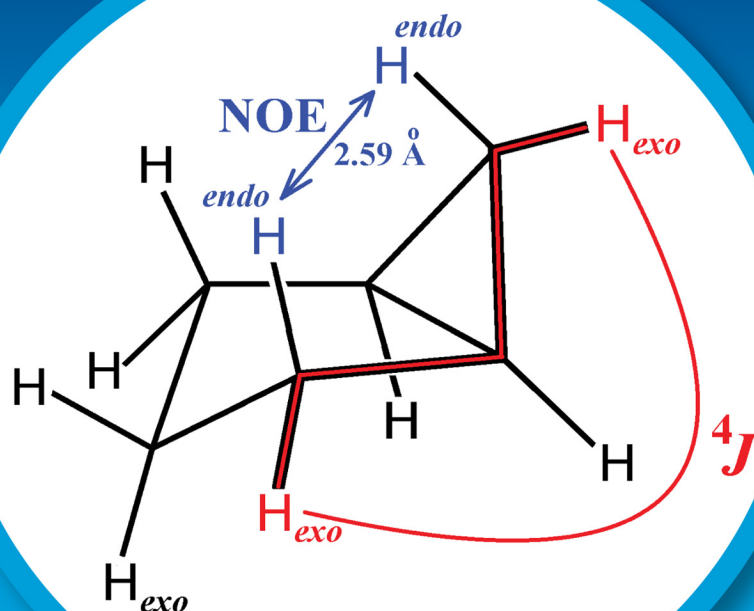
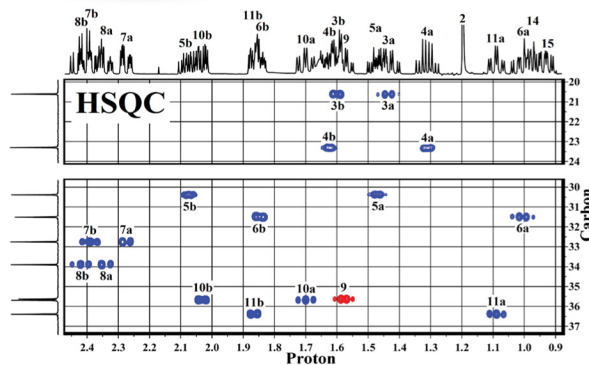
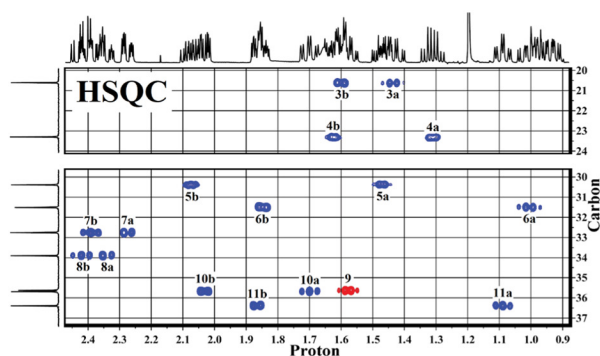


# NMR DATA INTERPRETATION EXPLAINED

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



Neil E. Jacobsen



WILEY



# **NMR DATA INTERPRETATION EXPLAINED**



# **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**

Copyright © 2017 by John Wiley & Sons, Inc. All rights reserved.

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

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at [www.copyright.com](http://www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

**Limit of Liability/Disclaimer of Warranty:** While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at [www.wiley.com](http://www.wiley.com).

***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

# Contents

*Examples xi*

*Preface xiii*

*Acknowledgments xv*

*About the Companion Website xvii*

## **Chapter 1 | Spectroscopy and the Proton NMR Experiment 1**

- 1 What is the Structure of a Molecule? 1**
- 2 Mass Spectrometry 3**
  - 2.1 Ionization Methods and Molecular Ions 4
    - 2.1.1 Electron Impact (EI) 4
    - 2.1.2 Soft Ionization 5
  - 2.2 High-Resolution Mass Spectrometry and Exact Mass 5
  - 2.3 Isotope Patterns and the Halogens Br and Cl 7
- 3 Infrared (IR) Spectroscopy 9**
- 4 Ultraviolet (UV) and Visible Spectroscopy 10**
- 5 A Highly Simplified View of the NMR Experiment 13**

## **Chapter 2 | Chemical Shifts and Splitting Patterns 17**

- 1 Chemical Shifts in the Proton Spectrum 17**
- 2 Splitting: The Effect of One Neighbor: A Doublet 21**
- 3 Splitting: The Effect of Two Neighbors: A Triplet 23**
- 4 Splitting: The Effect of Three Neighbors: A Quartet 25**
- 5 Splitting: The Effect of “n” Neighbors: A Multiplet 30**
- 6 Using Splitting Patterns to Choose from a Group of Isomers 34**
- 7 Peak Intensities (Peak Areas) and the Number of Protons in a Peak 37**
- 8 Publication Format for Proton NMR Data 39**
- 9 Recognizing Common Structure Fragments 41**
- 10 Overlap in Proton NMR Spectra. Example: 1-Methoxyhexane 45**
- 11 Protons Bound to Oxygen: OH Groups. Example: 2-Ethyl-1-Butanol 48**
- 12 Summary of Chemical Shifts and Splitting Patterns 50**

## **Chapter 3 | Proton (<sup>1</sup>H) NMR of Aromatic Compounds 51**

- 1 Benzene: The Aromatic Ring Current and the Shielding Cone 51**
- 2 Monosubstituted Benzene: X-C<sub>6</sub>H<sub>5</sub> 52**
  - 2.1 Toluene 52
  - 2.2 Aromatic Chemical Shifts: Resonance Structures 54
  - 2.3 Nitrobenzene 55
  - 2.4 Anisole 56

- 2.5 Substituent Effects on Aromatic Chemical Shifts 58
- 2.6 Long-Range  $J$  Couplings in Aromatic Rings: Protons 4 Bonds Apart 59
- 3 Disubstituted Benzene:  $X-C_6H_4-Y$  62**
  - 3.1 Symmetrical Disubstituted Benzene:  $X-C_6H_4-X$  62
  - 3.2 Unsymmetrical Disubstituted Benzene,  $X-C_6H_4-Y$  72
    - 3.2.1 *para* (1,4) Disubstituted Benzene:  $p-X-C_6H_4-Y$  73
    - 3.2.2 *meta* (1,3) Disubstituted Benzene:  $m-X-C_6H_4-Y$  78
    - 3.2.3 *ortho* (1,2) Disubstituted Benzene:  $o-X-C_6H_4-Y$  87
- 4 Coupling Between Aromatic Ring Protons and Substituent Protons; Homonuclear Decoupling 100**
  - 4.1 The Methyl Group ( $CH_3$ ) 100
  - 4.2 The Methoxy Substituent ( $OCH_3$ ) 102
  - 4.3 The Formyl ( $H-C=O$ ) Substituent 103
- 5 Trisubstituted Aromatic Rings: The  $AB_2$  System 106**
- 6 Other Aromatic Ring Systems: Heteroaromatics, Five-Membered Rings and Fused Rings 110**
  - 6.1 Pyridine ( $C_5H_5N$ ) 111
  - 6.2 Pyrrole ( $C_4H_5N$ ) 112
  - 6.3 Furan ( $C_4H_4O$ ) 113
  - 6.4 Naphthalene ( $C_{10}H_8$ ) 115
  - 6.5 Indole ( $C_8H_7N$ ) 117
  - 6.6 Quinoline and Isoquinoline ( $C_9H_7N$ ) 118
- 7 Summary of New Concepts: Proton NMR of Aromatic Compounds 120**

## Chapter 4 | Carbon-13 ( $^{13}C$ ) NMR

125

- 1 Natural Abundance and Sensitivity of  $^{13}C$  125**
- 2 Proton Decoupling—Removing the Splitting Effect of Nearby Protons 126**
- 3 Intensity of  $^{13}C$  Peaks—Symmetry and Relaxation 126**
- 4 Chemical Shifts of Carbon-13 ( $^{13}C$ ) Nuclei 129**
  - 4.1  $^{13}C$  Frequency and Chemical Shift Reference 129
  - 4.2 General Regions of the  $^{13}C$  Chemical Shift Scale 130
  - 4.3 Correlations between  $^1H$  and  $^{13}C$  Chemical Shift for a C-H Pair 132
  - 4.4 Quantitation of the Steric Effect for  $^{13}C$  Chemical Shifts 135
  - 4.5 Example of Steric Effects on  $^{13}C$  Chemical Shifts: The “Crowded CH” in Steroids 141
  - 4.6 The  $\gamma$ -*gauche* Effect: Steric Shifts That Give Stereochemical Information 143
  - 4.7 Inductive Effects in  $^{13}C$  Chemical Shifts: Electronegative Atoms 147
  - 4.8 The Effect of Ring Strain on  $^{13}C$  Chemical Shift of  $sp^3$ -Hybridized Carbons 150
- 5 Quaternary Carbons: the Carbonyl Group 151**
- 6 Simple Aromatic Compounds: Substituent Effects on  $^{13}C$  Chemical Shifts 156**
- 7 Highly Oxygenated Benzene Rings and Coumarin 161**
- 8 Fused Rings and Heteroaromatic Compounds 165**
  - 8.1 Pyridine ( $C_5H_5N$ ) 165
  - 8.2 Pyrrole ( $C_4H_5N$ ) 167
  - 8.3 Furan ( $C_4H_4O$ ) 168
  - 8.4 Naphthalene ( $C_{10}H_8$ ) 168
  - 8.5 Indole ( $C_8H_7N$ ) 170
  - 8.6 Quinoline and Isoquinoline ( $C_9H_7N$ ) 173
- 9 Edited  $^{13}C$  Spectra: DEPT 174**
  - 9.1 Non-decoupled  $^{13}C$  Spectra 175
  - 9.2 Edited  $^{13}C$  Spectra 176
  - 9.3 Practical Details of the DEPT Experiment 181
    - 9.3.1 Sensitivity 181



9.3.2	Pulse Calibration	181
9.3.3	<i>J</i> Value Setting	182
9.3.4	Phase Correction	185
<b>10</b>	<b>The Effect of Other Magnetic Nuclei on the <sup>13</sup>C Spectrum: <sup>31</sup>P, <sup>19</sup>F, <sup>2</sup>H and <sup>14</sup>N</b>	<b>185</b>
10.1	Splitting of <sup>13</sup> C Peaks By Deuterium ( <sup>2</sup> H)	185
10.2	Splitting of <sup>13</sup> C Peaks by Phosphorus ( <sup>31</sup> P)	186
10.3	Splitting of <sup>13</sup> C Peaks by Fluorine ( <sup>19</sup> F)	188
10.4	Splitting and Broadening of <sup>13</sup> C Peaks by Nitrogen ( <sup>14</sup> N)	189
<b>11</b>	<b>Direct Observation of Nuclei Other Than Proton (<sup>1</sup>H) and Carbon (<sup>13</sup>C)</b>	<b>190</b>
11.1	Phosphorus-31 ( <sup>31</sup> P) NMR	192
11.2	Fluorine-19 ( <sup>19</sup> F) NMR	194
<b>Chapter 5</b>	<b>  Alkenes (Olefins)</b>	<b>198</b>
1	Proton Chemical Shifts of Simple Olefins	199
2	Short-Range (Two and Three Bond) Coupling Constants ( <i>J</i> Values) in Olefins	202
3	The Allylic Coupling: A Long-Range (Four-Bond) <i>J</i> Coupling	205
4	Long-Range Olefin Couplings in Cholesterol: The <i>bis</i> -Allylic Coupling ( <sup>5</sup> <i>J</i> )	209
5	Carbon-13 Chemical Shifts of Hydrocarbon Olefins (Alkenes)	210
6	Resonance Effects on Olefinic <sup>13</sup> C Chemical Shifts	214
7	Alkynes	225
<b>Chapter 6</b>	<b>  Chirality and Stereochemistry: Natural Products</b>	<b>227</b>
1	The Molecules of Nature	227
2	Chirality, Chiral Centers, Chiral Molecules, and the Chiral Environment	230
3	The AB System	232
4	Detailed Analysis of the AB Spectrum: Calculating the Chemical Shifts	234
5	The ABX System	237
6	Variations on the ABX Theme: ABX <sub>3</sub> , ABX <sub>2</sub> and ABXY	245
7	The Effect of Chirality on <sup>13</sup> C Spectra. Diastereotopic Carbons	249
8	A Closer Look at Chemical Shift Equivalence in an Asymmetric Environment	251
8.1	Chemical Shift Equivalence of CH <sub>3</sub> Group Protons	251
8.2	Non-Equivalence of CH <sub>2</sub> Group Protons	252
8.3	Chemical Shift Equivalence by Symmetry	252
9	<i>J</i> Couplings and Chemical Shifts in the Rigid Cyclohexane Chair System	255
9.1	Cyclohexene and Cyclohexenone	262
10	A Detailed Look at the Dependence of <sup>3</sup> <i>J</i> <sub>HH</sub> on Dihedral Angle: The Karplus Relation	266
11	Magnetic Non-Equivalence. The X-CH <sub>2</sub> -CH <sub>2</sub> -Y Spin System: A <sub>2</sub> B <sub>2</sub> and AA'BB' Patterns	276
12	Bicyclic Compounds and Small Rings (Three- and Four-Membered)	286
12.1	The Bicyclo[2.2.1] Ring System	286
12.2	The Bicyclo[3.1.0] Ring System	291
12.3	The Bicyclo[3.1.1] Ring System	294
	Reference	298
<b>Chapter 7</b>	<b>  Selective Proton Experiments: Biological Molecules</b>	<b>299</b>
1	Sugars: Monosaccharides and Oligosaccharides	299
2	Slowing of OH Exchange in Polar Aprotic Solvents Like DMSO	305
3	Selective TOCSY Applied to the Assignment of the <sup>1</sup> H Spectra of Sugars	307

- 4 **The Selective NOE (Nuclear Overhauser Effect) Experiment** 319
  - 4.1 Recognizing Artifacts in Selective NOE Spectra 320
  - 4.2 The Relationship Between NOE Intensity and Distance 320
  - 4.3 Magnetization Transfer in the Selective TOCSY and Selective NOE Experiments 321
- 5 **Amino Acids and Peptides** 331
- 6 **Nucleic Acids** 348
- 7 **Parameter Settings for NMR Experiment Setup and NMR Data Processing** 357
- Bibliography** 358

## Chapter 8 | *Homonuclear Two-Dimensional NMR: Correlation of One Hydrogen (<sup>1</sup>H) to Another* 359

- 1 **Selective TOCSY Experiments Displayed as a Stacked Plot** 359
- 2 **The Two-Dimensional COSY Experiment** 365
- 3 **Shape and Fine Structure of COSY Crosspeaks; Contour Plots** 370
- 4 **2D-COSY Spectra of Sugars** 376
- 5 **2D-COSY Spectra of Aromatic Compounds** 391
- 6 **Parameter Settings in the 2D COSY Experiment; The DQF-COSY Experiment** 397
- 7 **COSY Spectra of Peptides** 399
- 8 **COSY Spectra of Natural Products** 405
- 9 **Two-Dimensional (2D) TOCSY (Total Correlation Spectroscopy)** 412
- 10 **Two-Dimensional (2D) NOESY (Nuclear Overhauser Effect Spectroscopy)** 423
- Parameter Settings Used for 2D Spectra in this Chapter** 429

## Chapter 9 | *Heteronuclear Two-Dimensional NMR: Correlation of One Hydrogen (<sup>1</sup>H) to One Carbon (<sup>13</sup>C)* 430

- 1 **3-Heptanone: A Thought Experiment** 430
- 2 **Edited HSQC: Making the CH<sub>2</sub> Protons Stand Out** 436
- 3 **The 2D-HSQC Spectrum of Cholesterol** 443
- 4 **A Detailed Look at the HSQC Experiment** 455
- 5 **Parameters and Settings for the 2D-HSQC Experiment** 458
  - 5.1 Spectral Window 458
  - 5.2 Acquisition Time 458
  - 5.3 One-Bond *J* Coupling Value 459
  - 5.4 Number of 1D Spectra Acquired: *F*<sub>1</sub> Resolution 460
  - 5.5 Number of Scans: Sensitivity 460
- 6 **Data Processing: Phase Correction in Two Dimensions** 460
- 7 **Long-Range Couplings between <sup>1</sup>H and <sup>13</sup>C** 463
- 8 **2D-HMBC (Heteronuclear Multiple-Bond Correlation)** 465
  - 8.1 2D-HMBC Spectra of Aromatic Compounds 467
  - 8.2 HMBC Spectra of Natural Products: Using the Methyl Correlations 475
  - 8.3 HMBC Spectra of Sugars 491
- 9 **Parameters and Settings for the 2D-HMBC Experiment** 495
  - 9.1 Spectral Window 495
  - 9.2 Acquisition Time 496
  - 9.3 One-Bond and Long-Range *J*<sub>CH</sub> Coupling Values 496
  - 9.4 Number of Scans 496
- 10 **Comparison of HSQC and HMBC** 496
- 11 **HMBC Variants** 497

---

Parameter Settings Used for 2D Spectra in this Chapter 497  
References 498

**Chapter 10 | *Structure Elucidation Using 2D NMR* 499**

1 Literature Structure Problems 500  
2 Sesquiterpenoids 501  
3 Steroids 522  
4 Oligosaccharides 552  
5 Alkaloids 574  
6 Triterpenes 597  
Reference 615

*Index* 617



# Examples

<i>n</i> -Propyl Acetate	32
1-Methoxyhexane	45
2-Ethyl-1-Butanol	48
<i>n</i> -Propyl Acetate: Mismatch of Nearly Equal <i>J</i> Values.	69
Acetaminophen ( <i>para</i> -acetaminophenol)	74
Methyl <i>m</i> -Nitrobenzoate	79
Acetylsalicylic Acid (Aspirin)	88
<i>ortho</i> -Chloroacetophenone ( <i>o</i> -Cl-C <sub>6</sub> H <sub>4</sub> -CO-CH <sub>3</sub> ): Strong Coupling Effects	95
2,6-Di-( <i>o</i> -anisyl)-anisole, a <i>meta</i> -terphenyl: the AB <sub>2</sub> System	106
Saturated Fatty Acids: Methyl Stearate	138
Methyl β-D-Riboside 2,3-Acetonide	147
N-( <i>tert</i> -Butoxycarbonyl)-L-Glutamine	152
Anisole (methoxybenzene): <sup>13</sup> C Substituent Effects	157
Acetaminophen ( <i>para</i> -acetamidophenol): <sup>13</sup> C Substituent Effects	158
Methyl <i>meta</i> -Nitrobenzoate (methyl 3-nitrobenzoate): <sup>13</sup> C Substituent Effects	159
<i>cis</i> -6a-Hydroxymaackiain Analog: <sup>13</sup> C Substituent Effects	160
Phlorizin	162
4-Methoxy-1-Naphthaldehyde: <sup>13</sup> C Spectrum	169
Indole-2-Carboxylic Acid	171
Cinchonidine	174
1,1,1-Trifluoro-2-Propanol	188
Fluorobenzene	188
Triphenylphosphine Oxide	188
Adenosine Triphosphate (ATP): <sup>31</sup> P Spectrum	192
4-Fluorobenzonitrile (C <sub>7</sub> H <sub>4</sub> NF): <sup>19</sup> F Spectrum	194
Perfluorooctanesulfonic acid (PFOS, C <sub>8</sub> F <sub>17</sub> SO <sub>3</sub> H): <sup>19</sup> F Spectrum	196
Cinnamic Acid	203
<i>n</i> -Butyl Acrylate (CH <sub>3</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -O-CO-CH=CH <sub>2</sub> )	204
1-Octene (CH <sub>3</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH=CH <sub>2</sub> )	206
Limonene (C <sub>10</sub> H <sub>16</sub> )	212
Linalool (C <sub>10</sub> H <sub>18</sub> O)	215
6α-Methylprednisolone (C <sub>22</sub> H <sub>30</sub> O <sub>5</sub> )	217
Elemol (C <sub>15</sub> H <sub>26</sub> O)	220
2-Hydroxy-2-methylsuccinic acid: the AB System	232
2-Hydroxysuccinic acid: the ABX System	241
Diethyl Ethyl(phenyl)malonate: ABX <sub>3</sub>	245
2-Butanol ( <i>sec</i> -butanol): ABXY <sub>3</sub>	248
Cholesterol (A Ring)	259
3- <i>n</i> -Butylcyclohexanone	262
1- <i>n</i> -Butylcyclohex-2-en-1-ol	262
A Synthetic Pyrrolidine: Determining the Major and Minor Diastereomers	267
Menthol	269
Hydrocinnamic Acid: AA'BB'	276
3-Phenyl-1-propanol: AA'BB'	279
<i>trans,trans</i> -1,4-diphenyl-1,3-butadiene: AA'BB'	281
<i>cis,trans</i> -1,4-diphenyl-1,3-butadiene: First Order	284

<i>endo</i> -Borneol	287
Synthetic Bicyclo[2.2.1] Compounds	288
$\alpha$ -Thujone and $\beta$ -Thujone	292
$\alpha$ -Pinene (Unknown Monoterpene)	295
D-Glucose in D <sub>2</sub> O	301
Methyl $\alpha$ -D-Glucoside in CD <sub>3</sub> OD	304
Methyl $\alpha$ -D-Glucoside in DMSO-d <sub>6</sub>	306
Methyl $\alpha$ -D-Glucoside (1D TOCSY): Selection of H1 Peak	309
Methyl $\alpha$ -D-Glucoside (1D TOCSY): Sequential Selection of All Peaks	311
$\beta$ -D-Glucose in D <sub>2</sub> O: Assignment of the <sup>1</sup> H Spectrum Using Selective TOCSY	313
$\alpha$ -D-Glucose in D <sub>2</sub> O: Assignment of the <sup>1</sup> H Spectrum Using Selective TOCSY	314
5-Ethyl-2-Methylpyridine: Regiochemistry of Alkyl Substituents by NOE	322
Thujone Isomers ( $\alpha$ and $\beta$ ): Stereochemistry of CH <sub>3</sub> Group by NOE	324
Sucralose: Selective NOE Experiment	329
Aspartame (L-Asp-L-Phe-OCH <sub>3</sub> )	333
Leucine Enkephalin Acetate (Tyr-Gly-Gly-Phe-Leu) in DMSO-d <sub>6</sub>	336
Adenosine Triphosphate (ATP): <sup>1</sup> H Spectrum	351
Methyl $\beta$ -D-Glucopyranoside in D <sub>2</sub> O: COSY Spectrum	377
D-Glucose in D <sub>2</sub> O: COSY Spectrum	387
2-Naphthol (2-hydroxynaphthalene) in CDCl <sub>3</sub> : COSY Spectrum	392
Cholesterol in CDCl <sub>3</sub> : COSY Spectrum	406
3-Heptanone: 2D-TOCSY Spectrum	416
Leucine Enkephalin Acetate (Tyr-Gly-Gly-Phe-Leu): 2D-TOCSY Spectrum	418
$\beta$ -D-Lactose Peracetate ( $\beta$ -D-Lactose Octaacetate): 2D-TOCSY Spectrum	419
$\alpha$ -Thujone in CDCl <sub>3</sub> : 2D-NOESY Spectrum	424
Leucine Enkephalin Acetate (Tyr-Gly-Gly-Phe-Leu): 2D-NOESY Spectrum	426
D-Glucose in D <sub>2</sub> O. HSQC Spectrum	440
Cholesterol in CDCl <sub>3</sub> : HSQC Spectrum	443
Ethyl Acetate in CDCl <sub>3</sub> : <sup>13</sup> C Spectrum without Decoupling	463
Ethyl Acetate in CDCl <sub>3</sub> : HMBC Spectrum	466
Oxybenzone: HMBC Spectrum	468
Testosterone: HMBC Spectrum	475
$\beta$ -D-Lactose Peracetate in CDCl <sub>3</sub> : HMBC Spectrum	492
Telekin: Unknown A, a Sesquiterpene Natural Product	502
Ergosterol Analog: Unknown B, a Steroid Natural Product	522
Raffinose: Unknown C, a Trisaccharide	553
Monocrotaline: Unknown D, an Alkaloid Natural Product	575
Pristimerin Analog: Unknown E, a Triterpene Natural Product	597

# 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 or 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](http://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 ( $^1\text{H}$  and  $^{13}\text{C}$ ), edited (DEPT)  $^{13}\text{C}$ , selective one-dimensional  $^1\text{H}$  (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 cryogenics (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  $\alpha$ - and  $\beta$ -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  $^1\text{H}$  and  $^{13}\text{C}$  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  $^1\text{H}$  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](http://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



# 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.,  $^1H = 1.007825$  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).

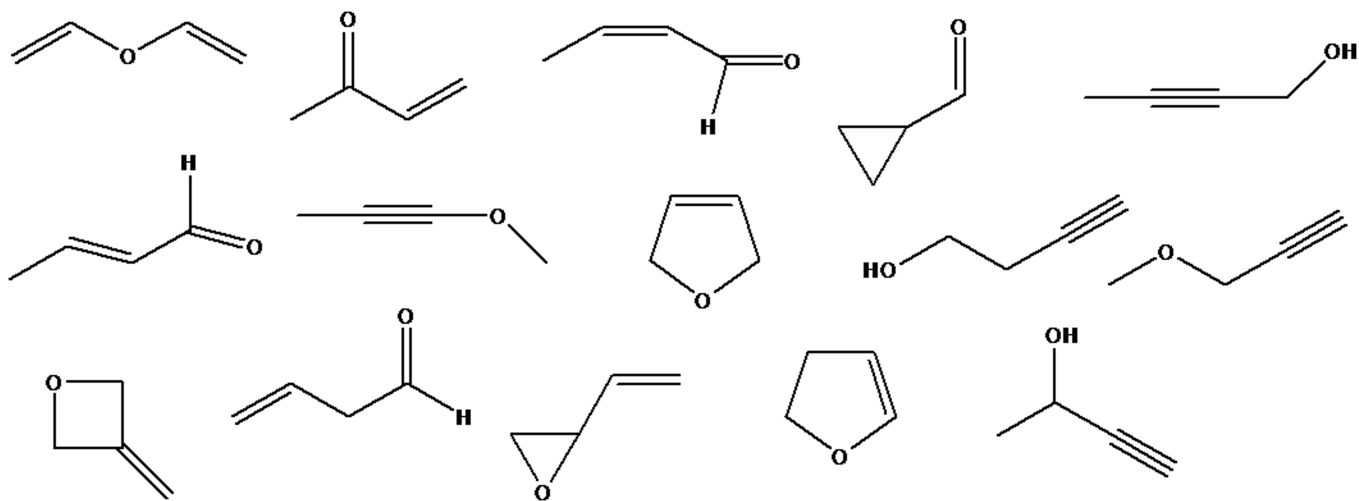


FIGURE 1.1

Note that all of the  $C_4H_6O$  structures in Figure 1.1 have one thing in common: the total of the number of  $\pi$  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:  $m = (n \times 2) + 2 = (4 \times 2) + 2 = 10$ .
5. Subtract the number of hydrogens in the modified molecular formula (6) from this saturated hydrocarbon value and divide the result by 2:  $m - 6 = 10 - 6 = 4$ ;  $u = 4/2 = 2$ .

This result ( $u$ ) is equal to the number of  $\pi$  bonds in the molecule *plus* the number of rings. Note that a triple bond ( $C \equiv C$ ) is really one  $\sigma$  bond and two  $\pi$  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  $\sigma$  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,  $C \equiv N$  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  $\geq$  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

the same type of information from each of the atoms in the molecule in turn, especially from the hydrogens. Adding up all of this information (chemical environments, distances, and angles) can give us a covalent structure (which atoms are connected to which by covalent bonds) and a conformation (shape of the molecule in three dimensions).

Determining the structure of a molecule by NMR is a puzzle-solving exercise, and to date it still requires a lot of human judgment and intuition; you don't just feed it into a computer and out pops a structure. The exercise can be exciting and challenging, and it gives the rare human experience of looking straight into the molecular world and getting unambiguous answers to our questions. But it must be emphasized that NMR does not give a **picture** of the molecule. In spite of its close relationship to MRI (magnetic resonance imaging), NMR spectroscopy is not an imaging experiment and it does not give any kind of image or picture of the molecule. You, the person interpreting the NMR data, must put all of these simple pieces of evidence together, along with whatever other information you have, to *propose* a structure of the molecule. Then you have to go over the evidence to make sure all of it is *consistent* with your proposed structure. As in all science, we can gather more and more evidence and be more and more sure of our conclusion, but we can never be absolutely sure. One of the advantages of NMR is that the sheer volume of complimentary information that can be gathered from multiple vantage points (the different atoms in the molecule) makes it a technique with a very high degree of confidence in the conclusions. For small molecules (molecular weight below 500 Da), this confidence comes very close to certainty for experienced users willing to do a number of NMR experiments.

There is another technique for molecular structure determination that *does* generate a picture or image of the molecule. **X-ray crystallography** measures the pattern of scattering of X-rays from a solid crystal of the molecule. By analyzing the intensities of thousands of spots from the scattered X-rays, a computer can create a three-dimensional map of the electron density of the molecule. Since atoms are basically dense clouds of electrons, the atoms can be accurately located and you get a three-dimensional structure of the molecule. The main drawback of this technique is that you need a crystal, and even then the crystal may not have the right properties to give good X-ray diffraction. Once you have a good crystal, the process is time consuming and requires a great deal of calculation and refinement of the data by an expert. In contrast, an NMR spectrum can be acquired in a few minutes if a pure sample can be dissolved in a solvent. The analysis of NMR data, as we shall see, is straightforward and can be learned by anyone with a basic understanding of organic chemistry.

Before we look at the NMR experiment in more detail, some of the other tools for organic structure determination will be briefly explained. These give information which is complementary to the NMR data and help to provide the complete picture of the molecule.

---

## 2 MASS SPECTROMETRY

---

Mass spectrometry is essentially a method for weighing individual molecules to determine their mass. Knowing the masses of individual atoms that make up the molecule ( $H = 1$ ,  $C = 12$ ,  $N = 14$ ,  $O = 16$ , *etc.*), we can narrow down the possibilities to a small number of possible molecular formulae. For example, for an integer mass of 120 units, we can have the following molecular formulae:

$$C_9H_{12}: [9 \times 12] + [12 \times 1] = 120$$

$$C_8H_8O: [8 \times 12] + [8 \times 1] + [1 \times 16] = 120$$

$$C_7H_4O_2: [7 \times 12] + [4 \times 1] + [2 \times 16] = 120$$

$$C_7H_8N_2: [7 \times 12] + [8 \times 1] + [2 \times 14] = 120$$

---

**Exercise 1.1:** Calculate the number of unsaturations (number of  $\pi$  bonds + number of rings) for each of the above molecular formulae. Explain why  $C_6H_{16}O_2$  is not a possible molecular formula for a molecular mass of 120.

---

We will see shortly that with a more accurate molecular mass, like 120.0687 for  $C_7H_8N_2$ , we can narrow down the possible molecular formulae to a single one.

NMR focuses on the hydrogens and carbons within a molecule, so it has a hard time counting the oxygen and nitrogen atoms, and other atoms like sulfur and halogens can be “invisible” in the NMR data. This makes mass spectrometry an essential complement to NMR data for determination of structure.

## 2.1 Ionization Methods and Molecular Ions

The basic experiment of a mass spectrometer is to convert a molecule into a charged species, an ion, and move it around in a vacuum using electric and magnetic fields, to determine its mass by the nature of its motion. There are three basic steps in this process:

1. **Ionization:** Convert the neutral molecule into an ion (usually positive).
2. **Mass Analysis:** Separate ions on the basis of their mass.
3. **Detection:** Detect the ion to generate an electrical signal.

From the point of view of the organic chemist, the first step is the most important. There are two main methods of ionization:

### 2.1.1 Electron Impact (EI)

The spectrometer gets the molecule into the gas phase and hits it with a high energy electron, knocking out an electron. This is a “hard” ionization process because it imparts a lot of energy to the molecule. The result is a radical cation ( $M^{+\bullet}$ ), a very unstable species that quickly fragments to generate more stable pieces of the molecule. This is the oldest and simplest ionization method and is usually used in conjunction with a gas chromatograph (GC). The sample is injected into the GC, the components (if it is not pure) are separated and the peaks emerging from the GC column go directly into the high vacuum of the mass spectrometer, where the electron beam ionizes the molecules. There are a number of disadvantages to this technique:

- The molecule must be at least somewhat volatile. This limits the technique to fairly simple, non-polar molecules.
- The molecular ion ( $M^{+\bullet}$ ) can be a very weak peak in the mass spectrum, which is dominated by the molecular fragments. While this provides useful information about the molecular structure, it limits the usefulness of mass spectrometry for determining the molecular formula.

The EI mass spectrum of caffeine ( $C_8H_{10}N_4O_2$ ) is shown in Figure 1.2.

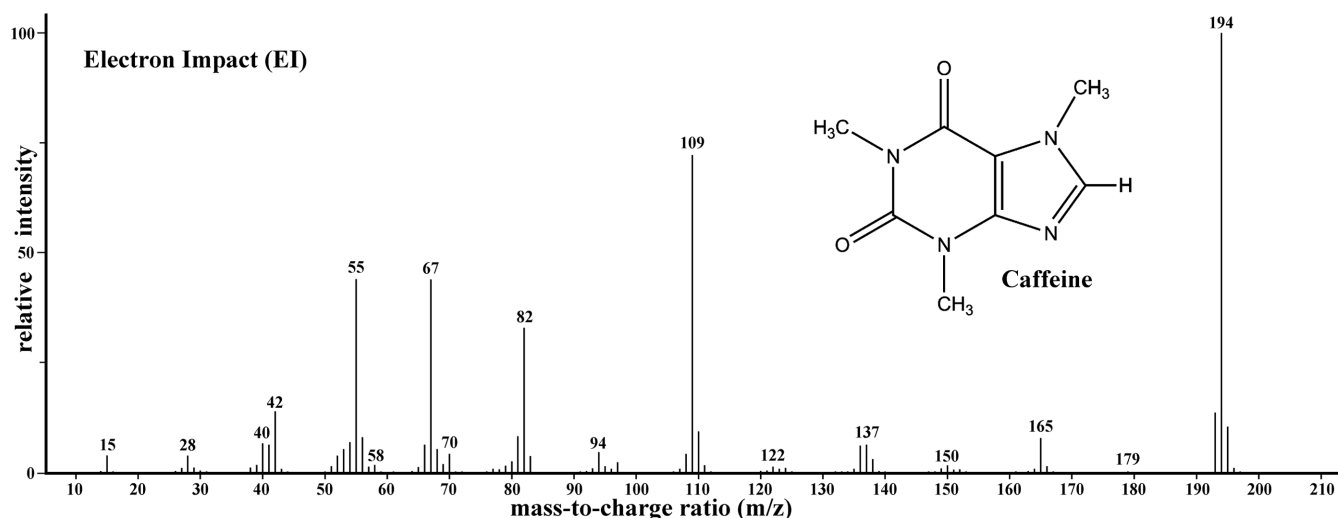


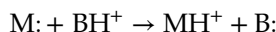
FIGURE 1.2 Courtesy of National Institute of Standards and Technology (NIST).



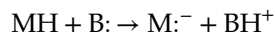
The horizontal scale is the mass-to-charge ratio ( $m/z$ ), which is essentially the mass in atomic mass units (amu), since virtually all ions are singly-charged ( $z = 1$ ). The vertical scale is the relative intensity of the peaks, relative to the most intense peak, known as the parent ion, as 100%. Note that the peaks are separated between consecutive integer masses. Because caffeine is a very stable aromatic compound, the molecular ion ( $M^+$  at  $m/z$  194) is also the most intense peak (the parent ion). The peak at 193 ( $M - 1$ ) is due to loss of hydrogen ( $H^\bullet$ ), and the peak at 194 ( $M + 1$ ) is due to the presence of one  $^{13}C$  atom in the molecule (9% intensity). These isotope peaks will be discussed in detail in the next section.

### 2.1.2 Soft Ionization

This is a general term for low energy ionization methods that essentially just protonate the molecule:



where  $BH^+$  is a proton donor that is supplied to the sample or the mass spectrometer. The molecule has to have at least some basicity, meaning that it needs to have a lone pair that can accept a proton. Some molecules have no basic sites, but these can usually be ionized by negative mode mass spectrometry:



where B: is a proton acceptor that is supplied to the sample or the mass spectrometer. This mode is less commonly used but is available if positive ion mode gives poor results. Even very weakly basic or acidic sites are amenable to one of these two modes, so the only molecules that would not work would be hydrocarbons.

One consequence of soft ionization is that it's possible to get multiply charged ions. For example, in positive ion mode there may be multiple basic sites on the molecule so there can be a number of different charge states (e.g.,  $[MH]^+$ ,  $[MH_2]^{+2}$ ,  $[MH_3]^{+3}$ , and so forth). This brings up an interesting point about mass spectrometry: it is not actually the *mass* that is measured, but rather the *mass-to-charge ratio* ( $m/z$ ). For small molecules, the organic molecules of this book, this is usually not a big issue, but for biological molecules like peptides and proteins multiply charged ions are very common. For example, if the molecular mass is 1000 (M) and the charge is +6 ( $[MH_6]^{+6}$ ), the measured mass-to-charge ratio ( $m/z$ ) would be  $1006/6 = 167.7$ . In this book, we will give mass spectral data ( $m/z$ ) for singly-charged positive ions only:

$M^+$	Molecular ion (radical cation)	Electron impact
$[MH]^+$	Protonated molecular ion	Soft ionization

In this case, the mass is the same as the mass-to-charge ratio.

## 2.2 High-Resolution Mass Spectrometry and Exact Mass

Some mass analyzers are low-resolution, essentially giving only the integer mass. In the example above ( $m/z$  120), it would be impossible to distinguish between the possible molecular formulae:



All of these give an integer mass of 120 for the molecular ion (or 121 for  $[MH]^+$ ). But high resolution mass analyzers, such as FTICR (Fourier Transform Ion Cyclotron Resonance) and TOF (Time of Flight), can give  $m/z$  measurement accurate to less than 1 ppm (1 part per million of the measured  $m/z$ ). For example, for a mass of 120, 1 ppm is 0.00012 mass units, or about one unit in the fourth decimal place. At this level of accuracy it's possible to distinguish between

different molecular formulae with the same integer mass. The exact masses of the major isotopes of common atoms are given below:

$^1\text{H}$	1.007825	$^{32}\text{S}$	31.972072
$^{12}\text{C}$	12.000000	$^{31}\text{P}$	30.973763
$^{16}\text{O}$	15.994915	$^{28}\text{Si}$	27.976928
$^{14}\text{N}$	14.003074	$^{19}\text{F}$	18.998403

One isotope is exactly equal to the integer mass ( $^{12}\text{C}$ , by definition), others are slightly more than the integer value ( $^1\text{H}$ ,  $^{14}\text{N}$ ), and others are slightly below the integer values ( $^{16}\text{O}$ ,  $^{32}\text{S}$ ,  $^{31}\text{P}$ ,  $^{28}\text{Si}$  and  $^{19}\text{F}$ ). These slight differences from the integer mass allow us to distinguish different molecular formulae if the mass measurement is made with very high accuracy.

It's important to understand that in mass spectrometry we are observing specific isotopic species. For example, for  $\text{C}_8\text{H}_8\text{O}$  the major molecular ion peak ( $\text{M}^+$ ) is really  $^{12}\text{C}_8^1\text{H}_8^{16}\text{O}$ :

$$m/z = (8 \times 12.000000) + (8 \times 1.007825) + (1 \times 15.994915) = 120.057515$$

This is a different mass from  $\text{C}_7\text{H}_8\text{N}_2$ , which is really  $^{12}\text{C}_7^1\text{H}_8^{14}\text{N}_2$ :

$$m/z = (7 \times 12.000000) + (8 \times 1.007825) + (2 \times 14.003074) = 120.068748$$

The difference in mass between these two formulae is 0.011233 mass units, or 94 ppm ( $0.011233/120 = 94 \times 10^{-6}$ ). These two formulae can easily be distinguished with a high resolution mass analyzer. High resolution not only requires special equipment, but takes more time and costs more money. Careful calibration using calibrant molecules is required to get this kind of accuracy.

Exact mass calculations can be made easily using this calculator from Scientific Instrument Services:

<http://www.sisweb.com/referenc/tools/exactmass.htm>

In this book, molecular masses ( $\text{M}^+$  or  $[\text{MH}]^+$ ) are given in many of the problems and examples to simplify the structural problem of unknowns. Most of these values are not experimental values; they are calculated and, in the case of exact masses, a random error is added or subtracted to give a simulation of actual data. Always remember to subtract 1 mass unit (1.007825 for exact mass) from the protonated molecular ion ( $[\text{MH}]^+$ )  $m/z$  value to get the mass that corresponds to the (neutral) molecular formula. Sometimes with chemical ionization methods a sodium or potassium ion can take the place of  $\text{H}^+$  in creating a positive ion:  $[\text{M}\cdot\text{Na}]^+$  or  $[\text{M}\cdot\text{K}]^+$ . In this case, instead of subtracting the mass of hydrogen to obtain the molecular mass, one has to subtract the mass of sodium ( $^{23}\text{Na} = 22.989770$ ) or potassium ( $^{39}\text{K} = 38.963708$ ).

---

**Exercise 1.2:** For each of the following protonated molecular ion ( $[\text{MH}]^+$ ) exact mass values, find the molecular formula, using only H, C, N and O. The number of carbons (obtained by NMR) is given to limit the number of possibilities. Compare the observed  $m/z$  value to the calculated value and give the error in ppm. Calculate the number of unsaturations ( $\pi$  bonds plus rings) in the molecule.

a. 167.1075 ( $\text{C}_{10}$ )

b. 136.1124 ( $\text{C}_9$ )

c. 210.1497 ( $\text{C}_{12}$ )

d. 195.1500 ( $\text{C}_{11}$ )

---

The **nitrogen rule** (or odd/even rule) is a simple consequence of the odd number of valences of nitrogen (3 bonds) combined with its even mass (14). The other common atoms have either an odd mass and odd number of bonds (H, Br, Cl) or an even mass and even number of bonds (C, O, S). The rule can be stated simply:

- If the neutral molecular mass ( $\text{M}$ ) is **even**, there is an **even** number of nitrogens in the molecular formula: 0, 2, 4, 6, . . .

- If the neutral molecular mass (M) is **odd**, there is an **odd** number of nitrogens in the molecular formula: 1, 3, 5, 7, . . .

Most importantly, an odd molecular mass means that we probably have nitrogen. An even mass means either we have no nitrogen, or we have at least two nitrogens.

Because the calculation of the number of unsaturations is based on the valencies of the various atoms (C = 4, N = 3, O = 2, H = 1, *etc.*), any molecular formula that violates the nitrogen rule will give a half-integer number of unsaturations. For example, in Exercise 1.2a a formula of C<sub>12</sub>H<sub>8</sub>N, with mass of 120, would violate the rule because the mass is even and the number of nitrogens (1) is odd. The calculated number of unsaturations (C<sub>12</sub>H<sub>8</sub>N → C<sub>13</sub>H<sub>9</sub>, [28 - 9]/2 = 9.5) is not an integer.

### 2.3 Isotope Patterns and the Halogens Br and Cl

So far we have dealt with atoms that have one isotope with almost 100% abundance (<sup>1</sup>H, 99.99%, <sup>12</sup>C, 98.9%, <sup>16</sup>O, 99.76%, <sup>14</sup>N, 99.63%). Of these only carbon gives a significant M + 1 isotope peak (<sup>13</sup>C = 1.11% of the <sup>12</sup>C abundance). The intensity of this isotope peak depends on the probability of an ion containing one <sup>13</sup>C atom:

<sup>12</sup> C <sub>8</sub> H <sub>8</sub> O:	Probability of <sup>12</sup> C <sub>8</sub> = 0.989 <sup>8</sup> = 0.915	<i>m/z</i> 120 (100%)
<sup>13</sup> C <sup>12</sup> C <sub>7</sub> H <sub>8</sub> O:	Probability = 8 × 0.011 × 0.989 <sup>7</sup> = 0.081	<i>m/z</i> 121 (8.9%)

The probability of all 8 carbons being <sup>12</sup>C is the product of all 8 individual probabilities: 0.989 × 0.989 × 0.989 × 0.989 × 0.989 × 0.989 × 0.989 × 0.989 = 0.989<sup>8</sup> (0.989 to the eighth power)

Here we are ignoring the slightly less than one probabilities for <sup>16</sup>O and <sup>1</sup>H, and multiplying by 8 in the second case because there are 8 different ways we can have one <sup>13</sup>C and seven <sup>12</sup>C atoms in a molecule with 8 carbons. In the mass spectrum, there will be the major molecular ion peak (M<sup>+</sup>) at 120 and another peak one mass unit higher (M + 1) with intensity 8.9% of the main (M<sup>+</sup>) peak. If there were a large enough number of carbons in the molecule, the probability of one <sup>13</sup>C would be larger than the probability of all carbon atoms being <sup>12</sup>C. For example, with 100 carbons (C<sub>100</sub>), the M + 1 ion would have an intensity of 111% relative to the molecular (all <sup>12</sup>C) ion. Organic molecules are small enough that this is never a problem. In general, the molecular ion is defined as the lowest mass ion in a cluster of isotope peaks, the ion in which all atoms have their lowest mass: <sup>1</sup>H, <sup>12</sup>C, <sup>14</sup>N, <sup>16</sup>O, <sup>35</sup>Cl, <sup>79</sup>Br, *etc.* In our discussion, this will be defined as 100% intensity, even though there may be other isotope peaks, or fragment ion peaks, that are more intense.

Carbon can be described as an **M + 1 atom** because of the significant isotope with one extra mass unit (<sup>13</sup>C *vs.* <sup>12</sup>C). Similarly, oxygen, sulfur, chlorine and bromine can be described as **M + 2 atoms**:

O	<sup>16</sup> O	100%	<sup>18</sup> O	0.21%
S	<sup>32</sup> S	100%	<sup>34</sup> S	4.52%
Cl	<sup>35</sup> Cl	100%	<sup>37</sup> Cl	31.96%
Br	<sup>79</sup> Br	100%	<sup>81</sup> Br	97.28%

As always with mass spectral intensities, the most abundant isotope is expressed as 100% intensity, leading to the perplexing fact that the total is more than 100%. Remember that these are not percentages of the total intensity, but only percent of the parent ion intensity. In low resolution mass spectrometry, the relative intensity of the M + 2 ion can be used to estimate the number of oxygens or, if it is greater than 4%, the number of sulfur atoms.

The halogens chlorine and bromine have huge M + 2 ions, making them very easy to pick out in a mass spectrum. Bromine occurs naturally as <sup>79</sup>Br (50.7% of total) and <sup>81</sup>Br (49.3% of total). Converting to percentages relative to the lower mass isotope (100%) makes it easier to calculate relative probabilities: <sup>79</sup>Br = 100%; <sup>81</sup>Br = 97.28%. Just convert the percentage to a fractional intensity (100% = 1.00) and multiply:

$$\text{Br}_2 = \begin{array}{ccc} \text{}^{79}\text{Br}_2(1.00 \times 1.00) & + & \text{}^{79}\text{Br}^{81}\text{Br}(2 \times 1.00 \times 0.9728) & + & \text{}^{81}\text{Br}_2(0.9728 \times 0.9728) \\ \text{M (100\%)} & & \text{M + 2 (194.6\%)} & & \text{M + 4 (94.6\%)} \end{array}$$

Fractional intensity is converted back to percent at the end. For the M + 2 ion there are two ways that <sup>79</sup>Br<sup>81</sup>Br can happen, so we multiply by two for this statistical factor. An isotope pattern like this (M: M + 2: M + 4 = 100:195:95) can

be directly read out as two bromine atoms in the molecule. This pattern is clearly seen in the EI mass spectrum of 3,5-dibromotoluene ( $C_7H_6Br_2$ , Figure 1.3), particularly in the expansion of the region around the molecular ion.

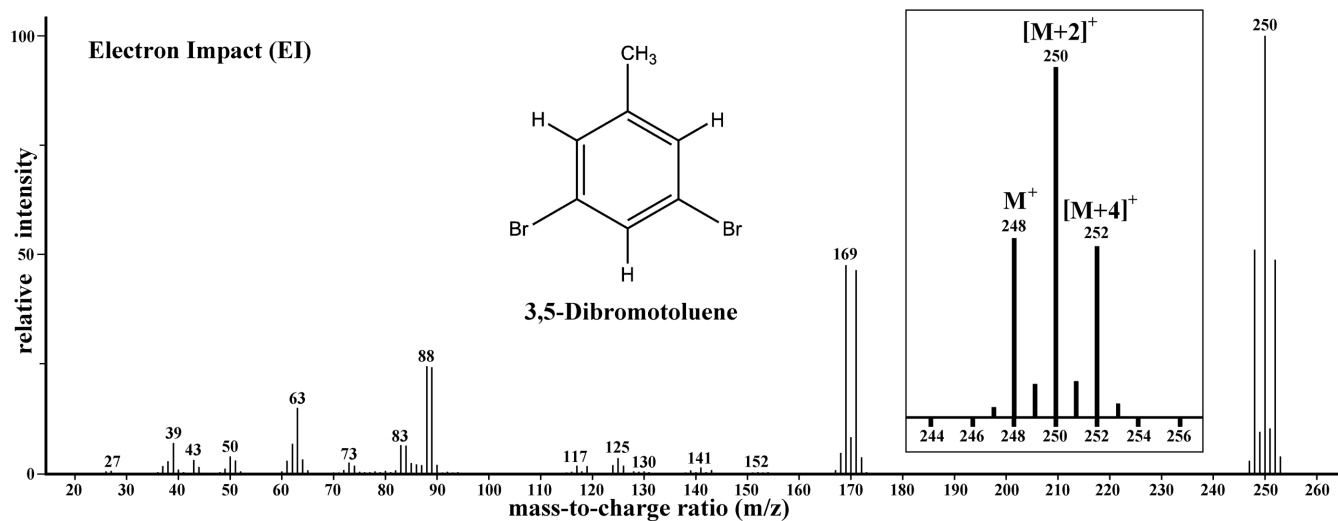


FIGURE 1.3 Courtesy of NIST.

The molecular ion ( $M^+ = 248$ ) is the lowest mass ion in the cluster of isotope peaks, representing the isotopic composition  $^{79}Br_2$  for the bromines. The roughly 1:2:1 ratio of peaks, each one two mass units higher than the previous peak, is a dead giveaway that there are two bromines in the molecule. The fragment ion at  $m/z$  169 has a companion peak at  $m/z$  171 of nearly equal intensity (1:1 ratio), so it contains only one bromine atom. In fact, this fragment represents the loss of Br from the molecular ion ( $248 - 79 = 169$ ). This fragment can be described as “M-79” (molecular ion minus 79 mass units) or “M-Br” (loss of bromine).

**Exercise 1.3:** Calculate the number of unsaturations in the caffeine and 3,5-dibromotoluene molecular formulae, and count the number of  $\pi$  bonds and the number of rings in each structure.

**Exercise 1.4:** For the following isotope combinations, calculate the  $M + 2$  (and  $M + 4$ ,  $M + 6$ , etc.) intensities in the mass spectral isotope pattern.

a.  $Cl_2$

b.  $BrCl$

c.  $Cl_2Br$

d.  $Br_3$

All of these isotope abundances are expressed as percent of the lowest mass isotopic species. In mass spectrometry, the peak intensities are usually expressed in percent of the most intense peak (the parent peak), which can easily be the  $M + 2$  or  $M + 4$  peak if there are multiple bromine or chlorine atoms present. As long as the peak intensities are expressed as ratios this should not cause any confusion. For example:

$$m/z\ 210/212/214\ (100:195:95)$$

implies that there are two bromines in the molecular formula, with the 210 mass corresponding to  $^{79}Br_2$ , the 212 mass corresponding to  $^{79}Br^{81}Br$ , and the 214 mass corresponding to  $^{81}Br_2$ . Alternatively, this could be written as:

$$m/z\ 210/212/214\ (51:100:49)$$

All intensity ratios in this book will be expressed using 100 for the lowest mass isotopic species in the molecular ion.

### 3 INFRARED (IR) SPECTROSCOPY

This is an important tool for the organic chemist that measures the absorption of electromagnetic radiation (light) in the frequency range of 500 to 4000 wavenumbers ( $\text{cm}^{-1}$ ), corresponding to a wavelength range of 2.5 to 20 microns (1 micron or  $\mu = 10^{-6}$  meters). This is beyond the low energy (red) edge of the visible light spectrum, hence the term “infra – red” (below red). Absorption of light in this frequency range corresponds to stretching vibrations and bending motions of the chemical bonds in a molecule. The IR spectrum can give useful information about the functional groups of a molecule: carbonyl ( $\text{C}=\text{O}$ ), hydroxyl ( $\text{OH}$ ), nitrile ( $\text{C}\equiv\text{N}$ ), olefin ( $\text{C}=\text{C}$ ), *etc.* The IR spectrum is presented with a frequency scale at the bottom and the baseline at the top, with absorption bands appearing as dips in the baseline towards the bottom of the display. Some of these bands can be narrow (“sharp”) and well-defined, and others can be very wide (“broad”) and amorphous. We are concerned here only with a few of the most useful absorption bands in the IR spectrum, those that are easily interpreted and give information that may be difficult to obtain from the NMR spectrum.

The IR spectrum of *para*-acetyl-benzonitrile ( $p\text{-CH}_3\text{CO-C}_6\text{H}_4\text{-CN}$ ) is shown in Figure 1.4.

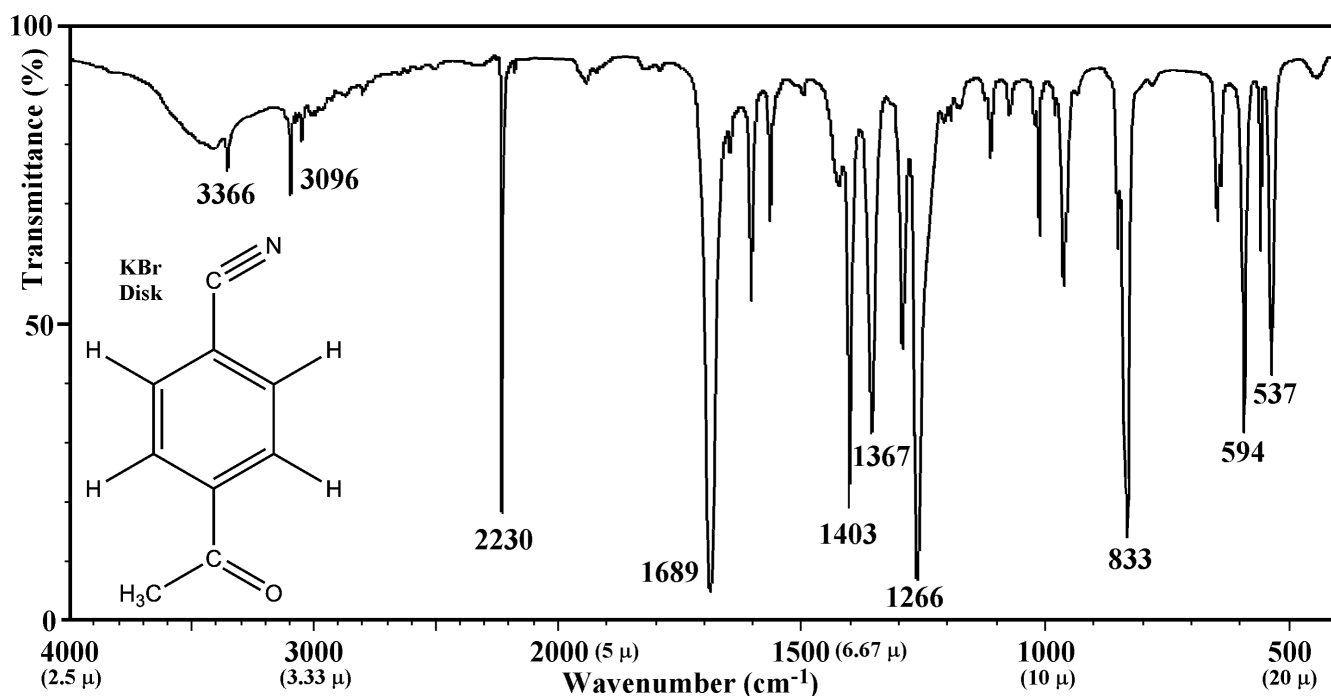


FIGURE 1.4 Courtesy of National Institute of Advanced Industrial Science and Technology (AIST), Japan.

The frequency (horizontal) scale is expanded in the 1500–500  $\text{cm}^{-1}$  range. Wavenumbers are the reciprocal of wavelength ( $10^4/\lambda$  in microns), so they are proportional to frequency ( $\text{Hz} = \text{speed of light}/\lambda$ ). The vertical scale is transmittance, the percentage of light that makes it through the sample without being absorbed. The solid sample was ground with solid potassium bromide (KBr) in a mortar and pestle and then pressed into a thin disc. The infrared light passes through this disc and ends up at a detector. Solid KBr does not absorb infrared light, so this is a very clean way to record the IR spectrum. Liquid samples (neat liquids without solvent) can be pressed in a thin film between two salt plates. There are also solution sample cells using solid salt, but in solution some regions will be wiped out by the IR absorption bands of the solvent.

The most important single band in an IR spectrum is the **carbonyl ( $\text{C}=\text{O}$ ) stretching** vibration, a strong band in the region of 1700  $\text{cm}^{-1}$ . This is particularly useful because this is a “quiet” region of the spectrum with little interference from other stretching or bending motions. The generic value for a ketone is 1715  $\text{cm}^{-1}$ . In Figure 1.4 this band is at 1689  $\text{cm}^{-1}$ , shifted to lower frequency from the generic value due to conjugation of the ketone with the aromatic ring. Esters ( $\text{R}-\text{C}(=\text{O})-\text{OR}'$ ) show this band shifted to higher frequency ( $\sim 1750 \text{ cm}^{-1}$ ).

The nitrile triple bond gives rise to a very distinctive band at 2230 in *p*-acetylbenzonitrile, due to the  **$\text{C}\equiv\text{N}$  stretching** vibration (Figure 1.4). This strong, narrow band is also in a quiet region of the spectrum. The only other

band in this region is the weak  $C\equiv C$  stretching vibration of alkynes ( $2260\text{--}2100\text{ cm}^{-1}$ ). Beware of contamination with deuterated chloroform (the most common solvent used in NMR) because the C-D stretching vibration is at  $2256\text{ cm}^{-1}$ .

The **C—H bond stretching** vibrations give rise to bands in the region near  $3000\text{ cm}^{-1}$ . These are fairly weak bands, and the general rule is that aliphatic ( $sp^3$ -hybridized) C—H bands occur to the right (lower frequency) of  $3000\text{ cm}^{-1}$  and aromatic and olefinic ( $sp^2$ -hybridized) C—H bands occur to the left (higher frequency) of  $3000\text{ cm}^{-1}$ . Note the band at  $3096\text{ cm}^{-1}$  in Figure 1.4. The distinction between aliphatic, aromatic and olefinic hydrogens is easily made by NMR.

The **O—H bond stretching** gives rise to a very broad band around  $3300\text{ cm}^{-1}$ . If salt plates are not protected from moisture this band will appear due to  $H_2O$  even if the sample contains no O—H bonds. The N—H stretching band of amines and amides appears in the same region and is also quite broad.

Since the advent of NMR spectroscopy in the 1960s, infrared spectroscopy has gradually diminished in importance in the elucidation of organic structures. The important difference between IR and NMR is that the frequency of infrared absorption bands depends on the vibrational modes of the molecule as a whole, whereas the resonant frequency of each nucleus (*e.g.*, a specific H or C in the molecule) is a local phenomenon that responds to the immediate environment (within 3–4 bonds or within  $5\text{ \AA}$ ) of that particular atom within the molecule.

## 4 ULTRAVIOLET (UV) AND VISIBLE SPECTROSCOPY

The visible light spectrum extends from a wavelength of  $390\text{ nm}$  ( $0.39\text{ microns}$ , violet) to  $700\text{ nm}$  ( $0.7\text{ }\mu$ , red). The infrared spectrum is lower energy (longer wavelength) than the low energy (red) side of the visible light spectrum, and the ultraviolet (UV) spectrum is higher energy (shorter wavelength) than the high energy (violet) side of the visible spectrum. The recorded ranges of spectrometers are shown below:

$$20\text{ }\mu \xrightarrow{\text{IR}} 2.5\text{ }\mu \quad 700\text{ nm} \xrightarrow{\text{VIS}} 390\text{ nm} \quad 390\text{ nm} \xrightarrow{\text{UV}} 210\text{ nm}$$

An example of an ultraviolet (UV) absorption spectrum is shown in Figure 1.5 for oxybenzone, a major ingredient in sunscreen creams.

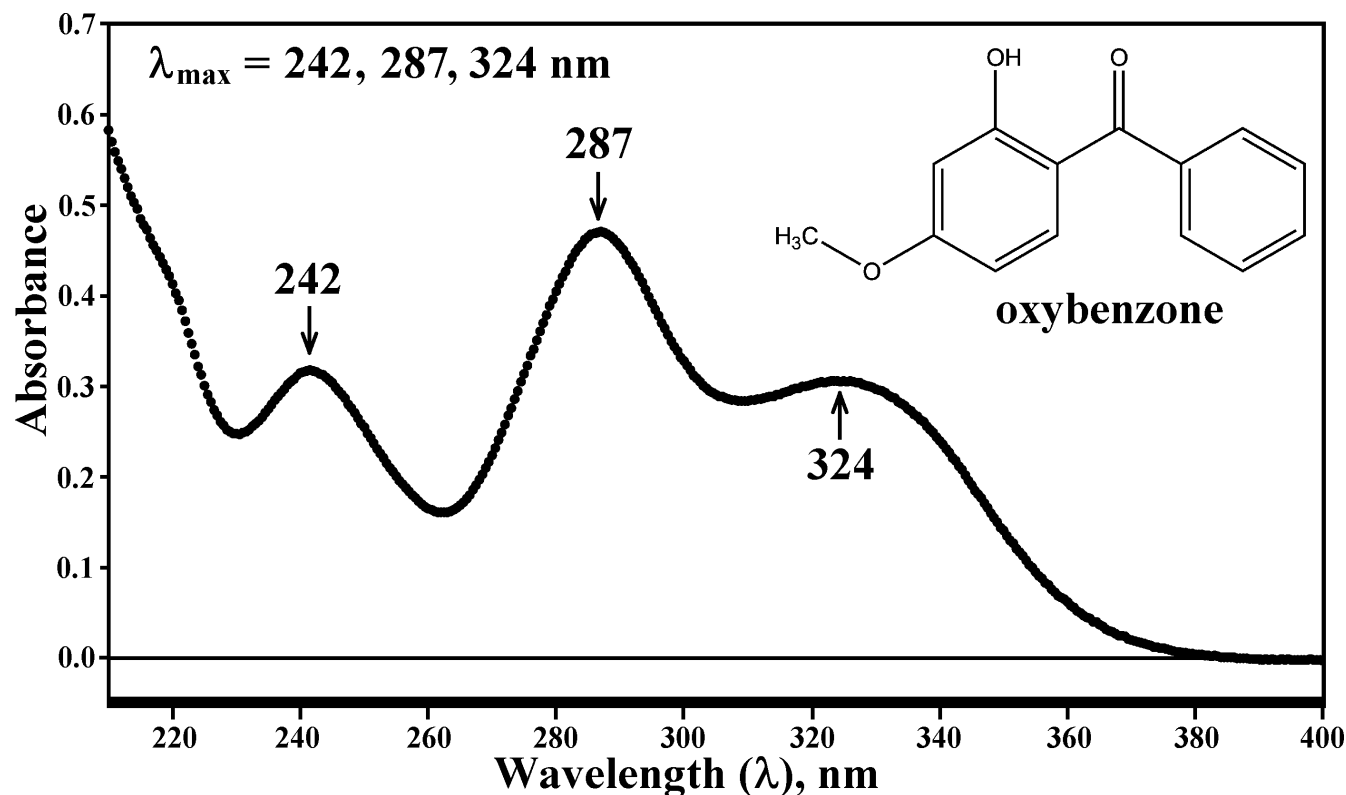


FIGURE 1.5