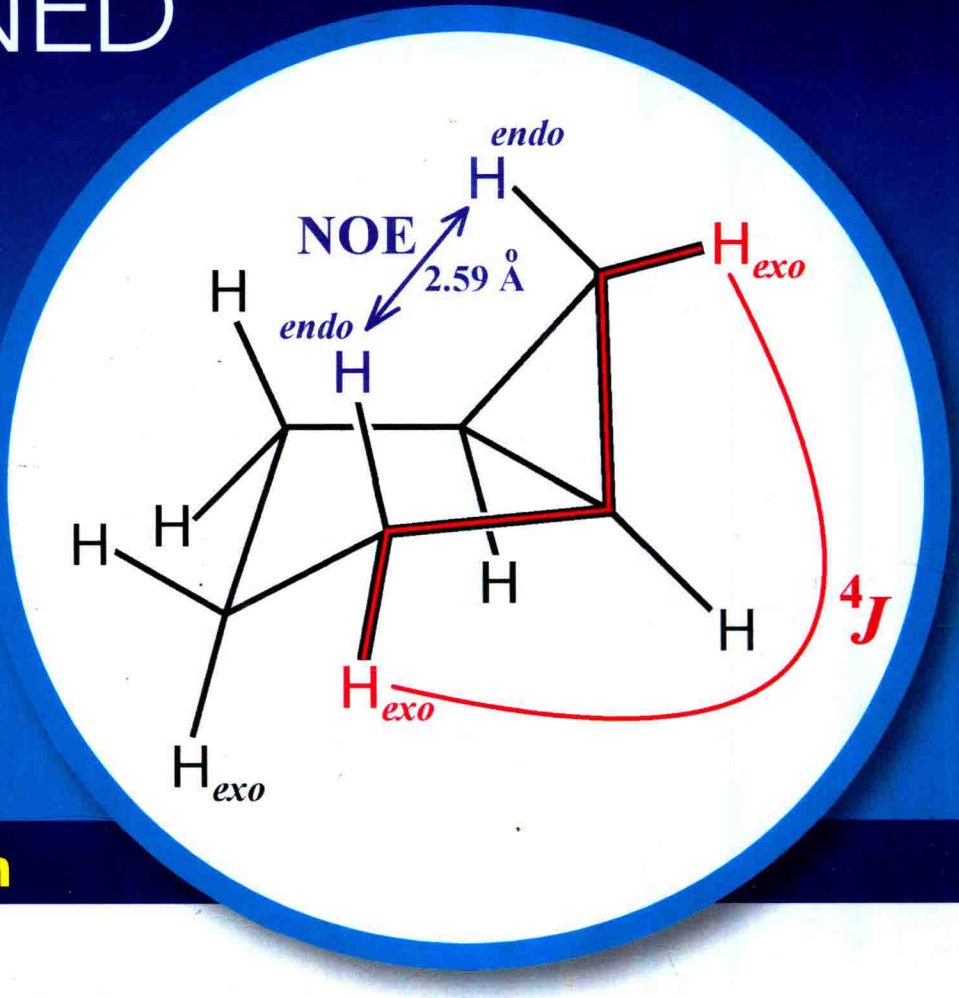
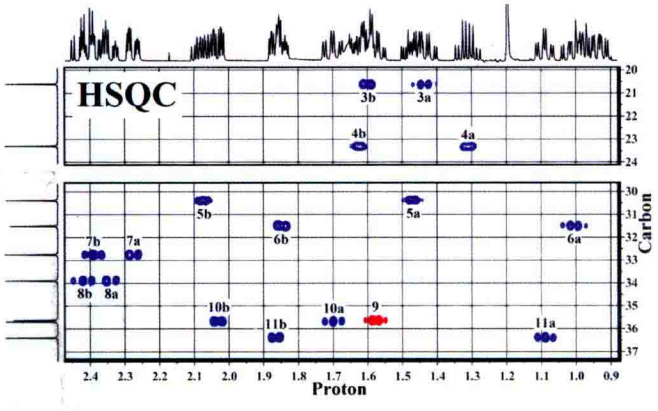
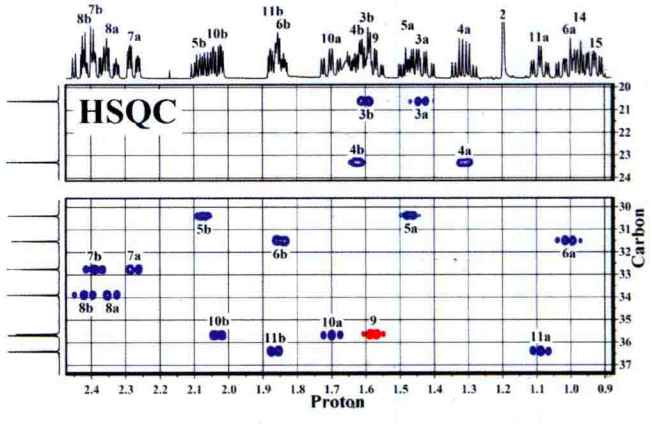
NMR DATA INTERPRETATION EXPLAINED

Understanding 1D and 2D NMR Spectra of Organic Compounds and Natural Products



Neil E. Jacobsen





with website

WILEY

NMR DATA INTERPRETATION EXPLAINED

Understanding 1D and 2D NMR Spectra of Organic Compounds and Natural Products

Neil E. Jacobsen, Ph.D.
University of Arizona

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Names: Jacobsen, Neil E.
Title: NMR data interpretation explained : understanding 1D and 2D NMR spectra of organic compounds and natural products / Neil E. Jacobsen, Ph.D., University of Arizona.
Other titles: Nuclear magnetic resonance data interpretation explained
Description: Hoboken, New Jersey : John Wiley & Sons, Inc., 2016. | Includes bibliographical references and index.
Identifiers: LCCN 2016009193 | ISBN 9781118370223 (cloth)
Subjects: LCSH: Nuclear magnetic resonance spectroscopy. | Nuclear magnetic resonance. | Organic compounds–Spectra. | Spectrum analysis.
Classification: LCC QD96.N8 J325 2016 | DDC 543/.66–dc23 LC record available at https://lccn.loc.gov/2016009193

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

NMR DATA INTERPRETATION EXPLAINED

Examples

Preface

Nuclear magnetic resonance (NMR) spectroscopy is a technique used to determine the structure of molecules at the level of individual atoms and covalent bonds. While it does not provide a direct picture or image of the molecule, the NMR data can be interpreted to determine which atoms in a molecule are connected to which atoms, and whether these bonds connecting them are single, double, or triple bonds. Further information can be obtained from this data about the distances between atoms that are not bonded, and the angles between bonds, leading to a complete three-dimensional model of the molecule.

The field of NMR can be divided into three categories: imaging (MRI), solid-state NMR, and solution-state (liquids) NMR. NMR imaging is familiar to anyone who has gone to a hospital or clinic for an MRI "scan," which yields a picture of "slices" through the human body that is extremely useful in medical diagnosis. Solid-state NMR is the analysis of solid materials, usually ground into a powder; this is applied primarily to the analysis of materials such as polymers, but it can also be applied to biological membranes. Solution-state NMR looks at molecules dissolved in a solvent, which can be water or an organic solvent such as acetone or chloroform. This book is focused on solution-state NMR, the primary tool used by organic chemists and biochemists to determine molecular structure.

A further distinction is made between "small molecules" and "large molecules" in solution. In the context of solution-state NMR, a large molecule is a biological molecule such as a protein or nucleic acid, made up of many repeating units that all have similar structures. A small molecule has a molecular weight less than 1000 Da and is usually made up of diverse structural elements (carbon chains, rings, and functional groups) rather than a repeating pattern. Small molecules are the domain of the organic chemist: natural products, drugs, and the intermediates and products of organic synthesis. Also included in this category are the short chains of biological molecules: peptides, oligonucleotides, and oligosaccharides (sugars). This book will focus on the use of NMR data to determine the covalent structure (which atoms are connected to which atoms) and three-dimensional shape (stereochemistry and conformation) of these small molecules.

This book is different from most books on NMR because it is focused on *examples* and *exercises*. Each topic is introduced with one of more examples of NMR data with detailed explanations of the interpretation of that data. Examples are then followed by a number of exercises using detailed images of NMR data, and these are followed by solutions, again with detailed explanation of the step-by-step reasoning used to solve the exercise. The title, *NMR Data Interpretation Explained*, is an indication of this focus on example and explanation. Every detail and aspect of the NMR data is explained, not just the simple and beautiful spectra but also the complex and surprising spectra. A large number of additional exercises, almost all of them showing detailed graphics of NMR data, have been provided at www.wiley.com/go/jacobsen/nmrdata. Solutions with detailed explanations are provided for half of the exercises, with the remaining solutions provided to instructors on the same website in a forum accessible by instructors only. All of the commonly used techniques of small-molecule solution-state NMR are covered: simple one-dimensional (^{1}H and ^{13}C), edited (DEPT) ^{13}C, selective one-dimensional ^{1}H (NOE, ROE, and TOCSY), and two-dimensional (COSY, TOCSY, NOESY, ROESY, HSQC, and HMBC). The final chapter puts all of these techniques together to solve the structures of a number of complex natural products: sesquiterpenes, steroids, alkaloids, sugars, and triterpenes. Many exercises are provided for each of these molecule types.

Another unique aspect of this book is that it does not attempt to explain the theory of NMR. Other books, including my own book (*NMR Spectroscopy Explained*, Wiley-Interscience, 2007), do an excellent job of explaining the theoretical basis of NMR and how the experiments actually work to give the NMR data. In my experience, the actual users of NMR spectrometers are more interested in solving a chemical problem using NMR data, and have little interest in how the spectrometer works or how the nuclei respond to magnetic fields and radio frequency pulses. It is for these NMR users, industry researchers as well as undergraduates, graduate students, and postdoctoral researchers in chemistry, biochemistry, medicinal chemistry, and pharmacy, that this book was written.

The NMR data used in this book came primarily from the NMR facility in the Department of Chemistry and Biochemistry at the University of Arizona. The instruments used include a Bruker Avance-III (400.13 MHz), a Bruker DRX-500 (499.28 MHz), a Bruker DRX-600 (600.13 MHz), and a Varian Inova-600 (599.7 MHz) with cryogenic probe.

Every attempt was made to obtain the highest-quality NMR data from pure samples. Data was processed using the Felix software package (Felix NMR, Inc., San Diego, CA) and the MestReNova software package (MestReLab Research, Santiago de Compostela, Spain). Literature data was also used, downloaded from the Japanese database SDBS (Spectral Database for Organic Compounds, National Institute of Advanced Industrial Science and Technology, AIST). In a few cases, NMR spectra were simulated using parameters (chemical shifts and *J* values) obtained from the literature.

NMR spectrometers are expensive (around $800,000 for a 600 MHz instrument), and require specialized expertise and expensive cryogens (liquid nitrogen and liquid helium) to operate, so many teaching institutions are unable to obtain a high-field NMR instrument. It was also with these colleges and universities in mind, all over the world, that this book was written, so that students can learn the technique using high-quality data from a wide variety of samples.

Acknowledgments

The idea for this book came from a Chemistry course created by Professor Eugene Mash at the University of Arizona. The course, Chemistry 447, is a laboratory course in the identification of organic compounds, and over the years the technique used by students has become almost exclusively NMR. Prof. Mash gathered together an amazing collection of unknown samples, including a large number of simple aromatics and monoterpenes, and more than 50 different steroids. I began giving a series of lectures on two-dimensional NMR in this course in 2006, and gradually acquired complete 1D and 2D data sets at 600 MHz for all of the steroid unknowns. Prof. Mash encouraged me to write a book that would include this data as well as data on a large number of organic compounds, so that students all over the world, especially in small colleges and in developing countries, would have access to high-quality 600 MHz NMR data.

In 2012, a new graduate course was created by Professor Hamish Christie at the University of Arizona, aimed at preparing new graduate students in Organic Chemistry for their research work. The course, Chemistry 545, teaches all of the latest laboratory techniques in organic synthesis while using the synthetic intermediates and products to teach students to use our NMR instruments and to interpret the NMR data. In this course I developed a deeper look at one-dimensional proton NMR data, beyond the simple spectra found in most undergraduate courses. Two of these laboratory experiments—isolation of the α- and β-isomers of the monoterpene thujone from cedar leaf oil, and preparation of a Shi oxidation catalyst from fructose—adapted well to teaching selective NOE and 2D NMR experiments, forming the core of the more advanced portions of this book.

I would like to thank Prof. Mash and Prof. Christie for these unique opportunities to develop an NMR curriculum and to gain years of experience in explaining and discussing NMR data with undergraduate and graduate students.

I also thank Prof. Robert Bates and Prof. Leslie Gunatilaka, both experts in natural product isolation and structure elucidation, for many exciting collaborations that ignited my fascination with using NMR to solve these complex structures. In the course of these studies, I developed the systematic method outlined in this book for solving structure problems using NMR data.

Dr. Jixun Dai, Assistant Director of the NMR Facility at the University of Arizona, prepared a large number of samples and ran the NMR experiments for those samples. He optimized many of the experiments on the Bruker DRX-500 and DRX-600 instruments, doing especially difficult work of implementing the most modern versions of the selective TOCSY and selective NOE experiments. His programming and data handling skills also saved me more than once from challenging issues in using old NMR data from obsolete platforms, and in simulation of NMR data. I thank him for the significant contribution he made to this book.

A large number of 1D ^{1}H and ^{13}C exercises in this book came from literature data provided by the National Institute of Advanced Industrial Science and Technology (AIST, Japan). Their website (SDBSWeb: http://sdbs.riodb.aist.go.jp) is a goldmine of NMR data for a wide variety of organic compounds. Their line lists (lists of NMR line frequencies) were used to reconstruct the literature spectra used in these exercises (*e.g.*, 300 and 399.65 MHz ^{1}H spectra). I am grateful for being able to use this data for educational purposes.

Finally, I would like to thank my wife, Dr. Linda Breci, for her unwavering support and patience, especially in the last year, as I completed this enormously time-consuming project. She also taught me what little I know about mass spectrometry (MS) and helped me with the section on MS, and she compiled the index of this book.

ABOUT THE COMPANION WEBSITE

This book is accompanied by a companion website:

www.wiley.com/go/jacobsen/nmrdata

The Student's website includes:

- Additional Chapter Exercises
  - A large number of exercises are provided, many showing detailed graphics of NMR data
- Solutions to Exercises
  - With detailed explanations are provided for half of the exercises

The Instructor's website includes:

- Instructor's Solutions Manual
  - Provides remaining solutions to exercises

Contents

CHAPTER 1

Spectroscopy and the Proton NMR Experiment

1 WHAT IS THE STRUCTURE OF A MOLECULE?

There are several levels of understanding what a molecule "looks like" on the scale of individual atoms. The first step is to understand how many of each type of atom make up the collection of atoms that are bonded together to form a molecule. The *molecular formula* is an accounting of the types of atoms in a molecule and the number of each type of atom (*e.g.*, $C_6H_8N_2O_4$). **Mass spectrometry** is used to "weigh" molecules and obtain their exact mass, in atomic mass units (amu). Because atoms have masses that can differ slightly from integer values (*e.g.*, $^{1}H = 1.007825$ amu, $^{12}C = 12.000000$, $^{16}O = 15.994915$, $^{14}N = 14.003074$), a very precise measurement of the mass of a molecule allows us to determine the molecular formula. With a molecular formula, we can start to think about how this group of atoms is connected together. For example, for C_4H_6O (Figure 1.1) we can think of many ways to connect the atoms, while satisfying the valence rules (four bonds to C, two to O, one to H).

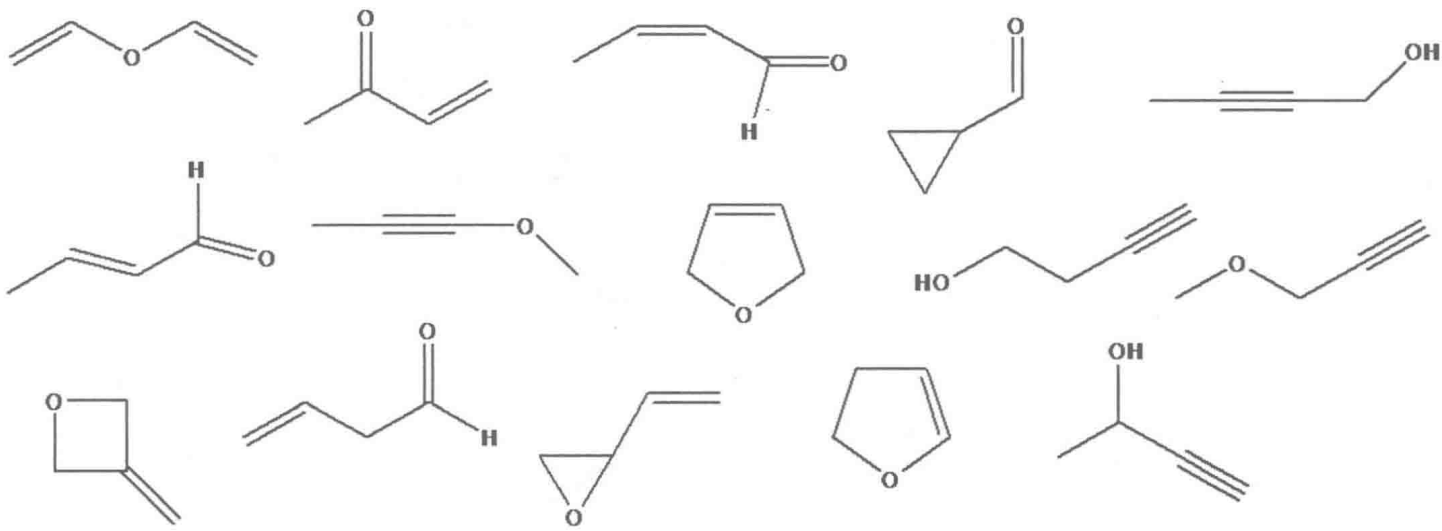


FIGURE 1.1

NMR Data Interpretation Explained: Understanding 1D and 2D NMR Spectra of Organic Compounds and Natural Products, First Edition. Neil E. Jacobsen.

Companion website: www.wiley.com/go/jacobsen/nmrdata

Note that all of the C_4H_6O structures in Figure 1.1 have one thing in common: the total of the number of π bonds plus the number of rings is two in each case. These two "unsaturations" can be determined from the molecular formula by a simple calculation:

1. Discard the oxygen(s): $C_4H_6O \rightarrow C_4H_6$.
2. Any halogens (F, Cl, Br, I) are converted to hydrogens.
3. Any nitrogens (N) are converted to CH (one C and one H for each N). You now have the modified molecular formula: C_4H_6.
4. If **n** is the number of carbon atoms in the modified molecular formula (C_n), calculate the number of hydrogens expected in a saturated hydrocarbon with this number of carbons: $\mathbf{m} = (\mathbf{n} \times 2) + 2 = (4 \times 2) + 2 = 10$.
5. Subtract the number of hydrogens in the modified molecular formula (6) from this saturated hydrocarbon value and divide the result by 2: $\mathbf{m} - 6 = 10 - 6 = 4$; $\mathbf{u} = 4/2 = 2$.

This result (**u**) is equal to the number of π bonds in the molecule *plus* the number of rings. Note that a triple bond (C≡C) is really one σ bond and two π bonds, so it counts as two "unsaturations".

For larger molecules the number of isomers (structures with the same molecular formula) increases very rapidly with the number of atoms. For the formula $C_8H_{11}NO_3$ there are 383 different commercially available compounds! NMR is especially useful for distinguishing between these many possibilities.

In the NMR instrument, each atom (actually the nucleus of each atom) has a precise resonant frequency in the radio frequency spectrum. We can "tune in to the radio channel" of each of these atoms in turn and gather information about the immediate surroundings of that atom in the molecule. There are several kinds of information we can get from each atom:

1. Nearby functional groups change the resonant frequency in predictable ways, so the exact resonant frequency can be used to determine the "chemical environment" of that atom. There are two types of these frequency-shifting effects:
   a. Nearby electronegative atoms (O, N, Br, *etc.*). This effect acts through σ bonds and dies off quickly after 2 or 3 bonds. This is similar to the well-known inductive effect that modifies reactivity in organic chemistry reactions.
   b. Nearby double bonds (C=C or olefin/aromatic, C=O or carbonyl, C≡N or nitrile, *etc.*). This effect acts directly through space and dies off after about 5 Ångstroms (one Ångstrom or Å is approximately the length of a C—H bond). The orientation of the plane of the double bond relative to the atom being observed is also important.
2. Hydrogen atoms are affected by the proximity of other hydrogen atoms in the molecule. So we can look around the immediate vicinity of *our* hydrogen (the one whose radio channel we are tuned to) and see the number and proximity of other hydrogens or groups of hydrogens. This effect manifests itself in two ways:
   a. "Splitting" of the resonant frequency of *our* hydrogen (the one being observed) by a nearby hydrogen into two resonant frequencies very close to each other. The stronger the effect, the wider is the separation of the two frequencies. This effect travels through the bonds and dies off quickly as the number of bonds separating the two hydrogens increases: 2 bonds ≥ 3 bonds > 4 bonds. This effect is sensitive to the angles formed by the bonds connecting the two hydrogens, so we can get information about the relative orientation of groups connected by single bonds. These can either be fixed orientations determined by rigid bonding in rings (stereochemistry) or preferred orientations in a flexible molecule (conformation).
   b. Enhancement of the NMR radio signal received from one hydrogen when we hit the other hydrogen with a radio signal at its precise radio frequency. This enhancement is called an **NOE** and it operates directly through space between hydrogens. The effect dies off quickly with increasing separation and is not seen at all for distances greater than 5 Å. The NOE gives us a molecular ruler for measuring distances between specific pairs of hydrogens in the molecule.

Note that the NMR experiment gives us lots of specific information from the point of view of one atom in the molecule: nearby functional groups and nearby hydrogens, through bonds or directly through space. We can get