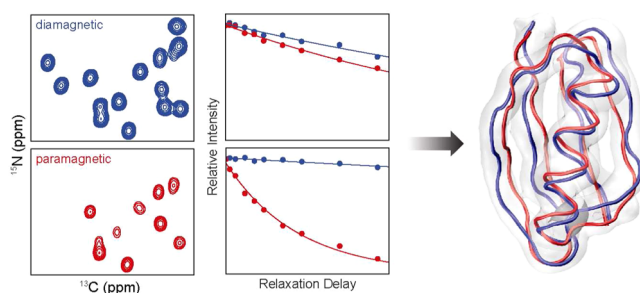# Protein Structure Determination with Paramagnetic Solid-State NMR Spectroscopy

ISHITA SENGUPTA, PHILIPPE S. NADAUD, AND
CHRISTOPHER P. JARONIEC*

*Department of Chemistry and Biochemistry, The Ohio State University,
Columbus, Ohio 43210, United States*

RECEIVED ON DECEMBER 30, 2012

CONSPECTUS

Many structures of the proteins and protein assemblies that play central roles in fundamental biological processes and disease pathogenesis are not readily accessible via the conventional techniques of single-crystal X-ray diffraction and solution-state nuclear magnetic resonance (NMR). On the other hand, many of these challenging biological systems are suitable targets for atomic-level structural and dynamic analysis by magic-angle spinning (MAS) solid-state NMR spectroscopy, a technique that has far less stringent limitations on the molecular size and crystalline state.

Over the past decade, major advances in instrumentation and methodology have prompted rapid growth in the field of biological solid-state NMR. However, despite this progress, one challenge for the elucidation of three-dimensional (3D) protein structures via conventional MAS NMR methods is the relative lack of long-distance data. Specifically, extracting unambiguous interatomic distance restraints larger than $\sim$5 Å from through-space magnetic dipole–dipole couplings among the protein ^1H, ^{13}C, and ^{15}N nuclei has proven to be a considerable challenge for researchers. It is possible to circumvent this problem by extending the structural studies to include several analogs of the protein of interest, intentionally modified to contain covalently attached paramagnetic tags at selected sites. In these paramagnetic proteins, the hyperfine couplings between the nuclei and unpaired electrons can manifest themselves in NMR spectra in the form of relaxation enhancements of the nuclear spins that depend on the electron–nucleus distance. These effects can be significant for nuclei located up to $\sim$20 Å away from the paramagnetic center.

In this Account, we discuss MAS NMR structural studies of nitroxide and EDTA-Cu^{2+} labeled variants of a model 56 amino acid globular protein, B1 immunoglobulin-binding domain of protein G (GB1), in the microcrystalline solid phase. We used a set of six EDTA-Cu^{2+}-tagged GB1 mutants to rapidly determine the global protein fold in a *de novo* fashion. Remarkably, these studies required quantitative measurements of only approximately four or five backbone amide ^{15}N longitudinal paramagnetic relaxation enhancements per residue, in the complete absence of the usual internuclear distance restraints. Importantly, this paramagnetic solid-state NMR methodology is general and can be directly applied to larger proteins and protein complexes for which a significant fraction of the signals can be assigned in standard 2D and 3D MAS NMR chemical shift correlation spectra.

Introduction

Many proteins implicated in fundamental biological processes or development of disease are an integral part of high molecular weight complexes that contain other proteins, nucleic acids, carbohydrates, or lipids. Notable examples of such macromolecular assemblies include membrane proteins,

amyloids, microtubules, and chromatin. While biological systems of this type frequently defy analysis using the classical tools of structural biology (X-ray crystallography and solution-state nuclear magnetic resonance (NMR) spectroscopy), they can be probed with atomic-resolution detail by magic-angle spinning (MAS) solid-state NMR, which suffers from fewer

limitations on molecular size and crystalline state. Historically, structural and dynamic studies of biological solids by solid-state NMR have employed tailored one- or two-dimensional (1D or 2D) radio-frequency (RF) pulse schemes in combination with samples containing isolated NMR-active nuclei, such as ^{13}C, ^{15}N, ^{2}H, or ^{31}P, located at a few specific sites.[1,2] Over the past decade, however, thanks to major strides in high-field magnet (≥ 20 T) and probe technologies, multidimensional RF pulse sequences, and methods for preparation of homogeneous samples, MAS solid-state NMR has been utilized with great success to yield high-resolution structural models of peptides and proteins in microcrystalline, fibrous, or membrane-bound states, as well as insights into backbone and side-chain conformational dynamics.[3−10] For a discussion of structural studies of proteins using complementary static solid-state NMR methods, the reader is referred to recent reviews.[11,12]

The remarkable recent progress notwithstanding, the elucidation of three-dimensional protein structures by conventional MAS solid-state NMR techniques continues to be a formidable undertaking. More specifically, this methodology generally relies on the availability of numerous site-resolved structural restraints in the form of distance-dependent magnetic dipole−dipole couplings among the ^{13}C, ^{15}N, and ^{1}H nuclei. Given the magnitudes of the relevant nuclear gyromagnetic ratios and the inverse third power dependence of the dipolar coupling constants on internuclear separation, the couplings of interest become exceedingly small (typically on the order of few to tens of hertz) for interatomic distances in the range of $\sim$5−10 Å and above as shown in Figure 1. Consequently, the detection and quantification by the established dipolar recoupling techniques of multiple ^{13}C−^{15}N, ^{13}C−^{13}C, and ^{1}H−^{1}H couplings corresponding to interatomic distances longer than $\sim$5 Å in highly ^{13}C,^{15}N-enriched proteins[13−17] constitutes a major challenge due to the presence of many stronger couplings associated with shorter, structurally less-interesting contacts as well as nuclear spin relaxation phenomena. While this limit can be pushed to $\sim$10 Å by detecting ^{1}H−^{1}H contacts in proteins prepared with sparse methyl and amide proton labeling[18,19] or couplings between ^{19}F and ^{13}C or ^{15}N nuclei in proteins containing a ^{19}F-labeled amino acid at a specific position,[20] the majority of MAS NMR structural studies that employ the usual dipolar-coupling-based methods nevertheless suffer from a relative shortage of critical unambiguous long-range distances that constrain the three-dimensional protein fold.
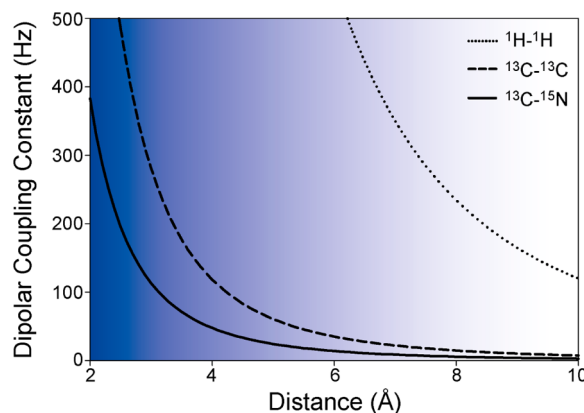

FIGURE 1. Dependence of ^{1}H−^{1}H, ^{13}C−^{13}C, and ^{13}C−^{15}N dipolar coupling constant magnitudes on the internuclear distance. The variable shading of the plot area serves as a rough representation of the relative difficulty with which a particular distance can be accessed by using conventional MAS NMR techniques; the lighter the shading, the more difficult the dipolar couplings corresponding to that distance are to detect and quantify.

Fundamentals of Paramagnetic Solid-State NMR Spectroscopy

The quantity of long-range distance restraints can be dramatically boosted by including in the NMR study one or more isostructural paramagnetic variants of the protein of interest. Indeed, paramagnetic NMR spectroscopy has been broadly utilized over the past several decades to investigate the structures of soluble proteins,[21−23] and, much more recently, analogous MAS based approaches have also been demonstrated for protein molecules in the solid phase.[23−28] In metalloproteins, which possess an intrinsic high-affinity metal binding site, a suitable paramagnetic center may already be present in the wild-type form of the protein or otherwise may be introduced through metal exchange. On the other hand for natively diamagnetic proteins, which are the central focus of this Account, a variety of covalent chemical tags containing stable radicals or capable of binding paramagnetic transition metal or lanthanide ions[29] may be incorporated at specific locations, most commonly via cysteine site-directed mutagenesis followed by thiol-disulfide chemistry. Figure 2 shows two representative tags of this type, which are also particularly relevant to the studies discussed here.

Unpaired electrons, which possess magnetic moments that exceed nuclear magnetic moments by several orders of magnitude, couple strongly to the surrounding nuclei through the hyperfine interaction. These couplings can appear in NMR spectra in the form of paramagnetic shifts or relaxation enhancements of the nuclear resonances, both of which have well-defined dependencies on the distance
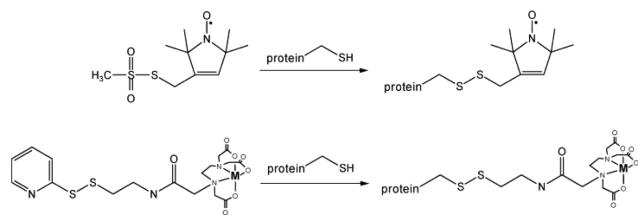
FIGURE 2. Incorporation of paramagnetic nitroxide spin label (top) and EDTA−metal tags (bottom) into proteins using a thiol-specific reaction with a cysteine side-chain introduced into the protein by using site-directed mutagenesis. The EDTA moiety strongly chelates various transition metal ions (M) including Cu^{2+} and Mn^{2+}.
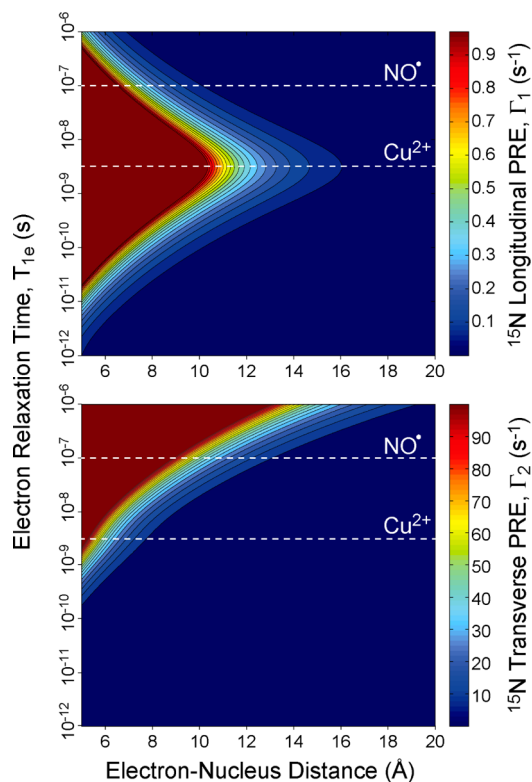

FIGURE 3. Longitudinal (top) and transverse (bottom) paramagnetic relaxation enhancements calculated using eq 1 as a function of the electron−nucleus distance and the electron spin relaxation time for a ^{15}N nucleus and a $S = 1/2$ paramagnetic center in a static magnetic field of 11.7 T (500 MHz ^1H frequency). The dashed white lines indicate the T_{1e} values typical for a nitroxide spin label and a Cu^{2+} center.

between the nucleus and the paramagnetic center.[21−25] Most importantly, these effects can be considerable for electron−nucleus separations of up to ∼20 Å or beyond, which exceeds by at least 3−4-fold the length scale of structural restraints accessible via traditional NMR methods. In this Account, we will concentrate mainly on the use of nuclear paramagnetic relaxation enhancements (PREs), which are the most significant paramagnetic effect for nitroxide radicals, Cu^{2+} ions, and other paramagnetic centers characterized by isotropic or weakly anisotropic magnetic susceptibility tensors, in the structural studies of proteins by MAS solid-state NMR, but note that paramagnetic pseudo-contact shifts are also a valuable source of long-range restraints on biomolecular structure.[26,30−34] In the context of such studies, Solomon dipolar relaxation,[35] which arises from the fluctuations of local magnetic fields generated by the paramagnetic center and experienced by the nuclear spins, is the primary mechanism responsible for causing longitudinal (Γ_1) and transverse (Γ_2 or $\Gamma_{1\rho}$) nuclear PREs. Assuming that in the solid phase the electron−nucleus couplings are modulated largely by rapid electron spin relaxation, with negligible contributions from overall molecular tumbling and slow chemical exchange type motions, the longitudinal and transverse nuclear PREs can be approximated as follows:[21,24,25]

$$\Gamma_1 \approx \frac{2C}{r^6}\left(\frac{3T_{1e}}{1+\omega_I^2 T_{1e}^2} + \frac{7T_{1e}}{1+\omega_e^2 T_{1e}^2}\right)$$
$$\Gamma_2 \approx \Gamma_{1\rho} \approx \frac{C}{r^6}\left(4T_{1e} + \frac{3T_{1e}}{1+\omega_I^2 T_{1e}^2} + \frac{13T_{1e}}{1+\omega_e^2 T_{1e}^2}\right) \quad (1)$$

where r is the electron−nucleus distance, T_{1e} is the relaxation time constant for the paramagnetic center, ω_I and ω_e are the nuclear and electron Larmor frequencies, respectively, and C is a prefactor that depends on fundamental constants, the nuclear gyromagnetic ratio, and the spin quantum number (S) for the paramagnetic center. The variation of longitudinal and transverse PREs with the electron−nucleus distance and electron relaxation time is illustrated in Figure 3 for the case of a ^{15}N nucleus interacting with a $S = 1/2$ paramagnetic center (e.g., nitroxide radical or Cu^{2+}) at 11.7 T. Notably, these plots show that the magnitudes of longitudinal and transverse nuclear PREs strongly depend on the identities of both the nuclear spin and the paramagnetic center, as well as the external magnetic field, and that under typical experimental conditions one or both of these PRE effects can be sizable for nuclei located ∼10−20 Å or farther away from the paramagnetic center. Moreover, given that reasonable relaxation time estimates are available in the literature for various paramagnetic species[21] (if necessary these values can be further refined during the course of the protein structure calculation), the experimentally determined PRE values can in principle be directly translated into electron−nucleus distances in a quantitative manner.
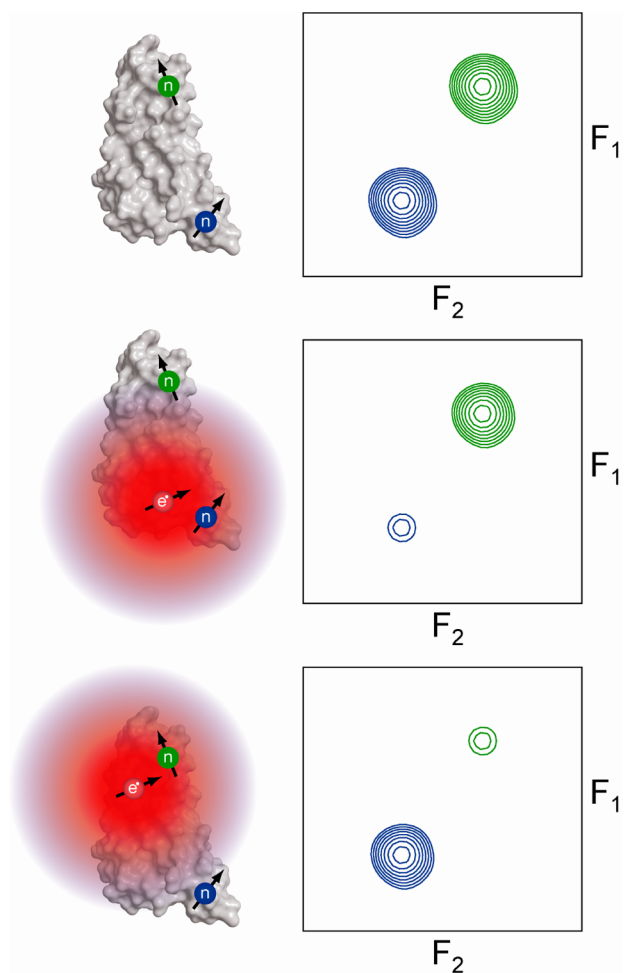
FIGURE 4. Schematic 2D NMR spectra for a diamagnetic protein (top) and two of its variants containing covalent paramagnetic tags (middle and bottom), illustrating site-resolved measurements of nuclear PREs by monitoring the intensities of the nuclear spin signals (green and blue contours) based on their proximity to the paramagnetic center (red sphere). The magnitude of the PRE effect is represented by the large sphere around the electron spin and decreases from red to blue.

In Figure 4, we illustrate the basic concept for how these nuclear PRE phenomena can be harnessed to yield a large number of site-resolved distance restraints for a natively diamagnetic protein. Namely, this is attainable by recording of a set of conventional 2D NMR chemical shift correlation spectra (e.g., $^{15}N-^{13}C$ or $^{13}C-^{13}C$, or analogous 3D data sets) for a reference diamagnetic protein and several of its variants containing covalently linked paramagnetic tags at different positions. The PREs are then extracted by monitoring individual cross-peak amplitudes in spectra that contain correlations for directly bonded backbone or side-chain nuclei and are similar to those normally used for resonance assignment purposes. In the schematic 2D NMR spectra in Figure 4, we consider signals for two nuclei of a particular

type (e.g., ^{15}N), located in different regions of the protein and colored accordingly in green and blue. While for the diamagnetic protein sample both cross-peaks have similar amplitudes and linewidths, their relative intensities differ substantially in the spectra for the two paramagnetic analogs based on the proximity of each nucleus to the unpaired electron spin.

Qualitative PRE Measurements in Spin-Labeled Proteins

Nitroxide radicals, which in the context of NMR experiments on hydrated proteins at room temperature are characterized by relatively long electron spin relaxation times (on the order of ∼100 ns),[21,36] are expected to generate significant transverse PREs for the neighboring nuclear spins. Indeed, these PRE effects were successfully observed in a recent MAS NMR study of uniformly $^{13}C,^{15}N$-enriched analogs of the model 56-residue B1 immunoglobulin binding domain of protein G (GB1) in the microcrystalline solid phase, containing nitroxide spin labels at two solvent-exposed surface residues.[37] The key results of that study are summarized in Figure 5, which shows a comparison of 2D $^{15}N-^{13}C\alpha$ chemical shift correlation spectra for the spin-labeled T53C mutant of GB1 and a reference diamagnetic protein tagged at the same residue with an analog of the nitroxide radical containing an acetyl group in place of the paramagnetic oxyl moiety. The foremost feature of these spectra is that a number of signals present for the control sample are severely attenuated or altogether missing in the spectrum of the spin-labeled protein. The most strongly affected resonances are invariably associated with residues found in strands $\beta 1-\beta 4$ and the intervening loops, predicted to be in the spatial proximity (within ∼10–12 Å) of the radical, based on the structural model of GB1. In contrast, the least affected residues are located in the α-helix and removed from the spin label by ∼15–20 Å or more.[37] These variations in signal intensities are generally consistent with the PRE values calculated for ^1H, ^{13}C, and ^{15}N nuclei. Specifically, these calculations predict that cross-peaks for amino acids located within ∼10 Å of the nitroxide radical should be effectively fully suppressed, owing mainly to the large transverse PRE for the amide proton that leads to rapid coherence decay during the initial $^1H-^{15}N$ cross-polarization period of the 2D experiment (e.g., the transverse magnetization for a ^1H nucleus located 5 Å away from the radical is expected to decay completely within a period of only ∼20 μs, while for a 10 Å electron–^1H distance the magnetization decays to 10% of its initial value in ∼500 μs). It is important to note here that for the studies described in this section
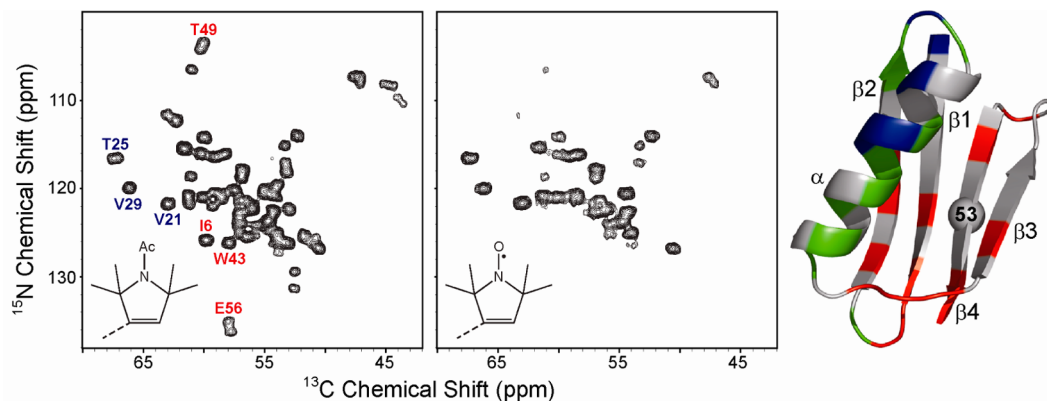
FIGURE 5. Two-dimensional $^{15}N-^{13}C\alpha$ MAS solid-state NMR spectra of nitroxide spin-labeled T53C mutant of GB1 (middle) and its diamagnetic analog (left). A few selected resonance assignments for residues proximal and distal to the tag are labeled in red and blue font, respectively. The tertiary structure of GB1 (right) showing the tag location as a gray sphere and the relative intensity of cross-peaks in spectra of the paramagnetic and diamagnetic variants color coded as follows: strongly attenuated (red, relative intensity 0−0.33), moderately attenuated (green, relative intensity 0.33−0.66), and least attenuated (blue, relative intensity 0.66−1). Figure adapted from ref 37.

(and other analogous experiments aimed at the detection of intramolecular electron−nucleus contacts discussed later in this Account), the $^{13}C,^{15}N$-enriched paramagnetic proteins are routinely diluted within the microcrystalline lattice to a mole fraction of ∼25% or less by coprecipitation with a diamagnetic, natural abundance protein in order to minimize the effects of intermolecular PREs. Furthermore, comparisons of chemical shifts for the backbone ^{15}N and ^{13}C atoms are used to confirm that the different mutants exhibit the same overall fold as the wild-type protein.

While somewhat qualitative in nature, MAS solid-state NMR experiments of proteins modified with spin label tags are capable of yielding vital information about intramolecular protein architecture or intermolecular protein−protein interactions on length scales that elude conventional techniques. Indeed, this spin-labeling methodology has very recently been employed by Ladizhansky and co-workers to probe long-range intermolecular contacts and determine the oligomerization interface for a membrane protein in lipid bilayers.[38]

Quantitative PRE Measurements in EDTA−Cu^{2+}-Tagged Proteins

Although clearly beneficial for certain types of applications, the use of nitroxides as relaxation agents in the context of high-resolution structural analysis of uniformly $^{13}C,^{15}N$-labeled proteins by solid-state NMR suffers from a major drawback. Namely, the presence of large transverse PREs, which lead to severely attenuated signal intensities for numerous residues, precludes quantitative PRE and electron−nucleus distance measurements for those residues. As suggested by the results of PRE calculations shown in Figure 3,

this problem may be overcome by using tags containing more rapidly relaxing paramagnetic centers, and Cu^{2+} ions appear to be especially promising in this regard. Specifically, in analogy to spin labels, Cu^{2+} centers have $S = 1/2$ and do not elicit significant paramagnetic shifts (due to a near isotropic magnetic susceptibility tensor).[21] However, with typical T_{1e}'s of ∼1−5 ns, in contrast to nitroxides, Cu^{2+} ions are expected to cause transverse PREs that are roughly 1−2 orders of magnitude lower, while concurrently generating substantial longitudinal PREs for the common biological low-γ nuclei such as ^{15}N and ^{13}C at moderate to high (∼10−20 T) external magnetic fields. The latter is related to the fact that the inverse of the electron relaxation time is on the order of the relevant nuclear Larmor frequencies in angular units.

The viability of this approach was demonstrated for several GB1 cysteine mutants, each modified with an EDTA-based Cu^{2+} binding tag[39,40] shown in Figure 2. The initial study by Nadaud et al.[41] showed that the combination of relatively small transverse PREs and sizable longitudinal PREs associated with the Cu^{2+} center enables quantitative measurements of the longitudinal PRE values for backbone ^{15}N nuclei at moderate (∼10 kHz) MAS frequencies. This was achieved by using standard 2D $^{15}N-^{13}C$ chemical shift correlation based methods to determine, in a residue-specific manner, the ^{15}N longitudinal relaxation rate constants (R_1) for proteins containing an EDTA−Cu^{2+} side chain, as well their diamagnetic EDTA−Zn^{2+} counterparts; for each ^{15}N site, the longitudinal PRE is obtained by taking the difference between the R_1 values obtained for the Cu^{2+} and Zn^{2+} proteins. Most importantly, measurable longitudinal PRE effects could be detected for ^{15}N nuclei located up to ∼20 Å from the Cu^{2+} center, and the experimental PRE
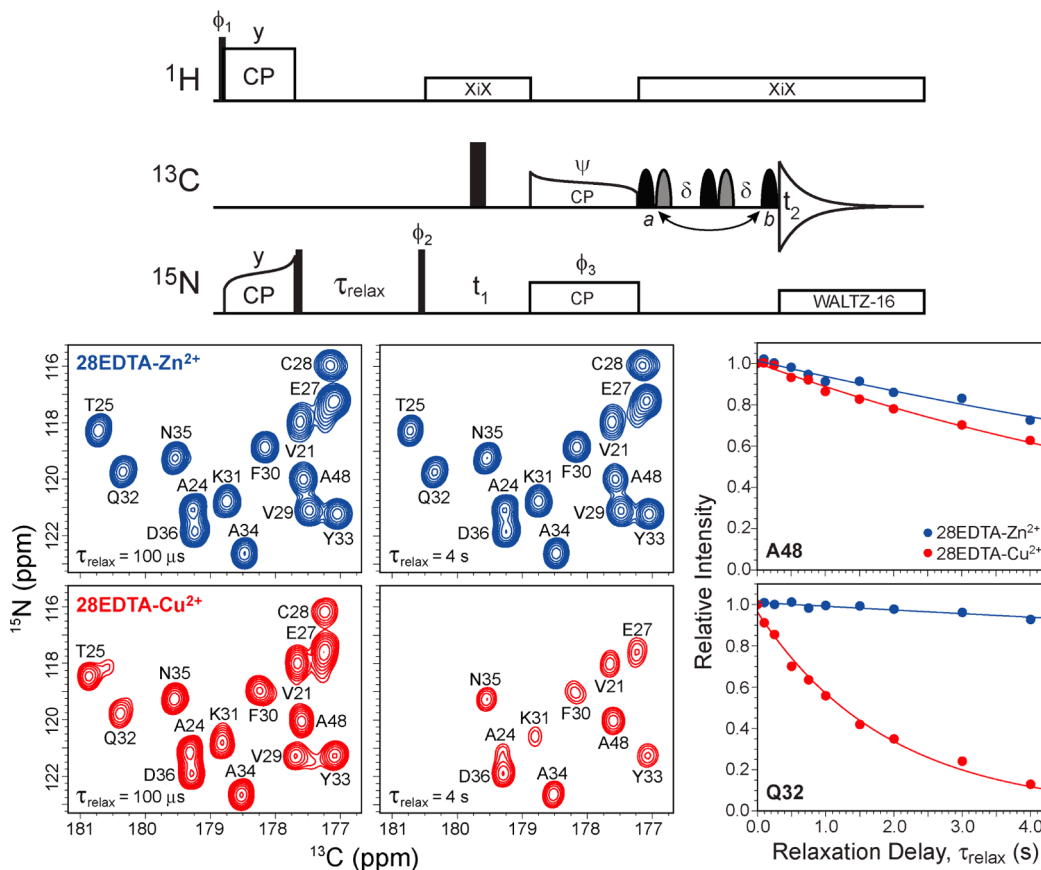
FIGURE 6. (top) Pulse scheme used to determine ^{15}N longitudinal relaxation rate constants in proteins in a residue specific manner, by recording a series of 2D $^{15}N-^{13}CO$ chemical shift correlation spectra as a function of the delay τ_{relax}. (bottom) Representative spectra and longitudinal ^{15}N relaxation trajectories for the K28C–EDTA–Zn^{2+} (blue contours and symbols) and K28C–EDTA–Cu^{2+} (red contours and symbols) GB1 mutants. Simulated trajectories used to extract the ^{15}N R_1 values were obtained by fitting experimental data to decaying single exponentials and are shown as solid lines of the corresponding color. Figure adapted from ref 28.

magnitudes and $^{15}N-Cu^{2+}$ distances were found to be in good agreement with those predicted based on the known protein structure.[41] In subsequent studies,[28,42,43] taking advantage of the intrinsic relaxation properties of proteins containing covalent paramagnetic tags, it was shown that paramagnetic relaxation-assisted condensed data collection type approaches[44−47] can be employed at high (∼40 kHz and greater) MAS rates to rapidly determine residue-specific longitudinal ^{15}N PREs for samples containing as little as ∼100 nmol of $^{13}C,^{15}N$-labeled protein. Note that the use of fast MAS comes with an additional benefit of allowing longitudinal ^{13}C PREs to be accessed in quantitative fashion by analogous methods,[27] which is not possible at lower MAS frequencies due to interference from proton-driven ^{13}C spin diffusion phenomena.

Figure 6 shows representative 2D $^{15}N-^{13}C$ correlation spectra recorded with short and long longitudinal ^{15}N relaxation delays (labeled as τ_{relax} in the pulse scheme shown in the figure) for the K28C–EDTA–Cu^{2+} and –Zn^{2+} mutants

of GB1. These data clearly illustrate that for most ^{15}N nuclei the intrinsic relaxation rates are relatively low for the diamagnetic protein. On the other hand, in the presence of Cu^{2+}, the individual amide ^{15}N relaxation rates vary widely as a function of the distance from the paramagnetic center as seen qualitatively in the spectra and quantitatively in the complete relaxation trajectories for selected residues. The summary of these measurements (Figure 7) shows that the magnitudes of the experimental longitudinal ^{15}N PREs are indeed strongly correlated with predictions based on the fold of GB1,[28,41−43] with the largest relaxation effects observed for the α-helix residues found in closest proximity to the EDTA−Cu^{2+} tag at position 28. Additionally, these quantitative data also enable the direct visualization of finer protein structural features, such as the individual turns of the helix (as seen from the modulation of the PRE magnitudes for residues ∼24−33 with a period of ∼3−4 amino acids).

It is noteworthy that several residues in K28C−EDTA−Cu^{2+} GB1 that are not proximal to the Cu^{2+} center, particularly
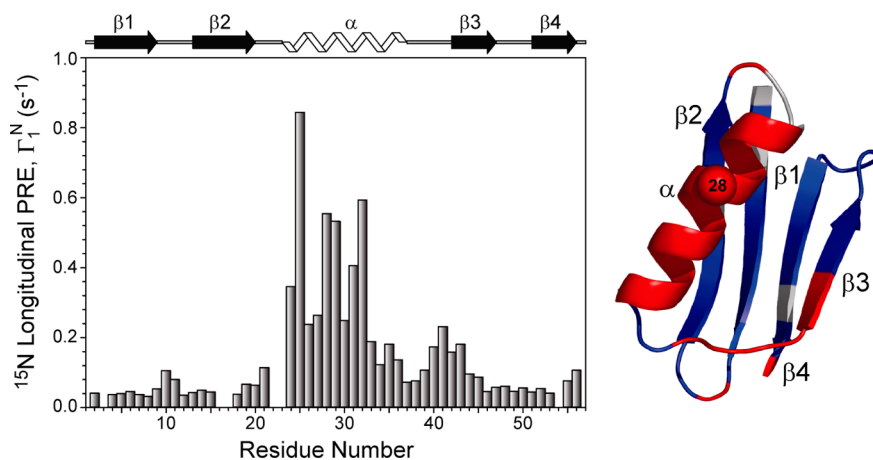
FIGURE 7. Longitudinal PREs for backbone amide ^{15}N nuclei for the K28C−EDTA−Cu^{2+} GB1 mutant plotted as a function of residue number (left) and mapped onto the tertiary structure of GB1 (right). The PREs are calculated by taking the difference between the ^{15}N R_1 values obtained for the Cu^{2+} and Zn^{2+} proteins as described in the Figure 6 caption. Residues with PRE values >0.1 s^{-1} (^{15}N−Cu^{2+} distances <15 Å) and <0.1 s^{-1} (^{15}N−Cu^{2+} distances >15 Å) are colored in red and blue, respectively. Figure adapted from ref 42.

amino acids D40−W43 in loop 3 and strand β3, appear to show somewhat elevated ^{15}N PREs. A study by Nadaud et al.[43] demonstrated that these relatively minor effects can be attributed mostly to the presence of native aspartate and glutamate residues in GB1, the side-chain carboxylate groups of which act as intrinsic low-affinity metal binding sites capable of interacting with any slight excess of Cu^{2+} ions used for the sample preparation and not chelated by the EDTA tag, with a smaller additional contribution from residual intermolecular ^{15}N−Cu^{2+} couplings due to insufficient dilution of the paramagnetic protein in the diamagnetic matrix. The influence of these effects on quantitative intramolecular PRE measurements can be minimized by loading the high-affinity metal binding tag with stoichiometric or slightly substoichiometric amounts of the metal and diluting the paramagnetic protein to a mole fraction of ∼15−20%, with the latter providing a reasonable compromise between the spectral sensitivity and the suppression of intermolecular PREs.

Protein Fold Determination Using PRE Restraints

In this section, we illustrate how ^{15}N longitudinal PREs can be used as structural restraints to elucidate the global three-dimensional fold of GB1, without having to rely on the availability of internuclear distances. In a recent study,[28] Sengupta et al. determined a collection of 231 ^{15}N PREs, as described above, for six GB1 cysteine point mutants containing EDTA−Cu^{2+} tags at solvent-exposed residues distributed throughout the protein at positions 8 (β1 strand), 19 (β2 strand), 28 (α-helix), 42 (β3 strand), 46 (β3 strand), and 53 (β4 strand). This corresponds to approximately 38 out of
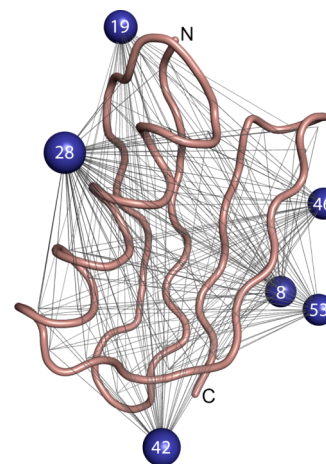

FIGURE 8. Summary of ^{15}N longitudinal PRE restraints determined for six GB1 variants containing cysteine EDTA−Cu^{2+} tags at positions 8, 19, 28, 42, 46, and 53 superimposed onto the GB1 X-ray structure. The Cu^{2+} centers (blue spheres) are shown at their approximate positions, and for clarity, the EDTA−side chains have been omitted.

55 possible backbone ^{15}N PREs per EDTA−Cu^{2+} variant or about 4−5 restraints per residue on average. Moreover, we note that roughly half of the measured PREs were smaller than 0.1 s^{-1}. To improve the convergence of the protein structure calculations, these PREs were converted to repulsive "NOE-type" distance restraints that prevented the ^{15}N and Cu^{2+} atoms in question from approaching each other closer than 15.1 Å. The larger PRE values were not translated to ^{15}N−Cu^{2+} distances, but rather used directly in the calculations. Figure 8 shows the summary of all ^{15}N longitudinal PRE restraints for the six EDTA−Cu^{2+} GB1 analogs superimposed onto the X-ray structure of GB1 (PDB entry 2GI9). It is evident that even with a relatively modest number
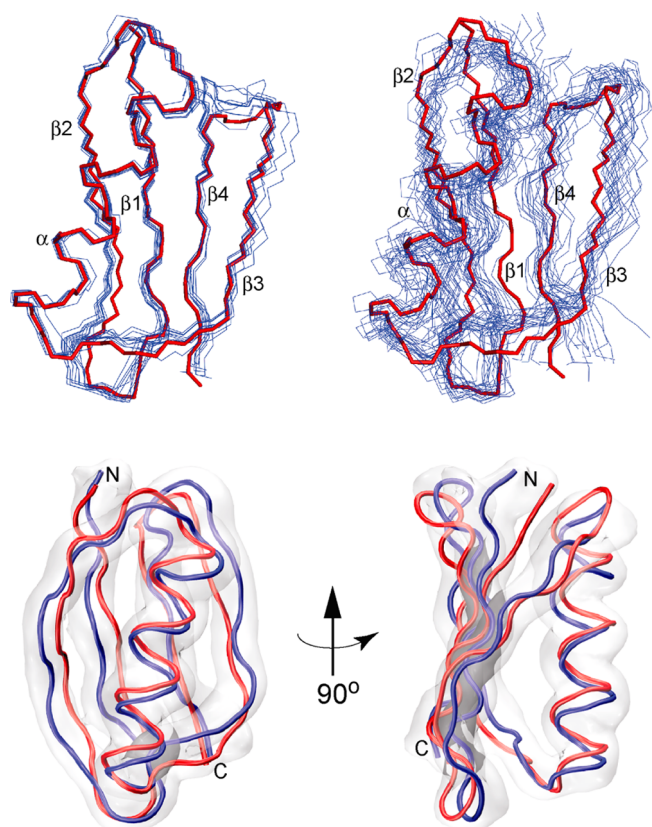
FIGURE 9. Summary of structure calculations for GB1 using solid-state NMR longitudinal ^{15}N PREs. (top left) Backbone traces (blue) for the ten lowest energy structures calculated using a protocol where the residues in regular secondary structure elements were frozen as rigid bodies according to the GB1 crystal structure, and the backbone conformations of the remaining residues were randomized. The reference GB1 X-ray structure is shown in red. (top right) Similar to top left, but showing 20 lowest energy structures calculated in a *de novo* manner using a largely unrestricted protocol with experimental restraints consisting of ^{15}N PREs and backbone torsion angles based on NMR chemical shifts. (bottom) Comparison of the GB1 X-ray structure (red) with the regularized mean *de novo* solid-state NMR structure (blue) obtained from the 20 lowest energy structures shown in the top right panel. The gray cloud represents the conformational space occupied by backbone atoms in the 20 lowest energy NMR structures. Figure adapted from ref 28.

of restraints, their long-range nature provides a wealth of information about the protein tertiary structure.

In order to assess the utility of solid-state NMR PRE restraints for protein structure determination, we first performed a set of idealized calculations for which the conformations of the backbone atoms in regular secondary structure elements (α-helix and the four β-strands) were fixed to those found in the crystal structure of GB1, while the backbone torsions for the remaining residues were randomized. The side chains of all residues, except ones decorated with the EDTA–Cu^{2+} tags, were fixed according to the crystal structure, and the EDTA–Cu^{2+} tags were also

frozen, in conformations optimized using the ^{15}N longitudinal PREs and GB1 X-ray coordinates. A representative set of low-energy protein structures calculated in this manner is shown in Figure 9 (top left panel). These data indicate that inclusion of the set of 231 ^{15}N PRE restraints enables the calculations to converge to an ensemble of structures that cluster tightly together and, more importantly, correctly reproduce the GB1 fold with coordinate root-mean-square deviations (rmsd's) for the backbone atoms ranging from 0.9 to 1.8 Å relative to the crystal structure. In contrast, structure calculations using the same protocol but with ^{15}N PRE restraints omitted do not converge, yielding numerous low-energy structures with incorrect folds characterized by backbone atom rmsd's in the range of $\sim$3−9 Å relative to the X-ray structure of GB1.[28]

With the worth of PRE based restraints in the context of protein structure calculations clearly established, we explored the possibility of obtaining the structure of GB1 in a *de novo* manner, by combining the long-distance ^{15}N PRE restraints with the usual torsion angle restraints based on NMR chemical shifts to define the local secondary structure. Using a largely unrestricted two-stage calculation protocol that, remarkably, did not make use of any conventional internuclear distance data, we were indeed able to elucidate a backbone fold for GB1, which was found to be in good agreement with the high-resolution X-ray protein structure (Figure 9).[28] The mean solid-state NMR structure determined in this fashion showed a rmsd of 1.8 Å for the backbone atom coordinates and a 2.7 Å all-heavy-atom rmsd relative to the crystal structure. Moreover, the conformational space spanned by the ensemble of the 20 lowest energy solid-state NMR conformers engulfed most of the protein backbone defined by the X-ray structure.

Concluding Remarks

The determination by MAS solid-state NMR techniques of numerous dipolar coupling based distance restraints significantly in excess of $\sim$5 Å is complicated for large biomolecules due to the inherent limitations of this methodology. However, this lack of long-range structural data, which acts as a bottleneck for protein structure determination, can be circumvented by using paramagnetic solid-state NMR methods that yield structural restraints in the $\sim$20 Å regime based upon dipolar interactions between the nuclear spins and a paramagnetic center embedded in the protein.

In this Account, we have outlined an approach to protein structure determination in the solid state, rooted in the measurements of site-specific nuclear paramagnetic relaxation

enhancements in variants of the protein of interest modified to contain covalent paramagnetic tags at specific locations. This approach is general and expected to become widely applicable to protein molecules that are compatible with the introduction of non-native side-chain tags or other linkers containing para-magnetic moieties.[29] A key advantageous feature of this methodology is that the extraction of site-resolved PRE data is relatively straightforward, consisting of monitoring of signal intensities in standard 2D or 3D NMR chemical shift correlation spectra displaying a minimal number of strong cross-peaks. These paramagnetic solid-state NMR methods are therefore relevant for larger protein systems, for which partial or complete resonance assignments can be established. Indeed, such applications are already beginning to emerge, as highlighted by a study of a seven-helix transmembrane sensory rhodopsin.[38] In a recent series of studies, we have applied paramagnetic solid-state NMR to the structural analysis of microcrystalline model protein GB1.[28,37,41−43] Most remark-ably, we have been able to show that even as few as four to five longitudinal ^{15}N PRE restraints per residue, recorded for several EDTA−Cu^{2+} GB1 mutants and supplemented only by chemical shift based torsional restraints, are adequate for deriving, in a *de novo* fashion, a protein backbone fold that agrees closely with the X-ray structure.[28] Although the resolu-tion of the resulting structure does not correspond to the highest achievable one, such structures are likely to be of sufficient quality for many applications and, importantly, can be obtained rapidly using rudimentary data analysis proce-dures and samples containing limited amounts of isotopically labeled paramagnetic protein. Moreover, if necessary, these moderate resolution structures can be further refined by incorporating additional paramagnetic or conventional dis-tance or dihedral angle restraints or by combining the existing PRE data with more sophisticated fragment based chemical shift−molecular mechanics structure calculation protocols.[48,49] Finally, we note that while the main focus of this Account was on the use of nuclear PREs as the sole long-range restraints in the structure determination of paramagnetically tagged pro-teins, the incorporation of paramagnetic restraints in the form of both nuclear PREs[27] and dipolar shifts[26,30−34] has also been recently demonstrated to significantly improve the quality of solid-state NMR structures for native metalloproteins derived by conventional internuclear distance based methods.[26,27]

BIOGRAPHICAL INFORMATION

Ishita Sengupta received her B.S. (2005) from Presidency Col-lege, Kolkata, India, and M.S. (2007) from University of Calcutta, India. She graduated with her Ph.D. from The Ohio State University in 2012.

Philippe Nadaud received his A.S. (2001) from University of Paris, France, and M.S. (2004) from University of Toulouse, France. He graduated with his Ph.D. from The Ohio State University in 2010 and is currently a postdoctoral researcher in the Jaroniec group.

Christopher Jaroniec received his B.S. (1997) from Kent State University and Ph.D. (2003) from the Massachusetts Institute of Technology, followed by a postdoctoral fellowship at the National Institutes of Health. He joined the faculty at The Ohio State University in 2006, where he is currently an Associate Professor of Chemistry.

FOOTNOTES

REFERENCES

1 McDowell, L. M.; Schaefer, J. High resolution NMR of biological solids. *Curr. Opin. Struct. Biol.* **1996**, *6*, 624–629.

2 Griffin, R. G. Dipolar recoupling in MAS spectra of biological solids. *Nat. Struct. Biol.* **1998**, *5*, 508–512.

3 McDermott, A. E. Structural and dynamic studies of proteins by solid-state NMR spectroscopy: rapid movement forward. *Curr. Opin. Struct. Biol.* **2004**, *14*, 554–561.

4 Böckmann, A. 3D protein structures by solid-state NMR spectroscopy: Ready for high resolution. *Angew. Chem., Int. Ed.* **2008**, *47*, 6110–6113.

5 McDermott, A. Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. *Annu. Rev. Biophys.* **2009**, *38*, 385–403.

6 Renault, M.; Cukkemane, A.; Baldus, M. Solid-state NMR spectroscopy on complex biomolecules. *Angew. Chem., Int. Ed.* **2010**, *49*, 8346–8357.

7 Böckmann, A.; Meier, B. H. Prions: En route from structural models to structures. *Prion* **2010**, *4*, 72–79.

8 Tycko, R. Solid-state NMR studies of amyloid fibril structure. *Annu. Rev. Phys. Chem.* **2011**, *62*, 279–299.

9 Reif, B. Ultra-high resolution in MAS solid-state NMR of perdeuterated proteins: Implications for structure and dynamics. *J. Magn. Reson.* **2012**, *216*, 1–12.

10 Shahid, S. A.; Bardiaux, B.; Franks, W. T.; Krabben, L.; Habeck, M.; van Rossum, B. J.; Linke, D. Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nat. Methods* **2012**, *9*, 1212–1217.

11 Opella, S. J.; Marassi, F. M. Structure determination of membrane proteins by NMR spectroscopy. *Chem. Rev.* **2004**, *104*, 3587–3606.

12 Dürr, U. H.; Gildenberg, M.; Ramamoorthy, A. The magic of bicelles lights up membrane protein structure. *Chem. Rev.* **2012**, *112*, 6054–6074.

13 Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. Structure of a protein determined by solid-state magic-angle spinning NMR spectroscopy. *Nature* **2002**, *420*, 98–102.

14 Lange, A.; Luca, S.; Baldus, M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* **2002**, *124*, 9704–9705.

15 Helmus, J. J.; Nadaud, P. S.; Höfer, N.; Jaroniec, C. P. Determination of methyl ^{13}C-^{15}N dipolar couplings in peptides and proteins by three-dimensional and four-dimensional magic-angle spinning solid-state NMR spectroscopy. *J. Chem. Phys.* **2008**, *128*, No. 052314.

16 De Paëpe, G.; Lewandowski, J. R.; Loquet, A.; Böckmann, A.; Griffin, R. G. Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* **2008**, *129*, No. 245101.

17 Nieuwkoop, A. J.; Wylie, B. J.; Franks, W. T.; Shah, G. J.; Rienstra, C. M. Atomic resolution protein structure determination by three-dimensional transferred echo double resonance solid-state nuclear magnetic resonance spectroscopy. *J. Chem. Phys.* **2009**, *131*, No. 095101.

18 Linser, R.; Bardiaux, B.; Higman, V.; Fink, U.; Reif, B. Structure calculation from unambiguous long-range amide and methyl ^{1}H-^{1}H distance restraints for a microcrystalline protein with MAS solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2011**, *133*, 5905–5912.

19 Huber, M.; Hiller, S.; Schanda, P.; Ernst, M.; Böckmann, A.; Verel, R.; Meier, B. H. A proton-detected 4D solid-state NMR experiment for protein structure determination. *ChemPhysChem* **2011**, *12*, 915–918.

20 Graesser, D. T.; Wylie, B. J.; Nieuwkoop, A. J.; Franks, W. T.; Rienstra, C. M. Long-range ^{19}F-^{15}N distance measurements in highly-^{13}C,^{15}N-enriched solid proteins with ^{19}F-dephased REDOR shift (FRESH) spectroscopy. *Magn. Reson. Chem.* **2007**, *45*, S129–S134.

21 Bertini, I.; Luchinat, C.; Parigi, G. *Solution NMR of Paramagnetic Molecules: Applications to Metallobiomolecules and Models*; Elsevier: Amsterdam, 2001.

22 Otting, G. Protein NMR using paramagnetic ions. *Annu. Rev. Biophys.* **2010**, *39*, 387–405.

23 Banci, L.; Bertini, I.; Luchinat, C.; Mori, M. NMR in structural proteomics and beyond. *Prog. Nucl. Magn. Reson. Spectrosc.* **2010**, *56*, 247–266.

24 Pintacuda, G.; Kervern, G. Paramagnetic solid-state magic-angle spinning NMR spectroscopy. *Top. Curr. Chem.* **2012**, DOI: 10.1007/128_2011_312.

25 Jaroniec, C. P. Solid-state nuclear magnetic resonance structural studies of proteins using paramagnetic probes. *Solid State Nucl. Magn. Reson.* **2012**, *43–44*, 1–13.

26 Bertini, I.; Bhaumik, A.; De Paëpe, G.; Griffin, R. G.; Lelli, M.; Lewandowski, J. R.; Luchinat, C. High-resolution solid-state NMR structure of a 17.6 kDa protein. *J. Am. Chem. Soc.* **2010**, *132*, 1032–1040.

27 Knight, M. J.; Pell, A. J.; Bertini, I.; Felli, I. C.; Gonnelli, L.; Pierattelli, R.; Herrmann, T.; Emsley, L.; Pintacuda, G. Structure and backbone dynamics of a microcrystalline metalloprotein by solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 11095–11100.

28 Sengupta, I.; Nadaud, P. S.; Helmus, J. J.; Schwieters, C. D.; Jaroniec, C. P. Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy. *Nat. Chem.* **2012**, *4*, 410–417.

29 Su, X. C.; Otting, G. Paramagnetic labelling of proteins and oligonucleotides for NMR. *J. Biomol. NMR* **2010**, *46*, 101–112.

30 Balayssac, S.; Bertini, I.; Lelli, M.; Luchinat, C.; Maletta, M. Paramagnetic ions provide structural restraints in solid-state NMR of proteins. *J. Am. Chem. Soc.* **2007**, *129*, 2218–2219.

31 Balayssac, S.; Bertini, I.; Bhaumik, A.; Lelli, M.; Luchinat, C. Paramagnetic shifts in solid-state NMR of proteins to elicit structural information. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17284–17289.

32 Bertini, I.; Emsley, L.; Lelli, M.; Luchinat, C.; Mao, J.; Pintacuda, G. Ultrafast MAS solid-state NMR permits extensive ^{13}C and ^{1}H detection in paramagnetic metalloproteins. *J. Am. Chem. Soc.* **2010**, *132*, 5558–5559.

33 Luchinat, C.; Parigi, G.; Ravera, E.; Rinaldelli, M. Solid-state NMR crystallography through paramagnetic restraints. *J. Am. Chem. Soc.* **2012**, *134*, 5006–5009.

34 Knight, M. J.; Felli, I. C.; Pierattelli, R.; Bertini, I.; Emsley, L.; Herrmann, T.; Pintacuda, G. Rapid measurement of pseudocontact shifts in metalloproteins by proton-detected solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2012**, *134*, 14730–14733.

35 Solomon, I. Relaxation processes in a system of two spins. *Phys. Rev.* **1955**, *99*, 559–565.

36 Kosen, P. A. Spin labeling of proteins. *Methods Enzymol.* **1989**, *177*, 86–121.

37 Nadaud, P. S.; Helmus, J. J.; Höfer, N.; Jaroniec, C. P. Long-range structural restraints in spin-labeled proteins probed by solid-state nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* **2007**, *129*, 7502–7503.

38 Wang, S.; Munro, R. A.; Kim, S. Y.; Jung, K.-H.; Brown, L. S.; Ladizhansky, V. Paramagnetic relaxation enhancement reveals oligomerization interface of a membrane protein. *J. Am. Chem. Soc.* **2012**, *134*, 16995–16998.

39 Ermácora, M. R.; Delfino, J. M.; Cuenoud, B.; Schepartz, A.; Fox, R. O. Conformation-dependent cleavage of staphylococcal nuclease with a disulfide-linked iron chelate. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6383–6387.

40 Ebright, Y. W.; Chen, Y.; Pendergrast, S.; Ebright, R. H. Incorporation of an EDTA—metal complex at a rationally selected site within a protein: Application to EDTA—iron DNA affinity cleaving with catabolite gene activator protein (CAP) and Cro. *Biochemistry* **1992**, *31*, 10664–10670.

41 Nadaud, P. S.; Helmus, J. J.; Kall, S. L.; Jaroniec, C. P. Paramagnetic ions enable tuning of nuclear relaxation rates and provide long-range structural restraints in solid-state NMR of proteins. *J. Am. Chem. Soc.* **2009**, *131*, 8108–8120.

42 Nadaud, P. S.; Helmus, J. J.; Sengupta, I.; Jaroniec, C. P. Rapid acquisition of multidimensional solid-state NMR spectra of proteins facilitated by covalently bound paramagnetic tags. *J. Am. Chem. Soc.* **2010**, *132*, 9561–9563.

43 Nadaud, P. S.; Sengupta, I.; Helmus, J. J.; Jaroniec, C. P. Evaluation of the influence of intermolecular electron-nucleus couplings and intrinsic metal binding sites on the measurement of ^{15}N longitudinal paramagnetic relaxation enhancements in proteins by solid-state NMR. *J. Biomol. NMR* **2011**, *51*, 293–302.

44 Wickramasinghe, N. P.; Parthasarathy, S.; Jones, C. R.; Bhardwaj, C.; Long, F.; Kotecha, M.; Mehboob, S.; Fung, L. W. M.; Past, J.; Samoson, A.; Ishii, Y. Nanomole-scale protein solid-state NMR by breaking intrinsic ^{1}H T$_1$ boundaries. *Nat. Methods* **2009**, *6*, 215–218.

45 Laage, S.; Sachleben, J. R.; Steuernagel, S.; Pierattelli, R.; Pintacuda, G.; Emsley, L. Fast acquisition of multi-dimensional spectra in solid-state NMR enabled by ultra-fast MAS. *J. Magn. Reson.* **2009**, *196*, 133–141.

46 Yamamoto, K.; Xu, J.; Kawulka, K. E.; Vederas, J. C.; Ramamoorthy, A. Use of a copper-chelated lipid speeds up NMR measurements from membrane proteins. *J. Am. Chem. Soc.* **2010**, *132*, 6929–6931.

47 Tang, M.; Berthold, D. A.; Rienstra, C. M. Solid-state NMR of a large membrane protein by paramagnetic relaxation enhancement. *J. Phys. Chem. Lett.* **2011**, *2*, 1836–1841.

48 Robustelli, P.; Cavalli, A.; Vendruscolo, M. Determination of protein structures in the solid state from NMR chemical shifts. *Structure* **2008**, *16*, 1764–1769.

49 Shen, Y.; Vernon, R.; Baker, D.; Bax, A. De novo protein structure generation from incomplete chemical shift assignments. *J. Biomol. NMR* **2009**, *43*, 63–78.