the FEBS Journal

REVIEW ARTICLE

Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination

Michael Bieri[1], Ann H. Kwan[2], Mehdi Mobli[3], Glenn F. King[3], Joel P. Mackay[2] and Paul R. Gooley[1]

1 Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Australia
2 School of Molecular Bioscience, The University of Sydney, Australia
3 Institute for Molecular Bioscience, The University of Queensland, St. Lucia, Australia

Keywords
chemical shift mapping; NMR; NOE; protein; protein complex; protein dynamics; protein folding; protein interaction; protein mutagenesis; saturation difference

Correspondence
P. R. Gooley, Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, VIC 3010, Australia
Fax: +61 3 9348 1421
Tel: +61 3 8344 2273
E-mail: prg@unimelb.edu.au

(Received 20 July 2010, revised 7 November 2010, accepted 5 January 2011)

doi:10.1111/j.1742-4658.2011.08005.x

A strength of NMR spectroscopy is its ability to monitor, on an atomic level, molecular changes and interactions. In this review, which is intended for non-spectroscopists, we describe major uses of NMR in protein science beyond solution structure determination. After first touching on how NMR can be used to quickly determine whether a mutation induces structural perturbations in a protein, we describe the unparalleled ability of NMR to monitor binding interactions over a wide range of affinities, molecular masses and solution conditions. We discuss the use of NMR to measure the dynamics of proteins at the atomic level and over a wide range of timescales. Finally, we outline new and expanding areas such as macromolecular structure determination in multicomponent systems, as well as in the solid state and *in vivo*.

Introduction

As detailed in the accompanying review [1], simple NMR spectra such as the 1D ^{1}H NMR spectrum and ^{15}N-HSQC can rapidly provide a great deal of information about a protein using a relatively small amount of sample. For example, one can quickly assess the degree of folding of the protein, its thermal and temporal stability, and its aggregation propensity. Over the last 20 years, the technology for protein structure determination using NMR methods has developed such that high-quality structures can be determined comparable with those determined using X-ray crystallography.

One of the traditional strengths of NMR spectroscopy is its versatility. NMR methods have been used to provide structural and functional information on materials as diverse as human tumours [2], spider silk [3] and soil [4]. This versatility is also reflected in the world of biomacromolecular NMR, and approaches have been developed to probe proteins that are challenging for X-ray diffraction methods, including inherently flexible [5] and integral membrane proteins [6]. The analysis of protein interactions – both with other macromolecules and with small molecule ligands – is a

Abbreviations
pKID, phosphorylated kinase-inducible activation domain; STD, saturation transfer difference.

particular strength of NMR, and the ability to probe protein dynamics has also developed into a mature discipline. The purpose of this review is to outline the range of different applications, other than solution structure determination, for which NMR can be used in the analysis of proteins. The focus is on relatively easy to perform experiments that at most require simple isotope labelling strategies.

Assessment of protein folding by NMR

Screening site-directed mutants

NMR spectroscopy can be an enormously helpful tool for functional mapping of a protein. It is common biochemical practice to use site-directed mutagenesis as a means to determine the functional importance of a residue. However, such an experiment always prompts the question: 'How localized are the effects of the mutation and has it affected the protein fold?' Because each peak in an NMR spectrum is a sensitive probe of the local chemical environment experienced by the nucleus, easily recorded spectra such as ^{1}H 1D (Fig. 1) and ^{15}N-HSQC NMR spectra can provide a simple and rapid assessment of whether a mutant protein is properly folded or not. The extent and location of any structural perturbations can also be monitored in ^{15}N-HSQC spectra if resonance assignments are available for the protein. Small chemical shift changes for residues in the immediate vicinity of the mutation would be expected, but more substantial changes might indicate that the mutation causes a more significant change to the structure or dynamics of the molecule [7]. In general, we would recommend that an analysis of this type be considered an essential aspect of any mutagenesis study, especially when no other assay is available to confirm proper folding of the protein.

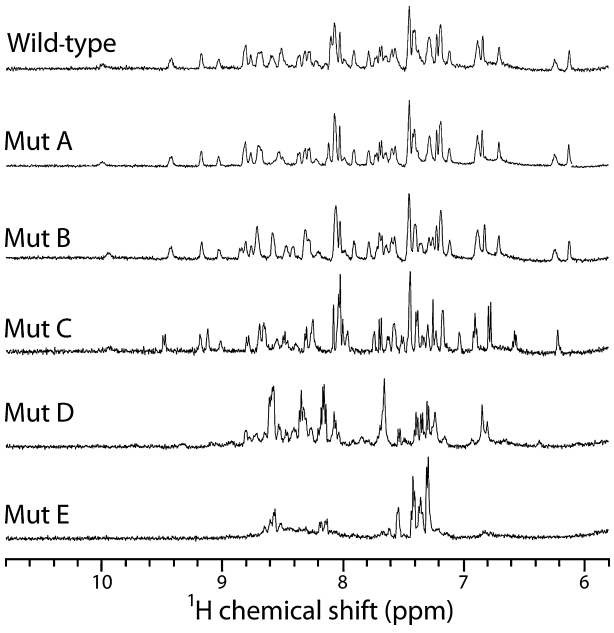

Fig. 1. 1D ^{1}H NMR spectra of the wild-type and several point mutants of a classical zinc finger from a mammalian transcription factor. The amide proton region of the spectrum is shown, and the data clearly indicate that the mutants A, B and C are well-folded and have conformations that closely resemble the wild-type domain, whereas D and E are not correctly folded and therefore should not be used for any subsequent functional experiments.

Probing folding pathways

Akin to observing the effect of a mutation on the structure of a protein, 1D and 2D NMR experiments have been used to monitor the process of protein folding [8]. The ^{15}N-HSQC experiment is especially powerful in following protein folding because a large number of amide resonances can be monitored simultaneously; thus, a probe exists for almost every residue in the protein. Recently, a combination of rapid sample mixing and fast NMR acquisition approaches have provided the ability to monitor folding in real time, at least for slow-folding proteins [9,10]. Valuable information about protein folding intermediates can be gained from such analyses. For example, monitoring of the folding of α-lactalbumin in this fashion revealed that all amide sites appear to undergo a single transition between the molten globule and folded states [10]; whereas a similar analysis of β-microglobulin indicated that this protein does not fold in a simple two-state manner [11].

Probing protein complexes

Assessing protein interactions by chemical shift and intensity changes

An important aspect of biology is the formation of bimolecular and higher order complexes. NMR experiments can detect such binding events even when the strength of the interaction is very weak (with K_d in the mM range) and difficult to detect by other biophysical techniques. The potential also exists to determine dissociation constants, and to discern the number and location of binding sites. With the development of high-field NMR spectrometers and cryogenically cooled probes, ligand binding experiments can be performed easily on 50 μM samples using volumes as small as 100 μL.

First impressions of ligand binding can be obtained without resonance assignments: one can simply ask whether the chemical shifts or resonance intensities are affected upon the addition of ligand (that is, does the ligand bind *at all*?). More detailed information, such as the ligand binding site, is obtained after complete assignment of all peaks in the ^{15}N-HSQC spectrum. To identify the amino acid residues involved in ligand binding, the labelled protein sample is titrated with ligand and ^{15}N-HSQC spectra recorded following each addition. Because the ligand is unlabelled, it will be effectively invisible in the experiment, although its effects on the protein can be observed. Peaks from residues involved in ligand binding will experience a change to their chemical environment that will be manifested as changes in peak position and/or intensity. The nature of these changes is effectively dependent on the off-rate of the interaction and the chemical shift difference between the free and bound states. Suffice to say, three broad conditions are observed (Fig. 2): (a) when the off-rate is much less than the frequency (chemical shift) difference between the free and bound states, the signal of the free state disappears in stages and reappears elsewhere in the spectrum as a resonance reflecting the bound state; (b) when the off-rate is much greater than the chemical shift difference between the two states, the peak gradually shifts position in the spectrum towards the chemical shift of the bound state; and (c) when the off-rate is about equal to the chemical shift difference between the two states, the peak broadens substantially and shifts; it may even vanish during the titration. These three conditions are referred to as slow, fast and intermediate exchange and they are typically associated with dissociation constants of submicromolar, high micromolar to millimolar, and micromolar, respectively.

Binding events typically affect nuclei within or near the binding site and so determination of which signals undergo the largest changes will reveal the location(s) of binding sites. Plots of chemical shift change during the titration (for fast exchange data) or intensity changes (for slow exchange data) can also be fitted to determine the K_d [12]. This approach is most straightforward in the former case, and it has been applied to a wide range of systems, including carbohydrate-binding proteins with one or multiple binding sites [13,14], protein–protein complexes [15,16] and nucleic acid-binding proteins [17–19]. In each case, resonance assignments were already available and binding caused significant chemical shift changes that could be mapped to discrete binding sites on the protein. If the affinity of an interaction is strong and the exchange rate is slow, reassignment of the protein signals must be carried out independently for the bound form and to obtain a structure of the complex [20]. However, even in the absence of bound-state assignments, intensity changes allow some mapping of the binding site [21,22].

The formation of encounter complexes is a subtle phenomenon that has only recently become accessible through the high-resolution of NMR measurements. For example, the interaction between phosphorylated

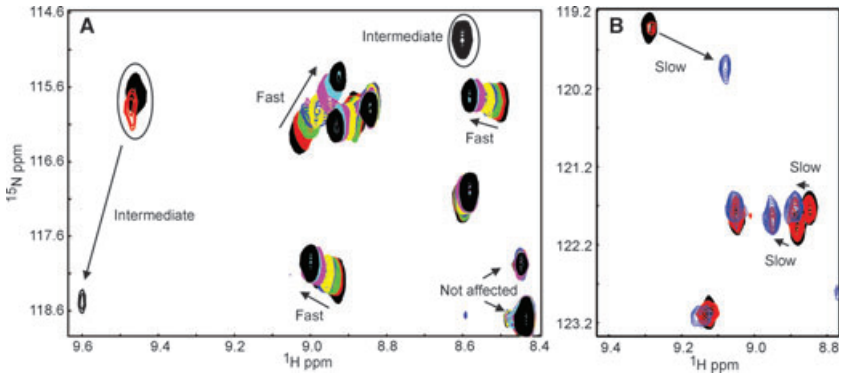

Fig. 2. Different chemical-exchange regimes observed in 2D ^{1}H,^{15}N-HSQC titrations of two proteins with a small oligosaccharide. (A) Series of eight ^{15}N-HSQC spectra of protein A in the presence of increasing amounts of oligosaccharide. In these spectra, several resonances show exchange on either the fast or intermediate timescale. Two peaks are marked as not affected. As examples of fast exchange, the three peaks indicated by arrows are shown to shift gradually through the spectrum in a 'straight line', indicating that a single interaction event occurs. Two peaks are circled that show intermediate exchange. The peak on the right broadens and disappears at the first titration point and does not appear to return; the peak on the left broadens and begins to reappear at the end of the titration. (B) Series of three ^{15}N-HSQC spectra of protein B during the addition of oligosaccharide. Chemical exchange in this titration is substantially slower. Shown in black are resonances in the absence of ligand. The red spectrum is taken at a 0.5 : 1 molar ratio of ligand:protein and the blue spectrum shows the protein saturated with ligand. For each of the resonances indicated with arrows, two peaks are observed at 0.5 : 1, one for the free protein and another for the complex.

kinase-inducible activation domain (pKID) and the KIX domain of the transcription factor CREB [23] was analysed by the addition of KIX to ^{15}N-labelled pKID. At sub-stoichiometric levels of KIX, signals in the ^{15}N-HSQC spectrum of pKID shift, as expected, in a 'straight line'; however, the direction of change is not initially towards the fully bound pKID-KIX positions, and overall the signals trace out non-linear paths during the titration. Such behaviour suggests that an initial (or encounter) complex is formed at sub-stoichiometric levels of KIX and that this complex is distinct from and converts to the final fully bound state.

Finally, in these chemical-shift mapping experiments, one must also be alert to the possibility of substantial conformational change. If the ^{15}N-labelled protein does undergo a significant change in conformation (e.g. an unfolded-to-folded transition), then chemical shift changes can extend across a much larger region than simply the binding interface.

Saturation difference spectroscopy

Chemical shift mapping is relatively straightforward providing that the protein of interest can be isotopically labelled and is of a molecular mass that yields a reasonable ^{15}N-HSQC spectrum. For large proteins and supramolecular complexes that do not give workable NMR spectra a method called saturation transfer difference (STD) can provide a great deal of useful information [24]; STD-NMR is widely used in drug discovery for finding weak binding compounds that can serve as drug leads [25]. A simple 1D experiment can take just a few minutes to acquire and so screening for 'hits' among a compound library containing thousands of members in a high-throughput manner is easily achieved. The experiment is also useful for mapping the binding epitope of a ligand that recognizes its receptor [26], even though no information can be obtained about the site on the protein receptor to which the ligand binds. In STD-NMR, the interaction is monitored by observing the NMR spectrum of the ligand, following saturation of resonances that correspond to nuclei in the protein (Fig. 3A). Only nuclei in the ligand that directly contact the protein will be perturbed. A relatively fast off-rate for complex formation (corresponding to $K_d > \sim 0.1$ μM) is a prerequisite for this approach; this is generally not a restriction in the screening of compound libraries, which typically yield 'hits' with affinities in this range. Unlike most NMR approaches, interactions in which one of the binding partners is extremely large can be probed. The approach has been applied to integral membrane proteins [27] and even viral particles [28] and whole cells [29].

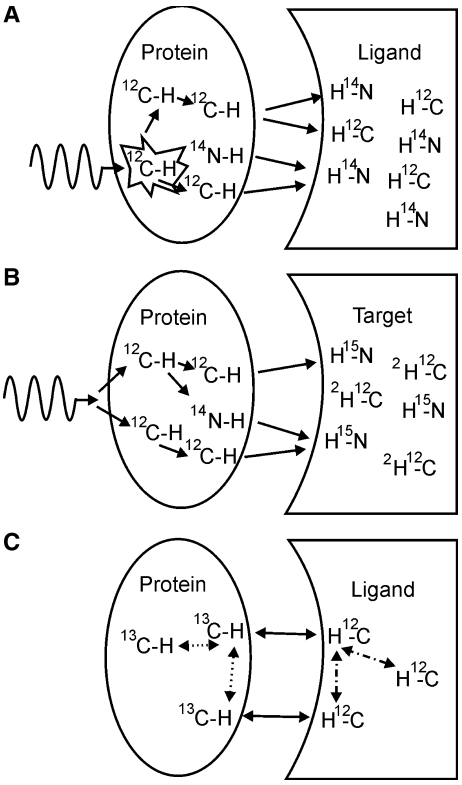

Fig. 3. Using saturation and NOE experiments to determine protein–ligand interactions. (A) Schematic of an STD experiment. In this case, neither the protein nor the ligand are labelled and a well-resolved resonance of the protein is excited. Excitation is passed throughout the protein and transferred to the protein–ligand interface, resulting in changes to the intensity of signals from the bound ligand. (B) Schematic for cross-saturation experiments. The protein is unlabelled, whereas the target (usually another protein) is labelled with ^{2}H and ^{15}N. The target is back-exchanged into $^{1}H_2O$ so that the ^{15}N nuclei are at least partially protonated. The protein is nonselectively excited, but the excitation is transferred from the target to only residues at the protein–target interface. (C) Schematic for using ^{13}C-half-filtered NOESY experiments. The protein is labelled with ^{13}C, whereas the 'ligand', irrespective of whether it is a protein, nucleic acid or any other molecule, is unlabelled. Protein–protein (dots), ligand–ligand (dot-dash) and protein–ligand (solid arrows) NOEs can now be separated.

In a related approach, information about protein–protein interfaces can be obtained from cross-saturation spectroscopy. If the two interacting partners in a complex are made with different isotopic labelling patterns, nuclei on one partner (the unlabelled partner) can be selectively irradiated and the effects on the ^{15}N-HSQC spectrum of the other (labelled) partner observed [30] (Fig. 3B). An advantage of this approach over the chemical-shift mapping experiments mentioned above is that cross-saturation will only be observed for nearby nuclei – chemical shift changes

can also be observed for residues that are distant to the binding interface but undergo a conformational change upon complex formation. However, cross-saturation has seen relatively little use to date, because of the requirement that one partner needs to be uniformly labelled with ^{2}H so that effects are confined to nuclei at the binding interface.

Defining a complex by NOEs

Returning to the analysis of smaller protein complexes, one can move beyond mapping of binding surfaces to the determination of a high-resolution structure of a complex using experiments based on the NOE (see the accompanying review [1] for an explanation). For a protein-protein complex, one could simply label both components with ^{13}C and ^{15}N and then determine the structure as if it were a single polypeptide chain. However, a cleverer strategy is available: the preparation of several NMR samples in which only one or the other component is labelled allows both simplification of the spectra and the application of more sophisticated NOE experiments in which one can selectively observe only the NOEs *within* one subunit or those *between* the two subunits (Fig. 3C) [31,32]. These latter NOEs are, of course, particularly valuable in defining the binding mode and they can be difficult to track down among the thousands of intrasubunit NOEs in a traditional NOESY experiment. Many structures have been determined using this strategy, including those of protein–protein [33], protein–nucleic acid and protein–small molecule [34] complexes, as well as oligomeric proteins [35]. Even very low-affinity complexes are amenable to this approach: the SH3 domain of PINCH-1 binds to the LIM domain of Nck2 with an affinity of ~ 3 mM and yet sufficient NOEs could be observed to define the interface [36].

Low-resolution modelling of macromolecular complexes

Full structure determination for a protein by NMR can be labour intensive, and several strategies have consequently been developed to allow the construction of lower resolution (but still very useful) models of protein complexes. These approaches take advantage of the knowledge gained from mutagenesis and chemical-shift mapping experiments described above, in combination with the structures of the two interacting partners (which could be NMR or X-ray structures, or even high-quality homology models). The structures and interaction restraints are fed into programs such as HADDOCK [37], which bring the two partners together to create models that are consistent with the experimental interaction data. A strength of HADDOCK is that it can allow for conformational changes in both the backbone and side chains of the two partners, which is a substantial improvement over rigid-body docking protocols. We have used HADDOCK to obtain models of several protein–protein and protein-DNA complexes [38–40], with or without intermolecular NOE data from half-filtered NOESY experiments. As an example of the robustness of the protocol, the interaction between DNA and the DNA-binding domain of the transcription factor MED-1 involved a substantial conformational change of the protein, which formed an additional 10-residue α-helix upon binding DNA. HADDOCK was able to account for this change to create a structural model that was consistent with all NMR and biochemical data [40] (Fig. 4).

Probing protein dynamics

It has long been recognized that proteins and other biological macromolecules are dynamic, displaying motions from the picosecond all the way through to the second timescale (Fig. 5). In many cases, these motions are considered important for biological processes such as catalysis, allosteric regulation, ligand binding and protein folding, and NMR is the most powerful technique for deconvoluting these motions. In all NMR experiments, the behaviour of the excited state of a nucleus is followed as it returns towards its ground state. The NMR signal is lost through both the nuclei returning to their equilibrium energy state (a loss of enthalpy) and by the excited state losing its coherence (or organization, a gain in entropy). These processes are collectively referred to as relaxation, and they are brought about primarily by interactions between the nucleus in question and nearby NMR-active nuclei (e.g. other protons, ^{15}N or ^{13}C nuclei). The strength of these interactions depends strongly on molecular motion. Thus, the time taken for each nucleus to either return to its equilibrium state or to lose its coherence can provide information on the local dynamics of the protein, and such motion can occur on timescales ranging from picoseconds to seconds or more.

Fast motions (ps–ns)

These motions are accessible through the measurement of the relaxation rate constants for individual ^{13}C or ^{15}N heteronuclei in a protein, using HSQC-type experiments. Three different parameters are typically obtained for each nucleus: the R_1 and R_2 relaxation rate constants and the magnitude of the NOE between

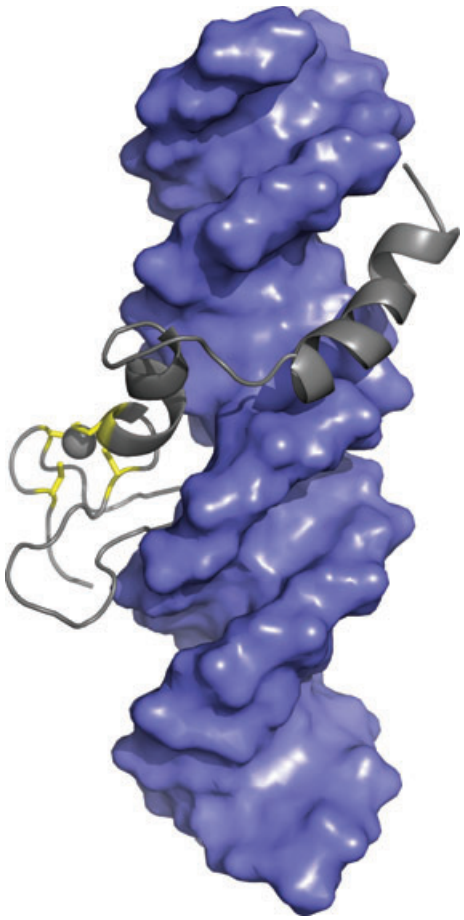

Fig. 4. Model of the complex formed between the GATA-type zinc finger of the transcription factor MED-1 and its DNA target [39]. The protein is shown as a grey ribbon (with zinc-ligating residues in yellow and the zinc ion as a grey sphere) and the DNA is shown as a surface representation. The model was created in HADDOCK using a combination of chemical-shift mapping, mutagenesis and intermolecular NOE data. The helix on the right-hand side of the structure as shown forms only upon DNA binding.

each amide proton and its attached nitrogen. It is not important to know exactly what these parameters measure, but suffice to say that R_1 and the ^{1}H-^{15}N NOE report directly on the existence of motions on the ps to ns timescale, whereas R_2 additionally depends on slower motions on the μs to ms timescale. These data can be analysed [41–43] to separate contributions from internal motion and overall Brownian diffusion. Internal motion on the ps–ns timescale is described by the generalized order parameter S^2, often called an 'entropy meter'. S^2 describes the rigidity of each residue and it can have a value between zero, for a nucleus undergoing completely unrestricted motion, and one, for a nucleus that moves only with the whole molecule. It is typically observed that S^2 is lower for

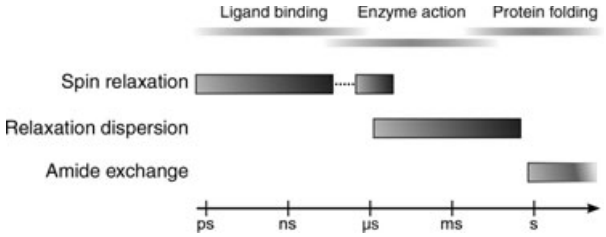

Fig. 5. NMR experiments used to investigate processes occurring on different timescales. Spin-relaxation experiments are used to investigate events that occur between the ps–ns timescale up to the low-μs timescale (fast exchange) [91]. More detailed insights are gained by using relaxation dispersion experiments, which map μs–ms timescale motions. The slowest processes, such as protein folding, range from seconds to days and can be monitored using amide-exchange experiments.

the N- and C-termini of the protein, reflecting their flexibility (Fig. 6).

The correlation between dynamics and complex formation has also been probed using these experiments. Intuitively, one might expect the binding of a ligand to a protein or enzyme to reduce motion in a protein and therefore we would expect to see an increase in S^2; indeed, many studies have observed this trend [44,45]. However, there are also many exceptions [46], arguing that our understanding and ability to predict dynamics is not yet fully developed. For example, the sterile alpha motif domain of VTS1p forms a tight complex with a RNA hairpin target, without any significant structural changes appearing to occur [47,48]. Analysis of ^{15}N relaxation data for the free domain shows a rigid structure with high S^2, whereas a general decrease in S^2 is surprisingly observed upon complex formation [48]. These data suggest that the VTS1p–RNA interaction is driven by an increase in conformational entropy.

Slower motions (μs-ms)

As noted above, R_2 relaxation rates can provide information on slower internal motions in proteins. Related, but more sophisticated, relaxation dispersion measurements developed by Loria *et al.* [49] and Mulder *et al.* [50] provide entry into far more detailed understanding of protein dynamics. This technique allows one to probe the kinetic, thermodynamic and structural parameters that define conformational fluctuations [51,52], even permitting the characterization of almost 'invisible' states that are populated at levels so low (down to a few per cent) that they cannot be directly detected [53,54].

These experiments have provided mechanistic insight into a number of systems. Analysis of the proteins NtrC [55,56], adenylate kinase [57] and Fyn SH3 [58]

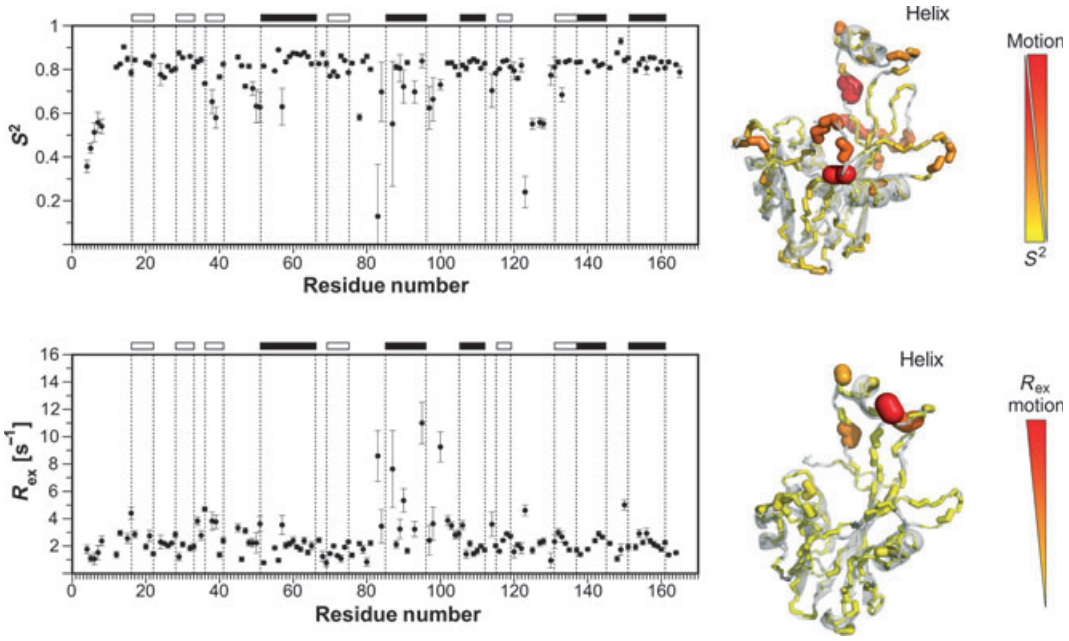

Fig. 6. ^{15}N spin-relaxation data for a small nucleotide binding protein. S^2 is a measure of backbone conformational entropy and it can take values from 0 (disorder) to 1 (rigid). Low S^2 values are typically observed for the N- and C-termini and for loops. R_{ex} describes μs–ms motion, which is often present in loops and observed in situations in which conformational change takes place. In this example, R_{ex} is particularly pronounced in a helix that extends away from the body of the protein. Structures are colour- and width-coded according to increasing motion, such that smaller S^2 or larger R_{ex} values are represented by thicker lines and increasing red colouring.

have demonstrated that ligand binding and enzymatic activity can reflect a shift of a pre-existing equilibrium within an ensemble, rather than the formation of a new structure (induced fit). Analysis of multiple species in the catalytic cycle of dihydrofolate reductase (DHFR), which is important for cell growth and proliferation [59], and a target of both anticancer and antibacterial drugs, has similarly shown that slow timescale motions are important for catalysis [60–62]. These motions again comprise pre-existing equilibria with species that are important at subsequent or preceding steps of the cycle [63]. Importantly, motions of the cofactor-binding site are coupled with those of the substrate-binding site. So how are these data useful? Similar studies of DHFR bound to the anticancer agent methotrexate or the antibiotic trimethorprim show slow-motion dynamics in the drug (or substrate)-binding site that resemble those of the holoenzyme, although much slower [64]. However, in both protein–drug complexes, the slow motions in the cofactor-binding site appear quenched and thus the motions of the two sites have been decoupled. These data show that inhibition is not simply a competition for a binding site, but can involve disruption of motion at a distal site. NMR approaches to understanding protein flexibility may thus offer new opportunities in drug design.

Really slow motion (> ms)

It is straightforward to use hydrogen–deuterium exchange experiments to monitor very slow processes in proteins. A sample of fully protonated protein is dissolved in a buffer made up in 100% deuterated water and ^{15}N-HSQC spectra are acquired at different time intervals [65,66], allowing the measurement of exchange rates for each amide proton. These rates are correlated with the structural stability of each part of a protein and are valuable for understanding the mechanisms underlying protein folding. Similar to the relaxation dispersion measurements described above, amide exchange rates can report on rare conformations [67] that cannot be observed directly and yet must be accessed for exchange to occur. The experiment hinges on the idea that for an amide hydrogen to exchange with deuterium it requires the breaking of all hydrogen bonds in which the amide proton is involved. Because most amides from structured parts of a folded protein are involved in intramolecular hydrogen bonds, this requires an unfolding event to break the hydrogen bond and thereby allow hydrogen–deuterium exchange to occur. These unfolding events are relatively rare and may be local, sub-global (involving a substantial portion of the protein) or global. By following the dependence of amide–proton exchange

on low levels of denaturant, regions that fold independently of the rest of the protein can be discerned [66,68].

ZZ-exchange experiments have also been used to characterize slow exchange processes [69], including ligand binding, enzyme activity and *cis–trans* peptide bond isomerism [70,71]. Resonances from a residue that is involved in a slow exchange process may give rise to two discrete signals and the exchange between the two states can be observed in the experiment, yielding an exchange rate constant.

Hot off the press – new NMR approaches

Out of the tube into the cell

A common source of frustration for structural biologists is the need to convincingly show that the structure that they have solved is biologically relevant, because it could be argued that most NMR and X-ray samples are not representative of *in vivo* conditions. However, the noninvasive nature of NMR enables the acquisition of spectra of isotopically labelled proteins inside cells [72]. 'In-cell' NMR [73] was initially used to study the folding of proteins in the bacterium *Escherichia coli* [74] and to monitor protein–protein interactions [75]. Later studies were conducted using *Xenopus laevis* oocytes [76] and most recently this approach has been extended to human cells [77], allowing changes in the structure of a protein in the cell to be probed by monitoring chemical shifts in a ^{15}N-HSQC spectrum of the protein. Recently, the full 3D structure was determined for a protein in the cytoplasm of *E. coli* (albeit at relatively low resolution); this structure agreed well with the structure determined *in vitro* [78], which is very reassuring after two and a half decades of *in vitro* structure determination using NMR! In-cell NMR is an exciting new avenue for NMR structure determination and this new technology is likely to find wide application in probing protein–drug interactions, protein folding and the in-cell dynamics of macromolecules.

The rise of protein solid-state NMR spectroscopy

Solving the structures of hard-to-crystallize proteins, such as integral membrane proteins, or naturally aggregating proteins, such as fibrillar or amyloid proteins, is a significant challenge, but nevertheless a highly desirable goal considering, for example, the many important functions carried out by membrane proteins and their prevalence as drug targets. In recent years, solid-state NMR approaches have been developed with the goal of obtaining structural and dynamic information from such systems [79–83]. As for solution-state NMR, the development of higher field magnets, improved instrumentation and the production of isotopically enriched samples has driven the development of this cutting-edge field. 'Raw' solid-state NMR spectra are typically very broad due to the effects of several nuclear interactions that are averaged to zero in the free-tumbling solution state. Numerous spectroscopic and hardware advances over the last 20 years (and particularly over the last 5 years) have substantially alleviated these problems, giving rise to extremely high-quality NMR spectra of proteins in the solid state. These advances have led to the determination of 3D structures of several small proteins [84–86], and these reports have ushered in a new era in the analysis of protein structure. More reports of chemical shift assignments made for larger proteins and protein complexes such as the tetrameric integral membrane KcsA potassium channel (70 kDa) [87] and DsbA (21 kDa) [88] demonstrate the potential of this field for obtaining structural and dynamic information of large macromolecular complexes that are not amenable to traditional solution NMR approaches or X-ray crystallography. Amyloid and other fibrillar proteins are similarly revealing their secrets to solid-state NMR approaches. Chemical shift assignments have been made for the Alzheimer's disease-related peptide Aβ(1–40) [89] and the Het-S(218–289) prion protein [90], among others, permitting determination of the conformation of the monomers that make up these otherwise recalcitrant fibrillar structures.

Summary

In this and the accompanying review [1], we have tried to provide an introduction to modern macromolecular NMR spectroscopy that is accessible to all life scientists. We hope to have demonstrated that NMR is a powerful and versatile tool that can provide insight into protein structure and function on many levels, and need not involve too much quantum mechanics (although that is always an option if one is so inclined). We further hope that these reviews provide researchers with the ability to interpret and critique the existing NMR literature as well as inspiration for scientists to make use of NMR spectroscopy to complement their own research. Chances are that your local NMR spectroscopist will be more than happy to help!

Acknowledgements

The authors acknowledge financial support from the Discovery grants DP0774245, DP0879065, DP1095728

and DP110103161 from the Australian Research Council. MB is a recipient of a Swiss National Science Foundation fellowship.

References

1 Kwan AH, Mobli M, Gooley PR, King GF & MacKay JP (2011) Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J* **278**, 687–703.

2 Negendank W (1992) Studies of human tumors by MRS – a review. *NMR Biomed* **5**, 303–324.

3 van Beek JD, Hess S, Vollrath F & Meier BH (2002) The molecular structure of spider dragline silk: folding and orientation of the protein backbone. *Proc Natl Acad Sci USA* **99**, 10266–10271.

4 Kogel-Knabner I (1997) C-13 and N-15 NMR spectroscopy as a tool in soil organic matter studies. *Geoderma* **80**, 243–270.

5 Dyson HJ & Wright PE (2004) Unfolded proteins and protein folding studied by NMR. *Chem Rev* **104**, 3607–3622.

6 Hiller S, Garces RG, Malia TJ, Orekhov VY, Colombini M & Wagner G (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* **321**, 1206–1210.

7 Hopkins EJ, Layfield S, Ferraro T, Bathgate RAD & Gooley PR (2007) The NMR solution structure of the relaxin (RXFP1) receptor lipoprotein receptor class A module and identification of key residues in the N-terminal region of the module that mediate receptor activation. *J Biol Chem* **282**, 4172–4184.

8 van Nuland NA, Dobson CM & Regan L (2008) Characterization of folding the four-helix bundle protein Rop by real-time NMR. *Protein Eng Des Sel* **21**, 165–170.

9 Schanda P, Forge V & Brutscher B (2006) HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magn Reson Chem* **44**, S177–184.

10 Schanda P, Forge V & Brutscher B (2007) Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc Natl Acad Sci USA* **104**, 11257–11262.

11 Corazza A, Rennella E, Schanda P, Mimmi MC, Cutuil T, Raimondi S, Giorgetti S, Fogolari F, Viglino P, Frydman L *et al.* (2010) Native-unlike long-lived intermediates along the folding pathway of the amyloidogenic protein beta2-microglobulin revealed by real-time two-dimensional NMR. *J Biol Chem* **285**, 5827–5835.

12 Roberts GCK (2003) The determination of equilibrium dissociation constants of protein–ligand complexes by NMR. In *BioNMR in Drug Research* (Zerbe O ed), pp 309–319. Wiley-VCH Verlag, GmbH & Co. KGaA, Weinheim, FRG.

13 Liu Y-N, Lai Y-T, Chou W-I, Chang MD-T & Lyu P-C (2007) Solution structure of family 21 carbohydrate-binding module from *Rhizopus oryzae* glucoamylase. *Biochem J* **403**, 21–30.

14 Simpson PJ, Xie H, Bolam DN, Gilbert HJ & Williamson MP (2000) The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. *J Biol Chem* **275**, 41137–41142.

15 Radhakrishnan I, Perez-Alvarado GC, Parker D, Dyson HJ, Montminy MR & Wright PE (1999) Structural analyses of CREB–CBP transcriptional activator–coactivator complexes by NMR spectroscopy: implications for mapping the boundaries of structural domains. *J Mol Biol* **287**, 859–865.

16 Velyvis A, Vaynberg J, Yang YW, Vinogradova O, Zhang YJ, Wu CY & Qin J (2003) Structural and functional insights into PINCH LIM4 domain-mediated integrin signaling. *Nat Struct Biol* **10**, 558–564.

17 Croy JE, Altschuler SE, Grimm NE & Wuttke DS (2009) Nonadditivity in the recognition of single-stranded DNA by the *Schizosaccharomyces pombe* protection of telomeres 1 DNA-binding domain, Pot1-DBD. *Biochemistry* **48**, 6864–6875.

18 Sette M, Spurio R, Trotta E, Brandizi C, Brandi A, Pon CL, Barbato G, Boelens R & Gualerzi CO (2009) Sequence-specific recognition of DNA by the C-terminal domain of nucleoid-associated protein H-NS. *J Biol Chem* **284**, 30453–30462.

19 Lingel A, Simon B, Izaurralde E & Sattler M (2003) Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. *Nature* **426**, 465–469.

20 Fletcher JI, Swarbrick JD, Maksel D, Gayler KR & Gooley PR (2002) The structure of Ap(4)A hydrolase complexed with ATP–MgFx reveals the basis of substrate binding. *Structure* **10**, 205–213.

21 Branson KM, Mertens HD, Swarbrick JD, Fletcher JI, Kedzierski L, Gayler KR, Gooley PR & Smith BJ (2009) Discovery of inhibitors of lupin diadenosine 5′,5″-P(1),P(4)-tetraphosphate hydrolase by virtual screening. *Biochemistry* **48**, 7614–7620.

22 de Oliveira JF, Sforca ML, Blumenschein TMA, Goldfeder MB, Guimaraes BG, Oliveira CC, Zanchin NIT & Zeri A-C (2010) Structure, dynamics, and RNA interaction analysis of the human SBDS protein. *J Mol Biol* **396**, 1053–1069.

23 Sugase K, Dyson HJ & Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025.

24 Meyer B & Peters T (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed* **42**, 864–890.

25 Pellecchia M, Bertini I, Cowburn D, Dalvit C, Giralt E, Jahnke W, James TL, Homans SW, Kessler H, Luchinat C *et al.* (2008) Perspectives on NMR in drug

discovery: a technique comes of age. *Nat Rev Drug Discov* **7**, 738–745.
26 Mayer M & Meyer B (2001) Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J Am Chem Soc* **123**, 6108–6117.
27 Assadi-Porter FM, Tonelli M, Maillet E, Hallenga K, Benard O, Max M & Markley JL (2008) Direct NMR detection of the binding of functional ligands to the human sweet receptor, a heterodimeric family 3 GPCR. *J Am Chem Soc* **130**, 7212–7213.
28 Rademacher C & Peters T (2008) Molecular recognition of ligands by native viruses and virus-like particles as studied by NMR experiments. In *Bioactive Conformation II*, Topics in Current Chemistry 273, pp. 183–202. Springer-Verlag, Berlin.
29 Claasen B, Axmann M, Meinecke R & Meyer B (2005) Direct observation of ligand binding to membrane proteins in living cells by a saturation transfer double difference (STDD) NMR spectroscopy method shows a significantly higher affinity of integrin alpha(IIb)beta(3) in native platelets than in liposomes. *J Am Chem Soc* **127**, 916–919.
30 Takahashi H, Nakanishi T, Kami K, Arata Y & Shimada I (2000) A novel NMR method for determining the interfaces of large protein–protein complexes. *Nat Struct Biol* **7**, 220–223.
31 Otting G & Wuthrich K (1990) Heteronuclear filters in 2-dimensional [H-1, H-1] NMR-spectroscopy – combined use with isotope labeling for studies of macromolecular conformation and intermolecular interactions. *Q Rev Biophys* **23**, 39–96.
32 Breeze AL (2000) Isotope-filtered NMR methods for the study of biomolecular structure and interactions. *Prog Nucl Magnet Reson Spectrosc* **36**, 323–372.
33 Zhang NX, Wang QH, Ehlinger A, Randles L, Lary JW, Kang Y, Haririnia A, Storaska AJ, Cole JL, Fushman D *et al.* (2009) Structure of the S5a:K48-linked diubiquitin complex and its interactions with Rpn13. *Mol Cell* **35**, 280–290.
34 Banci L, Bertini I, Cantini F, Inagaki S, Migliardi M & Rosato A (2010) The binding mode of ATP revealed by the solution structure of the N-domain of human ATP7A. *J Biol Chem* **285**, 2537–2544.
35 King GF, Shih YL, Maciejewski MW, Bains NPS, Pan BL, Rowland SL, Mullen GP & Rothfield LI (2000) Structural basis topological specificity function of MinE. *Nat Struct Biol* **7**, 1013–1017.
36 Vaynberg J, Fukuda T, Chen K, Vinogradova O, Velyvis A, Tu Y, Ng L, Wu C & Qin J (2005) Structure of an ultraweak protein–protein complex and its crucial role in regulation of cell morphology and motility. *Mol Cell* **17**, 513–523.
37 Dominguez C, Boelens R & Bonvin AM (2003) HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125**, 1731–1737.
38 Gamsjaeger R, Swanton MK, Kobus FJ, Lehtomaki E, Lowry JA, Kwan AH, Matthews JM & Mackay JP (2008) Structural and biophysical analysis of the DNA binding properties of myelin transcription factor 1. *J Biol Chem* **283**, 5158–5167.
39 Liew CK, Simpson RJ, Kwan AH, Crofts LA, Loughlin FE, Matthews JM, Crossley M & Mackay JP (2005) Zinc fingers as protein recognition motifs: structural basis for the GATA-1/friend of GATA interaction. *Proc Natl Acad Sci USA* **102**, 583–588.
40 Lowry JA, Gamsjaeger R, Thong SY, Hung W, Kwan AH, Broitman-Maduro G, Matthews JM, Maduro M & Mackay JP (2009) Structural analysis of MED-1 reveals unexpected diversity in the mechanism of DNA recognition by GATA-type zinc finger domains. *J Biol Chem* **284**, 5827–5835.
41 Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC & Gronenborn AM (1990) Deviations from the simple 2-parameter model-free approach to the interpretation of N-15 nuclear magnetic-relaxation of proteins. *J Am Chem Soc* **112**, 4989–4991.
42 Lipari G & Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic-relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* **104**, 4546–4559.
43 Lipari G & Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic-relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc* **104**, 4559–4570.
44 Eletsky A, Kienhofer A, Hilvert D & Pervushin K (2005) Investigation of ligand binding and protein dynamics in *Bacillus subtilis* chorismate mutase by transverse relaxation optimized spectroscopy–nuclear magnetic resonance. *Biochemistry* **44**, 6788–6799.
45 Shapiro YE, Kahana E, Tugarinov V, Liang ZC, Freed JH & Meirovitch E (2002) Domain flexibility in ligand-free and inhibitor-bound *Escherichia coli* adenylate kinase based on a mode-coupling analysis of N-15 spin relaxation. *Biochemistry* **41**, 6271–6281.
46 Yun SG, Jang DS, Kim DH, Choi KY & Lee HC (2001) N-15 NMR relaxation studies of backbone dynamics in free and steroid-bound Delta(5)-3-ketosteroid isomerase from *Pseudomonas testosteroni*. *Biochemistry* **40**, 3967–3973.
47 Edwards TA, Butterwick JA, Zeng L, Gupta YK, Wang X, Wharton RP, Palmer AG & Aggarwal AK (2006) Solution structure of the Vts1 SAM domain in the presence of RNA. *J Mol Biol* **356**, 1065–1072.
48 Ravindranathan S, Oberstrass FC & Allain FHT (2010) Increase in backbone mobility of the VTS1p-SAM domain on binding to SRE-RNA. *J Mol Biol* **396**, 732–746.

49 Loria JP, Rance M & Palmer AG III (1999) A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J Biomol NMR* **15**, 151–155.

50 Mulder FA, Skrynnikov NR, Hon B, Dahlquist FW & Kay LE (2001) Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J Am Chem Soc* **123**, 967–975.

51 Auer R, Hansen DF, Neudecker P, Korzhnev DM, Muhandiram DR, Konrat R & Kay LE (2010) Measurement of signs of chemical shift differences between ground and excited protein states: a comparison between H(S/M)QC and R1rho methods. *J Biomol NMR* **46**, 205–216.

52 Jee J, Ishima R & Gronenborn AM (2008) Characterization of specific protein association by ^{15}N CPMG relaxation dispersion NMR: the GB1(A34F) monomer–dimer equilibrium. *J Phys Chem B* **112**, 6008–6012.

53 Hansen DF, Vallurupalli P & Kay LE (2008) Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J Biomol NMR* **41**, 113–120.

54 Vallurupalli P, Hansen DF & Kay LE (2008) Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proc Natl Acad Sci USA* **105**, 11766–11771.

55 Gardino AK, Villali J, Kivenson A, Lei M, Liu CF, Steindel P, Eisenmesser EZ, Labeikovsky W, Wolf-Watz M, Clarkson MW *et al.* (2009) Transient non-native hydrogen bonds promote activation of a signaling protein. *Cell* **139**, 1109–1118.

56 Volkman BF, Lipson D, Wemmer DE & Kern D (2001) Two-state allosteric behavior in a single-domain signaling protein. *Science* **291**, 2429–2433.

57 Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M *et al.* (2007) Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838–844.

58 Korzhnev DM, Salvatella X, Vendruscolo M, Di Nardo AA, Davidson AR, Dobson CM & Kay LE (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586–590.

59 Schnell JR, Dyson HJ & Wright PE (2004) Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu Rev Biophys Biomol Struct* **33**, 119–140.

60 McElheny D, Schnell JR, Lansing JC, Dyson HJ & Wright PE (2005) Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. *Proc Natl Acad Sci USA* **102**, 5032–5037.

61 Osborne MJ, Schnell J, Benkovic SJ, Dyson HJ & Wright PE (2001) Backbone dynamics in dihydrofolate reductase complexes: role of loop flexibility in the catalytic mechanism. *Biochemistry* **40**, 9846–9859.

62 Schnell JR, Dyson HJ & Wright PE (2004) Effect of cofactor binding and loop conformation on side chain methyl dynamics in dihydrofolate reductase. *Biochemistry* **43**, 374–383.

63 Boehr DD, McElheny D, Dyson HJ & Wright PE (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642.

64 Mauldin RV, Carroll MJ & Lee AL (2009) Dynamic dysfunction in dihydrofolate reductase results from antifolate drug binding: modulation of dynamics within a structural state. *Structure* **17**, 386–394.

65 Bai Y, Sosnick TR, Mayne L & Englander SW (1995) Protein folding intermediates: native-state hydrogen exchange. *Science* **269**, 192–197.

66 Bedard S, Mayne LC, Peterson RW, Wand AJ & Englander SW (2008) The foldon substructure of staphylococcal nuclease. *J Mol Biol* **376**, 1142–1154.

67 Vadrevu R, Wu Y & Matthews CR (2008) NMR analysis of partially folded states and persistent structure in the alpha subunit of tryptophan synthase: implications for the equilibrium folding mechanism of a 29-kDa TIM barrel protein. *J Mol Biol* **377**, 294–306.

68 Michel LV & Bren KL (2008) Submolecular unfolding units of *Pseudomonas aeruginosa* cytochrome c-551. *J Biol Inorg Chem* **13**, 837–845.

69 Li Y & Palmer AG III (2009) TROSY-selected ZZ-exchange experiment for characterizing slow chemical exchange in large proteins. *J Biomol NMR* **45**, 357–360.

70 Bosco DA, Eisenmesser EZ, Pochapsky S, Sundquist WI & Kern D (2002) Catalysis of *cis/trans* isomerization in native HIV-1 capsid by human cyclophilin A. *Proc Natl Acad Sci USA* **99**, 5247–5252.

71 Demers JP & Mittermaier A (2009) Binding mechanism of an SH3 domain studied by NMR and ITC. *J Am Chem Soc* **131**, 4355–4367.

72 Ito Y & Selenko P (2010) Cellular structural biology. *Curr Opin Struct Biol* **20**, 640–648.

73 Serber Z & Dotsch V (2001) In-cell NMR spectroscopy. *Biochemistry* **40**, 14317–14323.

74 Dedmon MM, Patel CN, Young GB & Pielak GJ (2002) FlgM gains structure in living cells. *Proc Natl Acad Sci USA* **99**, 12681–12684.

75 Burz DS, Dutta K, Cowburn D & Shekhtman A (2006) Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat Methods* **3**, 91–93.

76 Sakai T, Tochio H, Tenno T, Ito Y, Kokubo T, Hiroaki H & Shirakawa M (2006) In-cell NMR spectroscopy of proteins inside *Xenopus laevis* oocytes. *J Biomol NMR* **36**, 179–188.

77 Inomata K, Ohno A, Tochio H, Isogai S, Tenno T, Nakase I, Takeuchi T, Futaki S, Ito Y, Hiroaki H *et al.*

(2009) High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458**, 106–111.
78 Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Walchli M *et al.* (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458**, 102–110.
79 Ader C, Frey S, Maas W, Schmidt HB, Gorlich D & Baldus M (2010) Amyloid-like interactions within nucleoporin FG hydrogels. *Proc Natl Acad Sci USA* **107**, 6281–6285.
80 Cady SD, Schmidt-Rohr K, Wang J, Soto CS, DeGrado WF & Hong M (2010) Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* **463**, 689–692.
81 Hiller M, Krabben L, Vinothkumar KR, Castellani F, van Rossum BJ, Kuhlbrandt W & Oschkinat H (2005) Solid-state magic-angle spinning NMR of outer-membrane protein G from *Escherichia coli*. *Chembiochem* **6**, 1679–1684.
82 Shewmaker F, Wickner RB & Tycko R (2006) Amyloid of the prion domain of Sup35p has an in-register parallel beta-sheet structure. *Proc Natl Acad Sci USA* **103**, 19754–19759.
83 Wasmer C, Lange A, Van Melckebeke H, Siemer AB, Riek R & Meier BH (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* **319**, 1523–1526.
84 Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K & Oschkinat H (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 98–102.
85 Franks WT, Wylie BJ, Schmidt HLF, Nieuwkoop AJ, Mayrhofer RM, Shah GJ, Graesser DT & Rienstra CM (2008) Dipole tensor-based atomic-resolution structure determination of a nanocrystalline protein by solid-state NMR. *Proc Natl Acad Sci USA* **105**, 4621–4626.
86 Manolikas T, Herrmann T & Meier BH (2008) Protein structure determination from C-13 spin-diffusion solid-state NMR spectroscopy. *J Am Chem Soc* **130**, 3959–3966.
87 Varga K, Tian L & McDermott AE (2007) Solid-state NMR study and assignments of the KcsA potassium ion channel of *S. lividans*. *Biochim Biophys Acta* **1774**, 1604–1613.
88 Sperling LJ, Berthold DA, Sasser TL, Jeisy-Scott V & Rienstra CM (2010) Assignment strategies for large proteins by magic-angle spinning NMR: the 21-kDa disulfide-bond-forming enzyme DsbA. *J Mol Biol* **399**, 268–282.
89 Petkova AT, Ishii Y, Balbach JJ, Antzutkin ON, Leapman RD, Delaglio F & Tycko R (2002) A structural model for Alzheimer's beta-amyloid fibrils based on experimental constraints from solid state NMR. *Proc Natl Acad Sci USA* **99**, 16742–16747.
90 Ritter C, Maddelein ML, Siemer AB, Luhrs T, Ernst M, Meier BH, Saupe SJ & Riek R (2005) Correlation of structural elements and infectivity of the HET-s prion. *Nature* **435**, 844–848.
91 Korzhnev DM, Billeter M, Arseniev AS & Orekhov VY (2001) NMR studies of Brownian tumbling and internal motions in proteins. *Prog Nucl Magn Reson Spectrosc* **38**, 197–266.