Recent contributions from solid-state NMR to the understanding of membrane protein structure and function

Peter J Judge and Anthony Watts

The plasma membrane functions as a semi-permeable barrier, defining the interior (or cytoplasm) of an individual cell. This highly dynamic and complex macromolecular assembly comprises predominantly lipids and proteins held together by entropic forces and provide the interface through which a cell interacts with its immediate environment. The extended sheet-like bilayer structure formed by the phospholipids is a highly adaptable platform whose structure and composition may be tuned to provide specialised functionality. Although a number of biophysical techniques including X-ray crystallography have been used to determine membrane protein structures, these methods are unable to replicate and accommodate the complexity and diversity of natural membranes. Solid state NMR (ssNMR) is a versatile method for structural biology and can be used to provide new insights into the structures of membrane components and their mutual interactions. The extensive variety of sample forms amenable for study by ssNMR, allows data to be collected from proteins in conditions that more faithfully resemble those of native environment, and therefore is much closer to a functional state.

Address
Biomembrane Structure Unit, Biochemistry Dept., University of Oxford, Oxford, OX1 3QU, UK

Corresponding author: Watts, Anthony (anthony.watts@bioch.ox.ac.uk)

Biophysical characterisation of membrane proteins

Many biophysical studies of proteins begin with expression in a suitable vector and isolation from other cellular components. Membrane protein purification is complicated by the presence of the lipid component of the membrane, which is frequently removed by detergent solubilisation in order to increase the concentration of the desired protein [1,2]. Many membrane proteins require specific lipids to be present in the membrane to be fully active [3**] and lipid species such as sphingolipids and cholesterol, may form into enriched domains that facilitate short-term, local organisation of parts of a natural membrane [4,5]. Biophysical experiments which require detergent-solubilised and delipidated membrane protein samples, are unable to replicate the complexity of natural membranes and may occasionally produce misleading descriptions of little functional meaning.

Around 200 unique membrane protein structures have been obtained by X-ray crystallography (http://blanco-biomol.uci.edu/Membrane_Proteins_xtal.html) and most are crystallised from detergent-solubilised preparations from which much of the native lipid has been removed [6]; a notable exception is the 7-transmembrane helical photoreceptor, bacteriorhodopsin (bR), which is routinely crystallised directly from the membranes in which it occurs naturally with only minor purification steps. Although tight binding of lipids to proteins is often reported in structural models from X-ray crystallography studies, the electron density corresponding to non-protein molecules is typically ill-defined and the lipid or detergent species may not be unambiguously assignable [7,8]. Crystallisation of membrane proteins remains a major hurdle for X-ray diffraction studies and many small bioactive peptides are elusive to crystallographic methods. Most X-ray diffraction data are acquired at cryo-temperatures at which the dynamic motions of peptides and proteins present under physiological conditions are suppressed and other biophysical approaches must be used to provide this detail.

Solid state NMR

Solid state NMR is a methodology commonly applied to a range of macromolecular (\(M_r \gg 100\) kDa) complexes which can be regarded as solid or solid-like, including membranes and membrane proteins [9]. Samples of this type cannot be studied by solution state NMR methods, as the anisotropic interactions are not averaged by molecular tumbling on the NMR time-scales of \(\sim\mu s\), resulting in the broadening of individual resonances. Solid state NMR is uniquely able to exploit the intrinsic anisotropy of macromolecular assemblies, and is broadly able to provide orientation information, distance restraints and torsion constraints to provide structural detail at sub-Å resolution. The same technique is also able to provide information about protein and lipid dynamics, allowing a more complete picture of the behaviour of a transmembrane protein [10*]. A key advantage of solid state NMR is the flexibility of sample compositions and lipids may be included to mimic more closely the characteristics of the native membrane in which a given protein would reside.
Lipids in membranes
Solid state (and previously, wide-line) NMR methods have been long used to reveal detailed order, dynamic and clustering details about membrane lipids, most often through the exploitation of $^{31}$P, $^2$H and $^{19}$F NMR [11]. In the recent context of domain formation, lateral exchange rate information for domain-forming lipids [12], and the influence of bilayer curvature on functional intermediate formation (in rhodopsin) of a membrane protein, have been resolved from $^2$H NMR [13*]. In addition, environmental factors such as temperature, hydration, and lipid bilayer properties are tightly coupled to the dynamics of membrane proteins, as has recently been shown for proteorhodopsin [14**].

Transmembrane protein structure determination via ssNMR orientational constraints
Except for a small family of β barrel proteins, all integral membrane proteins span the bilayer with α helices [15]. A key question is the orientation of those helices within the membrane and this information can be a first step towards obtaining structural information about a given protein [16]. Oriented sample methods exploit the large anisotropy of the electric field gradient of $^{15}$N (chemical shift) or $^2$H (quadrupolar) nuclei, and a key advantage of studying NMR labelled proteins in oriented bilayer samples of this type, is that all restraints are determined relative to a single external axis (the direction of the magnetic field), rather than to internal references within the protein (Figure 1).

Conventionally membranes containing a limited mixture of lipids are oriented mechanically using glass slides, which are then stacked before being placed into the magnet. The Influenza M2 ion channel was the first membrane protein structure to be deposited in the protein databank (PDB) (1MP6) and was resolved from orientational constraints alone [17]. Since then several similar single pass peptides have been solved using these approaches (e.g., phospholamban, Pf1, and amantadine-blocked M2 — for a complete list, see: www.drorlist.com/nmr/MPNM.R.html) [18**]. The technique has also been demonstrated for larger polytopic photoreceptor proteins from the retina [19,20].

Mechanical alignment of model membranes may be accomplished for a limited range of lipid mixtures, and a common cause for concern is the sample hydration level. Imperfect alignment and macroscopic disorder (mosaic spread) also result in spectral broadening, decreasing the precision of the orientational restraints [20]. A recent innovation in this field has been the introduction of magnetically aligned bicelles-based and nanodisc-based bilayers [21–23]. Bicelles may not only contain a more diverse range of lipids including cholesterol and those with unsaturated acyl chains [24] but bicelle-based samples also have a higher filling factor for the rf coil inside the NMR probe allowing increases in sample concentration, with a concomitant reduction in acquisition times relative to mechanically oriented systems [23].

High concentration detergent (and lipid) natural abundance $^{13}$C signals can obscure lower concentration protein resonances with micellar and bicellar systems, particularly in the C=O, CH$_2$ and CH$_3$ regions of the spectrum. Non-hydrolysable, ether-linked lipids have therefore been used to avoid protein spectral overlap; however, their phase behaviour differs from ester-linked lipids and protein function may be impaired [25].

A further potential disadvantage of bicelle-based systems has been that their alignment is sensitive to the temperature although the effect appears to be significantly reduced in the presence of biphenylated lipids [26]. Detergents may also significantly reduce the temperature dependence of the alignment [27] and, although the effect of adding non-physiological lipids and detergents to the sample on protein structure and functionality is unclear, the approach holds promise [28*,29].

High resolution solid state NMR
High resolution solid state NMR spectra are obtained by averaging the anisotropically broadened spectra of macromolecular complexes by mechanically spinning the sample at the magic angle (magic angle sample spinning, MAS); rates of 4–15 kHz are commonly used for $^{13}$C and $^{15}$N labelled systems [9]. Determining dipolar interactions and resolving correlation spectra, as in solution state NMR, are the basis of high resolution structural approaches.

Although MAS averages dipolar interactions, recoupling these interactions can be accomplished through rotational resonance (R$^2$) or REDOR (rotational echo double-resonance), or related methods, for homonuclear and heteronuclear spins respectively. A successful implementation of these approaches requires the choice of appropriate label positions. Not only must the labels be close (e.g., <~0.7 nm for $^{13}$C pairs) to allow significant dipolar recoupling to be reintroduced between the two spins, but additional structural information is often required when interpreting the spectra. A recent example is the use of REDOR to determine the conformation of residue His37 from the membrane embedded M2 viral ion channel (see above), which confers selectivity for H$^+$ ions on the channel [30*], but also adds new information about channel occlusion by adamantane [31**].

There has been an increasing trend to use nuclei other than $^{13}$C and $^{15}$N in REDOR and R$^2$ experiments, although in general these have focused primarily on synthetically synthesized peptides or hydrophobic species which interact with the membrane. Recent examples include the study of the interaction of the polyphenol epicatechin

References
Solid state NMR, both high resolution and wide line, can be used to gain detailed molecular insights into membrane-embedded proteins and their interaction with the membrane lipids, helping to arrive at functionally relevant descriptions. (a) Differential dynamics within membrane-embedded proteins have been demonstrated for helices and loops in U-[13C,15N] labelled, membrane-embedded bacteriorhodopsin [20], and for selectively reverse labelled sensory rhodopsin II from Natronomonas pharaonis (NpSRII) in liposomes [58]. Dipolar [13C,15N] correlation and water edited spectra of U-[13C,15N](V,L,F,Y) NpSRII in proteoliposomes were used, with different mixing times. Sequential Cα,Cβ and Cα,Cγ correlations were obtained from the [13C,15N] spectra recorded under weak coupling conditions shown here (left). Differential residues mobility (right) is described, with mobile residues in red, and rigid protein segments in blue — some residues are not labelled in the U-[13C,15N](V,L,F,Y) NpSRII sample (dark gray), and some cannot be assigned sequentially as a result of the reverse labelling (light gray). (b) Wide line deuterium NMR spectra of chain deuterated lipids in bilayers; (left); a typical spectrum) have been analyzed to deduce order parameters that enable the change in the energy of interaction (ΔΔG) of various lipid types with an integral membrane protein (left) in domains, or at the protein interface, in this case, retinal rhodopsin, a GPCR [13]. Figure courtesy of Klaus Gawrich, NIH, Bethesda, USA. (c) Oriented lipid (DOPC:DOPE) bilayers containing [13N labelled membrane peptides give rise to wide line NMR spectra (as a 2D representation of the [13N] chemical shift with the N-H dipolar coupling, so-called PISEMA spectra; see text) from which the orientational constraints of individual peptide residues can be assigned and structurally resolved [18**]. Spectral assignments were made for experimental data using selective labelling of Ile and Val (i), Leu and Phe (ii) and reversed labelled (iii) residues, and then simulated for two ideal helices with tilt angles of 32° and 105° are shown to indicate the transmembrane (TM) and amphipathic helices (red and green), respectively. This information then leads to models for the peptide packing to form the tetrameric, membrane-embedded M2 ion channel from the solid state NMR constraints as a ribbon representation of the TM and amphipathic helices (a, side view; b, 50° tilted view). To model the full channel pore and its constriction (c), the HOLE software was used [59], and several key residues at the junction between the TM and amphipathic helices which facilitates the close approach of adjacent monomers (d) to help in understanding the channel blocking mechanism. (d) Structural model of a polytopic (7-helical), membrane-embedded photoreceptor protein (sensory rhodopsin from Anabaena (Nostoc) sp. PCC 7120 (ASRI)) resolved from high-resolution solid state NMR data [46,47**]. 2D DARR [13C,15N] correlation spectra (at 800 MHz for 1H, 5C) with a long spin diffusion mixing times, gave long-range [13C,15N] distance restraints (703 medium to long range inter-residue cross peaks) were identified which were consistent with the X-ray structure, and then used for structure calculation together with the TALOS torsional restraints derived from the chemical shifts. The resulting model, including the identified water exposed fragments (red/orange) and water inaccessible residues (light/dark green), giv an approximate positioning of the protein in the lipid bilayer derived from H/D exchange, is shown (right) — unassigned residues are in gray.
gallate with DMPC bilayers using $^{31}$P–$^{13}$C REDOR, and $^{19}$F–$^{19}$F rotational resonances had been used to give distance constraints for an antimicrobial arylamide in contact with the lipid bilayer [32**]. $^{19}$F–$^{19}$F measurements have also been successfully demonstrated using the CODEX spin recoupling technique to measure distances of up to 11 Å within the M2 ion channel [33].

$^{19}$F, which does not occur naturally within biological molecules and has a high sensitivity in NMR experiments (high $\gamma$), is therefore well resolved, and may therefore be used for whole cell and native membrane experiments which would otherwise be inaccessible. A recent demonstration is the binding of a $^{19}$F labelled peptide PGLa to intact human erythrocytes and bacterial protoplasts [34**]. Fluorine-19 labels can also be introduced into short transmembrane peptides [35] and peptide ligands [36] by organic synthesis, but an exciting recent development has been the incorporation of $^{19}$F labelled amino acids into a soluble protein (retinol binding protein) using a cell-free peptide synthesis system [37]. Cell-free expression systems have the advantage of reducing the complexity of purification protocols and thereby sample preparation times [38*] as shown recently for the uniformly $^{13}$C and $^{15}$N labelled 75 kDa ion channel MscL produced by cell-free expression and Escherichia coli cell culture gave identical resolution [39**,40]. Again, the protein is in membranes, and MscL has been further studied by a range of methods in the less inaccessible open state by solid state NMR to help resolve membrane-determined mechanistic details [41].

Protein–protein interactions are still a challenge in structural biology for membrane-embedded proteins, but helix–helix interactions for glycoporin A [42], and more recently with signal enhancements by dynamic nuclear polarisation (DNP) (see below) [23,24] have been elucidated by solid state NMR, with generic conclusions about the potential mechanism in lipids for such interactions being resolved.

**Torsion angle constraints — protein structure determinations**

Extending and developing solution state NMR structural methods has led to the use of two-dimensional and three-dimensional spectra to assign resonances, whose chemical shifts may then be used for the basis of structure calculations. The first successful demonstration of this technique has been the partial assignment of resonances from the b-barrel protein OmpX from *E. coli* [43]. It should be noted that most other membrane proteins have a significant $\alpha$-helical content and that the chemical shifts of backbone carbon and nitrogen resulting in significant spectral overlap and assignment difficulties [44,45]. Nevertheless structural information has been obtained for proteorhodopsin using this method, importantly, in a lipid environment [46,47**].

Although many membrane proteins are more stable in a lipid environment, sample stability for multi-dimensional ssNMR is of concern, not only because of the long experimental times required (typically days to weeks), but also because of the lengthy proton decoupling times required [48]. An innovative study on crystalline samples of the soluble GB1 protein employing G-matrix Fourier Transform (GFT) projection spectroscopy has resulted in a reduction in acquisition times of up to 90% for 3D spectra [49], although this has yet to be applied to membrane proteins. Non-uniform sampling methods and selective or reverse labelling methods are also being explored to aid assignments [50].

One approach to address sensitivity issues, dynamic nuclear polarisation (DNP) in which microwave radiation is used to transfer spin polarisation from electrons to nuclei, has been successfully demonstrated for bacteriorhodopsin allowing a signal enhancement of up to 90-fold [51,52]. Full or partial deuteration of membrane proteins lengthens $T_2$ to improve peak resolution [53,54*]. Although MAS itself increases spectral resolution, orientational information from dipolar couplings is lost; however, spinning oriented membranes using magic angle oriented sample spinning (MAOSS) has the potential for both high resolution and orientational information content to be realized [55,56*,57].

**Future outlook — complementarity and the membrane environment**

Key goals for solid state NMR, therefore, are the reduction of sample preparation, data acquisition and analysis times through improvements in protein expression and labelling, instrumentation and computational analysis. Understanding structure and function with respect to the essential lipid environment is crucial [3**, and utilizing a combination of information from many methods, not just one in isolation, is the way forward, as recently shown in the combination of molecular dynamics (MD) and toxin–membrane interactions [30*] and distance constraints to support MD in rhodopsin [50]. Solid state NMR can be used to study both lipids and proteins, so it has a pivotal role at the interface of a holistic approach to membrane studies and is one more tool in the structural biologist’s arsenal.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

Here, the notion that membrane protein structures are best studied in the membrane environment is presented, and the potential for structural perturbations in non-membrane environments expounded.


29. Understanding channel blocking at the mechanistic level requires a bilayer, as shown convincingly here using NMR.


Technical details of how to obtain labelled proteins in high yield for NMR.


A further use of cell-free expression system for channel protein production and structural study.


A first example of a close-to-complete structural description for a polytopic protein in membranes, using clever labelling strategies.


Enhancing the sensitivity of NMR using DNP, and how deuteration can add to the enhancement, is reported here.


A first for DNP signal enhancement in oriented membrane systems.

