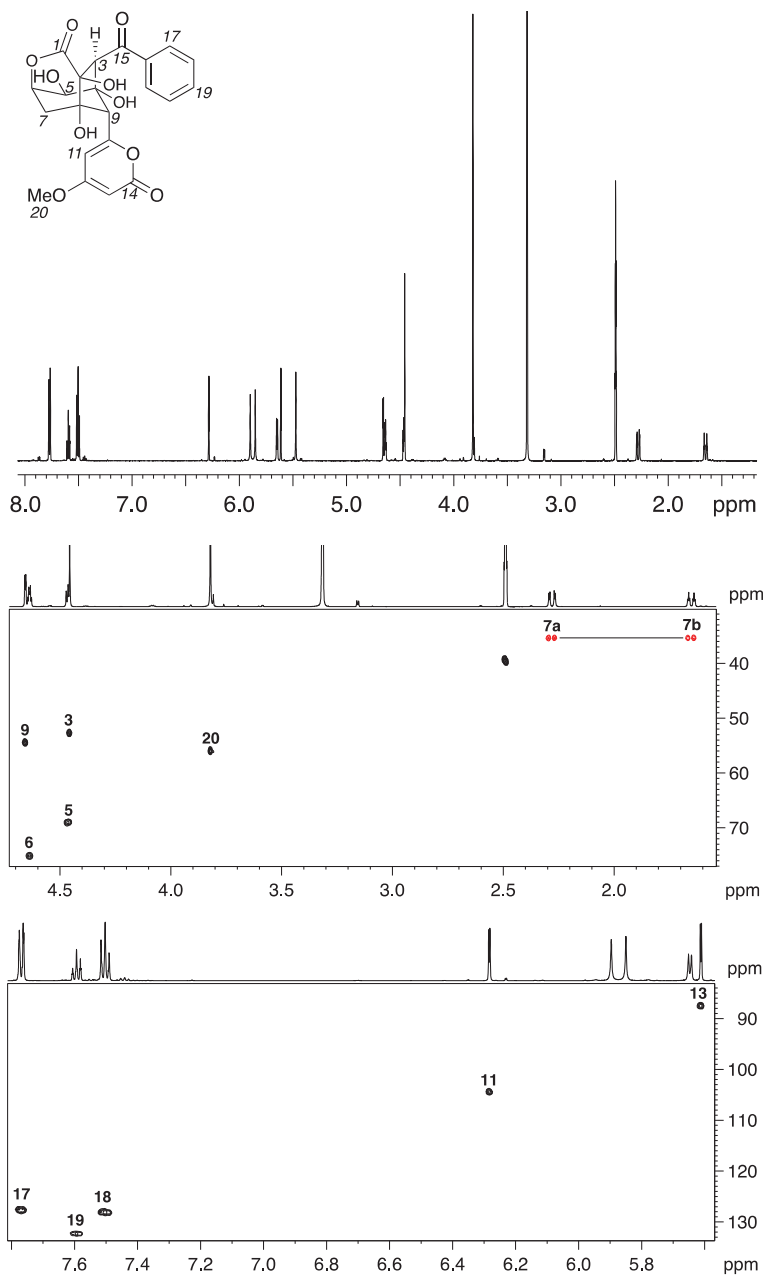


Figure 12.37 ^1H NMR and HSQC (600 MHz, $\text{DMSO}-d_6$) spectra of enterocin.

pyruvate carboxylase,^{41–43} and also inhibits both mitochondrial and cytoplasmic pig heart malate dehydrogenase.⁴⁴ The biosynthesis of chlorothricin proceeds *via* the polyketide spirotetronic acid aglycone, which is glycosylated

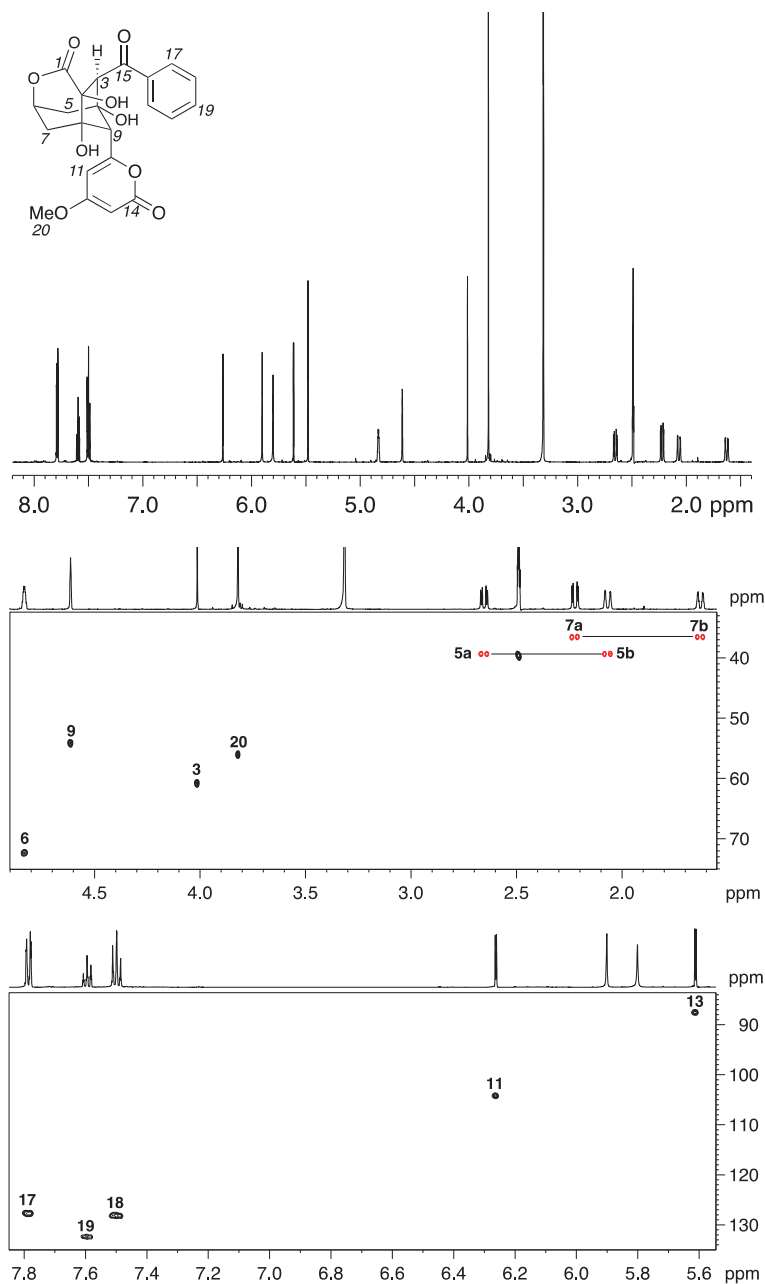


Figure 12.38 ^1H NMR and HSQC (600 MHz, $\text{DMSO}-d_6$) spectra of 5-deoxyenterocin.

with two D-olivose units. The second D-olivose unit is then O-acylated at C-3' with a 5-chloro-6-methyl-O-methylsalicylic acid residue by an iterative type I polyketide synthase.^{45,46}

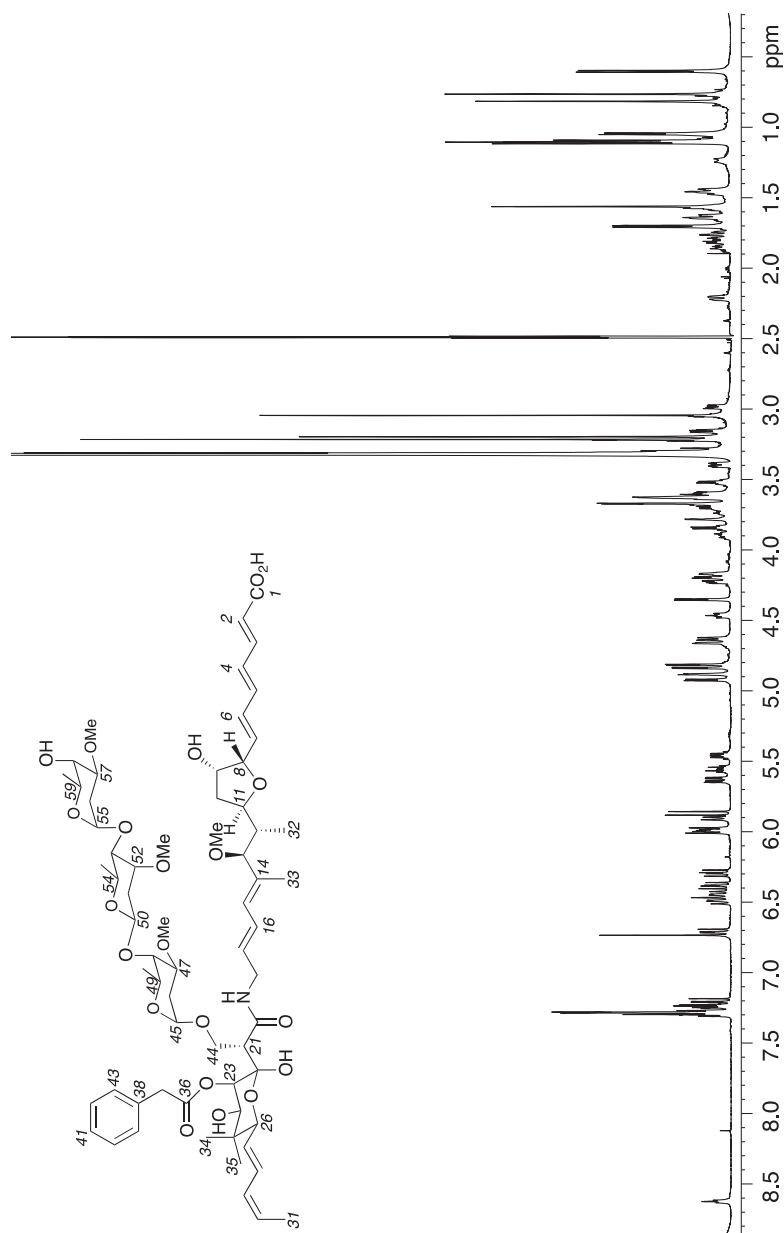


Figure 12.39 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of phenelfamycin E.

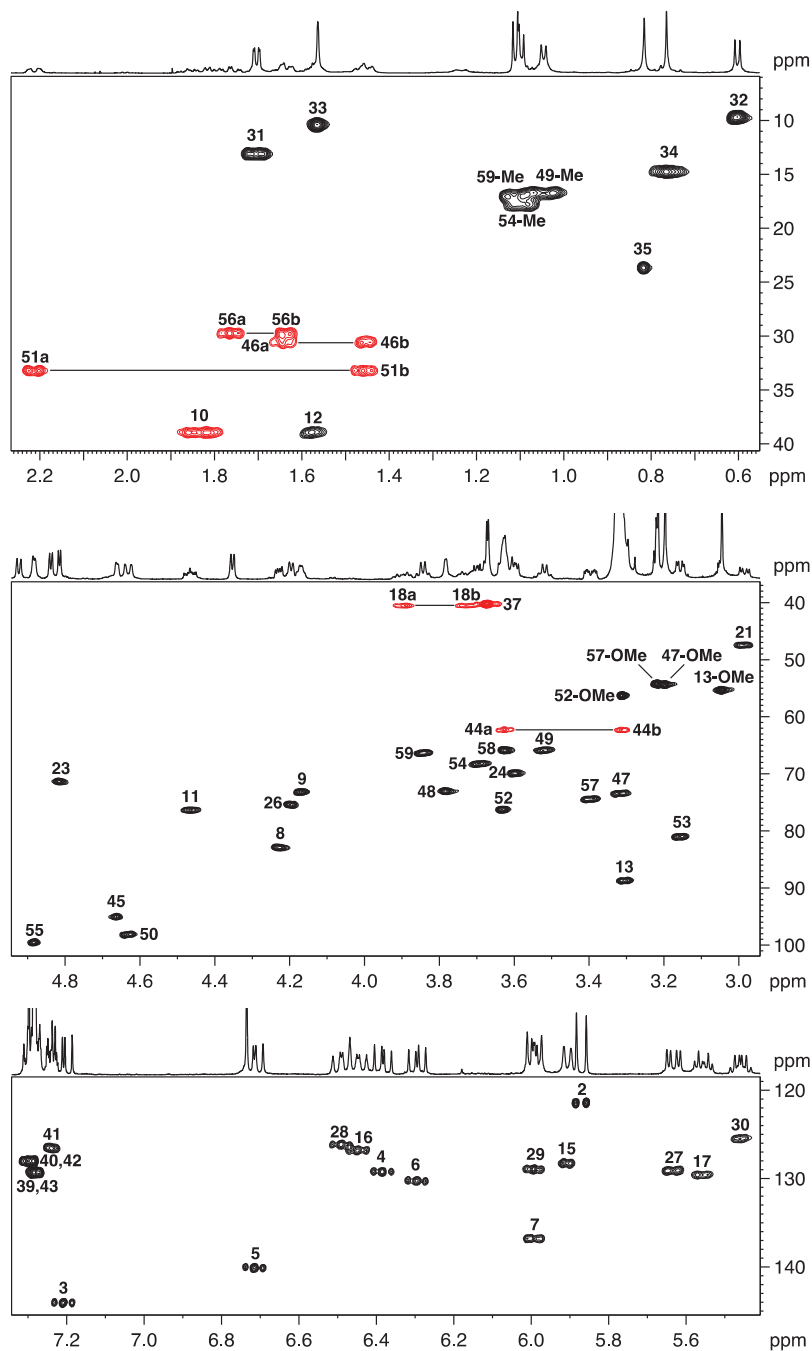


Figure 12.40 HSQC NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of phenelfamycin E.



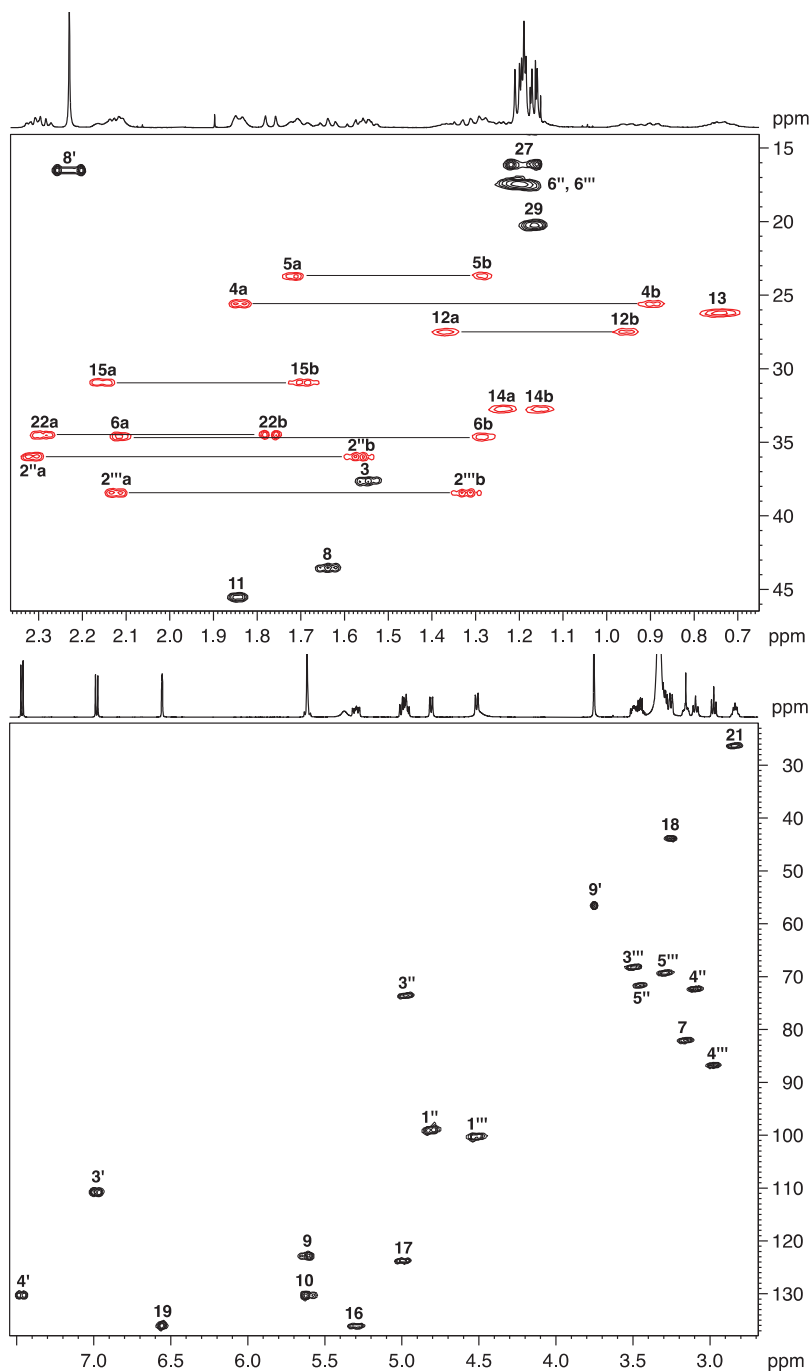


Figure 12.42 HSQC NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of chlorothricin.

12.4 Conclusion

The study of microbial natural products continues to inform our understanding of the natural world, while simultaneously delivering valuable new knowledge and molecular probes that inspire the development of new drugs, which in turn underpin advances in human and animal health, agriculture, and many fields of science. Key to this cascade of knowledge delivering past, current, and future benefits is the enabling technology of NMR spectroscopy. NMR spectroscopy is the premier approach used to characterize and identify microbial natural products. Two take-home messages that might be drawn from this chapter are: (i) time invested in the acquisition and careful analysis of high-quality NMR data is well spent; and (ii) even the most basic 1D and 2D NMR experiments can be remarkably informative.

References

1. R. J. Capon, R. Ratnayake, M. Stewart, E. Lacey, S. Tennant and J. H. Gill, *Org. Biomol. Chem.*, 2005, **3**, 123.
2. Y. Lin, Z. Shao, G. Jiang, S. Zhou, J. Cai, L. L. P. Vrijmoed and E. B. G. Jones, *Tetrahedron*, 2000, **56**, 9607.
3. E. Garo, C. M. Starks, P. R. Jensen, W. Fenical, E. Lobkovsky and J. Clardy, *J. Nat. Prod.*, 2003, **66**, 423.
4. R. Raju, A. M. Piggott, X.-C. Huang and R. J. Capon, *Org. Lett.*, 2011, **13**, 2770.
5. R. Raju, A. M. Piggott, M. M. Conte and R. J. Capon, *Org. Biomol. Chem.*, 2010, **8**, 4682.
6. R. Ratnayake, E. Lacey, S. Tennant, J. H. Gill and R. J. Capon, *Chem. – Eur. J.*, 2007, **13**, 1610.
7. R. Ratnayake, E. Lacey, S. Tennant, J. H. Gill and R. J. Capon, *Org. Lett.*, 2006, **8**, 5267.
8. H. Osada, H. Koshino, K. Isono, H. Takahashi and G. Kawanishi, *J. Antibiot.*, 1991, **44**, 259.
9. H. Takahashi, H. Osada, H. Koshino, T. Kudo, S. Amano, S. Shimizu, M. Yoshihama and K. Isono, *J. Antibiot.*, 1992, **45**, 1409.
10. H. Takahashi, H. Osada, H. Koshino, M. Sasaki, R. Onose, M. Nakakoshi, M. Yoshihama and K. Isono, *J. Antibiot.*, 1992, **45**, 1414.
11. H. Koshino, H. Takahashi, H. Osada and K. Isono, *J. Antibiot.*, 1992, **45**, 1420.
12. L. Fremlin, M. Farrugia, A. M. Piggott, Z. Khalil, E. Lacey and R. J. Capon, *Org. Biomol. Chem.*, 2011, **9**, 1201.
13. J. Uri, R. Bogner, I. Bekesi and B. Varga, *Nature*, 1958, **182**, 401.
14. A. Bax, A. Aszalos, Z. Dinya and K. Sudo, *J. Am. Chem. Soc.*, 1986, **108**, 8056.
15. R. J. Theriault, J. P. Karwowski, M. Jackson, R. L. Girolami, G. N. Sunga, C. M. Vojtko and L. J. Coen, *J. Antibiot.*, 1987, **40**, 567.

16. J. E. Hochlowski, S. J. Swanson, L. M. Renfranz, D. N. Whittern, A. M. Buko and J. B. McAlpine, *J. Antibiot.*, 1987, **40**, 575.
17. R. M. Smith, W. H. Peterson and E. McCoy, *Antibiot. Chemother.*, 1954, **4**, 962.
18. S. Masamune, J. M. Sehgal, T. E. E. Van, F. M. Strong and W. H. Peterson, *J. Am. Chem. Soc.*, 1958, **80**, 6092.
19. G. A. Morris and M. S. Richards, *Magn. Reson. Chem.*, 1985, **23**, 676.
20. G. T. Carter, *J. Org. Chem.*, 1986, **51**, 4264.
21. H. S. Penefsky, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**, 1589.
22. M. Misiek, O. B. Fardig, A. Gourevitch, D. L. Johnson, I. R. Hooper and J. Lein, *Antibiot. Annu.*, 1958, 852.
23. J. C. Sheehan, P. E. Drummond, J. N. Gardner, K. Maeda, D. Mania, S. Nakamura, A. K. Sen and J. A. Stock, *J. Am. Chem. Soc.*, 1963, **85**, 2867.
24. J. C. Sheehan, D. Mania, S. Nakamura, J. A. Stock and K. Maeda, *J. Am. Chem. Soc.*, 1968, **90**, 462.
25. D. E. Tisch, J. B. Huftalen and H. L. Dickison, *Antibiot. Annu.*, 1958, 863.
26. S. Miao, M. R. Anstee, K. LaMarco, J. Matthew, L. H. T. Huang and M. M. Brasseur, *J. Nat. Prod.*, 1997, **60**, 858.
27. Q. Wang, F. Song, X. Xiao, P. Huang, L. Li, A. Monte, W. M. Abdel-Mageed, J. Wang, H. Guo, W. He, F. Xie, H. Dai, M. Liu, C. Chen, H. Xu, M. Liu, A. M. Piggott, X. Liu, R. J. Capon and L. Zhang, *Angew. Chem., Int. Ed.*, 2013, **52**, 1231.
28. J. Wang, S. M. Soisson, K. Young, W. Shoop, S. Kodali, A. Galgoci, R. Painter, G. Parthasarathy, Y. S. Tang, R. Cummings, S. Ha, K. Dorso, M. Motyl, H. Jayasuriya, J. Ondeyka, K. Herath, C. Zhang, L. Hernandez, J. Allocco, A. Basilio, J. R. Tormo, O. Genilloud, F. Vicente, F. Pelaez, L. Colwell, S. H. Lee, B. Michael, T. Felcetto, C. Gill, L. L. Silver, J. D. Hermes, K. Bartizal, J. Barrett, D. Schmatz, J. W. Becker, D. Cully and S. B. Singh, *Nature*, 2006, **441**, 358.
29. S. B. Singh, H. Jayasuriya, J. G. Ondeyka, K. B. Herath, C. Zhang, D. L. Zink, N. N. Tsou, R. G. Ball, A. Basilio, O. Genilloud, M. T. Diez, F. Vicente, F. Pelaez, K. Young and J. Wang, *J. Am. Chem. Soc.*, 2006, **128**, 11916.
30. J. Wang, S. Kodali, S. H. Lee, A. Galgoci, R. Painter, K. Dorso, F. Racine, M. Motyl, L. Hernandez, E. Tinney, S. L. Colletti, K. Herath, R. Cummings, O. Salazar, I. Gonzalez, A. Basilio, F. Vicente, O. Genilloud, F. Pelaez, H. Jayasuriya, K. Young, D. F. Cully and S. B. Singh, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 7612.
31. K. B. Herath, A. B. Attygalle and S. B. Singh, *J. Am. Chem. Soc.*, 2007, **129**, 15422.
32. N. Miyairi, H. Sakai, T. Konomi and H. Imanaka, *J. Antibiot.*, 1976, **29**, 227.
33. Y. Tokuma, N. Miyairi and Y. Morimoto, *J. Antibiot.*, 1976, **29**, 1114.
34. S. Yaginuma, T. Asahi, N. Muto and M. Hayashi, Novel antibiotic 5-deoxyenterocin production by *Streptomyces* species and its antimicrobial activities, 1987, Toyo Jozo Co., Ltd., Japan, JP62056487A.

35. J. Piel, K. Hoang and B. S. Moore, *J. Am. Chem. Soc.*, 2000, **122**, 5415.
36. M. Jackson, J. P. Karwowski, R. J. Theriault, M. L. Maus and W. L. Kohl, *J. Antibiot.*, 1988, **41**, 1293.
37. J. E. Hochlowski, M. H. Buytendorp, D. N. Whittern, A. M. Buko, R. H. Chen and J. B. McAlpine, *J. Antibiot.*, 1988, **41**, 1300.
38. R. N. Swanson, D. J. Hardy, N. L. Shipkowitz, C. W. Hanson, N. R. Ramer, L. J. Coen and P. B. Fernandes, *J. Antibiot.*, 1989, **42**, 94.
39. W. Keller-Schierlein, R. Muntwyler, W. Pache and H. Zaehner, *Helv. Chim. Acta*, 1969, **52**, 127.
40. R. Muntwyler and W. Keller-Schierlein, *Helv. Chim. Acta*, 1972, **55**, 2071.
41. P. W. Schindler and H. Zaehner, *Arch. Mikrobiol.*, 1972, **82**, 66.
42. P. W. Schindler and H. Zaehner, *Eur. J. Biochem.*, 1973, **39**, 591.
43. P. W. Schindler and M. C. Scrutton, *Eur. J. Biochem.*, 1975, **55**, 543.
44. P. W. Schindler, *Eur. J. Biochem.*, 1975, **51**, 579.
45. Q.-L. He, X.-Y. Jia, M.-C. Tang, Z.-H. Tian, G.-L. Tang and W. Liu, *ChemBioChem*, 2009, **10**, 813.
46. W. Ding, C. Lei, Q. He, Q. Zhang, Y. Bi and W. Liu, *Chem. Biol.*, 2010, **17**, 495.

CHAPTER 13

Nuclear Magnetic Resonance in Saponin Structure Elucidation

ANTONIO JORGE RIBEIRO DA SILVA,*
RICARDO MOREIRA BORGES AND VITOR SOARES

Instituto de Pesquisas de Produtos Naturais Walter B. Mors, Avenida
Carlos Chagas Filho, 373, Edifício do CCS, Bloco H CEP 21941-902
Rio de Janeiro, RJ, Brazil

*Email: ajorge@nppn.ufrj.br

13.1 Introduction

Saponins are natural glycosides occurring primarily in plants.^{1,2} These natural compounds are present in many traditional medicines, attracting researchers to investigate their chemistry and pharmacological/biological activity.³ Industry applications of saponins include food, beverages, pharmaceutical, and cosmetics. Saponins are important raw materials for the synthesis of steroidal hormones³ and have been used as adjuvants in vaccine formulations.⁴

Saponins form abundant and persistent foams that are stable even to the action of mineral acids when in aqueous solution. The combination of a lipophilic triterpene or steroid moiety (aglycone or sapogenin) with a hydrophilic portion comprising one or more oligosaccharide units in their molecules explains this soap-like behavior. The steroidal aglycones have a common cholestane skeleton and may be generally classified as spirostanes or furostanes. The corresponding general structures for steroidal alkaloids are solanidanes and spirosolanes. Triterpenoidal aglycones are, in general, pentacyclic triterpenes with ursane, oleanane, or lupane skeletons. The hydrophilic moiety is a linear or branched oligosaccharide chain of variable

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

length. D-glucopyranosyl, D-galactopyranosyl, L-rhamnopyranosyl, L-arabinopyranosyl, D-xylopyranosyl, D-fucopyranosyl, D-apiofuranosyl and D-glucuronyranosyl, or D-galacturonyranosyl acids are common component units of the sugar chain. Monodesmosidic saponins have a sugar chain bound to the aglycone, while in bisdesmosidic saponins, there are two sugar chains.³

Isolation and structure elucidation of saponins can be a formidable task in view of their structural diversity, amphiphilic character, and the fact that they occur in complex mixtures.³ Traditionally, structure elucidation of these compounds has been made by using extensive hydrolytic or enzymatic degradation and chemical derivatization, together with GCMS and ¹H and ¹³C 1D nuclear magnetic resonance (NMR) methods in order to identify the aglycone, the number and type of monosaccharide residues present, and their mutual connections.⁵ NMR methods, especially 2D experiments, are increasingly replacing chemical methods to determine: (i) the structure and relative stereochemistry of the aglycone; (ii) the identity of component monosaccharides; and (iii) the aglycone linkage position and sequence of the oligosaccharide chain.

Oleanane triterpene and steroidal saponins along with steroidal alkaloids (also included here under the broad saponin designation) illustrate the structural discussions throughout the present text. The examples included in the following discussion were taken from results obtained by the research efforts of our group in recent years.

13.2 General Considerations

The 1D ¹H NMR spectra of saponins have characteristic features that are useful to help recognize the class of compound analyzed (triterpene or steroidal saponin, spirostan, or solanidane alkaloid) and other structural features. For example, two methyl singlets and two methyl doublets appearing between 0.5 and 1.5 ppm in a saponin ¹H NMR spectrum are indicative of a steroid skeleton, while seven (or six) methyl singlets appearing in the same region point to the presence of oleanane-type saponins. The majority of proton resonances in the oligosaccharide moiety appear between 3.5 and 4.2 ppm, generally displaying massive signal overlap. Anomeric protons and methyl resonances of 6-deoxy monosaccharide residues generally fall at 4.4–5.5 ppm and 1.1–1.3 ppm, respectively. Counting the number of aglycone carbons in ¹³C NMR spectra provides a simple distinction between steroidal and triterpene saponins, and the full number of carbon resonances helps in estimating the number and types of monosaccharide residues.

Figure 13.1 shows ¹H NMR spectra of three saponins: methylprotodioscin (1) (a furostane steroidal saponin), 3-O-β-D-glucuronyranosyl-hydroxiolean-12,15-dien-28-oic acid 28-O-α-D-apiofuranosyl (1 → 3)-(α-D-apiofuranosyl[1 → 4])-α-L-rhamnopyranosyl (1 → 2)-α-L-arabinopyranosyl ester (2) (a triterpene saponin),⁴ and tomatine (3), a steroidal glycoalkaloid. Note the spectral dispersion provided by the higher magnetic field (at 800 MHz), as compared to the spectra taken at 400 and 500 MHz.

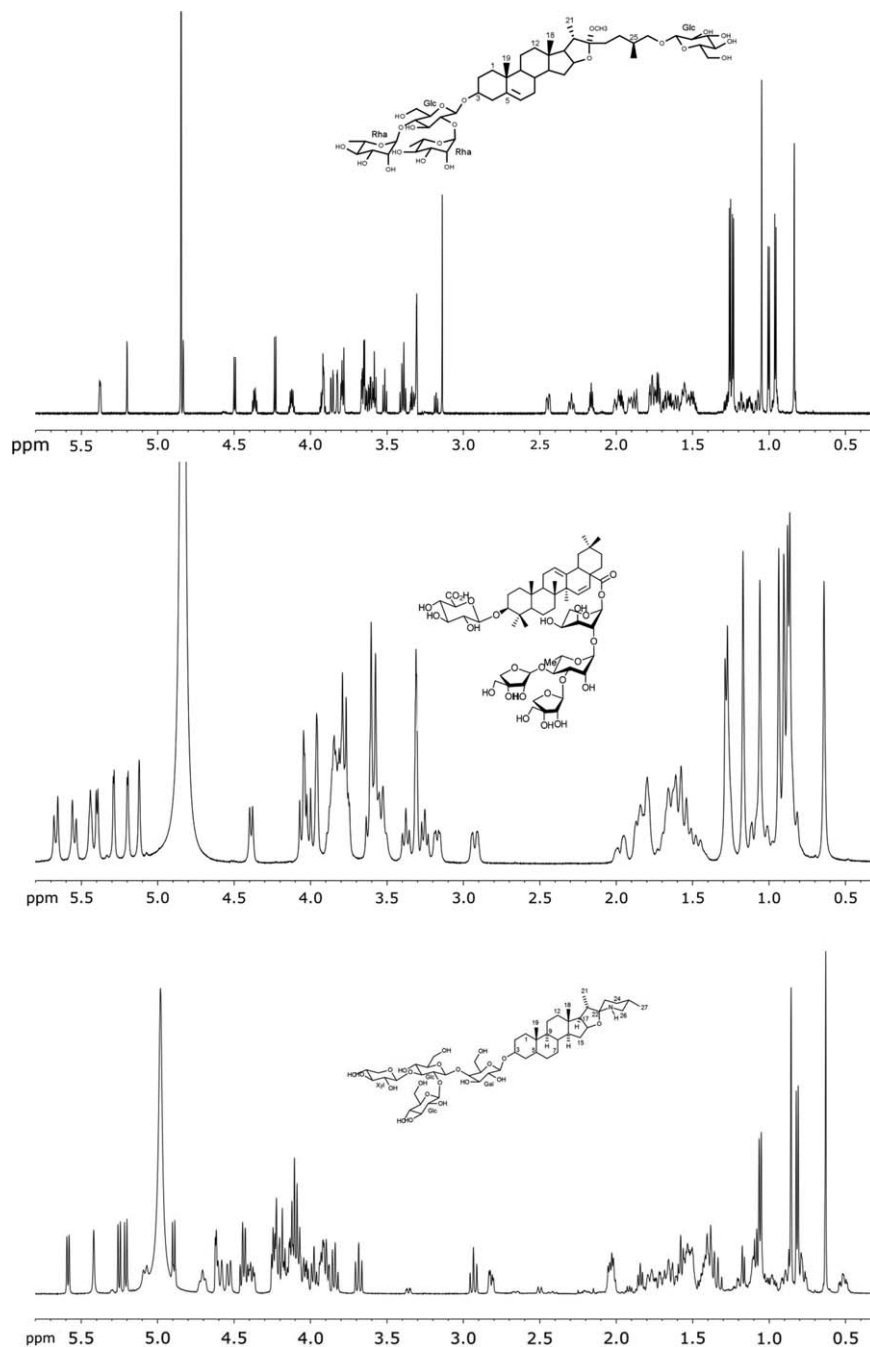


Figure 13.1 ^1H NMR spectra of saponins, top to bottom: methylprotodioscin (1) (MeOD, 800 MHz); 3-*O*- β -D-glucopyranosyl-3 β -hydroxyolean-12,15-dien-28-oic acid 28-*O*- α -D-apiofuranosyl (1 $\rightarrow$ 3)-(α -D-apiofuranosyl [1 $\rightarrow$ 4])- α -L-rhamnopyranosyl (1 $\rightarrow$ 2)- α -L-arabinopyranosyl ester (2) (MeOD, 400 MHz), and tomatine (3) (pyridine- d_5 , 500 MHz).

NMR spectra of saponins are normally measured using deuterated methanol, dimethyl sulfoxide, and pyridine as solvents. Care should be taken when comparing chemical shifts between spectra of different solvents, especially when one of them is pyridine, as dispersion of proton signals can be influenced by this solvent. On the other hand, these aromatic solvent-induced shifts may simplify the analysis of spectra and enable the extraction of J_{HH} values in the regions where signal overlap is a problem. The shifts induced by pyridine are larger in the neighborhood of polar groups like hydroxyls.

13.3 Structure and Relative Stereochemistry of the Aglycone

Obtaining complete chemical shift assignments is fundamental in the strategy of NMR-based structure elucidation of saponins. The use of 2D NMR experiments such as COSY, TOCSY, edited HSQC, HMBC, and ROESY provides unambiguous assignments of all carbons and protons. However, one should not forget the great deal of proton/carbon chemical shift data already compiled in the literature. The review by Tana *et al.*⁶ on “glucuronide oleanane-type triterpene carboxylic acid 3,28-*O*-bisdesmosides” reports an extensive compilation of chemical shifts, mostly of aglycone carbons, as well as structure elucidation strategies for the type of saponins focused upon. The review on pentacyclic triterpene carbon chemical shifts by Mahato and Kundu⁷ is another valued source of chemical shift data. Steroidal saponins and glycoalkaloids share structural similarities as compounds that are biosynthetically derived from a common precursor (cholesterol). The reader will also find in the literature a number of interesting reviews discussing carbon chemical shifts and structures of steroidal saponins^{8,9} and steroidal glycoalkaloids.¹⁰

Both ^{13}C and ^1H chemical shifts can be assigned by using correlated experiments such as COSY, TOCSY, HSQC and HMBC/HMQC. For example, Figure 13.2 displays the 0.6–1.2 ppm region of the HMBC correlation map for the oleanane-type triterpene saponin 4 (3-*O*- β -D-glucopyranosyl-3 β -hydroxyolean-12,15-dien-28-oic acid 28-*O*- β -D-xylopyranosyl [1 $\rightarrow$ 4]- α -L-rhamnopyranosyl [1 $\rightarrow$ 2]- α -L-arabinopyranosyl ester).^{11,12} Here, we see how the correlations of the methyl groups in the HMBC spectrum may provide good starting points to help assign the remaining hydrogens and carbons in a triterpene backbone. The free rotation of these groups keeps the values for $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ at around 4.5–6 Hz, allowing the detection of all expected methyl correlations.¹³ Notice that the mutual $^2J_{\text{CH}}/^3J_{\text{CH}}$ carbon/hydrogen correlations led to the identification of a sequence including 16 carbons in four of the five rings in the triterpene skeleton. For example, the four carbons correlated with hydrogens of methyl group C-27 were assigned at: 41.2 (C-8); 141.6 (C-13); 45.5 (C-14), and 128.7 ppm (C-5), leading to the location of the two double bonds in the vicinity of this methyl group. Additional entry points at the signals of H-3 (3.17 ppm) and the olefinic hydrogens H-12, H-15, and H-16 (5.44, 5.54, and 5.67 ppm, respectively) were used to locate the

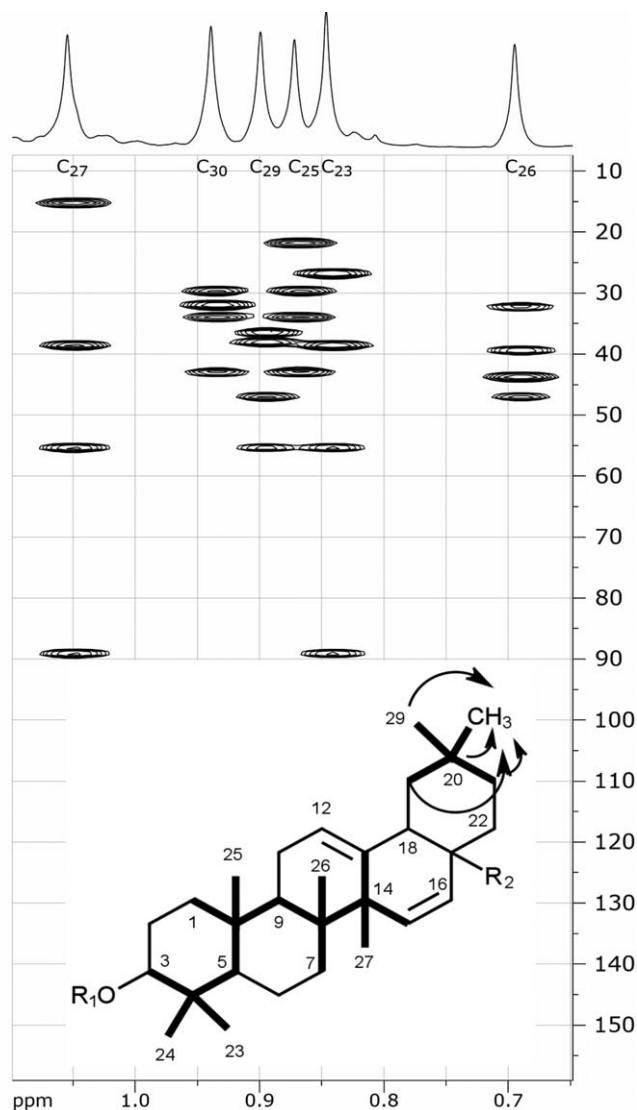


Figure 13.2 500 MHz HMBC of the oleanolic saponin (3-*O*- β -D-glucopyranosyl-3 β -hydroxyolean-12,15-dien-28-oic acid 28-*O*- β -D-xylopyranosyl (1 $\rightarrow$ 4)- α -L-rhamnopyranosyl (1 $\rightarrow$ 2)- α -L-arabinopyranosyl ester) (4) showing all expected aglycone methyl correlations. Solvent: pyridine-*d*5.

respective carbons by use of edited HSQC experiments: C-3 (91.1 ppm), C-12 (124.6 ppm), C-15 (128.7 ppm), and C-16 (137.5 ppm). An edited HSQC spectrum was also used to locate C/H pairs at positions 1 (C/H: 1.60 and 1.01/39.7 ppm), 5 (C/H: 0.82/57.0 ppm), 7 (C/H: 1.68 and 1.54/33.8 ppm), 9 (C/H: 1.66/48.7 ppm), 19 (C/H: 1.58 and 1.10/44.6 ppm), and 21 (C/H: 1.26/35.6 ppm). 2D homo- and hetero-nuclear correlated experiments such as edited HSQC, HMBC, and COSY spectra were used to assign the remaining

hydrogens and carbons. A straightforward application of the Karplus rule to the H3 signal at 3.14 ppm ($J = 11.8$ and 4.4 Hz) shows that this hydrogen is in an axial orientation, meaning that the oxygen in C-3 is β -oriented.

Figure 13.3 shows the results of the application of the same approach as before in the analysis of the steroidal glycoalkaloid tomatine. Note that the methyl groups attached to tertiary carbons show three correlations, unlike the methyl groups attached to quaternary carbons shown in the previous example. From the HMBC correlation map, 18 out of the 27 carbons in the

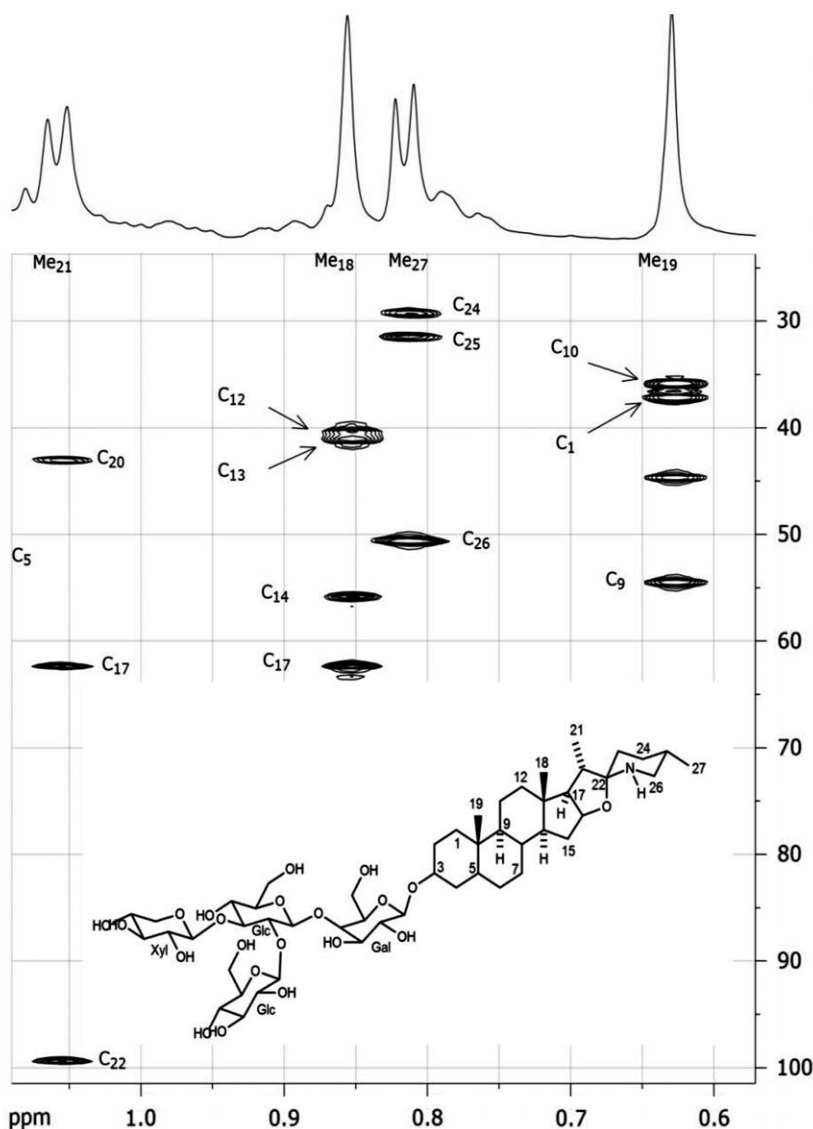


Figure 13.3 500 MHz HMBC of the glycoalkaloid tomatine (3). Solvent: pyridine-d₅.

aglycone could be identified. The remaining carbons/protons were identified by chemical shift correlated experiments as before.

The well-known dihedral angle dependence of $^3J_{\text{HH}}$ scalar couplings allows for the estimation of spatial relationships between protons. Additionally, although difficult to measure, vicinal ^{13}C – ^1H couplings may also be used to derive stereochemical information thanks to the same dependence. Thus, the measurement of vicinal coupling has been used as a powerful tool in the determination of the local stereochemistry of natural products.¹⁴

Overall pictures of molecular geometry can be obtained by measuring through-space interactions with NOESY or ROESY experiments. The ROESY sequence produces positive peaks, irrespective of the molecular size, contrary to NOESY; for this reason, ROESY is generally preferred for use with saponins, whose molecular weight range of 600–1700 can coincide with weak nuclear Overhauser effects.¹⁴

The 2D ROESY correlation map displayed in Figure 13.4 shows important through-space correlations used to confirm the overall stereochemistry of the steroidal glycoalkaloid tomatine.

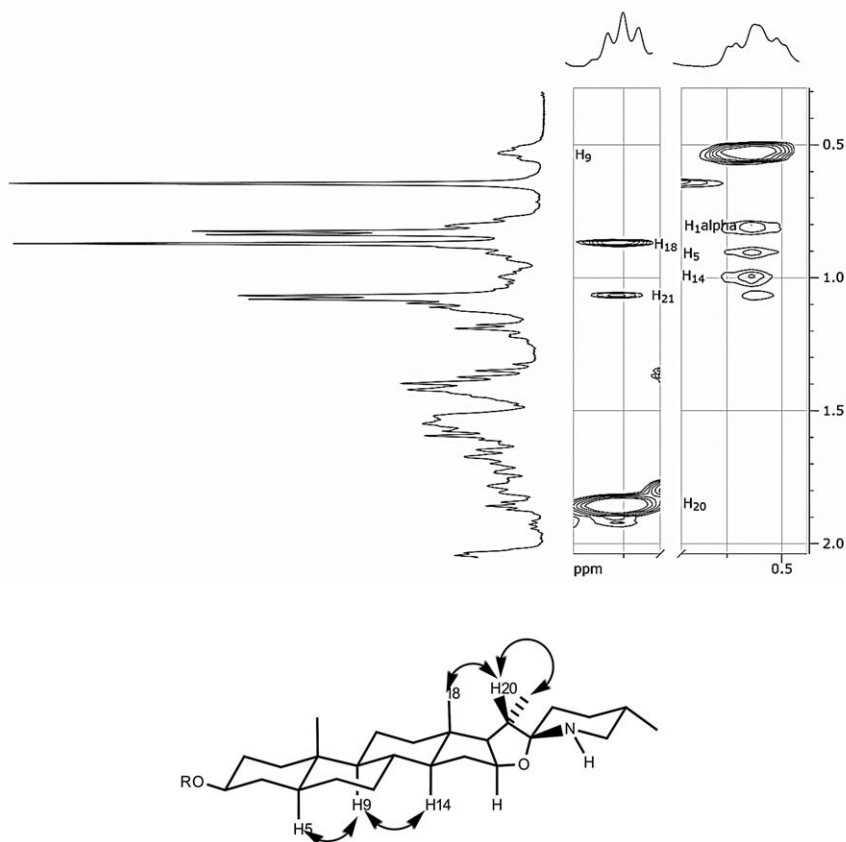


Figure 13.4 500 MHz slices of the 2D ROESY of tomatine. Solvent: pyridine- d_5 .

13.4 Identity of Component Monosaccharides

There are several experimental protocols that are used to determine the identity of component monosaccharides, as well as the structure of oligosaccharides and glycosides, most of them based on the analysis of the component monosaccharides after chemical or enzymatic degradation and derivatization. GCMS analysis of permethylated, peracetylated, or trimethylsilylated derivatives is routinely used for the structural elucidation of oligosaccharides and glycosides. Permethylation, acid hydrolysis, alditol reduction, and acetylation allow for the determination of sugar sequences and sites of interglycosidic linkage, but not anomeric carbon configurations.⁵

On the other hand, NMR spectroscopy provides extensive information on the natures of component monosaccharides and oligosaccharide chains, besides being an amenable analysis tool for those compounds that are very difficult to separate and purify.¹⁵ Proton and carbon NMR chemical shifts carry a wealth of structurally important information, but structure elucidation strictly based of NMR chemical shifts is a very difficult task. For example, anomeric proton resonances appear in the shift range of 4.4–5.8 ppm, while the signals from 6-deoxy sugar residues appear between 1.1 and 1.3 ppm. These signals are easy to correlate with structure. However, the other ring proton resonances are found in a very narrow spectral region (between 3 and 4.2 ppm) and generally display a large degree of overlap. Counting the number of anomeric proton/carbon resonances is a good starting point in the identification and estimation of the number of different monosaccharide residues present in the saponin.

Several 2D homonuclear correlation experiments such as DQF COSY, zero-quantum-filtered COSY (Z-COSY), and TOCSY may help define the coupled sequence of hydrogens in each of the monosaccharide residues, with TOCSY being the preferred sequence in this specific case.¹⁴ Figure 13.5 shows a Z-TOCSY correlation map of the oleanolic saponin 3-*O*- β -D-glucopyranosyl-3 β -hydroxyolean-12,15-dien-28-oic acid 28-*O*- α -D-apiofuranosyl (1 $\rightarrow$ 3)-[α -D-apiofuranosyl (1 $\rightarrow$ 4)]- α -L-rhamnopyranosyl (1 $\rightarrow$ 2)- α -L-arabinopyranosyl ester (2). Note the clear correlations observed for the anomeric proton of the glucopyranosyl residue (doublet at 4.39 ppm). The Z-TOCSY sequence suppresses distortions in proton multiplets that are artifacts due to zero-quantum coherence.¹⁴

In cases of heavy crowding of saccharide proton signals, such as in the tetrasaccharide linked to C3 in tomatine, the identification of the component monosaccharides may become a hard task. The 2D HSQC-TOCSY experiment is very useful in such cases where the wide chemical shift signal dispersion that is typical of carbon helps the identification of the monosaccharides. Figure 13.6 displays cross-peaks of the monosaccharide residues of tomatine measured in a 500 MHz 2D HSQC-TOCSY experiment. The sequence used was a GHSQC-TOCSY with a mixing time of 150 ms. Note the absence of one of the expected cross-peaks, indicating that further optimization of experimental parameters is required.

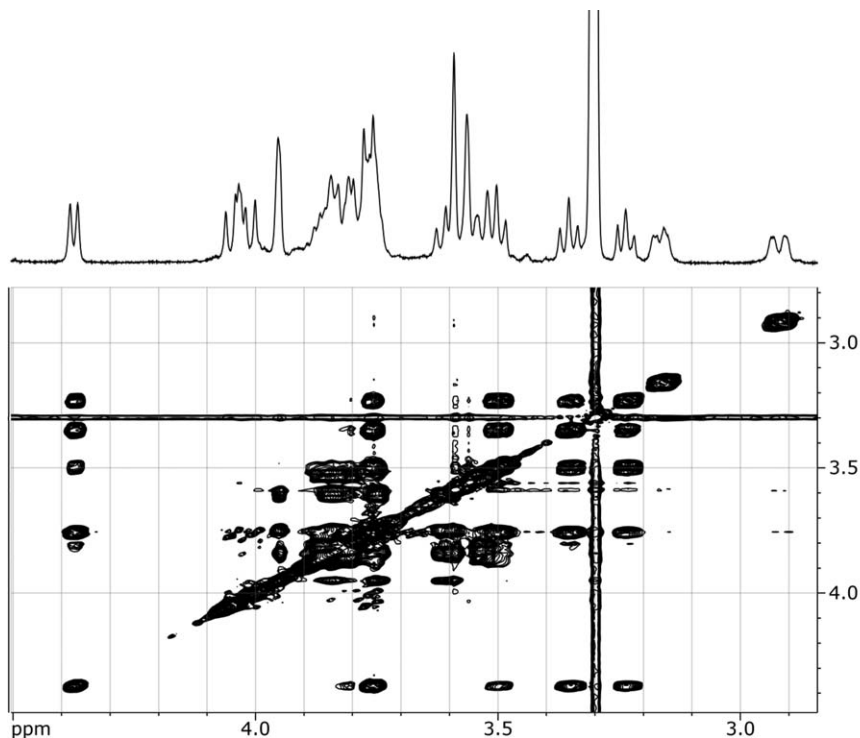


Figure 13.5 500 MHz Z-TOCSY of the oleanolic saponin 3-*O*- β -D-glucopyranosyl-3 β -hydroxyolean-12,15-dien-28-oic acid 28-*O*- α -D-apiofuranosyl (1 $\rightarrow$ 3)-(α -D-apiofuranosyl [1 $\rightarrow$ 4])- α -L-rhamnopyranosyl (1 $\rightarrow$ 2)- α -L-arabinopyranosyl ester (2).

Selective 1D TOCSY is an excellent alternative to the 2D experiment when we want to determine a sequence of coupled protons in a molecule. These experiments are generally set up as an array with increasing mixing times after selection of a well-resolved signal to be excited by using shaped pulses.¹⁴ Figure 13.7 shows the results of a 1D TOCSY-arrayed experiment where the anomeric hydrogen signal at 4.38 ppm was selected and, by using five different MLEV mixing times from 10 to 100 ms, we could identify the sugar unit (a 1-substituted glucopyranosyl unit) in 3-*O*- β -D-glucopyranosyl-3 β -hydroxy-olean-12,15-dien-28-oic acid. The large axial-axial coupling constant values are typical; the multiplicities displayed by the hydrogens and the lack of a methylene group at position 6 unequivocally characterize the monosaccharide residue as a glucopyranosyl residue. These experiments are fast, dependable, and offer clear data to help identify the nature of the carbohydrate residue, in addition to taking much less time than the corresponding 2D experiment. Additionally, this approach is especially useful in cases where severe signal overlapping is present.

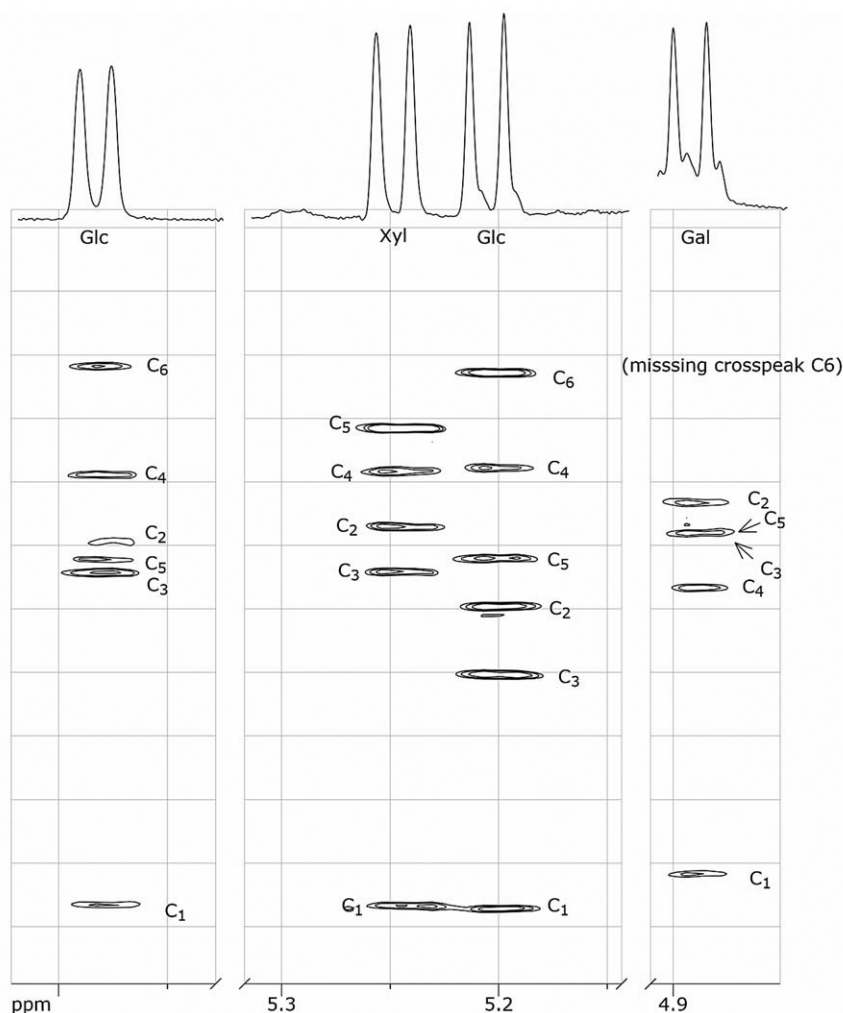


Figure 13.6 500 MHz HSQC-TOCSY of tomatine. Solvent: pyridine-d₅.

Next, we analyze 3-*O*-β-D-glucopyranosyl-3β-hydroxyolean-12,15-dien-28-oic acid 28-*O*-β-D-xylopyranosyl (1 → 4)-α-L-rhamnopyranosyl (1 → 2)-α-L-arabinopyranosyl ester, a saponin bearing the same glucopyranosyl unit linked to C-3, but with a trisaccharide at C-28 (see Figure 13.7). 1D TOCSY experiments were used to identify each of the component monosaccharides in the oligosaccharide moiety of the saponin. In the present case, the ¹H NMR spectrum displays four signals corresponding to four anomeric sugar residues at 5.43/94.30, 5.09/101.23, 4.48/107.21, and 4.39/107.10 ppm (¹H/¹³C chemical shifts; correlations obtained from a HSQC spectrum). These signals were selected and arrays of five 1D TOCSY spectra were acquired with the same mixing time values as in the previous example (Figure 13.7).

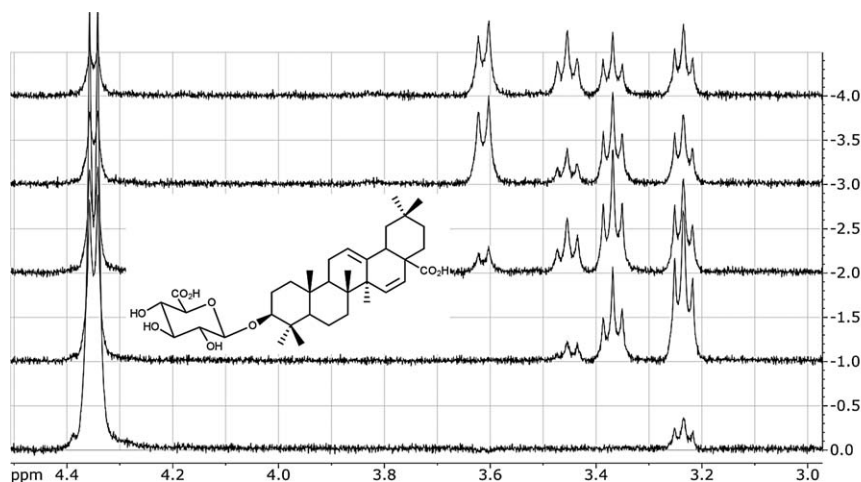


Figure 13.7 1D TOCSY of 3-O-β-D-glucopyranosyl-3β-hydroxyolean-12,15-dien-28-oic acid, selected signal: 4.38 ppm. Mixing times from 10 to 100 ms.

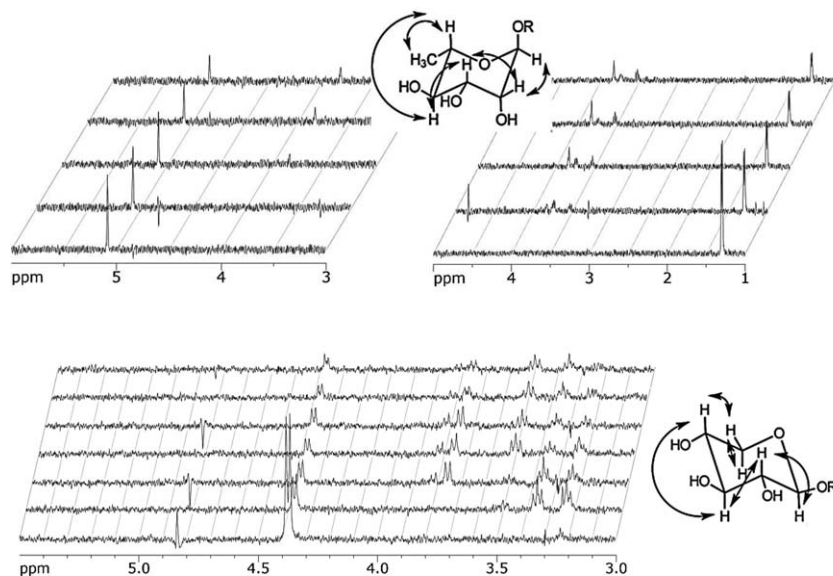


Figure 13.8 500 MHz 1D TOCSY of 3-O-β-D-glucopyranosyl-3β-hydroxyolean-12,15-dien-28-oic acid 28-O-β-D-xylopyranosyl (1→4)-α-L-rhamnopyranosyl (1→2)-α-L-arabinopyranosyl ester. Top: selected signals: 1.31 ppm (right, rhamnose methyl) and 5.09 ppm (left, rhamnose anomeric hydrogen). Bottom: selected signal: 4.48 ppm (xylose residue). Mixing times: from 10 to 100 ms.

Under these conditions, 1D sub-spectra for each of the monosaccharide residues were obtained, leading to the identification of the monosaccharide units as arabinofuranosyl, rhamnopyranosyl, xylopyranosyl, and glucopyranosyl, respectively. Figure 13.8 shows the 1D TOCSY spectral arrays for the

rhamnopyranosyl and xylopyranosyl units, respectively. Since the protons in the xylopyranosyl residue (a pentose) show strong couplings between their hydrogens (axial-axial, $^3J_{\text{HH}} \sim 7$ Hz), selective excitation of H-1 followed by successive transfer to H-2, H-3, H-4, and H-5 delineates a sequence of coupled protons within the xylopyranosyl structure. On the other hand, the rhamnopyranosyl residue lacks the strong coupling between H1 and H2, resulting in a weaker transfer of coupling information to the entire monosaccharide ^1H spin system. In this specific case, it is useful to select the signal of the methyl group doublet at ~ 1.3 ppm in order to provide full characterization of the entire spin system. After the assignment of the monosaccharide hydrogens, their carbons were assigned by means of a HSQC experiment.

13.5 Linkage Position and Sequence of the Oligosaccharide Chain

Following the identification of the component monosaccharide residues, the determination of the aglycone linkage site is made. In the HMBC spectrum of 3-*O*- β -D-glucopyranosyl-3 β -hydroxyolean-12,15-dien-28-oic acid 28-*O*- β -D-xylopyranosyl (1 $\rightarrow$ 4)- α -L-rhamnopyranosyl (1 $\rightarrow$ 2)- α -L-arabinopyranosyl ester, a cross-peak is observed between the signal of the anomeric proton of the glucopyranosyl residue at 4.39 ppm and C-3 in the aglycone (91.2 ppm), leading to the location of one of the glycosylation sites at this position (C-3). This location was confirmed by a 1D ROESY experiment (Figure 13.9, bottom) where the selection of the same anomeric proton at 4.39 ppm led to the detection of a ROE enhancement in the proton signal at 3.16 ppm (H-3). The second site of glycosylation was determined as follows: the correlation

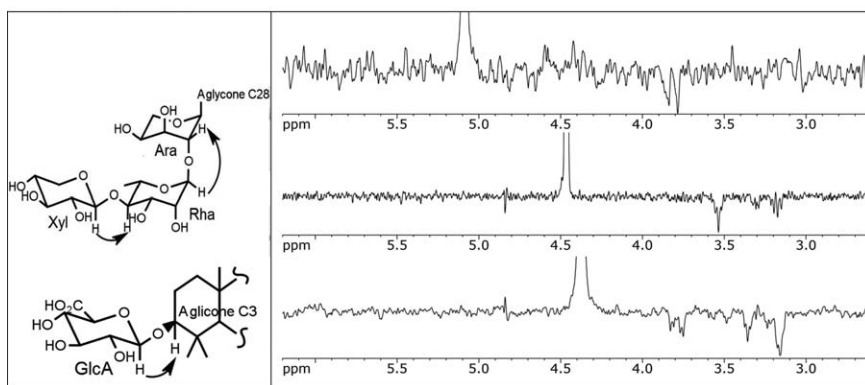


Figure 13.9 500 MHz 1D ROESY of 3-*O*- β -D-glucopyranosyl-3 β -hydroxyolean-12,15-dien-28-oic acid 28-*O*- β -D-xylopyranosyl (1 $\rightarrow$ 4)- α -L-rhamnopyranosyl (1 $\rightarrow$ 2)- α -L-arabinopyranosyl ester (**4**). Selected signals, bottom to top: 4.38 (H-1 glucopyranosyl acid residue); 4.48 ppm (H-1 xylopyranosyl residue); and 5.09 ppm (H-1 rhamnopyranosyl residue).

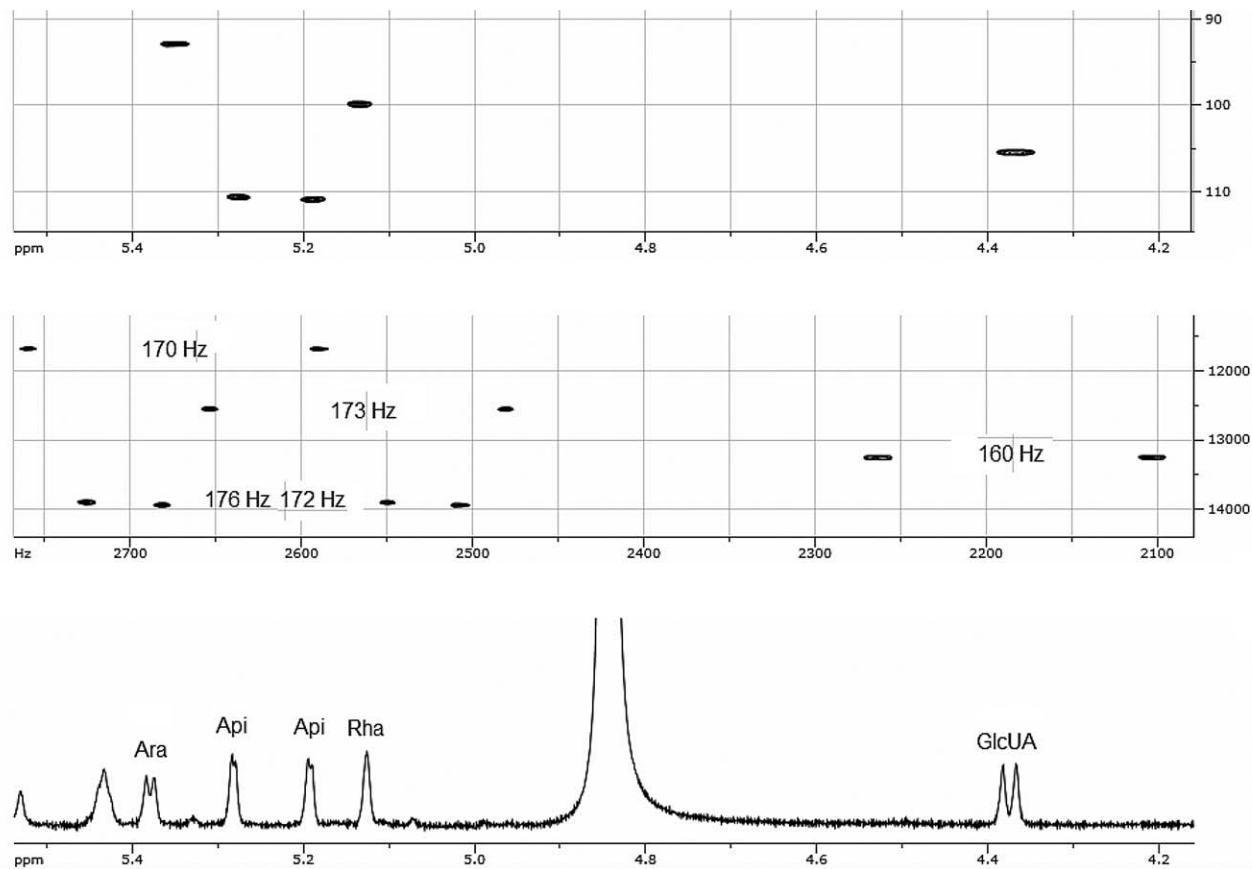


Figure 13.10 Anomeric carbons configurations of 3-O-β-D-glucopyranosyl-3β-hydroxyolean-12,15-dien-28-oic acid 28-O-α-D-apiofuranosyl (1→3)-(α-D-apiofuranosyl [1→4])-α-L-rhamnopyranosyl (1→2)-α-L-arabinopyranosyl ester (2). Bottom: ^1H NMR; middle: F_2 -coupled HSQC; top: conventional HSQC.

observed in the HMBC spectrum between the proton at 4.97 ppm (arabinopyranosyl anomeric proton) and the carbonyl carbon at 176.00 ppm (C-28 aglycone) located the arabinopyranosyl residue at C-28.

The sequence of monosaccharides in the oligosaccharide chain was obtained by means of a series of 1D ROE experiments (Figure 13.9, middle and top): selecting the signal at 4.48 ppm (xylopyranosyl residue anomeric proton); selecting the signal at 5.09 ppm (rhamnopyranosyl residue anomeric proton).

The measurement of $^1J_{\text{CH}}$ values for anomeric carbons is a reliable criterion for determining the anomeric configuration of saccharides occurring in pyranose form. In this case, the difference in scalar coupling magnitude is around 10 Hz, with greater values for equatorial C–H coupling (α -anomer). The values for these coupling constants are between 160 and 170 Hz. There are several methods for measuring $^1J_{\text{CH}}$; for example, the direct C–H coupling of anomeric carbons may be measured with the help of a 2D J -resolved heteronuclear experiment. However, 2D J -resolved heteronuclear experiments are X-detected experiments and, as such, lack the sensitivity offered by ^1H -detected experiments. Recently, Mazzola *et al.*¹⁶ used F_2 -coupled HSQC spectra for determining anomeric C–H coupling constants. Figure 13.10 shows the result of applying this approach in order to measure direct C–H coupling constants for the anomeric carbons of the oleanolic saponin 2. Based on these results, the configuration of the glucopyranosyl at C-3 was assigned as β , while the anomeric configuration in the remaining monosaccharide residues was assigned as α . It is worth mentioning that, in the case of the two pentofuranosides (apiose residues), this assignment may be ambiguous due to the similarity of the $^1J_{\text{H}}$ values for α or β epimers at the anomeric carbon.¹⁴

13.6 Concluding Remarks

Rather than providing an extensive coverage of saponin structural elucidation examples that are scattered in the literature, we chose to present parallel objective examples on how to get structural information by using specific NMR techniques, as well as the ever-growing proton and carbon chemical shifts database that is now available in the literature. The already-mentioned reviews by Tana *et al.*,⁶ Mahato and Kundu,⁷ Agrawal *et al.*,⁸ Sahu *et al.*,⁹ and Agrawal *et al.*¹⁰ are excellent sources of data and information on strategies for the structure elucidation of saponins. Agrawal also published a text about a systematic approach to NMR structure elucidation of steroidal saponins, now slightly outdated with regard to the employed NMR techniques, but containing a wealth of information on structures/spectral parameters (J and δ), while also establishing all essential steps to be followed in the structure elucidation of these compounds.¹⁷

Mass spectrometry and chemical degradation are still invaluable tools in the structure elucidation of saponins. However, the increased availability of higher-field NMR spectrometers,¹⁸ together with the development of new

pulse sequences,¹⁶ are helping researchers to uncover the complexities of saponin structures, especially in the case of subtle stereochemical problems that would be otherwise difficult to solve (e.g. the stereochemistry at C-25 in furostanol steroidal saponins).¹⁹

13.6.1 Experimental

All cited oleanane saponins were isolated from *Chiococca alba*.^{11,12} Methylprotodioscin was isolated from *Solanum paniculatum* (Vitor Soares, private communication). Tomatine (standard) was purchased from Sigma-Aldrich (St Louis, MO, USA).

The ¹H/¹³C NMR spectra were obtained in 400/100 MHz (Bruker DRX 400), 500/125 MHz (Varian VNMRSYS-500), and 800/200 MHz (Bruker Avance-III-800) spectrometers in the indicated solvents, with TMS as internal standard, using 5 and 3 mm tubes and the indicated solvents. 1D and 2D correlation experiments (1D TOCSY, HSQC, HMBC, and 1D ROESY) were performed with standard proprietary software. For 2D heteronuclear experiments (edited HSQC, HMBC, and HSQC-TOCSY), the typical acquisition parameters included 1K or 2K×512 data points, zero-filled to 1K or 2K×2K points and processed using linear prediction in F₁. All spectral data were processed using INMR for Windows (<http://www.inmr.net>).

References

1. A. Osbourn, R. J. M. Goss and Robert A. Field, *Nat. Prod. Rep.*, 2011, **28**, 126.
2. J. P. Vincken, L. Heng, A. de Groot and H. Gruppen, *Phytochemistry*, 2007, **68**, 275.
3. K. Hostettmann and A. Marston *Saponins*, ed. J. D. Phillipson, D. C. Ayres and H. Baxter, Cambridge University Press, Cambridge, 2005.
4. D. Nico, R. M. Borges, L. M. Brandão, F. D. Ferreira, D. C. Gomes, M. Palatnik, M. M. Rodrigues, A. J. R. da Silva and E C. Palatnik-de-Sousa, *Vaccine*, 2012, **30**, 3169.
5. T. Schöpke, *Saponins in Food, Feedstuffs and Medicinal Plants*, ed. W. Oleszek and A. Marston, Kluwer Academic Publishers, 2000, ch. 10, pp. 95–106.
6. N. Tana, J. Zhou and S. Zhao, *Phytochemistry*, 1999, **52**, 153.
7. S. B. Mahato and A. P. Kundu, *Phytochemistry*, 1994, **37**, 1517.
8. P. K. Agrawal, D. C. Jain, R. K. Gupta and R. S. Thakur, *Phytochemistry*, 1985, **24**, 2479.
9. N. P. Sahu, S. Banerjee, N. B. Mondal and D. Mandal, *Progress in the Chemistry of Organic Natural Products*, ed. D. Mandal, A. D. Kinghorn, H. Falk and J. Kobayashi, Springer International Publishing, Switzerland, 2008, ch. 2, vol. 89, pp. 45–141.

10. P. K. Agrawal, S. K. Srivastava and W. Gaffield, *Alkaloids: Chemical and Biological Perspectives*, ed. S. W. Pelletier, Elsevier, Oxford, 1998, ch. 2, vol. 7, pp. 47–295.
11. R. M. Borges, L. W. Tinoco, J. D. de Souza Filho, N. S. Barbi and A. J. R. da Silva, *J. Braz. Chem. Soc.*, 2009, **20**, 1738.
12. R. M. Borges, S. S. Valenca, A. A. Lopes, N. S. Barbi and A. J. R. da Silva, *Phytochem. Lett.*, 2013, **6**, 96.
13. A. Mangoni, *Handbook of Marine Natural Products*, ed. E. Fattorusso, W. H. Gerwick, O. Tagliatela-Scafati, Springer, Switzerland, 2012, ch. 8, pp. 519–544.
14. R. C. Breton and W. F. Reynolds, *Nat. Prod. Rep.*, 2013, **30**, 501.
15. P. K. Agrawal, *Phytochemistry*, 1992, **31**, 3307.
16. E. P. Mazzola, A. Parkinson, E. J. Kennelly, B. Coxon, L. S. Einbond and D. I. Freedberg, *Carbohydr. Res.*, 2011, **346**, 759.
17. P. K. Agrawal, *Advances in Experimental Medicine and Biology*, ed. G. R. Waller and K. Yamazaki, Plenum Press, New York, 1996, vol. 405, pp. 299–315.
18. R. R. King and L. A. Calhoun, *Magn. Reson. Chem.*, 2012, **50**, 627.
19. V. L. Challinor, S. Piacente and J. J. de Voss, *Steroids*, 2012, **77**, 602.

CHAPTER 14

Increasing the Adoption of Advanced Techniques for the Structure Elucidation of Natural Products

ANTONY J. WILLIAMS,*^a GARY E. MARTIN*^b AND
DAVID ROVNYAK*^c

^a ChemConnector Inc., Wake Forest, NC 27587, USA; ^b Merck Sharp and Dohme, Process & Analytical Chemistry, NMR Structure Elucidation, Rahway, NJ 07065, USA; ^c Bucknell University, Department of Chemistry, Lewisburg, PA 17837, USA

*Email: tony27587@gmail.com; gary.martin2@merck.com; drovnyak@bucknell.edu

The changes in technology that we experience in any given year are not slowing. “Smartphones” were initially used for receiving phone calls and primitive text messages, but now lead the charge in mobile computing, serving as music players, cameras, GPS receivers, gaming devices, and our primary means of managing email. The field of nuclear magnetic resonance (NMR) structure elucidation has witnessed important changes in instrumentation technology with time, and many of these have been discussed in these two volumes. Generally, the *primary* foci of evolving NMR technology are improving sensitivity (*e.g.* magnetic field strength and probe development) and providing new ways to manipulate the magnetization in order to influence the information that can be inferred from the data (*e.g.* pulsed

Modern NMR Approaches to the Structure Elucidation of Natural Products:

Volume 2: Data Acquisition and Applications to Compound Classes

Edited by Antony J. Williams, Gary E. Martin and David Rovnyak

© The Royal Society of Chemistry 2017

Published by the Royal Society of Chemistry, www.rsc.org

field gradients for coherence pathway selection). Even the fundamental procedure of data sampling has been reinvented, and has matured into vendor adoption to either improve the quality (*e.g.* resolution and interpretability) of data contained in a spectral matrix or to speed up the acquisition of data (*e.g.* covariance processing and non-uniform sampling, respectively). But in contrast to the less consequential decision of whether to switch to a smartphone or not, the very serious result of not utilizing powerful new advances in NMR is to greatly hinder the progress of any NMR-based inquiry. The judgment of the editors of this series is that, in general, there is an unacceptable and inexplicable lag in the adoption and application of new techniques.

Historically, heteronuclear shift correlation experiments were performed using heteronucleus detection. The 1979 report by Mueller of proton-detected heteronuclear shift correlation represented a major advance in sensitivity.¹ Performing proton- or “inverse”-detected heteronuclear shift correlation experiments required improvements in instrumentation because of the need to specifically select only the ^1H – ^{13}C heteronucleide pairs while rejecting 99% of the ^1H signal that arises from ^1H – ^{12}C . Vibration from tubes being spun to improve resolution was an early issue. RF stability and phase shifter accuracy represented yet another challenge. Nevertheless, it is instructive to look at the frequency of natural product reports in the literature for the two types of proton-detected heteronuclear shift correlation experiments, HMQC² and HSQC.³ Table 14.1 is based on SciFinder^{®4} searches combining the terms “HMQC” or “HSQC” and “natural products”. In part, the lag between the early reports of “inverse” or reversed detection versus heteronucleus detection was based on the time that it took for laboratories to obtain and install what can be termed inverse detection probes, in which the ^1H coil is nested closer to the sample inside of the $^{13}\text{C}/^{15}\text{N}$ or X-coil. This is not to say that inverse-detected experiments cannot be performed in “normal” geometry probes with the coil for the heteronucleus inboard of the ^1H decoupling coil—they can, and all of the editors have done those types of experiments.

Table 14.1 Citation frequency of the HMQC and HSQC heteronuclear shift correlation techniques over the interval from 1990 to the present.^a

Year range	HMQC articles (<i>n</i>)	HSQC articles (<i>n</i>)
1990–1994	52	10
1995–1999	177	39
2000–2004	346	111
2005–2009	358	266
2010–2015	345	423
Totals	1278	849

^aThis table was prepared using a SciFinder[®] search on the combined terms “HMQC” or “HSQC” with the term “natural product”. The columns reflect the number of times those terms appeared in the same paper in the 5-year intervals shown.

Even the most cursory examination of Table 14.1 reveals that the HMQC experiment was adopted much more quickly than the HSQC experiment. The former has also essentially plateaued, while the number of reports of the HSQC experiment in natural products publications has steadily climbed over the 25 years since the first reports. There are considerable advantages inherent to HSQC yet, perhaps inexplicably, the HMQC experiment continues to enjoy a high level of popularity among natural products investigators.

Long-range heteronuclear shift correlation is now routinely done using the HMBC experiment first reported by Bax and Summers in 1986.⁵ Indeed, a SciFinder[®] search of the terms HMBC and natural products unearthed 2492 reports when this chapter was being written. The HMBC experiment can be used for both ^1H – ^{13}C and ^1H – ^{15}N long-range correlation experiments with a simple adjustment of the gradient ratio. HMBC data typically provide $^2J_{\text{XH}}$ and $^3J_{\text{XH}}$ correlations, with occasionally still-longer-range correlations observed, which are almost uniformly limited to $^4J_{\text{XH}}$. In large part, the number of bonds across which the HMBC experiment can “reach” is limited by the fact that, at the end of the experiment, the detected magnetization is anti-phase and the data are magnitude calculated for presentation. In 2014, a new experiment, LR-HSQMBC, was reported by Williamson, Buevich, and Martin,⁶ which offers the potential to span as far as $^6J_{\text{CH}}$ for ^1H – ^{13}C and $^4J_{\text{NH}}$ for ^1H – ^{15}N experiments.⁷ In the initial presentation of the technique at the American Society of Pharmacognosy meeting held in Oxford, Mississippi, in August 2014, the potential of LR-HSQMBC to unravel proton-poor connectivities in natural products was enthusiastically received by a nearly full auditorium of attendees.⁸ Well into 2015, the only reported applications of the technique have emanated from the laboratories of the innovators, who have found it to be an essential tool in accelerating and disambiguating the structure-solving process. It will be interesting to see how long the adoption of the LR-HSQMBC experiment, which fills a need long vacant with the HMBC experiment, takes in the natural products community.

All of the editors have experienced the challenges associated with developing new technologies in our laboratories and attempting to have our colleagues implement and utilize them. Despite the obvious advantages of new experiments for the acquisition of improved data, we judge that the adoption of new technologies into mainstream usage generally, and perhaps unfortunately, resides with the instrument vendor providing new technologies in a “canned form”, either implemented in a hardware advance (e.g. cryogenic NMR probes) and easily addressed in an automated fashion, or by direct inclusion in the instrumentation software (or updates) for data processing at the instrument. For example, the adoption of linear prediction as a data-processing technique likely only grew in popularity when the instrument vendors delivered it as a console processing option, since moving the data offline to a third-party software tool would have had a significant activation barrier. Covariance processing, as reported in Volume 1, Chapter 10, is already available as an option in offline processing tools and now in instrument vendor software, yet there have only been three reports of the use

of any form of covariance processing in the *Journal of Natural Products*, all by G. E. Martin, one of the primary proponents and innovators of the technology, and his colleagues.^{9–11}

So how will this situation change? What is necessary in order to enable scientists to move quickly from simply reading peer-reviewed articles (or attending conference presentations) about the latest applied experimental approaches to applying them in their own laboratories? As discussed earlier, it appears that in the majority of cases, new techniques gain popularity only when the instrument vendors deploy the experiments as part of their standard pulse libraries. In 2014, we sadly saw Agilent withdraw from the delivery of NMR systems to the scientific community, having acquired Varian NMR in 2010.¹² At present, there are only two primary NMR vendors serving the research community—Bruker and JEOL—with Bruker holding the majority of the market share outside of the Asian market. Further, the analysis of the broader NMR instrumentation market space that has been made public so far is not optimistic. Assuming that there will not be any new entrants into the research-grade high-field NMR market, existing Varian/Agilent console systems will be replaced likely within the next decade as they reach obsolescence and parts for them become unavailable. Other popular ways to deliver details of pulse sequences is *via* the vendors' online discussion forums, as well as *via* online communities such as that exemplified by the NMRWiki.¹³ In principle, as a silver lining, deploying new innovations in acquisition, processing, and pulse sequences across all platforms will be simplified now that the vendor space is reduced. However, this does not mean that the new experiments will be utilized even if they are available to the laboratory scientist.

What else impedes the adoption of new technologies? In academia, faculty time for scholarship is continually reduced as universities need more non-research services from faculty to maintain the sustainability of these institutions. Universities also have turned increasingly to shared, fee-based spectrometers, discouraging students and their PIs from conducting those activities at the spectrometer that are deemed “non-essential”. Yet the editors remember well that long hours of “non-essential” time at the spectrometer were crucial for learning and experimentation and, notably, for coding new pulse sequences almost the day they were published. But now there is too much pressure to save the expensive hourly fees, too much competition for precious spectrometer time, and too much temptation to wait for the vendor or another laboratory to do it. In industry, the destabilization of career paths for PhD scientists has affected both the continuity and the building of expertise. The editors notice a trend that those industry laboratories driving, and benefitting from, innovation and its adoption are those with the greatest continuity of their PhD scientists. In other words, the contrast between transformative technologies at the cutting edge and the relative few who are using them draws attention to the changing fundamental economic models in academia and industry that harm innovation and its adoption. It should be noted, however, that there

remain bastions of creativity in industry where NMR scientists are *strongly* encouraged to develop new methods in order to solve problems.

We are hopeful that the time from which experiments reported in the literature can be moved to the laboratory and into general use can be reduced. Making the transfer of technology more rapid and accessible implies the need for changes in current practices. It would make sense that publishers mandate that pulse sequences be made available, not only *via* supplementary info in PDF form, but also linked *via* the appropriate instrument vendor forum or at a site such as NMRWiki. More didactic approaches to sharing innovations are needed in order to enhance adoption. With the advent of numerous social and professional networking platforms, scientists can even share the details of experiments themselves. Such a platform, whether simply for a scientist to deliver their own experimental details or even as a platform for general community sharing, allows for community feedback and dialoging regarding how the experiments can be improved or how they differ from platform to platform. Ultimately, such trends will require participation, time, and resources from institutions and individual scientists in order to remain up to date with the latest capabilities.

There are solutions and, much like the reality of the growing consequences of global warming, there is still time to act. Yet the loss of one major vendor and the weakening of the broader market can be seen as symptomatic of stagnation in the adoption of new technologies. The demonstrated, not hypothetical, ability to solve the complex structures of submilligram-scale, proton-deficient natural products in as little as a day when combining the technologies reported in this two-volume series justifies greater investment and growth in NMR. We hope it will happen.

References

1. L. Mueller, Sensitivity Enhanced Detection of Weak Nuclei Using Heteronuclear Multiple Quantum Coherence, *J. Am. Chem. Soc.*, 1979, **101**, 4481–4484.
2. A. Bax and S. Subramanian, Sensitivity-enhanced two-dimensional heteronuclear shift correlation NMR spectroscopy, *J. Magn. Reson.*, 1986, **69**, 525–569.
3. G. Bodenhausen and D. J. Ruben, Natural Abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, *Chem. Phys. Lett.*, 1980, **69**, 185–189.
4. SciFinder[®], CAS, Division of the American Chemical Society.
5. A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, 1986, **106**, 2093–2094.
6. R. T. Williamson, A. V. Buevich and G. E. Martin, *J. Org. Chem.*, 2014, **79**, 3887–3894.
7. R. T. Williamson, A. V. Buevich and G. E. Martin, *Tetrahedron Lett.*, 2014, **55**, 3365–3366.
8. R. T. Williamson, A. V. Buevich and G. E. Martin, LR-HSQMBC – A High-Sensitivity ‘Longer-Range’ Big Brother of HMBC on Steroids, invited

presentation, Bruker NMR Users Meeting held in conjunction with the American Society of Pharmacognosy (ASP) annual meeting. Oxford, Mississippi, August 2, 2014.

9. G. E. Martin, B. D. Hilton and K. A. Blinov, *J. Nat. Prod.*, 2011, **74**, 2400–2407.
10. G. E. Martin, B. D. Hilton, K. A. Blinov and A. J. Williams, *J. Nat. Prod.*, 2007, **70**, 1966–1970.
11. G. E. Martin, B. D. Hilton, P. A. Irish, K. A. Blinov and A. J. Williams, *J. Nat. Prod.*, 2007, **70**, 1393–1396.
12. R. Freeman and G. A. Morris, The Varian Story, *J. Magn. Reson.*, 2015, **250**, 80–84.
13. <http://nmrwiki.org/>.

Subject Index

- abietic acid (terpene) 255
- abyssomicin J (polyketide) 463–5, 474
- ACD/NNMR prediction program 76–7, 80, 82, 84
- 9-(5'-deoxy-5'-thio- β -D-xylofuranosyl)adenine disulfide 436
- ADEQUATE experiments 109, 223, 226–7, 240–3, 271–2, 280, 303, 306, 318
 - see also* alkaloids NMR (part II)
- adjacent polarization (ASAP) 202–3
- aglycone 486, 489–92
- alkaloids NMR (part I) ^1H - ^{13}C
 - heteronuclear shift correlation and contiguous protonated carbon spin systems
 - 1D reference spectra 319–21
 - “first-tier” methods 316–19
 - fundamental/“first-tier” 2D spectra 322–9
 - intentionally folding HSQC spectra 329–32
 - introduction 315–16
 - long-range heteronuclear shift correlation 335–41
 - NMR spectra 319
 - non-uniform sampling 332–5
 - proton–proton connectivity network 342–53
 - strategy for hyphenated 2D experiments 354
- alkaloids NMR (part I) ^1H - ^{13}C
 - heteronuclear shift correlation and contiguous protonated carbon spin systems – “first-tier” methods
 - parameter choices 318–19
 - probe selection 317–18
 - sample preparation 317
- alkaloids NMR (part I) ^1H - ^{13}C
 - heteronuclear shift correlation and contiguous protonated carbon spin systems – fundamental/“first-tier” 2D spectra
 - BIRD-based homonuclear decoupling in HSQC experiment: PS-HSQC 324–7
 - conventional HSQC 322–3
 - description 322
 - multiplicity-edited HSQC 324
 - one-bond heteronuclear correlation methods 327–9
- alkaloids NMR (part I) ^1H - ^{13}C
 - heteronuclear shift correlation and contiguous protonated carbon spin systems – proton–proton connectivity network
 - COSY/TOCSY spectra 342–6
 - description 342
 - HSQC COSY/TOCSY spectra 346–9
 - HSQMBC-COSY and HSQMBC-TOCSY spectra 349–53

- alkaloids NMR (part II) adjacent carbons using H2BC, 1,1 ADEQUATE and variants – 1,*n*-ADEQUATE and advanced variants 1,*n*-ADEQUATE 373 1,*n*-HD-ADEQUATE 374–7 description 371–3 edited 1,*n*-HD-ADEQUATE 374–80
- alkaloids NMR (part II) adjacent carbons using H2BC, 1,1 ADEQUATE and variants 1,*n*-ADEQUATE and advanced variants 371–80 conclusions 399–400 data 358–71 structure elucidation 380–99
- alkaloids NMR (part II) adjacent carbons using H2BC, 1,1 ADEQUATE and variants – data ADEQUATE experiment 361–71 description 358–9 protonated carbons and H2BC experiment 359–61
- alkaloids NMR (part II) adjacent carbons using H2BC, 1,1 ADEQUATE and variants – structure elucidation cryptospirolepine (structural ambiguity and 1,*n*-HD-ADEQUATE spectra) 386–96 description 380–3 eudistidine-C 397–9 staurosporine (Crews rule) 383–6
- allylestranol 277
- American Society of Natural Products 92
- liquid ammonia 73
- anabolic steroids 276
- androstanes 276
- Angium* sp. 452
- aqualene (terpene) 255
- archazolid A 162
- aromatic solvent-induced shifts (ASIS) 267
- asperdinin 166–7
- Aspergillus unilateralis* 442
- band-selective homodecoupling (BASHD) 208
- “benzene titration” 267–8
- biogenetic isoprene rule 264
- brassinosteroids 276, 298
- Butenandt, Adolf 282
- Candida albicans* 461
- β-carotene 236
- β-caryophyllene 255–6
- cervinomycin 237, 239, 340
- chemical shift anisotropy (CSA) 8–9
- chlorothricin 477–8, 481–2
- cholestanes 276
- cholesterol steroids 276–7 terpenes 255
- Claridge, Timothy 21
- clathria abietina* (marine sponge) 190
- computer-assisted structure elucidation (CASE) eudistidine-C 382 cryptospirolepine 388, 392–6, 399–400 H2BC experiment 360 LR-HSQMBC 237 ¹⁵N data 84–6, 113 small-molecule NMR 200, 237, 244
- conicasterol F 28–30
- contraception (steroids) 282
- cortisone 278
- COSY (correlation spectroscopy) experiments 20, 63, 128, 148, 200, 204, 208–9, 258–60, 284–5, 288, 293, 300–2, 306–7, 342–5, 349–54, 369, 386, 397, 399, 410, 415, 417, 448, 452, 489–90
- Crews rule (hydrogen/heavy atom ratio) 383–6, 399
- Cryptoyolepsis* sp. 386

- cryptospirolepine
 H2BC experiment 359–60
 LR-HSQMBC 237
 structure 82, 85
- cryptospirolepine (structural
 ambiguity and 1,*n*-HD-ADEQUATE
 spectra)
 1,1 and 1,*n*-HD-ADEQUATE
 391–2
 ¹H-¹⁵N HMBC and
 LR-HSQMBC 391
 advanced NMR data and CASE
 program 392–6
 description 386–8
 high-sensitivity experiments
 389–90
 resolution time 399–400
- cryptomisine 85
- cyclosporine 86–7, 87–8, 92–4, 99,
 100, 161
- N*³,5'-cycloxanthosine 426, 435
- density functional theory (DFT)
 28, 40, 55, 133, 152–4, 169, 171,
 237, 388
- desertomycin A 467
- desotamide 472–3
- dexamethasone 277
- dibromopalau'amine 165
- dictyodendrin J 433
- diketopiperazine (DKP) 441, 445
- dioxapyrrolomycin 194–5
- discodermolide 33–4
- DISCON program (NAMFIS) 34
- discorhabdin X 429
- donnazole A/B 80–1
- double-PFG echo (DPFGE) 229, 231
- doxorubicin disaccharide 33–4
- drugs (steroids) 282–3
- dynamic nuclear polarization
 (DNP) – NMR spectroscopy
 3–4
- electric field gradient (EFG) 136
- electron capture dissociation (OCD)
 181
- electron transfer dissociation (ETD)
 181
- enigmazole 44
- enterocin 477
- 5-deoxyenterocin 478
- 13-episteroids 298
- epothilone A 33–4
- estradiol 276
- estrane 276
- estrone 278
- 17 α -ethinylestradiol 277
- Eudisoma* sp. 397
- eudistidine-C 382, 397–9, 400
- fibrosterol sulfate A 35, 167–9
- fidaxomicin 466–7
- fluorine NMR spectroscopy 290–4
- FT-ICR (Fourier transform-ion
 cyclotron resonance) 181, 184–5,
 185–9, 194–5, 195–7
- FT-NMR (Fourier transform-nuclear
 magnetic resonance) 278–80
- gauge-independent atomic orbital
 (GIAO) 55
- geraniol (rose oil) 254
- ginkgolide C (NOE) 5
- gonane (steroids) 276
- Halichondria* sp. 406
- Hassel, Odd 282
- Hassnoot–Altona equation 32–3
- HCNMBBC experiments 87, 108–12,
 113, 243, 318
- Hench, Philip S. 282
- heterofibrin A2 437
- heteronuclear multiple bond
 correlation (HMBC)
 ADEQUATE experiments 241–2,
 243
 carbon–carbon correlation
 data 242
 cryptospirolepine 391, 392
 eudistidine-C 397
 fast NMR 204, 206
 ¹H-¹³C 85–6

- ^1H - ^{13}C shift correlation 320, 335–8, 354
- ^1H - ^{15}N shift correlation 100, 101–4, 107–8, 113
- heronamides 448
- increased adoption 503–4
- kibdelones 452–3
- perfect spin-echo 231–2
- RDCs 159
- saponins 489–90, 491
- steroids 285, 303, 306, 307
- structure determination 262–4
- terpenes 258, 262–3, 267–8
- ultra-long-range correlation NMR 234, 236
- “heteronuclear multiple bond correlation” (“HMQC”) term 503
- heteronuclear multiple quantum coherence (HMQC)
 - cryptospirolepine 390
 - NMR spectra 319
 - small molecule NMR spectroscopy 232
 - terpenes 258, 263, 268
 - ultra-long-range correlation NMR 234
- heteronuclear single quantum coherence (HSQC)
 - alkaloids and ^1H - ^{13}C shift correlation 317, 319–20, 322–7, 328–9, 329–32, 332–5
 - alkaloids 399
 - fast NMR 203–6
 - ^1H - ^{13}C 85–6
 - increased adoption 504
 - lupeol 266
 - multiplicity-edited 399
 - perfect spin-echo 231–2
 - proton-detected experiments 222
 - RDCs 145–8, 162, 166, 168, 171, 172
 - saponins 489–90
 - small molecule NMR spectroscopy 224–6, 229, 232, 243
 - steroids 299, 305–6, 307
 - terpenes 257–8, 261, 267–9, 271–2
 - ultra-long-range correlation NMR 234
 - ZQFs 233–4
- “heteronuclear single quantum coherence” (“HSQC”) term 503
- hit quality index (HQI) 83
- HOBS (homodecoupled bandwidth selective) methodology 216, 218, 220–2, 224, 226, 229–30, 243
- HOBS-TOCSY experiments 218–20
- HSQC-NOESY experiments 24–6, 36, 118, 184–5, 305
- HSQC-TOCSY experiments 224, 233–4, 234–5, 236–9, 270, 284–6, 285, 287, 306, 342, 346–53, 354, 399, 493
- LR (long range) HSQMBC 72, 82, 85, 94, 99–102, 104–6, 236–9, 320, 335, 338–343, 361, 369, 371, 382, 386, 390, 391–2, 396–8, 399, 504
- HSQMBC-TOCSY experiments 87, 94, 98–100, 105, 204, 233, 235, 237, 239–40, 349–53, 386
- ianthellidone F 432
- O*-methyl- α -L-iduronate sodium salt 32
- INADEQUATE experiments 128, 148, 159, 204, 272, 280, 285, 287–9, 302, 306
- inflammation (steroids) 282
- infrared multiphoton dissociation (IRMPD) 181
- isoprene units in terpenes 253, 255–6
- isoquinoline 22
- jaborosalactol 24, 164
- Jaborosa parviflora* 164
- Journal of Natural Products* 21, 505

Kaiser, Reinhold 4

Karplus

equation 33, 118, 282, 295

relationship 40, 45, 67, 491

kauradienoic acid (terpene) 255–6,
260

Kendall, Edward C. 282

lamellarin S 431

lanosterol (terpene) 255

“lattice” term (relaxation in solids)
7–8

laudiamide 33–4

Levenberg–Marquardt procedure 32

Lissodendoryx fibrosa (sponge)
35, 167

ludartin 152, 163–4

lupeol (terpene) 266

lyotropic liquid crystalline (LLC)
solution 131–2, 143

marine natural products NMR

case studies 404–26

introduction 403–4

snapshots 405, 426–38

marine natural products NMR – case
studies

bistelletazines 404, 409–10,
413–16

description 404–5

fascioquinols 405, 420–6

franklinolides 406–11, 424

ircinialactams 405, 417–23

phorbasins 412–120

trachycladindoles 404,
405–6, 409

massadine-14-O-sulfate 428

mass spectrometry (high-resolving
power, high accuracy) for
structural products elucidation
conclusions 197

FT-ICR 181, 184–5, 185–9,
194–5, 195–7

introduction 177–89

literature examples 190–7

unknowns 177–85

mass spectrometry (high-resolving
power, high accuracy) for
structural products elucidation –
literature examples

biosynthesis processes

by measuring stable

isotopic-labelled precursors

(FT-ICR MS) 194–7

triterpene saponins (metabolic
profiling) by accurate mass

measurement and MSⁿ

using LC FT-ICR MS

195–7

ultra-high resolving power to
separate isotopic fine
structures 190–4

mass spectrometry for structural
elucidation of unknowns

accurate mass and NMR

179–80

complexity/difficulty 177–9

isotopic fine-structure and

ultra-high-resolution MS

183–5

MS/MS 180–2

ultra-high resolution and mass
accuracy 183–4

maximum absolute error (MAE) 28

Medicago truncatula 195

menthol 131

methoximes 72–3

methyl orange 75–6

α -trifluoromethylphenylacetic acid
(MTPA) 60–6

methylspirastrellolide A 191–4

microbial natural products NMR

case studies 442–62

introduction 440–2

snapshots 441–2, 463–82

microbial natural products

NMR – case studies

aspergillazines 441, 442–4

heronamides 447–52

kibdelones 441, 452–6

nocardioazines 441, 444–7

reveromycins 441, 457–62

- microbial natural products
 - NMR – snapshots
 - caged structures 463–75
 - cyclic peptides 463
 - macrocyclic lactones 463
 - miscellaneous 475–82
- microcioamide A 190–1
- mirabilin I. 430
- molecular configuration by NMR
 - absolute configuration 60–6
 - chiral anisotropic reagents 60–6
 - conclusions 66–8
 - relative configuration 39–60
- molecular configuration by NMR – relative configuration
 - acetanides 43–4
 - computational methods 55
 - coupling constants and NOEs 40–3
 - description 39–40
 - J*-based analysis – Murata method 44–51
 - residual chemical shift anisotropy 56–60
 - universal database 51–5
- “molecular zipper” in terpenes 268
- ningalin G 433
- nitromethane 72
- NMR analysis of molecular flexibility in solution (NAMFIS) 30–5, 36
- NMR and nitrogenous natural products: ^3H - ^{15}N heteronuclear shift correlation
 - computer-assisted structure elucidation – ^{15}N data 84–6
 - conclusions 113
 - direct/long range 71–2
 - $^1\text{J}_{\text{NH}}/^n\text{J}_{\text{NH}}$ coupling constants 86
 - ^{15}N chemical shift prediction 76–84
 - ^{15}N chemical shift range 73–4
 - ^{15}N chemical shift referencing 72–3
 - ^{15}N pulse widths 74–6
 - results 86–112
- NMR and nitrogenous natural products: ^3H - ^{15}N heteronuclear shift correlation – results
 - ^1H - ^{15}N HSQMBC 95–8
 - ^1H - ^{15}N HSQMBC-TOCSY (long range experiment) 98–107
 - $^1\text{J}_{\text{NH}}$ magnetization transfer 87–91
 - $^n\text{J}_{\text{NH}}$ magnetization transfer 91–5
 - triple-resonance ^1H - ^{15}N experiments 107–12
- NMR and nitrogenous natural products: ^3H - ^{15}N heteronuclear shift correlation – ^{15}N chemical shift prediction
 - description 76–80
 - F_1 spectral windows 82–3
 - ^{15}N content database 83
 - “user trained” database 83–4
 - validation 80–2
- liquid-state NMR spectroscopy 122
- solid-state NMR spectroscopy 126
- Nocardiosis* sp. 444
- NOESY (nuclear Overhauser enhancement spectrometry)
 - experiments 17–18, 24–6, 27, 218, 232, 233, 258, 266, 285, 300–1, 303, 305, 307, 492
- norethisterone 280
- “nuclear Overhauser effect” *see* nuclear Overhauser enhancement
- nuclear Overhauser enhancement (NOE)
 - conclusions 35–6
 - deconvolution – NAMFIS 30–5
 - dimeric compounds 24–6
 - history 4–5
 - kinetic 16–17
 - multispin systems 15–16
 - perfect NOESY experiment 232

- nuclear Overhauser enhancement (NOE) (*continued*)
- protons 118
 - RDCs 117, 119
 - relative configuration 40
 - ROE and ROESY 18–21
 - scalar couplings – zero-quantum filters 31–4
 - small molecules 26–30
 - steroids 280
 - theory 6–15
 - transient 4, 17–18
- “nuclear Overhauser enhancement” term – history 3–5
- oleanic acid (terpene) 255
- oligomycin A 468–9
- Overhauser, Albert W. 3
- PALES program 143, 151, 159, 160–1
- parthenolide 160
- PEG spin-echo (PGSE) 230
- Penicillium* sp. 441, 442
- “Perfect echo CPMG” experiments 229
- “perfect-HMBC” experiments 229
- perfect HSQC Pulse sequence 231–2
- “perfect NMR” spectra 202, 227
- “perfect NOESY” experiments 229
- PFG spin-echo (PGSE) 230
- phenylfamycin E 479–80
- Phorbas* sp. 412
- plakilactones G/H 28–30, 38–40
- platensimycin 467, 475
- platensin 467, 476
- poly- γ -benzyl-L-glutamate (PBLG) 131, 133, 158, 159–60
- poly- γ -ethyl-L-glutamate (PELG) 131, 133
- polydimethylsiloxane (PDMS) 295
- poly(methylmethacrylate) (PMMA) 134, 137–40, 143, 163, 171
- polyprenol-12 268, 271
- poly(vinyl acetate) (PVC) 160
- progesterone 276, 278
- PSYCHE method 213–14, 220–1
- pulsed-field gradients (PSGs) 18, 288
- pulsed-field magnetic gradients (PFGs) 202–3
- pulse sequence developments in small-molecule NMR spectroscopy
- fast NMR 202–6
 - future perspectives 243–4
 - introduction 199–202
 - perfect NMR 227–32
 - pure shift NMR 206–9
 - ultra-long correlation NMR 234–43
 - ZQFs 233–4
 - ZS experiment 209–27
- pulse sequence developments in small-molecule NMR spectroscopy – perfect NMR description 237
- perfect spin-echo 228–32
- ZQFs 233–4
- pulse sequence developments in small-molecule NMR spectroscopy – pure shift NMR description 206–9
- homodecoupling in 1D/2D NMR 218–27
- sensitivity, strong coupling enhancements and spectral quality 210–16
- ZS applications 216–18
- ZS experiment 209–10
- pulse sequence developments in small-molecule NMR spectroscopy – ultra-long correlation NMR
- ADEQUATE 240–3
 - description 234–6
 - HSQMBC-TOCSY 237–40
 - long-range HSQMBC 236–2
- purealin 426–7
- “pure shift NMR” 202
- Reichstein, Tadeus 282
- residual chemical shift anisotropy (RCSA) 50, 56–60, 67

- residual dipolar couplings (RDCs)
 constants of small molecules 200–1
 equation 9
 relative configuration 40
 steroids 282, 294–8
- residual dipolar couplings (RDCs)
 and structural analysis
 alignment media 130–41
 alignment process 142–3
 conclusions 172
 data analysis – small molecules 148–59
 introduction 117–30
 measurement 143–8
 natural products 159–71
 NMR parameters 118, 120–2
- residual dipolar couplings (RDCs)
 and structural analysis (data) – small molecules
 conformational problem 156–8
 experimental uncertainties in RDC analysis 153–5
 possible scenarios 148–53
 software 159
 symmetrical rotors 158–9
- residual dipolar couplings (RDCs)
 and structural analysis – natural products
 description 159–66
 dimeric products 166–71
- retinal (terpene) 255
- ROESY (rotating-frame Overhauser effect spectrometry) experiments 18–21, 28, 218, 258, 266, 300–1, 303, 445, 452, 489, 492
- saggitamide A 162
- SAGs (alignment media) 133
- sapogenin *see* aglycone
- saponins (terpenes) 269
- saponin structure elucidation NMR
 aglycone 489–92
 conclusions 499–500
 general considerations 487–9
 introduction 486–7
 linkage position and sequence (oligosaccharide chain) 497–9
 monosaccharides 493–7
- self-aligned media 135
- sex steroids 276
- small-molecule NMR spectroscopy
see pulse sequence developments
- sodium cholate 130–1
- Solomon equation 11, 16–17
- Solomon, I. 4
- spectral aliased pure shift (SAPS)-HSQC spectra 222–3
- sphaeropsidin A 160
- Spirastrella coccinea* (marine sponge) 191
- spoginstantin-1 34–5
- staurosporine 78–9, 161, 385
- Steroids* (journal) 276
- steroids (NMR)
 conclusions 307
 considerations 298–307
 drugs 282–3
 history 278–80
 introduction 275–83
 modern and “rare” methods 284–98
 prologue 275
- steroids (NMR) – modern and “rare” methods
 1D and high-resolution 1D ^1H methods 288–90
 ^{13}C -detected experiments 287–8
 ^{19}F NMR 290–4
 covariance NMR 284–7
 HSQC-TOCSY experiment 287
 recent developments 284
 residual dipolar couplings 294–8
- Streptomyces antibioticus* 476
- Streptomyces candidus* 473
- Streptomyces platensis* 467
- Streptomyces* sp. 447, 457
- Streptomyces violaceoniger* 475

- Streptomyces viridochromogenes* 473–4
- structure elucidation of natural products (increased adoption) 502–6
- structure–activity relationship (SAR) 406
- strychnine
- aliphatic portion 360–1
 - conformation 28–9, 236–7
 - ^1H - ^{15}N 86–7
 - ^1H - ^{15}N HMBC 72, 94–5
 - ^1H - ^{15}N HSQMBC-TOCSY 99–101, 240–2
 - ^1H - ^{15}N LR-HSQMBC 106, 205, 225, 238
 - H2BC spectrum 360
 - HCNMBC 108–12
 - heteronuclear shift
 - correlation 319–21
 - HSQC spectra 330–1
 - HSQC-TOCSY spectrum 347–8
 - HSQMBC-TOCSY spectrum 240
 - long-range heteronuclear responses 237
 - new NMR methods 316, 320–1
 - orienting properties 133
 - PBLG alignment 160
 - residual chemical shift
 - anisotropy 57–9
 - triple-resonance ^1H - ^{15}N 107
- sucro-neolambertellin 162–3
- sungocine (*Strychnos*) alkaloid 225
- surface induced dissociation 181
- taxol and derivatives 33–4
- telomycin 470
- triterpenes 268, 270
- terpenes: mono, sesqui, and higher terpenes
- conclusions 272
 - introduction 253–4
 - specialized techniques 267–72
 - structure elucidation by NMR spectroscopy 257–61
 - types and related compounds 254–6
 - wrong structure avoidance 261–6
- monoterpenes and cyclohexane rings 259
- testosterone 277–8, 282
- tetramethylsilane 72
- tetanortriterpenoids 256
- thelepamide 50
- Theonella* (marine sponge) 28
- tibolone 280–1, 298, 303
- time of flight (TOF) 179, 181
- TOCSY (total correlation spectroscopy) experiments 218–20, 222, 233, 269–70, 300–1, 305, 306, 307, 342, 489, 493–4
- tomatine 491–2, 495
- Trachyladus laevispirulifer* 405, 426
- Trichoderma virens* 442
- tricyclocohumol 165
- ultra-high resolving power to separate isotopic fine structures
- high-resolution, highly accurate MS and MS^n in methylspirastrellolide A 191–2
 - S atoms determination in microcionamide A 190–1
- vatiparol 170–1
- Vinca* alkaloid derivative 171
- “virtual coupling” 265
- vitamin A (terpene) 255
- WATERGATE (perfect echo) experiments 229–31
- Wieland, Heinrich 282
- Windaus, Adolf 282
- Zangger-Sterk (ZS) experiment 208, 209–10, 213–16, 216–18, 222, 290
- ZQ coherence (ZQC) 233