Biological Magnetic Resonance
Volume 9

A Continuation Order Plan is available for this series. A continuation order will bring delivery of each new volume immediately upon publication. Volumes are billed only upon actual shipment. For further information please contact the publisher.

Biological Magnetic Resonance Volume 9

Edited by

Lawrence J. Berliner

Ohio State University
Columbus, Ohio

and

Jacques Reuben

Hercules Incorporated
Research Center
Wilmington, Delaware

PLENUM PRESS • NEW YORK AND LONDON

The Library of Congress has cataloged the first volume of this series as follows:

Library of Congress Cataloging in Publication Data

Main entry under title:

Biological magnetic resonance:

Includes bibliographies and indexes.
1. Magnetic resonance. 2. Biology—Technique. I. Berliner, Lawrence, J. II. Reuben, Jacques.
QH324.9.M28B56 574.19′285 78-16035
AACR1

ISBN 978-1-4615-6551-2 ISBN 978-1-4615-6549-9 (eBook)
DOI 10.1007/978-1-4615-6549-9

© 1990 Plenum Press, New York
Softcover reprint of the hardcover 1st edition 1990
A Division of Plenum Publishing Corporation
233 Spring Street, New York, N.Y. 10013

All rights reserved

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher

Contributors

Brandan A. Borgias • Cray Research, Inc., San Ramon, California 94583

Thomas L. James • Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143

John L. Markley • Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706

Alan G. Marshall • Department of Biochemistry and Department of Chemistry, The Ohio State University, Columbus, Ohio 43210

Joseph E. Meier • Department of Chemistry, The Ohio State University, Columbus, Ohio 43210

Stanley J. Opella • Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Andrew D. Robertson • Department of Biochemistry, College of Agricultural and Life Science, University of Wisconsin-Madison, Madison, Wisconsin 53706. *Present address*: Department of Biochemistry, Beckman Center, Stanford University, Stanford, California 94305

Philip L. Yeagle • Department of Biochemistry, School of Medicine, State University of New York at Buffalo, Buffalo, New York 14214

Jiejun Wu • Department of Biochemistry, The Ohio State University, Columbus, Ohio 43210

Preface

We are pleased to present Volume 9 of our highly successful series, which now celebrates 12 years of providing the magnetic resonance community with topical, authoritative chapters on new aspects of biological magnetic resonance. As always, we try to present a diversity of topic coverage in each volume, ranging from applications of *in vivo* magnetic resonance to more fundamental aspects of electron spin resonance and nuclear magnetic resonance.

Philip Yeagle presents an eagerly awaited chapter on ^{31}P NMR studies of membranes and membrane protein interactions. Alan Marshall has contributed two chapters to the volume: one, with Jiejun Wu, describes magnetic resonance studies of 5S-RNA as probes of its structure and conformation; the second chapter, with Joseph Meier, reviews methods for the suppression of water resonances in proton NMR—a much-needed and highly informative chapter. The problem of relaxation and spin diffusion is examined in great detail by Brandan Borgias and Thomas James in a discussion of the use of a complete relaxation matrix analysis (CORMA) as applied to DNA fragments in 2D NOE spectra. Proton NMR methods in spectral assignments of proteins are covered with excellent examples by Andrew Robertson and John Markley. Finally, Stanley Opella has contributed a superb chapter on the use of solid-state NMR techniques to understand the structure of large macromolecular assemblies.

After the success of our special collection (Volume 8, *Spin Labeling: Theory and Applications*), we plan more special topics volumes in the future. Our current plans involve volumes on *in vivo* magnetic resonance spectroscopy and on techniques and applications in protein structure determination, as well as volumes specifically devoted to magnetic resonance approaches to, e.g., complex carbohydrate analysis and nucleic acid structure. We would

greatly appreciate input and suggestions from our readers regarding the quality and composition of this and further volumes.

Lawrence J. Berliner
Jacques Reuben

Contents

Chapter 1

Phosphorus NMR of Membranes

Philip L. Yeagle

1. Introduction 1
   1.1. Model and Biological Membrane Structures 1
   1.2. ^{31}P NMR of Membranes—Theoretical and Practical Aspects 6
2. ^{31}P NMR and Membrane Morphology 17
   2.1. Introduction 17
   2.2. Model Membranes 22
   2.3. Biological Membranes 28
3. ^{31}P NMR and Membrane Structure and Dynamics 30
   3.1. Phospholipid Headgroup Conformation 30
   3.2. Phospholipid Headgroup Dynamics 37
   3.3. Phospholipid-Protein Interactions 38
   3.4. Transmembrane Phospholipid Distribution 48
4. New Areas for Investigation with ^{31}P NMR 49
   References 50

Chapter 2

Investigation of Ribosomal 5S Ribonucleic Acid Solution Structure and Dynamics by Means of High-Resolution Nuclear Magnetic Resonance Spectroscopy

Alan G. Marshall and Jiejun Wu

1. Introduction 55
   1.1. Non-NMR Methods 56
   1.2. NMR Methods 57

2. NMR Strategies for RNA Solution Structure Determination ... 58
2.1. Proton FT-NMR in ^1H_2O ... 58
2.2. Adjunct Experiments for Assignment of Proton Resonances 70
2.3. Nuclei Other than Protons ... 77
3. Structural Analysis of Specific 5S rRNAs ... 85
3.1. *E. Coli* 5S rRNA and Its Enzyme-Cleaved Fragments ... 86
3.2. *B. subtilis* 5S rRNA ... 89
3.3. Wheat Germ 5S rRNA and Its Enzyme-Cleaved Fragments 91
3.4. Yeast 5S rRNA and Its Enzyme-Cleaved Fragments ... 94
3.5. Other 5S rRNAs ... 98
3.6. Yeast 5.8S rRNA ... 98
4. NMR Analysis of rRNA Structural Dynamics ... 102
4.1. Base-Pair Proton Exchange with Solvent ... 102
4.2. Heat-Induced Melting ... 106
4.3. Effects of Added Cations (Especially Mg^{2+}) ... 109
5. Conclusion and Perspectives for Future Work ... 112
5.1. Sequence Variation ... 112
5.2. Ribosomal Protein–RNA Interactions ... 113
5.3. Two-Dimensional Proton FT-NMR ... 113
References ... 114

Chapter 3

Structure Determination via Complete Relaxation Matrix Analysis (CORMA) of Two-Dimensional Nuclear Overhauser Effect Spectra: DNA Fragments

Brandan A. Borgias and Thomas L. James

1. Introduction ... 119
2. Structure Evaluation Based on 2D NOE Spectroscopy ... 121
2.1. Basic 2D NOE Experiment: $90° - t_1 - 90° - \tau_m - 90° - t_2$ 122
2.2. Complete Relaxation Matrix Analysis—CORMA ... 122
2.3. Structure Evaluation by CORMA ... 130
2.4. Direct Determination of Distances by CORMA ... 134
2.5. Refinement of Structures Based on CORMA ... 135
3. Experimental Details of 2D NOE Spectroscopy ... 143
3.1. Special Techniques ... 144
3.2. Resonance Assignment ... 146
3.3. Intensity Measurement ... 147
4. Detailed Molecular Structure from 2D NOE Intensities ... 149
References ... 151

Chapter 4

Methods of Proton Resonance Assignment for Proteins

Andrew D. Robertson and John L. Markley

1. Introduction 155
   1.1. Features of the Proton NMR Spectrum of a Protein 157
   1.2. Proton NMR Assignments in Proteins 159
2. Protein Requirements 160
3. Instrumental Requirements 161
4. Sequential Assignments 162
   4.1. Amino Acid Spin System Identification 164
   4.2. Sequence-Specific Resonance Assignments: Connecting Spin Systems 168
   4.3. Resolving Resonance Overlap 170
5. Protein Structure Analysis with Sequential Assignment Data ... 172
   References 175

Chapter 5

Solid-State NMR Spectroscopy of Proteins

Stanley J. Opella

1. Introduction 177
2. Protein Dynamics 179
   2.1. Backbone Motions 181
   2.2. Sidechain Motions 185
3. Protein Structure 187
   3.1. Oriented Samples 188
   3.2. Unoriented Samples 193
4. Perspective 195
   References 196

Chapter 6

Methods for Suppression of the H_2O Signal in Proton FT/NMR Spectroscopy: A Review

Joseph E. Meier and Alan G. Marshall

1. Introduction 199
2. Frequency-Selective Excitation 200
   2.1. Continuous-Wave and Rapid-Scan Techniques 201
   2.2. Tunable Notch Filter 202
   2.3. Frequency-Selective Hard Pulse Sequences 202

2.4. Frequency-Selective Soft Pulse Sequences 209
2.5. Tailored Excitation 211
2.6. Nonlinear Excitation 213
3. Selective Presaturation 217
4. Software-Based Methods 218
5. Relaxation Methods 220
5.1. T_1-Based Methods 221
5.2. T_2-Based Methods 225
5.3. $T_{1\rho}$-Based (Spin-Lock) 228
6. Polarization Transfer Methods 229
6.1. Heteronulcear Polarization Transfer 229
6.2. Homonuclear Polarization Transfer 232
7. Multiple-Quantum Filtering 233
8. Experimental Comparisons 235
8.1. Redfield $2\underline{1}4\underline{1}2$ Soft-Pulse Sequence 235
8.2. $1\underline{1}$ Hard-Pulse Sequence 236
8.3. $1\underline{3}3\underline{1}$ Hard-Pulse Sequence 236
References 238

Contents of Previous Volumes 241

Index 249

1

Phosphorus NMR of Membranes

Philip L. Yeagle

1. INTRODUCTION

This review describes the use of ^{31}P nuclear magnetic resonance (NMR) to study the structure and dynamics of model and biological membranes. ^{31}P NMR has proven to be a powerful, nonperturbing probe of membranes. This review begins with an introduction to the structure of the membranes under consideration. Other reviews in the area include Seelig (1978), Yeagle (1978), and Smith and Ekiel (1984).

1.1. Model and Biological Membrane Structures

Living cells derive much of their structure and function from their cell membranes. The plasma membrane, for example, constitutes the outer boundary of many cells, controlling the permeability of the cell and specifically pumping essential materials into and out of the cell. Hormonal stimulation of cells is mediated at the plasma membrane, as are cell-cell interactions that promote tissue organization. Many important metabolic functions are located at the plasma membrane. The plasma membrane is also involved in the structural organization of the rest of the cell through interactions with cytoskeletal components.

These and other functions echo throughout the architecture of living cells, as a myriad of membranes work to promote basic cell function, cell growth, and cell differentiation. Included in these functions are synthesis

Philip L. Yeagle • Department of Biochemistry, School of Medicine, State University of New York at Buffalo, Buffalo, New York 14214.

of new membrane material by the endoplasmic reticulum and Golgi bodies receptor cycling via plasma membranes and endosomes, oxidative phosphorylation and steroid hormone synthesis on mitochondrial membranes, and control of mRNA exit from the nucleus by the nuclear membrane. While the individual functions of cellular membranes are too extensive to review here, it is nonetheless clear that the variety of critical cell functions controlled and made possible by cell membranes has stimulated intense research effort toward understanding cell membrane structure and function.

An understanding of cell membrane structure and function begins with an appreciation of the basic structural and functional components of membranes. The two major classes of membrane components are lipids and proteins. The structures of these components and the thermodynamic consequences of those structures determine the structure of biological membranes, which in turn controls their function.

The lipids of biological membranes exhibit incredible variety. Only a small portion of these lipids have been examined. However, overall lipid structural themes lead to some shared behavioral characteristics. Lipids of biological membranes are in general amphipathic. They have polar, or hydrophilic, headgroups and nonpolar, or hydrophobic, tails. In this review, the behavior of phospholipids will be emphasized. In general, each phospholipid contains a glycerol as the "backbone" with a phosphate esterified at position 3′ of the glycerol. At positions 1′ and 2′ hydrocarbon chains are found. Often the bonding is that of an ester or an ether. If the structure of the phospholipid is restricted to these entities one obtains the structure of phosphatidic acid.

Most phospholipids have an alcohol esterified to the phosphate. For example, if choline is esterified to the phosphate, phosphatidylcholine results. If ethanolamine is esterified to the phosphate, phosphatidylethanolamine results. More details of these structures can be found elsewhere. Differences in headgroup structure lead, in some cases, to dramatic differences in the behavior of phospholipids.

Differences in the structure of the hydrocarbon chains also lead to differences in the behavior of the phospholipids. The hydrocarbon chains may be fully saturated straight-chain hydrocarbons, or they may be unsaturated. The hydrocarbon chains may be branched, or they may be hydroxylated. The hydrocarbon chains may also vary considerably in length. The consequences of some of these structural variations will be observed in this review. Figure 1 shows schematically the structural elements of phospholipids.

An interaction between the hydrophobic tails of the phospholipid and the aqueous phase is entropically unfavorable. In contrast, the hydrophilic headgroups of the lipids are usually well able to interact with the aqueous phase. These principles lead directly to the structure of the lipid bilayer

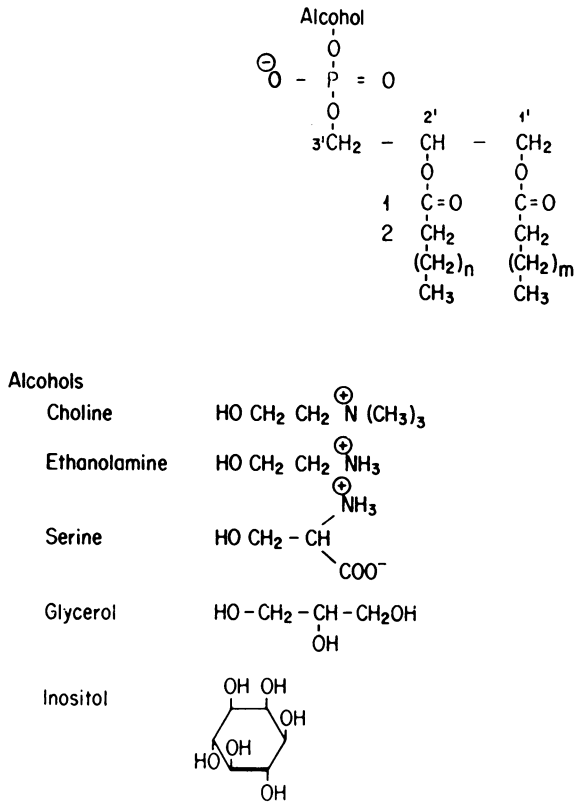


Figure 1. Structure of the phospholipid classes, according to headgroup.

commonly found in biological membranes in aqueous media, such as in the cell. It is worth noting that it is the thermodynamics of the interactions between the structural elements described in general above and the aqueous phase that leads to the lipid bilayer. For this reason, water is properly considered the third crucial element of membranes.

Figure 2 shows the structure of the lipid bilayer. This structure sequesters the hydrophobic portions of the lipids from the aqueous media. Only the polar headgroups of the lipids are left to encounter the aqueous phase. This prevents the entropic problems mentioned above and thus satisfies the hydrophobic effect.

Many pure phospholipids, dispersed in aqueous media, spontaneously form lipid bilayers. Extensive studies have led to the conclusion that the lipid bilayer is the structure formed by lipids in cell membranes, also. For this reason, a vast literature exists of studies of lipid bilayers, consisting only of lipids, as models for lipid behavior in biological membranes. Three

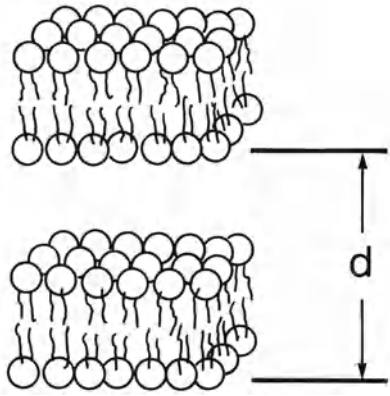


Figure 2. Schematic representation of the phospholipid bilayer. The balls represent the phospholipid headgroups and the wavy lines represent the hydrocarbon chains of the phospholipids.

such model membrane systems which have been extensively employed in NMR studies will be briefly described here.

The smallest and most uniform preparation is the sonicated, unilamellar phospholipid vesicle (SUV) (Huang, 1969). Such a vesicle consists of a single phospholipid bilayer enclosing a small aqueous space. Depending upon the preparation and lipid content, each vesicle is 25–30 nm in diameter. Each contains 2000–3000 lipid molecules, with more phospholipids in the outer half of the bilayer leaflet (on the outside surface of the vesicle) than in the inner half of the bilayer leaflet (on the inside surface of the vesicle). The rotational correlation time for these vesicles is in the range of 10^{-6} sec.

A larger, unilamellar model membrane system can be made by detergent dialysis, ethanolic injection, or other methods. Such a large unilamellar vesicle (LUV) contains a substantially larger aqueous space than the SUV. It has a diameter of from 150 to 300 nm. This increase in diameter is sufficient to increase the rotational correlation time beyond the msec range.

Probably the most common model membrane system to be studied by magnetic resonance is the multilamellar liposome (MLV). This is a large system, in excess of 1000 nm diameter. Rotational correlation times here are in the sec range. An MLV consists of concentric shells, each shell consisting of a phospholipid bilayer. Often 10 to 12 shells are found in a MLV.

Biological membranes are usually isolated from tissue or cell suspension and consist of vesicles that are heterogeneous in size. Usually these vesicles are large enough to have rotational correlation times longer than 1 msec. However, it is not unusual for smaller vesicles to form during isolation procedures. These smaller vesicles have much shorter rotational correlation times than the large vesicles. The effects on NMR lineshapes can be profound, and the researcher should always be alert for such effects. Usually, this requires electron microscopy to examine the samples that are being investigated by NMR.

Biological membranes, as noted above, contain protein as their other major component. Figure 3 shows schematically the nature of the incorporation of proteins into the lipid bilayer. As one might suspect, the portions of the proteins which must span the hydrophobic portion of the lipid bilayer are themselves hydrophobic. Often they consist of linear sequences of amino acids with hydrophobic side chains. In some cases, these transmembrane sequences have been shown to be in alpha helices.

The membrane describes an essentially two-dimensional world, with only a limited third dimension (compared to the extent of the other two dimensions). This structure influences the nature of the NMR spectra derived from the membranes. The membrane components diffuse laterally in the plane of the membrane as long as the membrane is liquid crystalline, unless the proteins are anchored by external cytoskeletal elements. Membrane components undergo rotational diffusion, often with their director oriented approximately perpendicular to the membrane surface. Time dependence of the orientation of the director (or wobble) is also sometimes measurable. Internal molecular motions are also often observed with membrane components. However, translational diffusion in the third dimension (i.e., perpendicular to the plane of the membrane) is severely restricted by the hydrophobic effect.

The interior of a membrane is highly anisotropic. This anisotropy arises from the orientation of the membrane components. This ordering effect

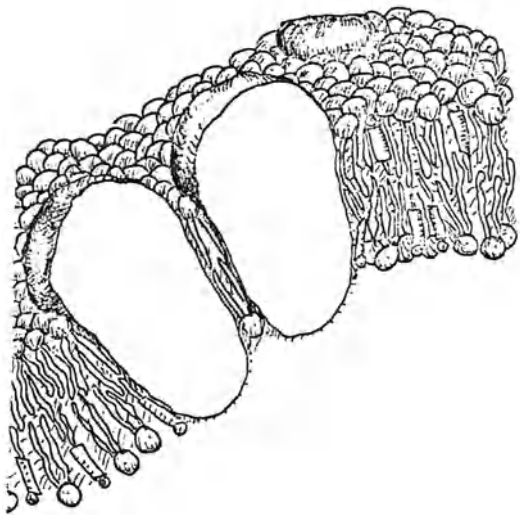


Figure 3. Schematic representation of a membrane containing integral membrane proteins in addition to the phospholipid bilayer.

gives rise to the liquid crystal nature of the membrane. Lowering the temperature can often induce a phase transition of the membrane to a gel state. All motion is severely restricted in the gel state, including translational diffusion and internal motions as well as rotational diffusion. These changes are reflected in the properties of the NMR spectra obtained from the phospholipids in the membrane.

1.2. ^{31}P NMR of Membranes—Theoretical and Practical Aspects

^{31}P NMR is a powerful, nonperturbing means of examining the structure and dynamics of membranes. The ^{31}P nucleus is 100% naturally abundant. Each phospholipid contains at least one phosphate in its headgroup. Therefore, ^{31}P NMR spectra can be accumulated from membranes without the addition of any probes to the membranes under study. This almost uniquely allows nonperturbing studies of biological membranes. Most other NMR experiments require the enrichment of the isotope of interest in the membrane. To enrich the membranes with a particular isotope requires the introduction of synthetic lipids into the membrane under study, which by definition prohibits the nonperturbing study of the native biological membrane.

^{31}P is one of the more sensitive nuclei for the NMR experiment. ^{31}P NMR exhibits 6.6% of the sensitivity of ^{1}H NMR. For this reason, not only have model membranes been extensively studied, but many preparations of biological membranes have also been examined. Because of these factors, ^{31}P NMR has permitted a broad study of membranes that has contributed significantly to our understanding of membrane structure and dynamics.

^{31}P NMR primarily senses the behavior and environment of the phosphorus atom in the phospholipid headgroup. The phospholipid headgroup in part defines the surface of the membrane. Therefore, ^{31}P NMR measures most directly the properties of the membrane surface; other properties are sensed indirectly. The surface is an important part of the membrane, where much of the functional activity is located. Thus it is advantageous to have such an excellent means of determining membrane surface properties. However, one should keep in mind what is being measured when attempting to integrate conclusions obtained from ^{31}P NMR data with data obtained from techniques that are primarily sensitive to other parts of the membrane structure. For example, calcium binding to phosphatidic acid immobilizes the headgroups, but has little effect on the hydrocarbon chains of the same phospholipid (Smaal *et al.*, 1987).

^{31}P NMR is equally useful when applied to SUV, LUV, MLV, biological membrane preparations, tissues, and whole cells. Therefore, ^{31}P NMR is more useful than techniques that can only be performed in solution phases

such as solution, or SUV only. ^{31}P NMR is blind to any materials in the sample that do not contain phosphorus. ^{31}P NMR is considerably more sensitive than ^{13}C NMR or ^{2}H NMR.

1.2.1. Theoretical Aspects

1.2.1.1. Chemical Shift Tensor. The ^{31}P chemical shift tensor of phosphodiesters is anisotropic. Measurements of the principle elements of the chemical shift tensor have been made for both simple organic phosphodiesters and for phospholipids in dry powders. The powder patterns show a clear manifestation of the three principle elements of the chemical shift tensor, σ_{11}, σ_{22}, σ_{33}. See Figure 4 for an example. The results of these measurements have been tabulated in Table 1. The full range of the chemical shift tensor is over 200 ppm ($\sigma_{33} - \sigma_{11}$).

The relationship between the three principle elements of the chemical shift tensor and the molecular coordinate system of the phosphate has been determined in crystals of organic phosphodiesters (Kohler and Klein, 1977; Herzfeld *et al.*, 1978). This orientation is represented in Figure 5. No such determinations have been made for phospholipids because of the unsuitability of phospholipid crystals for ^{31}P NMR experiments.

The trace of the chemical shift tensor represents the isotropic chemical shift. Although there is considerable similarity among the phospholipid chemical shifts tensors, there are measurable differences in the isotropic chemical shifts of the various phospholipids. High resolution ^{31}P NMR measurements of phospholipids in solution or in detergent micelles reveals that the isotropic chemical shifts of the phospholipids range over a few ppm, most within nearly 1 ppm of each other (London and Feigenson, 1979; Sotirhos *et al.*, 1986). Table 2 presents a summary of the results.

These differences in isotropic chemical shifts can be observed in SUV (Yeagle *et al.*, 1977a) and in serum lipoproteins (Yeagle *et al.*, 1977b). In

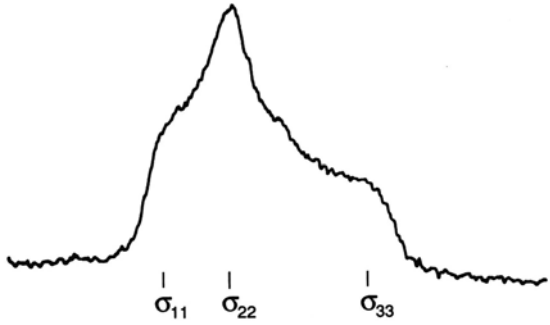


Figure 4. ^{31}P NMR powder pattern obtained from a dry powder of phosphatidylethanolamine using crosspolarization NMR.

TABLE 1
Principal Values of the ^{31}P Shielding Tensor of Some Phospholipids and Phospholipid Constituents Relative to 85% H_3PO_4[a]

	σ_{11}	σ_{22}	σ_{33}	Temp (°C)
A. Phospholipid constituents				
Urea phosphoric acid				
crystal	−26.6	−2.5	44.6	20
powder	−24	−5	45	20
Phosphoethanolamine				
crystal	−67	−13	69	25 ± 5
powder	−63	−8	65	25 ± 5
	−63	−8	72	−72
Serine phosphate				
crystal	−48	−2	51	25 ± 5
powder	−46	−3	52	25 ± 5
	−42	1	61	−110
Glycerol phosphate	−63	5	44	25 ± 5
	−63	5	44	−100
Barium diethylphosphate				
crystal	−75.9	−17.5	109.8	20
powder	−79	−19	113	20
	−80	−22	108	−110
B. Phospholipids				
DPPC				
anhydrous	−98	−34	134	20
monohydrate	−81	−25	108	25 ± 5
	−87	−25	119	25 ± 5
	−81	−25	110	20
50% (by wt.) H_2O	−81	−21	108	−110
DMPC				
anhydrous	−97	−34	133	20
monohydrate	−81	−22	110	20
DPLC				
anhydrous	−96	−32	133	20
monohydrate	−79	−19	108	20
DPPE				
anhydrous	−85	−14	87	25 ± 5
	−81	−20	105	25 ± 5
	−83	−21	103	20
50% (by wt.) H_2O,				
pH = 5.4	−81	−22	104	−110
pH = 10.8	−80	−21	104	−110
DLPE (anhydrous)	−84	−23	100	20
DSPA (anhydrous)	−40	−4	48	25 ± 5
PS (anhydrous)	−80	−20	112	25 ± 5

[a] From Seelig (1978).

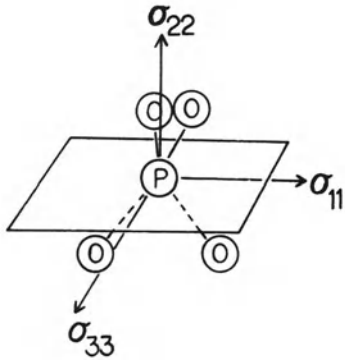


Figure 5. Relationship between the principal elements of the chemical shift tensor and the structural elements of the phosphate group.

these systems, it is therefore possible to study the behavior of individual phospholipids in the same system independently, yet simultaneously. As shown in Figure 6 the same possibility exists in membranes exhibiting ^{31}P powder patterns rather than isotropic resonances. The use of magic angle spinning (MAS) brings to effect the scaling factor:

$$(3 \cos^2 \theta - 1)$$

where θ is the angle between the axis of rotation and the external magnetic field. This factor scales chemical shift anisotropy as well as dipolar interactions. When $\theta = 54.7°$ then this scaling factor approaches zero, provided the spinning is fast enough. Therefore, MAS causes the collapse of the powder patterns of PE and PC, in this case to their isotropic resonances, and the individual phospholipids can again be studied independently yet simultaneously in the same membrane.

TABLE 2
^{31}P Chemical Shifts of Phospholipids[a]

Species	Chemical shift from external 85% H_3PO_4 (ppm) in cholate
PC	+0.65
PI	+0.40
lyso PC	+0.15
PS	+0.12
PE	+0.000
SM	+0.00
CL	−0.31
PG	−0.43
PA	−3.8

[a] From London and Feigenson (1979).

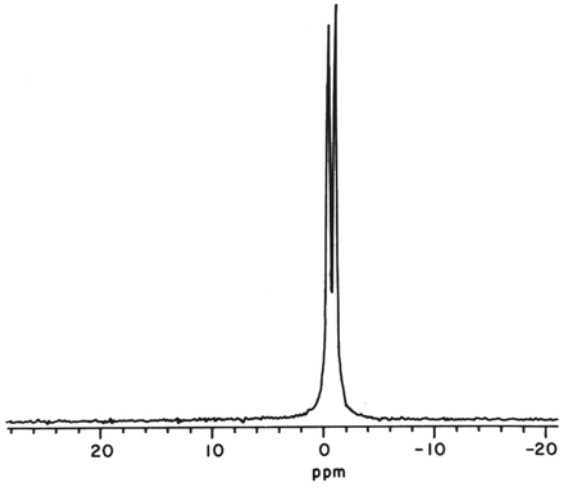


Figure 6. MAS ^{31}P NMR spectra of a mixture of PE and PC in MLV. Because of the averaging produced by the MAS, the difference in the isotropic chemical shift of PE and PC is evident.

The origin of the differences in the isotropic chemical shifts is not well understood. A dependence of the isotropic chemical shift on the geometry of the phsophate has been proposed (Wigfield *et al.*, 1975). The interaction of charged ions with the phosphate will affect the ^{31}P isotropic chemical shifts. For example, the binding of calcium to phosphatidylcholine causes an upfield shift in the ^{31}P NMR resonance of the phospholipid phosphate (Hutton *et al.*, 1977). Changes in the chemical bonding around the phosphate will also cause shifts.

Two examples of the latter effect are available. If the ester bond on one side of the phosphate is replaced by a carbon–phosphorus bond, a 20-ppm change in chemical shift is observed (Ferguson *et al.*, 1982; Jarrell *et al.*, 1981). This has been measured for the naturally occurring phospholipids. If one of the nonesterified oxygens of the phosphate is replaced with sulfur, a larger effect is observed (Chupin *et al.*, 1979; Vasilenko *et al.*, 1982; Bruzik *et al.*, 1983). In this case, a 50-ppm downfield shift in the isotropic chemical shift is observed. This observation has a useful side effect. The difference in the chemical shift tensors is sufficient to separate the axially symmetric powder patterns of the sulfur-labelled phospholipid from the normal phospholipid in phospholipid bilayers. This again offers the opportunity to measure the behavior of two different phospholipids in the same membrane, independently yet simultaneously without sonication or MAS.

1.2.1.2. Relaxation. Considerable study has been made of phosphorus relaxation, although as in the question of phosphorus chemical

shifts, the understanding of phosphorus relaxation is as yet incomplete. Spin lattice relaxation T_1 has been the most extensively studied. ^{31}P T_1 has been studied in most model membrane systems as well as in biological membrane preparations. For representative values, see Table 3.

The question of relaxation mechanism has been investigated carefully. In principle, the spin lattice relaxation of the phosphorus can be represented as follows:

$$\frac{1}{T_{1(\mathrm{obs})}} = \frac{1}{T_{1(\mathrm{DD})}} + \frac{1}{T_{1(\mathrm{CSA})}} \tag{1}$$

where the subscript DD refers to dipolar interactions with protons and the subscript CSA refers to relaxation due to a time dependent modulation of the chemical shift anisotropy. The data presently available suggest that ^{31}P T_1 is dominated by dipolar interactions with the protons at moderate field strengths. The static interaction between the phospholipid methylene protons (no directly bonded protons) and the phosphorus has been calculated to be about 6 kHz (Chan *et al.*, 1981). In hydrated membranes, this interaction is substantially motionally averaged. However, it is apparently of sufficient strength to dominate T_1.

The dominance of the dipolar interaction in ^{31}P T_1 at moderate field strengths has been demonstrated in two ways. First, at moderate field, a substantial ^{31}P $\{^1H\}$ nuclear Overhauser effect (NOE) has been observed for phospholipids in SUV and MLV (Yeagle *et al.*, 1975b). Since this effect arises from ^1H–^{31}P dipolar interactions, the presence of such dipolar interactions is indicated. The ^{31}P $\{^1H\}$ NOE measured in SUV and MLV is 1.4 at 40 MHz. The maximum possible is 2.24 (Yeagle *et al.*, 1975a).

$$\mathrm{NOE} = \frac{\rho_{dd}}{\rho_{dd} + \rho^*} \cdot \frac{\gamma_H}{2_{\gamma P}} \tag{2}$$

TABLE 3
^{31}P Spin Lattice Relaxation Times of Phospholipids in Membranes (23 °C)

Membranes	T_1 (sec)
Unsonicated egg PC	1.4 ± 0.1
Human erythrocyte ghost	1.5 ± 0.1
Sarcoplasmic reticulum	1.3 ± 0.1
ROS disk	1.4 ± 0.1
Vesicular stomatitis virus	0.1 ± 0.1

where γ represents the respective magnetogyric ratios, ρ_{dd} is the relaxation rate of ^{31}P due to proton-phosphorus dipole-dipole interactions, and ρ^* is the relaxation rate due to other relaxation mechanisms. Measurement of the NOE as a function of temperature showed an increase with temperature (Yeagle *et al.*, 1977a). These data suggested that the NOE observed was motionally limited and not limited by alternate relaxation mechanisms such as chemical shift anisotropy. The dependence upon correlation time is expressed in the following relation:

$$\mathrm{NOE} = 1 + \frac{6J_2(\omega_H + \omega_P) - J_0(\omega_H - \omega_P)}{J_0(\omega_H - \omega_P) + 3J_1(\omega_P) + 6J_2(\omega_H + \omega_P)} \tag{3}$$

where $J_i(\omega)$ are spectral density functions for anisotropic motion as experienced by the phospholipid headgroup.

Second, the field dependence of the ^{31}P T_1 and ^{31}P $\{^1H\}$ NOE was measured from 25 MHz to 109 MHz. The ^{31}P T_1 has little or no field dependence (Yeagle *et al.*, 1984). A strong field dependence would be expected if chemical shift anisotropy were making a large contribution to the spin lattice relaxation, since the chemical shift anisotropy relaxation is proportional to the square of the field strength, H.

$$\frac{1}{T_{1(\mathrm{CSA})}} \propto H^2 f(\tau_c) \tag{4}$$

This effect of chemical shift anisotropy will be noted where the relevant correlation time is in the fast motion regime (see below). Finally the measured field dependence of the ^{31}P $\{^1H\}$ NOE is exactly what is predicted from the theoretical expression describing the field dependence of the NOE. At 40 MHz, the NOE is 1.4. At 109 MHz, the NOE is 1.1 (both determinations at 22 °C). This difference is predicted by the frequency dependence in equation (3), because $1/\tau_c$ is near ω_0 for ^{31}P NMR of phospholipids in membranes. Therefore, at moderate field strengths, the dominance of the dipolar relaxation term is demonstrated. At high fields, the question requires further investigation.

An interesting outgrowth of the studies just described is the conclusion that the ^{31}P T_1 relaxation for phospholipids in membranes is dominated by a correlation time in the fast motion regime ($\omega_0\tau_c < 1$). The NOE data most directly lead to this conclusion because in the slow motion regime no ^{31}P $\{^1H\}$ NOE can be observed. That is, equation (3) predicts essentially no NOE when $\omega_0\tau_c \gg 1$. Therefore one is faced with something of an enigma in the ^{31}P NMR of LUV, MLV, and biological membranes. The T_1 is in the

fast motion regime, which is therefore "liquid-like." Yet the ^{31}P NMR spectra of the same systems show powder patterns, which are "solid-like." This is perhaps less surprising when one recalls that phospholipid bilayers are liquid crystalline (in most of the studies reviewed here), exhibiting highly anisotropic motion, which may be rapid but also is ordered, or motionally restircted. Therefore the investigator is faced with the challenge in studying ^{31}P NMR of phospholipid bilayers and biological membranes of deciding whether high resolution NMR techniques or solid state NMR techniques are most appropriate for the question at hand.

Given the ordered environment of the membrane, one might question whether ^{31}P T_1 is anisotropic across the powder pattern observed in LUV, in MLV, and in biological membrane preparations. No indication has been obtained to data to indicate any anisotropy in ^{31}P T_1, which is likely due to the relatively long ^{31}P T_1 normally observed, allowing for motional averaging.

The motion is explicity demonstrated in magnetization transfer experiments. Using a rapid pulse train, it is possible to generate discrete frequencies for irradiation (DANTE sequence). Using this pulse sequence, de Kruijff *et al.* (1980) selectively irradiated a portion of the ^{31}P powder pattern for MLV. Time evolution of their measurements showed a relatively rapid distribution of magnetization among all regions of the powder pattern, compared to the ^{31}P T_1. Therefore the T_1 measurements would not be expected to show any anisotropy. T_2, which is on the msec time scale, would, however, be expected to show anisotropy. The fulfillment of this expectation is discussed below.

A model for ^{31}P T_1 in phospholipid bilayers has been described. This model incorporates anisotropic headgroup rotational diffusion, with the director oriented perpendicularly to the membrane surface. Relaxation is dominated by interactions with headgroup protons and is at least in part intermolecular (Yeagle *et al.*, 1975a; 1977a). In this model, consistency is achieved between predictions of the strength of the ^{31}P $\{^1$H$\}$ dipolar interactions based on the molecular model developed and the T_1 observed.

^{31}P T_2 relaxation for phospholipids in membranes is more complex. This suggests that at the longer timescales to which T_2 is sensitive, motions in the membrane are more complicated. T_2 is anisotropic across the powder pattern. Values for T_2 are in the range of 1 to 20 msec. T_2 proves to be much more sensitive to changes in membrane properties than does T_1. Unfortunately, the details of the mechanisms and the important motions and correlation times involved in ^{31}P T_2 relaxation for phospholipids in membranes have not been resolved.

$T_{P\text{-}H}$ is the relaxation parameter characterizing the transfer of spin polarization from protons to phosphorus in a crosspolarization experiment. $T_{P\text{-}H}$ data have been reported (Frye *et al.*, 1985). However, even less is

known about the mechanisms for this potentially important parameter in phospholipid bilayers than about T_2.

1.2.2. Practical Aspects

The most prevalent problem in the literature of ^{31}P NMR of membranes is spectral distortion. Most ^{31}P NMR powder patterns reported in the literature do not represent the true lineshape. These spectral distortions make interpretations of the lineshapes difficult and risky.

Inadequate ^{1}H decoupling is one source of such distortions. ^{1}H decoupling must be employed to obtain the sharp edges of powder patterns necessary for the proper interpretation of those powder patterns. Lack of ^{1}H decoupling, distorting the ^{31}P powder pattern of dry phospholipids, led some early investigators to the erroneous conclusion that the ^{31}P chemical shift tensor of the phospholipids was axially symmetric (McLaughlin *et al.*, 1975a). However, the conditions for adequate decoupling are not difficult to achieve in most cases. Although the static interaction is of the order of 6 kHz (Chan *et al.*, 1981), in the liquid crystalline state the phospholipids experience rapid axial diffusion which partially motionally averages the ^{31}P–^{1}H dipolar interactions. Furthermore, at moderate and high magnetic field, the chemical shift anisotropy readily dominates the lineshape (at frequencies of 40 MHz or less, the dipolar interactions can alter the lineshape beyond recognition). Therefore obtaining adequate ^{1}H decoupling conditions on most commercial spectrometers in the liquid crystalline state is not difficult in ^{31}P NMR of phospholipids in membranes.

Serious problems concerning ^{1}H decoupling arise in anhydrous powders where the powder patterns are over 200 ppm in width. However, the best means to obtain powder patterns in anhydrous powders is to employ crosspolarization, which requires strong ^{1}H decoupling anyway. This technique exploits the strong ^{1}H–^{31}P dipolar interactions found in anhydrous powders. Difficulties arise in performing the crosspolarization experiment in the liquid crystal state of phospholipid bilayers. Because of the rapid axial diffusion, the dipolar interactions are partially motionally averaged. The weakened dipolar interactions make the Hartman–Hahn match difficult to achieve (Frye *et al.*, 1985).

The most common source of spectral distortions arises from the methods employed to collect the ^{31}P NMR spectra. The powder patterns obtained from model and biological membranes are broad. As a result, the decay of the signal in the time domain (FID) is rapid. To illustrate the seriousness of this phenomenon for data collection, a typical example will be given. When employing a frequency width of 50 kHz, the majority of the information concerning the lineshape is contained in the first few points of the FID; for a typical powder pattern 45 ppm in width, the most crucial region

is the first 50 or so points of the FID. In the case where broader components are observed, the first 10–15 data points may contain the majority of the information concerning lineshape.

On most commercial spectrometers designed primarily for high resolution spectra, the rf gates are not optimized for rapid turnoff and turnon. This introduces distortions in the collection of the data. Furthermore, because of the high Q of the coils in such instruments and consequent long "ring down," substantial delays must be introduced between the initial excitation pulse and gating on the receiver to eliminate from the FID lingering artifacts due to the excitation pulse. For these and other reasons, normal single-pulse-acquisition experiments invariably lead to the loss and/or distortion of the first portion of the FID. As noted above, the most crucial information for defining ^{31}P powder patterns of membranes is found in the first portion of the FID. As a result, most ^{31}P NMR spectra reported in the literature are plagued with distortions of the powder patterns. When the powder patterns contain features broader than the normal bilayer powder patterns, the broader features will be severely attenuated in most single-pulse experiments, because the early portion of the FID is effectively lost.

Fortunately there is a solution to these problems. These distortions can be avoided by the use of a completely phase cycled Hahn echo (see Figure 7). Unless technology of this sort is employed, ^{31}P powder patterns that are obtained from membranes must be viewed with suspicion. The fully phase-cycled Hahn echo permits accumulation of undistorted spectra with flat baselines (Rance and Byrd, 1983). The latter are also crucial for the proper interpretation of ^{31}P powder patterns.

When using the Hahn echo, one must be careful about two further issues. One is that the echo time must be short relative to the T_2 of the resonances being observed. Secondly, the pulse widths used should be short enough to properly excite the full extent of the powder pattern expected. Pulse widths for 90° pulses on some spectrometers are too long to permit proper excitation of powder patterns that are broader than the normal bilayer powder pattern.

One of the easiest things to have happen in the accumulation of ^{31}P NMR spectra of membranes is to lose broad features or distinct broad

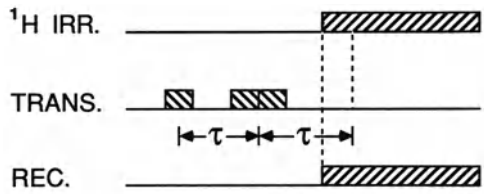


Figure 7. Hahn echo sequence.

components. A good example of the former can be seen in the reported spectra of biological membranes. The T_2 for the ^{31}P NMR powder pattern of biological membranes is an order of magnitude shorter than for many pure phospholipid bilayers. As a consequence the powder pattern of biological membranes is not identical with pure phospholipid bilayers. It is broader; the middle portion of the powder pattern appears to be "filled in" relative to the powder pattern of pure phospholipid bilayers. The height of the powder pattern at $\sigma_\perp$ relative to the height at $\sigma_\parallel$ is less in the biological membrane than in pure phospholipid bilayers. Spectral simulations demonstrate that this is a consequence of the shortened T_2 in the biological membrane. Yet many reported spectra of biological membranes appear to strive for the "idealized" lineshape. When that is obtained from a biological membrane, it is likely due to the removal of the very first portion of the FID arising due to a lack of adequate data collection at the very beginning of the decay (in the single pulse, immediately after the excitation or even slightly sooner, depending upon the duration of the pulse; in the Hahn echo, at the refocussing of the signal in the $x'y'$ plane of the rotating frame). The effect of these distortions in biological membranes is to make the powder patterns similar to the idealized lineshape for pure phospholipid bilayers. This artifact denies the difference in T_2 between the systems which is readily established experimentally. The effect of these distortions in pure phospholipid bilayers is to remove intensity from the middle of the spectrum, giving the powder pattern a swaybacked look. Figure 8 provides an example. Examination of the literature shows that these artifacts are quite commonly published.

1.2.3. History

The first ^{31}P NMR studies of membranes were reported in 1972 (Barker *et al.*, 1972; Horwitz and Klein, 1972; Davis, 1972; Davis and Inesi, 1972; Sheetz and Chan, 1972). These studies collectively demonstrated the viability

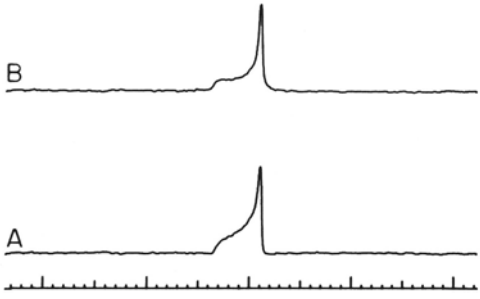


Figure 8. Effects of loss of the first portion of the FID on the transformed ^{31}P NMR spectra from MLV of PC. Spectrum A is properly obtained using fully phase cycled Hahn echos. Spectrum B is transformed with a loss of only 60 μsec of the FID. The serious distortions in lineshape are clearly evident.

of ^{31}P NMR studies on membranes. Most early work involved model membrane systems. However, in a study that was to point the way to much future work and the real power of ^{31}P NMR, ^{31}P NMR spectra were reported from *Escherichia coli* and rabbit muscle sarcoplasmic reticulum (McLaughlin *et al.*, 1975b). These early studies also gave indication of the complexity and richness of the ^{31}P NMR spectra, showing that anisotropic powder patterns could be expected in unsonicated membrane systems. For the sake of correctness, one should also note that early suggestions that the ^{31}P chemical shift tensor for phospholipids was symmetric (McLaughlin *et al.*, 1975a) and that ^{31}P T_1 was dependent upon vesicle size (Horwitz and Klein, 1972) are incorrect.

2. ^{31}P NMR AND MEMBRANE MORPHOLOGY

2.1. Introduction

^{31}P NMR has proven to be a highly useful means of examining the morphology of membranes. One usually conceives of phospholipids inhabiting a bilayer. Because of the highly anisotropic nature of the bilayer, one might expect a characteristic powder pattern that is uniquely identifiable with the bilayer. To some extent this is true. However, phospholipids can inhabit other phase structures in addition to the bilayer. Furthermore, even in the bilayer phase, there are several lineshapes depending upon the structure of the bilayer phase.

In this section, the use of ^{31}P NMR to assess the morphology of phospholipid structures will be reviewed. There is a substantial literature on this subject, some of it repetitive. Examples will be chosen to exemplify the spectroscopy and to highlight the conclusions reached. However, this review will not attempt to be comprehensive in this area. For a review of the older literature, the reader is referred to Cullis and de Kruijff (1979).

One important thought should accompany any consideration of ^{31}P NMR studies of phospholipid morphology. The ^{31}P NMR lineshapes which are observed for the various phase structures of the phospholipids are readily predictable from the motional averaging that obtains in the particular structure. However, since the lineshapes are the result of motional averaging, the lineshapes are not theoretically in a one-to-one correspondence with the particular phase structure. Thus, for example, whenever a hexagonal II phase is present, one can expect a particular ^{31}P NMR powder pattern. It is not theoretically required that observation of the "hexagonal" powder pattern indicate the presence of true hexagonal (II) phase in the sample. Thus the concurrent use of other techniques to corroborate the ^{31}P NMR studies is required. Yet it should be noted that studies to date have not provided a clear example of a "hexagonal" lineshape when

hexagonal II phase was not present (Gruner and Jain, 1985; Tilcock *et al.*, 1986).

In the following discussion, the lamellar, hexagonal (II) and some of the isotropic phases will be examined. The nature of the ^{31}P NMR spectra obtained will be described. Examples will be provided to illustrate the major points.

2.1.1. Lamellar

The phospholipid bilayer is a highly anisotropic structure. The dominant motion of the phospholipids for the ^{31}P NMR lineshapes is axial diffusion in the bilayer. The director for axial diffusion is ordered. Therefore, the effect of phospholipid motion in the membrane on the ^{31}P NMR lineshape can be described, to a first approximation, by application of a rotation matrix to the chemical shift tensor of the phosphorus. Such rotation leads directly to the axially symmetric powder pattern observed in the ^{31}P spectra from phospholipid bilayers. Figure 8 shows an example. However, the width of the powder pattern in the liquid crystalline state is less than that predicted from simple rotation of the ^{31}P chemical shift tensor. Therefore, to explain the width of the liquid crystalline powder pattern, an additional motion has been suggested (Kohler and Klein, 1977). A time dependence of the orientation of the director for axial diffusion would be adequate to explain the additional motional averaging in the liquid crystal. When phospholipids are in the gel state, this motional averaging (referred to as "wobble") is damped, and the observed width of the powder pattern increases.

In the dry powder, where there is no rapid motion, the three principal elements σ_{11}, σ_{22}, and σ_{33} of the chemical shift tensor are readily discernable (see Figure 4). The powder pattern obtained in the presence of rapid axial diffusion needs only two elements for adequate definition, $\sigma_{\perp}$ and $\sigma_{\parallel}$. Thus the upfield maximum of the powder pattern corresponds to the chemical shift $\sigma_{\perp}$ perpendicular to the axis of axial diffusion, while the downfield shoulder corresponds to the chemical shift $\sigma_{\parallel}$ parallel to the diffusion axis. In phospholipid bilayers, using this analysis, the director for axial diffusion has been shown to be perpendicular to the membrane surface by using the angular dependence of the ^{31}P chemical shift arising from oriented bilayers (Seelig and Gally, 1976).

To collapse the anisotropic ^{31}P powder pattern to the axially symmetrical powder pattern observed in phospholipid bilayers, the axial diffusion must be rapid relative to the full width $(\sigma_{33} - \sigma_{11})$ of the ^{31}P chemical shift tensor in Hz. If the rate of axial diffusion is not rapid by this definition, then alterations in the lineshape will be observed. These lineshape changes have been modelled, and a progression of lineshapes has been predicted that depends upon the rate of axial diffusion. This modelling

was used to investigate the effect of a transition from the liquid crystal to the gel state for a phospholipid bilayer. The authors concluded that a substantial reduction in the rate of axial diffusion occurs below the phase transition temperature (Campbell *et al.*, 1979).

2.1.2. Hexagonal (II)

The hexagonal (II) phase has attracted considerable interest in the past few years because it is a nonlamellar structure readily formed by phosphatidylethanolamine. A diagram of the hexagonal (II) phase can be found in Figure 9. It consists of tubes, each containing a small aqueous channel, with the phospholipid headgroups directed inward toward the aqueous channel. The hydrocarbon chains are directed towards the exterior of the tubes. The tubes of the hexagonal (II) phase are stacked with the hydrocarbon surfaces of the tubes in close apposition. This phase structure tends to separate from excess water, so that the primary exposure to the aqueous phase is in the aqueous pores.

An important difference is observed in the ^{31}P NMR powder patterns arising from the hexagonal (II) phase structure compared to the lamellar structure. The cylinders are small in diameter. Typically the diameter of the aqueous pore is about 2 nm. With a common diffusion coefficient for lateral diffusion of about $10^{-8}\,cm^2\,s^{-1}$, circumnavigation of the cylinder occurs with a correlation time of about 10^{-6} sec. This is rapid compared to the width in Hz of the ^{31}P powder pattern in the lamellar phase. Therefore, an additional motional averaging of the observed expression of the chemical shift tensor occurs, compared to the motional averaging that occurs in the lamellar phase. All other factors remaining constant, the transition from

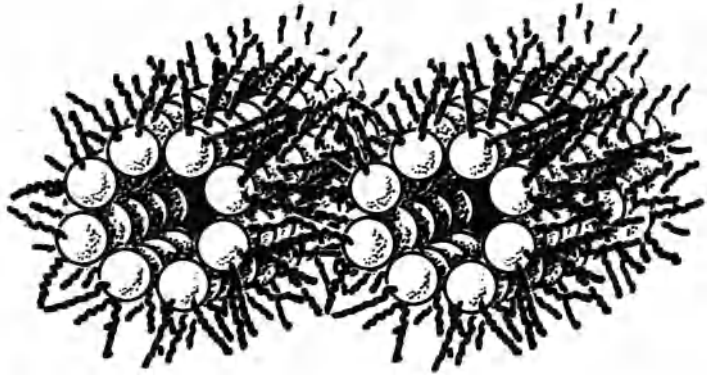


Figure 9. Schematic representation of hexagonal (II) phase structure.

the lamellar to the hexagonal (II) phase produces a reduction by half of the width of the powder pattern and an apparent reversal of the shape of the powder pattern. An example for 1-palmitoyl-2-oleoyl-phosphatidylethanolamine is shown in Figure 10.

2.1.3. Isotropic

The term isotropic phase is really a catchall phrase to encompass several different phase structures. The property they all have in common is that the ^{31}P NMR resonance produced is isotropic. That is, no powder pattern is observed. All the intensity is collapsed into one symmetrical, relatively narrow resonance.

The requirement for an isotropic phase in ^{31}P NMR is that the ^{31}P chemical shift tensor have 90° of rotational freedom with a correlation time short compared to 10^{-4} sec. At least five different means have been described to satisfy this requirement for phospholipids.

1. Solution phase, where the phospholipid is dissolved in organic solvent, satisfies the requirement (Klose, 1975).
2. Micellar phase satisfies the requirement since the micelles are small enough to have a correlation time for rotation shorter than 10^{-6} sec. Phospholipids dissolved in detergent micelles are one example (London and Feigenson, 1979).
3. SUV provide another example since their rotation is fast enough ($\tau_c = 10^{-6}$ sec) to satisfy the requirement (Yeagle *et al.*, 1975b). However, any vesicle larger than SUV produces ^{31}P linewidths too

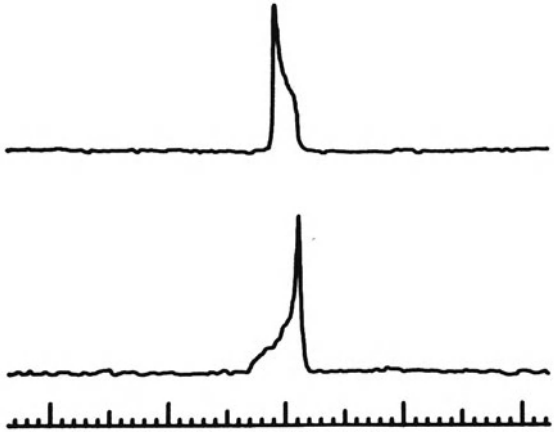


Figure 10. Example of change in ^{31}P powder pattern for PE. Lower spectrum is of lamellar phase and upper spectrum is of hexagonal (II) phase.

broad to distinguish the isotropic chemical shifts of individual phospholipid classes.

4. Cubic phase also produces an isotropic resonance. Cubic phase consists of a series of short rods connected at right angles, each rod resembling a short piece of hexagonal (II) phase. Phospholipids diffusing from one segment to another will experience the 90° of rotational freedom required for the isotropic phase. Since the length of the segments is similar to the diameter of the segments, normal lateral diffusion rates of phospholipids are adequate to provide sufficiently rapid movement through the 90° of rotational freedom to produce the isotropic resonance (Boni and Hui, 1983).
5. Finally, an inverted micellar phase has been suggested to produce an isotropic resonance. This structure is postulated to coexist within the phospholipid bilayer and is used to explain the lipidic particles observed in freeze fracture electron micrographs of some lipid mixtures at some temperatures. The structure offers rapid rotational freedom through lateral diffusion around the inverted micelle. This structure has been widely postulated and the concurrent observation of lipidic particles and isotropic ^{31}P NMR resonances has been documented. However, the details of this phase structure have not yet been adequately demonstrated experimentally (Cullis and deKruijff, 1979).

One more means is available to obtain an isotropic ^{31}P NMR resonance. It has been demonstrated that changing the orientation of the phosphorus chemical shift tensor relative to the axis of rotational diffusion of the phospholipid will produce significant changes in the observed lineshape (Thayer and Kohler, 1981). At particular orientations of the headgroup, isotropic resonances may result even from a lamellar phase. Furthermore, even "hexagonal" lineshapes can result. Although documented cases of these possibilities have not yet been presented for particular phospholipid systems, one should be cautious in interpretations until possibilities of changes in phospholipid headgroup conformation can be ruled out.

Returning to SUV, the size of the vesicles influences their rate of rotation. This, in turn, influences the linewidth of the ^{31}P NMR resonance from such a vesicle. The slower the rotation, the less able that motion is to average the chemical shift tensor, and chemical shift anisotropy then contributes to the linewidth. This has been demonstrated in a study of ^{31}P NMR spectra obtained from vesicles of different size (Burnell *et al.*, 1980).

If instead of increasing vesicle size, solution viscosity is increased, vesicle tumbling is also reduced. In this case, phospholipid lateral diffusion and the radius of curvature of the vesicles determine the extent of motional averaging of the chemical shift tensor. Using these parameters, a published

study arrived at a determination of a lateral diffusion coefficient (Cullis, 1976). A value of $D_t = 2.6 \times 10^{-8}\ \mathrm{cm^2\,s^{-1}}$ was obtained which is close to the values obtained by other techniques. Effects of a gel-to-liquid phase transition are likely underestimated in this study. Vesicle fusion and increase in dipolar interactions complicate the measurement.

2.2. Model Membranes

In the following, some selected examples will be discussed which will illustrate the use of ^{31}P NMR to study transitions between the various phase structures described above. ^{31}P NMR has some distinct advantages in studying the phase behavior of phospholipids. First, it is a reasonably sensitive technique. Therefore, relatively modest amounts of lipids can be employed, and a survey of a full temperature range can be readily obtained. Surveying a full temperature range with techniques like X-ray diffraction and freeze fracture electron microscopy presents a formidable task.

In addition, ^{31}P NMR is sensitive to all the phospholipids in a sample and can be made representative of the relative populations of each of the phase structures present. X-ray diffraction is not sensitive to disordered structures, such as some of the isotropic phases described above. Electron microscopy is very difficult to make truly representative of the relative abundance of various phase structures present.

The major weakness of the ^{31}P NMR approach is observed when isotropic phases are present. ^{31}P NMR does not offer an effective means to distinguish among the various possible sources of the isotropic resonance. In this case, other techniques are required to establish the structure of the phase.

Much of the following discussion will center on phosphatidylethanolamine, since that is the most extensively studied phospholipid which is capable of forming nonlamellar phases. However, we will begin with the example of phosphatidylcholine, which commonly forms the lamellar phase.

2.2.1. Phosphatidylcholine

Phosphatidylcholine (PC) provides a good example for examining the effects of a gel-to-liquid crystalline phase transition on the ^{31}P NMR powder pattern from MLV. It has been shown by several investigators that both the intrinsic linewidth of the ^{31}P NMR resonance and the width of the powder pattern $(\sigma_\perp - \sigma_\parallel)$ increase when the phospholipid enters the gel state (McLaughlin *et al.*, 1975b; Uhing, 1975; Gawritsch *et al.*, 1977; Cullis *et al.*, 1976). An increase in $(\sigma_\perp - \sigma_\parallel)$ indicates a general restriction in the motion of the phospholipid headgroup in the gel state. In one case such a study was performed in sonicated vesicles (Uhing, 1975). However, in such

studies one must be aware that sonicated vesicles exhibit a propensity to fuse when at the phase transition temperature. Fusion to form larger vesicles will also increase the apparent linewidth.

^{31}P NMR was also instrumental in identifying a subphase at about 11 °C for dipalmitoylphosphatidylcholine (Fuldner, 1981). The headgroup conformation has been shown both by ^{31}P NMR and by neutron diffraction, not to be greatly altered by the phase transition (Buldt *et al.*, 1978).

The crosspolarization (CP) ^{31}P NMR lineshape is particularly sensitive to the motional changes occurring in this transition (Frye *et al.*, 1985). Figure 11 shows the ^{31}P NMR lineshape for phosphatidylcholine in the liquid crystal state. A minimum in intensity occurs at an orientation of the director (for headgroup axial diffusion) at the "magic angle" with respect to the external magnetic field. The "magic angle" satisfies the expression $(3 \cos^2 \theta - 1) = 0$ and scales the dipolar interactions which permit the CP experiment. The headgroup is essentially engaging in magic angle spinning.

As the temperature is decreased, several changes occur. One is that $(\sigma_\perp - \sigma_\parallel)$ increases as the gel state is entered, but the powder pattern remains axially symmetric. Another change is the loss of the local minimum in the CP ^{31}P powder patterns. Figure 12 shows that this change occurs in the pretransition-to-main-transition temperature region. This change corresponds to a reduction in the axial diffusion rate of the headgroup.

2.2.2. Phosphatidylethanolamine

Phosphatidylethanolamine (PE) is an excellent example of a phospholipid that is capable of existing in nonlamellar phases (Marsh and

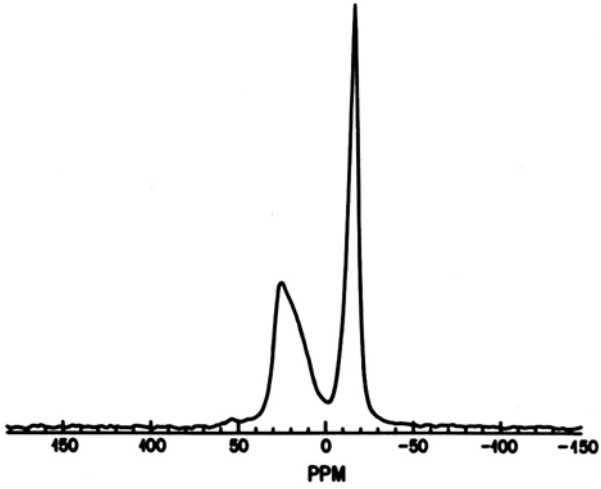


Figure 11. CP ^{31}P NMR lineshape from liquid crystalline MLV of PC.

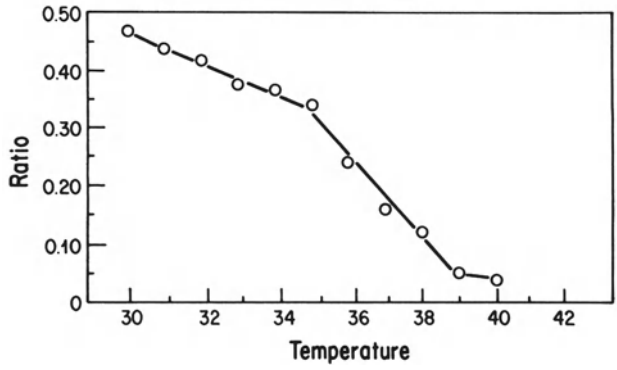


Figure 12. Ratio of intensity at $\sigma_\perp$ to intensity at orientation of director for headgroup rotation at magic angle with magnetic field (local minimum in spectra) as a function of temperature (Frye *et al.*, 1985).

Seddon, 1982). ^{31}P NMR studies of the phase behavior of pure phosphatidylethanolamine dispersions will be reviewed first, followed by studies of mixed phospholipid systems containing PE.

2.2.2.1. Pure Phosphatidylethanolamine. Pure PE is readily able to adopt the hexagonal (II) phase structure. The transition from the lamellar phase to the hexagonal (II) phase is induced by an increase in temperature (Cullis and deKruijff, 1978). The sensitivity of ^{31}P NMR to this phase change is demonstrated in Figure 10. This represents a simple transition in which only two phases are present: lamellar and hexagonal (II). The temperature of this transition is dependent upon the hydrocarbon chains of the PE. The more unsaturated the hydrocarbon chains, the lower the lamellar-to-hexagonal (II) phase transition temperature.

Studies have been performed which have provided a thermodynamic basis for the lamellar to hexagonal (II) phase transition of PE (Yeagle and Sen, 1986). As seen in Figure 13, the lamellar state of PE can be stabilized by chaotropic agents. These and other data have led to the conclusion that the PE headgroup forms a membrane surface that is relatively hydrophobic compared to the surface of a PC bilayer; see, for example, Yeagle and Sen (1986). The intermolecular headgroup interactions, described in Section 3.1, that characterize the surface of these membranes are manifest in the PE headgroup as intermolecular hydrogen bonding between headgroups. For this and other reasons, the PE headgroup interacts poorly with the aqueous phase and adopts the hexagonal (II) phase as a means of limiting its exposure to the aqueous phase. The chaotropic agents alter the structure of the aqueous phase, making it more compatible with the PE headgroup and removing the driving force for formation of the hexagonal II phase.

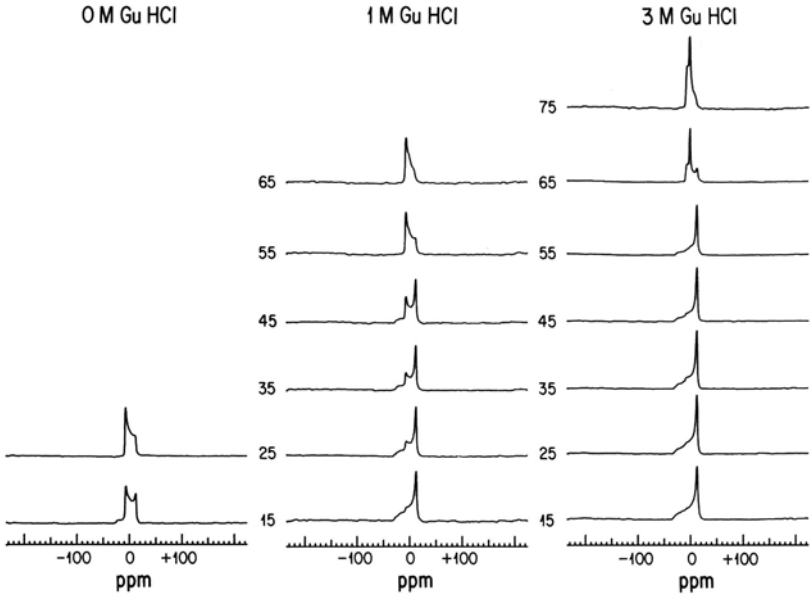


Figure 13. Effect of GuHCl on the phase behavior of soy PE (Yeagle and Sen, 1986).

2.2.2.2. PE Mixed with Other Phospholipids. When PE is mixed with other phospholipids, the phase behavior of the system becomes yet more interesting. Mixing PC with PE will stabilize the lamellar phase, dependent upon the PC content. The higher the PC content, the more stable the lamellar phase.

However, these spectra show the presence of another morphology for these PE-PC mixtures. An isotropic resonance is observed in some of the regions where lamellar and hexagonal (II) phase coexist. Freeze fracture electron microscopy of these phospholipid dispersions show the presence of lipidic particles when the isotropic resonance is present (Hui *et al.*, 1981). From these data a phase diagram can be constructed (see Figure 14).

Care must be taken when studying such systems, however. The lamellar-to-hexagonal (II) phase transition exhibits considerable thermal hysteresis. This makes the cooling curves considerably different from the heating curves. Often, cooling the sample from the hexagonal (II) phase may cause the lipids to be kinetically trapped in an isotropic phase. In addition, the starting point for a thermal study must be carefully established. Usually this requires several freeze-thaw cycles of the lipid sample before starting the study at the temperature of interest. In systems that should be lamellar, this procedure tends to provide a reproducible starting phase structure.

Recently the difference in the axially symmetric powder patterns of PE and PC has been exploited to resolve differences in lipid composition in

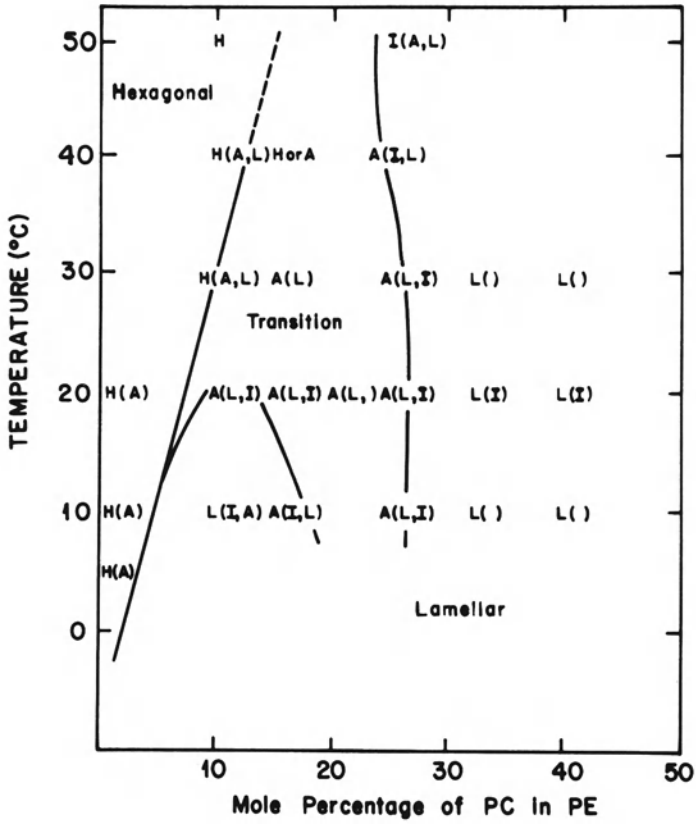


Figure 14. Phase diagram of PE/PC mixtures, determined in part by ^{31}P NMR (Hui *et al.*, 1981).

the lamellar and hexagonal phases when they coexist (Eriksson *et al.*, 1985). The lack of resolution at $\sigma_{\parallel}$ in their observed spectra should be explored further, however, with a Hahn echo.

PE mixed with phosphatidylserine will undergo a lamellar-to-hexagonal (II) phase transition under the influence of calcium. This can be modulated by calcium and cholesterol (Bally *et al.*, 1982). PE mixed with phosphatidylglycerol also is capable of forming nonlamellar structures (Farren and Cullis, 1980).

2.2.3. Phosphatidylethanolamine Derivatives

These studies show that the structure of the headgroup is the governing factor in the phase behavior of these phospholipids. PE, with three hydrogens on the amine, readily forms hexagonal (II) phases. PC, with three

methyl groups on the amine, is stable in the lamellar configuration. Some studies have therefore been performed on the effect of headgroup structure on phospholipid phase behavior.

The effect of sequential methylation of the amine was explored. As one might expect, the addition of methyl groups to the amine of PE made the behavior of the modified PE more like PC, with its three methyl groups (Gagne *et al.*, 1985).

An interesting study compared the effect of adding bulk to the PE headgroup without disturbing the hydrogen bonding capability of the amine of PE. Adding methyl or ethyl groups to the C_2 of the ethanolamine was compared to adding the same substituents to the amine. ^{31}P NMR studies showed that the L_α lamellar phase was destabilized in the former case, whereas in the latter case it was stabilized. Furthermore, the C_2 substituents promoted the formation of dehydrated gels, as indicated by the appearance of very broad ^{31}P powder patterns.

2.2.4. Other Phospholipids

Diphosphatidylglycerol (DPG) is normally lamellar in aqueous dispersion, free of divalent cations (Gerritsen *et al.*, 1980). However, the addition of calcium induces the formation of the hexagonal (II) phase (Cullis *et al.*, 1978). ^{31}P NMR powder patterns indicated that adriamycin inhibits the ability of calcium to induce the hexagonal (II) phase in DPG (Nicolay *et al.*, 1985).

The phase behavior of sphingomyelin has been studied (Cullis and Hope, 1980; Hui *et al.*, 1980). Sphingomyelin forms lamellar phases under all conditions studied. However, some isotropic resonances were detected, corresponding to the spontaneous formation of small vesicles. Care in interpretation of such isotropic components is emphasized by the misinterpretation of such components in an earlier study (Yeagle *et al.*, 1978a). Even mixtures containing sphingomyelin only as one of the components exhibited spontaneous vesicularization (Fraser *et al.*, 1986).

The effects of polyethylene glycol on phospholipids have been measured and bilayer defects noted (Boni *et al.*, 1981). Further studies of the effects of calcium on dipalmitoylphosphatidylglycerol have been observed. The influence of lysophosphatidylcholine on the phase behavior has been documented, as well as those of some curious synthetic phospholipid derivatives containing a cholesterol moety. These findings (Noggle *et al.*, 1982) were later questioned by Gruner and Jain (1985), who investigated the question. The structures of bilayers composed of lysophosphatidylcholine and fatty acids have been explored with ^{31}P NMR (van Echteld, *et al.*, 1981; Allegrini *et al.*, 1983).

Dolicyl phosphate also induces nonbilayer structures (van Duijn *et al.*, 1986).

Molecules such as gramicidin have been studied for their effects on the phase behavior of phospholipids as detected by ^{31}P NMR. At high concentrations, gramicidin will alter phospholipid phase structure (van Echteld *et al.*, 1981b; Killian *et al.*, 1983; 1986; 1987).

Proteins have been shown to affect the lipid polymorphism as detected by ^{31}P NMR. Proteins studied include glycophorin (Taraschi *et al.*, 1982a), polylysine (de Kruijff and Cullis, 1980), cytochrome oxidase (Rietveld *et al.*, 1987), rhodopsin (Albert *et al.*, 1984), and wheat germ agglutinin (Taraschi *et al.*, 1982b). Saturation transfer techniques have been used to demonstrate slow exchange between the phases (de Kruijff *et al.*, 1980).

Phosphatidic acid can be induced to form hexagonal (II) phase by myelin basic protein (Smith and Cornell, 1985). In the presence of calcium, however, phosphatidic acid exhibits interestingly different behavior (Smaal *et al.*, 1987): lamellar structures are formed. However, the calcium evidently produces a selective immobilization of the phospholipid headgroup. A ^{31}P powder pattern is observed equivalent to that arising from a dehydrated powder of phospholipid. At the same time, the hydrocarbon chains of the lipid produce ^{2}H NMR powder patterns equivalent to liquid crystalline phospholipids.

2.3. Biological Membranes

The lipids extracted from biological membranes can also exhibit nonlamellar structures. For example, PE from the human erythrocyte membrane can readily form the hexagonal (II) phase.

More interesting is the observation that total lipid extracts from some biological membranes are capable of forming nonlamellar structures. A good example is found in the rod outer segment disk membrane. The lipid extract from that membrane readily forms the hexagonal (II) phase in the presence of calcium (de Grip *et al.*, 1979; Albert *et al.*, 1984).

One group (Deese *et al.*, 1981b; Brown *et al.*, 1982) reported an inability to repeat an earlier study (de Grip *et al.*, 1979) in which the hexagonal (II) phase had been observed for disk membrane lipids. However, Brown and his colleagues did not repeat the experiments that they set out to repeat, having left out the divalent cations which were included in the previous study (de Grip *et al.*, 1979). These authors (Brown *et al.*, 1982) also suggested that the previous study (de Grip *et al.*, 1979) had worked with oxidized lipids and that those oxidized lipids promote the hexagonal (II) phase. However, oxidation inhibits the formation of the hexagonal (II) phase (Albert *et al.*, 1984). Subsequently, it was shown that the lipids form

hexagonal (II) phase in the presence of calcium and are stable in the lamellar phase in the absence of calcium (Albert *et al.*, 1984).

Interestingly, calcium is not capable of destabilizing the lamellar architecture of the intact disk membrane. Even in the presence of 50 mM calcium, the disk membrane remains lamellar. Apparently, the interactions between the disk membrane proteins, notably rhodopsin, and the disk membrane lipids stabilize the lamellar phase.

The normal erythrocyte membrane exhibits ^{31}P NMR spectra characteristic of phospholipid bilayer. However, if one isolates the phospholipids from the inner monolayer of that membrane, the phase behavior of liposomes of those lipids is dependent upon calcium, since the major lipids are phosphatidylethanolamine and phosphatidylserine (Hope and Cullis, 1979). Further addition of bilayer destabilizing agents to the erythrocyte membrane can induce the membrane to at least partially abandon its lamellar structure (Viti *et al.*, 1983).

Several studies of liver mitochondrial membranes have examined the phase behavior of the lipids using ^{31}P NMR (de Kruijff *et al.*, 1982; Cullis *et al.*, 1980). Phospholipids extracted from the mitochondrial membrane could be induced to form hexagonal (II) phase under the influence of calcium. At first, it was reported that the intact mitochondria could also be induced to form nonlamellar phases (Cullis *et al.*, 1980). However, a later study by the same investigators reported that lamellar phases were observed only if care was taken to preserve the biological activity of the membrane preparations (de Kruijff *et al.*, 1982).

Studies of liver microsomal membranes were reported to show a small isotropic component in the ^{31}P NMR spectra (Stier *et al.*, 1978; de Kruijff *et al.*, 1978). In light of the above results with mitochondria, it would appear to be important to establish the functional integrity of the membrane preparation. This was not reported. Furthermore, as mentioned previously, interpretation of ^{31}P NMR spectral shapes requires independent conformation of membrane morphology, such as through the use of freeze fracture electron microscopy. This was not reported. The presence of small vesicles in biological membrane preparations is common. Thus, the observation of an isotropic resonance in the ^{31}P NMR spectra cannot be confidently interpreted until other evidence is obtained. In the second study, narrow resonances are observed from the membrane, though not of the same shape as reported above. The same criticism of lack of description of membrane viability applies. Also, as above, no independent evidence is provided on the vesicles' sizes in the same preparations used for the NMR experiments. Attempts were made to rule out motional averaging due to small vesicles by increasing viscosity, but an increase in macroscopic solution viscosity will not reduce the motional averaging due to lateral diffusion around the surface of the vesicle. In fact, upon dehydration of the membrane, which

will greatly decrease lateral diffusion rates and consequent motional averaging, the spectrum substantially broadens and the ^{31}P NMR spectra clearly show that a majority of the phospholipids are in a standard bilayer configuration.

Two interesting studies have been reported on the behavior of membranes and phospholipids derived from *Tetrahymena* (Jarell *et al.*, 1981; Ferguson *et al.*, 1982). The interest here is that phosphonolipids, in which a carbon–phosphorus bond occurs, are common in these protozoa. As a result, the chemical shift tensor of the phosphorus is different. This results in about a 20-ppm downfield shift of the isotropic resonance. Therefore, the spectrum of membranes containing this lipid consist of two overlapping powder patterns, though both powder patterns reflect lamellar phase. This is in spite of high phosphatidylethanolamine content of the membrane and the tendency of that lipid and the phospholipid to form hexagonal (II) phase in pure form.

Studies on *E. coli* have shown that while the extracted phospholipids with their high phosphatidylethanolamine content can undergo a transition to hexagonal (II) phase, the intact membranes are predominantly in lamellar phase (Burnell *et al.*, 1980; Gally *et al.*, 1980). Finally, studies on human aorta also show evidence for multiple forms of motional averaging (Cushley *et al.*, 1979).

In summary, phospholipids can exist in a variety of phases, which can be easily monitored by ^{31}P NMR. In biological membranes, good evidence has been presented only for lamellar phase.

Detergents have considerable effects on the structure of membranes. As detergents solubilize the membrane, micelles of detergent and the membrane components are formed. This change results in a considerable alteration in membrane morphology as detected by ^{31}P NMR. Some investigators have exploited this observation to follow the effects of detergents on biological membrane structure (Madden and Cullis, 1982; Sugimoto *et al.*, 1984; Bayerl *et al.*, 1986).

Anesthetics are incorporated into membranes, and ^{31}P NMR might be useful in studying the effect of anesthetics on membrane structure. Studies of this sort have been published (Boulanger *et al.*, 1981; Hornby and Cullis, 1981).

3. ^{31}P NMR AND MEMBRANE STRUCTURE AND DYNAMICS

3.1. Phospholipid Headgroup Conformation

Extensive examination of phospholipid headgroups using ^{31}P NMR has been performed using both SUV and MLV. The studies on sonicated

vesicles will be considered first. Because of their homogeneity, the SUV make excellent subjects for measurements. When extrapolating conclusions obtained from these vesicles to biological membranes, however, the small size and consequently highly curved surface of the vesicles compared to most biological structures should be noted.

Valuable information on phospholipid headgroup conformation was obtained by exploiting the positive nuclear Overhauser effect (NOE) arising from the ^{1}H–^{31}P dipolar interactions. It was first demonstrated in 1975 that $^{31}P\{^{1}H\}$ NOE was experimentally observable in a variety of compounds (Yeagle *et al.*, 1975a). The observation of a 40% enhancement at 40 MHz for ^{31}P resonances of phospholipids in SUV (and MLV) immediately suggested several headgroup properties (Yeagle *et al.*, 1975b). One is that headgroup motion must be, at least modestly, in the fast motion regime, since at long correlation times, whether dominated by isotropic or anisotropic motions, the NOE nearly vanishes.

Information about headgroup conformation can be obtained from the NOE, since the NOE has an implicit contribution from the internuclear proton–phosphorus distance and the number of protons, as shown in equation (3). (This is because the effectiveness of dipolar interactions and thus of the cross-relaxation upon which the NOE depends is in turn dependent upon these factors, among others.)

The method employed in these studies was to determine which protons were contributing to the dipolar interactions causing the NOE (Yeagle *et al.*, 1976; 1977a). Several considerations are important in such studies. The strength and bandwidth of the ^{1}H decoupler must be tightly controlled. The position of the ^{1}H decoupler should be calibrated with a known compound. Any conclusions should be tested with compounds in which relevant positions of hydrogen are selectively deuterated. Multiple measurements at each decoupler position also appear to be necessary to accurately determine the magnitude of the NOE. The measurements cannot be done at high field because of the field dependence of the NOE referred to earlier. Finally the width of the ^{1}H resonances of the sample must be considered, because if they are broad, a resolved frequency dependence of the NOE will not be obtained. Therefore, these measurements must be done in sonicated vesicles and not in unsonicated systems.

The result obtained for egg phosphatidylcholine is seen in Figure 15. Clearly a maximum is observed in this plot at the position of the *N*-methyl proton resonance. As a crucial check of this result, phosphatidylcholine that was specifically deuterated in the *N*-methyl protons was prepared. Figure 15 also shows this result. With the *N*-methyl protons removed, the maximum shifts, as expected, to the methylene proton region for the headgroup. These data provide unequivocal evidence that the *N*-methyl protons are interacting closely with the phosphate group. Such interactions were

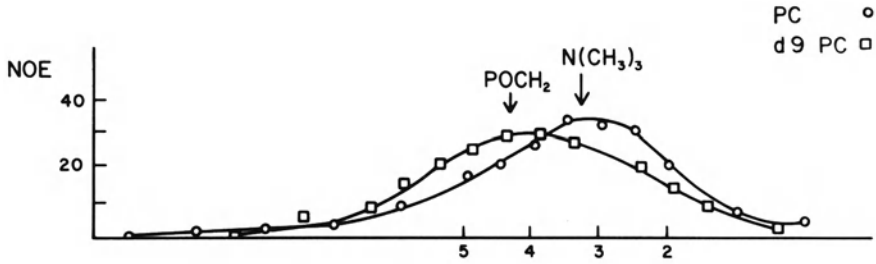


Figure 15. Dependence of the $^{31}P\{^1H\}$ NOE on the frequency of cw ^1H decoupling (Yeagle *et al.*, 1977a).

more recently confirmed for phosphatidylcholine by two other groups (Viti and Minetti, 1981; Shibata *et al.*, 1983).

In mixtures of phosphatidylcholine and noncholine-containing phospholipids, like phosphatidylethanolamine or diphosphatidylglycerol, the NOE could still be shown to arise from the *N*-methyl protons of PC. Therefore the *N*-methyl–phosphate interaction appears to be intermolecular.

These groups led to the conclusion that the phosphatidylcholine headgroup was aligned approximately parallel to the membrane surface. As reviewed elsewhere (Yeagle, 1978), this model fits all the available data and is now generally accepted.

A fourth group reported it was unable to detect these *N*-methyl–phosphate interactions (Burns *et al.*, 1983). Unfortunately, in the latter case, no calibration of the position of the proton decoupler was described, and no attempt to check results with readily synthesized specifically deuterated phospholipids was reported. These missing features to their study plus poor signal-to-noise ratios likely explain the discrepancy with the previously reviewed studies.

Sphingomyelin has been studied using the $^{31}P\{^1H\}$ NOE as described above for phosphatidylcholine. The data reported in these studies suggest that sphingomyelin has a headgroup conformation similar to that of phosphatidylcholine (Yeagle *et al.*, 1977a).

The ^{31}P NMR resonance from sonicated phosphatidylcholine vesicles consists of two overlapping isotropic resonances. Shift reagent studies demonstrate that the upfield resonance arises from phospholipids on the inside of the vesicle. The downfield, more intense resonance arises from phospholipids on the outside of the vesicle (Hutton *et al.*, 1977).

This result can be understood in terms of charge effects on the ^{31}P chemical shift. The phospholipid headgroups are more tightly packed on the inside of the sonicated vesicle than on the outside, due to the small radius of curvature of the SUV. ^{13}C NMR of the ester carbonyls of phos-

phatidylcholine vesicles reflect this packing difference (Yeagle and Martin, 1976). Binding of a cation to the phosphate causes an upfield shift (Hutton *et al.*, 1977). Since the phosphatidylcholine headgroups interact such that the positively charged *N*-methyl moiety closely interacts with the negatively charged phosphate on a neighboring lipid, a consistent interpretation is that the headgroup interactions just referred to are tighter on the inside of the vesicle than on the outside, causing the resonance from the phospholipid headgroups on the inner half of the vesicle bilayer leaflet to occur upfield of the resonance from the phospholipid headgroups in the outer half of the vesicle bilayer.

While the work just reviewed establishes the headgroup conformation for phosphatidylcholine in sonicated vesicles, it does not speak to the question of the headgroup conformation in systems with much larger radii of curvature. Much work has been done on this question using MLV. In MLV, no mode of isotropic motional averaging is adequate to produce isotropic resonances (i.e., the rotational correlation times are too long). Since the pseudo axially symmetric powder pattern described previously is observed for the ^{31}P NMR spectra of these systems, different approaches must be used to determine phospholipid headgroup conformation.

Three approaches to the question of phospholipid headgroup conformation in MLV have been explored. In one approach (Griffith *et al.*, 1978), oriented membranes were used. The lamellae of phospholipid were sandwiched between glass plates to orient the bilayers parallel to the surface of the glass plates. The material used in this study was a monohydrate of dipalmitoylphosphatidylcholine. Powder patterns were not observed from these oriented membranes. At certain orientations, the ^{31}P NMR spectra of this systems exhibit two resonances, at −80 ppm and −30 ppm. Because the value of −80 ppm corresponds to an edge of the powder pattern, it is possible to align the director of headgroup rotational diffusion with respect to the macroscopic surface of the membrane. Determination of this orientation was achieved through knowledge of the relationship of the coordinate system of the chemical shift tensor and the molecular coordinate system. The study concluded that the O(3)-P-O(4) plane is tilted at 50° with respect to the bilayer normal for many of the molecules in the bilayer.

In a second approach the chemical shift anisotropy was expressed in terms of order parameters. For dipalmitoylphosphatidylcholine the following expression has been suggested in which the principle elements of the chemical shift tensor are used:

$$\Delta\sigma = -56\, S_{11} + 135\, S_{33} \tag{5}$$

Here, S_{11} is the order parameter of the axis connecting the esterified oxygens

(of the phosphate group), and S_{33} refers to the connecting axis of the nonesterified oxygens. It was deduced that S_{33} must be negative and that the connecting axis of the nonesterified oxygen atoms must be approximately perpendicular to the membrane surface (Niederberger and Seelig, 1976).

A third approach combined order parameters for the phosphate group obtained from ^{31}P NMR measurements with order parameters for the methylene segments and *N*-methyls of the headgroup obtained from ^{2}H NMR measurements of specifically deuterated phosphatidylcholine. A fit was calculated between observed deuterium quadrupole splittings and splittings derived from molecular models. It was claimed that only a small number of conformations could fit the ^{2}H NMR data and the ^{31}P NMR data. The conformation that was selected aligned the headgroup parallel to the surface of the membrane, rapidly exchanging between two enantiomeric forms (Seelig *et al.*, 1977). This model agrees well with the model described above.

Unfortunately details of these calculations have not been provided. Therefore it is difficult to evaluate the above results in comparison with those of Skarjune and Oldfield (1979), who apparently used a similar approach. While the results of the latter investigators do not contradict the model of phospholipid headgroup conformation described above, questions are raised as to the uniqueness of the fit obtained previously. It is in fact suggested that a number of molecular models could fit the experimental data.

Perhaps the most definitive data concerning the conformation of the phosphatidylcholine headgroup was obtained using specifically deuterated phospholipids and neutron diffraction (Buldt *et al.*, 1978). Because the neutron scattering profiles of deuterated lipids are significantly different from those of native lipids, the positions of the deuterium labels can be accurately mapped. The results demonstrate that the choline methylene and methyl groups reside at approximately the same distance from the center of the bilayer. Therefore, the headgroup of the phosphatidylcholine is parallel to the surface of the membrane.

Several investigations examined the effects of changes in the structure of phosphatidylcholine on the ^{31}P NMR spectra and on the headgroup conformation of the phospholipid. One study examined the effect of changing the position of attachment of the phosphocholine moiety to the glycerol, and found little difference in headgroup behavior. The effective ^{31}P chemical shift anisotropy was similar, as were the ^{2}H quadrupole splittings for the methylene segments of the headgroup (Seelig *et al.*, 1980).

Alterations of headgroup structure can produce changes in the ^{31}P NMR powder pattern that is observed. The effects of substitution of methanol, ethanol, propanol, and butanol for choline were examined (Browning, 1981a, b). A decrease in the width of the ^{31}P NMR powder pattern was correlated with an increase in headgroup size. The authors

interpreted the reduction in terms of a change in the time-averaged headgroup conformation, rather than a change in the order parameter of the phosphate of the phospholipids. It should be noted that order parameters contain assumptions concerning conformation with respect to axes of rotational diffusion. A change in conformation will of itself, cause a change in the powder pattern observed. With bulkier headgroups, the headgroup is forced to assume a more vertical orientation with respect to the membrane surface because it cannot pack properly in the plane of the surface. A total change of about a factor of 2 was noted in $\sigma_\perp - \sigma_\parallel$ over the range of headgroups studied.

Analogous results were obtained when the distance between the phosphate and the *N*-methyls was increased (Siminovitch *et al.*, 1983). Increasing the size of the headgroup led to changes in the ^{31}P NMR spectra consistent with a headgroup conformational change. Accompanying these data were ^{14}N NMR spectra which supported the results obtained from ^{31}P NMR. In a companion study Browning (1981a, b) showed that introducing hydroxyls into the polar headgroup changed headgroup properties, as evidenced by a change in $\sigma_\perp - \sigma_\parallel$. Detergents have also been shown to change phospholipid headgroup conformation (Beyer, 1986).

Other common phospholipid classes have been examined by ^{31}P NMR. For Seelig and coworkers, the ^{31}P NMR data, in combination with the ^{2}H NMR data, suggested a headgroup conformation for phosphatidylethanolamine similar to the headgroup conformation derived for phosphatidylcholine (Seelig and Gally, 1976). This conclusion was deduced using the same calculation method described above for phosphatidylcholine, which has been questioned by Oldfield and coworkers (Skarjune and Oldfield, 1979). However, the conformation suggested agrees with the conformation seen in the crystal structure of dilaurylphosphatidylethanolamine (Hitchcock *et al.*, 1974). Phosphatidylglycerol has also been studied by the same methods, and again the same headgroup conformation is detected (Wolgemuth *et al.*, 1980).

Bilayers of phosphatidylserine (PS) reveal significantly different ^{31}P NMR powder patterns compared to PC, PE, and PG. The width of the ^{31}P powder pattern ($\sigma_\perp - \sigma_\parallel$) is greater for PS (and PI) than measured for PC, PE, and PG. The deuterium quadrupole splittings were also larger for PS than for PC, PE, and PG (Browning and Seelig, 1980). In addition, at one of the methylenes of the headgroup, the two deuteriums were found to be motionally inequivalent. These data suggested that the headgroup of phosphatidylserine is considerably more rigid than for the other phospholipids. However, one must be cautious in interpreting data involving phosphatidylserine, both because of the ease of oxidation of this phospholipid and because of the great affinity of the headgroup of phosphatidylserine for cations which, upon binding, are known to render the whole

membrane more rigid. Such an effect was noted earlier in studies of calcium binding to phosphatidylserine (Kohler and Klein, 1977).

Since biological membranes consist of complex mixtures of phospholipids it is important to examine the behavior of phospholipids in the presence of other phospholipid classes. The $^{31}P\{^{1}H\}$ NOE work cited above indicated that the ^{31}P NMR resonance of phosphatidylethanolamine in sonicated vesicles with phosphatidylcholine was influenced by the *N*-methyl protons of phosphatidylcholine. These results suggested that the two phospholipids were not phase-separated in the plane of the membrane. This work was performed with natural phospholipids containing a mixture of hydrocarbon chains with considerable unsaturation. Other workers have suggested that a phase separation occurs when the system is fully saturated (Arnold *et al.*, 1981). However, the conclusion is based on data from two overlapping powder patterns for the two phospholipids. Because of the manner in which the spectra were taken and the signal-to-noise ratio, taken together with the fact that the two phospholipid powder patterns are nearly superimposable, would suggest that the phase diagram must have been very difficult to construct.

Watts and coworkers have examined the effect on headgroup behavior of mixing two different phospholipids, phosphatidylcholine and phosphatidylglycerol, both in fully saturated form (Sixl and Watta, 1982). ^{31}P NMR spectra revealed no significant differences due to mixing, and the results were interpreted in terms of random mixing of the two phospholipids. However, ^{2}H NMR studies of specifically deuterated phospholipids revealed a number of effects on one phospholipid due to the presence of the other, probably reflecting some change in orientation of the headgroups. However, why the ^{31}P NMR spectra do not exhibit a sensitivity to such orientational changes is mysterious, especially since in studies of structural changes the ^{31}P NMR spectra did reflect the presumed orientational change. Again the importance of using more than one technique in examining a system is evident.

In a novel study, mixtures of dipalmitoylphosphatidylcholine and bovine brain phosphatidylserine were studied by ^{31}P NMR (Hui *et al.*, 1983). The phosphatidylcholine molecule was labelled by substitution of a sulfur for one of the non-esterified oxygens of the phosphate. This modification causes a large change in the chemical shift tensor, such that the powder pattern of the sulfur labelled phospholipid is shifted about 50 ppm downfield from the corresponding powder pattern of the unlabelled phospholipids. Since $(\sigma_{\perp} - \sigma_{\parallel})$ is less than or equal to 50 ppm in the liquid crystalline state, spectra containing both phospholipids consist of two pseudo-axially symmetric powder patterns that experience very little, if any, overlap. Therefore it is possible to study the behavior of the two phospholipids simultaneously, yet independently, in the same vesicle. The results of this

study indicated that in the presence of calcium, the phosphatidylserine phase separated.

The effects of cholesterol on the headgroup of phosphatidylcholine have been studied using both SUV and MLV. Cholesterol has been shown, using the NOE approach, to act as a spacer between the phospholipids, breaking up the intermolecular headgroup interactions by increasing the average separation of the phospholipids (Yeagle *et al.*, 1977a). This was manifest in a change in the source of the NOE from the *N*-methyl protons to the methylene protons. The analogous effect is unsonicated systems has been observed (Brown and Seelig, 1978).

Cations can bind to phospholipid headgroups, and there is some suggestion that such binding may alter the conformation of the phospholipid headgroup. ^{31}P NMR spectra and ^{2}H NMR spectra of specifically deuterated phospholipids were examined in the presence of lanthanides (Brown and Seelig, 1977; Akutsu and Seelig, 1981). While the ^{2}H NMR spectra showed considerable effect of the metal ion binding, the ^{31}P NMR powder patterns were not sensitive to the binding of lanthanide ion. This is somewhat surprising, in light of the substantial change in isotropic chemical shift induced by lanthanide ions. Perhaps the different packing for the polar headgroups in small vesicles enhances the interaction between the lanthanide ion and the phosphate. It is also not clear what kind of membraneous structure is formed by hydrating phospholipids with buffers containing molar concentrations of trivalent cations.

Divalent cations have a pronounced effect on negatively charged phospholipids (Cullis and de Kruijff, 1976; Lichtenberg *et al.*, 1979; Hope and Cullis, 1980). Substantial increases in the width of the ^{31}P NMR powder pattern are well documented for calcium binding to phosphatidylserine. The effect is apparently to dehydrate the surface of that membrane, thereby producing a spectrum similar to those of dehydrated phospholipids.

Using the NOE approach, no effect of a transmembrane potential was observed for phosphatidylcholine in sonicated vesicles (Yeagle, 1979). Decreasing the radius of curvature has the effect of decreasing the packing efficiency of the headgroups (Castellino and Violand, 1979; Yeagle, 1979). The headgroups can no longer interact is an intermolecular fashion as described above. This is manifest in a loss of the contribution to the NOE from the *N*-methyls, in favor of contributions from the methylenes when phospholipids are in micelles.

3.2. Phospholipid Headgroup Dynamics

The dynamics of the phospholipid headgroup have already been touched upon and the results from ^{31}P NMR studies will be summarized here. The correlation time for headgroup rotational diffusion is about 1 nsec.

This has been derived from two independent measurements. One is the $^{31}P\{^{1}H\}$ NOE measurement (Yeagle *et al.*, 1975b). It was previously described that the NOE is motionally limited. Thus the magnitude of the NOE is reflective of the effective correlation time. Theoretical modelling of the NOE, based on anisotropic motion, provides the basis for determining the effective correlation time.

The other determination of the correlation time was obtained from a temperature dependence of the ^{31}P T_1. A minimum in the plot of T_1 vs. temperature is observed. This result unequivocally shows that a correlation time in the neighborhood of 1 nsec is the dominant correlation time for spin lattice relaxation in the system (Tamm and Seelig, 1983).

3.3. Phospholipid–Protein Interactions

3.3.1. Introduction

Much of what has been discussed up to this point concerned phospholipids in phospholipid bilayers in the absence of the other major component of biological membranes, membrane proteins. Inclusion of membrane proteins in the lipid bilayer must necessarily create a different environment for the phospholipids. In phospholipid bilayers, phospholipids have other phospholipids as their nearest neighbors. In membranes containing membrane proteins, some of the phospholipids must have proteins as nearest neighbors. Therefore, lipid-protein interactions in membranes have been the subject of intense interest for a number of years. Interactions between these two major membrane components can be expected to be important to membrane function. For example, lipids are often required to activate membrane enzymes (Warren *et al.*, 1974; Mitchell *et al.*, 1983; Cheng *et al.*, 1986). Proteins in turn affect the behavior of membrane lipids, as evidenced by calorimetry data (van Zoelen *et al.*, 1978a), from electron spin resonance data (Griffith and Jost, 1976), and in the morphological behavior of membrane lipids (Taraschi *et al.*, 1982a; Albert *et al.*, 1984; Rietveld *et al.*, 1987). Before an adequate understanding of the potential roles of such interactions in cell membrane function can be obtained, a clear picture of the possible modes of interaction between lipids and proteins must be drawn.

Membrane proteins do not resemble phospholipids. They are generally very much larger than phospholipids. Membrane proteins are constructed of different constituents than are phospholipids. Thus, when phospholipids diffusing in the plane of the membrane encounter a membrane protein, they must experience a considerably different environment than when next to other phospholipids.

Because of the large size difference, the membrane protein will influence the motional properties of the phospholipid. Lateral diffusion of the phos-

pholipid will be limited by the membrane protein, in that the phospholipid cannot occupy sites within the volume of the membrane protein. Furthermore, limitations on the escape angle from the protein surface will limit lateral diffusion of a phospholipid occupying a site immediately adjacent to the protein. These limitations increase the residence time of phospholipids in sites next to the membrane protein, even in the absence of favorable interactions between these two membrane components.

The protein surface is constructed of amino acids. In the interior of the membrane, the protein generally exposes hydrophobic amino acid side chains to the hydrophobic interior of the membrane. Thus the hydrophobic hydrocarbon chains of the lipids encounter the hydrophobic residues of the protein. One can expect only weak dispersive forces to characterize the interaction between lipids and proteins in the hydrocarbon region of the membrane.

At the interface between hydrophobic domains and hydrophilic domains in integral membrane proteins, sequence analysis generally shows a concentration of charged amino acid residues. This is the region of the protein in contact with the phospholipid headgroups of the membrane. Thus, with the phospholipid headgroups more interesting and complex lipid–protein interactions are likely, in contrast with the hydrocarbon chains. ^{31}P NMR of phospholipids in membranes containing proteins therefore provides an important view of lipid–protein interactions by reporting on the effects of membrane proteins on bilayer surface properties.

3.3.2. Interfacial Lipids

The term "interfacial lipids" will be used to designate those phospholipids in sites immediately adjacent to a membrane protein. Those sites are characterized by residence times long relative to lateral diffusion. However, phospholipids in some of these interfacial sites have residence times sufficiently long ($\tau_c > 10^{-4}$ sec) that distinguishable spectral components are observed in the NMR spectra. These sites with relatively long residence times will be characterized by the term "bound" phospholipids. "Bound" phospholipids thus constitute a subclass of interfacial lipids in this definition.

3.3.2.1. "Bound" Phospholipids. ^{31}P NMR studies of glycophorin in lipid bilayers have provided considerable information concerning the properties of bound phospholipids. Human erythrocyte glycophorin as normally isolated has about four phospholipids tightly bound to it. The bound phospholipids most likely are phosphatidylinositol (PI), as indicated by ^{31}P NMR analysis of solubilized glycophorin (Armitage *et al.*, 1977).

Glycophorin with PI bound to it was reconstituted into a glycolipid bilayer. The ^{31}P NMR spectrum thus only contains resonances from the

phospholipids originally bound to glycophorin. ^{31}P NMR is blind to the bulk of the lipid making up the normal phospholipid bilayer. Figure 16 shows the result obtained. An axially symmetric powder pattern is observed. This indicates axial diffusion of the bound phospholipids (Yeagle and Kelsey, 1989).

Two features distinguish this powder pattern from the powder pattern normally observed from phospholipid bilayers. One feature is the width of the powder pattern: $\sigma_\perp - \sigma_\parallel$ is about 100 ppm. For phosphatidylcholine bilayers, $\sigma_\perp - \sigma_\parallel$ is normally about 45 ppm (Yeagle, 1982), while for PI bilayers it is about 57 ppm. The breadth of this powder pattern indicates that the headgroup of the bound (to glycophorin) phospholipid is motionally restricted relative to normal phospholipid bilayers (Campbell *et al.*, 1979).

The second feature distinguishing the powder pattern of phospholipids bound to glycophorin in a lipid bilayer from a normal phospholipid bilayer powder pattern is the inferred location of the isotropic chemical shift. This is at least 10 ppm downfield from the normal isotropic chemical shift.

Therefore, PI bound to glycophorin in a lipid bilayer is motionally restricted, although undergoing axial diffusion (perhaps coincident with axial diffusion of the protein). The immediate chemical environment of the phosphate is different than in phospholipid bilayers, perhaps due to an interaction between the phospholipid phosphate and amino acid residues on the protein.

^{31}P NMR spectra of phospholipids interacting with the proteins of the cytochrome oxidase complex have also been obtained (Seelig and Seelig, 1985). As in the case of glycophorin, the ^{31}P powder pattern is much broader

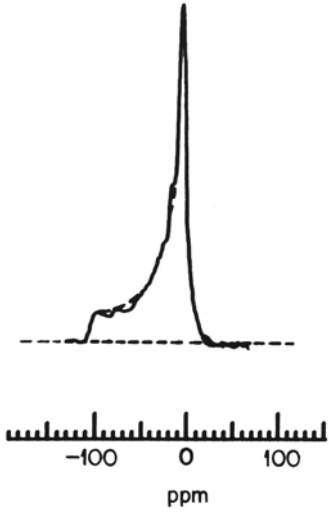


Figure 16. ^{31}P NMR spectra of human erythrocyte glycophorin, with three molecules of PI tightly bound, reconstituted into bilayers of digalactosyldiglyceride. The dashed line is the spectral simulation.

than observed from pure phospholipid bilayers. The distortions to the axially symmetric shape may arise from the low lipid–protein ratio of this system. All the phospholipid is likely in contact with the protein because of the high protein content of the membrane. Similar results with two spectral components were observed in reconstituted systems, although not adequately analyzed (Rajan *et al.*, 1981). When the powder patterns present are axially symmetric, the spectral simulations are particularly simple. The equations of Seelig (1978) can be employed.

$$S(v) = \int_{-\infty}^{\infty} I(v - v^*)\rho(v^*)\, dv^* \tag{5}$$

where $I(v - v^*)$ is given by:

$$I(v - v^*) = \frac{1}{\pi\Delta} \frac{1}{1 + (v - v^*)^2/\Delta} \tag{6}$$

and $\rho(v)$ is given by:

$$\rho(v) \propto \left[\frac{1}{3} - \frac{1}{3}\frac{v - v_i}{v_\perp - v_i}\right]^{-1/2} \tag{7}$$

It should be noted that the spectra reported in Figures 16–17 give clear indications of the shapes of both spectral components: the normal bilayer powder pattern and the motionally restricted powder pattern. Spectral simulations of the ^{31}P NMR spectra show the observed spectra can be simulated by the addition of two axially symmetric powder patterns, one about twice as broad as the other, with the broad powder pattern representing the protein-induced motionally restricted phospholipid domain.

Previous studies (Romans *et al.*, 1979; Yeagle and Romans, 1981; Albert and Yeagle, 1983; Selinsky and Yeagle, 1984) did not provide an adequate representation of the ^{31}P powder patterns to permit this kind of analysis, due primarily to low signal-to-noise ratio and to the small intensity and the breadth of the broad powder pattern. The technical requirements to capture the FID of the broad component are severe, and the broad component is easily lost. A recent report on sarcoplasmic reticulum (Ellena *et al.*, 1986), presents evidence for a broad component corresponding to as much as 10 percent of the total observed resonance intensity in the ^{31}P NMR spectra from sarcoplasmic reticulum. Unfortunately, the first part of the FID was removed before the Fourier transform was calculated. Because the broad powder pattern is most heavily represented in the first few points of the

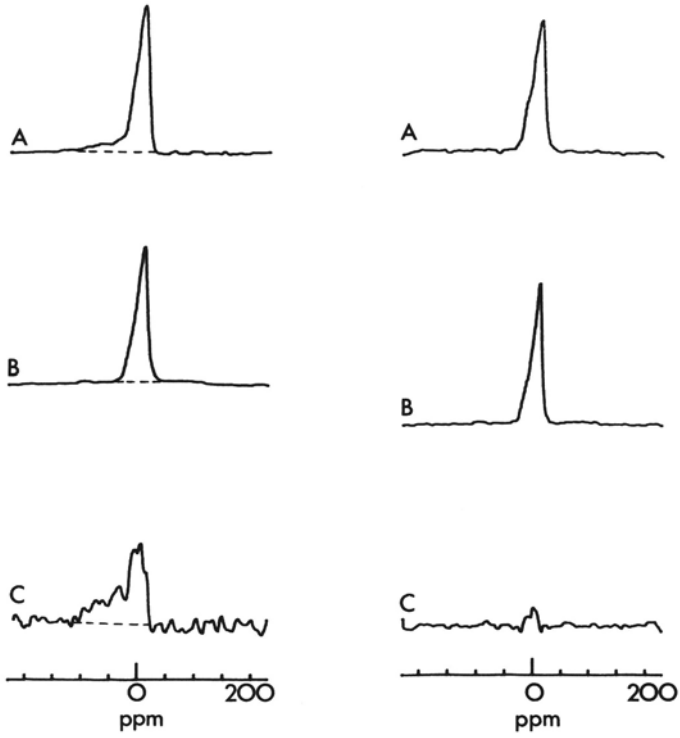


Figure 17. ^{31}P NMR spectra of light sarcoplasmic reticulum. Left panel in the presence of 110 mM K^+; right panel in the presence of 10 mM K^+. A is the spectrum of the membrane, B is the lipid extract from the membrane and C is the difference spectrum, A − B. In the presence of 110 mM K^+, C reveals a broad, axially symmetric powder pattern, very much like the spectrum in Figure 16. These results show the shape for the broad, protein-induced ^{31}P NMR spectral component in an active biological membrane.

FID, removal of any portion thereof will reduce the contribution of the broad component to the observed spectra.

The authors also neglected to include potassium in the sarcoplasmic reticulum preparation for the NMR measurements, which was included in Selinsky and Yeagle (1984), the work Ellena *et al.* (1986) claimed to repeat exactly. Removal of potassium makes the broad component difficult to observe (Selinsky and Yeagle, 1983; Selinsky, 1984). The spectra presented by these investigators for sarcoplasmic reticulum (Ellena *et al.*, 1986) resemble closely the spectra previously reported for sarcoplasmic reticulum in the absence of potassium (Selinsky and Yeagle, 1983; Selinsky, 1984). This issue is clarified in Figure 17. In the presence of potassium, a broad axially symmetric powder pattern is observed after subtracting the normal lipid bilayer powder pattern from the spectrum of the whole membrane.

This is the same kind of powder pattern seen in Figure 16. Figure 17 also shows that, in the absence of potassium, little evidence for a broad spectral component can be found. It is unfortunate that, as in their earlier work on lipid morphology (Section 2.3), Brown and coworkers pay inadequate attention to important buffer conditions when attempting to repeat exactly the experiments of other workers.

Seelig *et al.* (1981) also report a ^{31}P NMR spectrum for sarcoplasmic reticulum. However, the first portion of the FID is not included in the Fourier transform and the broad component is largely lost in a wavy baseline. McLaughlin *et al.* (1981) have even a more serious problem in that the experimental uncertainty in their total spectral intensity determination is larger than the proportion of the intensity in the broad component.

Using a different approach, Beyer and Klingenberg (1985) recently reported ^{31}P NMR data suggesting specificity of interaction of phospholipids with a different membrane protein, using a different technique. Differential broadening of the ^{31}P NMR resonance of cardiolipin in mixed phospholipid-protein-detergent micelles suggested preferential interactions between cardiolopin and the ATP-ADP exchange protein of mitochondria. Rotational diffusion of the micelles caused sufficient motional averaging of the chemical shift tensors to permit spectral resolution of the resonances arising from the individual phospholipid classes.

Returning to the results with glycophorin, the data indicate some specificity with respect to the phospholipid-protein interaction. The broad ^{31}P powder pattern and the absence of any normal bilayer powder pattern indicate that there is little exchange of glycolipid for phospholipid during the course of the experiment. If exchange had taken place, some phospholipid would give rise to a normal bilayer powder pattern. The exchange rate of phospholipids between sites with long residence times induced by membrane proteins and sites in the normal phospholipid bilayer has been measured with ^{31}P NMR. The observation of distinguishable ^{31}P powder patterns mandates an exchange rate slow relative to $10^4\ \mathrm{sec}^{-1}$. Using the technique of magnetization transfer, the exchange rate was explicitly measured in the case of membranes containing the Ca ATPase. A rate of $1\ \mathrm{sec}^{-1}$ was measured (Selinsky and Yeagle, 1985). In the case of PI bound to glycophorin, the exchange rate is much slower.

One other interesting system deserves special mention. ^{31}P NMR studies of Sendai virus envelope membrane show a remarkable effect of the viral coat protein. Virtually none of the phospholipids in the envelope is in a normal bilayer configuration, as sensed by ^{31}P NMR. Rather, all the phospholipids give rise to a very broad powder pattern, characteristic of dehydrated states. Tryptic cleavage of the F protein in the envelope frees some of the envelope membrane phospholipid to form a normal ^{31}P bilayer powder pattern (Abidi and Yeagle, 1984).

^{31}P T_1 measurements on vesicular stomatitis virus (VSV) show a pronounced effect of protein (Moore *et al.*, 1977). Temperature dependence studies of ^{31}P T_1 in other systems have shown that the shortest ^{31}P T_1 to be expected due solely to changing correlation time is about 1 sec or a little less (Tamm and Seelig, 1983). Therefore it was something of a surprise to observe a T_1 of about 0.1 sec for the envelope membrane of VSV. This ^{31}P T_1 was increased nearly an order of magnitude if the viral envelope glycoprotein G was cleaved by a protease or denatured by an increase in temperature.

A pronounced effect is observed in ^{31}P NMR of membranes when the protein content is so high (usually in reconstituted systems) that all the phospholipids are in contact with protein. When ^{31}P T_1 is measured as a function of lipid-protein ratio in a variety of reconstituted systems, the T_1 shows a sharp change when the lipid-protein ratio drops below 50:1 (mole:mole), as shown in Figure 18. This has been ascribed to a trapping of phospholipids between proteins at high lipid content (Yeagle *et al.*, 1984).

3.3.2.2. Interfacial Lipids. Interfacial phospholipids, other than bound phospholipids (as defined above), are not distinguishable from supraanular phospholipids with ^{31}P NMR because of rapid exchange

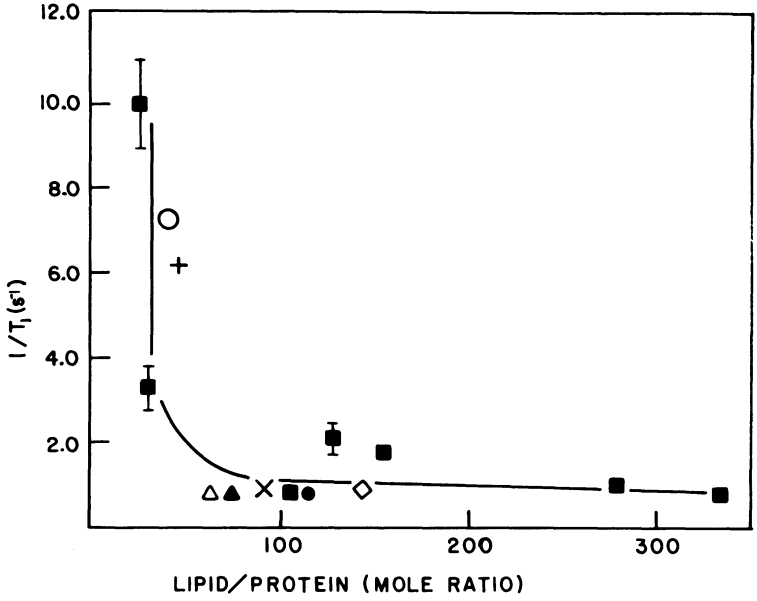


Figure 18. ^{31}P T_1 as a function of lipid-protein ratio for membranes containing protein. Included are data for glycophorin reconstituted membranes, rhodopsin reconstituted membranes, rod outer segment disk membranes, sarcoplasmic reticulum membranes, Ca ATPase reconstituted membranes, human erythrocyte ghosts, cytochrome oxidase reconstituted membranes, and pure phospholipids (Yeagle *et al.*, 1984).

between sites. The rapid exchange causes an averaging of the powder pattern observed. Only techniques with a short timescale of measurement are capable of detecting nonbound interfacial lipids. ESR of lipid spin labels is a good example (Griffith and Jost, 1976) because the threshold of the ESR experiment is about 10^{-8} sec, compared to 10^{-4} sec for ^{31}P NMR.

In contrast to the examples described above, the human erythrocyte membrane apparently does not have a spectrally distinct motionally restricted domain. All, or nearly all, the phospholipid in the membrane contributes to the normal phospholipid bilayer ^{31}P NMR resonance (Cullis and Grathwohl, 1977; Yeagle, 1984b). This may result because of the interference with or replacement by cholesterol of the motionally restricted phospholipids. This effect of cholesterol has been noted for glycophorin, one of the two major membrane proteins of the human erythrocyte membrane (Yeagle, 1984a). Alternatively, this result may be the effect of more rapid exchange between sites next to the protein and sites remote from the protein surface. Another system that does not show evidence from ^{31}P NMR data for more than one phospholipid domain is the lipovitellin phosvitin complex (Banaszak and Seelig, 1982).

3.3.3. Supraannular Phospholipids

Supraannular phospholipids are largely represented by the bilayer powder pattern observed for the majority of the phospholipids in biological membranes. This axially symmetric powder pattern is typical of that observed for phospholipids in a bilayer. However, some differences are observed in the presence of membrane proteins, compared to the bilayer resonance arising from pure phospholipid bilayers. The major difference detected is in the T_2. The T_2 is significantly shorter in the presence of membrane proteins than in their absence (Rajan *et al.*, 1981; Yeagle *et al.*, 1984). T_2 in pure phospholipid bilayers can be 10–20 msec, while in protein-containing membranes T_2 can be an order of magnitude shorter. This change in T_2 has an effect on the ^{31}P powder pattern, broadening portions of it so that the upfield maximum is less pronounced compared to the downfield shoulder. These effects are particularly apparent in the anisotropy of the T_2, which is highly anisotropic across the powder pattern (Rajan *et al.*, 1981). The dramatic reduction in ^{31}P T_2 in the presence of membrane proteins is, at low to moderate levels of membrane protein, not accompanied by a change in T_1. This indicates that headgroup rotational diffusion (the motion with the short correlation time) in the supraannular phospholipid is not significantly altered by the presence of membrane protein. Rather, an increase in the intensity of low frequency (relatively long correlation times) motions is indicated. One source of such low frequency motion is

protein rotational diffusion. Fluctuations in the lipid bilayer due to protein rotational diffusion would affect T_2 but not T_1, since these fluctuations have correlation times in the range of 10^{-5} to 10^{-4} sec.

3.3.4. Synthesis of All Magnetic Resonance Data

A variety of NMR studies have been published describing the presence of at least two spectral components in the presence of membrane proteins. In membranes containing the Ca ATPase, a motionally restricted component has been detected using ^{13}C NMR (Robinson *et al.*, 1972; Stoffel *et al.*, 1977) and ^{31}P NMR (Albert *et al.*, 1981; Selinsky and Yeagle, 1984; 1985). A motionally restricted phospholipid domain has been observed in membranes containing rhodopsin, using Hahn echos (Albert *et al.*, 1982; Albert and Yeagle, 1983) and crosspolarization ^{31}P NMR (Albert *et al.*, 1985). Motionally restricted phospholipids have been detected in the presence of human erythrocyte glycophorin using ^{31}P NMR (Armitage *et al.*, 1977; van Zoelen *et al.*, 1978; Romans *et al.*, 1979; Yeagle and Romans, 1981; Yeagle, 1984), ^{13}C NMR (Utsumi *et al.*, 1980) and ^{1}H NMR (Lau and Cowburn, 1981). Cytochrome oxidase has also been observed to create a motionally restricted phospholipid domain using ^{31}P NMR (Rajan *et al.*, 1981; Seelig and Seelig, 1985) and ^{19}F NMR (Longmuir *et al.*, 1977). The preferential motional restriction of cardiolipin has been detected in the ADP–ATP exchange protein using ^{31}P NMR (Beyer and Klingenberg, 1985). A motionally restricted phospholipid component has been observed in low density serum lipoproteins using ^{1}H NMR (Finer *et al.*, 1975), ^{31}P NMR (Yeagle *et al.*, 1977b; Yeagle *et al.*, 1978b) and ^{13}C NMR (Lund-Katz and Phillips, 1986). Protein-induced motional restriction of phospholipids in Sendai virus envelope membrane has been detected using ^{31}P NMR (Abidi and Yeagle, 1984). ^{31}P NMR T_1 has also shown strong effects of protein on phospholipid behavior in vesicular stomatitis virus (Moore *et al.*, 1977). There is, therefore, a concensus of NMR data that some membrane proteins induce the formation of a motionally restricted phospholipid domain that is in slow exchange, on the NMR timescale, with the supraannular phospholipid.

Furthermore, the relative intensity of the broad powder pattern is directly proportional to the amount of protein in the membrane, expressed as the protein–phospholipid mole ratio. This proportionality has been documented for cytochrome oxidase (Rajan *et al.*, 1981), Ca ATPase (Selinsky and Yeagle, 1984), rhodopsin (Albert and Yeagle, 1983), and glycophorin (van Zoelen *et al.*, 1978a; Utsumi *et al.*, 1980; Yeagle and Romans, 1981).

The population of these spectrally distinct motionally restricted domains detected with NMR is less than or equal to the number of phos-

pholipids required to form a complete annulus around the membrane protein. This result is in contrast to the results with ESR spin labels, where the argument has been well supported that the number of lipids in the motionally restricted domain as detected by ESR corresponds well with one complete annulus around the protein.

For example, consider the case of rhodopsin. NMR and ESR experiments quantitatively agree on 20–24 phospholipids in the motionally restricted domain, which agrees well with calculations of the number of phospholipids required for a complete annulus around rhodopsin (Albert and Yeagle, 1983; Ryba *et al.*, 1987). In contrast, consider Ca ATPase. ESR experiments indicate about 25 lipids in the motionally restricted domain (Silvius *et al.*, 1984), while NMR experiments show 11–15 phospholipids in the motionally restricted domain (see Table 4). This is understood simply on the basis of exchange rates. ESR is capable of detecting more motionally restricted lipid because lipids exchanging more rapidly than $10^4\,sec^{-1}$ are counted in the ESR experiment and not counted in the NMR experiment.

Most of the NMR experiments noted above involve measurements of the phospholipid headgroup, while ESR generally involves the use of spin labels on the hydrocarbon chains of the lipids. Thus it is important to keep track of which portion of the lipid is being measured. For example, it is possible that the lipid headgroup is bound to a membrane protein by interactions with the charged residues on the surface of the protein. At the same time, there is no attractive force of any preferential significance between the lipid hydrocarbon chains and the protein, compared to interactions with other lipid hydrocarbon chains. Thus the hydrocarbon chains could be moving on and off the surface of the protein, while the headgroup remains anchored. An extreme case was reported using covalently bound spin labels. The head was covalently anchored to the protein, while the tails were sensitive to two environments, one motionally restricted and the other like a normal bilayer (Favre *et al.*, 1979). This result was echoed in experiments showing that calcium immobilized the headgroup of PA but

TABLE 4
Quantitative Evaluation of Motionally Restricted Phospholipid Component Induced by Ca ATPase and Rhodopsin

System	Mole ratio	Total lipids motionally restricted (%)	Lipids per protein
Ca ATPase	(79:1)	19	15
	(101:1)	11	11
Rhodopsin	(40:1)	50	20
	(75:1)	25	19

did not immobilize the hydrocarbon chains of the same phospholipid (Smaal *et al.*, 1987).

Some of the NMR experiments reviewed above employed labels in the hydrocarbon chains (Longmuir *et al.*, 1977; Utsumi *et al.*, 1980). The lineshapes of the motionally restricted component were not well defined in these studies. Furthermore, the effects of heterogeneity in membrane vesicle size on the lineshapes were not adequately addressed. Therefore it is difficult to use these data to compare with the results of ESR measurements of hydrocarbon chains.

^{2}H NMR experiments on deuterated hydrocarbon chains of phospholipids have also been performed in the presence of membrane proteins. In contrast to all the other NMR studies, only one component is seen in the ^{2}H NMR spectra (Deese *et al.*, 1981a). This is understood (Albert *et al.*, 1985) in terms of protein motion in the membrane. Correlation times for protein axial diffusion are of the order of 10^{-4} to 10^{-5} sec. This has been shown to cause artificial attenuation of spectral intensity for ^{2}H labels experiencing this characteristic motion (Spiess and Sillescu, 1981). Protein rotation causes similar problems in the observed resonance intensity in other NMR experiments involving membrane proteins (Smith *et al.*, 1987). This is caused by a fast-decaying echo in the solid echo experiment used to accumulate the ^{2}H NMR data, which results in a loss of spectral intensity from any deuterated lipids interacting with the membrane protein. An experiment has been reported in which ^{31}P NMR and ^{2}H NMR were performed on the same rhodopsin-containing membrane. The ^{31}P NMR spectra showed two overlapping powder patterns, and the ^{2}H NMR spectra showed one powder pattern, characteristic of lipid bilayers (Albert *et al.*, 1985). Therefore ^{2}H NMR is not the method of choice for studying long-lived lipid–protein associations.

3.4. Transmembrane Phospholipid Distribution

Because small phospholipid vesicles produce high resolution ^{31}P NMR spectra, separable resonances are observed for many of the phospholipid headgroup classes. These vesicles are largely impermeable to cations. Therefore, the water-soluble shift reagent species of the lanthanide series can be selectively placed on one side of the vesicle membrane or the other. By so doing, it is possible to alter the chemical shift of the resonance from the phospholipids that inhabit the side of the membrane to which the lanthanide has access. This technique allows the separation of resonances due to phospholipids on the outside of the vesicle from the resonances due to phospholipids on the inside of the vesicle (Bystrov *et al.*, 1972). By this method, the distribution of phospholipid classes across the membrane bilayer can be determined. A number of measurements have been reported.

In many cases the distribution of phospholipids is not uniform (Michaelson *et al.*, 1973; Berden *et al.*, 1975; Chrzesczyk *et al.*, 1977; de Kruijff *et al.*, 1975; Castellino, 1978). Cation binding to the membrane surface can be studied by similar methods (Michaelson *et al.*, 1974; Grasdalen *et al.*, 1977; Westman and Goran-Eriksson, 1979).

One must be careful not to overinterpret such data. One study suggested that inside–outside ratios for vesicles could be used as a measure of vesicle size (de Kruijff *et al.*, 1976). However, an essential but undetermined variable, that is, bilayer thickness, plagues that analysis. Vesicle size cannot be determined from the outside–inside ratios alone.

4. NEW AREAS FOR INVESTIGATION WITH ^{31}P NMR

A very interesting study on magnetically oriented rod outer segments was recently reported (Mollevanger *et al.*, 1986). Orientation of a macroscopically ordered system should produce Lorentzian lineshapes centered at the appropriate position in the axially symmetric powder pattern of the membrane. The authors present spectra showing evidence of such an orientation effect. This potentially could reveal much interesting information. The data require further analysis and correlation, however. Assignment of resonances to morphologically distinct portions of the rod outer segment membranes remains largely hypothetical. Assignment of and $\sigma_{\parallel}$ require a closer look. For example, $(\sigma_{\perp} - \sigma_{\parallel}) = 27$ ppm should correspond to $\sigma_{\perp}$ of 13 rather than 11, for a $(\sigma_{\perp} - \sigma_{\parallel})$ of 40 ppm rather than 38 ppm. This study is one of the more interesting new contributions to the field and should lead to yet more new information about membrane structure.

Other studies show the value of using NMR to investigate the effects of other molecules on membrane structure (Bammel *et al.*, 1987; Beyer, 1986; Capuozzo and Bolard, 1985; Dea and Chan, 1986; Jones and Hunt, 1985; Roux and Champeil, 1984; Stidham *et al.*, 1985).

A new series of studies have introduced the use of ^{31}P NMR of hexane phosphonic acid diethyl ester to study phospholipid behavior (Bayerl *et al.*, 1985; 1986). It has not yet been clearly described, however, just what these measurements are reflecting on a molecular scale, since the location of the probing molecule is not well described.

Recently, ^{31}P NMR spectra of membranes have been obtained from intact cells and tissue. These include whole red cells and tissue from muscle biopsies. The phosphate metabolite spectra can be readily removed by spectral subtraction to reveal the membrane powder pattern underneath. The spectra reveal the same axially symmetric powder pattern obtained from phospholipid bilayers and biological membrane vesicle preparations.

REFERENCES

Abidi, T. F., and Yeagle, P. L., 1984, *Biochim. Biophys. Acta* **775**:419.

Akutsu, H., and Seelig, J., 1981, *Biochemistry* **20**:7366.

Albert, A. D., and Yeagle, P. L., 1983, *Proc. Natl. Acad. Sci. USA* **80**:7188.

Albert, A. D., Lund, M., and Yeagle, P. L., 1981, *Biophys. J.* **36**:393.

Albert, A. D., Yeagle, P. L., and Litman, B. J., 1982, *Biophys. J.* **37**:34.

Albert, A. D., Sen, A., and Yeagle, P. L., 1984, *Biochim. Biophys. Acta* **771**:28.

Albert, A. D., Lane, S. A., and Yeagle, P. L., 1985, *J. Memb. Biol.* **87**:211.

Allegrini, P. R., Van Scharrenburg, G., De Haas, G. H., and Seelig, J., 1983, *Biochim. Biophys. Acta* **731**:448.

Armitage, I. M., Shapiro, D. L., Furthmayr, H., and Marchesi, V. T., 1977, *Biochemistry* **16**:1317.

Arnold, K., Losche, A., and Gawrisch, K., 1981, *Biochim. Biophys. Acta.* **645**:143.

Bally, M. B., Tilcock, C. P. S., Hope, M. J., and Cullis, P. R., 1982, *Can. J. Biochem. Cell Biol.* **61**:346.

Bammel, B. P., Brand, J. A., Simmons, R. B., Evans, D., and Smith, J. C., 1987, *Biochim. Biophys. Acta* **896**:136.

Banaszak, L. J., and Seelig, J., 1982, *Biochemistry*, **21**:2436.

Barker, R. W., Bell, J. D., Radda, G. K., and Richards, R. E., 1972, *Biochim. Bioophys. Acta.* **260**:161.

Bayerl, T., Klose, G., Ruckpaul, K., and Schwarze, W., 1985, *Biochim. Biophys. Acta* **812**:437.

Bayerl, T., Klose, G., Blanck, J., and Ruckpaul, K., 1986, *Biochim. Biophys. Acta* **858**:285.

Berden, J. A., Barker, R. W., and Radda, G. K., 1975, *Biochim. Biophys. Acta* **375**:186.

Beyer, K., 1986 *Biochim. Biophys. Acta* **855**:365.

Beyer, K., and Klingenberg, M., 1985, *Biochemistry* **24**:3821.

Boni, L. T., and Hui, S. W., 1983, *Biochim. Biophys. Acta* **731**:177.

Boni, L. T., Stewart, T. P., Alderfer, J. L., and Hui, S. W., 1981, *J. Membrane Biol.* **62**:71.

Boulanger, Y., Schreier, S. Smith, I. C. P., 1981, *Biochemistry* **20**:6824.

Brown, M. F., and Seelig, J., 1977, *Nature.* **269**:721.

Brown, M. F., and Seelig, J., 1978, *Biochemistry* **17**:381.

Brown, M. F., Deese, A. J., and Dratz, E. A., 1982, *Methods in Enzymology* **81**:709.

Brown, P. M., Steers, J., Hui, S. W., Yeagle, P. L., and Silvius, J. R., 1986, *Biochemistry* **25**:4259.

Browning, J. L., 1981a, *Biochemistry* **20**:7123.

Browning, J. L., 1981b, *Biochemistry* **20**:7133.

Browning, J. L., and Seelig, J., 1980, *Biochemistry* **19**:1962.

Bruzik, K., Jiang, R. T., and Tsai, M. D., 1983, *Biochemistry* **22**:2478.

Buldt, G., Gally, H. U., Seelig, A., Seelig, J., and Zaccai, G., 1978, *Nature* **271**:182.

Burnell, E. E., Van Alphen, L., Verkleij, A., and De Kruijff, B., 1980, *Biochim. Biophys. Acta* **597**:492.

Burnell, E. E., Cullis, P. R., and DeKruijff, B., 1980, *Biochim. Biophys. Acta.* **603**:63.

Burns, R. A., Stark, R. E., Vidusek, D. A., and Roberts, M. F., 1983, *Biochemistry*, **22**:5084.

Bystrov, V. F., Shapiro, Y. E., Viktorov, A. V., Barsukov, L. I., and Bergelson, L. D., 1972, *FEBS Lett.* **25**:337.

Campbell, R. F., Meirovitch, E., and Freed, J. H., 1979, *J. Phys. Chem.* **83**:525.

Capuozzo, E., and Bolard, J., 1985, *Biochim. Biophys. Acta* **820**:63.

Castellino, F. J., 1978, *Arch. Biochem. Biophys.* **189**:465.

Castellino, F. J., and Violand, B. N., 1979, *Arch. Biochem. Biophys.* **193**:543.

Chan, S. I., Bocian, D. F., and Peterson, N. O., 1981, in *Membrane Spectroscopy* (E. Grell, ed., Springer-Verlag, New York., p. 1.

Cheng, K., Lepock, J. R., Hui, S. W., and Yeagle, P. L., 1986, *J. Biol. Chem.* **261**:5081.

Chrzeszczyk, A., Wishnia, A., and Springer, C. S. Jr., 1977, *Biochim. Biophys. Acta* **470**:161.

Chupin, V. V., Vasilenko, I. A., Predvoditelev, D. A., Serebrennikova, G. A., and Evstigneeva, R. P., 1979, *Doklady Akademii Nauk SSSR* **248**:235.
Cullis, P. R., 1976, *FEBS Lett.* **70**:223.
Cullis, P. R., and de Kruijff, B., 1976, *Biochim. Biophys. Acta* **43**:523.
Cullis, P. R., and de Kruijff, B., 1978, *Biochim. Biophys. Acta* **513**:31.
Cullis, P. R., and de Kruijff, B., 1979, *Biochim. Biophys. Acta* **559**:399.
Cullis, P. R., and Grathwohl, C., 1977, *Biochim. Biophys. Acta* **471**:213.
Cullis, P. R., and Hope, M. J., 1980, *Biochim. Biophys. Acta* **597**:533.
Cullis, P. R., de Kruijff, B., and Richards, R. E., 1976, *Biochim. Biophys. Acta* **426**:433.
Cullis, P. R., Verkleij, A. J., and Ververgaert, P. H. J. Th., 1978, *Biochim. Biophys. Acta* **513**:11.
Cullis, P. R., de Kruijff, B., Hope, M. J., Nayer, R., Rietveld, A., and Verkleij, A. J., 1980, *Biochim. Biophys. Acta* **600**:625.
Cushley, R. J., Forrest, B. J., Grover, A. K., and Wassall, S. R., 1979, *Can. J. Biochem.* **58**:206.
Davis, D. G., 1972, *Biochem. Biophys. Research Communication* **49**:1492.
Davis, D. G., and Inesi, G., 1972, *Biochim. Biophys. Acta* **282**:180.
Dea, P., and Chan, S. I., 1986, *Biochim. Biophys. Acta* **854**:1.
Deese, A. J., Dratz, E. A., Dahlquist, F. W., and Paddy, M. R., 1981a, *Biochemistry* **20**:6420.
Deese, A. J., Dratz, E. A., and Brown, M. F., 1981b, *FEBS Lett.* **124**:93.
deGrip, W. J., Drenthe, E. H. S., van Echteld, C. J. A., deKruijff, B., and Verkleij, A. J., 1979, *Biochim. Biophys. Acta* **558**:330.
de Kruijff, B., and Cullis, P. R., 1980, *Biochim. Biophys. Acta* **601**:235.
de Kruijff, B., Cullis, P. R., and Radda, G. K., 1975, *Biochim. Biophys. Acta* **406**:6.
de Kruijff, B., Cullis, P. R., and Radda, G. K., 1976, *Biochim. Biophys. Acta* **436**:729.
de Kruijff, B., van Den Besselaar, A. M. H. P., Cullis, P. R., van Den Bosch, H., and van Deenen, L. L. M., 1978, *Biochim. Biophys. Acta* **514**:1.
de Kruijff, B., Morris, G. A., and Cullis, P. R., 1980, *Biochim. Biophys. Acta* **598**:206.
de Kruijff, B., Nayer, R., and Cullis, P. R., 1982, *Biochim. Biophys. Acta* **684**:47.
Ellena, J. F., Pates, R. D., and Brown, M. F., 1986, *Biochemistry* **25**:3742.
Eriksson, P-O., Rilfors, L., Lindblom, G., and Arvidson, G., 1985, *Chem. and Phys. of Lipids* **37**:357.
Farren, S. B., and Cullis, P. R., 1980, *Biochem. Biophys. Res. Commun.* **97**:182.
Favre, E., Baroin, A., Bienvenue, A., and Devaux, P. F., 1979, *Biochemistry* **18**:1156.
Ferguson, K. A., Hui, S. W., Stewart, T. P., and Yeagle, P. L., 1982, *Biochim. Biophys. Acta* **684**:179.
Finer, E. G., Henry, R., Leslie, R. B., and Robertson, R. N., 1975, *Biochim. Biophys. Acta* **380**:320.
Fraser, P. E., Moscarello, M. A., Rand, R. P., and Deber, C. M., 1986, *Biochim. Biophys. Acta* **863**:282.
Frye, J., Albert, A. D., Selinsky, B. S., and Yeagle, P. L., 1985, *Biophys. J.* **48**:547.
Fuldner, H. H., 1981, *Biochemistry*, **20**:5707.
Gagne, J., Stamatatos, L., Diacovo, T., Hui, S. W., Yeagle, P. L., and Silvius, J. R., 1985, *Biochemistry* **24**:4400.
Gally, H. U., Pluschke, G., Overath, P., and Seelig, J., 1980, *Biochemistry* **19**:1638.
Gawrisch, K., Arnold, K., Ruger, H. J., Kertscher, P., and Nuhn, P., 1977, *Chem. and Phys. of Lipids* **20**:285.
Gerritsen, W. J., de Kruijff, B., Verkleij, A. J., De Gier, J., Van Deenen, L. L. M., 1980, *Biochim. Biophys. Acta* **598**:554.
Grasdalen, H., Goran Eriksson, L. E., Westman, J., and Ehrenberg, A., 1077, *Biochim. Biophys. Acta* **469**:151.
Griffin, R. G., Powers, L., and Pershan, P. S., 1978, *Biochemistry* **17**:2718.
Griffith, O. H., and Jost, P. C., 1976, in *Spin Labeling: Theory and Applications* (Berliner, L. J., ed.), Academic Press, New York, p. 250.

Gruner, S. M., and Jain, M. K., 1985, *Biochim. Biophys. Acta* **818**:352.
Herzfeld, J., Griffin, R. G., and Haberkorn, R. A., 1978, *Biochemistry* **17**:2711.
Hitchcock, P. B., Mason, R., Thomas, K. M., and Shipley, G. G., 1974, *Proc. Natl., Acad. Sci. USA* **71**:3036.
Hope, M. J., and Cullis, P. R., 1979, *FEBS Lett.* **107**:323.
Hope, M. J., and Cullis, P. R., 1980, *Biochem. and Biophys. Res. Commun.* **92**:846.
Hornby, A. P., and Cullis, P. R., 1981, *Biochim. Biophys. Acta* **647**:285.
Horwitz, A. F., and Klein, M. P., 1972, *J. Supramol. Struct.* **1**:19.
Huang, C., 1969, *Biochemistry*, **8**:344.
Hui, S. W., Stewart, T. P., and Yeagle, P. L., 1980, *Biochim. Biophys. Acta* **601**:271.
Hui, S. W., Stewart, T. P., Yeagle, P. L., and Albert, A. D., 1981, *Arch. Biochem. Biophys.* **207**:227.
Hui, S. W., Stewart, T. P., and Boni, L. T., 1983, *Chem. and Phys. of Lipids* **33**:113.
Hutton, W. C., Yeagle, P. L., and Martin, R. B., 1977, *Chem. and Phys. of Lipids.* **19**:255.
Jarell, H. C., Byrd, R. A., Deslauriers, R., Ekiel, I., and Smith, I. C. P., 1981, *Biochim. Biophys. Acta* **648**:80.
Jones, I. C., and Hunt, G. R. A., 1985, *Biochim. Biophys. Acta* **820**:48.
Killian, J. A., de Kruijff, B., Van Echteld, C. J. A., Verkleij, A. J., Leunissen-Bijvelt, J., and De Gier, J., 1983, *Biochim. Biophys. Acta* **728**:141.
Killian, J. A., Verkleijj, A. J., Leunissen-Bijvelt, J., and de Kruijff, B., 1985, *Biochim. Biophys. Acta* **812**:21.
Killian, J. A., Borle, F., de Kruijff, B., and Seelig, J., 1986, *Biochim. Biophys. Acta* **854**:133.
Killian, J. A., Burger, K. N. J., and de Kruijff, B., 1987, *Biochim. Biophys. Acta* **897**:269.
Klose, G., 1975, *Chem. and Phys. of Lipids.* **15**:9.
Kohler, S. J., and Klein, M. P., 1977, *Biochemistry* **16**:519.
Lau, A., and Cowburn, D., 1981, *Biophys. Chem.* **14**:267.
Lichtenberg, D., Amselem, S., and Tamir, I., 1979, *Packing and Head Group Conformation* **18**:4169.
London, E., Feigenson, and G. W., 1979, *J. Lipid Research.* **20**:408.
Longmuir, K. J., Capaldi, R. A., and Dahlquist, F. W., 1977, *Biochemistry* **16**:5746.
Lund-Katz, S., and Phillips, M. C., 1986, *Biochemistry* **25**:1562.
Madden, T. D., and Cullis, P. R., 1982, *Biochim. Biophys. Acta* **684**:149.
Marsh, D., and Seddon, J. M., 1982, *Biochim. Biophys. Acta* **690**:117.
McLaughlin, A. C., Cullis, P. R., Berden, J. A., and Richards, R. E., 1975a, *J. Mag. Res.*, **20**:146.
McLaughlin, A. C., Cullis, P. R., Hemminga, M. A., Hoult, D. I., Radda, G. K., Ritchie, G. A., Leeley, P. J., and Richards, R. E., 1975b, *FEBS Lett.* **57**:213.
McLaughlin, A. C., Herbette, L., Blasie, J. K., Wang, C. T., Hymel, L., and Fleischer, S., 1981, *Biochim. Biophys. Acta* **643**:1.
Michaelson, D. M., Horwitz, A. F., and Klein, M. P., 1973, *Biochemistry* **12**:2637.
Michaelson, D. M., Horwitz, A. F., and Klein, M. P., 1974, *Biochemistry* **13**:2605.
Mitchell, R. D., Volpe, P., Palade, P., and Fleischer, S., 1983, *J. Biol. Chem.* **258**:9867.
Mollevanger, L. C. P. J., Dratz, E. A., de Kruijff, B., Hilbers, C. W., and de Grip, W. J., 1986, *Eur. J. Biochem.* **156**:383.
Moore, N. F., Patzer, E. J., Wagner, R. R., Yeagle, P. L., and Hutton, W. C., 1977, *Biochim. Biophys. Acta* **464**:234.
Nicolay, K., van der Neut, R., Fok, J. J., and de Kruijff, B., 1985, *Biochim. Biophys. Acta* **819**:55.
Niederberger, W., and Seelig, J., 1976, *J. Am. Chem. Soc.* **98**:12.
Noggle, J. H., Marecek, J. F., Mandal, S. B., Van Venetie, R., Rogers, J., Jain, M. K., and Ramirez, F., 1982, *Biochim. Biophys. Acta* **691**:240.
Rajan, S., Kang, S. Y., Gutowsky, H. S., and Oldfield, E., 1981, *J. Biol. Chem.* **256**:1160.
Rance, M., and Byrd, A., 1983, *J. Mag. Res.* **52**:221.
Rietveld, A., van Kemenade, T. J. J. M., Hak, T., Verkleij, A., and de Kruijff, B., 1987, *Eur. J. Biochem.* **164**:137.

Robinson, J. D., Birdsall, N. J. M., Lee, A. G., and Metcalfe, J. C., 1972, *Biochemistry* **11**:2903.
Romans, A. Y., Yeagle, P. L., O'Connor, S. E., and Grisham, C. M., 1979, *J. Supramol. Struct.* **10**:241.
Rosier, F., M'Zali, H., and Giraud, F., 1986, *Biochim. Biophys. Acta* **863**:253.
Roux, M., and Champeil, P., 1984, *FEBS Lett.* **171**:169.
Ryba, N. J. P., Horvath, L. I., Watts, A., and Marsh, D., 1987, *Biochemistry* **26**:3234.
Seelig, A., and Seelig, J., 1985, *Biochim. Biophys. Acta* **815**:153.
Seelig, J., 1977, *Quart. Rev. Biophys.* **10**:353.
Seelig, J., and Gally, H. U., 1976, *Biochemistry*, **15**:5199.
Seelig, J., Gally, H., and Wohlgemuth, R., 1977, *Biochim. Biophys. Acta* **467**:109.
Seelig, J., Dijkman, R., and DeHaas, G. H., 1980, *Biochemistry*, **15**:2215.
Seelig, J., Tamm, L., Hymel, L., and Fleischer, S., 1981, *Biochemistry* **20**:3922.
Selinsky, B. S., 1984, Ph.D. Thesis. State University of New York at Buffalo, Buffalo, New York.
Selinsky, B. S., and Yeagle, P. L., 1983, *Biophys. J.* **41**:166a
Selinsky, B. S., and Yeagle, P. L., 1984, *Biochemistry* **23**:2281.
Selinsky, B. S., and Yeagle, P. L., 1985, *Biochim. Biophys. Acta* **813**:33.
Sheetz, M. P., and Chan, S. I., 1972, *Biochemistry* **11**:4573.
Shibata, T., Uzawa, J., and Sugiura, Y., 1983, *Chem. and Phys. of Lipids.* **33**:1.
Silvius, J. R., McMillen, D. A., Saley, N. D., Jost, P. C., and Griffith, O. H., 1984, *Biochemistry* **23**:538.
Siminovitch, D. J., Jeffrey, K. R., and Eibl, H., 1983, *Biochim. Biophys. Acta.* **727**:122.
Sixl, F., and Watts, A., 1982, *Biochemistry*, **21**:6446.
Skarjune, R., and Oldfield, E., 1979, *Biochemistry*, **18**:5903.
Smaal, E. B., Nicolay, K., Mandersloot, J. G., de Gier, J., and de Kruijff, B., 1987, *Biochim. Biophys. Acta* **897**:453.
Smith, I. C. P., and Ekiel, I. H., 1984, in *Phosphorus*-31 *NMR, Principles and Applications*, Academic Press, New York, p. 447.
Smith, R., and Cornell, B. A., 1985., *Biochim. Biophys. Acta* **818**:275.
Smith, R., Cornell, B. A., K[illegible]iry, M. A., and Separovic, F., 1983, *Biochim. Biophys. Acta* **732**:492.
Smith, S. O., Palings, I., Copie, V., Raleigh, D. P., Courtin, J., Pardoen, J. A., Lugtenberg, J., Mathies, R. A., and Griffin, R. G., 1987, *Biochemistry* **26**:1606.
Sotirhos, N., Herslof, B., and Kenne, L., 1986, *J. of Lipid Res.* **27**:386.
Spiess, H. W., and Sillescu, H., 1981, *J. Mag. Res.* **42**:381.
Stidham, M. A., Siedow, J. N., McIntosh, T. J., Porter, N. A., and Moreland, D. E., 1985, *Biochim. Biophys. Acta* **812**:721.
Stier, A., Finch, A. E., and Bosterling, B., 1978, *FEBS Lett.* **91**:109.
Stoffel, W., Zirenberg, O., and Scheefers, H., 1977, *Hoppe-Seyler's Z. Physiol. Chem.* **358**:865.
Sugimoto, Y., Saito, H., Tabeta, R., Kodama, M., Nagata, C., Itabashi, M., Hirota, T., and Toyoshima, S., 1984, *Gann* **75**:798.
Tamm, L. K., and Seelig, J., 1983, *Biochemistry*, **22**:1474.
Taraschi, T. F., de Kruijff, B., Verkleij, Ar. J., and van Echteld, C. J. A., 1982a, *Biochim. Biophys. Acta* **685**:153.
Taraschi, T. F., van Der Steen, A. T. M., de Kruijff, B., Tellier, C., and Verkleij, A. J., 1982b, *Biochemistry* **21**:5756.
Thayer, A. M., and Kohler, S. J., 1981, *Biochemistry*, **20**:6831.
Tilcock, C. P. S., Cullis, P. R., and Gruner, S. M., 1986, *Chem. Phys. Lipids* **40**:47.
Uhing, M. C., 1975, *Chem. Phys. Lipids*, **14**:303.
Utsumi, H., Tunggal, B. D., and Stoffel, W., 1980, *Biochemistry* **19**:2385.
van Echteld, C. J. A., de Kruijff, B., Mandersloot, and J. G., De Gier, 1981a, *Biochim. Biophys. Acta* **649**:211.
van Echteld, C. J. A., van Stigt, R., de Kruijff, B., Leunissen-Bijvelt, J., Verkleij, A. J., and De Gier, J., 1981b, *Biochim. Biophys. Acta* **648**:287.

van Zoelen, E. J. J., van Dijck, P. W. M., de Kruijff, B., Verkleij, A. J,, and van Deenen, L. L. M., 1978a, *Biochim. Biophys. Acta* **514**:9.

van Zoelen, E. J. J., van Dijck, P. W. M., de Kruijff, B., and Verkleij, A. J. D., 1978b, *Biochim. Biophys. Acta* **727**:39.

van Duijn, G., Valtersson, C., Chojnacki, T., Verkleij, A. J., Dallner, G., and de Kruijff, B., 1986, *Biochim. Biophys. Acta* **861**:211.

Vasilenko, I., de Kruijff, B., Verkleij, A. J., 1982, *Biochim. Biophys. Acta* **685**:144.

Viti, V., and Minetti, M., 1981, *Chem. Phys. Lipids*, **28**:215.

Viti, V., Cicero, R., Callari, D., Guidoni, L., Billitteri, A., and Sichel, G., 1983, *FEBS Lett.* **158**:36.

Warren, G. B., Toon, P. A., Birdsall, N. J. M., Lee, A. G., and Metcalfe, J. C., 1974, *Proc. Natl. Acad. Sci. USA* **71**:622.

Westman, J., and Goran-Eriksson, L. E., 1979, *Biochim. Biophys. Acta* **557**:62.

Wigfield, D. C., Phelps, D. J., Pottie, R. F., and Sander, R., 1975, *J. Am. Chem. Soc.* **97**:898.

Wohlgemuth, R., Waespe-Sarcevic, N., and Seelig, J., 1980, *Biochemistry*, **19**:3314.

Yeagle, P. L., 1978, *Accounts Chem. Res.* **11**:321.

Yeagle, P. L., 1979, *Arch. Biochem. and Biophys.* **198**:501.

Yeagle, P. L., 1982, *Biophys. J.*, **37**:227.

Yeagle, P. L., 1984a, *J. Membrane Biol.* **78**:201.

Yeagle, P. L., 1984b, in *Erythrocyte Membranes: Recent Clinical and Experimental Advances* (G. Brewer, ed.), A. Liss, New York.

Yeagle, P. L., and Martin, R. B., 1976, *Biochem. Biophys. Res. Commun.* **69**:775.

Yeagle, P. L., and Romans, A. Y., 1981, *Biophys. J.*, **33**:243.

Yeagle, P. L., and Kelsey, D., 1989, *Biochemistry* **28**:2210.

Yeagle, P. L., and Sen, A., 1986, *Biochemistry* **25**:7518.

Yeagle, P. L., Hutton, W. C., and Martin, R. B., 1975a, *J. Am. Chem. Soc.* **97**:7175.

Yeagle, P. L., Hutton, W. C., Huang, C., and Martin, R. B., 1975b, *Proc. Nat. Acad. Sci.*, **72**:3477.

Yeagle, P. L., Hutton, W. C., Huang, C., and Martin, R. B., 1976, *Biochemistry*, **15**:2121.

Yeagle, P. L., Hutton, W. C., Huang, C., and Martin, R. B., 1977a, *Biochemistry*, **16**:4344.

Yeagle, P. L., Langdon, R. G., and Martin, R. B., 1977b, *Biochemistry* **16**:3487.

Yeagle, P. L., Hutton, W. C., and Martin, R. B., 1978a, *Biochemistry* **17**:5745.

Yeagle, P. L., Martin, R. B., Pottenger, L., and Langdon, R. G., 1978b, *Biochemistry* **17**:2707.

Yeagle, P. L., Selinsky, B. S., and Albert, A. D., 1984, *Biophys. J.* **45**:1085.

2

Investigation of Ribosomal 5S Ribonucleic Acid Solution Structure and Dynamics by Means of High-Resolution Nuclear Magnetic Resonance Spectroscopy

Alan G. Marshall and Jiejun Wu

1. INTRODUCTION

5S rRNA, a single-chain ribonucleic acid of about 120 primary bases, is found universally in the large subunit of all but fungal and animal mitochondrial ribosomes (Erdmann and Wolters, 1986). 5S rRNA is essential for ribosomal structure, since ribosomes from which it has been removed cannot be structurally reconstituted without it (e.g., Nomura and Erdmann, 1970). 5S rRNA also appears to be essential for ribosomal function because ribosomes reconstituted with structurally intact but chemically modified 5S rRNA synthesize proteins at reduced rate (Pieler *et al.*, 1984). As might be anticipated from the essential and universal function of 5S rRNA, its more than 350 known primary base sequences (Erdmann and Wolters, 1986) are highly conserved. Moreover, it is strongly suspected that the secondary and tertiary structure of 5S rRNA is also highly conserved, since 5S rRNA isolated from several prokaryotes (but not eukaryotes) can be reconstituted with the remaining ribosomal components from another bacterium (*B. stearothermophilus*) to give a ribosome which is functional for protein

Alan G. Marshall and Jiejun Wu • Department of Biochemistry, The Ohio State University, Columbus, Ohio 43210. Alan G. Marshall is also a member of the Department of Chemistry.

synthesis (Pieler *et al.*, 1984). Finally, although the location of 5S rRNA in the large ribosomal subunit is well established (Brimacombe and Stiege, 1985), its specific function remains unknown.

1.1. Non-NMR Methods

Ribosomal 5S rRNA typically consists of a single chain of ~120 primary bases. As recently reviewed (Erdmann, 1976; Erdmann *et al.*, 1980; Pieler *et al.*, 1984; Delihas *et al.*, 1984), non-NMR methods for analysis of the secondary structure of 5S rRNA can be categorized as follows:

1. Single stranded oligonucleotide segments: partial ribonuclease digestion, base-specific chemical modification, complementary oligonucleotide binding, and intramolecular crosslinking;
2. Overall size and/or shape: low-angle X-ray scattering, neutron scattering, sedimentation, electron microscopy, circular dichroism (CD);
3. Thermal stability from heat-induced stepwise melting monitored by: differential scanning calorimetry (DSC) or ultraviolet (UV), infrared (IR), Raman, or electron spin resonance (ESR) spectroscopy;
4. Base-pair number and type (G·C, A·U, G·U): UV, IR, Raman, CD hyperchromism;
5. Primary sequence comparisons: often combined with minimal free energy calculations based on Tinoco rules (Tinoco *et al.*, 1971; Tinoco *et al.*, 1973);
6. X-ray crystal diffraction: not yet achieved, although preliminary crystals have been obtained (Morikawa *et al.*, 1982; Abdel-Meguid *et al.*, 1983); and
7. Functional tests: large ribosomal subunit reconstitution, binding of specific ribosomal proteins to 5S rRNAs.

Among the abovementioned techniques, comparative sequence analysis has proved to be the most powerful nonexperimental method for devising a common secondary model for 5S rRNA. In fact, the most widely accepted "universal" minimal base-pair model for both eukaryotic and prokaryotic 5S rRNA's was inferred in this way, and is now supported by a wide variety of NMR and non-NMR experiments (Erdmann, 1976; Erdmann *et al.*, 1980; Pieler *et al.*, 1984; Delihas *et al.*, 1984). Figure 1 shows the prokaryotic version of this secondary structural model, adapted to the primary base sequence of *E. coli* 5S rRNA. The model consists of five helical segments arranged in three distinct stems, designated as I (molecular stalk or terminal helix), II ("tuned" helix), III (common arm), IV (prokaryotic loop), and V (unnamed).

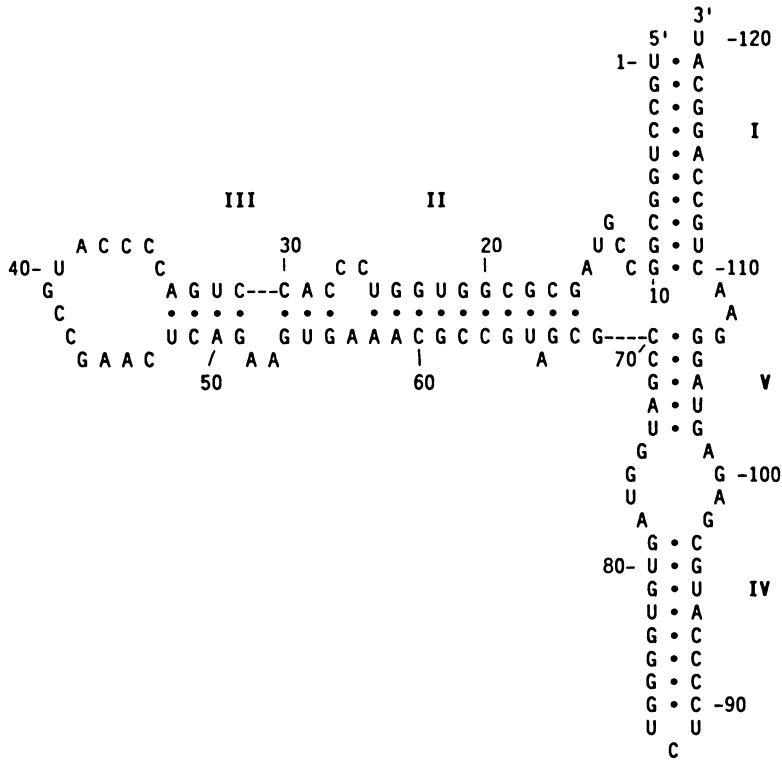


Figure 1. The most commonly accepted base-pairing scheme for 5S rRNA, adapted to the primary nucleotide sequence of *E. coli* 5S rRNA. The helixes of the three component stems are marked with Roman numerals. Helices V and IV are drawn coaxial to helix I.

1.2. NMR Methods

Provided that individual nuclear resonances can be resolved, NMR offers two main advantages for the determination of structure and dynamics of nucleic acids containing 10–160 bases. First, once the resonances are assigned, the base-pair sequence can be established, primarily from one- and two-dimensional ^{1}H homonuclear Overhauser enhancements (NOE) for base-pair protons in the downfield (9–15 ppm from DSS) region of the spectrum. Although base-pair proton resonances can often be relatively easily identified (i.e., G·C, A·U, G·U), a full assignment (e.g., G·C followed by A·U, etc.) generally requires additional experiments, e.g., enzymatic partial digestion, variation in salt concentration (especially [Mg^{2+}]), variation in temperature, spin labeling, comparison of RNA's of different primary sequence, etc. It is important to recognize that application of

modern ^{1}H NMR techniques, e.g., NOESY (Wüthrich, 1986) to the large (~40,000 Dalton) 5S rRNAs is often limited or even precluded by multiple overlap of broad (≥50 Hz) resonances. High-field (≥11.75 tesla; ≥500 MHz for ^{1}H) is absolutely essential. In spite of limited sample quantity, the complete secondary base-pair proton spectra of yeast tRNA^{phe} (Heerschap *et al.*, 1985) and *E. coli* tRNA^{val} (Hare *et al.*, 1985) have been assigned, and several partial assignments for helical segments of several 5S rRNAs are now available (see Section 3).

The second advantage of ^{1}H NMR is that, once the resonances have been assigned, a large number of base pairs may be monitored simultaneously as the molecule is subjected to various perturbations (e.g., denaturation by heat; changes in ionic strength, temperature, or pH; chemical denaturation; etc.). In this way, for example, the relative thermal stability of each of several helical segments can be determined.

At natural abundance, only ^{1}H and ^{31}P NMR are feasible for 5S rRNA. However, as discussed below, most of the 120-odd ^{31}P resonances overlap, so that relatively little can be learned from the ^{31}P spectrum. At major cost (in both effort and expense), ^{19}F, ^{13}C, and ^{15}N can be incorporated at various sites in 5S rRNA, and the results of those experiments will be discussed.

In summary, this article aims to review what has been learned about ribosomal 5S rRNA from NMR spectroscopy. Most of the methods discussed in Section 2 were originally calibrated on the much smaller (~75–90 bases) tRNAs, for which an X-ray crystal structure (tRNA^{phe}) was previously known. We shall discuss their application to the determination of several 5S rRNA helical segment base-pair sequences, and then review measurements of 5S rRNA structural dynamics (e.g., melting, base-pair proton exchange rate with solvent). Finally, directions for future work will be suggested.

2. NMR STRATEGIES FOR RNA SOLUTION STRUCTURE DETERMINATION

2.1. Proton FT-NMR in ^1H_2O

The immediate problem of conducting ^{1}H NMR of RNAs is that the most informative signals arise from A·U, G·C, and G·U imino base-pair protons which exchange relatively rapidly (≥10 sec^{-1}) with the solvent H_2O protons. Thus, ^{1}H NMR experiments must be conducted in H_2O (~100 M in protons), except for a small amount (~5%) of D_2O to provide a field-frequency lock signal. Since typical 5S rRNA concentrations fall in the 0.4–0.75 mM range, the resulting >100,000:1 dynamic range exceeds the

linear range of the analog detector as well as the dynamic range (typically ≤ 15 bit plus sign, or about 32 K) of the analog-to-digital converter (ADC) required to sample the time-domain free-induction decay (FID). Additional detection problems (spinning sidebands, curved spectral baseline, phasing, quantitation of relative peak areas, etc.) also result directly from the dynamic range problem. Therefore, we shall first consider methods for suppression of the huge H_2O signal in proton NMR spectra.

2.1.1. Solvent Suppression

The solvent dynamic range problem is especially severe in Fourier transform (FT) NMR, because the whole spectrum (including the H_2O signal) is typically excited and detected at once. For slow-sweep (continuous-wave or CW) or fast-sweep (correlation) NMR (e.g., Redfield, 1978), the rf frequency-sweep generator can be turned off as the excitation frequency passes through the H_2O resonance frequency so that the H_2O signal is not created in the first place. In fact, the earliest 5S rRNA NMR spectra were all acquired in frequency-sweep mode (Wong *et al.*, 1972; Kearns and Wong, 1974; Salemink *et al.*, 1981). However, FT operation offers hundredfold higher sensitivity as well as much simpler double-resonance capability. We will therefore consider the FT experiment from here on.

In an FT-NMR spectrometer, all of the spectral signals are excited simultaneously by means of a short ($\sim$2–400 μs) magnetic field pulse. The induced rf response voltage is detected, amplified, heterodyned to audio-frequency bandwidth, sampled by an ADC, and accumulated in computer core memory. Ideally, the time-domain FID signal amplitude is scaled such that the maximum FID magnitude is slightly less than the dynamic range of the ADC. For a noiseless signal, the maximum number of FIDs that can be acquired is determined by the difference between the length of one computer word and the ADC dynamic range (e.g., $24 - 16 = 8$ bits $= 256$ scans on a 16-bit ADC combined with a Bruker Aspect 3000 24 bit/word computer). However, since random noise accumulates only as the square root of the number of FIDs, the presence of noise in the signal increases significantly the maximum available number of FIDs without time-domain overflow. In practice, we find that signal-to-noise ratio continues to increase up to $>10^4$ accumulated FIDs.

As discussed in a recent review (Chapter 6 of this volume), more than 100 methods and variants have been proposed for solvent elimination in FT-NMR. Each method is based on some difference between the H_2O signal and the signal(s) of interest: e.g., chemical shift, T_1, T_2, $T_{1\rho}$, scalar coupling, etc. Since RNA imino proton resonances are shifted relatively far downfield (9–15 ppm) from H_2O (about 4.8 ppm), the most successful methods for solvent suppression in RNA proton NMR spectra have been based on

selective excitation designed to excite the imino proton spectral frequency range but not the H_2O resonance.

For example, the Alexander long pulse (Alexander, 1961; Redfield and Gupta, 1971), Redfield $2\underline{1}4\underline{1}2$ pulse (Redfield, 1976; 1978; Redfield *et al.*, 1975; Redfield and Kunz, 1979) and the binomial $1\underline{1}$ and $1\underline{1}3\underline{1}$ hard pulses (Sklenar and Starcuk, 1982; Plateau *et al.*, 1983; Turner, 1983; Clore *et al.*, 1983; Hore, 1983a; 1983b) are the most frequently used. The frequency-domain effect of such a pulse can usually be understood qualitatively from its Fourier transform, sometimes with a full Bloch equation analysis, particularly for low rf amplitudes. Figure 2 compares the frequency-domain spectra obtained by FT of the corresponding time-domain excitation pulse(s)

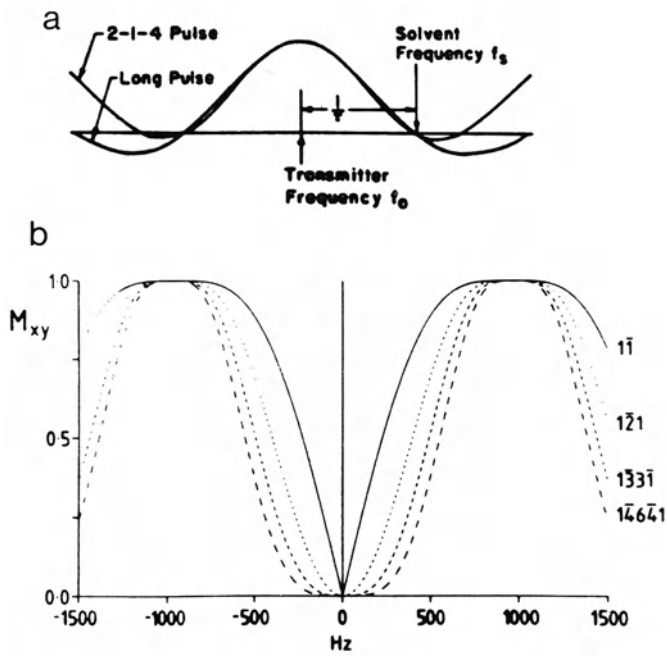


Figure 2. Frequency domain magnitude-mode spectra of various NMR rf pulse sequences designed to produce broad-band excitation except at or near the solvent Larmor frequency. (a) Long-pulse and Redfield $2\underline{1}4\underline{1}2$ pulse sequence: both spectra have been normalized to the same magnitude at the carrier (transmitter) frequency. Adapted from Redfield *et al.* (1975). (b) Pulse sequences for the binomial family. Transverse magnetization (M_{xy}) resulting from each pulse sequence is plotted as a function of offset from the carrier frequency. Full scale corresponds to complete conversion of longitudinal magnetization M_z into M_{xy}. Reproduced, with permission, from Hore (1983b).

for several common water-suppression pulse sequences. In each case, the rf carrier frequency and pulse durations are adjusted so as to place the frequency of zero excitation amplitude exactly at the frequency of the H_2O peak maximum. Although both the long-pulse and Redfield 2$\underline{1}$4$\underline{1}$2 excitations provide zero-magnitude power at the center of the water peak, the Redfield sequence provides superior nulling of the "wings" of the water resonance on either side of its maximum. (It is important to keep in mind that any experimental pulse sequence will produce nonflat excitation power over the spectral bandwidth of interest, as is evident from the examples in Figure 2.) In practice, it is generally possible to reduce the water resonance peak height by a factor of 1,000–5,000 by use of Redfield or binomial pulse sequences.

Additional manipulation of the stored discrete FID data sets can further enhance water suppression. For example, in alternating delay acquisition (ADA) (Roth *et al.*, 1980), every other FID is delayed by half of one period of the difference between the H_2O Larmor frequency and the rf carrier frequency. Thus, the H_2O FID is 180° out of phase in alternate scans, and addition of an even number of FIDs effectively cancels most of the remaining water signal. Similarly, in data shift accumulation (DSA), every other FID data set is left-shifted by the number of time-domain data points corresponding to half of one period of the heterodyned H_2O Larmor frequency (Roth *et al.*, 1980). In order to save time, DSA may be applied to large FID data blocks rather than to individual FIDs.

The above water suppression techniques can be applied to direct spectral acquisition and can also be incorporated into one-dimensional nuclear Overhauser enhancement (NOE) experiments (Redfield and Kunz, 1979; Johnson and Redfield, 1978), relaxation time measurements (Early *et al.*, 1980; Wright *et al.*, 1981), and two-dimensional experiments (Haasnoot and Hilbers, 1983). As shown in part from the data in Figure 3, we obtain the best overall results with 1$\underline{3}$3$\underline{1}$ sine or cosine excitation pulse sequences combined with ADA or DSA to give an overall water suppression of a factor of ~50,000. The effectiveness of various water suppression pulse sequences is shown in Figure 3 for wheat germ 5S rRNA.

2.1.2. Base-Pair Hydrogen Bond Proton Chemical Shifts

A typical 5S rRNA contains about 1300 protons. Fortunately, chemical shift ranges for the various protons have been established from library spectra of synthetic oligonucleotides and tRNAs (Reid, 1981b; Wüthrich, 1986): ribose protons, 3.7–6.3 ppm; nonlabile base protons, 5.3–8.4 ppm; base amino protons, 6.6–9.0 ppm; and base imino protons, 10–15 ppm; all are expressed relative to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). Unlike tRNAs, for which modified (e.g., methylated) bases are common,

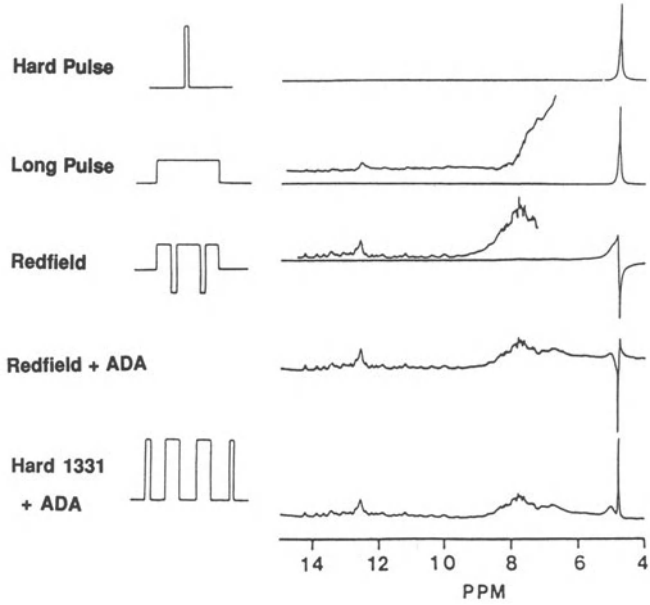


Figure 3. Comparison of several water-nulling NMR rf excitation pulse sequences applied to wheat germ 5S rRNA. ADA denotes alternating delayed acquisition, and a negative-amplitude pulse denotes a 180° phase shift. The best water-nulling and uniform phase is obtained with 1331 plus ADA. Adapted from Li (1985).

intact 5S rRNA offers few resolved characteristic resonances except in the downfield (10–15 ppm) part of the spectrum. Diagnostically useful features of resonances in the 10–15 ppm spectral range are listed below:

1. The resonances found between 10–15 ppm include base-paired imino protons, along with a few extremely deshielded aromatic CH protons (Reid, 1981b; Robillard and Reid, 1979) and/or imino protons which are not base-paired but which are nevertheless prevented from rapid exchange with the solvent by hydrogen bonding to the phosphate backbone or by being buried deep in the interior of the molecule.
2. There is only one imino proton for each A·U and G·C pair, and two for each G·U pair (Figure 4). Thus, interpretation of this spectral region usually reduces simply to identifying the correct base-pair type from among three possibilities, based on the dipolar coupling between a given imino proton and its nearest neighbor proton in that base pair (see Section 2.1.3a).
3. The imino protons from two adjacent base pairs are spatially proximal (0.5 nm), and their dipolar coupling can therefore be detected by T_1 or NOE measurements (see Section 2.1.3a). This

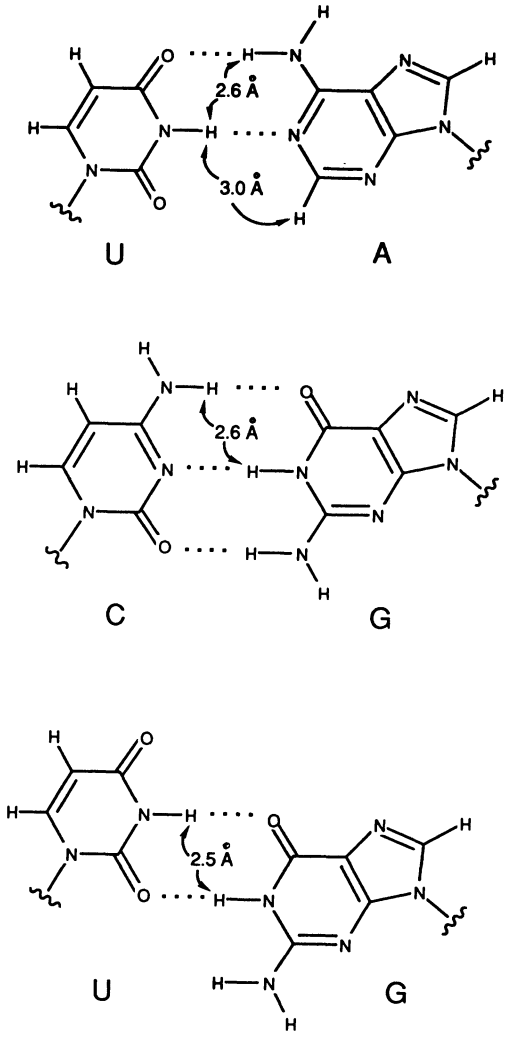


Figure 4. Molecular structures of the three most common base pairs found in RNA. The distances between each imino proton and its NOE observable neighbor proton(s) within each base pair are indicated.

feature provides for base pair sequencing, i.e., by stringing together pairs of dipole-connected base pairs.

4. The base-paired imino protons exchange with solvent sufficiently slowly that their ^1H NMR resonances are observable in H_2O, but sufficiently rapidly that the resonances rapidly disappear when the experiment is conducted in D_2O (Wong *et al.*, 1972; Kearns and

Wong, 1974). Thus, nonimino protons can be identified by comparing ^1H NMR spectra in H_2O vs. D_2O.

Compared to tRNAs, the identification and assignment of base-pair imino proton resonances in 5S rRNAs is more difficult for several reasons. First, 5S rRNA is 50% larger than tRNA, so that 50% more resonances (with larger linewidth, due to slower tumbling) are crowded into the same chemical shift range, leading to extensive peak overlap. Second, the only single crystals of 5S rRNA [*Thermus, thermophilus* (Morikawa *et al.*, 1982)] and the RNase A fragment from *E. coli* (Abdel-Meguid *et al.*, 1983) diffract X-rays to a resolution too low (2.5 and 0.4 nm, respectively) to define the secondary structure. Thus, there is no reference crystal structure of 5S rRNA to serve as a guide for NMR spectral assignments. In the following sections, we will discuss several methods for spectral assignments, including comparison to proposed secondary structural models (including ring-current calculated chemical shifts), enzymatic partial cleavage, thermal partial denaturation, and variation in ionic strength. In place of the modified bases which provided starting points for spectral assignments in tRNAs, characteristic G·U NOEs (Chang and Marshall, 1986; Chen and Marshall, 1986) or spin labeling (Lee and Marshall, 1987) can provide starting points for 5S rRNA assignments.

2.1.3. Detection of Spatially Proximal Protons by Means of Proton Homonuclear Steady-State NOEs

The nuclear Overhauser effect arises from dipolar coupling between spatially proximal magnetic nuclei. In a one-dimensional steady-state NOE experiment, one resonance (spin S) is saturated by intense radiofrequency irradiation and the change in intensity of another resonance (spin I) is measured. The nuclear Overhauser enhancement $f_I(S)$ is the fractional change in intensity (usually approximated by peak height) of I on saturation of S:

$$f_I(S) = \frac{I - I_0}{I_0} \tag{1}$$

in which I and I_0 represent the intensity for resonance I in the presence and absence of saturation of resonance S, respectively. The NOE can be regarded as the fraction of total direct relaxation rate due to cross-relaxation between spins I and S (Noggle and Schirmer, 1971):

$$f_I(S) = \frac{\gamma_S(S+1)\sigma}{\gamma_I(I+1)\rho} \tag{2}$$

in which γ_S and γ_I are the magnetogyric ratios of spins S and I, S and I are the spins of nuclei S and I, σ is the cross-relaxation rate between spins S and I, and ρ is the total direct relaxation rate for spin I.

Dipolar and chemical exchange relaxation mechanisms contribute to both σ and ρ. Other nondipolar relaxation mechanisms contribute to ρ but not σ in equation (2) and thus reduce the NOE. At sufficiently low temperature, chemical exchange is slow, and the NOE is dominated by dipolar relaxation whose rate is proportional to r^{-6}, in which r is the internuclear distance; τ_{rot}, the rotational correlation time of the I-S vector; and the Larmor frequencies of I and S.

As shown in Figure 5, for I = S = ^{1}H, the NOE as a function of τ_{rot} for two dipolar-coupled protons varies between these limits:

$$-1.0 \leq f_I(S) \leq +0.5 \tag{3}$$

As we will later show for the closely related case of I = ^{19}F and S = ^{1}H, $f_I(S)$ reaches its limiting value of -1.0 when the rotational correlation time τ_{rot} for diffusional reorientation of the I-S axis is long compared to a period of the proton Larmor frequency (about 500 MHz at 11.75 tesla). For 5S rRNA (about 40,000 dalton), $\tau_{rot} > 10$ nsec (Marshall and Smith, 1980), so that the full limiting (negative) NOE is expected for protons which are firmly bound to the macromolecular frame. Internal flexibility and/or nondipolar relaxation mechanisms act to reduce the magnitude of the NOE.

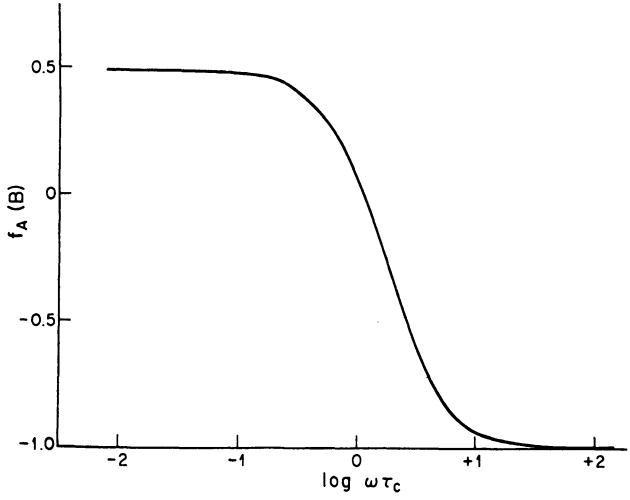


Figure 5. Theoretical homonuclear proton NOE as a function of rotational correlation time τ_c where ω is the proton Larmor (angular) frequency. f_I(S) represents the fractional change in peak area of proton I as a result of saturation of proton S. From Bothner-By (1979).

Also, if excess water molecules are spatially close to spin I, then the population difference transferred from S to I is quickly dissipated ($\rho \gg \sigma$), and $f_I(S)$ is again reduced.

When a third nucleus is nearby, the NOE is further reduced by spin diffusion, in which the energy level population difference for spin I changes the energy level population difference for the third spin (again by means of dipolar relaxation), and so on. Thus, the saturation of spin S is first "felt" by spin I and then propagated onward to other spins in the vicinity. Clearly, the longer the period of irradiation of spin S, the more spin diffusion will occur; NOEs will then be observed not only between spin S and spin I, but also between spin S and all other spins which are dipolar-coupled to spin I. Unfortunately, frequency-selectivity is directly related to irradiation period. Thus, every NOE experiment represents a compromise in which the irradiating power must be left on long enough to provide effective and selective saturation of the target resonance (spin S) without allowing extensive spin diffusion of the NOE to other spins which are not immediately proximal to S.

The nuclear Overhauser effect depends on dipolar coupling, which is relatively short-range. The homonuclear proton NOE drops below practical detectability for $r \geq 0.4$–0.5 nm interproton separation. Thus, the primary NOE for a base-pair imino proton will be to other protons in that base pair, with secondary NOEs to the nearest neighbor base pairs.

It is in principle possible (and, for small molecules, feasible) to determine molecular geometry from NOE measurements. In a two-spin system, with fixed nondipolar relaxation rate and kinown τ_{rot}, the internuclear distance r can be calculated from NOE measurement. However, for a multispin system, only the distance ratios can be determined and then only for irradiation that is sufficiently brief that spin diffusion can be neglected (Noggle and Schirmer, 1971).

NOE measurements were first applied to tRNA (Johnson and Redfield, 1978) and later to 5S rRNA. NOEs have since proved to be the most definitive technique for base-pair identification and sequencing of RNA and DNA duplexes in solution (Wüthrich, 1986; Bothner-By, 1979). Specifically, RNA base-pair type can be established from intra-base-pair NOEs, and base-pair sequence from inter-base-pair NOEs.

2.1.3.1. Base-Pair Identification from Primary NOEs. The NOE ^{1}H NMR difference spectrum (i.e., unirradiated minus irradiated) for each of the three standard RNA base pairs is shown in Figure 6. First, irradiation of the uracil N3H imino proton of an A·U pair produces a prominent (~4–7%) narrow NOE difference peak at 7–8 ppm arising from the nearby (0.29 nm) adenine C2H proton. The NOE to the (also proximal) adenine amino protons has generally not been observed, presumably due to rapid exchange of the amino protons with H_2O. Second, irradiation of the guanine

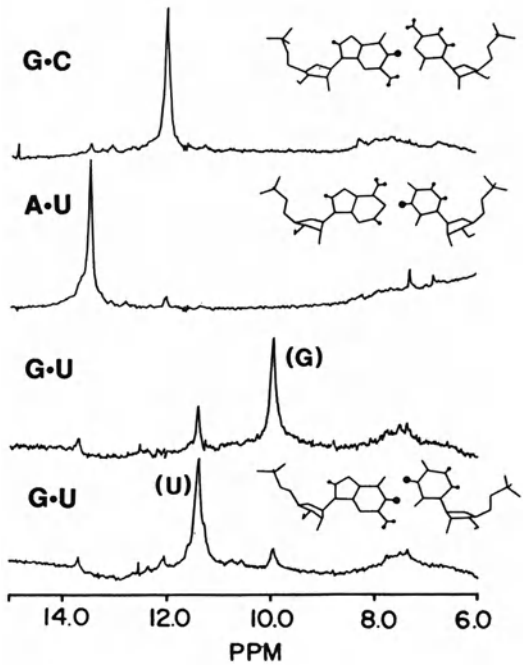


Figure 6. Identification of base-pair type (A·U, G·C, G·U) and base-pair sequence from homonuclear proton NOE patterns. The 500 MHz proton homonuclear NOE difference spectrum (unirradiated minus irradiated spectrum) resulting from irradiation of each of four low-field resonances of *B. subtilis* 5S RNA (40 mg/ml, 10 mM cacodylate, 0.1 M NaCl, 0.1 mM EDTA, pH 7.0, 95:5 $H_2O:D_2O$) is shown. Saturation of the proton designated by a large dot in the corresponding base pair diagram yields an intrabase NOE pattern characteristic of that base-pair type. Small interpair NOEs between imino protons of adjacent base-pairs then serve to establish base-pair sequence. Reproduced, with permission, from Chang and Marshall (1986b).

imino N1H proton of a G·C pair usually produces at best a weak broad NOE difference peak (Sanchez *et al.*, 1980) from the cytosine amino protons. However, in those experiments it is likely that the unobserved intra-base-pair imino-amino NOE was due to low digital resolution rather than the inherent rapid exchange of the amino protons with H_2O (Wu and Marshall, 1989). Third, irradiation of either of the two imino protons (guanine N1H or uracil N3H) of a G·U pair typically produces a strong (>15%) NOE for the other (Johnson and Redfield, 1978). These characteristic intra-base-pair NOEs thus provide a means for identifying base-pair type (A·U, G·C, or G·U). Moreover, the guanine N1H of a G·U pair usually resonates upfield (i.e., closer to DSS) from the uracil N3H of that pair, as confirmed by ^{15}N labeling (see Section 2.3.4). Finally, the NOE produced by irradiation

of the guanine N1H of a G·U pair is usually larger than the NOE produced by irradiation of the uracil N3H of the same pair (Chang and Marshall, 1986b), due to different T_1 for the two imino protons.

2.1.3.2. Base-Pair Sequences from Secondary NOEs. Inspection of Figure 6 reveals additional weak NOEs (2–4%) arising from dipolar coupling between imino proton(s) of the base pair immediately above or below the irradiated base-pair imino proton. For example, Figure 6 clearly shows a dipolar connectivity between the A·U at ~13.6 ppm and the G·C at ~11.9 ppm, evidenced by the NOE difference peak produced at either peak frequency by irradiation at the other. Since each interpair connection establishes a two-pair segment, one might expect to be able to map the full extent of any continuous base-paired helical RNA segment (Roy and Redfield, 1981; Hare and Reid 1982a; 1982b; Roy *et al.*, 1982). Such NOE-based base-pair sequencing experiments are in fact highly successful for relatively short RNA or DNA duplexes (10–20 base pairs). For the much larger tRNAs or 5S rRNAs, in which multiple base-paired segments are present, proton NMR and/or NOEs alone do not suffice to determine the secondary structure, for several reasons.

First, suppose that NOE experiments show that peaks A and C are each dipolar-coupled to peak B (which consists of two resonances which happen to be superimposed at the same chemical shift). Although the connectivities, A–B and B–C, are thereby established, we cannot be sure that the B proton coupled to A is the same B proton that is coupled to C. Thus, an NOE-determined base-pair sequence terminates at any peak which contains more than one imino proton resonance. For example, Figure 7 establishes an A·U–G·C segment and a G·C–G·U segment involving the same G·C chemical shift, but there is no way of knowing whether or not the two G·C signals originate from the same or different G·C base-pairs in the RNA molecule. Ambiguities arising from overlapped resonances pose the single greatest problem for determination of RNA structure from NMR and NOE experiments. If two resonances partially overlap (say, by 0.5–1.0 linewidths), then off-center lower-power (and therefore more frequency-selective) irradiation slightly to the left or right of the composite peak can sometimes be used to distinguish one resonance from the other.

A second problem arises from the inherent need for difference spectra for one-dimensional NOE measurements. If a given resonance in the irradiated spectrum is shifted even slightly from its position in the unirradiated spectrum, then the difference spectrum will exhibit a prominent signal at the frequency of that peak, even if no dipolar coupling is involved (Derome, 1987). Since chemical shifts are temperature-dependent, it is therefore especially important to maintain a constant temperature throughout the experiment, and to alternate frequently between irradiated and unirradiated mode.

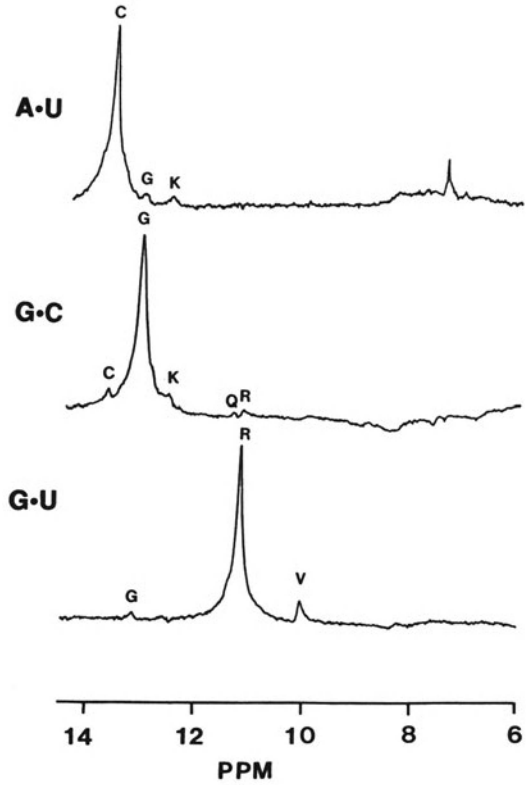


Figure 7. Ambiguity in inference of base-pair sequence from secondary NOEs in the presence of overlapping resonances. Shown are the 500 MHz proton homonuclear NOE difference spectra resulting from irradiation of each of three low-field resonances of wheat germ 5S RNA (35 mg/ml, 10 mM sodium cacodylate, 0.1 NaCl, 0.1 mM EDTA, pH 7.0, 95:5 $H_2O:D_2O$) at 23 °C. Saturation of each resonance gives (small) inter-base pair NOEs (peak C to peak G and peak G to peak R). (Peaks R and V are from a GU pair—see Figure 6.) However, the base-pair sequence, C-G-R/V cannot be inferred from this experiment, because peak G contains two superimposed resonances. Adapted from Li and Marshall (1986).

Third, NOE connectivity by itself distinguishes neither between 5′ → 3′ and 3′ → 5′ nor between the two primary nucleotide strands on which a given base is located. Additional experiments designed to resolve these ambiguities will be discussed below.

Fourth, an NOE experiment does not distinguish between secondary and tertiary base pairs. Furthermore, imino protons which are not paired but slowly exchange with solvent can be mistaken for G·C imino base-pair protons in the same chemical shift range.

Finally, even after short base-pair sequences are established by NOE experiments, their assignment to specific bases of the primary nucleotide

sequence is not straightforward, because of multiple possible base-pairing schemes which could be consistent with the NMR results. Therefore, NMR/NOE techniques are presently best used to distinguish between previously proposed secondary structural base-pairing schemes. Fortunately, assignment of even a few short base-paired segments can often suffice to decide between two or more proposed 5S rRNA secondary models (see Section 3).

In short, ^{1}H homonuclear NOE experiments offer the most powerful single available technique for determination of RNA structure in solution. Nevertheless, as discussed in the next section, NOE experiments must be combined with a variety of other physical, chemical, and biochemical measurements in order to obtain definitive results.

2.2. Adjunct Experiment for Assignments of Proton Resonances

2.2.1. Total Base-Pair Number from Spectral Integration and/or Simulation

5S rRNA secondary structure determination requires three successive determinations: the total number of base pairs; the relative numbers of A·U, G·C, and G·U pairs; and their primary sequence assignments. NMR studies afford a particularly direct approach to the first two determinations by means of simple integration and/or simulation of the downfield region. Integration is feasible when at least one resonance is sufficiently well-resolved (and unmelted) to provide a standard peak area corresponding to a single imino proton. Simulation is performed by fitting the absorption-mode spectrum to a sum of component peaks of common shape (e.g., Lorentz or Gauss, depending on prior time-domain apodization), height, and width, and of adjustable peak positions.

The reproducibility of a given integration method is about ±10%. However, the accuracy with which the number of base pairs can be determined from such integration or simulation depends upon several factors (Robillard and Reid, 1979; Kearns, 1976): the availability of well-resolved peaks for unit peak area standard, baseline curvature (and flattening algorithm, if employed), and selective excitation pulse sequence. It is assumed (as seen in Figure 6) that each G·U contributes two, and that A·U and G·C each contribute one NMR-visible proton resonance to the downfield ^{1}H NMR spectrum.

For tRNAs, for which the base-pair number is known imdependently from X-ray crystallography, the total base-pair number determined by proton NMR peak integration and simulation is accurate to about ±1 out of ~26 (Robillard and Reid, 1979).

Although the total number of base-pairs determined from such integrations or simulations can be relatively accurate, several caveats are in order: (1) noninteger relative peak areas may result when some resonances are partially "melted" due to rapid exchange with solvent; (2) because G·U pairs contribute two imino protons, each G·U peak must be identified (by primary NOEs) to make sure that it is not miscounted as two G·Cs: (3) some imino resonances in the downfield region may arise from nonpaired bases; (4) the number of base pairs can depend strongly on [Mg^{++}] and temperature; (5) more than one 5S rRNA conformation (and thus more than one base-pairing pattern) may be present; and (6) tertiary base-pair ^{1}H resonances may prove indistinguishable from secondary ones.

The choice of chemical shift range can also affect the total base-pair estimate. In the absence of NOE identification of base-pair type, it is usually assumed that G·C base-pair protons lie between 11–13 ppm and A·U protons between 13–15 ppm. However, those two chemical shift ranges can overlap, as can the ranges for G·U resonances (10–12.5 ppm) and G·Cs.

In spite of all the prospective problems with peak integration/simulation, the results shown in Table 1 for several 5S rRNAs studied in different

TABLE 1
Number of Base Pairs Determined by ^{1}H NMR for Various 5S rRNAs

Species	Range (ppm)	T (°C)	No Mg^{2+}	$+MG^{2+}$	G·C	Reference
Yeast	11–15	30	21 ± 2	28 ± 3	85%	Wong *et al.* (1972)[a]
	11.6–15.0	25	32 ± 2	36 ± 2		Luoma *et al.* (1980)[b]
E. coli	11–15	30	28 ± 2	28 ± 3	86%	Kearns and Wong (1974)[a]
A_H	9.5–15	26		34–38		Kime and Moore (1983b)[c]
A_L	9.5–15	26	28–32			Kime and Moore (1983b)
B	9.5–15	26	26–30			Kime and Moore (1983b)
B	11.6–14.5	26	≥33			Burns *et al.* (1980b)[b]
B. subtilis	9–15	23	38			Chang and Marshall (1986b)[d]
B. licheniformis	10.7–14.5	29	37			Salemink *et al.* (1981)[e]
Wheat germ	9–14.5	23	35			Li and Marshall (1986)[d]

[a] The integration standard was incorrect: error of −6 or −7 protons.
[b] Long pulse.
[c] 1$\underline{1}$ hard pulse.
[d] Redfield 2$\underline{1}$412 pulse.
[e] Correlation spectrum.

laboratories are remarkably self-consistent, with the exception of the earliest attempts (Wong *et al.*, 1972; Kearns and Wong, 1974) which were limited by poor resolution and (Kearns and Wong, 1974) by an incorrect choice of standard (19 rather than 26 resonances between 11–15 ppm for yeast $tRNA^{phe}$ (Reid and Hurd, 1977; Robillard and Reid, 1979). In general, the number of 5S rRNA base-pairs at room temperature is about 35, plus an additional 3–4 new base-pairs in the presence of Mg^{2+}. 5S rRNA is made up predominantly of G·C pairs. The NMR results are supported by a variety of non-NMR results from, e.g., ultraviolet and infrared hyperchromism (Burkey *et al.*, 1983; 1984; Chang *et al.*, 1984), and differential scanning microcalorimetry (Li and Marshall, 1985; Chang and Marshall, 1986a).

2.2.2. Computed Ring Current Induced Chemical Shifts

RNA imino base-pair hydrogen-bond protons are shifted far downfield, primarily as the result of electric currents induced by the applied magnetic field in the aromatic rings of the bases. Because each hydrogen-bond imino proton is flanked by an aromatic ring on each side, the ring current induced chemical shift is about twice as large (12–15 ppm) as for a proton on a single aromatic ring (7–8 ppm). Thus, for a given base-pair RNA helical segment, one might hope to be able to compute the chemical shift for each base-pair hydrogen-bond imino proton quantitatively (e.g., Arter and Schmidt, 1976; Geerdes and Hilbers, 1979).

Ring current induced chemical shifts have been tabulated for various RNA helical conformations, and their success in matching experimental imino proton chemical shifts (mainly in tRNAs) has been recently reviewed (Kearns, 1976; 1977; Robillard and Reid, 1979; Reid *et al.*, 1979; Reid, 1981a; Perkins, 1982). For a known base-pair sequence, the calculated ring current induced shifts are typically accurate to within ~0.3 ppm (Kearns, 1976; Reid *et al.*, 1979; Hilbers, 1979; Schimmel and Redfield, 1980). Unfortunately, ring current shifts have not proved particularly useful in evaluating experimental spectra of unknown base-pair sequences for several reasons.

First, the induced shift calculations are based on a static model corresponding to an average tRNA conformation. Second, irregularities in the helical backbone (e.g., from G·U pairs, which are rare in tRNA but common in 5S rRNA), tertiary base-pairing, and base twist and/or tilt are not included. Third, chemical shift contributions from sources other than ring currents are not considered: e.g., binding of counterions, pH, and temperature (by means of its effect on the helical pitch and tilt). Even for tRNA, imino proton chemical shift assignments based solely upon ring current estimates can be incorrect (e.g., Sanchez *et al.*, 1980; Redfield *et al.*, 1981).

Ring current analysis of some early 5S rRNA proton NMR spectra at different temperatures was interpreted to demonstrate the existence of a molecular stalk (Kearns and Wong, 1974; Salemink *et al.*, 1981) and a prokaryotic loop (Salemink *et al.*, 1981). However, those results are highly questionable, because (in addition to the above-cited difficulties) the spectral resolution was relatively poor and one of the methods used for ring current calculation (Kearns and Wong, 1974) has since been proved incorrect (Reid and Hurd, 1977).

2.2.3. Comparison of RNAs of Different Primary Sequences

The greatest difficulty in assigning base pairs by means of ^{1}H homonuclear NOE experiments is in establishing a starting point, i.e., the primary sequence location of the base-pair imino hydrogen-bond proton resonance of that nucleotide. As noted above, G·U pairs offer a possible starting point, because (1) there are usually only a few G·U pairs in the secondary structure, and (2) the two hydrogen-bond protons of each G·U pair are relatively easy to identify because of their characteristic strong mutual NOEs.

One of the conceptually simplest approaches is to compare the ^{1}H NMR spectra of two 5S rRNAs whose primary nucleotide sequences differ at only a few residues. If one (or two, for a G·U pair) resonances exhibit different chemical shifts in the two spectra, then one can immediately identify the primary sequence position for that resonance(s), provided that the two 5S rRNA molecules have otherwise identical secondary and tertiary structure. Since the 5S rRNA primary sequences for several hundred ogranisms are known, many such comparative opportunities exist. The principal difficulty with this approach is that the experimental conditions for optimal isolation of 5S rRNA can differ substantially for different organisms (Li *et al.*, 1984; Lee and Marshall, 1986a).

Two more general approaches which are both in their infancy with respect to 5S rRNA are site-specific mutagenesis and direct RNA synthesis. Those methods will be discussed in Section 3.

2.2.4. Nitroxide Spin Labeling

Another potential starting point for imino proton assignment is to shift and/or broaden the resonance arising from a base at known primary sequence location. Although paramagnetic metal ions (especially lanthanides other than La^{3+}) can shift and broaden proton resonances, binding of metal ions to RNA is notoriously nonspecific. Therefore, a better approach is to covalently label a given residue with a nitroxide radical

whose unpaired electron will then broaden any nearby (up to ~0.8 nm) proton resonance beyond detectability.

A convenient labeling site is afforded by the 2′,3′-diol linkage unique to the 3′-terminal ribose of 5S rRNA. Periodate cleavage of that diol can be followed by reaction with a nitroxide amine to yield a morpholino paramagnetic nitroxide radical at the 3′-terminus. For example, Figure 8 shows the imino proton downfield ^1H 500 MHz spectrum of yeast 5S rRNA before and after morpholino labeling at the 3′-terminus. Although it was not possible to achieve 100% labeling, it is clear that two resonances are markedly reduced in intensity and can therefore be assigned to the terminal $G_1 \cdot C_{120}$ and $G_2 \cdot U_{119}$ base pairs (Lee and Marshall, 1987). Independent ESR measurements, which are blind to unlabeled 5S rRNA, confirmed that the spin label was indeed attached firmly to the RNA.

2.2.5. Differential Thermal Melting

If an RNA molecule can be induced by heating to unfold in stages at well-separated temperatures, then the corresponding imino base-pair proton NMR spectrum can be simplified by selective "melting" of particular resonances. It is important to note that the RNA duplex need not break completely (as in optical "melting" experiments); once a given base-pair imino proton begins to exchange rapidly ($>100\ sec^{-1}$ or so) with solvent, the imino resonance will broaden below detectability.

Optimal use of selective thermal melting requires independent determination of the salt concentrations, e.g., [NaCl] and [$MgCl_2$] which produce optimally separated melting states. Differential scanning microcalorimetry (DSC) offers just such a control experiment, as shown in Figure 9 for wheat germ 5S rRNA at four different combinations of salt concentrations (Li and Marshall, 1985). Examples of thermal melting as an aid in spectral assignment and interpretation will be given in Section 3.

2.2.6. Variation in Ionic Strength

Changing the ionic strength of an RNA solution can also selectively shift particular imino proton resonances, as shown in Figure 10. Thus, the component resonances of an NMR peak composed of two or more overlapped resonances can sometimes be resolved by changing the salt concentration, particularly [Mg^{2+}]. Of course, a sufficiently large change in salt concentration can actually change the number of base pairs (see Table 1) and hence the conformation of the RNA molecule. This sort of dilemma besets all attempts to assign proton NMR spectra of RNAs—one must perturb the molecule in some way in order to distinguish between the various

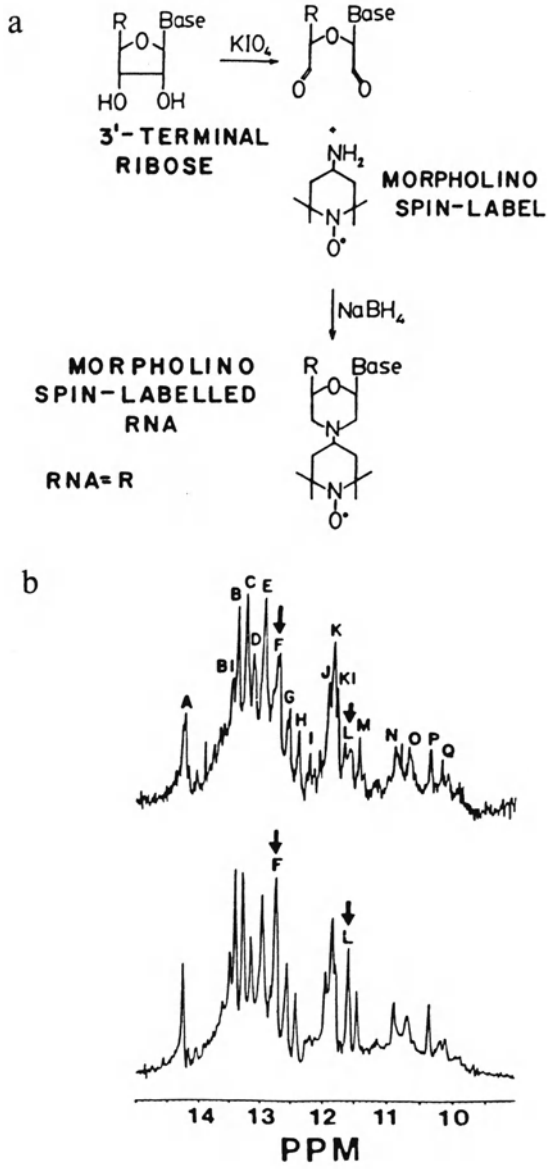


Figure 8. Use of spin labeling for assignment of yeast 5S rRNA ^{1}H 500 MHz low-field NMR spectrum. (a) Chemical reaction scheme for spin labeling of yeast 5S rRNA with TEMPO-NH_2 to give a morpholino spin label attached to the 3′ terminal ribose, (b) Proton 500 MHz NMR spectra of spin-labeled (top trace) and native (bottom trace) yeast 5S rRNA (34 mg/ml, 10 mM Tris, 0.1 M NaCl, 0.1 mM EDTA, pH 7.0, 95:5 $H_2O:D_2O$) at room temperature. There is a significant decrease in intensity of protons near the 5S rRNA 3′ terminus. In fact, in conjunction with NOE results, F/L can be assigned to $U_{119}{\cdot}G_2$. Adapted from Lee and Marshall (1987).

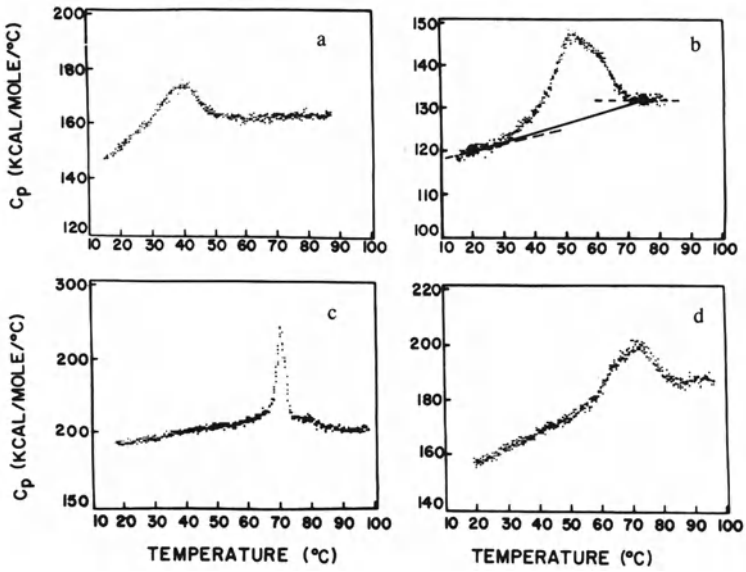


Figure 9. Differential scanning calorimetry melting curves (arbitrary *y*-scale zero) for wheat germ 5S rRNA [0.16, 0.17, 0.11, and 0.17 mM in (a), (b), (c), and (d), respectively] for four different salt conditions. Each buffer contained 10 mM sodium cacodylate, pH 7.0, with the addition of: (a) no added NaCl or $MgCl_2$; (b) 100 mM NaCl; (c) 10 mM $MgCl_2$; and (d) 100 mM NaCl and 10 mM $MgCl_2$. These results, combined with those of Figure 10, clearly show that wheat germ 5S rRNA exhibits different conformational stability under different salt conditions. Reproduced, with permission, from Li and Marshall (1985).

resonances, without perturbing the molecule so much that its structure is no longer the native conformation.

2.2.7. Enzymatic Partial Cleavage

Perhaps the most valuable single adjunct tool for assignment of 5S rRNA proton NMR spectra is the cleavage of the molecule into fragments of well-defined sequence by use of suitable enzymes. In this way, it becomes possible to assign certain resonances unambiguously—e.g., if a given base-paired segment contains only one G·U pair, then its imino protons provide an absolute starting point for peak assignments.

RNA cleavage specificities for three common ribonucleases (RNase T1, RNase T2, and RNase A) are illustrated in Figure 11. Although most 5S rRNAs contain more than one site susceptible to cleavage by a given ribonuclease, it is nevertheless possible to achieve relatively high yield of a given specific 5S rRNA fragment by careful choice of incubation conditions. Assignments based on partial enzymatic cleavage will be discussed in detail in Section 3.

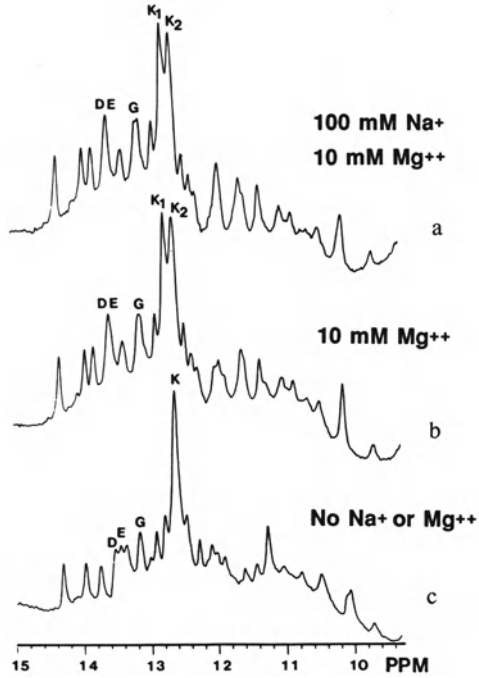


Figure 10. Use of salt-induced NMR frequency shifts to resolve overlapping low-field proton 500 MHz resonances of wheat germ 5S rRNA: (a) 30 mg RNA in buffer D; (b) 26 mg RNA in buffer C; (c) 24 mg RNA in buffer A. The buffers are as in Figure 9. Note that peaks D and E are resolved on removal of Mg^{2+} and Na^{+}, and peak K resolves into two components on addition of Mg^{2+}. Adapted from Li and Marshall (1986).

2.3. Nuclei Other than Protons

2.3.1. ^{31}P

Although a typical 5S rRNA contains about 1300 protons, there are only ~120 nucleotides and thus only ~120 phosphates. Therefore, one might hope that a ^{31}P NMR spectrum would be better resolved than a proton NMR spectrum. In fact, one of the earliest high-resolution NMR studies of 5S rRNA (Salemink *et al.*, 1981) reported a ^{31}P NMR spectrum. Unfortunately, although ^{31}P chemical shifts are known to be highly sensitive to phosphate diester bond angle and torsion angle (Gorenstein, 1975; 1984), Figure 12 shows that most of the phosphates of a typical 5S rRNA evidently share similar type of diester bonds and torsion angles and thus exhibit very similar chemical shifts.

The ^{31}P NMR spectrum in Figure 12 shows two other significant features. First, the 5′-triphosphate characteristic of eukaryotic 5S rRNAs is evident from the three phosphate peaks upfield of the main phosphodiester backbone envelope. Second, the presence of a cyclic phosphate resonance at about 20 ppm downfield offers clear evidence for cleavage (if present)

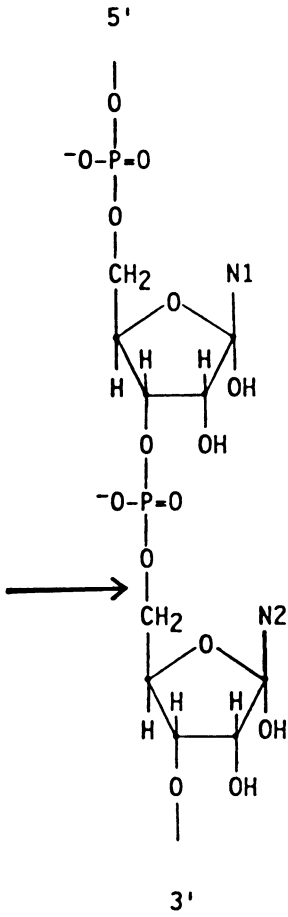


Figure 11. Specificities of ribonucleases T1, T2 and A. Each RNase cleaves at the position designated by an arrow when base(s) N1 = G (RNase T1); N1 = A (RNase T2); and N1 = C or U, N2 = A (RNase A).

of the phosphodiester backbone, and thus gives a built-in measure of RNA degradation during NMR experiments.

The phosphodiester ^{31}P resonances near -1 ppm offer two measures of RNA solution structure, as shown by variation of $[MgCl_2]$ (Figure 13) and of temperature (Figure 14). First, as the RNA A-helix winds tighter at higher $[MgCl_2]$ or lower temperature, the composite phosphodiester peak shifts upfield. The ^{31}P chemical shift of the main resonance peak thus offers a direct measure of the pitch of the helix. Second, the few ^{31}P peaks shifted downfield from the main phosphodiester resonance represent phosphates in conformations markedly different (e.g., sharp turns in the phosphodiester backbone, as in hairpin loops) from the g^-g^- conformation of the RNA

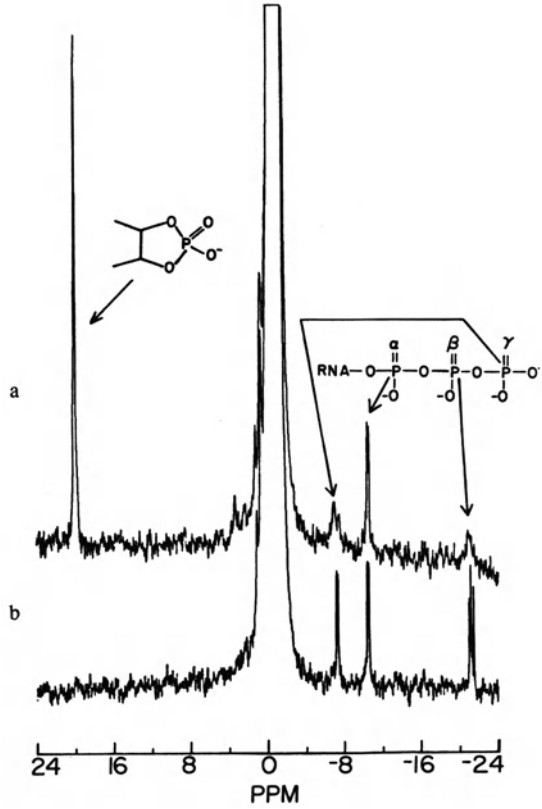


Figure 12. ^{31}P NMR spectra of yeast 5S rRNA with no added Mg^{2+} in 10 mM sodium cacodylate, pH meter reading 7.0, 1 mM EDTA, 100 mM NaCl, and 88:12 $H_2O:D_2O$. (a) Partially degraded yeast 5S rRNA (13 mg/ml = 0.3 mM). Note the large cyclic-phosphate peak at 20 ppm, resulting from cleavage of the phosphodiester backbone. (b) Native yeast 5S rRNA (21.9 mg/ ml = 0.55 mM). Note the 5′-triposphate signals and the absence of cyclic-phosphate. From K. O. Burkey, unpublished results.

A-helical segments. As the molecule relaxes its shape at higher temperature or lower [$MgCl_2$], these peaks merge back into the main phosphodiester resonance. Thus, ^{31}P offers simultaneous probes of both the main A-helical segments and other strained phosphates.

Finally, the ^{31}P NMR spectra of yeast 5S RNA (Figures 12–14) and that of *E. coli*, *B. licheniformis* $tRNA^{phe}$ (Salemink *et al.*, 1981), not shown here, show that 5S rRNAs present fewer downfield-shifted ^{31}P resonances than do typical tRNAs. Thus, it appears that the 5S tRNA phosphodiester backbone is less distorted from A-helicity than is the tRNA polynucleotide chain.

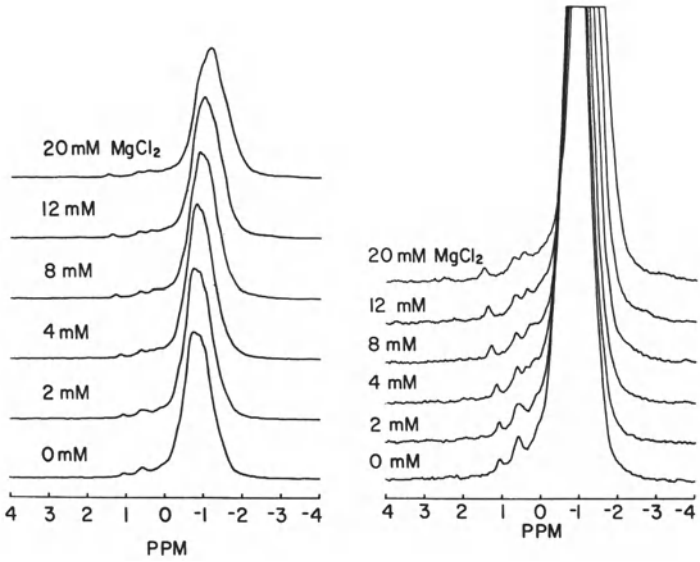


Figure 13. Titration of yeast 5S rRNA (0.55 mM) phosphodiester ^{31}P NMR spectrum at 34 °C with a solution of 2 M $MgCl_2$, 10 mM sodium cacodylate at pH meter reading 7.0, 1 mM EDTA, 100 mM NaCl, in 88:12 $H_2O:D_2O$ stored over chelating resin. Left: normal vertical scale, to show the main peak region. Right: expanded vertical scale, to show minor peaks. From K. O. Burkey, unpublished results.

2.3.2. ^{19}F

Although fluorine is obviously not present in native RNA, bacteria grown in a minimal medium containing 5-fluorouracil (5FU) will incorporate 5FU at virtually all primary sequence uracil position of 5S rRNA (Marshall and Smith, 1977; 1980) and tRNA (Harkin and Horowitz, 1987; Gollnick *et al.*, 1987). ^{19}F NMR of 5FU-labeled 5S rRNA therefore offers a means for simultaneous monitoring of all of the uracil sites.

Because ^{19}F chemical shifts span an extraordinarily wide range, 8–10 different uracil environments can be resolved in the ^{19}F NMR spectrum of 5FU-labeled *E. coli* 5S rRNA at 35 °C (Figure 15). At 72 °C, all of the ^{19}F resonances collapse to a single chemical shift near that of free 5-fluorodeoxy-UMP, indicating that all of the fluorinated uracils share a common environment (denatured, exposed to solution). Inspection of the room-temperature ^{19}F spectrum then shows that only 25–35% of the 5FU residues are exposed to solution when the molecule is in its native conformation (Marshall and Smith, 1977).

A difficulty with ^{19}F (or ^{31}P) as an RNA NMR label is that for a molecule the size of 5S rRNA, chemical shift anisotropy (CSA) contributes significantly to line-broadening at magnetic fields greater than ~2 tesla.

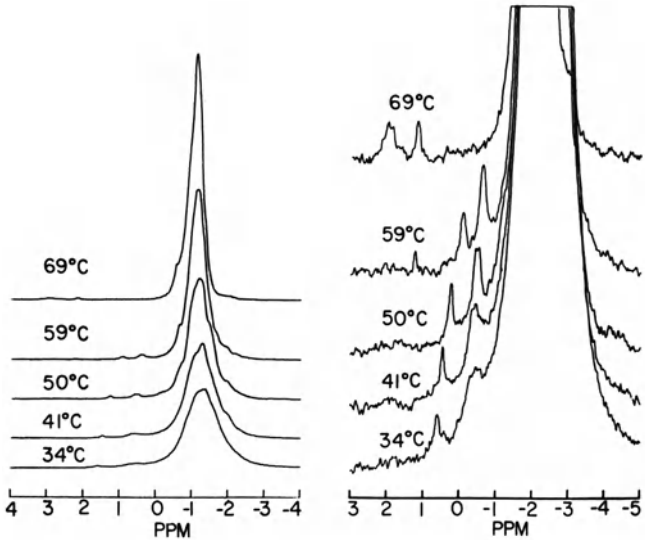


Figure 14. ^{31}P NMR spectra of yeast 5S rRNA (22.3 mg/ml = 0.56 mM), in the buffer of Figure 12 with 20 mM $MgCl_2$ added, as a function of temperature. Left: normal vertical scale, to show the main peak region. Right: expanded vertical scale, to show minor peaks. From K. O. Burkey, unpublished results.

Since CSA broadening scales as the square of magnetic field strength, whereas resolution increases linearly with field strength, resolution may actually grow worse as the magnetic field strength increases above ~6.35 tesla (~254 MHz for ^{19}F).

Because the ^{19}F NMR frequency at ~6.35 tesla is approximately the same as the reciprocal of the rotational correlation time for 5S rRNA, a heteronuclear NOE experiment in which ^{19}F peak height is observed in the presence or absence of saturation of the ^{1}H NMR spectrum is perfectly tuned to detect motional flexibility at each of the labeled uracil positions. In that experiment, saturation of the ^{1}H spectrum eliminated virtually all of the ^{19}F resonances (Marshall and Smith, 1980), providing the first evidence (NOE ≈ -1.0; see Figure 5) that the 5S rRNA molecule was relatively rigid (rotational correlation time $\geq$ 10 nsec) throughout, in contrast to early (misinterpreted) optical results. This experiment was a key step in showing that subsequent homonuclear ^{1}H NOE experiments would be feasible.

2.3.3. ^{13}C

E. coli 5S rRNA labeled (by bioincorporation, as for 5FU) with 4-^{13}C-uracil gave the ^{13}C NMR spectrum (Hamill *et al.*, 1978) shown in Figure

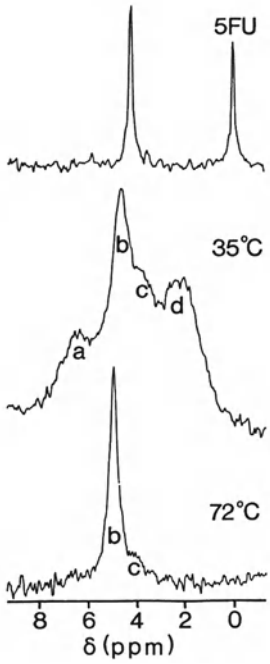


Figure 15. 254 MHz ^{19}F NMR spectra of: approximately equimolar (each 7.5 mM in 10 mM sodium cacodylate D_2O buffer, pH meter reading 7.0) mixture of 5-fluorouracil (5FU) and 5-fluoro-deoxy-UMP (top); 5FU labeled *E. coli* 5S rRNA (0.3 mM in 10 mM sodium cacodylate D_2O buffer, pH meter reading 7.6) at 35 °C (middle) and 72 °C (bottom). No proton decoupling was employed in any of the spectra. Chemical shifts are referenced to 5FU as zero. All spectra were phased identically to the top spectrum. Note that the various 5FU-5S rRNA ^{19}F NMR environments seen at 35 °C collapse to a common (unfolded) environment at 72 °C. Reproduced, with permission, from Marshall and Smith (1977).

16. As for fluorine-labeled 5S rRNA, the ^{13}C NMR spectrum revealed about eight resolved peaks corresponding to different uracil environments, and was interpreted to show that 75–95% of the uracils in *E. coli* 5S rRNA are involved in secondary or tertiary hydrogen-bonding. Compared to ^{19}F as a biolabel, ^{13}C has the advantages that (1) the labeled RNA is functionally identical to unlabeled native RNA, (2) the ^{13}C NMR linewidths are narrower than those of ^{19}F, and (3) chemical shift anisotropy does not significantly broaden the resonances at typical magnetic field strengths. However, ^{13}C labeling is quite expensive.

2.3.4. ^{15}N

^{15}N labeling of 5S rRNA offers several unique advantages. Most important, incorporated uniformly in all four bases, ^{15}N labeling allows for base type identification in proton NMR spectra. First, since each imino hydrogen-bond proton is directly bonded to a nitrogen, each ^{1}H imino downfield resonance is split into a doublet by scalar coupling ($J_{H-N} \sim 87$ Hz), as shown in Figure 17. Furthermore, the N1 of guanine and N3 of uracil occur in the ^{15}N NMR spectrum in two regions, each ~4 ppm wide, separated by ~16 ppm. The ^{15}N of uracil N3 resonates at slightly higher frequency in an

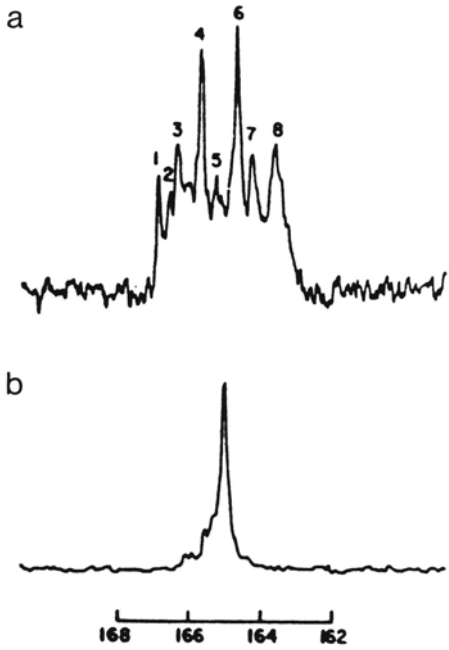


Figure 16. 25 MHz ^{13}C NMR spectra of ^{13}C-4-uridine carbons (46% isotopic enrichment) of *E. coli* 5S rRNA (68 mg/ml, 0.04 M $MgCl_2$, 5 mM dithiothreitol, 2 mM EDTA, 0.2 mM proteinase K, 2% dioxane, pH 7.4) at 37 °C (a) and 75 °C (b). Chemical shift is referenced to TMS as zero. The eight resolved low-temperature ^{13}C environments collapse to a single (unfolded) environment at 75 °C. Reproduced, with permission, from Hamill *et al.* (1978).

A·U pair than in a G·U pair (Gonnella *et al.*, 1982; Kime, 1984a). Thus, if the ^{15}N resonances from the U or G regions are selectively saturated, the corresponding ^{1}H NMR doublets will collapse, and the ^{1}H base-pair type can be identified according to which ^{15}N chemical shift was irradiated (Kaun *et al.*, 1982; Griffey *et al.*, 1982a; 1982b; 1983; Kime, 1984a), as shown in Figure 18. In this way, it has proved possible to identify all of the base-pair imino hydrogen-bond resonances from an RNase A cleavage fragment (Douthwaite *et al.*, 1979) of *E. coli* 5S rRNA (Kime, 1984a).

The ^{15}N-^{1}H shift correlation method is complementary to the proton homonuclear NOE technique described above for identification of base-pair type. However, the ^{15}N-^{1}H method detects imino protons exchanging slowly with solvent whether they are base-paired or not. Even for hydrogen-bonded imino base-pair protons, the ^{15}N-^{1}H method cannot distinguish between guanine imino N1 protons from G·U and those from G·C pairs. In contrast, the proton homonuclear NOE experiment does not typically detect slowly exchanging unpaired imino protons, but can usually distinguish between each of the three base-pair types, A·U, G·U, and G·U.

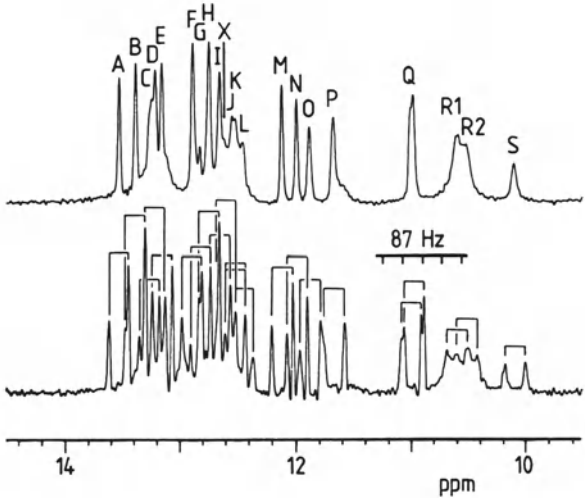


Figure 17. 500 MHz proton low-field NMR spectra of 1.2 mM *E. coli* rrnB 5S rRNA fragment 1 (100 mM KCl, 5 mM cacodylate, 4 mM $MgCl_2$, pH 7.1) at 303 °K. (top) native RNA. (bottom) ^{15}N-enriched RNA. Note that the imino protons are split (~87 Hz) into slightly asymmetrical doublets. Reproduced, with permission, from Kime (1984a).

The ^{15}N-^{1}H shift correlation method has proved useful for at least three additional purposes. First, in the ^{1}H NMR spectrum of a mixture of ^{15}N labeled RNA and protein(s), the method can obviously distinguish RNA protons from protein protons. Such experiments are especially useful for identifying new resonances specific to the RNA–protein complex, as in the complex formed between *E. coli* 5S rRNA fragment 1 and L25 protein (Kime, 1984b). Second, in hybrid 5S rRNA duplexes formed by combining an ^{15}N-labeled strand with an unlabeled complementary strand, the method can tell us which proton resonance arises from which strand, thereby facilitating resonance assignments (Gewirth *et al.*, 1987). Third, the method has been used to assign which imino proton is which in the ^{1}H NMR spectrum of a G·U base-pair (Griffey *et al.*, 1982b).

In summary of NMR from nuclei other than ^{1}H, ^{31}P is present at 100% natural abundance, but its NMR spectrum is so highly overlapped as to offer little assignable information. Moreover, in view of the relatively poor inherent ^{31}P signal-to-noise ratio, a large amount of sample (>30 mg per experiment) may be required. ^{19}F- and ^{13}C-labeling were historically important, but no longer justify the major effort (800-liter fermentation to yield a single ^{19}F NMR sample) and expense required for their bioincorporation. The effort and expense of ^{15}N-labeling, on the other hand, may well be warranted in view of its use for assigning proton NMR spectra in experiments conducted at the signal-to-noise ratio of proton NMR.

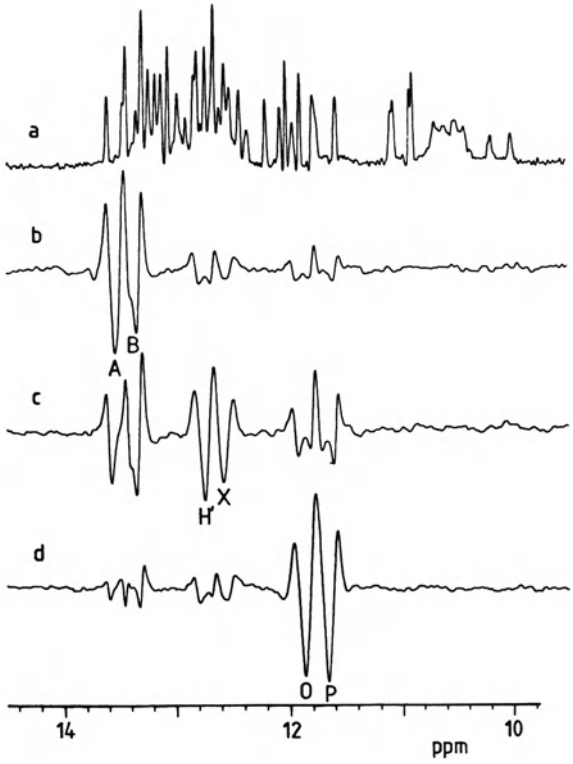


Figure 18. 500 MHz proton NMR difference (presence minus absence of irradiation) spectra (b, c, d) of the ^{15}N-enriched *E. coli* rrnB 5S rRNA fragment 1 (see Figure 19) in the presence of ^{15}N decoupling of uracil N3, along with Figure 17b shown for reference (a). Selective irradiation at 100 Hz intervals spanning the uracil N3 ^{15}N NMR frequency range identifies the uracil ^{15}N3H protons. Similarly, irradiation (not shown) at intervals spanning the guanine N1 ^{15}N frequency range reveals the guanine ^{15}N1H protons. Reproduced, with permission, from Kime (1984a).

3. STRUCTURAL ANALYSIS OF SPECIFIC 5S rRNAs

The various-sized ribosomal RNAs are listed in Table 2. Prokaryotic 5S rRNA binds to proteins L5, L18, and L25 of the large ribosomal subunit. Because eukaryotic 5S rRNA does not bind to those proteins, whereas eukaryotic 5.85 rRNA recognizes L18 and L25 from *E. coli* (Wrede and Erdmann, 1977), it is widely believed that eukaryotic 5.8S rRNA has the same function in eukaryotes as 5S rRNA does in prokaryotes. Interestingly, 5.8S rRNA is homologous to the 5′-end (and 4.5S rRNA is homologous to the 3′-end) of bacterial 23S rRNA. Thus, the detailed secondary structural analysis of small ribosomal RNAs from even a few organisms offers wide applicability to rRNAs from other species. With that motivation, we now

TABLE 2
RNAs from Small and Large Ribosomal Subunits[a]

Source	Small subunit	Large subunit
Mitochondria (mouse)	12S (956)[b]	16S (1582)[b] 5S ?
Bacteria (*E. coli*)	16S (1541)[b]	23S (2904)[b] 5S (120)[b]
Chloroplasts (*Zea mays*)	16S (1491)[b]	23S (2992)[b] 4.5S (95)[c] 5S (120)[c]
Eukaryotes (yeast)	18S (1789)[b]	28S 5.8S (158)[c] 5S (121)[c]

[a] Modified from Liljas (1982).
[b] Bracketed number is the number of nucleotides in each RNA.
[c] From Erdmann and Wolters (1986).

consider the structural conclusions based on proton NMR of a few 5S RNAs and one 5.8S RNA.

3.1. *E. coli* 5S rRNA and Its Enzyme-Cleaved Fragments

Although the first ^{1}H NMR spectrum (continuous-wave, at 300 MHz) of *E. coli* 5S rRNA was obtained 15 years ago (Kearns and Wong, 1974), the first ^{1}H NMR spectrum (Fourier transform, 400 MHz) at resolution sufficiently high to resolve individual base-pair imino protons was not obtained until 1980 (Burns *et al.*, 1980). Even better resolution became possible when 500 MHz ^{1}H FT-NMR instruments became widely available ~1982.

As noted in Section 2.2.7, the single most useful adjunct for ^{1}H NMR spectral peak assignments is the isolation of 5S rRNA fragments generated by means of controlled digestion by specific ribonuclease(s). The fragment ^{1}H NMR spectrum contains fewer and narrower (hence, better-resolved) imino proton resonances than that of intact 5S RNA. In order to show that the fragment retains the same conformation as in the intact 5S rRNA, it is necessary to show that both the imino proton resonance chemical shifts and homonuclear proton NOE patterns for the fragment match those of intact 5S rRNA.

E. coli 5S rRNA fragment 1 (bases 1–11, 69–87, and 89–120), shown in Figure 19, was the first 5S rRNA fragment used for ^{1}H NMR spectral assignment (Kime and Moore, 1983a). The fragment was produced from a cloned 5S rRNA gene rrnB to give a product which can be considered

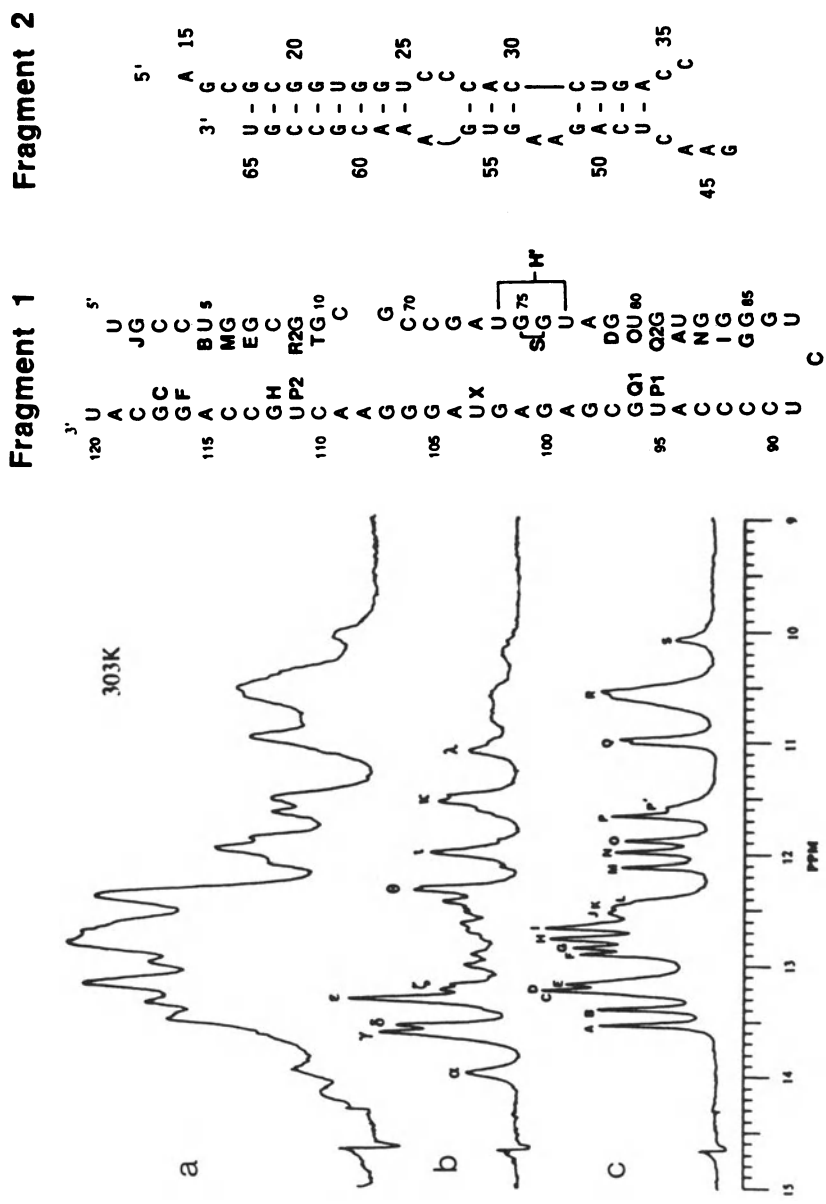


Figure 19. 490 MHz proton NMR spectra of *E. coli* rrnB 5S rRNA (right, top), fragment 2 (right, middle), and fragment 1 (right, bottom), taken under identical conditions: 303 K in 5 mM $MgCl_2$, 0.1 M NaCl, and 5 mM MES at pH 6.0. Assignments of downfield base-pair proton resonances are listed at the right for fragment 1 and for fragment 2 in conformation 1. Reproduced, with permission, from Gewirth and Moore (1987b) and Leocles and Moore (1986).

nearly identical to 5S rRNA isolated from ribosomes (Kime *et al.*, 1984). In the ^{1}H NMR spectrum of fragment 1 (Figure 19c), each of the resonances in the fragment can be found (at the same chemical shift) in the spectrum of the intact 5S rRNA (Figure 19a). Thus, the fragment appears to maintain its conformation when the rest of the RNA molecule is cleaved away, and can thus be regarded as a 5S rRNA domain by analogy to the domain concept for proteins (Kime *et al.*, 1983b).

Most of the downfield proton NMR spectrum of *E. coli* 5S rRNA fragment 1 has been assigned by homonuclear proton NOEs. In particular, the demonstration of NOE connectivity of G·C–G·C–A·U–G·U and $(G{\cdot}C)_3$–A·U–$(G{\cdot}C)_3$ of fragment 1 (Kime and Moore, 1983a) provided the first direct evidence for the existence of the proposed molecular stalk and prokaryotic loop in *E. coli* 5S rRNA: namely, $G_{81}{\cdot}U_{95}$–$U_{82}{\cdot}A_{94}$–$G_{83}{\cdot}C_{93}$–$G_{84}{\cdot}C_{92}$ and $G_2{\cdot}C_{118}$–$C_3{\cdot}G_{117}$–$C_4{\cdot}G_{116}$–$U_5{\cdot}A_{115}$–$G_6{\cdot}C_{114}$–$G_7{\cdot}C_{113}$–$C_8{\cdot}G_{112}$, respectively (Kime *et al.*, 1984). By this assignment alone, many proposed 5S rRNA secondary base-pairing schemes could be eliminated, leaving the modified Fox and Woese (Garrett *et al.*, 1981) and Luoma and Marshall (Luoma and Marshall, 1978a; 1978b) models. The same NOE connectivities were also found in intact 5S rRNA, including a G·U–G·C–G·C connectivity that is absent in the fragment.

It is fair to note that the interpretations (Kime and Moore, 1983a, 1983b; Kime *et al.*, 1984) are not definitive, because several spectral peaks appear to derive from more than one imino proton. Thus, the proposed assignments may represent a sufficient but not unique explanation.

The assignment of $(G{\cdot}C)_3$–A·U–$(G{\cdot}C)_3$ to the terminal stem base-paired segment poses the immediate difficulty (see Section 2.1.3b) that NOEs alone cannot resolve the twofold symmetry of that sequence. However, for an RNA A-helix in which an A·U pair is sandwiched between two G·C pairs, the amino proton of the G·C pair on the 5′ side of the U is closer to the uracil-N3H proton than is that of the G·C pair on the 3′side of the U. Thus, if the *E. coli* 5S rRNA terminal stem is assumed to take on the A-helix conformation, the G·Cs above and below the A·U can be distinguished by the relative distances between each of their imino base-pair protons (peaks F and M) and the imino proton of the central A·U (peak B). Although NOEs change with time following the initial preirradiation period, the initial NOE magnitude as a function of preirradiation period depends only upon the distance(s) to the nearest observed proton(s) from the irradiated proton (Noggle and Schirmer, 1971). For a given rotational correlation time and one known distance (0.29 nm between U-N3H and A-C2H protons in an A·U pair), other interproton distances can then be obtained from the time-dependent NOE experiment (0.38 nm for F to "B" and 0.45 nm M to B, respectively). Thus, the 5′-to-3′ sequence F-B-M could be inferred for the U-containing strand (Kime *et al.*, 1984).

The same authors also calculated the ring current shifts for the two possible assignment scheme and arrived at the same assignment. However, the best match between calculated and experimental shifts was for an RNA A′-helix (rather than the A-helix assumed for the NOE analysis), based on the method of Kearns and coworkers (Kearns, 1976). Moreover, the method of Kearns (1976) has been shown to give incorrect shifts in prior tRNA experiments, as noted earlier. Thus, ring current analysis cannot be considered sufficiently reliable as a method for peak assignment in RNAs of initially unknown conformation, as in this case.

A second *E. coli* 5S rRNA enzymatic cleavage fragment, fragment 2 containing bases 15–65, with more than one digestion site in the proposed single-stranded region (bases 35–44), gives the proton NMR spectrum shown in Figure 19b. For fragment 2, G·C–G·U–G·C, G·C–G·C, A·U–A·G, and two A·U–G·C base-paired segments were inferred from NOE connectivities. By comparison to the De Wachter model (De Wachter *et al.*, 1982) for this RNA segment, the A·U–A·G segment was assigned to A_{58}·U_{25}–A_{59}·G_{24}. The A·G pair identification was based on its resonance position (12.3 ppm) and an intrabase pair NOE at 7.51 ppm which was assumed to be an adenine-C2H proton. Definitive assignments for other short sequences could not be made from NOEs alone. The authors tentatively assigned all of the resonances based on ring current induced shift calculations. Of the two possible conformations for helices II and III in the De Wachter model, conformation 2, with only two stable A·U pairs, can be ruled out by the appearance of three A·Us in the downfield proton spectrum of fragment 2 (Leontis and Moore, 1986).

3.2. *B. subtilis* 5S rRNA

Escherichia coli is a Gram-negative bacterium. It is therefore interesting to consider 5S rRNA (116 nucleotide primary sequence) from *Bacillus subtilis*, a Gram-positive bacterium. Unfortunately, the downfield imino base-pair proton 500 MHz NMR spectrum of intact *B. subtilis* 5S rRNA exhibits extensive peak overlap, as shown in Figure 20. Fortunately, the proposed secondary structure of *B. subtilis* 5S rRNA is predicted to contain several G·U pairs, which facilitate spectral assignment by providing starting points for base-pair sequencing based upon homonuclear proton NOE connectivities: each G·U pair contributes two imino base-pair protons with characteristically strong mutual NOE, which makes a G·U pair easy to identify; G·U imino base-pair proton resonances are found in the otherwise relatively empty 10–12.5 ppm spectral band; and the expected total number of G·U pairs is small (about 3–4), so that assignment of a given G·U to its correct primary oligonucleotide sequence position is much simpler than for the more numerous A·U or G·C pairs.

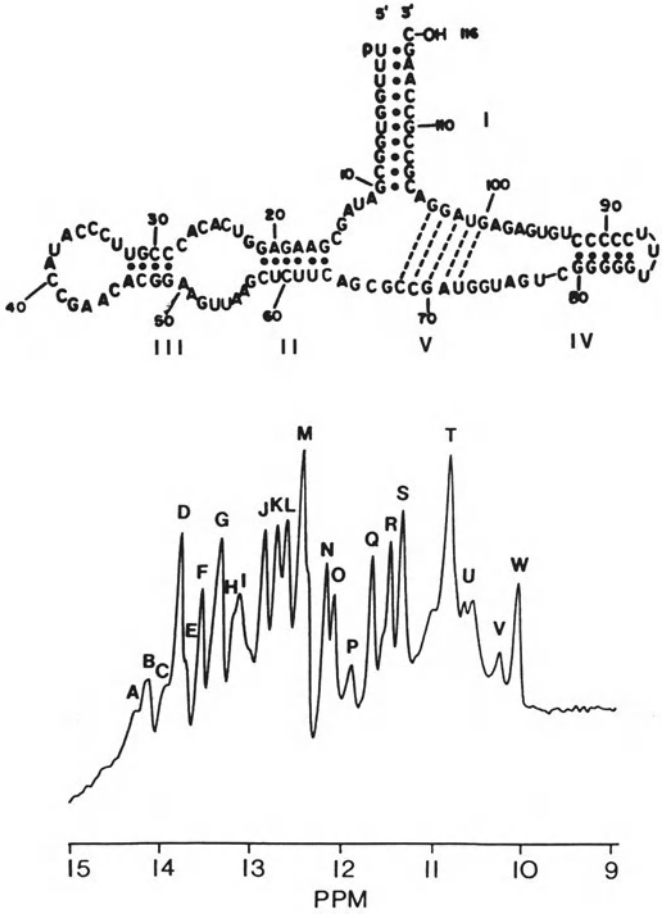


Figure 20. (Top) Fox and Woese model adapted to the *B. subtilis* 5S RNA primary nucleotide sequence. Helix V is depicted in dashed lines as proposed later (De Wachter *et al.*, 1982; Delihas *et al.*, 1984); (bottom). The 500 MHz ^{1}H NMR downfield spectrum of *B. subtilis* 5S rRNA (40 mg/ml, 10 mM cacodylate, 0.1 M NaCl, 0.1 mM EDTA, pH 7.0, 95:5 H_2O:D_2O) at 23 °C. Adapted from Chang and Marshall (1986b).

For a G·U base pair, the interpair NOE between a G·U imino proton and the imino proton of the base pair on the 5′ side of the G·U guanine is sometimes undetectable. Such "one-sided" NOE connectivity has been observed in several 5S rRNAs (Chang and Marshall, 1986b; Chen and Marshall, 1986, Kime and Moore, 1983a) and 16S rRNA (Heus *et al.*, 1983), but not in tRNAs (e.g., Roy and Redfield, 1983; Hare and Reid, 1982b; Heerschap *et al.*, 1982). It is conjectured that the "wobble" G·U base pair distorts an RNA helix from the A-form, so as to produce a more pronounced

overlap between the G·U pair and the base-pair on its 3′-side than between the G·U pair and the base-pair on its 5′ side (Mizuno and Sundaralingam, 1978).

Figure 20 shows the ^{1}H downfield 500 MHz NMR spectrum of *B. subtilis* 5S rRNA, along with a proposed secondary structural base-pairing model. Three short G·U-containing base-pair sequences have been identified by means of ^{1}H homonuclear NOEs (Chang and Marshall, 1986b). In particular, one G·C–G·C–G·U base-paired segment could be assigned to either $G_4 \cdot C_{112}$–$G_5 \cdot C_{111}$–$U_6 \cdot G_{110}$ or $U_6 \cdot G_{110}$–$G_7 \cdot C_{109}$–$G_8 \cdot C_{108}$ of the terminal helix. (As noted earlier, NOE connectivities alone do not establish the sense, whether 3′ to 5′ or 5′ to 3′, of a base-paired segment.) Further assignments were less definitive because of extensive peak overlap, but the primary sequence locations for two other sequences, G·U–A·U–G·C and G·U–G·C, were tentatively assigned. The most important result of the *B. subtilis* NMR experiments was to demonstrate that the terminal helix previously demonstrated for Gram-negative *E. coli* is also present in Gram-positive *B. subtilis*. Thus, it appears likely that the terminal helix segment is universal to all bacterial 5S rRNAs.

3.3. Wheat Germ 5S rRNA and Its Enzyme-Cleaved Fragments

Figure 21 shows proposed secondary structures and 500 MHz proton downfield NMR spectra of wheat germ (*Triticum aestivum*) 5S rRNA and two RNase T1 cleavage fragments. For intact wheat germ 5S rRNA, two NOE-connected base-pair short sequences were determined: G·C–G·U, corresponding to peaks G-R/V; and G·C–A·U–G·C, corresponding to peaks K-C-G (Li and Marshall, 1986). Although the G·C–G·U segment has not yet been assigned, the other segment was determined to be $C_{18} \cdot G_{60}$–$A_{19} \cdot U_{59}$–$C_{20} \cdot G_{58}$ and provided the first direct evidence for the existence of the "tuned helix" segment of 5S rRNA secondary structure. The protocols and interpretations in this case are illustrative of NOE base-pair sequencing in RNA, and proceed as follows.

3.3.1. "Tuned" Helix

First, although NOE connectivities established base-pair sequences corresponding to peaks K-C-G and G-R/V, peak G contains two overlapping resonances. Thus, the imino proton from resonance G in K-C-G need not be the same as the imino proton from resonance G in G-R/V, as noted in Figure 7 and associated discussion. At increased temperature (35 °C), the G peak resolves into two resonances, G1 and G2. Further heating results in melting of peak G2 together with peak C at ~50 °C, whereas peaks G1 and R/V remain visible in the spectrum up to 55 °C. The heating experiments

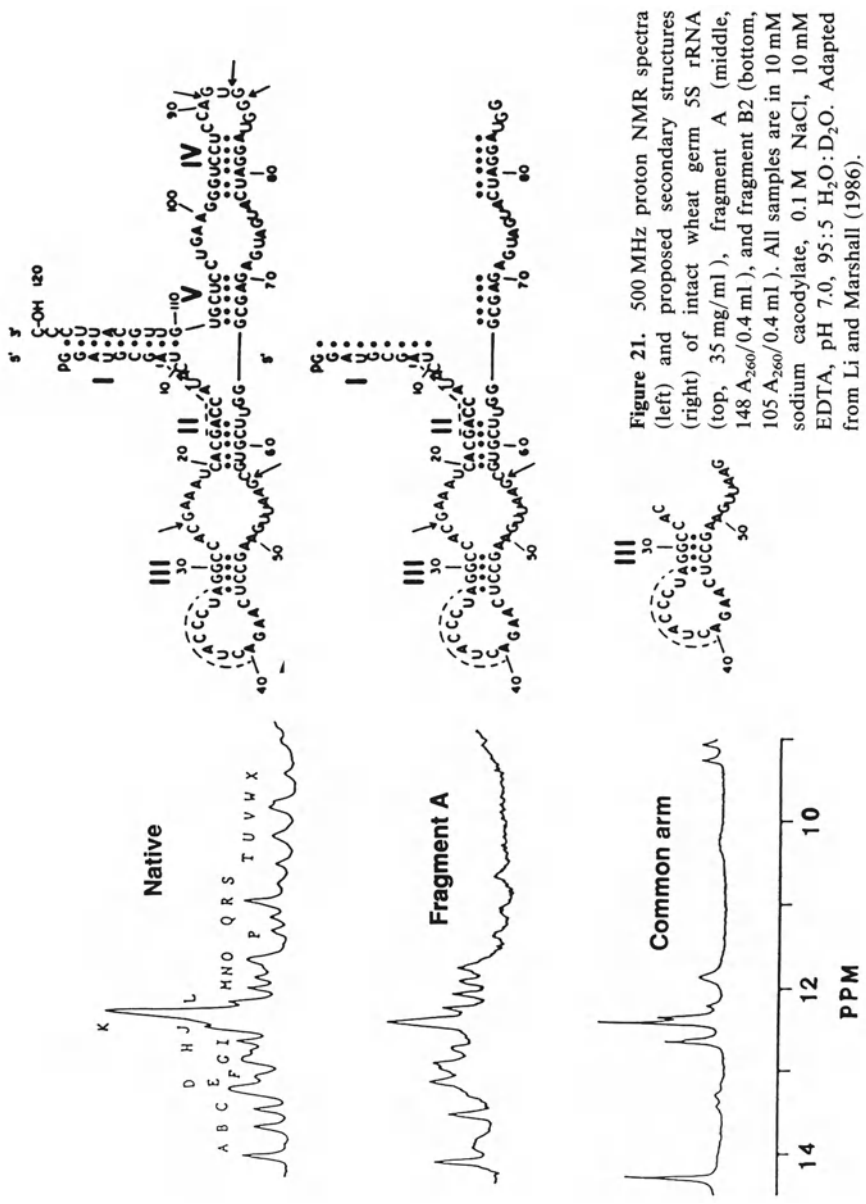


Figure 21. 500 MHz proton NMR spectra (left) and proposed secondary structures (right) of intact wheat germ 5S rRNA (top, 35 mg/ml), fragment A (middle, 148 A_{260}/0.4 ml), and fragment B2 (bottom, 105 A_{260}/0.4 ml). All samples are in 10 mM sodium cacodylate, 0.1 M NaCl, 10 mM EDTA, pH 7.0, 95:5 $H_2O:D_2O$. Adapted from Li and Marshall (1986).

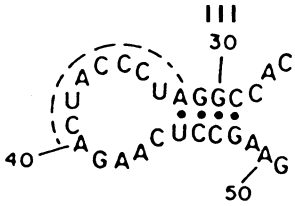
therefore strongly suggest that the two NOE-connected base-pair segments occur in different helical segments of the 5S rRNA (Li and Marshall, 1986).

Second, in the standard 5S rRNA base-pairing scheme shown in Figure 21a, there are two possible G·C–A·U–G·C sequences which could correspond to peaks K-C-G: namely, C_{18}·G_{60}–A_{19}·U_{59}–C_{20}·G_{58} and C_{105}·G_{70}–U_{106}·A_{69}–C_{107}·G_{68}. The two can be distinguished by considering RNase T1 cleavage fragment A (Figure 21b), consisting of primary sequence 1–86, 87, 88). Resonances K, C, G are clearly still observed (and are still NOE-connected, showing that the fragment retains its conformation after cleavage) in fragment A, even though the C_{105}·G_{70}–U_{106}·A_{69}–C_{107}·G_{68} can no longer be present. Finally, peak C cannot be in the "common arm" fragment (see Section 3.3.2). Thus, peaks K-C-G can be assigned to C_{18}·G_{60}–A_{19}·U_{59}–C_{20}·G_{58}, and constitute the first direct observation of the "tuned helix" segment of any 5S rRNA. (The assignment was consistent with ring current induced chemical shift calculations.) Furthermore, the four-stem 5S rRNA secondary structural model of Luoma and Marshall (Luoma and Marshall, 1978a; 1978b) can now be ruled out (at least for wheat germ 5S rRNA) because no such G·C–A·U–G·C segment occurs in primary sequence positions 1–88 of that model.

3.3.2. Base Pairing across the "Common Arm" Hairpin Loop

The proton 500 MHz NMR spectrum of wheat germ RNase T1 cleavage fragment B2, containing mainly primary oligonucleotides 26–51 (including the "common arm" helix) is shown in Figure 21c (Li *et al.*, 1987). Of the 6–7 resonances observed in the downfield region, two A·Us and at least four G·Cs could be identified from intra-base-pair NOE's. Inter-pair NOE connectivity between A·Us and two G·Cs were observed. The data cannot be accommodated by the standard base-pairing model (Luehrsen and Fox, 1981) unless two new base pairs are added, namely, G_{41}·C_{34}–A_{42}·U_{33} (Figure 22). In support of that assignment, G_{41} in wheat germ 5S rRNA is not attacked by RNase T1, suggesting that G_{41} may be base-paired in the secondary/tertiary structure; no additional downfield resonance was observed when $(rU)_5$ (which might be expected to bind to $A_{40}G_{41}A_{42}A_{43}$) was mixed with the "common arm" segment; ring current shift calculations are consistent with such an assignment; and comparative sequence analysis for eukaryotic 5S rRNAs conserves the G_{41}·C_{34}–A_{42}·U_{33} base-pairing. Alternative explanations such as spin diffusion, multiple conformers, and dimer formation were considered unlikely, as discussed at length by the authors (Li *et al.*, 1987).

The proposed G_{41}·C_{34}–A_{42}·U_{33} base-pairing shown in Figure 22 may constitute the most important single new conclusion resulting from proton NMR of 5S rRNAs, because it explains why Watson–Crick complementary



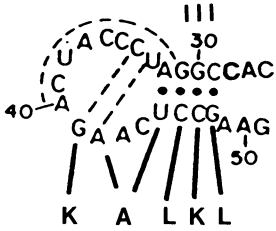


Figure 22. Fox and Woese model (top) and extended model (bottom) for the secondary structure of the common-arm portion of wheat germ 5S RNA. The new structure contains two more base pairs than the Fox and Woese model. Assigned proton NMR resonances are labeled as in Figure 21. Adapted from Li *et al.* (1987).

binding between the widely conserved AGAAC segment in the hairpin loop of the "common arm" of 5S rRNAs and the equally conserved GTΨC segment of tRNAs is not essential for protein synthesis (Pace *et al.*, 1982; Zagorska *et al.*, 1984).

3.4. Yeast 5S rRNA and Its Enzyme-Cleaved Fragments

Figure 23 shows proposed secondary structures adapted to the primary nucleotide sequences of 5S rRNAs from two yeasts, *Torulopsis utilis* and *Saccharomyces cerevisiae.* Although the two rRNAs differ at six primary nucleotide positions, only two are predicted to occur in a base-paired segment: $G_{18}{\cdot}C_{61}$ in *T. utilis* becomes $C_{18}{\cdot}G_{61}$ in *S. cerevisiae.* Consistent with that secondary structure, the 500 MHz proton NMR downfield spectra and NOE connectivity relations for both rRNAs are nearly identical, strongly suggesting a common secondary and tertiary structure for the two rRNAs (Chen and Marshall, 1986).

3.4.1. Terminal Helix

Figure 24 also shows proton 500 MHz NMR spectra of *T. utilis* 5S rRNA and two fragments derived directly or by reconstitution of RNase cleavage fragments. Proton homonuclear NOE connectivities of the intact 5S rRNA served to identify peaks K/L-A-C as a G·U-A·U-A·U segment

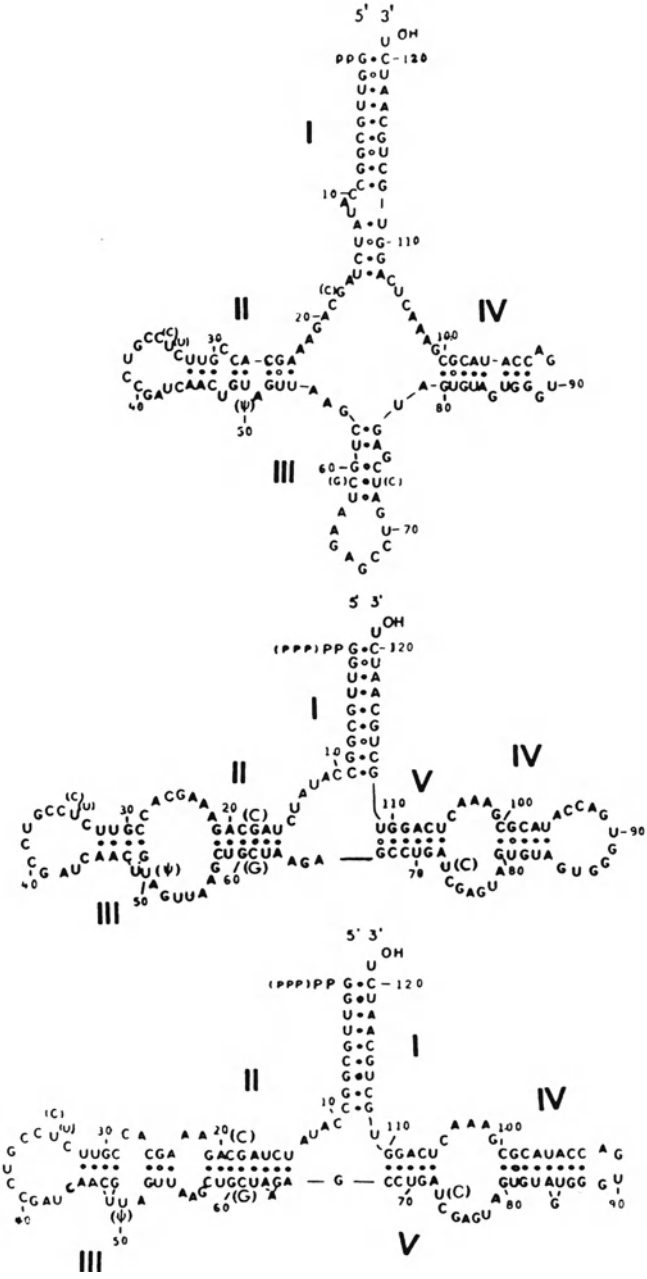


Figure 23. Three proposed 5S RNA secondary models for *T. utilis* and *S. cerevisiae* rRNA, (by: top, Luoma and Marshall; middle, Luehrsen and Fox; and bottom, Nishikawa and Takemura). Bases shown in parentheses represent changes in proceeding from *T. utilis* to *S. cerevisiae.* Reproduced, with permission, from Chen and Marshall (1986).

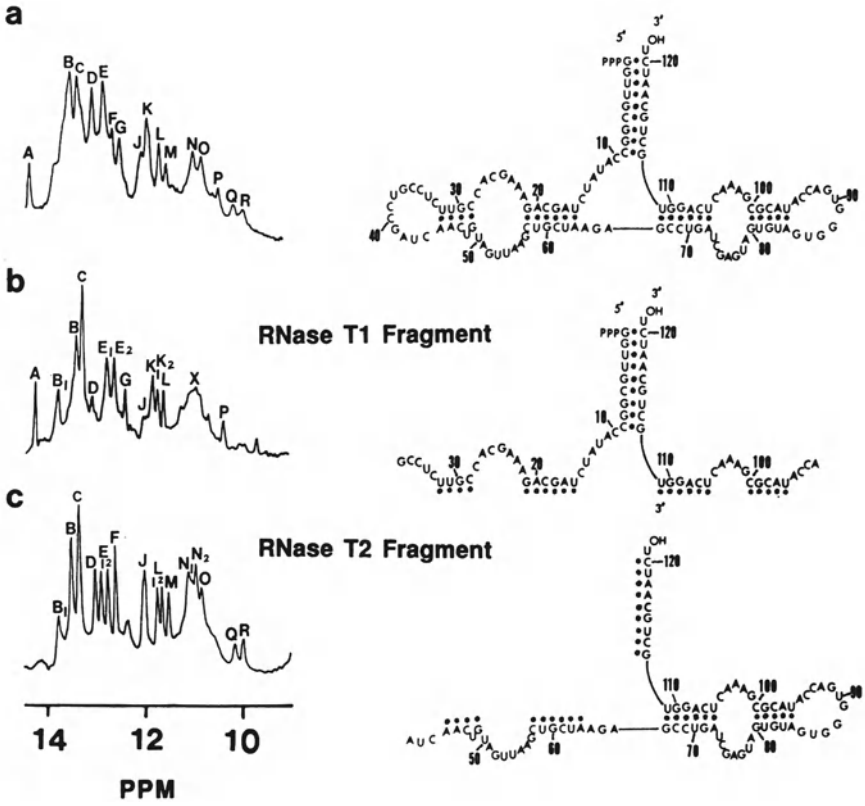


Figure 24. 500 MHz ^{1}H NMR low-field spectra of *T. utilis* 5S rRNA and its enzymatically cleaved fragments (left) along with their corresponding base-pairing schemes (right). (Top) intact 5S RNA (388 A_{260}/0.4 ml); (middle) 5S RNA RNase T1 fragment (210 A_{260}/0.4 ml); (bottom) 5S RNA RNase T2 fragment (200 A_{260}/0.4 ml). All samples are in the same buffer: 10 mM sodium cacodylate, 0.1 M NaCl, 1 mM EDTA, pH 7.0, 95:5 H_2O:D_2O. Reproduced, with permission, from Chen and Marshall (1986).

which could then be assigned to the only such sequence in the Luehrsen and Fox model, namely, $G_2 \cdot U_{119}$–$U_3 \cdot A_{118}$–$U_4 \cdot A_{117}$ of the terminal helix (Chen and Marshall, 1986). In addition, K/P-B and M/N-B were recognized as G·U–G·C sequences.

A terminal helix-containing fragment of *T. utilis* 5S rRNA was produced by reconstituting a 1:1 mixture of RNase T1 cleavage fragments containing primary oligonucleotides 1–37 and 92–121. Peak K in the intact 5S rRNA (Figure 24a) resolves into two resonances (K1 and K2) in the terminal helix fragment (Figure 24b), for which NOE-connected segments K1/L-A, K2/P-B, and possibly K1/L-B were shown. Because peak B melted at relatively low temperature and its chemical shift was [Mg^{2+}]-dependent, it was initially

assigned to be $G_1 \cdot C_{120}$, which was later shown by more direct spin-labeling experiments to be peak F (Lee and Marshall, 1987). This error illustrates the great difficulty of making definitive RNA primary sequence assignments based on NOEs of overlapped resonances. K2/P-B represents either $G_7 \cdot U_{114}$–$C_6 \cdot G_{115}$ or $G_7 \cdot U_{114}$–$G_8 \cdot C_{113}$: again, NOEs alone cannot decide the order of two dipole-connected base-pairs. Peak G was tentatively assigned to $G_5 \cdot C_{116}$ on the reasonable (but tenuous) basis that it melts together at about 30 °C with peak C which was definitively assigned to $U_4 \cdot A_{117}$.

As noted in Section 2.2.4., spin labeling at the 3′ terminus of yeast 5S rRNA established that resonances F and L are nearest to the 3′ terminus (Lee and Marshall, 1987). NOE experiments on intact *S. carlsbergensis* 5S rRNA demonstrated that resonances L/K-A correspond to a G·U–A·U segment. Moreover, resonance F contains at least one G·C and exhibits a weak NOE to resonance K. These three results combine to establish F-K/L-A NOE connectivity, which can then be assigned to $G_1 \cdot C_{120}$–$G_2 \cdot U_{119}$–$U_3 \cdot A_{118}$ in a Luehrsen and Fox model (Luehrsen and Fox, 1981). The otherwise conceivable alternative, $G_{26} \cdot C_{52}$–$G_{25} \cdot U_{53}$–$A_{24} \cdot U_{54}$ in helix III of the Nishikawa and Takemura model, can be ruled out because resonances F and L must arise from protons which adjoin the 3′ end of the molecule (where the spin label was covalently attached).

Finally, from additional NOE connectivities, the base-pair sequence corresponding to resonances F-K/L-A-C-HF-K/P-B was claimed and assigned (except for $C_9 \cdot G_{112}$) to the entire terminal helix (Lee and Marshall, 1987). However, this assignment must be viewed as tentative, since resonances K, A, and C each contain more than one proton intensity.

3.4.2. Helices IV and V

Figure 24c shows the 500 MHz proton downfield NMR spectrum of yet another fragment, bases 42–121, expected to contain helices IV and V of the Luehrsen and Fox model, and generated by means of RNase T2 digestion of *T. utilis* 5S rRNA. Because resonances M and N are missing from the spectrum of the RNase T1 fragment but remain in the RNase T2 fragment, NOE-connected resonances B-M/N2 were tentatively assigned to $G_{80} \cdot C_{100}$–$U_{81} \cdot G_{99}$, and L1/N1-C to $G_{67} \cdot U_{111}$–$C_{68} \cdot G_{110}$ (Chen and Marshall, 1986). It seems highly likely that such segments do in fact arise from helices IV and V. Again, however, the definitiveness of the assignment is reduced by (1) the presence of more resonances than could be identified or assigned—perhaps due to additional base-pairs (not present in native 5S rRNA) formed between the long single-stranded segments of the RNase T1 and T2 fragments: (2) the failure to identify the expected number of A·Us; and (3) the short lengths of the identifiable base-pair segments.

3.5. Other 5S rRNAs

3.5.1. *B. licheniformis* 5S rRNA

A proton 360 MHz correlation-mode NMR downfield spectrum of *B. licheniformis* 5S rRNA at 29 °C is shown in Figure 25 (Salemink *et al.*, 1981). In that early study, the experimental spectrum was compared to spectra simulated from ring current induced chemical shifts (Arter and Schmidt, 1976; Geerdes and Hilbers, 1979) for several proposed secondary base-pairing schemes. The number of base-paired hydrogen-bond protons was determined by both integration and simulation of the experimental spectrum. The above results, as well as temperature variation of the NMR spectrum, were consistent with the presence of the molecular stalk and prokaryotic loop base-paired segments of *B. licheniformis* 5S rRNA. Although some resonances were assigned to specific base pairs, no NOE data was available. Thus the assignments must be regarded as highly speculative.

3.5.2. *B. megaterium* 5S rRNA

A preliminary proton 500 MHz NMR spectrum of *B. megaterium* (another Gram-positive organism) is shown in Figure 26 (Kim and Marshall, unpublished). The primary base sequence of *B. megaterium* 5S RNA differs from that of *B. subtilis* by 10 base substitutions, of which seven are expected to occur in base-paired segments of the secondary structure, for a total of 5 base pairs between the two species (see Figure 26).

3.6. Yeast 5.8S rRNA

As noted in Table 2, 5.8S rRNA is found in the large subunit of eukaryotic ribosomes. The ~38 reported 5.8S rRNA primary base sequences

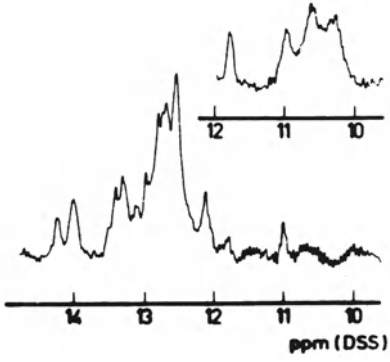


Figure 25. 360 MHz ^{1}H NMR spectrum of *B. licheniformis* 5S rRNA (1 mM, in 10 mM cacodylate, 10 mM EDTA, pH 7.0) at 40 °C. Estimated [Mg^{2+}] was 8 mM. The inset shows part of a spectrum recorded at 29 °C with the vertical scale expanded by a factor of 4. From Salemink *et al.* (1981).

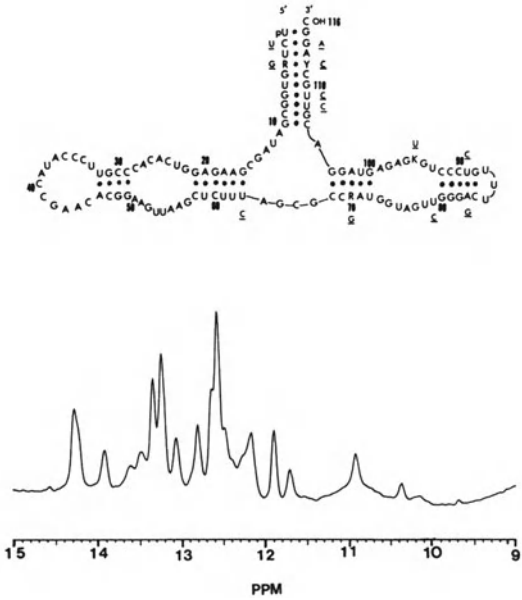


Figure 26. Top: Fox and Woese model of *B. megaterium* 5S rRNA. The nucleotide sequence inhomogeneity is indicated as follows: K = G or U, R = A or G, and Y = C or U. The underlined smaller letters along side show the corresponding nucleotide sequence of *B. subtilis* 5S rRNA. Bottom: 500 MHz ^{1}H NMR spectrum of *B. megaterium* 5S rRNA (28.7 mg/ml in 10 mM cacodylate, 0.1 M NaCl, 1 mM EDTA, pH 7.5, 95:5 H_2O:D_2O) at 23 °C. From Kim and Marshall, unpublished results.

contain 153–172 nucleotides and exhibit a high degree of sequence homology among themselves (Erdmann and Wolters, 1986) as well as to the 5′ end of prokaryotic 23S rRNA (Nazar, 1980). The biological function of 5.8S rRNA is unclear, although it may be involved in the binding of aminoacyl-tRNA to the ribosome (Wrede and Erdmann, 1077; Ulbrich *et al.*, 1979). 5.8S rRNA is known to bind to the terminal regions of 26S or 28S rRNA (Brimacombe and Stiege, 1985).

In ribosomal reconstitution experiments, yeast 5S rRNA did not bind to any of the 34 proteins from the 50S subunit of *E. coli* but yeast 5.8S rRNA did bind to proteins L18 and L25 (i.e., two of the three ribosomal proteins to which prokaryotic 5S rRNA binds), suggesting that 5.8S rRNA in eukaryotes may have the same function as 5S RNA in prokaryotes. In support of that view, the 5.8S rRNA/L18/L25 complex exhibits ATPase and GTPase activities, just as the *E. coli* 5S rRNA/protein complex does (Wrede and Erdmann, 1977). Moreover, (1) the eukaryotic 5.8S, 18S, and 28S rRNA genes are in the same transcription unit, separated from 5S rRNA genes, whereas prokaryotic 5S, 16S, and 23S rRNAs are transcribed together

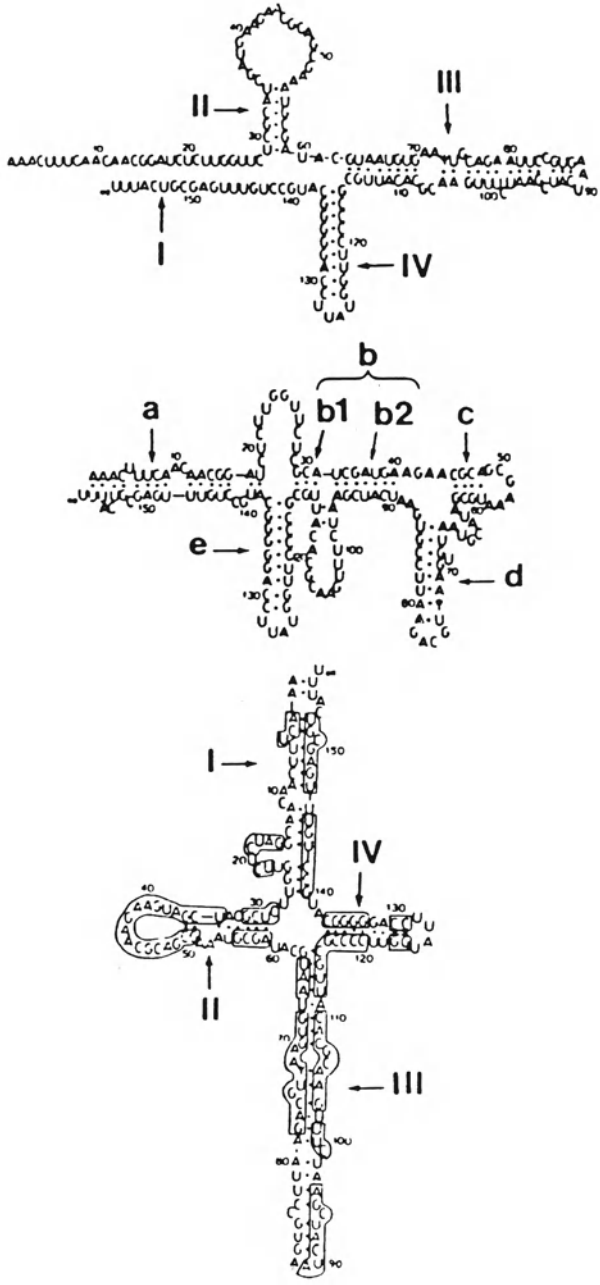


Figure 27. Three proposed secondary structural models for yeast 5.8S rRNA (by: top, Rubin; middle, Nazar; and bottom, Luoma and Marshall). Note that all of the models contain a GC-rich base-paired segment (bases #116–137) which is expected to be highly heat-stable. Reproduced, with permission, from Lee and Marshall (1986b).

(Freifelder, 1987); and (2) eukaryotic 5.8S rRNA and prokaryotic 5S rRNA are post-transcripts and have monophosphate 5′ termini, whereas eukaryotic 5S rRNA is a primary transcript and has a triphosphate 5′ terminus (Freifelder, 1987) as shown in Figure 12. Figure 27 shows three proposed secondary structures adapted to the primary nucleotide sequence of yeast 5.8S rRNA.

The most unique secondary structural feature of 5.8S rRNA is its GC-rich arm (bases 116–137 in the yeast 5.8S rRNA primary nucleotide sequence), containing a highly conserved sequence of several consecutive G·C base pairs. Because of the high thermodynamic stability of G·C pairs, the G·C-rich arm is expected to show the greatest thermal stability, −20.6 kcal/mol based on Tinoco rules (Tinoco *et al.*, 1973). As shown in Figure 28, Lee and Marshall (1986b) exploited that idea by heating yeast

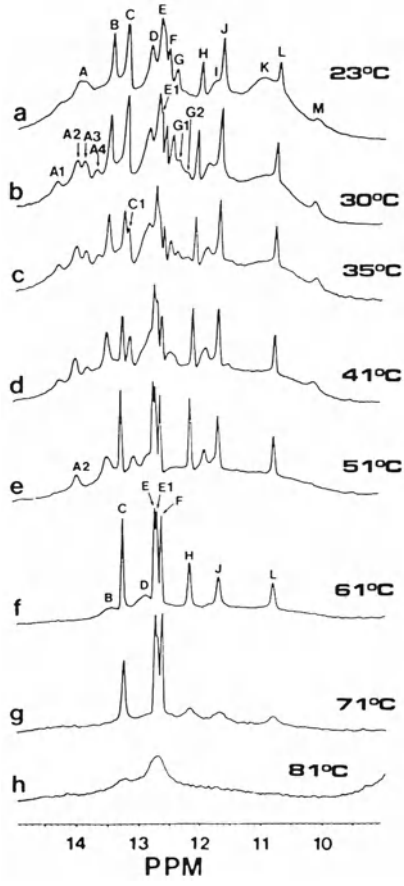


Figure 28. 500 MHz ^{1}H NMR spectra of yeast 5.8S rRNA (56 mg/ml in 10 mM sodium cacodylate, 0.1 M NaCl, 10 mM EDTA, pH 7.0, 95:5 $H_2O:D_2O$) at several different temperatures. The most striking feature is that several sharp resonances arising from the GC-rich base-paired segment persist even at 61 °C. Reproduced, with permission, from Lee and Marshall (1987).

5.8S rRNA to a temperature (~60 °C) high enough to melt away all other (proton 500 MHz) imino base-pair resonances, leaving a spectrum arising only from base pairs in the G·C-rich arm. Several of the G·C-rich arm base-pair imino proton resonances could be assigned by homonuclear proton NOEs. Their experiments constitute the first direct demonstration of the presence of a specific base-paired segment in a 5.8S rRNA.

4. NMR ANALYSIS OF rRNA STRUCTURAL DYNAMICS

4.1. Base-Pair Proton Exchange with Solvent

Once the imino base-pair proton NMR chemical shifts have been correctly assigned to specific base-pair sequences in the secondary structure of a given RNA, those resonances can then serve to report on changes induced at those sites by external perturbations (e.g., variation in temperature, solvent, or salt concentrations). Moreover, the rate of exchange of an imino base-pair proton with water provides a measure of accessibility (via steric clearance in a static RNA structure and/or flexing of the RNA molecule) to admit solvent molecules to a given base-pair hydrogen bond.

4.1.1. Exchange Rates from Real-Time Intensity Changes in NMR Spectra

An obvious means to observe the exchange of solvent-accessible imino base-pair protons with water is to compare ^1H NMR downfield spectra of an RNA equilibrated in H_2O and the same RNA quickly solvent-exchanged into D_2O. The exchange of a given imino proton with solvent can be monitored from the rate of decrease of its ^1H NMR intensity in D_2O, provided that the exchange lifetime is at least a few minutes. Although this direct method proved successful in identifying several long-lived slowly exchanging imino base-pair protons in tRNA (Johnston *et al.*, 1979; Figueroa *et al.*, 1983; Leroy *et al.*, 1985a), essentially all of the base-pair imino protons of *E. coli* 5S rRNA exchanged with solvent at rates too fast to observe directly (Leontis and Moore, 1986; Leontis *et al.*, 1986b).

4.1.2. Exchange Rates from NMR Relaxation Rates: Theory

Depending on the rate and mechanism of exchange with solvent, imino base-pair hydrogen-bond proton NMR peaks may be shifted, broadened, and/or reduced in intensity. The first-order rate constant for exchange can be determined as slow or fast compared to differences (in sec^{-1}) in chemical shift, $(1/T_1)$, or $(1/T_2)$, where T_1 and T_2 are the longitudinal (spin–lattice) and transverse (spin–spin) NMR relaxation times (Crothers *et al.*, 1974).

A simplified mechanism for intermolecular chemical exchange between RNA base-pair protons and water protons is given in equation (4), in which it is recognized that a base pair must first open before exchange can occur. If base-pair opening and closing is treated as a two-state equilibrium, then the chemical exchange process can be a three-state model (Crothers *et al.*, 1973; 1974; Tetelbaum and Englander, 1975a; 1975b; Hilbers, 1979),

$$\text{Closed} \underset{k_{\text{close}}}{\overset{k_{\text{open}}}{\rightleftharpoons}} \text{Open} \xrightarrow{k_{\text{ex}}} H_2O \tag{4}$$

in which k_{open} and k_{close} are the rate constants for the opening and closing of the base pair, and k_{ex} is the apparent first-order rate constant for chemical exchange with water. The rate constant k for the overall process then becomes

$$k = k_{\text{open}} \frac{k_{\text{ex}}}{k_{\text{close}} + k_{\text{ex}}} \tag{5}$$

The chemical exchange step is a proton transfer reaction whose rate constant is

$$k_{\text{ex}} = \alpha k_d B \frac{10(\Delta \text{p}K_a)}{1 + 10(\Delta \text{p}K_a)} \tag{6}$$

in which α is an accessibility factor ($0 \leq \alpha \leq 1$), k_d is the diffusion-limited collision rate constant, $\Delta \text{p}K_a$ is $\text{p}K_a$(acceptor) $-$ $\text{p}K_a$(donor), and B is base catalyst concentration (Eigen, 1964; Tetelbaum and Englander, 1975a; 1975b; Leroy *et al.*, 1985b; Leontis and Moore, 1986). For a free nucleotide in solution, $\alpha = 1$ and $k_d \approx 10^{10}\ \text{M}^{-1}\ \text{sec}^{-1}$. The $\text{p}K_a$ for an imino proton is ~9.3.

The first step (k_{open}) in the mechanism of equation (4) should be independent of buffer and pH, but not the second step. Thus, one can hope to distinguish proton exchange mechanisms which are rate-limited by the first or second stages of equation (4) (Tetelbaum and Englander, 1975a). The most interesting proton exchange parameter is k_{open}, which reports directly the lability of each 5S rRNA hydrogen bond whose base-pair imino proton resonance has been assigned. k_{open} can be determined from measurements of either T_1 or T_2, as described below.

In the slow-exchange limit, chemical exchange provides an additional pathway for nuclear spin relaxation rates ($1/T_1$) and ($1/T_2$) according to equation (7) (Johnston and Redfield, 1981)

$$\frac{1}{T_1} = \frac{1}{T_1'} + R_{\text{ex}}, \qquad \frac{1}{T_2} = \frac{1}{T_2'} + R_{\text{ex}} \tag{7}$$

in which T_1 and T_2 are the measured relaxation times, T_1' and T_2' are the "intrinsic" relaxation times for an imino hydrogen-bond base-pair proton in the absence of chemical exchange, and R_{ex} is the chemical exchange contribution to the NMR relaxation rate. T_1 can be obtained from saturation-recovery or inversion-recovery experiments (Johnston and Redfield, 1977; 1979; Patel, 1978; Pardi and Tinoco, 1982; Kearns, 1984). T_2 can be calculated from full NMR line width at half-maximum height, $\Delta\nu_{1/2}$, according to

$$\frac{1}{T_2} = \pi\Delta\nu_{1/2} \tag{8}$$

T_1 and T_2 measurements are then conducted at various buffer concentrations for each temperature. T_1' and T_2' can thus be determined when exchange is very slow (i.e., at low temperature and without buffer catalysts). More precisely, since T_1', T_2', and R_{ex} also depend on temperature (McConnell, 1958; Johnston and Redfield, 1981; Leroy *et al.*, 1985b), the relaxation rates must be fitted according to, e.g., equation (9) (Leontis and Moore, 1986)

$$\frac{1}{T_1} = \frac{C\eta(T)}{T} + R_{ex}(T) \tag{9}$$

in which C is a constant, $\eta(T)$ is sample viscosity, and T is temperature (K). In practice, R_{ex} is regarded as the overall exchange rate R. At each temperature,

$$\frac{1}{R} = \left(1 + \frac{k_{close}}{k_{ex}}\right) \Big/ k_{open} \tag{10}$$

Because k_{ex} is proportional to B, equation (10) can be expressed as

$$\frac{1}{R} = \left(1 + \frac{B_{eq}}{B}\right) \Big/ k_{open} \tag{11}$$

in which B_{eq} is the base concentration for which $k_{close} = k_{ex}$ (so that $R = k_{open}/2$). Thus, k_{open} can be obtained from a straight-line plot of $1/R$ vs. $1/B$ (Leontis and Moore, 1986).

With the above methods, imino proton exchange rates in the range $5 < R < 10\ \text{sec}^{-1}$ and $100 < R < 500\ \text{sec}^{-1}$ can be determined from T_1 and T_2 measurements, respectively (Johnston and Redfield, 1979). Furthermore, the activation energy E_a required to open a base pair can be determined from the slope of an Arrhenius plot of log k_{open} vs. $1/T$. From comparison

of the activation energies for several adjacent base pairs, the mechanism for base-pair opening may then be inferred. For example, if several base pairs open concertedly, then they should have similar k_{open} and E_a values; the converse is, however, not necessarily true (Pardi and Tinoco, 1982).

4.1.3. Exchange Rates from NMR Relaxation Rates: Experimental Results

The T_2 method has been applied to the determination of imino proton exchange rates for *E. coli* 5S rRNA fragment 1 (Leontis and Moore, 1986; Leontis *et al.*, 1986b), in which Tris buffer served as the base catalyst for the exchange step. Tris was found to broaden most but not all of the downfield resonances without affecting their chemical shifts. $1/T_2'$ was obtained from two imino proton resonances whose line widths were unaffected by changes in Tris concentration at the three temperatures studied (293, 303, 313 K). k_{open} values for several A·U and G·U pairs were determined (see Figure 29) to be 50–300 sec^{-1} at 303 K, depending on base-pair type and location in the primary nucleotide sequence. The results also showed that for 5 mM cacodylate buffer at pH 7.4, exchange in fragment 1 is in pre-equilibrium, i.e., is catalyst-limited. It should be noted that the choice of catalyst in the experiment can affect the results. For example, resonance S was insensitive to Tris but sensitive to spermidine concentration. But spermidine induced both broadening and chemical shift changes for several imino proton resonances, and even produced new downfield resonances, suggesting that spermidine alters the conformation of fragment 1, in

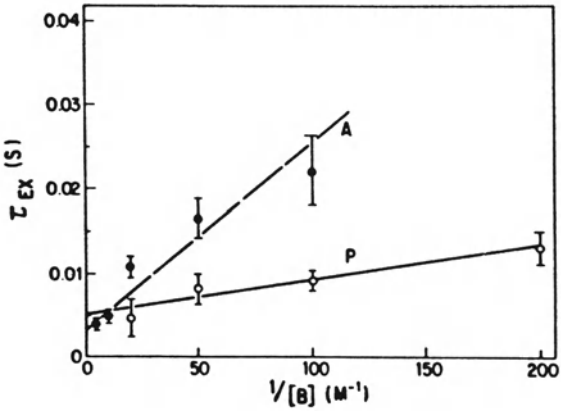


Figure 29. A plot of exchange lifetime (τ_{ex}) vs. reciprocal of Tris concentration for two 490 MHz ^{1}H nuclear resonances (A and P) of *E. coli* rrnB 5S rRNA fragment 1. The RNA sample (1 mM) was dissolved in 0.5 mM Tris, 0.1 M KCl, 5 mM $MgCl_2$, pH 7.4, 313 K. Reproduced, with permission, from Leontis and Moore (1986).

addition to any catalyzed increase in base-pair proton chemical exchange rates.

Based on T_1 measurements conducted on *E. coli* 5S rRNA fragment 1 in 5 mM cacodylate buffer in the absence of base catalyst, the exchange rates for several imino base-pair protons ranged from 0.5–15 sec^{-1} at 303 K. As seen from the Arrhenius plot of Figure 30 (Leontis and Moore, 1986), T_1 first increases with temperature (due to decrease in solution viscosity) and then begins to decrease at higher temperature (due to increasing chemical exchange rate). From the relatively linear high-temperature segments of Figure 30, the activation energies E_a for all of the imino protons appear to be essentially the same (~20 kcal/mol), suggesting that E_a is determined by the chemical exchange step, k_{ex} in equation (4), rather than the base-pair opening step, k_{open} in equation (4). Interestingly, further reports from the same research group (Leontis *et al.*, 1986b) listed $25 < E_a < 40$ kcal/mol, in which the T_1 values were in the order, U of G·U < G of G·U < U of A·U < G of G·C between 298–313 K. The discrepancy between the two sets of E_a values was not clearly explained.

4.2. Heat-Induced Melting

With increasing temperature, the lowfield resonances of 5S RNA generally first lose intensity and then broaden (e.g., Figure 31). At the threshold

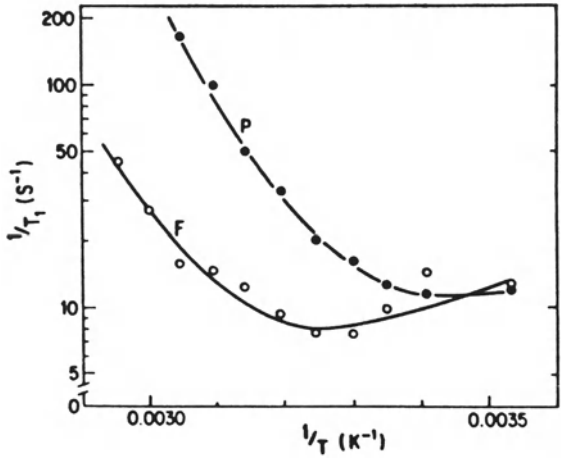


Figure 30. A plot of longitudinal proton NMR relaxation rate vs. reciprocal temperature for two 490 MHz ^1H nuclear resonances (P and F) of *E. coli* rrnB 5S rRNA fragment 1. The RNA sample was dissolved in 0.1 M KCl, 5 mM $MgCl_2$, 5 mM cacodylate, pH 7.2. The smooth curve represents a best fit of the data to equation (9). Reproduced, with permission, from Leontis and Moore (1986).

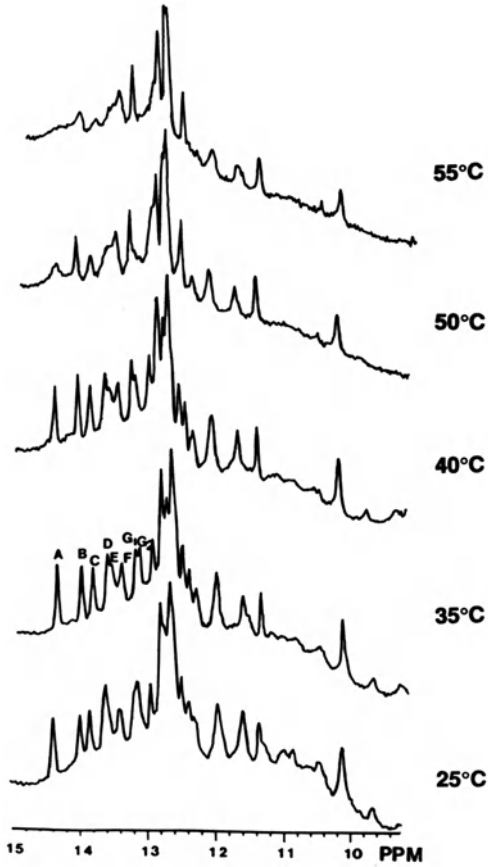


Figure 31. 500 MHz ^{1}H NMR spectra of wheat germ 5S rRNA (0.75 mM in 10 mM cacodylate, 0.1 M NaCl, 10 mM $MgCl_2$, pH 7.0, 95:5 $H_2O:D_2O$) at several temperatures. Adapted from Li and Marshall (1986).

temperature at which a resonance disappears, it is said to have "melted" (e.g., peak A at 55 °C in Figure 31). It is important to recognize that base-pair "melting" as determined from NMR is a kinetic process (proton exchange with solvent H_2O), whereas melting determined from calorimetry (Li and Marshall, 1985; Chang and Marshall, 1986a) or optical absorbance changes with temperature (Burkey *et al.*, 1983; 1984; Chang *et al.*, 1986a) is an equilibrium process (breaking and/or unstacking of base pairs) (Crothers *et al.*, 1973, 1974; Pardi and Tinoco, 1982). As might be expected, the lifetime for NMR base-pair proton chemical exchange corresponds directly with temperature-jump relaxation results (Crothers *et al.*, 1974; Kearns, 1977; Patel, 1978).

A base pair is half melted ($T = T_m$) in an optical absorbance measurement when, at equilibrium, half of the base pairs are broken (or unstacked)

and half are not. In NMR, on the other hand, an imino proton resonance might be broadened below detectability due to rapid exchange with solvent ("breathing" of base-pairs), even if that base pair were still 95% unmelted by the equilibrium criterion (i.e., only 5% of base pairs broken, on the average). Base-pair melting midpoint T_m values determined from optical melting experiments are therefore typically several °C higher than those obtained from proton NMR (Luoma *et al.*, 1980; Leontis *et al.*, 1986b). Although heat-induced melting monitored by NMR thus does not reflect the thermodynamic stability of base pairs, especially at high temperatures where chemical exchange is rapid, the NMR experiment offers the major advantage that the fate of each base pair whose imino proton resonance can be resolved can be monitored individually.

The imino base-pair proton NMR peak height can be reduced in two ways. First, the NMR peak height is directly related to the fraction of unbroken base pairs; thus, the peak area will decrease according to the fraction of base pairs broken. Second, the peak is broadened by rapid ($>30\ \mathrm{sec}^{-1}$) chemical exchange with solvent. Usually both effects contribute to peak height reduction on increase in temperature, but the NMR peak usually loses intensity before extensive broadening sets in (Crothers *et al.*, 1973; Hilbers, 1979).

Although it is generally presumed that tertiary base pairs melt (by kinetic or equilibrium criteria) at lower temperature than do secondary base pairs (Reid and Hurd, 1977; Hilbers, 1979; Robillard and Reid, 1979; Schimmel and Redfield, 1980), there are counterexamples (Reid *et al.*, 1979). The same is true for more rapid exchange of tertiary compared to secondary base-pair protons, based on T_1 measurements (Robillard and Reid, 1979; Hilbers, 1979).

Another working hypothesis for base-pair sequencing by proton NMR is that neighboring base pairs of the same type should melt at about the same temperature (Chen and Marshall, 1986; Li and Marshall, 1986). Although generally valid, this assumption can lead to incorrect base pair assignment, as pointed out by Chen and Marshall (1986) and Li and Marshall (1986). Finally, it is generally accurate to suppose that terminal base undergo relatively more rapid chemical exchange with solvent and thus "melt" (by NMR criteria) at lower remperature. This phenomenon is called the terminal "fraying effect" (Hilbers, 1979).

For most species studied so far, 5S rRNA melting appears to be reversible, as judged by both NMR and optical criteria (e.g., Wong *et al.*, 1972; Marshall and Smith, 1977). Thus, it appears that the 5S rRNA primary nucleotide sequence is sufficient to determine its secondary/tertiary folding pattern. Certain 5S rRNAs have been reported to change conformation after heating and cooling (e.g., Hamill *et al.*, 1978; Salemink *et al.*, 1981). In this context, an important practical point is that 5S rRNA can be degraded readily

at elevated temperatures in the presence of Mg^{2+} (e.g., Kearns and Wong, 1974; Salemink *et al.*, 1981).

4.3. Effects of Added Cations (Especially Mg^{2+})

Addition of Mg^{2+} to a 5S rRNA sample produces several related changes (usually continuous and reversible) in the proton downfield NMR spectrum (e.g., Figure 32):

1. The total integrated NMR downfield intensity may increase, due to an increased number of base pairs, in accordance with increased melting midpoint T_m from independent optical melting experiments.
2. Many imino base-pair proton resonances shift in position, presumably due to Mg^{2+}-induced conformational changes. In favorable

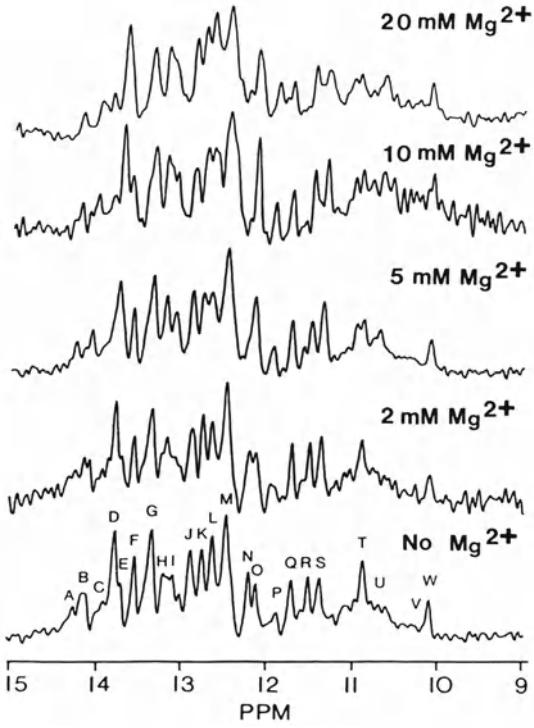


Figure 32. Effect of $[Mg^{2+}]$ upon the 500 MHz ^1H NMR down-field spectra of *B. subtilis* 5S rRNA. The lowermost spectrum is that of 0.4 mM RNA without Mg^{2+} in 10 mM cacodylate, 0.1 M NaCl, 1 mM EDTA, pH 7.0. Small aliquots of concentrated MgCl solution were added to give the indicated concentrations. Reproduced, with permission, from Chang and Marshall (1986b).

cases, such differential shifts can provide an aid in spectral assignment by resolving previously superimposed resonances (Li and Marshall, 1986), as illustrated in Figure 33.

3. NMR peak intensity or melting temperature may change (up or down) for an individual imino proton resonance.

Interest in the effects of Mg^{2+} binding arose initially from a search for alternate functional conformers of 5S rRNA, and from the speculation that 5S rRNA might have a direct role in protein biosynthesis based on switching between different conformations. In that context, three pairs of resolvably different 5S rRNA conformers have been demonstrated by non-NMR studies:

1. A native "A" form and a denatured "B" form in *E. coli* (Aubert *et al.*, 1968; Richards *et al.*, 1973; Erdmann, 1976; Weidner *et al.*, 1977; Weidner and Crothers, 1977; Noller and Garrett, 1979; Jagadeeswaran and Cherayil, 1980; Delihas *et al.*, 1984; Christensen *et al.*, 1985). The nonfunctional "B" form could be renatured into a functional "A" form by addition of Mg^{2+}.
2. High-temperature "A_H" and low-temperature "A_L" forms in *E. coli* (Delihas *et al.*, 1984; Christensen *et al.*, 1985; Kao and Crothers, 1980; Rabin *et al.*, 1983). The A_H form was stabilized by Mg^{2+}.

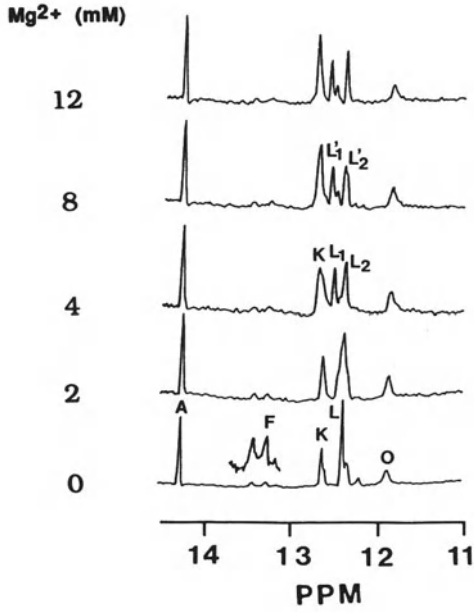


Figure 33. 500 MHz ^1H NMR spectra of the common-arm fragment of wheat germ 5S rRNA (105 A_{260}/0.4 ml) upon titration with Mg^{2+}. In the lowermost spectrum, Mg^{2+} was completely removed and the RNA was dissolved in a buffer containing 10 mM sodium cacodylate, 10 mM EDTA, pH 7.0. Note that at 4 mM [Mg^{2+}], peak L is split into L1 and L2, showing that the common arm contains more than four base pairs. Adapted from Li *et al.* (1987).

3. A high-Mg^{2+} "A" form and low-Mg^{2+} "A" form in rat liver (Muller *et al.*, 1985). Various proposed alternative 5S rRNA secondary structures based on primary nucleotide sequence comparisons have also been proposed (De Wachter *et al.*, 1982; 1984; Trifonov and Bolshoi, 1983).

Studies of *E. coli* 5S rRNA at different Mg^{2+} concentrations (Leontis *et al.*, 1986a; 1986b; Kime and Moore, 1982) concluded that the NMR spectral changes seen upon addition/removal of Mg^{2+} at low temperature are qualitatively different from the A_H-to-A_L transition. The B form of *E. coli* 5S rRNA has been shown to differ from the A form (Kime and Moore, 1982) and to contain at least 33 base pairs at 26 °C, as shown in Figure 34 (Burns *et al.*, 1980).

In oligonucleotides, both imino proton exchange kinetics and base pair lifetimes are found to be independent of Mg^{2+} concentration (Leroy *et al.*, 1985b). For tRNA, tertiary base pairs and terminal base pairs are believed to be more sensitive to Mg^{2+} than are base pairs in the interior of a helix (Reid and Hurd, 1977; Johnston and Redfield, 1977; Robillard and Reid, 1979; Schimmel and Redfield, 1980; Pardi and Tinoco, 1982). Although not definitive, a Mg^{2+} titration may be used to detect such base pairs.

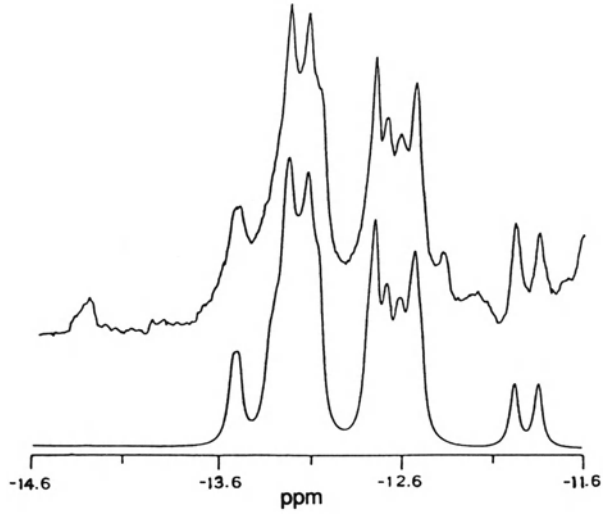


Figure 34. Experimental (A) and simulated (B) 400 MHz ^1H NMR spectrum of the denatured (B form) of *E. coli* 5S rRNA (33 mg/ml in 10 mM phosphate, 0.1 M NaCl, 1 mM EDTA, pH 6.5) at 26 °C. The simulated spectrum was constructed by coadding 31 Lorentzian lines of equal intensity and linewidth (30 Hz at half maximum height). Adapted from Burns *et al.* (1980).

5. CONCLUSION AND PERSPECTIVES FOR FUTURE WORK

NMR has proved to be the single most powerful technique for the investigation of secondary structure and dynamics of small RNAs in solution. Its strength over other techniques is that it can definitively determine secondary base-pair sequences and base-pair hydrogen-bond lability.

Several important secondary structural features have been established for both prokaryotic and eukaryotic 5S rRNAs: the terminal helix (*E. coli*, yeast), and intraloop base pairs which may explain the lack of binding of tRNA to the Watson-Crick complementary hairpin loop bases of the highly conserved "common arm" segment of 5S rRNA. In addition, the GC-rich arm of yeast 5.8S rRNA has been investigated.

A main point of this review is that correct interpretation of proton NMR high-resolution spectra still requires a variety of adjunct experiments designed to simplify (e.g., enzymatic cleavage) or differentially shift (e.g., Mg^{2+} addition) or broaden (e.g., temperature variation) the highly overlapped imino base-pair hydrogen-bond resonances. As outlined below, future work will be directed at such adjunct experiments and toward analysis of the biological function of 5S rRNA via its interaction with ribosomal proteins.

5.1. Sequence Variation

At least several consecutive base pairs in each of the five base-paired helical segments of 5S rRNA have been identified from proton 500 MHz NOEs. The most successful adjunct method for base-pair assignment and sequencing has been the direct production (Kime and Moore, 1983a; Chen and Marshall, 1986; Leocles and Moore, 1986; Li and Marshall, 1986; Li *et al.*, 1987) and/or reconstitution (Chen and Marshall, 1986) of fragments produced by selective enzymatic cleavage by RNase T1, T2 and A. The immediate task is to continue and complete the assignments of all secondary (and preferably tertiary) base pairs by a combination of the methods of Section 2.2. Spin labeling (Lee and Marshall, 1987) could prove especially useful in identifying tertiary base pairs, because the spin label will broaden spatially proximal base-pair proton resonances, whether the proximity arises from secondary or tertiary structure.

As noted in Section 2.2.3, one of the most definitive aids to assignment of specific base-pair imino proton resonances is to compare the spectra of two 5S rRNAs of slightly different known primary nucleotide sequences. For example, use has already been made of naturally occurring differences between 5S rRNAs from two closely related biological species (Chen and Marshall, 1986). However, it is becoming possible to produce artificial RNA sequences either by direct synthesis (Uhlenbeck and Lameron, 1977; Uhlenbeck *et al.*, 1982) or by genetic engineering (Gewirth and Moore,

1987b). Recently, two mutants of *E. coli* rrnB 5S rRNA were produced and examined by proton NMR (Gewirth and Moore, 1987). The wild type has a $G_{117}\cdot C_3$ base pair, and the two mutants have $G_{117}\cdot U_3$ and $A_{117}\cdot U_3$, respectively, as expected for mutations which conserve base pairing in the highly base-paired terminal RNA helix. As a result of the mutation, (1) the wild type resonance intensity decreases to 10–20% of its original value, due to transcription of the host gene; (2) new resonances can be assigned to the new (mutated) base pair; and (3) imino base-pair proton resonances of base pairs adjacent to the mutated base pair are shifted because of the different ring currents from the new base pair. Since the mutated terminal helix segment is no longer twofold symmetric (as the wild type 5S rRNA is), the original chemical shift assignments (based on ring-current arguments) for the terminal helix could also be confirmed from the spectra of the mutants.

5.2. Ribosomal Protein–RNA Interactions

Now that the resolved proton resonances of so many sites in the primary sequence of 5S rRNA have been assigned, it becomes possible to study the effect of ribosomal protein binding on the structure of 5S rRNA. Such experiments will be important for two reasons. First, it is vital to establish whether the structure of isolated 5S rRNA is changed by binding of the proteins (e.g., L5, L18, and L25 in *E. coli*) with which it is known to be associated in the ribosome. The protein-bound structure may also provide a better understanding of the 5S rRNA conformation *in vivo*. Second, the 5S rRNA–protein interaction provides a simple model for study of general RNA–protein interactions.

5.3. Two-Dimensional Proton FT-NMR

Although two-dimensional proton NMR has been applied with great success to the determination of three-dimensional structures of small proteins (Wüthrich, 1986) and short (≤30-mers) synthetic RNA and DNA duplexes (Wüthrich, 1986; Gronenborn and Clore, 1985), the methods have not worked well for RNAs as large (or larger than) tRNA (~80 bases). The 2D methods have failed mainly because of the broad (~50 Hz) and overlapping peaks, the short relaxation times, and the rapid polarization transfer from one proton to another. Nevertheless, the recent availability of relatively small base-paired 5S rRNA segments (either by direct synthesis or enzymatic cleavage of the native rRNA) should make 2D FT-NMR feasible. NOESY should be especially valuable for establishing NOE connectivities between partially overlapped resonances—a major difficulty for the one-dimensional FT-NMR NOE method.

Note added in proof. In Section 3.3.2, we described the observation of additional downfield resonances in the wheat germ 5S rRNA "common arm" fragment, and experiments which led to the proposal that two new base pairs may exist across the "common arm" hairpin loop. During the preparation of this review, we completed additional experiments (improved one-dimensional NOE as well as two-dimensional NOESY in H_2O) on the common arm fragment of wheat germ 5S rRNA. The "extra" downfield base-pair proton resonances can now be explained by the overlapping of two sets of resonances due to the presence of two distinct solution conformers of the common arm fragment (and its corresponding base-paired segment in intact 5S rRNA). In contrast, only a single conformation is manifested in the ^{1}H NMR behavior of all of the H6 and H5 pyrimidine and most of the H8/H2 purine protons under the same salt conditions. All of the imino protons, most of the amino protons, and the nonexchangeable protons near the helical segment have now been assigned. The difference between the two conformers identified is most likely due to a difference in helical tightness. Conformational features such as the presence of a cytidine-uridine (C·U) pair at the loop-helix junction and base stacking into the hairpin loop were also evaluated. Extension of base stacking into the common arm hairpin loop now appears to account for the observed lack of binding between the highly conserved single-stranded GTΨC segment of tRNA and the Watson–Crick complementary conserved GAAC segment of the 5S rRNA common arm (Wu and Marshall, 1989a; 1989b).

ACKNOWLEDGMENTS. The authors are deeply indebted to present and past co-workers whose efforts have produced much of what is known from NMR about small ribosomal RNAs: James O. Alben, Robert E. Bruce, Kent O. Burkey, Philip D. Burns, Shiow-Meei Chen, Lee-Hong Chang, Hsiau-Wen Huang, Jong-Hwa Kim, Kai Mon Lee, Shi-Jiang Li, Gregory A. Luoma, and James L. Smith. In addition, we thank K. O. Burkey and J.-H. Kim for generously making available their unpublished work on ^{31}P NMR of yeast 5S rRNA and ^{1}H NMR of *B. megaterium* 5S rRNA, respectively.

This work was supported by grants to Alan G. Marshall by the U.S. Public Health Service (NIH GM-29274 and NIH RR-01458) and The Ohio State University.

REFERENCES

Abdel-Meguid, S. S., Moore, P. B., and Steitz, T. A., 1983, *J. Mol. Biol.* **171**:207.
Alexander, S., 1961, *Rev. Sci. Instrum.* **32**:1066.
Arter, D. B., and Schmidt, P. G., 1976, *Nucl. Acids Res.* **3**:1437.

Aubert, M., Scott, J. F., Reynier, M., and Monier, R., 1968, *Proc. Natl. Acad. Sci. USA* **61**:292.
Bothner-By, A. A., 1979, in *Biological Applications of Magnetic Resonance* (Shulman, R. G., ed.), Academic Press, New York, pp. 177–219.
Brimacombe, R., and Stiege, W., 1985, *Biochem. J.* **229**:1.
Burkey, K. O., Marshall, A. G., and Alben, J. O., 1983, *Biochemistry* **22**:4223.
Burkey, K. O., Luoma, G. A., Alben, J. O., and Marshall, A. G., 1984, *Biochemistry* **23**:3652.
Burns, P. D., Luoma, G. A., and Marshall, A. G., 1980, *Biochem. Biophys. Res. Commun.* **96**:805.
Chang, L.-H., and Marshall, A. G., 1986a, *Biochemistry* **25**:1299.
Chang, L.-H., and Marshall, A. G., 1986b, *Biochemistry* **25**:3056.
Chang, L.-H., Burkey, K. O., Alben, J. O., and Marshall, A. G., 1984, *Biochemistry* **23**:3659.
Chen, S.-M., 1986, Ph.D. Thesis, The Ohio State University, Columbus, Ohio.
Chen, S.-M., and Marshall, A. G., 1986, *Biochemistry* **25**:5117.
Chou, S.-H., Hare, D. R., Wemmer, D. E., and Reid, B. R., 1983, *Biochemistry* **22**:3037.
Christensen, A., Mathiesen, M., Peattie, D., and Garrett, R. A., 1985, *Biochemistry* **24**:2284.
Clore, G. M., Kimber, B. J., and Gronenborn, A. M., 1983, *J. Mag. Res.* **54**:170.
Crothers, D. M., Hilbers, C. W., and Shulman, R. G., 1973, *Proc. Natl. Acad. Sci. USA* **70**:2899.
Crothers, D. M., Cole, P. E., Hilbers, C. W., and Shulman, R. G., 1974, *J. Mol. Biol.* **87**:63.
Delihas, N., Anderson, J., and Singhal, R. P., 1984, *Prog. Nucl. Acids Res.* **31**:161.
Derome, A. E., 1987, *Modern NMR Techniques for Chemistry Research*, Pergamon Press, New York.
De Wachter, R., Chen, M.-W., and Vandenberghe, A., 1982, *Biochimie* **64**:311.
De Wachter, R., Chen, M.-W., and Vandenberghe, A., 1984, *Eur. J. Biochem.* **143**:175.
Douthwaite, S., Garrett, R. A., Wagner, R., and Feunteun, J., 1979, *Nucl. Acids Res.* **6**:2453.
Early, T. A., Feigon, J., and Kearns, D. R., 1980, *J. Mag. Res.* **41**:343.
Eigen, M., 1964, *Angew Chem. Int. ed. Engl.* **3**:1.
Erdmann, V. A., 1976, *Prog. Nucl. Acids Res. and Mol. Biol.* **18**:45.
Erdmann, V. A., and Wolters, J., 1986, *Nucl. Acids Res.* **14**:r1.
Erdmann, V. A., Appel, B., Digweed, M., Kluwe, D., Lorenz, S., Luck, A., Schreiber, A., and Schuster, L., 1980, in *Genetics and Evolution of RNA Polymerase, tRNA and Ribosomes* (Osawa, S., ed.), U. Tokyo Press, Tokyo, pp. 553–568.
Figueroa, N., Keith, G., Leroy, J.-L., Plateau, P. Roy, S., and Greron, M., 1983, *Proc. Natl. Acad. Sci. USA* **80**:4330.
Freifelder, D., 1987, *Molecular Biology*, 2nd. ed., Jones and Bartlett Publishers, Inc., Boston.
Garrett, R. A., Douthwaite, S., and Noller, H. F., 1981, *Trends Biochem. Sci.* **6**:137.
Geerdes, H. A. M., and Hilbers, C. W., 1979, *FEBS Lett.* **107**:125.
Gewirth, D. T., and Moore, P. B., 1987, *Biochemistry* **26**:5657.
Gewirth, D. T., Abo, S. R., Leontis, N. B., and Moore, P. B., 1987, *Biochemistry* **26**:5213.
Gollnick, P., Harkin, C. C., and Horowitz, J., 1987, *J. Mol. Biol.* **197**:571.
Gonnella, N. C., Birdseye, T. R., Nee, M., and Roberts, J. D., 1982, *Proc. Natl. Acad. Sci. USA* **79**:4834.
Gorenstein, D. G., 1975, *J. Am. Chem. Soc.* **97**:898.
Gorenstein, D. G., 1984, *Phosphorus-31 NMR Principles and Applications*, Academic Press, Orlando, Florida.
Griffey, R. H., Poulter, C. D., Yamaizumi, Z., Nishimura, S., and Hurd, R. E., 1982a, *J. Am. Chem. Soc.* **104**:5810.
Griffey, R. H., Poulter, C. D., Yamaizumi, Z., Nishimura, S., and Hurd, R. E., 1982b, *J. Am. Chem. Soc.* **104**:5811.
Griffey, R. H., Poulter, C. D., Yamaizumi, Z., Nishimura, S., and Hawkins, B. L., 1983, *J. Am. Chem. Soc.* **105**:143.
Gronenborn, A. M., and Clore, G. M., 1985, *Prog. NMR Spectro.* **17**:1.
Haasnoot, C. A. G., and Hilbers, C. W., 1983, *Biopolymers* **22**:1259.

Hammill, W. D., Grant, D. M., Cooper, R. B., and Harmon, S. A., 1978, *J. Am. Chem. Soc.* **100**:633.
Hare, D. R., and Reid, B. R., 1982a, *Biochemistry* **21**:1835.
Hare, D. R., and Reid, B. R., 1982b, *Biochemistry* **21**:5129.
Hare, D. R., Ribeiro, N. S., Wemmer, D. E., and Reid, B. R., 1985, *Biochemistry* **24**:4300.
Harkin, C. C., and Horowitz, J., 1987, *J. Mol. Biol.* **197**:555.
Heerschap, A., Haassnoot, C. A. G., and Hilbers, C. W., 1982, *Nucl. Acids Res.* **10**:6981.
Heerschap, A., Mellema, J.-R., Janssen, H. G. J. M., Walters, J. A. L. I., Haasnoot, C. A. G., and Hilbers, C. W., 1985, *Eur. J. Biochem.* **149**:649.
Heus, H. A., Van Kimmenade, H. M. A., Van Knippenberg, P. H., Haasnoot, C. A. G., De Bruin, S. H., and Hilbers, C. W., 1983, *J. Mol. Biol.* **170**:939.
Hilbers, C. W., 1979, in *Biological Applications of Magnetic Resonance* (Shulman, R. G., ed.), Academic Press, New York, pp. 1-43.
Hore, P. J., 1983a, *J. Mag. Res.* **54**:539.
Hore, P. J., 1983b, *J. Mag. Res.* **55**:283.
Jagadeeswaran, P., and Cherayil, J. D., 1980, *J. Theor. Biol.* **83**:369.
Johnston, P. D., and Redfield, A. G., 1977, *Nucl. Acids Res.* **4**:3599.
Johnston, P. D., and Redfield, A. G., 1978, *Nucl. Acids Res.* **5**:3913.
Johnston, P. D., and Redfield, A. G., 1979, in *Transfer RNA: Structure, Properties, and Recognition* (Schimmel, P. R., Soll, D., and Abelson, J. N., eds.), Cold Spring Harbor Laboratory, New York, pp. 191-206.
Johnston, P. D., and Redfield, A. G., 1981, *Biochemistry* **20**:3996.
Johnston, P. D., Figueroa, N., and Redfield, A. G., 1979, *Proc. Natl. Acad. Sci. USA* **76**:3130.
Kao, T. H., and Corthers, D. M., 1980, *Proc. Natl. Acad. Sci. USA* **77**:3360.
Kaun, E., Ruterjans, H., and Hull, W. E., 1982, *FEBS Lett.* **141**:217.
Kearns, D. R., 1976, *Prog. Nucl. Acids Res. and Mol. Biol.* **18**:91.
Kearns, D. R., 1977, *Ann. Rev. Biophys. Bioeng.* **6**:477.
Kearns, D. R., 1984, *CRD Critical Reviews in Biochemistry* **15**:237.
Kearns, D. R., and Wong, Y., 1974, *J. Mol. Biol.* **87**:755.
Kearns, D. R., Patel, D. J., and Shulman, R. G., 1971a, *Nature* **229**:338.
Kearns, D. R., Patel, D. J., Shulman, R. G., and Yamane, T., 1971b, *J. Mol. Biol.* **61**:265.
Kim, J. H., and Marshall, J. H., 1989, *Biochemistry* (to be published).
Kime, M. J., 1984a, *FEBS Lett.* **173**:342.
Kime, M. J., 1984b, *FEBS Lett.* **175**:259.
Kime, M. J., and Moore, P. B., 1982, *Nucl. Acids Res.* **10**:4973.
Kime, M. J., and Moore, P. B., 1983a, *Biochemistry* **22**:2615.
Kime, M. J., and Moore, P. B., 1983b, *FEBS Lett.* **153**:199.
Kime, M. J., Gewirth, D. T., and Moore, P. B., 1984, *Biochemistry* **23**:3559.
Kuntzel, H., Heidrich, M., and Piechulla, B., 1981, *Nucl. Acids Res.* **9**:1451.
Lee, K. M., and Marshall, A. G., 1986a, *Prepar. Biochem.* **16**:247.
Lee, K. M., and Marshall, A. G., 1986b, *Biochemistry* **25**:8245.
Lee, K. M., and Marshall, A. G., 1987, *Biochemistry* **26**:5534.
Leocles, B. L., and Moore, P. B., 1986, *Biochemistry* **25**:3916.
Leontis, N. B., and Moore, P. B., 1986, *Biochemistry* **25**:5736.
Leontis, N. B., Ghosh, P., and Moore, P. B., 1986a, *Biochemistry* **25**:7386.
Leontis, N. B., Ghosh, P., and Moore, P. B., 1986b, in *Biomolecular Stereodynamics IV (Proceedings of the Fourth Conversation in the Discipline Biomolecular Stereodynamics, SUNY, Albany, NY, June 4-9, 1985)*, (Sarma, R. H., and Sarma, M. H., eds.), Adenine Press, New York, pp. 287-306.
Leroy, J.-L., Bolo, N., Figueroa, N., Plateau, P., and Gueron, M., 1985a, *J. Biomol. Struct. Dynam.* **2**:915.

Leroy, J.-L., Broseta, D., and Gueron, M., 1985b, *J. Mol. Biol.* **184**:165.
Li, S. J., 1985, Ph. D. Thesis, The Ohio State University, Columbus, Ohio.
Li, S. J., Chang, L.-H., Chen, S.-M., and Marshall, A. G., 1984, *Anal. Biochem.* **138**:465.
Li, S.-J., and Marshall, A. G., 1985, *Biochemistry* **24**:4047.
Li, S.-J., and Marshall, A. G., 1986, *Biochemistry* **25**:3673.
Li, S.-J., Wu, J., and Marshall, A. G., 1987, *Biochemistry* **26**:1578.
Liljas, A., 1982, *Prog. Biophys. Mol. Biol.* **40**:161.
Leuhrsen, K. R., and Fox, G. E., 1981, *Proc. Natl. Acad. Sci. USA* **78**:2150.
Luoma, G. A., and Marshall, A. G., 1978a, *Proc. Natl. Acad. Sci. USA* **75**:4901.
Luoma, G. A., and Marshall, A. G., 1978b, *J. Mol. Biol.* **125**:95.
Luoma, G. A., Burns, P. D., Bruce, R. E., and Marshall, A. G., 1980, *Biochemistry* **19**:5456.
Marshall, A. G., and Smith, J. L., 1977, *J. Am. Chem. Soc.* **99**:635.
Marshall, A. G., and Smith, J. L., 1980, *Biochemistry* **19**: 5955.
McConnell, H. M., 1958, *J. Chem. Phys.* **28**:430.
Meier, J., and Marshall, A. G., this volume, Chapter 6.
Mizuno, H., and Sundaralingam., 1978, *Nucl. Acids Res.* **5**:4451.
Morikawa, K., Kawakami, M., and Takemura, S., 1982, *FEBS Lett.* **145**:194.
Muller, J. J., Misselwitz, R., Zirwer, D., Damaschun, G., and Welfle, H., 1985, *Eur. J. Biochem.* **148**:89.
Nazar, R. N., 1980, *FEBS Lett.* **119**:212.
Noggle, J. H., and Schirmer, R. E., 1971, *The Nuclear Overhauser Effect, Chemical Applications*, Academic Press, New York.
Noller, H. F., and Garrett, R. A., 1979, *J. Mol. Biol.* **132**:621.
Nomura, M., and Erdmann, V. A., 1970, *Nature* **228**:744.
Pace, B., Matthews, E. A., Johnson, K. D., Cantor, C. R., and Pace, N. R., 1982, *Proc. Natl. Acad. Sci. USA* **79**:36.
Pardi, A., and Tinoco, I. Jr., 1982, *Biochemistry* **21**:4686.
Patel, D. J., 1978, *Ann. Rev. Phys. Chem.* **29**:337.
Perkins, S. J., 1982, in *Biological Magnetic Resonance:4* (Berliner, L. J. and Reuben, J., eds.), Plenum Press, New York, pp. 193-336.
Pieler, T., Digweed, M., and Erdmann, V. A., 1984, in *Gene Expression* (Clark, B. F. and Peterson, H. U., eds.), Muksgaard, Copenhagen, pp. 353-376.
Plateau, P., Dumas, C., and Gueron, M., 1983, *J. Mag. Res.* **54**:46.
Rabin, D., Kao, T-H., and Crothers, D. M., 1983, *J. Biol. Chem.* **25**:10813.
Redfield, A. G., 1976, in *NMR: Basic Principles and Progress* Vol. 13 (Diehl, P., *et al.* eds.), Springer-Verlag, New York, pp. 137-152.
Redfield, A. G., 1978, *Methods Enzymol.* **49**:253.
Redfield, A. G., and Gupta, R. K., 1971, *J. Chem. Phys.* **54**:1418.
Redfield, A. G., and Kunz, S. D., 1979, in *NMR and Biochemistry* (Opella. S. J., and Lu, P., eds.), M. Dekker, New York, pp. 225-239.
Redfield, A. G., Kunz, S. D., and Ralph, E. K., 1975, *J. Mag. Res.* **19**:114.
Redfield, A. G., Roy, S., Sanchez, V., Troop, J., and Figueroa, N., 1981, in *Proceedings of the Second SUNY Conversation in the Discipline Biomolecular Stereodynamics*, vol. 1, (Sarma, R. H., ed.), Adenine Press, New York, pp. 195-208.
Reid, B. R., 1981a, in *Topics in Nucleic Acid Structure* (Neidle, S., ed.), Wiley, New York, pp. 113-139.
Reid, B. R., and Hurd, R. E., 1977, *Acc. Chem. Res.* **10**:396.
Reid, B. R., and Robillard, G. T., 1975, *Nature* **257**:287.
Reid, B. R., 1981b, *Ann. Rev. Biochem.* **50**:969.
Reid, B. R., Azhderian, E., Hurd, R. E., 1979, in *NMR and Biochemistry* (Opella, S. J., and Lu, P., eds.), M. Dekker, New York, pp. 91-115.

Richards, E. G., Lecanidou, R., and Geroch, M. E., 1973, *Eur. J. Biochem.* **34**:262.
Robillard, G. T., and Reid, B. R., 1979, in *Biological Applications of Magnetic Resonance* (Shulman, R. G., ed.), Academic Press, New York, pp. 45-112.
Rordorf, B. F., and Kearns, D. R., 1975, *Biochem. Biophys. Res. Commun.* **65**:857.
Roth, K., Kimber, B. J., and Feeney, J., 1980, *J. Mag. Res.* **41**:302.
Roy, S., and Redfield, A. G., 1981, *Nucl. Acids Res.* **24**:7073.
Roy, S., and Redfield, A. G., 1983, *Biochemistry* **22**:1386.
Roy, S., Papastavros, M. Z., and Redfield, A. G., 1982, *Biochemistry* **21**:6081.
Salemink, P. J. M., Raue, H. A., Heerschap, A., Planta, R. J., and Hilbers, C. W., 1981, *Biochemistry* **20**:265.
Sanchez, V., Redfield, A. G., Johnston, P. D., and Tropp, J., 1980, *Proc. Natl. Acad. Sci. USA* **77**:5659.
Schimmel, P. R., and Redfield, A. G., 1980, *Ann. Rev. Biophys. Bioeng.* **9**:181.
Sklenar, V., and Starcuk, Z., 1982, *J. Mag. Res.* **50**:495.
Tetelbaum, H., and Englander, S. W., 1975a, *J. Mol. Biol.* **92**:55.
Tetelbaum, H., and Englander, S. W., 1975b, *J. Mol. Biol.* **92**:79.
Tinoco, Jr., I., Uhlenbeck, O. C., and Levine, M. D., 1971, *Nature*, **230**:362.
Tinoco, Jr., I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. and Gralla, J. 1973, *Nature New Biol.* **246**:40.
Trifonov, E. N., and Bolshoi, G., 1983, *J. Mol. Biol.* **169**:1.
Turner, D. L., 1983, *J. Mag. Res.* **54**:146.
Ulbrich, N., Lin, A., and Wool, J. G., 1979, *J. Biol. Chem.* **254**:8641.
Uhlenbeck, O. C., and Cameron, V., 1977, *Nucl. Acids Res.* **4**:85.
Uhlenbeck, O. C., Lowary, P. T., and Wittenberg, W. L., 1982, *Nucl. Acids Res.* **10**:3341.
Weidner, H., and Crothers, D. M., 1977, *Nucl. Acids Res.* **4**:3401.
Weidner, H., Yaun, R., and Crothers, D. M., 1977, *Nature* **266**:193.
Wemmer, D. E., Chou, S.-H., and Reid, B. R., 1984, *J. Mol. Biol.* **180**:41.
Wong, Y. P., Kearns, D. R., Reid, B. R., and Shulman, R. G., 1972, *J. Mol. Biol.* **72**:741.
Wrede, P., and Erdmann, V. A., 1977, *Proc. Natl. Acad. Sci. USA* **74**:2704.
Wright, J. M., Feigon, J., Denny, W., Leupin, W., and Kearns, D. R., 1981, *J. Mag. Res.* **45**:514.
Wu, J., and Marshall, A. G., 1989a, *Biochemistry* (to be published).
Wu, J., and Marshall, A. G., 1989b, *Biochemistry* (to be published).
Wüthrich, K., 1986, *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zagorska, L., Van Duin, J., Noller, H. F., Pace, B., Johnson, K. D., and Pace, N. R., 1984, *J. Biol. Chem.* **259**:2798.

3

Structure Determination via Complete Relaxation Matrix Analysis (CORMA) of Two-Dimensional Nuclear Overhauser Effect Spectra

DNA Fragments

Brandan A. Borgias and Thomas L. James

1. INTRODUCTION

Methods for the complete high-resolution determination of molecular structure in solution have been a goal of scientists since the development of X-ray diffraction techniques in the early part of this century. Methods that yield at least a part of the structural picture include hydrodynamic techniques, electron diffraction and other scattering techniques, circular dichroism, fluorescence decay, extended X-ray absorption fine structure, and magnetic resonance. The information obtainable from these methods ranges from rough estimates of the overall shape and dimension of a molecule to a gauge of the extent of secondary structure (local conformational properties such as α-helix and β-sheet). At best it is possible to obtain a measure of some of the distances within the molecule. However, with the possible exception of nuclear magnetic resonance, a high-resolution structural model is not yet attainable with any of these methods.

NMR has experienced an explosion of activity in recent years due largely to the development of two-dimensional spectroscopy (2D NMR).

Brandan A. Borgias • Cray Research, Inc., San Ramon, California 94583. **Thomas L. James** • Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143.

The 2D NMR experiment, first suggested by Jeener (1971), is now routinely used in a vast array of specialized experiments for spectrum assignment and structural characterization of macromolecules. The key attraction of 2D NMR methods is that they achieve resolution of crowded spectral regions by mapping the spectrum onto two frequency axes rather than the usual (chemical shift) plot of the spectrum (Figure 1). Typically, the 2D spectrum will have a ridge of high intensity running along the diagonal of the plot which has the general appearance of the ordinary 1D spectrum (in fact a projection onto the axis results in the ordinary 1D spectrum). The off-diagonal space in the 2D spectrum will have peaks at coordinates (ω_1, ω_2)

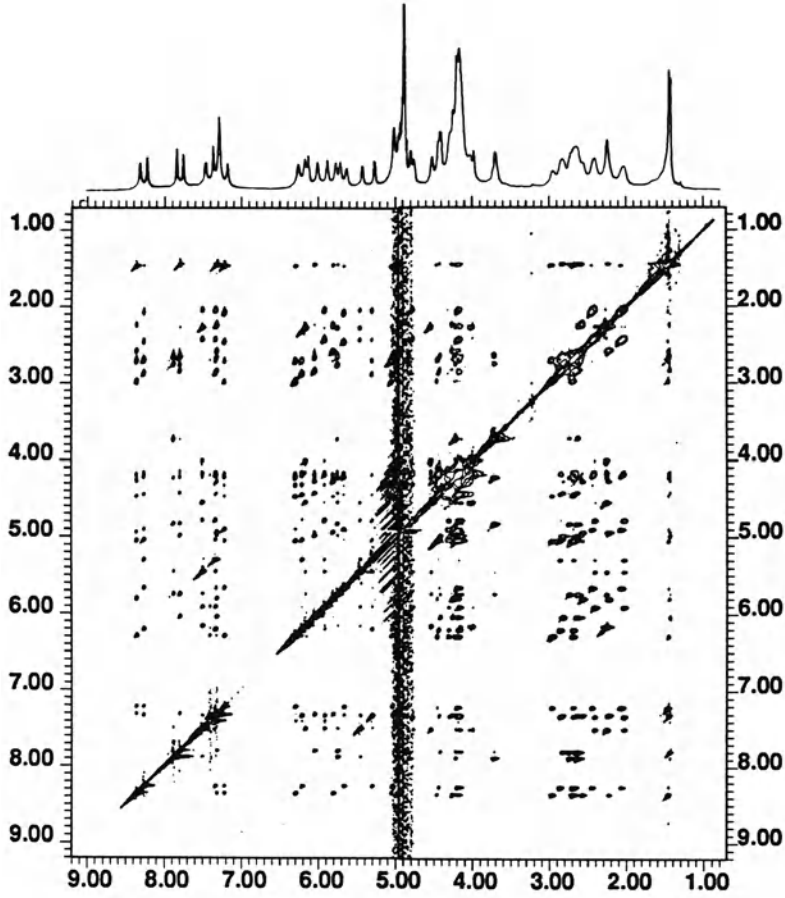


Figure 1. Pure absorption phase two-dimensional nuclear Overhauser effect (2D NOE) spectrum of [d(GGTATACC)$_2$]. Mixing time 250 msec. The usual 1D spectrum is also shown for comparison.

corresponding to the frequencies of two nuclei in communication with each other, and with intensities dictated by the particular internuclear interaction detected by the experiment. The off-diagonal (or cross-peaks) can be the result of scalar coupling (e.g., in correlated spectroscopy or COSY) or dipole-dipole coupling which is the case of 2D nuclear Overhauser effect spectroscopy (2D NOE or NOESY).

The homonuclear 2D NOE experiment readily yields structural information. This experiment relies on proton-proton dipolar relaxation to yield a correlation map between protons in close spatial proximity ($\leq$0.5 nm). The intensities of the cross-peaks in the 2D NOE spectrum are related to the distances between the protons and can be used to obtain estimates of those distances. While the focus of this review will be the application of 2D NOE spectroscopy to the determination of oligonucleotide structure, the general methodology has broader application. The details of the experiment and data analysis will be discussed, and the evaluation of macromolecular structure, once accurate proton-proton distances are obtained, will be described. The use of 2D NOE to determine macromolecular structure has become widespread in recent years. And, while considerable success has been achieved with this technique, there are certain limitations and precautions that are not always considered which we will elaborate below.

2. STRUCTURE EVALUATION BASED ON 2D NOE SPECTROSCOPY

The 2D NOE experiment was first described (along with the mathematically equivalent experiment for chemical exchange) by Jeener *et al.* (1979) and given firm theoretical foundation in Ernst's laboratory (Macura and Enrst, 1980). The essential concepts described in these seminal papers have been further elaborated by several other laboratories. A matrix method for the explicit prediction of the 2D NOE intensities expected for a known molecular structure was developed (Keepers and James, 1984). An early test of the method compared results obtained for a rigid test molecule in solution with its X-ray crystal structure (Young and James, 1984). We call this approach CORMA for complete relaxation matrix analysis. Others (Abel *et al.*, 1986; Macura *et al.*, 1986; Massefski and Bolton, 1985; Olejniczak *et al.*, 1986; and Mirau, 1987) have applied or further expanded upon this complete matrix approach. The critical feature in CORMA is the explicit treatment of the relaxation network in calculating the 2D NOE intensities. The alternative (and still widely used) approach assumes that the intensity correlating with a particular proton-proton pair depends on the distance separating those protons alone and can be called the isolated spin pair

approximation (ISPA). Most analyses of 2D NOE spectra to data have incorporated ISPA in the qualitative or semiquantitative extraction of distance limits based upon the relative cross-peak intensities. Gronenborn and Clore (1985) provide a good review of this approach along with all aspects of data collection and analysis.

2.1. Basic 2D NOE Experiment: $90° - t_1 - 90° - \tau_m - 90° - t_2$

The basic 2D NOE experiment centers around the three-pulse sequence: (delay time $- 90° - t_1 - 90° - \tau_m - 90° - t_2)_n$ (Kumar *et al.*, 1980). The effect of the first 90° pulse is to create transverse magnetization. During delay time t_1, free precession occurs which has the effect of frequency-labeling each of the spins. The second pulse achieves inversion of the spins. During the mixing time τ_m the effect of magnetization exchange due to cross-relaxation is recorded by the spins in the form of population changes. The final 90° observation pulse again generates transverse magnetization for signal detection during t_2. Double Fourier transformation with respect to times t_1 and t_2 results in the 2D NOE spectrum (Figure 1). The effect of cross-relaxation between neighboring protons during τ_m is to transfer magnetization between then. This results in off-diagonal intensities that are roughly inversely proportional to the sixth power of the distance between them. This approximate proportionality is the basis for the commonly used ISPA approach to the estimation of distances. This approach carries with it the implicit assumption that each proton–proton pair can be considered as an isolated system with no external influence from other neighboring protons. However, this is not the case. In fact, there are on average 3.4 neighboring protons within a 0.3 nm radius of each proton in B-DNA. In order to accurately determine distances from the 2D NOE spectrum, we choose to analyze the intensities by taking into account the complete relaxation network.

2.2. Complete Relaxation Matrix Analysis—CORMA

The cross-relaxation during the mixing time τ_m is described by the system of equations (Macura and Ernst; 1980)

$$\frac{\partial M}{\partial t} = -RM \tag{1}$$

In equation (1), **M** is the magnetization vector describing the deviation from thermal equilibrium ($\mathbf{M} = \mathbf{M}_z - \mathbf{M}_0$), and **R** is the matrix describing

the complete dipole–dipole relaxation network.* This is essentially an extension of the two-spin equations of Solomon (1955):

$$R_{ii} = 2(n_i - 1)(W_1^{ii} + W_2^{ii}) + \sum_{j \neq i} n_j(W_0^{ij} + 2W_1^{ij} + W_2^{ij}) + R_{1i} \quad (2a)$$

$$R_{ij} = n_i(W_2^{ij} - W_0^{ij}) \quad (2b)$$

Here n_i is the number of equivalent spins in a group such as a methyl rotor, and the zero, single, and double transition probabilities W_n^{ij} are given (for isotropic random reorientation of the molecule) by:

$$W_0^{ij} = \frac{q\tau_c}{r_{ij}^6} \quad (3a)$$

$$W_1^{ij} = 1.5\frac{q\tau_c}{r_{ij}^6}\frac{1}{1 + (\omega\tau_c)^2} \quad (3b)$$

$$W_2^{ij} = 6\frac{q\tau_c}{r_{ij}^6}\frac{1}{1 + 4(\omega\tau_c)^2} \quad (3c)$$

where $q = 0.1\gamma^4\hbar^2$. The term R_{1i} represents external sources of relaxation such as paramagnetic impurities and is generally ignored. The system of equation (1–3) has the solution

$$\mathbf{M}(\tau_m) = \mathbf{a}(\tau_m)\mathbf{M}(0) = \exp(-\mathbf{R}\tau_m)\mathbf{M}(0) \quad (4)$$

where **a** is the matrix of mixing coefficients which are proportional to the 2D NOE intensities. This matrix of mixing coefficients is what we wish to evaluate. The exponential dependence of the mixing coefficients on the cross-relaxation rates complicates the calculation of intensities (or the distances). But, as shown below in Section 2.2.2, the calculation of intensities is a relatively easily surmountable problem. Before continuing, we would note here that the expression for the rate matrix given above is actually still an approximation in that it neglects cross-correlation terms between separate pairwise and higher-order interactions (Werbelow and Grant, 1978; Bull, 1987). However, the importance of the cross-correlation terms appears to be small for $\tau_m \leq \tau_m^{\text{opt}}$, the optimal mixing time for maximum cross-peak intensity (Bull, 1987). The expressions given above also do not account for

* Throughout this chapter we will use boldface type to indicate matrices or vectors: **R**. Individual elements of matrices or vectors are indicated by a subscripted letter in normal type: R_{ij}.

second-order effects due to strong scalar coupling (Kay *et al.*, 1986a; 1986b). However, the magnitude of error due to the neglect of scalar coupling is small for macromolecules where $\omega\tau_c \gg 1$ and $\tau_m \leq \tau_m^{\text{opt}}$. Moreover, the effects of coupling can be suppressed by use of appropriate pulse sequences (Macura *et al.*, 1981; 1982; Rance *et al.*, 1985) as described in Section 3.1.3.

2.2.1. Approximate Distance Calculations

The exact evaluation of intensities (the mixing coefficients **a**) according to equation (3) is hindered by the absence of a general analytical expression for matrix exponentiation. Historically, this has led to the ISPA approach since the exponential can be recast into a Taylor series expansion:

$$\mathbf{a}(\tau_m) = \exp(-\mathbf{R}\tau_m) \approx 1 - \mathbf{R}\tau_m + \tfrac{1}{2}\mathbf{R}^2\tau_m^2 - \cdots + \frac{(-1)^n}{n!}\mathbf{R}^n\tau_m^n + \cdots \tag{5}$$

Truncating the series after the linear term results in a simple approximation for the mixing coefficients that is valid for short ($\tau_m \to 0$) mixing times:

$$a_{ij}(\tau_m \to 0) \approx (\delta_{ij} - R_{ij}\tau_m) \tag{6}$$

Incorporation of equations (3b), (4a), and (4c) into equation (6) gives, for $i \neq j$,

$$a_{ij} = n_i \frac{q\tau_m}{r_{ij}^6}\left(1 - \frac{6}{1 + 4(\omega\tau_c)^2}\right) \tag{7}$$

The qualitative result is that the intensities are inversely proportional to the sixth power of the distance between the correlated protons. The practical application of this approximation usually goes one step further to eliminate the dependence on correlation time. This is achieved by scaling all the distances with respect to a known reference distance which is assumed to have the same correlation time as the proton–proton pair of interest. For oligonucleotides the following rules are generally followed (Gronenborn and Clore, 1985):

1. The correlation times for sugar–sugar, sugar–methyl, and sugar–base (with the exception of H1′–base) interproton vectors are the same as the intranucleotide H2′–H2″ vector with a reference distance 0.177 nm (for nucleotide atom labeling see Figure 2).

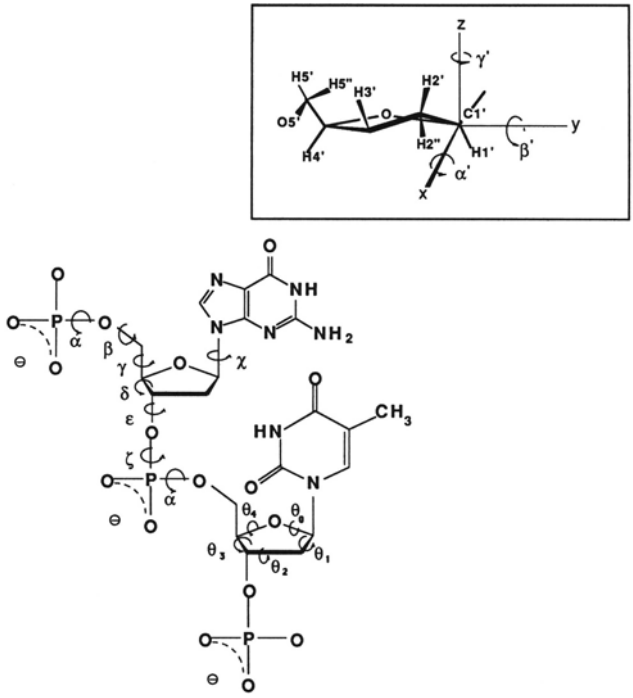


Figure 2. Oligonucleotide atom labeling and definition of COMATOSE conformational parameters. The oligonucleotide is defined by 10 × N parameters (N is the number of nucleotides in the oligomer). These are the exocyclic torsion angles χ and γ, and the sugar pucker parameters P and θ_{max}, which define the ring torsion angles θ_1, θ_2, and θ_3 and the group orientation parameters x, y, z, and α', β', γ' which define the translation and orientation of the nucleotide relative to the initial trial structure (inset). The chain torsion angles involving the phosphodiester bonds α, β, ε, and ζ are not explicitly included in COMATOSE.

2. The correlation times for base-base, base-methyl and H1′-base interproton vectors are the same as the intranucleotide H5-H6 vector of cytidine with a reference distance 0.246 nm.

Then distances are calculated according to equation (8):

$$r_{ij} = r_{\text{ref}}\left(\frac{a_{\text{ref}}}{a_{ij}}\right)^{1/6} \tag{8}$$

Aside from the errors associated with network relaxation (i.e., failure of the isolated spin pair assumption: $r_{ij} \ll r_{ik}$ and $r_{ij} \ll r_{jk}$ for all k atoms), this protocol references all distances (up to ~0.5 nm) to a single short distance that is not really representative of the majority of distances determined from the 2D NOE spectrum. For example, consider the results obtained

with a series of short oligonucleotide sequences (Gronenborn *et al.*, 1984a; Gronenborn *et al.*, 1984b; Clore and Gronenborn, 1984). The overall root-mean-square difference (RMSD) between distances derived by ISPA from NMR intensities and those derived from fiber diffraction data on B-DNA (Arnott and Hukins, 1972) is ~0.05 nm. The RMSD between the NMR distances and A-DNA was found to be ~0.11 nm. This clearly discriminated against A-DNA as a plausible solution structure. The low RMSDs between the NMR and B-DNA structures would seem to indicate good agreement, but note that 0.05 nm is also the magnitude of the RMSD between regular B-DNA and an analogous energy-minimized structure. These structures are generally very similar, but nonetheless are significantly different at the local level. The conclusion to be drawn is that for nucleic acids—for which it is possible to predict the structure to a moderately high level of accuracy *a priori*—the distances derived from ISPA are of only marginal utility in further defining the structure.

Keeping in mind the limitations of the ISPA approach, it is still worthwhile to consider the results obtained by its application to the determination of macromolecular structure. Examples of the application of ISPA distances in the determination of oligonucleotide structure can be found in structural studies on $[5'd(CGTACG)]_2$ and $[5'd(ACGCGCGT)]_2$ (Gronenborn *et al.*, 1984b); a study of the refinements of [5′(CACAG):5′(CUGUG)] (Clore *et al.*, 1985a) and a B-DNA sequence $[5'(CGTACG)]_2$ (Clore *et al.*, 1985b); the study of the structure of an RNA pentadecamer comprising the anticodon loop and stem of yeast $tRNA^{Phe}$ (Clore *et al.*, 1984); a study of the interaction between the gene-5 protein and DNA (King and Coleman, 1987); and a study of the duplex $[d(GAATTCGAATTC)]_2$, where scalar coupling interactions (COSY) were used in conjunction with 2D NOE to build a structural model (Chary *et al.*, 1987). In all these studies, general structural features can be obtained. But, because of the inherent errors incorporated into the ISPA distances, some details and structural variations are less certain. Nonetheless, their general conclusion are probably valid.

Perhaps more exciting are the complete structural studies which have been performed on proteins in Wüthrich's lab and in the labs of Clore and Gronenborn. Wüthrich's group has used the ISPA distances determined from 2D NOE intensities to set limits in distance geometry calculations (Kuntz *et al.*, 1979; Crippen, 1984; Havel and Wüthrich, 1985; Braun and Gō, 1985) for determining the structure of the protease inhibitor IIa (Williamson *et al.*, 1985), tendamistat (Kline *et al.*, 1986), and metallothionein-2 (Braun *et al.*, 1986). Gronenborn and Clore, on the other hand, have applied the complete battery of 2D NOE, distance geometry, and molecular mechanics and dynamics to the determination of the solution structures of phoratoxin and hirudin (Clore *et al.*, 1987a; 1987b). Again, while the distances determined from the 2D NOE spectra in these studies are subject

to the errors of the ISPA calculations, good three-dimensional pictures of the proteins are obtainable. This success is probably largely due to the fact that the distances used in these studies are between protons on amino acyl residues separated by at least five intervening residues. The effect of this judicious use of only these long range distances is that the overall structure is reasonably well constrained by important loop-closing interresidue contacts. The details of the local structure, however, are determined primarily by energetic, steric, and geometric constraints.

2.2.2. Exact Calculation of 2D NOE Intensities

The primary motivation for the ISPA approach is its computational simplicity. However, with currently available computer power the exact numerical solution for the matrix **a** is a relatively tractable problem. The problem of numerically evaluating the matrix of mixing coefficients is simply one of solving the eigenvalue problem for the rate matrix $\mathbf{R} = \boldsymbol{\chi}\boldsymbol{\lambda}\boldsymbol{\chi}^{-1}$ where $\boldsymbol{\chi}$ is the eigenvector matrix and $\boldsymbol{\lambda}$ is the diagonal eigenvalue matrix. Then, since $\boldsymbol{\lambda}$ is diagonal, the series expansion for the exponential collapses.

$$\mathbf{a} = 1 - \boldsymbol{\chi}\boldsymbol{\lambda}\boldsymbol{\chi}^{-1}\tau_m + \tfrac{1}{2}\boldsymbol{\chi}\boldsymbol{\lambda}\boldsymbol{\chi}^{-1}\boldsymbol{\chi}\boldsymbol{\lambda}\boldsymbol{\chi}^{-1}\tau_m^2 - \cdots \tag{9}$$

$$\mathbf{a} = \boldsymbol{\chi}\exp(-\boldsymbol{\lambda}\tau_m)\boldsymbol{\chi}^{-1} \tag{10}$$

We have developed a program, CORMA,* for calculating and displaying the 2D NOE intensities for any set of proton coordinates. This program is a direct translation of equations (2) and (3) for calculation of the rate matrix and calculates the exponential by diagonalizing **R** and evaluating equation (10). Calculations using CORMA allow direct determination of the contributions to cross-peak intensity due to spin diffusion effects, and assessment of the error associated with the ISPA determination of distances.

Spin diffusion effects can be easily measured by running parallel calculations in which intervening protons in the network have been removed, leaving behind only the direct dipole–dipole interaction. Similarly, the optimal mixing times for data collection can be estimated by running several calculations spanning a range of possible mixing times, and observing the buildup and decay of the total cross-peak intensity. A mixing time near the

* CORMA is written in FORTRAN and is compatible with both VMS and UNIX compilers. Full implementation of the program's plotting features requires a laser printer (e.g., Apple LaserWriter) which supports POSTSCRIPT (Adobe Systems, Inc.) language for graphic control. In the absence of such a printer, the program still writes output files containing all the calculated intensities in residue-by-residue submatrices. Also output are all the associated distances and relaxation rates. The program has been successfully run on molecules ranging up to 500 protons (e.g., BPTI) and is available from the authors.

end of the build-up phase (~80–90% of the maximum intensity) will yield the most usable information: greatest signal-to-noise ratio, large number of cross-peaks, minimal loss of information to averaging of intensities due to spin diffusion. Of course, the optimal mixing time for a particular cross-peak is a function of the distance between that pair of protons and of the local environment of the pair (i.e., if either proton has many near neighbors, the cross-peak will show the effects of spin diffusion at shorter mixing times). The buildup and decay of cross-peaks corresponding to distances between 0.23 and 0.25 nm in a dinucleotide are shown in Figure 3. This calculation assumed an isotropic correlation time of 5 nsec, and the error associated with simply relating distances to intensities is evident. It may also be inferred from this plot that the mixing times that will yield the most information are between 100–250 msec. In fact, the total cross-peak intensity has reached ~90% of its maximum by 100 msec. Mixing times longer than this allow the observation of a few additional long-distance peaks, but the peaks corresponding to the shortest distances will have already begun to decay.

In order to assess the magnitude of the ISPA error, we have generated simulated intensity sets (i.e., theoretical 2D NOE spectra) for a model

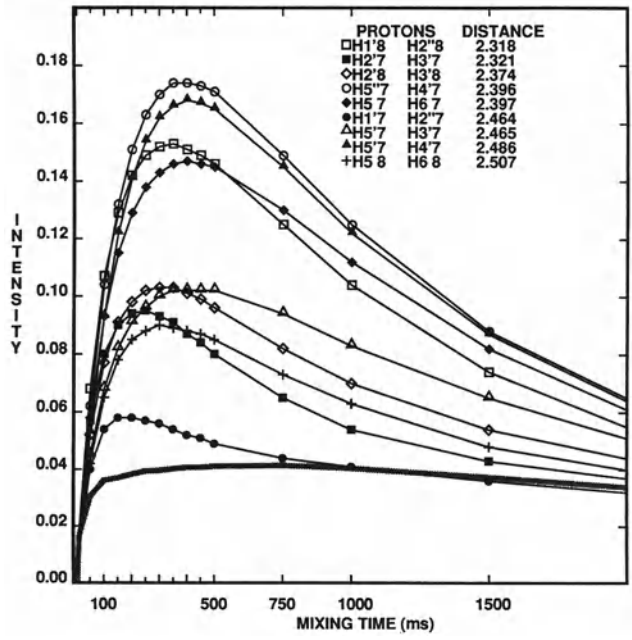


Figure 3. Buildup and decay of intensities for cross-peaks corresponding to proton–proton distances of 0.23–0.25 nm in dCC extracted from energy-minimized octamer. Intensities were calculated using CORMA assuming an isotropic correlation time of 5 nsec. The grey shaded line corresponds to the average of all cross-peak intensities.

structure, using CORMA. We then calculated distances from these intensities according to equation (8), i.e., using the isolated spin pair approximation. Comparison of the ISPA-estimated distances with the true distances from the defined model shows that ISPA systematically underestimates the distance even for relatively short mixing times where the approximation is presumed to be valid. As shown in Figure 4, which represents distances for the oligomer d(GGTATACC), even for moderate distances of 0.35 nm the error is considerable. The systematic nature of the error suggests the possibility of applying an empirical correction factor to the estimated distances. This factor will not only be a function of the correlation time and mixing time, but will also depend on the local environment of the protons. So, assuming a plausible structural model is available, the distances could be corrected. But note, the range of actual distances corresponding to an estimated distance of 0.2 nm at 250 msec is ~0.17 nm, so there remains considerable uncertainty in the distances. Errors of this magnitude preclude the detailed determination of structure where an RMSD of less than 0.05 nm is necessary to distinguish between equally plausible but different structures (e.g., Arnott-B vs. energy-minimized B-DNA).

We note here, and emphasize, that the ISPA calculation of distances is not without merit, the inherent errors notwithstanding. Considerable work has been accomplished with this approach, and for the overall determination of macromolecular structure, especially in proteins, it appears to be very successful. This approach, however, is limited to the determination of the general medium resolution structure. Many details of local structure will be lost or be in error due to the errors associated with the approximation. These same details are often of paramount importance in understanding the significance of seemingly minor structural or chemical variations with respect to the activity of an enzyme or protein.

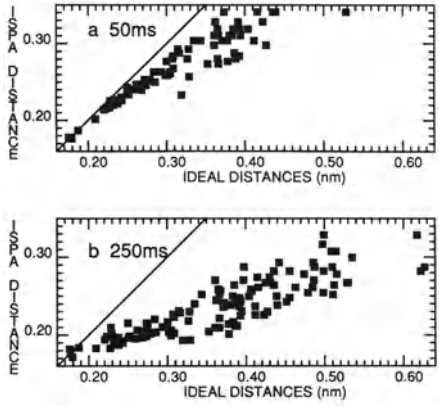


Figure 4. Comparison of distances calculated according to the isolated spin pair approximation (ISPA) with actual distances. The ISPA distances were calculated according to equation (8) of the text from ideal intensities generated by CORMA for d(GGTATACC) in energy-minimized B-DNA conformation assuming a correlation time of 4 nsec and mixing times of (a) 50 msec and (b) 250 msec.

2.3. Structure Evaluation by CORMA

In our laboratory, CORMA has been applied to the investigation of several oligonucleotide sequences. Over the course of these studies the sophistication of the analysis has developed from the evaluation of a few models on the basis of selected intensities to a more detailed analysis of many closely related models using all well-resolved intensities. This approach is limited, however, by the choice of structural models considered for analysis. To this end we are currently working on a program for refinement of structure based on the 2D NOE (see Section 2.5).

2.3.1. The Structure of $[d(GGAATTCC)]_2$

One of the first duplexes studied was $[d(GGAATTCC)]_2$ (Broido *et al.*, 1984; 1985), in which the central six base pairs comprise the EcoRI endonuclease recognition site. These six base pairs are also identical to those in the dodecamer $[d(CGCGAATTCGCG)]_2$, for which a crystal structure has been reported (Dickerson and Drew, 1981). The 2D NOE intensities involving the nonexchangeable base protons were compared with CORMA intensities calculated first for regular B-DNA and then for a structure obtained as a result of energy minimization using the program AMBER, assisted model building with energy refinement (Weiner and Kollman, 1981; Weiner *et al.*, 1984). Neither trial structure accurately reproduced all peak intensities used for comparison, but many features of the experimental data were represented by both the regular B form and the energy-minimized structure. Overall, the energy-minimized structure generated intensities which best matched the experimental intensities. It was noted that the effective correlation times involving the base protons were found to lie in the range 1.7–2.7 nsec with the A H8 protons exhibiting the longer correlation times. Also, while most of the intensities were best fit with the energy-minimized structure, the T(5)-H6 protons were better fit by the regular B structure. These observations suggested a sequence-dependent structural change at the purine/pyrimidine junction in the oligomer. The plausibility of this is supported by crystallographic results: when EcoRI endonuclease was cocrystallized with $[d(TCGCGAATTCGCG)]_2$ (Frederick *et al.*, 1984) a kink was found at the GAA–TTC junction. It was suggested that this kinked conformation, while clearly stabilized by the binding of the DNA to the protein, must also exist transiently for the isolated oligonucleotide alone in solution.

2.3.2. $[d(GGTATACC)]_2$—Anomalous Structure of the TATA Box

A second study focused on the duplex $[d(GGTATACC)]_2$, containing the TATA box which is commonly found in the promoter region of genes

and which binds to RNA polymerase preparatory to transcription. Twenty 2D NOE cross-peaks at each of four mixing times (100, 175, 250, and 400 msec) were compared with CORMA-generated intensities for regular B-DNA and for the analogous energy-minimized structure (Jamin *et al.*, 1985). In this case, neither model structure adequately modeled the solution structure; the GG and CC regions were well represented by regular B-DNA, but the TATA box exhibited large discrepancies with both models. In particular, an (interstrand?) interaction between A(6)-H8 and T(3)-CH_3 was detected experimentally, but it is much too long in either of the model structures to generate significant intensity.

2.3.3. $[d(AT)_5]_2$—a Wrinkled D Structure

An alternative structural model to account for these observations was suggested by a study on the decamer duplex $[d(AT)_5]_2$ (Suzuki *et al.*, 1986). The solution structure of this sequence has been subject to some controversy. It has been proposed to have a normal B-DNA structure (Assa-Munt and Kearns, 1984; Borah *et al.*, 1985) or a left-handed B conformation (Gupta *et al.*, 1983). On the other hand, no less than seven models have been proposed for its structure in the solid state: A-DNA (Davies and Baldwin, 1963), B-DNA (Arnott and Hukins, 1972), alternating B-DNA (Klug *et al.*, 1979), D-DNA (Marvin *et al.*, 1961), a right-handed D-DNA (Arnott and Selsing, 1974), left-handed D-DNA with Hoogsteen base pairing (Drew and Dickerson, 1982), and wrinkled D-DNA (Arnott *et al.*, 1983). Comparison of the experimental intensities with CORMA calculated intensities assuming a 7 nsec correlation time for structures A, B, alternating B, left-handed B, C, D, wrinkled D and Z derived from $[d(CGCGCG)]_2$ (Wang *et al.*, 1979) showed most favorable fits for B, D, and wrinkled D DNA structures, with the wrinkled D form providing the best fit to experimental data. The B, D, and wrinkled D structures (all members of the B family of DNA structures) were then subjected to energy minimization using AMBER in which counter ions and water of hydration were included. This yielded the interesting result that the minimized structures were all very similar to each other as well as to the wrinkled D starting structure. Moreover, a hydration tunnel running the length of the comparatively deep and narrow minor groove was discovered (Figure 5).

2.3.4. $[d(GGTATACC)]_2$ Revisited—Duplex Adopts a B-D-B Structure

Returning attention to the duplex $[(GGTATACC)]_2$ with knowledge of the wrinkled D structure adopted by oligo d(AT), resulted in a solution to the problem of its structure (Zhou *et al.*, 1987). A model was made by joining B-form $[d(GG)]_2$ and $[d(CC)]_2$ caps to the $[d(TATA)]_2$ sequence

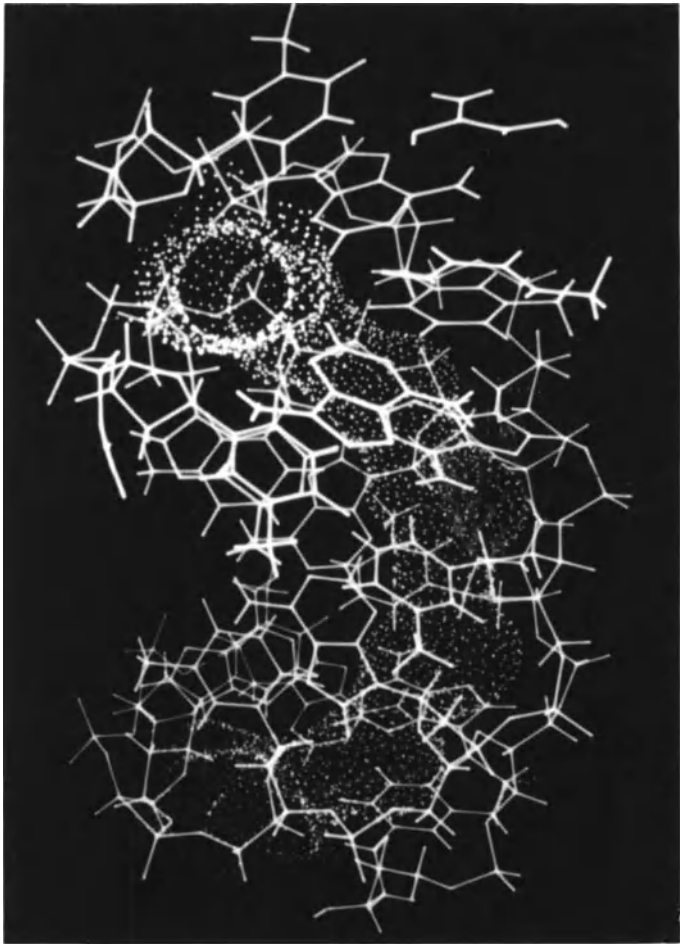


Figure 5. Wrinkled D structure of $[d(AT)_5]_2$ generated by energy minimization, showing the hydration tunnel in the minor groove. This structure is most compatible with the experimentally observed 2D NOE intensities.

in the wrinkled-D conformation. This BDB model was then subjected to energy-minimization which included counter ions and hydrogen-bonded water, as in the previous study (Suzuki *et al.*, 1986). The 2D NOE intensities corresponding to these models were then compared with the experimental intensities previously generated (Jamin *et al.*, 1985). Additionally, the 2D NOE intensities corresponding to four other models were also considered: A (Arnott and Hukins, 1972), crystalline A (Shakked *et al.*, 1983), B (Arnott and Hukins, 1973), and energy-minimized B (with the inclusion of counter ions and water of hydration. In this study (Zhou *et al.*, 1987), all nonoverlap-

ping cross-peaks were quantified from all four mixing times (100, 175, 250, and 400 msec) yielding a total of 542 cross-peak intensities. The energy-minimized BDB structure (Figure 6) have the best fit to the experimental intensities and was also most favorable from an energetic standpoint. Like the wrinkled-D $[d(AT)_5]_2$ decamer, it also contains a hydration tunnel. However, in this structure, the tunnel is sealed off by the $[d(GG)]_2$ and $[d(CC)]_2$ caps which are in a predominantly B-DNA conformation. The energy-minimized B structure generated in this study also gave a reasonable fit, as opposed to the energy-minimized B structue in the prior study (Jamin *et al.*, 1985). This is presumably due to the inclusion of counterions and hydration water during the minimization.

These studies illustrate the utility of applying CORMA to the analysis of oligonucleotide structure. It is important to bear in mind that relatively small variations in structure are translated into quantifiable differences in

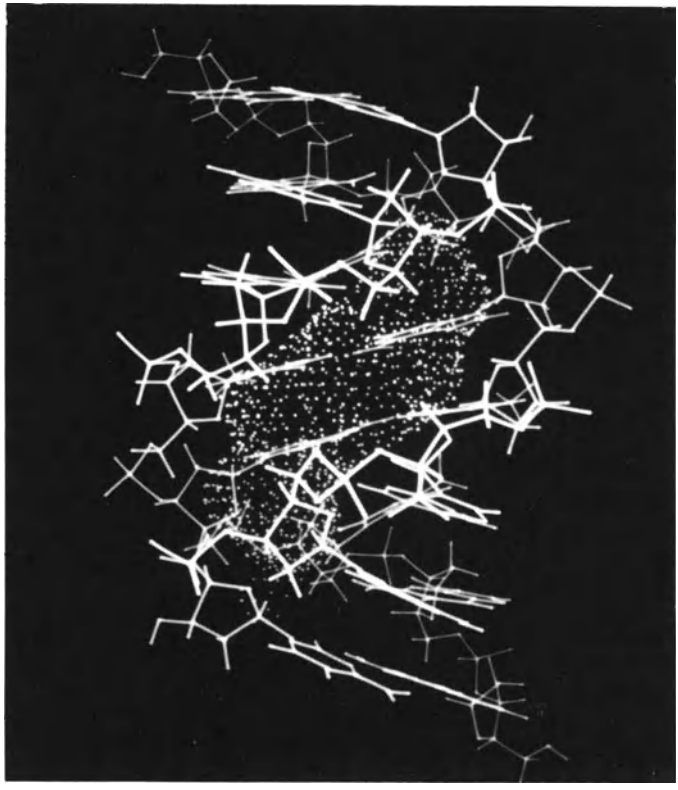


Figure 6. Energy-minimized B-D-B structure of $[d(GGTATACC)]_2$ which best fits the experimental 2D NOE intensity data. Note the interior hydration accessible surface throughout the TATA region.

intensities. These differences will more than likely be lost in a qualitative analysis of the 2D NOE spectrum. Careful attention to detail, and in particular explicit accounting for network relaxation effects using CORMA, allows us to discriminate between very similar structures and find sequence-dependent structural variations in the oligonucleotide.

2.4. Direct Determination of Distances by CORMA

In the preceding discussion on CORMA, we described the calculations necessary to evaluate intensities based on a known structure. In general then, the use of CORMA is limited to the evaluation of several trial structures to determine which of them best approximate the true structure in solution. We have applied this approach to several oligonucleotide systems in our laboratory, and have discovered some interesting structural features that are consistent with the 2D NOE data. However, the success of this approach is limited by our selection of trial structures. An obvious solution to this problem is to apply the same computational techniques used in CORMA to the direct calculation of proton–proton distances from the experimental intensities. In principle this approach should work, and has been discussed elsewhere (Olejniczak *et al.*, 1986). Rearrangement of equation (4) gives:

$$\frac{-\ln\left(\dfrac{\mathbf{a}(\tau_m)}{\mathbf{a}(0)}\right)}{\tau_m} = \mathbf{R} \tag{11}$$

So, assuming isotropic tumbling and using equations (2) and (3), distances can be calculated directly from the intensity matrix. For relatively small systems, in which most of the peaks (in particular the diagonals) can be resolved and accurately estimated, this is clearly the method of choice. However, the effects of peak overlap (low resolution) and low signal-to-noise can severely hamper the accurate estimation of distances.

We have tested the performance of the direct calculation of distances (referred to as the DIRECT method) in the presence of several different types of data errors: random noise, cross-peak overlap, and diagonal peak overlap. CORMA was used to calculate intensities at mixing times of 50 msec and 250 msec for a single strand of d(GGTATACC) in the energy-minimized BDB conformation. Intensities were calculated with a noise level of ±0.003 (which is experimentally typical), or to a precision of five significant figures with no noise. The effects of cross-peak overlap were modeled by selecting for analysis only those intensities which correspond to well-resolved and assignable peaks in an experimental spectrum. The effects of diagonal peak overlap were considered by assigning to each of the diagonals an average (over all the calculated diagonal intensities). The distances calculated by this approach compare favorably with ISPA distances. The results of these

tests are presented in Table I and Figures 4 and 7. Both methods accurately assign the distances for very short interactions and short mixing times. However, even with predicted distances of only 0.3–0.4 nm, the accuracy of the distance calculations is already questionable. ISPA severely underestimates all the distances due to the neglect of spin diffusion and relaxation from nearby protons, while the DIRECT calculation suffers from the neglect of significant relaxation pathways when all the cross-peaks or diagonals cannot be properly assigned or measured. It may still be possible to extract considerable usable information from the DIRECT calculation if a reasonably good idea of the structure is known so that estimated distances which are far off target can be discriminated against while retaining the reasonably accurate distances. The poor prospects for structure determination by this approach in cases with significant peak overlap have led us to pursue the less direct and more time-consuming approach of iterative least-squares refinement of structure based on the 2D NOE intensities.

2.5. Refinement of Structures Based on CORMA

In view of the inherent limitations on direct calculations of distances, and the relative ease with which we can calculate the intensity matrix from

TABLE 1
Comparison of ISPA-, DIRECT-, and COMATOSE-derived distance errors[a] in d(GGTATACC)

Intensities used in calculation[b]	ISPA[c]		DIRECT[d]		COMATOSE[e]	
	50 msec	250 msec	50 msec	250 msec	50 msec	250 msec
±0.003 (all + D)	0.064	0.124	0.045	0.075[f]	0.044	0.043
±0.003 (all + ⟨D⟩)	—	—	0.003	0.076	—	—
±0.003 (resolved + D)	0.066	0.133	0.046	0.067[f]	0.046	0.042
±0.003 (resolved + ⟨D⟩)	—	—	0.041	0.078	—	—
$\pm 5 \times 10^{6}$ (all + D)	0.055	0.125	0.000	0.001	0.048	0.039
$\pm 5 \times 10^{-6}$ (all + ⟨D⟩)	—	—	0.015	0.043	—	—
$\pm 5 \times 10^{-6}$ (resolved + D)	0.058	0.132	0.029	0.075[f]	0.040	0.042
$\pm 5 \times 10^{-6}$ (resolved + ⟨D⟩)	—	—	0.032	0.119	—	—

[a] Errors are RMS errors (nm) for distances that should be between 0.3–0.4 nm.
[b] Intensities were calculated by CORMA for d(GGTATACC) in energy-minimized BDB conformation with precisions of 10^{-3} and 10^{-5} with noise at the level indicated. The notations "all" and "resolved" refer, respectively, to the use of all calculated intensities or only those corresponding to well resolved and assignable intensities in experimental data. The notations "D" and "⟨D⟩" refer, respectively, to the use of calculated diagonal intensities (assuming perfect resolution) or of an average diagonal intensity (worst case assignment of diagonals in case of overlap).
[c] Isolated spin pair approximation. Note that the errors in ISPA are predominantly systematic; ISPA always underestimates distances due to neglect of spin diffusion. The range of actual distances corresponding to a calculated distance of 0.3 nm goes from ±0.05 nm at 50 msec, to ±0.08 nm at 250 msec.
[d] Distances calculated directly from diagonalization of intensities.
[e] Distances from complete relaxation matrix analysis torsion optimized structure. Only distances corresponding to intensities used in analysis are compared. The starting structure for optimization was Arnott B-DNA.
[f] These calculations resulted in negative values for the two smallest eivenvalues of the intensity matrix.

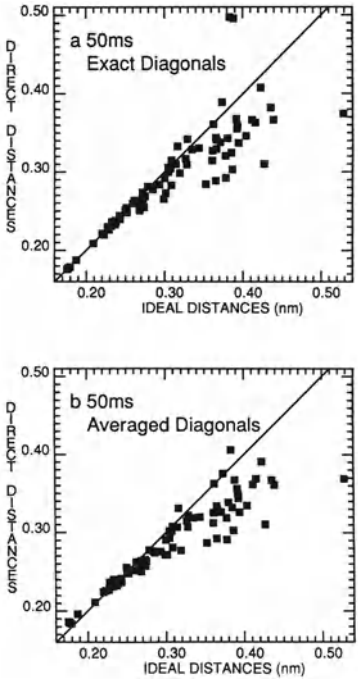


Figure 7. Comparison of distances calculated according to the direct diagonalization of the scaled intensity matrix, i.e., experimental 2D NOE spectrum (DIRECT method) with ideal distances. DIRECT distances were calculated according to equations (2)–(4) and (11). Intensities were calculated by CORMA for d(GGTATACC) in energy-mimimized B-DNA conformation assuming an isotropic correlation time of 4 nsec. (a) 50 msec mixing time, all calculated intensities ≥0.003. (b) 50 msec mixing time, intensities selected to correspond to experimentally resolved peaks. (c) 250 msec mixing time, all calculated intensities ≥0.003. (d) 250 msec mixing time, intensities selected to correspond to experimentally resolved peaks.

a known structure, we have begun developing a program for the refinement of macromolecular structure based on 2D NOE intensities. The goal is to optimize a trial structure while minimizing the error E between calculated and observed intensities:

$$E = \sum_i (I_0^i - kI_c^i)^2 \tag{12}$$

To this end, we have incorporated CORMA into a program with a Marquardt–Levenberg least-squares optimizer.

Several alternative algorithms which specified different variable parameters were tried. One promising algorithm makes use of torsion angles and group orientation parameters rather than the full set of Cartesian coordinates. For nucleic acids, the following parameters are defined for each nucleotide (see Figure 2):

Exocyclic torsion angles: χ and γ

Sugar pucker parameters: P and θ_{max}

Group orientation parameters: α', β', γ' and x, y, z

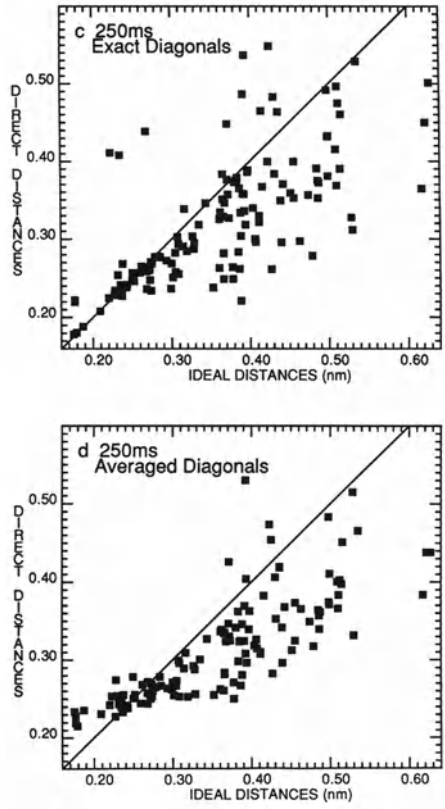


Figure 7. Continued.

The exocyclic torsion angles are obvious parameter choices. The sugar pucker parameters P and θ_{max} allow a reduction of parameters since the ring torsions are correlated and constrained by the condition of ring closure. P and θ_{max} are used to define each of the furanose ring torsions according to the equations (Altona and Sundaralingam; 1972)

$$\theta_j = \theta_{max} \cos[P + (j-2) \cdot 4\pi/5], \qquad j = 0, \cdots, 4 \tag{13a}$$

and

$$\tan P = \frac{\theta_4 + \theta_1 - \theta_3 - \theta_0}{2\theta_2\left(\sin\frac{\pi}{5} + \sin\frac{2\pi}{5}\right)} \tag{13b}$$

These equations are only approximate for the sugar ring in DNA since it is not a regular five-membered ring. But, it is noted by Altona and Sundaralingam (1972), in the majority of cases they studied that the error between input torsion angles for calculating P by equation (13b) and those generated according to equation (13a) is small ($\leq \pm 6°$). A more exact treatment of sugar pucker parameters is that proposed by Cremer and Pople (1975). In their calculations, the pucker is described by a phase and, instead of the maximum torsion, the amplitude of deviation from a mean plane through the sugar atoms. While the accuracy of parameterizing the pucker by this method is laudable, it is computationally cumbersome to reverse the process and generate atomic coordinates for a specified phase and amplitude. Hence, we have chosen not to calculate the parameters according to this latter approach.

Finally, each nucleotide is described as a freely floating entity, its location described by the group orientation parameters. Each nucleotide position is completely independent of the others, being held in place only by the constraints imposed by the fit between calculated and observed internucleotide NOE intensities. The internucleotide torsion angles (α, β, ε, and ζ) are not specifically included in the description of the structure, although constraints such as the $O5' - P$ or $O3' - P$ distances could be included during the refinement. This yields ten structural parameters for each nucleotide. An alternative might be to link the entire polymer together, resulting in only eight parameters per nucleotide. To do this, however, would require propagating the coordinate changes along the whole chain (affecting up to half of the atoms) for each change in α, γ, ε, ζ and P or θ_{max}. An alternative methodology for describing and optimizing oligonucleotide structures by torsion angle perturbations has been described (Sklenar *et al.*, 1986; Lavery *et al.*, 1986a; 1986b).

The two primary rewards for choosing torsion angles as variable parameters in the refinement are the following:

1. A reduction in the number of parameters is achieved. With the parameter set described above there are only ten structural parameters for each nucleotide versus 24–30 proton coordinates, and ~100 coordinates if all atoms are included.
2. Chemically meaningful structural constraints are automatically applied. By varying only the torsion angles, all bond distances and angles will remain at reasonable values (although no consideration is curently made for unfavorable energetic or van der Waals interactions).

As an illustration of these benefits, an initial algorithm which varied only the proton coordinates without any restraints or constraints suffered from a pathological tendency for all the protons to gravitate toward the center

of mass. This gravitation is due in part to the consideration of only observed (nonzero) intensities. The consequence of this is that the weakest included intensities will always tend to be too large rather than too small (in which case they would tend to be unobserved). This imposes a systematic positive bias on the weakest intensities included in the analysis. The end result is a tendency to underestimate all long distances. Moreover, the end result—a set of proton coordinates—was in no way related to any physically achievable structure.

The optimization of the structure based on torsion angles, called COMATOSE for complete matrix analysis torsion optimized structure, generally proceeds much more successfully. Tests with a tetranucleotide sequence d(TATA) showed considerable improvement in the estimated proton–proton distances. Further tests with the octanucleotide d(GGTATACC), the same sequence used in evaluating the ISPA and direct distances above, also reveal that some improvement in the determination of distances is obtainable (Figure 8 and Table I). There are limitations with use of this approach

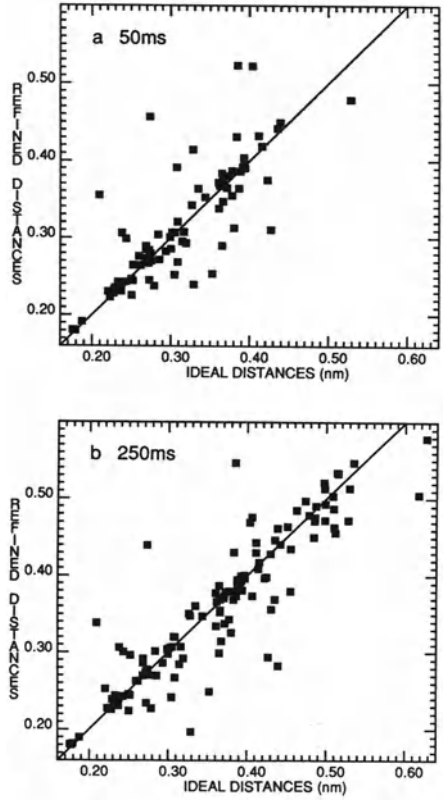


Figure 8. Comparison of distances obtained from COMATOSE refinement with ideal distances. Intensities were calculated by CORMA for d(GGTATACC) in energy minimized B-DNA conformation assuming an isotropic correlation time of 4 nsec. The starting conformation was the Arnott B-DNA conformation. (a) 50 msec mixing time, selected intensities corresponding to experimentally resolved peaks. (b) 250 msec mixing time, selected intensities corresponding to experimentally resolved peaks.

alone, however. In particular, clashes between atoms are not precluded (no van der Waals or energy terms are currently included), and in fact, clashes have occurred with the longer sequences. However, the proton–proton distances corresponding to the cross-peaks used in the analysis generally appear to be more reliable than those obtained from the other methods and can potentially be of considerable value as targets in molecular mechanics and dynamics calculations or for generating alternative trial structures with distance geometry. Moreover, no information is needed from the questionable diagonal peak intensities, and overlapping cross-peaks pose no problem (other than loss of information) since the intensities are simply not included in calculation of the error.

2.5.1. Effects of Internal Motion—Estimation of Correlation Times

The problem of accounting for internal motion in detail is not a trivial one (Hertz, 1983). Earlier work in our lab (Keepers and James, 1982) described the calculations necessary for explicit treatment of internal motion assuming two different jump models to describe the internal motion in oligonucleotides. While this study showed that it is possible to account for the dynamical properties of oligonucleotides in a more or less exact manner, we currently assume effective isotropic motion in our analyses. This assumption has been shown (Keepers and James, 1984) to introduce relatively small errors with respect to distances (distance errors of ~10%). We have estimated effective isotropic correlation times alternatively from nonselective T_1 measurements, or the ratio between T_1 and T_2 (Suzuki *et al.*, 1986):

$$\tau_c \approx \frac{2}{\omega}\left(\frac{T_1}{3T_2}\right)^{1/2} \tag{14}$$

which is valid in the limit $\omega\tau_c \gg 1$. Better estimates of the correlation time are obtainable from selective T_1 measurements. Ideally, the 2D NOE experiment supplies all the information necessary to obtain correlation times. Tests using synthetic intensity sets have shown that refinement of the overall effective isotropic correlation time is feasible along with the COMATOSE structure refinement. We are also pursuing the feasibility of refining local effective correlation or diffusion times for each nucleotide unit, in a manner reminiscent of the use of isotropic temperature factors in crystallography.

2.5.2. Multiple Exchanging Conformations

A related problem is the possibility of multiple conformations and fluctional behavior. It is well known that the energy barrier separating

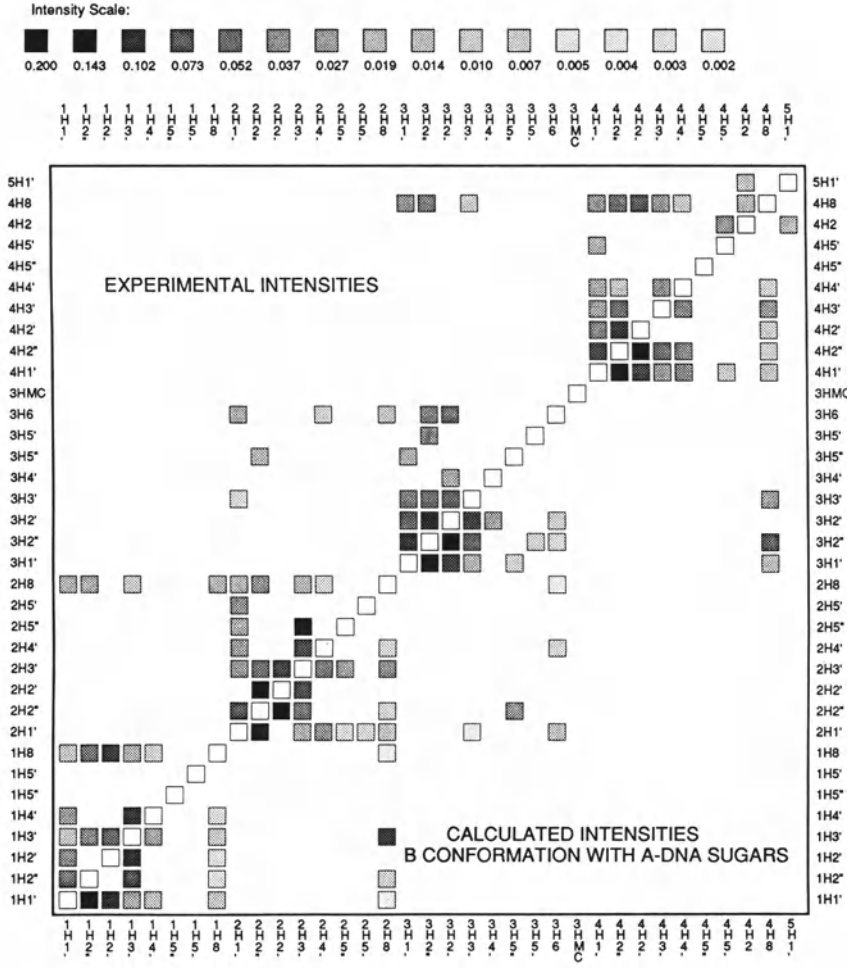


Figure 9. Schematic plot of intrastrand 2D NOE intensities for the first four nucleotides in the duplex $[d(GGTATACC)]_2$. Conditions: 250 msec mixing time, 4 nsec isotropic correlation time. Intensities above the diagonal correspond to experimentally obtained and well resolved intensities. Intensities below the diagonal correspond to CORMA-generated intensities for a model with overall B-DNA conformation but with A-DNA sugars. Only those intensities which correspond to experimental intensities are shown below the diagonal. Experimental and calculated intensities have been normalized by the ratio between the sums of the displayed cross-peak intensities (i.e., the sums of the scaled intensities in the two halves of the plot are the same).

C3′-endo and C2′-endo conformations of furanose rings is small; in fact, it is calculated to be ~0.6 Kcal mol^{-1} (Levitt and Warshel, 1978). Consequently, the possibility of rapid interconversion between multiple conformations is a real one in oligonucleotides (Keepers and James, 1982; Levy *et al.*, 1983). Since each nucleotide can act independently of all the rest, it is unlikely that the whole oligonucleotide strand fluctuates between C3′-endo and C2′-endo in concert, even though some torsional motions are apparently

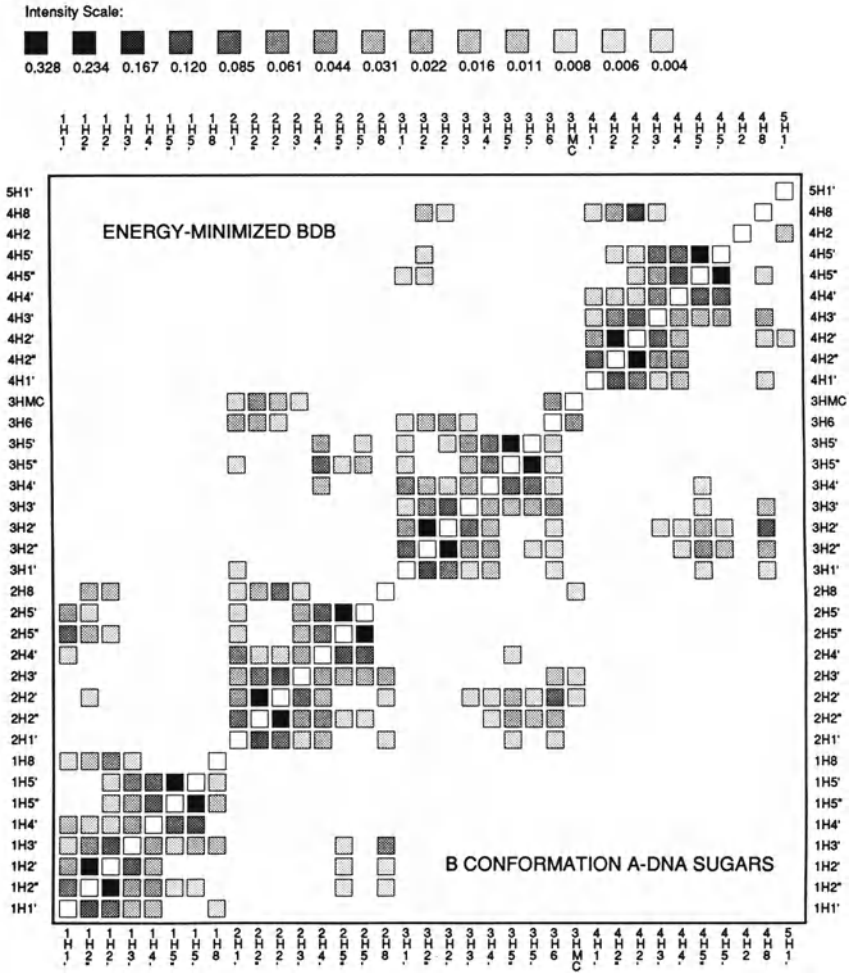


Figure 10. Schematic plot of CORMA-generated intensities for [d(GGTATACC)]$_2$ in energy-minimized BDB conformation and in B conformation with A-DNA sugars. Conditions: 110 msec mixing time, 4 nsec isotropic correlation time. All intensities greater than 0.003 are shown. Normalization is performed as in Figure 9.

correlated (Kitimura *et al.*, 1981). Consequently, there are $2^8 = 256$ possible ways of describing the conformation of the single strand octamer considering only the two limiting furanose conformations. The 2D NOE intensities actually reflect the time and ensemble average over all these conformations. So at worst, the problem is totally untractable; there are barely enough intensities to warrant refinement of a single structure to say nothing of 256 interconverting structures. Optimism must prevail. Since well-resolved peaks are observed (multiple conformations are not evident from the spectrum), it is reasonable to assume that the system can be defined by a single structure or a family of closely related structures. The structure obtained from analysis of the 2D NOE spectrum will then be an average over all members of this family. This is the best than can be expected. From the details of the structure so obtained, some additional information can possibly be extracted. If a structure is partially incompatible with reasonable energetic constraints, this is perhaps an indication that regions of the structure are poorly defined (the average structure is between energy minima associated with the true interconverting structures). These regions of incompatibility should correlate with known or anticipated structurally mobile regions of the molecule.

An example illustrating these difficulties follows. A method has recently been proposed for determining the conformation of the furanose rings in nucleic acids by means of graphical analysis of the sums of proton–proton scalar coupling constants (Rinkel and Altona, 1987). Subjecting coupling constants obtained primarily from double-quantum-filtered COSY experiments for the $[d(GGTATACC)]_2$ duplex to this analysis results in the prediction that the conformation of the furanose rings is ~80–90% C2′-endo and ~10–20% C3′-endo (Zhou *et al.*, 1988). Although this coupling constant analysis suggests more than one conformation, there is no observed chemical shift multiplicity, and well-resolved 2D NOE spectra are obtained which are compatible with the BDB structure described above (Zhou *et al.*, 1987). Moreover, a model structure built from nucleotides with C3′-endo sugars but with an overall B-DNA conformation yields intensities that are inconsistent (as are the intensities for A-DNA) with the experimental 2D NOE spectra (Figures 9 and 10).

3. EXPERIMENTAL DETAILS OF 2D NOE SPECTROSCOPY

The discussion so far has centered on the evaluation of oligonucleotide structure once 2D NOE spectra have been obtained and assigned and individual intensities accurately measured. While we have shown that oligonucleotide structure can be determined to a relatively high level of precision once data of high quality are obtained, the acquisition of that data poses significant problems in its own right.

3.1. Special Techniques

In this section we discuss some of the problems and artifacts inherent in 2D NOE spectroscopy which can preclude a detailed structure determination. And we describe some of the techniques and precautions that can be used to circumvent these obstacles.

3.1.1. Pure Absorption Spectra

To obtain the highest quality data, pure absorption phase 2D NOE spectra should be acquired (States *et al.*, 1982; Ohuchi *et al.*, 1987)). This technique incorporates the use of phase cycling of the pulses to separately collect the real and imaginary parts of the t_1 dimension. The advantage of this technique are improved resolution and an effective increase in signal-to-noise ratio. This procedure has become routine in recent years. In light of its advantages, there is little justification for not using the pure absorption phase technique in the acquisition of 2D NOE data, even for qualitative purposes.

3.1.2. Suppression of t_1 Noise and Diagonal Peak Intensities

A particularly disturbing artifact of 2D NOE spectra is the appearance of so-called t_1 noise. The effects of t_1 noise are manifested in the spectrum as ridges of intensity which traverse the spectrum parallel to the ω_1 frequency axis at the frequency of strong or sharp resonances. These ridges can obscure or otherwise interfere with the accurate determination of true cross-peak intensities. Instrumental instabilities such as variation of the carrier position or phase instabilities are presumed to be the source of t_1 noise (Mehlkopf *et al.*, 1984). An additional problem is that of so-called t_1 ridges, which probably arise from systematic digitization errors or base-line distortions in the spectra recorded with the first few t_1 values (Denk *et al.*, 1985). A variety of techniques have been proposed for the elimination of these ridges. The most direct is the simple subtraction of the average ridge intensity from the whole slice (Klevit, 1985). Various pulse sequences have also been proposed to eliminate the ridges. Denk *et al.* (1985) describe a pulse sequence that incorporates spectral co-addition of the free induction decay (FID) for a normal 2D NOE spectrum with the FID for a 2D NOE spectrum accumulated with a very short variable delay (VD) in place of the mixing time, and the FID obtained which includes the short VD but which is first inverted by a preparatory π pulse. The effect of this procedure is to suppress both the diagonal and t_1 ridges. The disadvantage is loss of diagonal peak intensity (necessary for complete quantitative analysis to obtain distances directly; see Section 2.4) and a decrease in signal-to-thermal-noise ratio.

Otting *et al.* (1986) have given a more complete description of the source of t_1 ridges. They propose that the ridges arise from improper weighting of just the first data point in the FID, and suggest that data be acquired with sine modulation in t_1 followed by sine Fourier transformation, or that data collected with cosine modulation be corrected by scaling the first point prior to cosine Fourier transformation. This technique is favorable to the indiscriminate suppression of diagonal peak intensity since all intensity relationships are maintained and correspond to those given in equation (4). Other schemes for the suppression of diagonal peak intensity have also been suggested (Bremer *et al.*, 1984; Harbison *et al.*, 1985). These schemes, while perhaps useful in qualitative applications, must be viewed with a critical eye if quantitative use of the 2D NOE intensities is to be made.*

3.1.3. Elimination of Coherent Transfer Effects

Scalar coupling between protons can give rise to J crosspeaks as in 2D correlated spectroscopy (COSY). These arise from coherent magnetization transfer via zero-, single- and double-quantum transitions. The suppression of these effects has been studied in Ernst's lab with the following conclusions:

1. Phase cycling can eliminate all but the zero-quantum coherence (Macura *et al.*, 1981).
2. To eliminate zero-quantum coherence, advantage is taken of the rapid oscillation of the coherence with respect to the mixing time (Macura *et al.*, 1981). A slight randomization of the mixing time ($\tau_m = \tau_m^0 + \tau_m^{\text{var}}$ where $\tau_m^{\text{var}} > 1/\omega_{ZQT}^{\min}$ and $\omega_{ZQT}^{\min}$ is the minimum chemical shift difference between coupled spins) along with the incrementing of the t_1 decay effectively eliminates the J cross-peaks by smearing them out along the ω_1 axis.
3. The zero-quantum coherence can also be partially refocused by insertion of a π pulse midway through τ_m.
4. An alternative method is to retain the zero-quantum coherence transfer, but to displace it along ω_1 by incrementing the mixing time by χt_1 (Macura *et al.*, 1982). This, followed by symmetrization, can effectively eliminate the zero-quantum coherence, although overlap with other true NOE cross-peaks can still be a problem.
5. An optimized procedure combining the above has also been described (Rance *et al.*, 1985). This has the advantage of cancelling the zero-quantum coherence in specific frequency ranges.

* In particular, the technique of Bremer *et al.* (1984) is claimed to be a method that projects diagonal peak intensity onto cross-peaks in a quantifiable manner. However, we have reservations about this claim due to a questionable commutation of matrix multiplication leading to their expression for the mixing coefficients their equations (7), (8), and (9).

Alternatively, an explicit account for J coupling can be included in the analysis of the 2D NOE intensities. Kay *et al.* (1986a) have developed a program for simulation of 2D NOE spectra in the presence of strong J coupling. This program accounts for cross-relaxation (as does CORMA) but also includes the contributions from scalar coupling as well. They have compared the magnitude of error due to neglect of spin diffusion and of scalar coupling (Kay *et al.*, 1986b). For typical conditions (2 nsec correlation time; 50–150 msec mixing times) the relative intensities calculated according to ISPA can be in error by as much as 50%. However, CORMA (which neglects J coupling) results in at most a 10% error except in the case of very strong coupling ($J/\delta \geq 0.5$) and short mixing times (50 msec), where the error was ~30%. The conclusion is that CORMA is probably sufficient in most practical cases where cross-peaks arising from strongly coupled protons are generally not resolvable anyway (Kay *et al.*, 1986b).

3.2. Resonance Assignment

For detailed structural analysis to be at all feasible, most if not all resonances need to be assigned. The assignment procedure usually begins with the first order assignment of the spin types, followed by more detailed analysis ultimately yielding the complete sequential resonance assignment. Typically, the oligonucleotide resonances can be categorized as follows: aromatic protons δ(G-H8, C-H6, A-H2, -H8, T-H6) ~ 8.5 – 7.2 ppm; δ(H1′, C-H5), ~6.3 – 5.4 ppm; δH3′, ~5.1 – 4.5 ppm; δH4′, ~4.5 – 4.2 ppm; δH5′, H5″, ~4.3 – 3.8 ppm; δH2′, H2″, ~3.0 – 1.8 ppm; δT-HM, ~1.6 – 1.2 ppm. Additional specific assignments can be postulated based on results from decoupling experiments and coupling constants (Tran-Dinh *et al.*, 1982), but a more expeditious approach is to obtain specific sequential assignments directly from 2D NMR experiments.

A very successful sequential assignment protocol which depends on the regular structure of right-handed DNA is that devised independently by Scheek *et al.* (1983; 1984), and Hare *et al.* (1983). In this protocol, the close spatial proximity of protons on adjacent nucleotides allows essentially complete assignment of all protons from the 2D NOE spectrum alone (see, for example Broido *et al.*, 1984). The assignment process goes as follows:

1. An aromatic proton (H8 or H6) is close to its own H1′ as well as to the H1′ of its 5′ neighbor. This allows the complete assignment of all the H1′ and H8(6) protons.
2. The H8(6) protons are also near the H2″ proton of the 5′ neighboring nucleotide. This allows an internal check on the assignments generated above, or provides an alternate pathway for assignment in the case of poorly resolved peaks or unobserved cross-peaks.

3. Finally, the remaining protons H3′, H4′, H5′, and H5″ can be assigned by finding crosspeaks with the base or anomeric protons. In general, the H4′ proton yields the lowest field cross-peak with H1′.

Assignments based on this approach have been found to be successful for both A- and B-DNA (Haasnoot *et al.*, 1984), where systematic differences were noted in distinguishing between these structural families. The sequential assignment strategy has been shown to fail in the case of left-handed (Z-type) helices (Feigon *et al.*, 1984).

Distinguishing between the H2′, H2″ and H5′, H5″ protons relies on their differences in chemical shift and relative intensities. For B-family DNA, the H1′ protons are nearer their own H2″ than their own H2′, which allows a check on the individual assignment of these two protons. These can also be distinguished on the basis of chemical shift: the H2′ proton tends to resonate at higher field than the H2″ proton (Jamin *et al.*, 1985; Broido *et al.*, 1985). Also, the H5″ proton is closer to the H3′ proton on its own nucleotide than is the H5′, while H5′ is closer to the H1′ of the 5′ neighbor than is H5″. The corresponding intensity differences can be used in conjunction with the chemical shifts of the tentatively assigned peaks to further aid the assignment process. For B-DNA, the H5″ proton is in a predominately hydrocarbon environment, while the H5′ proton is in a heteroatomic environment. Therefore, the H5′ protons should generally resonate downfield of the H5″ protons (Broido *et al.*, 1985). This is verified by a number of cases where these protons have been assigned by other means (e.g., Sarma, 1980; Tran-Dinh *et al.*, 1982; Frechet *et al.*, 1983).

The preceding arguments for distinguishing between the geminal protons are somewhat structure dependent. Chazin *et al.* (1986) suggest a protocol which calls for the combined use of COSY techniques, to establish the spin systems, along with the sequential assignment based on 2D NOE connectivities. This approach circumvents some of the circular arguments used in making the assignments from the 2D NOE alone and may prove especially useful for systems which deviate substantially from the A-DNA or B-DNA family of structures.

3.3. Intensity Measurement

The techniques for analysis of the 2D NOE intensities which we have been describing rely on accurately measured peak intensities. The accuracy of measured 2D NOE peaks depends on a number of factors, including the method for data acquisition (discussed above) as well as any signal apodization prior to the Fourier transform, and the method of peak evaluation. We briefly mention some aspects of the spectrum processing that should be kept in mind.

3.3.1. Effects of Signal Apodization Prior to Fourier Transform

Some type of apodization which results in resolution enhancement and improvement of signal-to-noise ratio is almost certainly preferable to no signal conditioning prior to the double Fourier transform. The advantages are the ability to resolve cross-peaks in crowded spectral regions and to observe relatively weak peaks with some level of confidence. We have previously discussed the importance of resolving as many peaks as possible (Sections 2.4, 2.5). The inclusion of weak peaks (~0.1% auto peak at mixing time 0) is also important in order to best define the structure. In their review of 2D NMR methods, Wider *et al.* (1984) have discussed the merits of Lorentz–Gauss transforms (Ernst, 1966), sine-bell and shifted sine-bell (Wagner *et al.*, 1978), and of the squared sine-bell functions. Clearly, these apodization procedures are necessary for the qualitative assignment and structure determination process. The pitfall with these forms of signal conditioning is that there will be intensity distortions that depend on the natural linewidth of each resonance. In particular, the unshifted sine-bell function results in sharp wings of negative intensity, which if included in the peak integration will yield zero net intensity. No systematic study of the errors associated with the use of these spectral conditioning techniques with respect to quantitative 2D NOE analysis has been presented. However, it is probably wise to compare intensities obtained using alternative schemes. The intensities measured for peaks of varying width obtained with and without resolution enhancement (sine-bell type functions) should be compared, and a determination regarding the suitability of that apodization can be made. There will probably always be a trade-off between the accuracy of measured intensities and the ability to clearly distinguish peaks.

3.3.2. Intensity Evaluation Protocol

In our early studies using CORMA for detailed structural analyses of oligonucleotides, an attempt was made to evaluate the accuracy of several alternative ways to measure the experimental 2D NOE peak intensities. Several possible alternatives include the assignment of intensities proportional to

1. Peak heights based on actual maxima or interpolated heights.
2. Areas calculated from peak heights and half widths.
3. Integrated peak intensities for the slice containing the maximum.
4. Weights of individual peaks cut from spectra.
5. Peak volumes obtained by adding areas from all slices contributing to peak.
6. The sum of all (ω_1, ω_2) intensities contributing to each peak.
7. Intensities derived from curve fitting.

Curve fitting is the method of choice for most accurate measurement of one-dimensional (1D) NOE effects (Weiss and Ferretti, 1983). However, with 2D NOE spectra which comprise ~500 1D spectra, ~200 slices containing up to 15 measurable peaks each would have to be fitted. This represents an excessive amount of calculation. With all methods, the digital resolution of the spectrum is a major consideration. Each peak is typically defined by only a few points in each dimension so that, depending on the registration between the peak maximum and the frequencies actually represented in the spectrum, the peak maximum will typically underestimate the intensity. Zero-filling is generally applied to the data prior to the Fourier transform to improve the digital resolution (Becker *et al.*, 1979). However, in a typical 2D NOE spectrum which is zero-filled to $1k \times 1k$ points, the resolution is still only ~5 Hz and this problem must be considered. Peak heights alone are a poor estimator of intensity.

Comparison between curve-fitting and area determinations from single slices and sums-of-slices showed that, for well-resolved peaks, agreement between the various intensity estimation methods was within 10% relative error (Broido *et al.* 1985; Jamin *et al.*, 1985). However, in the presence of overlap the error was as large as 25%, with the largest discrepancy being between single-slice and sum-of-slices measurements. Qualitatively, however, the relative ranking of intensities was preserved by each method for estimating the intensities.

With these considerations in mind we have evaluated structures on the basis of agreement between ranked intensities $[d(AT)_5]_2$ (Suzuki *et al.*, 1986) and with intensities estimated by summing all the points within a rectangular region defined around the peak (software written by Drs. S. Manogaran and R. Scheek at University of California, San Francisco) for $[d(GGTATACC)]_2$ (Zhou *et al.*, 1987). Similar conclusions were obtained in both cases.

4. DETAILED MOLECULAR STRUCTURE FROM 2D NOE INTENSITIES

In all the preceding discussion, we have glossed over the details of how a structure is obtained from the 2D NOE results. With either ISPA (Section 2.2.1), or the DIRECT calculation of distances (Section 2.4), the result of the analysis of intensities is only a set of proton–proton distances. With COMATOSE (Section 2.5), using ideal intensities, we might expect to arrive at a reasonable picture of the true structure without any additional steps other than generation of a suitable starting structure. However, even under the best possible experimental conditions, such a result is probably not attainable. The reasons for this include limited signal-to-noise ratios

and overlapping peaks which limit the amount of information which we can supply for analysis in the form of observables. There may also be systematic errors due to residual t_1 noise, coherent transfer effects, or cross-correlation effects that are not taken into account by CORMA. In addition, and perhaps most significant, is the fact that, depending on how close the trial structure is to the ideal structure, the trajectory of the least-squares optimization will usually lead to a local minimum that does not fully approximate the true structure. However, we have found that even in the presence of all these limitations, we are able to improve the RMSD error in distances between the refined and ideal structures. So again, the end result is a set of proton-proton distances that are reasonably accurate. And, if several models have been analyzed, some estimate of the errors on those distances is obtainable.

The routes to a complete structural model after obtaining an accurate set of distances are the same regardless of the source of the distances. Typically a set of interproton distances will be obtained which can be used as targets in energy refinement or molecular dynamics calculations or as distance limits for the generation of new structures by distance geometry. The success of these various structure completion approaches has already been illustrated when used in conjunction with distances obtained with the ISPA approximation. We envision a coordinated approach to the problem of structure determination which uses several, if not all of the techniques in an iterative manner to fully develop an accurate model of the macromolecular structure. The overall scheme is illustrated in Figure 11. Accordingly, ISPA distances might be useful in generating a suitable trial structure for further refinement by COMATOSE. Then, depending on the agreement between the refined distances obtained from several trial models, energy refinement or additional distance geometry calculations could be run using the more accurate distances obtained from COMATOSE as

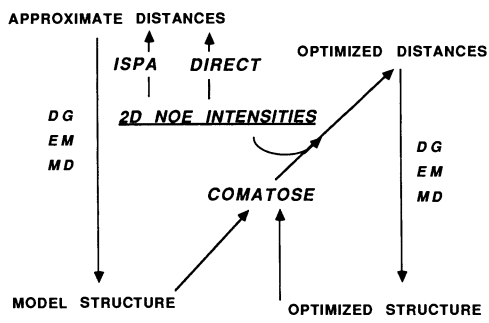


Figure 11. Scheme for the complete structure determination of a macromolecule. DG, EM, and MD refer to distance geometry, energy minimization, and molecular dynamics, respectively.

constraints. The final structure obtained by this process will not be overly biased by an arbitrary selection of trial structures and will satisfy the conditions imposed by energetic, steric, and geometric constraints including the specific conditions of being consistent with the experimentally observed 2D NOE intensities.

Note added in proof. Since this chapter was written, a number of articles have been published on various aspects described within. The CORMA (Borgias *et al.*, 1987/1989) and COMATOSE algorithms have been described in detail (Borgias and James, 1988), and some significant advances in the DIRECT calculation of distances have been made (Boelens *et al.*, 1988; 1989; Borgias and James, 1989a). More current reviews of relaxation matrix analysis can also be found (Borgias and James, 1989b; Borgias *et al.*, 1990).

REFERENCES

Abel, E. W., Coston, T. P. J., Orrell, K. G., Ŝik, V., and Stephenson, D., 1986, *J. Mag. Res.* **70**:34.

Altona, C., and Sundaralingam, M., 1972, *J. Am. Chem. Soc.* **94**:8205.

Arnott, S., and Hukins, D. W. L., 1972, *Biochem. Biophys. Res. Commun.* **47**:1504.

Arnott, S., and Hukins, D. W. L., 1973, *J. Mol. Biol.* **81**:93.

Arnott, S., and Selsing, E., 1974, *J. Mol. Biol.* **88**:509.

Arnott, S., Chandrasekaran, R., Puigjaner, L. C., Walker, J. K., Hall, I. H., Birdsall, D.L., and Ratliff, R. L., 1983, *Nucleic Acids Res.* **11**:1457.

Assa-Munt, N., and Kearns, D. R., 1984, *Biochemistry* **23**:791.

Becker, E. D., Ferretti, J. A., and Gambhir, P. N., 1979, *Anal. Chem.* **51**:1413.

Boelens, R., Scheek, R. M., Dijkstra, K., and Kaptein, R., 1985, *J. Mag. Res.* **62**:378.

Boelens, R., Koning, T. M. G., and Kaptein, R., 1988, *J. Mol. Struct.* **173**, 299–311.

Boelens, R., Koning, T. M. G., van der Marel, G. A., van Boom, J. H., and Kaptein, R., 1989, *J. Magn. Reson.* **82**, 290–308.

Borah, B., Cohen, J. S., and Bax, A., 1985, *Biopolymers* **24**:747.

Borgias, B. A., Thomas, P. D., and James, T. L., 1987/1989, COMPLETE RELAXATION MATRIX ANALYSIS (CORMA), University of California, San Francisco.

Borgias, B. A., and James, T. L., 1988, *J. Magn. Reson.* **79**, 493–512.

Borgias, B. A., and James, T. L., 1989a, Two-dimensional nuclear Overhauser effect: complete relaxation matrix analysis, in *Methods in Enzymology* (Oppenheimer, N. J., and James, T. L., eds.), vol. 176, Academic Press, San Diego.

Borgias, B. A., and James, T. L., 1989b, *J. Magn. Reson.* (to be published).

Borgias, B. A., Gochin, M., Kerwood, D. J., and James, T. L., 1990, Relaxation matrix analysis of 2D NMR data, in *Progress in Nuclear Magnetic Resonance Spectroscopy* (Emsley, J. W., Feeney, J., and Sutcliffe, L. H., eds.), vol. XX, Pergamon, Oxford (to be published).

Braun, W., and Gō, N., 1985, *J. Mol. Biol.* **186**:611.

Braun, W., Wagner, G., Wörgötter, E., Vaŝak, M., Kägi, J. H. R., and Wüthrich, K., 1986, *J. Mol. Biol.* **187**:125.

Bremer, J., Mendz, G. L., and Moore, W. J., 1984, *J. Am. Chem. Soc.* **106**:4691.

Broido, M. S., Zon, G., and James, T. L., 1984, *Biochem. Biophys. Res. Commun.* **19**:663.

Broido, M. S., James, T. L., Zon, G., and Keepers, J. W., 1985, *Eur. J. Biochem.* **150**:117.

Bull, T. E., 1987, *J. Mag. Res.* **72**:397.
Chary, K. V. R., Hosur, R. V., Govil, G., Zu-kun, T., and Miles, H. T., 1987, *Biochemistry* **26**:1315.
Chazin, W. J., Wüthrich, K., Hyberts, S., Rance, M., Denny, W. A., and Werner, L., 1986, *J. Mol. Biol.* **190**:439.
Clore, G. M., and Gronenborn, A. M., 1984, *FEBS Lett.* **175**:117.
Clore, G. M., Gronenborn, A. M., Piper, E. A., McLaughlin, L. W., Graeser, E., and Van Boom, J. H., 1984, *Biochem. J.* **221**:737.
Clore, G. M., Gronenborn, A. M., and McLaughlin, L. W., 1985a, *Eur. J. Biochem.* **151**:153.
Clore, G. M., Gronenborn, A. M., Moss, D. S., and Tickle, I. J., 1985b, *J. Mol. Biol.* **185**:219.
Clore, G. M., Sukumaran, D. K., Nilges, M., and Gronenborn, A. M., 1987a, *Biochemistry* **26**:1732.
Clore, G. M., Sukumaran, D. K., Nilges, M. Zarbock, J., and Gronenborn, A. M., 1987b, *EMBO J.* **6**:529.
Cremer, D., and Pople, J. A., 1975, *J. Am. Chem. Soc.* **97**:1354.
Crippen, G. M., 1984, *J. Comp. Chem.* **5**:548.
Davies, D. R., and Baldwin, R. L., 1963, *J. Mol. Biol.* **6**:251.
Denk, W., Wagner, G., Rance, M., and Wüthrich, K., 1985, *J. Mag. Res.* **62**:350.
Dickerson, R. E., and Drew, H. R., 1981, *J. Mol. Biol.* **149**:761.
Drew, H. R., and Dickerson, R. E., 1982, *EMBO J.* **1**:663.
Ernst, R. R., 1966, *Advan. Magn. Res.* **2**:1.
Feigon, J., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., and Rich, A., 1984, *Nucl. Acids Res.* **12**:1243.
Frechet, D., Clem, D. M., Kan, L.-S., and Ts'o, P. O. P., 1983, *Biochemistry* **22**:5194.
Frederick, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobsen, L., Wang, B.-C., Greene, P., Boyer, H. W., and Rosenberg, J. M., 1984, *Nature* **309**:327.
Gronenborn, A. M., and Clore, G. M., 1985, Investigation of the solution structures of short nucleic acid fragments by means of NOE measurements, in *Progress in Nuclear Magnetic Resonance Spectroscopy* (Emsley, J. W., Feeney, J., and Sutcliffe, L. H., eds.), vol. 17, Pergamon, Oxford, pp. 1–32.
Gronenborn, A. M., Clore, G. M., Jones, M. B., and Jiricny, J., 1984a, *FEBS Lett.* **165**:216.
Gronenborn, A. M., Clore, G. M., and Kimber, B. J., 1984b, *Biochem. J.* **221**:723.
Gupta, G., Sarma, M. H., Dhingra, M. M., Sarma, R. H., Rajagopalan, M., and Sasisekharan, V., 1983, *J. Biomol. Struct. Dyn.* **1**:395.
Haasnoot, C. A. G., Westerink, H. P., van der Marel, G. A., and van Boom, J. H., 1984, *J. Biomol. Struct. Dynam.* **2**:345.
Harbison, G. S., Feigon, J., Ruben, D. J., Herzfeld, J., and Griffin, R. G., 1985, *J. Am. Chem. Soc.* **107**:5567.
Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, B., and Reid, B. R., 1983, *J. Mol. Biol.* **171**:319.
Havel, T. F., and Wüthrich, K., 1985, *J. Mol. Biol.* **182**:281.
Hertz, H. G., 1983, The problem of intramolecular rotation in liquids and NMR, in *Progress in Nuclear Magnetic Resonance Spectroscopy* (Emsley, J. W., Feeney, J., and Sutcliffe, L. H., eds.), vol. 16, Pergamon, Oxford, pp. 115–162.
Jamin, N., James, T. L., and Zon, G., 1985, *Eur. J. Biochem.* **152**:157.
Jeener, J., 1971, lecture given at Ampere International Summer School II, Basko Polje, Yugoslavia.
Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R., 1979, *J. Chem. Phys.* **71**:4546.
Kay, L. E., Scarsdale, J. W., Hare, D. R., and Prestegard, J. H., 1986a, *J. Mag. Res.* **68**:515.
Kay, L. E., Holak, T. A., Johnson, B. A., Armitage, I. M., and Prestegard, J. H., 1986b, *J. Am. Chem. Soc.* **108**:4242.

Keepers, J. W., and James, T. L., 1982, *J. Am. Chem. Soc.* **104**:929.
Keepers, J. W., and James, T. L., 1984, *J. Mag. Res.* **57**:404.
King, G. C., and Coleman, J. E., 1987, *Biochemistry* **26**:2929.
Kitimura, K., Wakahara, A., Mizuno, H., Baba, H., and Tomita, K., 1981, *J. Am. Chem. Soc.* **103**: 3899.
Klevit, R., 1985, *J. Mag. Res.* **62**:551.
Kline, A. D., Braun, W., and Wüthrich, K., 1986, *J. Mol. Biol.* **189**:377.
Klug, A., Jack, A., Viswamitra, M. A., Kennard, O., Shakked, Z., and Steitz, T. A., 1979, *J. Mol. Biol.* **131**:669.
Kumar, A., Ernet, R. R., and Wüthrich, K., 1980, *Biochem. Biophys. Res. Commun.* **95**:1.
Kumar, A., Hosur, R. V., Chandrasekhar, K., and Murali, N., 1985, *J. Mag. Res.* **63**:107.
Kuntz, I. D., Crippen, G. M., and Kollman, P. A., 1979, *Biopolymers* **18**:939.
Lavery, R., Sklenar, H., Zakrzewska, K., and Pullman, B., 1986a, *J. Biomol. Struct. Dynam.* **3**:989.
Lavery, R., Sklenar, H., and Pullman, B., 1986b, *J. Biomol. Struct. Dynam.* **3**:1015.
Levitt, M., and Warshel, A., 1978, *J. Am. Chem. Soc.* **100**:2607.
Levy, G. C., Marchetti, P. S., Ejchart, A., Levy, L. F., Kumar, A., Hilliard, P. R., and Rill, R. L., 1983, *J. Biomol. Struct. Dynam.* **1**:795.
Macura, S., and Ernst, R. R., 1980, *Mol. Phys.* **41**:95.
Macura, S., Huang, Y., Suter, D., and Ernst, R. R., 1981, *J. Mag. Res.* **43**:259.
Macura, S., Wüthrich, K., and Ernst, R. R., 1982, *J. Mag. Res.* **46**:269.
Macura, S., Farmer, II, B. T., and Brown, L. R., 1986, *J. Mag. Res.* **70**:493.
Marvin, D. A., Spencer, M., Wilkins, M. H. F., and Hamilton, L. D., 1961, *J. Mol. Biol.* **3**:547.
Massefski, W., Jr., and Bolton, P. H., 1985, *J. Mag. Res.* **65**:526.
Mehlkopf, A. F., Korbee, D., Tiggelman, T. A., and Freeman, R., 1984, *J. Mag. Res.* **56**:315.
Mirau, P. A., 1988, *J. Mag. Res.* **80**:439.
Ohkubo, T., Kobayashi, Y., Shimonishi, Y., and Kyoguku, Y., 1986, *Biopolymers* **25**:S123.
Ohuchi, M., Hosono, M., Furihata, K., and Seto, H., 1987, *J. Mag. Res.* **72**:279.
Olejniczak, E. T., Gampe, R. T., Jr., and Fesik, S. W., 1986, *J. Mag. Res.* **67**:28.
Otting, G., Widmer, H., Wagner, G., and Wüthrich, K., 1986, *J. Mag. Res.* **66**:187.
Rance, M., Bodenhausen, G., Wagner, G., Wüthrich, K., and Ernst, R. R., 1985, *J. Mag. Res.* **62**:497.
Rinkel, L. J., and Altona, C., 1987, *J. Biomol. Struct. Dynam.* **4**:621.
Sarma, R. H., 1980, in *Nucleic Acid Geometry and Dynamics* (Sarma, R. H., ed.), Pergamon Press, Elmsford, N.Y., pp. 143–155.
Scheek, R. M., Boelens, R., Russo, N., Kaptein, R., and van Boom, J. H., 1983, *J. Am. Chem. Soc.* **105**:2914.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., and Kaptein, R., 1984, *Biochemistry* **23**:1371.
Shakked, Z., Rabinovich, D., Kennard, O., Cruse, W. B. T., Salisbury, S. A., and Viswamitra, M. A., 1983, *J. Mol. Biol.* **166**:183.
Sklenar, H., Lavery, R., and Pullman, B., 1986, *J. Biomol. Struct. Dynam.* **3**:967.
Solomon, I., 1955, *Phys. Rev.* **99**:559.
States, D. J., Haberkorn, R. A., and Ruben, D. J., 1982, *J. Mag. Res.* **48**:286.
Suzuki, E., Pattabiraman, N., Zon, G., and James, T. L., 1986, *Biochemistry* **25**:6854.
Tran-Dinh, S., Neumann, J. M., Huynh-Dinh, T., Igolen, J., and Kan, S. K., 1982, *Org. Mag. Res.* **18(3)**:148.
Wagner, G., Wüthrich, W., and Tshesche, H., 1978, *Eur. J. Biochem.* **86**:67.
Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van der Marel, G., and Rich, A., 1979, *Nature* **282**:680.
Weiner, P. K., and Kollman, P. A., 1981, *J. Comput. Chem.* **2**:287.

Weiner, P. K., Singh, U. C., Kollman, P. A., Caldwell, J., and Case, P. A., 1984, *AMBER: A Molecular Mechanics and Dynamics Program*, University of California, San Francisco.
Weiss, G. H., Ferretti, J. A., 1983, *J. Mag. Res.* **55**:397.
Werbelow, L., and Grant, D. M., 1978, *Adv. Mag. Res.* **9**:189.
Wider, G., Macura, S., Kumar, A., Ernst, R. R., and Wüthrich, K., 1984, *J. Mag. Res.* **56**:207.
Williamson, M. P., Havel, T. F., and Wüthrich, K., 1985, *J. Mol. Biol.* **182**:295.
Young, G. B., and James, T. L., 1984, *J. Am. Chem. Soc.* **106**:7986.
Zhou, N., Bianucci, A. M., Pattabiraman, N., and James, T. L., 1987, *Biochemistry* **26**:7905.
Zhou, N., Manogoran, M., Zon, G., James, T. L., 1988, *Biochemistry* **27**:6013.

4

Methods of Proton Resonance Assignment for Proteins

Andrew D. Robertson and John L. Markley

1. INTRODUCTION

Correlation of protein activity with structure is essential to researchers seeking an understanding of life processes at the molecular level. It is becoming important as well to commercial enterprises, where protein production is geared toward medical, industrial, and agricultural applications. The number of identified proteins is growing faster than the number of proteins for which structural details are known. This dilemma has been exacerbated by the advent of site-directed mutagenesis and the resulting ease with which variant proteins can be produced. In the last few years, NMR spectroscopy has been called upon with increasing frequency for the necessary structural information.

Modern ^{1}H NMR spectroscopy can provide atomic-level information about protein conformation and dynamics. The chemical shift is sensitive to the local electronic environment but, with the possible exception of aromatic ring effects (Perkins, 1982), is not interpretable in terms of peptide conformation. Other features of the spectrum are much more informative. Two- and three-bond scalar coupling, manifested as resonance splitting, can be used to determine dihedral angles (Karplus, 1959). Interproton

Andrew D. Robertson and John L. Markley • Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706. Andrew D. Robertson's present address is Department of Biochemistry, Beckman Center, Stanford University, Stanford, California 94305.

distances can be measured through observation of the nuclear Overhauser effect (NOE), which results from dipolar (through-space) coupling between protons (Noggle and Schirmer, 1971). In small proteins, dipolar interactions between protons separated by up to 5 Å can be detected (Wüthrich, 1986) and have been used to determine the solution structures of small globular proteins (Williamson *et al.*, 1985). The protein's electrostatic environment can be probed through determination of the pK'_a values for ionizing groups within the protein (Markley, 1975). Dynamical features of proteins can be monitored by lineshape analysis, measurement of spin relaxation rates, and determination of rates for the exchange of labile protons with solvent (Wagner and Wüthrich, 1986).

Much of the information in the ^{1}H NMR spectrum of a protein would be inaccessible were it not for the development of 2D NMR methodology (Aue *et al.*, 1976; Bax and Lerner, 1986; Wüthrich, 1986; Ernst *et al.*, 1987). Even with the high-field spectrometers now in use, the typical 1D spectrum of a protein is severely complicated by resonance overlap. With the availability of 2D NMR experiments, resolution is no longer limiting with small proteins ($M_r < 10{,}000$). The problem now lies in assigning resolved resonances to specific protons in the protein.

Our aim in this chapter is to provide a practical overview of current proton resonance assignment strategies for proteins, with an emphasis on refinements in the sequential assignment method originally described by Wüthrich and co-workers (1982). While mindful of the recent advances in the use of isotopic enrichment for proton resonance assignment (LeMaster and Richards, 1988), we limit our discussion to proteins containing natural abundance levels of NMR-sensitive isotopes; indeed most investigations of the structure of small proteins with NMR are likely to be conducted with such proteins. We address the investigator who is interested in applying NMR spectroscopy to the characterization of large peptides and small globular proteins but who may be daunted by the seemingly complex technical and theoretical aspects of NMR. Our experience indicates that a person with some background in 1D NMR spectroscopy and access to modern high-field spectrometers can obtain detailed information about a protein's structure in less than a year; prior experience with 2D NMR can decrease this time significantly. For discussion of the principles of 2D NMR, we refer the reader to the article by Bax and Lerner (1986) and the comprehensive treatise by Ernst *et al.* (1987). A more detailed treatment of some of the material covered in this chapter can be found in the book by Wüthrich (1986). The present discussion is illustrated with examples from work on turkey ovomucoid third domain (OMTKY3), a 6100-Dalton protein isolated after limited proteolysis of whole ovomucoid (Robertson *et al.*, 1988).

1.1. Features of the Proton NMR Spectrum of a Protein

Proton NMR resonances for proteins usually fall into one of five partially overlapping regions of the spectrum (Figure 1), depending on the nature of the chemical group in which the proton is located. Amide and aromatic proton resonances are found in the low-field portion of the spectrum (regions a and b in Figure 1); protons attached to peptide backbone α-carbons resonate at frequencies around the middle of the spectrum (region c); and resonances from the aliphatic protons of the amino acid side chains are generally found at higher field (regions d and e). The spectrum of heat-denatured protein is simpler than that of native protein and resembles the theoretical spectrum of the constituent free amino acids. The conformation-dependent dispersion of chemical shifts results from the electronic microenvironments created by the protein fold. Resolution and assignment of NMR resonances are facilitated by this dispersion, insofar as different residues of a given type have different chemical shifts. In spite of this, the 1D spectrum of a protein such as OMTKY3 remains severely complicated by overlap of signals from hundreds of protons. The overlap problem is best addressed by making use of a variety of 2D NMR experiments that

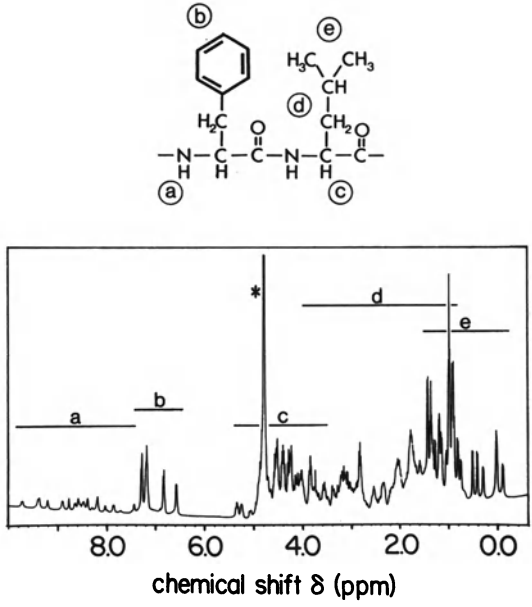


Figure 1. One-dimensional 500 MHz ^{1}H NMR spectrum of OMTKY3 freshly dissolved in ^2H_2O. See the text for a description of the regions of the spectrum.

spread the chemical shift information into two dimensions (Bax and Lerner, 1986). These methods generally permit resolution of a majority of the resonances in a protein whose mass is less than 10 kDaltons.

The results of 2D NMR experiments usually are presented as contour plots. Figure 2 shows a 2D *co*rrelated *s*pectroscop*y* (COSY) contour plot of OMTKY3. In most such representations, signals along the diagonal, running from the upper right to the lower left in Figure 2, correspond to the one-dimensional spectrum. Cross-peaks arranged symmetrically above and below the diagonal represent connectivities produced by the particular kind of 2D NMR experiment employed: through-bond scalar coupling (COSY, RELAY, HOHAHA) or through-space dipolar coupling (NOESY; see Section 5, below). In the COSY plot of Figure 2, the crosspeaks connect resonances of protons that are separated by two or three bonds; most of

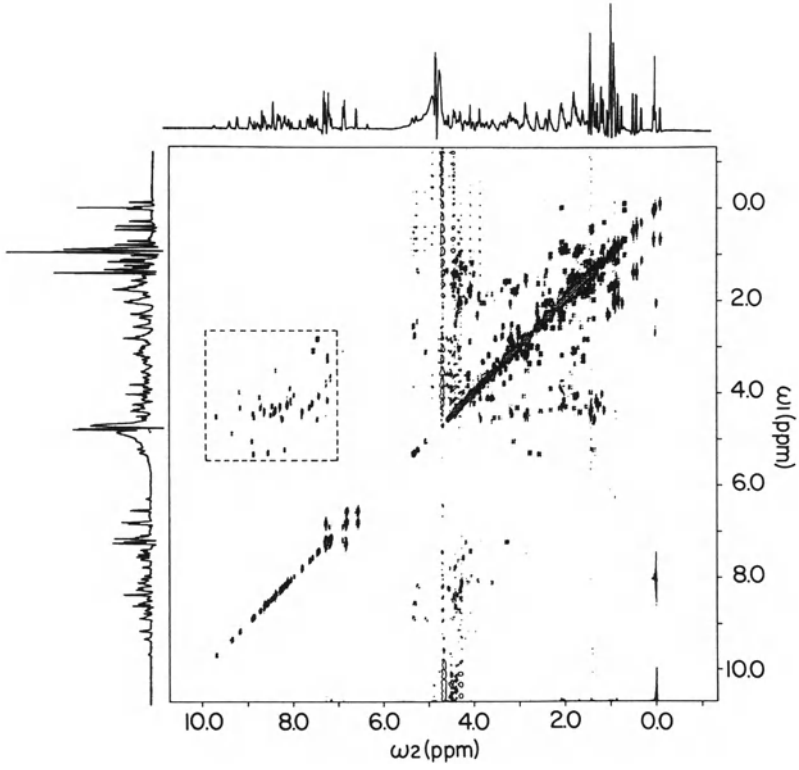


Figure 2. Phase-sensitive double-quantum-filtered COSY of OMTKY3 in H_2O. The region enclosed by the dashed box contains cross-peaks due to the coupling between backbone $C_\alpha\underline{H}$ and $N\underline{H}$ and is referred to as the "fingerprint" region of the COSY map. Much of the noise has been suppressed using the MAKEUP program (Zolnai *et al.*, 1986).

the peaks in the boxed region represent three-bond scalar coupling between the peptide backbone alpha and amide protons. Resonance intensities are indicated by the number of contours that make up each feature.

1.2. Proton NMR Assignments in Proteins

In some cases ^{1}H NMR assignments can be based on empirical rules for chemical shifts. The aromatic resonances of histidine provide one such example. The $C_\varepsilon\underline{H}$ of a histidine residue gives rise to a singlet peak in the low-field aromatic region whose chemical shift normally has a characteristic pH dependence. This region of the spectrum for staphylococcal nuclease (Figure 3) contains $C_\varepsilon\underline{H}$ resonances arising from four histidine residues.

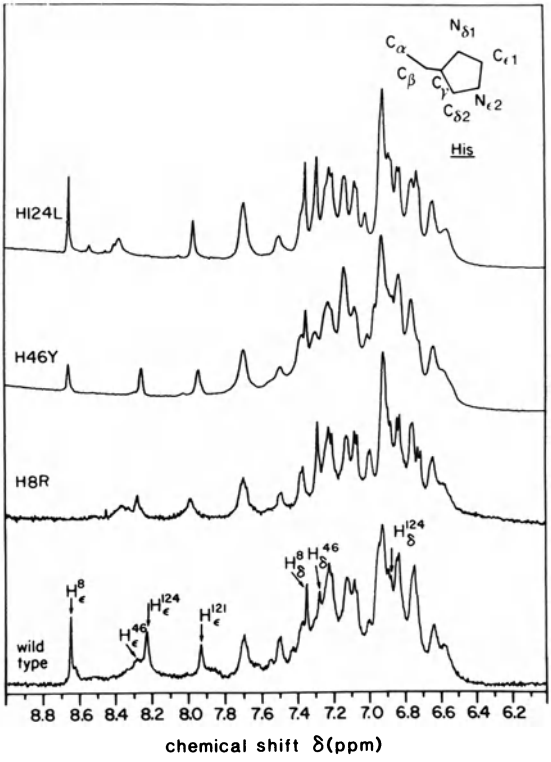


Figure 3. ^{1}H NMR spectra (500 MHz) of wild type staphylococcal nuclease and three mutants used to assign resonances of the four histidine residues in nuclease (Alexandrescu *et al.*, 1988). The wild-type nuclease has histidines at positions 8, 46, 121, and 124. In mutant H8R, histidine 8 is replaced by arginine; in mutant H46Y, histidine 46 is replaced by tyrosine; and in mutant H124L, histidine 124 is replaced by leucine.

Unambiguous assignment of these resonances to specific histidine residues in the nuclease structure was made using another assignment technique: comparison of spectra from proteins differing by single amino acid substitutions (Figure 3). An early assignment technique (Markley, 1972), which is fast regaining popularity, entails incorporation of isotopically enriched amino acids into the protein of interest. The NMR spectrum is simplified either by removal of NMR-active nuclei, as in the case of ^{1}H observation of a deuterated protein, or by increasing the concentration of NMR-active nuclei, as in the case of ^{13}C observation of ^{13}C-enriched protein. Assignments also can be made through correlation of NMR data (NOEs, pK determinations, paramagnetic shifts, ring-current shifts) with X-ray derived structures. A large number of assignments for OMTKY3 have been made in this way (Ortiz-Polo, 1985). Chemical modifications of amino acid side chains and their concomitant perturbation of the NMR spectrum have been used to identify resonances (Brown *et al.*, 1976). For small proteins, however, the current assignment method of choice is based on NMR data alone in conjunction with the amino acid sequence. Sequential assignment strategies are discussed in more detail in Section 4.

2. PROTEIN REQUIREMENTS

The amount of information that can be obtained from the NMR spectrum of a protein is dictated both by the properties of the protein of interest and by the capabilities of the NMR spectrometer used for data acquisition. The availability, solubility, size, and three-dimensional structure of the protein affect spectral sensitivity and resolution. In general, at least a micromole of pure protein, dissolved in 0.2 to 2.0 ml of solvent, is necessary for most 2D ^{1}H NMR experiments. Figure 4 outlines the limits imposed by protein size on the type of information that can be extracted from NMR experiments. Resonance linewidths generally depend on the correlation times for the random tumbling of the protein in solution (τ_c); larger proteins have longer correlation times and broader lines. This, coupled with the larger number of resonances, leads to lower resolution. To data, the largest monomeric protein subjected to extensive ^{1}H NMR sequential assignment is lysozyme (M_r = 14,000) (Redfield and Dobson, 1988). The protein also must have structure in solution, as resolution of resonances belonging to amino acids of a given type depends on the chemical shift dispersion resulting from the folded protein structure. An additional requirement for application of the sequential assignment method is knowledge of the protein's amino acid sequence; although errors have been detected in amino acid sequences by 2D NMR analysis (Štrop *et al.*, 1983; Wagner *et al.*, 1986), no extensive protein sequencing by NMR has been achieved.

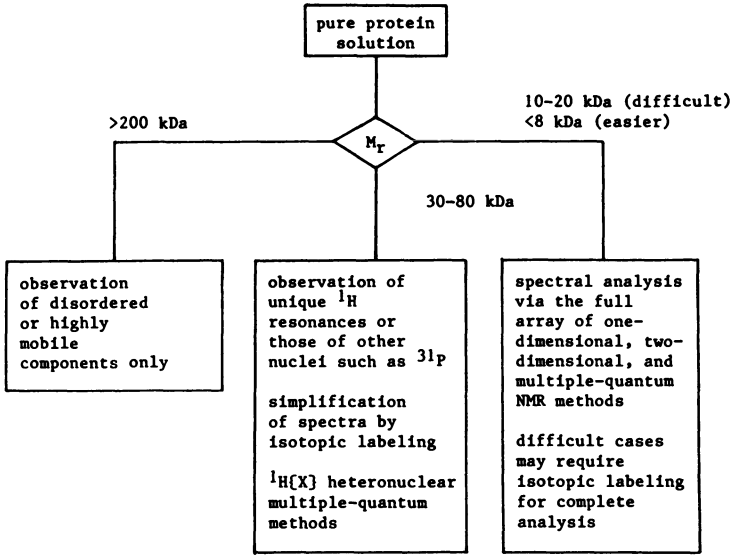


Figure 4. Flow chart describing the limitations imposed upon the NMR spectroscopist by protein size.

The sample conditions used for acquisition of all 2D NMR data should be reproduced as closely as possible; chemical shifts generally are sensitive to pH, temperature, ionic strength, and protein concentration. It is best, when possible, to use the same sample for all experiments and to conduct the 2D NMR experiments one after another without removing the sample from the magnet.

3. INSTRUMENTAL REQUIREMENTS

Data for protein sequential assignments usually have been collected on spectrometers operating at or above a proton frequency of 400 MHz. Assignments for peptides less than 30 residues in size have been achieved with data collected at 360 MHz (Wider *et al.*, 1982; Wemmer and Kallenbach, 1983). Most high-field NMR spectrometers delivered in the last 5 years are equipped for 2D experiments. The computer must have pulse programming and automated data acquisition capabilities and space for storage of the large 2D data sets (at least 256 × 1024 data points, but more typically 512 × 2048 data points). Fourier transformation of such data sets can take hours on a computer not equipped with an array processor. Offline data processing may increase data throughput by releasing the spectrometer for data acquisition, but this is not an absolute requirement;

on many modern spectrometers, data processing and data acquisition may be performed simultaneously. A critical requirement is the spectrometer time needed for data acquisition, which can run from hours to days. For TOCSY and HOHAHA experiments and for proton detection of ^{1}H–^{13}C correlation experiments, some adaptations of the spectrometer may be necessary.

4. SEQUENTIAL ASSIGNMENTS

There are three fundamentally different approaches to ^{1}H NMR sequential assignments. The first relies on NOE data to detect short ^{1}H–^{1}H distances between neighboring amino acids (Wüthrich *et al.*, 1982). The second depends on the observation of one-bond scalar coupling between the peptide carbonyl carbon of residue i and the amide nitrogen of residue $i + 1$; it requires dual ^{13}C and ^{15}N labeling (Kainosho and Tsuji, 1982). The third method detects multiple-bond heteronuclear coupling between the carbonyl of residue i and a proton of residue $i + 1$; it requires ^{13}C labeling (Westler *et al.*, 1988). Additional 2D NMR experiments are then used to attach the carbon and nitrogen to protons in their respective amino acid residues. All these methods rely on knowledge of the protein's primary structure; sequentially related di- and tripeptides are constructed with the NMR data and aligned with the protein sequence. The isotopic enrichment techniques should be highly reliable because they are based on unambiguous through-bond interactions. These methods are limited by the high cost of isotopes and their incorporation into protein. In spite of possible problems in distinguishing NOEs which are between sequentially related residues from those which are between residues more distant in the sequence, the NOE-based method has been applied successfully to the assignment of protons in over 50 large peptides and small proteins (for a compilation see Markley *et al.*, 1988).

The first stage of the NOE-based sequential assignment method is the identification of the scalar coupling networks, or spin systems, corresponding to different amino acid types (see Figure 5) (Wagner and Wüthrich, 1982). Particular emphasis is placed on identifying the backbone N$\underline{H}$ and $C_{\alpha}\underline{H}$ resonances, as these are used for sequence specific assignment in the second stage of assignment. None of the experiments based on proton–proton scalar coupling permits connection of sequentially neighboring residues, since each residue acts as an isolated spin system, i.e., the scalar coupling across the peptide bond is, in most cases, too small to be observed. These connections are made by combining NOESY data, first stage assignments, and protein sequence data to deduce sequence-specific backbone N$\underline{H}$ and $C_{\alpha}\underline{H}$ assignments. The NOESY experiment relies on dipolar

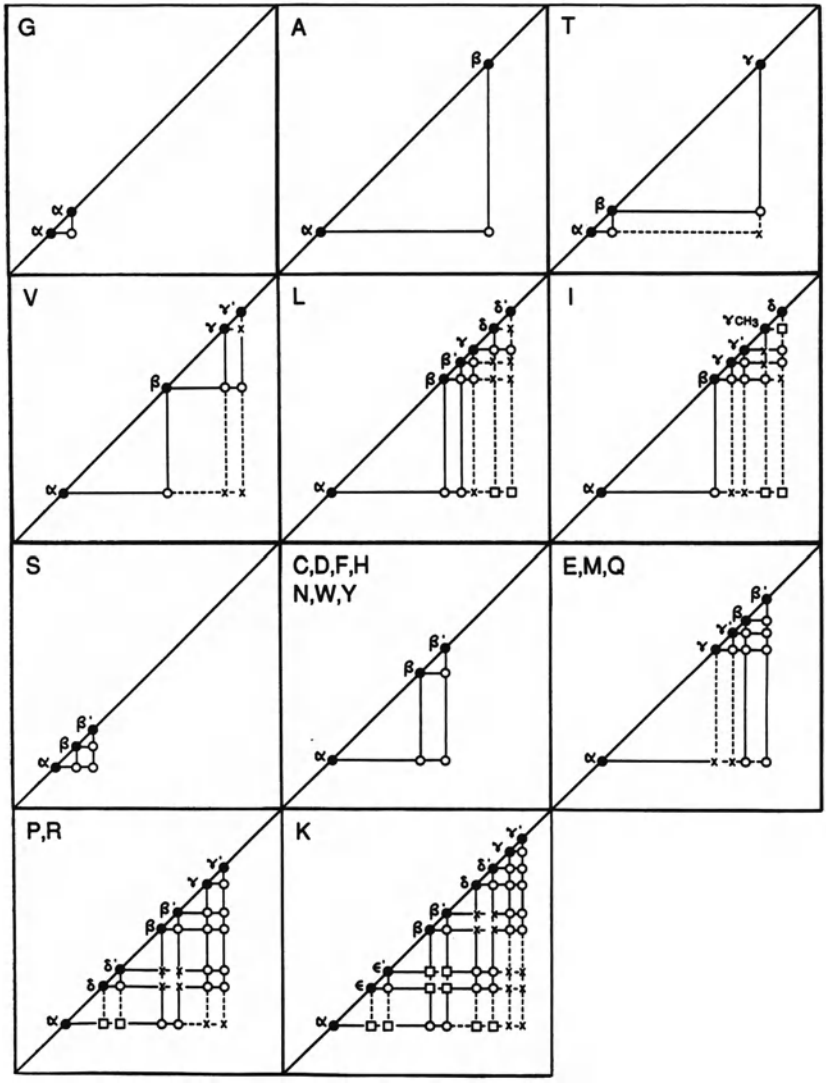


Figure 5. Proton NMR spin systems for the amino acid side chains. The axes are relative chemical shifts covering the range from approximately 1 to 5 ppm left to right and top to bottom. Solid circles (●) are diagonal peaks, open circles (○) are cross-peaks seen in the COSY spectrum, crosses (×) represent RELAY peaks, and open squares (□) are potential HOHAHA cross-peaks (after a figure in Wüthrich, 1986).

coupling connectivities, which are a function of through-space interproton distance. Proteins can exhibit a number of interresidue NOESY connectivities, some of which are diagrammed in Figure 6, linking elements of the spin systems of sequentially adjacent residues. Valid links must connect spin systems consistent with the primary amino acid sequence. In addition, these resonances and the NOESY data can be used to determine protein secondary and tertiary structure (Wüthrich, 1986).

4.1. Amino Acid Spin System Identification

Many amino acids have characteristic spin systems owing to both the chemical structure of the side chain (bonded arrangement of protons) and the chemical shifts of these protons (Figure 5). Most of the 2D NMR experiments used for the identification of amino acid spin systems are based on scalar (through-bond) coupling. The prototypical experiment of this type is the COSY experiment (Figure 2), where the off-diagonal cross-peaks represent coupling between protons separated by two and three bonds (Aue *et al.*, 1976). The results of the original experiment were presented as magnitude-calculated spectra; recent modifications of the original pulse scheme make it possible to preserve cross-peak fine structure that results from scalar coupling (States *et al.*, 1982; Marion and Wüthrich, 1983). This in turn results in improved spectral resolution and can provide information about the magnitude of scalar coupling constants (Hyberts *et al.*, 1987).

In theory, the results of the COSY experiment contain all of the information necessary for spin system identification. Partial identification of spin systems for methyl-containing residues is illustrated in an expansion of the COSY spectrum of OMTKY3 in Figure 7 (leftmost panel). Pairs of cross-peaks in the upper half of the figure represent the three-bond coupling between the two methyl groups of valine and leucine residues and the single methine proton attached to C_β (in valine) or C_γ (in leucine). The intense

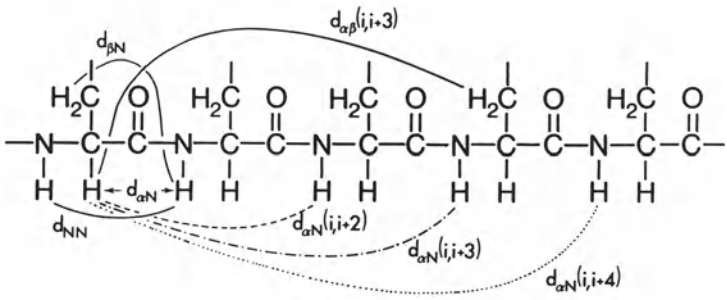


Figure 6. Short ^1H–^1H distances used for sequential assignment and the determination of protein secondary structure with ^1H NMR spectroscopy.

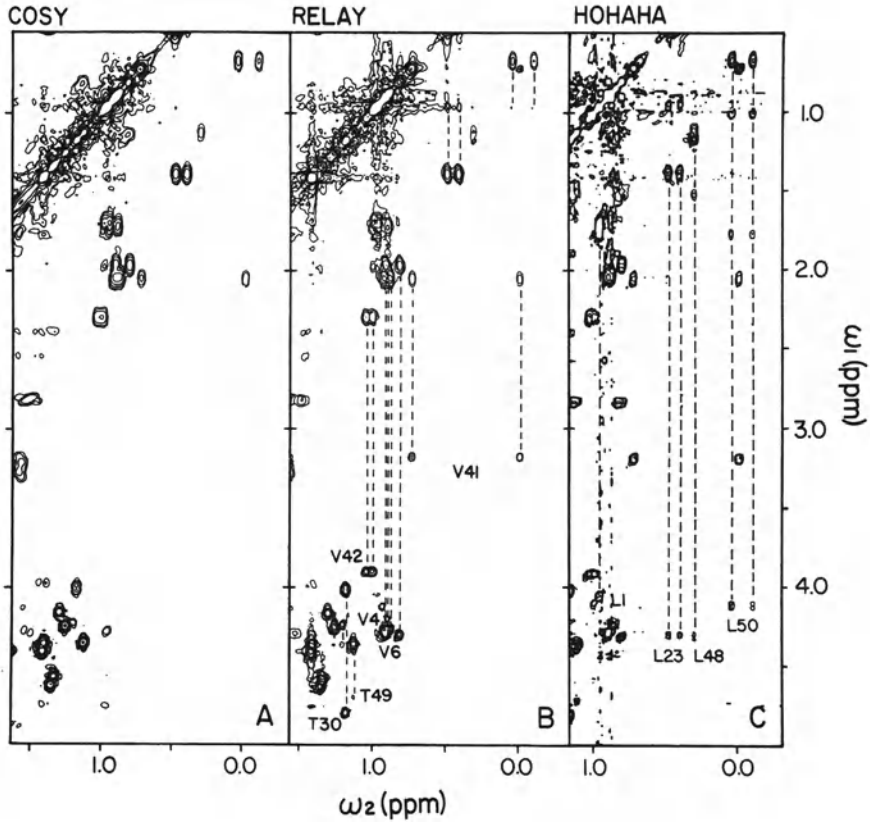


Figure 7. Expansions of the high-field portions of the COSY (A), RELAY (B), and HOHAHA (C) maps for OMTKY3 (Robertson *et al.*, 1988).

cross-peaks in the lower half of the spectrum represent the three-bond coupling between the single methyl groups of alanine and threonine side chains and the $C_\alpha\underline{H}$ in alanine or $C_\beta\underline{H}$ in threonine (there are no isoleucines in OMTKY3). Distinguishing valine from leucine or alanine from threonine, however, can be complicated by overlap of cross-peaks, particularly those between protons on long amino acid side chains.

The relayed correlation spectroscopy (RELAY) and double-RELAY experiments can go a long way toward resolving ambiguities in spin system identification (Eich *et al.*, 1982; Wagner, 1983; Bax and Dronby, 1985). Additional peaks in the RELAY map (compared to COSY) arise from pairs of protons that share scalar coupled protons but are not themselves coupled (Figure 7, center panel). For example, ($C_\alpha\underline{H}$, $C_\gamma\underline{H}$) and ($N\underline{H}$, $C_\beta\underline{H}$) cross-peaks are expected in the RELAY map, the connectivities being mediated by the $C_\beta\underline{H}$ and $C_\alpha\underline{H}$, respectively. For OMTKY3, the RELAY data in

Figure 7 suffice to identify all the valine $C_\alpha\underline{H}$, $C_\beta\underline{H}$, and $C_\gamma\underline{H}$ resonances along with ($C_\alpha\underline{H}$, $C_\gamma\underline{H}$) connectivities in the region around (1.00, 4.50) ppm from three threonine residues. Because leucine methyl resonances have RELAY connectivities only to $C_\beta\underline{H}$ resonances, leucine $C_\alpha\underline{H}$ resonances cannot be identified directly with this experiment. This information can be provided by the double-RELAY experiment, an extension of the RELAY experiment, or by the *to*tal *c*orrelation *s*pectroscop*y* (TOCSY) (Braunschweiler and Ernst, 1983) or 2D *ho*monuclear *Ha*rtmann–*Ha*hn (HOHAHA) experiments (Bax and Davis, 1985). In the TOCSY and HOHAHA experiments, magnetization is transferred from one spin to all other spins within a scalar coupling network; this often results in a two-dimensional map with strings of cross-peaks representing the entire spin system. In the HOHAHA map of OMTKY3, the $C_\delta\underline{H}$ for four of the five leucines show connectivities to their $C_\alpha\underline{H}$ resonances and to some $C_\beta\underline{H}$ resonances (Figure 7, rightmost panel). In favorable cases, the HOHAHA experiment can be used to identify spin systems for most of the long-side-chain amino acids (Figure 8).

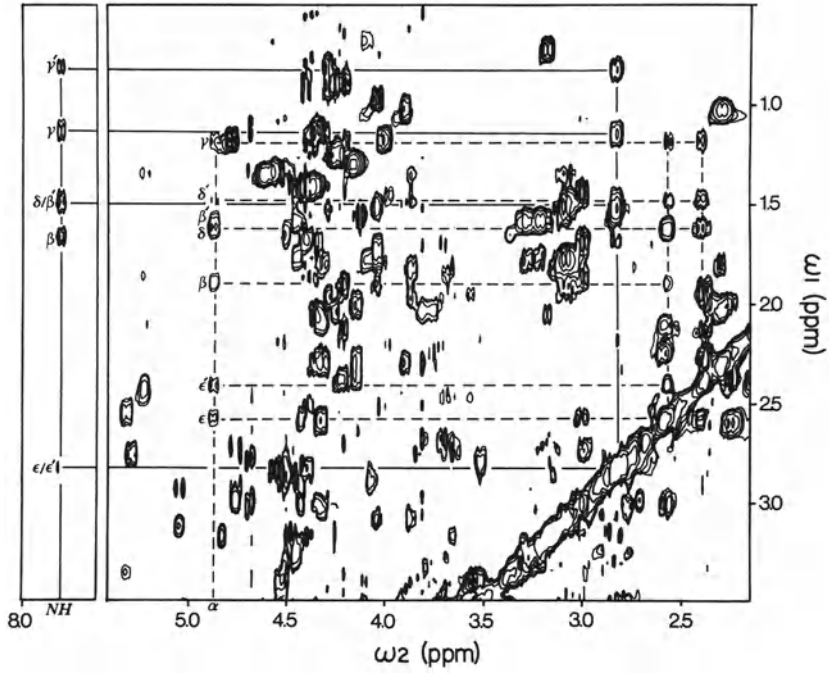


Figure 8. Lysine spin systems in the HOHAHA map of OMTKY3. The elements of the spin systems of lysines 13 and 29 are connected by dashed and solid lines, respectively.

The eight amino acids with the "simpler" AMX spin systems are probably the most difficult to distinguish from one another (Figure 5). Chemical shifts for the β protons of Ser residues typically fall near 4 ppm, which is 0.5 to 1 ppm further downfield (higher ppm) than the chemical shifts of β protons in other AMX spin systems. Additional partial resolution of AMX-type spin systems from one another may be achieved with the NOESY experiment. The NOE between the $C_\beta\underline{H}$ and $C_\delta\underline{H}$ of aromatic residues often can be observed in the NOESY spectrum (Figure 9). Because of the possible overlap of aromatic and amide proton resonances (Figure 1), it is essential that the data be collected on a protein sample in which

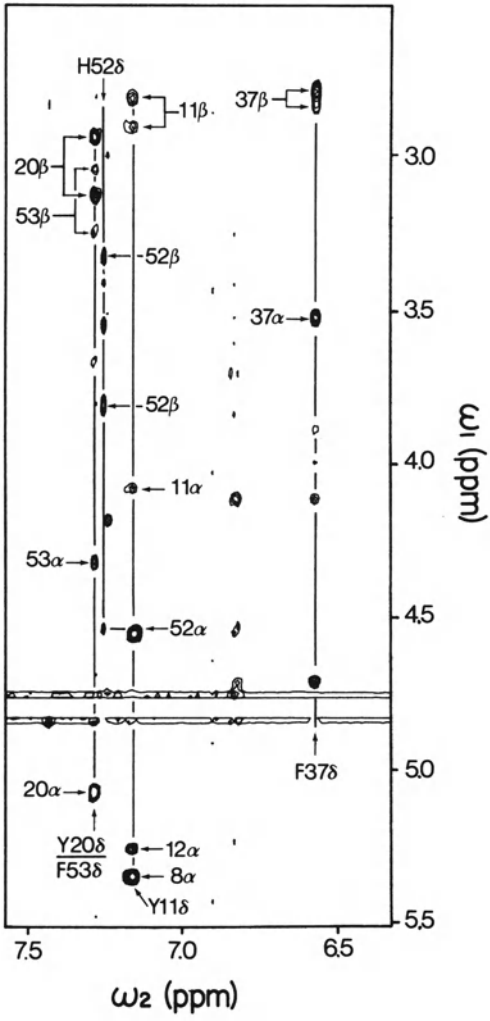


Figure 9. Expansion of the 500 MHz pure-phase NOESY map of OMTKY3 in H_2O. The connectivities between $C_\beta\underline{H}$ and $C_\delta\underline{H}$ resonances of aromatic residues are indicated.

all amide protons have been exchanged completely with solvent deuterons. NOESY data from protein dissolved in H_2O may be used to identify Asn and Gln side-chain resonances through observation of cross-peaks between $C_\beta\underline{H}$ or $C_\gamma\underline{H}$ resonances and the resonances of the primary amide protons.

The interpretation of proton NMR data is complicated by degeneracy of many of the chemical shifts. Proton resonances from aliphatic side chains are not always diagnostic of the type of amino acid to which they belong. ^{13}C chemical shifts, on the other hand, are dispersed over a broader chemical shift range (approximately 15 to 70 ppm for aliphatic carbon resonances versus 1 to 5 ppm for aliphatic proton resonances) and, in some cases, are nearly unique to a residue or class of residues. Access to the ^{13}C resonance information has been limited by the low natural abundance (~1%) and magnetogyric ratio (25% that of the proton) of this nucleus. Because relative sensitivity varies as the cube of the magnetogyric ratio, the two factors together lead to a roughly 10,000-fold decrease in sensitivity relative to protons. Recent developments in 2D NMR methodology take advantage of the superior sensitivity of the proton and detect the ^{13}C chemical shift information indirectly via attached protons (Sklenár and Bax, 1987). Although the sensitivity in such experiments is lower than that in the ^1H-^1H experiments, data for interpretable spectra can be acquired in about the time typically used for acquisition of the ^1H-^1H data. Figure 10 provides an example of the high information content of such experiments. Because of their distinctive chemical shifts, the δ carbons of proline and arginine, the ε carbons of lysine, and the γ carbons of glutamate can be readily resolved in the $^1H\{^{13}C\}$ *het*eronuclear *cor*relation (HETCOR) experiment. From the point of view of proton resonance assignments, these data permit unambiguous identification of resonances which are otherwise very difficult to resolve with the proton NMR data alone. Similar results are observed in the aromatic, α-carbon, and methyl regions of such plots. The $^1H\{^{13}C\}$ HETCOR experiment can be used to determine if resonances for gem methylene protons are degenerate, as is the case for the $C_\gamma\underline{H}$ of Glu-43 and the $C_\delta\underline{H}$ of Pro-22 in OMTKY3 (Figure 10).

4.2. Sequence-Specific Resonance Assignments: Connecting Spin Systems

The protein sequence and extent of spin system identification determine when NOESY data can be used to connect neighboring residues. Starting points for sequence-specific assignments are located if spin systems for unique residues in the protein are identified or when sequential NOESY connectivities (Figure 6) can be made between two or three identifiable spin systems, and the corresponding di- or tripeptide sequence occurs only once in the protein sequence. An example of a unique dipeptide sequence in OMTKY3 is Gly-Lys at positions 54 and 55. In the first stage of

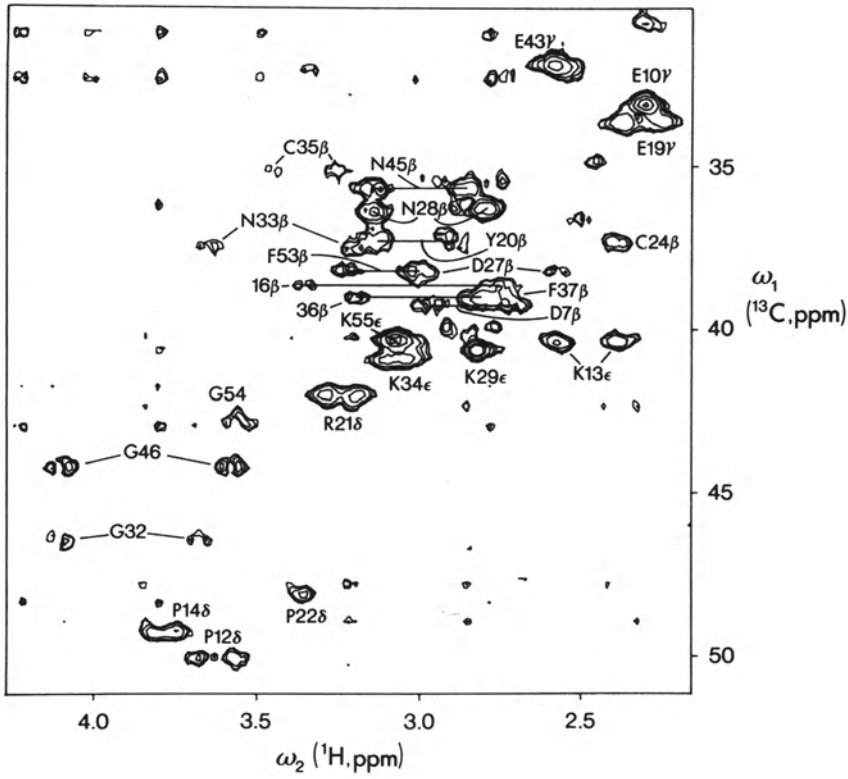


Figure 10. Expansion of the $^1H\{^{13}C\}$ heteronuclear correlated spectrum of OMTKY3. Resonance assignments are indicated.

assignment, a number of Gly and Lys spin systems were identified; apparent sequential $d_{\alpha N}$ connectivities were identified between a Gly and a Lys spin system (Figure 11). It follows then that these are the spin systems of Gly 54 and Lys 55. Assignments can be continued from such starting points in both the C- and N-terminal directions as long as unambiguous sequential connectivities are observed across each new peptide bond. As new backbone resonances are added sequentially, their spin systems, identified in the first stage of assignment or as a result of the sequential assignment, should be checked against the protein sequence for consistency.

The $d_{\alpha N}$ connectivity is probably the most informative and reliable of the three sequential NOESY connectivities diagrammed in Figure 6. A $d_{\alpha N}$ crosspeak, unlike a d_{NN} connectivity, establishes the direction (N- or C-terminal) in which one is moving and is usually located in a region of the spectrum which is less crowded than that containing $d_{\beta N}$ cross-peaks. In cases where $d_{\alpha N}$ connectivities are not observed or where there are

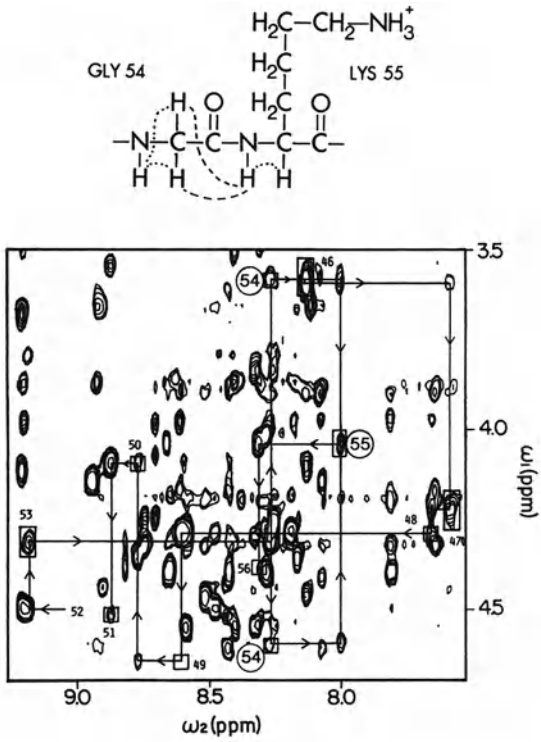


Figure 11. Expansion of the NOESY spectrum of OMTKY3 in H_2O, illustrating sequential $d_{\alpha N}$ connectivities. The intraresidue NOEs between the $C_\alpha\underline{H}$ and $N\underline{H}$ are boxed; these correspond to COSY cross-peaks found in the fingerprint region. The horizontal lines indicate NOEs used for sequential assignment. The method is illustrated (circled numbers) by the NOEs observed between glycine and lysine residues which were identified during the first stage of assignment. A single dipeptide sequence in OMTKY3 is consistent with this observation, Gly^{54}-Lys^{55}. The observed NOEs are indicated on the dipeptide above the NOESY map. Intraresidue contacts are dotted lines and interresidue contacts are dashed lines.

ambiguities due to resonance overlap, the d_{NN} and $d_{\beta N}$ connectivities can be used to make or confirm sequential assignments. NOESY connectivities between aliphatic protons also can be used for confirmation of sequential assignments but, because of crowding in the aliphatic region of the NOESY spectrum for most proteins, initial assignments with these data usually are only tentative and require further testing. In any case, efforts should be made to identify as many sequential NOEs as possible, as this increases the reliability of the assignment.

4.3. Resolving Resonance Overlap

Overlapping NMR signals may be distinguished through the use of more sophisticated NMR experiments or by alteration of the sample condi-

tions. RELAY and HOHAHA experiments permit resolution of signals by increasing the information content of the resulting spectra. Other experiments, such as the higher order multiple-quantum-filtered COSY experiments, provide simplified 2D spectra by removing resonances from particular amino acid types (Müller *et al.*, 1986).

One of the simplest and most popular spectral simplification techniques is to dissolve the protein in deuterated aqueous solvent. This step removes the large NMR signal from water; in addition, resonances from exchange-labile amide protons in the protein are no longer observed. With many

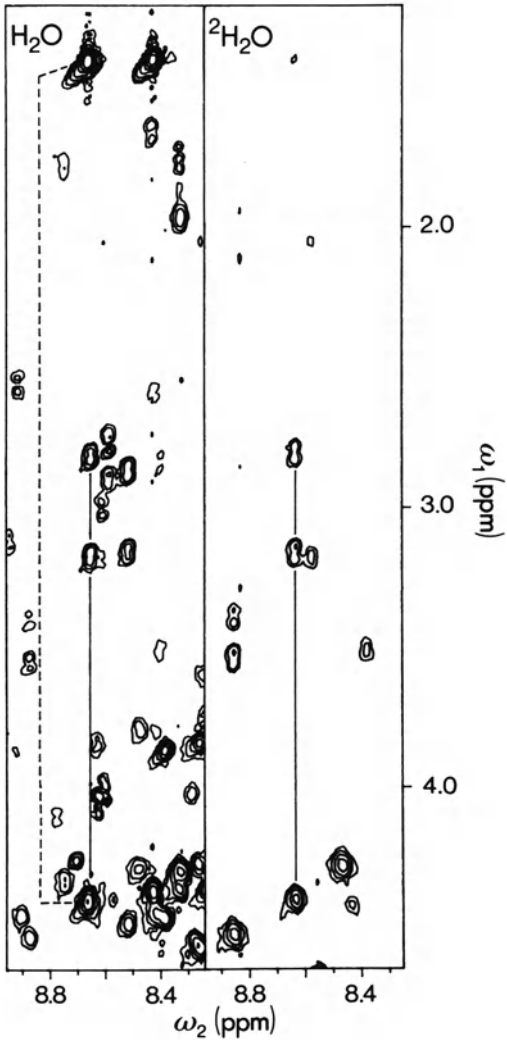


Figure 12. Expansion of the RELAY map of OMTKY3 in H_2O and ^2H_2O. The solid line connects cross-peaks belonging to the spin system of an asparagine residue. The backbone amide proton of this residue is in slow exchange with solvent and hence can be observed in protein dissolved in ^2H_2O. The dashed line indicates crosspeaks arising from an alanine spin system whose amide proton exchanges more quickly than that of the asparagine residue.

proteins, however, some of the amide protons are protected from exchange with solvent by features of the protein's structure, particularly by the intramolecular hydrogen bonding responsible for the stability of secondary structure (Englander and Kallenbach, 1984). The spectrum of freshly dissolved protein often contains resonances from such amide protons (Figure 12). Complete exchange usually is accomplished by subjecting the protein to conditions that favor exchange, such as high temperature and pH. A degree of control over the extent of exchange is afforded by differences in the exchange behavior (rate, pH, and temperature dependences) of amide protons within a protein. Appropriate choice of sample conditions leads to select populations of observed amide proton resonances.

Alterations in sample pH and temperature also can lead to changes in chemical shifts. Such changes can be used to resolve overlap, provided the overlapping resonances do not show identical pH or temperature dependent chemical shift perturbations.

5. PROTEIN STRUCTURE ANALYSIS WITH SEQUENTIAL ASSIGNMENT DATA

Once extensive ^{1}H NMR assignments have been obtained for a protein, it is possible to make use of NMR data (short distances derived from NOE data and dihedral angles derived from coupling constants) to model the secondary and tertiary structure (Wüthrich, 1986). The strength of the NOE has an r^{-6} dependence on the interproton separation r and an additional dependence on the product of the magnetic field strength and the correlation time (τ_c) of the vector connecting the dipole-coupled protons (Noggle and Schirmer, 1971). For most protons in a protein the most important motion-contributing to τ_c is the random tumbling of the protein in solution.

Normally, the upper limit on the NOE-derived interproton distance in a protein the size of OMTKY3 is 5 Å (Braun *et al.*, 1981). Despite this limitation, a number of interesting short (<5 Å) ^{1}H–^{1}H distances in proteins can be detected by the NOESY experiment. The NOE between protons on adjacent residues in the primary sequence is the basis for sequential resonance assignments. Wüthrich and co-workers conducted molecular modeling studies and surveyed X-ray derived protein structures in a effort to identify additional short ^{1}H–^{1}H distances that would permit determination of secondary structure with NOESY data (Wüthrich *et al.*, 1984). They concluded that features such as parallel and anitparallel β-sheet, reverse turns, and helices could be identified reliably with NOESY data using contacts such as those diagrammed in Figure 6 and shown in Figure 13. NOESY data collected for sequential assignments usually is sufficient also for determination of secondary structure.

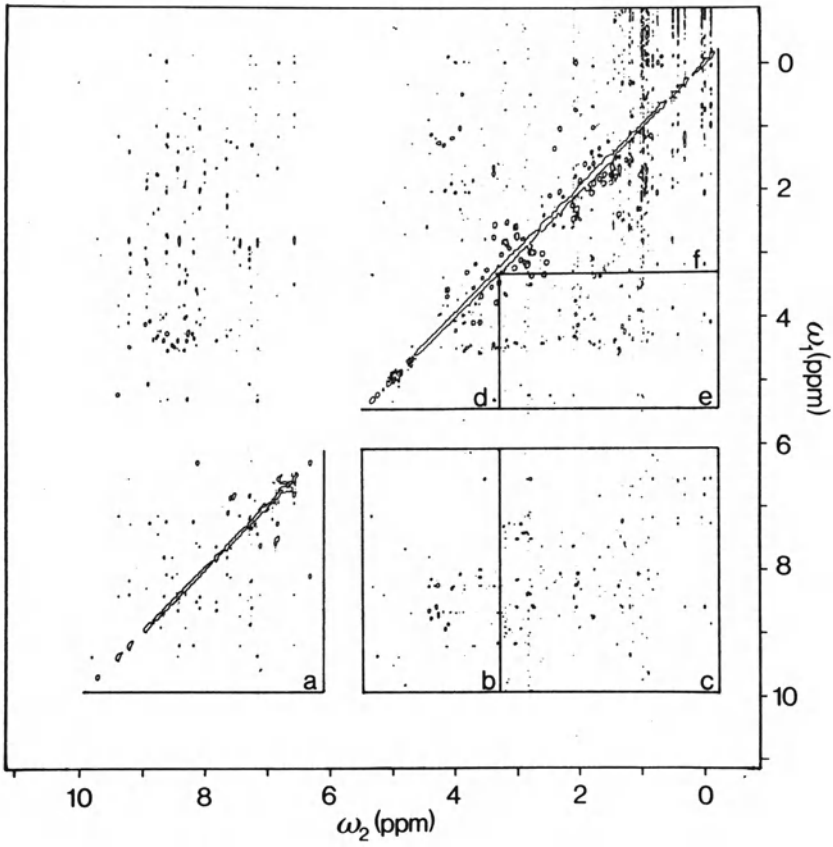


Figure 13. 500 MHz NOESY spectrum of OMTKY3 in H_2O. Regions of the spectrum containing contacts between particular types of protons are indicated: aromatics and amides (a); aromatics and amides to $C_\alpha\underline{H}$, $C_\beta\underline{H}$ of Ser and Thr and $C_\delta\underline{H}$ of Pro (b); aromatics and amides to side-chain aliphatic protons (c); $C_\alpha\underline{H}$, $C_\delta\underline{H}$ of Pro and $C_\beta\underline{H}$ of Ser and Thr (d); $C_\alpha\underline{H}$, $C_\delta\underline{H}$ of Pro, $C_\beta\underline{H}$ of Ser and Thr, aliphatic side-chains and aliphatic side chains (e).

Helices are perhaps the easiest structural features to identify. A string of sequential d_{NN} connectivities combined with intermediate range $d_{\alpha N}$ $(i, i+2)$, $(i, i+3)$, and $(i, i+4)$ connectivities is a reliable indication of helical structure (Figure 14). By contrast, β-sheet structure is located through the observation of a stretch of strong $d_{\alpha N}$ connectivities, indicative of extended chain conformations (Wüthrich *et al.*, 1984), and a series of interstrand NOEs, such as those indicated in Figure 15. Retarded exchange of amide protons with deuterated solvent is evidence for hydrogen bonding and possible involvement in secondary structure (Figure 14).

NMR data are being used to determine the tertiary structures of small proteins in solution (Williamson *et al.*, 1985). The two basic approaches to

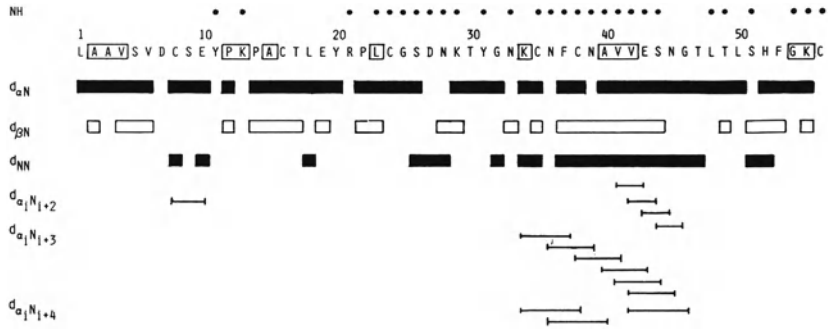


Figure 14. Sequence of OMTKY3, with sequential assignment starting sites (boxed). NOESY connectivities are indicated below the sequence. The dots above the sequence denote slowly exchanging NHs as defined by observation of a cross-peak in the fingerprint region of the COSY map for OMTKY3 freshly dissolved in H_2O at pH 3.84, 25 °C (Robertson *et al.*, 1988).

structure determination, distance geometry (Havel *et al.*, 1979) and molecular dynamics with energy minimization (Kaptein *et al.*, 1985), use short distances, estimated with NOE data, and dihedral angles, determined from three-bond scalar coupling constants. Distance geometry algorithms rely purely on geometrical constraints, while the energy minimization approach includes theoretical potentials for intramolecular interactions, particularly electrostatic interactions.

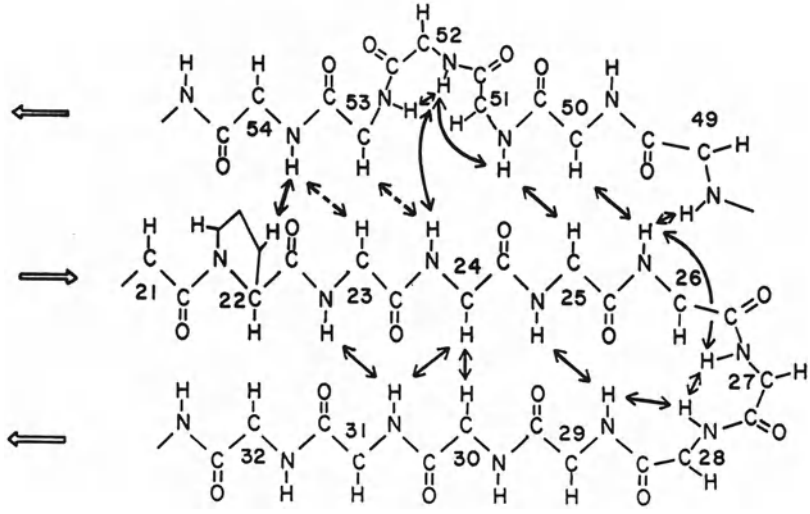


Figure 15. NOESY connectivities used to establish the relative orientation of the strands of β-sheet. The dashed lines indicate possible NOE connectivities that are observed by overlapping cross-peaks (Robertson *et al.*, 1988).

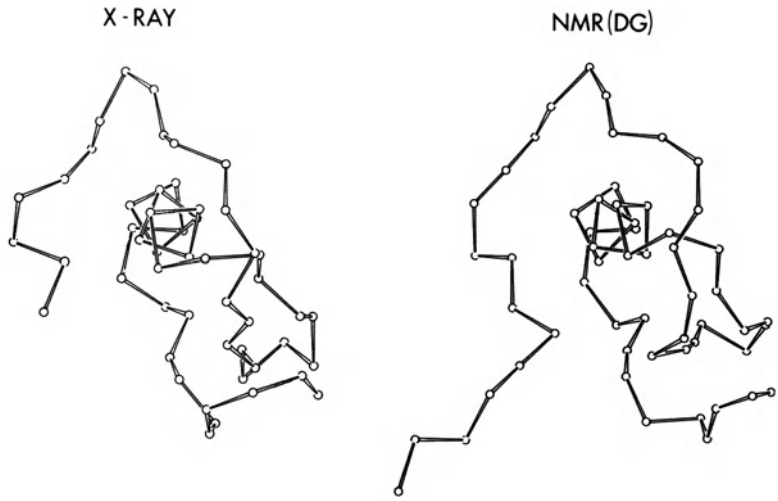


Figure 16. Models for the three-dimensional structure of OMTKY3 determined by X-ray crystallography (left) and ^{1}H NMR spectroscopy (right). The α-carbon backbone is shown in each case. The X-ray structure was determined by Read *et al.* (1983). For the NMR structure (Darba *et al.*, 1988), ^{1}H-^{1}H contacts observed in the NOESY spectrum of OMTKY3 were converted to distance constraints, which then were used as input for the DG900 program (Thomasson *et al.*, 1988) installed on the Cray XMP computer located at the San Diego Supercomputer Center.

The solution structure in Figure 16 was obtained after one run of a distance geometry program. The approximately 200 short ^{1}H-^{1}H distances used for the structure determination were extracted from the same NOESY data used for sequential assignment. The N-terminal seven residues of OMTKY3 were not observed in the X-ray studies, but the NMR data suggest an extended chain conformation for these residues.

REFERENCES

Alexandrescu, A. T., Mills, D. A., Ulrich, E. L., Chinami, M., and Markley, J. L., 1988, *Biochemistry*, **27**:2158.

Aue, W. P., Bartholdi, E., and Ernst, R. R., 1976, *J. Chem. Phys.* **64**:2229.

Bax, A., and Davis, D. G., 1985, *J. Magn. Reson.* **65**:355.

Bax, A., and Drobny, G. P., 1985, *J. Magn. Reson.* **61**:306.

Bax, A., and Lerner, L., 1986, *Science* **232**:960.

Braun, W., Bosch, C., Brown, L. R., Gō, N., and Wüthrich, K., 1981, *Biochim. Biophys. Acta* **667**:377.

Braunschweiler, L., and Ernst, R. R., 1983, *J. Magn. Reson.* **53**:521.

Brown, L. R., De Marco, A., Wagner, G., and Wüthrich, K., 1976, *Eur. J. Biochem.* **62**:103.

Darba, P., Krezel, A., Fejzo, J., Macura, S., Robertson, A. D., and Markley, J. L., 1988, 29th Experimental NMR Conference, Rochester, NY.

Eich, G., Bodenhausen, G., and Ernst, R. R., 1982, *J. Am. Chem. Soc.* **104**:3731.
Englander, S. W., and Kallenbach, N. R., 1984, *Q. Rev. Biophys.* **16**:521.
Ernst, R. R., Bodenhausen, G., and Wokaun, A., 1987, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.
Havel, T. R., Crippen, G. M., and Kuntz, 1979, *Biopolymers* **18**:73.
Hyberts, S. G., Märki, W., and Wagner, G., 1987, *Eur. J. Biochem.* **164**:625.
Kainosho, M., Tsuji, T., 1982, *Biochemistry* **21**:6273.
Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., and van Gunsteren, R., 1985, *J. Mol. Biol.* **182**:179.
Karplus, M., 1959, *J. Chem. Phys.* **30**:11.
LeMaster, D., and Richards, F., 1988, *Biochemistry* **27**:142.
Marion, D., and Wüthrich, K., 1983, *Biochem. Biophys. Res. Commun.* **113**:967.
Markley, J. L., 1972, *Meth. Enzymol.* **26**:605.
Markley, J. L., 1975, *Accts. Chem. Res.* **8**:70.
Markley, J. L., Ulrich, E. L., and Kyogoku, Y., 1989, *Protein Seq. Data Anal.* **2**:23.
Müller, N., Ernst, R. R., and Wüthrich, K., 1986, *J. Amer. Chem. Soc.* **108**:6482.
Noggle, J. H., and Schirmer, R. E., 1971, *The Nuclear Overhauser Effect*, Academic Press, New York.
Ortiz-Polo, G., 1985, M. S., Thesis, Purdue University, West Lafayette, Ind.
Perkins, S. J., 1982 "Applications of Ring Current Calculations to the Proton NMR of Proteins and Transfer RNA" in *Biological Magnetic Resonance*, Vol. 4 (L. J. Berliner and J. Reuben, eds.), Plenum, New York, pp. 193–336.
Read, R. J., Fujinaga, M., Sielecki, A. R., and James, M. N. G., 1983, *Biochemistry* **22**:4420.
Redfield, C., and Dobson, C. M., 1988, *Biochemistry* **27**:122.
Robertson, A. D., Westler, W. M., and Markley, J. L., 1988, *Biochemistry*, **27**:2519.
Sklenár, V., and Bax, A., 1987, *J. Magn. Reson.* **71**:379.
States, D. J., Haberkorn, R. A., and Ruben, D. J., 1982, *J. Magn. Reson.* **48**:286.
Štrop, P., Wider, G., and Wüthrich, K., 1983, *J. Mol. Biol.* **166**:641.
Thomasson, J., Day, M., and Kuntz, I. D., 1988, "A Vectorized Distance Geometry Program," to be published.
Wagner, G., 1983, *J. Magn. Reson.* **55**:151.
Wagner, G., and Wüthrich, K., 1982, *J. Mol. Biol.* **155**:347.
Wagner, G., and Wüthrich, K., *Meth. Enzymol.* **131**:307.
Wagner, G., Neuhaus, D., Wörgötter, E., Vasák, M., Kägi, J. H. R., and Wüthrich, K., 1986, *Eur. J. Biochem.* **157**:275.
Wemmer, D. E., and Kallenbach, N. R., 1983, *Biochemistry* **22**:1901.
Westler, W. M., Kainosho, M., Nagao, H., Tomonaga, N., and Markley, J. L., 1988, *J. Am. Chem. Soc.* **110**:4093.
Wider, G., Lee, K. H., and Wüthrich, K., 1982, *J. Mol. Biol.* **155**:367.
Williamson, M. P., Havel, T. F., and Wüthrich, K., 1985, *J. Mol. Biol.* **182**:295.
Wüthrich, K., 1986, *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
Wüthrich, K., Wider, F., Wagner, G., and Braun, W., 1982, *J. Mol. Biol.* **155**:311.
Wüthrich, K., Billeter, M., and Braun, W., 1984, *J. Mol. Biol.* **180**:715.
Zolnai, Z., Macura, S., and Markley, J. L., 1986, *Computer Enhanced Spectroscopy* **3**:141.

5

Solid-State NMR Spectroscopy of Proteins

Stanley J. Opella

1. INTRODUCTION

The methods and theories of solid-state NMR spectroscopy make it possible to study larger and more complex biological systems than with conventional NMR methods. They are also capable of providing unique information about the structure and dynamics of proteins. There are now a number of books and reviews that describe in detail solid-state NMR spectroscopy (e.g., Haeberlen, 1976; Mehring, 1983; Fyfe, 1983) and its applications to proteins (e.g., Oldfield *et al.*, 1982; Torchia, 1984; Opella 1986; Opella *et al.*, 1987).

An important definition concerns the difference between solid-state NMR spectroscopy and solution NMR spectroscopy. In spite of their names, these two approaches can not really be differentiated on the basis of the physical state of the samples which are investigated; although only solid-state NMR spectroscopy can be used to study crystalline proteins, both forms of spectroscopy are used to study nonviscous solutions of proteins. Rather, it is the dynamics of the molecules that dictate which approach is appropriate for a particular protein or complex, with solid-state NMR methods useful for samples where the proteins undergo limited or no motional averaging and solution NMR methods appropriate for samples where the proteins undergo rapid overall reorientation. Solid-state NMR spectroscopy and solution NMR spectroscopy differ in the way in which

Stanley J. Opella • Department of Chemistry, University of Pennsylvania, Phildelphia, Pennsylvania 19104.

the spin interactions are manifested in the spectra as well as in some aspects of the instrumentation and techniques used to obtain the spectra and perform experiments. Both are high resolution methods, in that resonances from individual sites can be identified and characterized, and they both have high sensitivity, in that resonances from individual sites can be observed with good signal-to-noise ratios in a matter of seconds or minutes in favorable cases.

Solid-state NMR spectroscopy has a long and illustrious history of studying biological or quasibiological materials, going back to the initial observations of resonances in paraffin (Purcell *et al.*, 1946) and early experiments with gypsum (Pake, 1948). However, it was the development and implementation of high resolution methods by Waugh and co-workers at M.I.T, first with proton multiple pulse experiments (Waugh *et al.*, 1968) and then with dilute spin double resonance experiments (Pines *et al.*, 1973), that made applications of solid-state NMR spectroscopy to substantial chemical and biological problems feasible. The ability to resolve resonances in immobile samples meant that, for the first time, NMR could be used for spectroscopic studies of proteins that were in the crystalline solid state or part of supramolecular structures in solution. The same spectroscopic developments that made it possible to narrow the resonances sufficiently to obtain high resolution spectra also made it possible to manipulate the system so that spectral parameters from a variety of spin interactions could be used to characterize molecular structure and dynamics.

The concept of dilute spins is important for the design and implementation of most high resolution solid-state NMR studies. ^{2}H, ^{13}C, ^{14}N, and ^{15}N in natural abundance or in labeled sites in proteins are generally well separated from like spins, and since these nuclei have small gyromagnetic ratios, the homonuclear dipole–dipole couplings are small. This makes it relatively easy to obtain spectroscopic results from identifiable individual sites. In contrast, ^{1}H spins with their large gyromagnetic ratio, high natural abundance, and high concentration have strong homonuclear dipole–dipole couplings. This makes high resolution solid-state ^{1}H NMR spectroscopy considerably more difficult than for dilute spins. The ^{1}H spins have a strong influence on the spectroscopy of the dilute spins; the strongly coupled network of ^{1}H spins provides a source of magnetization that can be usefully transferred to the dilute spins for sensitivity enhancement, and they participate in well-defined heteronuclear dipole–dipole couplings. ^{2}H and ^{14}N are spin $S = 1$ nuclei; therefore their spectral properties are influenced by the nuclear quadrupole interaction as well. All of the nuclei and their interactions present both spectroscopic problems in terms of their detection and characterization and opportunities for obtaining valuable information about properties of molecules. The nuclear spin interactions are highly anisotropic and the angular dependencies of these spin interactions are directly

manifested in all spectra of stationary samples with immobile sites. In unoriented samples, the spectra consist of powder patterns because all orientations are present. These powder pattern lineshapes are strongly influenced by motional averaging (Spiess, 1982); therefore, they provide a mechanism for characterizing the motions of proteins. Magic angle sample spinning averages the powder patterns to single isotropic resonances (Andrew *et al.*, 1958; Schaefer and Stejskal, 1976), giving high resolution spectra at the expense of most of the anisotropic information. In single-crystal or uniaxially oriented samples, each site yields a single line or well-defined multiplet depending on the spin interactions with resonance frequencies or splitting that is a function of the orientation of the site with respect to the applied magnetic field.

Most protein structures have been determined by X-ray diffraction studies of single-crystal samples (Wyckoff *et al.*, 1985) and NMR spectroscopy is now capable of determining the structures of small globular proteins in solution (Wüthrich, 1986). However, there are several compelling reasons for developing solid-state NMR spectroscopy as a method for determining the structures of proteins. Solid-state NMR methods permit the study of large ordered systems such as nucleoprotein or membrane–protein complexes which are difficult or impossible to study at high resolution by either X-ray diffraction or solution NMR methods. In addition, since uniaxially oriented samples can be studied as easily as single crystals and some experiments can use unoriented samples, solid-state NMR spectroscopy extends the range of samples that can be studied.

Although highly refined views of the time-average locations of protein sites are available from the diffraction and solution NMR studies, both of these methods, for the most part, provide only fairly indirect indications of intramolecular motions. In contrast, solid-state NMR spectroscopy is highly effective at describing the intramolecular motions of proteins. Since it is possible to describe protein structure and dynamics with the same samples and with related experiments, solid-state NMR spectroscopy offers the promise of providing a fully integrated biophysical view of protein function.

2. PROTEIN DYNAMICS

The current high level of activity in experimental and theoretical studies of protein dynamics reflects the importance of the motions of the sidechain and backbone sites in the functions of proteins (Gurd and Rothgeb, 1979; Hirs and Timasheff, 1986). Solid-state NMR spectroscopy is useful for studying protein dynamics because the effects of motional averaging on

solid-state NMR spectral parameters are often dramatic and are readily interpreted in immobile molecules, because only the intramolecular motions of interest are influencing the spectral parameters (Opella, 1986). The averaging effects of motions on the spin interactions have two consequences—one is an averaging of the spectral frequencies that contribute to powder patterns and the other is the induction of nuclear spin relaxation. The motions are always defined with respect to a timescale determined by the dominant spin interaction and its spectral parameters. The timescales range from around 10^4 Hz for the motional averaging of chemical shift anisotropy powder patterns to around 10^9 Hz for relaxation induced by fluctuations. An immobile site does not have large-amplitude motions that occur frequently compared to the relevant timescale. It is convenient to categorize sites, especially in the backbone of a protein, as immobile or mobile. In addition to the frequencies of the motions, the lineshapes are affected by the amplitudes, directions, and types of the motions, i.e., jump vs. diffusional. In order to understand the influence of these factors, specific models must be considered. At many sites, there is evidence of small-amplitude librational motions. If these motions have amplitudes less than about 10°, then they have little discernable effect on the powder patterns.

Large amplitude backbone and sidechain motions can be detected and characterized by comparing experimentally observed powder pattern lineshapes to those calculated for specific models of motional averaging of rigid lattice powder patterns. For large-amplitude motions it is possible to use the lineshape of the motionally averaged powder pattern to find out the type of motion, angles, amplitudes, etc., and then to use the relaxation parameters to determine the actual frequencies of the motions. Small-amplitude motions often do not alter powder pattern lineshapes sufficiently to be detected by a lineshape analysis, but the relaxation parameters are sensitive indicators of the presence of motions, even if the amplitudes are small when the frequencies are near the Larmor frequencies of the nuclei.

The calculated NMR spectra in Figure 1 demonstrate the dramatic effects of large-amplitude motions on powder pattern lineshapes. The bottom row of spectra are powder patterns from the ^{15}N chemical shift anisotropy interaction of immobile backbone amide sites (Figure 1B), the ^{2}H quadrupole interaction from immobile C–D bonds (Figure 1E), and the ^{2}H quadrupole interaction from C–D bonds of methyl groups undergoing threefold reorientation (Figure 1H). The comparison of the powder pattern breadths in the spectra in Figures 1E and 1H indicate the spectroscopic effects of the limited motional averaging present in methyl groups. The top row of spectra (Figures 1A, 1C, and 1F) are single-line resonances that result from complete motional averaging. The highly characteristic lineshapes shown in Figures 1D and 1G are useful in studies of proteins, since experimentally observed lineshapes similar to these indicate the presence

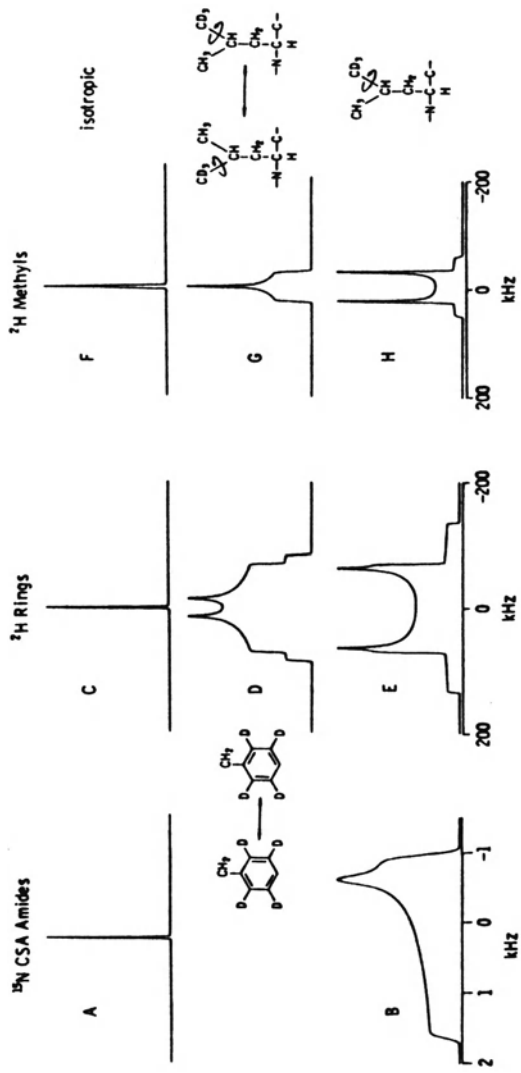


Figure 1. Calculated NMR spectra. (A) ^{15}N chemical shift anisotropy of amide group averaged by isotropic motion, (B) ^{15}N chemical shift anisotropy of immobile amide group, (C) ^{2}H quadrupole interaction of C–D bond averaged by isotropic motion, (D) ^{2}H quadrupole interaction of C–D bond averaged by twofold 180° ring flips, (E) ^{2}H quadrupole interaction of immobile C–D bond, (F) ^{2}H quadrupole interaction of a $-CD_3$ group averaged by isotropic motion, (G) ^{2}H quadrupole interaction interaction of a $-CD_3$ group averaged by two-site tetrahedral jump plus threefold methyl reorientation and, (H) ^{2}H quadrupole interaction of a $-CD_3$ group averaged by threefold methyl reorientation (from Colnago *et al.*, 1986).

of aromatic ring flips (Figure 1D) or hop motions of methyl groups about tetrahedral carbon sites (Figure 1G).

2.1. Backbone Motions

Proteins are highly structured polypeptides. Each can be thought of as a linear sequence of uniform rigid planar peptide groups linked together to form the backbone, with sidechains of widely varying size and mobility attached to the α carbons at the junction of two peptide planes. This oversimplified view focuses attention in studies of dynamics on identifying the exceptional few peptide backbone sites that are highly mobile, defining the small-amplitude motions of the rigid sites, and then using the backbone dynamics of the individual residues as a framework for analyzing their sidechain dynamics.

The main biological system that we are studying by solid-state NMR spectroscopy is the filamentous bacteriophage fd (Cross and Opella, 1985; Colnago *et al.*, 1987; Opella *et al.*, 1987). This virus serves the dual functions of providing a tractable model system for the development of new methods and presenting several interesting biological problems, especially with regard to its major coat protein. The virus particles are about 90% by weight coat protein, with several thousand copies of the small (5000 daltons) coat protein subunits arranged symmetrically around the single stranded circular DNA which is extended lengthwise within the filament (Marvin and Hohn, 1969; Makowski, 1984). Studies of the virus require solid-state NMR spectroscopy because its total mass is about 16×10^6 Daltons, which, combined with its shape, prevents it from reorienting rapidly in solution. The virus particles are immobile on the slowest timescales (10^3 Hz) of the NMR experiments. Since the individual subunits are small and the overall aggregate is large, this is a highly favorable system. Applications of solid-state NMR spectroscopy to proteins will be illustrated in this review with examples from our studies of the coat protein of the bacteriophage fd.

Since the individual coat protein subunits are immobilized by the protein–protein and protein–DNA interactions of the virus, and the virus particles are immobile in solution, all experimentally observed manifestations of motional averaging are a consequence of intramolecular motions. In general, large-amplitude rapid motions are readily recognized by their effects on powder pattern lineshapes as shown in Figure 1. If the motion is isotropic, the corresponding resonance is a single narrow line at the average frequency. While all of the available backbone sites and spin interactions are available for analysis, we have found ^{15}N NMR spectroscopy of labeled amide backbone sites to be an effective and reliable way of characterizing local motions of protein backbones (Colnago *et al.*, 1987; Bogusky *et al.*, 1987).

In its structural form, fd coat protein in the virus particle has several mobile residues, although the majority of the backbone sites are immobile and are rigidly held by the framework of the secondary and tertiary structure of the protein (Colnago *et al.*, 1987). This is readily seen with the solid-state ^{15}N NMR spectra presented in Figure 2. Two different ^{15}N labeled virus samples were used to obtain the experimental spectra shown. The uniformly labeled protein samples have ^{15}N in all nitrogen sites. There are two predominant chemical types of nitrogens in the protein giving observable resonances. The backbone amide sites have resonance intensity between 30 and 200 ppm with an isotropic frequency of about 100 ppm, and the amino groups from lysine sidechains and the amino terminus contribute to the narrow resonance near 15 ppm (Harbison *et al.*, 1984). The amino resonance is always narrow, regardless of the amino group dynamics, because of the small breadth of the amino nitrogen chemical shift tensor. In contrast, the breadth and the lineshape of the amide ^{15}N resonance is a sensitive monitor of the dynamics of the amide backbone sites. The calculated spectrum in Figure 1B is representative of a ^{15}N amide powder pattern for a site in a rigid lattice; it has a breadth of about 170 ppm (about 10 kHz at typical magnetic field strengths) and is nearly axially symmetric. Large-amplitude motions in all directions, which occur more frequently than the 10 kHz breadth of the powder pattern result in complete averaging of the powder pattern to a single narrow resonance at the isotropic position, as shown in the calculated spectrum in Figure 1A.

Experimental solid-state NMR spectra of ^{15}N labeled fd coat protein are shown in Figures 2A, 2B, and 2C. The uniformly labeled intact coat protein in the virus gives the spectrum in Figure 2C. There are clearly two

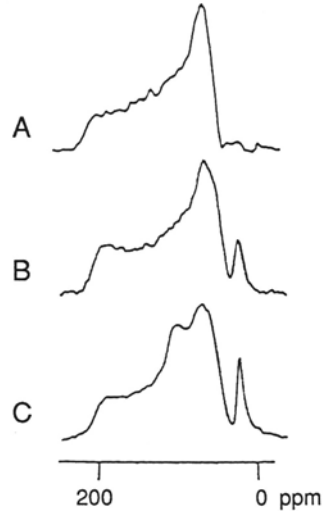


Figure 2. Experimental solid-state ^{15}N NMR spectra of ^{15}N labeled fd virus samples. (A) ^{15}N-Leu labeled fd, (B) uniformly ^{15}N labeled fd treated with the enzyme leucine amino peptidase, (C) uniformly ^{15}N labeled fd (from Colnago *et al.*, 1987; D. Y. Chen and S. J. Opella, unpublished results).

types of ^{15}N amide resonances; most of the backbone sites are immobile and contribute to the powder pattern that is very similar to the one in Figure 1B representative of an immobile amide site. However, some of the ^{15}N amide resonance intensity is in the form of a relatively narrow isotropic peak superimposed on the broad underlying powder pattern. This motionally averaged resonance intensity is from backbone sites undergoing effectively isotropic motion more rapid than the 10^4 Hz timescale of the chemical shift anisotropy interaction.

The N-terminal four residues of fd coat protein are mobile (Colnago *et al.*, 1987). Three of them contribute to the narrow amide resonance intensity at 100 ppm in Figure 2C (residues 2–4) while the terminal amino group has a resonance that overlaps with the other amino groups near 15 ppm. Two types of experiments assign the narrow amide resonance intensity to the N-terminal residues. They are illustrated with the spectra in Figures 2A and 2B. One approach is to use selective isotropic labeling. fd coat protein has two leucine residues (Leu-14 and Leu-41) that can be selectively labeled with ^{15}N. The powder pattern in Figure 2A is from an unoriented solution of fd virus labeled with ^{15}N-Leu; therefore it contains resonance intensity only from the two immobile Leu amide backbone sites (Cross and Opella, 1985). Ideally, the spectrum from only a single labeled site should be examined; however, there is little possibility for ambiguity with the qualitative assessment of only two overlapping powder patterns. The other approach is to use limited proteolytic cleavage, in this case with the enzyme leucine aminopeptidase. This enzyme cleaves solvent-exposed residues with free terminal amino groups. The spectrum in Figure 1B shows that the mobile residues are removed by this enzyme, and hence are at the N-terminus and are exposed to the solvent. This strategy has been successful in giving a fairly complete view of the backbone dynamics of the fd coat protein in its structural form in the virus (Colnago *et al.*, 1987), as well as in its membrane-bound form (Leo *et al.*, 1987). The transition between the membrane-bound form and the structural form is biologically important and, interestingly, the largest changes in the protein are in the dynamics of several C-terminal residues.

Even the backbone sites categorized as immobile do not have the properties of a rigid lattice. For example, the ^{15}N amide sites of the coat protein in the virus have $T_1 = 18$ sec which is much shorter than the $T_1 = 600$ sec for a polycrystalline model peptide (Cross and Opella, 1982). This drastic reduction in relaxation time reflects the presence of motions with frequencies near those of the Larmor frequency (10^9 Hz), yet the finding of a powder pattern unaveraged from that expected for a rigid lattice shows that there are no large-amplitude motions at these sites. Small-amplitude rapid motions account for both the powder pattern lineshape and relatively short T_1 for the amide nitrogen sites.

Highly mobile backbone sites of proteins may consist of one or a few residues at the N- or C-termini (Colnago *et al.*, 1987; Leo *et al.*, 1987) or several residues connecting structured domains such as α helices that span membranes (Keniry *et al.*, 1984). The vast majority of backbone sites are rigidly held as part of the protein secondary and tertiary structure, although they have small amplitude rapid motions. It may be possible to relate these small-amplitude motions observed in the NMR experiments to those discussed in terms of temperature factors from diffraction experiments and molecular dynamics calculations (Dobson and Karplus, 1986; Ringe and Petsko, 1986).

2.2. Sidechain Motions

The sidechains attached to the few mobile backbone sites in a protein are also mobile, since the entire residue or segment is undergoing rapid isotropic motion. Far more typical are residues with backbone sites that are immobile, but have independent sidechain motions. The study of sidechain motions is at a relatively early stage; however, examples of essentially all of the aliphatic and aromatic sidechain motions have been described.

The sidechains offer several nuclei and spin interactions for analysis. In most cases, the motions have been characterized by using the lineshapes of the ^{2}H-labeled sidechain sites. ^{2}H NMR is particularly convenient for these studies because most sidechain sites can be readily labeled with deuterium, the ^{2}H quadrupole interaction is relatively large and is well understood, and the motions of interest have large effects on the shape and breadth of the powder patterns.

Large-amplitude motions have been observed in both aromatic and aliphatic sidechains in residues with immobile backbone sites (Opella, 1985). Examples of large-amplitude motions have been observed for all sidechains, except for histidine and tryptophan, which are bulky asymmetric aromatic sidechains. In contrast, the large but symmetric aromatic sidechains of tyrosine and phenylaline have been shown to undergo rapid two-site 180° jump motions about the C_β-C_γ bond axis in solid-state NMR experiments. These aromatic ring flips dramatically illustrate the dynamic nature of proteins, especially since the flips must occur more often than 10^6 Hz in order to average the ^{2}H quadrupole interaction.

Aliphatic sidechain dynamics are potentially quite complex because of the opportunities for multiple modes of motion. However, despite all of the possibilities, most of the experimental measurements on aliphatic sidechains have been interpretable in terms of simple models. Sidechain sites, in particular the methyl groups, can be readily labeled and studied by ^{2}H NMR spectroscopy. Calculated ^{2}H NMR spectra for $-CD_3$ groups undergoing various types of motional averaging are shown in Figures 1F–H.

The calculated spectra are representative of those resulting from motional averaging of methyl groups bonded to a tetrahedral carbon. Rapid jumps among three or more equivalent sites result in the same averaging as seen of the powder pattern lineshape diffusional motion about the same axis (Spiess, 1982). The powder pattern in Figure 1H has a reduced breadth as compared to that in the Figure 1E, although the shape is unchanged. This reduction in breadth represents the averaging from the threefold reorientation of a $-CD_3$ group.

The interpretation of the lineshapes observed from $-CD_3$ groups on all sidechains other than alanine, where the methyl carbon is directly bonded to the α carbon, requires consideration of additional modes of motion, in particular two-site jump motions about tetrahedral carbon sites in the sidechain. Leucine sidechains have been shown to undergo twofold jump motions about the tetrahedral γ carbon in proteins, including collagen (Batchelder *et al.*, 1982) and fd coat protein (Colnago *et al.*, 1987) by solid-state ^2H NMR spectroscopy. The axially symmetric powder pattern from methyl groups undergoing threefold reorientation or the $\eta = 1$ powder pattern expected with twofold jumps is expected in the fast motional limits for the $-CD_3$ groups of leucine residues. Experimental ^2H NMR spectra for CD_3-Leu labeled coat protein obtained at temperatures between -10 °C and $+25$ °C are shown in Figures 3A–C. At low temperatures, the sidechains of both leucine residues appear to be immobile. Above about 10 °C both of the sidechains give powder patterns with $\eta = 1$, indicative of the twofold jump motions at the tetrahedral carbon site. The two leucine residues differ in their activation energy for the jump motion, since at 0 °C one of the sidechains is immobile and the other undergoes twofold jump motions.

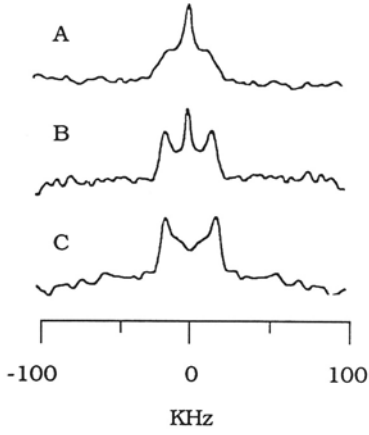


Figure 3. Experimental ^2H NMR spectra of ^2H-Leu labeled fd virus at (A) 25°C, (B) 0 °C, (C) −40 °C (from Colnago *et al.*, 1987).

Even at temperatures where the jump motions occur in both leucine residues, the backbone sites are immobile as seen in the ^{15}N chemical shift powder pattern in Figure 2A. The experimental ^{2}H quadrupole powder patterns at both low and high temperatures are very similar to those calculated for a methyl group undergoing only threefold reorientation (−10 °C) (Figure 1H) or threefold reorientation plus two-site tetrahedral jump motions (25 °C) (Figure 1G). For the experimental and calculated lineshapes of the sidechain sites to be so similar, the backbone sites must be immobile.

We have performed ^{2}H NMR experiments on essentially all of the aromatic and aliphatic sidechains of the structural (Colnago *et al.*, 1987) and membrane-bound (Leo *et al.*, 1987) forms of fd coat protein, which give, when combined with the results for the backbone sites, a fairly complete view of the dynamics of the protein.

3. PROTEIN STRUCTURE

The same anisotropic spin interactions that produce the rigid lattice and motionally averaged powder patterns used to describe protein dynamics (Opella, 1985) can also be used to describe protein structure (Opella *et al.*, 1987). The structural information can be in the form of distances or angles, depending on the sample and experimental conditions.

Structural information can be obtained from experiments with both oriented and unoriented samples. Quite different strategies are necessary for deriving structural information from oriented samples as compared to unoriented samples, with or without magic angle sample spinning. Although the same spin interactions are of interest, their spectral manifestations are very different. Figure 4 shows calculated spectra for ^{15}N amide chemical shift anisotropy in three different experimental regimes. The spectrum in Figure 4a is a rigid lattice powder pattern expected from an unoriented and stationary sample. If the sample is subjected to rapid rotation at the magic angle, then the isotropic resonance shown in Figure 4b results. The single-line resonance is equivalent to that obtained with rapid isotropic motion of the molecule. A single-crystal sample would yield a single-line spectrum, such as that in Figure 4c, although the actual resonance frequency depends on the orientation of the groups with respect to the direction of the applied magnetic field and can occur anywhere within the frequency limits defined by the breadth of the powder pattern in Figure 4a. The corresponding oriented spectrum for a heteronuclear dipole–dipole interaction between two spins, such as the ^{1}H–^{15}N coupling of the amide nitrogen, would be a doublet with the splitting a function of the orientation of the internuclear bond vector with respect to the direction of the magnetic field.

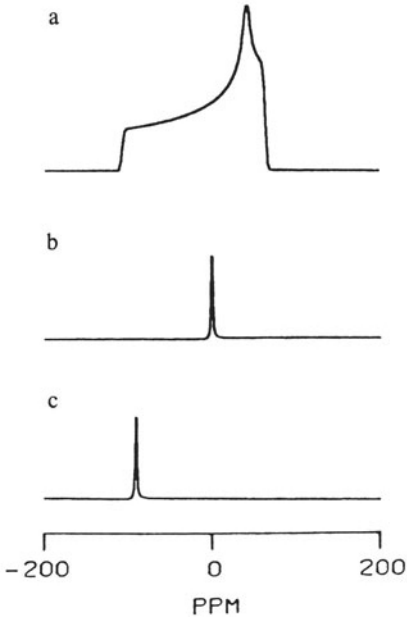


Figure 4. Calculated ^{15}N NMR spectra. (a) Powder pattern, (b) isotropic resonance, (c) oriented resonance (from Cross and Opella, 1985).

3.1. Oriented Samples

Oriented samples, whether in the form of single crystals, fibers, or oriented solutions, offer several advantages for NMR structural studies. Oriented samples give spectra that are relatively simple, yet contain directly interpretable structural information in the form of resonance frequencies and splittings (Opella and Waugh, 1977). These spectra have high resolution and can be obtained with high sensitivity. Many spectral parameters are a function of the orientation of the groups or bonds with respect to the applied magnetic field. Structural information can be obtained for both backbone and sidechain sites of proteins.

There are many common features to the use of all of the spin interactions to derive information about the orientations of the peptide backbone and sidechain groups. The initial step is to prepare an oriented sample of a protein. The experimental measurements involve the determination of the frequencies and splittings of the available spin interactions and the assignment of the resonances to specific sites in the protein. There are four functional elements in our analysis of the results from oriented samples (Stewart *et al.*, 1987; Opella *et al.*, 1987):

1. The experimental data are used to determine angular restriction, regardless of the chemical properties of the site.

2. The chemical and geometrical properties of the groups are taken into account. This often places strong restrictions on the orientations which are consistent with even a single experimental measurement.
3. A computer program is used to sort out the data in terms of protein structure. This utilizes the properties of peptide groups and any other restrictions that are established, such as the protein volume, the positions of disulfide bonds, or dominant types of secondary structure such as α helix and β sheet.
4. Molecular dynamics calculations and energy minimization are used to refine the NMR derived structures.

The backbone structure of a protein is fully described by the mutual orientation of the linked peptide planes. Therefore, if the orientation of the individual planes can be determined by the NMR experiments relative to a fixed axis, such as the applied magnetic field, then the mutual orientations of adjacent planes can be determined. This is the same information contained in Ramachandran conformational plots (Ramachandran and Sasisekharan, 1976) and is sufficient to describe the structure of the protein backbone. These orientations can be determined from the observable spectral parameters from the anisotropic spin interactions. The example presented here utilizes the heteronuclear dipole-dipole interactions defined in Figure 5. In the case of dipole-dipole couplings, it is possible to directly determine the angles of bonds relative to the direction of the applied magnetic field. However, we have found that it is better to treat all of the spin interactions in a consistent way by relying on the restrictions represented in the areas of plots of the two angles α and β. This gives a direct representation of the orientations consistent with the experimental measurements, as well as the

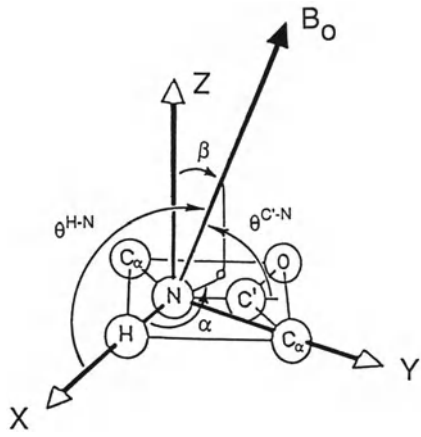


Figure 5. Peptide plane axis system, defining angles used in analysis.

uncertainties in experimental measurements and the magnitudes and orientations of the principal elements of the various interactions.

fd virus particles orient spontaneously in a strong magnetic field. The high degree of orientation of the virus filaments make structural studies feasible through experimental measurements of spectral parameters that yield angles between bonds or chemical groups and the filament axis, which is parallel to the direction of the magnetic field. It is possible to determine the structure of the protein from the orientations of the peptide planes that constitute the backbone of the protein. The heteronuclear ^{1}H-^{15}N and ^{13}C-^{15}N dipolar interactions are particularly valuable sources of angular information.

Figure 6 presents the data used to determine the orientation of the peptide plane of Leu-41. The one-dimensional ^{15}N NMR spectrum of ^{15}N-Leu labeled coat protein in an oriented fd sample is shown in Figure 6c. This spectrum is proton decoupled; therefore, the only spin interaction influencing the resonances is chemical shift anisotropy. There are two single line resonances, since there are two leucine residues in the protein and they are both labeled with ^{15}N. In principle, it is possible to use the chemical shift as a source of angular restriction; however, the ^{15}N amide chemical shift tensor has σ_{33} approximately colinear with the N-H bond, and we have found the heteronuclear dipolar interaction from this bond to be a more reliable source of angular restriction.

The magnitude of the N-H heteronuclear dipolar coupling can be measured for both of the leucine residues with a two-dimensional separated local field experiment. The contour plot from this experiment is shown in Figure 6d. The splittings of the doublets in the second dimension correspond to the individual N-H dipolar couplings. The C-N dipolar coupling can be measured in a one-dimensional spectrum on a double-labeled protein, where the leucine residues are labeled with ^{15}N and the lysine residues are labeled with ^{13}C at the carbonyl position. In this double labeled protein a single ^{13}C-^{15}N bond is formed between Lys-40 and Leu-41, resulting in a splitting of the ^{15}N resonance of Leu-41 shown in the experimental spectrum in Figure 6a and simulated in Figure 6b. This splitting serves two purposes. It assigns the ^{15}N resonances and their N-H dipolar splittings to specific leucine residues and it provides a measurement of the C-N dipolar coupling for Leu-41 (Cross and Opella, 1985).

Angular restriction plots, as presented in Figure 7, are the first step in the interpretation of the experimental data for the Leu-41 peptide plane. The dark areas represent the combination of the two angles which are consistent with the experimental NMR measurements and the geometry of the peptide plane. The left plot shows the combinations of angles consistent with the measured N-H dipolar splitting alone, the center plot shows the combinations of angles consistent with the measured C-N dipolar splitting

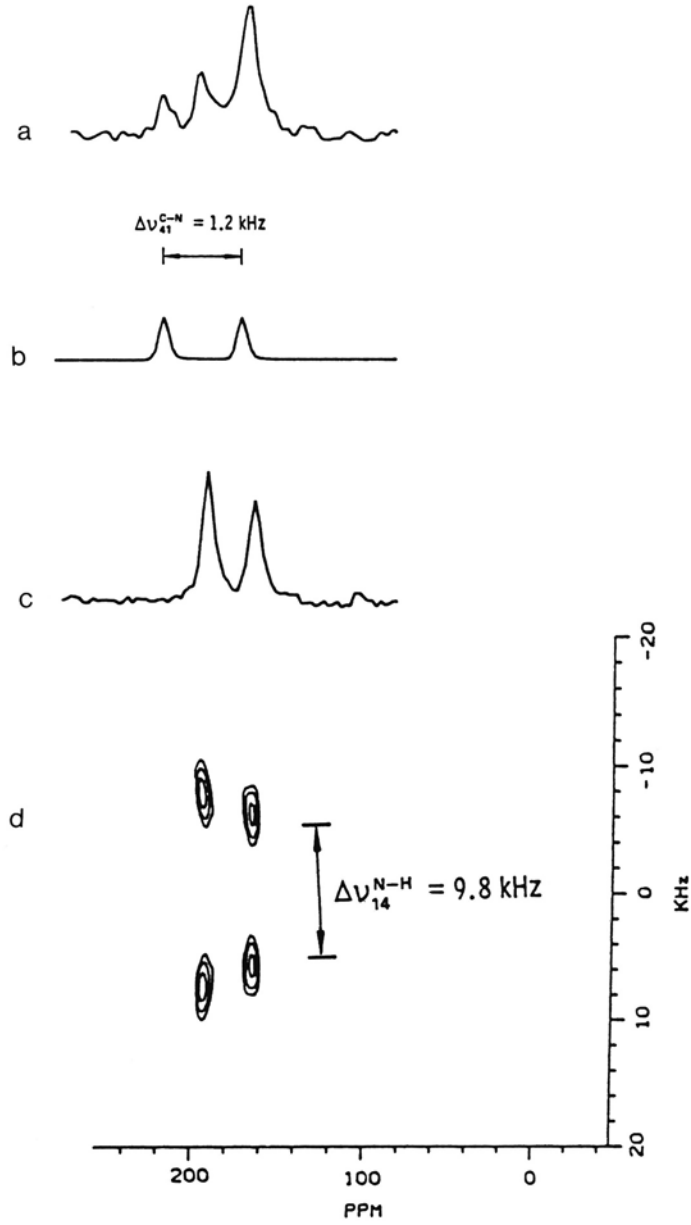


Figure 6. Experimental spectra of oriented ^{15}N-Leu labeled fd. (a) ^{15}N NMR spectrum, (b) ^{15}N NMR spectrum of double labeled sample, ^{15}N-Leu and ^{13}C-Lys, (c) simulated spectrum showing splitting due to ^{15}N-Leu-^{13}C-Lys dipole-dipole spectrum (from Cross and Opella, 1985).

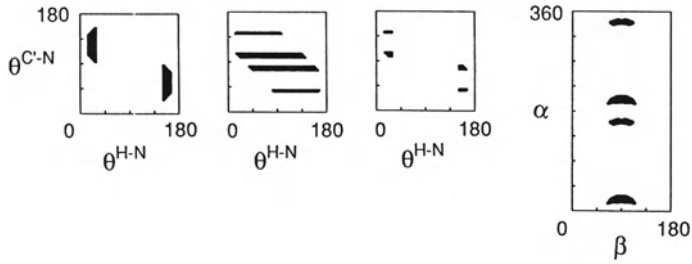


Figure 7. Angular restriction plots for the experimental data for Leu-41 shown in Figure 6.

alone, and the left plot shows the restrictions arising from the combination of the two experimental measurements. The combination of restrictions from multiple measurements determines the orientations of the planes more precisely and more reliably than possible with any single measurement. In order to employ a uniform axis system for all of the spin interactions, the angular restriction plots are converted to α vs. β plots. These angles are defined in the axis system in Figure 5. There are several orientations of the peptide plane for Leu-41 consistent with the NMR data. It is possible to choose the correct one using a computer program that selects for allowed conformational arrangements for adjacent peptide planes. Figure 8 shows the orientation of the peptide plane of Leu-41 with respect to the magnetic field vector.

It is also possible to determine the orientations of sidechains using this basic approach. The coat protein has a single tryptophan residue (Trp-26) and the orientation of its sidechain was determined using the spectroscopic

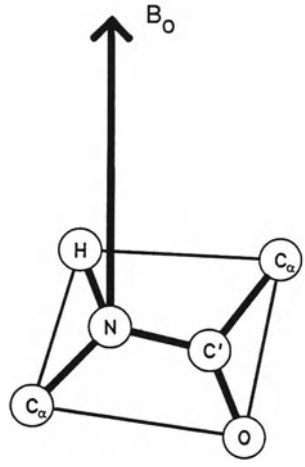


Figure 8. Orientation of the plane of Leu-41 derived from the data in Figure 6 and the analysis in Figure 7.

properties of the indole nitrogen (Cross and Opella, 1983). ^{2}H NMR is more generally applicable for sidechains. It also gives the opportunity to simultaneously describe the structure and dynamics of the sidechains. The powder pattern from the unoriented sample can be interpreted to describe the sidechain dynamics as discussed above. The oriented sample gives a doublet for each sidechain. The magnitude of the splitting can be used to describe the orientation of the sidechain.

We have carried out many measurements of dipolar couplings in oriented samples and of powder pattern lineshapes of unoriented samples of fd virus. The structure in Figure 9 summarizes the results, showing the structure of the backbone as determined from the analysis of parameters from the peptide backbone sites (Opella *et al.*, 1987). The dynamics of many of the sidechains have been determined and are indicated schematically (Colnago *et al.*, 1987).

Magnetic ordering of paramagnetic crystals of myoglobin, and possibly other proteins, results in uniaxial orientation of the molecules; therefore, these samples also give oriented spectra. Oldfield and co-workers (Oldfield and Rothgeb, 1980: Rothgeb and Oldfield, 1981; Lee and Oldfield, 1982) have used the magnetic ordering of myoglobin crystals to carry out structural studies using ^{2}H NMR spectroscopy. As shown for the leucine sidechains in Figure 8, each labeled site gives a ^{2}H NMR spectrum consisting of a doublet, the splitting of which is a function of the orientation of the group relative to the field direction. They applied these methods to $-CD_3$ groups in methionine sidechains and C–D bonds in the heme groups of myoblobin.

3.2. Unoriented Samples

Rapid diffusional or multisite hop motions result in averaging of spin interaction tensors. The averaging is highly characteristic in that an effectively axially symmetric powder pattern results, with the breadth and sign of the powder pattern reflecting the angle between the axis of molecular motion and the principal axis system of the spin interaction. Bacteriorhodopsin in reconstituted synthetic bilayers undergoes rotational diffusion about an axis perpendicular to the plane of the bilayer. ^{13}C-carbonyl-labeled bacteriorhodopsin reconstituted in synthetic lipids gives axially symmetric powder patterns (Lewis *et al.*, 1985). The analysis of the lineshape was able to limit the possible arrangements of secondary structure for bacteriorhodopsin. The filamentous bacteriophage Pf1 undergoes rapid diffusional rotation about its long axis in solution at high pH but not at low pH, presumably owing to differences in interparticle interactions. ^{13}C-carbonyl-labelled Pf1 coat protein showed motional averaging consistent with the protein being predominantly helical with the helix axis aligned approximately parallel to the filament axis (Tsang and Opella, 1986).

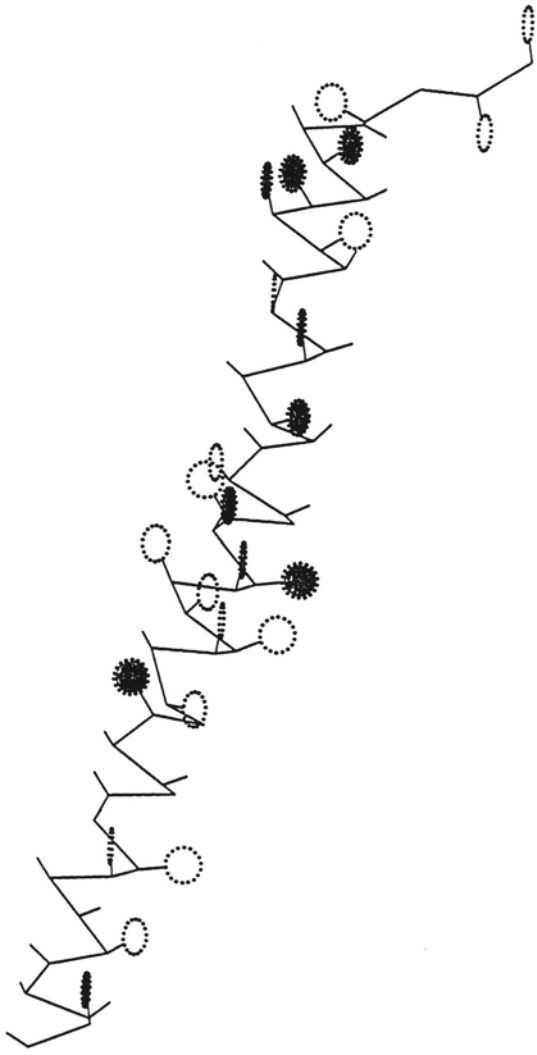


Figure 9. Schematic representation of the structure and dynamics of the coat protein in fd. The structure of the backbone atoms is as described in Opella *et al.* (1987). The dynamics of the sidechains are shown as dark for immobile groups and as white for mobile groups as described in Colnago *et al.* (1987).

Several types of solid-state NMR experiments are useful in obtaining structural information on immobile unoriented molecules subjected to magic angle sample spinning. Isotropic chemical shifts can be used as empirical parameters indicative of secondary structure, permitting direct comparisons

between solid-state and solution NMR spectra. The ^{13}C isotropic chemical shifts of some sites, including the carbonyl and α carbons, depend on local conformation. For rigid molecules, direct comparisons between the solid-state and solution structures are possible by obtaining ^{13}C NMR spectra of both states (Pease *et al.*, 1981). For flexible molecules, solid-state NMR studies permit the conformations to be analyzed without the complications of averaging and interconversion seen in solution (Saito *et al.*, 1985).

Polyalanine can adopt either α-helix or β-sheet conformations, depending upon the conditions. There are approximately 5 ppm differences in the chemical shifts of the carbon resonances for all sites between these two forms (Saito *et al.*, 1983; Kricheldorf and Muller, 1983). These shifts are typical of those seen in other homopolymers of amino acids that can exist in both α-helix and β-sheet conformations (Saito *et al.*, 1985). The correlations of chemical shifts with structure have been extended to structured proteins including silk fibroin (Saito *et al.*, 1983b; Kricheldorf *et al.*, 1983) and collagen (Saito *et al.*, 1984).

There are several other ways in which the NMR spectra generated in magic angle sample spinning experiments can be used to derive structural information from unoriented samples. The spectral effects of nitrogen on the resonances of bonded carbons are dependent on distance and angular factors; the lineshapes of the carbon resonances can be used as a source of structural information about peptide bonds (Hexem *et al.*, 1982). The heteronuclear dipole-dipole coupling can be used to determine C-H and N-H bond lengths in unoriented samples (Frey *et al.*, 1985). Spin exchange experiments are able to differentiate sites that are close in space from those that are more distant in peptides and proteins (Frey and Opella, 1984).

4. PERSPECTIVE

Solid-state NMR spectroscopy of proteins continues to offer great promise. This promise lies both in extending NMR spectroscopy to proteins and samples that can not be studied effectively by X-ray diffraction or solution NMR spectroscopy and in providing a high resolution of protein structure and dynamics. The successful studies to-data on model systems—bacteriorhodopsin (Oldfield *et al.*, 1982; Rice *et al.*, 1981), collagen (Torchia, 1982), and filamentous bacteriophage coat protein, as described in this review—indicate that this promise will be fulfilled.

The instrumentation and methods for solid-state NMR spectroscopy are continually being improved; therefore, the technical barriers to protein studies are being reduced. Of equal importance, stable isotopes are being used with increasing sophistication. The combinations of biochemical methods and spectroscopy that can be fruitfully applied to the study of

proteins are leading to nearly unlimited possibilities for understanding proteins.

ACKNOWLEDGMENTS. D. Bouchard, D.-Y. Chen, and P. Stewart helped prepare the figures for this review. The research performed at the University of Pennyslvania was supported by N.I.H. grants GM-24266, GM-29754, and GM-34343.

REFERENCES

Andrew, E. R., Bradbury, A., and Eades, R. G., 1958, *Nature* **182**:1659.

Batchelder, L. S., Sullivan, C. P., Jelinski, L. W., and Torchia, D. A., 1982, *Proc. Natl. Acad. Sci. USA* **79**:386.

Bogusky, M. J., Schiksnis, R. A., Leo, G. C., and Opella, S. J., 1987, *J. Mag. Res.* **72**:186.

Colnago, L. A., Leo, G. C., Valentine, K. G., and Opella, S. J., 1986, in *Biomolecular Stereodynamics III* (R. H. Sarma and M. H. Sarma, eds.), Adenine Press, Guilderland, pp. 147–158.

Colnago, L. A., Valentine, K. G., and Opella, S. J., 1987, *Biochemistry* **26**:847.

Cross, T. A., and Opella, S. J., 1982, *J. Mol. Biol.* **159**:543.

Cross, T. A., and Opella, S. J., 1983, *J. Am. Chem. Soc.* **105**:306.

Cross, T. A., and Opella, S. J., 1985, *J. Mol. Biol.* **182**:367.

Cross, T. A., Frey, M. A., and Opella, S. J., 1983, *J. Am. Chem. Soc.* **105**:7741.

Dobson, C., and Karplus, M., 1986, *Meth. Enzymol.* **131**:362.

Frey, M. H., and Opella, S. J., 1984, *J. Am. Chem. Soc.* **106**:4942.

Frey, M. H., Opella, S. J., Rockwell, A. L., and Gierasch, L. M., 1985, *J. Am. Chem. Soc.* **107**:1946.

Fyfe, C. A., 1983, *Solid State NMR for Chemists*, C.F.C. Press, Guelph, Ontario.

Gall, C. M., DiVerdi, J. A., and Opella, S. J., 1981, *J. Am. Chem. Soc.* **103**:5039.

Gall, C. M., Cross, T. A., DiVerdi, J. A., and Opella, S. J., 1982, *Proc. Natl. Acad. Sci. USA* **79**:101.

Gurd, F. R. N., and Rothgeb, T. M., 1979, *Adv. Prot. Chem.* **33**:73.

Haeberlen, U., 1976, *Adv. Magn. Res.* Suppl., Academic Press, New York.

Harbison, G. S., Jelinski, L. W., Stark, R. E., Torchia, D. A., Herzfield, J., and Griffin, R. G., 1984, *J. Mag. Res.* **60**:79.

Hexem, J. G., Frey, M. H., and Opella, S. J., 1982, *J. Chem. Phys.* **77**:3847.

Hirs, C. H., and Timasheff, S. N., *Meth. Enzymol.* **131**.

Keniry, M. A., Gutowsky, H. S., and Oldfield, E., 1984, *Nature* **307**:383.

Kinsey, R. A., Kintanar, and Oldfield, E., 1981, *J. Biol. Chem.* **256**:9028.

Kricheldorf, H. R., and Muller, D., 1983, *Macromolecules* **16**:615.

Kricheldorf, H. R., Muller, D., and Ziegkler, K., 1983, *Polymer Bull.* **9**:284.

Lee, R. W., and Oldfield, E., 1982, *J. Biol. Chem.* **257**:5023.

Leo, G. C., Colnago, L. A., Valentine, K. G., and Opella, S. J., 1987, *Biochemistry* **26**:854.

Lewis, B. A., Harbison, G. S., Herzfeld, J., and Griffin, R. G., 1985, *Biochemistry* **24**:4671.

Makowski, L., 1984, in *Biological Macromolecules and Assemblies*, Vol. I, *The Viruses*, (A. McPherson, ed.), Wiley, New York, pp. 203–253.

Marvin, D. A., and Hohn, B., 1969, *Bacteriol. Rev.* **33**:172.

Mehring, M., 1983, *High Resolution NMR in Solids*, 2nd ed. Springer-Verlag, Berlin.

Oldfield, E., and Rothbeb, T. M., 1980, *J. Am. Chem. Soc.* **102**:3635.
Oldfield, E., Kinsey, R. A., and Kintanar, A., 1982, *Meth. Enzymol.* **88**:310.
Opella, S. J., 1985, *Meth. Enzymol.* **17**:327.
Opella, S. J., 1986, *Meth. Enzymol.* **131**:327.
Opella, S. J., and Waugh, 1977, *J. Chem. Phys.* **66**:4919.
Opella, S. J., Stewart, P. L., and Valentine, K. G., 1987, *Quart, Rev. Biophys.* **19**:7.
Pake, G. E., 1948, *J. Chem. Phys.* **16**:327.
Pease, L. G., Frey, M. H., and Opella, S. J., 1981, *J. Amer. Chem. Soc.* **103**:467.
Pines, A., Gibby, M. G., and Waugh, J. S., 1973, *J. Chem. Phys.* **59**:569.
Purcell, E. M., Torrey, H. C., and Pound, R. V., 1946, *Phys. Rev.* **69**:37.
Ramachandran, G., and Sasisekharen, V., 1976, *Adv. Protein Chem.* **23**:283.
Rice, D. M., Blume, A., Herzfeld, J., Wittebort, R. J., Huang, T. H., Das Gupta, S. K., and Griffin, R. G., 1981, *Biomol. Stereodynam.* **2**:255.
Rice, D. M., Wittebort, R. J., Griffin, R. G., Meirovitch, E., Steinson, E. R., Meinwald, Y. C., Freed, J. H., and Scheraga, H. A., 1981b, *J. Am. Chem. Soc.* **103**:7707.
Ringe, D., and Petsko, G., 1986, *Meth. Enzymol.* **131**:389.
Rothgeb, T. M., and Oldfield, E., 1981, *J. Biol. Chem.* **256**:1432.
Saito, H., Tabeta, R., Shoji, A., Ozaki, T., and Ando, I., 1983a, *Macromolecules* **16**:1050.
Saito, H., Iwanaga, Y., Tabeta, R., Narito, M., and Asakura, T., 1983b, *Chem. Lett.* p. 427.
Saito, H., Tabeta, R., Shoji, A., Ozaki, T., and Ando, I., 1983a, *Macromolecules* **16**:1050.
Saito, H., Tabeta, R., Shoji, A., Ozaki, T., Ando, I., and Miyata, T., 1984, *Biopolymers* **23**:2279.
Saito, H., Tabeta, R., Shoji, A., Ozaki, T., Ando, I., and Asakura, T. 1985, in *Magnetic Resononance in Biology and Medicine* (G. Govil, C. L. Khetrapal, and A. Saron, eds.) Tata McGraw-Hill, New Delhi, 195–215.
Schaefer, J., and Stejskal, E. O., 1976, *J. Am. Chem. Soc.* **98**:1031.
Spiess, H. W., 1982, *NMR Basic Princ. Prog.* **15**:55.
Stewart, P. L., Valentine, K. G., and Opella, S. J., 1987, *J. Mag. Res.* **71**:45.
Torchia, D. A., 1982, *Meth. Enzym.* **82**:174.
Torchia, D. A., 1984, *Ann. Rev. Biophys. Bioeng.* **13**:124.
Tsang, P., and Opella, S. J., 1986, *Biopolymers* **25**:1859.
Waugh, J. S., Huber, L. M., and Haeberlen, U., 1968, *Phys. Rev. Lett.* **20**:180.
Wüthrich, K., 1986, *NMR of Proteins and Nucleic Acids*, Wiley-Interscience Publications, New York.
Wyckoff, H., Hirs, C., and Timasheff, S., 1985, *Meth. Enzymol.* Vols. **114, 115**.

6

Methods for Suppression of the H_2O Signal in Proton FT/NMR Spectroscopy

A Review

Joseph E. Meier and Alan G. Marshall

1. INTRODUCTION

Fourier transform NMR in dilute aqueous solutions is essential for the analysis of the solution structure and dynamics of biological molecules (e.g., proteins, carbohydrates, phospholipids, RNA, DNA), drugs, and a wide variety of polar molecules (Wüthrich, 1976; Wüthrich, 1986; Campbell-Burke and Shulman, 1987; Brindle and Campbell, 1987; Shulman *et al.*, 1979; Radda and Seeley, 1979; Kuchel *et al.*, 1984; Rabenstein, 1984; Scott and Baxter, 1981; Roberts and Jandetzky, 1981). Although much can be learned about the nonexchangeable protons from FT/NMR conducted in D_2O solution, exchangeable protons (e.g., base-pair protons in RNA and DNA) can obviously be observed only in H_2O. Even in D_2O, the residual HDO signal can be large enough to obscure signals at or near the HDO peak.

The principal problem posed by proton NMR in aqueous solution is dynamic range: the H_2O signal is approximately 100 M in protons, whereas solute concentrations may be 10^{-3} M or less. In the absence of noise, even a 16-bit analog-to-digital converter (ADC) can accommodate a dynamic range of only about 32 K (15 bits plus sign) without clipping the time-domain free-induction decay (FID). In the presence of noise, that dynamic range

Joseph E. Meier and Alan G. Marshall • Department of Chemistry, The Ohio State University, Columbus, OH 43210. Alan G. Marshall is also a member of the Department of Biochemistry.

can be extended to a limit determined by the word length (typically 24 or 32 bits) of the random-access memory in which the accumulated FID is stored, but at the cost of accumulating a very large required number of FIDs. Moreover, the analog electronic circuitry cannot be held linear over such a wide dynamic range. Additional detection problems (spinning sidebands, curved spectral baseline, phasing, quantitation of relative peak areas, etc.) also result directly from the dynamic range problem.

Even with perfect analog and digital detection, the H_2O macroscopic magnetic moment at high magnetic field strength ($\geq$10 tesla) is so large that it induces a large current in the detector coil. The resultant power loss due to I^2R heating of the coil must occur at the expense of a corresponding loss in the H_2O signal strength. The H_2O FID therefore decays faster at high than at low field strength, and the corresponding frequency-domain H_2O peak is broadened (e.g., to $\geq$10 Hz at 11.75 tesla, 500 MHz for ^1H). The broad solvent signal further increases the difficulty of observing proton NMR peaks near the chemical shift of H_2O, ~4.8 ppm downfield from 2,2-dimethyl-2-silapentane-5-sulfonate (DSS).

Solutions to the above dynamic range problem are all based on difference in one or more NMR parameters (e.g., Larmor frequency, T_1, T_2, J-coupling, etc.) of the solvent compared to the solute protons. In this review of more than 100 published methods for solvent suppression, we have tried to group related techniques together, based on a combination of the principle of suppression (e.g., difference in Larmor frequency) and the experimental method (e.g., hard-pulse, soft-pulse, software manipulation) by which that principle is implemented. For example, the H_2O magnetization may be reduced before detection by frequency-selective excitation (i.e., excitation of spectral signals at all frequencies except that of the solvent), or selective presaturation of only the H_2O signal. Similarly, after detection, various hardware and software manipulations may be used to cancel the solvent signal after a given cycle of FIDs accumulated with different timing or event sequences.

We have attempted to provide sufficient detail to enable the reader to understand the principle and practice of each solvent suppression method. In addition, we present a critical comparison, supported by several experimental examples, of the advantages and disadvantages of the various methods.

2. FREQUENCY-SELECTIVE EXCITATION

The most widely used water-suppression methods are based on the difference in Larmor frequency between H_2O and other protons in the sample of interest. In all of these methods, any resonances at or near the

H_2O chemical shift are also suppressed. Nevertheless, most of the remainder of the spectrum can be recovered with the water signal suppressed by a factor of up to 50,000 or more.

2.1. Continuous-Wave and Rapid-Scan Techniques

Prior to the advent of Fourier transform (FT) NMR (Ernst and Anderson, 1966), NMR spectra were obtained by continuous-wave (CW) scanning across the desired spectral range, slow so as to maintain a steady-state response, to excite and detect one resonance at a time. In CW/NMR, the solvent signal could therefore be eliminated simply by not scanning through the frequency range at or near the solvent Larmor frequency. Although effective in terms of solvent suppression, CW/NMR has long since been abandoned because FT/NMR offers the Fellgett multichannel advantage (factor of ~10,000 in speed) of acquiring the whole spectrum at once rather than one spectral element (channel) at a time (Marshall and Verdun, 1989).

FT/NMR is based on the response to excitation (usually one or more fixed-frequency radiofrequency pulses) during a period which is short compared to nuclear magnetic relaxation times T_1 and T_2. In CW/NMR, the excitation (usually a frequency sweep) varies slowly compared to T_1 and T_2. In rapid-scan NMR, the excitation frequency is swept adiabatically through the various Larmor frequencies, so that the macroscopic magnetization remains aligned along the effective field direction in the rotating frame throughout the excitation period (Arata and Ozawa, 1976; Dadok and Sprecher, 1974). When used for water suppression, two rapid-scan frequency sweeps that pass through the resonances of interest (but not through the water resonance) are employed (see Figure 1). The first sweep, at high rf power, produces an adiabatic inversion (equivalent to that from a 180° fixed-frequency rf pulse). The second sweep, at lower rf power, generates an observable transverse magnetization which is acquired and digitized in the usual way. An optional homogeneity-spoiling pulse (HSP) may be

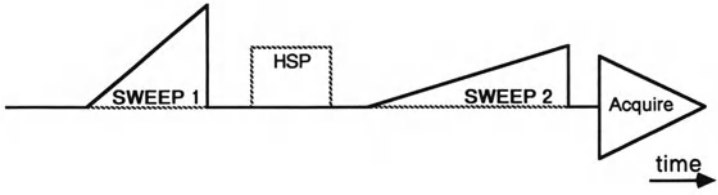


Figure 1. Time-domain event sequence for rapid-scan solvent suppression. An initial high-power frequency sweep (Sweep 1) which adiabatically inverts all solute resonances is followed by an optional homogeneity-spoiling pulse (HSP) to eliminate transverse magnetization. A second lower-power rf frequency sweep (Sweep 2) generates the observed transverse transient magnetization.

inserted between the two frequency sweeps, in order to eliminate any transverse magnetization remaining after the original adiabatic inversion. Since the frequency scale precision is determined by the precision in sweep rate, the sweep rate ($\sim$250 Hz s^{-1} for ^{1}H NMR) must be very accurate ($\leq \pm 0.1\%$). Compared to FT/NMR, rapid-scan NMR has lower signal-to-noise ratio, since not all resonances are excited simultaneously. Another disadvantage is that only a small part of the spectrum can be scanned at a time and each side of the water line must be scanned separately. Finally the spectrum can be perturbed considerably when there are spin–spin interactions of varying degree (Hoult, 1976).

2.2. Tuneable Notch Filter

Another early method for water suppression is the tunable notch filter (Marshall *et al.*, 1979). In this method the whole spectrum (including the H_2O resonances) is excited and detected, but signals at or near the Larmor frequency of H_2O are selectively removed before analog-to-digital conversion, by either a passive band-reject filter or an active tuneable band-reject filter (Marshall *et al.*, 1979). Advantages of the notch filter are that (a) neither complex pulse sequences nor double irradiation are needed; (b) the method may be applied to both FT and CW/NMR; (c) solvent suppression is relatively insensitive to variation in solvent relaxation times; and (d) the method is simple to implement. Disadvantages are that (a) suppression by a factor of more than about 25 is difficult; (b) the filter introduces phase and amplitude distortion for resonances near the solvent frequency; and (c) as in all frequency-selective methods, the frequency range at or near the solvent Larmor frequency is unobservable.

2.3. Frequency-Selective Hard Pulse Sequences

Frequency-selective solvent suppression produced by a precisely timed sequence of high-amplitude ("hard") fixed-frequency rf pulses presently offers the most popular method for water suppression. Since an rf pulse of duration T excites all resonances within a frequency range of $\sim(1/T)$ Hz of the rf carrier frequency, each hard pulse typically excites the solvent as well as the solute resonance. However, the pulse sequence is designed such that the solute magnetization components are tipped into the x–y plane for observation, whereas the solvent magnetization direction at the end of the pulse sequence is realigned approximately along its original orientation along the applied static magnetic field (z-) axis and is thus unobservable.

Most of the frequency-selective hard pulse sequences were originally devised from a linear-system approximation, for which the observed spectral

response magnitude produced by a given pulse sequence is simply proportional to the frequency-domain excitation magnitude obtained by FT of the time-domain excitation signal. The linear approximation is valid only when (a) the response is small (i.e., small tip angles); (b) the excitation period is short compared to T_1 and T_2; and (c) the Larmor frequency is at or very near to the rf carrier frequency, so that the direction of the effective magnetic field in the rotating frame is approximately the same for all chemical shifts in the excitation range. The nonlinear effects may be taken into account by a Bloch equation approach to be discussed in Section 2.6 (Warren and Silver, 1988). Nevertheless, the pulse sequences designed from the linear-system approximation can be remarkably effective, as we shall show below.

2.3.1. Sine Sequences

Hard (i.e., very short) rf pulses are designed to approximate δ-function time-domain "spikes." Since the FT of a sinusoid is a δ-function, a series of time-domain δ-function pulses offers a convenient means for generating frequency-domain sinusoids. For example, consider the frequency-domain $\sin^n$ functions (see Figure 2):

$$S_n(\nu) = \sin^n(\pi\nu\tau) \tag{1}$$

The frequency-domain functions of equation (1) may be generated by Fourier transformation from a sum of $n + 1$ time-domain δ-functions:

$$\sum_{k=0}^{n} (-1)^k \frac{n!}{k!(n-k)!} \delta\left(t + \frac{k-n}{2}\tau\right) \tag{2}$$

in which the δ-function "spikes" are seen to be equally spaced at intervals of $\tau/2$ seconds, with alternating signs and with amplitudes given by the binomial coefficients. For example, the pulse sequence corresponding to $n = 1$ is denoted $1\underline{1}$, and consists of equal-magnitude positive and negative time-domain spikes separated by a delay period, to give a frequency-domain excitation spectrum with envelope $\sin(\pi\nu\tau)$. The pulse sequence corresponding to $n = 2$, denoted $1\underline{2}1$, consists of three pulses of alternating sign and relative magnitudes 1:2:1, whose frequency-domain spectrum has the form $\sin^2(\pi\nu\tau)$. The $n = 3$ pulse sequence, $1\underline{3}3\underline{1}$, produces a $\sin^3(\pi\nu\tau)$ spectrum, and so on. In the shorthand notation for such "trigonometric" pulse sequences, the numbers represent the relative magnitudes of the pulses and each underline represents a 180° phase shift. In practice, it is more convenient to produce rf pulses of equal magnitude and spacing, but varying width rather than varying height, as shown in Figure 3.

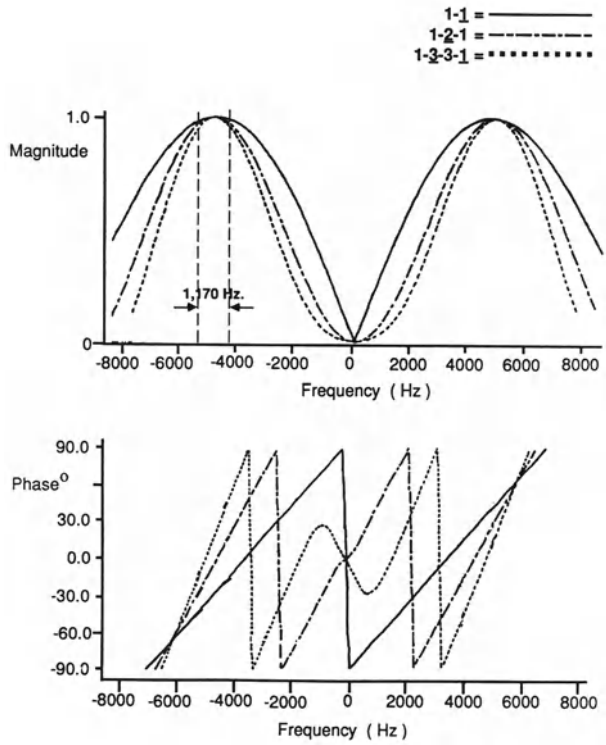


Figure 2. Frequency-domain magnitude (top) and phase (bottom) spectra, $\sin^n(\pi\nu\tau)$, resulting from $1\underline{1}$ ($n = 1$), $1\underline{2}1$ ($n = 2$), and $1\underline{3}3\underline{1}$ ($n = 3$) time-domain single-frequency hard pulse sequences, from equation (2) and Figure 3. The carrier frequency, ν, is set at the center of the solvent resonance. Note that as n increases, the "null" range of zero-excitation magnitude at or near the solvent frequency widens, but at the expense of narrowing the frequency range for maximal excitation of solute resonances at other frequencies. The $1\underline{3}3\underline{1}$ time-domain excitation waveform corresponding to the $\sin^3(\pi\nu\tau)$ frequency-domain excitation spectrum represents a compromise between those two effects. Note the approximately linear (and therefore correctable) phase variation with frequency in the phase spectra (bottom plot).

When used for solvent suppression, the carrier frequency ν of a sine-based hard pulse sequence is set equal to the solvent Larmor frequency, thereby producing zero excitation magnitude at the center of the solvent spectral peak (Hore 1983a, Sklenar and Starcuk, 1982; Hetherington *et al.*, 1985; Bleich and Wilde, 1984). Figure 2 shows that the excitation null widens (for more effective suppression of the full solvent peak width) as n increases, i.e., in proceeding from $\sin(\pi\nu\tau)$ to $\sin^2(\pi\nu\tau)$ to $\sin^3(\pi\nu\tau)$, etc. However, the same Figure shows that the range of (solute) frequencies over which the excitation magnitude is maximal decreases as n increases. Thus, the best choice of n in equation (2) represents a compromise between

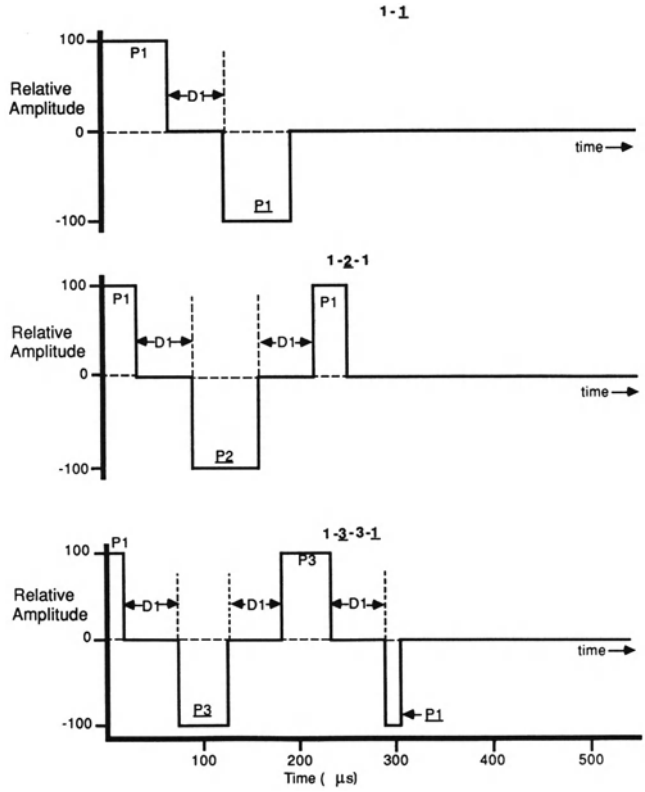


Figure 3. Time-domain hard pulse sequences for the 1$\underline{1}$ (top), 1$\underline{2}$1 (middle), and 1$\underline{3}$3$\underline{1}$ (bottom) excitation waveforms whose frequency-domain magnitude and phase spectra are shown in Figure 2. In each case, acquisition begins promptly after the designated series of pulses (P) and delays (D). The spectra of Figure 2 correspond to $D1 = 96$ μsec and $P1$ amplitudes which produce 45°, 22.5°, and 11.25° rotation of magnetization, respectively, where a 90° pulse requires 160 μsec duration.

optimal solvent suppression and optimal detection uniformity and magnitude for solute detection.

The excitation magnitude spectra in Figure 2 do not necessarily reflect the relative response spectra, due to off-resonance and relaxation effects discussed in Section 2.5. Also, as the number of pulses increases, timing and phasing errors become more important, and the increasingly steep variation of phase with frequency renders spectral phase correction more difficult (Turner, 1983).

In practice, the 1$\underline{3}$3$\underline{1}$ sequence (with the carrier frequency set to the H_2O Larmor frequency) gives excellent results when the solute resonances of interest are relatively far from H_2O and/or the desired range of solute

frequencies is small. $D1$ is then chosen so that the desired spectral frequency range is excited with approximately constant magnitude. Figure 4 graphically illustrates the effect of $D1$ duration on the excitation process, either from the time-domain trajectories of the solute and solvent magnetization vectors or from the frequency-domain excitation magnitude spectra. From

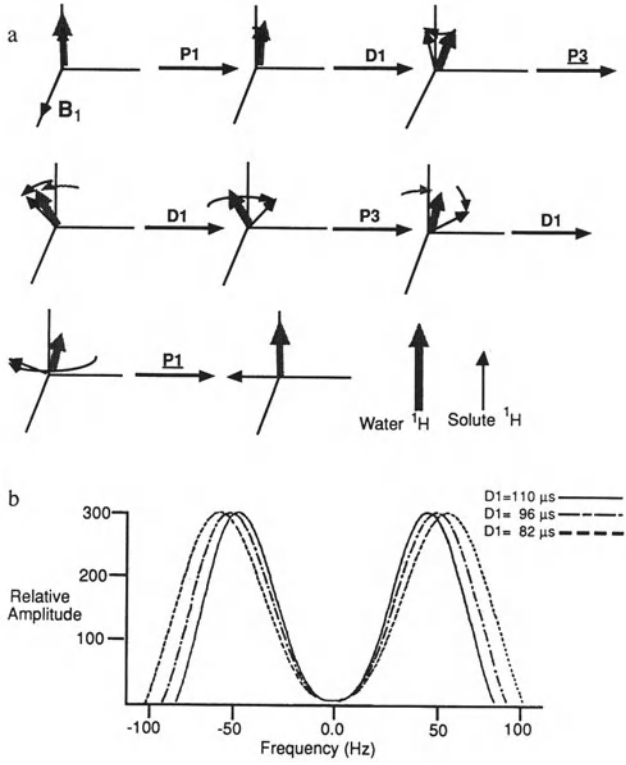


Figure 4. Time-domain (a) and frequency-domain (b) representations of the $1\underline{3}3\underline{1}$ pulse sequence of Figure 3. In the time-domain rotating-frame picture, the magnetization vectors for water protons and for solute protons of a different Larmor frequency are displayed at various stages of the $1\underline{3}3\underline{1}$ event sequence. The first rf pulse ($P1$) flips both solvent and solute magnetization vectors by 11.25° about the rotating-frame rf magnetic field, $\mathbf{B}_1$. During the first delay period ($D1$), the solute magnetization rotates by 180° about the z axis, but the water magnetization remains stationary. The next pulse ($P\underline{3}$) flips the solute magnetization by an additional 33° toward the xy plane while the water magnetization rotates by −33.75° back across the z axis. Similarly, the next two pulses and delays result in a net flip angle of $(11.25 + 33.75 + 33.75 + 11.25)° = 90°$ for the solute magnetization, with the water magnetization realigned $(11.25 - 33.75 + 33.75 - 11.25)° = 0°$ back along the positive z axis at the end of the final pulse ($P\underline{1}$). In the frequency-domain representation, the excitation frequency-domain magnitude as a function of frequency away from the carrier (solvent) frequency is plotted for three different time-domain pulse spacings, $D1 = 110$, 96, and 82 μsec.

either standpoint, if $D1$ is made longer, resonances with Larmor frequencies closer to water may be observed, but resonances farther from the water Larmor frequency no longer fall in the frequency range of maximum excitation magnitude and are at least partially suppressed.

Solvent suppression with sine pulse sequences poses several phasing problems. First, performance is degraded if the time-domain phase shifts are not exactly 180°. That problem can be largely cured by inverting the transmitter and receiver phase on alternate scans. Another improvement is to increment the phases of all four pulses and the phase of the receiver in 90° steps, as in the CYCLOPS scheme discussed in Section 3 (Hore, 1983b). With these modifications, the $1\underline{3}3\underline{1}$ sequence can reduce the solvent signal by a factor of ~1,000. However, Figure 2 shows that any trigonometric pulse sequence results in a substantial variation of phase with frequency in the final response spectrum: the more pulses (and hence the more delay periods between pulses), the larger the (linear) phase variation. For example, Figure 2 shows the phase variation over the range of approximately constant magnitude excitation (1,170 Hz, denoted by vertical dotted lines in Figure 2). Moreover, as the number of pulses increases, Figure 2 shows that the variation of phase with frequency near the null (carrier) frequency becomes markedly nonlinear (and therefore virtually impossible to phase-correct). As a result, phasing of a response spectrum produced by (e.g.) a $1\underline{3}3\underline{1}$ pulse sequence is usually limited to a narrow spectral range far from the solvent Larmor frequency. Recently, a binomial pulse sequence has been combined with an "exorcycle" phase-cycling scheme (e.g., $[1\underline{3}3\underline{1}]_{\pm x}$, add, $[1\underline{3}3\underline{1}]_{\pm y}$, subtract) to eliminate frequency-dependent phase shifts to leave only a zero-order phase shift (Kienlin *et al.*, 1988).

The $1\underline{3}3\underline{1}$ sequence is generally preferred over the $1\underline{2}1$ and $1\underline{1}$ sequences because its broader null makes it less sensitive to the placement of the carrier frequency at the exact center of the solvent spectral peak. In addition, the $1\underline{3}3\underline{1}$ sequence is less sensitive to off-resonance effects due to finite rf excitation magnitude. However, the $1\underline{1}$ and $1\underline{2}1$ sequences introduce smaller and more linear frequency-domain phase variation, and may be more useful in applications requiring wider spectral bandwidth and/or resonances nearer to the solvent Larmor frequency.

2.3.2. Cosine Sequence

The cosine hard pulse sequences are identical to the corresponding sine sequences except that there are no 180° phase shifts (e.g., 1331 for $n = 3$), and the null is offset by $1/2$ ($D1$) Hz from the rf carrier frequency (Plateau *et al.*, 1983; Clore *et al.*, 1983). Thus, the cosine sequence carrier frequency is set at the middle of the solute Larmor frequency range, with $D1$ adjusted to give zero excitation magnitude at the solvent Larmor

frequency. In practice, the cosine sequences have not been as widely used, probably because of their greater susceptibility to off-resonance effects which shift the null frequency away from $1/2(D1)$ Hz. Finally the "even" (e.g., 11 or 1331) cosine or sine sequences are preferred over their "odd" (121 or 14641) counterparts, because the "even" sequences always exhibit a null (at least to first order) even for pulses with incorrect relative amplitude (or width) ratio (e.g., $1:\underline{2.8}:2.8:\underline{1}$) (Hore, 1983b).

2.3.3. Variations of Binomial Pulse Sequences

The basic binomial pulse sequences have been modified in several minor ways for both one-dimensional (Starcuk and Sklenar, 1985; Starcuk and Sklenar, 1986; Hall and Hore, 1986; Sklenar and Bax, 1987) and two-dimensional (Santos *et al.*, 1984; Hore, 1984) NMR. For example, a 180° (π) pulse added after the normal sequence acts to decrease the phase shift errors introduced by the binomial sequences (Galloway *et al.*, 1987) (see, e.g., Figure 5 for its use in a $1\underline{3}3\underline{1}$ sequence), but at the cost of reduction in the degree of suppression of the solvent signal. The NERO (nonlinear excitation with rejection on-resonance) hard pulse sequence, consisting of six pulses and a series of delays, has been designed (from Bloch equation analysis) to yield no phase distortion across the spectral range of interest (Levitt and Roberts, 1987). In yet another variation, presaturation and partial relaxation (see Section 3) precede a binomial sequence (preferably $1\underline{1}$) in order to detect solvent-exchangeable solute protons with resonances close to the water Larmor frequency (Tong and Kotovych, 1986). Other variations are designed to produce more uniform excitation magnitude over the spectral range of interest (Morris *et al.*, 1986; Wang and Pardi, 1987), e.g., OBTUSE (offset binomial tailored for uniform spectral excitation) (Morris *et al.*, 1986), which has some advantages for two-dimensional NMR. Figure 6 shows the event sequence for one of the best modified binomial

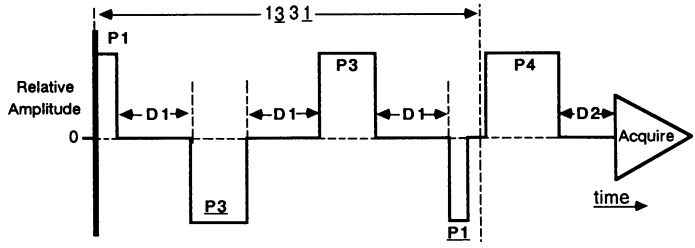


Figure 5. Frequency-selective hard-pulse sequence which partially compensates for the spectral phase variation introduced by binomial sequences. There is no delay between the last $P\underline{1}$ pulse of the normal $1\underline{3}3\underline{1}$ sequence and the $P4$ (180°) pulse. $D2$ is a short echo delay.

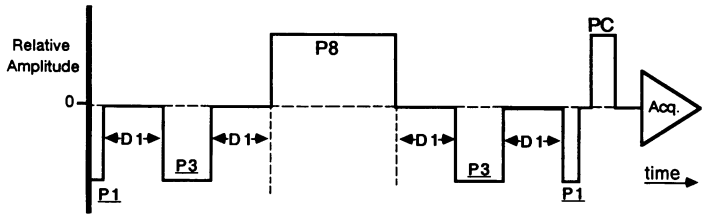


Figure 6. OBTUSE hard-pulse sequence. PC is a correction pulse that is 90° out of phase with respect to the remaining pulses. Pulse *P*8 is phase-shifted by 180° and pulse *PC* by 90° with respect to pulses *P*1 and *P*3.

hard pulse sequences, in which the null occurs at the carrier frequency and the widths of the *P*8 and correction (PC) pulses are empirically optimized to yield water suppression by a factor of >1000 (Morris *et al.*, 1986).

The "jump and return" pulse sequence is a $1\underline{1}$ sequence ($90°_y - D1 - 90°_y$) in which the solute magnetization components fan out in the x-y plane and the H_2O magnetization rotates 180° during *D*1, so that the second 90° pulse flips only the H_2O magnetization back to the z-axis (Plateau and Gueron, 1982; Guittet *et al.*, 1984). Although easy to implement, its degree of solvent suppression is not as high as the $1\underline{3}3\underline{1}$ binomial sine sequence, and it is relatively sensitive to time-domain phase errors.

2.4. Frequency-Selective Soft Pulse Sequences

"Hard" and "soft" rf pulses differ in their relative duration (~10 μsec vs. ≥ 200 μsec for ^1H) and (inversely) corresponding relative magnitude. The simplest "soft" pulse for water suppression is a single fixed-frequency rf pulse of duration T for which the frequency-domain magnitude spectrum gives a null at ~$0.61/T$ Hz away from the rf carrier frequency (Redfield and Gupta, 1971). Redfield *et al.* (1975) noted that a much improved suppression could be obtained by dividing such a single pulse into five segments. The resulting Redfield "214" sequence, which would be denoted $2\underline{1}4\underline{1}2$ in our former binomial pulse notation, has the time-domain waveform and frequency-domain magnitude and phase spectra (along with those for a $1\underline{3}3\underline{1}$ binomial sequence for comparison) shown in Figure 7. The Redfield 214 soft-pulse sequence has no delays between the component pulses, so that its frequency-domain spectrum spans a bandwidth of approximately $6/P1$ Hz, in which *P*1 is the duration of the shortest component pulse of the sequence. The two *P*1 pulses are phase shifted by 180° with respect to the other three pulses. The duration of each *P*1 pulse is $\frac{1}{6}$ of the duration of a single fixed-frequency 90° pulse of the same

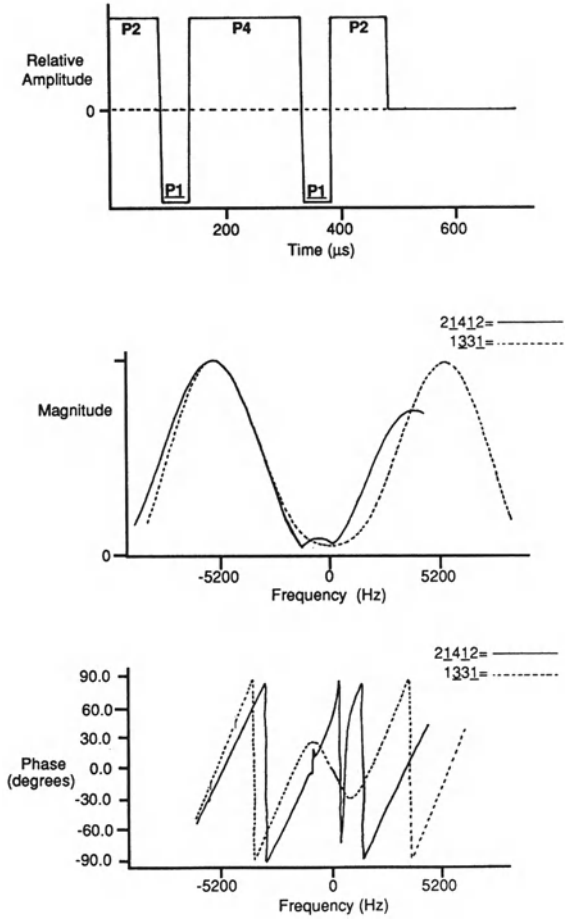


Figure 7. Time-domain (top) and frequency-domain magnitude (middle) and phase (bottom) spectra of a Redfield 214 soft rf pulse sequence. In this example, the duration of $P1$ was set at 38.4 μsec in order to excite approximately the same bandwidth as the $1\underline{3}3\underline{1}$ binomial sequence of Figures 2 and 3. The magnitude and phase spectra of the $1\underline{3}3\underline{1}$ sequence are also included for comparison.

time-domain magnitude. The Redfield 214 pulse sequence frequency-domain null occurs at $\sim 1/6(P1)$ Hz away from the carrier frequency (i.e., almost the same as for a single soft pulse of the same total excitation period.

Figure 7 shows that the null frequency range for the Redfield 214 sequence is somewhat narrower than that for the binomial $1\underline{3}3\underline{1}$ sequence. Moreover (see Figure 7), the Redfield 214 method is suitable for detection only above or below the Larmor frequency, but not both, because of its asymmetrical excitation magnitude. Also, the small hump at the Redfield

214 null frequency makes placement of the carrier frequency critical. Although independent adjustment of the durations of each of the component pulses can reduce the hump (Redfield *et al.*, 1975), optimization of the individual pulse lengths (and of the carrier frequency) can be tedious. Nevertheless, a properly optimized Redfield 214 soft pulse sequence can yield comparable water suppression (at least on one side of the water spectral peak) to that of the 1331 sequence. Finally, the phase spectrum of Figure 7 shows that the Redfield 214 sequence introduces phase variation comparable to that of the binomial sequences.

Bloch analysis has been used to find a soft pulse sequence with a flatter and broader frequency-domain null than the Redfield 214 sequence. One such sequence is the soft 1-2-1 (Yao *et al.*, 1986; Hall *et al.*, 1984). The actual flip angles produced by each component pulse were 34.2°-68.4°-34.2°. For a solution (e.g., a protein) with a broad water resonance, this 121 sequence appears to offer an attractive alternative to the Redfield 214.

Frequency-selective hard and soft pulse sequences can produce extremely good water suppression, at the obvious cost of suppressing any other resonances with chemical shifts at or near that of water. Soft pulses have the advantage that they do not require high rf power. However, hard pulses have several other advantages. First, because the duration of each component hard pulse is typically short compared to the reciprocal of the frequency range of interest, all spins with Larmor frequencies in that range are excited to a comparable extent, except for off-resonance effective field direction effects, mentioned above. Also, soft pulses require stable rf amplitude throughout the excitation period, and significant relaxation may occur during the relatively long excitation period required for soft pulse excitation. Short (hard) pulses also tend to be easier to generate for typical FT/NMR spectrometers, which are designed to produce short high-power rf pulses.

Finally hard pulses can be used to look at both sides of the solvent line at the same time. Hard pulses would not be the sequences to use if the region around water is important, because they leave it unobservable.

2.5. Tailored Excitation

Tailored excitation (Brandes and Kearns, 1985; Tomlinson and Hill, 1973; Hsu *et al.*, 1987) is the name given to a time-domain waveform or series of pulses designed in advance to have a specified frequency-domain magnitude spectrum (in this case, with zero magnitude at the solvent Larmor frequency). In this way, it is possible to obtain a frequency-domain excitation that is essentially flat over the (solute) NMR frequency range of interest. Typically, the "tailoring" is performed on a audio-frequency time-domain waveform which is mixed with a fixed-frequency rf carrier to produce selective excitation over the desired narrow rf bandwidth.

In an analog approach, a "tailored" audio-frequency signal may be generated by passing a broadband audiofrequency signal through a combination of filters designed to produce a null at what will become the H_2O Larmor frequency after rf mixing. Audiofrequency filtering is preferred to rf filtering because of the difficulty of producing a tunable high-Q rf filter (Brandes and Kearns, 1985). For example, a band-pass filter can be adjusted to pass excitation with maximum magnitude at either $+\nu_0$ or $-\nu_0$ with respect to the H_2O Larmor frequency; the width of the pass-band of frequencies then determines the "quality factor" of the filter and its degree of solvent suppression (Brandes and Kearns, 1985). Although capable of producing relative flat excitation power above and below the H_2O resonance frequency, this method has the disadvantages that (a) the experiment takes twice as long, if signals both above and below the H_2O frequency are to be detected; (b) additional filter hardware is required; and (c) nonlinear phase and/or amplitude distortion is introduced by the analog filters.

Much better results are obtained by use of Fourier synthesized pulses (Tomlinson and Hill, 1973; Hsu *et al.*, 1987). Originally proposed by Tomlinson and Hill (1973), the method consists of specifying a time-domain discrete waveform (whose FFT has the desired frequency-domain excitation magnitude spectrum) which is then passed through a digital-to-analog converter (DAC) to generate the desired time-domain analog excitation waveform. The method has been improved and extended to Fourier transform ion cyclotron resonance mass spectroscopy (Marshall, 1985) and magnetic resonance imaging (Hsu *et al.*, 1987) as SWIFT (stored waveform inverse fourier transform) excitation. In SWIFT (see Figure 8), the desired magnitude mode discrete frequency spectrum is specified and stored in computer memory. The corresponding time-domain discrete waveform is then produced by FFT. Since a broad frequency-domain spectrum corresponds to a narrow time-domain signal, the inverse FT step of the above procedure can produce a time-domain waveform of inconveniently large dynamic range. In order to reduce that dynamic range, it is useful to introduce a quadratic phase variation in the original stored discrete frequency-domain data before inverse FT (Chen and Marshall, 1987a; Chen and Marshall, 1987b). Finally, the discrete time-domain stored waveform is apodized (to reduce truncation effects) and clocked out through a DAC to the transmitter. Compared to the binomial pulse sequences, SWIFT has the advantages of much flatter excitation magnitude over the solute Larmor frequency range; more efficient use of excitation power, since the power is concentrated only in the frequency range(s) of interest; and correctable phase variation in the final response FT spectrum. Solute resonances on either side of the solvent may be excited and detected simultaneously. The required waveforms may be generated by the same computer that is used for FT of experimental FID's. As for all frequency-selective methods, the

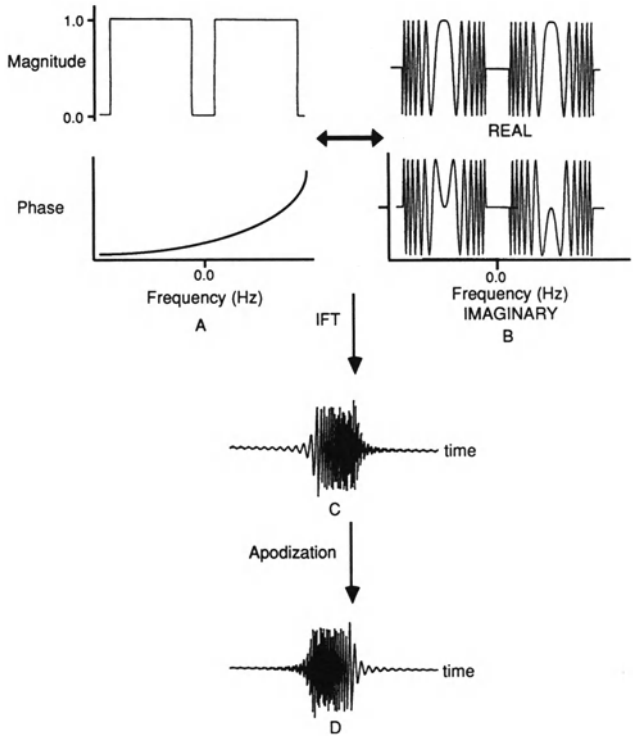


Figure 8. Schematic diagram for generation of a stored waveform inverse Fourier transform (SWIFT) excitation audiofrequency waveform for water suppression. The desired frequency-domain magnitude spectrum (A), containing a null at and near the solvent Larmor frequency, is quadratically modulated to reduce time-domain dynamic range. The magnitude/phase spectrum of (A) is displayed in its equivalent real/imaginary complex representation. Inverse Fourier transform of (B) yields the time-domain waveform (C). In this case, (C) is real only because of the symmetry of (B). Apodization (to reduce truncation effects) of (C) yields the time-domain waveform (D), which is then passed through a digital-to-analog converter, mixed with an rf fixed-frequency carrier, and passed to the transmitter amplifier.

SWIFT technique is not suitable for detection of signals at or very near the solvent Larmor frequency.

2.6. Nonlinear Excitation

In the preceding discussion, we have neglected three consequences of the Bloch equations. First, the magnitude of the spin system response (observed as the transverse magnetization) is proportional to the magnitude of the excitation only for very small tip angle ($\sin\theta \approx \theta$). Second, the

magnitude and direction of the effective rotating-frame magnetic field $\mathbf{B}_{\text{eff}}$ is different for nuclei of different Larmor frequency:

$$\mathbf{B}_{\text{eff}} = B_1\mathbf{i} + (B_0 - \omega/\gamma)\mathbf{k} \tag{3}$$

in which B_1 is the rf magnetic field magnitude, B_0 is the applied static magnetic field magnitude, γ is the magnetogyric ratio of the resonance of interest, and ω is the (angular) frequency of rotation of the rotating frame (i.e., $\omega = 2\pi\nu$, where ν is the frequency of the rf carrier in Hz). Third, longitudinal and transverse relaxation with rate constants $(1/T_1)$ and $(1/T_2)$ occur during the excitation period, particularly for the longer excitation periods needed for highly frequency-selective excitation. For all of the above reasons, we cannot simply assume that the frequency-domain response of a spin system will be proportional to the FT of a given excitation waveform. The problem is particularly acute for 180° pulses (see Figure 9).

As a result of the Bloch equation behavior, several time-domain excitation waveforms with near-identical Fourier transform magnitude spectra cam produce markedly different NMR spectral responses. Unfortunately, the responses for only a very few time-domain waveforms have been computed analytically under general conditions (e.g., as a function of frequency offset between carrier and Larmor frequency, time-domain amplitude, and initial condition of the spin system). One such time-domain waveform is the $(\text{sech}\ \alpha T)^{1+5i}$ function shown in Figure 9(a). Let the maximum time-domain magnitude of the $(\text{sech}\ \alpha T)^{1+5i}$ waveform which produces a flip angle of ~180° for on-resonance nuclei be denoted as $1X$. Figures 9(b–e) show that waveforms of the same shape, but with maximum time-domain amplitude $4X$, $16X$, and $64X$, respectively, also produce a near-perfect inversion of magnetization. Alternatively, suppose that we produce a flip angle of ~180° by use of a conventional (fixed-amplitude) rf time-domain rectangular pulse applied for approximately the same length of time. Figures 9(b–e) show that the $4X$, $16X$, and $64X$ versions of the square pulse do *not* produce good inversion when (as in this case) the original tip angle is not exactly 180°. Figure 9 gives one example of how nonlinear systems can produce some surprises to those who treat the system as if it were linear.

Amplitude- and/or phase-modulated waveforms of the type shown in Figure 9(a) are rapidly gaining popularity in NMR, especially in *M*agnetic *R*esonance *I*maging (MRI) (Warren and Silver, 1988; Gutow *et al.*, 1985; McCoy and Warren, 1987; Lurie, 1985). In MRI the improved slice profile and lowered power dissipation offered by such waveforms makes them highly successful. For NMR solvent suppression, time-domain waveforms have been designed (by analysis of their Bloch equation effects) to optimize one or more effects: frequency-selective excitation, uniform excitation, and

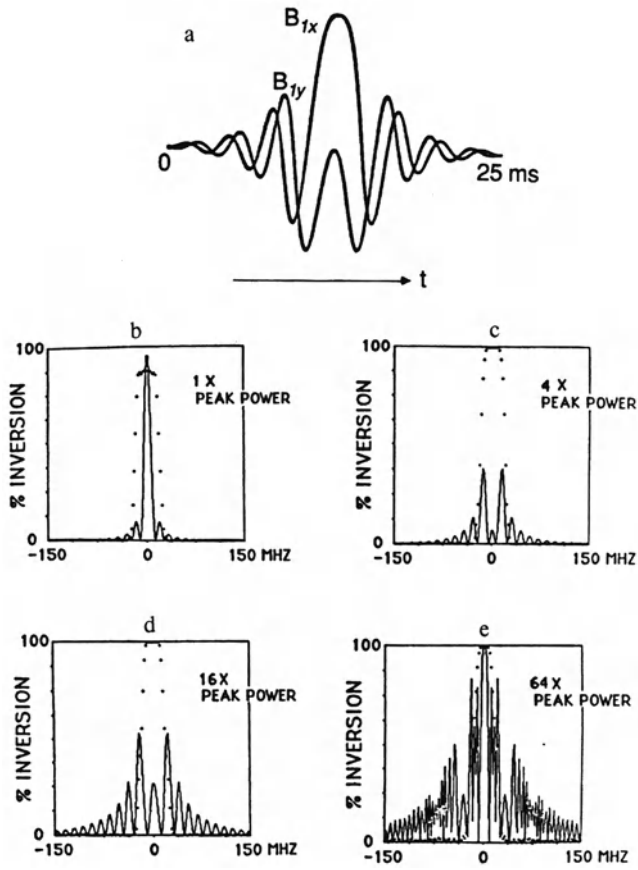


Figure 9. Simulated NMR frequency-domain magnitude response to two excitation waveforms. (a) Real and imaginary components of the complex time-domain waveform $(\mathrm{sech}\alpha T)^{1+5i}$. (b)-(e) Response to single-frequency pulse of rectangular envelope and duration, T (solid curves) and response to the waveform of (a) (dotted curves) for time-domain amplitudes, $1X$, $4X$, $16X$, and $64X$ peak power. A rectangular pulse of amplitude $1X$ peak power is scaled to produce complete inversion (i.e., ~180° pulse) at the Larmor frequency, labeled as 0 Hz in the Figure. (Because the original 180° rectangular pulse did not produce exactly 180° inversion on resonance, the rectangular pulses of $4X$, $16X$, and $64X$ amplitude produce finite magnitude on resonance.) (Adapted with permission from W. S. Warren, 1988, *Science* **242**:878.

reduced power dissipation (Warren and Silver, 1988; Warren, 1988; Brereton *et al.*, 1989). For example, one waveform used for water suppression consists of two parts: a frequency-selective waveform which achieves a +90° flip angle for a narrow band of frequencies centered at the solvent Larmor frequency followed by a second waveform which produces a broadband −90° flip angle (Gutow *et al.*, 1985; McCoy and Warren, 1987). Although

successful for both one-dimensional (Gutow *et al.*, 1985) and two-dimensional (McCoy and Warren, 1987) NMR experiments, that technique required a large dynamic range from the rf amplifier and created phase distortions arising from the extended duration of the selective pulse (Loaiza *et al.*, 1988).

A new narrowband-reject waveform devised by means of computer optimization (see Figure 10) proceeds by first constructing the desired frequency domain response. The desired response is then inverse Fourier transformed to give a time-domain representation. The true response produced by that time-domain waveform is then calculated by integration of the Bloch equations by fourth-order Runge–Kutta approximation. The difference between the desired and actual response is calculated and used to correct the frequency-domain representation of the pulse. The method proceeds iteratively by using a Powell's conjugate gradient minimization method to control the search until the desired accuracy is achieved (McCoy *et al.*, 1988). This procedure produces uniform and complete excitation of the solvent resonance and is self-refocusing (Loaiza *et al.*, 1988). The NMR response is obtained without the large phase distortions which characterize many other frequency-selective waveforms.

For a given time-domain rf magnitude and frequency-domain excitation bandwidth, tailored excitation (particularly when crafted by Bloch equation criteria) offers time-domain waveforms with lower dynamic range, less sensitivity to radiation damping and other relaxation effects (e.g., presence

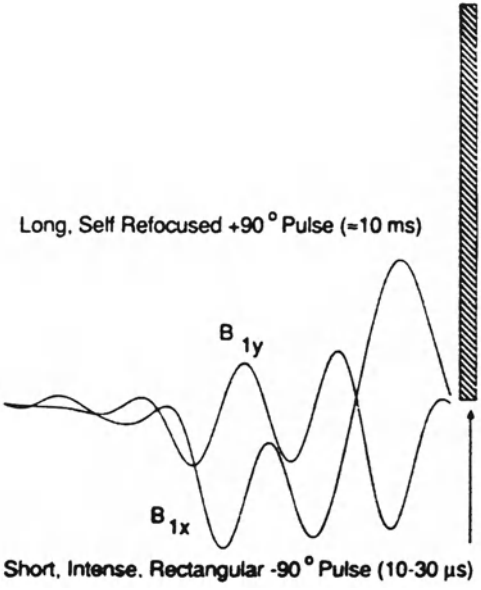


Figure 10. Phase and amplitude-modulated ERASER pulse shape used for self-refocused selective excitation. B_{1x} and B_{1y} denote the real (in phase) and imaginary (90° out of phase) components of the time-domain excitation waveform. (Taken, with permission, from M. A. McCoy, F. Loaiza, K., Valentine, and W. S. Warren, *J. Mag. Res.* **1988**, 155.)

of paramagnetic relaxation reagents), and reduced spectral phase distortion than with hard pulses (including composite pulses) (Loaiza *et al.*, 1988). However, Bloch-crafted waveforms do require hardware modifications to generate and apply the pulses (e.g., phase locking for successive narrowband and broadband waveforms). Moreover, optimization of the waveform shape and magnitude can be complicated and can be quite sensitive to the initial conditions and/or optimization criteria (Warren and Silver, 1988). Because many different waveforms may have similar effects, there is no single method for finding a "best" waveform. However, this promising avenue is under active investigation, and more general methods may soon be available.

3. SELECTIVE PRESATURATION

A general class of water suppression methods is based on selective saturation of the H_2O resonance, followed by broadband (nonselective) excitation of the solute resonances. In this case, saturation is achieved by continuous pre-irradiation at the H_2O Larmor frequency until the two magnetic energy levels are equally populated, so that no observable H_2O signal can be generated for a short time (compared to T_1 of H_2O) afterward. (We shall discuss pulsed T_1-based methods in Section 5.1.)

Selective presaturation is readily performed by means of time-shared gated homonuclear decoupling in the continuous-wave mode (Bleich and Glasel, 1975), as shown in Figure 11. Another somewhat more complicated method uses a double-resonance approach (Bleich and Glasel, 1975; Wider *et al.*, 1983), in which the selective rf-saturating irradiation from the decoupler is left on throughout the experiment (including the signal acquisition period). Even with gated decoupling, a residual rf spike at the H_2O Larmor frequency results from leakage between the decoupler transmitter and the receiver. The rf spike may be attenuated by phase synchronization of the

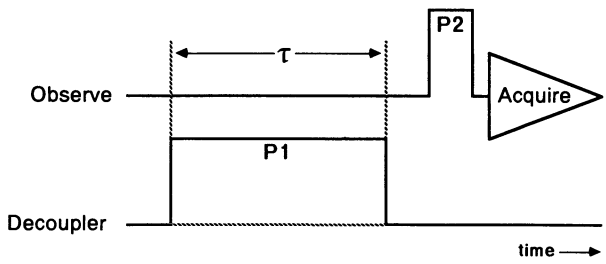


Figure 11. Event sequence for water suppression by means of continuous-wave selective presaturation. The first rf pulse *P*1 tuned to the H_2O Larmor frequency is left on for time τ to saturate the water protons. A second (nonselective) rf pulse *P*2 then generates an observable magnetization for the solute protons.

decoupler transmitter and the reference oscillator for the receiver channel, by 180° phase alteration of the $P2$ pulse in Figure 11, and by signal averaging (Krishna, 1976).

$$P1 - (\pi/2)_x - \text{add to memory} - 5T_1 - P1 - (\pi/2)_{-x} - \text{subtract from memory} \quad (4)$$

in which it is understood that T_1 in event sequence (4) refers to the solute protons.

It is important that the saturating irradiation frequency does not drift with respect to the water Larmor frequency during an irradiation period which may last several seconds. Further improvements include choosing the same phase for the saturation and observation pulses ($P1$ and $P2$ in Figure 11) and use of CYCLOPS (*cycl*ically *o*rdered *p*hase *s*equence) for the $P2$ pulse to minimize quadrature image peaks (Hoult, 1976).

Another variation is to use a long (soft) pulse for saturation and apply it through the same coil used for the nonselective excitation (Cutnell *et al.*, 1980), thereby freeing the decoupler for other uses. In this way, the saturating pulse can be made much longer in duration to produce more complete and frequency-selective saturation.

In practice, all of the presaturation methods leave a residual water signal. First, the water magnetization partially recovers during the period between $P1$ and $P2$ (Cutnell *et al.*, 1980). Second, perfect saturation is rendered more difficult by static magnetic field inhomogeneity, so that not all of the H_2O protons are equally affected by $P1$, leading to baseline anomalies near the H_2O peak (Hoult, 1976), particularly when the frequency-domain water peak is broad. On the other hand, presaturation methods are relatively easy to implement and use, and can in favorable cases suppress the water peak by a factor of >1000, without the spectral phasing problems introduced by the delays in the frequency-selective multiple-pulse techniques described above. A major disadvantage of presaturation is that the intensities of resonances from solvent-exchangeable protons will be reduced by cross-relaxation or chemical exchange. Finally, as for the methods of Section 2, presaturation methods render the spectral range near the water Larmor frequency unobservable.

4. SOFTWARE-BASED METHODS

In this section, we shall discuss those methods which reduce the relative magnitude of the water signal after it reaches the analog-to-digital converter. These techniques are generally designed to reduce the water signal by an additional factor of up to ≥ 100, when combined with another method which

reduces the water signal before it reaches the ADC. Hardware-based methods are therefore designed to reduce the analog dynamic range to fit into an ADC word. Software-based methods further reduce the digital dynamic range to avoid FFT roundoff errors and reduce baseline curvature and other anomalies.

The fast discrete Fourier transform (FFT) algorithm generates roundoff errors from multiplication and addition, as well as from the repeated rescaling required to prevent computer word overflow (Cooper, 1976; Cooper *et al.*, 1977). These errors effectively reduce both the maximum signal-to-noise ratio and dynamic range obtainable after the FFT procedure and the precision with which any spectral parameter (peak height, peak width, peak frequency) can be determined (Josef and Sprecher, 1974).

In general, floating-point FFT arithmetic gives substantially better precision than the fixed-point arithmetic of some co-processors. The FT roundoff error is greatest when the memory is essentially full at the start of the transform (Cooper, 1976). Thus, particularly in quadrature NMR, the rf carrier can be placed very near to the H_2O Larmor frequency, so that fewer memory locations near the beginning of the FID (where the H_2O signal is strongest) are full. FT roundoff error decreases with increasing number of time-domain data points (Josef and Sprecher, 1974). Zero-filling does not affect the dynamic range but does lower the signal-to-noise ratio (Cooper, 1976) and precision (Josef and Sprecher, 1974).

One trick for reducing the size of the H_2O signal is to adjust the rf carrier frequency and digitizing rate such that the H_2O Larmor frequency falls exactly at the Nyquist frequency. The time delay between excitation and acquisition is then adjusted until the H_2O signal is sampled (twice per cycle) exactly at the zero-crossings of the H_2O time-domain damped sinusoid, to yield zero signal at the H_2O frequency (Cooper, 1976). However, in practice it is very difficult to adjust the frequency and timing with sufficient precision to achieve large suppression of the H_2O.

In the block averaging method, the time-domain FID analog signal is amplified until it just fills the ADC word length (typically, 12 or 16 bits) without clipping. Discrete FID's are co-added in the computer memory until the accumulated FID just fills the length of a computer word (say, 24 bits), and an FFT is performed. The process is repeated, and the discrete FT spectra added together. In this way, the data points at or near the center of the solvent peak will overflow the memory, but the small solute signals will continue to accumulate. Thus, block-averaging does not increase dynamic range in the sense of ratio of largest to smallest peak, but does increase the signal-to-noise ratio of the small peaks, which is the end goal. Disadvantages of block averaging are (a) the need to interrupt acquisition repeatedly for the multiple FFT steps; (b) difficulty in observing small peaks which rest on the "shoulder" of the memory-overflowed H_2O peak; (c)

reduction of maximum available memory space by a factor of two (since two frequency-domain data sets must be stored before co-addition); and (d) the necessity for phase stability from one block to the next to be high if an absorption-mode spectrum is desired.

Double-precision arithmetic can also be used to increase dynamic range (provided that the precision of the sine lock-up table is also increased to double precision), at the expense of reducing the maximum available number of data points by a factor of two.

Absolute-value (magnitude-mode) spectral display is phase independent, but is undesirable because it produces peak shifts and distortions of overlapping resonances (Dadok and Sprecher, 1974; Roth *et al.*, 1980).

Data shift accumulation is a software method for reducing the discrete time-domain water signal in memory (Haasnoot and Hilbers, 1983; Benz *et al.*, 1972). The FID is copied, one of the copies is shifted by half of one period of the water sinusoid, and the two copies are added together. Since the water sinusoids in the two data sets are 180° out of phase with each other, any signal at or near the H_2O frequency cancels itself. Once the water signal has been reduced, more FIDs can be accumulated to improve solute spectral signal-to-noise ratio. Data shift accumulation is readily performed by most FT/NMR software packages. Although the water peak cancels itself, signals at other frequencies are phase shifted according to their (discrete) frequency, ν_k, $0 < k \leq N/2$,

$$\varphi_k = \frac{\pi m \nu_k}{\nu_{\text{Nyquist}}} \tag{5}$$

in which m is the number of time-domain points shifted, ν_{Nyquist} is the Nyquist frequency, $\nu_k = (2k/N)\nu_{\text{Nyquist}}$, and N is the number of time-domain data points (Haasnoot and Hilbers, 1983). The linear phase shift of equation (5) is readily corrected in the usual way. Finally, although it is in principle possible to make up for small ADC dynamic range by increasing the number of accumulated FIDs, in practice it is preferable to employ an ADC with the largest possible dynamic range.

5. RELAXATION METHODS

For solutes (e.g., carbohydrates) with resonances at or near the H_2O Larmor frequency, frequency-selective methods are obviously of no help. However, if the solvent and solute relaxation times (T_1 or T_2) are sufficiently different, then it is possible to suppress the H_2O signal so as to leave solute resonances of approximately the same Larmor frequency observable.

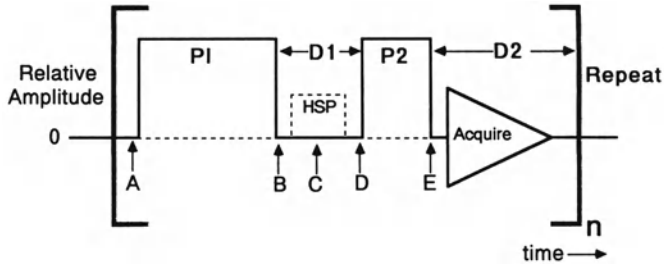


Figure 12. Event sequence for water suppression by means of an inversion-recovery (WEFT) experiment. $P1$ and $P2$ are nonselective $180°_x$ and $90°_x$ pulses, and $D1$ is adjusted such that the solvent magnetization reaches its null point just before the $P2$ pulse. $D2$ is a relaxation delay between successive event sequences, and HSP is an optional homogeneity-spoiling pulse.

5.1. T_1-Based Methods

5.1.1. Inversion/Recovery (WEFT)

If the longitudinal relaxation time of the solvent, T_1(solvent), is significantly longer than that of the solute, T_1(solute), then water magnetization may be suppressed by suitable choice of $D1$ in a simple $180°_x - D1 - 90°_x$ inversion/recovery pulse sequence shown in Figure 12 (Mooberry and Krugh, 1975), which has been renamed the WEFT (*w*ater *e*liminated *f*ourier *t*ransform spectrum) method (Gupta, 1976; Patt and Sykes, 1972; Inubushi and Becker, 1983; Lauterwein and Gerothanassis, 1983). Figure 13 shows the evolution of the solvent and solute magnetization components during the event sequence of Figure 12. The delay period, $D1$, is chosen so that the water magnetization has recovered to its null point (D in Figure 13) just before the (nonselective) $90°_x$ pulse which generates the observed FID, whereas the solute magnetizations have partially recovered at least some

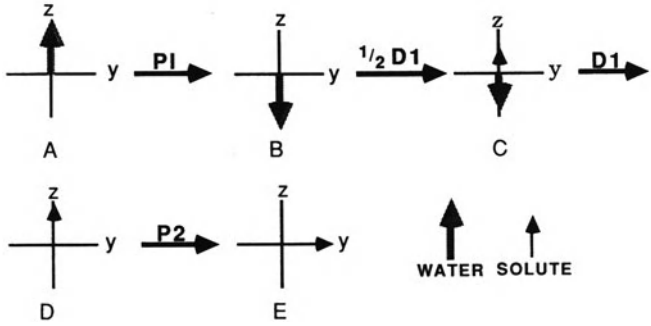


Figure 13. Evolution of the solvent and solute magnetizations at various stages, (A)-(E) of the WEFT event sequence of Figure 12.

of their equilibrium values. An optional homogeneity-spoiling pulse (i.e., a large static-field gradient, HSP in Figure 12) improves the degree of water suppression. WEFT works best for solutes in 99+% D_2O, for which the residual (but still large) HDO magnetization has a relatively long T_1.

In order to ensure near-complete recovery of the solvent magnetization between successive scans, the delay period $D2$ between the $90°_x$ pulse of a given event sequence and the $180°_x$ pulse of the next event sequence should be $\geq 5\ T_1$(solvent), so that $D1_{null} = T_1$(solvent) ln2 seconds for precise nulling of the water signal. However, a delay of $5T_1$ can be 10–15 sec for HDO, so that the experiment takes inordinately long to perform. Thus, it may be preferable to operate in a steady-state mode, in which $D2 < 5T_1$. For a given $D2$ and T_1(solvent), $D1_{null}$ is given by (Mooberry and Krugh, 1975; Krugh and Schaefer, 1975)

$$D1_{null} = T_1\left\{\ln\left[1 - \exp\left(-\frac{T_1 \ln 2 + D2}{T_1}\right)\right] + \ln 2\right\} \tag{6}$$

If the relative frequency-domain peak areas of the solute resonances are to reflect the true relative abundances of those protons, then $(D2 + D1)$ must be at least five times as long as the longest solute $T1$, or $T_1(\text{solute}) \leq 0.2T_1(\text{solvent}) \ln 2$. If equilibrium peak areas are not critical, then it is sufficient that $T_1(\text{solute}) < T_1(\text{solvent})$. Although the inversion/recovery method can successfully detect solute resonances whose frequencies are very near that of H_2O, disadvantages of the method are (a) lengthy data acquisition; (b) the need for careful measurement of the solvent T_1 (which varies from sample to sample) in order to set the best $D1$ value; (c) distortion of spectral peak relative areas by chemical exchange and/or cross-relaxation with solvent protons; and (d) reduction in sensitivity because of incomplete recovery of solute magnetization between scans.

5.1.2. VASE

Among the modified versions of WEFT (Patt and Sykes, 1972; Haasnoot, 1983; Kao and Hruby, 1986) is VASE (*v*ariable *a*ngle *s*ignal *e*limination), illustrated in Figure 14. As in WEFT, $D2$ is a relaxation delay ($\geq 5T_1$). *PS* is a long (soft) frequency-selective fixed-frequency rf pulse which flips the water magnetization by $\theta°$, $90° \leq \theta \leq 180°$. $D1_{null}$ is (as in WEFT) the evolution period required for water to reach its null, $P2$ is a nonselective detection pulse, and HSP is an optional homogeneity-spoiling pulse. Because the solvent magnetization is no longer tipped a full 180°, $D1_{null}$ can be shortened considerably. $D1_{null}$ can be determined from (Patt and Sykes, 1972)

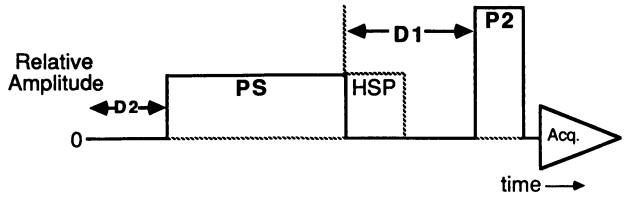


Figure 14. Variable angle signal elimination (VASE) event sequence. *PS* is a frequency-selective $\theta°$ pulse for water, $90° < \theta < 180°$. HSP is a homogeneity spoiling pulse to remove any residual solvent x-y magnetization, and $D1$ is adjusted to produce minimal water signal. $D2$ is a relaxation delay $\geq 5\ T_1$(solvent).

$$\cos\theta = 1 - \exp\left(\frac{t_{\text{null}}}{T_1}\right) \tag{7}$$

The advantages of VASE are that the selective water nulling should distort only those signals close to the water Larmor frequency; all the other signals will be recorded at full sensitivity since they begin each event sequence at their equilibrium magnetization values. Moreover, $D1_{\text{null}}$ can be made short enough to reduce cross-relaxation and saturation transfer effects. A final advantage of VASE is that T_1 measurements can be made on the solute protons.

5.1.3. DEFT

In the *d*riven *e*quilibrium *f*ourier *t*ransform (DEFT) T_1-based method (Figure 15, top) (Hochmann and Kellerhals, 1980; Becker *et al.*, 1969; Shoup and Becker 1972; Rabenstein *et al.*, 1988), the excited transverse solvent magnetization is driven back to its equilibrium z direction (where it is unobservable) at the beginning of the detection period. The evolution of transverse and longitudinal solvent and solute magnetizations during the DEFT event sequence is shown in Figure 15 (bottom). The net effect of the pulse sequence is that the water transverse magnetization created by the first 90° pulse $P1$ is returned back to the z direction after the "detection" pulse $P3$, whereas the water longitudinal magnetization is nulled after the second $D1$ delay period following the 180° pulse, $P2$. On the other hand, the faster-relaxing solute magnetization recovers to the positive z direction in time to be flipped by $P3$ into the x-y plane for detection.

The DEFT method may be further improved by the use of transmitter and receiver phase cycling (Hochmann and Kellerhals, 1980; Becker *et al.*, 1969). The DEFT experiment is normally conducted in steady-state mode (to save time), by adding a relaxation delay, RD $< T_1$(water) between

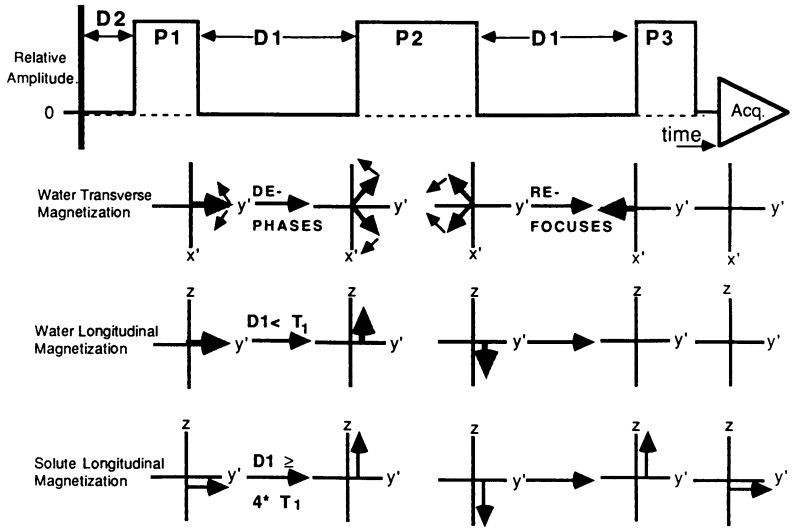


Figure 15. Event sequence (top) and evolution of solute and solvent transverse and longitudinal magnetization components (bottom) at various stages of the driven equilibrium Fourier transform (DEFT) experiment. *P*1 and *P*3 are nonselective 90° pulses and *P*2 is a nonselective 180° pulse. *D*1 is adjusted to produce minimal H_2O signal after the "detection" pulse, *P*3.

successive event sequences. In that case, *D*1 can be computed from specified values of RD, *P*1 tip angle φ, and T_1(water):

$$\cos\varphi = \frac{\left\{1 - \exp\left[-\dfrac{D1}{T_1(H_2O)}\right]\right\}^2}{\exp\left[-\dfrac{2D1}{T_1(H_2O)}\right]\left\{1 - \exp\left[-\dfrac{RD}{T_1(H_2O)}\right]\right\}} \tag{8}$$

DEFT has the advantages (over WEFT) that the specified $T_1(H_2O)$ value is not so critical, and that the experimental parameters (RD, ϕ, and *D*1) can be adjusted to suit the application. Best results have been obtained when *D*1 and RD are 2–4 times as long as the longest solute T_1 (Hochmann and Kellerhals, 1980; Shoup and Becker, 1972). DEFT does not appear to adversely affect the solute resonances near the H_2O Larmor frequency, and the relative solute peak areas approach their equilibrium values if *D*1 is sufficiently long. The method works best when (as for typical solutes in 99+% D_2O), T_1(water) $\gg$ T_1(solute). If the water T_2 value is not sufficiently short to minimize signal arising from water transverse magnetization (Hochmann and Kellerhals, 1980; Becker *et al.*, 1969), HSP may be required.

5.2. T_2-Based Methods

5.2.1. Natural T_2 Differences at High Applied Static Magnetic Field Strength

At very high magnetic field strength ($\geq$10 tesla) the resistive heating of the receiver coil by the large electrical current induced by the precessing, huge H_2O magnetization draws power from the H_2O signal and thereby shortens its T_2 value. However, the much smaller solute magnetization is unaffected. Rabenstein *et al.* (Rabenstein and Isab, 1979; Bryant and Eads, 1985) have exploited this natural difference in T_2 between water and solutes to eliminate the solvent signal in several ways. Most simply, one can delay acquisition by several $T_2(H_2O)$ after a nonselective 90° excitation pulse: the water FID will have largely died away, but the solute signal will have dropped only slightly by the beginning of the acquisition period. Spin-echo techniques are also useful for discriminating against the solvent based on its different T_2 value. The interested reader is referred to the Rabenstein review (Rabenstein and Isab, 1979; Bryant and Eads, 1985) for solvent suppression methods based on inherent differences in T_2 between solvent and solute at high applied static magnetic field strength.

5.2.2. WATR

When the natural differences in T_2 value between solvent and solute is not sufficiently high, a common strategy is to introduce paramagnetic relaxation reagents which complex and exchange preferentially with water and thereby preferentially shorten its T_2 value (Rabenstein and Fan, 1986; Rabenstein *et al.*, 1987; Mirau, 1987; Eads *et al.*, 1986).

This general approach is sometimes denoted as (WATR) *w*ater *a*ttenuation by T_2 *r*elaxation (Eads and Bryant, 1986). The efficacy of a given paramagnetic reagent depends on the rate (pseudo-first order rate constant preferably fast compared to T_2) of chemical exchange of protons between bulk and complexed water molecules. Moreover, selective relaxation enhancement of the solvent requires that the solute have low affinity for the paramagnetic species. Commonly used aqueous relaxation reagents include manganese(II) chloride, square planar copper(II) complexes, and tetraammine copper(II) complexes (Rabenstein *et al.*, 1985; Eads and Bryant, 1986). Other reagents include NH_4Cl and hydroxyl ammonium chloride, the effectiveness of which varies with pH (Rabenstein *et al.*, 1987). In general, the most effective reagents are those which induce a larger chemical shift separation between complexed and uncomplexed water protons (Rabenstein *et al.*, 1987).

Figure 16 shows the usual Meiboom-Gill modified Carr-Purcell spin-echo event sequence designed to remove inhomogeneous broadening of spectral peaks. The sequence begins with a nonselective $90°_x$ pulse $P1_x$, followed by a delay period $D1$ which is short (say, 1 msec) compared to the natural T_2 of solvent or solute. After a refocussing $180°_y$ pulse $P2_y$ and a second $D1$ delay, protons of different chemical shift refocus (except for natural T_2 decay) to give a symmetrical echo FID signal. $D2$ is a relaxation delay to allow the system to recover before the start of the next sequence. The $D1 - P2_y - D1$ sequence may be repeated n times, to generate a series of echoes of decreasing magnitude. Since the relaxation reagent gives H_2O the shorter natural T_2 value, the H_2O signal damps nearly to zero to yield a net suppression of H_2O by a factor of up to ~10,000 for an echo acquired after n = 40-400 repetitions of the $D1 - P2_y - D1$ sequence. Advantages of this sequence are that (a) phase modulation (and thence peak magnitude distortion) by scalar (J) coupling can be reduced by making $D1$ short and thereby spacing the 180° pulses closer together; and (b) phase anomalies associated with long delay times are reduced. A disadvantage of this sequence is that the signal-to-noise ratios of the solute resonances can be reduced considerably if $D1$ and/or n are too large. As an example, for a water T_2 of ~0.06 sec, a delay period of $D1$ = 0.5 sec would reduce the water signal strength by a factor of ~4000, compared to a factor of only ~64 for a solute with T_2 = 1 sec. Finally the spin-echo technique is not suitable for solutes with short spin-spin relaxation times, because the difference in T_2 between water and solute is then no longer large.

5.2.3. SWATTR

Protons on proteins or other large molecules typically have relatively short T_2 values, so that the aforementioned WATR method cannot easily

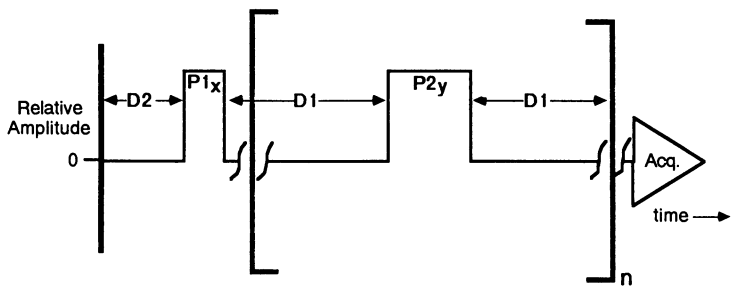


Figure 16. Meiboom-Gill modified Carr-Purcell spin-echo event sequence for discrimination between signals of different "natural" (rather than inhomogeneity-induced) T_2 values. $P1_x$ and $P2_y$ are nonselective 90° and 180° pulses which differ by a 90° phase shift. $D1$ is a refocussing delay and $D2$ is a relaxation delay between successive event sequences.

discriminate between solvent and solute protons from their T_2 values alone. However, a method which exploits both T_1 and T_2 differences between solvent and solute protons is SWATTR (*s*elective *w*ater *a*ttenuation by means of T_2 and T_1 *r*elaxation) (Eisenstadt, 1980; Piveteau *et al.*, 1987), shown in Figure 17. A relaxation reagent (e.g., ammonium chloride) selectively shortens the water T_2 but not the water T_1. The first frequency-selective 90° pulse flips only the water magnetization, which is then rapidly dephased (i.e., rapid T_2 decay of transverse magnetization) by the relaxation reagent during the first evolution period, $D1$. However, the same water protons recover their z magnetization much more slowly because of their long T_1. The next frequency-selective pulse, $P2$, flips any residual water z magnetization by 180°, and since the x-y magnetization component has dephased completely during $D1$, it is not refocused by the 180° pulse. At the end of the second $D1$ delay period, the water magnetization recovers back to its null point and the nonselective 90° pulse samples the remaining solute magnetizations only. The 90° pulse and the delay period before the 180° pulse shorten the time required to reach the solvent null point, compared to a method based on a selective 180° pulse alone. The nonselective observe pulse $P3$ has the advantage over a frequency selective pulse that all of the solute resonances are excited with uniform magnitude. Moreover, solute relative peak heights are not weighted according to their relative T_2 values. Finally, the nonselective pulse does not introduce spectral phase distortion (as most frequency-selective excitations do). Disadvantages of the SWATTR technique are that (a) the degree of solvent suppression is typically smaller than for most of the methods already mentioned; (b) if solute protons exchange with solvent with a pseudo-first order rate constant faster than

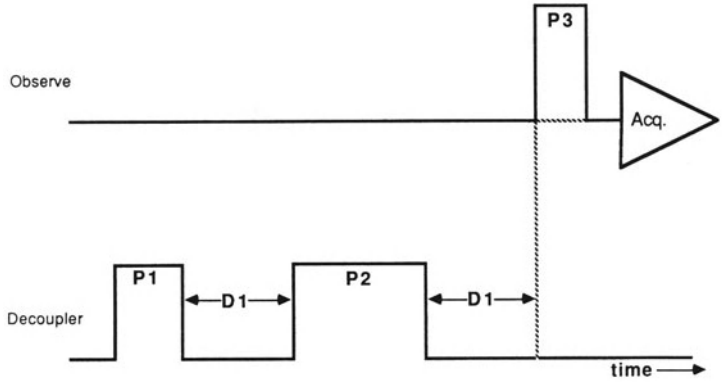


Figure 17. Event sequence for selective water attenuation by means of T_2 and T_1 relaxation (SWATTR). The decoupler irradiation $P1$ (90°) and $P2$ (180°) is frequency-selective for the water resonance, and the observe pulse, $P3$ (90°) is nonselective.

$1/T_1$, their apparent peak heights will be distorted from the equilibrium values; (c) resonances near the water peak can also be distorted. The SWATTR method is especially well suited for two-dimensional NMR experiments (in which uniform excitation is important) and for experiments for which phase distortion is otherwise a serious problem.

5.3. $T_{1\rho}$-Based (Spin-Lock)

Yet another relaxation-based technique employs the spin-locking technique illustrated in Figure 18. An initial $90°_x$ nonselective pulse flips all magnetization to the $+y$ axis in the rotating frame, after which all of the magnetization components evolve at their Larmor frequencies during $D1$. A frequency-selective pulse LP (fixed-frequency rf at the H_2O frequency and phase-shifted by 90° compared to $P1$) then creates an effective field along the $+y$ axis in the rotating frame, and the water magnetization remains locked along that effective field direction. The water magnetization relaxes with a characteristic relaxation time constant $T_{1\rho}$ which is generally longer than T_2: $T_2 < T_{1\rho} < T_1$. However, other (solute) magnetization components relax rapidly (in both transverse and longitudinal directions) during LP, as they try to precess about a rotating-frame effective field which is off resonance and therefore rotating about z. Finally, a third (nonselective) pulse $P2_{-x}$ rotates the H_2O magnetization from the $+y$ back to the $+z$ direction, where it is unobserved, and also flips the nearly recovered solute magnetizations from the $+z$ direction onto the x-y plane for detection (Blondet *et al.*, 1986; Bax *et al.*, 1987). In practice, transmitter and receiver phase-cycling is needed in order to correct for instrumental imperfections and other effects (Canet *et al.*, 1988).

A variation is to spin-lock all of the magnetization components except water, so that the water rapidly dephases and is not detected (see Figure 19) (Canet *et al.*, 1988); Starcuk *et al.*, 1988). In this case, an initial nonselective $90°_x$ pulse flips all magnetization components to the $+y$ axis, where they evolve during period $D1$. During LP_x, all but the water magnetization components are spin-locked along the x axis, except for the water

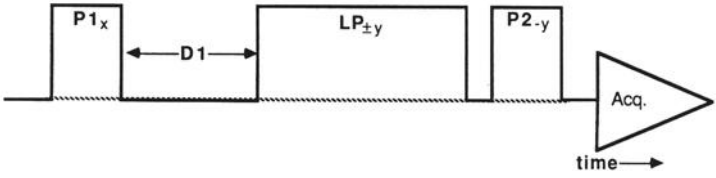


Figure 18. Event sequence for an experiment in which only the water magnetization is spin-locked. $P1_x$ and $P2_{-x}$ are nonselective 90° pulses, and LP_y is a spin-lock pulse frequency-selective for H_2O.

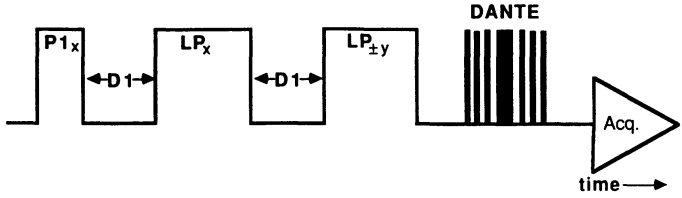


Figure 19. Event sequence for an experiment in which all but the water magnetization is spin-locked. $P1_x$ is a nonselective 90° pulse, LP_x and LP_y are frequency-selective (for all frequencies except H_2O) spin-lock pulses, and the final excitation (here shown as DANTE) is a frequency-selective 180° pulse with zero magnitude at the H_2O Larmor frequency.

resonance which remains along $+y$ axis. The water signal rapidly dephases under the influence of LP_x. At this stage, the water magnetization has canceled to first order. To achieve second-order cancellation, the system is allowed to evolve for a second $D1$ period and is then locked along the y-axis (Canet *et al.*, 1988; Starcuk *et al.*, 1988). Finally, a selective refocussing pulse (e.g., DANTE, with zero excitation magnitude at the H_2O Larmor frequency) generates a detectable FID for the solute magnetization components (Canet *et al.*, 1988). Yet another method combines a spin-locking pulse with a frequency-selective binomial hard pulse sequence in order to eliminate phase shifts resulting from the hard pulse sequence (Brooks *et al.*, 1984).

Spin-locking methods can offer a high degree of water suppression (comparable to the bionomial sequences), and they achieve it before the signal reaches the analog-to-digital converter. Moreover, the method is not especially sensitive to the choice of experimental parameters (Blondet *et al.*, 1986). However, the method of Figure 18 requires a rather complicated phase cycling to achieve adequate suppression. The method of Figure 19 suffers from reduced signal-to-noise ratio since the observed magnetization is partly relaxed at the beginning of the detection period. Both methods tend to introduce phase anomalies into the spectrum, and the corrective measures generally reduce the degree of solvent suppression and/or signal-to-noise ratio for solute signals (Blondet *et al.*, 1986). On the positive side, the spin-locking idea can be combined with some other suppression methods.

6. POLARIZATION TRANSFER METHODS

6.1. Heteronuclear Polarization Transfer

Perturbation of the equilibrium (Boltzmann) populations of one nucleus can change the relative populations of the energy levels of another

nucleus which is J-coupled to the first. Since observed macroscopic magnetization (corresponding to a "polarization" of the average directions of the microscopic magnetic moments) is directly related to energy level population difference, it is common to say that polarization has been transferred from one type of nucleus to another in such an experiment. Heteronuclear transfer of polarization from nuclei of high magnetogyric ratio (e.g., ^{1}H) to J-coupled nuclei of low magnetogyric ratio (e.g., ^{13}C) was originally introduced as a means for enhancement of signal-to-noise ratio for nuclei of low magnetogyric ratio.

In the "reverse" heteronuclear polarization transfer experiment, polarization is transferred from the low-γ spins (^{13}C) to high-γ spins (^{1}H). In this way, the observed signal originates only from protons which are J-coupled (i.e., connected through chemical bonds) with a particular J-value to ^{13}C atoms, and there can thus be no signal at all from H_2O. Although the signal-to-noise ratio for the solute protons detected via reverse polarization transfer is much lower than that for direct ^{1}H detection because of the hundredfold lower natural abundance and fourfold lower magnetogyric ratio of ^{13}C compared to ^{1}H, the complete elimination of the solvent signal may nevertheless make such an experiment attractive (Bulsing and Doddrell, 1986; Esposito *et al.*, 1988; Bornais *et al.*, 1983; Roth, 1980; Hardy and Dumoulin, 1983).

The (reverse) polarization transfer may be achieved by the application of any of several possible rf pulse sequences. For example, the event sequence for a reverse INEPT (*i*nsensitive *n*uclei *e*nhanced by *p*olarization *t*ransfer) experiment (Esposito *et al.*, 1988) is shown in Figure 20. The experiment can be improved by zeroing the proton magnetization at the beginning of the sequence by use of a spherically randomizing rf pulse, which also acts to enhance the ^{13}C polarization from the nuclear Overhauser

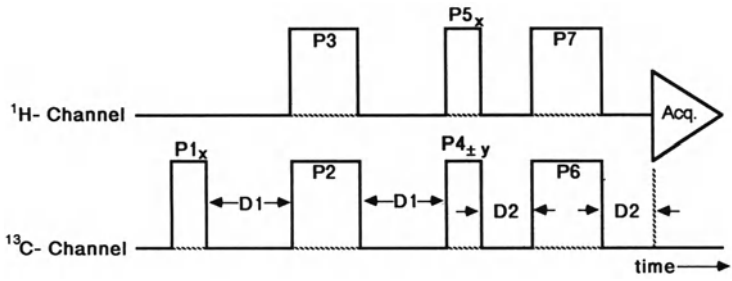


Figure 20. Event sequence for a heteronuclear "reverse" INEPT experiment. $P1$ and $P4$ are 90° pulses and $P2$ and $P6$ are 180° pulses for ^{13}C; $P7$ is a 180° and $P3$ and $P5$ are 90° ^{1}H pulses. Delays $D1 = J/2$ and $D2 = J/4$, where J is the scalar coupling constant (in Hz) between ^{1}H and ^{13}C. This experiment is tuned to detect only those ^{1}H which are J-coupled ($J = 2D1$) to ^{13}C atoms.

effect (NOE) (Esposito *et al.*, 1988). Because the basic INEPT sequence requires near-perfect 180° pulses, some sort of phase cycling is mandatory: for example, Bulsing and Doddrell (1986) used a 64-scan phase cycling algorithm. Solvent suppression by a factor of up to 800,000 has been claimed for the reverse INEPT method.

The INEPT experiment is inherently tuned to J-coupling of a particular J-value, and the resulting solute proton NMR peak areas can deviate significantly from the relative abundances of the various chemically different protons. Another method less sensitive to the J-value, known as reverse DEPT (*d*istortionless *e*nhancement by *p*olarization *t*ransfer) (Esposito *et al.*, 1988), is shown schematically in Figure 21. The evolution of the ^{13}C and ^{13}C-coupled and uncoupled (e.g., H_2O) ^1H magnetizations at various stages of the experiment are shown in Figure 22. The purpose of the initial randomizing field pulse has already been explained, and the final $P6$ "purging" pulse is designed to remove any water magnetization remaining at that point. The reverse DEPT experiment offers a high degree (factor of 100,000–800,000) of water suppression (Esposito *et al.*, 1988). Since the detected signal arises from heteronuclear scalar ^{13}C–^1H couplings, the experiment can also be used to edit the spectrum. For instance, by adjusting the length of pulse $P6$ in the DEPT sequence one can control whether the CH, CH_2, and CH_3 protons have positive or negative absorption-mode spectral amplitude. Moreover, spectral resolution is potentially enhanced over direct ^1H detection because the spectral line width is controlled by the longer ^{13}C relaxation times. Disadvantages of the reverse polarization transfer experiments are that (a) complex multiple-pulse sequences require rf channels of different frequency; (b) signal-to-noise ratios are poor, since the sensitivity is determined by ^{13}C rather than ^1H; (c) solvent-exchangeable protons (e.g., those attached to nitrogen or oxygen) are not observable unless a separate polarization transfer (e.g., ^{15}N–^1H or ^{17}O–^1H) is carried out for each type of exchangeable proton present; (d) relative apparent abundances of various solute protons are distorted.

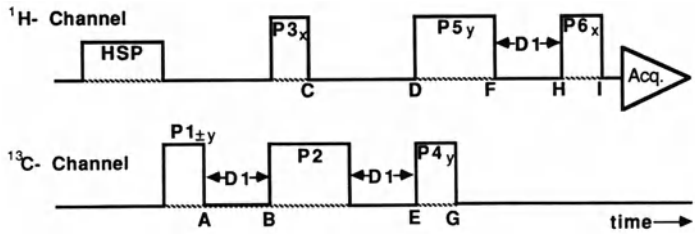


Figure 21. Event sequence for a reverse DEPT experiment. $P1 = P4 = 90°$ for ^{13}C; $P2 = 180°$ for ^{13}C; $P3 = 90°$ for ^1H; $P5 = 180°$ for ^1H; and $P6$ is a "purging" pulse that does not affect the buildup of polarization transfer signal formed on the $+x$ axis. Delay $D1 = J/2$, where J is the scalar ^{13}C–^1H coupling constant (in Hz).

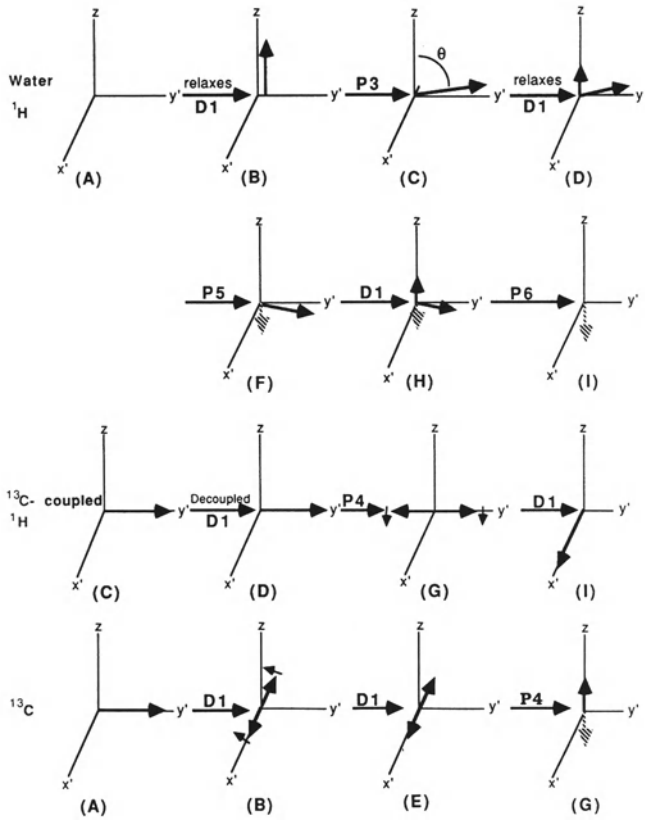


Figure 22. Evolution of magnetizations of protons that are not J-coupled (top) or are J-coupled (middle) to ^{13}C nuclei at various indicated stages of the reverse DEPT event sequence of Figure 21. At (E), the 90° ^{13}C pulse also polarizes the J-coupled ^{1}H magnetization which is refocused by the 180° ^{1}H pulse at (D).

6.2. Homonuclear Polarization Transfer

Polarization transfer may also be based on protons which are J-coupled to other protons rather than to ^{13}C atoms, in a method known as POTSHOT (*po*larization *t*ransfer by a *s*elective *ho*monuclear *t*echnique) (Dumoulin, 1985). The event sequence, involving only a single-frequency transmitter, is shown in Figure 23. The delay $D1$ is set approximately to $1/4J$, where J is the specified ^{1}H–^{1}H coupling constant. At the end of the second pass through this sequence, the detected proton signal, with $J \approx D1$, will be phase-shifted by 180° relative to the signal from the first pass. The signals from J-coupled protons will thus add when the second pass is subtracted

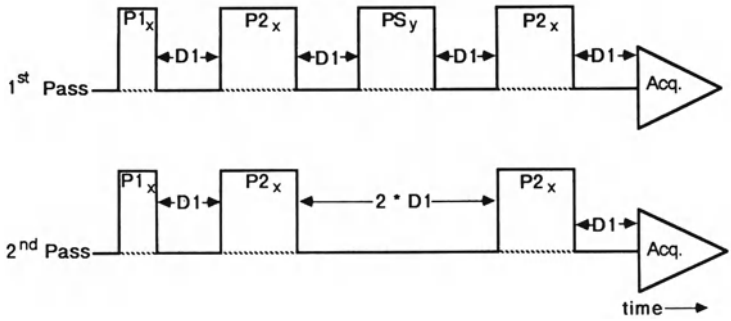


Figure 23. Event sequence for the POTSHOT experiment, based on detection of only those protons which are J-coupled to other protons. $P1$ and $P2$ are nonselective 90° and 180° ^1H pulses, and PS is a 180° pulse which is frequency-selective for a given transition of the J-coupled species. $D1 = 1/4J$, where J is a specified ^1H–^1H coupling constant (in Hz).

from the first. Because the uncoupled (e.g., water) protons do not experience the 180° phase shift, their signals will cancel during the same subtraction. This method has the advantages that only one transmitter is needed and that (as for reverse DEPT) the spectrum can be edited ("tuned") for specific J-values. Disadvantages are that instrumental imperfections and off-resonance effects tend to make the degree of suppression quite small (~40) compared to most other techniques discussed in this review. Also, the method of Figure 23 does nothing to reduce the magnitude of the water signal before it reaches the analog-to-digital converter. Thus, this method would normally be combined with another method (e.g., presaturation of T_2 based) (Dumoulin, 1986) which reduces the water signal strength before the ADC stage, in order to boost the water suppression to a suitable level.

7. MULTIPLE-QUANTUM FILTERING

In a multiple-quantum filtering experiment (Dumoulin and Vatis, 1986), all single-quantum signals as in Figure 24(a) cancel out on alternate scans, whereas zero- and double-quantum signals as in Figure 24(b) add on alternate scans. Water protons exhibit only single-quantum transitions, because the water protons are not J-coupled to protons of different Larmor frequency. Although only single-quantum coherences can be experimentally detected, it is possible to devise event sequences which create zero-quantum, as in Figure 25(b), or two-quantum, as in Figure 25(a) coherences which are then detected from their modulation of an (observable) single-quantum coherence. In both sequences of Figure 25, pulse $P2$ eliminates chemical shift variations, and pulse $P4$ allows for the detection of echoes in which all multiplets have been allowed to refocus (thereby simplifying the final

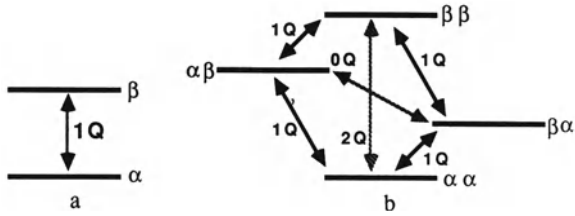


Figure 24. Magnetic energy level diagrams for uncoupled spins $\frac{1}{2}$ (a) and a system of two J-coupled spins $\frac{1}{2}$ (b). Zero-, single-, and double-quantum transitions are denoted as $0Q$, $1Q$ and $2Q$, respectively.

spectrum). The HSP removes both single and double-quantum signals. Single-quantum transitions are cancelled out by proper phase cycling of the transmitter and receiver channels (Dumoulin and Vatis, 1986). For water protons, either pulse sequence simply acts as a 180° pulse. Thus if the relative phase of the receiver and transmitter is changed by 180° from one acquisition to the next, and the second acquisition is subtracted from the first, the water signal will cancel to zero. For J-coupled nuclei, on the other hand, the second acquisition signal will be inverted from that of the first, so that the same subtraction step has the effect of doubling the signal. In practice, hardware instabilities and rf inhomogeneity limit the efficacy of a simple two-step phase cycling sequence, and a rather more complicated phase cycling is required (Dumoulin and Vatis, 1986).

Advantages of the multiple-quantum filtering technique are that (a) spectra are simplified, since there are fewer possible double-quantum than single-quantum transitions; (b) broadband excitation allows signals very

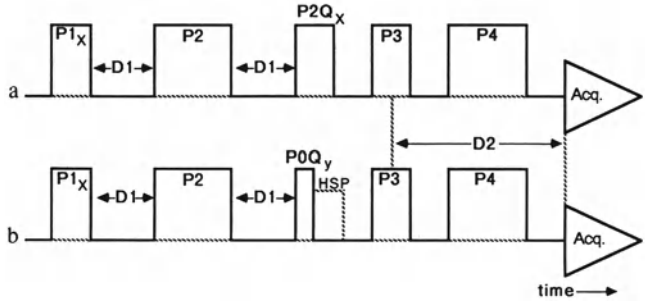


Figure 25. Event sequence for multiple quantum filtering. $P1$ is a nonselective 90° pulse, and $P2$ (nonselective 180° pulse) eliminates chemical shift variations. $P4$ (nonselective 180°) refocuses J-coupled multiplet signals. Delays $D1 = 1/4J$ and $D2 = 1/2J$ allow time for J-coupled multiplet components to rephase. In sequence (a), $P2Q$ (nonselective 90°) generates only double quantum coherences. In sequence (b), $P0Q$ (nonselective 45°) generates both zero and double quantum coherence. A homogeneity-spoiling pulse (HSP) removes all single and double quantum coherences.

close to the H_2O resonance to be observed; and (c) the degree of suppression is independent of T_1 and T_2. Disadvantages are that (a) the phase-cycled pulse sequence is complex and nontrivial to implement; (b) the observed spectral relative peak areas depend on the J-coupling, so that quantitation is difficult (Kao and Hruby, 1986); and (c) the method fails to reduce the water signal before it reaches the analog-to-digital converter, so that either the ADC must have very high dynamic range or the method must be combined with another method which does reduce the H_2O signal strength before the ADC stage.

8. EXPERIMENTAL COMPARISONS

In this section, we include experimental comparisons of some of the more commonly used water-suppression pulse sequences. All spectra were acquired from a Bruker AM-500 FT/NMR spectrometer operating at 11.75 tesla and equipped with an ASPECT-3000 computer. In each case, the sample was ~0.5 mM transfer RNA^{phe} from brewer's yeast, dissolved in 95%/5% H_2O/D_2O, in which D_2O was added to provide a field-frequency lock signal. In all plotted spectra, the time-domain data was processed so as to convert from Lorentzian to Gaussian spectral line shape (to reduce spectral peak overlap) with a Gaussian width of 0.1 Hz.

For this sample, the H_2O and RNA proton T_1 (and T_2) relaxation times were similar in magnitude, so that neither T_1- nor T_2-selective solvent suppression techniques are suitable. Because the base-pair protons exchange with H_2O, presaturation methods do not apply. J-coupling techniques are also not useful because the widths of the spectral lines of interest are at least as broad as their J-couplings. Furthermore, since the resonances of principal interest (namely, the base-pair hydrogen-bond imine proton resonances at 9–15 ppm downfield from DSS) are far from the H_2O Larmor frequency (~4.8 ppm), frequency-selective methods (both hard-pulse and soft-pulse) are suitable. The degree of suppression of each technique was determined according to the receiver gain level at which the time-domain FID signal was just clipped by the ADC. The total time required to obtain a given signal-to-noise ratio for the proton resonances of interest was also compared for different suppression methods.

8.1. Redfield 2$\underline{1}$4$\underline{1}$2 Soft-Pulse Sequence

The Redfield 214 sequence was the most time-consuming to set up, because of the strong variation in suppression with pulse durations. With the rf carrier placed 12,793.3 Hz from H_2O, the best suppression was observed for $P1 = 23.92$ μsec, $P2 = 46.30$ μsec, and $P4 = 88.22$ μsec. The

degree of water suppression can be judged from the increase in receiver gain compared to single-pulse excitation (20 db vs. 1 db, or a factor of ~10) at which the time-domain FID is just clipped by the ADC. 1415 FID's had to be accumulated to produce the spectrum the 8–15 ppm region of which is shown in Figure 26. Note that the severely curved baseline makes it difficult to phase the spectral peaks and determine their relative intensities.

8.2. 1$\underline{1}$ Hard-Pulse Sequence

A 1$\underline{1}$ hard-pulse sequence with the rf carrier centered at H_2O was easy to implement, with $P1 = 26.93$ μsec. The degree of water suppression was a factor of ~100 (receiver gain of ~40 db vs. 1 db for one-pulse excitation). 512 FID's were added to produce the spectrum shown in Figure 27. The 1$\underline{1}$ spectrum was considerably easier to phase than the 1$\underline{3}$3$\underline{1}$ spectrum (see Section 8.3), but the 1$\underline{1}$ method was more sensitive to the placement of the carrier frequency. Moreover, the maximum obtainable signal-to-noise ratio was lower for the 1$\underline{1}$ than for the 1$\underline{3}$3$\underline{1}$ sequence, because the memory filled up faster for the 1$\underline{1}$ method (due to its lower degree of water suppression than that of the 1$\underline{3}$3$\underline{1}$).

8.3. 1$\underline{3}$3$\underline{1}$ Hard-Pulse Sequence

The frequency-selective 1$\underline{3}$3$\underline{1}$ hard-pulse sequence, with the rf carrier centered at H_2O, was easy to implement (~10 min set-up time) and optimize.

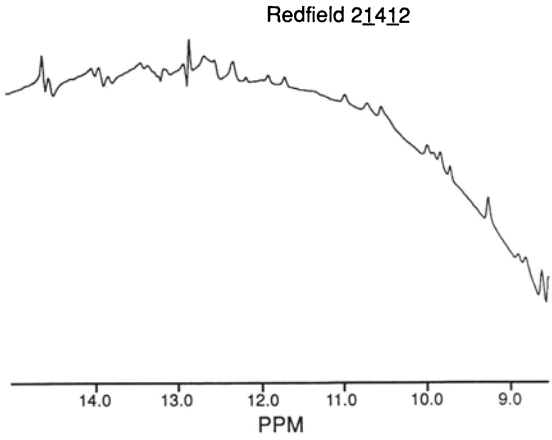


Figure 26. 9–15 ppm region of a 500 MHz ^{1}H FT/NMR spectrum of ~0.5 mM tRNA^{phe} in 95%/5% H_2O/D_2O, produced by a Redfield 314 soft-pulse sequence. Note the severe baseline curvature.

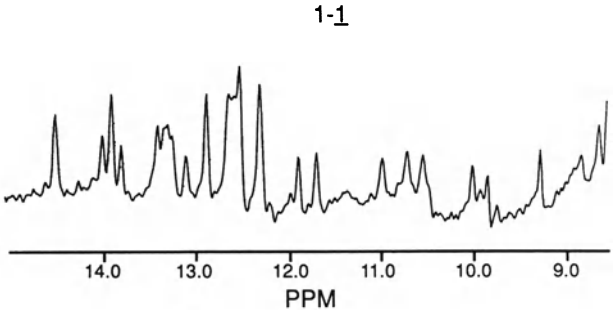


Figure 27. 9–15 ppm region of a 500 MHz ^{1}H FT/NMR spectrum of the same sample as in Figure 26, but produced by a hard-pulse 11 sequence.

The rf power was such that the (hard) pulse width for a 90° flip angle was 53.86 μsec. Thus, rf pulses $P1$ and $P3$ were set to 6.73 and 20.19 μsec, and delay $D1$ was set to 97.9 μsec. This method gave the best suppression (relative receiver gain 64 db vs. 1 db, or a factor of ~1,500). Only 128 FID's (i.e., factor of >50 shorter total acquisition period than for the Redfield 214 sequence) were needed to produce the spectrum shown in Figure 28. A plot of the entire spectrum (Figure 29) graphically shows the great reduction in the magnitude of the H_2O peak (which would otherwise be 4×10^5 times as large as the solute resonances), and also shows the minimal phase distortion in the baseline. (Only the 8–15 ppm region was phased, because that was the region of interest.) We attempted to reduce the phase variation by use of the pulse sequence of Figure 5, but then the degree of water suppression was no longer high enough to be useful. Overall, the best peformance for the $tRNA^{phe}$ sample was obtained with the 1331 hard-pulse sequence.

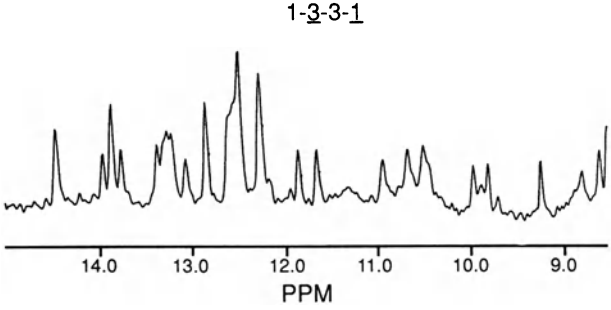


Figure 28. 9–15 ppm region of a 500 MHz ^{1}H FT/NMR spectrum of the same sample as in Figure 26, but produced by a hard-pulse 1331 sequence.

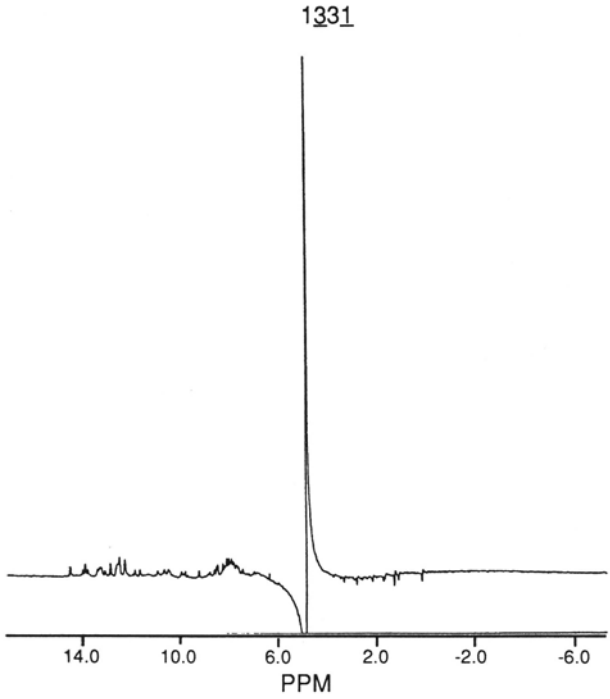


Figure 29. Plot of the full-range spectrum corresponding to Figure 28. Note the great reduction in the 4×10^5-fold more concentrated water proton signal.

ACKNOWLEDGMENTS. The authors wish to thank J. Wu for providing the $tRNA^{phe}$ sample and J. Wu and C. E. Cottrell for helpful discussions. This work was supported by grants (to A.G.M.) from the National Institutes of Health (GM-29274) and The Ohio State University.

REFERENCES

Arata, Y., and Ozawa, H., 1976, *J. Mag. Res.* **21**:67.

Bax, A., Sklenar, V., Clore, G. M., and Gronenborn, A. M., 1987, *J. Am. Chem. Soc.* **109**:6511.

Becker, E. D., Ferretti, J. A., and Farrar, T. C., 1969, *J. Am. Chem. Soc.* **91**:7784.

Benz, F. W., Feeney, J., and Roberts, G. C., 1972, *J. Mag. Res.* **8**:114.

Bleich, H. E., and Glasel, J. A., 1975, *J. Mag. Res.* **18**:401.

Bleich, H., and Wilde, J., 1984, *J. Mag. Res.* **56**:154.

Blondet, P., Decorps, M., Albrand, J. P., Benabid, A. C., and Remy, C., 1986, *J. Mag. Res.* **69**:403.

Bornais, J., Brownstein, S., and Bywater, S., 1983, *J. Mag. Res.* **52**:120.

Brandes, R., and Kearns, D. R., 1985, *J. Mag. Res.* **64**:506.

Brereton, I. M., Galloway, G. J., Field, J., Marshman, M. F., and Doddrell, D. M., 1989, *J. Mag. Res.* **81**:411.

Brindle, K. M., and Campbell, I. D., 1987, *Rev. Biophys.* **19**:159.

Brooks, W. M., Irving, M. G., Simpson, S. J., and Doddrell, D. M., 1984, *J. Mag. Res.* **56**:521.
Bryant, R. G., and Eads, T. M., 1985, *J. Mag. Res.* **64**:312.
Bulsing, J. M., and Doddrell, D. M., 1986, *J. Mag. Res.* **68**:52.
Campbell-Burke, S. L., and Shulman, R. G., 1987, *Annual Rev. Microbiol.* **41**:595.
Canet, D., Boudot, D., and Brondeau, J., 1988, *J. Mag. Res.* **79**:377.
Chen, L., and Marshall, A. G., 1987a, *Raoud Commun. Mass Spectrom.* **1**:39.
Chen, L., and Marshall, A. G., 1987b, *Int. J. Mass Spectrom. Ion Proc.* **79**:115.
Clore, G. M., Kimber, B. J., and Gronenborn, A. M., 1983, *J. Mag. Res.* **54**:170.
Cooper, J. W., 1976, *J. Mag. Res.* **22**:345.
Cooper, J. W., Machay, I. S., and Pawle, G. B., 1977, *J. Mag. Res.* **28**:405.
Cutnell, J. D., Dallas, J., and Lamar, G. N., 1980, *J. Mag. Res.* **41**:213.
Dadok, J., and Sprecher, R. F., 1974, *J. Mag. Res.* **13**:243.
Dumoulin, C. L., 1985, *J. Mag. Res.* **64**:38.
Dumoulin, C. L., 1986, *Mag. Res. Med.* **3**:90.
Dumoulin, C. L., and Vatis, D. D., 1986, *Mag. Reson. Med.* **3**:282.
Eads, T. M., and Bryant, R. G., 1986, *J. Agric. Food Chem.* **34**(5):834.
Eads, T. M., Kennedy, S. D., and Bryant, R. G., 1986, *Anal. Chem.* **58**:1752.
Eisenstadt, M., 1980, *J. Mag. Res.* **39**:263.
Ernst, R. R., and Anderson, W. A., 1966, *Rev. Sci. Instrum.* **37**:93.
Esposito, G., Gibbons, W. A., and Bazzo, R., 1988, *J. Mag. Res.* **80**:523.
Galloway, G. J., Haseler, L. J., Marshman, M. F., Williams, D. H., and Doddrell, D. M., 1987, *J. Mag. Res.* **74**:184.
Guittet, E., Delsuc, M. A., and Lallemand, J. Y. 1984, *J. Am. Chem. Soc.* **106**:4278.
Gupta, R. K., 1976, *J. Mag. Res.* **24**:461.
Gutow, J., McCoy, M., Sparo, F., and Warren, W. S., 1985, *Phys. Rev. Lett.* **55**:1090.
Haasnoot, C. A., 1983, *J. Mag. Res.* **52**:153.
Haasnoot, C. A., and Hilbers, C. W., 1983, *Biopolymers* **22**(5):1259.
Hall, L. D., Sukumar, S., and Talagala, S. L., 1984, *J. Mag. Res.* **56**:275.
Hall, M. P., and Hore, P. J., 1986, *J. Mag. Res.* **70**:350.
Hardy, C. J., and Dumoulin, C. L., 1983, *Mag. Reson. Med.* **5**:58.
Hetherington, H. P., Avison, M. J., and Shulman, R. G., 1985, *Biochem.* **82**:3115.
Hochmann, J., and Kellerhals, H., 1980, *J. Mag. Res.* **38**:23.
Hore, P. J., 1983a, *J. Mag. Res.* **55**:283.
Hore, P. J., 1983b *J. Mag. Res.* **54**:539.
Hore, P. J., 1984, *J. Mag. Res.* **56**:535.
Hoult, D. I., 1976, *J. Mag. Res.* **21**:337.
Hsu, A. T., Hunter, W. W., Schmalbrock, P., and Marshall, A. G., 1987, *J. Mag. Res.* **72**:75.
Inubishi, T., and Becker, E. D., 1983, *J. Mag. Res.* **51**:128.
Kao, L. F., and Hruby, V. J., 1986, *J. Mag. Res.* **70**:394.
Kienlin, M. V., Decorps, M., Albrand, J. P., Foray, M. F., and Blondet, P., 1988, *J. Mag. Res.* **76**:169.
Krishna, N. R., 1976, *J. Mag. Res.* **22**:555.
Krugh, T. R., and Schaefer, W. C., 1975, *J. Mag. Res.* **19**:99.
Kuchel, P. W., Chapman, B. E., Endre, Z. H., King, G. F., Thorburn, D. R., and York, M. J., 1984, *J. Biomed. Biochem. Acta* **43**:719.
Lauterwein, J., and Gerothanassis, I. P., 1983, *J. Mag. Res.* **51**:153.
Levitt, M. H., and Roberts, M. F., 1987, *J. Mag. Res.* **71**:576.
Loaiza, F., McCoy, M. A., Hammes, S. L., and Warren, W. S., 1988, *J. Mag. Res.* **77**:175.
Lurie, D. J., 1985, *Magn. Reson. Imaging* **3**:239.
Marshall, A. G., 1985, *Acc. Chem. Res.* **18**:316.
Marshall, A. G., and Verdun, F. R., 1989, *Fourier Transforms in Optical, NMR, and Mass Spectrometry: A User's Handbook*," Elsevier, Amsterdam.

Marshall, A. G., Marcus, T., and Sallos, J., 1979, *J. Mag. Res.* **35**:227.
McCoy, M. A., and Warren, W. S., 1987, *Chem. Phys. Lett.* **133**:165.
McCoy, M. A., Loaiza, F., Valentine, K., and Warren, W. S., 1988, *J. Mag. Res.* **80**:155.
Mirau, P. A., 1987, *J. Mag. Res.* **73**:123.
Mooberry, E. S., and Krugh, T. R., 1975, *J. Mag. Res.* **17**:128.
Morris, G. A., Smith, K. I., and Waterton, J. C., 1986, *J. Mag. Res.* **68**:526.
Patt, S. L., and Sykes, B. D., 1972, *J. Chem. Phys.* **56**(6):3182.
Piveteau, D., Delsuc, M. A., Guittet, E., Lallemand, J. Y., 1987, *J. Mag. Res.* **71**:347.
Plateau, P., and Gueron, M., 1982, *J. Am. Chem. Soc.* **104**:7310.
Plateau, P., Dumas, C., and Gueron, M., 1983, *J. Mag. Res.* **54**:46.
Rabenstein, D. L., 1984, *J. Biophys. Biochem. Method.* **9**:277.
Rabenstein, D. L., and Fan, S., 1986, *Anal. Chem.* **58**:3178.
Rabenstein, D. L., and Isab, A. A., 1979, *J. Mag. Res.* **36**:281.
Rabenstein, D. L., Fan, S., Nakashima, T. T., 1985, *J. Mag. Res.* **64**:541.
Rabenstein, D. L., Srivatsa, G. S., and Lee, R. W., 1987, *J. Mag. Res.* **71**:175.
Rabenstein, D. L., Mills, K. K., and Strauss, E. J., 1988, *Anal. Chem.* **60**:1380A.
Radda, G. K., and Seeley, P. J., 1979, *J. Annual Rev. Physiol.* **41**:749.
Redfield, A. G., and Gupta, R. K., 1971, *J. Chem. Phys.* **54**:1418.
Redfield, A. G., Kunz, S. D., and Ralph, E. K., 1975, *J. Mag. Res.* **19**:114.
Roberts, J. K. M., and Jandetzky, O., 1981, *Biochim. Biophys. Acta* **639**:53.
Roth, K., 1980, *J. Mag. Res.* **40**:489.
Roth, K., Kimber, B. J., and Feeney, J., 1980, *J. Mag. Res.* **41**:302.
Santos, H., Turner, D. L., and Xavier, A. V., 1984, *J. Mag. Res.* **58**:344.
Scott, A. I., and Baxter, R. L., 1981, *Annual Res. Biophys. Bioeng.* **10**:151.
Shoup, R. R., and Becker, E. D., 1972, *J. Mag. Res.* **8**:298.
Shulman, R. G., Brown, T. R., Ugurbil, K., Ogawa, S., Cohen, S. M., and den Hollander, J. A., 1979, *Science* **205**:160.
Sklenar, V., and Bax, A., 1987, *J. Mag. Res.* **74**:469.
Sklenar, V., and Starcuk, Z., 1982, *J. Mag. Res.* **50**:495.
Starcuk, Z., and Sklenar, V., 1985, *J. Mag. Res.* **61**:567.
Starcuk, Z., and Sklenar, V., 1986, *J. Mag. Res.* **66**:391.
Starcuk, Z., Pcek, L., Fiala, R., and Starcuk Jr., Z., 1988, *J. Mag. Res.* **80**:344.
Tomlinson, B. L., and Hill, H. D., 1973, *J. Chem. Phys.* **59**(4):1775.
Tong, J. P., and Kotovych, G., 1986, *J. Mag. Res.* **69**:511.
Turner, D. L., 1983, *J. Mag. Res.* **54**:146.
Wang, C., and Pardi, A., 1987, *J. Mag. Res.* **71**:154.
Warren, S. W., 1988, *Science* **242**:878.
Warren, S. W., and Silver, M. S., 1988, *Advan. Mag. Res.* **12**:247.
Wider, G., Hosur, R. V., and Wuthrich, K., 1983, *J. Mag. Res.* **52**:130.
Wüthrich, K., 1976, *NMR in Biological Research: Peptides and Proteins*, American Elsevier, New York.
Wüthrich, K., 1986, *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
Yao, C., Simplaceanu, V., Lin, A., and Ho, C., 1986, *J. Mag. Res.* **66**:43.

Contents of Previous Volumes

VOLUME 1

Chapter 1

NMR of Sodium-23 and Potassium-39 in Biological Systems
Mortimer M. Civan and Mordechai Shporer

Chapter 2

High-Resolution NMR Studies of Histones
C. Crane-Robinson

Chapter 3

PMR Studies of Secondary and Tertiary Structure of Transfer RNA in Solution
Philip H. Bolton and David R. Kearns

Chapter 4

Fluorine Magnetic Resonance in Biochemistry
J. T. Gerig

Chapter 5

ESR of Free Radicals in Enzymatic Systems
Dale E. Edmondson

Chapter 6

Paramagnetic Intermediates in Photosynthetic Systems
Joseph T. Warden

Chapter 7

ESR of Copper in Biological Systems
John F. Boas, John R. Pilbrow, and Thomas D. Smith

Index

VOLUME 2

Chapter 1

Phosphorus NMR of Cells, Tissues, and Organelles
Donald P. Hollis

Chapter 2

EPR of Molybdenum-Containing Enzymes
Robert C. Bray

Chapter 3

ESR of Iron Proteins
Thomas D. Smith and John R. Pilbrow

Chapter 4

Stable Imidazoline Nitroxides
Leonid B. Volodarsky, Igor A. Grigor'ev, and Renad Z. Sagdeev

Chapter 5

The Multinuclear NMR Approach to Peptides: Structures, Conformation, and Dynamics
Roxanne Deslauriers and Ian C. P. Smith

Index

VOLUME 3

Chapter 1

Multiple Irradiation ^{1}H NMR Experiments with Hemoproteins
Regula M. Keller and Kurt Wüthrich

Chapter 2

Vanadyl(IV) EPR Spin Probes: Inorganic and Biochemical Aspects
N. Dennis Chasteen

Chapter 3

ESR Studies of Calcium- and Proton-Induced Photon Separations in Phosphatidylserine-Phosphatidylcholine Mixed Membranes
Shun-ichi Ohnishi and Satoru Tokutomi

Chapter 4

EPR Crystallography of Metalloproteins and Spin-Labeled Enzymes
James C. W. Chien and L. Charles Dickinson

Chapter 5

Electron Spin Echo Spectroscopy and the Study of Metalloproteins
W. B. Mims and J. Peisach

Index

VOLUME 4

Chapter 1

Spin Labeling in Disease
D. Allan Butterfield

Chapter 2

Principles and Applications of ^{113}Cd NMR to Biological Systems
Ian M. Armitage and James D. Otvos

Chapter 3

Photo-CIDNP Studies of Proteins

Robert Kaptein

Chapter 4

Application of Ring Current Calculations to the Proton NMR of Proteins and Transfer RNA

Stephen J. Perkins

Index

VOLUME 5

Chapter 1

CMR as a Probe for Metabolic Pathways *in vivo*

R. L. Baxter, N. E. Mackenzie, and A. I. Scott

Chapter 2

Nitrogen-15 NMR in Biological Systems

Felix Blomberg and Heinz Rüterjans

Chapter 3

Phosphorus-31 Nuclear Magnetic Resonance Investigations of Enzyme Systems

B. D. Nageswara Rao

Chapter 4

NMR Methods Involving Oxygen Isotopes in Biophosphates

Ming-Daw Tsai and Larol Bruzik

Chapter 5

ESR and NMR Studies of Lipid-Protein Interactions in Membranes

Philippe F. Devaux

Index

VOLUME 6

Chapter 1

Two-Dimensional Spectroscopy as a Conformational Probe of Cellular Phosphates

Philip H. Bolton

Chapter 2

Lanthanide Complexes of Peptides and Proteins

Robert E. Lenkinski

Chapter 3

EPR of Mn(II) Complexes with Enzymes and Other Proteins

George H. Reed and George D. Markham

Chapter 4

Biological Applications of Time Domain ESR

Hans Thomann, Larry R. Dalton, and Lauraine A. Dalton

Chapter 5

Techniques, Theory, and Biological Applications of Optically Detected Magnetic Resonance (ODMR)

August H. Maki

Index

VOLUME 7

Chapter 1

NMR Spectroscopy of the Intact Heart

Gabriel A. Elgavish

Chapter 2

NMR Methods for Studying Enzyme Kinetics in Cells and Tissue
K. M. Brindle, I. D. Campbell, and R. J. Simpson

Chapter 3

Endor Spectroscopy in Photobiology and Biochemistry
Klaus Möbius and Wolfgang Lubitz

Chapter 4

NMR Studies of Calcium-Binding Proteins
Hans J. Vogel and Sture Forsén

Index

VOLUME 8

Chapter 1

Calculating Slow Motional Magnetic Resonance Spectra: A User's Guide
David J. Schneider and Jack H. Freed

Chapter 2

Inhomogeneously Broadened Spin-Label Spectra
Barney Bales

Chapter 3

Saturation Transfer Spectroscopy of Spin Labels: Techniques and Interpretation of Spectra
M. A. Hemminga and P. A. de Jager

Chapter 4

Nitrogen-15 and Deuterium Substituted Spin Labels for Studies of Very Slow Rotational Motion
Albert H. Beth and Bruce H. Robinson

Chapter 5

Experimental Methods in Spin-Label Spectral Analysis
Derek Marsh

Chapter 6

Electron–Electron Double Resonance
James S. Hyde and Jim B. Feix

Chapter 7

Resolved Electron–Electron Spin–Spin Splittings in EPR Spectra
Gareth R. Eaton and Sandra S. Eaton

Chapter 8

Spin-Label Oximetry
James S. Hyde and Witold K. Subczynski

Chapter 9

Chemistry of Spin-Labeled Amino Acids and Peptides: Some New Mono- and Bifunctionalized Nitroxide Free Radicals
Kálmán Hideg and Olga H. Hankovsky

Chapter 10

Nitroxide Radical Adducts in Biology: Chemistry, Applications, and Pitfalls
Carolyn Mottley and Ronald P. Mason

Chapter 11

Advantages of ^{15}N and Deuterium Spin Probes for Biomedical Electron Paramagnetic Resonance Investigations
Jane H. Park and Wolfgang E. Trommer

Chapter 12

Magnetic Resonance Study of the Combining Site Structure of a Monoclonal Anti-Spin-Label Antibody

Jacob Anglister

Appendix

Approaches to the Chemical Synthesis of ^{15}N and Deuterium Substituted Spin Labels

Jane H. Park and Wolfgang E. Trommer

Index

Index

Amide proton exchange, 171
Amino acid spin systems, 162–170
Amplitue modulation, 214
Analog-to-digital converter (ADC), 199, 219
Anesthetics, 30
Apodization, 148

B. subtilis 5S rRNA, 89–91, 109
Bacteriofage fd coat protein, 182, 183, 186, 191, 194
Base-pair identification, 66–68
Base-pair number, 70–72
Base-pair schemes, 63
Base-pair sequences of 5S rRNA, 68–70
Binomial pulse sequences, 203, 208, 209
Block averaging method, 219

Calcium binding, 10, 26, 29, 36, 37
Carbon-13 NMR of labeled 5S rRNA, 81–83
Cardiolipin, 43, 46
Carr–Purcell pulse sequence, 226
Chemical shift anisotropy, 80
Chemical shift ranges for 5S rRNA, 61
Cholesterol, 26, 37, 45
Coherent transfer effects, 145
Continuous-wave methods, 201
Cosine sequences, 207
COSY (2D correlated spectroscopy), 158
Cross-relaxation, 122
CYCLOPS pulse sequence, 218

DEFT method, 223
DEPT method, 231
Detergents, 30
Deuterium NMR, 34, 48
Deuterium powder patterns, 181, 186
Dilute spins, 178
Dipalmitoylphosphatidylcholine, 36
Diphosphatidylglycerol, 27, 32
Distance calculations, 124–127, 129, 134–139
Dolicyl phosphate, 28
Dynamic range, 199

E. coli 5S rRNA, 82–89, 105–106, 111
Effective field, 214
ERASER pulse shape, 216
Erythrocyte membrane, 45
Exchange rates, 43, 102–106
Exchanging conformations, 140–143

Fluorine-19 NMR of labeled 5S rRNA, 80
Frequency-selective pulse sequences, 200

Glycophorin, 39
Gramicidin, 28

Hard (high amplitude) pulses, 202, 236
Heteronuclear correlated spectrum (HETCOR), 169
Heteronuclear polarization transfer, 229
Homogeneity-spoiling pulse (HSP), 201, 222
Homonuclear Hartmann–Hahn experiment (HOHAHA), 165, 166
Homonuclear polarization transfer, 232
Hydration of oligonucleotides, 131, 133

Imino proton exchange rates, 104–106
INEPT experiment, 230
Interfacial lipids, 39–45
Internal motion, 140
Inversion-recovery sequence, 221
Isolated spin pair approximation (ISPA), 122, 126

Jump and return pulse sequence, 209

Lanthanides, 37
Lipid bilayers, 3
Lipid phases, 28
Liposomes, 4
LUV (large unilamelar vesicles), 4
Lysophosphatidylcholine, 27

Meiboom–Gill pulse sequence, 226
Melting of 5S rRNA, 106–109
Membrane proteins, 28, 38, 46
Membranes, 5
Mitochondrial membranes, 29
MLV (multilamelar liposomes), 4, 31, 33
Motional averaging, 180
Multiple-quantum filtering, 233

NERO pulse sequence, 208
Nitrogen-15 NMR of labeled 5S rRNA, 82–85
Nitrogen-15 powder patterns, 181, 183, 188, 191
NOE in 5S rRNA, 64–70, 91–97
NOESY: *see* Two-dimensional NOE spectrum
Noise suppression, 144
Nonlinear excitation, 213
Notch filter, 202

OBTUSE pulse sequence, 208
Oligonucleotide spectra, 120
Oligonucleotide structure, 130–133, 150
Ovomucoid third domain (OMTKY3), 156–158, 165–175

Peak intensities, 127–129, 148–149
Phase modulation, 214
Phase-cycling schemes, 207, 218
Phosphatidylcholine (PC), 10, 16, 22, 31, 34, 36
Phosphatidylethanolamine (PE), 7, 20, 23–27, 32, 35, 36
Phosphatidylglycerol, 35
Phosphatidylinositol, 39
Phosphatidylserine, 35, 36
Phospholipid phases, 18–20
Phospholipids, 2, 8, 9, 11, 13
Phosphorus-31
  isotropic chemical shifts, 9
  NOE, 11, 31, 32
  powder patterns, 7, 16, 20, 41
  relaxation times, 11, 13, 44
  shielding tensor, 8
Polarization transfer methods, 229–232
Poly(ethylene glycol), 27
POTSHOT method, 232
Presaturation, 217
Program for peak intensity calculation, 127–129

Rapid-scan techniques, 201
Redfield pulse sequences, 209, 235
Relaxation
  methods, 220
  reagents, 217, 225
Relayed correlation spectrum (RELAY), 165, 171
Resonance assignment protocol, 146–147
Rhodopsin, 46, 47
Ring-current induced shifts, 72
Rotational correlation time of 5S rRNA, 81

Sarcoplasmic reitculum, 42
Saturation, 217
Sequential assignments, 164, 168–170
Signal attenuation, 225
Signal-enhancement schemes, 231
Sine sequences, 203
Soft pulses, 209
Software-based methods, 218
Sphingomyelin, 27, 32
Spin labeling, 73
Spin systems, 162–170
Spin-lock, 228
Staphylococcal nuclease, 159
Supraannular phospholipids, 45
SUV (sonicated unilamelar vesicles), 4, 31
SWATTR method, 226
SWIFT excitation, 212

Tailored excitation, 211
Two-dimensional NOE spectrum (NOESY), 120, 122, 167, 170, 173

VASE method, 222
Vesicles, 4
Virus envelope membrane, 43

WATR approach, 225
Water signal, 200
WEFT method, 221
Wheat germ 5S rRNA, 62, 69, 76, 77, 91–93, 107, 110

Yeast 5S rRNA, 74, 75, 79, 80, 81, 94–98
Yeast 5.8S rRNA, 98–102