Annual reports on
NMR Spectroscopy

Volume 65

Academic Press is an imprint of Elsevier
84 Theobald's Road, London WC1X 8RR, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
Linacre House, Jordan Hill, Oxford OX2 8DP, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

First edition 2009

ISBN: 978-0-12-374734-1
ISSN: 0066-4103

Printed and bound in Great Britain
09 10 11 12 10 9 8 7 6 5 4 3 2 1

CONTRIBUTORS

Tomoyasu Aizawa
Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan

Makoto Demura
Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan

P. C. Griffiths
School of Chemistry, Cardiff University, Cardiff CF10 3AT, United Kingdom

Terry Gullion
Department of Chemistry, West Virginia University, Morgantown, WV 26505, USA

Evan F. Haney
Structural Biology Research Group, Department of Biological Sciences, University of Calgary, Alberta, Canada T2N 1N4

Keiichi Kawano
Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

Mineyuki Mizuguchi
Faculty of Pharmaceutical Sciences, University of Toyama, Toyama 930-0194, Japan

Carolyn E. Mountford
Centre for Clinical Spectroscopy, Department of Radiology, Brigham & Women's Hospital, Harvard Medical School, Suite C3-10CC, Boston, MA 02115, USA

Saadallah Ramadan
Centre for Clinical Spectroscopy, Department of Radiology, Brigham & Women's Hospital, Harvard Medical School, Suite C3-10CC, Boston, MA 02115, USA

Hans J. Vogel
Structural Biology Research Group, Department of Biological Sciences, University of Calgary, Alberta, Canada T2N 1N4

Mike P. Williamson
Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, United Kingdom

PREFACE

It is my pleasure to introduce Volume 65 of the Annual Reports on NMR Spectroscopy. Readers of this series will be familiar with the wide range of topics covered in each volume. The present one is an exception, in that almost all of the areas covered have a biological flavour. It begins with an account of 'NMR of Antimicrobial Peptides' by E. F. Haney and H. J. Vogel; this is followed by an account from M. Mizuguchi, T. Aizawa, K. Kawano and M. Demura on 'NMR Studies of Protein Folding: Folding Studies of Calcium-Binding Lysozyme and α-Lactalbumin'; 'Applications of the NOE in Molecular Biology' is covered by M. P. Williamson; T. Gullion reports on 'Recent Applications of REDOR to Biological Problems'; 'Electrophoretic NMR-Ions, Molecules, Mixtures, Pores and Complexes' are covered by P. C. Griffiths; finally, S. Ramadan and C. E. Mountford provide an account of 'Two-Dimensional NMR on Biopsy and In Vivo'.

My sincere thanks go to all of these reporters for their timely and well presented contributions.

G. A. Webb
Royal Society of Chemistry, Burlington House, Piccadilly, London, UK
February 2008

CHAPTER 1

NMR of Antimicrobial Peptides

Evan F. Haney and **Hans J. Vogel**

Contents

Abstract

The continuing increase in the resistance of pathogenic bacteria to conventional antibiotics has led to the emergence of new strains that are often referred to as 'superbugs'. There is a serious need for new anti-infective therapies as clinicians can no longer effectively combat these bacterial pathogens. In the search for new and effective antibiotic compounds, antimicrobial peptides, which are part of a larger group of host defence peptides and proteins, have emerged as a potential class of agents that may stem the tide of antibiotic resistance. However, the mechanism of action of antimicrobial peptides is currently poorly understood. Widely different and complex

Structural Biology Research Group, Department of Biological Sciences, University of Calgary, Alberta, Canada T2N 1N4

Annual Reports on NMR Spectroscopy, Volume 65
ISSN 0066-4103, DOI: 10.1016/S0066-4103(08)00201-9

models for the microbicidal action of antimicrobial peptides have been proposed. Consequently, in order to optimize the design of new synthetic antimicrobial peptides, each peptide must be examined on its own to determine which factors are most responsible for its bactericidal or bacteriostatic activity. Nuclear magnetic resonance (NMR) spectroscopy has become the tool of choice for determining the structures of the antimicrobial peptides and to characterize their interactions with their initial target, the negatively charged bacterial membrane. Moreover, interactions with zwitterionic eukaryotic membranes are also surveyed in an attempt to avoid host cytotoxicity of these peptides. Standard solution-state NMR techniques are now commonly used to determine the high-resolution structure of an antimicrobial peptide in aqueous solution, in detergent micelles or in membrane mimetic organic solvents. Solid-state NMR is increasingly being used to examine the orientation of the peptide in phospholipid bilayers as well as the oligomeric state of the peptide; it can also be used to characterize the changes in biological membranes that are induced upon peptide binding. Through the incorporation of ^{15}N, ^{13}C and ^{2}H nuclei into antimicrobial peptides by recombinant fusion-protein expression techniques, as well as advances in NMR technology and experimental procedures, antimicrobial peptides can now be studied in large membrane mimetic vesicles which more closely resemble biological membrane bilayers *in vivo*. This review paper will discuss many of the solution and solid-state NMR experiments that have been used to characterize antimicrobial peptides and that have provided researchers with unique insights into the diverse mechanisms of action of this class of biomolecules.

Key Words: Antimicrobial peptides, Nuclear magnetic resonance spectroscopy, Mechanism of antimicrobial activity, Detergent micelle, Peptide structure determination, Membrane perturbation.

1. INTRODUCTION

Small peptides with the ability to selectively kill bacterial cells were reported as early as 1981 when the primary structure of insect cecropin was published by Hans Boman and co-workers.[1] Cecropins were identified as important components of the insect immune system and it was assumed that this type of mechanism would not be present in the more sophisticated immune system of higher animals. However, research into antimicrobial peptides exploded when the first peptide from an animal was discovered on the skin of the African clawed frog, *Xenopus leavis*.[2] This peptide was discovered virtually by accident when Michael Zasloff was working on oocytes harvested from female frogs. Dr. Zasloff used a non-sterile surgical technique to harvest the oocytes from female frogs, sutured the wounds and then placed the animals back into a holding tank. In spite of the unsterile surgical procedures and the microbially rich environment present in the holding tanks, these frogs rarely developed infections and the wound healed normally. Zasloff hypothesized that the skin of the frogs had 'sterilizing' activity

and proceeded to isolate a small antimicrobial peptide from the skin of the frog that he named magainin.[2]

Since that time, antimicrobial peptides have been identified in insects,[3] plants,[4] animals[5] and humans.[6] Even bacterial cells have been shown to produce peptides with bactericidal activity[7] which gives them an evolutionary advantage against competing microorganisms in the same environment. Antimicrobial peptides are now recognized as an important part of the innate immune system and often are the first line of defence against invading pathogens. Research into antimicrobial peptides is fuelled by the increasing resistance of clinically relevant bacterial strains to commonly used antibiotics.[8] This creates a need to develop new bactericidal compounds to replace the ineffective conventional antibiotics. Antimicrobial peptides with different amino acid sequences are continually being discovered and synthetically altered in an attempt to generate more powerful antibacterial molecules and to understand the underlying mechanism of activity of this class of peptides. There are a few antimicrobial peptide databases that have been set up to catalogue the ever increasing number of peptide sequences with antibacterial effects. One of these was established in 2004 by Guangshun Wang and co-workers at the University of Nebraska[9] and it now contains information on over 700 peptide sequences with antibacterial, antifungal, antiviral and anticancer activities (http://aps.unmc.edu/AP/main.php). Another database was established even earlier by the research group of Allesandro Tossi at the University of Trieste in Italy (http://www.bbcm.units.it/~tossi/amsdb.html) with the same goal of cataloguing these antibacterial molecules as they are reported in the literature.

Antimicrobial peptides show a tremendous amount of variation in their primary sequences and structures. Typically, naturally occurring antimicrobial peptides range in size from 12 to 60 residues; however, synthetic peptides as short as six residues have been shown to have significant antimicrobial activity.[10] Most antimicrobial peptides have a net positive charge owing to a preponderance of lysine and arginine residues. The cationic nature of these peptides is believed to be responsible for the selectivity of antimicrobial peptides for bacterial membranes (discussed below). Antimicrobial peptides also possess a high proportion of hydrophobic residues such as phenylalanine, leucine, valine and alanine.[11] Tryptophan residues are found in unusually high proportions in some classes of antimicrobial peptides[12] and the fact that they have a preference for the interfacial region of lipid bilayers[13] also has implications for their mechanism of action.

Numerous structures of antimicrobial peptides are reported in the literature. They typically adopt amphipathic conformations with one face containing the cationic amino acids and the opposite face possessing the hydrophobic amino acids. We have previously attempted to classify antimicrobial peptides based on their structural characteristics and their amino acid composition.[14] Antimicrobial peptides can be separated into five classes: amphipathic α-helices, β-sheet peptides, cysteine-rich peptides, peptides with an unusually high content of a common amino acid and peptides with uncommon and/or chemically-modified amino acids.

Structural characterization of antimicrobial peptides is a vital step in understanding the underlying mechanism of antimicrobial activity of this class of

compounds. X-ray crystallography and Nuclear Magnetic Resonance (NMR) Spectroscopy are the two most important techniques for high resolution structure determination of biomolecules, such as proteins. However, in the case of antimicrobial peptides, the use of NMR has been much more prevalent than X-ray crystallography, as evidenced by the fact that the overwhelming majority of the high resolution structures of antimicrobial peptides deposited in the Protein Data Bank were solved using various NMR methods. Obtaining high quality crystals of antimicrobial peptides in a lipid environment is extremely difficult because of their small size, intrinsic flexibility and the inherent difficulties associated with crystallizing membrane proteins. On the other hand, NMR spectroscopy has proven to be a versatile tool for studying antimicrobial peptides. Recent reviews have examined the structural characterization of antimicrobial peptides and how they interact with phospholipids.[12,14–19]

This chapter is not meant to be a comprehensive review of all of the NMR studies reported for antimicrobial peptides, nor will it delve into detail regarding the theoretical considerations underlying the experimental techniques used. Instead, this review will examine the type of information that can be obtained from NMR spectroscopy when studying antimicrobial peptides and how this information can be applied to decipher their mechanisms of action when killing bacterial cells. Additionally, we will discuss some of the limitations of the techniques used and examine some key considerations when analyzing the results.

2. MECHANISM OF ACTION

Owing to the large diversity in sequence and structures, the mechanisms of action of antimicrobial peptides are also varied and diverse. Many antimicrobial peptides are believed to interact with the surface of bacterial cells and induce membrane perturbations. Peptides are thought to induce structural changes in the bacterial membrane that lead to the formation of pores or the solubilisation of the membrane, both of which can lead to the disruption of electrochemical gradients and the release of intracellular components to the environment. The most popular pore forming models are the barrel-stave model and the toroidal pore model in which the antimicrobial peptides oligomerize to form transmembrane pores. In the barrel-stave model (Figure 1A), the pore is completely lined by peptides with the hydrophobic face of the amphipathic molecule interacting with the acyl chains of the bilayer.[20] This model was originally proposed to explain the antimicrobial activity of alamethicin[21] and the structure of this pore was recently solved using X-ray diffraction studies and found to contain eight alamethicin helices per transmembrane pore.[22] However, the barrel-stave model has proven to be relatively ineffective at explaining the antimicrobial activity of the myriads of other antimicrobial peptides. It may well be that the bacterial anionic alamethicin peptide is simply not a good model for the cationic antimicrobial peptides found in eukaryotes.

In the toroidal pore model (Figure 1C), the antimicrobial peptides induce a curvature strain on the leaflets of the bilayer until the two sides of the bilayer fuse to form a peptide/lipid lined pore.[23] This mechanism has been invoked to explain

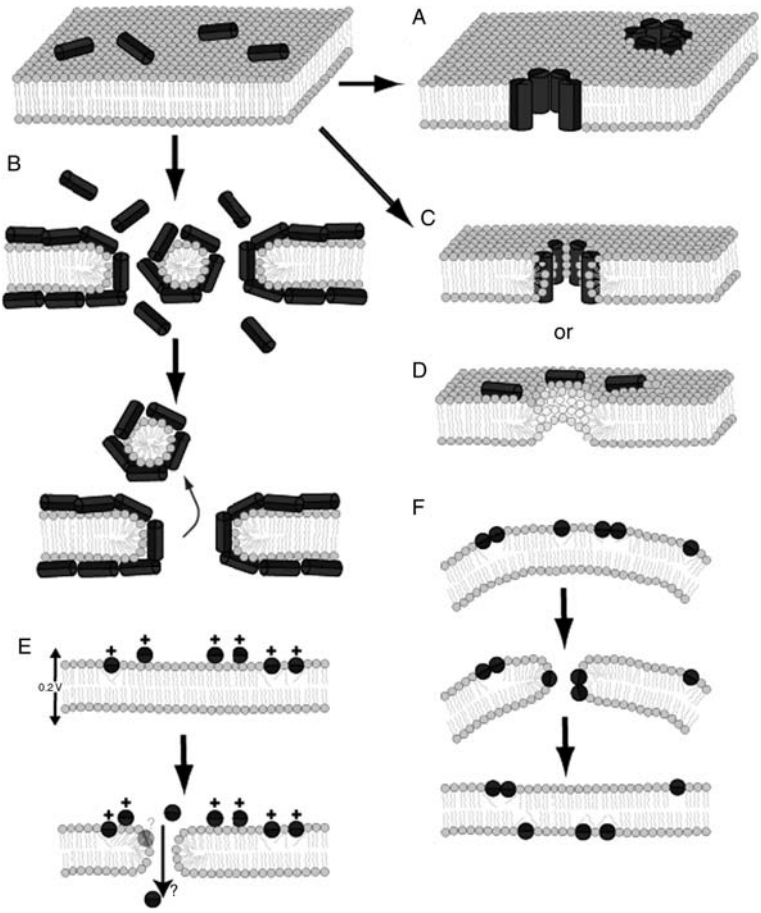


Figure 1 Proposed models for the mechanisms of antimicrobial activity of antimicrobial peptides. Following aggregation of the antimicrobial peptide on the surface of the bilayer, the peptide can disrupt the bilayer through; the barrel-stave model (A), the carpet mechanism (B), the toroidal pore model (C), the in plane diffusion model (D), the electroporation model (E) or the sinking raft model (F). (Figure adapted from Ref. 12, with permission).

the antimicrobial activity of a number of antimicrobial peptides including magainin 2,[24] piscidin[25] and LL-37.[26] A mechanism similar to the toroidal pore model is the in-plane diffusion model[27] (Figure 1D) which is applicable to peptides that induce local membrane perturbations close to the region where the peptide binds and causes the formation of a lipid lined pore that is stabilized by peptides bound to the surface of the bilayer. Experimental evidence for this type of mechanism is limited but molecular dynamics simulations of the magainin analog MG-H2 have demonstrated that a pore forms spontaneously when this peptide is mixed with DPPC bilayers. Of significant importance is that only one peptide is found inside the pore while the remainder stabilize the opening by binding to the rim of the pore.[28]

The carpet model of antimicrobial activity (Figure 1B) is also a popular mechanism to explain the antimicrobial activity of many antimicrobial peptides.[20] This

model suggests that the peptides accumulate on the surface of the bacterial membrane and once an effective concentration is reached, they induce the micellarization of the membrane and essentially dissolve part of the phospholipid bilayer, leading to cell death.

Other mechanisms for antimicrobial activity have been proposed and although they receive less attention than those already mentioned, they are nonetheless relevant when discussing the possible modes of action of this class of antibacterial molecules. The electroporation model (Figure 1E) explains the formation of transient membrane pores as a result of the electric potential that is created across a membrane due to the accumulation of a large number of cationic peptides on one side of the bilayer.[29,30] The sinking raft model (Figure 1F) also advocates the formation of transient pores in the membrane. In this model, the peptides aggregate on the outer leaflet of the membrane which creates a mass imbalance across the bilayer. This induces changes in the local membrane curvature around this peptide oligomer and these peptides eventually pass through the membrane and redistribute themselves along both sides of the bilayer.[31] While passing through the membrane these peptide bundles induce transient pores that have the same effect as the pores mentioned earlier.

Some antimicrobial peptides have relatively little effect on membrane stability but instead pass through the phospholipid bilayer and exert their antimicrobial effect on intracellular targets.[11] The short proline-rich family of antimicrobial peptides are an example of this type of peptide as they are thought to inhibit intracellular protein synthesis and/or protein folding. Most of the antimicrobial peptides in this family are produced by insects; however, there are also several mammalian examples of proline-rich antimicrobial peptides.[32] Proline residues in antimicrobial peptides often have distinct effects on the ability of these antibiotic molecules to disrupt the integrity of a bacterial membrane. Inserting a proline residue into an α-helical antimicrobial peptide with known membrane destabilizing properties decreases the membrane permeabilization ability of that peptide.[33] The opposite situation has also been observed for buforin II, a potent 21-residue antimicrobial peptide derived from a naturally occurring peptide found in the stomach of the Asian toad.[34] Buforin II is an α-helical antimicrobial peptide with a proline hinge that is capable of quickly killing bacterial cells but it has no effect on the integrity of the plasma membrane.[35] Interestingly, buforin II analogs that lack this proline hinge turn into membrane destabilizing peptides indicating that the proline residue is essential for the translocation mechanism across the membrane.[35]

Other antimicrobial peptides with different structural motifs also appear to act on intracellular targets within bacterial cells with no appreciable effect on the integrity of the phospholipid bilayer. Polyphemusin is a *β*-hairpin peptide that displays antimicrobial activity *in vitro* and causes lipid flip-flop but it does not induce significant vesicle leakage.[36] In fact, polyphemusin accumulates in the cytoplasm of *E. coli* cells without any measurable damage to the integrity of the bacterial membrane.[37] The proposed mechanism used by polyphemusin to cross the bacterial membrane is similar to the sinking raft model except that the peptide induces negative curvature strain on the bilayer and then translocates across the

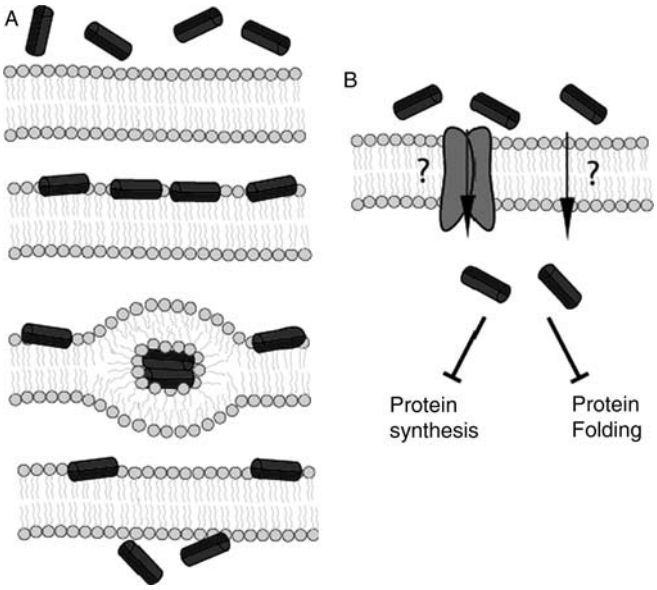


Figure 2 Proposed non-bilayer perturbing mechanisms of membrane translocation of antimicrobial peptides. Antimicrobial peptides may pass through the membrane through a non-bilayer intermediate like that proposed for polyphemusin[36] (A). Antimicrobial peptides could also reach the cytoplasm through integral membrane proteins or through some other unknown mechanism (B).

membrane through a non-bilayer intermediate without disrupting the integrity of the plasma membrane (Figure 2A).

Certain antimicrobial peptides do not interact directly with the membrane but instead they cross the bacterial membrane through an integral membrane receptor or by some other unknown mechanism (Figure 2B). For example, bacteriocins are a class of antimicrobial peptides that are ribosomally synthesized by bacterial cells as a means of out-competing other species in the local environment for nutrients and resources. Microcin E492 is a small bacteriocin produced by *Klebsiella pneumoniae* with activity against other Enterobacteriaceae species.[38–40] The mechanism of action of this peptide is not completely understood but there are various bacterial membrane receptors that have been implicated in its activity. Microcin E492 is recognized by the outer membrane receptors FepA, Fiu and Cir[41] which is likely due to a post-translational modification of microcin that causes it to mimic a cathechol-type siderophore.[39] Additional evidence indicates that microcin E492 inserts into the inner membrane of bacterial cells and induces the formation of pores, leading to membrane depolarization and increased permeability.[42,43] A recent mutagenesis study also found that an inner membrane protein complex responsible for mannose uptake is essential for the bactericidal activity of this antimicrobial peptide.[44] The mannose phosphotransferase system has also been implicated as a receptor for lactococcin A, lactococcin B and other pediocin-like bacteriocins.[45]

Bacteriocins are not the only antimicrobial peptides that take advantage of transmembrane receptor proteins to transport across the bacterial membrane.

The interaction between *E. coli* cells and the proline-rich antimicrobial peptide Bac-7 (1–35) was examined and peptide translocation appears to be mediated by an inner membrane receptor encoded by the *sbmA* gene.[46] The *sbmA* gene encodes a 407 residue inner membrane protein in Gram-negative bacteria[47] and is known to be important for the uptake of other microcins into *E. coli* cells.[48,49] In this particular study, *E. coli* mutants with a point mutation in the *sbmA* gene or an *sbmA* knockout were much less susceptible to the antimicrobial activity of Bac-7; in contrast, overexpression of the *sbmA* gene made these cells more sensitive to the effects of this peptide.[46] Although the exact mechanism through which the *sbmA* gene product translocates this proline-rich antimicrobial peptide is presently unclear, the role of transmembrane proteins in bacterial cells as a means of crossing the phospholipid bilayer needs to be considered when discussing which mechanisms of action antimicrobial peptides rely on to exert their antibiotic effects. Clearly, specific interactions mediated by bacterial membrane receptors can be responsible for membrane translocation while the other mechanisms of antimicrobial activity are considered non-specific in their mode of action.

Finally, we should point out that there is an increasing amount of evidence that points to intracellular targets for those antimicrobial peptides that do not cause serious destabilization of the phospholipid bilayer of bacterial cells. Buforin II is thought to act by spontaneously crossing the cell membrane and enter the cytoplasm of the cell where it is free to bind the bacterial DNA and RNA and consequently inhibit normal cellular functioning.[50] Direct evidence of an antimicrobial peptide binding to DNA has been observed for indolicidin, a peptide found in bovine neutrophils, where peptide binding was confirmed through Trp fluorescence quenching upon interaction with DNA and the observation that indolicidin inhibits DNA migration in agarose gels.[51] Additionally, an all D-amino acid form of the proline-rich antimicrobial peptide apidaecin is inactive,[52] suggesting that this peptide binds to an intracellular target in a stereospecific manner as opposed to the non-specific interactions with the bacterial membrane described earlier.

Because of the widely divergent proposed models of antimicrobial activity, numerous questions need to be answered to elucidate the antimicrobial mechanism for a given peptide. What is the structure of this molecule in aqueous solution? Does the conformation of the peptide change in the presence of lipids? Is there a preference for a charged lipid species versus a zwitterionic one? Is the peptide oriented along the surface of the bilayer or does it adopt a transmembrane conformation? How deeply does the peptide insert itself into the hydrophobic core of a bilayer? Does the peptide affect the structure and organization of bilayers? Answers to all these questions, and many more, can be obtained through the use of NMR experiments using both solution and solid-state NMR techniques.

3. SOLUTION-STATE NMR

The small size of antimicrobial peptides makes them attractive candidates for structural studies using solution-state NMR. The experimental approach for determining the solution structure of this class of peptides is well established

and involves common 2D homonuclear NMR techniques (described later in this chapter). Sample preparation is simple as the peptide is usually dissolved in a 9:1 mixture of H_2O:D_2O to a final peptide concentration of 1–3 mM. The D_2O is necessary as a field-locking signal during the NMR experiment. The high concentration of peptide is required because NMR is an inherently insensitive technique and increasing the peptide concentration simply provides a stronger NMR signal. The solubility of antimicrobial peptides at this high concentration is usually not an issue because of their large cationic charge. However, aggregation may occur in some cases depending on other factors like salt concentration or pH which may force the peptides into insoluble aggregates. If peptide aggregation is a problem, a lower peptide concentration can be used and then the loss in signal strength can be overcome through the use of higher field magnets and cryoprobes.

The pH value to be used for the NMR sample deserves some careful consideration. Many of the solution structures of antimicrobial peptides that are reported in the literature are obtained at acidic pH values between 3.5 and 4.5. At high pH values, there is increased exchange between the hydrogen atoms in the peptide and the surrounding water which decreases the intensity of NOE cross-peaks that are required for structure determination, namely the amide protons in the peptide backbone. For example, the solution structure of lactoferricin B (LFcinB), a 25-residue antimicrobial peptide obtained through pepsin digestion of the intact bovine lactoferrin protein,[53] was determined at a pH of 4.5.[54] The structure of tachystatin B, an antimicrobial and chitin-binding peptide from the Japanese horseshoe crab, was solved at an even lower pH of 3.5.[55]

Another aspect of the pH that is important is considering the solubility of any lipid molecules that are present in the NMR sample. A micelle forming phospholipid molecule, dioctanoyl phosphatidylglycerol (D8PG), was recently used in NMR studies of peptides derived from the human antimicrobial peptide LL-37.[56] The pH of the NMR sample had to be maintained between 5 and 6 because a lower pH leads to a decreased solubility of D8PG but the authors still wanted to take advantage of the lower hydrogen exchange between the peptide and the bulk water which allowed for the observation of intermolecular NOE between the peptide and the D8PG micelles. Finally, spontaneous oxidation of disulfide bonds can occur at higher pH values so structural characterization of disulfide stabilized antimicrobial peptides is usually done at acidic pH to avoid this potential complication.[57]

Most of the models used to explain antimicrobial activity involve an electrostatic interaction between the negatively charged head groups of the bacterial membranes and the positively charged antimicrobial peptide. This presents a challenge in antimicrobial peptide research because the acidic conditions of the NMR samples may not be representative of the state of the peptide at a physiological pH. It is conceivable that an acidic pH in the NMR sample, especially one containing a micellar environment, may change the protonation state of charged amino acid sidechains which could potentially disrupt the electrostatic interactions between the peptide and the micelle surface. Histatins are a class of histidine-rich antimicrobial peptides found in human saliva and they have been structurally characterized by NMR. Histatin 5 was examined in an aqueous

solution at a pH of 3.8 and it was determined to be relatively unstructured.[58] A later study looked at the structural characteristics of aqueous solutions of histatin 3 and histatin 5 at pH 7 and 7.5, respectively.[59] Histatin 5 at pH 7.5 was relatively unstructured which agreed with the previous study at pH 3.8 while histatin 3 exhibited some structure at pH 7. In the study at higher pH, the authors actually took advantage of the fast exchange rates of the amide protons to confirm a lack of secondary structure of the peptide: amides that are involved in stabilizing hydrogen bonds will exchange much more slowly than those in a random coil conformation.

It is necessary to reference all NMR spectra to a standard compound with a known chemical shift. A commonly used internal standard is 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) which we normally employ in our NMR studies with antimicrobial peptides.[60–64] DSS is highly water soluble and the resulting high intensity proton peaks are easily identifiable and are defined as 0.0 ppm by convention. However, DSS is negatively charged and has been shown to interact with some cationic peptides.[65] Our group has also observed intermolecular NOEs between LFcinB and DSS under certain conditions (unpublished results). Such interactions could potentially be troublesome if they affect the structure of the peptide or disrupt important electrostatic interactions.

An alternative to using DSS as an internal standard is to reference the NMR spectra to DSS as an external standard. This method has been applied to NMR studies of LL-37 derivatives[56] and bovine indolicidin.[51] Recently, a new internal standard, 4,4-dimethyl-4-silapentane-1-sulphonate (DSA) was developed that is similar in structure to DSS but the negatively charged sulphate group in DSS has been replaced with a positively charged ammonium group.[66] DSS and DSA both have trimethylsilyl groups that give identical chemical shifts which can be defined as 0.0 ppm in the NMR spectra. Six cationic peptides that are believed to adopt β-hairpin structures were tested with either DSS or DSA to examine if the chemical shifts of these standards were altered in the presence of the peptide. In most of the cases, the relative chemical shift of the DSS was significantly altered while the resonances corresponding to DSA were relatively consistent.[66]

4. SOLUTION-STATE NMR EXPERIMENTS

Structural characterization of antimicrobial peptides can usually be achieved through the use of three main NMR experiments; two-dimensional (2D) correlated spectroscopy (COSY), total correlation spectroscopy (TOCSY) and nuclear Overhauser enhancement spectroscopy (NOESY). COSY was one of the first 2D NMR experiments reported[67] and it is useful for studying relatively small biomolecules. Cross-peaks in a COSY experiment arise from coupled spins which in the case of peptides corresponds to protons separated by two or three bonds.[68] The pulse sequence for the COSY experiment is two 90° pulses separated by an evolution time followed by the acquisition of the NMR signal.[69] Further improvement of the COSY spectra is achieved when a double quantum filter (DQF) is

applied to the experiment.[70] Normally in a COSY experiment, protein cross-peaks have different line shapes than the diagonal which can overshadow important cross-peaks that reside close to the diagonal. The DQF-COSY experiment produces a phase-sensitive spectrum in which protein peaks and diagonal peaks have the same lineshapes and this makes cross-peaks close to the diagonal easier to identify.

The TOCSY experiment is also a through-bond NMR technique that allows for the identification of all of the protons within an amino acid spin system. This technique, also referred to as homonuclear Hartmann-Hahn (HOHAHA) spectroscopy, uses isotropic mixing that transfers in-phase magnetization through several couplings during the mixing period.[68] Cross-peaks generated from this experiment arise from all of the protons within a spin system and can be used in conjunction with the COSY experiments described above to assign all of the chemical shifts of the resonances for the protons in an antimicrobial peptide.

The most important experiment for structure determination of antimicrobial peptides is the NOESY experiment. Contrary to the COSY and TOCSY experiments that rely on scalar couplings and provide information on single amino acids, the NOESY experiment correlates protons that are close in space through dipole–dipole cross-relaxation. It is based on the nuclear Overhauser effect where the spin polarization of one spin population is transferred to a nearby spin population resulting in a cross peak in the spectra. This experiment correlates protons that are closely related to each other in three-dimensional (3D) space and forms the basis for structural calculations of these molecules. The intensity of a given cross-peak in a NOESY spectrum is directly related to the distance between the two protons in the structure of the peptide. NOE cross-peaks arise from protons that are within 5.5 Å of each other in space and the intensity of the cross-peak increases as the distance between the two spins decreases. Oftentimes, peaks in a NOESY spectrum will be separated by their intensities into strong, medium and weak ones which are then categorized into specific distance restraints between protons when structure calculations are performed.

Various computer programs exist that take the information obtained from the NOESY spectra and use it as distance restraints to determine the 3D structure of the antimicrobial peptide. We will briefly examine some of these programs here. TALOS (torsion angle likelihood obtained from shift and sequence similarity program) predicts the backbone angles in a polypeptide chain based on various chemical shift values.[71] This program takes advantage of the fact that the chemical shifts of backbone atoms are sensitive to the structure of the peptide and uses $^{1}H\alpha$, ^{15}N, ^{13}CO, $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts as input values. TALOS compares the chemical shift values to a database of chemical shifts of known structures and searches for a triplet of residues with similar sequence and chemical shift values. The central residue in the triplet can then be assigned φ and ψ values for the backbone angles which are used as restraints for structure determination. The application of the TALOS program requires that the peptide of interest be isotopically labelled with NMR active nuclei. Because most NMR studies of antimicrobial peptides use synthetic peptides that are not labelled, it is not commonly used in

structural characterization of antimicrobial peptides. With that in mind, some studies have taken advantage of natural abundance ^{15}N and ^{13}C in a peptide to obtain inputs values for the TALOS software[72,73] and advances in recombinant expression of antimicrobial peptides will also permit for isotopic labelling of this class of molecules (discussed below).

The Crystallography and NMR System (CNS) software[74,75] is an open source software package available to non-profit research groups. It can be used to analyze both crystallographic and NMR data. At the moment, available inputs for NMR data include NOEs, J coupling, chemical shift and dipolar coupling.

Xplor-NIH is another software package available for NMR structure determination that uses a variety of energy minimization steps to generate 3D structures of biomolecules from NMR data.[76] Numerous potential energy terms are included in Xplor-NIH which correspond to observable NMR data. The most important of these for peptide structure determination is NOE-derived interproton distances in which each NOE cross-peak is defined by the two nuclei that are interacting in the peptide. For ambiguous restraints, Xplor-NIH implements the ARIA protocol described below. In addition to the NOE distance restrains, Xplor-NIH also has the ability to input restraints based on residual dipolar couplings (RDCs), chemical shift anisotropy, torsion angle restraints, three-bond J coupling constant restraints and various other chemical shift values related to φ and ψ angles.[76]

Analysis of a NOESY spectrum involves two steps: assigning cross-peaks to magnetically interacting atoms and then calculating structures based on the restraints that were identified. The analysis of NOESY spectra becomes complex when there is more than one possible assignment for a single cross-peak. To get around this problem, an iterative approach that uses these ambiguous distance restraints has been developed to automatically assign NOEs and determine which distance restraints are compatible with the structure of the biomolecule obtained. This program is known as ARIA (ambiguous restraints in iterative assignments)[77,78] and as of 2003, more than 50 structures in the PDB used ARIA to determine the NMR solution structure.[79] A new version of this program, ARIA 2.1, was released in early 2007[80] with the ability to automatically assign NOEs and perform faster structure calculations.

Numerous other automated assignment programs are available for structure determination based on NMR data (reviewed in Refs. 81 and 82). Examples of these include, CYANA,[83,84] Autostructure,[85] NOAH[86] and PASD.[87] If a relatively complete chemical shift assignment is available and a peak list based on NOESY spectra has been generated for a given peptide in solution, then these programs can automatically assign the NOEs between protons and calculate a 3D structure using these restraints.

We have only highlighted a few of the software packages available for structure determination of proteins from NMR data. In practice, most research groups will use one program for all of their structure calculations because they become comfortable with a specific program and have developed strategies to streamline and facilitate faster structure calculations.

5. SOLVENT CONSIDERATIONS

Certain antimicrobial peptides are structured in aqueous solution; however, these peptides are usually comparatively large and often contain intramolecular disulfide bridges that help maintain their 3D structures in water. For example, tachystatin B, an antimicrobial peptide isolated from the horseshoe crab, folds in aqueous solution thanks to a characteristic cysteine-knot motif.[55] β-defensins also fall into this category of antimicrobial peptides that are folded in aqueous solution. β-defensins have 6 Cys residues that form intramolecular disulfide bonds between Cys1-Cys5, Cys2-Cys4 and Cys3-Cys6, respectively. Various solution structures of β-defensins in aqueous solution have been published including spheniscin from penguins[88] and human β-defensins[89,90] (Figure 3). An even more interesting example comes from an θ-defensin, RTD-1, isolated from the leukocytes of *Rhesus macaques*.[91] This peptide, like the β-defensins, contains three disulfide bonds but it is unique because the peptide backbone is naturally cyclized and the peptide forms a β-sheet stabilized by disulfide bridges.[92] Defensins are not the only disulfide-stabilized antimicrobial peptides that are structured in aqueous solution. LFcinB, a 25-residue antimicrobial peptide released from the intact lactoferrin protein upon digestion with pepsin, has a single disulfide bond that cyclizes this peptide and it folds into a twisted β-sheet in a low salt aqueous solution[54] (Figure 3).

Stable and structured cyclized peptides are also found in plants. These peptides fall into the class of cyclotides and structural examples of these peptides include the extremely stable kalata B1[93] and kalata B2[94] peptides. Many of these peptides can interact with phospholipid bilayers but they alter their conformations to facilitate their interaction with the lipid molecules. For example, arenicin-2 forms a twisted β-sheet in aqueous solution but CD data indicate that the conformation of the peptide changes in the presence of SDS and DPC micelles.[57] This would be

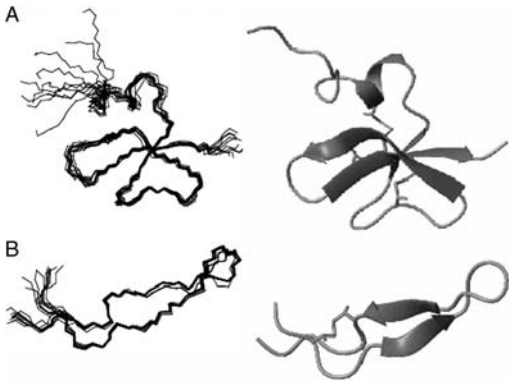


Figure 3 Structures of human β-defensin 3 (A) (PDB ID 1KJ6) and bovine lactoferricin (B) (PDB ID 1LFC) in aqueous solution. The backbone overlay of the 10 lowest energy structures are shown on the left while the ribbon trace is shown on the right.

due to a rearrangement of the peptide structure to allow for more favourable interactions with the detergent micelles.

For the most part, the structure determination of this type of antimicrobial peptide is fairly straightforward as the peptide readily dissolves in aqueous solution and relatively simple NMR experiments are sufficient to obtain the NOE restraints necessary for structural characterization. Additional information, such as chemical information about which Cys residues are involved in the disulfide bonds, can also be incorporated into the structure calculations to give high resolution solution structures from the information available.

Most linear antimicrobial peptide sequences are unstructured in an aqueous environment and they only fold into an amphipathic structure when they encounter the bacterial membrane. This complicates the structure determination of antimicrobial peptides bound to phospholipids because vesicles comprising phospholipids are too large to be used in conventional NMR experiments. Linewidths in modern NMR experiments are directly related to the rotational correlation time of the molecule in question and as a result, larger molecules will produce broader linewidths in an NMR experiment.[68] Large unilamellar vesicles (LUVs) are considered a useful model for studying membranes because they are fairly homogenous in size and structure and are similar to biological membranes because they have a relatively flat surface like a biological membrane. Unfortunately, depending on the phospholipid species used, LUVs are typically in the range of 60–120 nm in diameter[95] which means that they tumble slowly in solution and as a result they produce extremely broad NMR signals. Even small unilamellar vesicles (SUVs), with a diameter of 15–30 nm, are still too large for conventional NMR experiments.

To overcome the size restraints of solution-state NMR experiments, detergent micelles are commonly used to mimic a lipid environment. Many detergents form small, spherical micelles with the most common detergents used in NMR experiments being sodium dodecylsulphate (SDS) and dodecylphosphocholine (DPC). SDS micelles have a negatively charged head group and are thought to be an adequate model for the bacterial membrane. The phospholipid bilayer of bacterial cells is comprised of phosphatidylethanolamine (PE) and phosphatidylglycerol (PG). The ethanolamine head group of the PE molecule is zwitterionic but the glycerol head group of PG confers an overall negative charge to the membrane. As a result, the bacterial membrane also has a net negative charge which favourably attracts the cationic charges of antimicrobial peptides to the surface of the cell. The outer leaflet of eukaryotic membranes are mostly composed of lipids with a phosphatidylcholine head group, which is also zwitterionic, meaning that these membranes have no net negative charge. DPC is therefore considered to be a model for eukaryotic membranes because the choline head group of the DPC molecules is identical to that found in PC containing phospholipids. The chemical structure of SDS and DPC can be seen in Figure 4 along with the structure of certain phospholipids relevant to our discussion. The detergent molecules are usually perdeuterated to avoid signal overlap with peptide resonances,[18] especially in the aliphatic region of the spectra where most of the acyl chain chemical shifts would be observed.

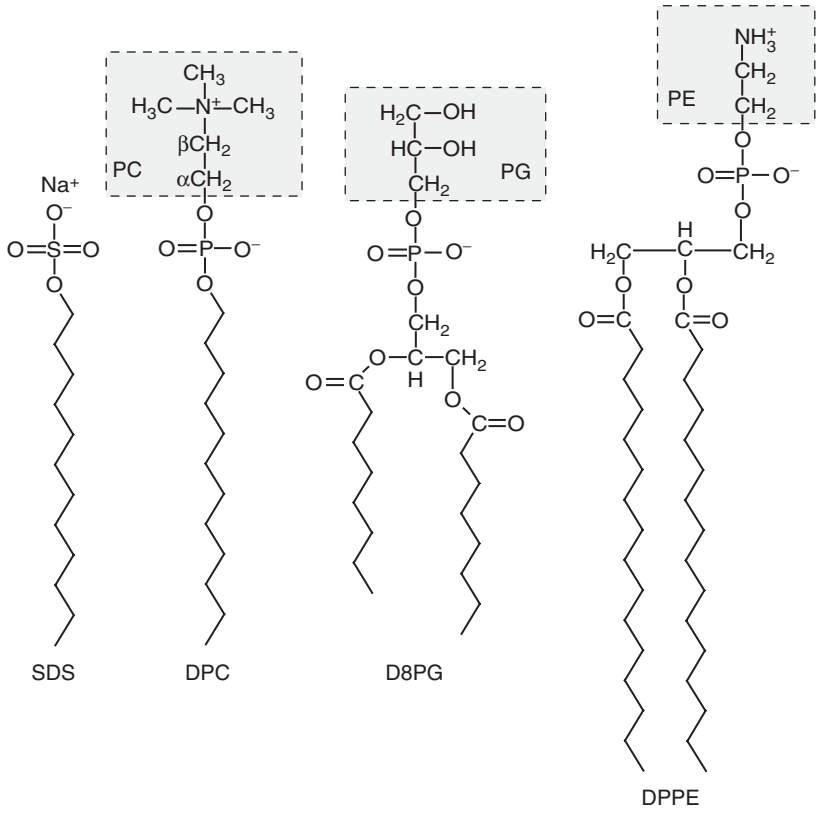


Figure 4 Chemical structure of SDS, DPC, and D8PG. The headgroups of phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) are also highlighted. These three headgroups can all be found on the phospholipid species, in this case DPPE is shown. It should also be noted that the acyl chains in the phospholipid species can be of different length and saturation, both of which will effect the resulting properties of the bilayer.

While it is generally accepted that the structures of antimicrobial peptides bound to detergent micelles are a good representation of the structure of these molecules when they are bound to the bacterial membrane, it is important to keep in mind the limitations of using detergent micelles. Micelles have a highly curved surface and often the structure or the antimicrobial peptide will also be curved to reflect the fact that the peptide wraps around the surface of the micelle.

The number of peptides bound to a detergent micelle is also an important consideration because most antimicrobial peptides require a threshold concentration of peptide to be reached in order to disrupt bacterial cell growth. However, it is generally assumed that one peptide molecule will bind to a single micelle and this peptide–lipid ratio may be smaller than that required for antimicrobial activity. In the case of DPC, ~30 detergent molecules are present in each micelle and only one peptide molecule is bound. This has been observed in NMR studies of mellitin[96] and nisin[97] in the presence of DPC micelles. The SDS aggregation

number is closer to 60 molecules per micelle[98] but it is still assumed that only one antimicrobial peptide is in complex with any given SDS micelle. The peptide to lipid molar ratio that is required to have any antibacterial effect is usually very high but increasing the peptide concentration in the presence of detergents is difficult.

6. STRUCTURES OF ANTIMICROBIAL PEPTIDES IN MICELLES

A large proportion of the high resolution structures of antimicrobial peptides have been determined in the presence of deuterated detergent micelles. The effect of micelles inducing structure in antimicrobial peptides can be seen in Figure 5 as an

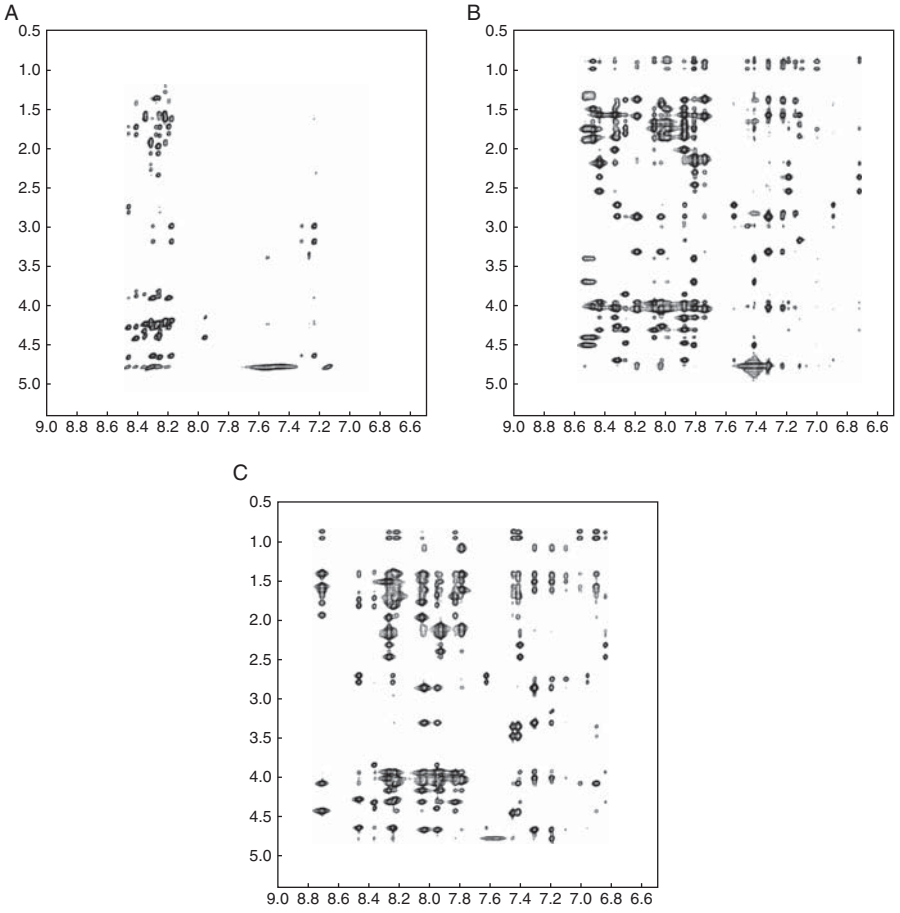


Figure 5 2D ^{1}H NOESY spectra of LFampinB in water (A), SDS (B) and DPC (C) micelles. The increased number of peaks and greater separation between peaks is indicative of increased structure in the presence of SDS and DPC micelles compared to the aqueous environment. The mixing times for the aqueous and micelle spectra are 500 and 100 ms, respectively. (Unpublished results from Haney, Lau and Vogel).

increase in the number of peaks and greater dispersion of NOE cross-peaks in a 2D NOESY spectrum. The 3D structures of antimicrobial peptides differ drastically from one molecule to the next. Some examples of solution structures of antimicrobial peptides bound to detergent micelles can be seen in Figure 6.

Antimicrobial peptides derived from LFcinB have received a lot of attention in an attempt to characterize the antimicrobial activity of this β-sheet peptide. As mentioned earlier, the 25-residue LFcinB peptide folds into an amphipathic β-sheet in aqueous solution[54] but when mixed with SDS or DPC micelles, the resulting spectra produces broad resonances that are unsuitable for structure determination (Unpublished results). The antimicrobially active centre of LFcinB was narrowed down to six residues (RRWQWR) and this shorter peptide, with an amidated C-terminus to remove the carboxy-terminal negative charge, maintains similar antimicrobial activity compared to the parent peptide.[99] The structure of

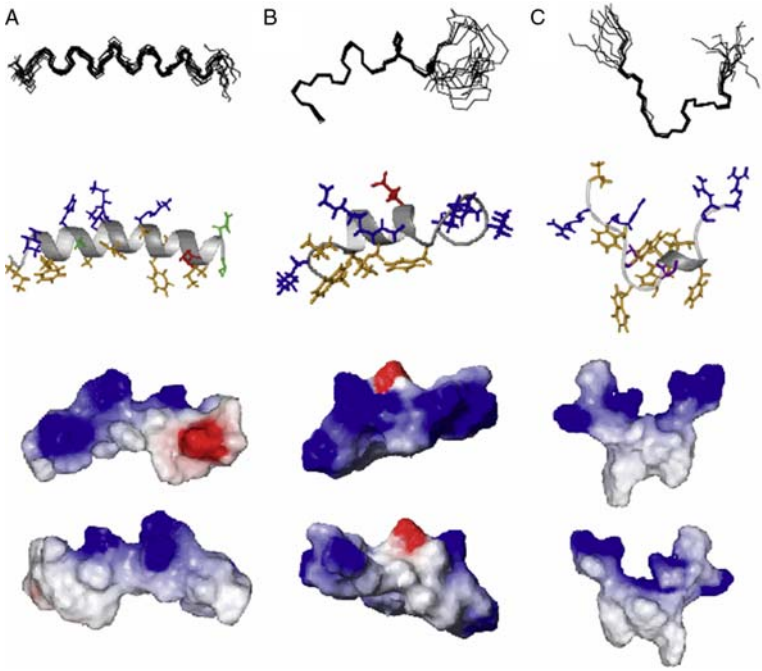


Figure 6 Solution structures of some antimicrobial peptides determined in the presence of detergent micelles. Magainin in DPC micelles (A) (PDB ID 2MAG), LFampinB in SDS (B) and tritrpticin in SDS (C) (PDB ID 1D6X). The top panel shows the overlay of the 10 lowest energy structures of each of the peptides. The next panel shows a representative ribbon diagram of each antimicrobial peptide with the side chains also indicated. The bottom two panels depict the surface charge distribution of each of the antimicrobial peptides with the positive (black), negative (gray) and uncharged (white) surfaces highlighted to emphasize the amphipathic nature of these peptides. The third panel shows the peptide in the same orientation as panels 1 and 2 while the bottom panel shows the opposite side of the peptide.

the RRWQWR-NH_2 hexapeptide was solved in the presence of SDS micelles and it was found to form a stable amphipathic structure with no regular secondary structure elements.[62] More recently, the solution structure of an 11-mer derivative of LFcinB, with similar antimicrobial activity to LFcinB but lower hemolytic activity,[100] was solved in the presence of SDS micelles.[63] In addition, the solution structure of a 15-mer of LFcinB bound to SDS micelles was determined.[101] In agreement with the hexamer structure, the 11-mer and the 15-mer adopted amphipathic structures in the presence of SDS micelles but there were no regular secondary structure elements. These results suggest that perhaps LFcinB undergoes a significant conformational change upon binding to the surface of a bacterial cell and that the β-sheet structure is disrupted upon interacting with the target cell. This hypothesis is supported by the fact that a reduced form of LFcinB, lacking the disulfide bond, maintains most of its antimicrobial activity[102] suggesting that the β-sheet structure observed in aqueous solution is not directly related to the antimicrobial mode of action of this peptide. An alternative explanation is that LFcinB has an intracellular target within the bacterial cell[103] and the structures observed in the short derivatives are intermediates involved in the internalization of this peptide.

Support for this mechanism is available from another study of peptides derived from human lysozyme. One of these peptides, a heptameric peptide, RAWVAWR-NH_2, had significant antimicrobial activity[104] and adopted a stable conformation in SDS and DPC micelles but demonstrated weak interactions with phospholipids.[105] However, a fluorescein labelled version of this peptide was readily taken up by *E. coli* and *S. aureus* cells[105] which indicates that the micelle-bound structure could again be an intermediate involved in membrane translocation. Similar results have been seen from studies of hexameric antimicrobial peptides discovered through combinatorial chemistry.[10] One of these peptides, Ac-FRWWHR-NH_2 (combi-2), preferentially interacted with anionic phospholipids and a well-defined amphipathic structure emerged in the presence of SDS and DPC micelles.[106] However, combi-2 and a similar peptide, combi-1 (Ac-RRWWRF-NH_2), did not induce leakage from vesicles and confocal microscopy of fluorescein-labelled versions of these two peptides indicated that both are internalized into bacterial cells.[106] All the studies of LFcinB and shorter derivatives performed to date demonstrated that these peptides do not induce leakage from LUVs[62,63,101] and considering the similarities between the sequence of the antimicrobial active centre of LFcinB and the sequences of the lysozyme and combinatorial Trp- and Arg-rich peptides, the internalization of LFcinB into the bacterial cell[107] is an attractive explanation for the observed antimicrobial activity.

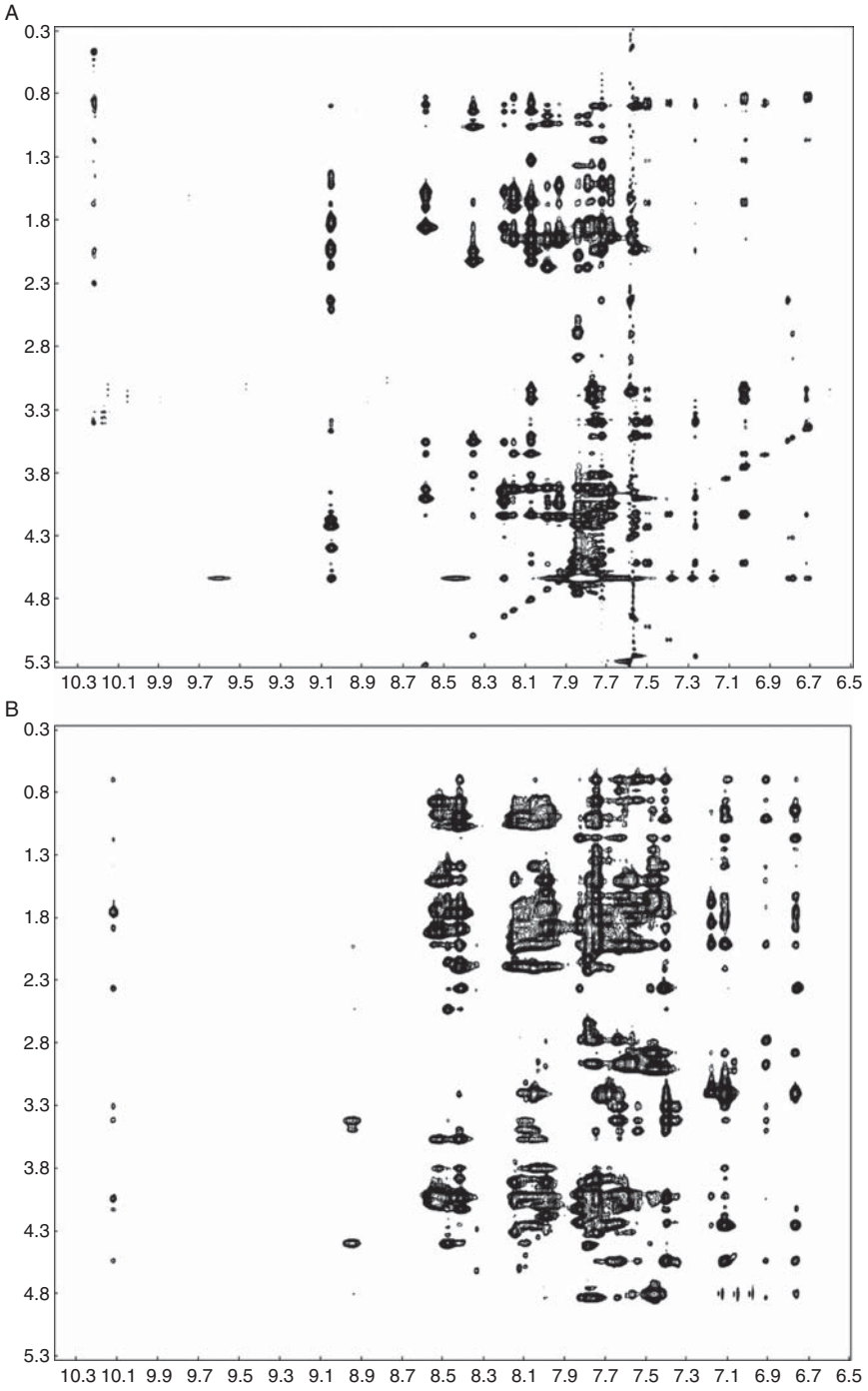
The iron binding protein lactoferrin has proven to be a rich source of antimicrobial peptides as evidenced by the recent discovery of yet another novel antimicrobial peptide, lactoferrampin (LFamp), in the cationic N-terminal lobe of bovine lactoferrin.[108] We have determined the solution structure of LFamp in the presence of SDS and DPC micelles and characterized its interaction with phospholipids using other biochemical techniques.[60] The peptide adopted a similar conformation in both detergent micelles with an amphipathic N-terminal helix stretching from residues 1–12 while the positively charged C-terminus remained

relatively flexible. Tryptophan fluorescence and quenching results suggest that the hydrophobic patch bordered by Trp1 and Phe11 inserts into the hydrophobic core of the bilayer and holds lactoferrampin at the surface of a phospholipid bilayer. However, the positive charges at the C-terminus of the peptide are absolutely essential for the activity of this peptide.[109] Thus, the structure of lactoferrampin bound to micelles reveals a unique feature compared to other antimicrobial peptides in which there is a clear separation between the residues responsible for membrane insertion and the residues that are implicated in attracting the peptide to the surface of a bacterial cell through electrostatic interactions.

7. ORGANIC MEMBRANE MIMETIC SOLVENTS

Many organic solvents promote the formation of secondary structure in peptides and they can usually induce the formation of amphipathic helices that are observed for many micelle bound structures of antimicrobial peptides. Circular dichroism experiments are commonly performed when studying antimicrobial peptides in aqueous solution, in organic solvent and in the presence of SDS and DPC micelles to demonstrate that the helical content of an antimicrobial peptide that is observed in the organic solvent and the micellar environment is identical. The rationale behind using organic solvents for structure determination, especially in cases where amphipathic helices are formed, is that this structure can be extended to the membrane bound structure of an antimicrobial peptide bound to a bacterial cell. An added advantage of using organic membrane mimetics is that the line broadening that can occur in spectra of peptides bound to detergent micelles is not seen for organic solvents. For instance, a 2D NOESY of the α-helical portion of the mammalian chemokine MIP-3α produces better defined cross-peaks in a membrane mimetic organic solvent compared to SDS micelles (Figure 7).

Trifluoroethanol (TFE) is a helix inducing solvent that promotes the formation of intramolecular hydrogen bonds and it has been frequently used to determine the structure of antimicrobial peptides. For example, NMR studies of the polypeptide hormone glucagon demonstrated that the conformation of this peptide was similar in DPC micelles[110] compared to a sample prepared with a TFE–water mixture.[111] These results suggested that TFE could serve as a solvent that mimics the properties of a bilayer. One of the earliest applications of this solvent in NMR studies of antimicrobial peptides was published in 1988 when conformational changes of magainin 2 were examined in aqueous solution and then after addition of TFE.[112] In this case, important inter-residue NOEs, indicative of the induction of helical structure, were only observed after the addition of TFE. The first structure of an antimicrobial peptide in aqueous TFE (50% w/v) was reported in 1996 when the solution structure of buforin II was determined in a 1:1 mixture of TFE–water.[113] Since this time, solvent mixtures of TFE–water have been used to examine the solution structures of caerin 1.1,[114] gaegurin 4[115] and histatin 3 and 5,[59] for example.


A
0.3
0.8
1.3
1.8
2.3
2.8
3.3
3.8
4.3
4.8
5.3
10.3 10.1 9.9 9.7 9.5 9.3 9.1 8.9 8.7 8.5 8.3 8.1 7.9 7.7 7.5 7.3 7.1 6.9 6.7 6.5
B
0.3
0.8
1.3
1.8
2.3
2.8
3.3
3.8
4.3
4.8
5.3
10.3 10.1 9.9 9.7 9.5 9.3 9.1 8.9 8.7 8.5 8.3 8.1 7.9 7.7 7.5 7.3 7.1 6.9 6.7 6.5

Another organic solvent mixture to study antimicrobial peptides was utilized by our group while examining the solution structure of human lactoferricin (LFcinH). The 49 residue LFcinH is a weak antimicrobial peptide but it binds strongly to bacterial endotoxin. It can be released from the 80 kDa human iron binding protein lactoferrin following pepsin hydrolysis in the stomach.[116] Unlike the previous studies mentioned, well resolved spectra of LFcinH could not be obtained in the presence of SDS or DPC micelles and only a poorly defined structure based on a limited number of observed NOEs could be determined in aqueous solution.[117] This is in stark contrast to the 25 residue bovine peptide, LFcinB, which forms a relatively stable β-sheet in low salt aqueous solution.[54] Molecular dynamics studies have also indicated that LFcinB does not even require the presence of a bilayer to fold into the amphipathic β-sheet conformation.[118] As a result, the solution structure of LFcinH was determined in a membrane mimetic solvent comprised of a 4:4:1 mixture of $CDCl_3$-methanol-d3-H_2O.[117] The chloroform–methanol–water system was previously used in NMR studies of the transmembrane H^+-transporting subunit c of the F_1F_0 ATP synthase complex[119] and the *E. coli* multidrug resistance protein, EmrE.[120] Both of these proteins were found to have similar biochemical properties in the chloroform–methanol–water mixture compared to a bilayer environment which serves as evidence that this membrane mimetic could be used as a substitute for the membrane environment in NMR studies. Other organic solvents that have been used for structural characterization using NMR include methanol,[121] dimethyl sulfoxide[59] and aqueous suspensions of hexafluoroisopropanol.[122,123] Examples of structures of some antimicrobial determined in organic solvents are shown in Figure 8.

It is also advantageous to examine the temperature coefficients of amide protons using NMR spectroscopy to determine which residues are directly involved in hydrogen bond formation. This additional structural information may be useful for determining which residues are found in regular secondary structure elements and which ones are highly mobile. This experiment was applied, for example, to two cationic helical peptides in the presence of 50% TFE.[124] In this particular case, the temperature coefficients indicated a lack of intramolecular hydrogen bonds but slow hydrogen isotope exchange suggested that the amide-proton chemical shifts were dominated by bifurcated hydrogen-bonds between carbonyl groups and TFE. Amide temperature dependence measurements are not limited to studies of peptides in organic solvent membrane mimetics. Amide exchange rates were used to study the structures of the defensins in aqueous solution and it was determined that there was a good correlation between the temperature coefficients and the NH exchange rates; those residues that were involved in intramolecular hydrogen bonding displayed lower

Figure 7 NOESY spectra of the MIP-3α peptide showing the fingerprint region acquired in membrane mimetic solvent (4:4:1 methanol:$CDCl_3$:H_2O) (A) and SDS micelles (B). The increased effective molecular weight of the peptide in SDS micelles clearly leads to broader linewidths and therefore greater spectral overlap compared to the membrane mimetic solvent. The mixing times for the organic solvent mixture and micelle spectra are 250 and 120 ms, respectively. (Unpublished results from Nguyen, Chan, Boszhard, De Boer, Zaat and Vogel).

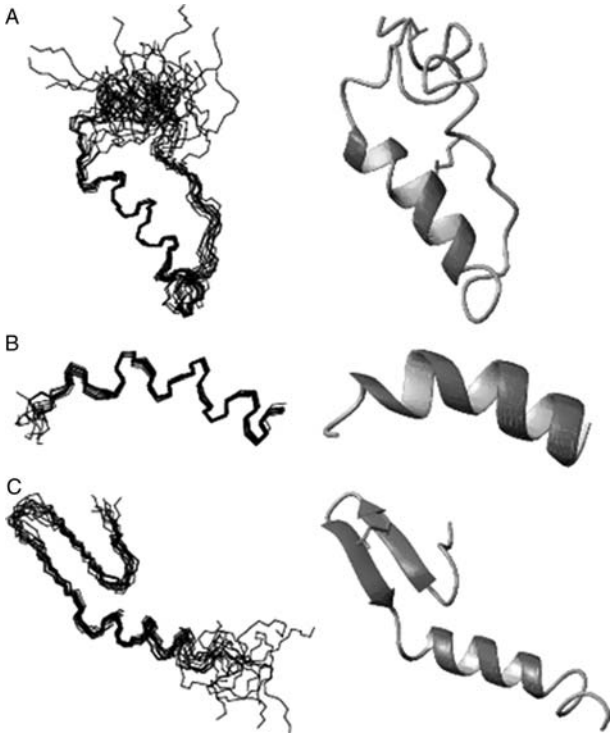


Figure 8 Structures of selected antimicrobial peptides determined in organic membrane mimetics. LFcinH in 4:4:1 $CDCl_3$:methanol-d3:H_2O (A) (PDB ID 1Z6V). Ovispirin in 50% H_2O:40% TFE-d3:10% D_2O (B) (PDB ID 1HU5). A type IIa bacteriocin, Leucocin A in 90% TFE (C) (PDB ID 2LEU). The backbone overlay of the 10 lowest energy structures are shown on the left while the ribbon trace is shown on the right.

temperature coefficients and slow amide exchange rates.[125] Similar NMR studies with small proteins provide some further guidelines on the proper interpretation of such experiments.[126]

It is generally accepted that the solution structures of antimicrobial peptides in organic solvents are similar to those that are determined in the presence of detergent micelles. For example, ^{1}H NMR experiments of magainin revealed that the peptide adopts an α-helical conformation in SDS, DPC and a TFE/water mixture.[127] However, it is important to recognize that these NMR solvents could have a large impact on the 3D structures. Some membrane active peptides adopt different conformations in TFE compared to detergent micelles. This is the case with a nine residue peptide derived from the cytotoxic α-sarcin protein produced by *Aspergillus giganteus*. This peptide adopts a *β*-sheet conformation in the presence of SDS micelles and lipid vesicles but it forms an α-helix in TFE.[128] This example underscores the need to ascertain through CD spectroscopy experiments[129] that the structure formed in organic solvents is the same as that in the presence of detergents or lipids.

8. ORIENTATION IN DETERGENT MICELLES

NOESY experiments are not limited to examining the structures of antimicrobial peptides. They can also be used to examine intermolecular interactions between antimicrobial peptides and detergent or lipid species. In this case, direct NOEs between the aromatic residues of the peptide and protons of the detergent molecules can give valuable information about the depth of insertion of the peptide and the orientation of the peptide in a micelle. Nisin, a polycyclic and cationic antimicrobial peptide from the lantibiotic class of antibiotic peptides, was shown to make direct contacts with micellar protons.[130] Van den Hooven and co-workers determined the chemical shift values of the protons in SDS and DPC micelles and subsequently identified the intramolecular NOEs between nisin and the micelles. This was accomplished by comparing the NOESY spectra of nisin in the presence of perdeuterated detergents to a sample containing peptide and protonated detergent.[130] Intermolecular NOEs were observed downfield from the water signal indicating that nisin interacts at the surface of the micelle. We have employed a similar method to examine the intermolecular interactions between the membrane proximal region of the HIV1 glycoprotein, gp41 and DPC micelles.[131] This peptide forms an amphipathic structure in the presence of the zwitterionic micelles and it is thought that its micellar interactions mimic those of antimicrobial peptides because of the unusually high Trp content in this short sequence. The charged and hydrophobic surfaces of the peptide are separated by a collar of four tryptophan residues and one tyrosine residue that are distributed throughout the peptide. NOE cross-peaks were observed between the aromatic tryptophan residues and the head group protons of the DPC micelles, indicating that the peptide resides in the interfacial region of the lipid bilayer.

The interactions between Mastoparan B, a cationic antimicrobial peptide isolated from the venom of the *Vespa basalis* hornet, and various detergent micelles and lipid bicelles have also been examined in this manner.[132] Mastoparan B folds into an amphipathic α-helix in the presence of lipids and NOESY experiments performed with undeuterated SDS indicate that the indole moiety of Trp9 is in close proximity to the methylene protons of the SDS molecules. The interactions observed between DPC micelles and mastoparan B were considered to be weaker. These results indicate that the preferential insertion of mastoparan B into negatively charged micelles, with Trp9 anchoring the peptide in the membrane, may be analogous to the antimicrobial mechanism of this peptide *in vivo*. Intermolecular NOE cross-peaks have also been used to examine the interaction between antimicrobial peptides derived from human LL-37 and the micelle forming phospholipid, D8PG. Three resonances unique to the D8PG lipid molecules were used to determine intermolecular NOEs between the peptide and the lipid molecules.[56]

Despite the relative simplicity of this type of experiment, intermolecular NOEs between antimicrobial peptides and detergents are rarely used in solution-NMR experiments. This is especially curious because this interaction is the most important when we consider that many antimicrobial peptides have been shown to exert their antimicrobial activities through a direct interaction with the phospholipid bilayer of a bacterial cell. The reason for this could be that it is difficult to

distinguish between intramolecular NOEs of the lipid and the peptide molecules individually and intermolecular NOEs between the two species. All of the proton chemical shifts of SDS and DPC are found upfield of the water signal[130] and will interfere with many of the aliphatic proton resonances from the peptide. Additionally, there is significant overlap between many of the proton resonances originating from the acyl chains of the detergent molecules[130] which make it difficult to assign a cross-peak to a specific methylene group. This means that only NOEs found downfield of the water signal are useful for determining intermolecular contacts. Fortunately, the aromatic amino acids tryptophan and phenylalanine have been shown to anchor antimicrobial peptides to phospholipid bilayers[12,72,133] and the proton chemical shifts of these side chains occur above the water signal. It is also important to consider that detergent concentrations in NMR samples are much higher than the antimicrobial peptide concentration. Knowing that the intensity of an NMR signal is directly proportional to the concentration of the molecule, these detergent signals could potentially abolish nearby NOEs simply because of their signal strength.

The orientation of antimicrobial peptide with respect to the membrane can be determined through means other than direct intermolecular NOEs. Paramagnetic spin labels can be incorporated into micelles with probes at different positions on the acyl chains that can be used to map the position of the peptide in the micelles. One of the most popular paramagnetic spin labels is doxyl-labelled stearic acids. Doxyl groups have an unpaired electron and this can be placed at various positions of an acyl chain. The doxyl groups induce paramagnetic relaxation in nearby protons so if the antimicrobial peptide is in close proximity to the spin-label, then there will be a broadening of the proton signals close to the doxyl probe.[98] Two popular relaxation probes that can be incorporated into detergent micelles are 5- and 16-doxylstearate because the different locations of the nitroxide moiety allow researchers to probe peptide interactions at the surface and the centre of the micelle respectively.[134] For instance, 5-doxylstearic acid and 16-doxylstearic acid were used to probe nisin interactions with the head groups of DPC and SDS micelles.[130] The presence of the 5-doxyl spin label led to a broadening of signals in the NH-Hα region of the TOCSY and demonstrated that nisin binds at the surface of the micelle. 16-doxylstearic acid broadened many of the same signals that the 5-doxyl species affected but this effect was weaker and 10–15% of the resonances affected by the 5-doxyl stearic acid were unaffected by the deeper 16-doxyl probe. This approach has also been used to map the surface orientation of the plant cyclotide, kalata B1, on the surface of DPC micelles.[93] Broadening of the kalata B1 proton signals was observed in the NOESY spectra for both the 5- and 16-doxylstearate probes but the 5-doxylstearate molecule attenuated cross-peak intensities more than the deeper 16-doxylstearate probe. Based on these results, two parts of the kalata B1 peptide (Gly26-Val29-Cysl-Gly7 and Ser18-Cys22) were identified as being immersed into the micelle and form the interaction surface for the peptide-micelle complex. It should be noted that stearic acid acyl chains are much longer than the hydrophobic tails of SDS and DPC micelles and as a result it is difficult to determine the exact depth of the doxyl group within the micelle. This is an inherent limitation of this technique because there is no way

of knowing the exact location of the peptide with respect to the doxyl group depending on where the stearic acid chain is found in the micelle.

A newer solution-state method to examine the orientation of a peptide in the presence of detergent micelles uses paramagnetic relaxation induced by a water-soluble relaxation agent gadolinium-diethylenetriamine pentaacetic acid-bis-methylamide Gd(DTPA-BMA).[135] This was first used to examine the surface residues of soluble proteins[136] and is analogous to the doxyl labelled spin probes described above; however, the stronger paramagnetic relaxation effects will be seen in amino acid residues that are protruding into the aqueous solution. To adequately define the orientation of an α-helix in a membrane mimetic, two parameters need to be established: the tilt angle, which is the angle between the peptide and the surface of the micelle, and the azimuthal angle, which describes the rotation of the helix and consequently dictates which amino acid side chains are oriented into the hydrophobic core of the micelles and which point out towards the aqueous solution. As a proof of concept, the orientation of a hybrid antimicrobial peptide, CM15, was examined using this technique. The CM15 peptide contains residues 1–7 of cecropin A and residues 2–9 of mellitin and displays antimicrobial activity *in vitro*.[137] The solution structure of CM15 bound to DPC-d38 micelles was determined in a similar fashion to the examples described earlier. To define the tilt and azimuthal angle, a titration was performed by adding Gd(DTPA-BMA) into the NMR sample containing CM15 and DPC micelles. After each addition of Gd(DTPA-BMA), a saturation-recovery 2D-NOESY was recorded and proton T1 relaxation times were calculated. The relaxation times were plotted as a function of Gd(DTPA-BMA) concentration and the slope of the resulting straight line is equal to the paramagnetic relaxation enhancement (PRE) of that nucleus (Figure 9A). A plot of the PRE of the Hα nuclei of CM15 as a function of residue number results in a sinusoidal curve which can be used to calculate the tilt angle and the azimuthal angle.[135] Another interesting application of this experiment involves mapping the PRE values onto the structure of the peptide bound to the detergent micelles. If the PREs are separated based on their intensities, then a visual representation of the orientation of the peptide with respect to the micelle surface can also be obtained (Figure 9B).

One of the most elegant techniques that has been developed to probe the depth of insertion of an antimicrobial peptide involves exploiting the paramagnetic relaxation induced by molecular oxygen. The premise of this technique is based on the fact that O_2 preferentially partitions to the centre of a bilayer at high pressures.[138] As a result, the paramagnetic relaxation observed in this case will be strongest for nuclei at the centre of the bilayer and weakest for those close to the head groups. The proof of concept was demonstrated using bicelles containing; ^{19}F labelled cholesterol, a semiperfluorinated test compound ($CF_3(CF_2)_5C_2H_4$-O-maltose (TFOM)) and a fluorinated antimicrobial peptide, indolicidin.[138] ^{19}F labels were chosen because the chemical shift range of ^{19}F spectra is large and multiple ^{19}F nuclei can be resolved in one spectrum. Additionally, ^{19}F nuclei are particularly sensitive to nearby paramagnetic species. Two fluorinated cholesterol species were synthesised, one with a ^{19}F atom attached to C6 which would be located close to the head group region of a bilayer. The other cholesterol species was fluorinated at position 25 which is found at

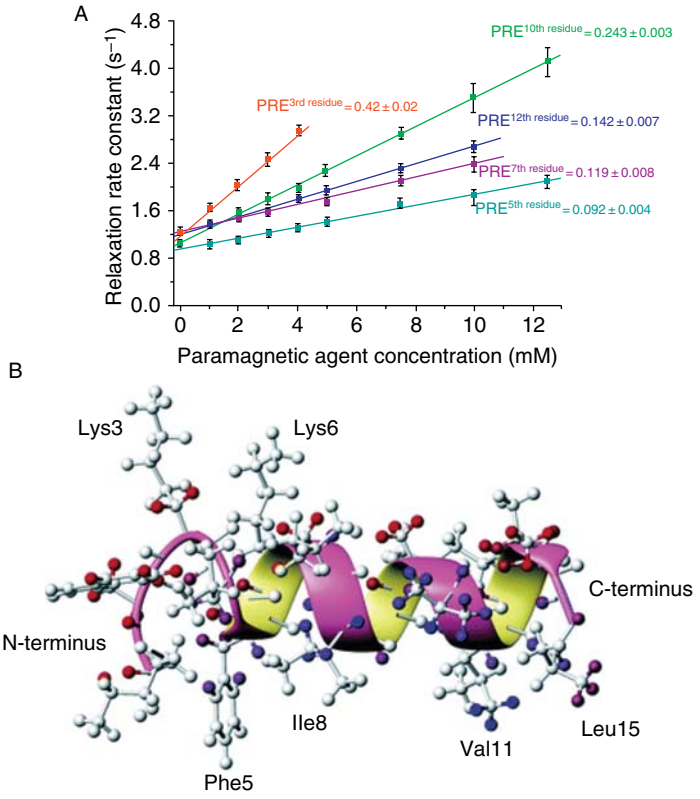


Figure 9 Paramagnetic relaxation rates (A) resulting from the titration of the paramagnetic relaxation agent Gd(DTPA-BMA). The resulting PREs can be mapped onto the surface of the micelle bound structure (B) of CM15 (residues with measureable PREs are shown in gray). Color versions of these figures are available from Ref. 135.

the centre of the bilayer. Additionally, the fluorinated acyl chain of TFOM gives multiple ^{19}F nuclei at various depths throughout the bilayer. Altogether, these molecules can be used to compare the paramagnetic relaxation rates induced by molecular oxygen throughout the bilayer (Figure 10). Two fluorinated indolicidin molecules were made with Trp-8 and Trp-9 being replaced by 5-fluorotryptophan and the resulting paramagnetic relaxation rates placed Trp-8 close to the centre of the bilayer while Trp-9 resides closer to the head groups.[138] These results help explain the solution structure of indolicidin which was found to be roughly banana shaped in the presence of SDS and DPC micelles[139] but whose orientation in a phospholipid bilayer could not be established based on the structure alone.

9. STRUCTURE DETERMINATION OF LARGER COMPLEXES

An acknowledged limitation of using micelles to examine the solution structure of antimicrobial peptides is that the curved surface of a micelle differs significantly from a planar phospholipid bilayer. Additionally, the structure of the detergent

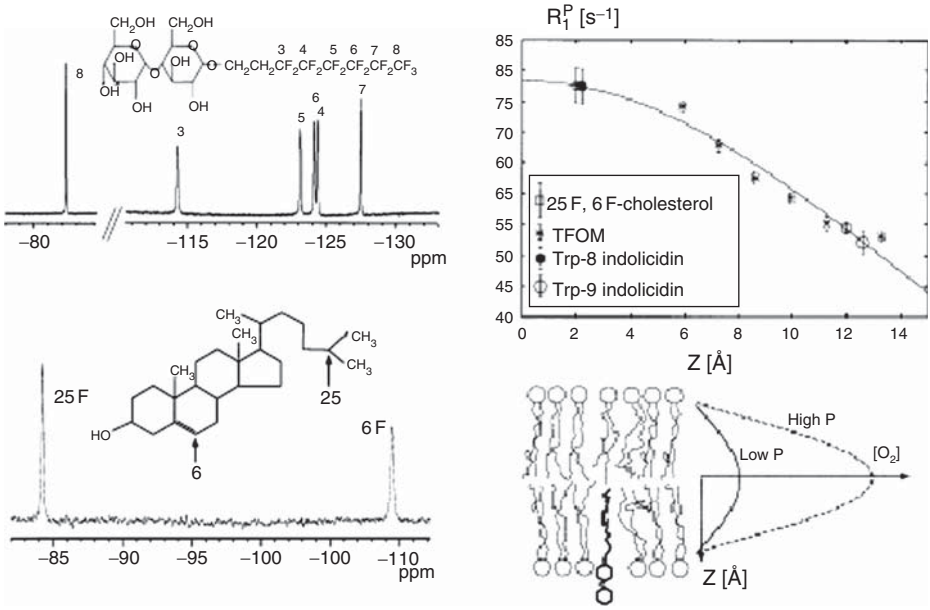


Figure 10 Paramagnetic relaxation rate as a function of membrane depth. On the left is the 1D ^{19}F spectra of TFOM and fluorinated cholesterol. Note the comparison between fluorinated indolicidin and how the depth of the Trp residue can be mapped according to the paramagnetic relaxation rate (R_1^P) observed from the cholesterol and TFOM species. For the panels on the right, Z = 0 Å defines the center of the bilayer. (Adapted from Ref. 138, with permission).

molecules, both in their head groups and their acyl chains differ significantly from the structure of a normal phospholipid (Figure 4) and the number of head groups available for detergent molecules pales in comparison to the variety of head groups that are available for phospholipid species. There are a few NMR studies that have used bilayer structures to examine the solution structure of antimicrobial peptides which might be considered to be more biologically relevant than peptides bound to detergent micelles.

The use of lipid bicelles is potentially useful for examining antimicrobial peptide structures using solution-state NMR techniques. Bicelles are formed when short chain lipid are mixed with long chain lipids to form a flat disk-like bilayer.[140] The short chain lipids form the edge of the bilayer to accommodate the large amount of positive curvature while the long chain lipids make up the planar surface of the bicelle. The size of bicelles can be altered by changing the ratio of the lipids involved[141] and small bicelles ranging 25–75 Å are small enough to be used for solution-state NMR experiments. For example, the solution structure of indolicidin was determined in DPC and SDS micelles[139] and when the ^{1}H NMR spectra of indolicidin in DPC micelles was compared to the spectra of indolicidin in the presence of various fast tumbling bicelles, all of them gave similar spectra with improved peak dispersion compared to the peptide in an aqueous environment.[141] Bicelles have also been used to examine Mastoparan X (MPX), a

membrane active peptide found in wasp venom. NOESY and chemical shift analysis indicated that MPX folded into a well-defined amphipathic helix in zwitterionic and negatively charged bicelles.[142] These results highlight the potential use of bicelles for structural determination of antimicrobial peptides; however, the use of this technique has, so far, been sparse in antimicrobial peptide research.

The transferred nuclear Overhauser effect (TRNOE) experiment[143,144] can sometimes be used to study the structure of antimicrobial peptides bound to the surface of vesicles. The TRNOE experiment requires relatively fast exchange between a small ligand bound to a much larger receptor, in this case an antimicrobial peptide binding to the surface of a micelle. The cross-relaxation of the bound peptide is much larger than the peptide freely floating in solution and these resonances are transferred to the unbound peptide through chemical exchange. The structure of a magainin 2 analog in the presence of perdeuterated DLPC vesicles was studied using the TRNOE experiment.[145] Interestingly, this magainin 2 analog forms a short antiparallel coiled-coil dimer which is different from the monomeric structures of magainin 2 seen in micelles and TFE[127] but it does support previous observations that magainin 2 can form dimers in PG containing membranes.[146] Vesicle bound structures of other membrane active peptides, like MPX[147] and melittin,[148] have also been reported using this technique. The requirement for fast-exchange in the TRNOE experiment limits the application of this experiment to antimicrobial peptides that bind relatively weakly to biologically relevant lipid environments.

The use of RDCs should also prove to be a useful technique for determining the solution structure of antimicrobial peptides bound to bicelles. RDCs arise from two spins in a protein, if the protein is weakly aligned in an NMR sample. This leads to an incomplete averaging of magnetic interactions and the residual signal from this phenomenon can be used as inputs in structural calculations. Structural characterization of an antimicrobial peptide using RDCs has yet to be reported but the potential application of this powerful technique was demonstrated with a membrane associated fragment of the HIV-1 envelope glycoprotein gp41.[149] In this study, the structure of gp41[282–304] bound to both dihexanoyl phosphatidylcholine (DHPC) micelles and bicelles made of DHPC and dimyristoyl phosphatidylcholine (DMPC) was determined. To obtain aligned micelles and bicelles, the samples were soaked in cylindrically shaped 6% polyacrylamide gel and then compressed to fit into an NMR tube. This is known as the strain-induced alignment in a gel (SAG) method[150,151] which is essential when studying bicelles and micelles because other alignment methods, such as liquid crystalline media,[152] are not compatible with detergents and lipids. The micelle bound structure displayed a high amount of curvature with the hydrophobic residues oriented on the inside while the bicelle bound structure was a nearly straight helix.[149] This study not only highlights the limitations of using micelles for structural determination of surface associated peptides but it also outlines the usefulness of this technique for future structural characterization of antimicrobial peptides in a lipid environment that more closely resembles the topology of a biological membrane.

It should be noted that RDC structure determination requires that the antimicrobial peptide in question is isotopically labelled with ^{15}N and ^{13}C nuclei.

Obtaining isotopically labelled peptides and proteins is usually done through recombinant expression in bacterial cells. Consequently, this may prove to be an indirect limitation for the application of RDCs because the recombinant expression of antimicrobial peptides is not a trivial process (see below).

10. ENDOTOXIN BINDING BY ANTIMICROBIAL PEPTIDES

Bacterial lipopolysaccharide (LPS) is an endotoxin found as a major component of the outer membranes of gram-negative bacteria. By binding to the Toll-like receptors on leukocytes, it induces an innate immune response through the release of inflammatory cytokines.[153] An overstimulation of the immune system by LPS induces a serious medical condition known as sepsis. Various antimicrobial peptides, including lactoferricin,[154,155] fowlicidin-3[156] and polymyxin B,[157] have been shown to bind LPS. It is unclear whether this interaction is directly related to the mechanism of antimicrobial activity of these peptides or whether these peptides are responsible for scavenging free LPS to protect against the severe immune response. NMR spectroscopy has been used to structurally characterise antimicrobial peptides bound to LPS. Future application of this technology might lead to the development of new endotoxin neutralising molecules with enhanced binding affinity for LPS that could be used to treat sepsis and septic shock.

The structure of polymyxin B bound to LPS was determined[158] using the TRNOE experiment described earlier. This strategy was also used to determine the LPS-bound structure of a LFcinH-based peptide.[159] In both cases, the conformation of the antimicrobial peptide bound to LPS was determined from the observed NOEs but this approach did not provide any information about the structure of the LPS molecule itself. Instead, the structures of the LPS bound peptides were used as inputs for molecular docking calculations, which resolved the binding interface between the lipid A component of LPS and the peptide and they also indicated which residues were directly involved in the binding to the endotoxin. A more recent study examined the interaction between melittin and LPS and again the TRNOE experiment was used to determine the structure of the peptide bound to the endotoxin.[160] However, instead of employing a molecular docking strategy, saturation transfer difference (STD) NMR was used to determine which residues of melittin were in direct contact with LPS.

11. RECOMBINANT EXPRESSION OF ANTIMICROBIAL PEPTIDES

For the most part, the antimicrobial peptides used in NMR studies are either purified from natural sources or they are chemically synthesised. For structural characterisation of these peptides by NMR, usually milligram quantities of highly purified peptides are needed. When purified from natural sources, a large amount of biological tissue is required to extract a sufficient amount of peptide. In the case of synthetic peptides, it is easy to obtain highly pure peptides but the price of peptide synthesis is steep for longer peptides and it increases dramatically if the

peptide requires special modifications. As mentioned above, most NMR studies of antimicrobial peptides have relied on proton NMR; however, other NMR active nuclei, such as ^{13}C and ^{15}N and ^{2}H, can provide important structural information in NMR studies through the use of well established triple resonance NMR experiments[161]. 3D ^{15}N or ^{13}C-separated NOESY spectra[162] is also used to separate the NOE cross-peaks into different planes which helps to alleviate most of the spectral overlap observed in 2D NOESY spectra. Unfortunately, the natural abundance of these atoms is inherently low and consequently their detection is insensitive. To overcome this limitation, recombinant expression of antimicrobial peptides has to be performed so that >99% of ^{13}C, ^{15}N and ^{2}H is incorporated into the peptide.

Another potential benefit of developing recombinant expression systems for antimicrobial peptides is that many studies focus on amino acid substitutions in the native peptide sequence to try and narrow down the residues that are essential for activity.[101,109,163] If a recombinant expression system for an antimicrobial peptide were available, then simple site-directed mutagenesis can be performed to generate the desired mutations at a fraction of the cost required to synthesise all of the individual peptides.

Recombinant expression of antimicrobial peptides in bacterial cells poses a number of inherent challenges. The most obvious difficulty arises from the simple fact that it is difficult to engineer a bacterial cell to produce a peptide that has evolved to kill it. Additionally, the difficulty of expressing short peptides is compounded by their tendency to aggregate with each other and complicated by the activity of the internal machinery of the bacterial cell that readily degrades small unfolded peptides in the cytoplasm. Despite these challenges, recombinant expression of some antimicrobial peptides has been accomplished through the use of various expression systems and fusion protein constructs.

One of the simplest methods for recombinant expression of proteins involves the addition of a His6-tag at the N or C terminus of the protein which facilitates purification on a nickel-sepharose column. Following purification, the His6 tag can then be removed from the antimicrobial peptide through CNBr digestion which will cut after a methionine residue that can also be engineered into the peptide sequence. Reverse-phase chromatography is then applied to obtain the purified antimicrobial peptide. Recombinant forms of cryptidin 4 have been expressed as N-terminal His6-tag fusion proteins[164] which facilitated the purification of these peptides. This method has also been used to obtain sufficient quantities of defensin peptides and mutants for NMR structure determination.[165] Additionally, a defensin from the oyster *Crassostrea gigas*, named Cg-Def, was expressed in *E. coli* cells and purified as an N-terminal His6-tagged fusion protein. The solution structure of this recombinant peptide was then solved in a 95:5 mixture of H_2O:D_2O.[166]

The recombinant expression of antimicrobial peptides linked to larger proteins that facilitate easier expression and purification has proven to be useful for obtaining large quantities of these peptides for NMR studies. The large fusion proteins are thought to stabilise the small peptides by allowing them to fold back onto the large protein molecule which protects them from bacterial proteases.

Additionally, they might mask the antimicrobial activity of these peptides by inhibiting the interaction between the antimicrobial peptide and the membrane or they block binding of the peptide to its intracellular target. These fusion expression systems still require significant optimisation but we will highlight here some examples of successful recombinant antimicrobial peptides. For example, this strategy was used to recombinantly express the antimicrobial peptide arenicin-2.[57] Arenicin-2 was placed at the N-terminus of thioredoxin and after purification the two were separated by a CNBr cleavage site. An octahistidine tag was engineered at the C-terminus of the thioredoxin protein to facilitate purification through a Ni-NTA column. Following CNBr digestion, the recombinant arenicin-2 was purified using reverse-phase chromatography and the solution structure as determined using conventional ^{1}H NMR techniques. Arenicin-2 was found to form a twisted β-sheet in water but it lacked a significant amount of amphipathicity despite the strong cationic nature of the peptide.[57]

Recombinant LL-37 was successfully expressed and purified from *E. coli* using the Glutathionine S-transferase (GST) fusion system.[167] This fusion system takes advantage of the binding ability of the GST tag to bind to a glutathione column to purify the fusion protein. The LL-37 peptide was then enzymatically cleaved from the fusion protein using factor Xa followed by reverse-phase HPLC to purify the peptide in large quantities. In the case of the GST-LL-37 fusion protein, the authors describe the production of ^{15}N-labelled LL-37 through the use of $^{15}NH_4Cl$ as the sole nitrogen source but they also acknowledge that the uniform labelling of LL-37 with other NMR active nuclei like ^{13}C, ^{19}F or ^{2}H are important for further NMR studies to examine the intricacies of the peptide–lipid interactions responsible for the activity of this peptide.

The ubiquitin-fusion protein system is another popular expression system because the ubiquitin protein can be easily cleaved by ubiquitin specific proteases from the fusion protein after the carboxy-terminal glycine residue.[168] Recently, the ubiquitin system was shown to be an efficient fusion partner for the recombinant expression of an antimicrobial peptide, piscidin.[169] Piscidins are 22 residue antimicrobial peptides found in the mast cells of fish with a highly conserved N-terminus rich in His and Phe residues.[170] Moon and co-workers set out to recombinantly express and purify piscidin at high concentrations with ^{13}C and/or ^{15}N labelled nuclei incorporated into the peptide. Their system contained a ubiquitin protein with a His6-tag at the N-terminus and the piscidin peptide attached to the C-terminus.[169] The His-tag facilitated the purification of the ubiquitin–piscidin fusion protein through the use of a metal chelating column. The purified fusion protein was then digested with yeast ubiquitin hydrolase to release the piscidin peptide from the ubiquitin protein. The recombinant ^{15}N-piscidin peptide was then purified using reverse-phase chromatography and the purity was assessed using MALDI-TOF. This study outlines the effectiveness of using ubiquitin as a fusion partner for the recombinant expression of antimicrobial peptides and demonstrates that it is feasible to generate isotopically labelled peptides for use in NMR studies.

Finally, maltose binding protein (MBP) is part of another fusion protein expression system that has proven to be useful for the recombinant expression

of antimicrobial peptides. It was used to mask the toxic effects of the antimicrobial peptide piscidin when it was expressed in *E. coli* as a fusion protein attached to the C-terminus of MBP.[171] Recently, the recombinant expression of a novel bacteriocin, lactococcin, was reported. In this study, overexpression of native recombinant lactococcin peptide led to the accumulation of this peptide in inclusion bodies while a lactococcin MBP fusion protein produced a soluble form of the peptide that was easily overexpressed and purified in high quantities.[172] The recombinant expression of antimicrobial peptides fused to MBP appears to overcome the difficulties associated with expressing these peptides on their own and it could also prove to be an efficient system to incorporate NMR active nuclei into the peptide of interest for NMR studies.

Other potential strategies for the recombinant expression of antimicrobial peptides include the use of other fusion protein systems such as the SUMO fusion system that claims to be useful for expressing difficult to express proteins.[173] Additionally, the use of a eukaryotic expression system such as yeast[174] could possibly counteract the membrane activities of the antimicrobial peptide because of the altered membrane composition compared to bacterial cells. This strategy has proven to be effective for the recombinant expression of penaeidins in *S. cerevisiae*.[175] Regardless of the recombinant expression system used, there are numerous obstacles to overcome in obtaining recombinant antimicrobial peptides. The diversity of strategies presented here only underscores that no single expression system appears to be capable of consistently producing antimicrobial peptides of varying amino acid composition. In spite of these limitations, the recombinant expression of antimicrobial peptides represents a significant advancement for research in this field and opens the door for new and more sophisticated NMR experiments in the future.

12. SOLID-STATE NMR OF ANTIMICROBIAL PEPTIDES

The major limitation of solution-state NMR studies is that it is currently not possible to examine the biologically significant interactions that occur between antimicrobial peptides and natural phospholipid bilayers. Yet, the interaction between antimicrobial peptides and the membrane of a bacterial cell is essential for defining its activity and in order to study this interaction by NMR, we must switch to solid-state NMR studies of antimicrobial peptides. Solid-state NMR is perhaps somewhat of a misnomer when discussing this type of experiment, as the name implies that the NMR samples are completely rigid and there is no degree of flexibility or motion that is allowed. However, biological membranes are flexible and dynamic structures and this intrinsic motion needs to be conserved during the experiment.

Solid-state NMR usually involves oriented bilayers which have a specific alignment to the external magnetic field.[18] Detailed 3D structures of antimicrobial peptides are difficult to obtain from solid-state NMR studies but there are numerous experiments that have been described that give great insights into the interactions between antimicrobial peptides and phospholipid bilayers. Solid-state

NMR has been used to examine the orientation of antimicrobial peptides in bilayers and to determine if they are surface bound or have a transmembrane orientation. Information about the depth of insertion into a bilayer and whether the peptide self associates into oligomeric structures can also be obtained. Additionally, solid-state NMR is useful for monitoring if an antimicrobial peptide destabilises the membrane and at what depth the perturbation is most pronounced. Consequently, when combined with the outcome of solution-state NMR experiments described above, it is possible to describe most of the interactions between antimicrobial peptides and lipids.

13. PEPTIDE ORIENTATION IN BILAYERS

Determining the orientation of an antimicrobial peptide in a phospholipid bilayer can give important insights into the mechanism of action used by the peptide to disrupt the bacterial cell. For example, a peptide that is found to orient itself parallel to the bilayer normal may not readily form toroidal pores and likely acts via another mechanism to kill the bacterial cell. Various solid-state NMR studies have been published that examine the orientation of antimicrobial peptides in bilayers.

The first technique to determine the orientation of a peptide with respect to the bilayer involves detecting the easily measurable ^{15}N chemical shift frequency in ^{15}N-labelled peptides. The theory behind this experiment is well understood and the results are easily interpreted by visual inspection of the resulting ^{15}N NMR spectra of the labelled peptide. The magnitudes and orientations of the ^{15}N amide chemical shift tensor are well established and allow one to distinguish between an N–H bond that is parallel or perpendicular to the direction of the applied magnetic field (Figure 11). Because we are working with a solid-state NMR sample that has the bilayer oriented perpendicular to the magnetic field, we can establish if the peptide is transmembrane or surface oriented from the resulting ^{15}N amide chemical-shift powder pattern. The proof of concept for this techniques was demonstrated using the antimicrobial peptide magainin 2 and a peptide from the d subunit of the torpedo acetylcholine receptor named M2.[176] Each of these peptides was chemically synthesised and a site-specific ^{15}N label was incorporated into the peptide. To prepare the samples, both the peptide and the lipid species are co-dissolved in an organic solvent which was then applied to a glass plate where the organic solvent was evaporated. These glass plates were then equilibrated at a specific humidity and stacked together to give oriented bilayers. The results of this experiment demonstrate that the orientation of the peptide with respect to the bilayer dictates the chemical shift observed from the ^{15}N label in the peptide. ^{15}N-Ala15 magainin 2 in 3:1 POPC/POPG produced a single-line resonance at approximately 30 ppm corresponding to a perpendicular orientation with respect to B_o and consequently parallel to the bilayer surface.[176] ^{15}N-Ala12 M2 peptide in 4:1 DMPC/DMPG gave a single-line resonance around 190 ppm which is indicative of a parallel orientation with respect to B_o implying that this peptide adopts a transmembrane configuration in a bilayer.[176]

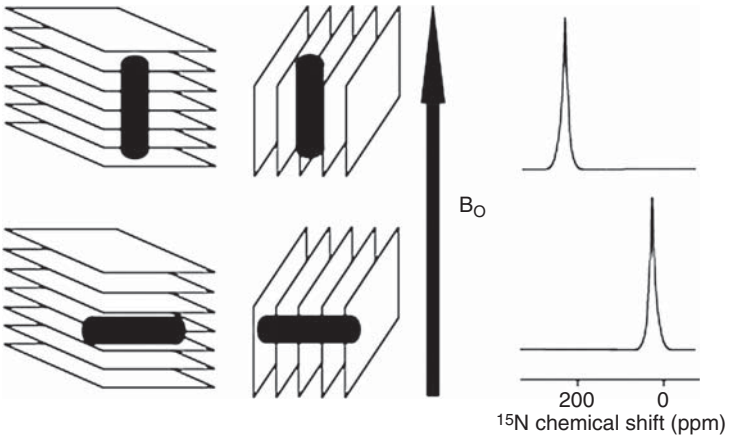


Figure 11 Using ^{15}N to examine the orientation of peptides in phospholipid bilayers using solid-state NMR. This schematic diagram depicts the bilayer (square planes) and peptides (cylinders) in relation to the magnetic field (B_0). As long as the orientation of the bilayer is known with respect to the magnetic field, then the resulting ^{15}N chemical shift resulting from the labeled peptide will indicate the orientation of the peptide in the magnetic field. (Figure adapted from Ref. 178, with permission).

PGLa is an antimicrobial peptide from the magainin family and solution-state ^{1}H NMR studies indicate that it adopts an α-helical structure in the presence of detergent micelles.[177] Consistent with this, the use of ^{15}N labelled amino acids in PGLa showed that the peptide was aligned parallel with the surface of the bilayer.[177,178] It should be noted that this measurement represents an average of all the ^{15}N labels in the sample and can only tell you if the helical axis of the peptide is more closely aligned with the bilayer normal or parallel to the bilayer surface. ^{13}C nuclei can also be used in a similar manner to determine the orientation of an antimicrobial peptide in phospholipid bilayers. Examples of studies that have examined antimicrobial peptide using ^{13}C chemical shift values include a study of the scorpion peptide, pin2,[179] and the β-sheet peptide, protegrin-1.[180]

Another application of ^{15}N nuclei to determine the orientation of antimicrobial peptides in bilayers is through the use of the 2D PISEMA (polarization inversion spin exchange at the magic angle) experiment.[181] The resonance frequencies in PISEMA spectra are dictated by the helix orientation, backbone dihedral angles, the magnitude and orientation of the principal elements of the amide ^{15}N chemical shift tensor and the NH bond length.[182] If the helix axis is parallel to the surface of the membrane then all of these amides have the same orientation in the magnetic field resulting in overlapping resonances from the same ^{1}H–^{15}N dipolar couplings and ^{15}N chemical shift frequencies. However, tilting the helix axis with respect to the magnetic field changes this symmetry and alters each of the amide NH bond vectors relative to the magnetic field[182] which manifests itself as a dispersion of the heteronuclear dipolar coupling and the chemical shift frequencies. The PISEMA experiment has been applied to solid-state NMR samples of magainin to examine the orientation of this peptide in the bilayer[183] and these results

support previous ^{15}N experiments that indicated a surface bound orientation for magainin.[184]

Other sensitive solid-state techniques have been developed to further characterize the orientation of antimicrobial peptides in phospholipid bilayers. For example, a number of studies describe highly sensitive ^{19}F NMR techniques to examine the orientation and dynamics of membrane bound peptides. ^{19}F (spin ½) is a highly sensitive NMR nucleus owing to its large gyromagnetic ratio[185] and dipolar interactions that can be detected over a large distance.[186] In addition, ^{19}F has a broad chemical shift range which can give rise to excellent resolution and since it is not commonly found in biological molecules, there is no concern for any background signals resulting from natural sources.[187] Homonuclear dipolar interactions between ^{19}F atoms or heteronuclear couplings to ^{31}P, ^{13}C or ^{15}N can be used to determine the intermolecular distances between the two atoms by measuring the dipolar couplings between the two nuclei. Additionally, the anisotropy of the dipolar interaction has been used to determine the angle between the two nuclei with respect to the magnetic field.[187]

^{19}F NMR studies have been performed on peptides containing fluorine substituted amino acid analogs which were substituted for nonpolar amino acids. This method has been applied, for example, to studies of the antimicrobial β-sheet peptide, gramicidin S (GS), to determine the structure and alignment of this antimicrobial peptide in DMPC bilayers.[188] GS is an interesting antimicrobial peptide because it is a cyclic antimicrobial peptide made by the gram positive bacterium, *Bacillus brevis*.[189] It is a cyclic peptide with the sequence (Val-Orn-Leu-DPhe-Pro)$_2$ and it displays antibacterial activity[190] as well as hemolytic activity.[191] The solution structure of GS was determined in DMSO using ^{1}H-NMR techniques and the peptide forms a well defined, antiparallel β-sheet with the two positively charged ornithine residues lying on one face of the peptide and the valine and leucine residues on the opposite face. The phenylalanine residues are found close to the plane formed by the backbone of the peptide.[192] To introduce ^{19}F atoms, two Leu residues in GS were replaced with 4-fluoro-phenylglycine (4F-Phg). Determining the correct distance between the ^{19}F nuclei was especially important in this case because NMR structural studies of GS analogs, with the sequence (Val-Lys-Leu-DTyr-Pro)$_2$, had suggested that the backbone conformation of this modified peptide differs significantly from the native peptide.[193] ^{19}F dipolar spectra were obtained using a modified Carr–Purcell–Meiboom–Gill (CPMG) multipulse sequence[187,194] and these generated distance restraints between the two ^{19}F nuclei. This demonstrated that the two ^{19}F nuclei in the 4F-Phg labelled GS were separated by a distance of ~6 Å, indicating that the ^{19}F labelled GS peptide had a similar conformation to the native GS and that the 4F-Phg probe could be incorporated into gramicidin S without largely disrupting the native structure of the peptide.[188] Similar to the ^{15}N solid-state NMR techniques described earlier, the oriented samples containing lipid and 4F-Phg labelled GS can be prepared so that the stacked bilayers are parallel or perpendicular to the magnetic field. The resulting chemical shift observed in the ^{19}F NMR spectra indicates whether the ^{19}F chemical shift anisotropy tensor is aligned with the magnetic field or perpendicular to it. The ^{19}F chemical shifts for labelled GS indicated that the peptide sits

at the surface of a DMPC bilayer with the 4F-Phg probes inserted in the hydrophobic core of the membrane.[188]

The use of two 4F-Phg labels presents some challenges in practice because the dipolar splitting can be affected by the local environment and the single ^{19}F nuclei can be difficult to reference. To overcome these challenges, another non natural amino acid 4-trifluoromethyl-phenylglycine (CF3-Phg) can be used as an alternative label because it has fast uniaxial rotation around the methyl axis.[195] CF3-Phg labels have been used for example to distinguish GS peptides containing either two L-CF3-Phg groups or one L-CF3-Phg and one D-CF3-Phg group.[195] CF3-Phg groups have also proven to be useful for studying the orientation of PGLa in phospholipid bilayers. In one study, four CF3-Phg analogs of PGLa were synthesized with the fluorine labeled amino acid replacing Ile9, Ala10, Ile 13 and Ala14, respectively.[196] At a peptide to lipid ratio of 1:200, PGLa was determined to be aligned along the surface of the DMPC bilayer and these results supported the previously mentioned studies with the ^{15}N labeled PGLa peptide.[177] Interestingly, at higher peptide–lipid concentrations ($\geq$1:50) there was an unexpected realignment of the PGLa peptide with respect to the bilayer surface and the helix axis assumed an oblique tilt angle in the middle of the membrane.[197] Increasing the peptide–lipid ratio up to and beyond 1:50 led to changes in the chemical shifts and dipolar couplings resulting from the CF3-Phg probes. The tilt angle of the PGLa peptide at high peptide–lipid ratios was calculated to be 123°, with the C-terminus of the peptide becoming inserted into the centre of the bilayer. The tilt angle was confirmed with ^{15}N NMR in which the ^{15}N signal from a ^{15}N-Gly11-PGLa peptide was seen to shift from 40 ppm at a peptide–lipid ratio of 1:200 to 68 ppm at a ratio of 1:50.[197]

This tilted state of PGLa represents an interesting result and demonstrates the strength of this approach for elucidating novel mechanisms of action for antimicrobial peptides. Many antimicrobial peptides become active only after a threshold concentration of peptide is reached and it is therefore important to study peptides at high concentrations to understand the underlying aspects of antimicrobial activity. It appears that PGLa self assembles as a dimer on the surface of a bilayer at high peptide concentrations which stabilizes the tilted state. It has been proposed that the tilted state of PGLa could be an intermediate between a monomeric surface bound state observed at low peptide concentrations and a transmembrane state which is highly cooperative among many PGLa peptides and which is ultimately responsible for the antimicrobial activity of this peptide.[197] Interestingly, PGLa was found to adopt a transmembrane alignment when mixed with an equal amount of magainin 2.[198] Both of these peptides are produced on the skin of *Xenopus laevis* and these results help explain previous studies that have found a synergistic relationship between these two peptides.[199,200] It has been proposed that PGLa associates at the surface of a bacterial membrane into tilted homodimers that occasionally adopt a transmembrane conformation which contributes to a small but significant antimicrobial activity. Upon addition of magainin 2, heterodimers form with the PGLa peptides and stable transmembrane pores appear in the bilayer resulting in significantly higher antimicrobial activity.[201]

So far, we have discussed the use of ^{15}N, ^{13}C and ^{19}F nuclei to examine the orientation of antimicrobial peptides with respect to a lipid bilayer but in principle, any spin interaction parallel to a helical axis can be used. The amide $^{2}H(D)$ quadrupolar coupling in a peptide backbone is capable of giving the same information and has been applied to studies of gramicidin S and gramicidin A.[202] The advantage of deuterium labelling versus ^{15}N is that the amide protons in the peptide backbone can be exchanged at many residues by a simple chemical exchange in D_2O while ^{15}N labels must be incorporated during chemical synthesis or biosynthetically. Additionally, the N-D and C-D groups produce large quadrupolar couplings which make them sensitive probes for this type of experiment. Recently, a proton inverse detected deuteron (PRIDE) NMR technique was described that has the ability to distinguish between surface bound and transmembrane orientation of peptides in bilayers.[203] Two peptides were used to demonstrate the efficacy of this method, a known transmembrane peptide, gramicidin D, and a surface bound peptide, ovispirin. The directly detected ^{2}H spectra of both peptides revealed only a strong D_2O signal and no large quadrupolar signals while the PRIDE spectrum showed large quadrupolar splittings and excellent suppression of the D_2O signal. Gramicidin D had a quadrupolar splitting of ~280 kHz which corresponded to a transmembrane orientation while the ovispirin PRIDE spectrum indicated a splitting of 104 kHz, which is consistent with an in-plane orientation of the antimicrobial peptide with respect to the bilayer.[203]

Deuterium labels have also been utilized in studies of PGLa in DMPC bilayers and it is interesting to compare the outcome to other available data. A series of eight PGLa derivatives were synthesized with Ala-d3 replacing Ala, Gly or Ile residues throughout the peptide.[198] This method is known as the geometric analysis of labelled alanines (GALA) and has been shown to be effective for examining the interactions between peptides and phospholipid bilayers.[142,204,205] In this fashion, PGLa was found to be bound to the surface of a bilayer at a peptide–lipid ratio of 1:200 which is consistent with the ^{15}N and ^{19}F studies described earlier.[177,196] The concentration-dependent realignment of PGLa was also seen with the Ala-d3 labels incorporated into the peptide, with a calculated tilt angle of 125°,[198] thereby confirming the previous observation made with CF3-Phg labelled PGLa.[197]

14. DEPTH OF MEMBRANE INSERTION

An alternative to determining the orientation of an antimicrobial peptide in a bilayer is to measure the depth of insertion of a peptide into the membrane.[15] Such information is important for interpreting the solution structure of a peptide and gives a better understanding of the interactions between an antimicrobial peptide and the phospholipid bilayer.

Many solid-state NMR techniques have been developed to determine the depth of membrane insertion by antimicrobial peptides. One of these techniques employs paramagnetic Mn^{2+}ions at the surface of the bilayer to enhance the T2 relaxation rates of nearby nuclear spins.[206,207] The paramagnetic Mn^{2+}ions bind to the surface of the membrane and dephase the ^{13}C signals from the peptide.

To determine the depth of protegrin 1 (PG-1) carbons, the dephasing is compared to the known values for the lipid functional groups which have been determined from neutron and x-ray diffraction studies.[208,209] PG-1 was found to be in the centre of the DLPC bilayer with Leu5 and Vall6 residing closest to the centre of the bilayer and the two most dephased residues Gly2 and Phe12 are found near the head groups of opposing leaflets of the bilayer.[206] These results are in agreement with previous studies of PG-1 in DLPC where the β-strand of PG-1 was shown to be in the centre of the bilayer but tilted from the bilayer normal.[210]

Another solid-state NMR technique involves measuring the ^{1}H spin diffusion between the phospholipids and the peptide or the peptide and the surrounding bulk water. In the spin-diffusion, a specific spin magnetization is selected based on its mobility and then it is transferred to a nearby target through dipolar couplings between the two nuclei.[211] This method has been used to characterize the orientation of colicin in a phospholipid bilayer. First, water was used as the source of magnetization to examine the membrane bound conformation of the colicin E1 channel domain.[212] In this case, the phospholipid bilayer containing the peptide was kept at a low temperature to hold it in a frozen gel state to limit the motion of these molecules. The water at the surface was still able to move at this temperature and a ^{1}H T2 filter was applied to select for the mobile ^{1}H nuclei in the water. The magnetization from the water molecules was transferred into the bilayer containing colicin and from the measured ^{15}N amide signals it was determined that there was a significant portion of the peptide that was found at the surface of the bilayer.[212] Unfortunately, this experiment had to be done at low temperatures to restrict the movement of the lipid bilayers and this is very different from the dynamic fluid state of biological membranes. Additionally, using water as the source of magnetization only allows one to probe regions of the peptide that are bound to the surface of the membrane but gives little information about peptides that insert into the hydrophobic core of the phospholipid bilayer.

To overcome these challenges, the methyl groups on the acyl chains were used as the magnetization source.[211] The methyl ^{1}H signal can be selected from its chemical shift using a 2D ^{1}H-^{13}C correlation experiment. The difference in this case is that the bilayer is kept at a temperature that allows it to remain fluid and consequently the diffusion coefficient of the lipids is smaller than that of the protein. As a result, the build-up curve only shows the shortest lipid-peptide distance but it cannot be assigned in a sequence specific manner.[15] When colicin Ia channel domain was examined in this manner, a small fraction of the peptide was determined to be within 2–4 Å of the centre of the bilayer.[211] The combination of these two magnetization transfer studies indicated that the membrane-bound colicin channel domain contains a transmembrane segment and a surface bound segment.

15. OLIGOMERIZATION OF ANTIMICROBIAL PEPTIDES

It is well known that many antimicrobial peptides can aggregate on the surface of bacterial cells and once an effective concentration is reached, they then disrupt the stability of the membrane leading to cell death. Oligomerisation may be induced

between antimicrobial peptides when they interact at the surface of a bacterial cell. Oftentimes, the solution structure of an antimicrobial peptide is determined in the presence of detergent micelles but since only one peptide is typically bound to a micelle at any given time, the possibility to study interactions between peptides is limited. For instance, a dimeric structure of a magainin based antimicrobial peptide, MSI-78, has been reported in DPC micelles.[213] Micelles have also been used to examine the oligomeric states of protegrin-1[214] and arenicin[215] but it becomes increasingly difficult to monitor these interactions using solution-state NMR experiments as the size of the oligomeric complex increases. Additionally, the structure of the peptide oligomers observed in detergent micelles may not be indicative of the structure that these peptides adopt in the presence of a phospholipid bilayer. Fortunately, solid-state NMR can monitor intermolecular interactions between antimicrobial peptides. This has led to a greater understanding of the importance of peptide aggregation on antimicrobial activity.

Two interesting solid-state NMR techniques have been described to examine the dimerisation of protegrin-1 in POPC bilayers. The first technique employs a ^{19}F spin diffusion technique to confirm the presence of PG-1 dimers in a phospholipid bilayer.[216] Strong ^{19}F–^{19}F dipolar couplings between 4-^{19}F-Phe labelled PG-1 were observed which indicated that PG-1 is dimerised within a 15 Å radius. This placed the C-terminal β-strands of PG-1 next to each other in the dimer structure but a parallel or antiparallel orientation of the two PG-1 peptides with respect to each other could not be definitively demonstrated.

A later study used rotational-echo double-resonance (REDOR) solid-state NMR to measure heteronuclear distances in isotopically labelled PG-1 and determined that β-strands of PG-1 form a parallel dimer in POPC bilayers.[217] REDOR experiments have also been employed to examine the dimeric structure of $(KIAGKIA)_3$ (K3),[218] an antimicrobial peptide related to magainin and PGLa.[219] In this case, the dimeric structure was found to be parallel, with two K3 peptides oriented 20° with respect to each other. Additional information from ^{19}F NMR of oriented bilayer samples indicated that these peptides adopt a transmembrane orientation. Altogether, these results suggest that K3 exerts its antimicrobial effect through the formation of a dimer/monomer lined pore through the membrane of a bacterial cell.[218]

16. BILAYER PERTURBATIONS

All of the NMR methods described to this point have been concerned with measuring parameters that are directly related to the antimicrobial peptide itself. Regardless of the mechanism of action of an antimicrobial peptide, the need for these molecules to interact at the surface of a bacterial membrane is absolutely essential. Whether the peptide induces the formation of pores, causes micellarization of the bacterial membrane or interacts with the bilayer as an intermediate step during cell penetration, all of these processes have a direct impact on the structural properties and spatial organization of the lipid molecules in a bilayer. Solid-state NMR has also proven to be useful for examining some changes that occur in phospholipid bilayers upon peptide binding.

The most common nucleus used for studying bilayers is ^{31}P. ^{31}P is a convenient nucleus for studying phospholipid bilayers because every phospholipid molecule found in a membrane contains a phosphate group in its head group. Add to that the 100% natural abundance of this spin 1/2 nucleus and a relatively high gyromagnetic ratio, ^{31}P becomes an extremely sensitive probe for studying membranes with NMR.[18] Probably the most widespread application of ^{31}P NMR is to check macroscopic alignment of oriented bilayer samples.[220] But this nucleus is capable of providing much more information. The ^{31}P chemical shift of the phosphorus atom in the head group of a phospholipid molecule is sensitive to the phase of the lipid, the conformation of the head group and perturbations to the surface of the membrane.[15] The addition of an antimicrobial peptide to a bilayer may alter the properties of the membrane and these changes can be measured with ^{31}P NMR.

The line shapes of ^{31}P spectra are heavily influenced by the lipid phase and characteristic spectra will arise for a lamellar phase, a hexagonal phase and a micellar phase respectively[221] (Figure 12). A common experiment used in studying antimicrobial peptides is looking for the formation of non-lamellar phases following the addition of peptide to a phospholipid bilayer. ^{31}P NMR cannot distinguish between normal and inverted lipid phases and several isotropic phases such as micelles, inverted micelles, certain cubic phases and SUVs will all produce single sharp peaks in a ^{31}P NMR spectra.[18] It is however, useful to examine the appearance of non-lamellar lipid phases.

^{31}P solid-state NMR has been used to examine the effect of a membrane-lysing peptide, pardaxin, on various lipid bilayers. Pardaxin is a mellitin-like peptide isolated from the Red Sea Moses sole, *Pardachirus marmoratus*, and it is released

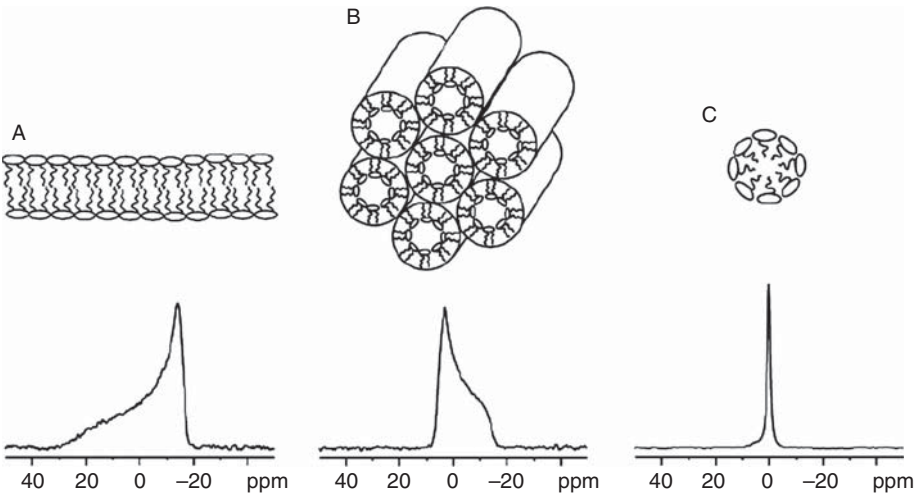


Figure 12 Characteristic ^{31}P line shapes from different lipid phases in unoriented samples. Panel (A) shows a bilayer structure, panel (B) is an inverted hexagonal phase and panel (C) is an isotropic micellar phase. (Figure adapted from Ref. 18. Copyright John Wiley and Sons, Ltd.).

as part of a shark repellent secretion to protect the fish from predation.[222] This is a cytotoxic peptide but it is useful to examine for our discussion of antimicrobial peptides because its mode of action is thought to involve a similar membrane disruption mechanism proposed for antimicrobial peptides. The ^{1}H NMR solution structure of pardaxin was solved in TFE:H_2O[223] and in the presence of DPC micelles.[224] Both structures are remarkably similar with two helical regions separated by a proline hinge, the presence of which is absolutely essential for the cytotoxic activity of this peptide.[225] A ^{31}P NMR study using uniaxially aligned samples demonstrated that the activity of pardaxin is dependent on the composition of the lipids in the bilayer. By monitoring the changes that occurred in the chemical shifts of the ^{31}P spectra, pardaxin was found to significantly disrupt bilayers containing zwitterionic lipids with choline or ethanolamine head groups. The ability of pardaxin to disrupt bilayers was inhibited by adding cholesterol or POPG to the samples.[226]

The effect of phospholipid head group composition on the membrane disrupting capabilities of PG-1 has also been studied.[227] In zwitterionic POPC bilayers, the addition of PG-1 increased the observed intensity of the ^{31}P spectra in the range of −15 to 20 ppm, indicating that the membrane is becoming unoriented. When tested with POPG, PG-1 induced the formation of an isotropic phase as indicated by the appearance of a strong peak at 1.5 ppm.[227] This isotropic peak was attributed to the formation of small vesicles caused by the fragmentation of the membrane by PG-1. In a later study, mutants of PG-1 were analyzed in a similar fashion to examine the effect of the disulfide bonds and the arginine residues of PG-1.[228] The changes induced in bilayers, monitored with ^{31}P NMR spectra, were not as pronounced for the mutated peptides compared to the native peptide. In fact, removing the disulfide bonds from PG-1 virtually abolished the membrane disrupting abilities of this peptide and highlighted the importance of the β-sheet structure for its activity. Removing key arginine residues resulted in increased disorder in the membrane but the micellar structures seen in PG-1 containing samples were not formed.[228]

The effect of acyl chain length on the ability of PG-1 to disrupt bilayers has also been examined with ^{31}P NMR.[210] Spectra of DLPC vesicles, with 12-C chains, mixed with PG-1 showed no significant change in the orientation of the lipids while increasing acyl chains to 14 (DMPC), 16 (DPPC) or a mix of 16 and 18 carbons (POPC) led to progressively larger perturbations in the bilayers as evidenced by significant changes in the ^{31}P spectra.[210] These results support the previous studies that PG-1 acts by dissolving the membrane of a target cell into small micelles but it also highlights the importance of hydrophobic mismatch in this interaction because the shorter acyl chain lipid vesicles are not fragmented by PG-1.

A complementary experiment to ^{31}P NMR for probing the membrane is ^{2}H NMR. Deuterium has a spin 1 nucleus and it possess a quadrupole moment.[229,230] The sensitivity of ^{2}H is low but this is compensated by a fast quadrupolar relaxation which can be exploited by using a short relaxation delay.[18] Selective or uniform labeling of phospholipids with deuterium gives a probe for examining the changes that occur in both the head group region as well as the hydrophobic core of a bilayer.[18]

The α- and β- segments of the choline head group are commonly used to probe lipid-peptide interactions at the surface of a membrane (Figure 4). The quadrupolar splittings of these groups change linearly and counter-directionally when charged molecules interact at the surface of a bilayer.[231] For example, changes in the quadrupolar splittings of these segments have been observed in studies of pardaxin[224] indicating that this peptide interacts with the choline headgroups of phospholipid bilayers.

The quadrupolar splitting from deuterated acyl chains are directly related to the motional order in the hydrophobic core of the bilayer. The acyl chain order parameter profile indicates the mobility of the chain at all the positions along the acyl chain and this information can be used to determine if a peptide interacts at the surface of a bilayer or becomes inserted into the hydrophobic core. A peptide that binds to the surface of a bilayer will increase the space between phospholipid molecules which should result in increased mobility of the acyl chains. Alternatively, a peptide that adopts a transmembrane orientation may either stabilize or destabilize the acyl chain packing depending on the interactions that occur with the hydrophobic acyl chains. The addition of the antimicrobial peptide chrysophsin to POPE/POPG-d31 and POPE/POPE-d31/POPG bilayers caused a significant decrease in the quadrupolar splitting while there was no significant change in bilayers composed of POPC/POPC-d31 cholesterol.[232] Static ^{15}N NMR spectra indicated that chrysophsin is bound at the/surface of the bilayer and these results demonstrate that this peptide inserts into the lipid–water interface of negatively charged bilayers.[232]

Similar increases in acyl chain disorder for surface bound antimicrobial peptides have been reported for pleurocidin.[233] In this example, ^{31}P and ^{15}N NMR experiments were used to determine the orientation of the bilayer and the peptide in the magnetic field and it was concluded that pleurocidin is bound to the surface of the membrane. Solid-state ^{2}H echo spectra of POPE-d31 or POPG-d31 in vesicles composed of POPE/POPE-d31/POPG or POPE/POPG-d31 were collected and the results were used to evaluate the ability of pleurocidin to disrupt the acyl chains[233] (Figure 13). Pleurocidin interacted with both lipid species but it disrupted the anionic PG acyl chains more strongly that the zwitterionic PE lipids. Pleurocidin contains three histidine residues which become protonated at acidic pH. However, the protonation of these residues did not affect the interaction between the peptide and the lipid vesicles as the spectra of POPE/POPE-d31/POPG (2:1:1) and POPE/POPG-d31 (3:1) in the presence of 2.5 mol% pleurocidin were nearly identical at pH 7.5 and 5.

High resolution ^{1}H–^{13}C heteronuclear dipolar solid-state NMR spectroscopy has also proven to be an effective means of observing changes in the membrane topology upon the binding of an antimicrobial peptide. In this case, the ^{1}H–^{13}C dipolar coupling profile of the phospholipid molecules can be measured to map out the region of the bilayer that the peptide interacts with. This approach has been used to examine the interactions between the antimicrobial peptide, $(KIGAKI)_3$, and aligned DMPC/DHPC bicelles.[234] In the presence of the $(KIGAKI)_3$ peptide, significant peak shifts in the 2D ^{1}H–^{13}C dipolar coupling/^{13}C chemical shift spectrum were seen for carbon atoms found close to the head

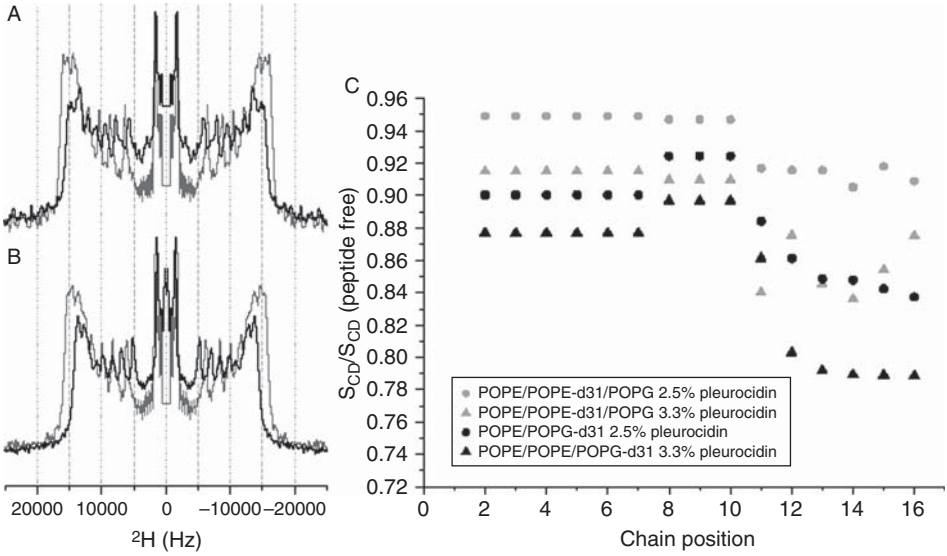


Figure 13 Effect of pleurocidin on vesicles containing either POPE-d31 or POPG-d31. The ^{2}H echo spectra of POPE/POPE-d31/POPG (2:1:1) vesicles (A) and POPE/POPG-d31 (3:1) vesicles (B) in the absence (grey lines) and presence (black lines) of 2.5 mol% pleurocidin. The ^{2}H order parameter profiles of the pleurocidin containing vesicles at pH 7.5 demonstrate that the peptide disrupts the acyl chains of both lipid species but this effect is more pronounced for the anionic PG molecules. (Figure from Ref. 233, with permission).

groups of the phospholipids. Based on these results, $(KIGAKI)_3$ was determined to be bound to the surface of the bicelle. These results are in agreement with previous ^{31}P and ^{2}H studies of $(KIGAKI)_3$ in phospholipid bilayers.[235]

^{31}P and ^{2}H NMR analyses are frequently used in combination to fully characterize the effect of an antimicrobial peptide on a phospholipid bilayer. The antimicrobial activity and membrane interactions of two synthetic peptides, MSI-78 and MSI-594, were examined using both of these techniques. These two peptides, based on the magainin and PGLa sequences, have enhanced antimicrobial activity over the natural peptides.[236] ^{15}N chemical shift spectra indicate that both peptides adopt a surface bound conformation. ^{31}P NMR spectra of aligned bilayers indicate that both peptides bind to the phosphate region of POPC and POPG bilayers. However, the disorder observed between the two lipid species is not the same, indicating that the lipid composition dictates the disorder induced by these peptides. Decreases in the quadrupolar splitting were also observed in de-Paked ^{2}H quadrupole coupling spectra of POPC-d31 bilayers in the presence of each peptide.[236] These results, coupled with differential scanning calorimetry experiments, point to MSI-78 and MSI-594 binding at the surface of a bacterial membrane where they induce positive curvature strain in the phospholipid bilayer and they mostly exert their antimicrobial activity through the carpet mechanism. In the case of MSI-78 and POPC bilayers, the membrane destabilization occurs because of the formation of a toroidal pore.[236]

17. CONCLUSIONS

Antimicrobial peptides are a large and very diverse group of biomolecules that exert their antimicrobial activity through a variety of mechanisms, most of which involve, at least in part, interactions with the plasma membrane of a bacterial cell. Over the last 10 years, NMR spectroscopy has proven to be a valuable research tool for studying antimicrobial peptides and their interactions with lipids. This review has outlined an array of NMR experiments that have been applied to determine the structural characteristics of antimicrobial peptides and to elucidate mechanistic properties associated with this class of biomolecules.

Solution-state NMR is the preferred approach for determining the high resolution structures of antimicrobial peptides in a variety of solvents. While few antimicrobial peptides are capable of folding in aqueous solution, many others are unfolded and only adopt an amphipathic conformation in the presence of detergent micelles or membrane mimetic organic solvents. As we have discussed, a common theme is that the peptides adopt amphipathic structures; however, different secondary structure elements have been identified. The detailed 3D structures, in combination with other biophysical experiments, can give an excellent insight into the mode of action of an antimicrobial peptide to selectively kill pathogenic bacteria. Through the application of recombinant expression methods to incorporate ^{13}C and ^{15}N isotopes into antimicrobial peptides, and with the constant technological advancements that are being made in NMR instrumentation, the size of complex that can be studied using solution-state NMR is expanding to include peptides bound to bicelles and even vesicles. One of the best examples that illustrates the potential of solution-state NMR is a study addressing the mechanism of membrane permeabilization for sapecin in vesicles. In this study, isotope labelling, transferred cross-saturation and hydrogen-deuterium exchange experiments were all used.[237] Using this novel approach, the authors were able to define the structure of the oligomeric state of sapecin and they could determine the region of this peptide complex that was in contact with the lipids (Figure 14).

In spite of the exciting advancements in solution-state NMR techniques, solid-state NMR remains the best method for examining the interactions between antimicrobial peptides and phospholipid bilayers. Through isotopic labelling with various NMR active nuclei, information regarding the orientation of the peptide in the bilayer, the depth of peptide insertion into the membrane and the oligomeric state of the antimicrobial peptide can all be established. Solid-state NMR has the added advantage of being able to monitor changes that occur in the phospholipid bilayer upon peptide binding which is an important consideration when examining the mode of action of antimicrobial peptides.

In the future, NMR studies of antimicrobial peptides will make use of a variety of solution and solid-state techniques that we have described to fully characterize the mode of action of an antimicrobial peptide. Indeed, examples of elegant studies that employ an array of NMR experiments are already beginning to surface in the literature.[238,239] NMR spectroscopy continues to give important insights into the mechanisms of antimicrobial activity and is contributing to the development of new synthetic analogs with enhanced activity and potential pharmaceutical applications.[240,241]

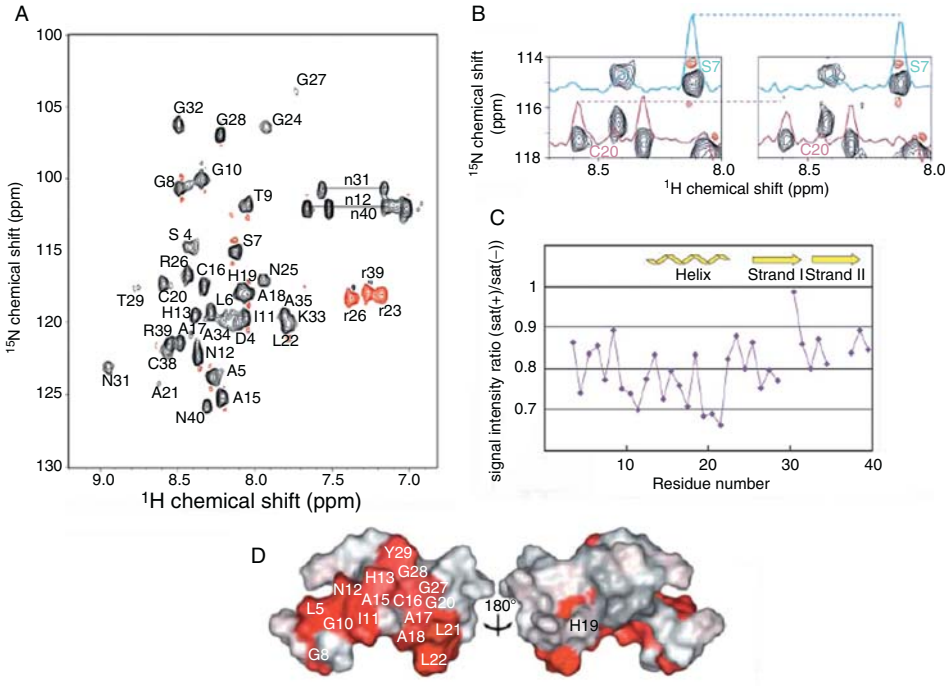


Figure 14 Transferred cross saturation (TCS) results for sapecin in PC containing vesicles. The ^{1}H, ^{15}N HSQC spectrum of L5A/V35A mutant of sapecin in the presence of PC vesicles (A) with the assignments indicated. A portion of the ^{1}H,^{15}N HSQC demonstrates the differences in signal intensity for labeled sapecin when the PC vesicles were irradiated (right) and when they were not (left) (B). Examining the reduction ratios of the peak intensities (C) and mapping the residues that are affected by the TCS experiment onto the structure of sapecin reveals which regions of the peptide form the binding interface with the PC vesicles (gray) (D). Color versions of these figures can be found in Ref. 237.

ACKNOWLEDGMENTS

The authors are indebted to Drs. David Schibli, Howard Hunter and Leo Nguyen for many stimulating discussions and sharing valuable insights on this topic in recent years. H. J. Vogel is a scientist of the Alberta Heritage Foundation for Medical Research.

REFERENCES

1. H. Steiner, D. Hultmark, A. Engstrom, H. Bennich and H. G. Boman, *Nature*, 1981, **292,** 246–248.
2. M. Zasloff, *Proc. Natl. Acad. Sci. USA*, 1987, **84,** 5449–5453.
3. P. Bulet, C. Hetru, J. L. Dimarcq and D. Hoffmann, *Dev. Comp. Immunol.*, 1999, **23,** 329–344.
4. F. Garcia-Olmedo, A. Molina, J. M. Alamillo and P. Rodriguez-Palenzuela, *Biopolymers*, 1998, **47,** 479–491.
5. K. A. Brogden, M. Ackermann, P. B. McCray and B. F. Tack, *Int. J. Antimicrob. Agents*, 2003, **22,** 465–478.
6. K. de Smet and R. Contreras, *Biotechnol. Lett.*, 2005, **27,** 1337–1347.

7. J. M. Willey and W. A. van der Donk, *Annu. Rev. Microbiol.*, 2007, **61,** 477–501.
8. S. B. Levy and B. Marshall, *Nat. Med.*, 2004, **10,** S122–S129.
9. Z. Wang and G. Wang, *Nucleic Acids Res.*, 2004, **32,** D590–D592.
10. S. E. Blondelle and R. A. Houghten, *Trends Biotechnol.*, 1996, **14,** 60–65.
11. K. A. Brogden, *Nat. Rev. Microbiol.*, 2005, **3,** 238–250.
12. D. I. Chan, E. J. Prenner and H. J. Vogel, *Biochim. Biophys. Acta*, 2006, **1758,** 1184–1202.
13. W. M. Yau, W. C. Wimley, K. Gawrisch and S. H. White, *Biochemistry*, 1998, **37,** 14713–14718.
14. P. M. Hwang and H. J. Vogel, *Biochem. Cell Biol.*, 1998, **76,** 235–246.
15. M. Hong, *Acc. Chem. Res.*, 2006, **39,** 176–183.
16. J. P. S. Powers and R. E. W. Hancock, *Peptides*, 2003, **24,** 1681–1691.
17. L. T. Nguyen, E. J. Prenner and H. J. Vogel, in: *Modern Magnetic Resonance*, G. A. Webb, ed., Springer, Berlin, 2006, pp. 1297–1305.
18. E. Strandberg and A. S. Ulrich, *Concepts Magn. Reson. Pt. A*, 2004, **23A,** 89–120.
19. G. S. Wang, *Curr. Org. Chem.*, 2006, **10,** 569–581.
20. Z. Oren and Y. Shai, *Biopolymers*, 1998, **47,** 451–463.
21. G. Baumann and P. Mueller, *J. Supramol. Struct.*, 1974, **2,** 538–557.
22. S. Qian, W. Wang, L. Yang and H. W. Huang, *Biophys. J.*, 2008, **94,** 3512–3522.
23. S. J. Ludtke, K. He, W. T. Heller, T. A. Harroun, L. Yang and H. W. Huang, *Biochemistry*, 1996, **35,** 13723–13728.
24. K. Matsuzaki, O. Murase, N. Fujii and K. Miyajima, *Biochemistry*, 1996, **35,** 11361–11368.
25. S. Campagna, N. Saint, G. Molle and A. Aumelas, *Biochemistry*, 2007, **46,** 1771–1778.
26. K. A. H. Wildman, D. K. Lee and A. Ramamoorthy, *Biochemistry*, 2003, **42,** 6545–6558.
27. B. Bechinger, *Biochim. Biophys. Acta*, 1999, **1462,** 157–183.
28. H. Leontiadou, A. E. Mark and S. J. Marrink, *J. Am. Chem. Soc.*, 2006, **128,** 12156–12161.
29. D. P. Tieleman, *Biophys. J.*, 2004, **86,** 371A–372A.
30. M. Miteva, M. Andersson, A. Karshikoff and G. Otting, *FEBS Lett.*, 1999, **462,** 155–158.
31. A. Pokorny and P. F. F. Almeida, *Biochemistry*, 2005, **44,** 9538–9544.
32. L. Otvos, *Cell. Mol. Life Sci.*, 2002, **59,** 1138–1150.
33. L. J. Zhang, R. Benz and R. E. W. Hancock, *Biochemistry*, 1999, **38,** 8102–8111.
34. C. B. Park, M. S. Kim and S. C. Kim, *Biochem. Biophys. Res. Commun.*, 1996, **218,** 408–413.
35. C. B. Park, K. S. Yi, K. Matsuzaki, M. S. Kim and S. C. Kim, *Proc. Natl. Acad. Sci. USA*, 2000, **97,** 8245–8250.
36. J. P. S. Powers, A. Tan, A. Ramamoorthy and R. E. W. Hancock, *Biochemistry*, 2005, **44,** 15504–15513.
37. J. P. S. Powers, M. M. Martin, D. L. Goosney and R. E. W. Hancock, *Antimicrob. Agents Chemother.*, 2006, **50,** 1522–1524.
38. V. Delorenzo, *Arch. Microbiol.*, 1984, **139,** 72–75.
39. X. Thomas, D. Destoumieux-Garzon, J. Peduzzi, C. Afonso, A. Blond, N. Birlirakis, C. Goulard, L. Dubost, R. Thai, J. C. Tabet and S. Rebuffat, *J. Biol. Chem.*, 2004, **279,** 28233–28242.
40. V. Delorenzo, J. L. Martinez and C. Asensio, *J. Gen. Microbiol.*, 1984, **130,** 391–400.
41. E. Strahsburger, M. Baeza, O. Monasterio and R. Lagos, *Antimicrob. Agents Chemother.*, 2005, **49,** 3083–3086.
42. R. Lagos, M. Wilkens, C. Vergara, X. Cecchi and O. Monasterio, *FEBS Lett.*, 1993, **321,** 145–148.
43. V. Delorenzo and A. P. Pugsley, *Antimicrob. Agents Chemother.*, 1985, **27,** 666–669.
44. S. Bieler, F. Silva, C. Soto and D. Belin, *J. Bacteriol.*, 2006, **188,** 7049–7061.
45. D. B. Diep, M. Skaugen, Z. Salehian, H. Holo and I. F. Nes, *Proc. Natl. Acad. Sci. USA*, 2007, **104,** 2384–2389.
46. M. Mattiuzzo, A. Bandiera, R. Gennaro, M. Benincasa, S. Pacor, N. Antcheva and M. Scocchi, *Mol. Microbiol.*, 2007, **66,** 151–163.
47. K. LeVier and G. C. Walker, *J. Bacteriol.*, 2001, **183,** 6444–6453.
48. M. Lavina, A. P. Pugsley and F. Moreno, *J. Gen. Microbiol.*, 1986, **132,** 1685–1693.
49. R. A. Salomon and R. N. Farias, *J. Bacteriol.*, 1995, **177,** 3323–3325.
50. C. B. Park, H. S. Kim and S. C. Kim, *Biochem. Biophys. Res. Commun.*, 1998, **244,** 253–257.
51. C. H. Hsu, C. P. Chen, M. L. Jou, A. Y. L. Lee, Y. C. Lin, Y. P. Yu, W. T. Huang and S. H. Wu, *Nucleic Acids Res.*, 2005, **33,** 4053–4064.

52. P. Casteels and P. Tempst, *Biochem. Biophys. Res. Commun.*, 1994, **199,** 339–345.
53. M. Tomita, W. Bellamy, M. Takase, K. Yamauchi, H. Wakabayashi and K. Kawase, *J. Dairy Sci.*, 1991, **74,** 4137–4142.
54. P. M. Hwang, N. Zhou, X. Shan, C. H. Arrowsmith and H. J. Vogel, *Biochemistry*, 1998, **37,** 4288–4298.
55. N. Fujitani, T. Kouno, T. Nakahara, K. Takaya, T. Osaki, S. I. Kawabata, M. Mizuguchi, T. Aizawa, M. Demura, S. I. Nishimura and K. Kawano, *J. Pept. Sci.*, 2007, **13,** 269–279.
56. G. Wang, *Biochim. Biophys. Acta*, 2007, **1768,** 3271–3281.
57. T. V. Ovchinnikova, Z. V. Shenkarev, K. D. Nadezhdin, S. V. Balandin, M. N. Zhmak, I. A. Kudelina, E. I. Finkina, V. N. Kokryakov and A. S. Arseniev, *Biochem. Biophys. Res. Commun.*, 2007, **360,** 156–162.
58. P. A. Raj, E. Marcus and D. K. Sukumaran, *Biopolymers*, 1998, **45,** 51–67.
59. D. Brewer, H. Hunter and G. Lajoie, *Biochem. Cell Biol.*, 1998, **76,** 247–256.
60. E. F. Haney, F. Lau and H. J. Vogel, *Biochim. Biophys. Acta*, 2007, **1768,** 2355–2364.
61. D. I. Chan, H. N. Hunter, B. F. Tack and H. J. Vogel, *Antimicrob. Agents Chemother*, 2008, **52,** 883–894.
62. D. J. Schibli, P. M. Hwang and H. J. Vogel, *FEBS Lett.*, 1999, **446,** 213–217.
63. L. T. Nguyen, D. J. Schibli and H. J. Vogel, *J. Pept. Sci.*, 2005, **11,** 379–389.
64. D. J. Schibli, L. T. Nguyen, S. D. Kernaghan, O. Rekdal and H. J. Vogel, *Biophys. J.*, 2006, **91,** 4413–4426.
65. G. S. Wang, P. A. Keifer and A. Peterkofsky, *Protein Sci.*, 2003, **12,** 1087–1096.
66. J. S. Nowick, O. Khakshoor, M. Hashemzadeh and J. O. Brower, *Org. Lett.*, 2003, **5,** 3511–3513.
67. W. P. Aue, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.*, 1976, **64,** 2229–2246.
68. J. Cavanagh, W. J. Fairbrother, A. G. Palmer, III, M. Rance and N. J. Kelton, *Protein NMR Spectroscopy – Principles and Practice*. Elsevier, Burlington, MA. 2007.
69. D. Marion and K. Wuthrich, *Biochem. Biophys. Res. Commun.*, 1983, **113,** 967–974.
70. M. Rance, O. W. Sorensen, G. Bodenhausen, G. Wagner, R. R. Ernst and K. Wuthrich, *Biochem. Biophys. Res. Commun.*, 1983, **117,** 479–485.
71. G. Cornilescu, F. Delaglio and A. Bax, *J. Biomol. NMR*, 1999, **13,** 289–302.
72. X. Li, Y. Li, A. Peterkofsky and G. S. Wang, *Biochim. Biophys. Acta*, 2006, **1758,** 1203–1214.
73. X. Li, Y. Li, H. Han, D. W. Miller and G. Wang, *J. Am. Chem. Soc.*, 2006, **128,** 5776–5785.
74. A. T. Brunger, P. D. Adams, G. M. Clore, W. L. Delano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr. D Biol. Crystallogr.*, 1998, **54,** 905–921.
75. A. T. Brunger, *Nat. Protoc.*, 2007, **2,** 2728–2733.
76. C. D. Schwieters, J. J. Kuszewski, N. Tjandra and G. Marius Clore, *J. Magn. Reson.*, 2003, **160,** 65–73.
77. M. Habeck, W. Rieping, J. P. Linge and M. Nilges, *Methods Mol. Biol.*, 2004, **278,** 379–402.
78. J. P. Linge, S. I. O'Donoghue and M. Nilges, *Methods Enzymol.*, 2001, **339,** 71–90.
79. J. P. Linge, M. Habeck, W. Rieping and M. Nilges, *Bioinformatics*, 2003, **19,** 315–316.
80. W. Rieping, M. Habeck, B. Bardiaux, A. Bernard, T. E. Malliavin and M. Nilges, *Bioinformatics*, 2007, **23,** 381–382.
81. M. C. Baran, Y. J. Huang, H. N. B. Moseley and G. T. Montelione, *Chem. Rev.*, 2004, **104,** 3541–3555.
82. A. S. Altieri and R. A. Byrd, *Curr. Opin. Struct. Biol.*, 2004, **14,** 547–553.
83. P. Guntert, C. Mumenthaler and K. Wuthrich, *J. Mol. Biol.*, 1997, **273,** 283–298.
84. P. Guntert, *Methods Mol. Biol.*, 2004, **278,** 353–378.
85. Y. P. J. Huang, G. V. T. Swapna, P. K. Rajan, H. P. Ke, B. Xia, K. Shukla, M. Inouye and G. T. Montelione, *J. Mol. Biol.*, 2003, **327,** 521–536.
86. N. Oezguen, L. Adamian, Y. Xu, K. Rajarathnam and W. Braun, *J. Biomol. NMR*, 2002, **22,** 249–263.
87. J. Kuszewski, C. D. Schwieters, D. S. Garrett, R. A. Byrd, N. Tjandra and G. M. Clore, *J. Am. Chem. Soc.*, 2004, **126,** 6258–6273.
88. C. Landon, C. Thouzeau, H. Labbe, P. Bulet and F. Vovelle, *J. Biol. Chem.*, 2004, **279,** 30433–30439.
89. D. J. Schibli, H. N. Hunter, V. Aseyev, T. D. Starner, J. M. Wiencek, P. B. Mccray, B. F. Tack and H. J. Vogel, *J. Biol. Chem.*, 2002, **277,** 8279–8289.
90. M. V. Sawai, H. P. Jia, L. D. Liu, V. Aseyev, J. M. Wiencek, P. B. Mccray, T. Ganz, W. R. Kearney and B. F. Tack, *Biochemistry*, 2001, **40,** 3810–3816.

91. Y. Q. Tang, J. Yuan, G. Osapay, K. Osapay, D. Tran, C. J. Miller, A. J. Ouellette and M. E. Selsted, *Science*, 1999, **286,** 498–502.
92. M. Trabi, H. J. Schirra and D. J. Craik, *Biochemistry*, 2001, **40,** 4211–4221.
93. Z. O. Shenkarev, K. D. Nadezhdin, V. A. Sobol, A. G. Sobol, L. Skjeldal and A. S. Arseniev, *FEBS J.*, 2006, **273,** 2658–2672.
94. C. V. Jennings, K. J. Rosengren, N. L. Daly, M. Plan, J. Stevens, M. J. Scanlon, C. Waine, D. G. Norman, M. A. Anderson and D. J. Craik, *Biochemistry*, 2005, **44,** 851–860.
95. R. C. Macdonald, R. I. Macdonald, B. P. M. Menco, K. Takeshita, N. K. Subbarao and L. R. Hu, *Biochim. Biophys. Acta*, 1991, **1061,** 297–303.
96. J. Lauterwein, C. Bosch, L. R. Brown and K. Wuthrich, *Biochim. Biophys. Acta*, 1979, **556,** 244–264.
97. H. W. vandenHooven, F. Fogolari, H. S. Rollema, R. N. H. Konings, C. W. Hilbers and F. J. M. vandeVen, *FEBS Lett.*, 1993, **319,** 189–194.
98. P. Damberg, J. Jarvet and A. Graslund, *Nucl. Magn. Reson. Biol. Macromol. Pt. B*, 2001, **339,** 271–285.
99. M. Tomita, M. Takase, W. Bellamy and S. Shimamura, *Acta Paediatr. Jpn.*, 1994, **36,** 585–591.
100. J. H. Kang, M. K. Lee, K. L. Kim and K. S. Hahm, *Int. J. Pept. Protein Res.*, 1996, **48,** 357–363.
101. W. Jing, J. S. Svendsen and H. J. Vogel, *Biochem. Cell Biol.*, 2006, **84,** 312–326.
102. K. S. Hoek, J. M. Milne, P. A. Grieve, D. A. Dionysius and R. Smith, *Antimicrob. Agents Chemother.*, 1997, **41,** 54–59.
103. H. Ulvatne, O. Samuelsen, H. H. Haukland, M. Kramer and L. H. Vorland, *FEMS Microbiol. Lett.*, 2004, **237,** 377–384.
104. A. Pellegrini, U. Thomas, N. Bramaz, S. Klauser and R. vonFellenberg, *J. Appl. Microbiol.*, 1997, **82,** 372–378.
105. H. N. Hunter, W. G. Jing, D. J. Schibli, T. Trinh, I. Y. Park, S. C. Kim and H. J. Vogel, *Biochim. Biophys. Acta*, 2005, **1668,** 175–189.
106. A. J. Rezansoff, H. N. Hunter, W. Jing, I. Y. Park, S. C. Kim and H. J. Vogel, *J. Pept. Res.*, 2005, **65,** 491–501.
107. H. H. Haukland, H. Ulvatne, K. Sandvik and L. H. Vorland, *FEBS Lett.*, 2001, **508,** 389–393.
108. M. I. A. van der Kraan, J. Groenink, K. Nazmi, E. C. I. Veerman, J. G. M. Bolscher and A. V. N. Amerongen, *Peptides*, 2004, **25,** 177–183.
109. M. I. van der Kraan, M. C. van der, K. Nazmi, W. van't Hof, J. Groenink, E. C. Veerman, J. G. Bolscher and A. V. Nieuw Amerongen, *Peptides*, 2005, **26,** 2093–2097.
110. W. Braun, G. Wider, K. H. Lee and K. Wuthrich, *J. Mol. Biol.*, 1983, **169,** 921–948.
111. G. M. Clore, S. R. Martin and A. M. Gronenborn, *J. Mol. Biol.*, 1986, **191,** 553–561.
112. D. Marion, M. Zasloff and A. Bax, *FEBS Lett.*, 1988, **227,** 21–26.
113. G. S. Yi, C. B. Park, S. C. Kim and C. Cheong, *FEBS Lett.*, 1996, **398,** 87–90.
114. H. Wong, J. H. Bowie and J. A. Carver, *Eur. J. Biochem.*, 1997, **247,** 545–557.
115. S. H. Park, Y. K. Kim, J. W. Park, B. J. Lee and B. J. Lee, *Eur. J. Biochem.*, 2000, **267,** 2695–2704.
116. J. L. Gifford, H. N. Hunter and H. J. Vogel, *Cell. Mol. Life Sci.*, 2005, **62,** 2588–2598.
117. H. N. Hunter, A. R. Demcoe, H. Jenssen, T. J. Gutteberg and H. J. Vogel, *Antimicrob. Agents Chemother.*, 2005, **49,** 3387–3395.
118. N. Zhou, D. P. Tieleman and H. J. Vogel, *Biometals.*, 2004, **17,** 217–223.
119. M. E. Girvin, V. K. Rastogi, F. Abildgaard, J. L. Markley and R. H. Fillingame, *Biochemistry*, 1998, **37,** 8817–8824.
120. M. Schwaiger, M. Lebendiker, H. Yerushalmi, M. Coles, A. Groger, C. Schwarz, S. Schuldiner and H. Kessler, *Eur. J. Biochem.*, 1998, **254,** 610–619.
121. T. Prasch, T. Naumann, R. L. M. Markert, M. Sattler, W. Schubert, S. Schaal, M. Bauch, H. Kogler and C. Griesinger, *Eur. J. Biochem.*, 1997, **244,** 501–512.
122. D. Sipos, M. Andersson and A. Ehrenberg, *Eur. J. Biochem.*, 1992, **209,** 163–169.
123. S. Srisailam, T. K. S. Kumar, A. I. Arunkumar, K. W. Leung, C. Yu and H. M. Chen, *Eur. J. Biochem.*, 2001, **268,** 4278–4284.
124. S. Rothemund, H. Weisshoff, M. Beyermann, E. Krause, M. Bienert, C. Mugge, B. D. Sykes and F. D. Sonnichsen, *J. Biomol. NMR*, 1996, **8,** 93–97.
125. J. J. Skalicky, M. E. Selsted and A. Pardi, *Proteins*, 1994, **20,** 52–67.
126. N. J. Baxter and M. P. Williamson, *J. Biomol. NMR*, 1997, **9,** 359–369.
127. J. Gesell, M. Zasloff and S. J. Opella, *J. Biomol. NMR*, 1997, **9,** 127–135.

128. J. M. Mancheno, A. M. Del Pozo, J. P. Albar, M. Onaderra and J. G. Gavilanes, *J. Pept. Res.*, 1998, **51,** 142–148.
129. S. R. Martin and M. J. Schilstra, in: *Biophysical Tools for Biologists: Vol 1 In Vitro Techniques*, Elsevier, San Diego, 2008, 263–293.
130. H. W. vandenHooven, C. A. E. M. Spronk, M. vandeKamp, R. N. H. Konings, C. W. Hilbers and F. J. M. vandeVen, *Eur. J. Biochem.*, 1996, **235,** 394–403.
131. D. J. Schibli, R. C. Montelaro and H. J. Vogel, *Biochemistry*, 2001, **40,** 9570–9578.
132. K. Yu, S. Kang, S. D. Kim, P. D. Ryu and Y. Kim, *J. Biomol. Struct. Dyn.*, 2001, **18,** 595–606.
133. H. J. Vogel, D. J. Schibli, W. G. Jing, E. M. Lohmeier-Vogel, R. F. Epand and R. M. Epand, *Biochem. Cell Biol.*, 2002, **80,** 49–63.
134. L. R. Brown, C. Bosch and K. Wuthrich, *Biochim. Biophys. Acta*, 1981, **642,** 296–312.
135. M. Respondek, T. Madl, C. Gobl, R. Golser and K. Zangger, *J. Am. Chem. Soc.*, 2007, **129,** 5228–5234.
136. G. Pintacuda and G. Otting, *J. Am. Chem. Soc.*, 2002, **124,** 372–373.
137. D. Andreu, J. Ubach, A. Boman, B. Wahlin, D. Wade, R. B. Merrifield and H. G. Boman, *FEBS Lett.*, 1992, **296,** 190–194.
138. R. S. Prosser, P. A. Luchette, P. W. Westerman, A. Rozek and R. E. W. Hancock, *Biophys. J.*, 2001, **80,** 1406–1416.
139. A. Rozek, C. L. Friedrich and R. E. W. Hancock, *Biochemistry*, 2000, **39,** 15765–15774.
140. M. N. Triba, D. E. Warschawski and P. F. Devaux, *Biophys. J.*, 2005, **88,** 1887–1901.
141. P. A. Luchette, T. N. Vetman, R. S. Prosser, R. E. W. Hancock, M. P. Nieh, C. J. Glinka, S. Krueger and J. Katsaras, *Biochim. Biophys. Acta*, 2001, **1513,** 83–94.
142. J. A. Whiles, R. Brasseur, K. J. Glover, G. Melacini, E. A. Komives and R. R. Vold, *Biophys. J.*, 2001, **80,** 280–293.
143. C. B. Post, *Curr. Opin. Struct. Biol.*, 2003, **13,** 581–588.
144. G. M. Clore and A. M. Gronenborn, *J. Magn. Reson.*, 1983, **53,** 423–442.
145. K. Wakamatsu, A. Takeda, T. Tachi and K. Matsuzaki, *Biopolymers*, 2002, **64,** 314–327.
146. K. Matsuzaki, O. Murase, H. Tokuda, S. Funakoshi, N. Fujii and K. Miyajima, *Biochemistry*, 1994, **33,** 3342–3349.
147. K. Wakamatsu, A. Okada, T. Miyazawa, M. Ohya and T. Higashijima, *Biochemistry*, 1992, **31,** 5654–5660.
148. A. Okada, K. Wakamatsu, T. Miyazawa and T. Higashijima, *Biochemistry*, 1994, **33,** 9438–9446.
149. J. J. Chou, J. D. Kaufman, S. J. Stahl, P. T. Wingfield and A. Bax, *J. Am. Chem. Soc.*, 2002, **124,** 2450–2451.
150. R. Tycko, F. J. Blanco and Y. Ishii, *J. Am. Chem. Soc.*, 2000, **122,** 9340–9341.
151. H. J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield and S. Grzesiek, *J. Biomol. NMR*, 2000, **18,** 303–309.
152. N. Tjandra and A. Bax, *Science*, 1997, **278,** 1111–1114.
153. J. Cohen, *Nature*, 2002, **420,** 885–891.
154. S. Farnaud, C. Spiller, L. C. Moriarty, A. Patel, V. Gant, E. W. Odell and R. W. Evans, *FEMS Microbiol. Lett.*, 2004, **233,** 193–199.
155. E. ElassRochard, A. Roseanu, D. Legrand, M. Trif, V. Salmon, C. Motas, J. Montreuil and G. Spik, *Biochem. J.*, 1995, **312,** 839–845.
156. Y. R. Bommineni, H. E. Dai, Y. X. Gong, J. L. Soulages, S. C. Fernando, U. DeSilva, O. Prakash and G. L. Zhang, *FEBS J.*, 2007, **274,** 418–428.
157. D. Rifkind, *J. Infect. Dis.*, 1967, **117,** 433–438.
158. P. Pristovsek and J. Kidric, *J. Med. Chem.*, 1999, **42,** 4604–4613.
159. B. Japelj, P. Pristovsek, A. Majerle and R. Jerala, *J. Biol. Chem.*, 2005, **280,** 16955–16961.
160. A. Bhunia, P. N. Domadia and S. Bhattacharjya, *Biochim. Biophys. Acta*, 2007, **1768,** 3282–3291.
161. A. Bax and S. Grzesiek, *Acc. Chem. Res.*, 1993, **26,** 131–138.
162. S. Talluri and G. Wagner, *J. Magn. Reson. B*, 1996, **112,** 200–205.
163. D. Ifrah, X. Doisy, T. S. Ryge and P. R. Hansen, *J. Pept. Sci.*, 2005, **11,** 113–121.
164. D. P. Satchell, T. Sheynis, S. Kolusheva, J. Cummings, T. K. Vanderlick, R. Jelinek, M. E. Selsted and A. J. Ouellette, *Peptides*, 2003, **24,** 1795–1805.
165. K. J. Rosengren, N. L. Daly, L. M. Fornander, L. M. H. Jonsson, Y. Shirafuji, X. Q. Qu, H. J. Vogel, A. J. Ouellette and D. J. Craik, *J. Biol. Chem.*, 2006, **281,** 28068–28078.

166. Y. Gueguen, A. Herpin, A. Aumelas, J. Garnier, J. Fievet, J. M. Escoubas, P. Bulet, M. Gonzalez, C. Lelong, P. Favrel and E. Bachere, *J. Biol. Chem.*, 2006, **281,** 313–323.
167. J. Y. Moon, K. A. Henzler-Wildman and A. Ramamoorthy, *Biochim. Biophys. Acta*, 2006, **1758,** 1351–1358.
168. R. T. Baker, *Curr. Opin. Biotechnol.*, 1996, **7,** 541–546.
169. W. J. Moon, D. K. Hwang, E. J. Park, Y. M. Kim and Y. K. Chae, *Protein Expr. Purif.*, 2007, **51,** 141–146.
170. U. Silphaduang and E. J. Noga, *Nature*, 2001, **414,** 268–269.
171. J. C. Pierce, W. L. Maloy, L. Salvador and C. F. Dungan, *Mol. Marine Biol. Biotechnol.*, 1997, **6,** 248–259.
172. Y. S. Kim, M. J. Kim, P. Kim and J. H. Kim, *Biotechnol. Lett.*, 2006, **28,** 357–362.
173. T. R. Butt, S. C. Edavettal, J. P. Hall and M. R. Mattern, *Protein Expr. Purif.*, 2005, **43,** 1–9.
174. S. Macauley-Patrick, M. L. Fazenda, B. McNeil and L. M. Harvey, *Yeast*, 2005, **22,** 249–270.
175. D. Destoumieux, P. Bulet, J. M. Strub, A. Van Dorsselaer and E. Bachere, *Eur. J. Biochem.*, 1999, **266,** 335–346.
176. B. Bechinger, Y. Kim, L. E. Chirlian, J. Gesell, J. M. Neumann, M. Montal, J. Tomich, M. Zasloff and S. J. Opella, *J. Biomol. NMR*, 1991, **1,** 167–173.
177. B. Bechinger, M. Zasloff and S. J. Opella, *Biophys. J.*, 1998, **74,** 981–987.
178. B. Bechinger, L. M. Gierasch, M. Montal, M. Zasloff and S. J. Opella, *Solid State Nucl. Magn. Reson.*, 1996, **7,** 185–191.
179. K. Nomura, G. Corzo, T. Nakajima and T. Iwashita, *Biophys. J.*, 2004, **87,** 2497–2507.
180. J. J. Buffy, A. J. Waring, R. I. Lehrer and M. Hong, *Biochemistry*, 2003, **42,** 13725–13734.
181. C. H. Wu, A. Ramamoorthy and S. J. Opella, *J. Magn. Reson., A*, 1994, **109,** 270–272.
182. F. M. Marassi and S. J. Opella, *J. Magn. Reson.*, 2000, **144,** 150–155.
183. F. M. Marassi, C. Ma, J. J. Gesell and S. J. Opella, *J. Magn. Reson.*, 2000, **144,** 156–161.
184. A. Ramamoorthy, F. M. Marassi, M. Zasloff and S. J. Opella, *J. Biomol. NMR*, 1995, **6,** 329–334.
185. T. C. Pochapsky and S. S. Pochapsky, in: *NMR for Physical and Biological Scientists*. Taylor and Francis Group, New York, 2000.
186. S. O. Smith, K. Aschheim and M. Groesbeek, *Q. Rev. Biophys.*, 1996, **29,** 395–449.
187. S. L. Grage and A. S. Ulrich, *J. Magn. Reson.*, 1999, **138,** 98–106.
188. J. Salgado, S. L. Grage, L. H. Kondejewski, R. S. Hodges, R. N. McElhaney and A. S. Ulrich, *J. Biomol. NMR*, 2001, **21,** 191–208.
189. G. F. Gause and M. G. Brazhnikova, *Nature*, 1944, **154,** 703.
190. L. H. Kondejewski, S. W. Farmer, D. S. Wishart, R. E. W. Hancock and R. S. Hodges, *Int. J. Pept. Protein Res.*, 1996, **47,** 460–466.
191. L. H. Kondejewski, S. W. Farmer, D. S. Wishart, C. M. Kay, R. E. W. Hancock and R. S. Hodges, *J. Biol. Chem.*, 1996, **271,** 25261–25268.
192. Y. Xu, I. P. Sugar and N. R. Krishna, *J. Biomol. NMR*, 1995, **5,** 37–48.
193. A. C. Gibbs, L. H. Kondejewski, W. Gronwald, A. M. Nip, R. S. Hodges, B. D. Sykes and D. S. Wishart, *Nat. Struct. Biol.*, 1998, **5,** 284–288.
194. S. L. Grage and A. S. Ulrich, *J. Magn. Reson.*, 2000, **146,** 81–88.
195. S. L. Grage, A. V. Suleymanova, S. Afonin, P. Wadhwani and A. S. Ulrich, *J. Magn. Reson.*, 2006, **183,** 77–86.
196. R. W. Glaser, C. Sachse, U. H. N. Durr, P. Wadhwani and A. S. Ulrich, *J. Magn. Reson.*, 2004, **168,** 153–163.
197. R. W. Glaser, C. Sachse, U. H. N. Durr, P. Wadhwani, S. Afonin, E. Strandberg and A. S. Ulrich, *Biophys. J.*, 2005, **88,** 3392–3397.
198. E. Strandberg, P. Wadhwani, P. Tremouilhac, U. H. N. Durr and A. S. Ulrich, *Biophys. J.*, 2006, **90,** 1676–1686.
199. R. W. Williams, R. Starman, K. M. P. Taylor, K. Gable, T. Beeler, M. Zasloff and D. Covell, *Biochemistry*, 1990, **29,** 4490–4496.
200. A. V. Gomes, A. Dewaal, J. A. Berden and H. V. Westerhoff, *Biochemistry*, 1993, **32,** 5365–5372.
201. P. Tremouilhac, E. Strandberg, P. Wadhwani and A. S. Ulrich, *J. Biol. Chem.*, 2006, **281,** 32089–32094.
202. K. P. Datema, K. P. Pauls and M. Bloom, *Biochemistry*, 1986, **25,** 3796–3803.

203. S. Yamaguchi and M. Hong, *J. Magn. Reson.*, 2002, **155,** 244–250.
204. P. C. A. van der Wel, E. Strandberg, J. A. Killian and R. E. Koeppe, *Biophys. J.*, 2002, **83,** 1479–1488.
205. E. Strandberg, S. Ozdirekcan, D. T. S. Rijkers, P. C. A. van der Wel, R. E. Koeppe, R. M. J. Liskamp and J. A. Killian, *Biophys. J.*, 2004, **86,** 3709–3721.
206. J. J. Buffy, T. Hong, S. Yamaguchi, A. J. Waring, R. I. Lehrer and M. Hong, *Biophys. J.*, 2003, **85,** 2363–2373.
207. S. Tuzi, J. Hasegawa, R. Kawaminami, A. Naito and H. Saito, *Biophys. J.*, 2001, **81,** 425–434.
208. S. H. White and W. C. Wimley, *Annu. Rev. Biophys. Biomol. Struct.*, 1999, **28,** 319–365.
209. M. C. Wiener and S. H. White, *Biophys. J.*, 1992, **61,** 434–447.
210. S. Yamaguchi, T. Hong, A. Waring, R. I. Lehrer and M. Hong, *Biochemistry*, 2002, **41,** 9852–9862.
211. D. Huster, X. L. Yao and M. Hong, *J. Am. Chem. Soc.*, 2002, **124,** 874–883.
212. K. K. Kumashiro, K. Schmidt-Rohr, O. J. Murphy, K. L. Ouellette, W. A. Cramer and L. K. Thompson, *J. Am. Chem. Soc.*, 1998, **120,** 5043–5051.
213. F. Porcelli, B. A. Buck-Koehntop, S. Thennarasu, A. Ramamoorthy and G. Veglia, *Biochemistry*, 2006, **45,** 5793–5799.
214. C. Roumestand, V. Louis, A. Aumelas, G. Grassy, B. Calas and A. Chavanieu, *FEBS Lett.*, 1998, **421,** 263–267.
215. T. V. Ovchinnikova, Z. O. Shenkarev, S. V. Balandin, K. D. Nadezhdin, A. S. Paramonov, V. N. Kokryakov and A. S. Arseniev, *Biopolymers*, 2007, **89,** 455–464.
216. J. J. Buffy, A. J. Waring and M. Hong, *J. Am. Chem. Soc.*, 2005, **127,** 4477–4483.
217. R. Mani, M. Tang, X. Wu, J. J. Buffy, A. J. Waring, M. A. Sherman and M. Hong, *Biochemistry*, 2006, **45,** 8341–8349.
218. O. Toke, R. D. O'Connor, T. K. Weldeghiorghis, W. L. Maloy, R. W. Glaser, A. S. Ulrich and J. Schaefer, *Biophys. J.*, 2004, **87,** 675–687.
219. W. L. Maloy and U. P. Kari, *Biopolymers*, 1995, **37,** 105–122.
220. M. Ouellet, J. D. Doucet, N. Voyer and M. Auger, *Biochemistry*, 2007, **46,** 6597–6606.
221. J. Seelig, *Biochim. Biophys. Acta*, 1978, **515,** 105–140.
222. S. A. Thompson, K. Tachibana, K. Nakanishi and I. Kubota, *Science*, 1986, **233,** 341–343.
223. M. G. Zagorski, D. G. Norman, C. J. Barrow, T. Iwashita, K. Tachibana and D. J. Patel, *Biochemistry*, 1991, **30,** 8009–8017.
224. F. Porcelli, B. Buck, D. K. Lee, K. J. Hallock, A. Ramamoorthy and G. Veglia, *J. Biol. Chem.*, 2004, **279,** 45815–45823.
225. Y. Shai, D. Bach and A. Yanovsky, *J. Biol. Chem.*, 1990, **265,** 20202–20209.
226. K. J. Hallock, D. K. Lee, J. Omnaas, H. I. Mosberg and A. Ramamoorthy, *Biophys. J.*, 2002, **83,** 1004–1013.
227. R. Mani, J. J. Buffy, A. J. Waring, R. I. Lehrer and M. Hong, *Biochemistry*, 2004, **43,** 13839–13848.
228. R. Mani, A. J. Waring, R. I. Lehrer and M. Hong, *Biochim. Biophys. Acta*, 2005, **1716,** 11–18.
229. J. Seelig, *Q. Rev. Biophys.*, 1977, **10,** 353–418.
230. J. H. Davis, *Biochim. Biophys. Acta*, 1983, **737,** 117–171.
231. J. Seelig, P. M. Macdonald and P. G. Scherer, *Biochemistry*, 1987, **26,** 7535–7541.
232. A. J. Mason, P. Bertani, G. Moulay, A. Marquette, B. Perrone, A. F. Drake, A. Kichler and B. Bechinger, *Biochemistry*, 2007, **46,** 15175–15187.
233. A. J. Mason, I. N. Chotimah, P. Bertani and B. Bechinger, *Mol. Membr. Biol.*, 2006, **23,** 185–194.
234. J. X. Lu, K. Damodaran and G. A. Lorigan, *J. Magn. Reson.*, 2006, **178,** 283–287.
235. J. X. Lu, K. Damodaran, J. Blazyk and G. A. Lorigan, *Biochemistry*, 2005, **44,** 10208–10217.
236. A. Ramamoorthy, S. Thennarasu, D. K. Lee, A. Tan and L. Maloy, *Biophys. J.*, 2006, **91,** 206–216.
237. K. Takeuchi, H. Takahashi, M. Sugai, H. Iwai, T. Kohno, K. Sekimizu, S. Natori and I. Shimada, *J. Biol. Chem.*, 2004, **279,** 4981–4987.
238. K. Nomura, G. Ferrat, T. Nakajima, H. Darbon, T. Iwashita and G. Corzo, *Biophys. J.*, 2005, **89,** 4067–4080.
239. Y. L. Pan, J. T. J. Cheng, J. Hale, J. H. Pan, R. E. W. Hancock and S. K. Straus, *Biophys. J.*, 2007, **92,** 2854–2864.
240. C. Landon, F. Barbault, M. Legrain, M. Guenneugues and F. Vovelle, *Proteins*, 2008, **72,** 229–239.
241. R. P. Hicks, J. B. Bhonsle, D. Venugopal, B. W. Koser and A. J. Magill, *J. Med. Chem.*, 2007, **50,** 3026–3036.

CHAPTER 2

NMR Studies of Protein Folding: Folding Studies of Calcium-Binding Lysozyme and α-Lactalbumin

Mineyuki Mizuguchi,[*] Tomoyasu Aizawa,[†] Keiichi Kawano,[‡] and Makoto Demura[†]

Contents

[*] Faculty of Pharmaceutical Sciences, University of Toyama, Toyama 930-0194, Japan
[†] Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan
[‡] Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

Annual Reports on NMR Spectroscopy, Volume 65
ISSN 0066-4103, DOI: 10.1016/S0066-4103(08)00202-0

Abstract NMR is a powerful tool for characterizing the structure and dynamics of proteins at the site-specific level. A variety of NMR techniques can be applied to study various conformational states during the folding process, including the highly unfolded, partially folded, and native states. One particularly powerful method is the combination of NMR with hydrogen–deuterium exchange to indicate the formation of hydrogen bonds in proteins. Examples of the application of a variety of NMR methods to the protein folding studies of α-lactalbumin and c-type lysozyme are presented.

Keywords: Alpha-lactalbumin, Calcium-binding lysozyme, Folding process, Molten globule, Hydrogen-deuterium exchange.

1. INTRODUCTION

Protein folding is the process by which an unfolded polypeptide chain folds into a specific native and functional structure under physiological conditions. Elucidation of the mechanism underlying protein folding is a fundamental problem in molecular structural biology. To elucidate protein folding, the structural changes occurring along folding pathways must first be understood in detail. Ideally, protein folding studies will provide information about the structural changes during the folding process at high structural resolution. NMR is particularly powerful in this regard, as it can provide high-resolution structural data on a variety of conformational states along the folding pathways, including native states, intermediate conformational states, and highly unfolded conformations.[1–6]

Although NMR has made particularly significant contributions to studies of protein folding, its application to folding studies is not easy because of its low intrinsic sensitivity. To obtain 2D and 3D NMR spectra, it is generally necessary to prepare protein solutions at high concentrations (about 1 mM), which is higher than those in other analytical methods such as circular dichroism and fluorescence spectroscopy. Therefore, low solubility and aggregation properties of proteins preclude the application of NMR to folding studies of a wide range of proteins. Recent progress in NMR technology (magnetic field strength, cryogenic probes, and electronics), however, has greatly enhanced the sensitivity of NMR experiments, allowing the acquisition of multidimensional data for more dilute protein samples that are less likely to aggregate. Furthermore, modern NMR technology is applicable to marginally stable and low-populated intermediates that accumulate transiently during the folding process.[7–9]

α-Lactalbumin (α-LA) and lysozyme are small globular proteins, and their folding has been extensively studied by a wide range of techniques, including NMR spectroscopy. α-Lactalbumin is a mammalian Ca^{2+}-binding protein that plays an important role in the biosynthesis of lactose.[10, 11] On the other hand, c-type lysozyme is a ubiquitous lytic enzyme that degrades the peptidoglycan of bacterial cell walls. The two functions are quite different, but α-LA and c-type lysozyme are homologous proteins that are presumed, by virtue of similarities in primary structure, gene

sequence, and 3D structure, to have diverged from a common ancestor.[12–14] This protein family is divided into three groups: α-LA, Ca^{2+}-binding lysozyme, and conventional (non-Ca^{2+}-binding) lysozyme.[15] Canine and equine lysozymes bind Ca^{2+} tightly and are considered an intermediate in the molecular evolution from non-Ca^{2+}-binding lysozyme to α-LA.[15, 16] α-LA and c-type lysozyme are comprised of two types of domains: the α- and β-domains.[14, 17–19] The α-domain contains four α-helices (A to D), while the β-domain has a β-sheet and a 3_{10}-helix (Figure 1).

One of the most interesting characteristics of α-LA and Ca^{2+}-binding lysozyme is that the intermediate conformational state accumulates transiently at an early stage of refolding prior to the formation of the native state[20–23] (Figure 2). The transient intermediate is very similar to the molten-globule state, which has been found to be stable at equilibrium under mildly denaturing conditions, such as in solutions at low pH or in the presence of low concentrations of denaturant.[20, 24, 25]

In this paper, we review the NMR studies on the folding and stability of α-LA and lysozyme. First, we briefly summarize the hydrogen–deuterium exchange of proteins and its applications to the studies of α-LA and lysozyme. Second, we summarize NMR techniques used to study the unfolded conformations of α-LA and lysozyme. Finally, we discuss the studies on folding kinetics by hydrogen exchange pulse labelling and real-time NMR.

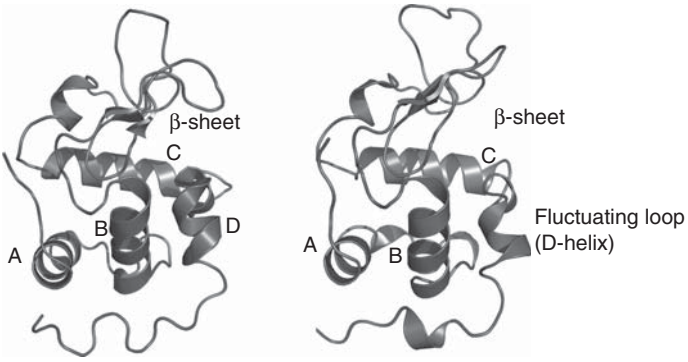


Figure 1 Structure of canine lysozyme (left) and bovine α-LA (right). The α-helices (A–D) and the β-sheet are labelled. The residues 105–110 of α-LA is referred to as a fluctuating loop. The figure was produced with a programme PyMOL (available at http://www.pymol.org).

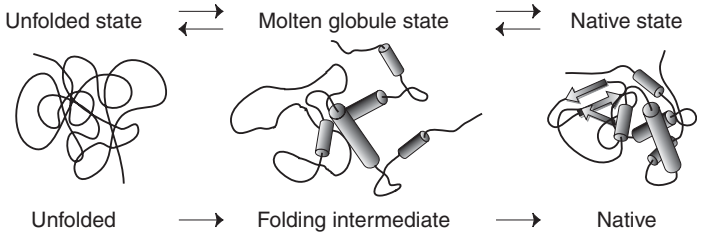


Figure 2 Schematic diagram of the three-state folding of proteins.

2. HYDROGEN–DEUTERIUM EXCHANGE

When proteins are dissolved in an aqueous solution, labile hydrogens in proteins, including amide hydrogens in the backbone and some of the polar side chains, are exchanged with the hydrogens from the solvent. Thus by changing the solvent from H_2O to D_2O, labile hydrogens will be exchanged with deuteriums. The process of the hydrogen–deuterium exchange can be followed by successive acquisitions of the NMR spectrum.

The rates of exchange for amide hydrogens in folded proteins can be slowed significantly compared to the corresponding intrinsic rate. The intrinsic rate is calculated from model peptide data and depends on pH, temperature, and the amino acid sequence.[26] The retardation of hydrogen exchange is strongly dependent on the environment of amide hydrogen in a protein; amide hydrogens in a hydrogen-bonded structure can be strongly protected from exchange, while those in the flexible part of the protein are exchanged rapidly. The retardation of the observed exchange rate for an individual amide hydrogen, k_{obs}, relative to the predicted intrinsic rate, k_{int}, is often reported as the protection factor, $P = k_{int}/k_{obs}$. Protection factors are frequently greater than 10^6 for amide hydrogens in rigid parts of the native protein.

The hydrogen–deuterium exchange can provide important information for understanding the principles of protein structure and structural dynamics. Hydrogen–deuterium exchange occurs when amide hydrogens are exposed to solvents where they are accessible for exchange. Therefore, it has been believed that the hydrogen exchange of proteins occurs as a result of structural unfolding and/or solvent penetration.[27] The vast majority of the slowly exchanging hydrogens is involved in hydrogen-bonded structures, whereas amide hydrogens not involved in hydrogen bonds are generally exchanged more rapidly, even if they are not in direct contact with the solvent as a result of burial in the interior of the protein. Thus, the hydrogen bond appears to be the primary factor in providing protection against exchange, while solvent accessibility plays a secondary role.[28–30]

Hydrogen exchange has made particularly significant contributions to the field of protein folding, since the technique can be applied to a great variety of conformational states of a protein, including completely unfolded proteins, partially folded states under equilibrium conditions, and transient kinetic intermediates formed during the folding process.[27, 31–35] In completely unfolded proteins, the observed rates of exchange are those predicted from the intrinsic rates.[26, 36] Molten globules contain significant levels of secondary structure, and the hydrogen bonds present in these can lead to protection from hydrogen–deuterium exchange. The protection factors found in equilibrium molten globules are much smaller than those in native proteins, indicating that the secondary structure in the molten globule is more labile than that in the native state.[33, 37–41]

2.1. Hydrogen–deuterium exchange of equine and canine lysozyme in the native state

Morozova-Roche et al. have reported the amide hydrogen exchange rates of equine lysozyme measured at pH 4.5 and 25°C.[40] Sixty-seven amide hydrogens of the Ca^{2+}-bound form were protected from hydrogen exchange. There were

significant differences in the hydrogen exchange behaviour of native equine and hen lysozymes. For instance, certain amides in the loop regions before and after the C-helix (the Ca^{2+}-binding site and the CD loop) showed considerably greater protection in equine than in hen lysozyme.[40, 42] In hen lysozyme, the corresponding residues showed little or no protection from hydrogen exchange.[42] These loops are located in a large cleft that divides the lysozyme into two structural domains.[18] In equine lysozyme, Ca^{2+}-binding promotes the formation of a hydrogen-bonded structure in the Ca^{2+}-binding region and the CD loops that link the α-domain to the β-domain.[40]

We have reported the hydrogen–deuterium exchange rate of canine lysozyme in the native state.[41] Hydrogen exchange measurements were carried out with uniformly ^{15}N-labelled protein at 0.6 mM. The hydrogen exchange reaction was initiated by dissolving lyophilized protein in D_2O buffer at pH 4.5 containing 10 mM $CaCl_2$. A series of HSQC spectra were collected at 25°C every 3 h for 3 days. The sample tube was stored at 25°C and occasionally subjected to a collection of HSQC spectra.

Eighty-one amide hydrogens were protected from hydrogen exchange at pH 4.5 and 25°C; the remaining amides were exchanged too rapidly to be detected by the HSQC experiments (Figure 3). The number of protected amides (81) in native canine lysozyme is larger than those reported for equine lysozyme (67) at 25°C and pH 4.5[40] and hen lysozyme (68) at 30°C and pH 7.5.[42] The protection factors for the protected residues in canine lysozyme are, however, lower than those in equine and hen lysozyme.[41] The maximum protection factor in hen lysozyme is more than 10^8, while the maximal protection factor is less than 10^6 in canine lysozyme and 10^7 in equine lysozyme.[40–42] This can be attributed primarily to the difference in thermal stability of their native states; hen lysozyme is more stable and highly cooperative than equine and canine lysozyme.[19, 43]

The exchange rate constants for tryptophan indole hydrogens can be used as probes for the protein dynamics. The locations of the five tryptophan residues (Trp 28, 63, 64, 108, and 111) are shown in Figure 4. In the native state of canine lysozyme, Trp 28, 108, and 111 are exchanged slowly, and the others are exchanged too rapidly to be observed.

2.2. NMR investigations of α-LA in the native state

A number of broad resonances have been found in the HSQC spectrum of bovine α-LA in the native state. For instance, cross-peaks arising from residues His107 and Cys111, located in the D-helix, were found to be much weaker than resonances from residues located in other regions. Furthermore, resonances arising from Leu105, Lys108, and Ala109, all located in the D-helix, were broadened beyond detection.[38] Resonances from these residues were also missing in the HSQC spectrum of human α-LA.[37] This observation indicates the conformational fluctuation on a millisecond-to-microsecond timescale in the D-helix region. In addition, Thr30, Thr33, and Ser34 from the B-helix, being in contact with the D-helix, also showed broadened resonances.[38]

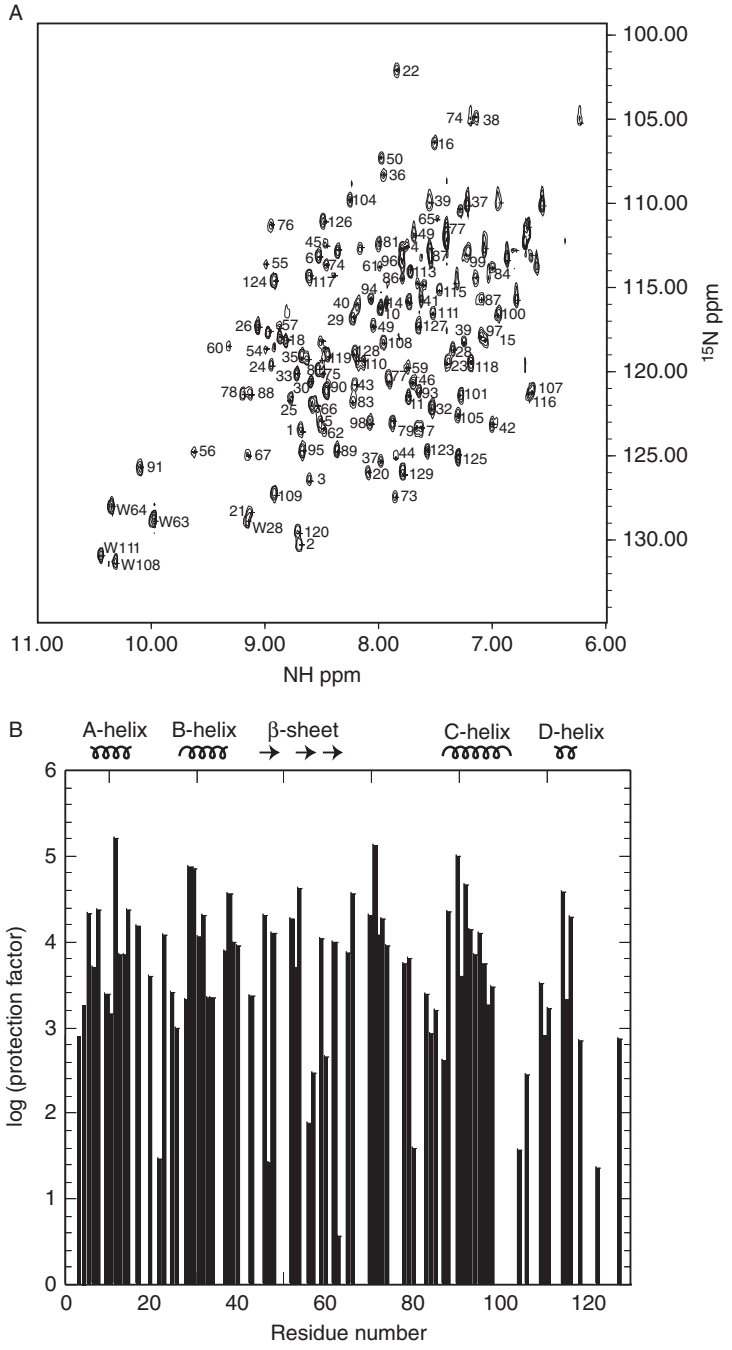


Figure 3 NMR spectrum and protection factor of canine lysozyme at pH 4.5 in the native state. (A) HSQC spectrum of holo canine lysozyme in the native state, indicating backbone assignment. (B) Protection factor versus amino acid residue number.

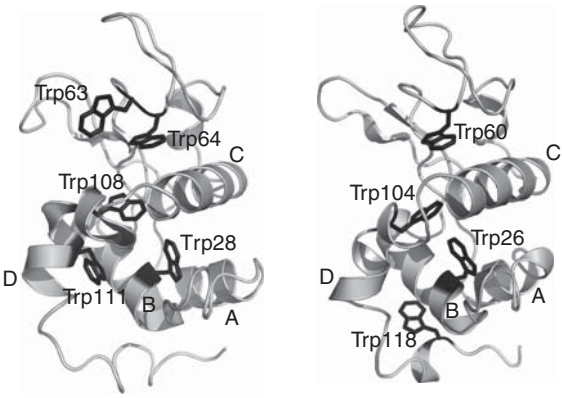


Figure 4 The locations of tryptophan residues in canine lysozyme (left) and bovine (right). The α-helices (A–D) are labelled. The figure was produced with a programme PyMOL (available at http://www.pymol.org).

The X-ray structures of α-LA have revealed the conformational fluctuation in the D-helix. The residues 105–110 adopt a distorted α-helical conformation (D-helix) in bovine α-LA, whereas they show a loop structure in goat and guinea pig α-LA.[17] The X-ray structures of human α-LA are both helical and loop in this region, indicating that the human α-LA structure fluctuates greatly in solution and that the helix-loop transition in the region of Trp104-Cys111 is in equilibrium.[44] Therefore, the D-helix (residues 105–110) of α-LA is referred to as a fluctuating loop (Figure 1).

Forge et al. studied the hydrogen–deuterium exchange of bovine α-LA in the native state at 20°C and pH 6.25.[38] The amide hydrogens of the helix D are exchanged too rapidly, suggesting the high flexibility of this region in native bovine α-LA. The authors also demonstrated that highly protected amide hydrogens are located around two disulphide bonds: Cys61–Cys77 within the *β*-domain of the protein, and Cys73–Cys91, which connects the *β*-domain to the C-helix. The same results were reported for native human α-LA.[37] The hydrogen exchange of native bovine α-LA also showed that the residues with the most highly protected amide hydrogens are clustered around the Ca^{2+}-binding region.[38] This effect is much more pronounced than in human α-LA where, in addition to the Ca^{2+}-binding region of the protein, much of the B-helix is well protected.[37]

2.3. NMR investigations of chimera between bovine α-LA and equine lysozyme

As discussed above, the D-helix of α-LA fluctuates in solution on a millisecond-to-microsecond timescale. In contrast, the corresponding region (residues 109–114; D-helix) in c-type lysozyme forms a rigid α-helix.[18, 19] Furthermore, the amide-hydrogen exchange studies of c-type lysozymes indicated that the amide hydrogens of the D-helix are exchanged slowly in contrast to the rapid amide hydrogen exchange in the fluctuating loop of α-LA.[37–41] We produced a chimera, BLAELZ,

by replacing the fluctuating loop (105–110 region) of bovine α-LA with the D-helix (109–114 region) of equine lysozyme, and studied the structure and dynamics of BLAELZ.[45, 46]

The ^{1}H-^{15}N HSQC spectrum of BLAELZ at 20°C is shown in Figure 5. The resonances are well dispersed, showing that BLAELZ has a well-defined tertiary structure. The backbone $^{1}H^{\alpha}$, $^{1}H^{N}$, and ^{15}N chemical shifts of BLAELZ were in excellent agreement with those of bovine α-LA, except for some residues, including those near the fluctuating loop.[38, 46] Signal broadening is not found for the resonances from residues in the 105–110 region.

We measured the rates of solvent exchange of amide hydrogens in the native state of BLAELZ by recording ^{1}H-^{15}N HSQC spectra as a function of time following the dissolution of the protein in D_2O at pH 6.3 and 20°C (Figure 5). Exchange rates were determined for 71 amide hydrogens of the protein main chain and for the side-chain indole hydrogen of Trp26, Trp104, and Trp107.[46] As has been observed for bovine and human α-LAs and equine lysozyme, residues with highly protected hydrogens were observed in both the α- and β-domains of BLAELZ (Figure 5). High protection against exchange was observed for amide hydrogens in the A- and B-helices. Amide hydrogens in the C-helix showed much higher protection factors compared with those in the rest of the α-domain.[46]

There were significant differences in the hydrogen exchange protection exhibited by BLAELZ and bovine α-LA (Figure 5). High protection from hydrogen exchange was observed for Ala106 and Trp107 located in a region corresponding to the fluctuating loop. The highest protection ($\log P > 7$) was observed at Trp104 just proximal to the fluctuating loop[46]; the protection factor of Trp104 in bovine α-LA was less than 10^5.[38] The high protection from exchange in the fluctuating loop of BLAELZ is similar to that observed in the D-helix of equine lysozyme.[39, 40] Therefore, the substitution of the fluctuating loop in bovine α-LA with the D-helix in equine lysozyme affects the hydrogen exchange behaviour of α-LA in the fluctuating loop.[46]

High protection from hydrogen exchange was observed for the indole hydrogens of three Trp residues, Trp26, 104, and 107, in BLAELZ.[46] Furthermore, we observed NOE cross-peaks between side-chain hydrogens (H^{β} and H^{γ}) of Val21 and $N^{\varepsilon 1}$ hydrogens of Trp26 and Trp107, suggesting that the B-helix comes in close proximity to the introduced D-helix (Figure 6). Three Trp residues and Val21 in BLAELZ correspond to Trp28, 108, 111, and Tyr23 in equine lysozyme, which are buried in the interior of the protein and form a rigid hydrophobic core.[18] Therefore, the overall structural stabilization of α-LA caused by the D-helix of equine lysozyme appears to be due to structural interactions of the inserted D-helix with the B-helix (Figure 6).

We investigated the dynamics of BLAELZ by model-free analysis derived from the relaxation rates R_1 ($=1/T_1$) and R_2 ($=1/T_2$) and the heteronuclear ^{1}H-^{15}N NOE.[46] The 105–110 region in BLAELZ exhibited R_2 values similar to those for other structural elements in the α-domain, and there was no significant R_{ex} term around the fluctuating loop (Figure 7). Therefore, the substitution of the fluctuating loop in bovine α-LA with the D-helix in equine lysozyme decreases the mobility on the millisecond-to-microsecond timescale in the residues 105–110.[46]

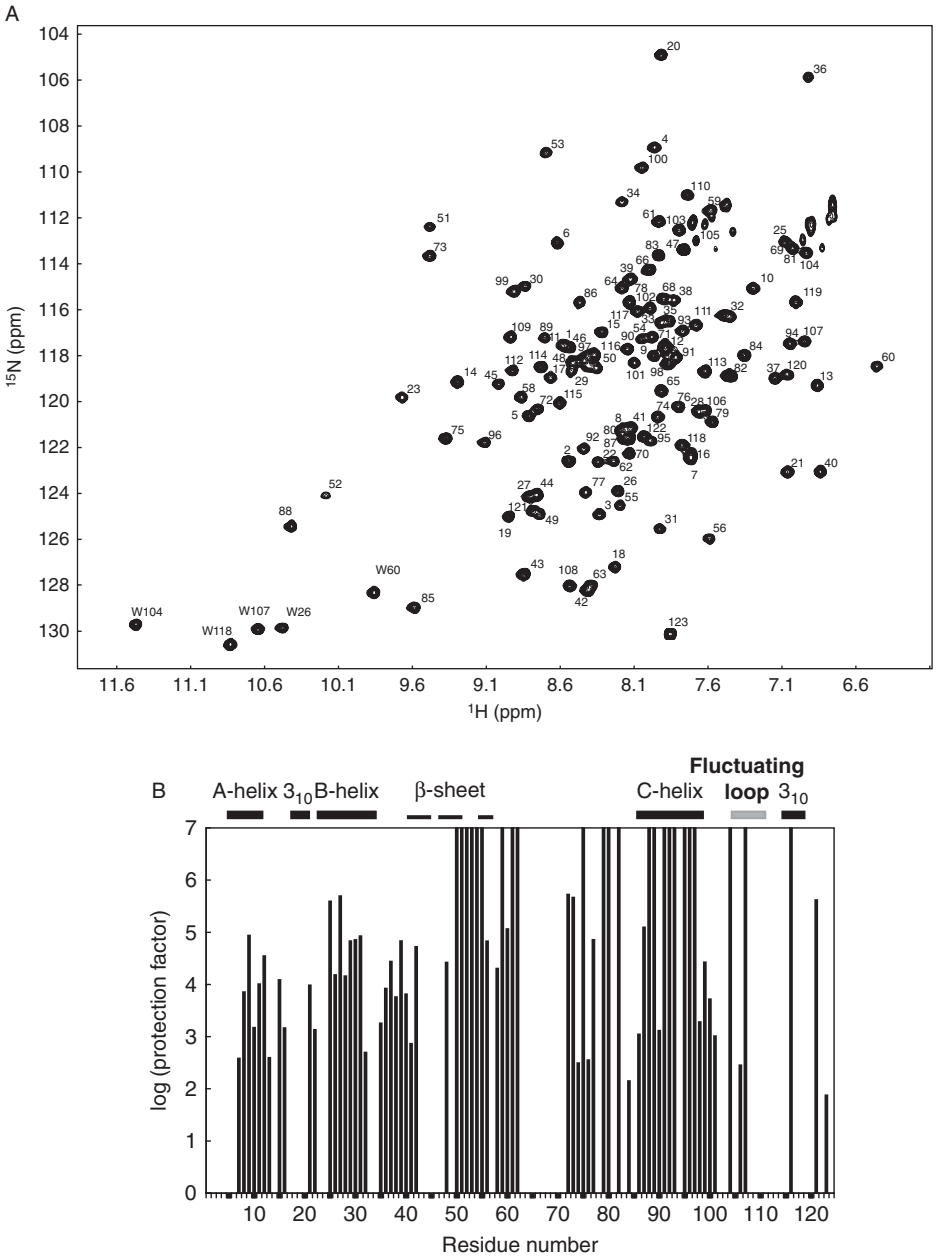


Figure 5 (A) HSQC spectrum of BLAELZ at pH 6.3 and 20°C. Assignments of the backbone resonances for most residues are indicated in the figure; others are omitted for clarity. The indole hydrogens of tryptophan residues are indicated by W. (B) Protection factor versus amino acid residue number. Locations secondary structure elements are indicated.

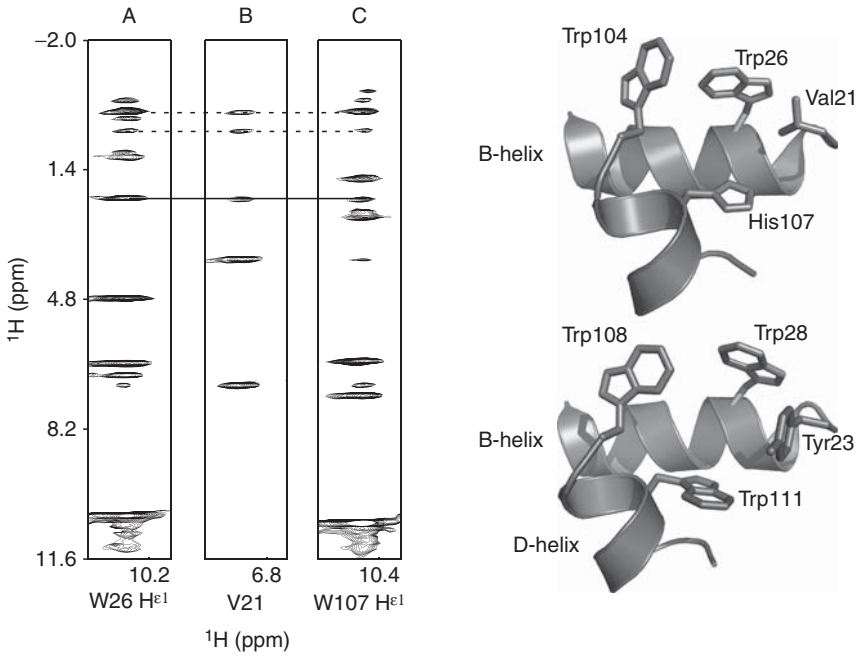


Figure 6 The left side represents the strips taken from the 3D ^{1}H-^{15}N NOESY-HSQC (A and C) and ^{1}H-^{15}N TOCSY-HSQC (B) spectra showing the NOEs between side-chain protons (H^{β} and H^{γ}) of Val21 and side-chain NH protons of Trp residues. H^{β}-$H^{\varepsilon 1}$ and H^{γ}-$H^{\varepsilon 1}$ NOE connectivities are shown by solid and dotted lines, respectively. A mixing time of NOESY spectrum is 100 ms. The right side shows the B- and D-helices of bovine α-LA (upper model) and equine lysozyme (lower model). The figure was produced with a programme PyMOL (available at http://www.pymol.org).

In addition, the order parameters (S^2) of the substituted region of BLAELZ were all higher than 0.8, suggesting a rigid structure in the 105–110 region on the picosecond-to-nanosecond timescale (Figure 7).[46] The S^2 values indicated that the fluctuating loop of BLAELZ is as rigid as the D-helix in hen lysozyme.[47]

There are other significant differences in the protection factors of BLAELZ and bovine α-LA. For instance, although no protection was observed around the C-terminal 3_{10} helix in bovine or human α-LA, Asp116 and Glu121 near the C-terminal 3_{10} helix showed high protection factors in BLAELZ (Figure 5).[37, 38, 46] Higher levels of protection were also observed for the amide protons in the antiparallel β-sheet and A- and B-helices of BLAELZ than in bovine α-LA.[38, 46] Moreover, the substitution results in increases in protection factors around the loop region between the B-helix and antiparallel β-sheet (Figure 5); bovine α-LA has no residue showing a protection factor around the same region.[38, 46] Studies of the hydrogen exchange behaviour of equine lysozyme have shown that the B-helix and the loop between the B-helix and β-sheet exhibit remarkable protection factors.[39, 40]

All of these results on BLAELZ indicate that the substitution of the fluctuating loop in bovine α-LA with the D-helix in equine lysozyme increases the protection

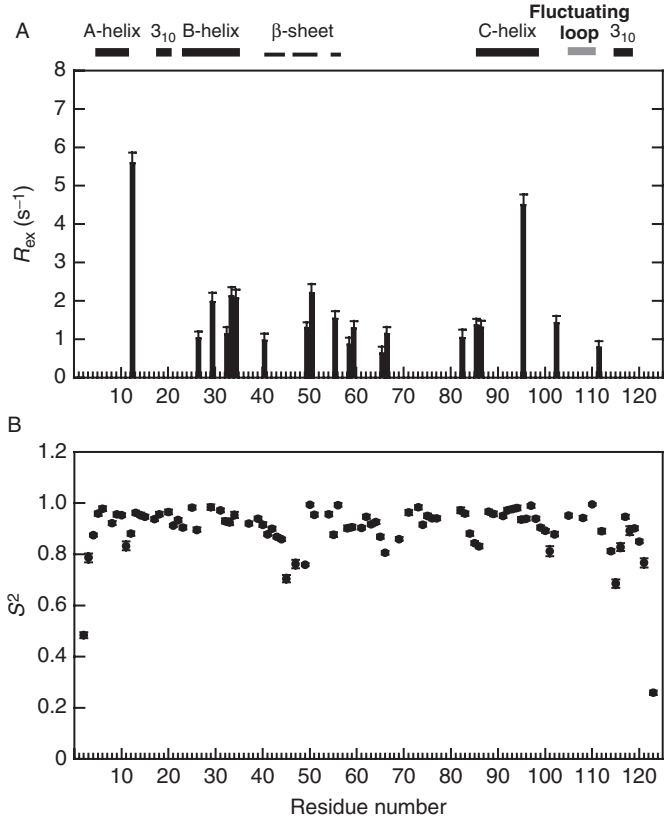


Figure 7 Model-free parameters for BLAELZ. The R_{ex} and S^2 values calculated from T_1, T_2, and NOE are plotted against amino acid residue numbers.

factors throughout the protein molecule (Figure 5). Therefore, the D-helix in equine lysozyme affects the stability of other structural elements, and this effect is responsible for the higher stability of BLAELZ than bovine α-LA. The replacement of a highly fluctuating loop in a protein with a rigid structural element in a homologous one may be one of the ways to construct stable proteins.[46]

2.4. The quenched hydrogen exchange methods for the study of the molten globule

Assignment of the spin systems in NMR spectra is an essential step in the NMR analysis of proteins. The development of indirect detection by heteronuclear NMR and an isotope-labelling technique allowed complex overlapped spectra to be unambiguously assigned for a relatively small protein with a rigid native structure. In the case of the molten globule intermediates, these conventional assignment strategies are usually insufficient because of the lack of resonance dispersion and broad line width. These phenomena arise from the fact that the molten

globule is a conformational ensemble, fluctuating on timescales of intermediate exchange broadening in NMR.[24, 25, 48, 49] The unfolded state is also a conformational ensemble, but the exchange rate is fast, giving rise to a single set of resonances corresponding to an ensemble average. Slow exchange between multiple states gives rise to multiple sets of resonances, which can be resolved or separated by the deconvolution of overlapping resonances if necessary. However, when states are in intermediate exchange, the resonances are broadened, sometimes beyond detection.[24, 25, 48, 49]

To investigate the equilibrium molten globule by the hydrogen–deuterium exchange, the exchanges have to be monitored by an indirect method.[33] The principle of the experiment is to allow exchange to proceed for variable periods of time in the molten-globule state at low pH and then to quench the exchange by raising the pH carefully, so that the protein is refolded to its native state. In this state, a substantial number of the amides that are slow to exchange with solvent give rise to resolved resonances in NMR spectra. By repeating the experiment with various exchange periods in D_2O at pH 2.0, the integrated intensity of these resonances decreases as the exchange period increases, giving an individual rate of exchange for these hydrogens in the molten globule (Figure 8). This method is called the quenched hydrogen exchange method and is widely used for the study of the folding intermediates.

A major advantage of the quenched hydrogen exchange method is that it enables the pattern of exchange kinetics in the molten globule to be probed after the protein has been refolded into its native state, where the resonance assignments are known (Figure 8). Thus, one can avoid the laborious task of reassigning the NMR spectrum for every new set of conditions studied. Moreover, it becomes possible to measure hydrogen exchange rates under conditions that are unsuitable for direct NMR investigation, such as the molten globules with poorly resolved spectra. However, the quenched hydrogen exchange method has two major limitations. First, it provides information only on the location of amide hydrogens that become protected from exchange in the molten globule, so that the nature of the structure that gives rise to protection must be deduced indirectly. For example, hydrogen bond acceptors can only be inferred by comparison with the structure of the folded protein, when we determine hydrogen bond donors from the data of protected amide hydrogens in the molten-globule state. In other words, it can provide information only on the native-like interactions in the molten-globule state, and no information is available on the non-native secondary structure. Second, only well-protected amide hydrogens in the native state can be observed, i.e. those exchanging slowly on the timescale of the 2D NMR measurement (typically 30–50% of the backbone amide hydrogens). A possible solution to this problem might be the use of the dimethyl sulphoxide water mixtures as NMR solvent.[50]

2.5. The molten globule of equine lysozyme

Hydrogen exchange rates have been reported for the molten-globule state of equine lysozyme at 5°C and 25°C.[39, 40] At 5°C, the 17 most protected amide hydrogens are in the regions of the sequence that corresponds to a hydrophobic

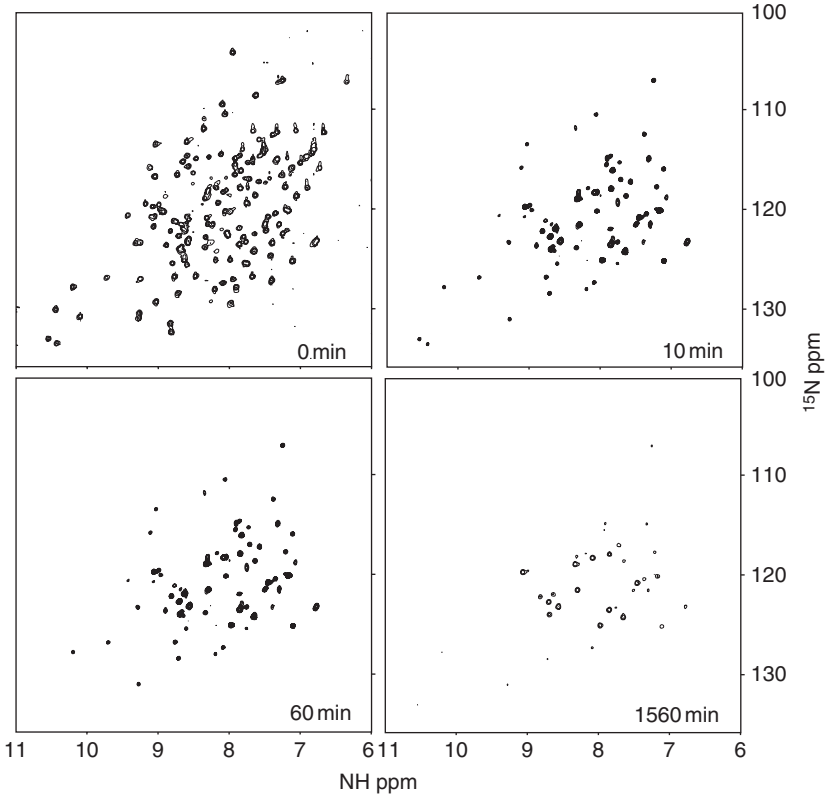


Figure 8 Hydrogen exchange measurements in the molten-globule state based on HSQC spectra in the native state of canine lysozyme as a function of time. After various period of time ranging, the hydrogen exchange reaction in the molten-globule state was quenched by freezing the samples.

cluster in the native state.[39] The most uniformly protected secondary structure element in the molten globule is the B-helix, as in the native state; all amides in the β-sheet region and the C-helix have protection factors less than 10, except for Lys46 ($P = 17.9$) and Ala95 ($P = 22.2$). The indole hydrogens of Trp28, Trp108, and Trp111 are exchanged slowly in the molten globule at both 5°C and 25°C. These three Trp residues are involved in the core of the native structure, indicating the formation of the native-like core in the molten globule.

2.6. The molten globule of canine lysozyme

The HSQC spectrum of the molten globule state of canine lysozyme at pH 2.0 and 30°C is shown in Figure 9. A large number of resolved and sharp resonances, compared with the molten globule of α-LA, are observed for the molten globule of canine lysozyme. This suggests that the molten globule of canine lysozyme does not fluctuate on a microsecond-to-millisecond timescale.

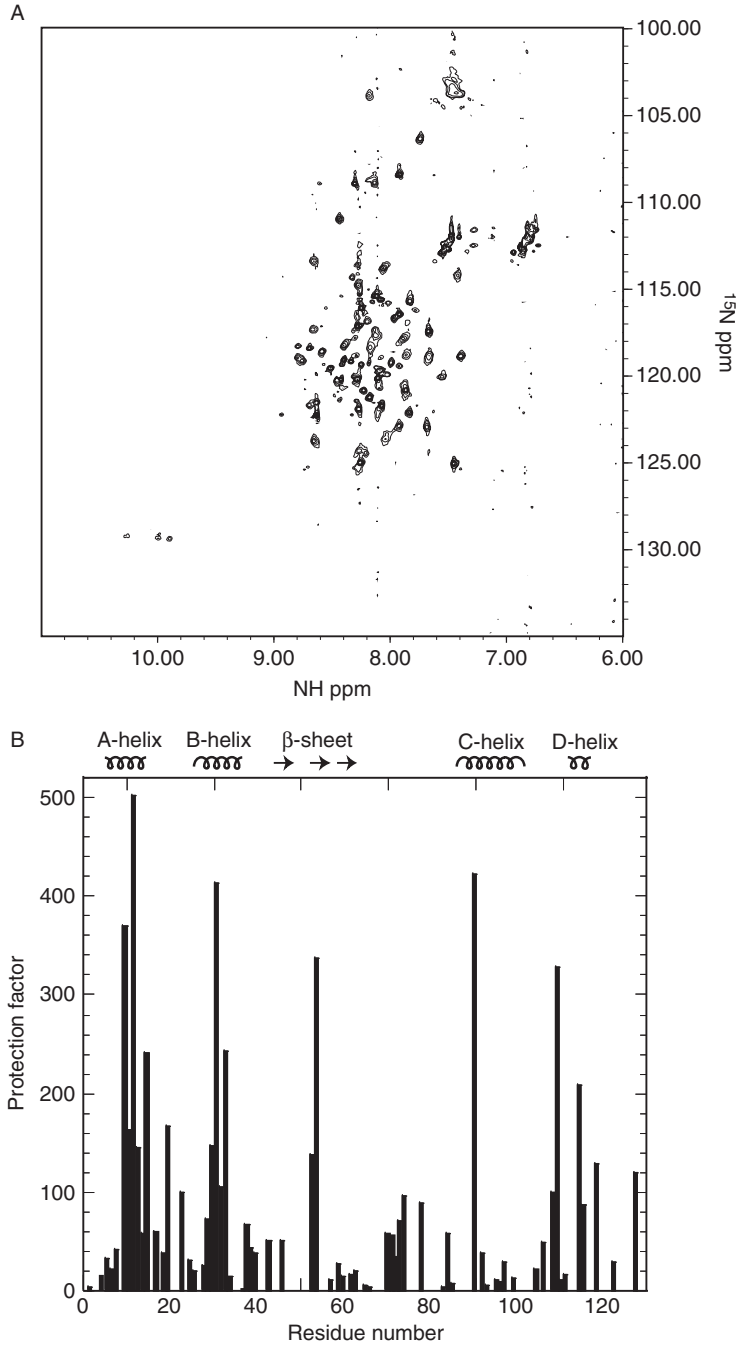


Figure 9 NMR spectrum and protection factor of canine lysozyme at pH 2.0 in the molten-globule state. (A) HSQC spectrum of the molten globule of canine lysozyme. (B) Protection factor versus amino acid residue number.

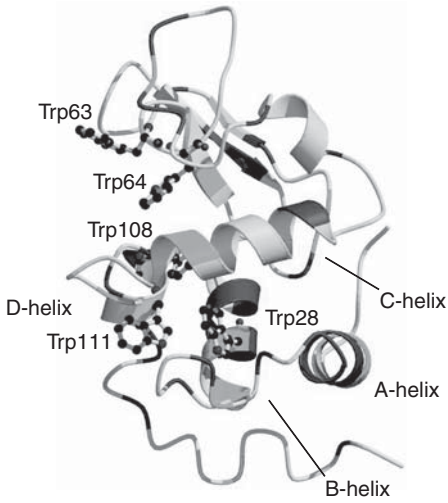


Figure 10 Most protected (shown in black: protection factor >50) and highly protected (shown in grey: protection factor >20) amide hydrogens in the molten-globule state of canine lysozyme.

We investigated the hydrogen exchange rates of the amide hydrogens of canine lysozyme in the molten-globule state at pH 2.0 and 25°C.[41] Sixty-four amide hydrogens were protected from exchange, and this number is much larger than that for equine lysozyme at 5°C (Figure 9). The A-helix is the most protected among secondary structure elements of canine lysozyme. In equine lysozyme, on the other hand, the B-helix is the most protected element in the native and molten-globule states, followed by the A-helix, D-helix, and C-helix. The protection factors are two orders of magnitude larger for canine lysozyme than for equine lysozyme. Figure 10 shows the protected amides in the molten-globule state of canine lysozyme. The most protected amides (protection factor >50) are located mainly in the hydrophobic core region in the native structure. These are the amides from Arg10–Met15 in the A-helix; from Met17 and Phe20 in the loop region between the A- and B-helices; from Val 29-Glu 33 in the B-helix; from Trp108 and Val109 in the D-helix; and from His114, Cys115, Lys118, and Cys127 in the C-terminal region (Figure 10). High protection from hydrogen exchange was also observed for the indole hydrogens of three Trp residues, Trp28, 108, and 111, in the molten globule, suggesting that these residues participate in the hydrophobic core formation as in the native state (Figure 10). The most protected residues of the A- and D-helices are located in the region that forms the interface with the B-helix, indicating that the topology of the A-, B-, and D-helices in the molten-globule state is similar to that in the native protein. Several amides in the β-sheet region also showed high protection factors in the molten globule of canine lysozyme. These are the amides of Phe43 in strand 1, those of Asp53 and Tyr54 in strand 2, and those of Leu59 and Asn60 in strand 3 (Figure 10). The most protected β-strand is strand 2, which is located at the centre of the β-sheet. These results suggest that the molten globule of canine lysozyme maintains the native-like

topology in the β-sheet domain. In the α-lactalbumin molten globule, the β-sheet domain remains largely unstructured, although the α-helical domain adopts a native-like backbone topology.[51, 52]

2.7. The molten globule of α-LA

The hydrogen exchange studies of α-lactalbumin molten globule have shown that amide hydrogens located in the α-domain are highly protected and those in the β-domain are exchanged rapidly.[37, 38] For instance, many residues in the β-domain of bovine α-LA, including Thr48, Glu49, Asn56, Cys77, and Asp82, which are protected in the native state, are no longer protected against exchange in the molten globule. The most protected hydrogen in the molten globule of bovine α-LA at 20°C is from Val92, with a protection factor 11.4. Although the C helix of bovine α-LA shows the highest level of protection, the protection factors for this helix in the molten globule are all small and only marginally higher than those for the rest of the protein. In the molten globule of human α-LA at 5°C, the B-helix contains the amide hydrogens most protected from exchange, with protection factors exceeding 100. The A-helix also persists in the molten globule of human α-LA, with a maximal protection factor of 50. The protection factors for the amide hydrogens in the C-helix of human α-LA are less than 20.

3. STRUCTURE AND DYNAMICS OF UNFOLDED STATES OF PROTEINS

The investigation of the structural information of the unfolded state of a protein is one of the important steps for understanding protein folding. Not only completely unfolded states but also partially unfolded structures of proteins are interesting topics of the protein folding problem. For example, information on residual structures, as well as on the size and dynamics of unfolded and partly folded states of proteins, is very important for understanding the protein folding mechanism from unfolded to folded states. The folded structures are well characterized for a large number of proteins by X-ray diffraction or NMR techniques. On the other hand, the structures of unfolded and partly unfolded states of proteins are not as well understood. Unfolded states of proteins, which are highly heterogeneous and dynamic, are not objects of X-ray crystallography. In the case of solution NMR analysis, one of the most serious obstacles for studying unfolded proteins is their very poor chemical shift dispersion. However, recent developments in high-field NMR spectrometers and new techniques have allowed the characterization of unfolded structures of several proteins.[53–56] In addition, recently intrinsically disordered proteins (also called natively unfolded proteins or intrinsically unstructured proteins) that lack an independently folded state have attracted the attention of researchers.[57, 58]

Unfortunately, it is extremely difficult to study many partially folded molten-globule proteins by NMR because their lines are broad as a result of exchange processes on intermediate timescales. Thus, only a few peaks appear in the ^{15}N-^{1}H

HSQC spectrum of molten-globule state samples, and it is not possible to apply many NMR methods to unfolded proteins directly. To observe the resonances of proteins in molten-globule states directly, methods of increasing the sample temperature or increasing the concentration of denaturant were applied in some studies.[59–62] The sharpening of peaks at higher temperatures in the molten-globule states can provide good information in some cases. With the method of gradually increasing concentrations of denaturant, the unfolding of individual residues in a molten globule also can be determined by using the assignment of unfolded states as a reference.

3.1. Signal assignments of unfolded proteins

Assignment of the spin systems in protein NMR spectra is the first step in the structural analysis of proteins. By the development of 2D NMR experimental methods in the 1970s and 1980s, small folded proteins can be unambiguously assigned by using only ^{1}H homonuclear resonances. However, the chemical shift dispersion of ^{1}H resonances decreases dramatically in the case of unfolded proteins. Thus, the introduction of methods for uniform isotope labelling with ^{13}C and ^{15}N is indispensable for assignment of unfolded proteins. Recently, uniform isotope labelling and triple-resonance pulse sequences have been widely used for assignment of folded proteins. Basically, unfolded states of proteins can be assigned by the same strategies used for assignment of folded proteins labelled with ^{15}N and ^{13}C.[63–65] Even in the unfolded states, the backbone ^{15}N and ^{13}CO chemical shifts are well dispersed because they are influenced by residue type and local amino acid sequence. Fortunately, the flexibility and longer T_2 of the unfolded states generally cause the resonances to be much narrower than those in the folded states. Thus, some pulse sequences that are insensitive for large folded proteins because of magnetization decay during pulse sequences can be utilized for the analysis of unfolded proteins.[66]

In addition to isotope-labelling methods, high sensitivity is also an important factor for the study of unfolded proteins. The solubility of unfolded proteins is generally very low because of the presence of exposed hydrophobic residues. Thus, NMR samples often must be prepared at very low concentrations, and the sensitivity of NMR experiments is crucial.

3.2. Structural analysis of unfolded states

Once NMR signals of unfolded proteins have been assigned, a number of NMR parameters can be used to analyse the structure in detail.[54, 55] A lot of structural and dynamic information about each residue of proteins is extracted from NMR resonances: chemical shifts, coupling constants, temperature coefficients, hydrogen exchange rates, NOE, and relaxation time. As for backbone structure, the deviations in chemical shifts from random coil values provide very convenient and sensitive probes of secondary structure without extra experiments in addition to resonance assignment. The secondary shifts of chemicals shifts of $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO, and $^{1}H\alpha$ from random values provide information on residual secondary

structure in unfolded states. Coupling constants also provide information about backbone conformation, although unfolded states of proteins are highly flexible and all parameters are a population-weighted average over all structural ensembles.[67] Remarkable non-averaged coupling constants clearly suggest residual structures in unfolded proteins.

One of the powerful methods for studying unfolded states is the aforementioned hydrogen exchange experiments. In general, the NMR measurements of hydrogen exchange methods are performed after the protein has been refolded into its native state. Thus, chemical shift assignments of unfolded states are not needed for this kind of experiment. However, recently developed DMSO-quenched hydrogen exchange methods require the assignment of denatured proteins in DMSO.[68, 69] This novel method makes it possible to probe the structured regions of water-insoluble proteins such as amyloid fibrils, and such information is very difficult to obtain by using other methods.

The NOE provides important distance information on highly structured formation in unfolded states, although the severe overlap of the ^{1}H resonances limits the number of NOE peaks that can be assigned. The NOEs between sequential amino acid residues can provide useful information on the local polypeptide backbone conformational preferences, particularly the presence of a secondary structure. Unfortunately it is extremely difficult to observe long-range interactions in unfolded proteins because of their intrinsic flexibility and conformational averaging. To solve this problem, paramagnetic relaxation enhancement (PRE) has been used to provide information on transient tertiary contacts.[70, 71] This method can detect ensemble-averaged long-range contacts of about 20 Å by using covalently attached paramagnetic nitroxide spin labels. The recent development of residual dipolar couplings (RDC) measured in partially aligned media has the potential to provide information on the overall structure of unfolded states without NOE data or spin-label experiments.[72, 73]

3.3. Dynamic studies of unfolded proteins

The information about the backbone motion of each residue in folded and unfolded proteins can be analysed by ^{15}N relaxation experiments. In the case of folded proteins, the measurement of T_1, T_2, and heteronuclear NOE and model-free analysis is widely used for studying polypeptide chain dynamics. However, the basic assumption of model-free analysis – isotropic tumbling with a single molecular correlation time – is invalid for unfolded proteins. Thus, an extension of the model-free analysis with a distribution of correlation times has been developed.[74] To avoid the assumptions of model-free analysis, reduced spectral density mapping has also been used to investigate the backbone dynamics of unfolded proteins.[75]

3.4. Studies of unfolded states of lysozyme and α -LA

In the cases of lysozyme and α-lactalbumin, the structures of different unfolded states have been investigated by NMR. Especially, hen lysozyme is constantly subjected to new methods also in this field as a model protein. Almost completely

sequential assignments for ^{1}H NMR resonances of hen lysozyme in the native state were reported by Redfield and Dobson in 1988.[76] On the other hand, structural studies of partially denatured states of hen lysozyme in trifluoroethanol (TFE) were reported in 1995 by using ^{15}N 3D NMR experiments.[77] In addition, the completely denatured structures of oxidized and reduced hen lysozyme in 8M urea at low pH were investigated extensively by Schwalbe et al. in 1997.[78] In the former study of partially folded states of lysozyme in 70% TFE solutions at pH 2.0, ^{15}N-filtered 3D NOESY and TOCSY measurements were applied for the sequential assignments, and all but 2 of the 126 main-chain amide nitrogen atoms were assigned. By using the pattern of medium-range NOEs, $^{1}H^{\alpha}$ chemical shift perturbations, coupling constants, amide hydrogen exchange rates,[77] and relaxation measurements,[79] six regions of polypeptide chain were observed to be stabilized in helical conformations. These regions include the four α-helices and 3_{10} helices formed in the native state. Moreover, the pattern of main-chain dynamics is similar to that of the native states despite the absence of extensive tertiary interactions. In the case of the completely denatured states of hen lysozyme in 8M urea, ^{15}N-resolved 3D NMR experiments were the first to provide almost complete sequential assignment for backbones ^{1}H and ^{15}N resonances, not only in the presence of the four native disulfide bridges but also in the absence of such bridges.[78] Structural analysis by NMR measurements showed that the local conformations of denatured states are largely determined by short-range interactions within the polypeptide chain. However, some conformational parameters and the relaxation data showed the possibility of the clustering of hydrophobic residues.

The more recent advances in NMR analysis have shown the long-range interactions in hen lysozyme even in the denatured states. In a detailed analysis using ^{13}C and ^{15}N-labelled hen lysozyme, the deviations in the torsion angle χ_1 of side chains were reported for many of the aromatic residues.[80] In addition, the combination of the NMR analysis and site-directed mutagenesis of the key hydrophobic residues for long-range interaction in denatured states showed that the non-native interactions stabilize native-like hydrophobic clusters.[81, 82] To clarify the hydrophobic interactions between clusters of tryptophane residues, new NMR experiments were developed to assign each tryptophane side-chain indole resonance and applied to hen lysozyme in the unfolded state.[83]

In the case of α-LA, the completely unfolded states were studied in the context of partially unfolded molten-globule states of this protein.[59–62] Although the partially folded molten-globule states of α-LA at low pH were shown by various methods to have a native-like secondary structure, high-resolution structural information of individual residues was restricted because of extreme line broadening in NMR spectra. In 1997, by utilizing the assignment of unfolded states of human α-LA, the unfolding of individual residues from a molten-globule state was reported.[62] All expected resonances of α-LA were reported to be clearly resolved and assigned in the NMR spectrum under extreme denaturing conditions. From the molten globule to the unfolded states, the sharpening of specific resonances as a function of denaturant concentration was observed. NMR resonances at each denaturant condition showed that the human α-lactalbumin molten globule unfolds non-cooperatively. This method was applied to studies of the

compact molten-globule structure of human α-LA in the absence of disulfide bonds[59, 61] and to a comparison of the denaturant unfolding of bovine and human α-LA.[60] Recently, the structures of unfolded states of α-LA and homologous lysozyme were compared.[84]

4. HYDROGEN EXCHANGE PULSE LABELLING AND REAL-TIME NMR

4.1. Hydrogen exchange pulse labelling

Hydrogen exchange pulse labelling is an especially powerful method for obtaining site-specific information on kinetic intermediates that are transiently populated in the course of refolding.[31, 32, 35, 42, 85] A typical experiment involves rapid dilution of denatured protein in H_2O buffer to initiate refolding; the protein is allowed to refold for a short period (typically milliseconds to seconds) before mixing with a high-pH labelling buffer in D_2O solution. Amide hydrogens involved in hydrogen-bonded structures during the refolding period are protected from exchange, whereas amide hydrogens in the unfolded regions are exchanged with deuteriums by the labelling pulse. After the exchange is quenched and the folding is completed, the protected amides are identified by using 2D NMR spectra. By repeating the experiment with various refolding periods, progressive stabilization of the hydrogen-bonded structure can be monitored. A major advantage of hydrogen exchange pulse labelling is its improved time resolution compared to real-time methods, although it is not without limitations, as noted in the quenched hydrogen exchange experiments mentioned above.

The hydrogen exchange pulse labelling method has successfully been applied to the folding studies of hen lysozyme.[34] More than 50 amides were analysed, giving rise to a detailed view of the development of the hydrogen-bonded structure throughout the protein. About half of the monitored amides were completely protected from exchange within 200 ms, while the other half were not completely protected until more than 1 s. The more rapidly protected amides are located in the α-domain, whereas the more slowly protected ones are in the β-domain. Therefore, the α-domain folds faster than the β-domain in the major pathway, indicating that each domain becomes stabilized with different kinetics. Furthermore, detailed analysis of the pulsed hydrogen exchange data suggests that the folding pathway of hen lysozyme involves multiple parallel pathways.[34]

4.2. Real-time NMR

Ideally, the folding of a protein could be followed by successive acquisition of NMR spectra while the protein refolds. Although this approach is possible in principle, in practice it will not occur in most cases, primarily because the timescale of data acquisition is much slower than the timescale of protein folding. Furthermore, 2D NMR is needed to monitor the structural changes at site-specific resolution. Therefore, reducing the rate of folding to the timescale of

minutes to hours is important for the real-time 2D NMR experiments. Folding coupled with proline isomerization is slow enough to apply the real-time 2D NMR technique.[6, 86–89]

In the pioneering work of real-time NMR, 1D NMR was utilized to monitor protein folding. In 1993, Koide et al. reported the kinetics of the refolding reaction of apoplastocyanin coupled with proline isomerization by continuous recording of 1D NMR.[86] In addition to the folding kinetics, real-time NMR enables characterization of a kinetic intermediate accumulated within the dead time of real-time NMR experiments. The intermediate of apoplastocyanin with a non-native proline-bond is characterized by non-native and specific tertiary structures in a localized region.

There have been a number of real-time NMR reports in which the refolding of proteins was directly monitored by successive acquisitions of 2D NMR spectra.[2, 6, 25, 87–90] The real-time 2D NMR method was particularly powerful for the characterization of the rate-limiting intermediate before refolding to the native state. The rate-limiting intermediates detected by the real-time 2D NMR can be classified into two categories: (i) a conformational ensemble with dynamic fluctuations occurring on a microsecond-to-millisecond timescale and (ii) a native-like intermediate having substantial rigid side-chain packing. The first type is the molten globule-like intermediate, which gives rise to poorly resolved resonances, the majority being broad and clustered between 7.8–8.5 ppm (^{1}H) and 117–122 ppm (^{15}N).[25, 88] The second type is a native-like intermediate, which gives rise to many well-dispersed resonances with native-like chemical shifts.[6, 89]

Balbach et al. utilized real-time 1D NMR to monitor the refolding of apo bovine α-LA.[24] Although bovine α-LA folds with a time constant of about 40 ms in the presence of Ca^{2+}, the folding rate is decreased by up to three orders of magnitude in the refolding of the apo form.[91] The NMR spectrum of the kinetic intermediate has been shown to resemble closely that of the equilibrium molten globule formed at low pH. In both cases, the chemical shift dispersion is small compared to that of the native state, indicating the lack of a persistent tertiary structure. More importantly, in both cases substantial line broadening is evident, indicating conformational fluctuations on a microsecond-to-millisecond timescale.[24]

By monitoring the well-resolved resonances in the α-domain and in the β-domain, Balbach et al. demonstrated that the native tertiary structure emerges simultaneously in both domains.[24] This is particularly interesting in light of the conclusion that the equilibrium molten globule of α-LA has an ordered secondary structure in the α-domain, with the β-domain being less organized.[51, 52]

Recently, Schanda et al. utilized band-selective optimized flip-angle short-transient (SOFAST) real-time 2D NMR spectroscopy to monitor the folding process of bovine α-LA.[92] The folding of bovine α-LA occurs cooperatively from the molten globule, which is a conformational ensemble interconverting on the microsecond-to-millisecond timescale. By combining SOFAST real-time 2D NMR and rapid sample mixing inside the NMR magnet to initiate the folding, resonances from a large number of amides throughout the sequence can be used as probes of the conformational transition from the molten-globule state to the native state.

5. CONCLUDING REMARKS

Our knowledge of protein folding has increased dramatically over the past two decades, and many of the underlying principles are becoming clear in outline. NMR is particularly powerful when applied to dynamic or flexible systems, such as molten globules and unfolded proteins, which are not usually amenable to X-ray crystallography. We have highlighted here the use of NMR to study the protein folding of α-LA and lysozyme. NMR will continue to be an essential tool for structural characterization during the folding process and will provide detailed contributions for joining experiment and theory.

REFERENCES

1. C. M. Dobson, P. A. Evans and S. E. Radford, *Trends Biochem. Sci.*, 1994, **19,** 31.
2. C. M. Dobson and P. J. Hore, *Nat. Struct. Biol.*, 1998, **5,** 504.
3. H. J. Dyson and P. E. Wright, *Adv. Protein Chem.*, 2002, **62,** 311.
4. H. J. Dyson and P. E. Wright, *Chem. Rev.*, 2004, **104,** 3607.
5. C. Redfield, *Methods*, 2004, **34,** 121.
6. M. Zeeb and J. Balbach, *Methods*, 2004, **34,** 65.
7. D. M. Korzhnev, X. Salvatella, M. Vendruscolo, A. A. Di Nardo, A. R. Davidson, C. M. Dobson and L. E. Kay, *Nature*, 2004, **430,** 586.
8. P. Neudecker, A. Zarrine-Afsar, W. Y. Choy, D. R. Muhandiram, A. R. Davidson and L. E. Kay, *J. Mol. Biol.*, 2006, **363,** 958.
9. P. Neudecker, A. Zarrine-Afsar, A. R. Davidson and L. E. Kay, *Proc. Natl. Acad. Sci. USA*, 2007, **104,** 15717.
10. B. S. Khatra, D. G. Herries and K. Brew, *Eur. J. Biochem.*, 1974, **44,** 537.
11. J. E. Bell, T. A. Beyer and R. L. Hill, *J. Biol. Chem.*, 1976, **251,** 3003.
12. K. Brew, T. C. Vanaman and R. L. Hill, *J. Biol. Chem.*, 1967, **242,** 3747.
13. L. Hall, R. K. Craig, M. R. Edbrooke and P. N. Campbell, *Nucleic Acids Res.*, 1982, **10,** 3503.
14. K. R. Acharya, D. I. Stuart, N. P. Walker, M. Lewis and D. C. Phillips, *J. Mol. Biol.*, 1989, **208,** 99.
15. K. Nitta and S. Sugai, *Eur. J. Biochem.*, 1989, **182,** 111.
16. K. Nitta, H. Tsuge, S. Sugai and K. Shimazaki, *FEBS Lett.*, 1987, **223,** 405.
17. A. C. Pike, K. Brew and K. R. Acharya, *Structure*, 1996, **4,** 691.
18. H. Tsuge, H. Ago, M. Noma, K. Nitta, S. Sugai and M. Miyano, *J. Biochem.*, 1992, **111,** 141.
19. T. Koshiba, M. Yao, Y. Kobashigawa, M. Demura, A. Nakagawa, I. Tanaka, K. Kuwajima and K. Nitta, *Biochemistry*, 2000, **39,** 3248.
20. K. Kuwajima, Y. Hiraoka, M. Ikeguchi and S. Sugai, *Biochemistry*, 1985, **24,** 874.
21. M. Arai and K. Kuwajima, *Fold. Des.*, 1996, **1,** 275.
22. M. Mizuguchi, M. Arai, Y. Ke, K. Nitta and K. Kuwajima, *J. Mol. Biol.*, 1998, **283,** 265.
23. M. Nakao, K. Maki, M. Arai, T. Koshiba, K. Nitta and K. Kuwajima, *Biochemistry*, 2005, **44,** 6685.
24. J. Balbach, V. Forge, N. A. van Nuland, S. L. Winder, P. J. Hore and C. M. Dobson, *Nat. Struct. Biol.*, 1995, **2,** 865.
25. J. Balbach, V. Forge, W. S. Lau, N. A. van Nuland, K. Brew and C. M. Dobson, *Science*, 1996, **274,** 1161.
26. Y. Bai, J. S. Milne, L. Mayne and S. W. Englander, *Proteins*, 1993, **17,** 75.
27. S. W. Englander and N. R. Kallenbach, *Q. Rev. Biophys.*, 1983, **16,** 521.
28. G. Wagner and K. Wuthrich, *J. Mol. Biol.*, 1982, **160,** 343.
29. T. G. Pedersen, B. W. Sigurskjold, K. V. Andersen, M. Kjaer, F. M. Poulsen, C. M. Dobson and C. Redfield, *J. Mol. Biol.*, 1991, **218,** 413.
30. J. S. Milne, L. Mayne, H. Roder, A. J. Wand and S. W. Englander, *Protein Sci.*, 1998, **7,** 739.
31. H. Roder, G. A. Elove and S. W. Englander, *Nature*, 1988, **335,** 700.

32. J. B. Udgaonkar and R. L. Baldwin, *Nature*, 1988, **335,** 694.
33. F. M. Hughson, P. E. Wright and R. L. Baldwin, *Science*, 1990, **249,** 1544.
34. S. E. Radford, C. M. Dobson and P. A. Evans, *Nature*, 1992, **358,** 302.
35. P. A. Jennings and P. E. Wright, *Science*, 1993, **262,** 892.
36. M. Buck, S. E. Radford and C. M. Dobson, *J. Mol. Biol.*, 1994, **237,** 247.
37. B. A. Schulman, C. Redfield, Z. Y. Peng, C. M. Dobson and P. S. Kim, *J. Mol. Biol.*, 1995, **253,** 651.
38. V. Forge, R. T. Wijesinha, J. Balbach, K. Brew, C. V. Robinson, C. Redfield and C. M. Dobson, *J. Mol. Biol.*, 1999, **288,** 673.
39. L. A. Morozova, D. T. Haynie, C. Arico-Muendel, H. Van Dael and C. M. Dobson, *Nat. Struct. Biol.*, 1995, **2,** 871.
40. L. A. Morozova-Roche, C. C. Arico-Muendel, D. T. Haynie, V. I. Emelyanenko, H. Van Dael and C. M. Dobson, *J. Mol. Biol.*, 1997, **268,** 903.
41. Y. Kobashigawa, M. Demura, T. Koshiba, Y. Kumaki, K. Kuwajima and K. Nitta, *Proteins*, 2000, **40,** 579.
42. S. E. Radford, M. Buck, K. D. Topping, C. M. Dobson and P. A. Evans, *Proteins*, 1992, **14,** 237.
43. Y. V. Griko, E. Freire, G. Privalov, H. van Dael and P. L. Privalov, *J. Mol. Biol.*, 1995, **252,** 447.
44. K. Harata and M. Muraki, *J. Biol. Chem.*, 1992, **267,** 1419.
45. M. Mizuguchi, K. Masaki and K. Nitta, *J. Mol. Biol.*, 1999, **292,** 1137.
46. M. Tada, Y. Kobashigawa, M. Mizuguchi, K. Miura, T. Kouno, Y. Kumaki, M. Demura, K. Nitta and K. Kawano, *Biochemistry*, 2002, **41,** 13807.
47. M. Buck, J. Boyd, C. Redfield, D. A. MacKenzie, D. J. Jeenes, D. B. Archer and C. M. Dobson, *Biochemistry*, 1995, **34,** 4041.
48. J. Baum, C. M. Dobson, P. A. Evans and C. Hanley, *Biochemistry*, 1989, **28,** 7.
49. A. T. Alexandrescu, P. A. Evans, M. Pitkeathly, J. Baum and C. M. Dobson, *Biochemistry*, 1993, **32,** 1707.
50. C. Nishimura, H. J. Dyson and P. E. Wright, *Proc. Natl. Acad. Sci USA*, 2005, **102,** 4765.
51. Z. Y. Peng and P. S. Kim, *Biochemistry*, 1994, **33,** 2136.
52. L. C. Wu, Z. Y. Peng and P. S. Kim, *Nat. Struct. Biol.*, 1995, **2,** 281.
53. D. R. Shortle, *Curr. Opin. Struct. Biol.*, 1996, **6,** 24.
54. H. J. Dyson and P. E. Wright, *Methods Enzymol.*, 2001, **339,** 258.
55. H. J. Dyson and P. E. Wright, *Methods Enzymol.*, 2005, **394,** 299.
56. T. Mittag and J. D. Forman-Kay, *Curr. Opin. Struct. Biol.*, 2007, **17,** 3.
57. H. J. Dyson and P. E. Wright, *Nat. Rev. Mol. Cell Biol.*, 2005, **6,** 197.
58. A. L. Fink, *Curr. Opin. Struct. Biol.*, 2005, **15,** 35.
59. C. M. Quezada, B. A. Schulman, J. J. Froggatt, C. M. Dobson and C. Redfield, *J. Mol. Biol.*, 2004, **338,** 149.
60. R. Wijesinha-Bettoni, C. M. Dobson and C. Redfield, *J. Mol. Biol.*, 2001, **312,** 261.
61. C. Redfield, B. A. Schulman, M. A. Milhollen, P. S. Kim and C. M. Dobson, *Nat. Struct. Biol.*, 1999, **6,** 948.
62. B. A. Schulman, P. S. Kim, C. M. Dobson and C. Redfield, *Nat. Struct. Biol.*, 1997, **4,** 630.
63. T. M. Logan, E. T. Olejniczak, R. X. Xu and S. W. Fesik, *J. Biomol. NMR*, 1993, **3,** 225.
64. A. T. Alexandrescu, C. Abeygunawardana and D. Shortle, *Biochemistry*, 1994, **33,** 1063.
65. V. L. Arcus, S. Vuilleumier, S. M. Freund, M. Bycroft and A. R. Fersht, *Proc. Natl. Acad. Sci. USA*, 1994, **91,** 9412.
66. D. Eliezer, J. Yao, H. J. Dyson and P. E. Wright, *Nat. Struct. Biol.*, 1998, **5,** 148.
67. H. J. Dyson and P. E. Wright, *Annu. Rev. Biophys. Biophys. Chem.*, 1991, **20,** 519.
68. M. Hoshino, H. Katou, Y. Hagihara, K. Hasegawa, H. Naiki and Y. Goto, *Nat. Struct. Biol.*, 2002, **9,** 332.
69. M. Hoshino, H. Katou, K. Yamaguchi and Y. Goto, *Biochim. Biophys. Acta*, 2007, **1768,** 1886.
70. J. R. Gillespie and D. Shortle, *J. Mol. Biol.*, 1997, **268,** 158.
71. J. R. Gillespie and D. Shortle, *J. Mol. Biol.*, 1997, **268,** 170.
72. D. Shortle and M. S. Ackerman, *Science*, 2001, **293,** 487.
73. M. S. Ackerman and D. Shortle, *Biochemistry*, 2002, **41,** 3089.
74. A. V. Buevich, U. P. Shinde, M. Inouye and J. Baum, *J. Biomol. NMR*, 2001, **20,** 233.
75. J. Yao, J. Chung, D. Eliezer, P. E. Wright and H. J. Dyson, *Biochemistry*, 2001, **40,** 3561.

76. C. Redfield and C. M. Dobson, *Biochemistry*, 1988, **27,** 122.
77. M. Buck, H. Schwalbe and C. M. Dobson, *Biochemistry*, 1995, **34,** 13219.
78. H. Schwalbe, K. M. Fiebig, M. Buck, J. A. Jones, S. B. Grimshaw, A. Spencer, S. J. Glaser, L. J. Smith and C. M. Dobson, *Biochemistry*, 1997, **36,** 8977.
79. M. Buck, H. Schwalbe and C. M. Dobson, *J. Mol. Biol.*, 1996, **257,** 669.
80. M. Hennig, W. Bermel, A. Spencer, C. M. Dobson, L. J. Smith and H. Schwalbe, *J. Mol. Biol.*, 1999, **288,** 705.
81. J. Klein-Seetharaman, M. Oikawa, S. B. Grimshaw, J. Wirmer, E. Duchardt, T. Ueda, T. Imoto, L. J. Smith, C. M. Dobson and H. Schwalbe, *Science*, 2002, **295,** 1719.
82. J. Wirmer, C. Schlorb, J. Klein-Seetharaman, R. Hirano, T. Ueda, T. Imoto and H. Schwalbe, *Angew Chem. Int. Ed. Engl.*, 2004, **43,** 5780.
83. C. Schlorb, K. Ackermann, C. Richter, J. Wirmer and H. Schwalbe, *J. Biomol. NMR*, 2005, **33,** 95.
84. J. Wirmer, H. Berk, R. Ugolini, C. Redfield and H. Schwalbe, *Protein Sci.*, 2006, **15,** 1397.
85. M. M. Krishna, L. Hoang, Y. Lin and S. W. Englander, *Methods*, 2004, **34,** 51.
86. S. Koide, H. J. Dyson and P. E. Wright, *Biochemistry*, 1993, **32,** 12299.
87. C. Steegborn, H. Schneider-Hassloff, M. Zeeb and J. Balbach, *Biochemistry*, 2000, **39,** 7910.
88. M. Mizuguchi, G. J. Kroon, P. E. Wright and H. J. Dyson, *J. Mol. Biol.*, 2003, **328,** 1161.
89. A. Kameda, M. Hoshino, T. Higurashi, S. Takahashi, H. Naiki and Y. Goto, *J. Mol. Biol.*, 2005, **348,** 383.
90. S. Iimura, H. Yagi, K. Ogasahara, H. Akutsu, Y. Noda, S. Segawa and K. Yutani, *Biochemistry*, 2004, **43,** 11906.
91. K. Kuwajima, M. Mitani and S. Sugai, *J. Mol. Biol.*, 1989, **206,** 547.
92. P. Schanda, V. Forge and B. Brutscher, *Proc. Natl. Acad. Sci. USA*, 2007, **104,** 11257.

CHAPTER 3

Applications of the NOE in Molecular Biology

Mike P. Williamson

Contents

Abstract

The nuclear Overhauser effect (NOE) arises from dipole–dipole relaxation between two spin-1/2 nuclei. It is therefore dependent on the distance between the nuclei and their motions. The NOE is used widely in molecular biology to measure distances, and thus to calculate structures; and also as one of a set of measurements to assess intermolecular motion. In this review, we concentrate on more recent developments of the NOE: particularly on

Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, United Kingdom

Annual Reports on NMR Spectroscopy, Volume 65
ISSN 0066-4103, DOI: 10.1016/S0066-4103(08)00203-2

progress towards automation of NOE collection and calculation of macromolecular structures; on applications to study complexes and ligand binding; and on applications to hydration of macromolecules.

Key Words: NOE, Protein, Relaxation, Structure calculation, Ligand binding, Solvation.

1. INTRODUCTION

The NOE is a widely used tool within molecular biology, and it is therefore not feasible or helpful to cover every application. The two main areas in which the NOE is applied are for the generation of distance restraints in structure calculation, and as one of several measurements used to study macromolecular dynamics. Undoubtedly, the main role of the NOE is in providing distance restraints. Despite a great deal of work on other sources of structural restraints in recent years, in particular residual dipolar couplings (RDCs) and chemical shifts (which at one point almost promised to supplant NOEs as the principal source of structural information in proteins[1]), NOEs remain the predominant source of structural information. We start from the assumption that all readers will have at least some familiarity with the NOE and its main applications, and we therefore begin by covering very briefly the more important aspects of NOE theory, emphasising those that will be important later. We then move on to cover each of the main areas of recent interest in turn.

In writing the review, we have had to consider where to draw the boundary of what is covered. The NOE is a phenomenon of dipolar relaxation, arising from the time dependence of dipolar coupling. We, however, felt that detailed discussion of dipolar coupling, either as applied to solids or as manifested in RDCs in liquids, was not warranted. Reviews on these topics can be found elsewhere, not least within this series.[2]

There have of course been several reviews[3–7] as well as two books[8,9] on the NOE. For more details particularly on the theory, readers are referred to the more recent of the two books,[8] or to a shorter review.[10] This review mainly covers developments since the publication of the book in 2000.

2. THE BASICS

2.1. NOE theory

Nuclei *in vacuo* relax extremely slowly. There are many possible sources of relaxation for molecules in a liquid or solid, but the most important of these is dipolar relaxation; that is, the magnetic field of a neighbouring nucleus acts as a source of relaxation. For many purposes, it helps to think of nearby moving nuclei producing a moving magnetic field, which effectively acts as a little local rf pulse, rotating the nuclear magnetisation and thereby causing it macroscopically to change (i.e., the summation of many little local pulses on individual magnetisation vectors averages out to a tendency for the macroscopic vector to move in a

particular way, often to relax towards equilibrium). This effect is commonly described as *cross-relaxation*, because it arises from one spin relaxing another: spin I relaxes spin S, and conversely, spin S also relaxes spin I.

It is, therefore, not surprising that this relaxation is distance-dependent: the rate of relaxation that produces the NOE is proportional to r^{-6}, where r is the internuclear distance. In more detail, for a situation in which I is initially at equilibrium and S is not, then the cross-relaxation rate is given by

$$\mathrm{d}I/\mathrm{d}t = -\sigma_{IS}(S - S^0) \tag{1}$$

and

$$\sigma_{IS} = \frac{1}{10}K^2\tau_c\left[\frac{6}{1+(\omega_I+\omega_S)^2\tau_c^2} - \frac{1}{1+(\omega_I-\omega_S)^2\tau_c^2}\right] \tag{2}$$

where $K = (\mu_0/4\pi)\gamma_I\gamma_S/r_{IS}^3$, and $(\mu_0/4\pi)$ is a proportionality constant to get everything into the correct units, $\hbar$ is Planck's constant divided by 2π, γ_I and γ_S are the gyromagnetic ratios for nuclei I and S, r_{IS} is the internuclear distance, ω_I and ω_S are the Larmor precession frequencies of nuclei I and S, and τ_c is the correlation time of the IS vector, meaning the characteristic time for reorientation of the vector and thus approximately the time taken for it to move by one radian. Thus the relaxation rate depends on the type of nuclei involved, on their distance, and on the correlation time for the IS vector. Because of the ω_I and ω_S terms, it also depends on the frequency of the NMR signals, and thus on the field of the spectrometer used. We will consider each of these effects.

The simplest case to consider is one where there is only a single correlation time throughout the molecule, or in other words where the molecule is rigid, and its motion is isotropic, that is, it tumbles at the same rate in all directions, which implies that it must be close to spherical. If we consider ^{1}H-^{1}H relaxation for proteins in a 600 MHz spectrometer (for which τ_c is of the order of 10 ns), then ω_I and ω_S are very similar and are both close to $6 \times 10^8 \times 2\pi$ or 4×10^9, implying that the first term in the bracket in Equation (2) is negligible compared to the second. This means that the cross-relaxation rate is negative (a 'negative NOE').[1] There are a number of consequences of this, but the most striking is that the magnetisation vectors of I and S relax together but in opposite directions, that is, the intensities of I and S converge. This is an intuitively easy situation to understand, and means that cross-relaxation leads to magnetisation being spread around the molecule essentially in a diffusive process. This process is known as *spin diffusion* and will be discussed in more detail below.

For all protons in the rigid protein, we can therefore simplify the above equation to

$$\sigma_{IS} = cr_{IS}^{-6} \tag{3}$$

[1] Somewhat confusingly, a negative cross-relaxation rate (a negative NOE, as found in proteins) corresponds to a peak in a 2D NOESY spectrum of the same sign as the diagonal (a positive peak as normally drawn) whereas a positive cross-relaxation rate (a positive NOE, as found in small molecules and also in ROESY) corresponds to a NOESY peak of opposite sign to the diagonal.

Thus, the cross-relaxation rate is very simply proportional to r^{-6}. The cross-relaxation rate is normally measured as the intensity of a NOESY cross peak at short mixing time (the time delay during the NOESY sequence during which NOEs build up), from which we conclude that the intensity of a NOESY cross peak is proportional to r^{-6}. This then allows us to measure distances within the molecule, as long as we have some reference calibration distance with a measured NOE intensity A_{ref}:

$$r_{IS} = r_{\text{ref}}[A_{IS}/A_{\text{ref}}]^{-1/6} \tag{4}$$

This is the basis on which all distance measurements are made using the NOE. The best reference distance to use is one that roughly matches the distance to be measured. In particular, use of a very short reference distance, such as that between two methylene protons in a CH_2 group, leads to errors and should be avoided.

Of course, if the protein is not rigid or not isotropic, then this equation is no longer valid. In fact, the isotropic requirement is usually not a problem, since most proteins rotate isotropically to within a factor of 2, and the $r^{-1/6}$ dependence means that this only introduces an error into the distance measurement of $2^{-1/6}$ or less, which is only 1.1. However, the rigidity is more problematic. Local internal motion means that different parts of a protein can have very different correlation times, which can introduce reasonably large errors into the distance calibration. Internal motion generally results in short local correlation times in mobile regions, which makes the cross-relaxation rate slower: in other words, NOEs become small or even disappear. This may result in a lack of restraints, but as long as the restraints are binned in the conventional way into strong/medium/weak (Section 3.2) then this merely results in a weaker restraint than necessary.

Nonetheless, these problems are as nothing compared to the problems introduced by spin diffusion, which form the biggest difficulty in relating peak intensity to distance. Spin diffusion means that in a chain of spins, the NOE spreads along the chain, in a manner which can be thought of conveniently as the spread of heat from one container to another.[8] It weakens as it goes, but the effect is that in the presence of spin diffusion, NOE intensities no longer reflect distance accurately: distant spins appear closer than they really are, and conversely to a lesser extent close spins appear farther than they really are. This is the reason why in almost all structure calculations made using restraints derived from NOE cross-peak intensities, the distance restraints do not correspond to the single distance predicted from Equation (4). This effect can be very severe, as illustrated in Figure 1, and there is no simple way to know whether a given NOE is affected by spin diffusion or not. In principle, one can measure NOEs using a series of mixing times, but this approach is time-consuming, especially as accurate intensities need to be obtained particularly for the least sensitive short mixing times, and in practice this is seldom done. The magnitude of the spin diffusion effect can be calculated exactly, so if one knows the geometry of the spins, then one can calculate exactly how the NOE will behave. However, it is usually the converse situation that applies: we have the measured NOEs and we want to use these to

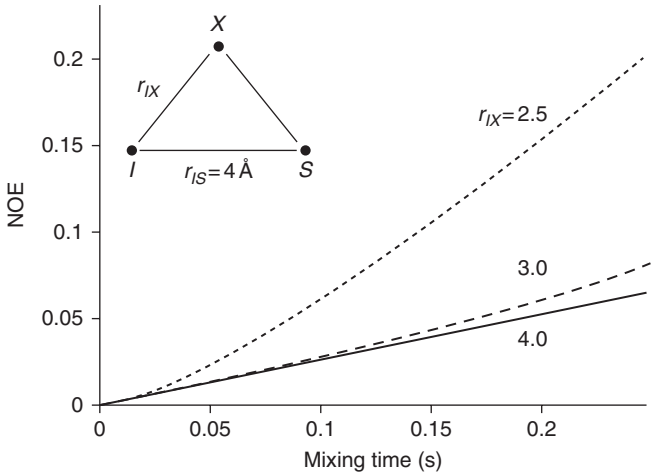


Figure 1 The effect of spin diffusion on NOE-derived distances. Protons *I* and *S* are 4.0 Å apart, and there is a third spin *X* equidistant between *S* and *I*. Spin diffusion from *I* to *S* via *X* leads to NOE buildup faster than would occur in the absence of *X*. The figure shows results for three positions of *X* with distances $r_{IX} = r_{XS}$ of 2.5 Å (dotted line), 3.0 Å (dashed line), and 4.0 Å (solid line). The time course of *IS* NOESY cross-peak buildup is shown, calculated for a protein with a correlation time of 10 ns on a 600 MHz spectrometer. For r_{IX} = 4 Å the *IS* NOE builds up smoothly and is hardly affected by the presence of *X*. However, for r_{IX} = 2.5 Å there is a 'lag time' of about 25 ms, during which *X* has little effect on the NOE between *I* and *S*, but then the *IS* NOE buildup rate increases rapidly because of spin diffusion via *X*. The NOE intensity at 100 ms was used to obtain an apparent distance r_{IS} using Equation (4): for r_{IX} = 2.5, 3.0 and 4.0 Å these apparent distances are respectively 3.5, 3.9 and 4.0 Å. Spin diffusion therefore has a marked effect on the apparent distance when a third spin comes between *I* and *S*. Calculated using iterative numerical integration.[11]

determine the molecular structure. A number of approaches have been used to overcome the problem, some of which are presented below. However, for almost all practical applications, no attempt is made to correct for spin diffusion other than by weakening the distance restraints.

There is a large scaling effect that arises from the correlation time. The correlation time scales approximately as the molecular weight, meaning that small molecules have very short correlation times, of around 10^{-11} s or less, whereas proteins have correlation times of 5 ns or longer. The effect on the cross-relaxation rate is shown in Figure 2. The main conclusion to be drawn from the figure is that as the correlation time (molecular weight) increases, cross-relaxation gets faster, and therefore NOEs build up faster, in principle permitting NOESY experiments of even very large molecules provided the NOESY mixing time is short enough.[12] This figure has an unusual appearance because it plots the absolute value of the log of σ_{IS} against the log of τ_c. The absolute value is used because at small $\omega\tau_c$, the cross-relaxation rate is positive (i.e., the first term of Equation (2) is larger than the second), while at large $\omega\tau_c$, it is negative, as we have seen already. At a value of

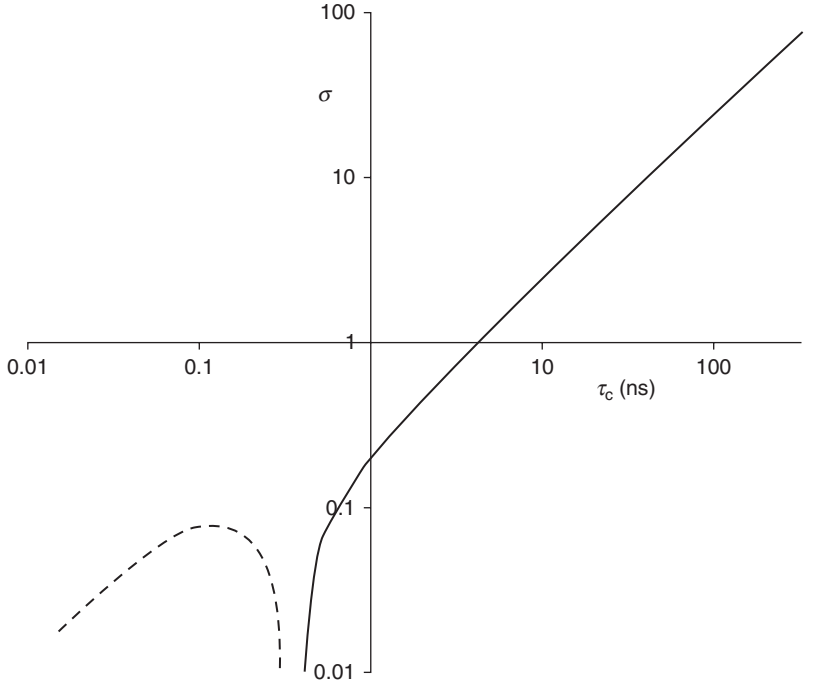


Figure 2 Variation of the cross-relaxation rate with correlation time, calculated for two protons 2.5 Å apart on a 500 MHz spectrometer. For $\omega\tau_c < 1.12$, the cross-relaxation rate σ is positive (dashed line), while for $\omega\tau_c > 1.12$, it is negative (solid line).

$\omega\tau_c$ of just greater than 1, the cross-relaxation rate goes from being positive to being negative (Figure 3). In other words, for small molecules, the NOE builds up very slowly and the cross-relaxation rate is positive, while for proteins the NOE builds up rapidly and the cross-relaxation rate is negative. For molecules of medium size (corresponding to small unstructured peptides in water), the NOE crosses over from positive to negative, and thus NOEs are very small. This effect is very important for studying both ligand binding and hydration. In both cases, we have situations where the intermolecular vector (protein to ligand or protein to water) has a very rapid correlation time, and so the cross-relaxation rate is small and positive. It should therefore be easily possible to distinguish between protein–protein NOEs (large and negative) and protein–ligand NOEs. The implications of this are discussed below.

The lack of any NOE for intermediate sized molecules, as discussed earlier, can be countered by using a technique called ROE (rotating frame NOE), in which cross-relaxation occurs not in the z (longitudinal) direction but in the transverse plane. ROEs are non-zero for all correlation times (Figure 3). ROE experiments are achieved by spin-locking magnetisation using pulse trains and therefore are prone to both COSY-type and TOCSY-type artefacts, as well as intensity variation because of off-resonance effects, which lead to cross peaks far away from the

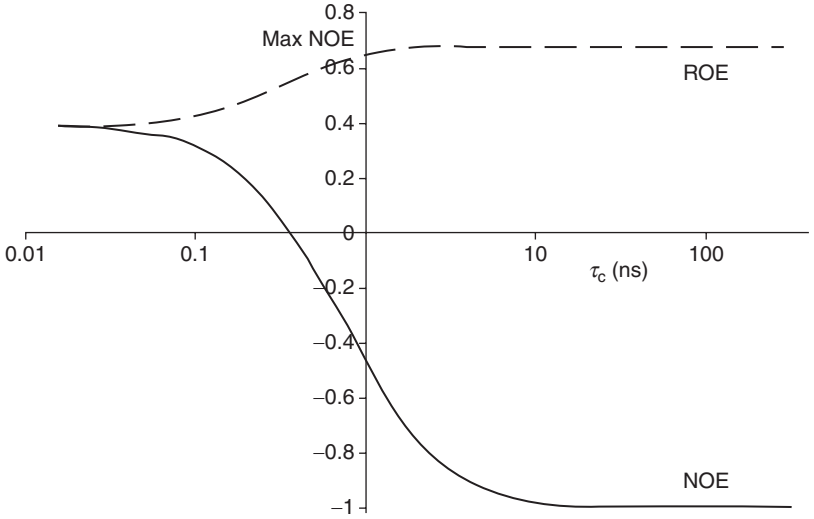


Figure 3 Maximum NOE obtainable in NOESY experiment (solid line) and ROESY experiment (dashed line) as a function of τ_c. The NOE is 0.38 for very small molecules, and goes to −1.0 for large molecules, while the ROE becomes 0.68 for large molecules.

carrier having a mixture of ROE and NOE.[8] In addition, because for proteins T_2 relaxation is faster than T_1 relaxation, signal-to-noise in ROESY spectra tends to be less than in NOESY. Therefore ROESY is not often used for proteins (whereas it is used extensively for saccharides, for example): the most common reason for using ROESY is that spin diffusion is essentially non-existent in ROESY.[8]

2.2. Macromolecular structure determination

Ever since the introduction of triple resonance spectra using ^{13}C and ^{15}N double labelled protein in the 1980s,[13] the standard procedure for calculation of protein structures (and very largely for nucleic acids also) has been:

- Use triple resonance backbone experiments (such as HNCO, HN(CA)CO, CBCA(CO)NH, HNCACB, HNCA) to assign the protein backbone
- Use HC(C)H and CCH TOCSY or similar experiments to assign sidechain signals
- Run NOESY spectra and peak pick to obtain a clean list of NOE cross peaks, from which noise and artefacts have been removed
- Assign the NOE peaks and use these as restraints for structure calculation

These two last steps are usually carried out iteratively: the initial peak list will contain a large number of ambiguous peaks (i.e., peaks where the assignment of one or both protons is not unique because of chemical shift degeneracy), which are set on one side initially, and very crude structures are calculated using only the unambiguous peaks. The crude structures are then used to filter the ambiguous peaks, because some of the possible ambiguous assignments are clearly not consistent with any possible structure, thereby allowing some of the previously

ambiguous peaks to be assigned. This then permits a better structure to be calculated and so on. In addition, it is invariably found that some of the confidently assigned NOEs used from the start violate consistently, which prompts detailed examination of these restraints, followed by possible reassignment; that is, the peak picking will also require some revision in the light of the structures and their restraint violations. The restraint set is then gradually refined, gradually increasing the number of NOEs used, after which a 'final' structure is calculated. There is no obvious end point to this process: typically it continues until the person carrying it out loses the will to live, this being a tedious process in the extreme. The structure is then validated.[14] As described in an excellent review by Nilges[14a] (quoting an earlier review by Thornton and colleagues[15]), there are four possible types of validation that could be carried out:

1. How well does the structure fit to the data used to calculate it? This could be done using an *R*-factor[16] (see Ref. 17 for a good review of this topic) but is more commonly done merely by listing restraint violations
2. How well does the structure predict restraints *not* used in the calculation? (routinely done in crystallographic studies using the free *R* factor but very rarely done in NMR, because of the difficulties inherent in predicting a spectrum from a structure, including problems in predicting chemical shifts and accurately calculating NOEs in the presence of spin diffusion and unknown internal motion; not to mention the fact that one cannot usually afford to set restraints aside merely for the purpose of cross-validation)
3. How well does the structure match to what we know about protein structures? (done mainly using programs such as PROCHECK[18] and WHATIF[19], but also using deviations of bond lengths and angles from standard distribution)
4. How well does the structure explain the biology? (done with very varying amounts of detail depending on what is known about the biology)

These steps have been refined a little in the 23 years since the first NMR structure was published[20] (where of course homonuclear spectra only were used, but this makes very little difference), but have remained essentially the same. Examples can be found widely in all the major journals in biochemistry and structural biology. We, therefore, present merely one protein example, the membrane protein OmpX, chosen to illustrate some interesting features.

Outer membrane protein X (OmpX) is, as its name suggests, a transmembrane protein: it is a 148-residue *E. coli* protein that promotes cell adhesion and invasion of host cells, and is involved in defence against the host complement system. This makes it an important target for structural determination, because of the very low number of structures of membrane proteins, but it also makes it a difficult target. Membrane proteins are usually difficult to purify, and need solubilisation using detergent micelles, which is often problematic, both because the spectral quality of reconstituted proteins can be poor and the solubility and stability can be limited, and because the solubilised protein micelle is of high molecular weight. Before the start of the NMR study, a crystal structure was published showing OmpX to have an eight-stranded β-barrel structure.

OmpX was overexpressed in *E. coli* as inclusion bodies (insoluble aggregates). It was grown either as fully ^{15}N, ^{13}C, ^{2}H labelled or else with selective ^{1}H labelling of isoleucine, valine and leucine methyls, by growth on medium that included α-ketobutyrate and α-ketovalerate.[21] The carbon source was $^{13}C,^{2}H$-glucose, employed in order to obtain as complete a level of deuteration as possible. It was resolubilised in 5 M guanidine hydrochloride, which denatures the protein, and then slowly diluted into 3% dihexanoylphosphatidylcholine (DHPC), dialysed and concentrated to about 2 mM protein.[22] This produced micelles of approximately 60 kDa, with a correlation time of approximately 25 ns. This detergent at this concentration produced the best spectrum of several tested, using TROSY HSQC spectra. TROSY experiments were used to obtain backbone and sidechain methyl assignments. NOEs were measured from 3D ^{13}C or ^{15}N-filtered NOESY spectra, with mixing times of 300 ms. Such a remarkably long mixing time can be used without excessive spin diffusion because of the extensive deuteration of the protein, though one suspects there must have been spin diffusion even so, and that use of such a long mixing time was a necessity because of weak NOEs. $^{13}C\alpha$ and $^{13}C\beta$ shifts provided information on secondary structure using TALOS,[23] while d_{NN} NOEs provided the register of the β-sheets. This provided 107 distance restraints. These restraints alone were sufficient to reproduce the β-barrel reasonably well, though the loops at both ends, and particularly some long loops at the extracellular end, were essentially undefined. One would not normally expect this small number of NOEs to produce a result even as good as this, but the fact that the protein is a continuous β-barrel, and therefore its secondary structure is completely defined by hydrogen bonds across the strands, is a big bonus. One would not expect an α-helical protein to be defined so well only by HN–HN NOEs. The global RMSD for the β-barrel was 3.1 Å (Figure 4A) but improved dramatically to 0.9 Å on addition of restraints on 100 hydrogen bonds across the β-sheets based on the limits to the sheets indicated by NOEs and carbon shifts. Again, this is rather a drastic step: one would not normally want to rely so heavily on hydrogen bonds to improve structure quality, but it was presumably a necessity in the absence of any other restraints. (Now, one would no doubt try measuring RDCs.) The resulting structure resembled closely the X-ray structure.[24,26]

The sidechain methyl-labelled sample was used to acquire separate ^{13}C and ^{15}N-resolved NOESY spectra, this time with the more respectable but still long mixing times of 200 ms, and the additional restraints improved the structure, giving NOEs between methyl groups that were up to 7 Å apart in the structure. There were a total of 526 NOEs (still a very small number compared to what would be expected for a water-soluble protein of comparable size). The NOEs were this time supplemented by only 34 hydrogen bonds, a much more respectable number, made possible by the more complete NOE set. The resultant RMSD was 0.93 Å for the best defined residues within the β-barrel, and 1.17 Å for all residues in secondary structure. The residues in loops were much less well defined (Figure 4B). By comparison to the RMSD for the HN-HN only structure, restrained using 100 hydrogen bonds (0.9 Å) it is clear that hydrogen bonds

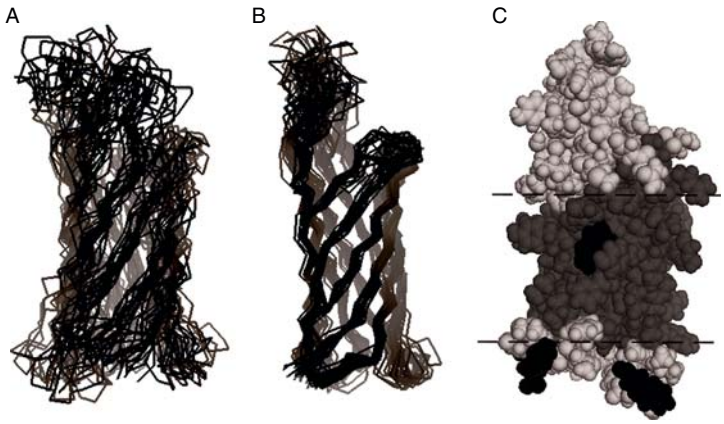


Figure 4 The structure of the *E. coli* membrane protein OmpX. (A) The structure as determined by NMR using only 107 HN–HN NOEs.[24] (B) The structure as determined using 526 NOEs to amide protons and methyl groups, plus restraints for 34 hydrogen bonds.[25] (C) A space-filling view of OmpX. Residues that have NOEs to the detergent hydrocarbon chain are shaded in dark grey. Residues with NOEs to the detergent choline headgroup are shaded black. The dashed lines indicate the expected position of the boundary between the hydrophobic and hydrophilic surfaces, corresponding to the expected membrane boundary. The lines are 28 Å apart.[22] Figures prepared using Pymol, using protein coordinates 1orm and 1q9f deposited in the Protein Data Bank.

achieve a drastic tightening of the structure, and therefore need to be used with caution.

In addition to the expected intramolecular NOEs, NOEs to the lipid detergent were also seen, both to the hydrocarbon tails and to the choline methyls. This elegantly defined the limits of the detergent-enclosed part of the protein, with a large number of NOEs from both surface methyls and HN protons to the hydrocarbon tails, and a small number of NOEs to the choline headgroup (Figure 4C). The observation of only a single set of resonances for all DHPC molecules additionally indicated rapid exchange of molecules in contact with the protein, but the negative cross-relaxation rate to DHPC indicated residence times longer than about 1 ns (see Section 5).

A very interesting category of protein structure determination has been the study of unfolded proteins: do they have any structure, and if so what is it? Long-range NOEs (in sequence, not distance) can be seen for unfolded proteins in solution, both for proteins unfolded by denaturant[27] and for natively unfolded proteins.[28] Clearly, this indicates contacts between residues far apart in the sequence, and the presence of hydrophobic clusters. The problem is that the r^{-6} dependence of the NOE means that transient contacts can give rise to significant NOEs, even when the corresponding structure is not highly populated. Simulations demonstrate that NOEs are not very good indicators of the distribution of conformations present, and remarkably that ensembles of highly non-native structures can give averaged NOEs very similar to those of a folded structure.[29,30] It is clear that structural studies of unfolded proteins need to consider a range of

experimental variables. A large fraction of expressed proteins contain sections that appear to be natively unfolded,[31] and further development of these methods is therefore important.

Structures of saccharides are in many ways more difficult than proteins.[8] The ^{1}H signals are crowded and close to the water; there are few long-range NOEs, implying that distances have to be measured extremely accurately to provide useful structural information; and correlation times are frequently variable (implying problems in applying Equation (4) directly[32]) and in the range where NOEs are close to the zero-crossing point. This tends to mean that ROE experiments are used much more widely for saccharides than they are for proteins or nucleic acids, and often using the 'T-ROESY' version that removes TOCSY cross-peak intensity.[33,34]

2.3. Dynamics from relaxation measurements

The second main application of the NOE in molecular biology is to studying internal dynamics in proteins. In most such applications, we are studying not ^{1}H–^{1}H NOEs but {^{1}H}-^{15}N NOEs in the H–N bond. Here also, the cross-relaxation rate varies markedly with the correlation time, as shown in Figure 5. In particular, for a short correlation time, the NOE to ^{15}N from ^{1}H is large and negative, while for long correlation time it is close to zero.[2] We still use an equation similar to Equation (2), but now the internuclear distance is fixed (because only the directly bonded proton contributes appreciably to the NOE), and what we are measuring is the effect of the correlation time on the cross-relaxation rate. In particular, the standard 'model-free' assumption[35] is that the NH bond is subject to two different motions: the overall tumbling of the protein, with correlation time τ_c and amplitude S^2 (where S^2 is a number ranging from 0 to 1), and a faster local motion, with correlation time τ_e and amplitude $(1 - S^2)$. The value of the NOE alone is not enough to determine both S^2 and τ_e, so the ^{15}N T_1 and T_2 are also measured. From a combination of these values, estimates of τ_c, τ_e and S^2 can be obtained.[36–38] It may be helpful here to note that the 'model free' assumption is model-free in that it does not require any particular model of the type of motion occurring. It does, however, implicitly make a number of assumptions which may not always be true (in particular independence of different motions and the separability of global and local motions), and the numbers derived are therefore more to be viewed as a simple numerical view of the degree of internal motion than an accurate portrayal of internal motion.

The NOE is a cross-relaxation phenomenon, and all NOEs are therefore sensitive to molecular motion. This creates some very interesting possibilities for experimental measurements. NOEs between any two nuclei can be measured, and are sensitive to motions at different frequencies, depending on the two nuclei studied. For example, ^{13}C-^{13}C cross-relaxation in fully labelled proteins turns out to be relatively easy to measure and to analyse, and provides information on

[2] In most publications dealing with ^{15}N NOEs, the long-correlation-time value is drawn as 1 rather than 0. The difference comes from whether you consider the change in intensity (giving a value of 0) or the actual intensity (giving a value of 1). This difference is of long standing and it is not worth trying to change the convention.

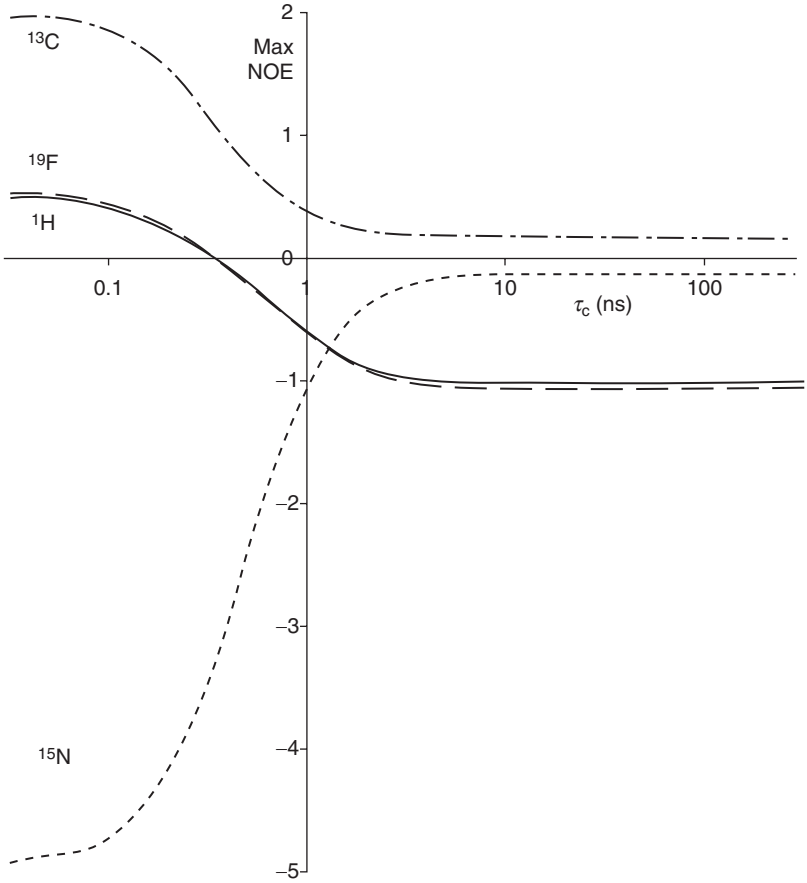


Figure 5 The dependence of the NOE on $\omega\tau_c$ for an ideal two-spin system following steady-state saturation of *S*. The zero crossing point for homonuclear NOE is at $\omega\tau_c = 1.12$, and the curves are calculated for a 500 MHz spectrometer ($\omega_H = 500 \times 2\pi \times 10^6$). The curves are shown for $X\{^1H\}$ with $X = {}^1H$ (homonuclear, solid line), ^{13}C (dashed and dotted line), ^{15}N (dotted line), and ^{19}F (dashed line). The curves for ^1H and ^{19}F are very similar because of the similarity of their gyromagnetic ratios.

sidechain mobility.[39] Even ^1H–^1H NOEs can be used to provide information on mobility, in a particularly elegant experiment.[40] As described earlier, for proteins the NOE (longitudinal) cross-relaxation rate is negative, while the ROE (transverse) cross-relaxation rate is positive. One can construct pulse sequences in which cross-relaxation occurs partly longitudinally and partly transverse, the relative proportions being determined by the spectroscopist. This is most simply done by using off-resonance ROESY, in which the effective tilt of the relaxation frame is described by an angle θ. For each internuclear vector, there must exist an angle θ_0 at which the total cross-relaxation rate (and therefore the observed NOE) is zero. This will depend only on the motion of the vector, and is most sensitive to

motions on a timescale of 100–300 ps. Different values of θ_0 are observed for different parts of the protein, thereby demonstrating differential mobility on a ps timescale. The method does not require an isotopically labelled molecule, making it suitable for analysis of motion in samples that are difficult to label, such as saccharides.

3. FASTER AND BETTER

In recent years, there has been a strong emphasis on structural genomics projects, which have received large amounts of funding. The aim of these projects is to determine protein structures rapidly, and to set up automated processes for doing so, so as to increase the throughput rate as much as possible. Both X-ray and NMR communities have risen to the challenge, and there has been a lot of activity.[17,41–47] Many authors have commented that the traditional manual methods for NOE assignment and structure refinement as described earlier tend to be subjective and difficult to validate, whereas automated methods are objective and reproducible – and of course if the automation is done properly, the structures should be at least as good as any manually derived structures and probably better. It is therefore of great importance that automated methods be developed and validated.[45]

In order to calculate a protein structure by NMR, one has first to assign the spectrum. This is normally done using a set of backbone (and Cβ/Hβ) triple resonance experiments, usually followed by sidechain assignment based on ^{13}C TOCSY, all of which experiments rely only on scalar couplings. All of the assignment methods are based on matching up frequencies in different triple resonance experiments and sorting them to obtain a self-consistent set of assignment. There have been many developments to speed up acquisition of these spectra, which do not fall within the remit of this review. Once the chemical shift assignment is complete, there is then conventionally a second and quite separate stage, in which ^{1}H–^{1}H NOESY experiments are used to obtain distance restraints. NOESY experiments are acquired, following which there is a process of peak picking: deciding which local maximum is a real peak and which is noise or artefact (caused, for example, by being close to the diagonal or to the residual water peak). After this, the NOE peaks must be assigned. The assignment of the NOEs is based again on matching chemical shifts: for each NOESY cross peak, its chemical shifts are measured and matched up to an assignment table, with some tolerance being imposed such that the chemical shift of the NOE peak must match the tabulated value within a tolerance. Clearly, too tight a tolerance may result in the correct assignment being missed because of small changes in the chemical shift due to different sample conditions, deuterium isotope effects, sample degradation, etc., while too loose a tolerance will result in very large numbers of possible assignments. Finally, the assigned NOEs are used as distance restraints to calculate a structure. By far the slowest, most tedious and most error-prone process in this list is the NOE assignment. There are opportunities for speeding up and improving all of these steps, but particularly NOE assignment.

Ever since the first structure calculations, it has been understood that this conceptually simple linear process is not in practice viable or sensible. A large majority of NOE cross peaks cannot be assigned unambiguously, and the small number of unambiguously assigned peaks are usually not enough to define the global protein fold, let alone to calculate the structure. One must therefore somehow use the information from the ambiguous peaks, most obviously by some kind of iteration between structure calculation and NOE assignment. Different groups have come up with different solutions to this problem, as outlined below. In some approaches, it is not even necessary to carry out chemical shift assignments before starting on analysis of NOEs. We look at these aspects in turn.

3.1. Faster acquisition

Most NOESY spectra of biomolecules have been and continue to be 3D spectra. Compared to 2D spectra, 3D spectra have better resolution and are thus easier to assign and to measure peak intensities (necessary for distance quantitation). It has been estimated that only a few percent of NOEs in medium-sized proteins can be assigned directly by chemical shift matching in 2D spectra, whereas 3D spectra allow *ca.* 20% to be assigned immediately.[48] We should note that already this creates some problems, since in order to obtain a 3D spectrum there must be two magnetisation transfer processes: one via the NOE, but the other via *J*-coupling. Thus, cross-peak intensities in a 3D spectrum are not determined entirely by the NOE but depend also on the size of *J*-couplings and on relaxation during the magnetisation transfer. They are also harder to measure than intensities in a 2D spectrum. This of course makes the relationship between peak intensity and distance even more ill-defined.

Acquisition of a 3D spectrum requires time incrementation in two indirect dimensions, and thus unavoidably takes longer, and/or implies worse resolution, than a 2D spectrum. Nevertheless, in 3D spectra, one can still obtain reasonably good resolution in all three dimensions. We wish to find NOEs between all pairs of protons in the protein, both those bonded to N and those bonded to C. A 3D NOESY spectrum is normally acquired as a NOESY-HSQC,[49] in which the coherence transfer process is

$$H_A(t_1) \xrightarrow{\text{NOE}} H_B \xrightarrow{^1J} X_B(t_2) \xrightarrow{^1J} H_B(t_3)$$

where X is either ^{13}C or ^{15}N, X_B is directly bonded to H_B, and the third transfer step is required merely for observation of the signal on ^1H. It is reasonably straightforward to set up an experiment that acquires data on both ^{13}C-H and ^{15}N-H systems simultaneously, speeding up the acquisition.[50] In a 3D spectrum, H_B is defined by both a ^1H frequency and a ^{13}C or ^{15}N frequency, and is thus reasonably unambiguous (meaning that for an average size protein there may be 4 or 5 possibilities within the chemical shift tolerance), while H_A only has a ^1H frequency, and is thus much less well defined. One may then try to find the 'matching' H_B-H_A-X_A peak elsewhere in the spectrum, but the process is neither simple nor unambiguous.

An obvious solution to the difficulty is to use a 4D spectrum, in which one would have the corresponding X frequency for both protons, and therefore a far less ambiguous restraint.[51] The problem here is that to acquire a 4D spectrum requires very long acquisition, and still results in rather low resolution in several dimensions, as well as low sensitivity because of magnetisation losses due to relaxation during the transfer steps. A number of ingenious solutions have been proposed in recent years.

One is to use reduced dimensionality, in particular the so-called G-matrix Fourier transform (GFT) method.[48,52] The idea here is to acquire several indirect dimensions jointly, as sums or differences of frequencies, and then work back to the original frequencies. Thus to acquire a standard 3D experiment, it is necessary to acquire, for example, 128 points in one indirect dimension ($\omega 1$) and 64 in the other ($\omega 2$), giving a total of 128 $\times$ 64 1D FIDS to be acquired. By contrast, reduced dimensionality requires the acquisition of only 128 $\times$ 7 FIDs, a time saving of 9. This is done by acquiring seven 2D spectra with indirect frequencies of 0, $\omega 1$, $\omega 2$, $\omega 1 + \omega 2$, $\omega 1 - \omega 2$, $-\omega 1 + \omega 2$, and $-\omega 1 - \omega 2$. These seven spectra then need to be combined together by adding and subtracting different combinations in order to obtain the contributing frequencies, $\omega 1$ and $\omega 2$. The resultant data have better definition of contributing chemical shifts than one would get from the 3D, because each frequency has been measured multiple times. The benefits are of course not completely unmixed: there is an added complexity in the processing, and the total acquisition time is shorter, implying that the resultant sensitivity is lower. However, the point has been made[48] that in a modern high-field spectrometer equipped with a cryoprobe, sensitivity is often excellent, and therefore one does not need long acquisition times for sensitivity reasons. The time savings are even greater for higher dimensionality (e.g., compressing five dimensions into two or three). However, for NOESY, probably the most useful application is for the compression of a 4D NOESY into three dimensions,[48] allowing a 4D spectrum to be acquired in as little as 30 h. The complexity of the processing in GFT (and its patenting) has so far deterred spectroscopists from applying GFT extensively, but it is starting to be taken up more widely.

There are a number of related concepts, of which the most popular are Hadamard techniques and spectral reconstruction from tilted planes. In Hadamard spectroscopy, one effectively only acquires indirect frequencies where data are expected: thus rather than acquiring 128 indirect frequencies, one only measures a lesser number of frequencies, namely those where (based, for example, on a quick 2D spectrum) one expects signals.[53,54] This is an efficient method for small molecules where there are a small number of indirect frequencies, but is less applicable to large molecules. Spectral reconstruction is a reworking of the projection reconstruction method used in X-ray tomography.[55] One acquires 2D projections of 3D spectra, and given enough suitably chosen projections at different angles one can back-calculate the peak positions in the original 3D spectrum. As with GFT, one needs far fewer projections than one would to acquire the standard 3D spectrum. Wüthrich and colleagues have developed the method into a technique named APSY, which has permitted a 6D spectrum to be acquired as a series of projections in two dimensions.[27,56] So far, it would be fair to say that the idea is in the category of 'interesting idea, let's wait and see if it works', rather than being an immediately applicable method.

A drawback of Fourier transforms as a way of extracting frequencies from FIDs is that one requires linear sampling of the data – for example, 128 regular increments of t_1, all with the same number of scans. A more efficient way of acquiring data is to acquire more points or more data at early increments, where the signal-to-noise is better, for example to apply an 'exponential weighting' to sampling. This however means that one cannot use an FT to process the data. But there are several alternatives available, particularly three-way decomposition[57,58] and maximum entropy,[59] both of which are promising ways of significantly speeding up acquisition. Four-dimensional decomposition has been applied to acquisition of 4D NOESY spectra, resulting in a 6-fold reduction in acquisition time.[60] Related methods involving sparse sampling have been proposed recently.[61–63]

3.2. NOE peak assignment and structure calculation

Particularly in a structural genomics era, a key question is, what is the most efficient way of carrying out the iterative process of protein structure calculation and NOE peak assignments? Most calculations use a highly ambiguous set of NOEs to calculate an imprecise and probably inaccurate initial structure, and then use the initial structure to weed out some possible assignments as being so far apart in the structure as to be impossible. This produces a less ambiguous set of NOEs, which makes a better second round structure, and so on. There is general agreement that an inaccurate initial structure biases most assignment procedures, and leads to an inaccurate final structure, implying that it is important to get the initial structure right. There is universal agreement that no initial set of NOE assignments (particularly one picked automatically) will be error-free: there will always be some assignments that turn out to be wrong (i.e., not present in the correct structure), because noise has been picked as a real peak, or there are missing chemical shift assignments, or because of chemical shift variation between samples etc. Is it worth spending time to unambiguously assign as many NOE peaks as possible at the start, so as to generate better initial structures? What is the best way of using ambiguous and possibly incorrect NOEs? There are of course several answers to this question.

A popular program for structure calculation is ARIA (Ambiguous Restraints for Iterative Assignment),[64] now in version 2.2. Like all the other popular software, it uses an unedited initial list of (mostly ambiguous) NOEs and gradually refines the list during the structure calculation. At the start of the calculation, it filters the list by rejecting any NOE peak that does not correspond to an assigned proton (or heteroatom) frequency within a certain user-defined chemical shift tolerance. It then uses the NOEs to calculate an initial structure. An important feature of ARIA is that ambiguous NOEs are applied using ambiguous distance restraints:[65] if there are two possible assignments for an NOE cross peak, then the restraint is given by

$$d_{\text{ambig}} = (\Sigma d^{-6})^{-1/6} \tag{5}$$

In other words, an ambiguous cross peak is treated as having contributions from all its possible assignments. If the ambiguous restraint is satisfied by one of the assignments within the list of possibilities, then it is satisfied, and the other possibilities have no effect. This idea has been used very widely and seems to work very well. Nilges and colleagues (the authors of ARIA) have also addressed the question posed earlier: is it worth working hard to assign as many NOEs as possible at the start of the alculation? They demonstrate[66,67] that the key requirement is that the ambiguous NOE list for each cross peak should contain the correct assignment: this is far more important than limiting the number of NOEs that are ambiguously defined, as long as the iterative assignment converges slowly (in a simulated annealing method, this means that it cools slowly). In other words, it is better to use chemical shift tolerances for cross-peak assignment that are large enough not to exclude the correct assignment, and then compensate for the highly ambiguous restraints by converging slowly. They present methods to estimate how large these parameters should be.[67] A similar conclusion was reached by other authors.[68] ARIA has a number of other interesting features, including a final refinement step in a shell of water molecules, which improves the geometry of the structure without any major detrimental effect on NOE violations.[14]

The standard procedure for calibration of distances from NOE intensities has been to use a range of reference distances within the structure, preferably ones of similar distance to those to be measured, and then to generate conservative (i.e., wide) upper and lower bound ranges based on the reference intensities,[8] the reason being that spin diffusion and internal dynamics mean that it is difficult if not impossible to accurately relate intensity to distance: spin diffusion in particular means that some long distances will have reasonably intense NOEs, whereas a variety of relaxation effects mean that some short distances will have small or zero NOEs. Typically, the lower bound is set equal to the closest possible approach (van der Waals contact or approximately 2 Å), while the upper bound is set to 3, 4, 5 or 6 Å depending on whether the NOE is classified as strong, medium, weak or very weak. There is no good reason for this strong/medium/weak classification – it is merely a simple scheme, based on the approach used in the first protein structure determination by NMR.[20] Automated methods permit a more sensible method of distance calibration. In ARIA for example, the overall calibration is based on an analysis of all NOE intensities observed together with statistical analysis of average distances expected in proteins,[5] while the upper and lower bounds are typically set to target (d) $\pm$ 0.125 d^2.[41] In fact, the errors in NOE-derived distances are not distributed normally (as implied by this equation) and it is more correct (and produces better structures) to use a log-normal distribution,[64,69] or at the very least to use a calibration error based on real distributions in proteins.[70] More recently, the possibility has been introduced into ARIA to correct distance restraints for spin diffusion, based on intermediate structures.[71] This approach has been used for a long time in connection with nucleic acid structures, where the low density of protons and their uneven distribution throughout the structure means that spin diffusion is a particular problem.[72]

The other major software for structure calculation is a suite of programs based on DYANA[73], incorporated both into an overarching environment called

KUJIRA[43] and into what is currently the only truly fully automated structure calculation program, FLYA.[74] This tackles the problem in a slightly different way. In particular, it uses two methods called network-anchoring and constraint-combination (now also available in ARIA). Network-anchoring makes use of the observation that a close contact between two residues is usually manifested as a set of NOEs between different protons on the two residues not just one: it therefore weighs NOEs so as to put more weight on NOEs that are supported by a network of others. Constraint-combination is an interesting and successful way of mitigating the effects of incorrect NOE assignments (arising from picked noise peaks, chemical shift variability, etc.): NOEs are selected randomly in pairs (or more) and the restraint is applied as a single ambiguous restraint. In this way, the probability that an incorrect restraint will bias the structure is greatly reduced. There follows an iterative procedure (as done also by ARIA), in which the restraints are used to calculate an initial structure, and the structure is used as a filter to reject NOE assignments that exceed the restraint distance in some defined fraction of the structure calculations by more than a defined cutoff. These slightly less ambiguous restraints are used as input for the next round, and so on. In this way, the degree of ambiguity of the NOE restraints is gradually reduced. In ARIA, the filter is made stronger on each iteration, on the grounds that the structures get better on each round.[5] The consequences of these filtering procedures in both ARIA and CANDID is that a significant proportion of NOE cross peaks are never used in the final round, because the possible assignments are all inconsistent with the structure. Güntert (the author of CANDID and CYANA) has suggested that the criterion for a good structure produced by CANDID is that fewer than 25% of long-range NOEs should have been rejected by the final round.[42] Programs differ in how easy it is for the user to find out which initial NOE cross peaks were rejected.

There are several other programs in use. NOAH[75,76] attacks the problem of ambiguous NOEs by ignoring any ambiguous NOEs with more than two possible assignments, randomly selecting individual assignments from the list, using them as restraints, and randomly exchanging between assignments throughout the calculation. This procedure works, essentially because correct NOE restraints work together to converge on a single structure, whereas incorrect restraints pull in different directions and therefore do not converge.[42] AUTOSTRUCTURE[77,78] uses an approach similar to network-anchoring in that it searches for patterns of NOEs typical of different secondary structure elements, which are more reliable than individual assignments – an idea that has been around for a long time,[79] and demonstrates that good ideas never go away, they just reappear in a different form! The simulated annealing program XPLOR-NIH has been modified to include this concept, into the program PASD, where it has been shown to produce structures that are not influenced by errors in the global fold of structures generated in earlier iterations (unlike ARIA and CANDID for example).[80] KNOWNOE[16] uses a Bayesian approach to eliminate NOE restraints from a list of ambiguous restraints, based on a distance cutoff. This makes the first-pass structure rather accurate, compared to other approaches, but it suffers from the problem that chemical shifts in the picked list need to be very accurate. SANE[81] is very similar to ARIA, except that it uses the AMBER forcefield, and deliberately requires the user to reject violated NOEs.

One of the problems confronted by all automated assignment programs is that chemical shift assignment of the spectrum is never complete. That is, there are always some nuclei for which there is no assignment, and therefore any procedure that rejects NOEs by matching to frequencies of known assignments will automatically reject the NOE, even though it may be a perfectly valid NOE, and may well be of value to chemical shift assignment. Indeed, CANDID requires ~90% complete assignment to work reliably,[42] and AUTOSTRUCTURE requires about 85%.[17] A recent paper shows that unassigned resonances can be represented as unattached atoms, which can subsequently be assigned and thereby attached to the protein, generating better structures and allowing structure calculations using only backbone and Hβ assignments.[82]

Another thorny question that has long been debated is the value of including 'non-NOEs' as restraints: if an NOE is *not* observed, is it valid to include it with a lower bound of approximately 4 Å, to stop the distance becoming too short? This is a risky thing to do, because NOEs can be missing for all sorts of reasons, in particular line broadening. Nevertheless, if NOEs are observed to other protons, then there cannot be line broadening and it should be justified to use non-NOEs as restraints. So far the issue has been raised several times,[5,83,84] but has not been seriously examined in automation procedures.

Oligomeric proteins generate particular problems with automated structure calculation, because in general one does not know whether any NOEs seen are intramolecular or intermolecular. The most general solution to the problem appears to be to use ambiguous NOE restraints in conjunction with symmetry restraints,[5] and/or (for dimers only) to use differential labelling of the monomers.[85] There have however been some recent developments, including a method to identify which NOEs are important for distinguishing between different intersubunit assignments, and hence which ones need to be treated carefully.[86,87]

3.3. Using NOEs without prior peak assignment

The backbone assignment step is now rapid and almost or completely automatic. However, sidechain assignment remains at best semiautomatic, and can involve considerable human intervention.[17] Therefore, for rapid and maximally automated structure calculation, is it necessary to carry out peak assignment (particularly of sidechains) before starting on the structure calculation? Do NOE spectra (supported by other knowledge such as assignments of homologous proteins) contain enough information to both assign and calculate structures? There have been several attempts in this direction, with some hope that they may succeed.[45]

Grishaev and Llinas[88–91] have done a good deal of work in this area. The idea of their programs, named BACUS, ABACUS and CLOUDS, is that it should be possible to use the NOE spectra in the absence of a complete assignment to generate distance maps ('clouds') between protons. Of course, it helps to have more information, so backbone triple resonance spectra are very helpful.[90] The assignment of peaks and structure calculation is done iteratively, crucially using Bayesian statistics to make the most probable assignments.[89] The method works on spectra containing little or no peak overlap, and perhaps in consequence has

been quoted as working well in conjunction with reduced dimensionality (GFT) data.[46]

A very interesting approach is to use homology information together with other easily available information, typically backbone assignments and RDCs to provide an initial and rather crude structure calculation, and use this to help assign otherwise unassigned NOE spectra. Meiler and Baker[92] used their structure prediction program ROSETTA to calculate a large ensemble of possible structures. The structures were ranked for their fit to experimental NOE cross peaks and RDCs. The best structure(s) were then used as the initial structure, from which NOEs were assigned iteratively. The method was tested on nine small proteins, and gave the correct fold in all cases, although success was heavily dependent on the quality of the initial assignments. In four of the nine, it produced good structures. A related application uses ROSETTA in combination with chemical shift values, with remarkable success.[93] The combination of shifts with limited NOE information is particularly promising. Clearly, as modelling and *ab initio* methods for predicting structures improve, the generation of initial structures from very limited NOE sets (and other experimental information) will also improve, possibly leading to major improvements in automated structure calculations.[94–97]

A Monte Carlo program has been developed for protein assignment that uses NOEs together with triple resonance *J*-coupling experiments.[98] Importantly, it also uses assignments from structurally related proteins (e.g., assignments of the free protein to study complexes). Another program has been described that uses only ^{15}N-labelled protein with RDCs and HN–HN NOEs.[99,100] This is likely to become an increasingly common situation. The programs were only used for assignment, but clearly one could integrate this with the other methods described in this section for automatic structure calculation in the absence of sidechain assignments. In principle, NOEs alone could provide all the information needed for assignment,[101] though one suspects that this would only work on very small proteins.

A related topic is structure calculation of very large proteins. For large proteins, it is necessary to perdeuterate the protein almost entirely, to reduce relaxation and permit the successful application of triple resonance experiments for backbone assignment.[102] The only remaining protons are HN protons, necessary for sensitive observation, but HN alone do not generally provide enough density of protons for structure calculation (although crude structures can be generated this way,[103] and it has been demonstrated that HN NOEs plus a good set of RDCs can produce good structures, at least for small proteins[104,105]). It was therefore suggested to protonate methyl groups from isoleucine, leucine and valine, by growing the protein on ^{13}C-labelled α-isoketobutyrate and α-ketovalerate.[106] This strategy has proved very successful. Interestingly, proteins containing only HN and methyl protons have few relaxation routes and therefore rather little spin diffusion. This permits NOEs to be transmitted over very large distances, of up to 12 Å, but obviously requiring a longer mixing time than normal.[107] It was shown that the NOEs from methyl groups and HN alone are sufficient to calculate rather crude structures, in a largely automated manner, but only with the aid of RDCs.[44]

In smaller proteins, methyl–methyl NOEs can be observed with good sensitivity without perdeuteration and used to generate global folds rapidly.[108]

This review largely ignores nucleic acids. This is not because nucleic acids are any less interesting or important than proteins, but merely because there has been little development on the application of NOEs to nucleic acids: by and large assignments and structure calculations carry on being done in roughly the same, low-throughput manner as they have always been. RNA creates particular problems because the structure is not predictably linear in the way that almost all DNA structures are, and therefore standard methods for assignment based on the expected sequential NOEs between the sugar protons of one nucleotide and the H8/H6 proton of the next are not always reliable. There have been some developments in this area.[109,110]

4. COMPLEXES

Although the main application of the NOE is to calculation of the structures of biomolecules, principally proteins, the NOE has also been used extensively to study complexes: mainly those of proteins with other proteins or with low molecular weight ligands, but also complexes of nucleic acids. Tightly bound complexes behave in exactly the same way as single proteins, but weaker binding presents additional opportunities and problems. There is a recent review, dealing mainly with small molecules and including a discussion of diffusion-based methods.[111]

4.1. The transferred NOE

Proteins are large molecules and therefore have short correlation times, rapid NOE buildup and extensive spin diffusion. By contrast, ligands are usually small molecules and therefore have slow NOE buildup and no spin diffusion. This implies that if a NOESY experiment is carried out on a system of protein plus ligand, where the ligand is in excess and (as is commonly the case) the off-rate is fast (implying binding in the μM range or weaker),[8] then NOEs within the free ligand develop very slowly, whereas NOEs within the bound ligand develop much more rapidly. Exchange of the ligand between bound and free will then produce free ligand (with sharp signals and chemical shifts at the positions of free ligand) displaying NOEs characteristic of the bound state. This is a very useful experiment, since it provides conformational information on the bound ligand but the information is measured from the easily observed and assigned free ligand signals, and is known as the transferred NOE (trNOE)[4,8] or exchange-transferred NOE.[112] It has been widely used. A particular beauty of the trNOE is that because the cross-relaxation rate increases with correlation time, it actually works better for larger proteins. Larger proteins have the additional benefit that their signals tend to be very broad due to rapid T_2 relaxation, and therefore interfere with the ligand signals less than they do in small proteins. The trNOE has been applied to

ligand binding to systems as large as ribosomes,[113] tRNA[114] or the acetyl choline receptor solubilised in a micelle.[115] More commonly, it is applied to protein/ligand complexes. In favourable cases, the ligand conformation can be defined reasonably accurately, which may allow the structure of the complex to be modelled using simple docking methods.[116] This is presumably because structural complementarity and energy calculations can often define the structure of a protein/ligand complex rather well given even limited experimental data, an observation made in the context of STD.[117]

Initial applications of the trNOE were (in retrospect) overinterpreted, claiming accuracies for the bound conformation that are not realistic. The pendulum then swung the other way, leading many people to avoid the trNOE. However, in recent years interest in the trNOE has reawoken, not least through the availability of programs for simulating the trNOE,[4,118] and the trNOE has once again emerged as an extremely useful tool for determining the structure of bound ligands.[3] There are several precautions that need to be taken to ensure that the data are realistic. Spin diffusion within the bound ligand is just as much of a problem as is spin diffusion within a protein, and similar methods need to be taken, such as using short mixing times in the NOE experiment. The 'dilution' of bound signals by the excess of free ligand does, however, permit a longer mixing time than one would use for the protein alone.[3] Spin diffusion from bound ligand to protein and back again does affect the intensities of trNOEs, but the effect is usually not large.[119] This is the case in particular for 'proton-rich' ligands such as peptides, whereas carbohydrates can be more problematic. The trROESY experiment has been proposed for this purpose (e.g., see Ref. 120), but is not generally useful because the difference in cross-relaxation rate between ligand and protein is small, and therefore the excess of free ligand has to be very small.[3] A particularly insidious problem for the trNOE is secondary weak binding to alternative sites, and it is necessary to check for this, most obviously by the addition of strong competitive ligands that should displace the ligand and remove any observed trNOE.[112]

In several cases, observations have been made on proteins that bind two ligands simultaneously, in which case it is often possible to observe trNOEs between one ligand and the other, thereby of course not only showing that both ligands can bind but giving structural information on their relative orientation.[121,122] This is a powerful technique and should be of significant benefit to 'SAR by NMR' strategies, in which a tight-binding ligand is conceptually assembled by joining together two weak-binding ligands that bind at adjacent sites.[123,124] More remarkably, it is even possible to observe NOEs between ligands when both ligands bind at the same site, but competitively (i.e., they cannot both bind together). In this case, magnetisation is transferred from one ligand to the other *via* the protein.[94]

As described earlier, trNOE works because cross-relaxation in the bound state is much faster than cross-relaxation in the free state. The same is also true of cross-correlated relaxation, which can provide information on dihedral angles in the bound state. Joint use of both methods can in favourable cases give better bound structures, but the range of exchange rates that are suitable is rather limited.[125]

4.2. STD and cross saturation[126]

Saturation of a protein resonance followed by spin diffusion means that after a few seconds of saturation, the protein resonances are almost completely saturated. It does not matter much at what chemical shift value the protein resonances are irradiated, since spin diffusion will distribute saturation efficiently, at least for larger proteins. This saturation can then be transferred to a bound ligand. In a similar way as the trNOE, exchange of the ligand to the free state will then result in a ligand signal that is attenuated because of the spin diffusion. Therefore, if a difference spectrum is calculated (off-resonance protein saturation minus on-resonance), the result is that signals can be seen for any ligand that is bound to the protein and that has an off-rate sufficiently fast for the free signal to be observed (i.e., an off-rate faster than T_1 relaxation of the free ligand). This effect is known as saturation transfer difference or STD,[127] and has proved very effective as a method for screening libraries of small molecules for binding to a protein target[128] and for detecting weak binding.[129] A very similar method has been used for assigning bound ligand signals.[130]

Somewhat more controversial is the use of STD for epitope mapping, that is, to define which part of the ligand is in contact with the protein. The concept is very simple: the parts of the ligand that are closest to the protein will have the largest NOEs. Therefore by measuring the percentage of ligand signal lost on saturation of the protein (or more commonly the percentage of signal seen in the difference spectrum), it should be possible to work out where the ligand binds. The problem is that T_1 relaxation of the ligand resonance is in competition with spin diffusion: the intensity of the STD therefore depends markedly on the relaxation rate of the ligand proton.[131] Therefore if the ligand signals all have similar relaxation rates, as in many saccharides[132] and peptides,[133] epitope mapping is probably reasonably secure: otherwise it would be prudent to use great care, in particular to measure T_1 for free ligand and use simulations to help interpret the spectra.[134]

A related method has been used for studying protein–protein binding, and has been named cross-saturation.[135] Saturation of any resonance in protein *A* leads, by spin diffusion, to saturation of all signals in the protein. If *A* binds to another protein *B*, then saturation will spread to this second protein also. This could provide a way to identify a binding partner, and characterise its binding site. However, there are two problems to overcome: how to saturate resonances on *A* without also saturating *B*; and how to avoid spin diffusion on *B*, which would spread saturation all over *B*. Both these problems were neatly overcome by using ^{15}N labelled and perdeuterated *B*, and unlabelled *A*, and detecting the saturation on *B* using a difference HSQC or TROSY. This means that saturation of *A* is specific to *A* and that the proton density on *B* is low enough that spin diffusion is severely restricted, thereby allowing identification of the binding site. There is a requirement for almost 100% perdeuteration of *B*, which is not easy to achieve given that growth of proteins in D_2O but on *U*-^{13}C glucose as sole carbon source usually results in retention of a reasonably high fraction of the glucose protons as ^{1}H[136] (as little as 3% residual ^{1}H is enough to compromise the experiment[137]). In addition, spin diffusion is preferably kept small by using up to 50%

D_2O as solvent, resulting in lower sensitivity. It is probably this requirement for very high levels of deuteration that has meant that the technique has not been widely adopted. The method has also been applied to protein–nucleic acid binding.[138] Here it is in some respects simpler, because there are regions of the spectrum (in particular the high-field end) where it is possible to saturate the protein without risk of saturating nucleic acid also.

4.3. Labelling

The trNOE and STD require dissociation of the ligand from the protein, and are thus limited to off-rates of about 100 s^{-1} or faster, and therefore effectively to dissociation constants of micromolar or weaker. To study more tightly bound complexes, one has to find some way to observe the bound ligand signal selectively, of which the most obvious is by a differential isotope labelling strategy. Such a strategy is also applicable to more weakly bound complexes and to protein–protein complexes, as we have seen already with cross-saturation.

The obvious way to do this is to label the protein with ^{15}N and/or ^{13}C, and observe the NOE to the ligand using an isotope filtering scheme for detection of NOEs to ^{14}N-H and/or ^{12}C-H. Better results are obtained by selective inversion of the water resonance in the middle of the mixing time, to suppress indirect relaxation pathways.[139] This would ideally require selective excitation of protein signals, which is difficult to do because $^1J_{CH}$ varies quite widely, implying that simple isotope-selective spin echoes are not very effective. However, it is possible to achieve reasonably good inversion of protein resonances across a range of $^1J_{CH}$ using adiabatic ^{13}C pulses.[140] An elegant STD experiment has been described that uses perdeuterated but methyl-protonated protein, and works even in the absence of detailed protein assignments.[141] An application of selective protein labelling has been described that characterises the ligand conformation and requires a known structure of the free protein but does not require assignment of the protein spectrum.[142]

Protein–protein complexes are in many ways simpler than protein–ligand complexes, since both components are rather rigid, and the contact area is large: hence there are relatively few orientations in which the complex can be assembled. They can therefore be studied using rather few intermolecular NOEs (as little as one), preferably in combination with RDCs to define the relative orientation of the proteins;[143] or by cross-saturation and RDCs.[144] Alternatively, and in the spirit of some of the automated joint assignment/structure calculation methods described earlier, it is possible to model the protein–protein complex, predict intermolecular NOEs, and then iteratively assign the experimentally observed NOEs aided by the prediction.[145] For protein–protein complexes, the reverse approach (detection on ^{13}C-H only) is also possible, and of course provides one way of studying otherwise symmetrical protein dimers.[146] A labelling strategy using a uniformly ^{13}C-labelled protein in complex with a fully deuterated but methyl-protonated partner has also been described.[147] Selective labelling provides a good approach to studying RNA–protein complexes, in this case using base-specific labelling.[148,149]

4.4. Exchange regimes

Reference has been made to exchange rates several times already, because clearly these are vital to studies of molecular complexes.[3] The NOE is a consequence of cross-relaxation, and in order for an NOE to be observable on a free component, cross-relaxation has to occur faster than the off-rate. This implies that trNOE techniques such as trNOE and STD are only applicable for relatively weakly bound complexes with dissociation constants weaker than low micromolar. Very often, this is the sort of dissociation constant where binding is just starting to get biologically interesting, and annoyingly trNOE and STD disappear for strongly bound ligands. Competition with even tighter binding ligands has been used to 'recover' trNOE and STD signals.[150]

Exchange can affect the success of experiments in other ways too, most significantly because intermediate exchange (i.e., exchange where the rate of exchange is comparable to the difference in chemical shift between free and bound species) causes very significant line broadening. Again, this tends to occur at low μM dissociation constant, and can be a serious problem in characterising complexes by NMR. The most obvious way to solve the problem is to modify the ligand (or the protein) so that binding is weaker and exchange is faster.[151] Another solution is to use an excess of ligand. In this case, the broadening of the ligand signal is reduced, but on the other hand the net NOE on the ligand is smaller.[152] A solution to this is to use one-dimensional driven NOEs, in conjunction with perdeuteration/methyl protonation, necessary to minimise spin diffusion during the long saturation period.

5. INTERACTIONS WITH WATER

Bulk water, of course, has a very short correlation time because of its rapid translational diffusion, while water buried within the protein interior often has a much longer correlation time. This means that NOEs from protein to bulk water also have a short correlation time while NOEs to tightly bound waters have long correlation times, and implies that it should be possible to characterise water bound to proteins, and to measure its lifetime, within certain limits.[153] This is important, since proteins only function in water, and most functionally important interactions occur within the thin hydration layer around the protein.

Bulk water exchanges rapidly with water molecules bound to proteins, and it also exchanges rapidly with most hydroxyl groups in proteins, meaning that all waters and almost all hydroxyls resonate at the same chemical shift. On the very reasonable assumption that proteins are surrounded by a hydration layer that has different properties than bulk water, then there are three main mechanisms by which cross peaks can be observed between protein and water in NOESY or ROESY spectra (Figure 6), consisting of different combinations of exchange and/or NOE. Chemical exchange between two protons (e.g., Figure 6D) invariably gives rise to positive peaks in both NOESY and ROESY spectra. An NOE between two protein protons, as in Figure 6C, gives a positive peak in NOESY but

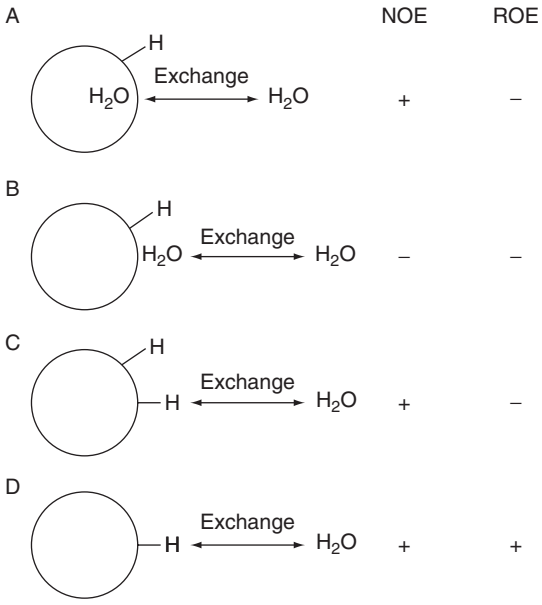


Figure 6 Origin of intermolecular water-protein cross peaks in NOESY and ROESY spectra. In each case, the sign of the cross peak in NOESY and ROESY spectra is indicated, where + indicates a peak of the same phase as the diagonal. (A) Magnetisation transfer by direct NOE between the protons of a tightly bound (long residence time, likely buried) hydration water molecule and protein. The chemical exchange between the hydration water molecules and the bulk water leads to a single averaged water signal. (B) Similar to (A) except that the bound water has a short residence time and is therefore more likely to be on the surface. (C) Exchange-relayed NOE. The NOE between a non-labile and an exchangeable proton (e.g., a hydroxyl proton) of the protein appears at the water chemical shift, because of rapid exchange between the labile proton and the water. (D) Direct chemical exchange between an exchangeable proton of the proton and the water proton. Modified from Ref. 153.

a negative peak in ROESY (i.e., negative and positive cross-relaxation rates, respectively). The interesting case is an NOE between a protein proton and a water molecule, as in Figure 6A, B. If the correlation time for the interaction is long (Figure 6A), then it behaves like a protein–protein NOE and has a positive peak in NOESY and a negative peak in ROESY, whereas if the correlation time is short (Figure 6B), then it has a negative peak in both. Therefore (as summarised in Figure 6), a direct NOE between protein and water always gives negative peaks in ROESY, but gives positive peaks in NOESY if the residence time of the water is long, and negative peaks if the residence time is short. Situation (c) (an exchange-relayed NOE) has positive peaks in NOESY and negative peaks in ROESY. And situation (d) (direct exchange) has positive peaks in both. Therefore, it is easy to identify situation (d), since this is the only one with positive ROESY peaks, and also to identify direct NOEs to rapidly exchanging water (b), since this is the only one with negative peaks in both. The difficult one to characterise is positive NOESY and negative ROESY, since this could arise in (a) or (c): unfortunately

this is also the most interesting, since if (a) it corresponds to long-residence time water. The only way to identify this safely is to be able to eliminate case (c) by looking at the protein structure and discounting any cases where there is a potential exchanging proton close by. The interpretation of long residence and short residence depends on the model used, and the common interpretation is that 'long residence' corresponds to waters with a residence time longer than about 1 ns.[153]

Using this theoretical basis, the hydration of biological molecules has been studied, and a large number of NOES to waters have been observed. A small number of water molecules, most of which are completely buried within the protein interior and are also observed in crystal structures, were shown to have residence times longer than 1 ns, while most water molecules were much more mobile.[153] Studies on protein–ligand complexes often identified water molecules bound in the interface, many of these having long residence times.[154,155] A variety of methods have been used to probe water residence times in more detail. However, three important papers were published by Halle and colleagues in 2003 and 2004,[156–158] demonstrating that the residence times for water obtained from NOESY/ROESY ratios are very dependent on the model used, and that unfortunately the standard model used up to that time was not adequate, because cross-relaxation to bulk water (by contrast to individual waters in the hydration layer) turns out to be very significant. More recent work by others supports this general conclusion.[159] The main consequence of this is that although the lifetimes derived for the small number of tightly bound waters were reasonable, lifetimes for more weakly bound waters were not, and in fact the large majority of hydration waters have dynamics retarded by a factor of less than 2 as compared to bulk water. Moreover, the relatively subtle differences in NOE/ROE ratio interpreted as differences in residence time are mainly due to different exposure of the site to bulk water. This result appears to invalidate many of the studies carried out on macromolecular hydration. There is a hydration shell of approximately 2 Å with a viscosity approximately two times greater than bulk water.[156] The conclusion is thus that water lifetime depends mainly on geometric factors, with water being more retarded the deeper and more isolated it is within the protein.

Despite the rather limited range of information available on water residence times, the large number of publications on protein–water interactions had the effect of stimulating interest in pulse sequences and experiments for studying water interactions. One outcome of this was the experiment known as WaterLOGSY,[160] which has proved extremely useful for screening ligand binding. The idea is that when a ligand binds to a protein (in a productive way, i.e., in a well-formed binding site) then it traps or retards one or more water molecules in the ligand–protein interface, such that there is at least one water molecule with a residence time longer than about 1 ns. On saturation of the water signal, there is a negative NOE to the ligand. By contrast, non-interacting ligands have positive NOEs. The difference is immediately obvious, since all the components have negative signal except for interacting molecules, which have positive signals, and the experiment is very sensitive, allowing experiments to be conducted with protein concentrations down to a few μM.

6. NOES INVOLVING NUCLEI OTHER THAN ^{1}H

6.1. NOEs involving fluorine

^{19}F has a gyromagnetic ratio close to that of ^{1}H. This means the NOE between ^{19}F and ^{1}H is large, even for molecules with long correlation time such as proteins (behaving much like the ^{1}H–^{1}H NOE, going from positive to negative as the correlation time increases), and the cross-relaxation rate between ^{19}F and ^{1}H is relatively fast, though nothing like as fast as that between ^{1}H and ^{1}H. Consequently, NOEs from ^{19}F to ^{1}H in proteins build up rapidly, but dissipate through the protein by spin diffusion even more rapidly. Therefore for detection of ^{19}F NOEs it may be necessary to perdeuterate the protein, leaving only ^{15}N–H protons. An elegant application of this method was described[161] in which MgF_3^- (a transition state analog of a transferring phosphate group) was bound into the active site of a perdeuterated and ^{15}N-labelled phosphoryl transferase, giving three fluorine signals. Saturation of each fluorine in turn gave NOEs to HN, detected by a difference HSQC experiment, thereby identifying the neighbours of each fluorine.

When fluorine is present in the solvent (e.g., as trifluoroethanol) then the concentration of fluorine is high enough that NOEs from fluorine to protein are large enough to be seen even without deuteration of the protein, for example by 1D gradient-detected NOE[162] or 2D methods.[163] As discussed in section 5, the NOE from fluorine to proton can be either positive or negative depending on the lifetime of bound trifluoroethanol, and both cases were observed.

NOEs from ^{19}F to ^{19}F can also be observed, and do not suffer from spin diffusion as badly. It is therefore possible to label proteins site-specifically with fluorines and detect close contacts using ^{19}F–^{19}F NOEs. This was done using rhodopsin in micelles, with fluorine labelling achieved by attachment of trifluoroethyl thio groups to genetically engineered cysteines, thereby presenting evidence for the arrangement of the transmembrane helices.[164]

6.2. Dynamic nuclear polarisation

The main reason for the low sensitivity of NMR is that the population difference between the energy levels is small, implying that the magnetisation corresponding to the transition between the energy levels is also small. There are a variety of ways of hyperpolarising spins, vastly increasing the population difference, and therefore making the signal much bigger. This is a very useful way of increasing sensitivity; and the population difference can then be passed on to other spins via cross-relaxation, thereby producing potentially very large signals from neighbouring spins.

One way of doing this is known as Chemically Induced Dynamic Nuclear Polarisation (CIDNP) and involves irradiation of a flavin photosensitiser with light. This produces a radical, which polarises nearby aromatic systems such as tyrosines and tryptophans. This polarisation can then be passed on to

neighbouring spins by cross-relaxation, giving large NOEs.[165] It has been used recently to enhance signals and thereby provide real-time information on protein folding.[166]

Another way of polarising nuclei is by microwave irradiation at very low temperature (of the order of 1 K). For study of proteins, the sample then has to be transferred rapidly to a protein solution and warmed up, before the polarisation can dissipate via relaxation. This does not usually provide enough time for cross-relaxation before the signal is lost, although the potential is certainly there.[167]

^{129}Xenon is a spin-1/2 nucleus, with a chemical shift that changes significantly depending on its environment. More importantly, it can be hyperpolarised by optical pumping with a laser, thereby amplifying the magnetisation by a factor of up to 10^5. In favourable cases this can be passed on to neighbouring protons by cross-relaxation,[168] an experiment known as SPINOE.[169] Therefore experiments have been carried out in which Xe is added to protein solutions. Obvious disadvantages to the method are the limited solubility of Xe in water, and the weak binding of Xe to proteins (mM in both cases), but the biggest problem turns out to be the rather long correlation time of Xe, which in practice makes the NOEs very small.[170]

ABBREVIATIONS

DHPC	dihexanoylphosphatidylcholine
NOE	nuclear Overhauser effect
RDC	residual dipolar coupling
rf	radiofrequency
STD	saturation transfer difference
trNOE	transferred NOE

ACKNOWLEDGMENT

I thank David Neuhaus for careful reading of the manuscript.

REFERENCES

1. J. C. Hus, D. Marion and M. Blackledge, *J. Mol. Biol.*, 2000, **298,** 927.
2. W. Hu and L. Wang, in: *Annual Reports on NMR Spectroscopy*, G. A. Webb, ed., Vol. **58,** Elsevier, New York, 2006, p. 232.
3. M. P. Williamson, in: *Handbook of Modern Magnetic Resonance: Applications in the Chemical and Related Sciences*, D. J. Craik, ed., Vol. **3,** Kluwer, London, 2005, p. 1339.
4. F. Ni, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1994, **26,** 517.
5. A. T. Brunger, P. D. Adams, G. M. Clore W. L. DeLano, *et al*., *Acta Crystallogr. Sect. D-Biol. Crystallogr.*, 1998, **54,** 905.
6. M. P. Williamson, in: *Encyclopedia of Analytical Science*, P. J. Worsfold, A. Townshend and C. F. Poole, eds. Elsevier, Amsterdam, 2005, p. 342.
7. M. P. Williamson, in: *Modern Magnetic Resonance*, T. Asakura, H. Saito and I. Ando, eds. Kluwer, London, 2006, p. 405.

8. D. Neuhaus and M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. 2nd ed, Wiley-VCH, New York, 2000.
9. J. H. Noggle and R. E. Schirmer, *The Nuclear Overhauser Effect: Chemical Applications*, Academic Press, New York, 1971.
10. H. P. Mo and T. C. Pochapsky, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1997, **30,** 1.
11. M. P. Williamson, *Magn. Reson. Chem.*, 1987, **25,** 356.
12. R. Horst, G. Wider, J. Fiaux, E. B. Bertelsen, A. L. Horwich and K. Wüthrich, *Proc. Natl. Acad. Sci. USA*, 2006, **103,** 15445.
13. L. E. Kay, M. Ikura, R. Tschudin and A. Bax, *J. Magn. Reson.*, 1990, **89,** 496.
14. J. P. Linge, M. A. Williams, C. Spronk, A. Bonvin and M. Nilges, *Proteins*, 2003, **50,** 496.

14a. M. Nilges and S. I. O'Donoghue, *Prog. Nucl. Magn. Spectrosc.*, 1998, **32,** 107.

15. M. W. MacArthur, R. A. Laskowski and J. M. Thornton, *Curr. Opin. Struct. Biol.*, 1994, **4,** 731.
16. W. Gronwald, K. Brunner, R. Kirchofer, J. Trenner, K. P. Neidig and H. R. Kalbitzer, *J. Biomol. NMR*, 2007, **37,** 15.
17. M. C. Baran, Y. J. Huang, H. N. B. Moseley and G. T. Montelione, *Chem. Rev.*, 2004, **104,** 3541.
18. R. A. Laskowski, J. A. C. Rullmann, M. W. MacArthur, R. Kaptein and J. M. Thornton, *J. Biomol. NMR*, 1996, **8,** 477.
19. G. Vriend and C. Sander, *J. Appl. Crystallogr.*, 1993, **26,** 47.
20. M. P. Williamson, T. F. Havel and K. Wüthrich, *J. Mol. Biol.*, 1985, **182,** 295.
21. C. Hilty, C. Fernández, G. Wider and K. Wüthrich, *J. Biomol. NMR*, 2002, **23,** 289.
22. C. Fernández, C. Hilty, G. Wider and K. Wüthrich, *Proc. Natl. Acad. Sci. USA*, 2002, **99,** 13533.
23. G. Cornilescu, F. Delaglio and A. Bax, *J. Biomol. NMR*, 1999, **13,** 289.
24. C. Fernández, C. Hilty, S. Bonjour, K. Adeishvili, K. Pervushin and K. Wüthrich, *FEBS Lett.*, 2001, **504,** 173.
25. C. Fernández, C. Hilty, G. Wider, P. Güntert and K. Wüthrich, *J. Mol. Biol.*, 2004, **336,** 1211.
26. C. Fernández and K. Wüthrich, *FEBS Lett.*, 2003, **555,** 144.
27. H. Tafer, S. Hiller, C. Hilty, C. Fernández and K. Wüthrich, *Biochemistry*, 2004, **43,** 860.
28. K. A. Crowhurst and J. D. Forman-Kay, *Biochemistry*, 2003, **42,** 8687.
29. R. Burgi, J. Pitera and W. F. van Gunsteren, *J. Biomol. NMR*, 2001, **19,** 305.
30. B. Zagrovic and W. F. van Gunsteren, *Proteins*, 2006, **63,** 210.
31. H. J. Dyson and P. E. Wright, *Chem. Rev.*, 2004, **104,** 3607.
32. M. Mackeen, A. Almond, I. Cumpstey, S. C. Enis, E. Kupče, T. D. Butters, A. J. Fairbanks, R. A. Dwek and M. R. Wormald, *Org. Biomol. Chem.*, 2006, **4,** 2241.
33. K. Lycknert, A. Helander, S. Oscarson, L. Kenne and G. Widmalm, *Carbohydr. Res.*, 2004, **339,** 1331.
34. E. A. Larsson, M. Staaf, P. Soderman, C. Hoog and G. Widmalm, *J. Phys. Chem. A*, 2004, **108,** 3932.
35. G. Lipari and A. Szabo, *J. Am. Chem. Soc.*, 1982, **104,** 4546.
36. P. Dosset, J. C. Hus, M. Blackledge and D. Marion, *J. Biomol. NMR*, 2000, **16,** 23.
37. A. G. Palmer, *Curr. Opin. Struct. Biol.*, 1997, **7,** 732.
38. P. Barthe, V. Ropars and C. Roumestand, *Compt. Rendus Chim.*, 2006, **9,** 503.
39. K. Houben and R. Boelens, *J. Biomol. NMR*, 2004, **29,** 151.
40. J. Schleucher and S. S. Wijmenga, *J. Am. Chem. Soc.*, 2002, **124,** 5881.
41. M. Nilges, M. J. Macias, S. I. O'Donoghue and H. Oschkinat, *J. Mol. Biol.*, 1997, **269,** 408.
42. P. Güntert, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2003, **43,** 105.
43. N. Kobayashi, J. Iwahara, S. Koshiba, T. Tomizawa, N. Tochio, P. Güntert, T. Kigawa and S. Yokoyama, *J. Biomol. NMR*, 2007, **39,** 31.
44. G. T. Montelione, D. Y. Zheng, Y. P. J. Huang, K. C. Gunsalus and T. Szyperski, *Nat. Struct. Biol.*, 2000, **7,** 982.
45. M. P. Williamson and C. J. Craven, *J. Biomol. NMR*, 2009, **43,** 131.
46. A. Yee, A. Gutmanas and C. H. Arrowsmith, *Curr. Opin. Struct. Biol.*, 2006, **16,** 611.
47. A. S. Altieri and R. A. Byrd, *Curr. Opin. Struct. Biol.*, 2004, **14,** 547.
48. Y. Shen, H. S. Atreya, G. H. Liu and T. Szyperski, *J. Am. Chem. Soc.*, 2005, **127,** 9085.
49. J. Cavanagh, W. J. Fairbrother, A. G. Palmer and N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice*. Academic Press, San Diego, 1996.
50. Y. L. Xia, A. Yee, C. H. Arrowsmith and X. L. Gao, *J. Biomol. NMR*, 2003, **27,** 193.
51. L. E. Kay, G. M. Clore, A. Bax and A. M. Gronenborn, *Science*, 1990, **249,** 411.

52. H. S. Atreya and T. Szyperski, *Proc. Natl. Acad. Sci. USA*, 2004, **101,** 9642.
53. E. Kupče and R. Freeman, *J. Magn. Reson.*, 2003, **162,** 300.
54. E. Kupče, T. Nishida and R. Freeman, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2003, **42,** 95.
55. E. Kupče and R. Freeman, *J. Biomol. NMR*, 2004, **28,** 391.
56. F. Fiorito, S. Hiller, G. Wider and K. Wüthrich, *J. Biomol. NMR*, 2006, **35,** 27.
57. S. B. Nolde, A. S. Arseniev, V. Y. Orekhov and M. Billeter, *Proteins*, 2002, **46,** 250.
58. A. Gutmanas, P. Jarvoll, V. Y. Orekhov and M. Billeter, *J. Biomol. NMR*, 2002, **24,** 191.
59. P. Schmieder, A. S. Stern, G. Wagner and J. C. Hoch, *J. Magn. Reson.*, 1997, **125,** 332.
60. T. Luan, V. Jaravine, A. Yee, C. H. Arrowsmith and V. Y. Orekhov, *J. Biomol. NMR*, 2005, **33,** 1.
61. B. E. Coggins and P. Zhou, *J. Magn. Reson.*, 2006, **182,** 84.
62. B. E. Coggins and P. Zhou, *J. Magn. Reson.*, 2007, **184,** 207.
63. K. Kazimierczuk, W. Kozminski and I. Zhukov, *J. Magn. Reson.*, 2006, **179,** 323.
64. J. P. Linge, M. Habeck, W. Rieping and M. Nilges, *Bioinformatics*, 2003, **19,** 315.
65. M. Nilges, *J. Mol. Biol.*, 1995, **245,** 645.
66. M. Fossi, J. Linge, D. Labudde, D. Leitner, M. Nilges and H. Oschkinat, *J. Biomol. NMR*, 2005, **31,** 21.
67. M. Fossi, H. Oschkinat, M. Nilges and L. J. Ball, *J. Magn. Reson.*, 2005, **175,** 92.
68. P. Savarin, S. Zinn-Justin and B. Gilquin, *J. Biomol. NMR*, 2001, **19,** 49.
69. M. Nilges, M. Habeck and W. Rieping, *Compt. Rendus Chim.*, 2008, **11,** 356–369.
70. M. Nilges, M. Habeck, S. I. O'Donoghue and W. Rieping, *Proteins*, 2006, **64,** 652.
71. J. P. Linge, M. Habeck, W. Rieping and M. Nilges, *J. Magn. Reson.*, 2004, **167,** 334.
72. M. Tonelli, N. B. Ulyanov, T. M. Billeci, B. Karwowski, P. Guga, W. J. Stec and T. L. James, *Biophys. J.*, 2003, **85,** 2525.
73. P. Güntert, C. Mumenthaler and K. Wüthrich, *J. Mol. Biol.*, 1997, **273,** 283.
74. P. Güntert, *Eur. Biophys. J.*, 2009, **38,** 129.
75. C. Mumenthaler and W. Braun, *J. Mol. Biol.*, 1995, **254,** 465.
76. Y. Xu, M. J. Jablonsky, P. L. Jackson, W. Braun and N. R. Krishna, *J. Magn. Reson.*, 2001, **148,** 35.
77. Y. P. J. Huang, H. N. B. Moseley, M. C. Baran, C. Arrowsmith, R. Powers, R. Tejero, T. Szyperski and G. T. Montelione, *Nucl. Magn. Reson. Biol. Macromol. Part C*, 111.
78. Y. J. Huang, R. Tejero, R. Powers and G. T. Montelione, *Proteins*, 2006, **62,** 587.
79. S. W. Englander and A. J. Wand, *Biochemistry*, 1987, **26,** 5953.
80. J. Kuszewski, C. D. Schwieters, D. S. Garrett, R. A. Byrd, N. Tjandra and G. M. Clore, *J. Am. Chem. Soc.*, 2004, **126,** 6258.
81. B. M. Duggan, G. B. Legge, H. J. Dyson and P. E. Wright, *J. Biomol. NMR*, 2001, **19,** 321.
82. E. Ab, D. J. R. Pugh, R. Kaptein, R. Boelens and A. Bonvin, *J. Am. Chem. Soc.*, 2006, **128,** 7566.
83. M. F. Summers, T. L. South, B. Kim and D. R. Hare, *Biochemistry*, 1990, **29,** 329.
84. S. Farr-Jones, G. P. Miljanich, L. Nadasdi, J. Ramachandran and V. J. Basus, *J. Mol. Biol.*, 1995, **248,** 106.
85. K. J. Walters, H. Matsuo and G. Wagner, *J. Am. Chem. Soc.*, 1997, **119,** 5958.
86. S. Potluri, A. K. Yan, J. J. Chou, B. R. Donald and C. Bailey-Kellogg, *Proteins*, 2006, **65,** 203.
87. S. Potluri, A. K. Yan, B. R. Donald and C. Bailey-Kellogg, *Protein Sci.*, 2007, **16,** 69.
88. A. Grishaev and M. Llinas, *Proc. Natl. Acad. Sci. USA*, 2002, **99,** 6707.
89. A. Grishaev and M. Llinas, *J. Biomol. NMR*, 2004, **28,** 1.
90. A. Grishaev, C. A. Steren, B. Wu, A. Pineda-Lucena, C. Arrowsmith and M. Llinas, *Proteins*, 2005, **61,** 36.
91. A. Grishaev and M. Llinas, *Nucl. Magn. Reson. Biol. Macromol. Part C*, 261.
92. J. Meiler and D. Baker, *Proc. Natl. Acad. Sci. USA*, 2003, **100,** 15404.
93. Y. Shen, O. Lange, F. Delaglio, P. Rossi, J. M. Aramini, G. H. Liu, A. Eletsky, Y. B. Wu, K. K. Singarapu, A. Lemak, A. Ignatchenko, C. H. Arrowsmith, *et al.*, *Proc. Natl. Acad. Sci. USA*, 2008, **105,** 4685.
94. V. M. Sánchez-Pedregal, M. Reese, J. Meiler, M. J. J. Blommers, C. Griesinger and T. Carlomagno, *Angew. Chem.*, 2005, **44,** 4172.
95. W. Li, Y. Zhang and J. Skolnick, *Biophys. J.*, 2004, **87,** 1241.
96. P. M. Bowers, C. E. M. Strauss and D. Baker, *J. Biomol. NMR*, 2000, **18,** 311.
97. J. Korukottu, M. Bayrhuber, P. Montaville, V. Vijayan, Y. S. Jung, S. Becker and M. Zweckstetter, *Angew. Chem.*, 2007, **46,** 1176.
98. T. K. Hitchens, J. A. Lukin, Y. P. Zhan, S. A. McCallum and G. S. Rule, *J. Biomol. NMR*, 2003, **25,** 1.

99. C. J. Langmead and B. R. Donald, *J. Biomol. NMR*, 2004, **29,** 111.
100. C. J. Langmead, A. Yan, R. Lilien, L. C. Wang and B. R. Donald, *J. Comp. Biol.*, 2004, **11,** 277.
101. H. Kamisetty, C. Bailey-Kellogg and G. Pandurangan, *Bioinformatics*, 2006, **22,** 172.
102. K. H. Gardner and L. E. Kay, *Annu. Rev. Biophys. Biomol. Struct.*, 1998, **27,** 357.
103. L. M. I. Koharudin, A. Bonvin, R. Kaptein and R. Boelens, *J. Magn. Reson.*, 2003, **163,** 228.
104. P. R. Jensen, J. B. Axelsen, M. H. Lerche and F. M. Poulsen, *J. Biomol. NMR*, 2004, **28,** 31.
105. A. W. Giesen, S. W. Homans and J. M. Brown, *J. Biomol. NMR*, 2003, **25,** 63.
106. N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis and L. E. Kay, *J. Biomol. NMR*, 1999, **13,** 369.
107. R. Sounier, L. Blanchard, Z. R. Wu and J. Boisbouvier, *J. Am. Chem. Soc.*, 2007, **129,** 472.
108. N. Van Melckebeke, J. P. Simorre and B. Brutscher, *J. Am. Chem. Soc.*, 2004, **126,** 9584.
109. R. W. Adamiak, J. Blazewicz, P. Formanowicz, Z. Gdaniec, M. Kasprzak, M. Popenda and M. Szachniuk, *J. Comp. Biol.*, 2004, **11,** 163.
110. J. Blazewicz, M. Szachniuk and A. Wojtowicz, *Bioinformatics*, 2005, **21,** 2356.
111. T. Brand, E. J. Cabrita and S. Berger, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2005, **46,** 159.
112. C. B. Post, *Curr. Opin. Struct. Biol.*, 2003, **13,** 581.
113. L. Verdier, J. Gharbi-Benarous, G. Bertho, P. Mauvais and J. P. Girault, *Biochemistry*, 2002, **41,** 4218.
114. R. Szilaghi, S. Shahzad-ul-Hussan and T. Weimar, *ChemBioChem*, 2005, **6,** 1270.
115. H. Furukawa, T. Hamada, M. K. Hayashi, T. Haga, Y. Muto, H. Hirota, S. Yokoyama, K. Nagasawa and M. Ishiguro, *Mol. Pharm.*, 2002, **62,** 778.
116. A. P. R. Zabell and C. B. Post, *Proteins*, 2002, **46,** 295.
117. U. Schieborr, M. Vogtherr, B. Elshorst, M. Betz, S. Grimme, B. Pescatore, T. Langer, K. Saxena and H. Schwalbe, *ChemBioChem*, 2005, **6,** 1891.
118. E. Z. Eisenmesser, A. P. R. Zabell and C. B. Post, *J. Biomol. NMR*, 2000, **17,** 17.
119. A. P. R. Zabell and C. B. Post, *J. Biomol. NMR*, 2002, **22,** 303.
120. M. A. Johnson and B. M. Pinto, *Carbohydr. Res.*, 2004, **339,** 907.
121. L. H. Lucas, K. E. Price and C. K. Larive, *J. Am. Chem. Soc.*, 2004, **126,** 14258.
122. D. W. Li, L. A. Levy, S. A. Gabel, M. S. Lebetkin, E. F. DeRose, M. J. Wall, E. E. Howell and R. E. London, *Biochemistry*, 2001, **40,** 4242.
123. S. B. Shuker, P. J. Hajduk, R. P. Meadows and S. W. Fesik, *Science*, 1996, **274,** 1531.
124. B. Becattini, C. Culmsee, M. Leone D. Y. Zhai, *et al.*, *Proc. Natl. Acad. Sci. USA*, 2006, **103,** 12602.
125. S. Ravindranathan, J. M. Mallet, P. Sinay and G. Bodenhausen, *J. Magn. Reson.*, 2003, **163,** 199.
126. N. R. Krishna and V. Jayalakshmi, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2006, **49,** 1.
127. M. Mayer and B. Meyer, *Angew. Chem.*, 1999, **38,** 1784.
128. J. Klein, R. Meinecke, M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 1999, **121,** 5336.
129. J. H. Streiff, N. O. Juranic, S. I. Macura, D. O. Warner, K. A. Jones and W. J. Perkins, *Mol. Pharm.*, 2004, **66,** 929.
130. H. Deng, A. Lewandowicz, S. M. Cahill, R. H. Furneaux, P. C. Tyler, M. E. Girvin, R. H. Callender and V. L. Schramm, *Biochemistry*, 2004, **43,** 1980.
131. J. L. Yan, A. D. Kline, H. P. Mo, M. J. Shapiro and E. R. Zartler, *J. Magn. Reson.*, 2003, **163,** 270.
132. M. Rinnbauer, B. Ernst, B. Wagner, J. Magnani, A. J. Benie and T. Peters, *Glycobiology*, 2003, **13,** 435.
133. S. Megy, G. Bertho, J. Gharbi-Benarous, F. Baleux, R. Benarous and J. P. Girault, *FEBS Lett.*, 2006, **580,** 5411.
134. V. Jayalakshmi and N. R. Krishna, *J. Magn. Reson.*, 2002, **155,** 106.
135. H. Takahashi, T. Nakanishi, K. Kami, Y. Arata and I. Shimada, *Nat. Struct. Biol.*, 2000, **7,** 220.
136. R. A. Venters, C. C. Huang, B. T. Farmer, R. Trolard, L. D. Spicer and C. A. Fierke, *J. Biomol. NMR*, 1995, **5,** 339.
137. J. Morrison, J. C. Yang, M. Stewart and D. Neuhaus, *J. Mol. Biol.*, 2003, **333,** 587.
138. A. N. Lane, G. Kelly, A. Ramos and T. A. Frenkiel, *J. Biomol. NMR*, 2001, **21,** 127.
139. C. Eichmuller, M. Tollinger, B. Krautler and R. Konrat, *J. Biomol. NMR*, 2001, **20,** 195.
140. C. Eichmuller, W. Schuler, R. Konrat and B. Krautler, *J. Biomol. NMR*, 2001, **21,** 107.
141. P. J. Hajduk, J. C. Mack, E. T. Olejniczak, C. Park, P. J. Dandliker and B. A. Beutel, *J. Am. Chem. Soc.*, 2004, **126,** 2390.
142. K. L. Constantine, M. E. Davis, W. J. Metzler, L. Mueller and B. L. Claus, *J. Am. Chem. Soc.*, 2006, **128,** 7252.
143. G. M. Clore, *Proc. Natl. Acad. Sci. USA*, 2000, **97,** 9021.

144. T. Matsuda, T. Ikegami, N. Nakajima, T. Yamazaki and H. Nakamura, *J. Biomol. NMR*, 2004, **29,** 325.
145. I. Bertini, M. Fragai, A. Giachetti, C. Luchinat, M. Maletta, G. Parigi and K. J. Yeo, *J. Med. Chem.*, 2005, **48,** 7544.
146. K. Zangger, M. Oberer, W. Keller and H. Sterk, *J. Magn. Reson.*, 2003, **160,** 97.
147. A. Ludwig, M. Lorenz, N. Grimbo, F. Steinle, S. Meiners, C. Bartsch, K. Stangl, G. Baumann and V. Stangl, *Biochem. Biophys. Res. Comm.*, 2004, **316,** 659.
148. R. D. Peterson, C. A. Theimer, H. H. Wu and J. Feigon, *J. Biomol. NMR*, 2004, **28,** 59.
149. H. H. Wu, L. D. Finger and J. Feigon, *Nucl. Magn. Reson. Biol. Macromol. Part C*, 525.
150. Y. S. Wang, D. J. Liu and D. F. Wyss, *Magn. Reson. Chem.*, 2004, **42,** 485.
151. S. D. Auweter, F. C. Oberstrass and F. H. T. Allain, *J. Mol. Biol.*, 2007, **367,** 174.
152. M. Reibarkh, T. J. Malia, B. T. Hopkins and G. Wagner, *J. Biomol. NMR*, 2006, **36,** 1.
153. G. Otting, *Prog. NMR Spectrosc.*, 1997, **31,** 259.
154. V. Tsui, I. Radhakrishnan, P. E. Wright and D. A. Case, *J. Mol. Biol.*, 2000, **302,** 1101.
155. S. Cai, S. Y. Stevens, A. P. Budor and E. R. P. Zuiderweg, *Biochemistry*, 2003, **42,** 11100.
156. B. Halle, *J. Chem. Phys.*, 2003, **119,** 12372.
157. B. Halle, *Philos. Trans. R. Soc. Lond. B*, 2004, **359,** 1207.
158. K. Modig, E. Liepinsh, G. Otting and B. Halle, *J. Am. Chem. Soc.*, 2004, **126,** 102.
159. D. Frezzato, F. Rastrelli and A. Bagno, *J. Phys. Chem. B*, 2006, **110,** 5676.
160. C. Dalvit, P. Pevarello, M. Tato, M. Veronesi, A. Vulpetti and M. Sundstrom, *J. Biomol. NMR*, 2000, **18,** 65.
161. N. J. Baxter, L. F. Olguin, M. Goličnik, G. Feng, A. M. Hounslow, W. Bermel, G. M. Blackburn, F. Hollfelder, J. P. Waltho and N. H. Williams, *Proc. Natl. Acad. Sci. USA*, 2006, **103,** 14732.
162. D. Martinez and J. T. Gerig, *J. Magn. Reson.*, 2001, **152,** 269.
163. K. Shikii, S. Sakurai, H. Utsumi, H. Seki and M. Tashiro, *Anal. Sci.*, 2004, **20,** 1475.
164. M. C. Loewen, J. Klein-Seetharaman, E. V. Getmanova, P. J. Reeves, H. Schwalbe and H. G. Khorana, *Proc. Natl. Acad. Sci. USA*, 2001, **98,** 4888.
165. P. J. Hore, M. R. Egmond, H. T. Edzes and R. Kaptein, *J. Magn. Reson.*, 1982, **49,** 122.
166. K. H. Mok, L. T. Kuhn, M. Goez, I. J. Day, J. C. Lin, N. H. Andersen and P. J. Hore, *Nature*, 2007, **447,** 106.
167. M. E. Merritt, C. Harrison, W. Mander, C. R. Malloy and A. D. Sherry, *J. Magn. Reson.*, 2007, **189,** 280.
168. S. M. Rubin, M. M. Spence, B. M. Goodson, D. E. Wemmer and A. Pines, *Proc. Natl. Acad. Sci. USA*, 2000, **97,** 9472.
169. G. Navon, Y. Q. Song, T. Room, S. Appelt, R. E. Taylor and A. Pines, *Science*, 1996, **271,** 1848.
170. H. Desvaux, L. Dubois, G. Huber, M. L. Quillin, P. Berthault and B. W. Matthews, *J. Am. Chem. Soc.*, 2005, **127,** 11676.

CHAPTER 4

Recent Applications of REDOR to Biological Problems

Terry Gullion

Contents

Abstract

Rotational-echo, double-resonance NMR (REDOR) was introduced approximately twenty years ago and has been widely used as a tool for structural analysis. The experiment is easy to implement and data analysis is straightforward; these properties have made the experiment useful. REDOR has been applied to a many chemical systems, including proteins and peptides, membranes, polymers, guest-host systems and inorganic glasses. This article describes some recent applications of REDOR to biological systems during the past five years. While not an exhaustive review of applications, the intent is to illustrate a variety of biological problems that have benefited from the use of REDOR.

Key Words: REDOR, Rotational-echo double-resonance, NMR, Dipolar recoupling.

Department of Chemistry, West Virginia University, Morgantown, WV 26505, USA

Annual Reports on NMR Spectroscopy, Volume 65
ISSN 0066-4103, DOI: 10.1016/S0066-4103(08)00204-4

1. INTRODUCTION

Rotational-echo, double-resonance NMR (REDOR) was introduced approximately 20 years ago as a magic-angle spinning technique for measuring heteronuclear dipolar interactions.[1,2] Since the dipolar interaction depends on the distance between the dipolar-coupled spins, inter-nuclear distances can be determined.[3] The REDOR experiment is a good structural tool for several reasons. First, the strength of the dipolar interaction is measured while maintaining high-resolution conditions. That is, the sharp liquid-like resonances provided by magic-angle spinning are preserved. Second, net evolution of the spin system is caused by only the dipolar interaction, and the generally stronger chemical shift interaction produces no net effect in REDOR. This property makes data analysis straightforward. Third, the experiment is relatively easy to implement, especially since it is (in most cases) based on trains of rotor-synchronized π-pulses.

REDOR has been applied to a wide variety of structural problems over the years. There have been many biological applications of REDOR, and the experiment has been very useful for characterizing materials such as glasses and polymers as well. In addition, modifications and improvements of the experiment continue.

The focus of this chapter is to show examples where REDOR has been useful in solving biological problems with the hope that new researchers will benefit from seeing the wide variety of biological systems and problems that have been studied. Specific examples are provided that have been reported during the past 5 years (2003–2007). In addition, a table summarizing earlier applications is provided as a reference. Most of the applications presented here are on very complicated systems that required information provided by other NMR experiments, but only the REDOR contributions to solving the structural problems are presented. The original papers should be read to get full details. I apologize if my search capabilities have missed articles, but I would appreciate them being brought to my attention.

2. APPLICATIONS WITH PEPTIDES

Peptides lend themselves to labelling with ^{2}H, ^{13}C and ^{15}N isotopes because of the availability of commercially available isotopically enriched amino acids. Many peptides and proteins cannot be characterized by solution-state NMR because they are either insoluble or their molecular weight is too large. Others cannot be characterized sufficiently by scattering techniques because of the lack suitable crystals. Solid-state NMR does not require crystals and even very high molecular weight proteins have been studied by REDOR. This section shows examples where structural details of peptides and proteins in the solid state have been provided by REDOR experiments.

The binding of citrate to *Aa*PurE, a large enzyme (151 kDa) that mutates ribonucleotides, was studied by Schaefer et al.[4] They used ^{13}C–^{15}N REDOR to

determine the protonation state of His59. [6-^{13}C]citrate was complexed with [U-^{15}N]*Aa*PurE. Their results indicate that equimolar amounts of the two forms of histidine shown in Figure 1 are present in the complex and that the ^{13}C label on the citrate is ~4 Å from the N$^{\delta}$ of His59 for both forms.

Petkova et al.[5] studied Alzheimer's β-amyloid fibrils in their ongoing efforts to determine the structure of its fibrils at the molecular level. Samples were made with certain residues uniformly ^{13}C, ^{15}N labelled. Frequency-selective ^{13}C–^{15}N REDOR experiments were used to measure specific ^{13}C–^{15}N dipolar couplings. One example of their structural work is the characterization of the intermolecular salt bridge formed between residues K28 and D23. They measured the dipolar interaction between the ^{13}C and ^{15}N labels illustrated in Figure 2 to show that the salt bridge is present.

Chan et al.[6] studied the intermolecular arrangement of glutamine residues in amyloid fibrils formed by residues 10–39 of the yeast prion protein Ure2p. Their work supported the model of ''polar zippers'' formed by hydrogen bonding between intermolecular glutamine residues. ^{13}C–^{15}N REDOR experiments were used to measure the intermolecular distance between the ^{15}N labels and ^{13}C labels at the carbonyl position. The sample contained glutamine that was uniformly ^{13}C,

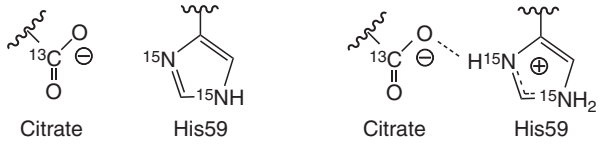


Figure 1 Two forms of histidine in the complex formed by citrate and *Aa*PurE. Citrate is close to His59.

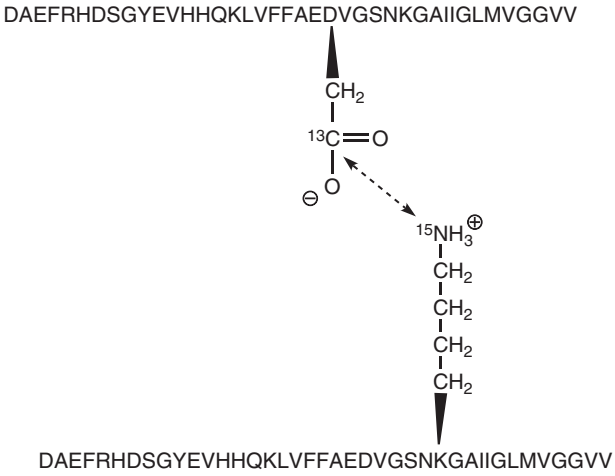


Figure 2 Aβ_{1-40} peptide. The distance (indicated by the arrow) between the ^{13}C and ^{15}N labels on residues D23 and K28 was measured with REDOR.

^{15}N labelled. By using a novel double–single quantum REDOR experiment, they were able to eliminate the strong intra-residue ^{13}C–^{15}N dipolar interactions and measure the weaker ^{13}C–^{15}N intermolecular interaction. The peptide and labelled glutamine side chains are shown in Figure 3.

The bend structure in Alzheimer's β-amyloid fibrils formed by Aβ_{1-40} peptide occurs in the region containing residues 23–29 and includes Asp23 and Lys28. Sciarretta et al.[7] prepared a sample with Asp23 and Lys28 residues uniformly labelled with ^{13}C and ^{15}N isotopes. The labelled side chains are illustrated in Figure 4. In a frequency-selective ^{13}C–^{15}N REDOR experiment, they showed that a salt bridge forms between Asp23 and Lys28 by determining the distance between the C_γ carbon of Asp23 and the N_ϵ nitrogen of Lys28.

Paci et al.[8] performed molecular dynamics simulations of amyloid aggregation of the 11-residue peptide transthyretin. The backbone dihedral angles used by

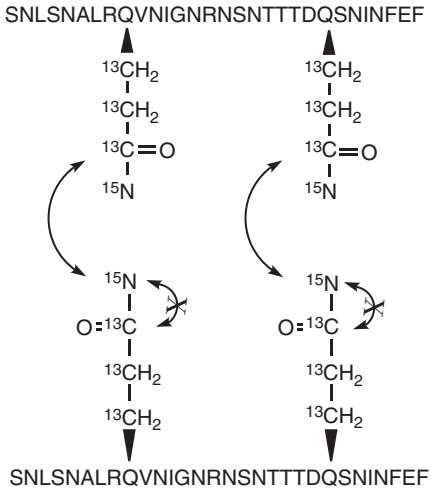


Figure 3 Residues 10–39 of the yeast prion protein Ure2p. Side chains of uniformly ^{13}C, ^{15}N glutamine residues are shown. The curved arrows indicate specific ^{13}C–^{15}N dipolar couplings measured by frequency-selective REDOR. The crossed out curved arrows indicate the intramolecular dipolar couplings not selected for measurement.

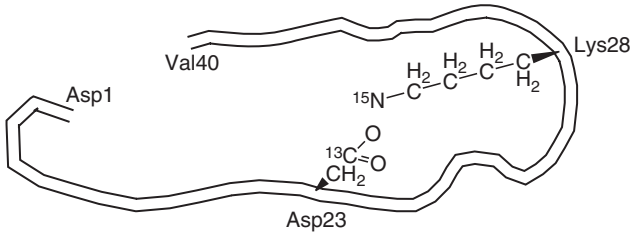


Figure 4 Illustration showing the bend in Aβ_{1-40} and proximity of Asp23 and Lys28 side chains.

Paci et al. in their simulation were determined by the REDOR experiments of Jaroniec et al.[9]

A special 17-residue peptide was designed to study amyloid formation. The peptide has the ability to go from a coiled-coil conformation at room temperature to amyloid fibrils at higher temperatures. Kammerer et al.[10] prepared the peptide with ^{13}C labels at Leu14 and ^{15}N labels at Ala7 for ^{13}C–^{15}N REDOR experiments. The Ala7 (small circles) and Leu14 residues (large circles) and peptide sequence are shown in Figure 5. ^{13}C–^{15}N REDOR measurements showed the peptide formed anti-parallel β-sheets (indicated by the arrows) with the Ala7 and Leu14 residues on adjacent molecules in close proximity.

Different peptide fragments based on various residue components of Alzheimer's β-amyloid peptide can form different types of fibrils. Amphiphilicity may determine the type of fibril formed. Gordon et al.[11] studied two β-amyloid peptides made from residues 16 to 22 (Aβ(16–22)). One of the peptides was modified by acylating the N-terminus with octanoic acid to increase its amphiphilicity. The unmodified peptide Aβ(16–22) is not amphiphilic and served as a control. Samples were made with some peptide ^{13}C labelled at the carbonyl position of either Leu17 or Phe20 and remaining peptide ^{15}N labelled at Ala21. For in-register parallel sheets, the distance between ^{13}C and ^{15}N labels was expected to be short for the ^{13}C label on Phe20. For in-register anti-parallel sheets, the distance between ^{13}C and ^{15}N labels was expected to be short for the ^{13}C label on Leu17. The approximate distances for the two types of sheets are shown in Figure 6. ^{13}C–^{15}N REDOR experiments showed that the unmodified Aβ(16–22) formed an in-register anti-parallel β-sheet and octanoyl-modified Aβ(16–22) formed an in-register parallel β-sheet.

In related experiments, Petkova et al.[12] studied Aβ(11–25) fibrils prepared at pH 7.4 and pH 2.4. Both samples produced anti-parallel β-sheets. However, the registry of the intermolecular hydrogen bonds was found to be pH dependent.

Human calcitonin (hCT) is a 32-residue fibril forming peptide. Naito et al.[13] prepared a series of ^{13}C-, ^{15}N-labelled hCT peptides (labelled residues

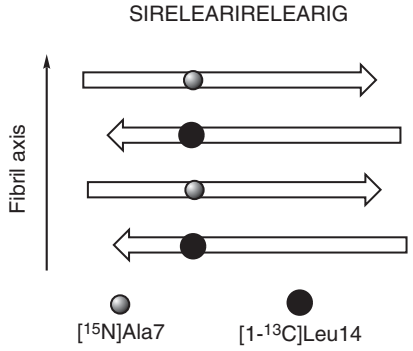


Figure 5 Schematic of anti-parallel sheet structure with Ala7 and Leu14 residues on different molecules in close proximity. The peptide sequence is shown at the top.

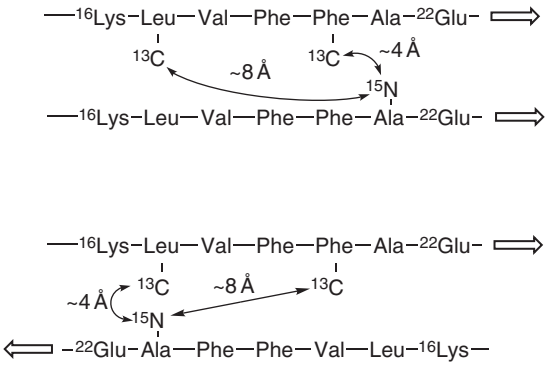


Figure 6 Illustrations of parallel and anti-parallel sheet structures showing approximate ^{13}C–^{15}N distances between nuclear spin labels.

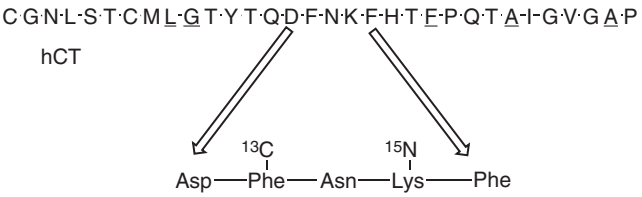


Figure 7 32-residue hCT sequence. Underlined residues are sites for ^{13}C and ^{15}N labelling. The penta-peptide shown at the bottom has the same sequence as the middle part of hCT (arrows indicate the piece duplicated).

Figure 8 Peptide sequence of statherin. The bold residues are potential sites of mutation.

underlined) along with a ^{13}C–^{15}N-labelled penta-peptide with sequence taken from the middle part of hCT. The peptide sequence and penta-peptide fragment are shown in Figure 7. From ^{13}C–^{15}N REDOR experiments, two features were proposed: first, fibrils form anti-parallel β-sheets and secondly, the phenyl rings of Phe16 are aligned on the same side of the sheet and are stabilized through π–π interactions.

The salivary protein statherin regulates the mineralization of hydroxyapatite. The structure of statherin absorbed on hydroxyapatite is helical near the N-terminal domain. The preservation of this helical structure in single and multiple point mutated statherin peptides was examined by Goobes et al.[14] using ^{13}C–^{15}N REDOR. The distance, indicated by the curved arrow in Figure 8, between the ^{13}C-labelled C1 carbon of Leu8 and the ^{15}N-labelled nitrogen of Gly12 was measured by REDOR for several samples with various residues (bold) mutated to alanine.

The experiments showed that the helical structure involving residues 1–12 is preserved.

Orexin-B is a 28-residue peptide that consists of two helical structures separated by a bend. One helical structure spans residues 7–19 and the second helical structure spans residues 23–28. Bernard et al.[15] performed ^{13}C–^{15}N REDOR experiments on samples with ^{13}C labels at the amide carbonyl carbon of Leu11 and the carbonyl carbon of Ala17 and with ^{15}N labels at the amide nitrogen of Leu[15] and the nitrogen of Ala22. The ^{13}C- and ^{15}N-labelled residues are indicated in Figure 9. REDOR measurements between the ^{13}C and ^{15}N spin pair from Leu11 and Leu[15] showed that the helical structure is maintained in the solid state. REDOR measurements between the ^{13}C and ^{15}N spin pair from Ala17 and Ala22 are consistent with the presence of the bend occurring in the solid state as well.

Amelogenin proteins play an important role in controlling the growth of tooth enamel, which is composed of long hydroxyapatite crystals. Shaw et al.[16] used ^{13}C–^{31}P REDOR with a 59-residue amelogenin, LRAP, to obtain structural relationship between LRAP and hydroxyapatite. Figure 10 illustrates the C-terminus of LRAP interacting with hydroxyapatite (block) and A46 indicated by the circle. The methyl carbon of Ala46 was ^{13}C labelled and the REDOR experiment probed the interaction between this label and the ^{31}P nuclei in the hydroxyapatite. They concluded that the COOH-terminus of LRAP is on the hydroxyapatite surface.

Elastin is an amorphous protein that provides the elasticity of skin and connective tissues. In an effort to learn more about the structure of elastin, Yao and Hong[17] measured carbon–proton and carbon–nitrogen distances in the elastic mimic $(VPGVG)_3$. Prior work on elastin by Hong et al.[18] used REDOR in HETCOR

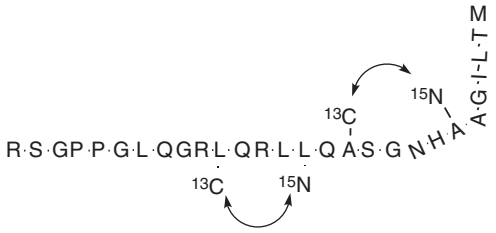


Figure 9 Sequence of orexin-B. The dipolar couplings measured between specific ^{13}C and ^{15}N spin pairs are indicated by the curved arrows.

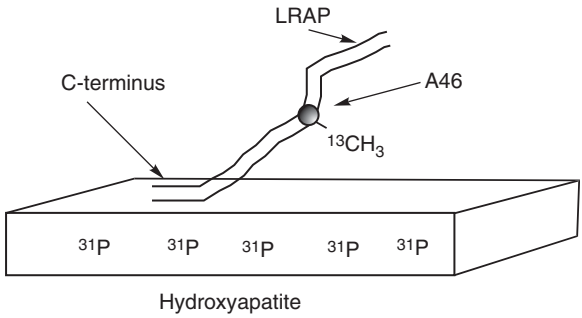


Figure 10 Schematic showing LRAP interacting with hydroxyapatite.

fashion for spectral assignments. A structural model of the peptide based on REDOR-measured distances suggested that two conformations of the peptide coexist. One structure is a turn in the region V6–V9, which is supported by distances measured between the carbonyl carbon of V6 and the nitrogen and the nitrogen proton of V9 (shown in Figure 11). The other structure is a more extended conformation of the peptide.

The structure of the silk I conformation of silk fibroin produced by *Bombyx mori* silkworms has remained elusive even though the silk II conformation is well understood. Gullion et al.[19] studied ^{13}C-, ^{2}H-labelled 30-residue $(AlaGly)_{15}$ synthetic silk I mimics to gain insight into the silk I conformation. By appropriate placing of ^{13}C and ^{2}H labels, the dihedral angles of Ala17 determined by ^{13}C–^{2}H REDOR are consistent with a β-turn. The REDOR measurements combined with chemical shift observations suggest that silk I is a repeated β-turn structure. Figure 12(left) shows the ^{13}C label on Gly16 and the ^{2}H label on Ala17, and the distance between these two labels is determined by Φ. The structure on the right has a ^{2}H label on Ala17 and a ^{13}C label on Gly18, and the distance between these two labels is determined by ψ of Ala17.

^{13}C–^{15}N REDOR experiments were done on the same peptide by Kameda et al.[20] to confirm hydrogen bonding between Gly14 and Ala17.

The conformation of the cyclic peptide cyclo(–D–Pro–Ala_4–) was studied with ^{13}C–^{15}N REDOR by Heller et al.[21] Figure 13 shows two selectively labelled cyclic peptides used in the work. The carbon–nitrogen distances determined by the REDOR experiments are 4.03 Å for the sample illustrated on the left and 3.39 Å for the sample shown on the right. These distance constraints help determine the conformation of the molecule.

Biomolecules highly labelled with ^{13}C and ^{15}N have been easier to study in many cases using frequency-selected REDOR experiments. However, Sharpe et al.[22] performed such REDOR experiments on uniformly ^{13}C-, ^{15}N-labelled valine (Figure 14) and found that the dipolar dephasing did not necessarily

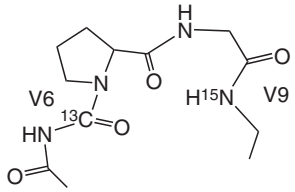


Figure 11 Sketch of the turn in the vicinity of residues 6–9 for elastic. The measured distance between the ^{13}C label on V6 and the ^{15}N label on V9 supports the turn motif.

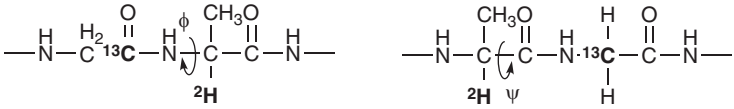


Figure 12 Placement of ^{13}C and ^{2}H labels to determine the dihedral angles of Ala17.

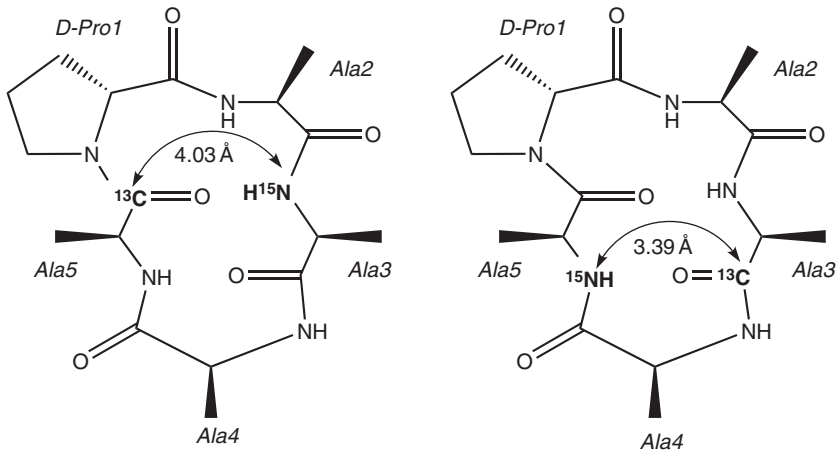


Figure 13 Placement of ^{13}C and ^{15}N labels in the cyclic peptide. REDOR-determined distances are indicated.

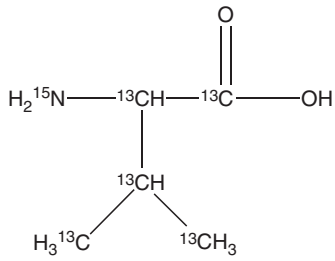


Figure 14 Uniformly ^{13}C-, ^{15}N-labelled valine.

behave as expected. They demonstrated this interesting problem further with an 18-residue peptide and suggest that the problem may depend upon the relaxation rates of the isotopes.

3. APPLICATIONS WITH PEPTIDES IN MEMBRANES

Solid-state NMR has become an important tool for the structural characterization of membrane-bound proteins. Membrane-bound proteins are difficult to characterize by solution-state NMR because the total mass is large, and they are difficult to characterize by scattering experiments because there is no long-range order. Fortunately, membrane-bound proteins are amenable for characterization by solid-state NMR experiments. Described below are a variety of structural characterizations of membrane-bound proteins by REDOR. The experiments have provided structural details of proteins and structural relationships between lipid and protein.

Liu et al.[23] performed ^{13}C–^{15}N REDOR experiments with membrane-bound phospholamban pentamers. Experiments performed with some Gln29 residues having ^{15}N labels and some Gln29 residues having ^{13}C labels showed that the glutamine side chains of the complex form a hydrogen-bonded group oriented towards the pore as is indicated in Figure 15.

Protegrin-1 (PG-1) is an 18-residue, arginine-rich, antimicrobial β-hairpin peptide with sequence RGGRLCYCRRRFCVCVGR. The peptide's ability to form pores in membranes is the root of its effectiveness in causing cell death. Using ^{13}C–^{31}P REDOR, Tang et al.[24] determined the proximity of the phosphate head group of the lipid to the carbons of various arginines in PG-1. They found the carbon atoms of arginine (Arg11 shown in Figure 16 relative to the phosphate head group) to be relatively close to the phosphorus atoms of the lipid (less than 6.5 Å) and thus proposed that the guanidinium–phosphate complex is responsible for pore formation.

The dimer structure of PG-1 in POPC bilayers was studied by Mani et al.[25] using ^{13}C–^{15}N REDOR and ^{13}C–^{19}F REDOR. Two samples were examined: one was made with a 1:1 mol ratio of [^{15}N-Cys$_{15}$]PG-1 and [^{13}C′-Cys$_{15}$]PG-1 and the other was made with a 1:1 mol ratio of [4-^{19}F-Phe$_{12}$]PG-1 and [^{13}C′-Val$_{16}$]PG-1. Figure 17 shows a sketch illustrating the labelling sites and the intermolecular distances measured by the REDOR experiment. Their work showed the two peptides dimerize in a parallel fashion when inserted in the bilayer.

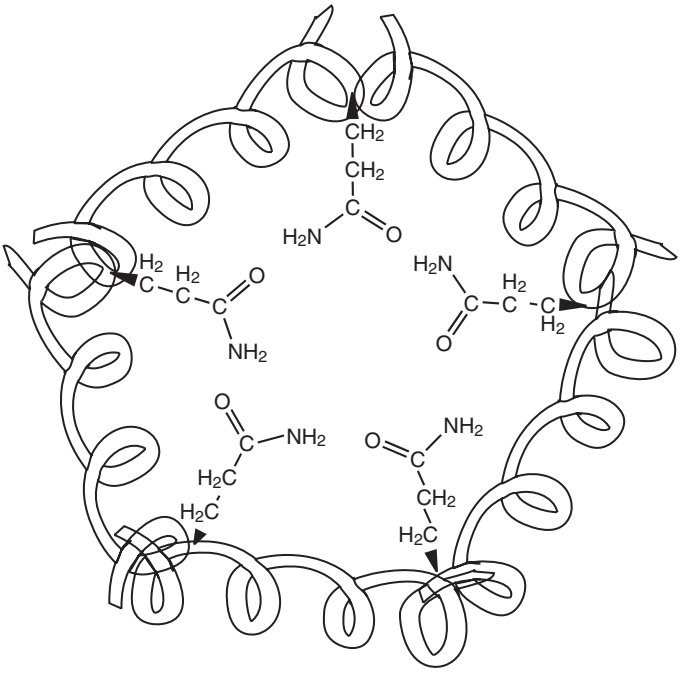


Figure 15 Illustration of the phospholamban pentamers (shown as the five helices) with glutamine side chains oriented towards the pore.

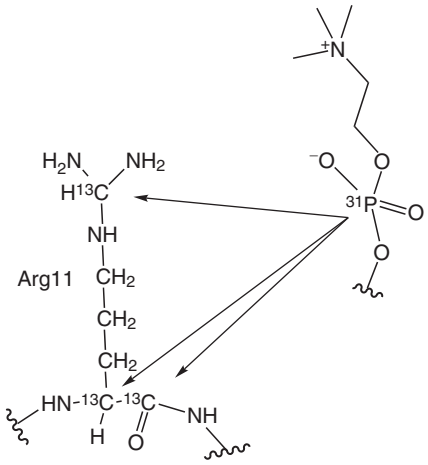


Figure 16 Arrows indicate phosphorus–carbon distances measured between the phosphate head group and three ^{13}C-labelled sites of Arg11.

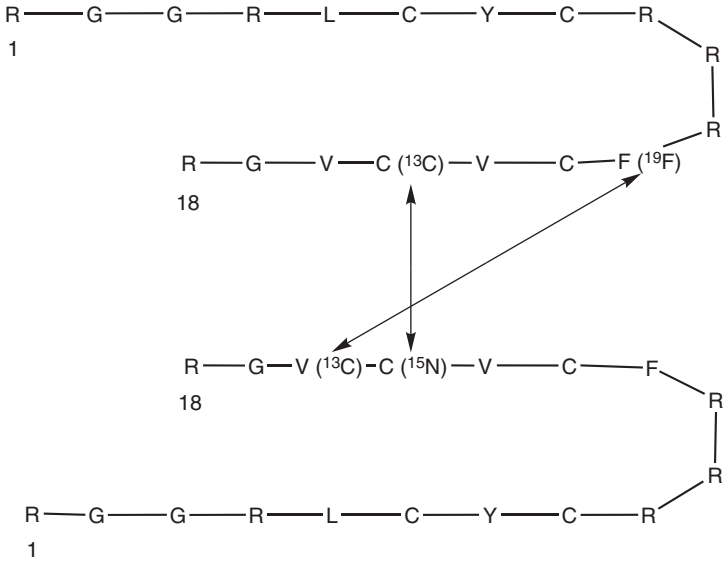


Figure 17 ^{13}C-, ^{15}N- and ^{19}F-labelled residues with the distances between specific labelled residues indicated by the arrows.

Tachyplesin I (TP) is a 17-residue antimicrobial peptide, and Doherty et al.[26] studied its structure and interaction with lipid bilayers using ^{13}C–^{1}H and ^{13}C–^{31}P REDOR. A sample with a ^{13}C label at the carbonyl carbon of Val7 labelled was used to measure the distance between the carbonyl carbon of Val7 and the proton

attached to the nitrogen of Phe8 (shown by arrow in Figure 18, top). More importantly, this distance was used to determine the dihedral angles of Cys7 (illustrated in Figure 18, bottom) and to show that its torsion angles are consistent with an anti-parallel β-sheet structure. In addition, the carbon–phosphorus distances measured between ^{13}C labels on Val7 and Gly10 and the ^{31}P nuclei of DMPC/DMPG bilayers suggests that the peptide is oriented parallel to the membrane surface and near the glycerol part of the membrane.

Porcelli et al.[27] studied the structures of the magainin antimicrobial peptides MSI-78 and MSI-594 in lipids. The sequence of MSI-78 is shown in Figure 19 and was prepared with ^{13}C label at the C′ position of Ala_{15} and an ^{15}N label at Phe_{16} (indicated by curved arrow). A similar ^{13}C-, ^{15}N-labelled sample of MSI-594 was also prepared. The ^{13}C–^{15}N REDOR experiments demonstrated that the structure of MSI-78 is the more compact structure of the two and supported the notion that MSI-78 is oligomerized in the lipid.

In a pair of papers, Toke et al.[28,29] examined the structure of $(KIAGKIA)_3$, K3, in phospholipid bilayers. K3 is an antimicrobial peptide with several proposed models of its interaction with the lipid. Using a wide array of REDOR experiments, including ^{13}C–^{31}P, ^{15}N–^{31}P and ^{13}C–^{19}F, a model of how K3 resides in lipid bilayers has emerged by measuring distances from the phosphorous nuclei of the lipid head group to the various ^{13}C- or ^{15}N-labelled residues on K3 and by measuring distances from a fluorine label located on the tail of the lipid molecules

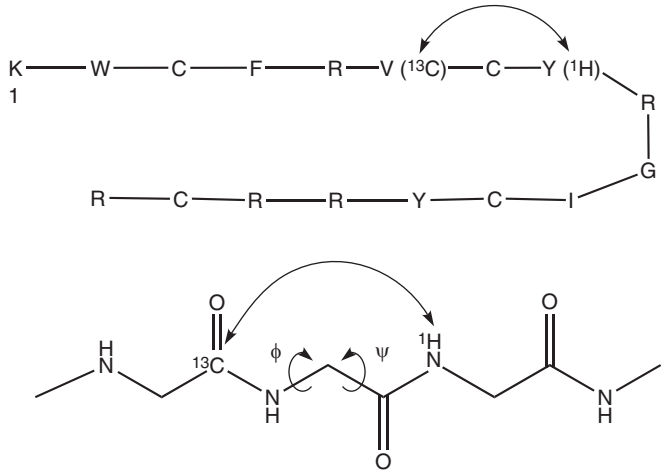


Figure 18 Sequence of tachyplesin I (top) with nuclear spin label sites connected by curved arrow. The distance between the indicated ^{13}C and ^{1}H is determined by the dihedral angles (φ, ψ) (bottom).

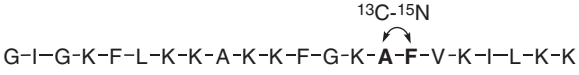


Figure 19 Sequence of antimicrobial peptide MSI-78.

to ^{13}C- or ^{15}N-labelled residues on K3 (see Figure 20). These studies show that K3 forms dimers in the lipid bilayer, the peptide is in contact with the lipid head group, and much of the peptide is surrounded by the lipid acyl chains.

Rhodopsin is a G protein-coupled receptor, which is found in the vertebrate rod photoreceptor cell. Werner et al.[30] used ^{15}N–^{13}C REDOR to identify the ^{15}N resonances coming from the backbone nitrogens of the five tryptophan residues in rhodopsin. The rhodopsin sample had all five tryptophan residues labelled with ^{15}N isotopes. Since each tryptophan residue is preceded by a unique amino acid, samples were prepared with ^{13}C labels on only one of the unique amino acids. An example is shown in Figure 21 where ^{13}C-labelled Gly(160) precedes ^{15}N-labelled Try(161). Since the ^{13}C and ^{15}N labels are so close together (1.3 Å), the ^{15}N resonance of the tryptophan next to the ^{13}C-labelled amino acid is strongly dephased by the REDOR experiment. The dephased ^{15}N resonance is then identified with a particular tryptophan (Trp(161) in this case).

The phosphates on the C-terminus of rhodopsin are important in causing arrestin to bind the cytoplasmic surface. The conformation of rhodopsin at residues 142 and 143 has been implicated in arrestin binding. Getmanova et al.[31] prepared rhodopsin samples with ^{19}F labels at residues 140 and 316 and performed ^{31}P–^{19}F REDOR in an effort to determine the proximity of the ^{31}P-containing rhodopsin C-terminus to the helix 8 (location of residue 316) and the second cytoplasmic loop (location of residue 140). Their results showed that the fluorinated residues must be farther than 12 Å from the phosphate, and thus phosphorylated rhodopsin does not have the same structure as that found for the crystal form of unphosphorylated rhodopsin.

In their work on bacteriorhodopsin, Mason et al.[32] have used ^{15}N–^{13}C REDOR to assign some of the nine ^{15}N methionine resonances that had not been assigned by other NMR methods. All methionine residues were labelled with the ^{15}N isotope, and several other select residues based on their proximity to the methionine residues were labelled with the ^{13}C isotope. Figure 22 shows an example where [$^{13}C_1$]Phe54 and [^{15}N]Met20 are close enough together to allow an ^{15}N-detected REDOR experiment to distinguish the ^{15}N resonance arising from Met20 versus the ^{15}N resonances arising from the other eight methionine residues (which are too far away to be dephased by the ^{13}C label on F54).

Kawamura et al.[33] also used ^{13}C–^{15}N REDOR as a filter for spectral assignment of NMR signals coming from specific valine and proline residues in bacteriorhodopsin.

HFP is a fusion peptide that is responsible for the fusion of the HIV virus membrane with targeted cell membranes. Zheng et al.[34] performed ^{13}C–^{15}N REDOR structural studies of trimers of ^{13}C-, ^{15}N-labelled HFP peptides prepared in different lipid/cholesterol mixtures. Their results are consistent with parallel strands of peptides that are shifted by two residues and with the presence of strands arranged in an anti-parallel fashion. These two motifs are shown in Figure 23 [only the first 21 residues of the peptide are shown; parallel (top) and anti-parallel (bottom)] along with the intra-peptide distances being measured between the ^{13}C label on Leu7 and the ^{15}N label on Phe11. They were able to rule out an in-register parallel strand arrangement.

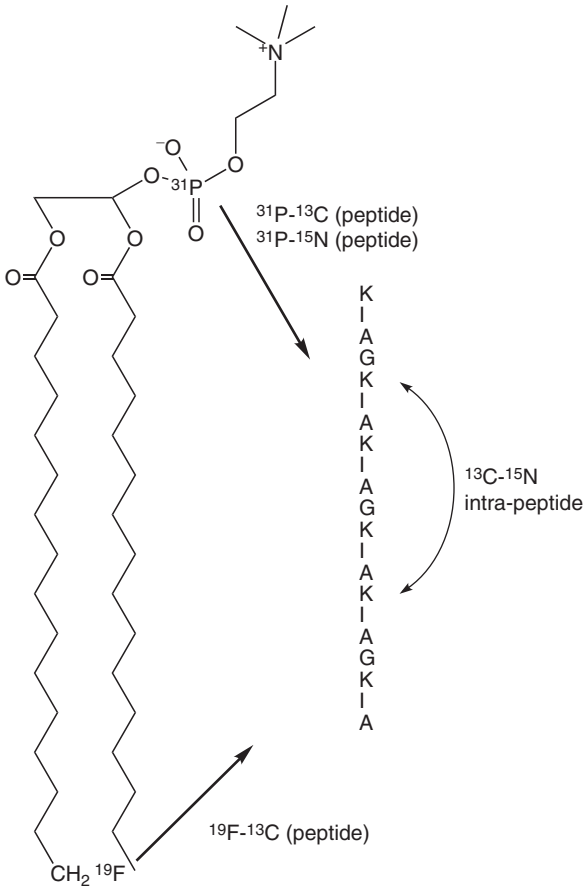


Figure 20 Diagram showing ^{31}P and ^{19}F of the lipid. Straight arrows indicate distances measured between ^{31}P (or ^{19}F) of the lipid and ^{13}C or ^{15}N labels on the peptide. The curved arrow represents intra-peptide distances between ^{13}C and ^{15}N labels.

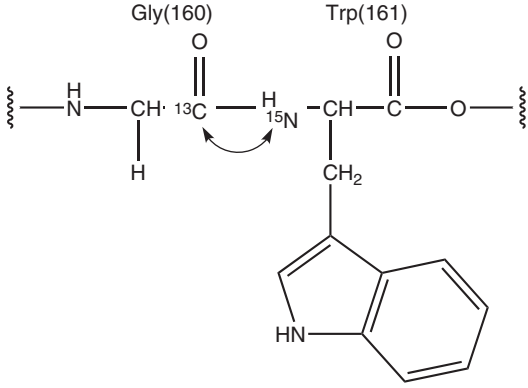


Figure 21 Directly bonded ^{13}C–^{15}N spin pair for Gly(160)–Trp(161).

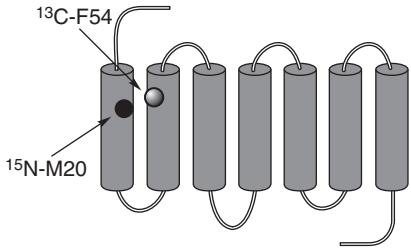


Figure 22 Model showing the proximity of M20 (large black circle) and F54 (small circle).

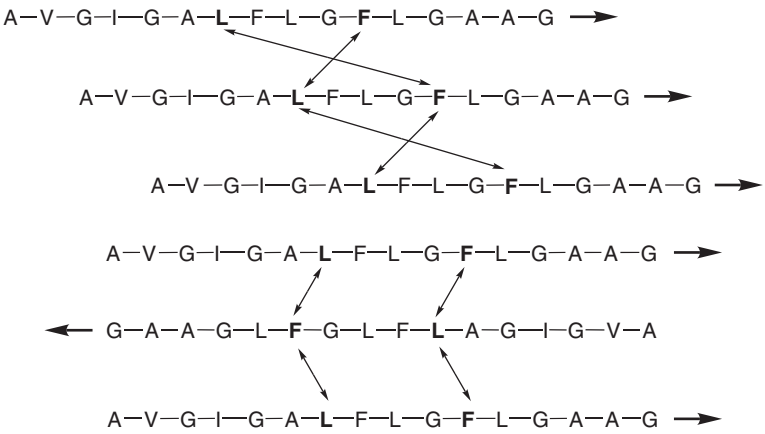


Figure 23 Parallel strands shifted by two residues (top) and anti-parallel strands (bottom). The distances between nuclear spin labels on the selected L and F residues are indicated by the connecting arrows.

Figure 24 ^{13}C- and ^{15}N-labelled fusion peptide.

In ^{13}C NMR studies of HIV and influenza fusion peptide in lipid bilayers, the simplification of ^{13}C NMR spectra in the carbonyl region was obtained by using ^{13}C–^{15}N REDOR as a filter. Such filtering can result in the elimination of the background natural-abundance ^{13}C signal. Wasniewski et al.[35] used this filtering method as a way to cleanly use the carbonyl ^{13}C chemical shifts as a tool to determine local backbone conformation of specific residues. An example is shown in Figure 24 where the F8 residue of a 20-residue fusion peptide is ^{13}C labelled and the neighboring residue is ^{15}N labelled. Hence, the ^{13}C carbonyl carbon of F8 can be recorded without interference from all the remaining natural-abundance signals with the REDOR filter.

Yang et al.,[36] Curtis-Fisk et al.[37] and Bodner et al.[38] used a similar filtering strategy in related papers.

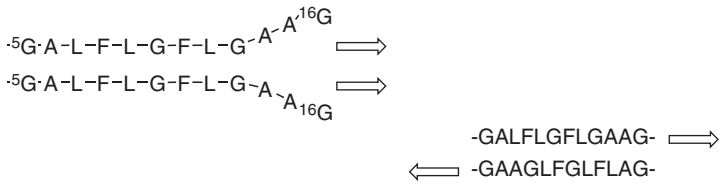


Figure 25 Parallel (top, left) and anti-parallel (bottom, right) strands of the fusion peptide.

Viral infection depends on the fusion of the virus and the cellular membrane. Yang and Weliky[39] performed ^{13}C–^{15}N REDOR experiments on a 23-residue fusion peptide (AVGIGALFLGFLGAAGSTMGARS, shown below for residues 5–16) and a 26-residue fusion peptide. The peptides are related to a strain of HIV-1. The peptides were added to lipid/cholesterol mixtures, which mimic the lipid and cholesterol composition of host cells. Their experiments showed that peptides adopt a non-helical conformation that is consistent with the presence of two equal populations of parallel and anti-parallel strands. Intra-strand hydrogen bonding is present between residues 5 and 16 in the anti-parallel form. Intra-strand hydrogen bonding is only present between residues 5 and 13 in the parallel form, which allows the remaining residues across from one another on adjacent strands to move farther apart as indicated in Figure 25(top).

Qiang et al.[40] used the same fusion peptides to study the relationship with the phosphate head group of the lipid. They prepared a series of ^{13}C-labelled (carbonyl carbons) peptides and performed ^{13}C–^{31}P REDOR to probe specific carbon (peptide)–phosphorous distances. Their data showed that the peptide region $^{14}A^{15}A^{16}G$ is near the phosphorus of the lipid.

In some of the Weliky group's other NMR work on dimers and trimers of fusion peptides, Yang et al.[41] used REDOR as a filter to simplify ^{13}C spectra in the carbonyl region.

Melittin is found in bee venom and works by disrupting cell membranes. Melittin consists of two α-helical segments, and the orientation of the helical axes relative to the lipid membrane has been proposed to be important in its action. Toraya et al.[42] have performed ^{13}C–^{15}N REDOR experiments on ^{13}C-, ^{15}N-labelled 26-residue melittin (labels at the V8 and L13 positions) in DLPC and DPPC membranes to determine the angle between the two helical segments. They found the angles to be ∼120° (illustrated in Figure 26) and concluded that melittin does not completely traverse the lipid bilayers.

4. OTHER APPLICATIONS

4.1. Compounds in membranes

Another area that solid-state NMR has proven useful is the characterization of compounds in membranes, including antibiotics and drugs. The effectiveness of many antibiotics and drugs relies on their interaction with membranes or cell

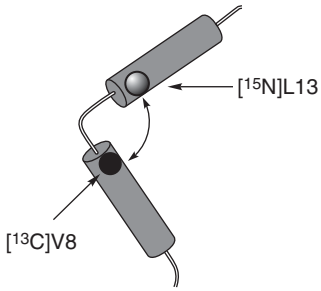


Figure 26 Model of melittin showing a bend separating two helical regions. The black circle represents the ^{13}C-labelled valine and the light circle represents the ^{15}N-labelled leucine.

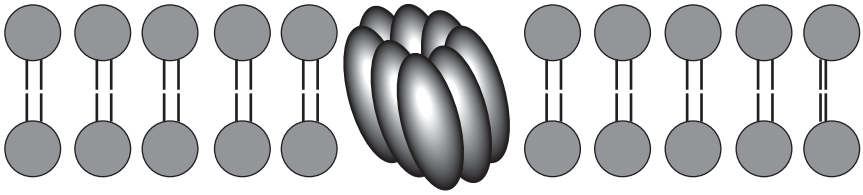


Figure 27 Barrel–stave model of AmB.

walls, so it is important to determine how these molecules interact with membranes. REDOR has proven useful in the characterization of such systems that have ^{13}C and/or ^{15}N labelling.

Matsuoka et al.[43] used ^{13}C–^{31}P REDOR and investigated the interaction between the antibiotic amphotericin B (AmB) and the membrane formed by DLPC. AmB was uniformly labelled with ^{13}C, but the extraction of carbon–phosphorus distances had proved difficult. However, the REDOR experiment was able to provide relative proximities between the carbons of AmB and the phosphorus atoms in the DLPC head groups. These REDOR results thus supported the model of the antibiotic spanning the membrane in a barrel–stave fashion as illustrated in Figure 27.

In earlier REDOR work by Matsuoka et al.,[44] AmB was shown to span the membrane formed by DMPC. For DSPC, either a double-length complex of AmB spans the membrane or a single-length complex forms in the bilayer without spanning the bilayer.

Matsumori et al.[45] performed ^{13}C–^{15}N REDOR experiments on a chemical derivative of AmB in ergosterol-containing DMPC and sterol-free DMPC. Part of the isotopically labelled AmB derivative is illustrated in Figure 28. This NMR study focused on measuring the ^{13}C–^{15}N dipolar coupling to gain insight into the motional dynamics of the AmB derivative in the membranes. The distance between the ^{13}C and ^{15}N labels is known and fixed. Hence, a measure of the dipolar coupling provides information on molecular dynamics. Comparison of dipolar couplings between the two membrane-bound samples and a pure powder sample of the AmB derivative found that the derivative is immobile in the ergosterol-containing membrane and has significant motion in the sterol-free membrane.

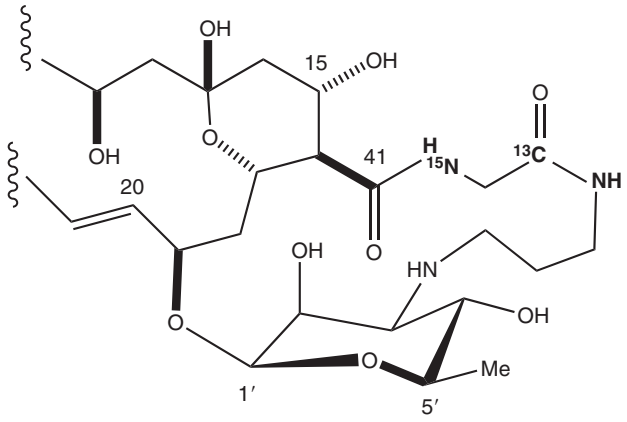


Figure 28 Part of the AmB derivative. The ^{15}N and ^{13}C labels are shown.

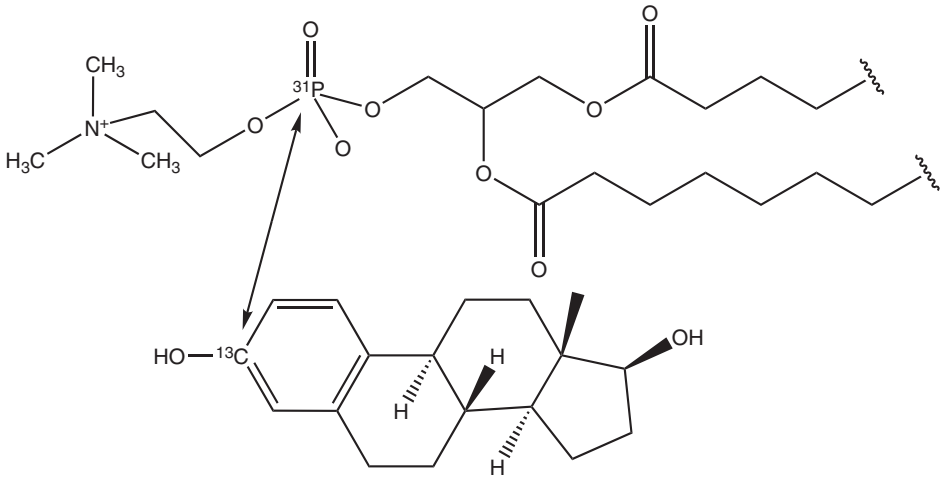


Figure 29 ^{13}C-labelled estradiol and fragment of DPPC.

Cegelski et al.[46] studied the placement of estradiol (and other sterols) in lipid bilayers using ^{13}C–^{31}P REDOR. [3-^{13}C]estradiol and DPPC are shown below. Measurement of the dipolar coupling between the ^{13}C label and the ^{31}P nucleus showed that the ^{13}C label is near the lipid head group (see Figure 29). ^{13}C–^{31}P REDOR experiments with [3-^{13}C]estradiol benzoate and [1-^{13}C]estradiol benzoate showed that the 3-^{13}C label is near the lipid head group while the 1-^{13}C is far from the head group. In addition, ^{13}C–^{19}F REDOR experiments using ^{19}F-labelled lipid and the ^{13}C-labelled sterols were performed to gain further insight into the location and orientation of the sterols with respect to the lipid.

Anandamide is an endogenous ligand for two G protein-coupled receptors. The conformation of anandamide in a membrane bilayer was studied with ^{13}C–^{31}P REDOR and ^{13}C–^{2}H REDOR by Tian et al.[47] Their ^{13}C–^{31}P REDOR experiment showed that the head group of anandamide is located near the water–lipid

interface. The ^{13}C–^{2}H REDOR experiments showed that anandamide adopts an extended conformation in the lipid bilayer. Dipolar-recoupled ^{13}C–^{31}P and ^{13}C–^{2}H interactions between lipid and anandamide are indicated by arrowed lines and curves in Figure 30.

Large amplitude motions of small molecules dispersed in liquid crystals average anisotropic interactions such as the chemical shift and dipolar interactions. Often, the motion is not isotropic so complete averaging does not occur. Trempe et al.[48] used ^{1}H–^{15}N REDOR applied during ^{1}H acquisition to recover the ^{1}H–^{15}N residual dipolar coupling in benzamide (^{15}N-labelled compound shown in Figure 31) dispersed in polyacrylamide-stabilized Pf1 phage medium.

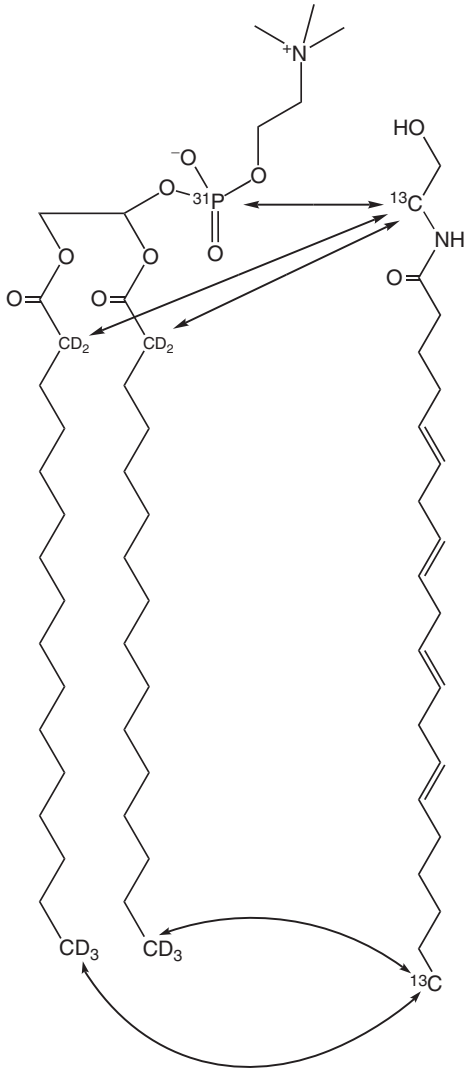


Figure 30 ^{2}H-labelled lipid and ^{13}C-labelled anandamide (shown as an extended conformation).

Human factor Xa (Fxa), which acts during blood coagulation, can be inhibited by pharmacological agents. The enzyme-bound conformation of Berlex Biosciences human factor Xa (Fxa) inhibitor, ZK-816042, was studied by Studelska et al.[49] using ^{13}C–^{15}N and ^{13}C–^{19}F REDOR. Multiple conformations of the molecule are possible because of the many single bonds linking the various ring structures as shown in Figure 32. The results of the REDOR experiments suggested a U-shape structure of ZK-816042.

Similar work on Berlex Biosciences inhibitor ZK-806299 was also examined by McDowell et al.[50] using ^{13}C–^{15}N and ^{13}C–^{19}F REDOR.

4.2. Nucleotides

Grant et al.[51] combined ^{23}Na–^{31}P REDOR with ^{23}Na MQMAS to determine the proximity of sodium ions to the phosphorus sites in the sodium–nucleotide complex disodium adenosine-5′-triphosphate trihydrate (structure shown in Figure 33). They were able to determine proximities to phosphorus for two of

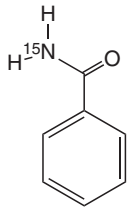


Figure 31 ^{15}N-labelled benzamide.

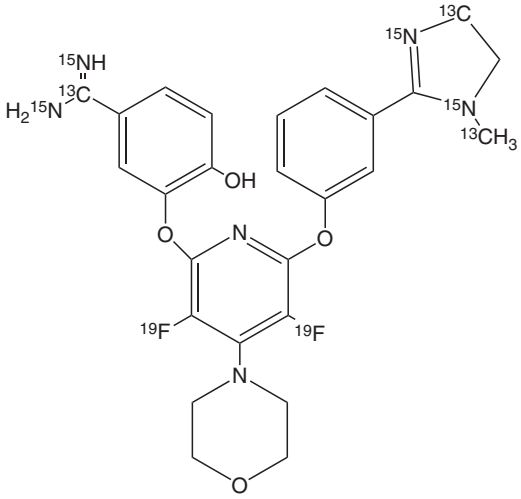


Figure 32 ^{13}C- and ^{15}N-labelled ZK-816042 (Berlex Biosciences).

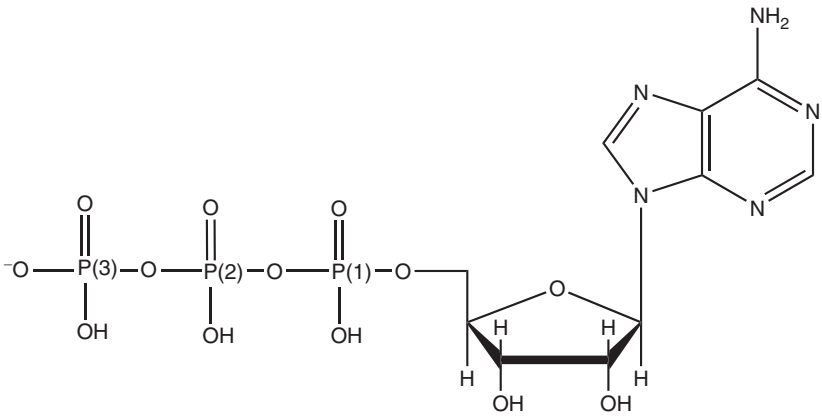


Figure 33 Adenosine-5′-triphosphate. Sodium ions not shown.

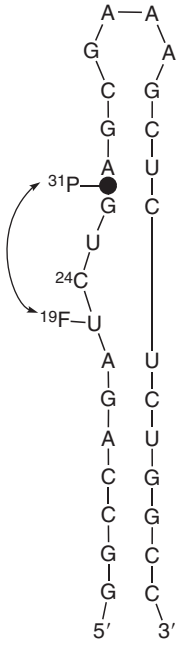


Figure 34 Modified RNA. The ^{31}P and ^{19}F sites are shown.

the four sodium ions (as assigned by the ^{23}Na MQMAS spectrum) but were not able to clearly determine the structural arrangement of the remaining two sodium ions relative to the nucleotide.

^{31}P–^{19}F REDOR experiments were performed by Olsen et al.[52] on an HIV-1 trans-activation response (TAR) RNA element (29-mer) to probe conformational changes which may occur upon binding of the viral protein tat. The TAR RNA was modified to include phosphorothioate (indicated by the dot) and U23 was modified as $^{2'F}$U (2′-fluoro-2′-deoxyuridine) and is shown in Figure 34. These

modifications provided ^{31}P and ^{19}F spins, which were used to probe the conformation in the binding region. Their results showed a significant increase in the phosphorous–fluorine distance upon binding of tat.

Uracil DNA glycosylase is one of the enzymes responsible for DNA repair of damaged bases. In an effort to understand the structural aspects of DNA bound to glycosylases, Jiang et al.[53] prepared DNA analogues with 2,4-difluorophenyl nucleotide (shown in Figure 35). The ^{19}F nuclei contained in this analogue were used to determine structural aspects of the fluorine-containing nucleotide relative to the DNA phosphorus-containing framework by using ^{31}P–^{19}F REDOR. This work on DNA duplexes bound to uracil DNA glycosylase suggested that the fluorinated substituent adopts a structure mimicking that found in a previously discovered intermediate state for the uracil-flipping pathway.

^{31}P–^{19}F REDOR was used to determine the effect of chain spacing upon binding with an antibiotic. Distamycin A is an antibiotic that binds non-covalently in the minor groove of B-form DNA rich in A and T bases. Olsen et al.[54] prepared DNA with one oligomer which had a T replaced by 2′-fluoro-2′-deoxyuridine ($^{2\prime F}$U) and another modified with a phosphorothioate (dot in Figure 36). These modifications provided a cross-chain ^{19}F–^{31}P spin pair that was used to probe the chain separation upon binding with distamycin. Large changes in chain spacing were found upon binding.

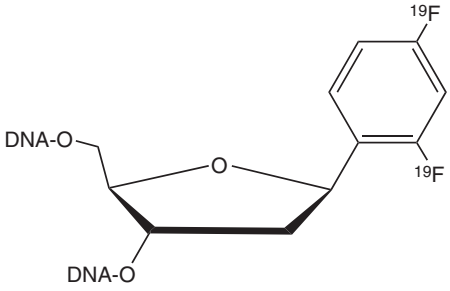


Figure 35 2,4-difluorophenyl nucleotide.

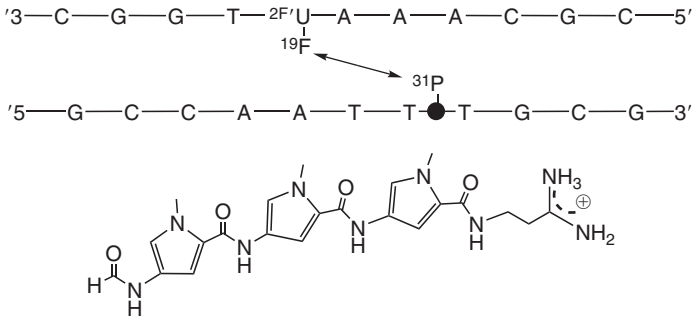


Figure 36 Distamycin A and DNA to provide specific ^{19}F and ^{31}P labels for distance measurements.

4.3. Membranes

Tang et al.[55] compared the conformation of crystalline POPC to that of POPC-containing trehalose. The structure of POPC is provided in Figure 37 with the arrowed curves indicating the two phosphorus–carbon distances measured by ^{13}C–^{31}P REDOR. Their data showed that the lipid head group and backbone conformation of POPC is not affected by the addition of trehalose.

Additionally, Tang et al.[55] measured the ^{13}C–^{31}P dipolar couplings between the lipid phosphorus and the C′ carbon of Val_{16} in the protein PG-1 placed in a lipid made of hydrated POPE/POPG and in a lipid POPE/POPG with trehalose. They demonstrated that the dipolar coupling was motionally averaged for the hydrated sample for temperatures as low as 226 K; whereas, the dipolar coupling produced the rigid-limit value at 253 K for the trehalose-protected system.

4.4. Carbohydrates

The two anomers of glucosamine hydrochloride (shown in Figure 38 with the β-form on the left and the α-form on the right) and *N*-acetylglucosamine were each studied by Chen et al.[56] to determine if there were any significant structural differences between anomeric forms. Samples with ^{15}N labels and ^{13}C labels at the C1 position were studied by ^{13}C–^{15}N REDOR. The resulting carbon–nitrogen distance measurements showed a small difference in internuclear distance (~0.1 Å) between the β- and α-forms for each carbohydrate. Such small distance differences suggested that both anomers have the chair conformation but that small differences do exist between the chair conformations.

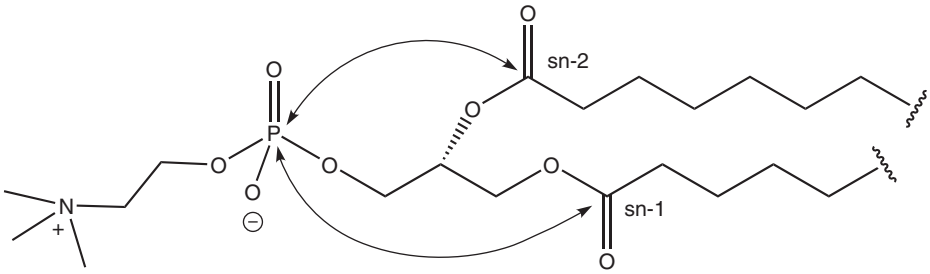


Figure 37 Partial structure of POPC.

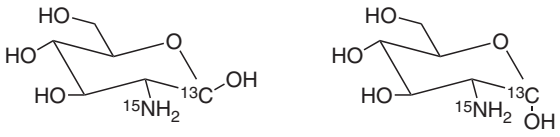


Figure 38 Glucosamine with ^{13}C and ^{15}N labels.

5. CONCLUSIONS AND SUMMARY OF EARLIER APPLICATIONS

The brief descriptions of REDOR applications presented here show that a wide variety of problems benefit from the experiment. The complexity of many of the biological systems studied is especially notable. NMR provides important structural details that would otherwise be very difficult to obtain. In most of the examples presented above, many varied NMR techniques were used to characterize the materials. The continued advancement of solid-state NMR methodology will serve to contribute to solving important structural problems in biological materials.

A partial summary of biological applications of REDOR from 1998 to 2002 is given in Table 1. This table is provided as a reference to more macromolecular

TABLE 1 Summary of biological applications from 1998 to 2002

Keywords	Nuclei	References
HIV-1, influenza fusion peptide	^{13}C, ^{15}N	57
Alzheimer's β-amyloid peptide	^{13}C, ^{15}N	58
M2-TMP, influenza A	^{13}C, ^{15}N	59
Heptapeptide	^{13}C, ^{19}F	60
Chemotactic tripeptide	^{13}C, ^{15}N	61
Neu receptor tyrosine kinase	^{13}C, ^{15}N	62
Staphylococcus aureus, chloroeremomycin	^{13}C, ^{15}N, ^{19}F, ^{31}P	63
Glycophorin A	^{13}C, ^{15}N	64
Bacteriorhodopsin	^{13}C, ^{15}N	65
Statherin, hydroxyapatite	^{13}C, ^{15}N	66
Silk, *Bombyx mori*	^{13}C, ^{15}N	67
Protein hydrogel	^{13}C, ^{15}N	68
Phospholamban	^{13}C, ^{15}N	69
Hydrogen bond lengths, polypeptides, α-helix	^{13}C, ^{15}N	70
Tripeptides, torsion angles	^{2}H, ^{13}C	71,72
Ligand, chemoreceptor	^{13}C, ^{19}F	73
Amyloid fibril	^{13}C, ^{15}N	74
Ubiquitin	^{13}C, ^{15}N	75
Adenosine-5′-triphosphate, magnesium	^{1}H, ^{25}Mg	76
Bacteriorhodopsin	^{13}C, ^{15}N	77
HIV-1	^{13}C, ^{15}N	78
Polymorph, gluconamide	^{13}C, ^{15}N	79
Paclitaxel	^{13}C, ^{15}N, ^{19}F	80
CO_2 fixation, bolaamphiphile	^{13}C, ^{15}N	81
Nucleic acids, phosphodiester	^{19}F, ^{31}P	82
Cross links, mussel byssal plaques	^{2}H, ^{13}C	83
Pheromone	^{13}C, ^{15}N	84
Fluoro-lumazine, lumazine synthase	^{13}C, ^{19}F	85
Inhibitor, blood coagulant factor	^{13}C, ^{15}N, ^{19}F	86
Enkephalin	^{13}C, ^{15}N	87,88
α-spectrin	^{13}C, ^{15}N	89
Spider dragline silk	^{13}C, ^{15}N	90

systems studied by REDOR. It is also provided to help newcomers to the field find examples that are relevant to their research interest. The table lists the nuclear spins used in the REDOR experiment and keywords describing the systems investigated.

ACKNOWLEDGMENTS

I thank Ms. Heather Lotz for assistance in organizing the material for this chapter. Support for this project was provided by NSF Grant CHE-0091663.

REFERENCES

1. T. Gullion and J. Schaefer, *J. Magn. Reson.*, 1989, **81,** 196.
2. T. Gullion and J. Schaefer, in: *Advances in Magnetic Resonance*, Vol. 13, W. S. Warren, ed., Academic Press, San Diego, 1989.
3. C. P. Slichter, *Principles of Magnetic Resonance*. Springer-Verlag, New York, 1989.
4. J. Schaefer, H. Jiang, A. E. Ransome and T. J. Kappock, *Biochemistry*, 2007, **46,** 9507.
5. A. T. Petkova, W.-M. Yau and R. Tycko, *Biochemistry*, 2006, **45,** 498.
6. J. C. C. Chan, N. A. Oyler, W.-M. Yau and R. Tycko, *Biochemistry*, 2005, **44,** 10669.
7. K. L. Sciarretta, D. J. Gordon, A. T. Petkova, R. Tycko and S. C. Meredith, *Biochemistry*, 2005, **44,** 6003.
8. E. Paci, J. Gsponer, X. Salvatella and M. Vendruscolo, *J. Mol. Biol.*, 2004, **340,** 555.
9. C. P. Jaroniec, C. E. MacPhee, N. S. Astrof, C. M. Dobson and R. G. Griffin, *Proc. Natl. Acad. Sci. USA*, 2002, **99,** 16052.
10. R. A. Kammerer, D. Kostrewa, J. Zurdo, A. Detken, C. Garcia-Echeverria, J. D. Green, S. A. Muller, B. H. Meier, F. K. Winkler, C. M Dobson and M. O. Steinmetz, *Proc. Natl. Acad. Sci. USA*, 2004, **101,** 4435.
11. D. J. Gordon, J. J. Balbach, R. Tycko and S. C. Meredith, *Biophys. J.*, 2004, **86,** 428.
12. A. T. Petkova, G. Buntkowsky, F. Dyda, R. D. Leapman, W.-M. Yau and R. Tycko, *J. Mol. Biol.*, 2004, **335,** 247.
13. A. Naito, M. Kamihira, R. Inoue and H. Saito, *Magn. Reson. Chem.*, 2004, **42,** 247.
14. R. Goobes, G. Goobes, W. J. Shaw, G. P. Drobny, C. T. Campbell and P. S. Stayton, *Biochemistry*, 2007, **46,** 4725.
15. G. M. Bernard, M. Miskolzie, G. Kotovych and R. E. Wasylishen, *Can. J. Chem.*, 2004, **82,** 1554.
16. W. J. Shaw, A. A. Campbell, M. L. Paine and M. L. Snead, *J. Biol. Chem.*, 2004, **279,** 40263.
17. X. L. Yao and M. Hong, *J. Am. Chem. Soc.*, 2004, **126,** 4199.
18. M. Hong, D. Isailovic, R. A. McMillan and V. P. Conticello, *Biopolymers*, 2003, **70,** 158.
19. T. Gullion, R. Kishore and T. Asakura, *J. Am. Chem. Soc.*, 2003, **125,** 7510.
20. T. Kameda, C. Zhao, J. Ashida and T. Asakura, *J. Magn. Reson.*, 2003, **160,** 91.
21. M. Heller, M. Sukopp, N. Tsomaia, M. John, D. F. Mierke, B. Reif and H. Kessler, *J. Am. Chem. Soc.*, 2006, **128,** 13806.
22. S. Sharpe, N. Kessler, J. A. Anglister, W.-M. Yau and R. Tycko, *J. Am. Chem. Soc.*, 2004, **126,** 4979.
23. W. Liu, J. Z. Fei, T. Kawakami and S. O. Smith, *Biochim. Biophys. Acta*, 2007, **1768,** 2971.
24. M. Tang, A. J. Waring and M. Hong, *J. Am. Chem. Soc.*, 2007, **129,** 11438.
25. R. Mani, M. Tang, X. Wu, J. J. Buffy, A. J. Waring, M. A. Sherman and M. Hong, *Biochemistry*, 2006, **45,** 8341.
26. T. Doherty, A. J. Waring and M. Hong, *Biochemistry*, 2006, **45,** 13323.
27. F. Porcelli, B. A. Buck-Koehntop, S. Thennarasu, A. Ramamoorthy and G. Veglia, *Biochemistry*, 2006, **45,** 5793.
28. O. Toke, *Biopolymers*, 2005, **80,** 717.
29. O. Toke, W. L. Maloy, S. J. Kim, J. Blazyk and J. Schaefer, *Biophys. J.*, 2004, **87,** 662.
30. K. Werner, I. Lehner, H. K. Dhiman, C. Richter, C. Glaubitz, H. Schwalbe, J. Klein-Seetharaman and H. G. Khorana, *J. Biomol. NMR*, 2007, **37,** 303.

31. E. Getmanova, A. B. Patel, J. Klein-Seetharaman, M. C. Loewen, P. J. Reeves, N. Friedman, M. Sheves, S. O. Smith and H. G. Khorana, *Biochemistry*, 2004, **43,** 1126.
32. A. J. Mason, G. J. Turner and C. Glaubitz, *FEBS J.*, 2005, **272,** 2152.
33. I. Kawamura, M. Ohmine, J. Tanabe, S. Tuzi, H. Saito and A. Naito, *Biochim. Biophys. Acta*, 2007, **1768,** 3090.
34. Z. Zheng, R. Yang, M. L. Bodner and D. P. Weliky, *Biochemistry*, 2006, **45,** 12960.
35. C. M. Wasniewski, P. D. Parkanzky, M. L. Bodner and D. P. Weliky, *Chem. Phys. Lipids*, 2004, **132,** 89.
36. J. Yang, M. Prorok, F. J. Castellino and D. P. Weliky, *Biophys. J.*, 2004, **87,** 1951.
37. J. Curtis-Fisk, C. Preston, Z. Zheng, R. M. Worden and D. P. Weliky, *J. Am. Chem. Soc.*, 2007, **129,** 11320.
38. M. L. Bodner, C. M. Gabrys, P. D. Parkanzky, J. Yang, C. A. Duskin and D. P. Weliky, *Magn. Reson. Chem.*, 2004, **42,** 187.
39. J. Yang and D. P. Weliky, *Biochemistry*, 2003, **42,** 11879.
40. W. Qiang, J. Yang and D. P. Weliky, *Biochemistry*, 2007, **46,** 4997.
41. R. Yang, J. Yang and D. P. Weliky, *Biochemistry*, 2003, **42,** 3527.
42. S. Toraya, K. Nishimura and A. Naito, *Biophys. J.*, 2004, **87,** 3323.
43. S. Matsuoka, H. Ikeuchi, Y. Umegawa, N. Matsumori and M. Murata, *Bioorg. Med. Chem.*, 2006, **14,** 6608.
44. S. Matsuoka, H. Ikeuchi, N. Matsumori and M. Murata, *Biochemistry*, 2005, **44,** 704.
45. N. Matsumori, Y. Sawada and M. Murata, *J. Am. Chem. Soc.*, 2006, **128,** 11977.
46. L. Cegelski, C. V. Rice, R. D. O'Connor, A. L. Caruano, G. P. Tochtrop, Z. Y. Cai, D. F. Covey and J. Schaefer, *Drug Dev. Res.*, 2006, **66,** 93.
47. X. Tian, J. Guo, F. Yao, D.-P. Yang and A. Makriyannis, *J. Biol. Chem.*, 2005, **280,** 29788.
48. J.-F. Trempe, A. Denisov and K. Gehring, *J. Magn. Reson.*, 2003, **164,** 329.
49. D. R. Studelska, L. M. McDowell, M. Adler, R. D. O'Conner, A. K. Mehta, W. J. Guilford, J. L. Dallas, D. Arnaiz, D. R. Light and J. Schaefer, *Biochemistry*, 2003, **42,** 7942.
50. L. M. McDowell, M. A. McCarrick, D. R. Studelska, R. D. O'Connor, D. R. Light, W. J. Guilford, D. Arnaiz, M. Adler, J. L. Dallas, B. Poliks and J. Schaefer, *J. Med. Chem.*, 2003, **46,** 359.
51. C. V. Grant, D. McElheny, V. Frydman and L. Frydman, *Magn. Reson. Chem.*, 2006, **44,** 366.
52. G. L. Olsen, T. E. Edwards, P. Deka, G. Varani, S. Th. Sigurdsson and G. P. Drobny, *Nucleic Acids Res.*, 2005, **33,** 3447.
53. Y. L. Jiang, L. M. McDowell, B. Poliks, D. R. Studelska, C. Cao, G. S. Potter, J. Schaefer, F. Song and J. T. Stivers, *Biochemistry*, 2004, **43,** 15429.
54. G. L. Olsen, E. A. Louie, G. P. Drobny and S. Th. Sigurdsson, *Nucleic Acids Res.*, 2003, **31,** 5084.
55. M. Tang, A. J. Waring and M. Hong, *J. Magn. Reson.*, 2007, **184,** 222.
56. Y.-Y. Chen, S.-Y. Luo, S.-C. Hung, S. I. Chan and D.-L. M. Tzou, *Carbohydr. Res.*, 2005, **340,** 723.
57. J. Yang, P. D. Parkanzky, M. L. Bodner, C. A. Duskin and D. P. Weliky, *J. Magn. Reson.*, 2002, **159,** 101.
58. O. N. Antzutkin, R. D. Leapman, J. J. Balbach and R. Tycko, *Biochemistry*, 2002, **41,** 15436.
59. K. Nishimura, S. Kim, L. Zhang and T. A. Cross, *Biochemistry*, 2002, **41,** 13170.
60. D. Seebach, T. Sifferlen, D. J. Bierbaum, M. Rueping, B. Jaun, J. Schaefer, A. K. Mehta, R. D. O'Connor, B. H. Meier, M. Ernst and A. Glatti, *Helv. Chim. Acta*, 2002, **85,** 2877.
61. C. M. Rienstra, M. Hohwy, L. J. Mueller, C. P. Jaroniec, B. Reif and R. G. Griffin, *J. Am. Chem. Soc.*, 2002, **124,** 11908.
62. S. O. Smith, C. Smith, S. Shekar, O. Peersen, M. Ziliox and S. Aimoto, *Biochemistry*, 2002, **41,** 9321.
63. S. J. Kim, L. Cegelski, D. R. Studelska, R. D. O'Connor, A. K. Mehta and J. Schaefer, *Biochemistry*, 2002, **41,** 6967.
64. S. O. Smith, M. Eilers, D. Song, E. Crocker, W. Ying, M. Groesbeek, G. Metz, M. Ziliox and S. Aimoto, *Biophys. J.*, 2002, **82,** 2476.
65. A. T. Petkova, M. Hatanaka, C. P. Jaroniec, J. G. Hu, M. Belenky, M. Verhoeven, J. Lugtenburg, R. G. Griffin and J. Herzfeld, *Biochemistry*, 2002, **41,** 2429.
66. J. R. Long, W. J. Shaw, P. S. Stayton and G. P. Drobny, *Biochemistry*, 2001, **40,** 15451.
67. C. Zhao and T. Asakura, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2001, **39,** 301.
68. S. B. Kennedy, E. R. deAzevedo, W. A. Petka, T. P. Russel, D. A. Tirrell and M. Hong, *Macromolecules*, 2001, **34,** 8675.

69. S. O. Smith, T. Kawakami, W. Liu, M. Ziliox and S. Aimoto, *J. Mol. Biol.*, 2001, **313,** 1139.
70. S. Kimura, A. Naito, H. Saito, K. Ogawa and A. Shoji, *J. Mol. Struct.*, 2001, **562,** 197.
71. I. Sack, Y. S. Balazs, S. Rahimipour and S. Vega, *J. Am. Chem. Soc.*, 2000, **122,** 12263.
72. I. Sack, Y. S. Balazs, S. Rahimipour and S. Vega, *J. Magn. Reson.*, 2001, **148,** 104.
73. O. J. Murphy, F. A. Kovacs, E. L. Sicard and L. K. Thompson, *Biochemistry*, 2001, **40,** 1358.
74. J. J. Balbach, Y. Ishii, O. N. Antzutkin, R. D. Leapman, N. W. Rizzo, F. Dyda, J. Reed and R. Tycko, *Biochemistry*, 2000, **39,** 13748.
75. D. Huster, S. Yamaguchi and M. Hong, *J. Am. Chem. Soc.*, 2000, **122,** 11320.
76. C. V. Grant, V. Frydman and L. Frydman, *J. Am. Chem. Soc.*, 2000, **122,** 11743.
77. M. Helme, H. Patzelt, A. Ockenfels, W. Gartner, D. Oesterhelt and B. Bechinger, *Biochemistry*, 2000, **39,** 10066.
78. J. J. Balbach, J. Yang, D. P. Weliky, P. J. Steinbach, V. Tugarinov, J. Anglister and R. Tycko, *J. Biomol. NMR*, 2000, **16,** 313.
79. I. Sack, S. Macholl, J. H. Fuhrhop and G. Buntkowsky, *Phys. Chem. Chem. Phys.*, 2000, **2,** 1781.
80. Y. Li, B. Poliks, L. Cegelski, M. Poliks, Z. Gryczynski, G. Psizczek, P. G. Jagtap, D. R. Studelska, D. G. I. Kingston, J. Schaefer and S. Bane, *Biochemistry*, 2000, **39,** 281.
81. V. Schimming, C.-G. Hoelger, G. Buntkowsky, I. Sack, J.-H. Fuhrhop, S. Rocchetti and H.-H. Limbach, *J. Am. Chem. Soc.*, 1999, **121,** 4892.
82. M. E. Meritt, S. Th. Sigurdsson and G. P. Drobny, *J. Am. Chem. Soc.*, 1999, **121,** 6070.
83. L. M. McDowell, L. A. Burzio, J. H. Waite and J. Schaefer, *J. Biol. Chem.*, 1999, **274,** 20293.
84. B. Arshava, M. Breslav, O. Antohi, R. E. Stark, J. R. Garbow, J. M. Becker and F. Naider, *Solid State Nucl. Magn. Reson.*, 1999, **14,** 117.
85. J. M. Goetz, B. Poliks, D. R. Studelska, M. Fischer, K. Kugelbrey, A. Bacher, M. Cushman and J. Schaefer, *J. Am. Chem. Soc.*, 1999, **121,** 7500.
86. L. M. McDowell, M. A. McCarrick, D. R. Studelska, W. J. Guilford, D. Arnaiz, J. L. Dallas, D. R. Light, M. Whitlow and J. Schaefer, *J. Med. Chem.*, 1999, **42,** 3910.
87. K. Nishimura, A. Naito, S. Tuzi, H. Saito, C. Hashimoto and M. Aida, *J. Phys. Chem. B*, 1998, **102,** 7476.
88. M. Kamihira, A. Naito, K. Nishimura, S. Tuzi and H. Saito, *J. Phys. Chem. B*, 1998, **102,** 2826.
89. S. Macholl, I. Sack, H.-H. Limbach, J. Pauli, M. Kelly and G. Buntkowsky, *Magn. Reson. Chem.*, 2000, **38,** 596.
90. L. W. Jelinski, A. Blye, O. Liivak, C. Michal, G. LaVerde, A. Seidel, N. Shah and Z. Yang, *Int. J. Biol. Macromol.*, 1999, **24,** 197.

CHAPTER 5

Electrophoretic NMR—Ions, Molecules, Mixtures, Pores and Complexes

P. C. Griffiths

Contents

Abstract

Electrophoretic NMR couples the chemical specificity of NMR, the ability of magnetic field gradients to quantify displacement and the coherent motion of charged particles induced by electric fields to allow the measurement of the electrophoretic mobility, a particularly interesting parameter in heterogeneous, multi-component solutions. Like the much more frequently studied self-diffusion coefficient, the electrophoretic mobility is sensitive to the size and shape of the ionic species, modulated by the presence of any structure or interactions experienced within the solution. This article reviews the more recent literature that utilises electrophoretic NMR to probe the nature of a range of nanotechnology related systems, specifically polyelectrolyte and surfactant containing solutions, membranes and porous media.

Key Words: Electrophoretic mobility, Charge, Ions, Surfactants, Polyelectrolytes, Micelles, Porous media, Membranes.

School of Chemistry, Cardiff University, Cardiff CF10 3AT, United Kingdom

Annual Reports on NMR Spectroscopy, Volume 65
ISSN 0066-4103, DOI: 10.1016/S0066-4103(08)00205-6

1. THEORETICAL AND HISTORICAL PERSPECTIVES

Under the effect of a constant electric field $E_{dc,}$ ions attain a steady-state velocity $v_e = \mu E_{dc}$, where μ is the electrophoretic mobility. The electrophoretic mobility is determined by the charge on the ion, its size and the structure of the medium in which it moves. Quantifying or harnessing the migration of ions in electric fields forms the basis of several analytical techniques since the electric force causing the motion balances the viscous drag acting on the ion, a force that depends on the hydrodynamic radius (size and shape) of the ion and the viscosity and structure within the medium.

Electrophoresis is an established field but the novelty of electrophoretic NMR (eNMR) arises due to the use of magnetic field gradients to quantify the drift velocities of the NMR active ions in the electric field.[1–5] The chemical specificity of the NMR signal therefore enables the study of small ions and complex mixtures. On most NMR spectrometers, the magnetic field gradient $g(z)$ is applied co-axially to the main magnetic field B_0, such that $B_z = B_0 + g(z)z$. The spins viewed within a rotating frame of reference $\omega_0 = -\gamma B_0$ about the z axis, are wound into a helix with pitch $2\pi/\gamma\delta g(z)$ by the first of two magnetic field gradient pulses of duration δ and amplitude $\gamma g(z)$. This helix is unwound by the second identical but of opposite phase field gradient pulse leading to a spin-echo. Diffusion of the ions leads to a loss of phase coherence and the associated attenuation of the spin-echo intensity. An electric field of duration t_1 applied in the z-direction during the time interval between these two gradient pulses, induces a phase, $\theta = qv_e t_1$ in the signal related to the drift velocity, v_e. Accordingly, the drift velocity may be quantified via changes in phase shift (Figure 1A) or the frequency of a cosine modulation, $\cos(E_{dc}q\mu t_1)$ imposed on the signal intensity (Figure 1B).

Sometimes known as DCNMR[6] or magnetic resonance electrophoresis MREP,[7,8] several groups[8–16] have contributed to the development of electrophoretic NMR following its initial conception by Packer,[17] most notably Holz[18] and Johnson, Jr.[9,10] The field has been reviewed and a number of excellent articles have been written.[6,10,19–23] This review builds on those, but with some inevitable overlap.

2. PRACTICAL ASPECTS

Electrophoretic NMR (eNMR) has evolved significantly since its conception but is still far from a 'standard' technique. Many of the common artefacts have been recognized and solutions proposed[24] and more recently, comprehensively studied by Stilbs et al.[22,23,25]; the more important of these are outlined here. Figure 2 displays some of the more common electrophoretic cell geometries—the counter-flow of ions in the U-tube (Figure 2A) results in a loss of the signs of the mobilities and can lead to more problematic shimming; cylindrical geometries (Figure 2B) minimise these aspects but the required gel plug introduces a further limitation; the cell presented in Figure 2C) benefits from a good filling factor and reduced

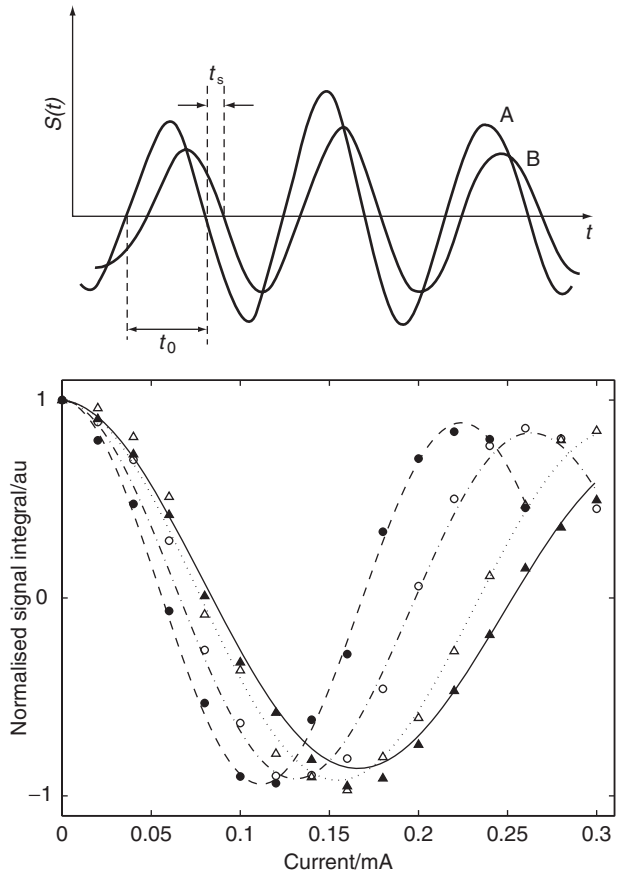


Figure 1 Determination of the electrophoretic mobility from phase shifts or cosine modulation to spin-echo intensities. *Top*: Two off-resonance signals in the presence of flow (A) spin echo detected when the gradient is switched off (B) spin echo detected when the gradient is switched on. Source: Electrophoretic NMR by C. S. Johnson, in Encyclopedia of Nuclear Magnetic Resonance, ed. D. M. Grant and R. K. Harris. 1996. Copyright John Wiley & Sons. Reproduced with permission. *Bottom:* Normalised signal intensity as a function of electric field current for a series of 50 mM binary surfactant solutions with various ratios of anionic/non-ionic surfactant.The lines through the data correspond to a fit to the cosine modulation. Reproduced with permission from *Langmuir*, 2001, **17**, 7178–7181. Copyright 2001 American Chemical Society.

electro-osmosis but gas generation is an issue, although this may be minimised by the use of suitable electrodes.

The biggest problem in eNMR—electro-osmosis—occurs when counter-ions adsorbed to the charged glass surface induce an electric double-layer which on application of the electric field, leads to the migration of the ions driving the flow of the solvent. A simple solution to suppress electro-osmosis is to coat the glass surface with a polymer such as poly(acrylamide) PAA, and further to suspend this arrangement in an NMR tube filled with D_2O to aid temperature control.

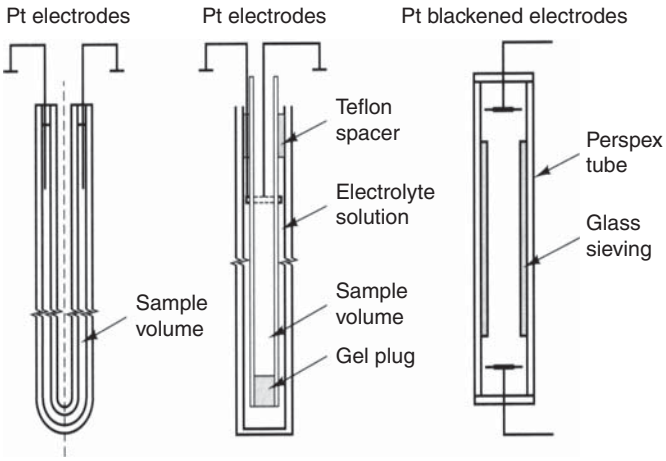


Figure 2 Common sample environments. Electrophoretic NMR by C. S. Johnson, in Encyclopedia of Nuclear Magnetic Resonance, ed. D. M. Grant and R. K. Harris. 1996. Copyright John Wiley & Sons. Reproduced with permission.

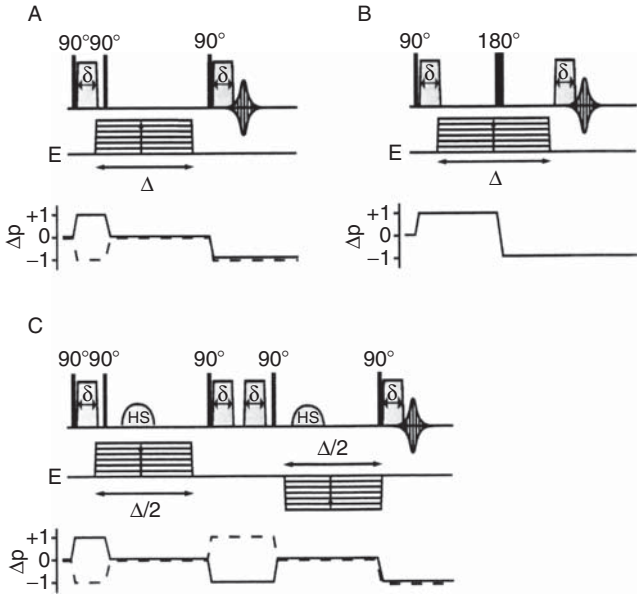


Figure 3 Pulse sequences often used in electrophoretic NMR experiments; (A) stimulated echo (B) spin echo and (C) double stimulated echo. Reproduced with permission from the Doctoral Thesis of Erik Pettersson Royal Institute of Technology 2005.

A version of the electrophoretic double stimulated echo pulse sequence, (EDSTE) Figure 3[22,23] was shown to be a facile approach to solve many of the drawbacks associated with electro-osmosis. In this sequence, since the electrophoretic velocity

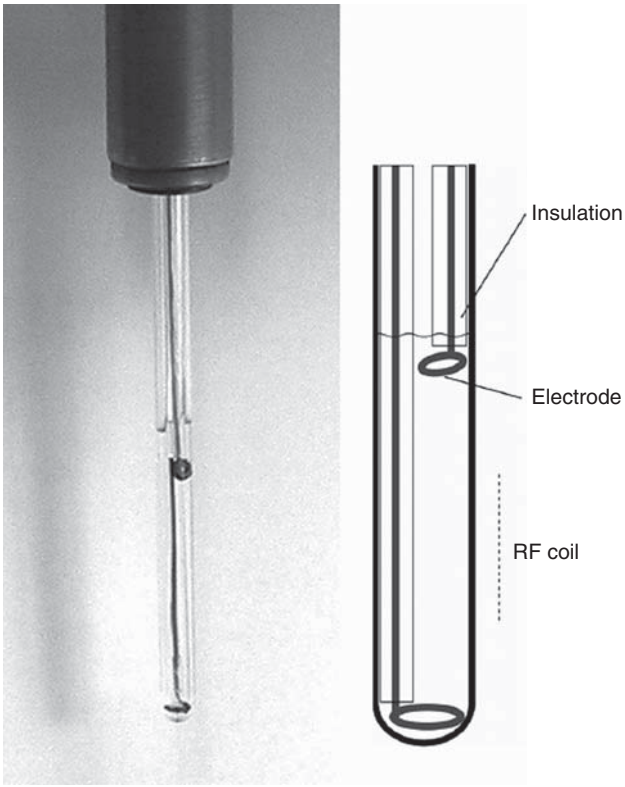


Figure 4 Photograph and schematic of the electrophoretic cell developed by Hallberg et al.[25] Photograph copyright F. Hallberg. The electrophoretic sample cell is based on a conventional 5 mm NMR tube. The distance between the electrodes is roughly 3 cm.

field for ions changes sign on a nanosecond timescale when the electric field is switched, and coherent displacements (natural convection) proceed unchanged, the constant velocity effects are filtered out leaving the electrophoretic phase modulation unaffected. Additional advantages of this sequence include the suppression of electro-osmotic flow and electrolysis.

Recently, Stilbs et al.[25] significantly furthered electrophoretic NMR experiments by moving away from the usual U-tube assembly; a new, much simpler sample cell has been designed consisting of a conventional 5-mm o.d. NMR tube into which two internal electrodes are glued, Figure 4. The improved filling factor of this assembly results in a significant improvement in signal-to-noise, approximately an order of magnitude, when combined with a low-pass filter that suppresses rf noise, Figure 5. This unidirectional sample cell—unlike the U-tube arrangement—permits the sign of the electrophoretic mobility to be determined, when combined with a Carr-Purcell-Meiboom-Gill (CPMG)-like pulse sequence, Figure 6.

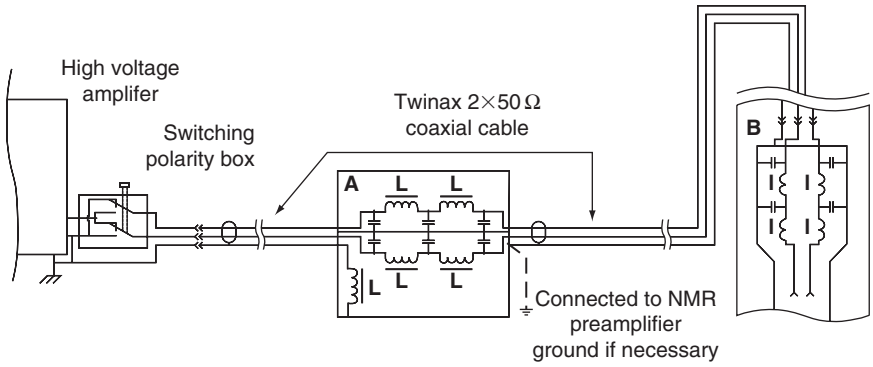


Figure 5 Schematic picture of the low-pass filter combination for suppressing rf noise in the electrophoretic setup. Reprinted from Journal of Magnetic Resonance 1992, Hallberg Furó, Ysuhmanov, Stilbs, Sensitive and Robust Electrophoretic BNR: Instrumentation and Experiments 69–77, Copyright (2008), with permission from Elsevier.

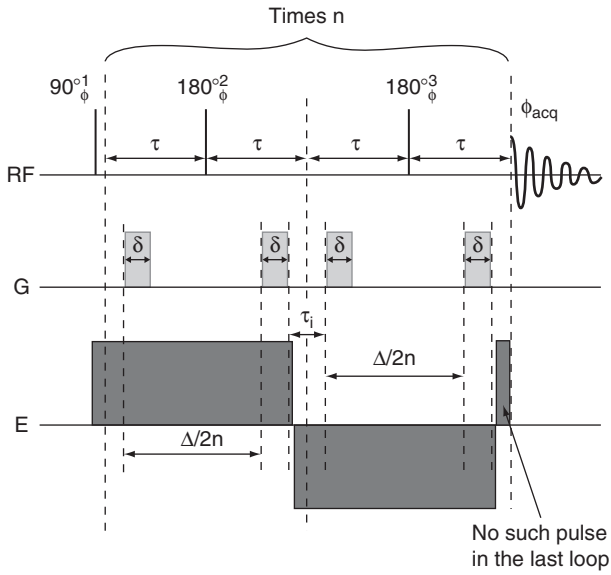


Figure 6 A CPMG-like eNMR pulse sequence. Reprinted from Journal of Magnetic Resonance 1992, Hallberg, Furó, Ysuhmanov, Stilbs, Sensitive and Robust Electrophoretic BNR: Instrumentation and Experiments 69–77, Copyright (2008), with permission from Elsevier.

3. INVESTIGATIONS INTO SMALL AND LARGE MOLECULE MOBILITIES, AND COMPLEX FORMATION

There have been a number of papers that elegantly demonstrate the capability and potential of eNMR. Morris and Johnson, Jr.[9,16] have shown that by combining both eNMR and PGSE-NMR experiments into a 2D experiment called MOSY

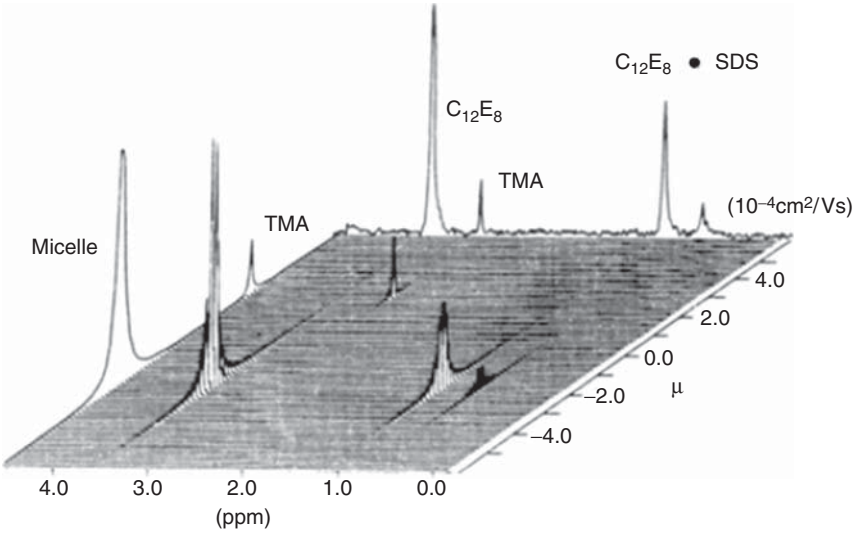


Figure 7 Mobility-ordered 2D NMR spectrum of a sample containing TMA counterions and mixed anionic (SDS)—non-ionic ($C_{12}E_8$) micelles. Electrophoretic NMR by C. S. Johnson, in Encyclopedia of Nuclear Magnetic Resonance, ed. D. M. Grant and R. K. Harris. 1996. Copyright John Wiley & Sons. Reproduced with permission.

(mobility ordered 2D NMR spectroscopy), Figure 7, the electrophoretic mobility and the identity of the ion may be simultaneously determined.

In a similar vein, He et al. demonstrated how the MOSY approach[12,13,26–28] and subsequently, in combination with COSY; that is, EP-COSY[29,30] may be applied to the challenge of resolving protein mixtures, where the 2D COSY spectra of different species within a mixture are separated on the basis of their electrophoretic mobilities (Figure 8).

3.1. Polymeric charge carriers

The study of polyelectrolyte systems by eNMR has been pioneered by the group of Ulrich Scheler who have examined a range of polymeric systems, viz. poly(acrylamide) copolymers,[31] poly(styrene sulphonate) (PSS),[31–34] bovine serum albumin (BSA),[35] mixtures of poly(diallyl dimethyl ammonium chloride) (PDADMAC) and perfluorinated succinic acid (PSA),[36] mixtures of PDADMAC and PSS,[37] rigid rod poly(aramide)s,[38] as well as theoretical considerations.[39]

Böhme and Scheler[32] and Wong and Scheler[31] demonstrated that the charge on the polymer, Z, may be determined by equating the electrostatic force driving the flow of the polymer to its frictional drag:

$$ZeE_{dc} = \frac{k_b T}{D_s} v. \quad (1)$$

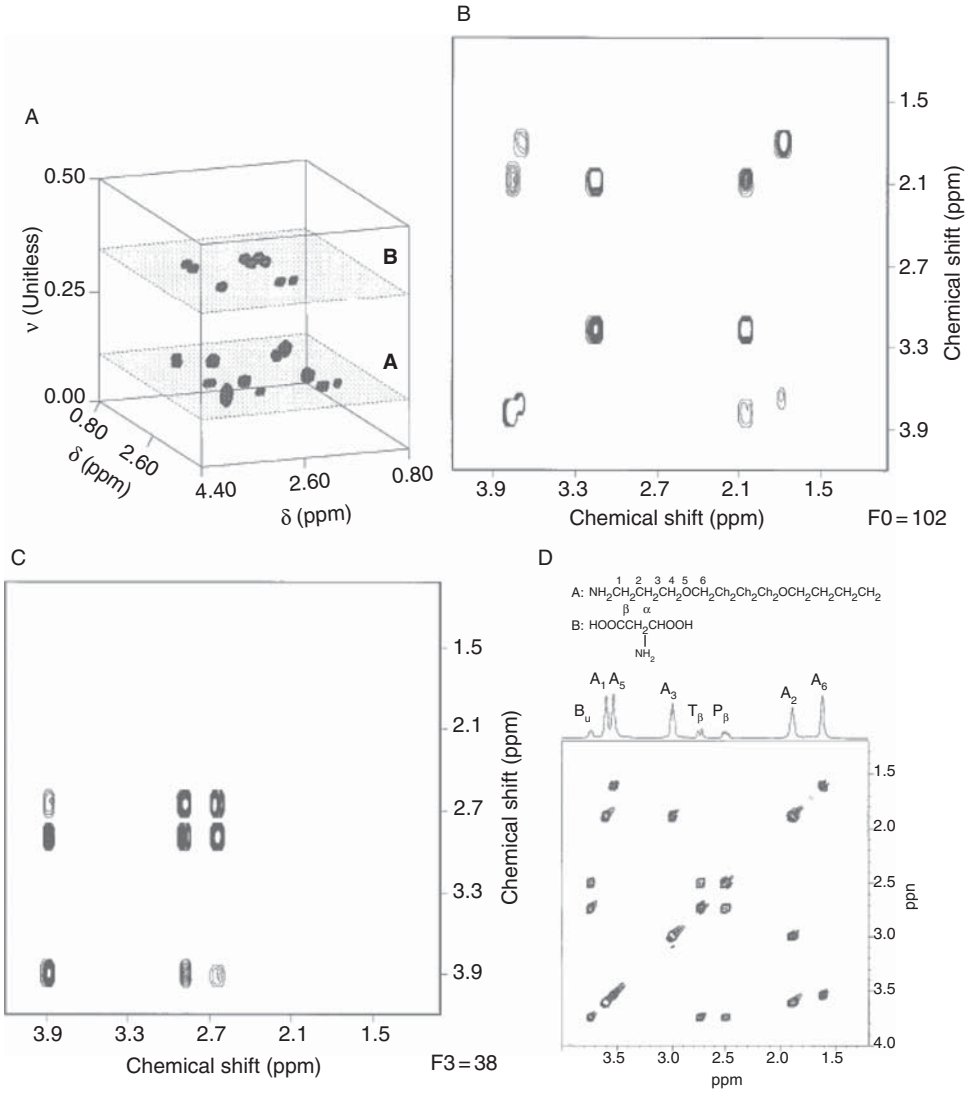


Figure 8 (A) Spectroscopic separation of COSY planes of (A) 4,9-dioxa-1,12-dodecanediamine (148 mM) and (B) L-aspartic acid (150 mM) in a 3D EP-COSY matrix. Using the cosinusoidal oscillation frequencies that were scaled up 100 times, the truncated oscillation curves were extended from 28 to 256 data points before the third Fourier transformation. (Signals were folded in the flow dimension.) Superposition of the component 2D-COSY spectra (B and C) gave the same spectral resonance pattern of a double-quantum-filtered (DQF) COSY experiment (D). The DQF-COSY data was acquired from a mixture solution of L-aspartic acid (100 mM) and 4,9-dioxa-1,12-dodecanediamine (100 mM) in D2O on a Bruker DMX 500 spectrometer. Reprinted from Journal of Magnetic Resonance, 147, He, Lin, Liu and Li, Three-dimensional electrophoretic NMR correlation spectroscopy, 361–365 Copyright (2000), with permission from Elsevier.

where k_b is the Boltzmann constant. Given that D_s—the self-diffusion coefficient—is easily determined;

$$Z = \frac{k_b T \mu}{e D_s}. \tag{2}$$

From this, the ionic strength dependence of the self-diffusion coefficient, the fractal dimension and electrophoretic mobility were determined. For the PSS with molecular weight $M_w = 77{,}000$ g mol^{-1},[32] the self-diffusion coefficient and fractal dimension increased whilst the electrophoretic mobility initially increased but ultimately dropped with increasing ionic strength. The effective charge on the PSS decreased from ~37 to 20 over the salt range $0 < [NaCl] < 0.04$ mol l^{-1}, reflecting the increased condensation of the counterions onto the polymer. A similar approach probed the charge on BSA as a function of pH,[35] Figure 9.

The simultaneous measurement of the electrophoretic mobilities and diffusion coefficients of both components in PDADMAC and PSA mixtures[36] demonstrated that at low PSA concentration, the succinic acid merely displaces the condensed polymer counterions, having little effect on the effective electrophoretic mobility of the polymer. At higher concentrations, the effective mobility of the polymer increases to that of the free acid, Figure 10.

Combining PGSE-NMR and eNMR experiments have proved to be very insightful. Böhme and Scheler[37] also quantified the charge (via the electrophoretic mobility) and hydrodynamic size (via the diffusion coefficient) of the complexes formed between PDADMAC and PSS; the binding of the smaller PDADMAC ($M_w = 5000$ g mol^{-1}) to the larger PSS ($M_w = 77{,}000$ g mol^{-1}) led to complexes that were smaller in size and more weakly charged compared with the PSS, rationalised by a neutralisation of charge reducing the electrostatic expansion of the PSS coil. Subsequently, the same authors[33] quantified the charge and hydrodynamic size of this PSS sample ($M_w = 77{,}000$ g mol^{-1}) as a function of the solution dielectric constant by adding methanol to the aqueous phase. The increased concentration of methanol induced a reduction in the effective charge on the polymer, and a concomitant more compact solution conformation. The molecular weight scaling exponent of the electrophoretic mobility and hydrodynamic radius have been measured[34] for a series of PSS samples spanning a molecular weight range represented by the degree of polymerization (DP), $1 < DP < \sim 500$. A scaling exponent of 1.56 was observed for the hydrodynamic radius vs. molecular weight behaviour, less than the Gaussian coil prediction (2) but greater than that of a stiff linear polymer (1). The electrophoretic mobility increases over the low DP region ($1 < DP < 20$), attaining a constant value for $DP > 20$. The charge on the molecule calculated from the inter-relation of these two quantities show that at low DP the effective charge is comparable to the nominal charge, that is, full dissociation, whereas at high DP, the effective charge is more consistent with Manning's theory of counterion condensation—only a fraction of the nominal charge is accessible, and this fraction decreases with increasing molecular weight, Figure 11.

Hyperbranched polyelectrolytes—poly(ethylene imine) (PEI) ($M_w = 2{,}000$ g mol^{-1} and 25,000 g mol^{-1})—have been examined by Griffiths et al.[40] within

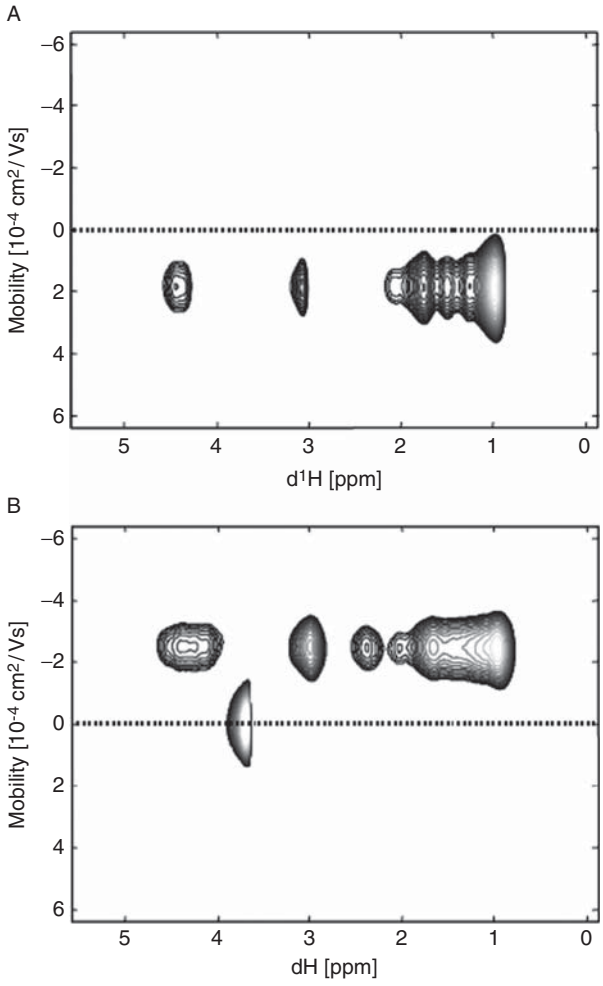


Figure 9 Two-dimensional electrophoresis NMR spectra of BSA at (A) pH = 2.4 and (B) pH = 10.5. Reprinted from Chemical Physics Letters 435, Böhme and Scheler, Effective charge on bovine serum albumin determined by electrophoresis NMR, 342–345, Copyright (2007), with permission from Elsevier.

the same framework adopted by Scheler. The branched globular architecture of PEI results in very little coil expansion over the pH range $4 < pH < 13$, that is, the range over which the fractional charge on the polymer varies from 1 to 0, and therefore lends itself to the study of the molecular weight dependence of the electrophoretic mobility of a branched polymer as a function of pH, without the complicating factors of significant pH induced changes in conformation. There was little change in the self-diffusion coefficient for both polymers over this pH range, but the expected molecular weight dependence was observed. The electrophoretic mobility showed no dependence on molecular weight, but a significant

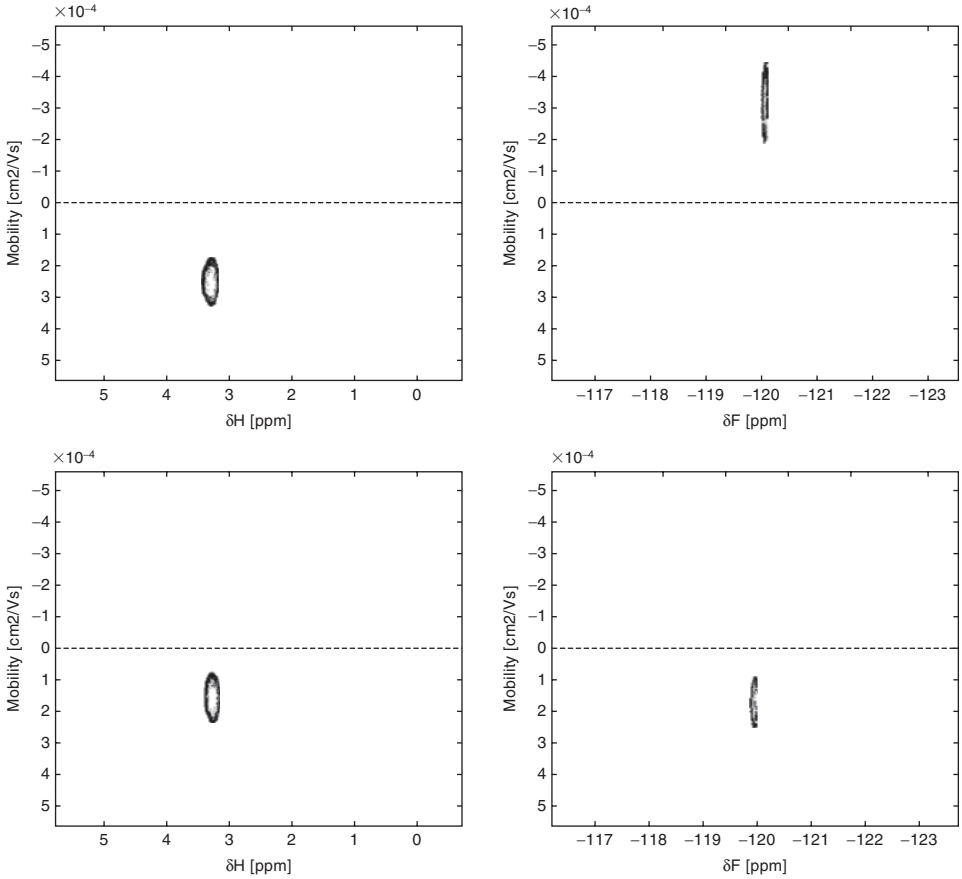


Figure 10 Two-dimensional electrophoresis NMR spectra of 5.2 mmol/l (monomer) PDADMAC (left) and 2.6 mmol/l perfluorinated succinic acid (PSA) (right) for the free components (top) and a mixture of 5.2 mmol/l (monomer) PDADMAC with 1.3 mmol/l , that represents compensation of half the charges of the polycation by the acid (bottom). Böhme and Scheler: Counterion Mobility and Effective Charge of Polyelectrolytes in Solution. Macromol. Symp. 2004, **211**, 87–92. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

step-change was observed around the pK_a of the PEI (p$K_a = \sim 10$). The similarity of the electrophoretic mobility for these two quite different molecular weights may be rationalised by noting that the charge on the polymer scales linearly with the number of ionisable groups, and therefore to a first approximation, molecular weight. Since the self-diffusion coefficient scaled linearly with molecular weight $D_s \propto M_w^{-1}$ and since $z \propto M_w^{+1}$, the electrophoretic mobility, being proportional to the product $z.D_s$ scales as M_w^{0} is independent of molecular weight, as shown in Figure 12. As found by Böhme and Scheler,[34] the nominal charge as perceived by a pH titration was found to be rather different to that determined in the eNMR 'titration'. Correcting the electrophoretic mobility data based on Manning's model

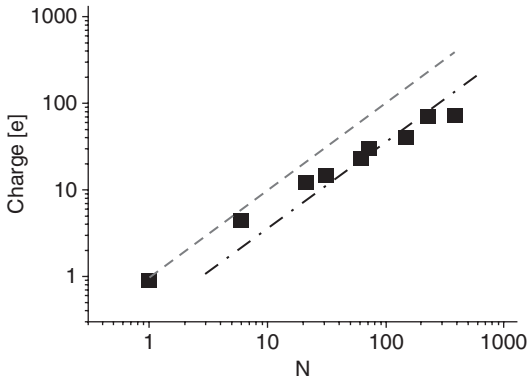


Figure 11 Experimental effective charge of PSS as function of the degree of polymerization. The dashed line represents the nominal charge and the chain dotted line indicates the charge calculated according to Manning's theory of counter ion condensation. Böhme and Scheler. Hydrodynamic Size and Electrophoretic Mobility of Poly(styrene sulphonate) vs. Molecular Weight. Macromolecular Chemistry and Physics 2007, **208**, 2254–2257. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

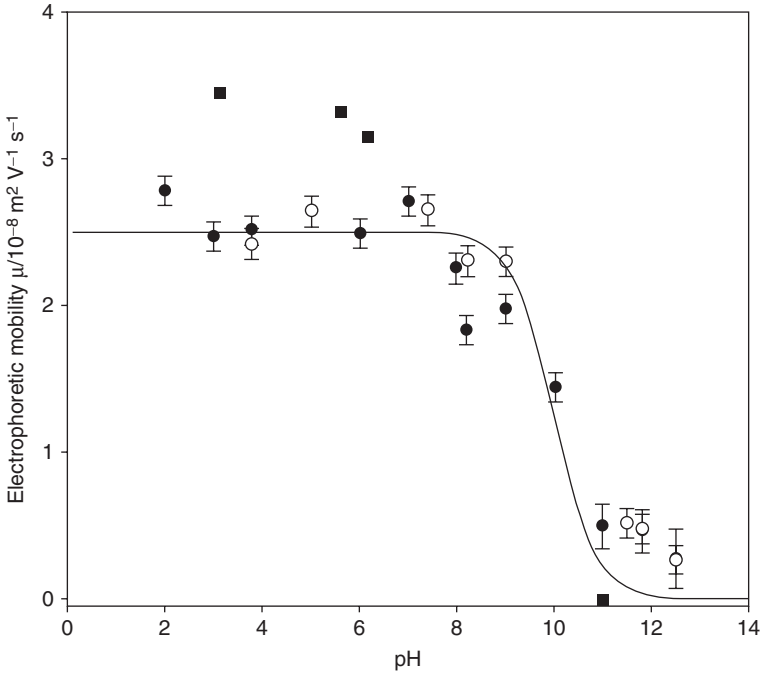


Figure 12 The electrophoretic mobility of two samples of branched poly(ethylene imine) (PEI) as a function of pH; closed symbols $M_w = 2000$ g/mol open symbols $M_w = 25{,}000$ g/mol. Shown for comparison is the predicted charge (line) based on $pK_a \sim 10$. The solid squares are data for PEI of $M_w = 7200$ g mol^{-1} in 0.1N NaCl (Lindquist, G.H.; Stratton, R.A.; *J. Coll. Int. Sci.* **1976**, 55, 45). Reproduced with permission from Griffiths et al., Macromolecules 2005; **38**(8); 3539–3542. Copyright 2005 American Chemical Society.

of counterion condensation does improve the agreement between the two approaches.

3.2. Electrophoretic mobility as a measure of structure

3.2.1. Surfactant self-assembly

At low concentrations, ionic surfactants behave as simple salts—the ionic headgroup and its counterion are largely dissociated. As the surfactant concentration increases through its critical micelle concentration, the surfactants self-associate to form complex structures such as micelles that maintain the hydration of the ionic groups and exclude the hydrophobic moieties from the aqueous phase. The counterion population partitions between micelle surface-associated ions and ions released into the bulk solution. Accurate determination of the degree of counterion dissociation α is rather problematic, and indeed whether α determines or is a consequence of the size and shape of the surfactant self-assembled structure. The charge on an ionic micelle is derived from this distribution of counterions, and maybe determined from a combination of PGSE-NMR and eNMR.

In a series of papers, Griffiths et al. have studied the degree of counterion binding in ionic surfactants in the presence of similarly structured electrolytes[41] and binary mixtures of anionic/non-ionic surfactant solutions,[42–44] particularly interesting systems for eNMR (and PGSE-NMR). A judicious choice of system—surfactant 1, surfactant 2, counterions, etc.—renders it is possible to study the individual mobilities (diffusion and electrophoretic mobility) in these complex mixtures without the need to dilute the sample, making NMR arguably the only technique that can provide this insight.

PGSE-NMR coupled with eNMR was used to study surfactant counterion binding in the presence of identical counterions viz., tetramethylammonium dodecylsulphate (TMADS)/tetramethylammonium chloride (TMACl) system.[41] The large difference in effective size between surfactant in the monomeric and micellised forms, and any associated counterions, enables the diffusion coefficient and electrophoretic mobility of each species to be de-convoluted in terms of the equilibrium between micellised and non-micellised states, Figure 13.

Invoking the common two-state mobility model, the attenuation function of the methylene resonance associated with dodecyl moiety reflects the concentration-weighted average mobility of monomeric and micellised surfactant:

$$\langle M_{(\text{measured})} \rangle = M_{(micelle)}(1 - x_{(\text{monomer})}) + x_{(\text{monomer})} M_{(\text{monomer})}. \tag{3}$$

where M is the mobility parameter (electrophoretic mobility μ or self-diffusion coefficient D_s) and x_{monomer} is a monomer mole fraction, whereas the behaviour of the methyl resonance arising from the TMA^+ ion reflects contributions from both bound and free counterions:

$$\langle M_{(TMA^+)} \rangle = x_{(TMA^+)} M_{(\text{TMACl})} + (1 - x_{(TMA^+)}) M_{(\text{micelle})}. \tag{4}$$

where $x_{(TMA^+)}$ is the mole fraction of TMA^+ counterions in solution, including contributions from monomeric surfactant and those dissociated from the micelles,

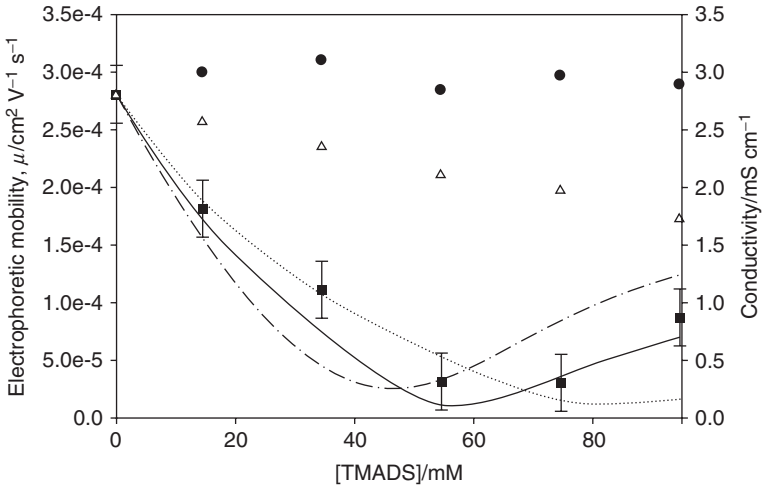


Figure 13 eNMR and conductivity data for TMADS constant Caq series [TMA+]=39.9 mM. Cosine fit to equation TMA+ ions (filled squares) and micelles (filled circles). Also shown are conductivity data (open triangles). Lines are predictions based on various values of α; 0.25 (dash-dot), 0.35 (solid), 0.45 (dotted). Reproduced with permission from Paul et al., J. Phys. Chem. 2005, **109**, 15775. Copyright 2005 American Chemical Society.

$$x_{\mathrm{TMA}^+} = \frac{[\text{monomer}] + \alpha([\text{surfactant}] - [\text{monomer}])}{[\text{surfactant}]}. \tag{5}$$

Hence, a simultaneous fit to eNMR and PGSE-NMR data yield α. This combination of NMR techniques is potentially more informative than other more commonly employed (but simpler) techniques (e.g., conductivity) due to the chemical specificity of NMR.

The electrophoretic mobility has been used to quantify counterion binding in binary mixtures of (anionic) sodium dodecylsulphate (SDS)/(non-ionic) dodecylmalono–*bis*-*N*-methylglucamide (C_{12}BNMG)[42] and SDS/tetra(ethylene oxide) dodecyl ether ($C_{12}EO_4$)[44] systems, both as a function of micelle composition, x_{SDS}, especially when benchmarked against independently determined aggregation numbers, micelle size and shape. Representative data for such mixed micelles are shown in Figure 14, with the *z*-potentials having been determined from the micellar mobility within the framework of the large-particle limit. For the SDS/C_{12}BNMG case,[42] over the mole fraction range $0.2<x_{SDS}<0.8$, these *z*-potentials are greater than a simple linear interpolation between the two limiting compositions, indicating that a significantly higher fraction of the sodium ions must be dissociated from the micelle surface compared with the case of simple SDS micelles. At higher SDS mole fractions, the *z*-potential tends towards the linear interpolation between the two limiting cases, indicating that the sodium counterion dissociation must tend towards the simple SDS case as the surface becomes more SDS-like. The re-adsorption of sodium ions reduces the

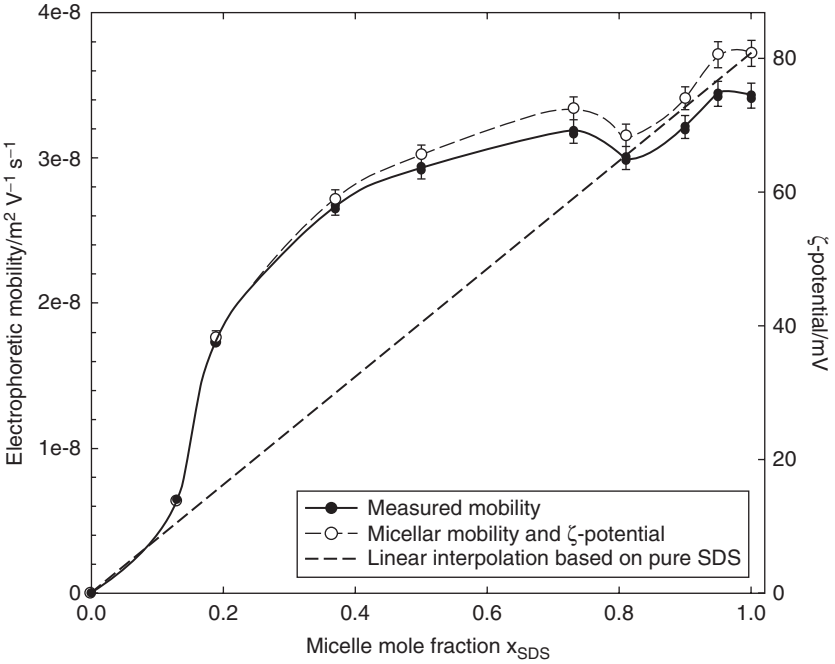


Figure 14 Measured (•) and corrected (○) electrophoretic mobilities (left hand scale) as a function of SDS mole fraction x_{SDS} for 50 mM surfactant solutions. The corresponding ζ-potentials are shown on the right hand scale. Reproduced with permission from Griffiths et al., *Langmuir* 2001, **17**, 7178–7181. Copyright 2001 American Chemical Society.

electrostatic repulsion between the headgroups in order that they can attain the separation required by the volume of the hydrophobic tail. In the case of SDS/$C_{12}E_4$,[44] significant micelle growth occurs with changes in micelle composition, coupled with a reduction in hydration of the surfactant headgroups. After an initial small decrease with decreasing SDS micelle mole fraction, the degree of sodium counterion dissociation (and ζ-potential) subsequently increases reflecting the interplay between the electrostatic character of the micelle surface and the micelle curvature.

3.2.2. Polymer—surfactant complexation

Pettersson et al.,[45] and to a lesser extent, Griffiths et al.[46] have demonstrated how eNMR studies complement the existing portfolio of techniques to study polymer/surfactant mixtures. Pettersson et al.[45] measured the self-diffusion coefficient and electrophoretic mobility of the non-ionic homopolymer poly(ethylene oxide) in the presence of (the anionic surfactant) SDS. The onset of an interaction between polymer and surfactant induced changes in the self-diffusion coefficient and the electrophoretic mobility of both species, Figure 15. Binding of the ionic surfactant to the non-ionic polymer above the critical association concentration resulted in a measurable electrophoretic mobility for the nominally uncharged polymer.

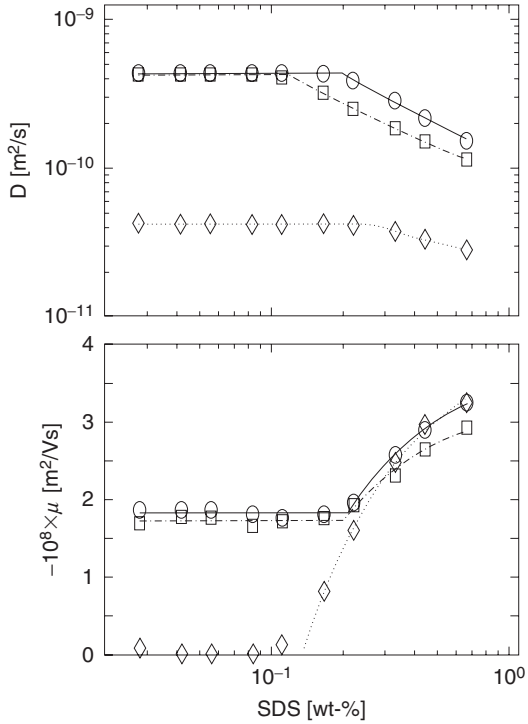


Figure 15 Behaviour of the self-diffusion coefficient (upper panel) and electrophoretic mobility (lower panel) for SDS in the absence (open circles) and presence (open squares) of poly(ethylene oxide) (PEO) and the PEO itself (open diamonds). Reproduced with permission from Stilbs et al., Langmuir 2004, **20**, 1138. Copyright 2004 American Chemical Society.

An increase in the electrophoretic mobility of the surfactant, and reductions in the diffusion coefficients of both surfactant and polymer could be interpreted in a likewise fashion. This study again reinforces the complementarity of eNMR and PGSE-NMR.

4. DETERMINATION OF SMALL MOLECULE AND ION MOBILITY

Electrophoretic NMR has also made in-roads into the study of ion mobility in a range of systems; metal ion—crown ether binding,[47] electrolytic assemblies,[48–55] polymer electrolyte membranes[56,57] and porous systems[14,57,58]. Furthermore, Stilbs et al.[59] have demonstrated a novel approach to quantifying binding in general, exemplified with non-ionic cyclodextrin-ionic surfactant assemblies. This approach is much more sensitive than the traditional analysis of diffusion coefficients where changes in the measured diffusion coefficient associated with any binding between similar-sized molecules may be small. The change in the effective charge may go from zero to a finite value, which is much easier to

quantify accurately in terms of binding than a marginal change in time-averaged self-diffusion upon binding.

Gel phase polyelectrolytes swollen with salt solutions find use in battery electrodes, with the desire that the transport properties of the ions approach that of liquid based systems. However, the transport properties of the ions—quantified via transference numbers—are not easily measured, without the use of a chemically selective detection technique. Dai and Zawodzinski[50] used eNMR to measure the lithium transference numbers in propylene carbonate-swollen poly(vinylidene fluoride)–hexafluoropropylene copolymer gel electrolytes with $CF_3SO_3^-$, $N(CF_3SO_2)_2^-$, $N(C_2F_5SO_2)_2^-$ and $(C(CF_3SO_2)_3^-$. Higher transference numbers were observed for larger anions that showed different salt concentration dependence compared with Li^+ $CF_3SO_3^-$ containing electrolytes. The results were discussed in the context of ion–ion interactions and the degree of anion charge delocalisation.

Walls et al.[54] measured the effects of fumed silica and salt concentration on the ion-transport properties of composite electrolytes comprising oligomeric poly (ethylene glycol) dimethyl ether, hydrophobic fumed silica and Li^+ $CF_3SO_2)_2N$. The fumed silica network had little effect on the ion-transport characteristics, whereas the ionic conductivity exhibited a maximum at ~1.06 M, rationalized in terms of a competing process between adding more charge carriers and a reduced mobility due to increased ion–ion interaction. The lithium transference number derived from PGSE-NMR was consistently greater than that from eNMR. Hayamizu et al.[51] also demonstrated the advantages of using a chemically selective detection method to the study of cation (^{7}Li), anion (^{19}F) and solvent (^{1}H) within a propylene carbonate solution doped with $LiN(SO_2CF_3)_2$.

Heil and Holz[14] and Holz et al.[58] demonstrated how ion diffusion through 3D fluid-filled porous media could be examined by a combination of PGSE-NMR and eNMR. The diffusion and electrophoretic mobility of tetrabutylammonium ($(C_2H_5)_4N^+$) cations through porous Sephadex (chromatographic) beads[14] and a series of cations—$(C_2H_5)_4N^+$; $(C_4H_9)_4N^+$ and $C_2H_5COO^-$)—through micron sized glass beads[58] followed a qualitatively expected dependence on measurement time. At long times, the two mobility parameters provided independent measures of the tortuosity of the systems (provided the effects of salt concentration had been accounted for), with the electrophoretic mobility and diffusion coefficient showing very good agreement with the Einstein relation ($D^{\pm} \propto \mu^{\pm}$), Figure 16.

At short times, where the mobility characteristic showed a pronounced time dependence, it was shown that specific interactions between the ions and the surface of the porous media followed

$$\frac{D(t)}{D_o} = 1 - A\sqrt{\left(D_o(t)\right)} + B\left(D_o(t)\right) + \ldots. \tag{6}$$

where,

$$A = \frac{4}{9}\sqrt{\pi}\left(\frac{S}{V_p}\right),$$

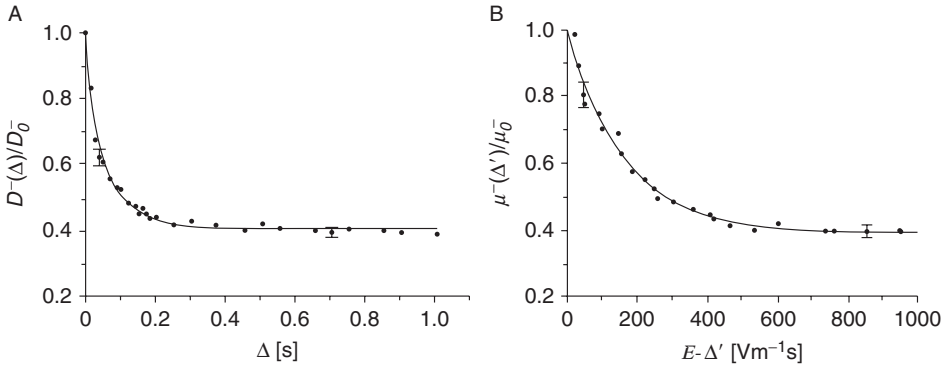


Figure 16 Observation time-dependent relative ionic self-diffusion $D^-(\Delta)/D_o^-$(a) and Δ'E—dependence of the relative ionic mobilities $\mu^-(\Delta')/\mu_o^-$(b)) for the propionate anion in the system 0.2 m C_2H_5COONa/D_2O in Sephadex G25. Reprinted from Magnetic Resonance Imaging 19, Holz, Heil and Schwab, Electrophoretic NMR studies of electrical transport in fluid-filled porous systems, 457–463. Copyright (2001), with permission from Elsevier.

S/V_p being the surface-to-pore volume and

$$\frac{\mu^{\pm}(tE)}{\mu_o^{\pm}} = 1 - A_\mu(\mu_o^{\pm}Et) + B_\mu(\mu_o^{\pm}Et)^2 + \ldots. \tag{7}$$

thereby illustrating how eNMR measurements may be used to probe porous media, Figure 17.

Porous media also comprise the principal material in membranes. Ise et al.[56] studied the mobility (electro-osmotic drag) of protons in polymer electrolyte membranes formed from Nafion© 117 and suplonated polyetherketones (s-PEEKK). Expressed in terms of the electro-osmotic drag coefficients

$$K_{\text{drag}} = \frac{1}{2}\left(\frac{\Delta\Phi_{\text{exp}}}{\Delta\Phi_0} - 1\right)$$

where $\Delta\Phi$ refer to the experimentally observed phase shifts. K_{drag} at a given water content expressed in terms of $n = [H_2O]/[SO_3H]$ was lower for s-PEEKK compared with the Nafion membrane, although under conditions of highest conductivity (n = 20 Nafion; n = 40 s-PEEKK) K_{drag} was comparable. Ise et al. elaborated the hydrodynamic model of Breslan and Miller, correcting for proton transfer between water molecules

$$K_{\text{drag}} = \frac{\Gamma_{\text{hydro}}}{\Gamma_{\text{hydro}} + \Gamma_{\text{trans}}}\left(n_{\text{hydrate}} + (n - n_{\text{hydrate}})\frac{\overline{\upsilon_{H_2O}}}{\upsilon_p}\right). \tag{8}$$

where Γ_{hydro} is the rate of hydrodynamic transport of hydrated protons, Γ_{trans} is the rate for proton transfer, υ_p the drift velocity of the hydrated protons, $\bar{\upsilon}_{H_2O}$ the

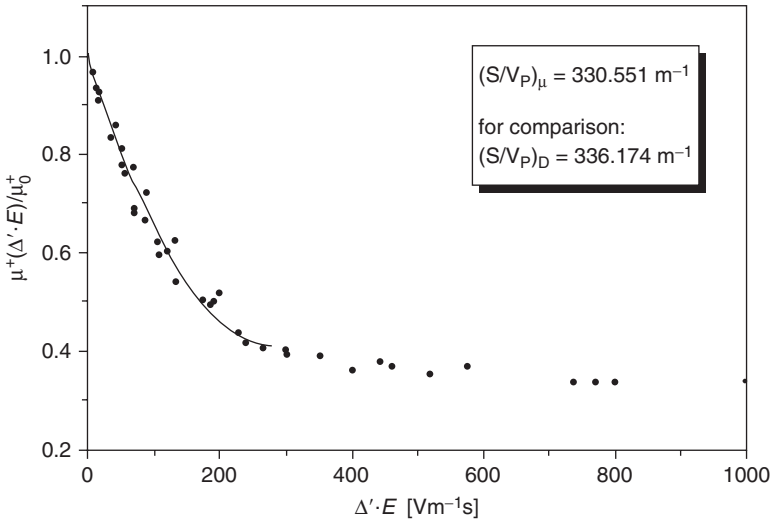


Figure 17 Example for a fit to Eq.(7) to the experimentally obtained △′E—dependence of the relative ionic mobilities $\mu^{-}(\Delta')/\mu_o^-$ of the tetrabutylammonium cation for the determination of the specific surface (S/V_P) of the porous medium. At low △′E values the linear relation can be clearly observed. System: 0.07 m $(C_4H_9)_4NCl/D_2O$ in Sephadex LH 20. Reprinted from Magnetic Resonance Imaging 19, Holz, Heil and Schwab, Electrophoretic NMR studies of electrical transport in fluid-filled porous systems, 457–463. Copyright (2001), with permission from Elsevier.

mean velocity of H_2O outside of the proton hydration spheres. K_{drag} may be further refined in terms of the local interaction

$$K_{\text{drag}} = \frac{1}{R^2} \int_{-R}^{R} K_{\text{drag}}^{\text{local}}(r) r \, \mathrm{d}r. \tag{9}$$

where R corresponds to the channel dimension and r the distance from the wall of polymeric surface.

Summarising the effects of temperature and water content on K_{drag}, two opposing factors were shown to determine the overall behaviour of K_{drag} —at low temperatures, proton transfer dominates the proton flux with little contribution from water flux and an increased water-polymer interaction leads to low values of K_{drag} with decreasing channel size, R. The water content dependence was shown to be most sensitive to the hydrodynamic pumping term, $(n - n_{\text{hydrate}})\upsilon_{H_2O}/\upsilon_p$, Figure 18.

5. OUTLOOK

Clearly, eNMR is maturing into a very useful experimental approach, especially when combined with PGSE-NMR. The insights gained by eNMR—as outlined very briefly here—are unique, especially when applied to complex mixtures or

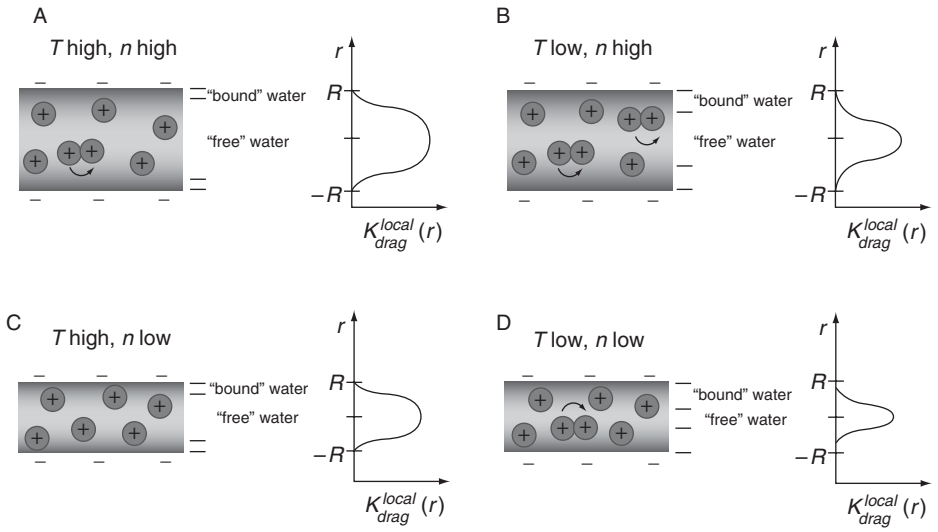


Figure 18 Qualitative description of the influence of temperature and water content on K_{drag}. Reprinted from Solid State Ionics 125, Ise, Kreuer and Maier, Electro-osmotic drag in polymer electrolyte membranes: an electrophoretic NMR study, 213–223. Copyright (1999) with permission from Elsevier.

colloidal systems, thereby harnessing the chemical specificity of NMR. The practical difficulties are being addressed, notably by Stilbs et al. and the expectation is that eNMR measurements will soon become routine.

REFERENCES

1. P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford University Press, Oxford, 1991.
2. W. S. Price, in: *Annual Reports on NMR Spectroscopy* Academic Press, New York, p. 51, 1996.
3. W. S. Price, *Concepts Magn. Reson.*, 1997, **9**(5) 299.
4. W. S. Price, *Concepts Magn. Reson.*, 1998, **10**(4) 197.
5. D. H. Wu, A. Chen and C. S. Johnson, *J. Magn. Reson., Ser A*, 1995, **115**(1) 123.
6. M. Holz, *Chem. Soc. Rev.*, 1994, **23**(3) 165.
7. F. M. Coveney, J. H. Strange and E. G. Smith, *Mol. Phys.*, 1992, **75**(1) 127.
8. F. M. Coveney, J. H. Strange, A. L. Smith and E. G. Smith, *Colloids and Surfaces*, 1989, **36**(2) 193.
9. K. F. Morris and C. S. Johnson, *J. Am. Chem. Soc.*, 1992, **114**(2) 776.
10. T. R. Saarinen and C. S. Johnson, *J. Am. Chem. Soc.*, 1988, **110**(10) 3332.
11. Q. H. He and C. S. Johnson, *J. Magn. Reson.*, 1989, **85**(1) 181.
12. Q. H. He and C. S. Johnson, *J. Magn. Reson.*, 1989, **81**(2) 435.
13. Q. H. He, D. P. Hinton and C. S. Johnson, *J. Magn. Reson.*, 1991, **91**(3) 654.
14. S. R. Heil and M. Holz, *J. Magn. Reson.*, 1998, **135**(1) 17.
15. M. Holz, D. Seiferling and X. A. Mao, *J. Magn. Reson., Ser A*, 1993, **105**(1) 90.
16. K. F. Morris and C. S. Johnson, *J. Magn. Reson., Ser A*, 1993, **101**(1) 67.
17. K. J. Packer, *Mol. Phys.*, 1969, **17**(4) 355.
18. M. Holz, O. Lucas and C. Muller, *J. Magn. Reson.*, 1984, **58**(2) 294.

19. I. Furo, *J. Mol. Liq.*, 2005, **117**(1–3) 117.
20. P. C. Griffiths, A. Paul and N. Hirst, *Chem. Soc. Rev.*, 2006, **35**(2) 134.
21. C. S. Johnson, Jr. and Q. He, in: *Advances in Magnetic Resonance*, Vol. 13, W. S. Warren, ed., *Conference on ''High Resolution NMR in Solids'' January 19–21, 1989*, Xiv+277P. *Academic Press, Inc., San Diego, CA, USA/London, England, UK*, 1989, (Illus. p. 131–160).
22. E. Pettersson, I. Furo and P. Stilbs, *Concepts in Magnetic Resonance Part A*, 2004, **22A**(2) 61.
23. P. Stilbs and I. Furo, *Current Opinion in Colloid & Interface Science*, 2006, **11**(1) 3.
24. B. Manz, P. Stilbs, B. Jönsson, O. Söderman and P. T. Callaghan, *J. Phys. Chem.*, 1995, **99**(29) 11297.
25. F. Hallberg, I. Furo, P. V. Yushmanov and P. Stilbs, *J. Magn. Reson.*, 2008, **192,** 69.
26. Q. H. He, Y. M. Liu and T. Nixon, *J. Am. Chem. Soc.*, 1998, **120**(6) 1341.
27. Q. H. He, Y. M. Liu, H. H. Sun and E. C. Li, *J. Magn. Reson.*, 1999, **141**(2) 355.
28. S. B. Thakur and Q. H. He, *J. Magn. Reson.*, 2006, **183**(1) 32.
29. Q. H. He, W. Lin, Y. M. Liu and E. C. Li, *J. Magn. Reson.*, 2000, **147**(2) 361.
30. E. C. Li and Q. H. He, *J. Magn. Reson.*, 2002, **156**(2) 181.
31. S. Wong and U. Scheler, *Colloids Surf. A-Physicochem. Eng. Asp.*, 2001, **195**(1–3) 253.
32. U. Bohme and U. Scheler, *Colloids Surf. A-Physicochem. Eng. Asp.*, 2003, **222**(1–3) 35.
33. U. Bohme and U. Scheler, *J. Colloid Interface Sci.*, 2007, **309**(2) 231.
34. U. Bohme and U. Scheler, *Macromol. Chem. Phys.*, 2007, **208,** 2254.
35. U. Bohme and U. Scheler, *Chem. Phys. Lett.*, 2007, **435**(4–6) 342.
36. U. Bohme and U. Scheler, *Macromolecular Symposia*, 2004, **211,** 87.
37. U. Bohme and U. Scheler, *J. Phys. Chem. B*, 2007, **111**(29) 8348.
38. U. Bohme, C. Vogel, J. Meier-Haack and U. Scheler, *J. Phys. Chem. B*, 2007, **111**(29) 8344.
39. K. Grass, U. Bohme, U. Scheler, H. Cottet and C. Holm, *Phys. Rev. Lett.*, 2008, **100**(9) 096104.
40. P. C. Griffiths, A. Paul, P. Stilbs and E. Petterson, *Macromolecules*, 2005, **38**(8) 3539.
41. A. Paul, P. C. Griffiths, E. Pettersson, P. Stilbs, B. L. Bales, R. Zana and R. K. Heenan, *J. Phys. Chem. B*, 2005, **109**(33) 15775.
42. P. C. Griffiths, E. Pettersson, P. Stilbs, A. Y. F. Cheung, A. M. Howe and A. R. Pitt, *Langmuir*, 2001, **17**(22) 7178.
43. P. C. Griffiths, A. Paul, P. Stilbs and E. Pettersson, *Langmuir*, 2003, **19**(20) 8605.
44. P. C. Griffiths, A. Y. F. Cheung, C. Farley, A. Paul, R. K. Heenan, S. M. King, E. Pettersson, P. Stilbs and R. Ranganathan, *J. Phys. Chem. B*, 2004, **108**(4) 1351.
45. E. Pettersson, D. Topgaard, P. Stilbs and O. Soderman, *Langmuir*, 2004, **20**(4) 1138.
46. P. C. Griffiths, N. Hirst, A. Paul, S. M. King, R. K. Heenan and R. Farley, *Langmuir*, 2004, **20**(16) 6904.
47. W. S. Price, F. Hallberg and P. Stilbs, *Magn. Reson. Chem.*, 2007, **45**(2) 152.
48. H. L. Dai and T. A. Zawodzinski, *J. Electrochem. Soc.*, 1996, **143**(6) L107.
49. H. L. Dai, S. Sanderson, J. Davey, F. Uribe and T. A. Zawodzinski, *Proceedings of the Symposium on Lithium Polymer Batteries*, 1997, **96**(17) 111.
50. H. L. Dai and T. A. Zawodzinski, *J. Electroanal. Chem.*, 1998, **459**(1) 111.
51. K. Hayamizu, S. Seki, H. Miyashiro and Y. Kobayashi, *J. Phys. Chem. B*, 2006, **110**(45) 22302.
52. H. J. Walls and T. A. Zawodzinski, *Electrochemical and Solid State Letters*, 2000, **3**(7) 321.
53. H. J. Walls, P. S. Fedkiw, S. A. Khan and T. A. Zawodzinski, *Lithium Batteries, Proc.*, 2000, **99**(25) 524.
54. H. J. Walls, P. S. Fedkiw, T. A. Zawodzinski and S. A. Khan, *J. Electrochem. Soc.*, 2003, **150**(3) E165.
55. T. A. Zawodzinski, H. L. Dai, S. Sanderson, J. Davey and F. Uribe, *Proc. Symp. Lithium Polymer Batteries*, 1997, **96**(17) 103.
56. M. Ise, K. D. Kreuer and J. Maier, *Solid State Ionics*, 1999, **125**(1–4) 213.
57. A. Sagidullin, J. Meier-Haack and U. Scheler, *Magn. Reson. Imaging*, 2007, **25**(4) 541.
58. M. Holz, S. R. Heil and I. A. Schwab, *Magn. Reson. Imaging*, 2001, **19**(3–4) 457.
59. F. Hallberg, C. F. Weise, P. V. Yushmanov, E. Petterson, P. Stilbs and I. Furo, *J. Am. Chem. Soc.*, 2008, **130**(24) 7550.

CHAPTER 6

Two-Dimensional Magnetic Resonance Spectroscopy on Biopsy and *In Vivo*

Saadallah Ramadan and **Carolyn E. Mountford**

Contents

Centre for Clinical Spectroscopy, Department of Radiology, Brigham & Women's Hospital, Harvard Medical School, Suite C3-10CC, Boston, MA 02115, USA

Annual Reports on NMR Spectroscopy, Volume 65
ISSN 0066-4103, DOI: 10.1016/S0066-4103(08)00206-8

Abstract Two-dimensional (2D) magnetic resonance spectroscopy (MRS) allows those resonances that are composite in a one-dimensional spectrum, to be examined usually as individual components, in a second frequency. The 2D method has allowed the assignment of molecules that are both diagnostic and prognostic, in the spectra from human cells, biopsies and stool. The acquisition and post-processing parameters required attention due to the very wide range of relaxation and J-coupling values present in these specimens. 2D MRS data can also be collected *in vivo*. Again the acquisition sequences needed to be developed, to facilitate the localization at the lower field strengths available, for *in vivo* studies for brain, breast, prostate, muscle and bone marrow. *In vivo* 2D MRS is in its infancy but promises to be a valuable tool in the future for monitoring disease and response to therapy, particularly using the higher field strength clinical magnets now available.

Key Words: Dimensional MR spectroscopy, Methods, Post processing, Cells, Biopsies, *In vivo*.

1. INTRODUCTION

1.1. History of two-dimensional magnetic resonance spectroscopy

At the Ampere Summer School in 1971, Jean Jeener[1] introduced two-dimensional (2D) Fourier NMR spectroscopy. Jeener proposed a simple two-pulse sequence:

$$90_x - t_1 - 90_x - \mathrm{ACQ},$$

which, after 2D Fourier transformation, produced a 2D magnetic resonance spectrum (MRS) now known as COrrelated SpectroscopY (COSY). The importance of this discovery was realized by Ernst[2] who developed it further and as a consequence received the Nobel Prize for Chemistry some 20 years later.

The COSY methodology was developed to solve the three-dimensional structure of proteins in solution. The method allowed for mainly unambiguous resonance assignments for folded proteins in solution. From these studies, the structures and functions of biological macromolecules could be determined.[3]

1.2. Cultured cells and biopsies

The same vertical bore magnets that are used for protein studies were also used for the development of the technique for study of human tissues. The frequency of the magnet used was mainly kept at 8.5 T (360 MHz) to make the eventual transition to *in vivo* studies less difficult.

The proton (^{1}H) MR spectra from cultured cells, biopsy material and humans *in vivo*, provides detailed information on the complex inter-relationship of the various biochemical pathways. For review, see Mountford et al.[4,5] The pools of chemicals, mobile on the MR time scale, alter with biological status, disease(s) and with ageing.[6] It has now been established that the MRS method is capable of monitoring biochemical information that is diagnostic and prognostic with a wide literature attesting to this. However, in the one-dimensional (1D) MR spectra, many of the resonances overlap forming composite resonances. In order for those resonances that represent changes in biochemical pathways to be unambiguously assigned, 2D MR spectroscopy needed to be employed.

The 2D COSY method was first applied to cultured cells in 1984,[7] healthy tissues in 1988,[8] diseased tissue in 1993 (Figure 3),[9] healthy human brain *in vivo* in 1994[10] and human immune deficiency (HIV)-affected human brain *in vivo* in 2008.[11]

In a COSY spectrum a cross-peak, off the diagonal, indicates scalar coupling between the two protons it connects on the diagonal. This can be seen in Figure 1A where the lipid cross-peaks C, D and G are denoted. For nomenclature, see Structure 1. In the COSY spectra from cells and tissue biopsies (recorded at frequencies of 8.5 T and above), there are up to 60 different chemical species all of which have the potential to alter simultaneously with the onset and development of disease.[12] Fewer cross-peaks are seen from those spectra obtained *in vivo*, due to the lower field strengths employed (1.5 and 3 T) and the inherent difficulties associated with MRS *in vivo*.[13–16]

The COSY method needed modification to be used for cultured cells, *ex vivo* biopsies and *in vivo*. This was because there was limited time available for experimental acquisition due to metabolic degradation of biopsies or patient comfort in the whole body magnet. The time available limits the number of acquisitions in the F_1 domain to less than a quarter of those collected in protein studies. Again in contrast to protein studies, there are some molecules, such as lipid (fats) that are present in such high concentrations in biopsies, which mask other important molecules and need to be filtered. There are two means of filtering such molecules from the resultant spectrum; during the acquisition of the data or by post-acquisitional processing. Both are effective but in the case of the former the data are irretrievably lost.

The development of 2D MRS on cells and biopsies was by need undertaken on cultured cells and diseased tissues that were being biopsied for clinical reasons. Here, the correlation with cell biology and histopathology was relatively straightforward, despite the need for serial-sectioned histopathology which is extremely time consuming. In contrast, the development of *in vivo* 2D MRS was initiated on healthy individuals and then those with a range of diseases. Correlation of *in vivo* 2D MRS relies on the patient going to surgery or biopsy for correlative histopathology.

2. ASSIGNMENTS

A summary of cross-peaks assigned for cells and biopsies are listed in Table 1.

Triglyceride. The initial assignments of lipid and amino acid cross-peaks in the COSY spectra of cultured cells were reported in 1984[7,17] and substantiated by a

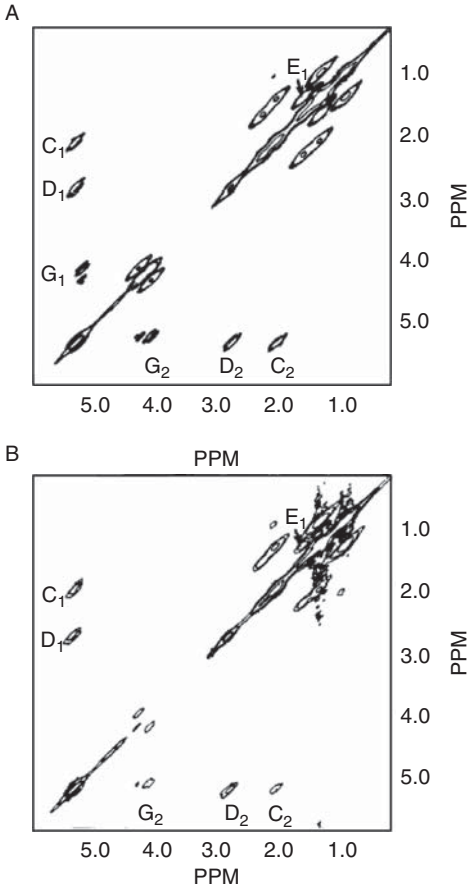


Figure 1 The first 8.5 T COSY of mouse abdominal fat in phosphate-buffered saline in D_2O using (A) a standard COSY pulse sequence $(90_x - t_1 - 90_x)$ and (B) a T2-filtered COSY pulse sequence $(90_\varphi - (\tau - 180_{\varphi-90} - \tau-)_n - t_1 - 90_\psi)$, where n = 240. Data were obtained at 37 °C with the sample spinning and with suppression of residual HOD signals by gated irradiation. A sine-bell window function was applied to the t_1 and t_2 domains. Assignments are as seen in Structure 1.[17] The goal was to modify the COSY pulse sequence to filter out resonances on the basis of their T2 relaxation rate in this case the lipid cross-peaks C, D and G. Taken from Ref. 8 with permission.

series of studies by May et al.[17,18] May et al. also provided chemical evidence for the presence of neutral lipid domains in the plasma membrane of cultured leukaemic lymphoblast. It was considered, following a study led by Bloom,[19] that the ''NMR signal (from cultured cells and biopsies could in part) can be accounted for by isotropically tumbling lipid which is not in diffusive exchange with the more conventionally structured membrane lipid.''

Ether-linked neutral lipids. Further chemical analyses of purified plasma membranes lipid from cultured human leukaemic T cells demonstrated the presence of significant quantities of neutral lipids, 1-*O*-alkyl-2,3-diacyl-*sn*-glycerol, an ether-linked triglyceride. Thus, it was concluded that ''the type of lipid, identified by

Table 1 Assignment of major resonances in 2D proton MR spectra of human cells, tissues and *in vivo*

Molecules	Abbreviation	Coupling	Chemical shift (ppm) F_2–F_1
Lipid fatty acyl chains	A	$(CH_2)_n$–$\mathbf{CH_2}$–$\mathbf{CH_3}$	0.87–1.33
	B	$\mathbf{CH_2}$–$\mathbf{CH_2}$–CH_3	1.33–2.08
	C	–CH_2–$\mathbf{CH_2}$–$\mathbf{CH}$=CH	2.00–5.31
	D	=CH–$\mathbf{CH_2}$–$\mathbf{CH}$=CH	2.82–5.32
	E	–C(O)–CH_2–$\mathbf{CH_2}$–$\mathbf{CH_2}$	1.33–1.62
	F	–O–C(O)–$\mathbf{CH_2}$–$\mathbf{CH_2}$	1.62–2.30
	G (vicinyl)	RO–$\mathbf{CH_2}$–$\mathbf{CH}$(OR)	4.12–5.26
		$\mathbf{CH_2}$–OR	4.26–5.26
	G′ (geminal)	RO–CH$\mathbf{H}'$–CH–OR	4.09–4.29
Cholesterol	Z	Methyl (C26–27) and methine (C25)	0.9–1.5
Amino acids			
Alanine	Ala	$\mathbf{CH}$–$\mathbf{CH_3}$	1.47–3.77
Glutamate/glutamine	Glu/Gln	–CH–$\mathbf{CH_2}$–$\mathbf{CH_2}$–COO–/NH_3	2.14–2.45
Glutamate	Glu	–$\mathbf{CH_2}$–$\mathbf{CH}$(NH_3)(COO)	2.14–3.78
Isoleucine	Ile	–CH–$\mathbf{CH}$(CH_3)–$\mathbf{CH_2}$–CH_3	0.95–2.04
Leucine	Leu	–$\mathbf{CH}$–$\mathbf{CH_3}$	0.95–1.71
Lysine/polyamines	Lys	–$\mathbf{CH_2}$–$\mathbf{CH_2}$–NH_3^+	1.67–3.00
Lysine	Lys	–$\mathbf{CH}$–$\mathbf{CH_2}$	3.75–1.91
Threonine	Thr	–CH–$\mathbf{CH}$(OH)–$\mathbf{CH_3}$	1.31–4.26
Valine	Val	–$\mathbf{CH}$–($\mathbf{CH_3}$)$_2$	1.01–2.28
Phospholipid metabolites			
Choline	Cho	–N^+–$\mathbf{CH_2}$–$\mathbf{CH_2}$–OH	3.51–4.05
Ethanolamine	Eth	H_3N–$\mathbf{CH_2}$–$\mathbf{CH_2}$–OH	3.15–3.87
Glycerophosphocholine	GPC	(CH_3)$_3$–N–$\mathbf{CH_2}$–$\mathbf{CH_2}$–CH_2–OPO_3^-	3.67–4.33
Phosphocholine	PC	(CH_3)$_3$–N–$\mathbf{CH_2}$–$\mathbf{CH_2}$–OPO_3	3.56–4.17
Phosphoethanolamine	PE	H_3N–$\mathbf{CH_2}$–$\mathbf{CH_2}$–OPO_3	3.22–3.98

(*continued*)

Table 1 (*continued*)

Molecules	Abbreviation	Coupling	Chemical shift (ppm) F_2–F_1
Others			
Citrate	Cit	–COO–C**H**A**H**B	2.52–2.66
			2.64–2.78
Inositol	Inos	–C**H**(OH)–C**H**(OH)	3.28–3.64
Spermine/Spermidine	Spd, Spm	H_3N–C$\mathbf{H}_2$–C$\mathbf{H}_2$	2.10–3.10
Lactate	Lac	–C**H**(OH)–C$\mathbf{H}_3$	1.30–4.10
Taurine	Tau	H_3N–C$\mathbf{H}_2$–C$\mathbf{H}_2$	3.25–3.42
Fucose		**H**5–**H**6 couplings	
	Fuc I		1.33–4.27
	Fuc II		1.25–4.28
	Fuc IIb		1.19–4.20
	Fuc III		1.41–4.30
	Fuc IV		1.31–4.38
	α-Fuc		1.22–4.21
	β-Fuc		1.26–3.81
UDP-GlnNAc	Ribose	**H**1′, **H**2	5.99–4.38
	Glucose	**H**1′, **H**2	5.52–3.99
		H2, **H**3	3.99–3.82
		H3′, **H**4	3.82–3.55
		H4′, **H**5	3.55–3.94
		H5′, **H**6	3.94–3.88

chemical analysis and two-dimensional NMR, is triglyceride (and diglyceride) and free or esterified cholesterol.''[18] The use of lipid, particularly unsaturated lipid, as one of the markers of disease would become important for many diseases in the years to come.

Cholesterol and cholesterol ester. The chemical composition of highly purified plasma membrane preparations from a series of biologically characterized Chinese hamster ovary (CHO) cell lines were undertaken to ascertain if neutral lipids, including cholesteryl ester and triacylglycerol, were present. In doing so, it was demonstrated that some cross-peaks in the COSY were from esterified cholesterol and triglyceride and were present in plasma membranes of these cells.[20] The cross-peak connecting resonances 0.9–1.5 ppm from the methyl (C26–27) and methine (C25) protons of the alkyl side chain of the cholesterol-ring system was assigned and is denoted Z (Table 1). It is still not possible to determine from the ^{1}H COSY spectrum if cross-peak Z originated from free or esterified cholesterol, or both. The presence of the cholesterol ring is also indicated by another resonance which is visible on the diagonal at 0.7 ppm. Cross-peak Z, which has previously been assigned to cholesterol C18 in human high-density lipoprotein (HDL),[21] can

have rapid molecular motion even when the remainder of the molecule is relatively restricted.

Amino acids. These are assigned from the existing literature and are included in Table 1.[22]

3. SPECTRAL FILTERING

3.1. Filtering during acquisition

Williams et al.[8] developed a modification of the COSY pulse sequence that filtered out resonances on the basis of their T2 relaxation rate. The replacement of the first 90° pulse in a COSY experiment, with a CPMG sequence, filters out components with T2 relaxation values less than the length of the pulse train. This is followed by an evolution period (t_1) and a terminal 90° pulse (reading pulse), resulting in the sequence

$$90_{\varphi} - (\tau - 180_{\varphi-90}\tau-)_n - t_1 - 90_{\psi} - \mathrm{ACQ}(t_2).$$

The effect of this pulse sequence is seen in Figure 1B where lipid crosses peaks C, D and G (see Structure 1) are reduced in intensity. When applied to cultured malignant cells or biopsies, this technique has helped to identify molecules that were present in small quantities but known to be the origin of a molecule(s) with long T2 relaxation value (around 1.3 ppm) and are associated with tumour spread.[23,24] The technique can also be used to obtain T2 relaxation rates for individual resonances in a broad envelope of lines. The general use of filtering in the acquisition phase is now widely used and there are many pulse sequences devised to address a plethora of biological questions both on biopsy.[25,26]

3.2. Post-processing of 2D spectra from cells, biopsies and *in vivo*

The wide range of molecules, present in cells, tissues and *in vivo*, with equally diverse T1 and T2 relaxation values meant that any one 2D data set needed to be post-processed using different mathematical criteria in four or five different ways to inspect all molecules active on the MR time scale.[27] An example of this can be seen in Figure 2A and B where the post-processing applied to the same COSY raw data set is quite different. The COSY data were obtained from cultured cells. A full data set consists of 512 FIDs of 32 transients each. The spectral differences in Figure 2A and B arise only from the choice of window functions and the number of FIDs used in processing. Schematic representations of the window functions applied are shown beside each spectrum. Figure 2A shows the result of processing 200 FIDs of 2K points using a sine bell in the t_1 domain and a Lorentzian–Gaussian in the t_2 domain (LB = −40, GB = 0.12) and is typical of many ^{1}H COSY spectra of intact cells published. Resonance assignments are as described in May et al.[17,18] with cross-peaks arising from triglyceride denoted A–G and G′ (see Structure 1). Cross-peaks C and D from the olefinic protons on the acyl chains and cross-peak G from the methylene–methine coupling on the glycerol backbone are generally the weakest of the triglyceride cross-peaks.

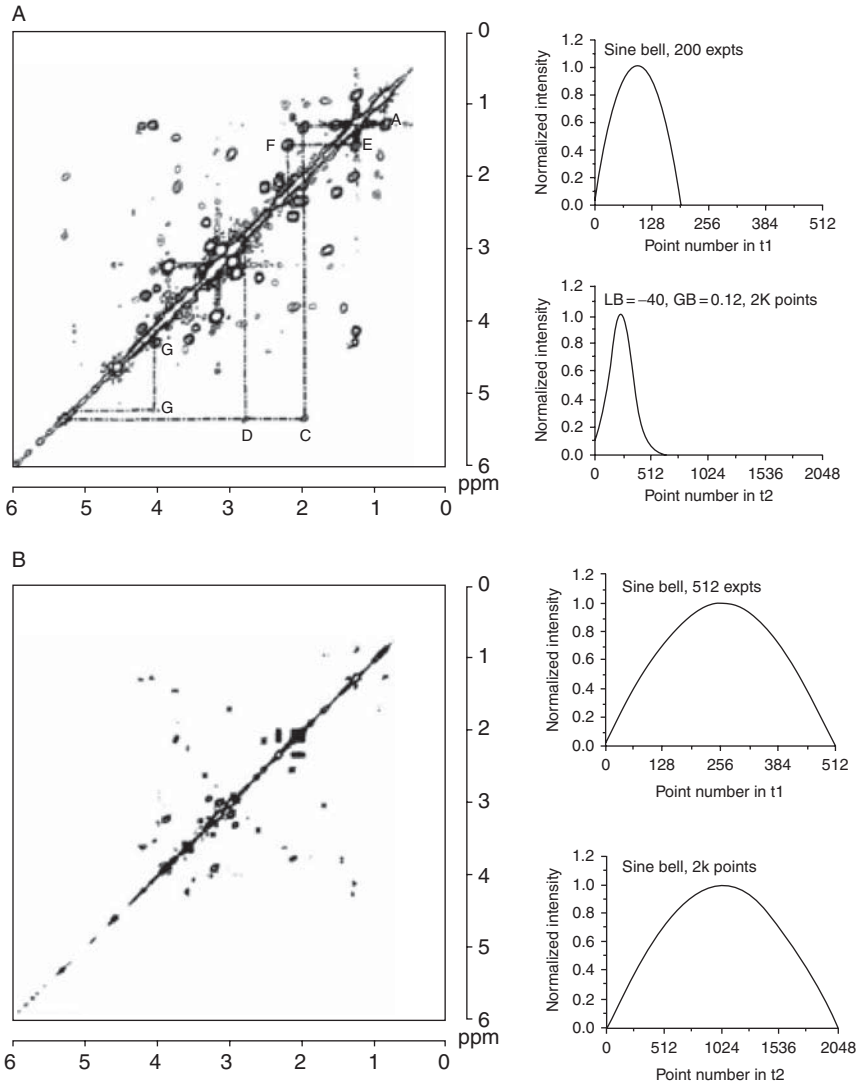


Figure 2 COSY of human MM138 malignant melanoma cells. (A). The 8.5 T magnitude-mode COSY spectrum (1K × 1K symmetrized matrix) was obtained by processing the first 200 2K FIDs with a sine bell in t_1, and a Lorentzian–Gaussian in t_2. LB = −40, GB = 0.12 ($t_1^{\max} = 69\text{ms}$, effective $t_2^{\max} = 89\text{ms}$). Schematic representation of the window functions on the two time domains is shown to the right of the contour plot. The *x*-axes represent complex point pairs in t_1 and alternating quadrature pairs in t_2, but in both cases the corresponding time is 0 to $t_1^{\max}$. Structure 1 shows the structure of a triglyceride molecule with the protons assigned to resonances in the spectrum. (B). The same data set as in (A), but all 512 FIDs (2K points) were processed with a sine bell in both t_1 and t_2 ($t_1^{\max} = t_{2/2} = 178\text{ms}$). Schematic representation of the window functions on the two time domains is shown to the right of the contour plot. Taken from Ref. 27 with permission.

A remarkable change is observed when the complete 512 FIDs of the 2K data set is processed with a sine bell in both domains (Figure 2B). The expected optimal resolution is obtained, but only a small number of cross-peaks with some J-coupling structure appear from amino acids and other small metabolites. Only methyl groups and mobile $-CH_2-$ groups in amino acid side chains are observed, that is, those spins expected to have the longest T2 values. All triglyceride cross-peaks disappear with the exception of the methyl–methylene coupling which could come from the amino acid isoleucine. Thus, triglyceride and amino acid cross-peaks which have relatively small T2 values can be decreased in intensity or even removed entirely by functions which emphasize a later part of the time domain. There are a series of intermediate situations that can be created and for a full description of the various methods of post-acquisitional processing, see Delikatny et al.[27]

4. ANALYSIS OF HUMAN BIOPSIES BY COSY

The first 2D study of human biopsy specimens was undertaken on cervical tissues.[9] The concept that MR spectroscopy could be used as an adjunct histopathology was tested. The tissue type was chosen by the pathologist (Prof. Peter Russell) because it was well established that histopathology could make the distinction between malignant and non-malignant cervical tissue with an accuracy of 99%.[28] The specimen used was a cervical punch biopsy approximately 6 mm^3. The study showed that a high-resolution spectrum (8.5 T) from lipid in 39 of 40 invasive carcinomas. In contrast, the 119 non-malignant biopsies showed little or no lipid spectrum but were characterized by strong unresolved resonances between 3.8 and 4.2 ppm. Resonance ratios of the methylene/methyl and the unresolved methylene resonances allowed accurate distinction between invasive and pre-invasive epithelial malignancy ($p < 0.0001$) with an accuracy of 99%. Since MR spectroscopic examination does not destroy the specimen, the specimens remained intact for further testing and histopathologic analysis.

This is a good illustration of the basis of MRS to monitor disease. The COSY on the left of Figure 3 is from a specimen containing a human papilloma virus infection alone and shows very little chemical activity on the MR time scale. In contrast, the COSY on the right is from a frankly invasive cancer, contains multiple cross-peaks now recognized as an adenocarcinoma spectral pattern. See Table 1 for assignments.

4.1. Loss of cellular differentiation monitored by COSY in colorectal cell model

Once a cell has malignant capabilities and becomes a tumour, a further series of events take place known as *tumour progression*. These include an increase in cellular de-differentiation and anaplasia, infiltration by the tumour into surrounding tissues (invasion) and the establishment of tumours at secondary sites (metastasis). An early observation[24] was the presence of a molecule(s) with a long T2 relaxation, at

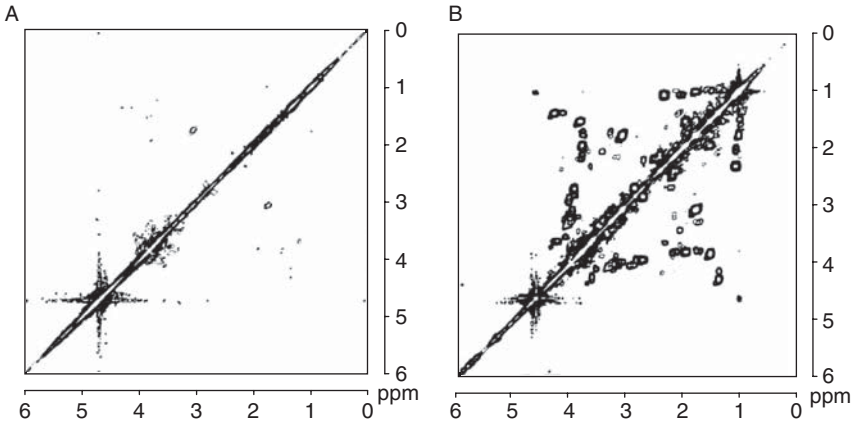


Figure 3 Typical 8.5 T symmetrized COSY from human cervical biopsy specimens is shown. Left is a specimen containing a human papilloma virus infection alone; right contains a frankly invasive carcinoma. Data were processed by zero filling to 1024 data points in the t_1 domain, applying appropriate window functions [sine bell in the t_1 domain and a Lorentzian–Gaussian function (with Lorentzian and Gaussian parameters of -30 Hz and 0.12, respectively) in the t_2 domain], followed by Fourier transformation. The resulting 1024 $\times$ 1024 data matrix was symmetrized with the minimum value algorithm of the spectrometer and plotted as a contour plot, with contour levels increasing as powers of 2. For correlative histological examination the samples were fixed, embedded and processed. Reprinted with permission from Ref. 9.

1.33 ppm in the spectra from tumours with the capacity to spread.[24] This molecule(s) was not recorded in tumours that were unable to form secondary cancers. This resonance(s) thus appeared to be an important diagnostic marker for malignancy and tumour spread. Its assignment and identification took over a decade and was only possible by the use of 2D spectroscopy. The candidates for this assignment were in the spectral region 1.3–4.3 ppm (F_2–F_1). Also since the resonances had a long T2 relaxation value, they remained after a T2-filtered COSY experiment.[8]

A possible candidate for this assignment was covalently linked fucose (see Structure 2). In the COSY spectrum, the H5–H6 cross-peaks from bound fucose are well clear of most others except for the methyl–methine coupling of threonine (1.31–4.26), whereas the remaining fucose cross-peaks may well be overlaid by other resonances.

In the first model, two malignant human colon cell lines of epithelial origin were studied by the COSY method.[29,30] The high tumorigenic line (LIM1215) was compared with one of low tumorigenic capacity (LIM1863). The LIM1215 cell line was capable of some differentiation and highly tumorigenic in immuno-deficient mice (6/6). In contrast, LIM1863 cells derived from a poorly differentiated, ulcerated and invasive colon carcinoma were deemed lowly tumorigenic with only 1/6 tumours formed in immuno-deficient mice and unlike most cultured colorectal cells retained a high degree of functional and morphological organization similar to that of normal colonic mucosa.

The COSY spectrum from the highly tumorigenic and malignant cell line is shown in Figure 4 and the assignments are as in Table 1. Differences were recorded between the highly (Figure 4) and lowly tumorigenic (not shown) cell lines particularly in the F_2 domain (1.00–1.70 ppm) and F_1 domain (3.70–4.50 ppm) regions shown in Figure 5. The methyl–methine coupling region (F_2 = 1.00–1.70 ppm and F_1 = 3.70–4.50 ppm) from highly tumorigenic (LIM1215) cell COSY (Figure 5A) contains cross-peaks from the couplings of lactate (1.33–4.12 ppm) and alanine (1.49–3.79 ppm) and a cross-peak denoted Thr/Fuc I (1.33–4.27 ppm) consistent with both fucose when in a covalent linkage and the amino acid threonine.[29,30] Cross-peaks not previously reported, and tentatively assigned to fucose moieties, of different species or shielding, were denoted Fuc II (1.25–4.28 ppm) and Fuc III

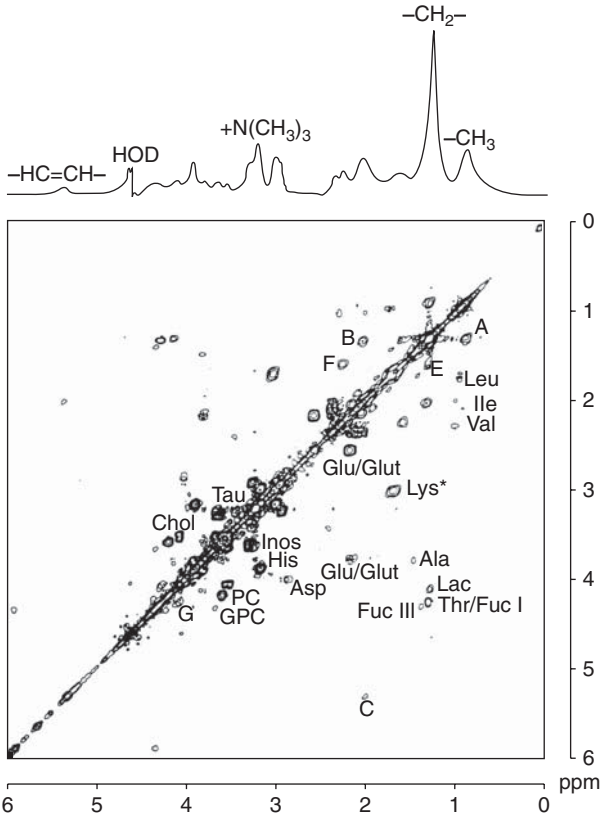


Figure 4 8.5 T proton MR 1D (100 scans) and symmetrized COSY (32 scans, 220 FIDs) spectra LIM1215 (highly tumorigenic) malignant colorectal cells in 400 μl of phosphate-buffered saline in D20. Data were obtained at 37 °C with the sample spinning. A line broadening of 3 Hz was applied to the 1D MR spectrum, and sine-bell and Lorentzian–Gaussian (LB = −30.0, GB = 0.20) window functions were used in t_1 and t_2 domains, respectively, for the COSY spectrum. Contour plots were generated with the lowest level set close to the noise level and subsequent levels increasing in powers of 2. Reprinted with permission from Ref. 29.

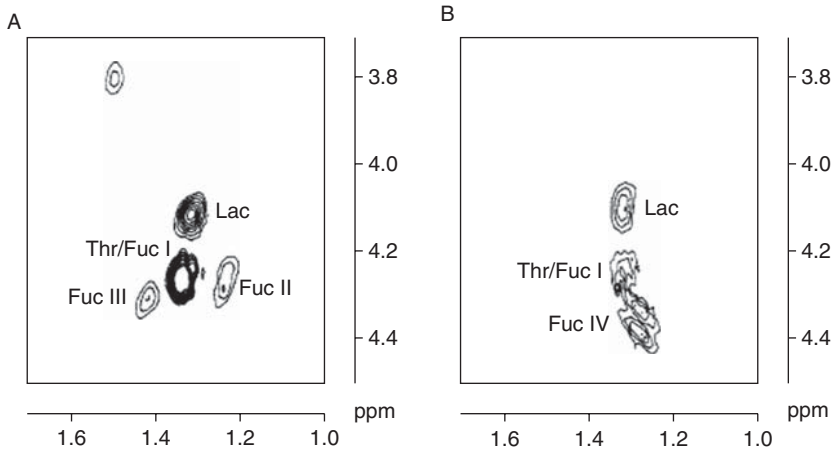


Figure 5 Expanded methyl–methine coupling region (F_2 = 1.00–1.70 ppm and F_1 = 3.70–4.50 ppm) of unsymmetrized COSY data of malignant colorectal cells. Contour plots were generated with the lowest level set close to the noise level and subsequent levels increasing linearly by increments of 1500. (A) Highly tumorigenic (LIM1215) cell line. Data processed using a Lorentzian–Gaussian in t_2 (LB = −30.0, GB = 0.20). (B) Lowly tumorigenic (LIM1863) cell line. Data processed using a Lorentzian–Gaussian in F_2 (LB = −50.0, GB = 0.25). Reprinted with permission from Ref. 29.

(1.41–4.30 ppm). In the COSY from the lowly tumorigenic cell line, no cross-peaks are observed at 1.25–4.28 ppm (Fuc II) or 1.41–4.30 ppm (Fuc III), but a new cross-peak is present at 1.31–4.38 ppm and denoted Fuc IV.

These assignments to covalently linked fucose were confirmed by the addition of free fucose to the same two cell lines. Four methyl–methine cross-peaks were assigned (Table 2) to covalently linked fucose on the basis of increases in volume following the addition of free fucose. Both cell lines incorporated the same amount of exogenous free fucose as determined chemically, but the COSY spectra indicated that the fucose was distributed differently by each cell line. Of the four sites containing MR-visible bound fucose, one was common to both cell lines, two characteristic of the highly tumorigenic line, and the remaining site unique to the lowly tumorigenic cells (Table 2).

4.2. Plasma membrane fragments

Membrane macromolecules and structures, described by others to be shed from the surface of malignant cells without compromise to cell viability,[32] are referred to as plasma membrane fragments (PMF) by Lean et al.[30]

Material released (shed) from the highly tumorigenic cells (LIM1215) in response to increased cell density was also found to be fucosylated whereas shed material from lowly tumorigenic cells (LIM1863) was not. This suggested a biological role for shed fucosylated antigens in tumour aggression (see screening test for colorectal cancer below). The subtle but highly reproducible differences in

Table 2 Colorectal cell models[29,30,37]

Cell line	Tumorigenicity	Differentiation	Fuc I	Fuc II	Fuc IIb	Fuc III	Fuc IV
PC/AA	Non	Poor					
PC/CCC1	Non	Poor					
LIM1863	Low	Poor	X				X
PCJW	Moderate	Moderate/good	X			X	
LIM1215	Strong	Moderate	X	X	X	X	
SW1222	Strong	Moderate	X	X	X	X	
SW480	Strong	Poor	X	X	X	X	
SW620	Strong	Poor	X	X	X	X	

the cell-surface fucosylation of these two cell lines, measured by MRS, had important connotations. If these assignments were correct, they were likely to be part of the cell-surface oligosaccharide antigens. Fucose molecules are incorporated as terminal sugars, internally, or as final entities.[33] This assignment was initially confirmed by acid hydrolysis which led to a 28% reduction in the intensity of the cross-peak and the appearance of the α- and β-forms of fucose.[30] Chemical analysis confirmed the release of free fucose.

Elevated fucosyltransferase activity accompanying malignant transformation is a well-documented phenomenon.[34] A fucosyltransferase is an enzyme that transfers an L-fucose sugar from a GDP–fucose (guanosine diphosphate–fucose) donor substrate to an acceptor substrate. The acceptor substrate can be another sugar such as the transfer of a fucose to a core *N*-acetylglucosamine (GlcNAc) sugar as in the case of *N*-linked glycosylation, or to a protein, as in the case of *O*-linked glycosylation produced by *O*-fucosyltransferase. These enzymes catalyse the transfer reactions chains and as a family are noted for their specificity, exhibiting preference for not only the position of covalent linkage but also the type of acceptor carbohydrate chain and in some cases very particular patterns of fucosylation and sialylation.[35]

4.3. Colorectal cell model identifies role for fucose in adenoma–carcinoma sequence

A second human colorectal cell model was investigated by 2D MRS which included six cell lines with increasing loss of cellular differentiation. The goal was to assign the resonance(s) to a specific biological function(s).[36] The cell lines denoted PC/AA, PC/AAC1, PC/JW, SW1222, SW480, SW620 are in order of decreased cellular differentiation (Table 2). The model comprised two

non-tumorigenic adenoma lines (PC/AA and PC/AAC1) and four carcinoma lines of increasing tumorigenicity (PC/JW, SW1222, SW480 and SW620). A gradual reduction in cellular differentiation and an accumulation of genetic alterations from adenoma to carcinoma characterized the selected cell lines.[37]

Four methyl–methine cross-peaks were again assigned to covalently linked fucose (Fuc I, Fuc II, Fuc III and Fuc IV) on the basis of increases in volume following the addition of free fucose.[37] These results (1) allowed assignment of the four cross-peaks to the protons of covalently linked fucose, (2) confirmed that fucosyltransferases are available for the enhanced fucosylation of vacant sites in response to an elevated concentration of surrounding free fucose in both cell lines and (3) showed from the MR spectra that the incorporated fucose is distributed into at least one common site (Fuc I) but two sites unique to the highly tumorigenic cells (Fuc II and Fuc III) and one site unique to the lowly tumorigenic cells (Fuc IV) (Table 2).

1D and 2D MRS showed that reduced differentiation in the cell model correlated with an increase in the levels of lipid, metabolites, and the glycosylation intermediate uridine diphospho-*N*-acetylglucosamine (Table 1) and cell-surface fucosylation. Mutations involving the K-raps, APC and DCC genes are present both in adenoma- and in carcinoma-derived lines in this model, but the first evidence of an abnormality in the p53 gene was concomitant with the cells ability to grow as a tumour in ethylic nude mice.[36] This genetic change coincided with the detection, by MRS, of UDP–hexose (ribose moiety, 2D MRS cross-peak between H2 at 4.38 ppm and H1 at 5.99 ppm) and the appearance of an additional fucosyl resonance at 1.41–4.30 ppm (Fuc III) in the least tumorigenic of the carcinoma cell lines. An increase in complexity of the fucosylation spectral pattern (Figure 6) was observed with further cellular de-differentiation and increased tumorigenicity. Collectively, these data support the existence of an adenoma–carcinoma sequence. The assignments of both colorectal cell models are listed in Table 2.

It is known that many polyfucosylated antigens exist on the surface of colorectal cells. One clue as to the origin of these fucose molecules is their chemical shift. The H5 protons of Fuc I–Fuc IV in the colorectal cell line and biopsy spectra resonate between 4.27 and 4.38 ppm. This is consistent with the fucose moiety being linked to Gal rather than α-GlcNAc[31]. Of the known antigens characteristic of colorectal cancer, the α-L-Fuc-(1 → 2)-α-D-Gal linkage is common to both Lewis antigen ''b'' and ''y'' (Leb and Ley) cells. This assignment remains speculative.

The proposed biological function(s) of the cell-surface fucosylated molecules suggested so far include immuno-suppression of the host,[34] that is, the mediation of cell-recognition and adhesion-signalling pathways. Listinsky et al.[38] showed evidence that α-L-fucose is important for cell–cell and cell–matrix adhesion in a variety of normal and pathologic processes, particularly neoplasia. It was also suggested that α-L-fucose provides the essential structure that enables carbohydrate ligands to bind to selectins and to carbohydrate counter-ligands and thereby alter cellular homeostasis.

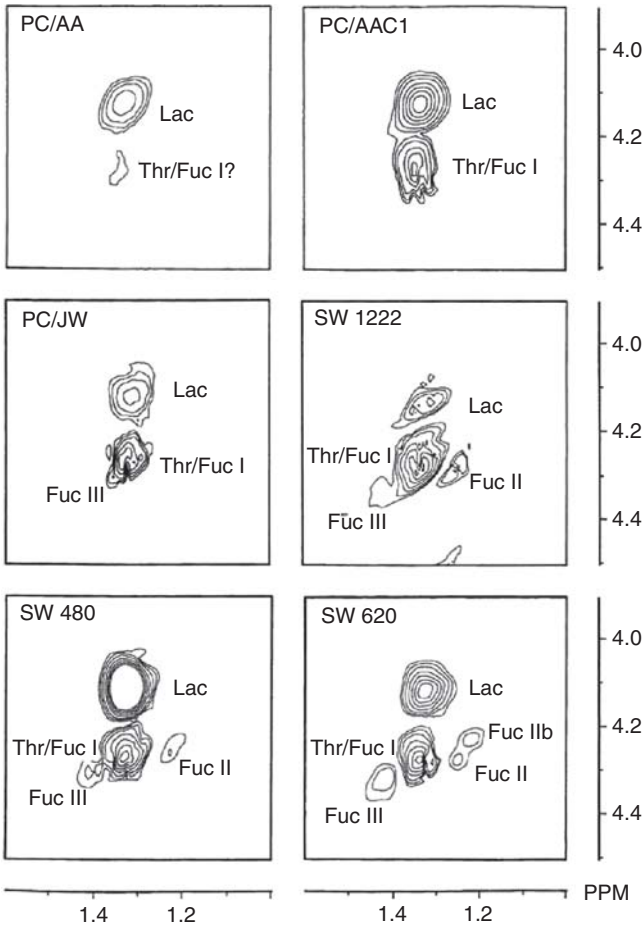


Figure 6 Expansions of the $-CH(OH)-CH_3$ region (F_2 = 1.00–1.60 ppm and F_1 = 3.90–4.50 ppm) from the ^1H MR (8.5 T) symmetrized COSY data from colorectal cells grown *in vitro* PC/AA, PC/AAC1, PC/JW, SW1222, SW480, SW620 are in order of decreased cellular differentiation. The absolute-intensity contour plots were generated with the lowest level = 500 (set close to the noise level); for levels 2 and 3 the increment was 250 and for subsequent levels 500. Taken from Ref. 37 with permission.

5. COLON

5.1. Colorectal biopsies

The differences in the cell-surface fucosylation recorded for the highly and lowly tumorigenic cell lines grown *in vitro* can be observed in excised human biopsies.[29] The presence of cross-peaks Thr/Fuc I, Fuc II and Fuc III in the spectrum of an invasive colorectal primary tumour from a patient staged by the Australian

ClinicoPathological Staging System to have spread to the bowel (not shown) illustrated that with careful choice of data acquisition and processing parameters information on the degree of cell-surface fucosylation may be obtained directly from a biopsy specimen. Furthermore, it was confirmed that the fucose cross-peaks were not observed in spectra of normal colon tissue (not shown), supporting a role for fucosylated species in tumour development and progression.[39]

5.1.1. Screening for colon cancer

Colorectal cancer is one of the most common cancers in the USA and Canada accounting for approximately 146,000 new cases in 1999. The lifetime risk that an individual will develop colorectal cancer is about 5–6% associated with colorectal cancer, including blood in the stool, anaemia, abdominal pain and alteration of bowel habits often become apparent only when the disease has advanced significantly. Prognosis depends largely on the stage of the disease at the time of diagnosis. The 5-year survival for a patient whose colorectal cancer is detected at an early stage is 92%, survival decreases to about 60% in patients with regional spread, and to about 6% in those with distant metastases. It is important to detect the precursor adenomas and cancer as early as possible to increase the chances of successful therapeutic intervention.

As reported above, material released from tumorigenic cells was also fucosylated, suggesting a biological role for shed fucosylated antigens in tumour aggression. This area of glycobiology has now flourished and there is a large literature.[30]

The team at the Institute for Biodiagnostics (National Research Council of Canada) have used the knowledge supplied by the COSY method and developed a screening test for colorectal cancer.[40] The method involves placing stool in the MR tube, and undertaking the COSY method (Figure 7). It can be seen in Figure 7B the stool from a patient with clinically verified colorectal cancer (F_2 = 1.00–1.70 ppm and F_1 = 3.70–4.50 ppm) that there is a recognizable series of fucosylated species apparent that are not recorded in the spectrum from the stool of a healthy person shown in Figure 7A. A large percentage (88%) of patients with colorectal disease gave positive MRS results (I. C. P. Smith and T. Bezabeh, private communication). This screening test for colorectal cancer is made robust by the use of a pattern recognition method to analyse 1D MRS data from the stool.

6. PRE-INVASIVE OVARIAN LESIONS IDENTIFIED BY COSY

Ovarian cancers occur in one out of 57 women and often remains asymptomatic until it has spread (metastasized) with around 70% of women presenting with advanced disease; the 5-year survival for these women is 20%. This poor outcome has created a bias towards aggressive therapy for all women with ovarian disease. However, this assumes that all ovarian tumours are life threatening. Conservative surgical management of early-stage epithelial ovarian cancer[41] is important for a patient. For those patients with benign disease or those with ''low malignant potential'' (LMP), sometimes known as borderline tumours, which may or may not be malignant, a more accurate method of identifying the pathology would be

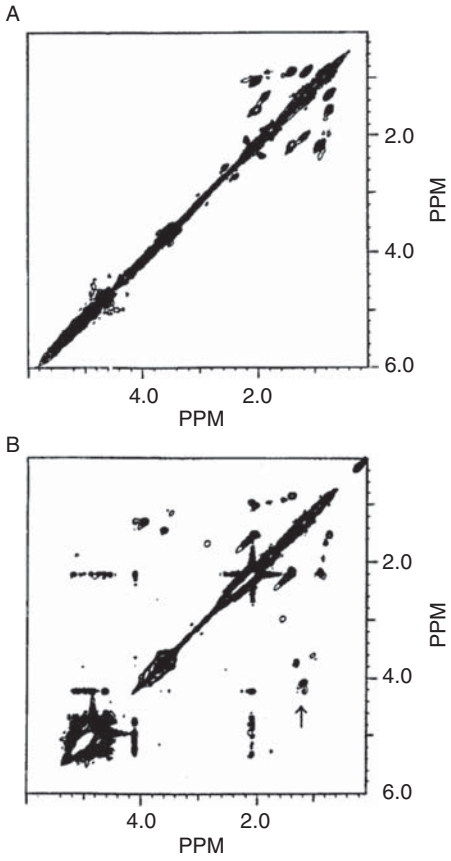


Figure 7 8.5 T MHz symmetrized COSY spectra of human stool from (A) normal subject and (B) colon cancer patient. An arrow in (B) denotes the area of cell-surface fucosylation that present in the spectrum from the stool of a patient with colon cancer. Reprinted with permission from Ref. 40.

important. This would be particularly true if their diagnosis could be established prior to surgery, thus allowing more confident triaging for appropriate therapy.[42,43]

The LMP tumours, which frequently occur in young women, account for 14% of ovarian surface epithelial/stromal neoplasms. These common ''surface epithelial stromal tumours''[41] are thought to have benign and malignant counterparts that are separated by arbitrary histological criteria, which may not correlate with the clinical outcome for the patient. Many women undergo unnecessary surgery and therapy for tumours that have no malignant potential.[41]

In a study of approximately 70 cases, it was shown that MR spectral differences distinguished between subsets of ovarian epithelial/stromal tumours.[44] Typical 1D MR spectra from normal, benign, LMP and malignant ovarian tumours of serous and mucinous types were compared. Using 1D MRS visual inspection

showed that while the malignant serous and mucinous tumours had very similar 1D MR profiles, the benign category showed slight differences. In the LMP category, mucinous tumours had a far more active chemical profile than serous LMP tumours, with the mucin being clearly visible at around 3.8 ppm.[45] The possibility was considered that some mucinous LMP tumours were exhibiting MR evidence of an adenoma–carcinoma sequence, a process that is no longer postulated for serous tumours.[46,47]

Compared in Figure 8 are COSY (8.5 T) from normal ovarian tissue and a poorly differentiated serous carcinoma. The assignments are as listed in Tables 1 and 2. The expanded methyl–methine coupling region (F_1 = 3.9–4.5 ppm and

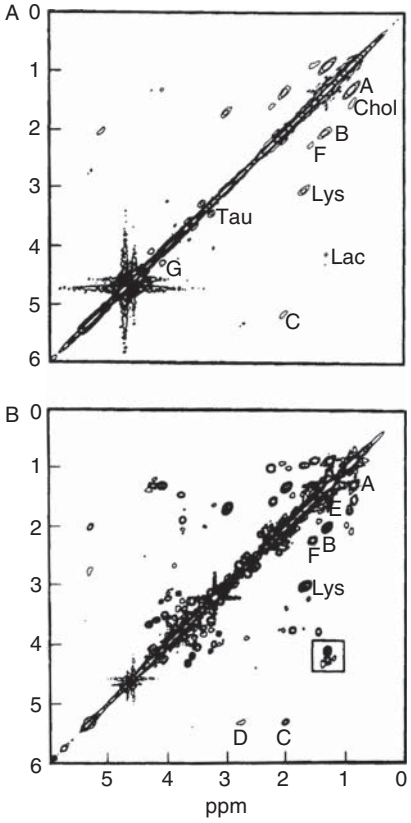


Figure 8 Ovarian tissue biopsies. (A) Normal ovarian tissue. (B) Poorly differentiated serous carcinoma. ^{1}H MR magnitude-mode COSY spectra: 64 accumulations, 200 experiments. Spectra processed using a sine-bell and Lorentzian–Gaussian (LB = −40, GB = 0.12) window function, in the t_1 and t_2 domain, respectively. A–F, from fatty acyl chains; G′, from geminal protons on the glycerol backbone of triglycerides; Chol, cholesterol; Lys, lysine; Lac, lactate. Reprinted from Ref. 45. Characterization of human ovarian epithelial tumours (*ex vivo*) by proton magnetic resonance spectroscopy with permission from Blackwell Publishing.

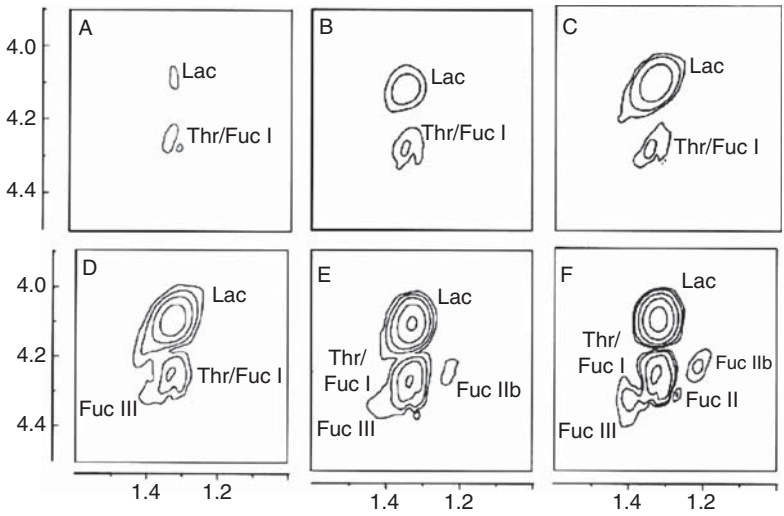


Figure 9 Ovarian tissue biopsies. Expanded methyl–methine coupling region (F_1 = 3.9–4.5 ppm and F_2 = 1.0–1.6 ppm) from the 2D COSY spectra (Figure 15) showing increasing complexity of fucosylation pattern with histologic grade for serous tumours. Contour levels started at 6000 and increased exponentially. (A) Normal ovary, (B) benign tumour, (C) proliferating tumour (LMP), (D) well-differentiated carcinoma, (E) moderately differentiated carcinoma and (F) poorly differentiated carcinoma. Thr/Fuc I, Fuc III, II, IIb: resonances attributable to bound fucose, with a contribution from threonine to Thr/Fuc I. Reprinted from Ref. 45 with permission from Blackwell Publishing.

F_2 = 1.0–1.6 ppm) from the COSY of the following: (A) normal ovary, (B) benign tumour, (C) LMP, (D) well-differentiated carcinoma, (E) moderately differentiated carcinoma and (F) poorly differentiated carcinoma are shown in Figure 9. The Thr/Fuc I, Fuc III, II, IIb: resonances attributable to bound fucose, with a contribution from threonine to Thr/Fuc I, shows increasing complexity of fucosylation pattern with histologic grade. These spectra are remarkably similar to those seen in Figure 6 where the same spectral region is shown for the colorectal cells, grown *in vitro*, showing loss of cellular differentiation. By comparison, the LMP specimen shown is comparable to the PC/JW, that is, poorly tumorigenic with moderate to good differentiation. It remains to be determined if there are subclasses in the LMP category.

7. *IN VIVO* 2D MRS

7.1. History

The first localized *in vivo* human brain 2D COSY experiment was collected in 1994, with an acquisition time of 102 min and region of interest (ROI) of 240 cm^3 [10] on a 2 T whole body magnet. A stimulated-echo volume-selective spectroscopy

(VOSY) sequence was used to generate three-dimensionally localized signal, which was followed by an incremental delay (t_1) and a hard 90° RF pulse for coherence transfer (CT), sandwiched between two equivalent gradient pulses. The signal-to-noise ratio (SNR) was low, but it proved that the COSY method could be undertaken *in vivo*.

In 1995, Ryner et al.[15] implemented a J point-resolved spectroscopy (JPRESS) in the white matter of human brain in 27 min (ROI 27 cm^3), using a ''90°–180°–t_1–180°–t_1–ACQ'' pulse sequence, where t_1 is the incremental delay, and all RF pulses are spatially selective Shinnar-Le Roux pulses.[48] The same authors[16] followed with the application of more complicated 2D localized MRS experiments including 2D zero- and double-quantum, zero-quantum-filtered COSY and spin-echo correlated spectroscopy (SECSY), double-quantum-filtered (DQF)-COSY and -SECSY. In these experiments, the excitation, slice selection and coherence transfer were all undertaken with spatially selective, shaped RF pulses in the presence of slice-selective gradients, which explains the minimum number of three RF pulses in the above pulse sequences. Ziegler et al.[13,14] presented basics for optimizing experimental and theoretical conditions pertaining to *in vivo* and ex vivo 2D MRS.

In 2008, at the time of writing, most ''localized'' COSY (L-COSY) are acquired with three slice-selective RF pulses, 90°–180°–t_1–90°–ACQ (t_2) as is explained in detail in Thomas et al.[49]

7.2. Hardware

The majority of current *in vivo* 2D data are collected on commercially available clinical MR scanners at a field strength of 1.5 or 3 T. L-COSY spectra also are now able to be collected at 7 T using a non-commercial coil (S. Ramadan, L. Wald, E.-M. Ratai and C. Mountford, unpublished data). There are more than one major MR manufactures and each has a spectroscopy package available. However, there is currently no commercial 2D software available for *in vivo* data processing.

The transmit-receive coils used for data acquisition depend on the organ in question. For the brain, a multi-channel coil can be used but a single-channel coil is usually better to improve SNR. There are commercially available breast coils that can be used for both imaging, and 1D and 2D MRS. For the prostate there are suitable, commercially available single-channel endorectal coils. For muscle and bone marrow, a single-channel loop coil or knee coil can be used.

7.3. Pulse sequences

In order for 2D to be applied *in vivo*, the RF pulses must be spatially selective. A pulse sequence needs to be applied that minimizes the number of pulses to create the required signal. Efficient pulse sequences use the minimum number of RF pulses to select a particular coherence transfer pathway.

The L-COSY sequence uses a minimum of three RF pulses (Figure 10), where the first two RF pulses generate a spin-echo, and then the third RF, a 90° RF pulse, acts as a spatially selective and coherence transfer pulse. Note that due to the gradient-assisted method of coherence selection that is used, only echo or

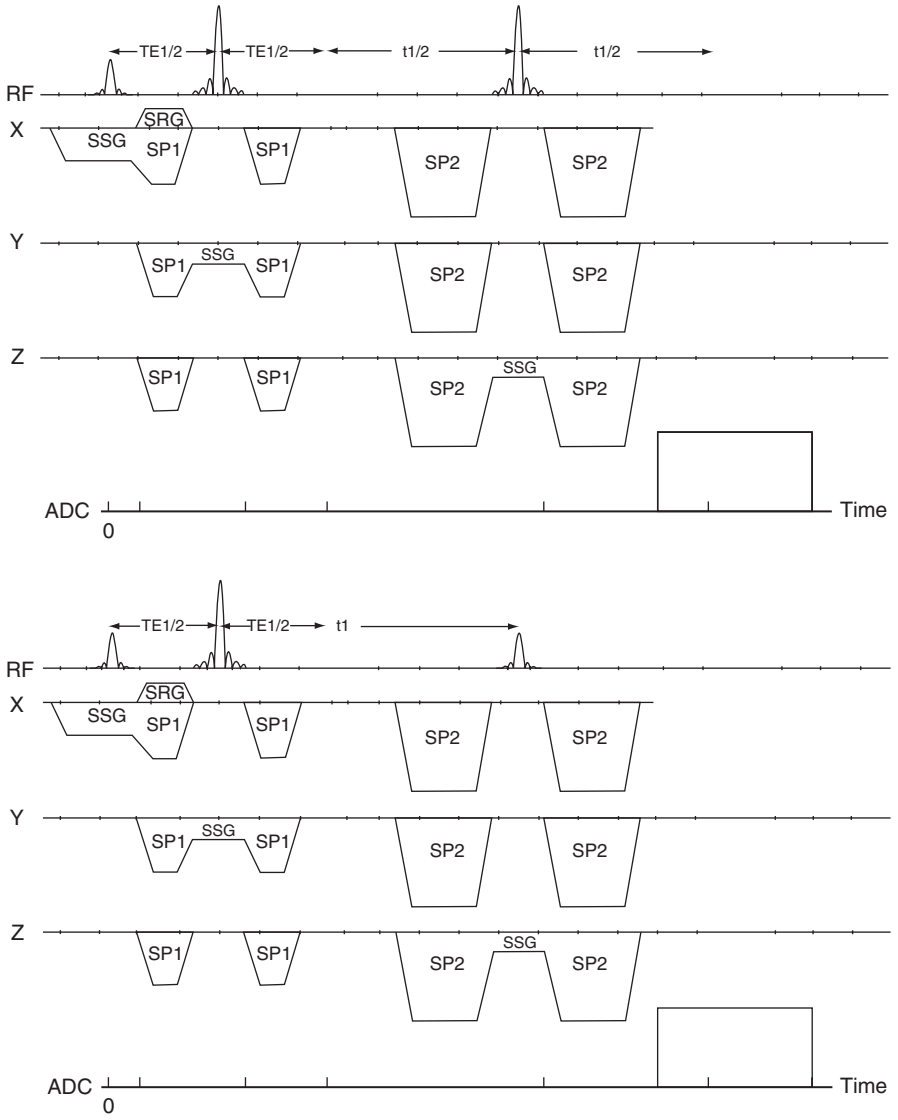


Figure 10 Top: JPRESS localized pulse sequence (90°–180°–180°). "SSG" are slice-selective gradients, "SP1" and "SP2" are spoiler gradient pulses, "SRG" is a slice refocusing gradient. "TE1" is constant and "t_1" is the incremented delay. Bottom: L-COSY localized pulse sequence (90°–180°–90°) with only one incremented delay, t_1, between the second and the third RF pulses.

anti-echo coherence type can be acquired.[26] Pulse sequences shown in Figure 10 select echo-type signal only.

In a different 2D pulse sequenced known as JPRESS, a minimum of three slice-selective RF pulses, applied in the presence of slice-selective gradients along each of the three orthogonal axes, are used to generate a double-refocused echo

(Figure 10). RF shapes can be sinc or Shinnar-Le Roux or any slice-selective pulse with suitable spectral bandwidth. Even though the sequence localizes in a single-shot, a phase cycle that eliminates all signal except double-refocused echo is recommended.[50] A typical t_1 increment in JPRESS spectrum is 10 ms, which causes signal loss due to T2 relaxation as the number of increments increase.

7.4. Processing of *in vivo* 2D MRS data

The raw matrix is usually linear predicted to double its size in F_1. This is followed by zero filling to double the size of the linear-predicted data. Apodization with phase shifted sine-bell squared functions along t_1 and t_2 is then applied before Fourier transform. The software package, known as *Felix*,[51] is used by many workers in this field, and is equipped with suitable processing functionalities. L-COSY spectra are usually asymmetrical about the diagonal due to the implementation of water suppression, which partially suppresses resonances close to water frequency. This in turn makes the process of coherence transfer between two coupled spins non-equivalent.[52] In the absence of water suppression, L-COSY spectra are symmetrical.

8. HUMAN MUSCLE

Localized MRS was applied successfully to human skeletal muscle *in vivo* to detect the two lipid pools referred to as intra-myocellular lipids (IMCL) and extra-myocellular lipids (EMCL).[53–56] It was proposed by Schick[54] that the resonances from the lipids in the muscle spectra are seen as two signals due to the geometrical arrangement and anisotropic susceptibility of these lipid compartments. Velan et al.[55,56] implemented L-COSY in 10 volunteers (Figure 11) and was able to determine IMCL/Cr and EMCL/Cr as 10.2 ± 1.9 and 15.4 ± 2.9, respectively. The measure of unsaturation was achieved by measuring the ratio of the C2/C1, where C2 and C1 are volume of cross-peaks due to diallylic and allylic protons, respectively. The measure of unsaturation for IMCL and EMCL were 1.1 ± 0.11 and 0.87 ± 0.12, respectively. Thomas et al.[57] used a STEAM-like pulse sequence, localized exchange spectroscopy (L-EXSY: 90°–t_1–90°–t_m–90°–ACQ, with all RF pulses being slice selective) with a long mixing time (t_m) to detect proton exchange between different molecular species or fragments of the same molecule. The detection of exchange in a 2D MRS approach can be grasped by considering the following. Assume that H_A and H_B are two exchanging protons that resonate at δ_A and δ_B, then during t_m, H_A and H_B will exchange, leading to the generation of cross-peaks at ($F_1 = \delta_A$, $F_2 = \delta_B$) and ($F_1 = \delta_B$, $F_2 = \delta_A$) after double-Fourier transform. Species that did not exchange during t_m appear on the diagonal at $F_1 = F_2$. Short t_m values (<100 ms) did not show exchange cross-peaks. At a t_m of 300 ms, two types of exchange cross-peaks were identified: (1) between creatine

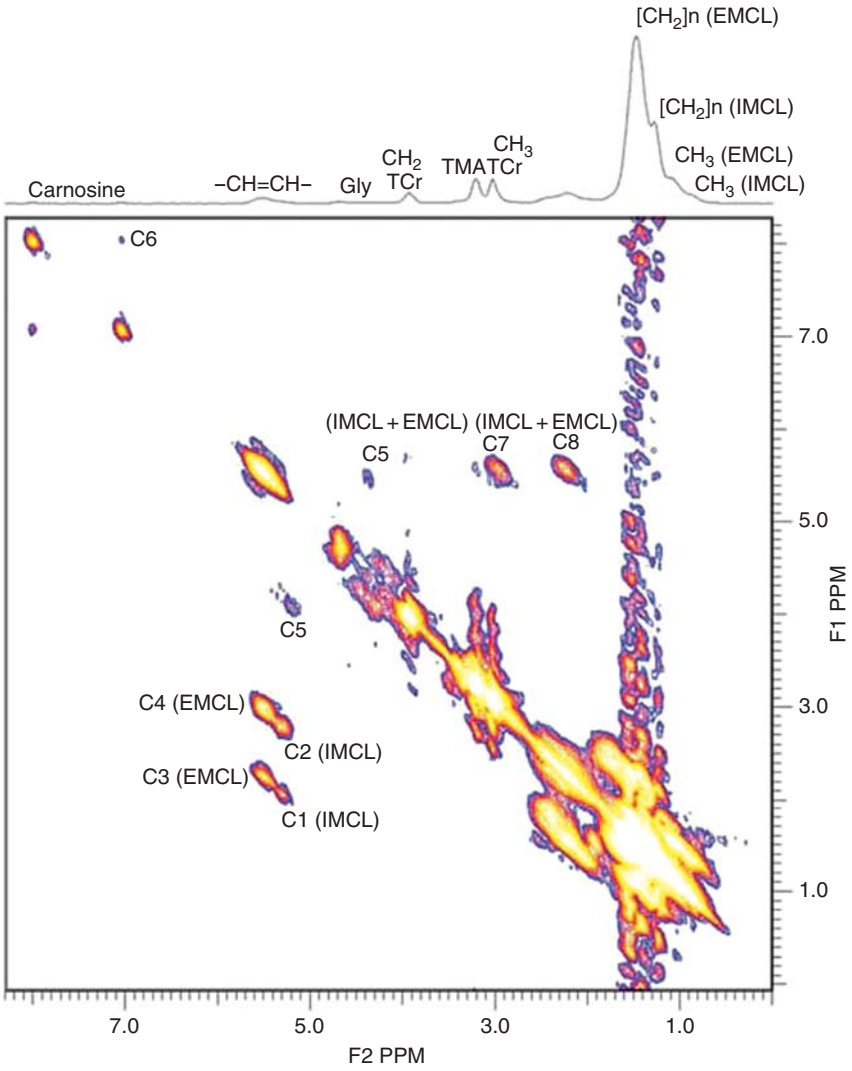


Figure 11 L-COSY spectrum of a healthy muscle with spectral assignments. C1, C3 and C8 are cross-peaks due to coupling between olefinic and allylic methylene protons of IMCL and EMCL, respectively. The cross-peaks C2, C4 and C7 are cross-peaks due to coupling between olefinic and diallylic methylene protons of IMCL and EMCL, respectively. C5 cross-peaks are due to the correlation between CH_2 and CH groups of the glycerol backbone protons. C6 are cross-peaks due to partially oriented imidazole protons of carnosine. TR = 2 s, $TE_{initial}$ = 30 ms, 50 increments in F_1, spectral width if F_1 = 1250 Hz, spectral with in F_2 = 5000 Hz, t_1 increment = 0.8 ms, 2048 points in F_2, averages per increment = 8, total acquisition time = 13 min. Linear prediction to 100 t_1 values, and zero filling to 256 points was done before Fourier transform (reproduced from Ref. 55 with permission).

methyl protons and water and (2) between olefinic and $(CH_2)_n$ of lipids molecule. These cross-peaks are probably due to dipolar coupling between these species. Exchange rates were not quantified. Experimental parameters were voxel size = 27 cm^3, TR = 2 s, $TE_{initial}$ = 22 ms, t_m = 7.5 or 300 ms, t_1 = 1.6 ms, spectral width = 2500 Hz in F_2, 1024 complex points along F_2 and 64 t_1 increments, and excitations per increment = 8.

To verify the identity of the additional peaks between 2.5 and 4.0 ppm of a 1D proton spectrum of muscle, more complex 2D pulse sequences were required to correctly assign them. 2D JPRESS, zero-quantum- and Zeeman-order filtering, double-quantum filtering, 2D-constant-time COSY and dipolar-order filtering were implemented on a 1.5 T whole body scanner to assign some unassigned, orientation-dependent resonance peaks in localized ^1H calf-muscle spectra.[58] It was found that residual dipolar coupling does not vanish in the anisotropic muscle structure, as reported in isotropic solutions. Many dipolar multiplets were identified between 2.7 and 4.5 ppm of the muscle spectrum. Thus, each of the methyl and methylene groups of Cr, which are usually detected as singlets, are found as doublets due to residual dipolar coupling.

Thus, 2D spectroscopy method offers a new opportunity for inspection of lipid alterations in muscle associated with a range of diseases. One such example is the effect of the cholesterol-lowering drugs, statins, that commonly cause muscle pain or weakness.[59]

9. HUMAN BONE MARROW

Many 1D spectroscopic studies reported that marrow lipid metabolism might provide important diagnostic information for tumours.[60–62] Single-voxel L-COSY and DQF-COSY were implemented on a 1.5 T whole body magnet in healthy and diseased tibial bone marrow.[63] Six healthy and six patients with acute leukaemia were studied using this technique. An unsaturated lipid proton index was obtained from the DQF-COSY by dividing the sum of all unsaturated cross-peaks in L-COSY by the $(CH_2)_n$ diagonal peak, but was unable to discriminate between healthy marrow and marrow of a leukaemic patient. Increased water content was detected in leukaemia patients. Representative 2D MRS spectra from healthy person, patient with leukemia and a patient who had leukemia but in remission are shown in Figure 12. The method as presented failed to make a distinction based on altered chemical composition. It remains to be seen if diagnostic information is available at the higher field strengths of 3, 4 and 7 T.

10. PROSTATE

Due to the location of prostate gland in the body, MRS studies are carried out with an endorectal coil as receiver and body coil as RF transmitter to optimize SNR.[64] Metabolites such as polyamines (spermidine and spermine), citrate and creatine are prominent markers in healthy prostate whereas these are reduced in

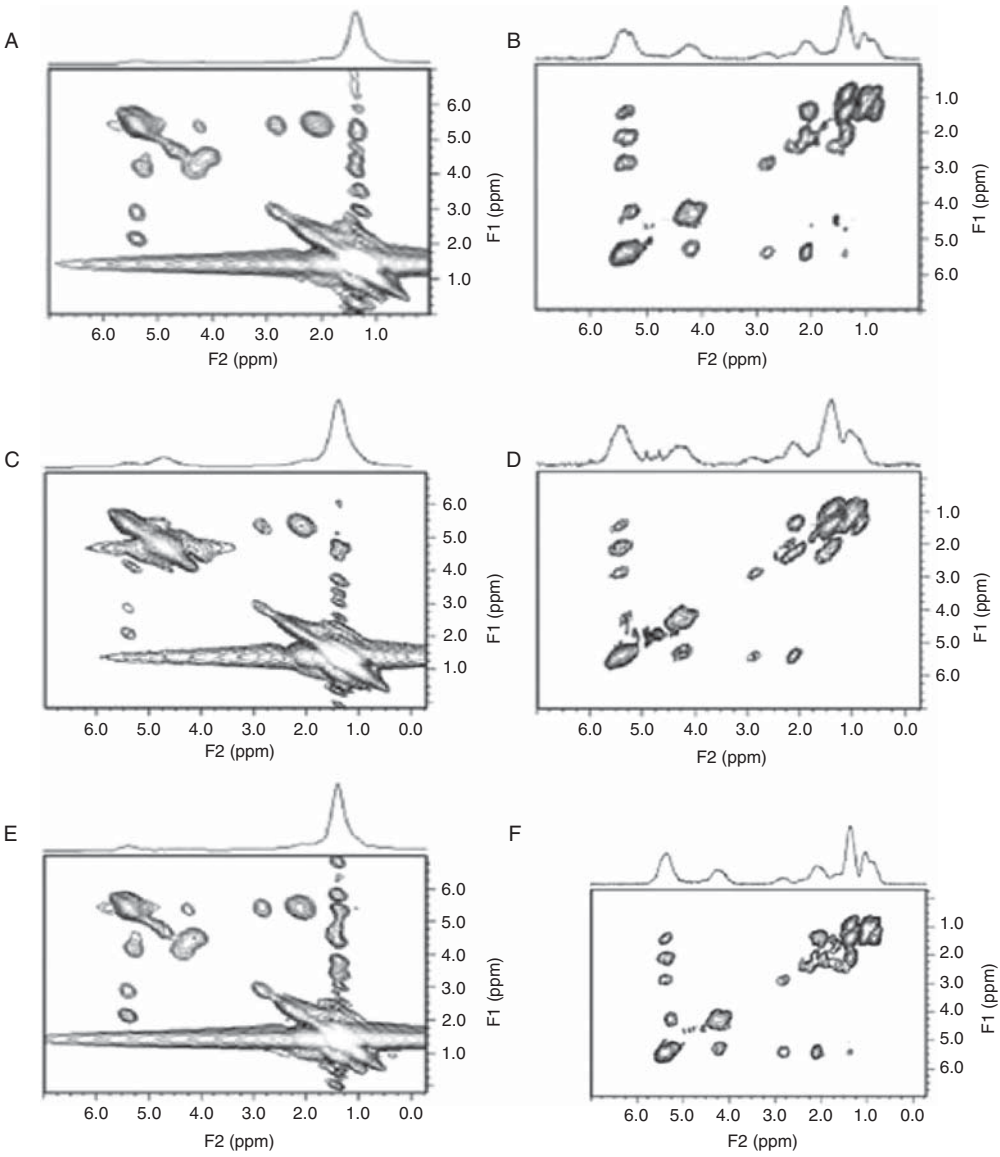


Figure 12 1.5 T L-COSY (left) and DQF-COSY (right) from the bone marrow of healthy person (A, B), patient with leukaemia (C, D) and leukaemia patient in remission (E, F). Experimental parameters – L-COSY: averages = 4, t_1 increment in F_1 = 1.6 ms, spectral width in F_1 = 625 Hz, acquisition duration = 256 ms in F_2, spectral width in F_2 = 1000 Hz, TR = 730 ms, $TE_{initial}$ = 35 ms, total experimental time = 2.5 min; DQF-COSY: averages = 8, t_1 increment in F_1 = 1.2 ms, spectral width in F_1 = 625 Hz, acquisition duration = 1024 ms in F_2, spectral width in F_2 = 1000 Hz, TR = 1700 ms, $TE_{initial}$ = 35 ms, total experimental time = 11 min (reproduced from Ref. 63 with permission).

malignant tissue and replaced by a strong choline and lipid resonances[65–67] and their relative concentrations reflect disease and pathology. The assignments for COSY spectra of biopsies from the different anatomical zones in the prostate are shown in Swindle et al.[68]

In Yue et al.,[69] air-filled endorectal coil is inserted into the patient's rectum. The air brings the coil closer to the prostate gland. Note that air is not the best medium that can be used to fill the endorectal coil to bring it closer to the prostate gland. Workers in this field used many solutions including D_2O and Barium sulphate suspension. Localized 2D JPRESS[69] was implemented on a 1.5 T whole body scanner using TR = 2 s, $TE_{initial}$ = 30 ms, number of t_1 increments = 40–64, data collection = 11–17 min, number of averages per t_1 increment = 8, size of t_1 increment = 5 ms added before and after the terminal 180° RF pulse. The prostate spectrum acquired from 2 cm^3 voxel from a healthy volunteer is shown in Figure 13 where citrate, polyamines and lipid are recorded. The JPRESS spectrum from the prostate of a patient with BPH is shown in Figure 14. Here, the spermine/spermidine levels remain high as does the citrate. In the 2D spectrum from a prostate cancer (Figure 15), there is a sharp reduction in polyamines and citrate, along with an increase of choline.

Thus, the feasibility 2D MRS has been shown but a more systematic study on a larger cohort, along with quantification of the cross-peaks,[70] is required.

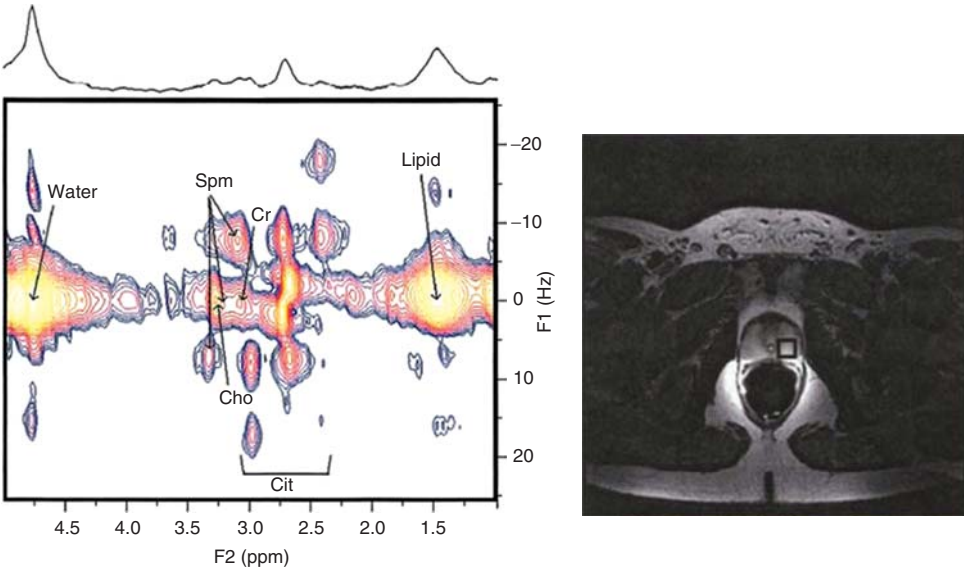


Figure 13 Left: 2D JPRESS spectrum (1.5 T) from a healthy prostate (28-year-old volunteer) (TR = 2 s, $TE_{initial}$ = 30 ms, 64 increments in F_1, 1024 points in F_2, averages per increment = 8). The cross-peaks from spermine/spermidine (Spm), citrate (Cit) and lipid are strong. Right: cross-sectional anatomical image showing the 2 cm^3 voxel position in the peripheral zone (reproduced from Ref. 69 with permission).

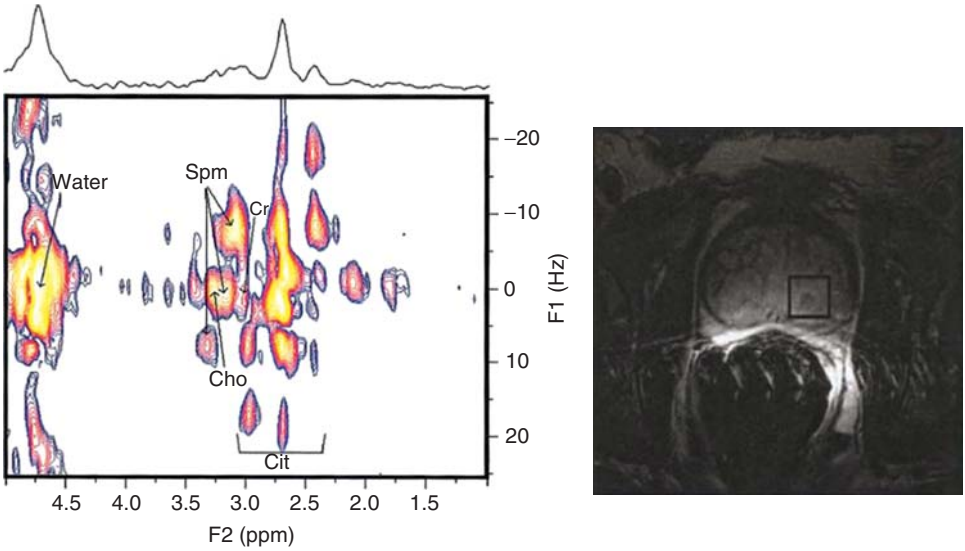


Figure 14 Left: 2D JPRESS spectrum (1.5 T) of the prostate of a 60-year-old with BPH (TR = 2 s, $TE_{initial}$ = 30 ms, 64 increments in F_1, 1024 points in F_2, averages per increment = 8). The spectrum shows an intense cross-peak from spermine/spermidine (Spm) and citrate (Cit) with lipid reduced in intensity. Right: cross-sectional anatomical image showing the 2.25 cm^3 voxel position in the peripheral zone (reproduced from Ref. 69 with permission).

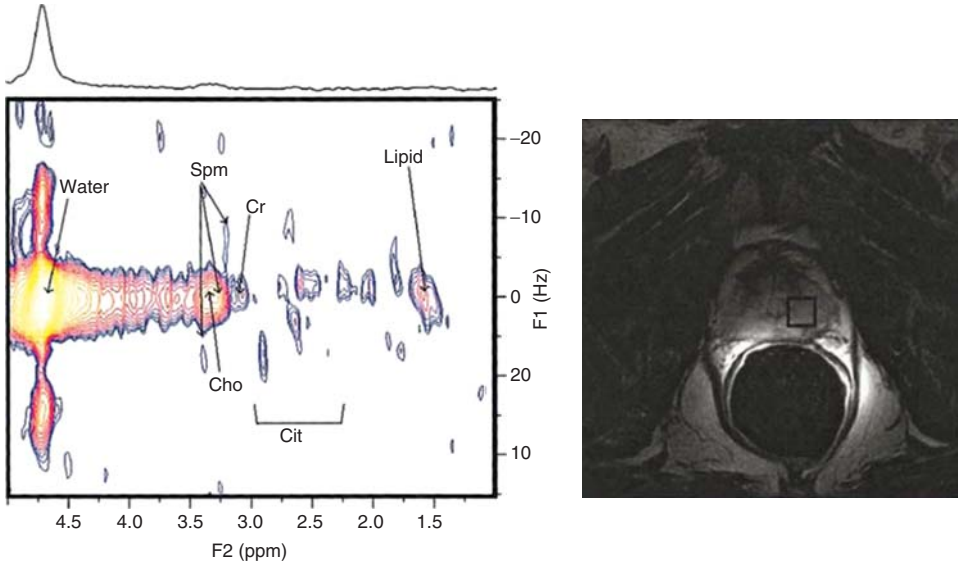


Figure 15 Left: 2D JPRESS spectrum (1.5 T) of a 79-year-old patient with prostate cancer (TR = 2 s, $TE_{initial}$ = 30 ms, 64 increments in F_1, 1024 points in F_2, averages per increment = 8). There is a strong choline (Cho) and lipid resonance on the diagonal. However, the cross-peaks from citrate and spermine/spermidine very much reduced in intensity. Right: cross-sectional anatomical image showing the 2 cm^3 voxel in the peripheral zone (reproduced from Ref. 69 with permission).

11. BRAIN

In 2001, Thomas et al.[49] reported localized 2D correlation spectroscopy (L-COSY) on a 1.5 T scanner with data acquired from the frontal and occipital grey/white matter regions of healthy volunteers. The data were acquired from a 27 cm^3 region in 34 min.

In the same year, Ziegler et al.[52] reported a similar experiment on a 3 T whole body scanner with data acquired from a similar voxel size in the left parieto-occipital lobe of a healthy volunteer in about the same time. Cross-peaks identified included *N*-acetyl aspartate (NAA), aspartate (Asp), myo-inositol (mI), taurine (Tau), glutamine/glutamate (Glx), gamma-aminobutyric acid (GABA) and threonine (Thr).

An *in vivo* L-COSY spectrum acquired on a 3 T magnet (Figure 16) shows (1) the improved spectral resolution at 3 T compared to similar spectra acquired at 1.5 T,[49] (2) more reliable detection of glutathione (GSH), (3) better separation of cross-peaks close to diagonal, for example, cerebral glucose (Glc) and the cross-peaks of NAA and Glx, (4) weaker effect of water suppression on close-by peaks (e.g., NAA and GSH) and (5) better separation of threonine (Thr) and lactate (Lac).

The wide variability of metabolites spin systems and their J-coupling values makes t_1 optimization a challenge in L-COSY. A diagram showing the coherence transfer amplitude as a function of t_1 values for different spin systems is shown in Figure 17, which shows that t_1 range must be wide enough to cover all metabolites that exist *in vivo*. The reproducibility of brain L-COSY[71] was shown by acquiring forty *in vivo* L-COSY spectra from 10 healthy volunteers, and scanning each one four times. The coefficients of variation of measured brain metabolites were less than 9% for Cr, Cho and NAA, and less than 17% for mI, Glx and Thr/Lac. The diagonal and cross-peak integrals were highly reproducible, with an inter-class correlation coefficient greater than 0.4 ($p < 0.05$).

In another study where L-COSY was used to study the neurochemistry of late-life depression,[72] the mean [Cho]/[Cr] in a sample of 33 elderly subjects was found to be higher in men than women by 10% ($p < 0.05$). Metabolites like mI, phosphoethanolamine (PE) and Glx were also found to be higher in depressed elderly females than controls, although these differences were not statistically significant. Cerebral metabolites were also measured using L-COSY technique in HIV-infected patients.[11] The concentration of mI and GABA were found to be elevated in the left frontal brain of HIV patients.

2D J-resolved spectroscopy (JPRESS) *in vivo* was first implemented in 1994 by Ryner et al.[15,16] on a 1.5 T magnet. More work followed[73] to produce a better quality JPRESS *in vivo* spectrum at 3 T, where NAA, Glx, Asp, Cr, Cho and mI were identified (Figure 18).

12. BREAST

An L-COSY spectrum can be acquired from ~1 cm^3 of breast tissue. *in vivo* breast L-COSY was implemented on eight cancer patients and nine healthy volunteers to evaluate water to lipid (saturated and unsaturated) ratio and to try to correlate

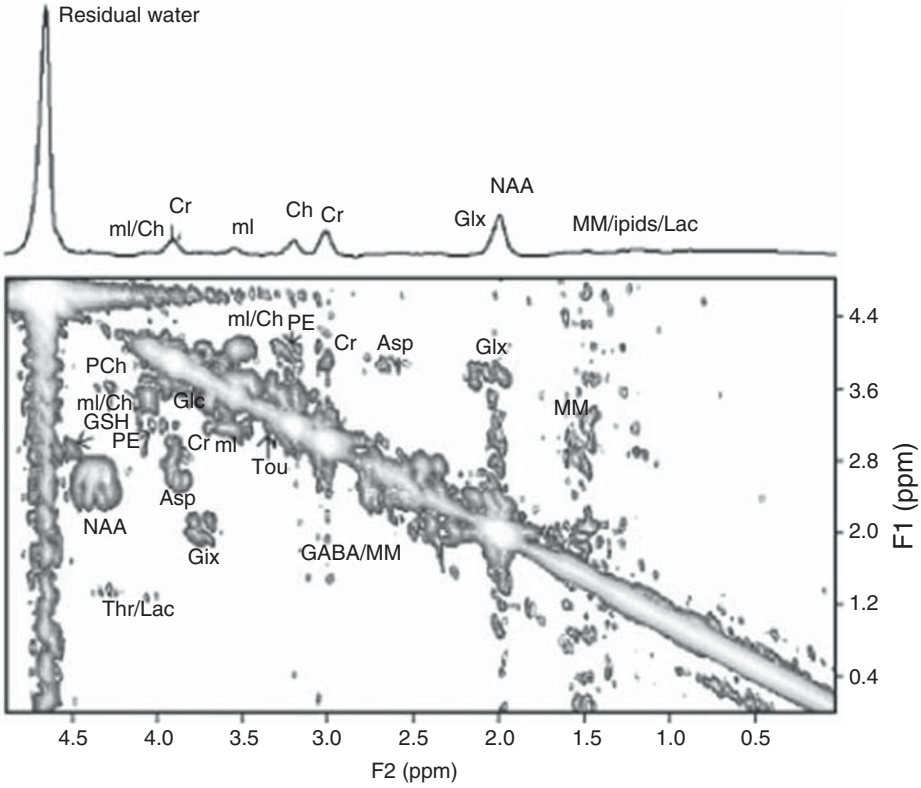


Figure 16 L-COSY spectrum of human brain *in vivo* at 3 T, acquired from 18 cm^3 voxel in the occipitoparietal region of a 38-year-old volunteer. Cross-peaks due to Glx(3.7,2.2), Asp(3.8,2.8), Et/PEt(3.9,3.2), Thr(4.3,1.3), PCH/GPC(4.3,3.8), GABA(3.0,2.0), GSH(4.5,2.9), MM(1.5,2.8), Tau(3.15,3.4), Glc (3.8,3.6), NAA(4.3,2.6), Cr(3.9,3.0), mI/Ch(4.0,3.5), mI(3.5,3.15) and Lac(4.0,1.3) are identified. Coefficients of variation of metabolite ratios with respect to Cr diagonal peaks varied between 8% and 19%. A head coil and a 4-in. surface coils were used for transmit and receive, respectively. Acquisition parameters: 2048 points in F_2, t_1 = 0.8 ms, 128 increments in F_1, TR = 2s, $TE_{initial}$ = 30 ms, averages = 8 per t_1 increment. A projected 1D spectrum is on top of 2D spectrum. (Reproduced from Ref. 73 with permission).

this ratio with disease.[74] Experimental parameters for beast L-COSY spectra were voxel size = 1 cm^3, experimental time = 10 min, $TE_{initial}$ = 30 ms, number of increments = 40, number averages per increment = 8, 1024 points and 2500 Hz in F_2, and 625 Hz in F_1. L-COSY spectra from healthy breast at three different locations were acquired and are shown in Figure 19. The olefinic–methylene protons coupling is evident by cross-peaks at (F_2 (ppm), F_1 (ppm)) (5.4,2.1) and (5.4,2.85), and the glyceryl backbone cross-peak at (5.2,4.4) was also detected. Cross-peaks from low molecular weight metabolites, for example, choline were not seen. Water/fat ratios were calculated for the three tissue types of a healthy subject (Figure 19) by evaluating diagonal peaks for water and polymethylene

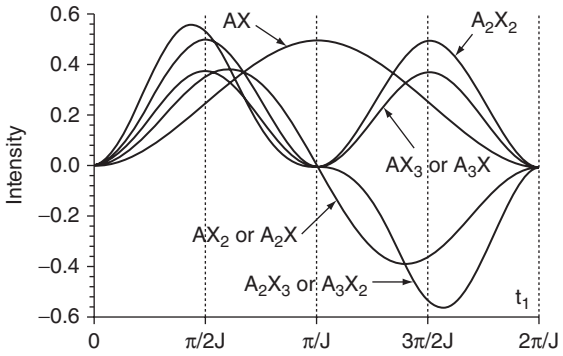


Figure 17 Amplitude of coherence transfer in a COSY pulse sequence (90°–t_1–90°–ACQ) for spin systems AX, A2X2, AX3, AX2 and A2X3, expressed as a function of t_1. The intensity is normalized to the signal of a one-spin system. Note that the intensity of AX is maximum when that of other spin systems is zero (Reproduced from Ref. 14 with permission).

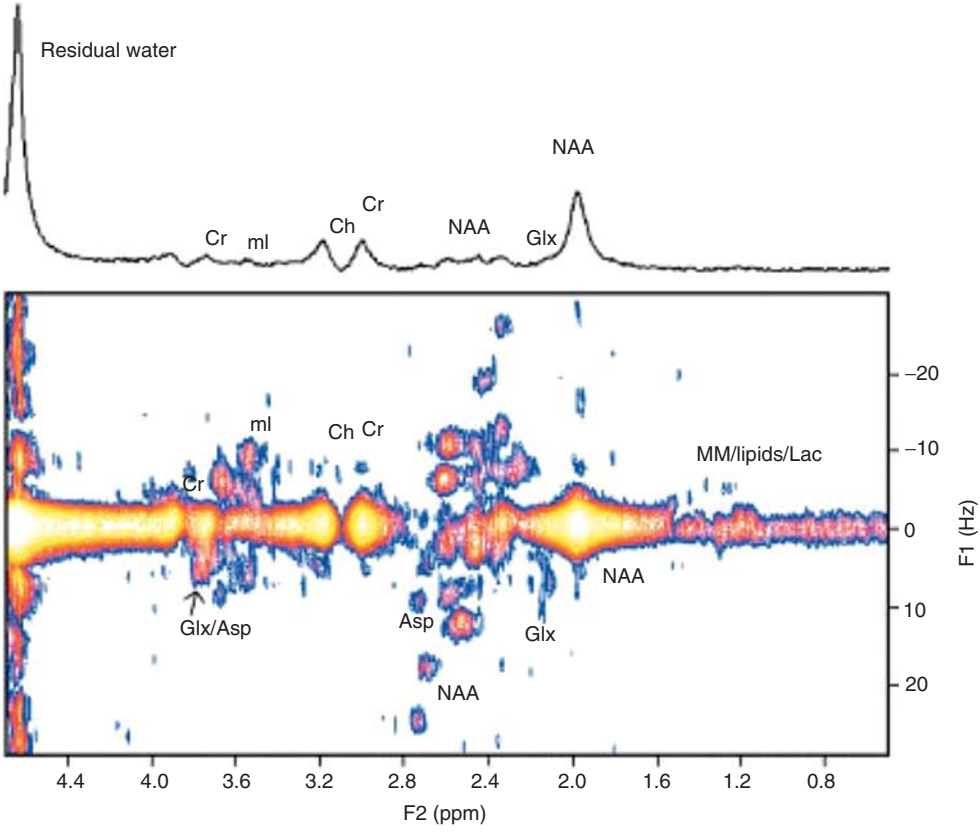


Figure 18 JPRESS spectrum of human brain *in vivo* at 3 T, acquired from 18 cm^3 voxel in the occipitoparietal region of a 21-year-old volunteer. Peaks due to NAA, Glx, mI, and Ch are identified. Coefficients of variation of metabolite ratios with respect to Cr diagonal peaks were all less than 11%. A head coil and a 4-in. surface coils were used for transmit and receive, respectively. Acquisition parameters: 2048 points in F_2, t_1 = 10 ms, 64 increments in F_1, TR = 2 s, $TE_{initial}$ = 30 ms, averages = 8 per t_1 increment. A projected 1D spectrum is on top of 2D spectrum (reproduced from Ref. 73 with permission).

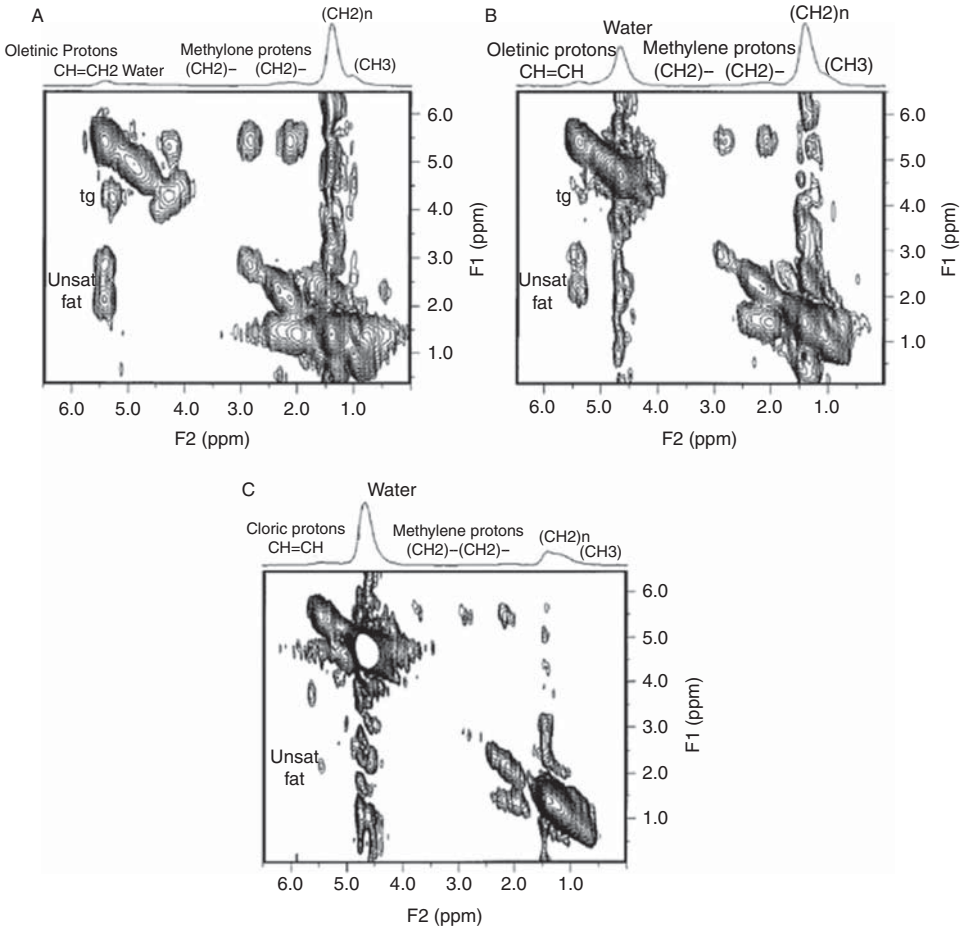


Figure 19 2D *in vivo* L-COSY spectra of a healthy breast from three different regions. Diagonal peaks of water (4.8 ppm) and lipid (0.9, 1.4, 2.1, 2.4, 2.85 and 5.4 ppm) are detected. Cross-peaks for (CH_2–CH=) at (5.4, 2.1) and triglyceride back bone protons (CH_2–CH) at (5.2, 4.4) were also observable. The above 2D are (A) 1 cm^3 in mammary fat tissue with water/fat = 0.056, (B) 1 cm^3 at the junction of fat and glandular tissues with water/fat = 0.65 and (C) 1 cm^3 in the glandular tissue with water/fat = 3.32. Acquisition parameters: voxel size = 1 cm^3, experimental time = 10 min, TR = 2 s, $TE_{initial}$ = 30 ms, number of increments = 40, number averages per increment = 8, 2500 Hz and 1024 points in F_2, 625 Hz in F_1 and t_1 increment = 1.6 ms. 1D spectrum is on top of 2D spectrum (reproduced from Ref. 74 with permission).

lipid ($(CH_2)_n$), and were found to be 0.056, 0.65 and 3.32 for fatty, mixture of fatty and glandular and glandular tissues, respectively. No choline resonances were detected in healthy tissue.

A 2D L-COSY from a breast benign mass of a 56-year-old patient is shown in Figure 20, with a more prominent water peak and weaker cross-peaks of the olefinic protons, when compared to the healthy spectra that are shown in

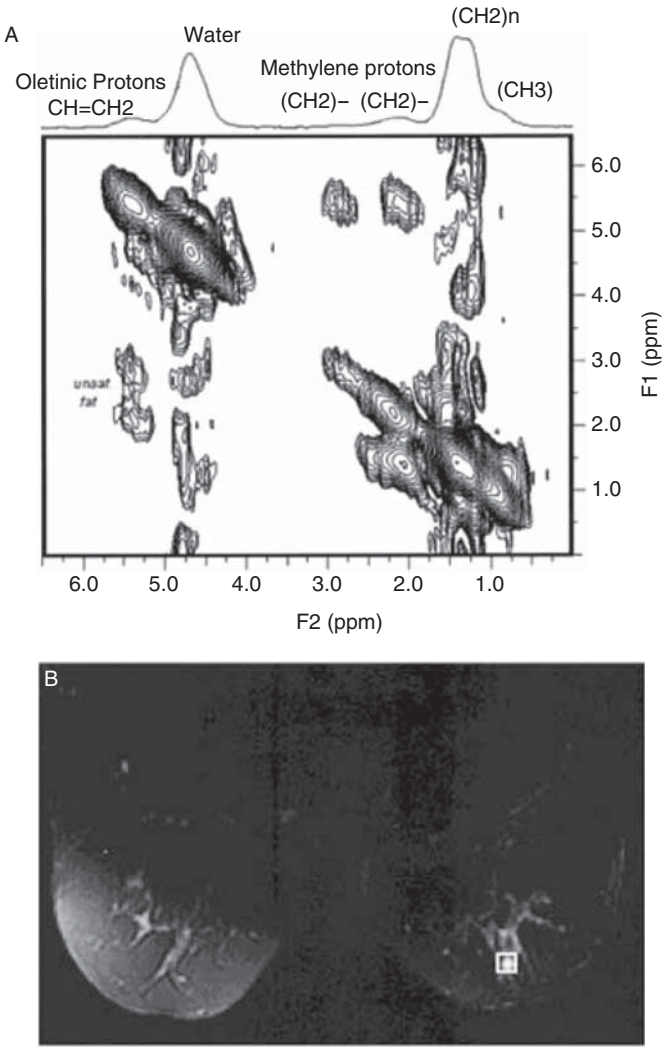


Figure 20 (A) 2D *in vivo* L-COSY spectra of a benign solid tumour and (B) the image showing where the voxel is located within the breast. Stronger water peak and weaker lipid peaks are observed when compared to healthy 2D spectra. Acquisition parameters: voxel size = 1 cm^3, experimental time = 10 min, TR = 2 s, $TE_{initial}$ = 30 ms, number of increments = 40, number averages per increment = 8, 2500 Hz and 1024 points in F_2, 625 Hz in F_1 and t_1 increment = 1.6 ms. 1D spectrum is on top of 2D spectrum (reproduced from Ref. 74 with permission).

Figure 19. Choline resonances were not observed in the spectra from benign breast tumours.[75]

The L-COSY spectrum from a patient with invasive ductal/adenocarcinoma is shown in Figure 21. Few observation can be made about Figure 21; (1) an elevated water diagonal peak, (2) a reduced saturated and olefinic lipid

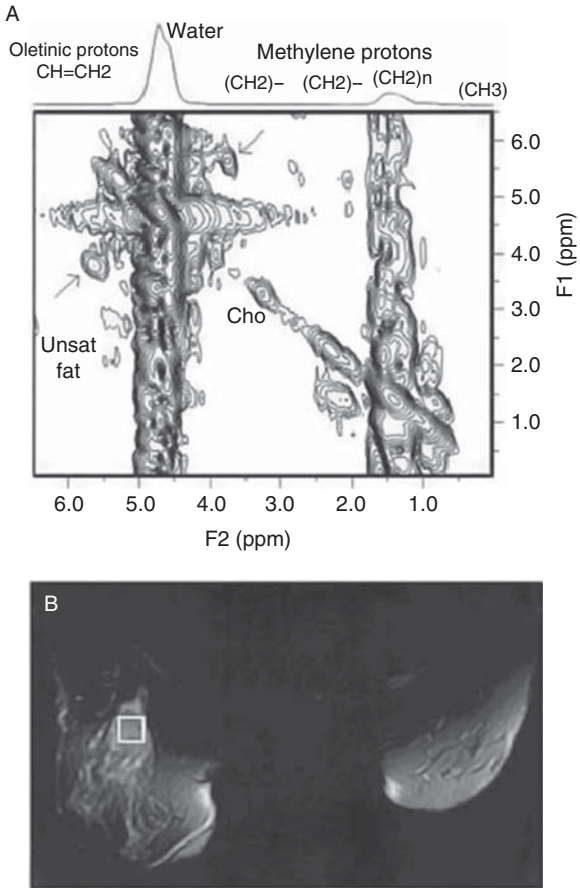


Figure 21 (A) 2D *in vivo* L-COSY spectra of an invasive ductal carcinoma and (B) the image showing where the voxel is located within the breast. Significantly stronger water peak and weaker lipid peaks are observed when compared to healthy 2D spectra. Note the Cho diagonal peaks at 3.25 ppm which is typical of ductal carcinoma spectra. Acquisition parameters: voxel size = 3.4 cm^3, experimental time = 10 min, TR = 2 s, $TE_{initial}$ = 30 ms, number of increments = 40, number averages per increment = 8, 2500 Hz and 1024 points in F_2, 625 Hz in F_1 and t_1 increment = 1.6 ms. 1D spectrum is on top of 2D spectrum (reproduced from Ref. 74 with permission).

resonances, (3) a prominent choline diagonal peak at F1= F2 = 3.25 ppm, and (4) a weak choline cross peak at (4.0, 3.5). The presence of choline and the higher water/fat ratio in cancer patients than in healthy volunteers is in agreement with 1D spectroscopy results.[76]

In 2005, a statistical classification report on the diagnostic efficiency of breast L-COSY in human breast carcinoma was published.[77] Twenty-one healthy women had their breasts examined as did 14 patients with invasive breast cancer. This cohort was studied by 2D L-COSY at 1.5. Eighteen diagonal and cross-peak ratios were evaluated for all 43 spectra. Significant decrease of diagonal and cross-lipid

peaks volumes were observed in carcinoma patients. Choline was only observed in spectra from the carcinomas.

The metabolite peak ratios significantly increased in carcinomas ($p < 0.05$) as described in Thomas et al.[77] Using these criteria, a classification and regression tree analysis (CART) was attempted. CART was able to improve the accuracy by analysing different combinations of peak ratios to correlate with carcinoma. Different prediction accuracies were obtained by considering different peak ratios. L-COSY improved peak characterization compared to 1D spectra from the breast.

13. 2D SPECTRAL FITTING (PROFIT)

ProFit (*prio*r knowledge *fit*ting)[70,78] is a 2D MRS fitting program with an approach that combines ''linear combination of model spectra'' (LCModel)[79] and ''variable projection'' (VAPRO)[80] methods. Thus, the combination of frequency domain fitting with prior knowledge constraints as well as a parameterized time domain fitting with linear and non-linear fitting algorithms is implemented in ProFit. The ProFit software produces concentration ratios to resonances in the spectrum, with Cramer-Rao lower bounds (CRLBs). In addition, the number of identifiable and quantifiable metabolites in human brain *in vivo* is 17 using JPRESS/ProFit compared to 8 using 1D/LCModel, mostly due to data spread on another spectral domain in JPRESS.[81]

ProFit was first attempted[70] on a 2D JPRESS phantom spectrum containing standard brain metabolites at physiological concentrations.[22] Reproducibility was checked by acquiring the JPRESS data at different times on a 3 T scanner, and CRLBs were calculated for each fit yielding acceptably low CRLBs.

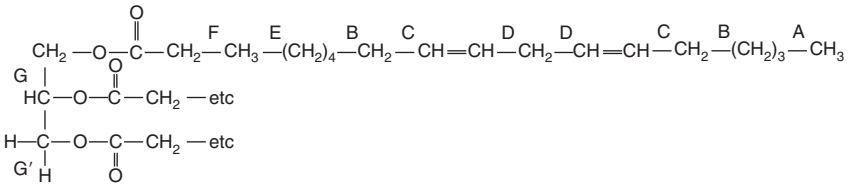
In addition, JPRESS and PRESS[82] data were collected from 27 healthy volunteers to evaluate ProFit *in vivo* and to compare to LCModel. The concentration ratios with respect to creatine, and the CRLBs were calculated using ProFit and LCModel. Results obtained by ProFit were slightly overestimated, probably due to macromolecular contamination. Intra-subject variability, obtained by acquiring and analysing data more than once, was lower in ProFit than LCModel. Recently, *in vivo* data acquired with L-COSY sequence[49] was analysed with ProFit and the resulting CRLBs compared to CRLBs obtained using JPRESS sequence.[83] It was found that the mean CRLBs are lower with L-COSY than with JPRESS.

14. CONCLUSIONS

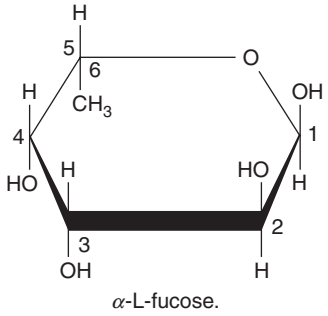
Two-dimensional MR spectroscopy has played an integral role in assigning the resonances in human cells, tissues and stool as they alter with disease and ageing. With an understanding of the wide range of T1 and T2 relaxation values present molecular species present in large quantities may be filtered, and those potentially diagnostic, but a lower concentration monitored. Many of the biological roles in tumour development and loss of cellular differentiation can be demonstrated using the 2D method. This is not possible using 1D MRS alone. The 2D

method also provided strong supporting evidence for the presence of an adenoma–carcinoma sequence in several organs including colon and ovary.

The progress of *in vivo* 2D MRS has been slow. Fourteen years after the first 2D *in vivo* COSY was collected, the clinical application of the technique is still in its infancy. There is no doubt about the utility and usefulness of the technique, however, more development is required to reduce the experimental acquisition time. The higher field strength magnets now available for *in vivo* studies are providing better spectral resolution and reducing scan time.



Structure 1



Structure 2

ABBREVIATIONS

1D	one-dimensional
2D	two-dimensional
Ala	alanine
Asp	aspartate
BPH	benign prostatic hyperplasia
CART	classification and regression tree
CHO	Chinese hamster ovary
Cho	choline
Cit	citrate
COSY	correlation spectroscopy
CPMG	Carr, Purcell, Meiboom, Gill
Cr	creatine

CRLB	Cramer-Rao lower bounds
CT	coherence transfer
DQF	double-quantum filter
EMCL	extra-myocellular lipid
FID	free induction decay
Fuc	fucose
GABA	gamma-aminobutyric acid
GDP	guanosine diphosphate
Glc	glucose
GlcNAc	*N*-acetylglucosamine
Gln	glutamine
Glu	glutamic acid
Glx	glutamine and glutamate combined
GSH	glutathione
HDL	high-density lipoprotein
IMCL	intra-myocellular lipid
JPRESS	J point-resolved spectroscopy
Lac	lactate
LCModel	linear combination of model spectra
L-COSY	localized correlation spectroscopy
Lex	Lewis antigen type x
L-EXSY	localized exchange spectroscopy
Ley	Lewis antigen type y
LIM1215	high human colon tumorigenic cell line
LIM1863	low human colon tumorigenic cell line
LMP	low malignant potential
mI	myo-inositol
MR	magnetic resonance
MRS	magnetic resonance spectroscopy
NAA	*N*-acetyl aspartate
PCa	prostate cancer
PE	phosphoethanolamine
PMF	plasma membrane fragments
ProFit	prior knowledge fitting
RF	radio frequency
SECSY	spin-echo correlated spectroscopy
SLR	Shinnar-Le Roux
SNR	signal-to-noise ratio
Spd	spermidine
Spm	spermine
STEAM	stimulated-echo acquisition mode
Tau	taurine
Thr	threonine
VAPRO	variable projection
VOSY	volume-selective spectroscopy

REFERENCES

1. J. Jeener, *Ampere International Summer School*. Basko Polje, Yugoslavia, 1971.
2. W. P. Aue, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.*, 1976, **64,** 2229.
3. R. L. Somorjai, M. Alexander, R. Baumgartner, S. Booth, C. Bowman, A. Demko, B. Dolenko, M. Mandelzweig, A. E. Nikulin, N. Pizzi, E. Pranckeviciene, R. Summers and P. Zhilkin, in: *Artificial Intelligence Methods and Tools for Systems Biology*, W Dubitzky and F Azuaje, eds., Springer/Kluwer, Boston, 2004.
4. C. Mountford, C. Lean, P. Malycha and P. Russell, *J. Magn. Reson. Imaging*, 2006, **24,** 459.
5. C. Mountford, S. Doran, C. Lean and P. Russell, *Chem. Rev.*, 2004, **104,** 3677.
6. A. Rutter, W. B. Mackinnon, L. I. Huschtscha and C. E. Mountford, *Exp. Gerontol.*, 1996, **31,** 669.
7. K. J. Cross, K. T. Holmes, C. E. Mountford and P. E. Wright, *Biochemistry*, 1984, **23,** 5895.
8. P. G. Williams, J. K. Saunders, M. Dyne, C. E. Mountford and K. T. Holmes, *Magn. Reson. Med.*, 1988, **7,** 463–471.
9. E. J. Delikatny, P. Russell, J. C. Hunter, R. Hancock, K. H. Atkinson, C. van Haaften-Day and C. E. Mountford, *Radiology*, 1993, **188,** 791–796.
10. I. M. Brereton, G. J. Galloway, S. E. Rose and D. M. Doddrell, *Magn. Reson. Med.*, 1994, **32,** 251.
11. S. Banakar, M. A. Thomas, A. Deveikis, J. Q. Y. Watzl, J. Hayes and M. A. Keller, *J. Magn. Reson. Imaging*, 2008, **27,** 710.
12. C. E. Mountford, S. Doran, C. L. Lean and P. Russell, *Chem. Rev.*, 2004, **104,** 3677.
13. A. Ziegler and M. Decorps, *J. Chim. Phys.*, 1998, **95,** 241.
14. A. Ziegler, M. Izquierdo, C. Remy and M. Decorps, *J. Magn. Reson. Ser. B*, 1995, **107,** 10.
15. L. N. Ryner, J. A. Sorenson and M. A. Thomas, *Magn. Reson. Imaging*, 1995, **13,** 853.
16. L. N. Ryner, J. A. Sorenson and M. A. Thomas, *J. Magn. Reson. Ser. B*, 1995, **107,** 126.
17. G. L. May, L. C. Wright, K. T. Holmes, P. G. Williams, I. C. Smith, P. E. Wright, R. M. Fox and C. E. Mountford, *J. Biol. Chem.*, 1986, **261,** 3048.
18. G. L. May, L. C. Wright, C. L. Lean and C. E. Mountford, *J. Magn. Reson.*, 1992, **98,** 622.
19. C. E. Mountford, W. B. Mackinnon, M. Bloom, E. E. Burnell and I. C. Smith, *J. Biochem. Biophys. Methods*, 1984, **9,** 323.
20. W. B. Mackinnon, G. L. May and C. E. Mountford, *Eur. J. Biochem.*, 1992, **205,** 827.
21. J. Hamilton and J. Morrisett, *Methods Enzymol.*, 1986, **128,** 412.
22. V. Govindaraju, K. Young and A. A. Maudsley, *NMR Biomed.*, 2000, **13,** 129.
23. C. E. Mountford, J. K. Saunders, G. L. May, K. T. Holmes, P. G. Williams, R. M. Fox, M. H. Tattersall, J. R. Barr, P. Russell and I. C. Smith, *Lancet*, 1986, **1,** 651.
24. C. E. Mountford, L. C. Wright, K. T. Holmes, W. B. Mackinnon, P. Gregory and R. M. Fox, *Science*, 1984, **226,** 1415.
25. R. R. Ernst, G. Bodenhausen and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford University Press/Clarendon Press, New York, 1987.
26. M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. Wiley & Sons, Chichester, 2001.
27. E. J. Delikatny, W. E. Hull and C. E. Mountford, *J. Magn. Reson.*, 1991, **94,** 563–573.
28. M. Coppleson, K. H. Atkinson and J. C. Dalrymple, eds., *Cervical Squamous and Glandular Intraepithelial Neoplasia: Clinical Features and Review of Management*, Churchill Livingstone, New York, 1992, p. 571.
29. C. L. Lean, W. B. Mackinnon, E. J. Delikatny, R. H. Whitehead and C. E. Mountford, *Biochemistry*, 1992, **31,** 11095–11105.
30. C. L. Lean, W. B. Mackinnon and C. E. Mountford, *Magn. Reson. Med.*, 1991, **20,** 306.
31. R. U. Lemieux, K. Bock, L. T. J. Delbaere, S. Koto and V. S. Rao, *Can J. Chem.*, 1980, **58,** 631–653.
32. A. Rosi, L. Guidoni, A. M. Luciani, G. Mariutti and V. Viti, *Cancer Lett.*, 1988, **39,** 153.
33. V. Schirrmacher, P. Altevogt, M. Fogel, J. Dennis, C. A. Waller, D. Barz, R. Schwartz, R. Cheingsong-Popov, G. Springer, P. J. Robinson, T. Nebe W. Brossmer, *et al.*, *Invasion Metastasis*, 1982, **2,** 13.
34. S. Hakomori, *Adv. Cancer Res.*, 1989, **12,** 301.
35. C. Campbell and P. Stanley, *J. Biol. Chem.*, 1984, **259,** 1208.
36. C. Pareskevca, P. Corfelda, S. Harper, A. Hague, K. Audcent and C. Williams, *Anticancer Res.*, 1990, **10,** 1189.

37. W. B. Mackinnon, L. Huschtscha, K. Dent, R. Hancock, C. Paraskeva and C. E. Mountford, *Int. J. Cancer*, 1994, **59,** 248–261.
38. J. J. Listinsky, G. P. Siegal and C. M. Listinsky, *Am. J. Clin. Pathol.*, 1998, **110,** 425.
39. T. Bezabeh, I. C. Smith, E. Krupnik, R. L. Somorjai, D. G. Kitchen, C. N. Bernstein, N. M. Pettigrew, R. P. Bird, K. J. Lewin and K. M. Briere, *Anticancer Res.*, 1996, **16,** 1553.
40. T. Bezabeh, NATIONAL RESEARCH COUNCIL OF CANADA, assignee, Method of diagnosing colorectal adenomas and cancer using proton magnetic resonance spectroscopy, World patent, WO/2004/027419, 2004.
41. P. Morice, F. Wicart-Poque, A. Rey, J. El-Hassan, P. Pautier, C. Lhomme, R. de Crevosier, C. Haie-Meder, P. Duvillard and D. Castaigne, *Cancer*, 2001, **92,** 2412.
42. P. Russell, in: *Ovarian Epithelial Tumours with Atypical Proliferation*, D. Lowe and H. Fox, eds., Churchill Livingstone, Edinburgh, 1992, p. 299.
43. P. Russell, S. J. Robboy and M. C. Anderson, in: *Ovarian Tumours: Classification and Clinical Perspective*, S. J. Robboy, M. C. Anderson and P. Russell, eds., Churchill Livingstone, London, 2002, p. 527.
44. J. C. Wallace, G. P. Raaphorst, R. L. Somorjai, C. E. Ng, M. Fung Kee Fung, M. Senterman and I. C. Smith, *Magn. Reson. Med.*, 1997, **38,** 569.
45. W. B. Mackinnon, P. Russell, G. L. May and C. E. Mountford, *Int. J. Gynecol. Cancer*, 1995, **5,** 211–221.
46. J. Diebold, F. Seemuller and U. Lohrs, *Lab. Invest.*, 2003, **83,** 251.
47. S. Hauptmann and M. Dietel, *Virchows Arch.*, 2001, **438,** 539.
48. J. Pauly, P. Leroux, D. Nishimura and A. Macovski, *IEEE Trans. Med. Imaging*, 1991, **10,** 53.
49. M. A. Thomas, K. Yue, N. Binesh, P. Davanzo, A. Kumar, B. Siegel, M. Frye, J. Curran, R. Lufkin, P. Martin and B. Guze, *Magn. Reson. Med.*, 2001, **46,** 58.
50. J. Hennig, *J. Magn. Reson.*, 1992, **96,** 40.
51. Felix, Felix NMR, Version 2007.
52. A. Ziegler, B. Gillet, J. C. Belloeil, J. P. Macher, M. Decorps and J. C. Nedelec, *Magn. Reson. Mater. Phys. Biol. Med.*, 2001, **14,** 45.
53. C. Boesch, J. Slotboom, H. Hoppeler and R. Kreis, *Magn. Reson. Med.*, 1997, **37,** 484.
54. F. Schick, B. Eismann, W. I. Jung, H. Bongers, M. Bunse and O. Lutz, *Magn. Reson. Med.*, 1993, **29,** 158.
55. S. S. Velan, C. Durst, S. K. Lemieux, R. R. Raylman, R. Sridhar, R. G. Spencer, G. R. Hobbs and M. A. Thomas, *J. Magn. Reson. Imaging*, 2007, **25,** 192.
56. S. S. Velan, S. Ramamurthy, S. Ainala, C. Durst, S. K. Lemieux, R. R. Raylman, R. G. Spencer and M. A. Thomas, *J. Magn. Reson. Imaging*, 2007, **26,** 410.
57. M. A. Thomas, H. K. Chung and H. Middlekauff, *Magn. Reson. Med.*, 2005, **53,** 495.
58. R. Kreis and C. Boesch, *J. Magn. Reson. Ser. B*, 1996, **113,** 103.
59. P. D. Thompson, P. Clarkson and R. H. Karas, *JAMA*, 2003, **289,** 1681.
60. A. Oriol, D. Valverde, J. Capellades, M. E. Cabañas, J. Ribera and C. Arús, *Magn. Reson. Mater. Phys. Biol. Med.*, 2007, **20,** 93.
61. T. Shih, C. Chang, C. Hsu, S. Wei, K. Su and H. Chung, *Spine*, 2004, **29,** 2844.
62. D. K. W. Yeung, J. F. Griffith, G. E. Antonio, F. K. H. Lee, J. Woo and P. C. Leung, *J. Magn. Reson. Imaging*, 2005, **22,** 279.
63. A. P. Prescot, A. S. K. Dzik-Jurasz, M. O. Leach, B. Sirohi, R. Powles and D. J. Collins, *J. Magn. Reson. Imaging*, 2005, **22,** 541.
64. M. Schnall, T. Connick, C. Hayes, R. Lenkinski and H. Kressel, *J. Magn. Reson. Imaging*, 1992, **2,** 229.
65. J. Kurhanewicz, R. Dahiya, J. M. Macdonald, L. H. Chang, T. L. James and P. Narayan, *Magn. Reson. Med.*, 1993, **29,** 149.
66. J. Kurhanewicz, D. B. Vigneron, H. Hricak, P. Narayan, P. Carroll and S. J. Nelson, *Radiology*, 1996, **198,** 795.
67. J. Kurhanewicz, D. B. Vigneron, R. G. Males, M. G. Swanson, K. K. Yu and H. Hricak, *Radiol. Clin. North Am.*, 2000, **38,** 115.
68. P. Swindle, S. Ramadan, P. Stanwell, S. McCredie, P. Russell and C. Mountford, *Magn. Reson. Mater. Phy.*, 2008, **21, 423**.
69. K. Yue, A. marumoto, N. Binesh and M. A. Thomas, *Magn. Reson. Med.*, 2002, **47,** 1059.
70. R. F. Schulte and P. Boesiger, *NMR Biomed.*, 2006, **19,** 255.

71. N. Binesh, K. Yue, L. Fairbanks and M. A. Thomas, *Magn. Reson. Med.*, 2002, **48,** 942.
72. N. Binesh, A. Kumar, S. Hwang, J. Mintz and M. A. Thomas, *J. Magn. Reson. Imaging*, 2004, **20,** 1039.
73. M. A. Thomas, N. Hattori, M. Umeda, T. Sawada and S. Naruse, *NMR Biomed.*, 2003, **16,** 245.
74. M. A. Thomas, N. Binesh, K. Yue and N. DeBruhl, *J. Magn. Reson. Imaging*, 2001, **14,** 181.
75. P. Stanwell and C. Mountford, *RadioGraphics*, 2007, **27,** S253.
76. N. R. Jagannathan, M. Singh, V. Govindaraju, P. Raghunathan, O. Coshic, P. K. Julka and G. K. Rath, *NMR Biomed.*, 1998, **11,** 414.
77. M. A. Thomas, N. Wyckoff, K. Yue, N. Binesh, S. Banakar, H. K. Chung, J. Sayre and N. DeBruhl, *Technol. Cancer Res. Treat.*, 2005, **4,** 99.
78. T. Lange, R. F. Schulte and P. Boesiger, *Proc. Intl. Soc. Mag. Reson. Med.*, 2007, 1401.
79. S. W. Provencher, *NMR Biomed.*, 2001, **14,** 260.
80. J. W. van der Veen, R. de Beer, P. R. Luyten and D. van Ormondt, *Magn. Reson. Med.*, 1988, **6,** 92.
81. E. Frias-Martinez, N. Rajakumar, S. Ramadan, S. Banakar, X. Liu, A. Singhal and M. A. Thomas, *Proc. Intl. Soc. Mag. Reson. Med.*, 2008, 1603.
82. P. A. Bottomley, *Ann. N.Y. Acad. Sci.*, 1987, **508,** 333.
83. E. Frias-Martinez, N. Rajakumar, X. Liu, A. Singhal, S. Banakar, S. Lipnick, G. Verma, S. Ramadan, A. Kumar and M. A. Thomas, *Proc. Intl. Soc. Mag. Reson. Med.*, 2008, 691.

SUBJECT INDEX

Page numbers followed by f indicate figures, t indicate tables

C