Quarterly Reviews of Biophysics **48**, 1 (2015), pp. 35–116. © Cambridge University Press 2015
doi:10.1017/S0033583514000122 Printed in the United States of America

Visualizing transient dark states by NMR spectroscopy

Nicholas J. Anthis and G. Marius Clore*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA

Abstract. Myriad biological processes proceed through states that defy characterization by conventional atomic-resolution structural biological methods. The invisibility of these 'dark' states can arise from their transient nature, low equilibrium population, large molecular weight, and/or heterogeneity. Although they are invisible, these dark states underlie a range of processes, acting as encounter complexes between proteins and as intermediates in protein folding and aggregation. New methods have made these states accessible to high-resolution analysis by nuclear magnetic resonance (NMR) spectroscopy, as long as the dark state is in dynamic equilibrium with an NMR-visible species. These methods – paramagnetic NMR, relaxation dispersion, saturation transfer, lifetime line broadening, and hydrogen exchange – allow the exploration of otherwise invisible states in exchange with a visible species over a range of timescales, each taking advantage of some unique property of the dark state to amplify its effect on a particular NMR observable. In this review, we introduce these methods and explore two specific techniques – paramagnetic relaxation enhancement and dark state exchange saturation transfer – in greater detail.

1. Introduction 36

2. Survey of NMR methods for studying invisible states 40
2.1. Paramagnetic NMR 40
2.1.1. Types of paramagnetic observables 40
2.1.2. Practical aspects of PRE 42
2.1.3. PRE and sparsely populated states 44
2.2. Relaxation dispersion 46
2.2.1. Chemical exchange 46
2.2.2. CPMG and rotating frame relaxation dispersion 48
2.2.3. Practical aspects of CPMG 52
2.2.4. Applications of CPMG 53
2.3. Saturation transfer 54
2.3.1. Principles and applications of saturation transfer 54
2.3.2. The DEST experiment 56
2.3.3. The CEST experiment 57

* Author for correspondence: G. M. Clore, Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA. Tel.: 301 496 0782; Email: mariusc@intra.niddk.nih.gov

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

2.4. Lifetime line broadening 58
2.5. Hydrogen exchange 60

3. Characterizing sparsely populated states by PRE 62
3.1. Theory 62
3.1.1. Quantitative description of PRE 62
3.1.2. PRE from minor states 65
3.2. Data acquisition 67
3.2.1. Producing paramagnetic samples 67
3.2.2. Measuring PRE data 72
3.3. Data analysis 77
3.3.1. Fitting PRE data 77
3.3.2. Ensemble simulated annealing calculations 80
3.4. Other approaches to treating PRE data 83
3.4.1. Maximum occurrence (MO), ensemble selection, and MD 83
3.4.2. PRE studies on intrinsically disordered systems 84
3.4.3. Solvent PRE 85
3.5. PRE in practice: application to CaM 86

4. Characterizing large transient assemblies by DEST and lifetime line broadening 89
4.1. Theory 89
4.1.1. Lifetime line broadening 90
4.1.2. DEST 91
4.2. Data acquisition 94
4.2.1. Measuring lifetime line broadening 94
4.2.2. Acquiring DEST data 94
4.3. Data analysis 95
4.4. Combined approaches involving DEST 97
4.5. DEST in practice: application to Aβ 98

5. Conclusion 100

6. Acknowledgements 101

7. References 101

1. Introduction

Modern structural biology techniques, in particular X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, have revealed much of the inner workings of biology at atomic resolution. However, these views have been largely restricted to states that are highly populated, homogeneous, static (especially for crystallography), and of relatively modest size (especially for NMR). However, many biological processes proceed through intermediate states that are short lived and comprise only a small fraction of the overall population of a molecular system at equilibrium; thus, they are invisible to conventional structural methods. These states, which constitute local minima above the global minimum on the free-energy landscape of a macromolecule (Bryngelson & Wolynes, 1987; Bryngelson *et al.* 1995; Miyashita *et al.* 2003, 2005; Onuchic *et al.* 1997; Wolynes, 2005), include macromolecular encounter complexes, protein-folding intermediates, and states involved in conformational selection and enzyme catalysis (Clore, 2011; Sekhar & Kay, 2013). Other elusive states include large (~1 MDa or larger) macromolecular

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

assemblies, especially those that are dynamic and/or heterogeneous, which include intermediates in some protein aggregation processes.

The methods available for characterizing biological molecules at atomic resolution, X-ray crystallography and conventional NMR spectroscopy, are generally incapable of probing the structures of sparsely populated or large, heterogeneous states. If well diffracting crystals are available, X-ray crystallography offers the most precise view into the structure of a protein at atomic resolution, although it offers only limited insight into its dynamics; the result is generally a single structure – or a family of closely related structures – which represents the most stable conformation within the constraints of the crystal lattice and the major population of the system of interest. Slight variations in structure can be seen between crystals and between protein chains within a crystal, and some side chains display multiple rotamers, although these can reflect crystal-packing artifacts. Crystallographic temperature factors (*B*-factors) can be interpreted as probes of local picosecond/nanosecond dynamics (Rueda *et al.* 2007), although they also reflect crystallographic heterogeneity. More recently, new approaches to systematically interpret lower intensity signals in electron density maps (0·3–1σ) (Fraser *et al.* 2009) or to fit X-ray diffraction data to an ensemble of structures (Burnley *et al.* 2012) have offered more rigorous means to explore protein dynamics and minor states by X-ray crystallography.

Conventional structure determination by NMR relies on interpretation of interproton distance restraints derived from nuclear Overhauser effects (NOEs) and dihedral angle restraints derived from *J* couplings (Clore & Gronenborn, 1998). The addition of long-range orientational restraints from residual dipolar couplings (RDCs) to this repertoire has further improved the breadth and accuracy of structure determination by NMR (Tjandra *et al.* 2000). Although these methods offer some insight into the dynamics of the molecule under study, they report almost solely upon the major population. More recently, paramagnetic relaxation enhancement (PRE) has been employed to provide long-range (up to ∼35 Å under favorable circumstances) distance restraints by NMR as well as to gain insight into sparsely populated states (Bertini *et al.* 2001b; Clore & Iwahara, 2009; Iwahara *et al.* 2004a; Keizers & Ubbink, 2011; Otting, 2010).

Other structural or biophysical methods are also capable of providing long-range distance information, although they often lack resolution and are generally incapable of observing minor states. Electron microscopy (EM) allows for lower-resolution structural characterization, generally of larger molecules, although modern single-particle cryo-EM has achieved resolutions better than 5 Å in very favorable cases (Baker *et al.* 2013; Bartesaghi *et al.* 2013; Lyumkis *et al.* 2013; Van Heel *et al.* 2000; Zhou, 2008). Although cryo-EM techniques can potentially 'freeze out' a range of molecular conformations, detection of minor species is made almost impossible by the need to average many individual particles. Atomic force microscopy (AFM) suffers from similar limitations. Other techniques for gaining structural insight include solution X-ray and neutron scattering, electron paramagnetic resonance (EPR) (in particular double electron–electron resonance (DEER)), and fluorescence (in particular Förster resonance energy transfer (FRET)). The latter results from the dipole–dipole coupling between two fluorophores and has an r^{-6} dependence on the distance between them, providing structural information in the range of 10–100 Å (Hillisch *et al.* 2001). DEER, which results from the dipole–dipole coupling between two paramagnetic centers, has an r^{-3} dependence on the distance between them, providing structural information in the range of 20–60 Å (Altenbach *et al.* 2008; Klug & Feix, 2008). These techniques do not suffer from the same molecular size limitations as NMR, but they have the additional major limitation that only a single pairwise distance can be measured per sample, whereas NMR offers hundreds to thousands of individual measurements per sample per experiment (Clore, 2008).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

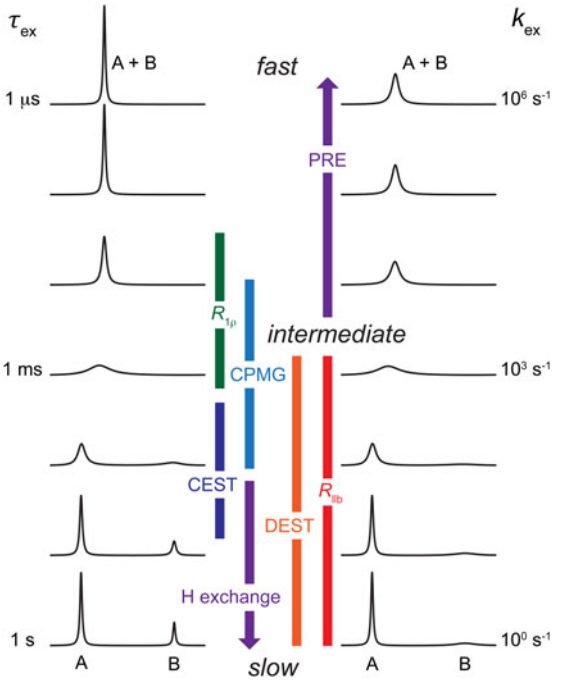


Fig. 1. NMR methods for characterizing a 'dark' state in exchange with a visible species. Simulated NMR spectra are shown for a resonance in exchange between two states (A and B) over a range of timescales. Shown from top to bottom are exchange lifetimes $\tau_{ex} = 1\ \mu s$, 10 μs, 100 μs, 1 ms, 10 ms, 100 ms, and 1 s ($k_{ex} = 10^6$–$10^0\ s^{-1}$). Spectra were simulated by solving the McConnell equations (Eqs. (16)–(20)) (McConnell, 1958) in MATLAB for a chemical shift difference between states of $\Delta\omega/2\pi = 120$ Hz (e.g. 2 ppm for ^{15}N on a 600 MHz spectrometer). In this example, state A is the major state ($p_A = 75\%$), and state B is the minor state ($p_B = 25\%$). On the left, states A and B have identical intrinsic rates of transverse relaxation ($R_2^0 = 10\ s^{-1}$). On the right, state B has an enhanced intrinsic relaxation rate ($R_{2,B}^0 = 100\ s^{-1}$), which could be due to a paramagnetic center or a slow rate of tumbling. In practice, p_B could be much lower and $R_{2,B}^0$ much larger, although less extreme values are used here for illustrative purposes. In the slow exchange regime (bottom), states A and B give rise to two separate peaks, although in practice state B may be invisible due to a low signal/noise ratio. In the fast exchange regime, a single peak is observed with a population-averaged chemical shift and apparent R_2. In the intermediate exchange regime (middle), the peaks undergo extreme chemical shift broadening, and the apparent R_2 is greatly enhanced due to R_{ex}. Down the center of the figure, NMR methods for characterizing dark states are shown, with a rough range of timescales over which they can be applied. The three methods shown toward the right (PRE, lifetime line broadening (R_{IIb}), and DEST) are used to visualize a dark state with a greatly enhanced R_2 compared with the major state. The other methods shown, toward the left, depend on either a difference in chemical shift between states A and B (rotating frame relaxation dispersion ($R_{1\rho}$), CPMG relaxation dispersion, and CEST) or a difference in rates of hydrogen exchange.

Although conventional NMR techniques for biomolecular structure determination report on an ensemble of closely related structures describing a major state, a range of additional NMR techniques have been developed to characterize processes on timescales from picoseconds to hours/days (Fig. 1). Rapid local motions (ps–ns) can be probed by NMR relaxation methods (Clore *et al.* 1990a, b; Ishima & Torchia, 2000; Kay *et al.* 1989; Peng & Wagner, 1994); often, ^{15}N is used to monitor backbone motion, measuring T_1 and T_2 relaxation and ^{1}H–^{15}N heteronuclear NOE values (Kay *et al.* 1989), and the framework for analyzing these measurements by the model-free formalism is well established (Clore *et al.* 1990a; Kay *et al.* 1989; Lipari & Szabo,

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

1982a, b; Xia *et al.* 2013). This methodology can be extended to describe larger-scale slower motions on the nanosecond timescale that are still faster than the reorientational time (Clore *et al.* 1990a, b). For larger proteins, the favorable spectral properties of methyl groups make the methyl ^{13}C, ^{1}H, and ^{2}H nuclei good probes for dynamic processes in larger molecules (Igumenova *et al.* 2006; Ruschak & Kay, 2010). Regardless, these experiments largely describe the rapid motion of the major state (or states) of a molecule. However, PRE offers the ability to characterize minor states in dynamic exchange with a major state on these fast timescales, as well as longer timescales up to ~250–500 μs.

Additional NMR techniques have been developed to monitor dynamics on the intermediate and slow timescales. Relaxation dispersion, using either Carr–Purcell–Meiboom–Gill (CPMG) (Carr & Purcell, 1954; Meiboom & Gill, 1958) or rotating frame ($R_{1\rho}$) experiments, can characterize systems on the intermediate exchange timescale (~50–100 μs to 10 ms) (Korzhnev & Kay, 2008; Mittermaier & Kay, 2006; Palmer *et al.* 2001). The dynamics of systems in the intermediate exchange regime can also be characterized by analysis of resonance line shapes (Rao, 1989). RDCs may also serve as a probe of dynamics on the μs timescale (Lange *et al.* 2008). Exchange processes taking place on slower timescales (ms–s) can be characterized by additional techniques, such as heteronuclear zz-exchange spectroscopy (Farrow *et al.* 1994; Iwahara & Clore, 2006b; Montelione *et al.* 1989; Sahu *et al.* 2007; Wider *et al.* 1991), saturation transfer (Fawzi *et al.* 2010b, 2011b, 2012; Vallurupalli *et al.* 2012) and hydrogen exchange (Bai, 2006; Krishna *et al.* 2004b). Processes that are even slower still (>s) can be studied by real-time NMR, where changes are monitored in sequentially acquired NMR spectra (Schanda & Brutscher, 2005).

Thus, a wide range of techniques have been developed to study motions by NMR over a wide range of timescales; some of these techniques are applicable to characterizing exchange involving states that are invisible to conventional structural methods owing to their transient nature coupled with a low equilibrium population and/or large molecular weight. Detection and characterization of an otherwise invisible species is accomplished by amplifying the effect of the dark state on some NMR observable, so that its footprint is observed in measurements on the visible species. In many cases, the minor state is detected through an enhancement of spin relaxation. Two types of relaxation (the exponential return to equilibrium of the bulk magnetization in a sample after perturbation) occur in NMR experiments. Longitudinal relaxation describes the return to the equilibrium distribution of spin *up* and *down* states along the z-axis (the axis of the B_0 magnetic field) and is characterized by an exponential rate constant R_1 (or time constant $T_1 = 1/R_1$). Transverse relaxation describes a loss of coherence on the transverse xy-plane and is characterized by an exponential rate constant R_2 (or time constant $T_2 = 1/R_2$). Since this coherent xy signal is detected in an NMR experiment, the rate of transverse relaxation (R_2) determines the height and linewidth of each peak in the NMR spectrum (larger R_2 leads to broader peaks and lower peak heights):

$$\Delta\nu_{1/2} = \frac{R_2}{\pi}, \tag{1}$$

where $\Delta\nu_{1/2}$ is the linewidth (full-width at half-height) given in Hz. R_1, on the other hand, determines the necessary delay between individual scans in an NMR experiment (larger R_1 allows a faster rate of repetition).

Broadly speaking, NMR relaxation is of interest as an atomic-level probe of molecular dynamics; additionally, if a visible species is in exchange with an invisible species, under certain circumstances this can lead to an enhancement of the apparent R_2 relaxation rate observed for the

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

visible species, allowing detection and characterization of the otherwise invisible state. The source of relaxation enhancement can vary, and the applicability of a given NMR experiment will depend on the timescale of the exchange event in question. Properties of the invisible state that can lead to such relaxation enhancements include slower tumbling due to a larger molecular weight (allowing study by dark state exchange saturation transfer (DEST) if exchange takes place on a slow (1 ms–1 s) timescale; Fawzi *et al.* 2011b, 2012), different resonance frequencies (allowing study by relaxation dispersion if exchange takes place on an intermediate (100 μs–10 ms) timescale; Sekhar & Kay, 2013), or closer proximity to an unpaired electron (allowing study by PRE if exchange takes place on a fast (<250–500 μs) timescale; Clore & Iwahara, 2009). In the next section, we will survey some of these available NMR techniques for studying sparsely populated states, and in the following two sections, we will discuss two of these methods, PRE and DEST, in greater detail.

2. Survey of NMR methods for studying invisible states

Several different NMR experiments are available for studying invisible states over a range of different timescales (Clore, 2011; Palmer *et al.* 2005; Sekhar & Kay, 2013; Fig. 1); these include PRE (Clore & Iwahara, 2009) and pseudocontact shifts (PCSs; Bertini *et al.* 2010) (fast); rotating-frame ($R_{1\rho}$) (Palmer & Massi, 2006) and CPMG relaxation dispersion (intermediate) (Korzhnev & Kay, 2008); chemical exchange saturation transfer (CEST; Vallurupalli *et al.* 2012); DEST and lifetime line broadening ($R_{1\mathrm{b}}$) (slow) (Fawzi *et al.* 2011b, 2012); and hydrogen exchange (very slow) (Englander *et al.* 2007). The ability of these various methods to characterize an otherwise invisible species relies on enhancement of relaxation rates due to exchange with the invisible species. A major advantage of these NMR methods is that they offer high resolution (at least at the level of an individual residue – e.g. backbone amide – but often at atomic resolution). Many of these methods are built upon phenomena that were described decades ago (e.g. PRE (Solomon, 1955), CPMG (Carr & Purcell, 1954; Meiboom & Gill, 1958), saturation transfer (Forsén & Hoffman, 1963)), but only through more recent methodological and technical advances have they become more widely applicable and has their utility for characterizing these otherwise invisible states been demonstrated.

2.1 Paramagnetic NMR

2.1.1 Types of paramagnetic observables

Unpaired electrons have dramatic effects on NMR spectra, and these paramagnetic effects are useful both as long-distance structural restraints and as probes of sparsely populated species. Here we will discuss two fundamental types of paramagnetic phenomena in NMR, PRE and hyperfine shifts (i.e. PCSs), as well as additional NMR observables that are influenced by the presence of a paramagnetic center, including RDCs, relaxation dispersion, and CEST.

PRE arises from dipolar coupling between a nucleus (primarily ^{1}H in solution NMR) and an unpaired electron (or collection of unpaired electrons). The PRE is proportional to the population-averaged $\langle r^{-6} \rangle$ distance between the paramagnetic center and the nucleus. This distance dependence, coupled with the large magnitude of the PRE (owing to the large magnetic moment of an electron), makes it a useful long-distance structural restraint, providing distance information of up to ~20 Å for a nitroxide tag and ~35 Å for some metal ions (such as Mn^{2+} or Gd^{3+}) with multiple unpaired electrons. The magnitude of the PRE is proportional

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

to the spin quantum number s ($= 1/2$ for each unpaired electron). Whereas PRE is an isotropic effect, PCSs, paramagnetic RDCs, and paramagnetic contributions to relaxation dispersion and CEST arise from a paramagnetic center with an anisotropic g-tensor. For measuring PRE data, it is thus desirable to employ a paramagnetic group with an isotropic g-tensor, eliminating these other paramagnetic effects. An isotropic g-tensor results from a symmetric arrangement of unpaired electrons; examples include nitroxides ($s = 1/2$), Mn^{2+} ($s = 5/2$, i.e. half-filled 3d-subshell), and Gd^{3+} ($s = 7/2$, i.e. half-filled 4f-subshell).

Owing to the large magnitude of the paramagnetic effect and its $\langle r^{-6} \rangle$ distance dependence, PRE values are extremely large at short distances. Because of this, under the right circumstances, the PRE profile from a major (visible) species will include contributions from minor (otherwise invisible) species, making PRE a powerful tool for characterizing such sparsely populated states (Anthis & Clore, 2013; Anthis *et al.* 2011; Bashir *et al.* 2009; Bertini *et al.* 2012b; Fawzi *et al.* 2010a; Iwahara & Clore, 2006a; Iwahara *et al.* 2004a; Tang *et al.* 2006, 2007, 2008b; Volkov *et al.* 2006). Although the NOE, which is the basis of conventional NMR structure determination, also has an r^{-6} distance dependence, the magnitude of the NOE is much smaller (limited to within ~6 Å); thus, NOE data do not sufficiently amplify minor state information, and sparsely populated states (<30% population) are not apparent in NOE data. The quantitative theoretical and practical aspects of PRE and its use as a probe of minor states are described further in Section 3.

In the presence of a paramagnetic center, the chemical shift of a nucleus may undergo a hyperfine shift, which is the sum of Fermi contact shift (FCS) and PCS contributions. The FCS is a through-bond effect resulting from the delocalization of unpaired electrons, and it occurs through direct contact (within a small number of chemical bonds) with the paramagnetic center (Bertini *et al.* 2001a). Thus, the FCS is analogous to diamagnetic J coupling between nuclei. PCSs are much longer-range through-space dipolar effects that arise from paramagnetic centers with an anisotropic g-tensor. The PCS has an $\langle r^{-3} \rangle$ distance dependence and a strong orientational dependence. Thus, additional orientational information can be derived from PCSs, although the analysis of PCSs adds additional complications over the analysis of PRE data. PCSs can also be used in the same manner as PRE data to characterize minor states, although PCS data require a paramagnetic center with a fixed orientation relative to the observed nucleus due to the orientational dependence of the PCS. Whereas PCSs are only observed in the presence of a paramagnetic center with an anisotropic g-tensor, FCSs are also observed for isotropic g-tensors. However, due to their short range, FCSs are generally not useful for protein NMR.

The magnitude of the PCS is given by (Bertini *et al.* 2002):

$$\delta_{\mathrm{PCS}} = \frac{1}{12} r^{-3} \left\{ \Delta\chi_{\mathrm{ax}} (3\cos^2\theta - 1) + \frac{3}{2} \Delta\chi_{\mathrm{rh}} \sin^2\theta \cos 2\phi \right\}, \tag{2}$$

where r is the distance between the paramagnetic metal and the observed nucleus, θ the angle between the metal/nucleus vector and the z-axis of the magnetic susceptibility tensor (χ, centered at the paramagnetic metal), and ϕ the angle between the projection of the metal/nucleus vector on the xy-plane and the x-axis of the χ-tensor. $\Delta\chi_{\mathrm{ax}}$ and $\Delta\chi_{\mathrm{rh}}$ are the axial and rhombic components, respectively, of the χ-tensor, which are defined as:

$$\Delta\chi_{\mathrm{ax}} = \chi_{zz} - \frac{1}{2}(\chi_{xx} + \chi_{yy}), \tag{3}$$

$$\Delta\chi_{\mathrm{rh}} = \chi_{xx} - \chi_{yy}. \tag{4}$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

The χ-tensor is related to the g-tensor by (Bertini *et al.* 2002):

$$\chi_{kk} = \frac{\mu_0 N_A \mu_0 S(S+1)}{3k_B T} g_{kk}^2, \tag{5}$$

where s is the electron-spin quantum number, μ_0 the permeability of free space, μ_B the Bohr magneton (the magnetic moment of a free electron), N_A Avogadro's number, k_B the Boltzmann constant, T the temperature, and g_{kk} an element of the g-tensor ($k = x$, y, or z).

The presence of a paramagnetic center with an anisotropic g-tensor will also cause partial alignment with the magnetic field. This gives rise to RDCs, given by (Bertini *et al.* 2002):

$$D_{AB} = \frac{hB_0^2 \gamma_A \gamma_B}{240\pi^3 kT} r_{AB}^{-3} \left\{ \Delta\chi_{ax}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{rh}\sin^2\theta\cos 2\phi \right\}, \tag{6}$$

where D_{AB} is the RDC between nuclei A and B, r_{AB} the distance between the two nuclei, θ the angle between the internuclear vector and the z-axis of the χ-tensor, ϕ the angle between the projection of the internuclear vector on the xy-plane and the x-axis of the χ-tensor, B_0 the static magnetic field strength, and h the Planck constant. Note that both the PCS and the paramagnetic RDC have an orientational dependence, whereas the PRE does not. Also, the PCS has an r^{-3} dependence on the distance between the observed nucleus and the paramagnetic center (compared with r^{-6} for PRE). Thus use of a combination of PCS, RDC, and PRE data (in addition to data from different paramagnetic metals) can help resolve any degeneracy in paramagnetic data (Fragai *et al.* 2013).

Owing to the large magnitude of hyperfine chemical shifts induced by paramagnetic groups (up to ~10 ppm for ^{1}H PCSs in proteins, >100 ppm for atoms directly bound to paramagnetic metal ions), they can facilitate the measurement of other chemical shift-dependent observables. In the case of intermediate chemical exchange (see the next section), PCSs can lead to exchange broadening, which can be characterized by relaxation dispersion (Eichmuller & Skrynnikov, 2007; Hass *et al.* 2010; Wang *et al.* 2007). Hyperfine shifts also facilitate the measurement of CEST (Vinogradov *et al.* 2013), an effect that has proven useful in magnetic resonance imaging (MRI), where compounds have been developed that cause hyperfine shifts and can be used as contrast agents in paramagnetic CEST (paraCEST) imaging (Hancu *et al.* 2010; Vinogradov *et al.* 2013). These CEST and paraCEST agents exist in addition to the more extensively used collection of PRE-based MRI contrast agents, where the PRE from a small soluble paramagnetic molecule or ion (e.g. chelated Gd^{3+}) locally enhances signal (most commonly, in T_1-weighted measurements; or, the PRE decreases the signal in T_2-weighted measurements) (Caravan *et al.* 1999).

2.1.2 Practical aspects of PRE

The use of paramagnetic restraints in biomolecular NMR was previously restricted primarily to metalloproteins, where a paramagnetic cation can easily replace the natural metal ligand (Bertini *et al.* 2001b; Cheng & Markley, 1995; Ubbink *et al.* 2002). Although the underlying theory of PRE was originally developed in the 1950s (Bloembergen & Morgan, 1961; Solomon, 1955) and the utility of PRE restraints for macromolecular structure determination was demonstrated in the 1980s (Kosen *et al.* 1986; Schmidt & Kuntz, 1984), only recently have PRE measurements become practically useful and commonly performed due to the development of straightforward biochemical methods for site-specific introduction of paramagnetic labels in proteins (Battiste & Wagner, 2000; Donaldson *et al.* 2001; Gaponenko *et al.* 2000). A wide range of chemistries are available for conjugating a paramagnetic tag (either a nitroxide or a metal ion-bound chelator)

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

to a protein (Hubbell *et al.* 2013; Jeschke, 2013; Keizers & Ubbink, 2011); most common are those that involve a nitroxide group or a paramagnetic metal ion chelated by EDTA and are conjugated to the protein via an exposed cysteine (Kosen, 1989). The inherent flexibility of these tags was previously a hurdle for the quantitative analysis of PRE data, although these issues have largely been overcome through improved computational methods (Iwahara *et al.* 2004b). More rigid tags have also been developed (Fawzi *et al.* 2011a; Keizers *et al.* 2007), and the use of such rigid tags is essential for measuring PCSs. Although PRE values are also affected by internal motion, they do not, like PCSs, have an orientational dependence, thus simplifying analysis and enabling the use of tags with flexible linkers.

The PRE is measured by taking the difference in relaxation rate between a paramagnetic state and a control diamagnetic state. PRE can be measured as either an enhancement of the longitudinal (Γ_1) or transverse (Γ_2) relaxation rate, although the measurement of Γ_2 is advantageous (Clore & Iwahara, 2009; Clore *et al.* 2007) because its larger magnitude makes it a more sensitive probe and because it exhibits a simple dependence of the order parameter (S^2) and is largely independent of the correlation time (τ_i) of internal motion (Iwahara *et al.* 2004b). Moreover, ^{1}H transverse relaxation is a single-exponential process, whereas ^{1}H longitudinal relaxation is multi-exponential due to cross-relaxation and, in the case of ^1H_N, water exchange, which results in perturbation of the observed ^{1}H–Γ_1 values. PRE values in a dynamic system cannot be directly interpreted as a single distance r, since both Γ_1 and Γ_2 depend upon the population-averaged value of $\langle r^{-6} \rangle$. This difficulty with the quantitative analysis of PRE data has been addressed by the development of new theoretical and computational methods that use an ensemble approach to take into account the different conformations sampled by a paramagnetic label on the surface of a protein, calculating PRE order parameters on the fly from atomic coordinates as a simulation proceeds (Iwahara *et al.* 2004b).

PRE data provide useful measurements of long distances in macromolecular structures (up to ~35 Å), and hundreds of individual distances can be measured from a single spin-label site (compared with FRET or DEER, where each doubly labeled sample only yields a single pairwise distance). As such, PRE data have been used to characterize the structures of a variety of macromolecules (Clore & Iwahara, 2009), including soluble proteins (Battiste & Wagner, 2000; Donaldson *et al.* 2001; Dvoretsky *et al.* 2002; Gaponenko *et al.* 2000), membrane proteins (Liang *et al.* 2006; Roosild *et al.* 2005), and proteins in complex with other proteins (Card *et al.* 2005; Gross *et al.* 2003; Mahoney *et al.* 2000; Mal *et al.* 2002; Rumpel *et al.* 2008), nucleic acids (Iwahara *et al.* 2003, 2004b; Ramos & Varani, 1998; Ueda *et al.* 2004; Varani *et al.* 2000), and oligosaccharides (Jain *et al.* 2001; Johnson *et al.* 1999; Macnaughtan *et al.* 2007). Long-distance PRE restraints do not have the same power to constrain a molecular structure as a series of short-distance NOE measurements between atoms near in three-dimensional (3D) space (but not linear sequence), but these long-distance measurements complement these NOE-based short-distance measurements, and they are particularly important in scenarios where measuring many short-range distances by NOEs is difficult (e.g. membrane proteins) (Gottstein *et al.* 2012).

PRE has also found applications in solid-state NMR, where additional challenges have had to be overcome, but where successes have included the *de novo* calculation of the fold of the protein GB1 (Sengupta *et al.* 2012). Owing to extremely rapid transverse relaxation rates in the solid state, it is technically difficult to measure Γ_2, so Γ_1 is generally measured, requiring a paramagnetic center that maximizes the Γ_1/Γ_2 ratio. Thus, while paramagnetic centers with a very slow electron relaxation times (large values of τ_s; e.g. nitroxides), are unsuitable due to a small Γ_1/Γ_2 ratio,

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

paramagnetic centers with a more intermediate value of τ_s give a more favorable Γ_1/Γ_2 ratio and are more appropriate for PRE measurements in the solid state (Sengupta *et al.* 2013). The large magnitude of intermolecular PREs due to molecular packing is another concern that arises in the solid state and must be considered (Nadaud *et al.* 2011). Beyond their utility for generating distance restraints, the paramagnetic groups can be used as doping agents to increase the signal-to-noise (S/N) ratio and decrease the total acquisition time needed for a solid-state NMR experiment (Wickramasinghe *et al.* 2007).

2.1.3 PRE and sparsely populated states

PRE is uniquely able to characterize otherwise invisible sparsely populated states (Anthis & Clore, 2013; Anthis *et al.* 2011; Bashir *et al.* 2009; Bertini *et al.* 2012b; Fawzi *et al.* 2010a; Iwahara & Clore, 2006a; Iwahara *et al.* 2004a; Tang *et al.* 2006, 2007, 2008b; Volkov *et al.* 2006) if two criteria are met: (1) the minor state must undergo rapid exchange; and (2) the distance between the paramagnetic center and some of the nuclei being measured must be shorter in the minor state than in the major state (Clore, 2008; Clore & Iwahara, 2009; Clore *et al.* 2007). This is possible because a nucleus in close proximity to a paramagnetic center will exhibit an extremely large PRE (>1000 s^{-1}), and the contribution from the close-proximity state will dominate the overall PRE measured for that nucleus, even if it constitutes a minor percentage of the total population (Fig. 2).

With respect to the first (kinetic) requirement, the exchange must be fast on the PRE timescale, the latter meaning that the minor state exchanges with the major state at a rate that is much greater than the difference in PRE rates between the two states ($k_{ex} \gg \Gamma_2^{minor} - \Gamma_2^{major}$). In practice, this generally requires a timescale of $\tau_{ex} < 250\ \mu s$ ($k_{ex} > 1000\ s^{-1}$). With respect to the second (structural) requirement, a shorter distance in the minor state could arise from two molecules forming a transient complex in the minor state or a conformational change that brings two parts of a molecule closer together in the dark state. In such a system, the overall PRE observed is an $\langle r^{-6} \rangle$ population-weighted average of the PREs for all conformations in exchange. Because the PRE can be extremely large at close distances, even a very small population can cause a PRE to be much greater than the very small near-zero value that would be observed for a nucleus located far from a paramagnetic center. Under these circumstances, one can not only detect, but also characterize, the structural details of any minor state(s) giving rise to PREs, even those with populations of <1%. In fact, PREs are the most robust and comprehensive method for providing structural information on minor states. No other technique provides direct structural information on minor states, although others can provide some indirect information (through chemical shifts or RDCs) and can provide more kinetic information than PRE. A limitation, however, is that PRE data offer less insight into the kinetics of an exchange process, due to the difficulty of deconvoluting the contribution of different parameters (distance, exchange rate, and population) in systems undergoing exchange on the intermediate PRE timescale.

PRE studies have been used to gain insight into the transient states underlying a range of biomolecular processes, including protein/DNA interactions, protein/protein interactions, and interdomain interactions within proteins. For example, PRE data have revealed that the A-box of HMGB-1 interacts non-specifically with DNA, relevant to its role as a DNA-bending protein (Iwahara *et al.* 2004a). PRE studies have characterized the inter-DNA-strand and intra-DNA-strand ('sliding') translocation of the transcription factors HoxD9 (Iwahara & Clore, 2006a; Iwahara *et al.* 2006), Oct1 (Takayama & Clore, 2011, 2012), and Sox2

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

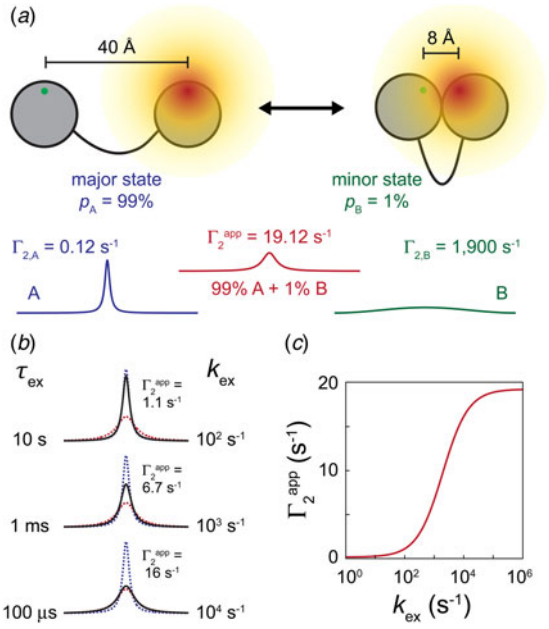


Fig. 2. Characterizing a sparsely populated state by PRE. If a major state (A) is in exchange with a minor state (B), the effect of the minor state can be amplified by PRE if two points (corresponding to the paramagnetic center and observed nucleus) are in closer proximity in the minor state than in the major state. (a) In this example, the protein is in exchange between an extended state (A, $r = 40$ Å) and a compact state (B, $r = 8$ Å). If a nitroxide spin label is attached to one point, the PRE measured for a proton at the other point would be virtually zero in state A ($\Gamma_{2,A} = 0{\cdot}12\ s^{-1}$), but very large in state B ($\Gamma_{2,B} = 1900\ s^{-1}$) causing broadening of the observed NMR signal. PRE values were calculated assuming a correlation time of $\tau_c = 10$ ns (corresponding to a protein at room temperature with a molecular weight of ~20 kDa). Spectra were calculated by solving the McConnell equations (Eqs. (16)–(20) and (57)) in MATLAB, assuming $R_{2,dia} = 10\ s^{-1}$ in each state (the height of the spectrum for state B [green] has been magnified 10^7 times for the purposes of visualization). If states A and B are in fast exchange on the PRE timescale (in this example: $k_{ex} \gg \Gamma_{2,A} - \Gamma_{2,B} = 1900\ s^{-1}$; $\tau_{ex} \ll 500\ \mu s$), the apparent PRE will be the population-averaged PRE of the two states ($\Gamma_2^{app} = 0{\cdot}99 \times 0{\cdot}12\ s^{-1} + 0{\cdot}01 \times 1900\ s^{-1} = 19{\cdot}12\ s^{-1}$). Thus, even although state B comprises only 1% of the total population, Γ_2^{app} arises almost exclusively from this minor state. (b) If the exchange rate is slower, the apparent PRE measured for A will be less than the population-weighted averaged PRE. Shown in black are simulated spectra for the system undergoing exchange with a rate of $k_{ex} = 10^2$, 10^3, or $10^4\ s^{-1}$. The dashed blue spectrum corresponds to the paramagnetic spectrum of A alone, and the dashed red spectrum corresponds to the population-weighted averaged paramagnetic spectrum under very fast exchange conditions ($10^6\ s^{-1}$). (c) Apparent PRE plotted against k_{ex}. The PRE increases from $0{\cdot}12\ s^{-1}$ for a system undergoing very slow exchange (the value for state A alone) to $19{\cdot}12\ s^{-1}$ for a system undergoing very fast exchange.

(Takayama & Clore, 2012), revealing transient interactions with non-specific DNA sequences that facilitate the formation of the specific complex. The molecular details of the encounter complex between Enzyme I (EI) and the histidine-containing phosphocarrier protein (HPr), which plays a central role in the bacterial phosphotransferase system, have been elucidated (Fawzi *et al.* 2010a; Tang *et al.* 2006). PRE measurements have also demonstrated that cytochrome *c* and cytochrome *c* peroxidase form an encounter complex that samples conformational space near, although not identical to, the major species observed by X-ray crystallography, offering fundamental insight into the structural basis of electron transport (Volkov *et al.* 2006, 2010). PRE studies have demonstrated that in the absence of a sugar ligand, maltose-binding protein (MBP) samples a

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

partially closed state ~5% of the time that is similar – although not identical – to its closed sugar-bound state (Tang *et al.* 2007) and in the absence of a peptide target, the calcium-sensing protein calmodulin (CaM) adopts a wide range of conformations, including structures similar in conformation to the more rigid peptide-bound complex (Anthis *et al.* 2011). These studies have offered a novel view into otherwise invisible states, providing fundamental insights into the process of molecular recognition and the complementary roles of conformational selection and induced fit.

2.2 Relaxation dispersion

2.2.1 Chemical exchange

If a molecule is in exchange between two states:

$$A \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} B, \tag{7}$$

then the number of NMR peaks (as well as their height and linewidth) observed for a given nucleus will depend on the chemical shift difference ($\Delta\omega$) between the two states and on the exchange rate constant ($k_{ex} = k_{AB} + k_{BA}$; the kinetics of exchange can also be described by the exchange lifetime, defined by $\tau_{ex} = 1/k_{ex}$.) (Fig. 1). If $\Delta\omega$ (given in angular frequency (rad·s^{-1}), related to $\Delta\nu$ (Hz) by $\Delta\omega = 2\pi\Delta\nu$) is much (>10 times) larger than k_{ex} (the slow chemical shift exchange regime), two distinct peaks will be observed for the resonance, at the chemical shifts of the individual states, with peak volumes proportional to the populations of the individual states. If k_{ex} is instead much (>10×) larger than $\Delta\omega$, a single peak will be observed at the population-weighted mean chemical shift; this is the fast chemical shift exchange regime. In the intermediate chemical shift exchange regime, however, the ratio of $k_{ex}/\Delta\omega$ takes a value close to 1 (±one order of magnitude), and line broadening is observed due to an increase in the observed transverse relaxation rate (R_2^{obs}), which is the sum of the intrinsic transverse relaxation rate (R_2^0) and an exchange contribution (R_{ex}) that depends upon k_{ex} and $\Delta\omega$ (Palmer & Massi, 2006; Palmer *et al.* 2001, 2005):

$$R_2^{obs} = R_2^0 + R_{ex}. \tag{8}$$

The value of R_{ex} becomes insignificant in the extreme limits of the fast and slow exchange regimes. For certain conditions, simple expressions of R_{ex} can be constructed for systems undergoing two-state exchange. In the slow exchange regime ($k_{ex} \ll \Delta\omega$), R_{ex} is given by (Palmer *et al.* 2001):

$$R_{ex,A} = k_{AB} = (1 - p_A)k_{ex} = p_B k_{ex}, \tag{9}$$
$$R_{ex,B} = k_{BA} = (1 - p_B)k_{ex} = p_A k_{ex}, \tag{10}$$

where $R_{ex,A}$ is R_{ex} measured for state A, $R_{ex,B}$ is R_{ex} measured for state B, p_A the population of state A (the observed state), and $p_B = 1 - p_A$ the population of state B (the state in exchange with state A). There are two different values of R_{ex} in the slow exchange regime because a separate peak is seen for each state. In the fast exchange regime ($k_{ex} \gg \Delta\omega$), R_{ex} is given by (Palmer *et al.* 2001):

$$R_{ex} = \frac{p_A p_B \Delta\omega^2}{k_{ex}}. \tag{11}$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Thus, while R_{ex} is independent of magnetic field strength (B_0) in the slow exchange regime, R_{ex} depends on the square of the magnetic field strength (i.e. $\Delta\omega^2$) in the fast exchange regime.

If the population of the minor state is very small ($p_A \gg p_B$), the following holds for all timescales (Swift & Connick, 1962):

$$R_{ex} = \frac{p_A p_B k_{ex}}{1 + (k_{ex}/\Delta\omega)^2} = \frac{p_A p_B k_{ex} \Delta\omega^2}{\Delta\omega^2 + k_{ex}^2}. \tag{12}$$

Under these conditions, the exchange-induced chemical shift (ω_{ex}) of state A – which is the difference in position of the peak in the exchanging system compared with its position in the absence of exchange – is given by:

$$\omega_{ex} = \frac{p_B k_{ex}^2 \Delta\omega}{\Delta\omega^2 + k_{ex}^2}. \tag{13}$$

If states A and B have different intrinsic transverse relaxation rates ($\Delta R_2^0 = R_{2,B}^0 - R_{2,A}^0 > 0$), then, for the case where $p_A \gg p_B$, R_{ex} is given by (Skrynnikov *et al.* 2002):

$$R_{ex} = k_{AB} \frac{\Delta R_2^0 \left(k_{BA} + \Delta R_2^0\right) + \Delta\omega^2}{\left(k_{BA} + \Delta R_2^0\right)^2 + \Delta\omega^2}. \tag{14}$$

Under these same conditions, the exchange-induced chemical shift is (Skrynnikov *et al.* 2002):

$$\omega_{ex} = k_{AB} \frac{k_{BA} \Delta\omega}{\left(k_{BA} + \Delta R_2^0\right)^2 + \Delta\omega^2}. \tag{15}$$

For a complete description of an exchanging system, it is necessary to invoke the McConnell equations (McConnell, 1958), a modified form of the Bloch equations (Bloch, 1946) extended to incorporate chemical exchange. Transverse relaxation in an exchanging system can be described by the following matrix equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M} = -(\mathbf{R} + \mathbf{K} - i\mathbf{\Omega})\mathbf{M}, \tag{16}$$

where $\mathbf{M}$ is the vector describing the transverse magnetization of each state, $\mathbf{R}$ the matrix describing the transverse relaxation rates of each state, $\mathbf{K}$ the matrix describing the kinetic rates of exchange between each state, and $\mathbf{\Omega}$ the matrix describing the chemical shift of each state. For a system undergoing two-site exchange between states A and B, these are given by:

$$\mathbf{M} = \begin{bmatrix} M_A^+ \\ M_B^+ \end{bmatrix}, \tag{17}$$

$$\mathbf{R} = \begin{bmatrix} R_{2,A} & 0 \\ 0 & R_{2,B} \end{bmatrix}, \tag{18}$$

$$\mathbf{K} = \begin{bmatrix} k_{AB} & -k_{BA} \\ -k_{AB} & k_{BA} \end{bmatrix}, \tag{19}$$

$$\mathbf{\Omega} = \begin{bmatrix} \Omega_A & 0 \\ 0 & \Omega_B \end{bmatrix}, \tag{20}$$

where M^+ is the transverse magnetization of each state (represented by a complex variable $Mx + iMy$) and $\Omega/2\pi$ is the chemical shift of each state. The McConnell equations can be solved (analytically for simple two-state cases (Palmer *et al.* 2001) or numerically for all cases by

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

integration or by calculating the matrix exponential) to provide a time course of the magnetization (i.e. free induction decay):

$$\mathbf{M}(t) = \mathbf{A} \times \mathbf{M}(0), \tag{21}$$

where $\mathbf{M}(t)$ is the transverse magnetization vector $\mathbf{M}$ at time t, $\mathbf{M}(0)$ is the initial value of $\mathbf{M}$ at $t = 0$, and $\mathbf{A}(t)$ is a matrix of time-dependent elements determined by solving the McConnell equations. Fourier transformation of Eq. (21) yields a frequency-domain NMR spectrum, from which apparent relaxation rates can be extracted by line shape fitting.

2.2.2 CPMG and rotating frame relaxation dispersion

Even a very small population of an otherwise invisible second state (as long as $p_b \geq 0.5\%$) can cause significant broadening of NMR peaks under conditions of intermediate exchange (in general τ_{ex} ~100 μs–10 ms), allowing characterization by relaxation dispersion NMR and insight into the otherwise invisible minor state (Baldwin & Kay, 2009; Korzhnev & Kay, 2008; Loria *et al.* 2008; Mittermaier & Kay, 2009; Neudecker *et al.* 2009; Palmer *et al.* 2001), using CPMG or $R_{1\rho}$ experiments. These experiments are performed by measuring the apparent R_2^{obs} of a nucleus (usually ^{15}N or ^{13}C, but also ^{1}H) while varying the repetition rate of 180° pulses (CPMG) or the strength of a spin lock field ($R_{1\rho}$) to attenuate the contribution of R_{ex} to the overall measured R_2 rate (Palmer *et al.* 2005).

In a CPMG relaxation dispersion experiment (Fig. 3), the effective observed R_2 ($R_{2,CP}^{obs}$) is measured as a function of the CPMG field (ν_{CP} [Hz]), which is the frequency of repetition pairs of 180° CPMG pulses and is related to τ_{CP} [s], the delay between 180° CPMG pulses, by:

$$\nu_{CP} = \frac{1}{2\tau_{CP}}. \tag{22}$$

At each CPMG field, the observed R_2 is given by:

$$R_{2,CP}^{obs} = R_2^0 + R_{ex,CP}, \tag{23}$$

where R_2^0 is the intrinsic transverse relaxation rate in the absence of exchange and $R_{ex,CP}$ is the contribution to transverse relaxation from chemical exchange broadening, which has been attenuated to some degree by CPMG pulses.

The parameters of interest are determined by fitting the CPMG data to a series of equations. Although the equations describing complicated systems involving three or more states can only be solved numerically, a system undergoing two-state exchange can be solved analytically (Carver & Richards, 1972; Davis *et al.* 1994; Jen, 1978; Palmer *et al.* 2001; Rule & Hitchens, 2006):

$$R_{2,CP}^{obs} = \frac{1}{2}\left(R_{2,A}^0 + R_{2,B}^0 + k_{ex} - [2\nu_{CP}]\ \cosh^{-1}\left[D_+ \cosh(\eta_+) - D_- \cos(\eta_-)\right]\right), \tag{24}$$

where

$$D_\pm = \frac{1}{2}\left(\pm 1 + \frac{\psi + 2\Delta\omega^2}{[\psi^2 + \zeta^2]^{1/2}}\right), \tag{25}$$

$$\eta_\pm = \frac{1}{(2\nu_{CP})\sqrt{2}}\left(\pm\psi + \left[\psi^2 + \zeta^2\right]^{1/2}\right)^{1/2}, \tag{26}$$

$$\psi = \left(R_{2,A}^0 - R_{2,B}^0 - p_A k_{ex} + p_B k_{ex}\right)^2 - \Delta\omega^2 + 4p_A p_B k_{ex}^2, \tag{27}$$

$$\zeta = 2\Delta\omega\left(R_{2,A}^0 - R_{2,B}^0 - p_A k_{ex} + p_B k_{ex}\right), \tag{28}$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

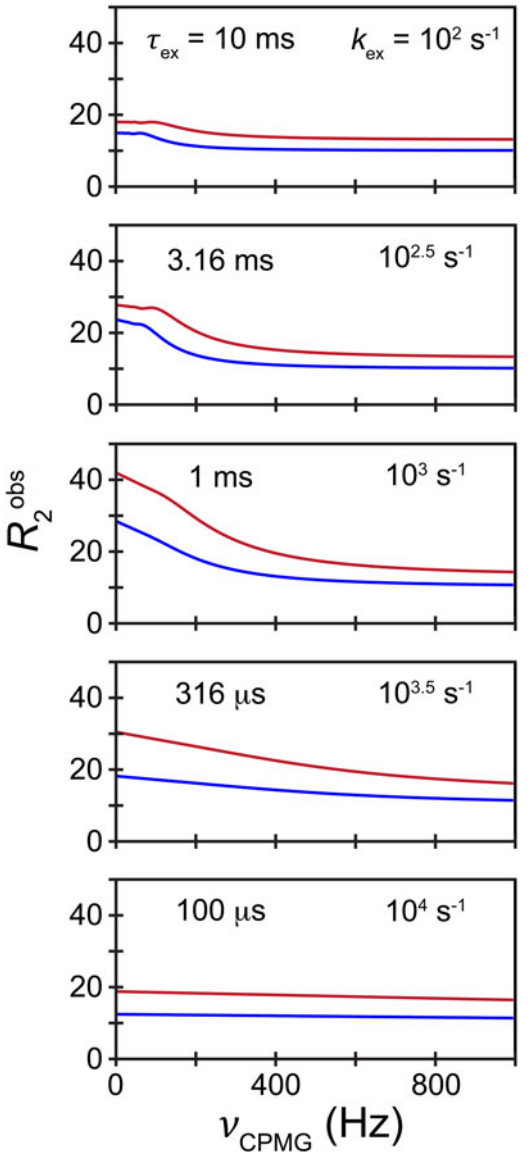


Fig. 3. Characterizing exchange with a 'dark' state by CPMG relaxation dispersion. Relaxation dispersion curves were simulated in MATLAB, showing the apparent observed transverse relaxation rate (R_2^{obs}) as a function of the repetition rate of 180° CPMG pulses (ν_{CPMG}, 0–1 kHz). Curves were calculated by solving the McConnell equations in MATLAB using Eq. (31). Curves are shown for range of exchange rates ($k_{ex} = 10^2$–10^4 s^{-1}) for a system where a visible species ($p_A = 95\%$) is in exchange with a dark state ($p_B = 5\%$), analogous to a ^{15}N resonance with a chemical shift difference between states A and B of $\Delta\delta = 2$ ppm. The blue curves correspond to a static B_0 field strength of 600 MHz ($\Delta\omega/2\pi = 120$ Hz), and the red curves correspond to 900 MHz ($\Delta\omega/2\pi = 180$ Hz). At a large ν_{CPMG}, R_2^{obs} approaches the intrinsic transverse relaxation rate R_2^0 (set here for both states to 10 s^{-1} at 600 MHz and 13 s^{-1} at 900 MHz). CPMG relaxation dispersion is most useful for characterizing exchange with a minor species when R_{ex} ($= R_2^{obs}[\nu_{CPMG} = 0] - R_2^{obs}[\nu_{CPMG} \rightarrow \infty]$) is largest, which occurs when $k_{ex} = \Delta\omega$ (~1000 s^{-1}, middle panel).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

and where $R^0_{2,A}$ is the R_2 of state A in the absence of chemical exchange (i.e. $\nu_{CP} = \infty$), $R^0_{2,B}$ is the R_2 of state B in the absence of chemical exchange, p_A is the population of state A, p_B is the population of state B, and $\Delta\omega$ is the difference in chemical shift between states A and B.

Simplified expressions can be written for $R^{obs}_{2,CP}$ under certain scenarios. In the fast exchange regime, the expression for $R^{obs}_{2,CP}$ can be simplified to (Palmer *et al.* 2001):

$$R^{obs}_{2,CP} = p_A R^0_{2,A} + p_B R^0_{2,B} + \left(\frac{p_A p_B \Delta\omega^2}{k_{ex}}\right)\left(1 - \frac{2[2\nu_{CP}]}{k_{ex}} \tanh\left[\frac{k_{ex}}{2(2\nu_{CP})}\right]\right). \quad (29)$$

If $p_A \gg p_B$, regardless of the timescale, the expression for R^{obs}_2 can be simplified to (Ishima & Torchia, 1999):

$$R^{obs}_{2,CP} = R^0_2 + \frac{p_A p_B \Delta\omega^2 k_{ex}}{k^2_{ex} + \left(p^2_A \Delta\omega^4 + 48[2\nu_{CP}]^4\right)^{1/2}}, \quad (30)$$

where R^0_2 is equal to $R^0_{2,A}$ in the slow exchange regime and $p_A R^0_{2,A} + p_B R^0_{2,B}$ in the fast exchange regime.

For systems more complex than a two-state system, R^{obs}_2 is not given by an analytical expression, but can only be calculated by numerical solution of the McConnell equations. For any system, the transverse magnetization at the end of a CPMG period (T_{CP}) consisting of an even number ($2n$) of 180° pulses, each separated by time τ_{CP}, is given by (Palmer *et al.* 2001):

$$\mathbf{M}(2n\tau_{CP}) = (\mathbf{A}\mathbf{A}^*\mathbf{A}^*\mathbf{A})^n \, \mathbf{M}(0), \quad (31)$$

where $\mathbf{M}(2n\tau_{CP})$ is the transverse magnetization vector at time $T_{CP} = 2n\tau_{CP}$ and $\mathbf{A}^*$ is the complex conjugate of $\mathbf{A}$ (Eq. (21)) when the McConnell equations have been solved for $t = \tau_{CP}/2$ (Palmer *et al.* 2001).

The CPMG relaxation dispersion curve has a complex shape that is a function of k_{ex}, the populations of the states undergoing exchange, and the chemical shift difference between these states. If the system is in exchange between two states, a simple model can be used to fit the data; more complicated multi-state models can then be employed if the two-site fit is not adequate, although one is generally limited to three-site exchange models to improve the fit without overfitting the data (Korzhnev & Kay, 2008; Li *et al.* 2011; Mittermaier & Kay, 2009; Neudecker *et al.* 2009; Trott & Palmer, 2004). The contribution of the populations and chemical shift differences to R_{ex} cannot be separated *a priori* for data recorded at a single magnetic field. Thus, unless the populations or chemical shift differences are already known, one must record data at different B_0 magnetic field strengths, because the populations of the species are independent of the magnetic field, while the differences in frequency are linearly proportional to the magnetic field.

A system in intermediate chemical exchange can also be characterized by rotating frame $R_{1\rho}$ relaxation dispersion, fitting the effective $R_{1\rho}$ ($= 1/T_{1\rho}$) as a function of the effective spin-lock field strength (ω_e). The observed $R_{1\rho}$ for a given spin-lock field ($R^{obs}_{1\rho}$) is given by the sum of $R_{1\rho}$ in the absence of exchange ($R^0_{1\rho}$) plus the contribution to $R_{1\rho}$ from chemical exchange broadening ($R_{ex,1\rho}$), which has been attenuated to some degree by the spin-lock field:

$$R^{obs}_{1\rho} = R^0_{1\rho} + R_{ex,1\rho}. \quad (32)$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

For two-state exchange in the fast exchange regime, the expression for $R_{1\rho}^{\mathrm{obs}}$ is given by (Davis *et al.* 1994; Deverell *et al.* 1970; Meiboom, 1961; Palmer *et al.* 2001; Rule & Hitchens, 2006):

$$R_{1\rho}^{\mathrm{obs}} = R_{1\rho}^{0} + \frac{p_{\mathrm{A}} p_{\mathrm{B}} \Delta\omega^2 k_{\mathrm{ex}}}{k_{\mathrm{ex}}^2 + \omega_{\mathrm{e}}^2} \sin^2\theta, \tag{33}$$

where $R_{1\rho}^{0}$ is $R_{1\rho}$ in the absence of chemical exchange (i.e. $\omega_{\mathrm{e}} = \infty$) and the angle between the B_1 spin-lock field and the static B_0 field is given by:

$$\theta = \tan^{-1}\frac{\omega_1}{\delta\omega}, \tag{34}$$

where ω_1 (rad·s^{-1}) is the strength of the spin-lock field and $\delta\omega$ (rad·s^{-1}) is the frequency offset (the difference between the resonance frequency of the nucleus and the transmitter frequency). The effective spin-lock field strength is given by:

$$\omega_{\mathrm{e}} = \sqrt{\omega_1^2 + \delta\omega^2}. \tag{35}$$

If $p_{\mathrm{A}} \gg p_{\mathrm{B}}$ and the spin-lock field is on resonance ($\delta\omega = 0$), $R_{1\rho}^{\mathrm{obs}}$ can be expressed regardless of the timescale as (Meiboom, 1961):

$$R_{1\rho}^{\mathrm{obs}} = R_{1\rho}^{0} + \frac{p_{\mathrm{A}}^2 p_{\mathrm{B}} \Delta\omega^2 k_{\mathrm{ex}}}{k_{\mathrm{ex}}^2 + p_{\mathrm{A}}^2 \Delta\omega^2 + \omega_1^2}, \tag{36}$$

where $R_{1\rho}$ in the absence of chemical exchange in the slow exchange regime is given by:

$$R_{1\rho}^{0} = R_{1,\mathrm{A}}^{0} \cos^2\theta + R_{2,\mathrm{A}}^{0} \sin^2\theta \tag{37}$$

and in the fast exchange regime is given by:

$$R_{1\rho}^{0} = \left(p_{\mathrm{A}} R_{1,\mathrm{A}}^{0} + p_{\mathrm{B}} R_{1,\mathrm{B}}^{0}\right)\cos^2\theta + \left(p_{\mathrm{A}} R_{2,\mathrm{A}}^{0} + p_{\mathrm{B}} R_{2,\mathrm{B}}^{0}\right)\sin^2\theta. \tag{38}$$

For off-resonance spin-lock fields, as long as the population of the minor state is very small ($p_{\mathrm{A}} \gg p_{\mathrm{B}}$), $R_{1\rho}^{\mathrm{obs}}$ can be expressed in a form analogous to Eq. (30) (Palmer, 2014):

$$R_{1\rho}^{\mathrm{obs}} = R_{1\rho}^{0} + \frac{p_{\mathrm{A}} p_{\mathrm{B}} \Delta\omega^2 k_{\mathrm{ex}}}{k_{\mathrm{ex}}^2 + \omega_{\mathrm{e,A}}^2 \omega_{\mathrm{e,B}}^2 / \omega_{\mathrm{e,ave}}^2}, \tag{39}$$

where $\omega_{\mathrm{e,A}}$ is the effective spin-lock field (Eq. (35)) for the chemical shift of state A, $\omega_{\mathrm{e,B}}$ for state B, and $\omega_{\mathrm{e,ave}}$ is the effective spin-lock field for the population-averaged chemical shift of states A + B. If the intrinsic transverse relaxation rates differ between the two states ($\Delta R_2^0 = R_{2,\mathrm{A}} - R_{2,\mathrm{B}} > 0$), $R_{1\rho}^{\mathrm{obs}}$ can be given by the following expression, which is accurate as long as $p_{\mathrm{A}} \gg p_{\mathrm{B}}$ (Palmer, 2014):

$$R_{1\rho}^{\mathrm{obs}} = R_{1\rho}^{0} + k_{\mathrm{AB}} \frac{k_{\mathrm{BA}}\left(\Delta\omega^2 + \left[\Delta R_2^0\right]^2\right) + \Delta R_2^0\left(\omega_{\mathrm{e,A}}^2 + k_{\mathrm{BA}}^2\right)}{k_{\mathrm{BA}}\left(\Delta\omega_{\mathrm{e,B}}^2 + k_{\mathrm{BA}}^2 + \left[\Delta R_2^0\right]^2\right) + \Delta R_2^0 \omega_1^2}. \tag{40}$$

A more comprehensive analytical expression is given by Baldwin & Kay (2013). Under all circumstances, though, values of $R_{1\rho}^{\mathrm{obs}}$ can be calculated by solving the McConnell equations using the same matrices as for a saturation transfer experiment (cf. Eqs. (82)–(84) below).

An advantage of $R_{1\rho}$ experiments is that they can characterize faster exchange processes than CPMG experiments. Whereas CPMG data alone can be used to characterize exchange on a timescale of roughly ~500 μs–5 ms, $R_{1\rho}$ relaxation dispersion experiments can measure exchange

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

processes as fast as $k_{ex} < 40\,000\ s^{-1}$ ($\tau_{ex} > 25\ \mu s$) (Ban *et al.* 2012; Hansen *et al.* 2009; Lundstrom & Akke, 2005; Weininger *et al.* 2013). The remainder of this section, however, focuses on CPMG relaxation dispersion.

2.2.3 Practical aspects of CPMG

From a series of CPMG experiments, one can determine the populations of the major and minor states, the rate of exchange between the states, and the magnitude of the chemical shift difference between the states, which offers indirect structural insight into the minor state. However, CPMG data alone only directly provide information on the magnitude, but not direction, of chemical shift differences. The direction of the chemical shift differences can be determined by either comparing the chemical shift of the resonance between heteronuclear single (HSQC) and multiple (HMQC) quantum coherence spectra or between spectra acquired at different B_0 fields (Bouvignies *et al.* 2010; Skrynnikov *et al.* 2002). As $\Delta\omega$ is directly proportional to B_0, moving from a lower to a higher B_0 field has the effect of decreasing $k_{ex}/\Delta\omega$, thus slowing the exchange regime and moving the frequency of the resonance closer to its intrinsic chemical shift (slow exchange) and farther from the population-averaged chemical shift (fast exchange). The difference in chemical shift observed between HSQC and HMQC spectra arises from the fact that, for ^{15}N for example, the exchange regime in the HSQC experiment is characterized by $k_{ex}/\Delta\omega_N$, but in the HMQC, exchange of zero-quantum and double-quantum magnetization are characterized by $k_{ex}/(\Delta\omega_N - \Delta\omega_H)$ and $k_{ex}/(\Delta\omega_N + \Delta\omega_H)$, respectively, which do not average linearly, thus altering the exchange regime and giving rise to a small difference in chemical shift between the HSQC and HMQC spectra (Skrynnikov *et al.* 2002). Likewise, if there is any direct knowledge of the chemical shift of the minor state from other measurements (e.g. ligand-binding studies, protein-folding/unfolding studies, concentration dependence, etc.), such information on the magnitude of chemical shift change can assist the CPMG fitting process.

The fitting of CPMG data is prone to difficulties, and distinguishing between k_{ex}, $\Delta\omega$, and p_B requires resonances undergoing exchange in the intermediate regime, with at least some satisfying the condition $k_{ex}/\Delta\omega < 2$; this limits the range of k_{ex} values measurable from CPMG alone to between ~200 and ~2000 s^{-1}. However, the incorporation of additional chemical shift information (e.g. from HSQC *versus* HMQC spectra) can extend this range further into the fast timescale, allowing fitting of data in the fast-intermediate exchange regime, up to $k_{ex}/\Delta\omega \sim 4$ (in practice, exchange rates up to $k_{ex} \sim 6000\ s^{-1}$; $\tau_{ex} \sim 170\ \mu s$) (Vallurupalli *et al.* 2011). The study of fast exchange processes by CPMG relaxation dispersion can be problematic in general, leading to large errors when the repetition rate ($\nu_{CP} = 1/\tau_{CP}$) is not much greater than k_{ex}. However, it has been demonstrated that if this condition is not satisfied, these errors can be reduced dramatically (from ~50 to ~5%) by measuring the intrinsic relaxation rate R_2^0 using an $R_{1\rho}$ experiment with a high spin-lock field (Ban *et al.* 2013).

Relaxation dispersion experiments on proteins are most commonly performed on the backbone ^{15}N nucleus, although measurements on ^1H_N, $^{13}C_\alpha$, $^1H_\alpha$, $^{13}C_\beta$, and $^{13}C'$ backbone nuclei and side chain methyl ^1H and ^{13}C nuclei (Hansen *et al.* 2008b; Ishima & Torchia, 2003; Lundstrom *et al.* 2007, 2009a, 2009b) are also feasible. Minor state chemical shifts derived from CPMG studies can be used to calculate backbone torsion angle restraints for the minor state using a program such as TALOS-N (Shen & Bax, 2013), and it is possible to calculate the structure of a small protein from backbone chemical shifts alone, using programs such as CHESHIRE (Cavalli *et al.* 2007) or CS-Rosetta (Shen *et al.* 2008, 2009), which rely upon matching

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

the empirical chemical shifts with a peptide fragment database. Such an approach can provide direct insight into the structure of a minor state (Korzhnev *et al.* 2010), although complications can arise if the minor state is partially unfolded, as the databases used for calculation of structures from chemical shifts are derived from folded proteins. It has also been demonstrated that the orientation of bond vectors in minor states can be measured through relaxation dispersion experiments (Hansen *et al.* 2008a; Vallurupalli *et al.* 2007, 2008a, b). By performing relaxation dispersion experiments in weakly aligned media, one can measure RDCs and residual chemical shift anisotropy (RCSA) for the minor state, although highly accurate measurements are required to detect the small effect of the minor state on these observables. Additionally, the introduction of a paramagnetic group with an anisotropic *g*-tensor will magnify $\Delta\omega$ if different PCSs are observed between the different states, potentially causing additional exchange broadening and facilitating CPMG relaxation dispersion studies. This strategy has been applied to intrinsic metalloproteins (Wang *et al.* 2007) and to otherwise diamagnetic proteins by incorporating a rigid metal-chelating paramagnetic tag to induce relaxation dispersion when none was apparent in the diamagnetic state of the protein (Eichmuller & Skrynnikov, 2007; Hass *et al.* 2010).

For larger proteins, the favorable spectral properties of methyl groups make methyl ^{13}C, ^{1}H, and ^{2}H (D) nuclei useful probes of dynamics (Igumenova *et al.* 2006; Ruschak & Kay, 2010). Methods have been developed to express recombinant proteins in *Escherichia coli* with various methyl isotopic labeling schemes (Tugarinov *et al.* 2006). Favorable relaxation properties are seen in $^{13}CDH_2$-labeled methyl groups, and, for $^{13}CH_3$-labeled methyl groups on a perdeuterated background, acquiring data with an HMQC experiment instead of an HSQC experiment allows one to take advantage of the slower multiple quantum relaxation due to cancellation of dipolar interactions within the methyl group (i.e. the 'methyl-TROSY' effect), helping make large systems accessible to analysis by NMR (Tugarinov *et al.* 2003). Using these methods, ^{13}C CPMG relaxation dispersion experiments have been performed to characterize intermediate dynamics in systems as large as the 82-kDa enzyme malate synthase G (Korzhnev *et al.* 2004a), the 230-kDa HsIV component of the bacterial proteasome (Shi & Kay, 2014), and the 670-kDa 20S core particle of the archaeal proteasome (Sprangers & Kay, 2007). Using a labeling scheme that isolates ^{1}H spins, ^{1}H CPMG relaxation dispersion experiments can also be performed on methyl groups (Hansen *et al.* 2012; Otten *et al.* 2010).

2.2.4 Applications of CPMG

CPMG relaxation dispersion is well suited for studying protein folding, due to its ability to identify minor states (i.e. folding intermediates) in exchange with a visible (folded or unfolded) species on an intermediate timescale (Fig. 3). These folding intermediates occupy local minima (i.e. 'frustrations') (Ferreiro *et al.* 2005) on an otherwise smooth, minimally frustrated protein-folding energy landscape (Bryngelson & Wolynes, 1987; Onuchic *et al.* 1997). CPMG relaxation dispersion studies on the folding of the SH3 domain from Fyn have identified an intermediate in the folding pathway (populated at ~1–2% at equilibrium), which has evaded detection by other methods and contains both native (central β-sheet) and non-native structural elements (Korzhnev *et al.* 2004b, 2006; Neudecker *et al.* 2007, 2009). Backbone ^{1}H, ^{13}C, and ^{15}N chemical shifts and ^{1}H–^{15}N RDCs from CPMG relaxation dispersion were used to determine the structure, using CS-Rosetta, of an intermediate in the folding pathway of the FF domain from HYPA/FBP11, revealing a folding core comprising helices H1 and H2, and also other non-native interactions that slow the kinetics of folding (Korzhnev *et al.* 2010). A similar approach was used

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

to solve the structure of a folding intermediate of a mutant of the Fyn SH3 domain, revealing an exposed β-strand that could potentially participate in protein misfolding and aggregation (Neudecker *et al.* 2012). The structure of a transient state of T4 lysozyme with an occluded binding cavity was also revealed in this manner (Bouvignies *et al.* 2011). ^{15}N CPMG relaxation dispersion data, in conjunction with chemical shift titrations, have been used to explore how an intrinsically unstructured peptide (pKID) undergoes folding upon binding to its target, the KIX domain of the CREB-binding protein (Sugase *et al.* 2007); binding proceeds through an intermediate state consisting of partially folded pKID bound to KIX, consistent with an induced fit mechanism of binding (Wright & Dyson, 2009).

CPMG relaxation dispersion has proven to be a valuable tool for exploring the connection in enzymes between the catalytic rate and internal conformational dynamics (Villali & Kern, 2010), an area that is the subject of a lively scientific debate (Kamerlin & Warshel, 2010; Karplus, 2010; Pisliakov *et al.* 2009). Results from CPMG relaxation dispersion studies have demonstrated close correspondence between the k_{ex} of conformational exchange and catalytic turnover (k_{cat}) for a variety of enzymes, including HIV protease (Ishima *et al.* 1999), RNase A (Cole & Loria, 2002; Watt *et al.* 2007), adenylate kinase (Adk) (Wolf-Watz *et al.* 2004), cyclophilin A (Eisenmesser *et al.* 2005), dihydrofolate reductase (DHFR) (Boehr *et al.* 2006, 2008, 2010), flavin oxidoreductase (Vallurupalli & Kay, 2006), and triosphosphate isomerase (Kempf *et al.* 2007). DHFR undergoes exchange with a minor open conformation at a rate that is identical to product dissociation, which is the rate-limiting step in catalysis (Boehr *et al.* 2006); its hydride transfer step is also closely tied to the dynamics of conformational exchange (Boehr *et al.* 2008). Likewise, it has been demonstrated by CPMG relaxation dispersion that the catalytic tunover (k_{cat}) of Adk is limited by the rate of opening of the nucleotide-binding lid (Wolf-Watz *et al.* 2004). Thus, for processes that occur on an intermediate timescale, relaxation dispersion offers exceptional insight into molecular dynamics, and in some cases into the structures of otherwise invisible minor states.

2.3 Saturation transfer

2.3.1 Principles and applications of saturation transfer

States that are invisible because they are sparsely populated or because they involve large, slowly tumbling molecules can be detected by saturation transfer if there are resonances in the minor state that can be saturated by off-resonance radiofrequency (RF) radiation that does not saturate the resonances of the major state; this saturation can then be transferred to the visible species by chemical exchange. Resonances can be selectively saturated in the dark state if their chemical shifts are sufficiently different from those in the visible state or if the dark state resonances are broadened considerably compared with those in the visible state. The DEST experiment can be used to characterize larger assemblies and aggregates that are invisible because they tumble slowly (causing large transverse relaxation rates and significant line broadening), as long as the large invisible state is in exchange with a smaller visible state on an intermediate-to-slow timescale (500 μs–1 s) (Fawzi *et al.* 2010b, 2011b, 2012; Libich *et al.* 2013). The CEST experiment, on the other hand, can be used to visualize a minor state in exchange with a major state, as long as the lifetime of the minor state is ~2–100 ms, its rotational correlation time is comparable to the major species, and it exhibits significantly different chemical shifts from the major state (Vallurupalli *et al.* 2012). Both experiments involve the application of off-resonance RF radiation to saturate a nucleus (usually ^{15}N, but also ^{13}C and ^{1}H) in the minor state (Fig. 4).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

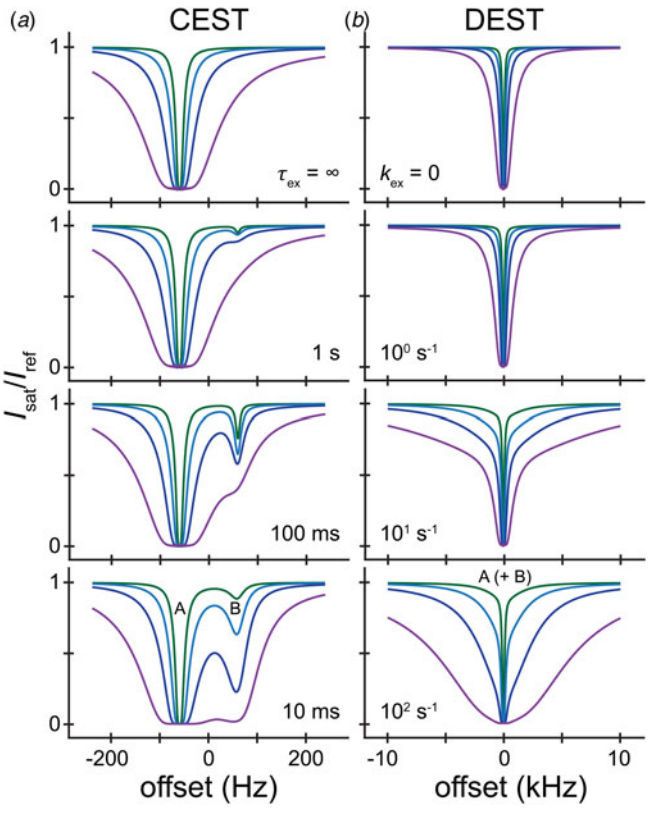


Fig. 4. Characterizing 'dark' states by exchange saturation transfer. Saturation profiles are plotted for (a) CEST and (b) DEST experiments by solving the homogeneous McConnell equations (Eqs. (82)–(84)) in MATLAB. The *y*-axis gives the ratio of the peak intensity measured in the presence of a CW RF B_1 field (I_{sat}) over the peak intensity measured for a reference experiment in the absence of the saturation field (I_{ref}). The *x*-axis gives the frequency offset of the applied B_1 field. Note that the units are (a) in Hz for the CEST profiles, but are (b) in kHz for the DEST profiles, where the *x*-axis has been expanded almost 50 fold to accommodate the broader DEST profiles. B_1 field strengths (applied in the *x*-dimension) are (a) 2·5 Hz (green), 5 Hz (cyan), 10 Hz (blue), and 25 Hz (purple) for CEST and (b) 25 Hz (green), 50 Hz (cyan), 100 Hz (blue), and 250 kHz (purple) for DEST (i.e. the DEST experiment is performed at about one order of magnitude higher B_1 field strength than CEST). In this figure, a major visible state A ($p_A = 95\%$) is separated from a minor dark state B ($p_B = 5\%$) by a chemical shift difference of $\Delta\omega/2\pi = 120$ Hz (e.g. 2 ppm for ^{15}N in a 600 MHz spectrometer). For the CEST profiles (a), $R_2^0 = 10\ s^{-1}$ and $R_1 = 1\ s^{-1}$ for both states. For the DEST profiles (b), $R_{2,B}^0 = 10\,000\ s^{-1}$. CW saturation was applied for 1 s ($= 1 \times T_1$). Results are shown for a system undergoing no exchange (top; i.e. the saturation profile of the major visible species A alone), and exchange at a rate of $k_{ex} = 1$, 10, and $100\ s^{-1}$ (bottom). Full saturation of the signal ($I_{sat}/I_{ref} = 0$) occurs when the B_1 field is applied on resonance to state A, and saturation can be transferred from states B to A by exchange either from (a) on-resonance saturation of state B due to a different chemical shift (CEST) or (b) off-resonance saturation of the broader profile of state B due to its faster rate of transverse relaxation (DEST).

A DEST or CEST experiment yields a plot of signal intensity *versus* the chemical shift offset of the B_1 saturating field for each resonance being monitored. These profiles can then be fit to a model (based on the McConnell equations (McConnell, 1958)) to determine the populations of and rates of exchange between each state. In addition to these kinetic and thermodynamic parameters, the DEST experiment also provides information on the transverse relaxation rates

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

(R_2) of the resonances in the dark state (Fawzi *et al.* 2011b, 2012). The CEST experiment provides direct information on the chemical shifts of the minor state (Bouvignies & Kay, 2012a, b; Vallurupalli *et al.* 2012), which are difficult to extract from DEST data due to the broadness of the saturation transfer profiles (Fawzi *et al.* 2012).

NMR saturation transfer experiments were originally developed in the 1960s (Forsén & Hoffman, 1963), and the DEST and CEST experiments are related to small molecule saturation transfer difference (STD) NMR spectroscopy, where the broader resonances in a larger macromolecule are saturated, and saturation is transferred to a bound ligand via cross-relaxation and ultimately to the NMR-visible unbound ligand via chemical exchange (Mayer & Meyer, 1999). Saturation transfer found some early biomolecular applications (Cayley *et al.* 1979; Gupta & Redfield, 1970), although CEST was first employed in MRI, where exchangeable protons with a chemical shift distinct from water (either in endogenous molecules or introduced CEST contrast agents) are saturated, and this saturation is transferred to bulk water via slow chemical exchange, thus amplifying these otherwise invisible signals to provide additional contrast (Van Zijl & Yadav, 2011; Ward *et al.* 2000). More recently, PCSs from paramagnetic groups (chelated lanthanides) have been used to generate even larger chemical shifts in paraCEST imaging, enabling faster timescales for which $k_{ex} \ll \Delta\omega$ remains true (Hancu *et al.* 2010; Vinogradov *et al.* 2013). Interestingly, one of the earliest biomolecular applications of saturation transfer involved using the hyperfine shifts in Fe^{3+}-bound cytochrome *c* to study less-well-resolved Fe^{2+}-bound cytochrome *c* via chemical exchange (Gupta & Redfield, 1970). More recently, the applications of saturation transfer have been expanded to observe invisible states in macromolecules by CEST (Bouvignies & Kay, 2012a, b; Vallurupalli *et al.* 2012) and DEST (Fawzi *et al.* 2011b, 2012, 2014; Libich *et al.* 2013).

2.3.2 The DEST experiment

The DEST experiment is used to characterize a visible species in exchange with a dark state that is invisible due to its large size (Fawzi *et al.* 2010b, 2011b, 2012, 2014; Libich *et al.* 2013). Owing to the rapid transverse relaxation rate of the signals in a slowly tumbling large dark state, these resonances can be saturated far off resonance at frequencies that do not excite the narrow signals of the rapidly tumbling visible species in the absence of the larger dark state. If the invisible species converts to the visible species at a rate that is of the order of or faster than its longitudinal relaxation rate, saturation will be transferred to the NMR-visible species and observed as a loss of intensity of cross-peaks for the visible species. Thus, it becomes possible to detect and characterize large aggregates and macromolecular assemblies that would otherwise be invisible. The DEST experiment is explored in greater detail in Section 3.

DEST is complementary to the transferred nuclear Overhauser effect (TRNOE) (Clore & Gronenborn, 1982, 1983), which offers structural restraints (in the form of short interproton distances) for a bound state in fast exchange with the free state. The TRNOE, however, is really only applicable in the case of exchanging systems involving small (low molecular weight) ligands that exhibit positive (or negligibly small) NOEs in the free state and large negative NOEs in the bound (high-molecular weight) state. DEST is qualitatively similar to other saturation-based experiments, including CEST and STD (Post, 2003). However, unlike STD, DEST offers quantitative results when performed on a heteronucleus (e.g. ^{15}N and ^{13}C), as it is not complicated by spin diffusion, which occurs in ^{1}H saturation experiments (Bodner *et al.* 2009; Fawzi *et al.* 2010b).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

The DEST experiment has revealed details of the aggregation of amyloid-β (Aβ) peptides into protofibrils (Fawzi *et al.* 2010b, 2011b, 2014), demonstrating that Aβ reaches a pseudo-equilibrium with protofibrils, with free peptides transiently binding to the surface of the protofibrils with a lifetime of ∼20 ms. During this process, the more hydrophobic segments of the peptide (residues 16–25 and 28–37) remain more tightly associated with the surface of the protofibrils than the other portions. These protofibrils were shown to have R_2 relaxation rates of ∼19 000 s^{-1}, consistent with the size distribution of 1·8–85 MDa observed by AFM, EM, analytical ultracentrifugation, and dynamic light scattering (Fawzi *et al.* 2011b). DEST has also been used to characterize the interaction of the Aβ peptide with the 780-kDa GroEL chaperonin (Libich *et al.* 2013), revealing that Aβ transiently interacts with the surface of the large accessible internal cavity of the chaperone on a faster (∼700 μs) timescale, using the hydrophobic portions of its sequence.

DEST experiments are performed with relatively large (>100 Hz) B_1 saturation field strengths to allow maximal sensitivity to the dark state. Unlike in CEST, where the need to resolve a difference in chemical shift between the visible and dark states puts an upper limit on the saturation field strength, no such upper limit exists for DEST. Higher saturation fields are needed to monitor slower exchange processes, and those involving a smaller dark state population or lower R_2^{dark}. In the example shown in Fig. 4, DEST is effective at revealing the dark state at exchange rates of 10 and 100 s^{-1}. DEST could be measured on a more slowly exchanging system if the population of the dark state were larger. DEST can also be measured on a more rapidly exchanging system; however, measurement of fast exchanging systems ($k_{\text{ex}} > 1000\ s^{-1}$) is only possible for smaller values R_2^{dark}. Otherwise, R_2^{obs} would be too fast and there would be no measurable NMR signal.

2.3.3 The CEST experiment

The CEST experiment (Bouvignies & Kay, 2012a, b; Vallurupalli & Kay, 2013; Vallurupalli *et al.* 2012) provides information similar to relaxation dispersion, but for processes on a slower timescale. If the visible species is in exchange with a dark state that has a significantly different chemical shift, application of a continuous-wave (CW) RF field at the invisible state resonance will cause a loss in intensity of the corresponding resonance in the NMR-visible major species (as long as the minor state has a lifetime of ∼2–100 ms and is populated ≥1%; Vallurupalli *et al.* 2012). By varying the offset of the saturating CW RF field, one can acquire, effectively, a ^{1}H, ^{15}N, or ^{13}C spectrum of the minor state, depending on which nucleus is used for saturation. The CEST data can be fit to the McConnell equations, providing kinetic parameters and populations for each state (thus occupying a niche similar to that of CPMG relaxation dispersion but for processes undergoing slower exchange).

The observation of CEST is dependent upon a chemical shift difference between the two states, and the maximum field strength that can be used in CEST is limited by the chemical shift difference between the two states. A higher B_1 field improves the sensitivity of CEST to a minor state, but decreases chemical shift resolution between the two states. In the example, in Fig. 4, a field strength of 10 Hz is best suited for observing the minor state B, which can be characterized if the exchange rate is 10 or 100 s^{-1}. However, the sensitivity is too low under conditions of slower exchange ($k_{\text{ex}} = 1\ s^{-1}$) to accurately observe state B. Characterization of exchange much faster than $k_{\text{ex}} = 100\ s^{-1}$ would not be possible because distinct chemical shifts are not observed for each state in the intermediate or fast exchange regimes. Higher saturation field strengths can be used in the presence of greater chemical shift differences between the major and minor states, and, in practice, most CEST studies have

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

used field strengths of ∼20–30 Hz (Bouvignies *et al.* 2014). Note that much larger field strengths can be used in DEST experiments because the results are not dependent upon resolving chemical shift differences between the light and dark states.

CEST experiments on proteins have been performed on ^{15}N and ^{13}C nuclei. If ^{13}C-CEST is performed on uniformly ^{13}C-labeled samples, the strong ^{13}C–^{13}C scalar couplings must be taken into account (Bouvignies *et al.* 2014; Vallurupalli *et al.* 2013). Owing to the narrow linewidths of the peaks in the minor state, CEST experiments require a large number of saturation field offset frequencies (e.g. ^{15}N CEST experiments have been performed with offsets at uniform increments of 15–25 Hz, requiring ≥60 total offsets to cover the ^{15}N chemical shift range of ∼30 ppm (Vallurupalli *et al.* 2012)). Owing to the broader linewidths, DEST experiments require a larger range of saturation frequencies, but a smaller total number (e.g. ^{15}N-DEST experiments on Aβ were performed with non-uniform increments of 2–7 kHz between offsets requiring just 15 total offsets to cover a large frequency range of 70 kHz, which is >1000 ppm in a 600 MHz static ^{1}H field (Fawzi *et al.* 2011b)).

The CEST experiment is applicable to studying slow exchange processes, which include protein folding. Exchange between the visible folded species of a mutant of the HYPA/FBP11 FF domain and an intermediate along its folding pathway has been characterized by ^{15}N-CEST (Vallurupalli *et al.* 2012) and carbonyl ^{13}C-CEST (Vallurupalli & Kay, 2013), yielding a best-fit exchange rate of $k_{ex} \approx 60\ s^{-1}$ and a minor state population of 1·4%. ^{15}N-DEST and methyl ^{13}C-DEST studies have also characterized the exchange between a mutant of the Fyn SH3 domain and a folding intermediate (Bouvignies & Kay, 2012a). FF domain and SH3 domain folding have also been characterized by aliphatic ^{13}C-CEST, allowing the measurement of most side chain ^{13}C chemical shifts in the sparsely populated folding intermediates (Bouvignies *et al.* 2014). Additionally, ^{15}N-CEST has been used to characterize slow conformational dynamics in active Ras-GTPases under low-temperature conditions (used to slow intrinsic GTP hydrolysis), where exchange was too slow and the population of the minor state too small to give reliable CPMG relaxation dispersion data (Long *et al.* 2013).

2.4 Lifetime line broadening

If a smaller NMR-visible species is in exchange with a much larger NMR-invisible species, an enhancement to the measured transverse relaxation rate (R_2^{obs}) of the visible species will be observed (Fig. 5). Qualitatively, this lifetime line broadening (R_{llb}) indicates the presence of an interaction with a large assembly or aggregate, but R_{llb} can also be used to quantitatively analyze the interaction under certain conditions, especially when combined with DEST measurements (Fawzi *et al.* 2010b, 2011b, 2012, 2014; Libich *et al.* 2013). It should be noted that R_{llb}, which arises from exchange between two states with a large difference in transverse relaxation rate, is distinct from chemical exchange broadening (R_{ex}), which arises from exchange between two states with a large difference in chemical shift, and both effects can be present in a given system, depending on the timescale of the exchange process (Fig. 5).

If the transverse relaxation rate in the dark state (R_2^{dark}) is sufficiently fast (due to a very large size), any transverse magnetization that is transferred from the visible state (A) to the dark state (B) will effectively be destroyed – as long as the rate of exchange from the dark to light state (k_{off}) is sufficiently slow (i.e. $R_2^{dark} \gg k_{off}$). If the rate of exchange from A to B is of the order of, or faster than the rate of longitudinal relaxation – but of the order of, or slower than the rate of transverse relaxation of the visible species (i.e. $R_1^{visible} \leq k_{on}^{app} \leq R_2^{visible}$) – then the apparent

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

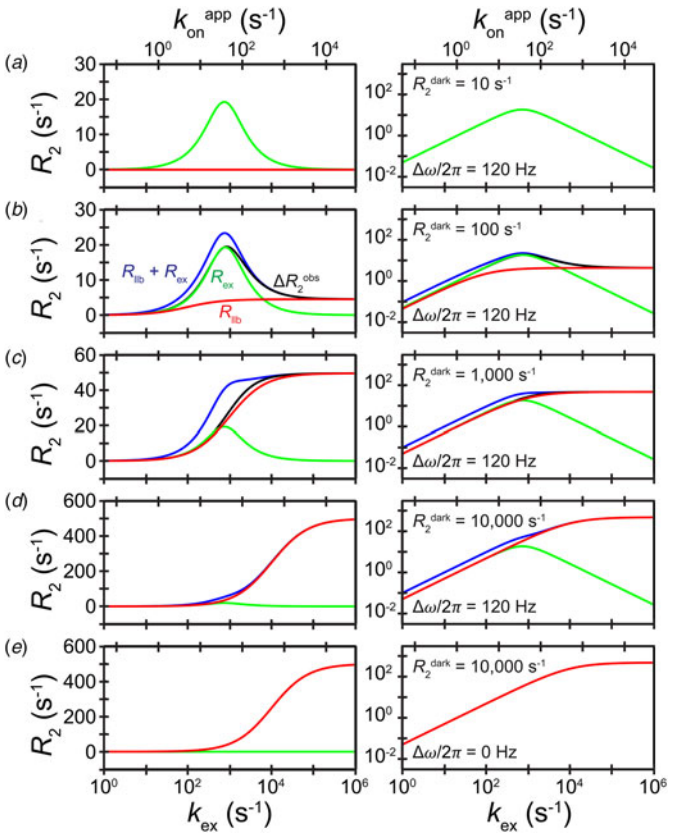


Fig. 5. Enhanced apparent observed transverse relaxation rate (R_2^{obs}) due to exchange with a minor species. Lifetime line broadening (R_{llb}, red), chemical shift exchange broadening (R_{ex}, green), their sum ($R_{llb} + R_{ex}$, blue), and the observed enhancement in apparent transverse relaxation rate (ΔR_2^{obs}, black) compared with a system in the absence of exchange are plotted *versus* exchange rate. Curves are shown for a small, rapidly tumbling visible species ($p_A = 95\%$, $R_{2,A}^0 = 10\ s^{-1}$) in exchange with a sparsely populated 'dark' state ($p_B = 5\%$). On the left, simulated data are plotted using a linear scale for the *y*-axis (note that the scale varies between plots); on the right, a logarithmic scale is used for the *y*-axis (same scale in all plots); the *x*-axis is plotted with a logarithmic scale in all instances. The *x*-axis on the top of the plot shows the apparent first-order rate constant (k_{on}^{app} or k_{AB}) for molecules of state A going to state B. The *x*-axis on the bottom of the plot shows the overall exchange rate ($k_{ex} = k_{on}^{app} + k_{off} = k_{AB} + k_{BA}$). (a) In the absence of a difference in R_2^0 between the two states, only R_{ex} is observed for a system with a difference in chemical shifts between the two states, which has a maximum value when $k_{ex} = \Delta\omega$. Here, $\Delta\omega/2\pi = 120$ Hz (e.g. 2 ppm for ^{15}N in a 600 MHz static field). (b–d) If the two states differ in chemical shift and state B exhibits a faster transverse relaxation time – i.e. for a high-molecular-weight, slowly tumbling dark state – then ΔR_2^{obs} will have contributions from both R_{ex} and R_{llb}. In the slow exchange regime ($k_{ex} \ll R_{2,B} - R_{2,A}$ and $k_{ex} \ll \Delta\omega$), $\Delta R_2^{obs} = R_{ex} = R_{llb} = k_{on}^{app}$. In the fast exchange regime ($k_{ex} \gg R_{2,B} - R_{2,A}$ and $k_{ex} \gg \Delta\omega$), $\Delta R_2^{obs} = R_{ex} + R_{llb}$, and under very fast exchange, the value of R_{ex} approaches zero, and R_{llb} approaches a population-weighted maximum value (e.g. in panel (d), $\Delta R_2^{max} = 500\ s^{-1} = 5\% \times 10\,000\ s^{-1}$). Note that the same curve for R_{ex} is shown in panels (a–d). (e) In the absence of a difference in chemical shift between the two states, only R_{llb} is observed for a system with a difference in R_2^0 between the two states. For all panels, R_2^{obs} was determined by first solving the McConnell equations to generate a magnetization time course, which was then fit to the free induction decay of a single resonance with R_2^{obs}. A similar procedure was carried out to determine R_2^{obs} in the absence of R_{llb} ($R_2^{obs}[R_{2,B} = R_{2,A}]$) and in the absence of R_{ex} ($R_2^{obs}[\Delta\omega = 0]$). Values were calculated as follows for each exchange rate: $\Delta R_2^{obs} = R_2^{obs} - R_{2,A}$, $R_{llb} = R_2^{obs}[\Delta\omega = 0] - R_{2,A}$, and $R_{ex} = R_2^{obs}[R_{2,B} = R_{2,A}] - R_{2,A}$. All calculations were performed in MATLAB.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

transverse relaxation rate of the visible species will be enhanced by a factor equivalent to the apparent forward rate of the exchange process ($R_{1\mathrm{b}} = k_{\mathrm{on}}^{\mathrm{app}}$). (Note that if $k_{\mathrm{on}}^{\mathrm{app}}$ is too fast, the NMR signal of the visible species will decay too rapidly for measurement of an NMR spectrum.) If the transverse relaxation rate in the dark state is not sufficiently fast to destroy all magnetization (i.e. if some magnetization transferred to the dark state is returned to the visible state by the reverse exchange process), then the enhancement to the transverse relaxation rate will be less than the forward rate of the exchange process ($R_{1\mathrm{b}} < k_{\mathrm{on}}^{\mathrm{app}}$). Because lifetime line broadening generally arises in systems where DEST can also be measured, the two measurements are complementary and can be fit simultaneously to better define the parameters describing the system (Fawzi *et al.* 2010b, 2011b, 2012, 2014).

Lifetime line broadening has been observed for Aβ under pseudo-equilibrium conditions, where Aβ monomers are in exchange with high-molecular-weight protofibrils. These experiments showed that 3–4% of the exchangeable Aβ peptides were associated with the protofibrils (the rest being free monomers), illustrating the sensitivity of lifetime line broadening to a small population of a minor state (Fawzi *et al.* 2010b). Lifetime line broadening has also been used to characterize transient association between Aβ and the chaperonin GroEL, which takes place on a faster timescale of $k_{\mathrm{ex}} \approx 1500\ \mathrm{s}^{-1}$ (Libich *et al.* 2013) (compared with $k_{\mathrm{ex}} \approx 50\ \mathrm{s}^{-1}$ for Aβ self-association into protofibrils; Fawzi *et al.* 2010b, 2011b). This phenomenon has also been observed for α-synuclein, where the addition of lipid vesicles causes an enhancement of 2–5 s^{-1} to $^{15}\mathrm{N}$ R_2, which was interpreted as the pseudo-first-order association rate constant $k_{\mathrm{on}}^{\mathrm{app}}$ for the binding of α-synuclein to lipid vesicles (Bodner *et al.* 2009, 2010).

2.5 Hydrogen exchange

The rate at which a hydrogen atom exchanges with the water solvent depends on how accessible it is to the solvent (and on temperature and pH as well as the chemical composition of neighboring groups). An exchangeable group (generally the backbone amide proton) that is buried in the core of a protein and engaged in a hydrogen bond (i.e. part of a secondary structure element) will have a much slower exchange rate than an exposed exchangeable group. This can be quantified by calculating a protection factor (PF) for each exchangeable group (i.e. for each backbone amide of a protein; Englander & Kallenbach, 1983; Redfield, 2004):

$$\mathrm{PF} = \frac{k_{\mathrm{intr}}}{k_{\mathrm{obs}}}, \tag{41}$$

where k_{obs} is the observed exchange rate and k_{intr} is the intrinsic exchange rate for the fully denatured protein. Protection factors for regions of secondary structure in natively folded proteins can be of the order of 10^5–10^7 (Redfield, 2004).

If the major visible species observed for a system is the folded state of a protein (H exchange rates near zero) even a rare excursion into a more unfolded state with faster H exchange will have a dramatic effect on the measured exchange rate, amplifying by many fold the information from the minor state. Owing to the dramatic differences between folded and partially folded or unfolded states, hydrogen exchange has been particularly useful for studying protein folding, providing single-residue resolution insight into which exchangeable groups become partially exposed over the course of the experiment (Englander *et al.* 2007). Whereas most of the NMR methods for characterizing sparsely populated states described in previous sections require a dark state population of at least ~0·5% (Baldwin & Kay,

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

2009; Clore & Iwahara, 2009; Fawzi *et al.* 2012), hydrogen exchange can be used to characterize minor states with much lower populations (Englander *et al.* 2007).

NMR offers the ability to measure hydrogen exchange at the resolution of individual residues over a wide range of timescales (Fitzkee *et al.* 2011; Friedman *et al.* 2010; Schwartz & Cutnell, 1983; Spera *et al.* 1991; Wagner & Wuthrich, 1982; Wand & Englander, 1996). Mass spectrometry is also a useful tool for measuring hydrogen exchange at high resolution, especially for systems intractable to study by NMR (Englander, 2006; Englander *et al.* 2003; Eyles & Kaltashov, 2004; Zhang & Smith, 1993). Although mass spectrometry is primarily applicable to measuring hydrogen exchange on slower timescales (Hitchens & Bryant, 1998; Konermann *et al.* 2008; Zhang & Smith, 1993), flow-quench methods have made timescales $\ll$1 s accessible (Balasubramaniam & Komives, 2013; Dharmasiri & Smith, 1996). Methods for measuring hydrogen exchange come in two flavors: monitoring of hydrogen–deuterium exchange (HDX) with the solvent by either NMR or mass spectrometry (Hitchens & Bryant, 1998; Konermann *et al.* 2008; Zhang & Smith, 1993) or monitoring transfer of ^{1}H magnetization from the solvent to the protein (HMX) by NMR (Chevelkov *et al.* 2010; Fitzkee *et al.* 2011; Friedman *et al.* 2010; Schwartz & Cutnell, 1983; Spera *et al.* 1991). HDX can be measured either by transferring a protonated protein sample to a $^{2}H_2O$ solution and then monitoring the loss of ^{1}H signal intensity over time or by taking a protein with all exchangeable groups deuterated and then transferring it to a $^{1}H_2O$ solution and measuring the signal buildup of over time. HDX can also be measured by pulse-labeling methods (PHDX) (Roder *et al.* 1988; Udgaonkar & Baldwin, 1988) or by more recently developed native state hydrogen exchange (NHDX) (Bai *et al.* 1995; Englander *et al.* 2007).

In PHDX, the protein is placed in destabilizing fast-hydrogen-exchange conditions for a brief amount of time and exchange is then quenched by transferring back to native conditions, where ^{1}H signal intensity can then be measured by NMR under conditions where exchange is much slower (Krishna *et al.* 2003, 2004a; Roder *et al.* 1988; Udgaonkar & Baldwin, 1988). HDX on backbone amides in partially unstructured states can be used to identify elements of residual secondary structure, which exhibit higher protection factors due to intramolecular hydrogen bonds (Redfield, 2004). Some insight into protein-folding pathways has been gained from hydrogen exchange studies on molten globules, which are partially folded states with well-defined secondary structure but ill-defined tertiary structure that have been isolated from some folding pathways, often by partial denaturation at low pH (Ohgushi & Wada, 1983). Highly characterized molten globule states include those of α-lactalbumin (Baum *et al.* 1989; Forge *et al.* 1999) and apomyoglobin (Hughson *et al.* 1990; Jennings & Wright, 1993).

NHDX provides the ability to detect sparsely populated intermediates through studies on proteins under native conditions. As with the other methods for studying minor states by NMR discussed in this review, it relies on brief excursions from the major state to a sparsely populated intermediate state, especially intermediate states in protein-folding pathways (Bai & Englander, 1996; Bai *et al.* 1995; Englander *et al.* 1997, 2007). Such NHDX studies have identified protein-folding intermediates, which can be understood in terms of foldons, which are small units of secondary structure that fold cooperatively (Englander *et al.* 2007). Examples of protein-folding pathways that have been explored by NHDX include cytochrome *c* (Maity *et al.* 2005), RNase H (Chamberlain & Marqusee, 2000; Chamberlain *et al.* 1996), apocytochrome b_{562} (Chu *et al.* 2002; Fuentes & Wand, 1998a, b), and outer surface protein A (OspA) (Yan *et al.* 2002, 2004).

NMR offers additional avenues for measuring faster hydrogen exchange rates, including by monitoring the exchange of ^{1}H magnetization from the solvent to the protein of interest or by monitoring NMR line shapes of exchanging hydrogens or scalar-coupled ^{15}N nuclei, amongst

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

other methods (Fitzkee *et al.* 2011; Forsén & Hoffman, 1963; Friedman *et al.* 2010; Kateb *et al.* 2007; Schwartz & Cutnell, 1983; Spera *et al.* 1991). This allows one to measure faster exchange processes (such as allostery and ligand binding) without the complications of having to switch between $^1H_2O/^2H_2O$ solvents. These methods have been used to explore a range of exchange processes, including the dynamic nature of the catalytic site and peptide-binding pockets of HIV-1 integrase (Fitzkee *et al.* 2011). Depending on exchange rates, a combination of different hydrogen exchange methods can be used to measure a wider range of exchange timescales, as has been done on studies of the formation of the complex between CaM and a target peptide (Ehrhardt *et al.* 1995; Seeholzer & Wand, 1989; Spera *et al.* 1991). More recently, hydrogen exchange was combined with CEST to directly compare hydrogen exchange rates between the visible folded Fyn SH3 domain and its sparsely populated unfolded state (Long *et al.* 2014). Thus, hydrogen exchange is one of many NMR-based methods, including paramagnetic NMR, relaxation dispersion, and saturation transfer, that can be used to gain insight into states that are invisible to conventional structural methods. The remainder of this review, however, focuses in greater detail on two of these methods: PRE and DEST.

3. Characterizing sparsely populated states by PRE

3.1 Theory

3.1.1 Quantitative description of PRE

PRE is an enhancement of the relaxation rate of a nucleus caused by a dipolar interaction with a paramagnetic center consisting of one or more unpaired electrons. Two other NMR observables can yield long-range distances in paramagnetic systems: PCSs and RDCs. Whereas PCSs and paramagnetic RDCs require an anisotropic electron g-tensor, PRE can be measured in any paramagnetic system, though, as discussed below, it is advantageous to use a system with an isotropic g-tensor. Examples of paramagnetic groups with isotropic g-tensors include nitroxides, Mn^{2+}, and Gd^{3+}. Many other paramagnetic metal ions, including Fe^{3+} and Dy^{3+}, have anisotropic g-tensors (Otting, 2010).

The longitudinal (Γ_1) and transverse (Γ_2) PRE rates arising from direct dipole–dipole interactions are described by the following form of the Solomon–Bloembergen (SB) equations (Bloembergen & Morgan, 1961; Solomon, 1955):

$$\Gamma_1 = \frac{2}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_I^2 g^2 \mu_B^2 s(s+1)\left\{3J_{SB}(\omega_I) + J_{SB}(\omega_I - \omega_S) + 6J_{SB}(\omega_I + \omega_S)\right\}, \tag{42}$$

$$\Gamma_2 = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_I^2 g^2 \mu_B^2 s(s+1)\left\{4J_{SB}(0) + 3J_{SB}(\omega_I) + J_{SB}(\omega_I - \omega_S) + 6J_{SB}(\omega_S) + 6J_{SB}(\omega_I + \omega_S)\right\}, \tag{43}$$

where s is the electron spin quantum number, μ_0 the permeability of free space ($\sim 1{\cdot}257 \times 10^{-6}$ m kg s^{-2} $Å^{-2}$), μ_B the Bohr magneton (the magnetic moment of a free electron; $\sim 9{\cdot}274 \times 10^{-24}$ J T^{-1}), g the electron g-factor ($\sim 2{\cdot}0$), γ_I the proton gyromagnetic ratio ($\sim 2{\cdot}675 \times 10^8$ rad s^{-1} T^{-1}), $\omega_I/2\pi$ the Larmor frequency of the proton (e.g. 600 MHz in a 14·1-T magnetic field), $\omega_S/2\pi$ the Larmor frequency of the electron (e.g. 395 GHz in a 14·1-T magnetic field), and $J_{SB}(\omega)$ is the generalized spectral density function for the reduced correlation function, given by:

$$J_{SB}(\omega) = r^{-6}\frac{\tau_c}{1 + (\omega\tau_c)^2}, \tag{44}$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

where r is the distance between the paramagnetic center and the reporter nucleus. The overall correlation time, τ_c, is defined as:

$$\tau_c = \left(\tau_r^{-1} + \tau_s^{-1}\right)^{-1}, \tag{45}$$

where τ_r is the rotational correlation time of the macromolecule and τ_s is the effective electron relaxation time. Values of τ_s depend on the nature of the paramagnetic group. Paramagnetic metal ions with anisotropic g-tensors generally exhibit very small values (e.g. $\tau_s \sim 10^{-13}$ for anisotropic lanthanides; Otting, 2010), thus making $\tau_c \approx \tau_s$ in this case. Nitroxides general exhibit much larger τ_s values ($\sim 10^{-6}$; Bernini *et al.* 2009), thus making $\tau_c \approx \tau_r$ in that case. Paramagnetic metal ions with isotropic g-tensors, however, exhibit intermediate values on the same order as τ_r (e.g. $\tau_s \sim 10^{-9}$ for Mn^{2+}, $\tau_s \sim 10^{-8}$ for Gd^{3+} (Bertini *et al.* 2005; Iwahara *et al.* 2004b; Pintacuda *et al.* 2004a)), preventing simplification of the expression for τ_c.

Under the conditions of biomolecular NMR (where $\omega_s^2\tau_c^2 \gg 1$), the high-frequency components of the spectral density function (those dependent on ω_S) can be ignored, and the longitudinal and transverse PREs can be expressed by the simplified formulas (Clore & Iwahara, 2009):

$$\Gamma_1 = \frac{2}{5}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_I^2 g^2 \mu_B^2 s(s+1) J_{SB}(\omega_I), \tag{46}$$

$$\Gamma_2 = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_I^2 g^2 \mu_B^2 s(s+1)\left\{4J_{SB}(0) + 3J_{SB}(\omega_I)\right\}. \tag{47}$$

A paramagnetic group can also cause enhanced relaxation through Curie spin relaxation, which arises from the interaction between the nucleus and the time-averaged magnetization of the electrons (due to the Boltzmann population difference in electron spins states; Bertini *et al.* 2001a). The Γ_2 rate due to Curie spin relaxation is given by (Gueron, 1975):

$$\Gamma_{2,\mathrm{Curie}} = \frac{1}{5}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\left(\omega_I^2 g^4 \mu_B^4 s^2 (s+1)^2\right)}{(3k_B T)^2 r^6}\left\{4\tau_r + \frac{3\tau_r}{1+(\omega_I\tau_r)^2} - 4\tau_c - \frac{3\tau_c}{1+(\omega_I\tau_c)^2}\right\}. \tag{48}$$

Note that Curie spin relaxation is independent of the rate of electron relaxation (τ_s).

Equations (46)–(48) show that the PRE has an r^{-6} dependence on the distance between the observed nucleus and the paramagnetic center. The dependence on the magnetic field and on the size of the molecule, however, is slightly more complicated. The dipole–dipole Γ_2 rate is approximately independent of field, whereas the Curie Γ_2 rate is approximately proportional to the square of the magnetic field. (The dipole–dipole Γ_1 rate decreases with increasing field strength.) In cases where $\tau_c \approx \tau_r$ – due to a relatively long τ_s (e.g. nitroxide spin radicals) – Curie spin relaxation is negligible, and the Γ_2 rate is therefore independent of field. In these cases, the dipole–dipole Γ_2 rate is directly proportional to τ_r (and to molecular weight). On the other hand, for metal ions with a very short τ_s (especially those with an anisotropic g-tensor, e.g. Fe^{3+}, Dy^{3+}), the Γ_2 rate in macromolecules is dominated by Curie-spin relaxation. In these cases, the Curie-spin Γ_2 rate is also directly proportional to τ_r, while the dipole–dipole Γ_2 rate is independent of molecular weight because $\tau_c \approx \tau_s$. For isotropic metal ions, though, such as Mn^{2+} and Gd^{3+}, the contribution from Curie-spin relaxation is minor, and negligible for medium-sized macromolecules (<40 kDa). For example, in the 20 kDa complex of SRY and DNA–EDTA–Mn^{2+}, Curie-spin relaxation contributed 2% of the Γ_2 rates at a ^{1}H-frequency of 800 MHz and a temperature of 308 K (Iwahara *et al.* 2007). However, for a larger molecule (>100 kDa), Curie-spin relaxation can contribute more than 20% of the total value of Γ_2 under these conditions (Clore & Iwahara, 2009).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Thus, for measuring PREs, it is advantageous to use a paramagnetic group with an isotropic *g*-tensor because the Curie spin contribution to the Γ_2 rate will be much less (often negligible) compared with one with an anisotropic *g*-tensor. In addition, for an anisotropic paramagnetic group, the generation of PCSs means that the chemical shifts (and therefore resonance assignments) are altered and that exchange contributions are different between the diamagnetic and paramagnetic samples, which complicates analysis, as described below (Clore & Iwahara, 2009).

Artificially introduced paramagnetic groups are generally attached to the macromolecule of interest through linkers that have several rotatable bonds. Hence, these paramagnetic centers are intrinsically flexible. In addition, the observed nuclei may be located in mobile regions of the macromolecule. The effect of fast dynamics in the picosecond to nanosecond timescale on the PRE for paramagnetic systems with an isotropic *g*-tensor can be readily taken into account by modifying the SB equations described above by incorporating a 'model-free' formalism (SBMF) (Lipari & Szabo, 1982a). This transforms the spectral density function in Eq. (44) into (Iwahara *et al.* 2004b):

$$J_{\mathrm{SBMF}}(\omega) = \langle r^{-6}\rangle\left\{\frac{S^2\tau_{\mathrm{c}}}{1+(\omega\tau_{\mathrm{c}})^2} + \frac{(1-S^2)\tau_{\mathrm{t}}}{1+(\omega\tau_{\mathrm{t}})^2}\right\}, \tag{49}$$

where S^2 is the square of the generalized order parameter, $\langle r^{-6}\rangle$ is the population-averaged distance inverse sixth power of the distance between the paramagnetic center and the observed nucleus, and τ_{t} is the total correlation time, defined as:

$$\tau_{\mathrm{t}} = \left(\tau_{\mathrm{r}}^{-1} + \tau_{\mathrm{s}}^{-1} + \tau_{\mathrm{i}}^{-1}\right)^{-1}, \tag{50}$$

where τ_{i} is the correlation time for internal motion. These equations are valid for isotropically tumbling systems; otherwise, the angles between the principal axes of the diffusion tensor and PRE vector must be taken into account (Benetis & Kowalewski, 1985; Woessner, 1962).

The model-free formalism is most commonly employed for analyzing relaxation due to fixed-length dipole–dipole interactions, such as those in ^{1}H–^{15}N or ^{1}H–^{13}C bond vectors. However, it can also be applied to variable-length vectors (e.g. PRE; Iwahara *et al.* 2004b), where it takes the form (Olejniczak *et al.* 1984):

$$S^2 = \frac{4\pi}{5}\langle r^{-6}\rangle^{-1}\sum_{m=-2}^{2}\left|\left\langle\frac{Y_2^m\Omega^{\mathrm{mol}}}{r^3}\right\rangle\right|^2, \tag{51}$$

where $Y_2^{\mathrm{m}}(\Omega)$ are second-order spherical harmonics and Ω^{mol} are Euler angles in the molecular frame. The order parameter can be approximated by decomposing it into its radial and angular components, as has been demonstrated for ^{1}H–^{1}H dipolar interactions (Bruschweiler *et al.* 1992):

$$S^2 \approx S^2_{\mathrm{angular}}S^2_{\mathrm{radial}}, \tag{52}$$

where the angular (S^2_{angular}) and radial (S^2_{radial}) order parameters are defined as:

$$S^2_{\mathrm{angular}} = \frac{4\pi}{5}\sum_{m=-2}^{2}\left|\langle Y_2^m\Omega^{\mathrm{mol}}\rangle\right|^2, \tag{53}$$

$$S^2_{\mathrm{radial}} = \langle r^{-6}\rangle^{-1}\langle r^{-3}\rangle^2. \tag{54}$$

These relationships are used directly for back-calculating PRE values from known structures (and for refining against PRE data) as described below. The PRE measured for a given nucleus has a population-weighted $\langle r^{-6}\rangle$ dependence on the distance between the nucleus and the paramagnetic

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

center, hence its utility as a structural restraint (and as a probe of minor states) emerges. However, because most paramagnetic tags used for measuring PRE are highly flexible (and the molecules of interest are also often quite flexible), it is important in structure calculations to refine directly against the PRE data (e.g. using prePot in Xplor-NIH (Iwahara *et al.* 2004b) as discussed below) rather than converting PRE data into a single distance, which is generally incorrect (although PRE calculators have been made available online at spin.niddk.nih.gov/clore for qualitative purposes).

Both the transverse and longitudinal PREs are influenced by internal motion (Iwahara *et al.* 2004b). However, while Γ_1 has a complicated dependence on the internal correlation time, the Γ_2 shows a simpler dependence. As the internal correlation time is increased from 0 (infinitely fast internal motion) to a value greater than the overall rotational correlation time, the ratio $\Gamma_{2,\mathrm{SBMF}}/\Gamma_{2,\mathrm{SB}}$ (where $\Gamma_{2,\mathrm{SBMF}}$ is Γ_2 calculated from the SBMF equation, and $\Gamma_{2,\mathrm{SB}}$ is Γ_2 calculated from the SB equation) increases from the value of the order parameter ($0 \le \boldsymbol{S}^2 \le 1$) to 1. For the longitudinal PRE, however, $\Gamma_{1,\mathrm{SBMF}}/\Gamma_{1,\mathrm{SB}}$ first increases from the value of $\boldsymbol{S}^2$ when $\tau_i = 0$, up to a maximum value $\gg 1$ when $\tau_i \approx 1/\omega$ (where ω is the resonance frequency of the observed nucleus in rad s^{-1}), before eventually approaching 1 (Iwahara *et al.* 2004b). Thus, whereas a simple estimation of τ_i is sufficient for structural analysis of ^{1}H–Γ_2 data, such treatment of ^{1}H–Γ_1 data requires accurate information on τ_i for individual PRE vectors (Iwahara & Clore, 2010). Another complication is that, due to cross-relaxation and water exchange, ^{1}H R_1 relaxation does not follow a single-exponential curve. Thus, it is advantageous to measure the transverse PRE instead of the longitudinal PRE. For very fast motion ($\tau_i \ll \tau_r$; i.e. picosecond timescale motion), the transverse PRE is virtually independent of the internal correlation time, and the order parameter acts as a scaling factor. In either case, though, the SBMF equations reduce to the SB equations under conditions where the internal motion is either very slow ($\tau_i \approx \tau_c$) or highly restricted in space ($\boldsymbol{S}^2 \approx 1$).

3.1.2 PRE from minor states

In a system undergoing exchange between states on a fast timescale, the overall PRE profile observed for the ensemble of states is the population-weighted average of all the members of the ensemble. Owing to the $\langle r^{-6} \rangle$ distance dependence of the PRE, such a profile will be dominated by the states in which the paramagnetic center more closely approaches the observed nuclei. Thus, even if a state occupies a small (<10%) portion of the total population, it can be readily detected if it involves a shorter distance between a paramagnetic probe and reporter than the other state(s) (Fig. 2). Thus, an otherwise invisible minor state can be detected and structurally characterized using PRE (Clore & Iwahara, 2009).

Under fast exchange conditions on the PRE timescale (i.e. the rate of exchange between the species is much greater than the difference in PRE rates between states) the overall PRE observed for a nucleus is given by:

$$\Gamma_2^{\mathrm{app}} = \sum p_i \Gamma_{2,i}, \tag{55}$$

where Γ_2^{app} is the apparent transverse PRE measured for the entire ensemble, $\Gamma_{2,i}$ is the transverse PRE for state i, and p_i is the population of state i. Eq. (55) holds when the rate of exchange is slower than the overall correlation time ($k_{\mathrm{ex}} \ll 1/\tau_c$); in the limit of ultrafast exchange, however, where $k_{\mathrm{ex}} \gg 1/\tau_c$, in Eq. (55) would be replaced by $\Gamma_2^{\mathrm{app}} = \left(\sum p_i \sqrt{\Gamma_{2,i}}\right)^2$ (i.e. under ultrafast exchange conditions the contribution of each state to Γ_2^{app} is roughly proportional to the square of its population due to the fact that individual components of Γ_2^{app} are averaged as $<r^{-3}>$ as opposed to $<r^{-6}>$ under the more usual conditions of Eq. (55)).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

In the case of two-state exchange between a major state (A) and a minor state (B), the apparent observed PRE is given by:

$$\Gamma_2^{\text{app}} = p_A\Gamma_{2,A} + p_B\Gamma_{2,B}. \tag{56}$$

If the exchange between states A and B is not fast on the PRE timescale (i.e. if $k_{ex} \gg \Gamma_{2,B} - \Gamma_{2,A}$ is not true), the apparent PRE is not simply a population-averaged value, but is instead intermediate between $\Gamma_{2,A}$ and the population-averaged value (Fig. 2c). Under these conditions of slower exchange, the apparent PRE can be calculated using the McConnell equations (McConnell, 1958). These necessary equations are identical to the general equations for chemical exchange (Eqs. (16)–(20)), except the relaxation matrix **R** (Eq. (18)) includes the PRE from each state, taking the form (for two-site exchange):

$$\mathbf{R} = \begin{bmatrix} R_{2,A} + \Gamma_{2,A} & 0 \\ 0 & R_{2,B} + \Gamma_{2,B} \end{bmatrix}. \tag{57}$$

Apparent relaxation rates can be extracted by fitting the line shape of the Fourier-transformed solution of the McConnell equations. Under such cases, where exchange takes place on the intermediate PRE timescale ($k_{ex} \approx \Gamma_{2,B} - \Gamma_{2,A}$), the apparent PRE measured for a system in exchange between two states, in addition to encoding distance and population information, also encodes kinetic information, although in practice it is difficult to extract this information from PRE data alone.

To visualize how PRE can be used to detect and characterize a minor state, consider a protein in exchange between two states (Fig. 2). In state A, the protein is in a more extended conformation, and the distance between the two selected points is 40 Å. In state B, however, the protein is in a more compact conformation, and the distance between these same two points is 8 Å. If one of the points is a paramagnetic center, and the other is a ^{1}H nucleus, we can calculate the PRE (after specifying the correlation times and the order parameter of the vector between the paramagnetic center and the observed nucleus). In this example, we will consider a single unpaired electron (e.g. a nitroxide spin label), and we will set τ_c = 10 ns (corresponding to a protein at room temperature with a molecular weight of ~20 kDa). Also, we will assume that the paramagnetic center/observed nucleus vector is rigid, although this is not the case for most conventional spin labels (e.g. MTSL). For state A, with the nucleus far from the unpaired electron, $\Gamma_{2,A} = 0{\cdot}12\ s^{-1}$ (effectively zero). However, in state B, where they are in much closer proximity, $\Gamma_{2,B} = 1900\ s^{-1}$.

If the protein in this example is in fast exchange between states A and B ($k_{ex} \gg \Gamma_{2,A} - \Gamma_{2,B} = 1900\ s^{-1}$; $\tau_{ex} \ll 500\ \mu s$), the apparent PRE would be the population-weighted average of $\Gamma_{2,A}$ and $\Gamma_{2,B}$ (Eq. (56)). Thus, even if state A comprises 99% of the total population, and state B only comprises 1% of the total population, $\Gamma_2^{\text{app}} = 0{\cdot}99 \times 0{\cdot}12\ s^{-1} + 0{\cdot}01 \times 1900\ s^{-1} \approx 19\ s^{-1}$. Thus, even though state B comprises only 1% of the total population, Γ_2^{app} arises almost exclusively from the minor state.

Two conditions are necessary for detection of a minor state by PRE: (1) a shorter distance between probe and reporter in the minor state, and (2) fast exchange between the states. Having a shorter distance to measure in the minor state is a necessary requirement for PRE measurements on minor states, and detecting such a state is most straightforward for the transient association between two molecules or for a minor state that involves forming a more compact structure than in the major state (e.g. CaM; Anthis *et al.* 2011). However, other states can also be measured, as long as a paramagnetic probe position can be found where at least some measurable nuclei are in closer proximity to the probe in the minor state than in the major state.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

For quantitative PRE measurements (i.e. for the overall measured PRE profile to be exactly the population-weighted average of the PRE profiles from the individual states), the rate of exchange ($k_{ex} = k_{AB} + k_{BA}$) must be fast on the PRE timescale, meaning that $k_{ex} \gg \Gamma_{2,B} - \Gamma_{2,A}$. Otherwise, the apparent PRE becomes highly dependent on k_{ex}, as seen in Fig. 2. That being said, under conditions where $k_{ex} \sim \Gamma_{2,B} - \Gamma_{2,A}$, the Γ_2^{obs}rate is only decreased by a factor of about 2 relative to the limiting value in the fast exchange limit (Fig. 2*c*); given the $\langle r^{-6} \rangle$ dependence of the PRE, the assumption of fast exchange in the intermediate PRE regime would only result in a ~10% overestimation in the corresponding distance. As the exchange rate decreases, the apparent PRE decreases, eventually becoming equal to that of state A, providing no information on the minor state. If exchange is not fast on the PRE timescale, the apparent PRE also becomes dependent on the chemical shift difference between states A and B ($\Omega_A - \Omega_B$), decreasing with decreasing k_{ex} or increasing $\Omega_A - \Omega_B$ (unless $k_{ex} \gg |\Omega_A - \Omega_B|$) (Clore & Iwahara, 2009). However, under these conditions, the peak will also be subject to additional line broadening, making measurements difficult or impossible as $|\Omega_A - \Omega_B|$ approaches k_{ex}. If $k_{ex} \ll |\Omega_A - \Omega_B|$, then the two states will display two separate peaks; the minor state peak will likely be invisible, and the major state peak will not contain PRE information on the minor state (Clore & Iwahara, 2009).

3.2 Data acquisition

PRE data are acquired by measuring transverse relaxation rates for a paramagnetic sample and a diamagnetic sample. Subtraction of the diamagnetic R_2 from the paramagnetic R_2 for each nucleus yields its PRE rate Γ_2. The molecule of interest must be either recombinantly expressed (often in *E. coli*) or chemically synthesized and then tagged with a paramagnetic label (or, if a natural paramagnetic ion ligand will be used, it must be added to the protein via a chelating group or an intrinsic metal-binding site). The protein will usually be universally or selectively ^{15}N- and/or ^{13}C-labeled to perform measurements using modern multidimensional heteronuclear NMR spectroscopy. For intermolecular PRE measurements, one can use a ^{15}N/^{13}C-labeled protein without a paramagnetic tag in the presence and absence of a non-isotopically labeled (natural abundance) protein with a paramagnetic tag. It is advisable to perform experiments on perdeuterated proteins because the $^1H_\alpha$–^1H_N coupling can influence the PRE measurements, but if using perdeuterated proteins is not feasible, this coupling can largely be cancelled by appropriate design of the pulse sequence (Iwahara *et al.* 2003, 2007).

3.2.1 Producing paramagnetic samples

Broadly speaking, two types of paramagnetic labels are available for measuring PREs: those containing a nitroxide group and those that act as a ligand for a paramagnetic cation (e.g. Mn^{2+}, Cu^{2+}, and Gd^{3+}). If a protein has an intrinsic cation-binding site, a paramagnetic cation can often be introduced in place of the natural metal ligand. For proteins that are not native metalloproteins, short metal-binding sequences can be genetically encoded. Examples of such sequences include the N-terminal ATCUN motif (XXH), which binds paramagnetic Cu^{2+} (Donaldson *et al.* 2001), the sequence YVDTNNDGAYEGDEL, which specifically binds lanthanides (Wohnert *et al.* 2003), and the N-terminal sequence HHP, which forms a dimer with two peptides bound to paramagnetic Ni^{2+} (Jensen *et al.* 2004). A paramagnetic Cu^{2+} ion can also be introduced into a protein by making a double histidine mutation at i and $i + 4$ positions in an α-helix and then

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

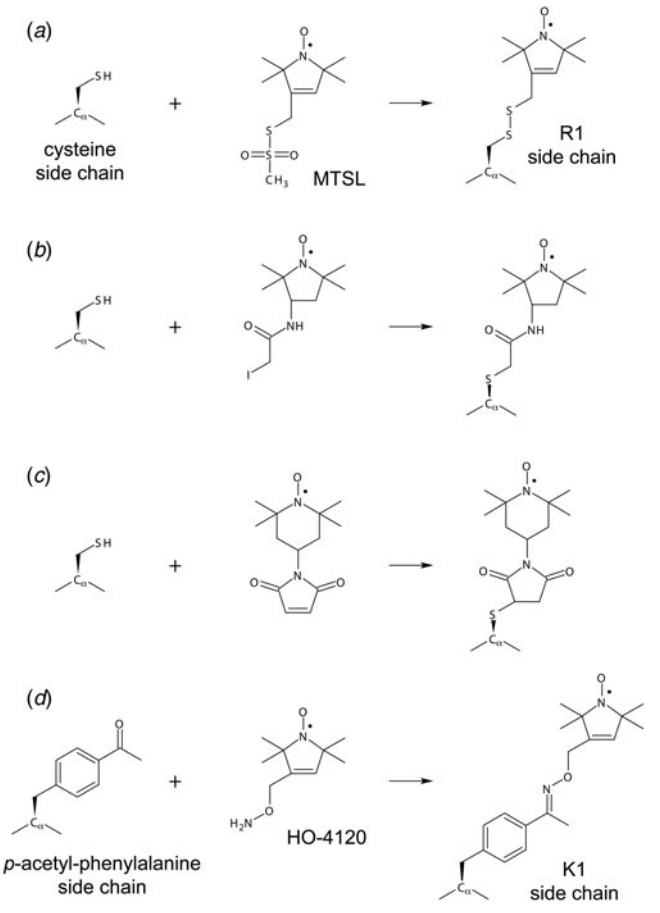


Fig. 6. Site-directed spin labeling. Shown are four different chemistries for conjugating a spin label to a cysteine residue (a–c) or to an unnatural amino acid (d). All four examples are nitroxide spin labels (with the unpaired electron denoted with a dot), but the same chemistry can be used for spin labels that chelate a paramagnetic metal ion. The C_α of the peptide backbone is labeled. (a) Reaction of a methanethiosulfonate with cysteine to form a disulfide bond. The compound (1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL; left) reacts with cysteine to form the R1 side chain (right) (Battiste & Wagner, 2000; Gaponenko *et al.* 2000). (b) Reaction of an iodoacetamide with cysteine to form an S–C bond. Shown on the left is 3-(2-iodoacetamido)-2,2,5,5-tetramethyl-1-pyrolidinyloxy (iodoacetamido-PROXYL) (Gillespie & Shortle, 1997a). (c) Reaction of a maleimide with cysteine to form an S–C bond. Shown on the left is N-(1-oxyl-2,2,6,6-tetramethyl-4-piperidinyl) maleimide (maleimide-TEMPO) (Griffith & McConnell, 1966; Tang *et al.* 2008a). (d) Reaction of 3-aminooxymethyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxyl radical (HO-4120) with the unnatural amino acid *p*-acetyl-L-phenylalanine to form the side chain K1 (Fleissner *et al.* 2009).

adding the chelating agent nitrilotriacetic acid in a 1:1 ratio with Cu^{2+} (which prevents Cu^{2+}-mediated dimerization) (Liu *et al.* 2014).

The most common method for attaching a paramagnetic probe to a protein of interest is site-directed spin labeling (Hubbell *et al.* 2000; Todd *et al.* 1989), where a paramagnetic tag is conjugated to an exposed cysteine residue introduced by site-directed mutagenesis (Fig. 6). Conjugation by disulfide chemistry (Battiste & Wagner, 2000; Dvoretsky *et al.* 2002; Gaponenko *et al.* 2000) offers high specificity for cysteine and is straightforward to perform

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

using a spin label molecule with a methanethiosulfonate or pyridylthiol functional group (Fig. 6*a*). Disulfide linkages are somewhat unstable, and over time some of the paramagnetically tagged protein is replaced by disulfide cross-linked protein dimers, although over a short time period (up to ∼1 month, often longer), the disulfide-conjugated paramagnetic protein sample is pure and stable enough to perform PRE measurements. Paramagnetic probes can also be conjugated using iodo- or bromoacetamide derivatives, which irreversibly form a C–S bond with cysteine (Gillespie & Shortle, 1997b; Ogawa & McConnel, 1967) (Fig. 6*b*); however, although they are advantageous for their exceptional stability, they exhibit lower specificity (also reacting at a slower rate with histidine, methionine, and lysine). Maleimide derivatives, however, are not recommended due to their instability (Kosen, 1989) (Fig. 6*c*). As these tags conjugate to exposed cysteine residues, any natural exposed cysteine residues will need to be mutated to serine or alanine. Even buried cysteines can become transiently exposed by the protein's intrinsic motions, so one will need to ensure that they do not react with the tag if they are to remain unmutated.

An alternative method, which is especially useful when the removal of natural cysteines is difficult or impossible, involves the incorporation of an unnatural amino acid using an orthogonal-tRNA/aminoacyl-tRNA synthetase pair (Wang *et al.* 2001). The spin label site is generated by introducing an amber (TAG) codon into the DNA sequence of the protein of interest by site-directed mutagenesis. The protein is then coexpressed with the unnatural tRNA/aminoacyl-tRNA synthetase pair. This method can be used to produce a protein containing the unnatural amino acid *p*-acetyl-L-phenylalanine, which can be conjugated to HO-4120, a hydroxylamine-functionalized nitroxide spin-label (Fleissner *et al.* 2009) (Fig. 6*d*). This method has been used to produce spin-labeled proteins at high yields, even in perdeuterated media, where it has been demonstrated that the key for sufficient yields is to induce expression of the tRNA/aminoacyl-tRNA synthetase pair while still in rich, protonated media before changing to deuterated media for expression of the target protein (Venditti *et al.* 2012). For shorter sequences (up to ∼50 residues), a spin-labeled amino acid can be directly incorporated through solid-phase peptide synthesis (Hanson *et al.* 1996; Lindfors *et al.* 2008; Milov *et al.* 2002; Monaco *et al.* 1999a, b; Shenkarev *et al.* 2004).

Some spin label advances have been summarized in recent reviews (Hubbell *et al.* 2013; Jeschke, 2013; Keizers & Ubbink, 2011), and Fig. 7 shows some nitroxide and metal-chelating paramagnetic tags. The paramagnetic nitroxide side chain R1, which is formed by disulfide conjugation of the spin label MTSL to cysteine (Fig. 7*a*; Battiste & Wagner, 2000; Gaponenko *et al.* 2000), is commonly used in EPR studies and has been used to measure PREs in several studies (Anthis & Clore, 2013; Anthis *et al.* 2011; Bermejo *et al.* 2009; Fawzi *et al.* 2011a; Volkov *et al.* 2006). For studies involving R1, an acylated diamagnetic control is available (Fig. 7*b*). Alternatively, for PRE measurements, an EDTA-based tag (Fig. 7*e*), which can chelate a paramagnetic ion (e.g. Mn^{2+}) or a diamagnetic control (e.g. Ca^{2+}), can be used (Dvoretsky *et al.* 2002; Ebright *et al.* 1992; Ermacora *et al.* 1992). Such tags have the advantage of giving larger PRE values due to larger electron spin number. However, these metal-bound tags are problematic for PCS studies because their flexibility greatly complicates analysis of PCSs, which are orientation dependent. Additionally, metal-bound EDTA forms two stereoisomers, each with a slightly different χ-tensor, causing the appearance of multiple peaks due to slow metal/tag exchange (Ikegami *et al.* 2004; Pintacuda *et al.* 2004b; Vlasie *et al.* 2007). Metal-binding tags have also been developed with multiple functionalities, for example 4′-mercapto-2,2′:6′,2-terpyridine-6,6-dicarboxylic acid (4MTDA), which binds to lanthanides and can be

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

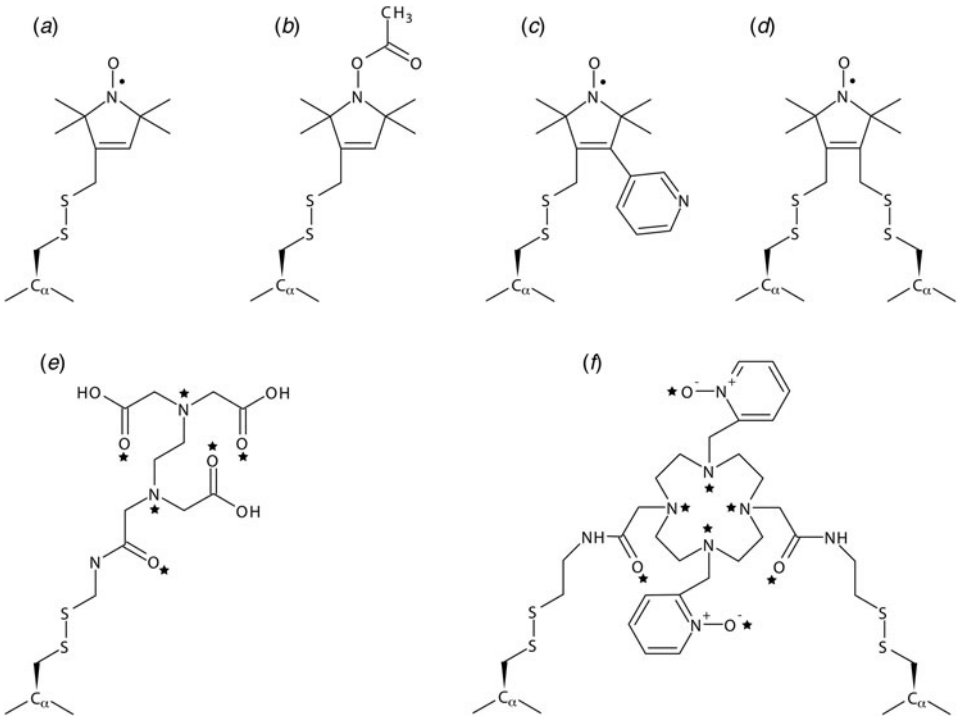


Fig. 7. Spin label side chains for measuring PRE. Shown are three nitroxide spin label side chains (a, c, d), a diamagnetic control (b), and two side chains for chelating paramagnetic metal ions (e, f). The unpaired electron in each nitroxide side is denoted with a dot, and metal chelating sites are denoted with stars. The C_α of the peptide backbone is labeled. (a) The side chain R1, introduced by the reaction of (1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL) with a cysteine side chain (Battiste & Wagner, 2000; Gaponenko *et al.* 2000). (b) An acylated derivative of R1, which can be used as a diamagnetic control for R1/MTSL PRE studies, introduced by the reaction of (1-acetoxy-2,2,5,5-tetramethyl-*δ*-3-pyrroline-3-methyl) methanethiosulfonate with a cysteine side chain (Altenbach *et al.* 2001). (c) The side- chain R1p, a derivative of R1 that adopts a narrower range of conformations and simplifies analysis, introduced by the reaction of 3-methanesulfonilthiomethyl-4-(pyridin-3-yl)-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxyl radical (HO-3606) with a cysteine side chain (Fawzi *et al.* 2011a). (d) The side chain RX, which also is more rigid than R1, introduced by the reaction of 3,4-bis-(methanethiosulfonylmethyl)-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxy radical (HO-1944) with two cysteine side chains located within close proximity (specifically at i and $i+3$ or $i+4$ in a helix or at i and $i+2$ in a β-strand) (Fleissner *et al.* 2011). (e) A side chain for chelating paramagnetic metal ions (e.g. Mn^{2+}) or diamagnetic controls (e.g. Ca^{2+}), introduced by the reaction of S-(2-pyridylthio) cysteaminyl-EDTA with a cysteine side chain (Dvoretsky *et al.* 2002; Ebright *et al.* 1992; Ermacora *et al.* 1992). (f) The CLaNP (caged lanthanide NMR probe) tags chelate lanthanide ions and offer a more rigid paramagnetic center, suitable for both PRE and PCS measurements. Shown here is the side chain introduced by conjugating 1,4,7,10-tetraazacyclododecane-1,7-[di-(N-oxido-pyridine-2-yl)methyl]-4,10-bis (2-(acetylamino)ethylmethanesulfonothioate) (CLaNP-5) to two cysteine side chains located in close proximity to one another (Keizers *et al.* 2007, 2008).

used for the measurement of PRE and PCSs, but also fluoresces in the UV–vis range (Huang *et al.* 2013).

A cysteine-conjugated MTSL tag (i.e. residue R1; Fig. 7*a*), has five torsional angles about which rotation is possible. This allows the tag to sample a large range of conformational space and makes it highly dynamic (Bridges *et al.* 2010; Chou *et al.* 2003) (overall correlation time of the

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

order of ∼0·7 ns (Oganesyan, 2011)). However, the use of a shorter spin label increases the likelihood of undesired perturbations to the structure or dynamics of the molecule under study (Jeschke, 2013). Thus, whereas the dynamics of the spin label have complicated analysis and impeded progress in the past, newer computational methods described below have made it possible to take these dynamics into account, making it preferable to use the more flexible tag that is less likely to disrupt the protein under study. NMR studies attempting to quantitatively interpret the PREs arising from R1 (Bermejo *et al.* 2009), like any other label with internal motions, must take into account the ensemble of nitroxide positions, their interconversion dynamics, and the order parameters for the interspin (nitroxide–nucleus) vectors (Clore & Iwahara, 2009; Iwahara *et al.* 2004b).

Various analogs of R1 (Columbus *et al.* 2001) have shown reduced flexibility, thus potentially simplifying analysis. A series of R1 analogs aromatically substituted at position 4 of the paramagnetic pyrroline ring exhibit restricted motion, and the 4-pyridyl analog of R1 (R1p; Fig. 7*c*) shows particular promise. In one study, multiple tag conformations were required to fit PRE data from R1, but a single conformation could be used for the R1p tag (Fawzi *et al.* 2011a). A different nitroxide side chain (V1), formed by the reaction between a free cysteine and bis(2,2,5,5-tetramethyl-3-imidazoline-1-oxyl-4-il)-disulfide (IDSL), also displays more constrained geometry than R1 (Toledo Warshaviak *et al.* 2013). If two cysteines are available in close proximity (at i and $i+3$ or $i+4$ in a helix or at i and $i+2$ in a β-strand), the side chain RX (Fig. 7*d*) can be introduced, forming two disulfide linkages and maintaining a more rigid conformation than R1 (Fleissner *et al.* 2011). An alternative strategy for a more rigid tag is the incorporation of the unnatural amino acid TOAC into the peptide of interest for PRE studies (Lindfors *et al.* 2011). TOAC contains a nitroxide radical that is contained within a single piperidine ring that includes the backbone C_α. Its downsides are that it can only be incorporated into synthetic peptides (not recombinant proteins), and it can still adopt multiple conformations (Crisma *et al.* 2005; Flippen-Anderson *et al.* 1996).

More rigid tags have also been designed to introduce paramagnetic lanthanides into proteins via chelating groups. These are particularly valuable for PCS studies, where rigid electron–nucleus vectors are needed due to the orientational dependence of the PCS. A series of CLaNP (caged lanthanide NMR probe) tags have been developed, which form a much more rigid structure. After chelating a lanthanide ion, the CLaNP tag can then be conjugated to a protein via two cysteine residues in close proximity on the surface of the protein. One of these, CLaNP-5 (Keizers *et al.* 2007, 2008) is shown in Fig. 7*f*. CLaNP-7, has a pH-dependent χ-tensor and has a lower charge (+1 *versus* +3 for CLaNP-5), lessening the potential for electrostatic interactions with negatively charged molecules (Liu *et al.* 2012). Another tag, 4-mercaptomethyl-dipicolinic acid, adopts a rigid conformation after conjugation to a surface cysteine located near a carboxyl-bearing side chain, followed by addition of a lanthanide ion (Shen *et al.* 2009). Another strategy for introducing a paramagnetic lanthanide uses stereospecifically methylated DOTA (DOTA-M8), which is conjugated to the protein via a single surface cysteine, and adopts a more rigid conformation due to hydrophobic interactions on the protein surface (Haussinger *et al.* 2009).

It is essential to ensure that the introduction of the mutation and the introduction of the tag do not affect the structure or function of the protein. This can preferably be done through a functional assay if one is available. Biophysical measures (e.g. CD, melting temperature) can be used to ensure that there are no changes to protein structure or stability. If the protein is not expressed

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

or becomes insoluble, this likely indicates that the mutation has disrupted the structure of the protein and made it unstable. If it is possible to acquire an NMR spectrum (e.g. ^{1}H–^{15}N HSQC) of the mutant protein and compare it to the spectrum of the wild-type protein, there should not be any broadened or disappearing peaks, and only peaks near the site of the mutation should exhibit changes in chemical shift. One can then check a spectrum of the protein conjugated to a diamagnetic tag. These results should also be the same. Some very localized disruption of the protein structure may be inevitable and they can be considered acceptable if they are restricted to adjacent residues. When PRE data are acquired, the R_2 rate of the diamagnetic control should not vary from that of the wild type. Enhancements to the diamagnetic R_2 can indicate some disruption of local structure. Any significant disruption needs to be taken into account and may indicate that a particular tag position is not appropriate. Finding the best sites for a paramagnetic tag often involves trial and error, as tags at various different positions may cause disruptions to protein structure.

3.2.2 Measuring PRE data

The transverse PRE rate (Γ_2) of a nucleus is given by the difference in its transverse relaxation rate between the paramagnetic ($R_{2,\mathrm{para}}$) and diamagnetic ($R_{2,\mathrm{dia}}$) states of the molecule:

$$\Gamma_2 = R_{2,\mathrm{para}} - R_{2,\mathrm{dia}} \tag{58}$$

The PRE is determined by measuring R_2 values for a paramagnetic sample and for a diamagnetic sample and then subtracting the diamagnetic R_2 values from the paramagnetic R_2 values. The most appropriate diamagnetic sample will depend on the nature of the paramagnetic sample. If the PREs arise from a paramagnetic ion, an appropriate diamagnetic control would be a similar sample with the paramagnetic ion replaced with a diamagnetic ion of similar properties. For a metalloprotein, this might be the natural metal ligand. Or, for a chelating extrinsic tag for example, one could use Mn^{2+} for the paramagnetic sample and Ca^{2+} for the diamagnetic sample. For a protein tagged with a nitroxide label, a good diamagnetic sample is a protein tagged with a similar diamagnetic label. For MTSL, one could use an acylated analogous tag ((1-acetoxy-2,2,5,5-tetramethyl-δ-3-pyrroline-3-methyl) methanethiosulfonate) for the diamagnetic control. An alternative is to add ascorbic acid directly to the paramagnetic sample, which reversibly reduces nitroxides, forming a diamagnetic product. This method has the advantage of only requiring one sample, but has the downside of requiring a large excess of ascorbic acid, which must remain present during acquisition (Venditti *et al.* 2011). These methods are generally preferable, however, to using a wild-type or untagged protein as the diamagnetic control, as the tag can have subtle local effects that will not be taken into account in the wild-type protein.

Commonly, one measures the PRE of the backbone amide protons ($^{1}H_N$) using a 2D ^{1}H–^{15}N HSQC-based pulse sequence (Iwahara *et al.* 2007). Two examples are shown in Fig. 8. In the first example (Fig. 8*a*), the period for the ^{1}H transverse relaxation measurement (T) is incorporated in the first INEPT scheme. Identical experiments are performed for the paramagnetic and diamagnetic samples to obtain Γ_2. This pulse sequence can readily be modified to a transverse relaxation optimized (TROSY) version, although Boltzmann ^{15}N-magnetization should be destroyed at the beginning the pulse sequence so that all observables are modulated solely by ^{1}H transverse relaxation during the time interval T (Clore & Iwahara, 2009). If a TROSY pulse sequence is being used, it is advantageous to measure the slower TROSY R_2 rate for each sample by placing the delay T at the end of the sequence (Fig. 8*b*). Additionally, 3D experiments can also be employed

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

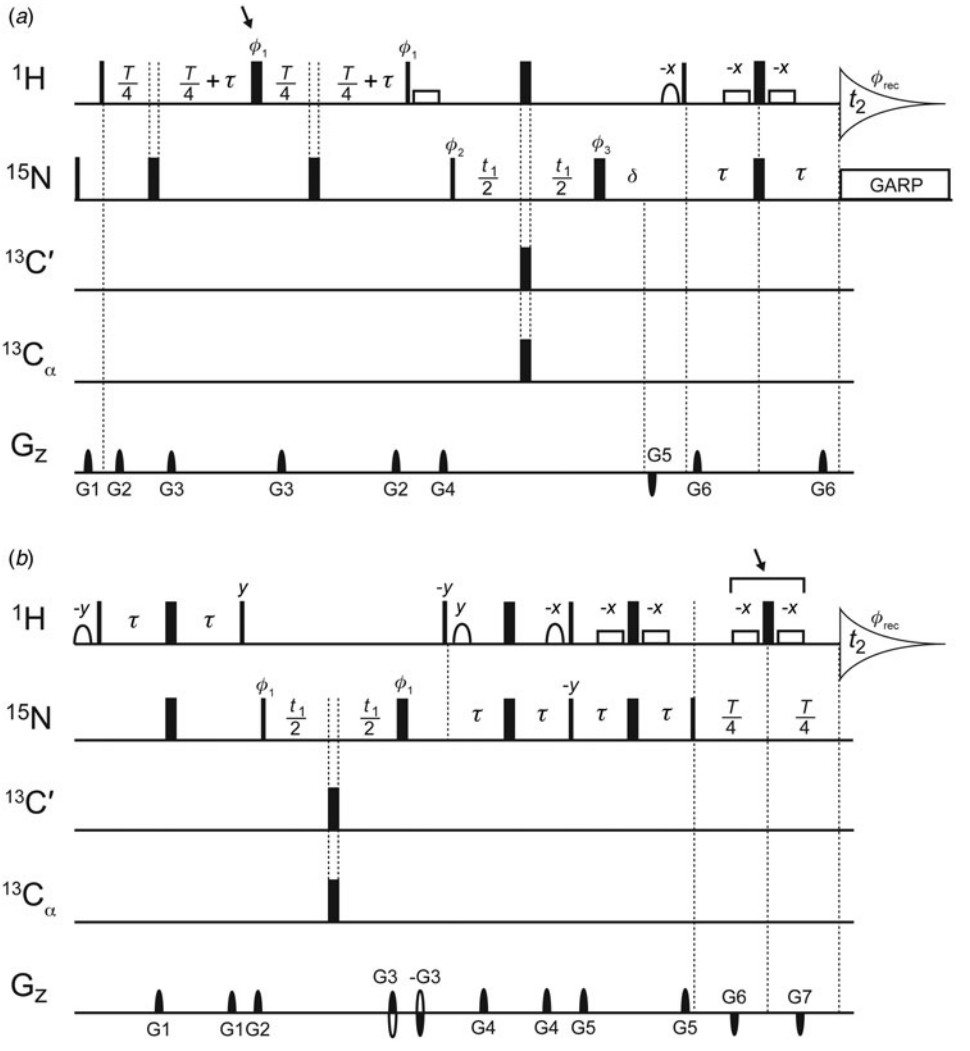


Fig. 8. Pulse sequences for 2D ^1H_N–Γ_2 PRE experiments. (a) HSQC-based pulse sequence for measuring ^1H_N–Γ_2. Full details are given in Iwahara *et al.* (Iwahara *et al.* 2007). The delay T is incremented to measure ^1H_N transverse relaxation. Two 180° ^{15}N pulses are applied to eliminate longitudinal cross-correlated relaxation interference between the ^{15}N–^{1}H dipolar interaction and ^{15}N chemical shift anisotropy. Although $^3J_{HN-H\alpha}$ is active for non-deuterated proteins during the period T, the resulting modulation is cancelled out when Γ_2 is measured by the two-time-point method, using an equal delay T in both the paramagnetic experiment and the diamagnetic control. Alternatively, $^3J_{HN-H\alpha}$ can be directly cancelled by replacing the 180° ^{1}H pulse separating the two halves of delay T (arrow) with an IBURP2 pulse that selectively excites only the amide region of the ^{1}H spectrum. (b) TROSY pulse sequence for measuring ^1H_N–Γ_2. Note that in this pulse sequence the delay T (located at the end of the pulse sequence) is incremented to measure the slower-relaxing TROSY component of the ^1H_N transverse relaxation rate (although ^1H_N–Γ_2 is the same as for R_2 measurements). For proteins that are not deuterated, $^3J_{HN-H\alpha}$ can be cancelled out either by employing the two-time-point method or by replacing the final 180° ^{1}H pulse and adjacent soft water flip-back pulses separating the two halves of delay T (arrow) with an IBURP2 pulse that selectively excites only the amide region of the ^{1}H spectrum. Unless otherwise indicated, all pulses have phase x; $\phi_1 = y, -y, -x, x$; $\phi_{rec} = y, -y, x, -x$. Quadrature detection in the indirect ^{15}N dimension is achieved using the Rance-Kay echo/anti-echo scheme (Kay *et al.* 1992) by inverting gradients G3 and −G3 and using phase $\phi_1 = y, -y, x, -x$ for the second free induction decay (FID) of each quadrature pair.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

to overcome peak overlap. HNCO-based experiments measuring $^{1}H_{N}$–Γ_2 have been developed for this purpose (Anthis *et al.* 2011; Hu *et al.* 2007).

The ^{1}H–Γ_2 PRE can also be measured for protons directly bonded to ^{13}C (Iwahara *et al.* 2004b), and the measurement of methyl ^{1}H–Γ_2 can be particularly useful for larger proteins (Fig. 8). (Note that although R_2 for methyl CH_3 protons is biexponential due to cross-correlation (Ollerenshaw *et al.* 2003), the two-time point procedure described below still gives reasonably accurate Γ_2 values as long as the PRE affects both components of the biexponential decay equally (Clore & Iwahara, 2009)). A scheme has been developed to automatically assign the methyl resonances of a protein given only the measured PRE values and the 3D atomic coordinates of the molecule (Venditti *et al.* 2011).

PRE rates are measured by recording spectra with different time delays (T) in an interleaved manner. R_2 rates are traditionally measured using multiple time points (~8), from $0\times$ to $\sim 1{\cdot}3 \times T_2$. For a paramagnetic experiment, however, more time points may be required to cover the slow and fast relaxing peaks. For each peak, a plot of its intensity *versus* time is fit to a single-exponential decay to determine R_2. Errors can be determined by Monte Carlo analysis. Peaks with a very high PRE will have very low intensity or will have completely disappeared, and it will not be possible to accurately measure their R_2 values.

For measuring PRE rates, however, one can also employ a two-time-point measurement. The merits of such a strategy have been discussed at length previously (Clore & Iwahara, 2009; Iwahara *et al.* 2007), and the use of two time points can offer several-fold time savings over multiple time points. Two-time-point measurements are particularly suitable for PRE measurements because, by subtracting the diamagnetic R_2 from the paramagnetic R_2, relaxation mechanisms common to both states are cancelled out, leaving only the PRE. It is important to note that the same delay (ΔT) must be used for both the diamagnetic and paramagnetic samples to cancel artifacts present in both experiments. (In the case of paramagnetic systems with an anisotropic g-tensor, however, exchange contributions can be significantly different between the paramagnetic and diamagnetic states; Clore & Iwahara, 2009).

In addition to the time savings of reduced data acquisition times, the two-time-point measurement offers the added advantage of allowing one to calculate the PRE rate directly from the intensity of a peak in just four spectra without requiring any curve fitting:

$$\Gamma_2 = \frac{1}{T_b - T_a} \ln \frac{I_{\mathrm{dia}}(T_b) I_{\mathrm{para}}(T_a)}{I_{\mathrm{dia}}(T_a) I_{\mathrm{para}}(T_b)}, \tag{59}$$

where $I_{\mathrm{dia}}(T)$ and $I_{\mathrm{para}}(T)$ are the peak intensities for the diamagnetic and paramagnetic states, respectively, at time T. The total transverse relaxation time, ΔT, is given by $T_b - T_a$. The error in PRE can be propagated from the spectral noise using:

$$\sigma(\Gamma_2) = \frac{1}{T_b - T_a} \sqrt{\left(\frac{\sigma_{\mathrm{dia}}}{I_{\mathrm{dia}}(T_a)}\right)^2 + \left(\frac{\sigma_{\mathrm{dia}}}{I_{\mathrm{dia}}(T_b)}\right)^2 + \left(\frac{\sigma_{\mathrm{para}}}{I_{\mathrm{para}}(T_a)}\right)^2 + \left(\frac{\sigma_{\mathrm{para}}}{I_{\mathrm{para}}(T_b)}\right)^2}, \tag{60}$$

where σ_{dia} and σ_{para} are the standard deviations of the noise in the spectra of the diamagnetic and paramagnetic states, respectively. When using the two-time-point method, the relative error in the recorded PRE value can be minimized by carefully choosing the ΔT value between the two time points. The exact choice of time points will depend on the intrinsic R_2 of the molecule and the anticipated PRE values. The value of ΔT that gives the minimum percent error in a

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

two-time-point measurement is given by:

$$\Delta T \approx \frac{1{\cdot}15}{R_2 + \Gamma_2}, \tag{61}$$

where the exact value of the numerator in Eq. (61) depends on the specific values of R_2 and Γ_2 (ranging from ∼1·1 to ∼1·2, but 1·15 is a useful rule of thumb). The Γ_2 value will vary significantly between different ^1H_N nuclei (as will R_2 to a lesser extent), so one must choose the PRE value that one will want to measure with the greatest precision, giving larger relative errors for values of $R_2 + \Gamma_2$ higher and lower than the chosen value. For a protein with average ^1H_N–R_2 values of 20 s^{-1}, to optimally measure a PRE of 40 s^{-1} (in the middle of the generally useful range of 0 to ∼80 s^{-1}), the optimal value of ΔT would be ∼ 1·15/(20 s^{-1} + 40 s^{-1}) = 19 ms.

A limitation of the two-time-point method is that the results can be heavily influenced by the presence of diamagnetic contamination (e.g. incomplete tagging or dimerization in aged samples). This effect increases for higher amounts of contamination and for larger PRE values, and under extreme cases of diamagnetic contamination causes an apparent PRE much smaller than the true PRE. If one is not careful, peaks close to the noise in cases with large PREs can also suffer from this problem. If diamagnetic contamination is greater than ∼3%, smaller values of ΔT are required for accurate determination of the PRE than predicted by Eq. (61). Diamagnetic contamination can be caused by a variety of factors, including: (a) incomplete conjugation of the extrinsic paramagnetic group and insufficient purification of the conjugated species; (b) the presence of diamagnetic impurities in the paramagnetic stock solution (e.g. trace amounts of diamagnetic metals such as Zn^{2+} and Ca^{2+} in stock solutions of Mn^{2+} or the reduced species of a nitroxide spin label); (c) chemical instability of the conjugated states, especially for disulfide linkages. Thus, it is important to determine the extent of such contamination and what effect it may have on the measured ^1H–Γ_2 values (Clore & Iwahara, 2009; Iwahara *et al.* 2007). On the other hand, intermolecular PRE measurements are not as drastically affected by diamagnetic contamination. As long as the intermolecular association process occurs on a fast timescale, the apparent PRE is simply the true PRE scaled by a factor of $(1 - p_{dia})$ (Clore & Iwahara, 2009).

For a sample with no diamagnetic contamination and an infinite S/N ratio, a plot of apparent PRE *versus* true PRE is a straight line with a slope of 1. In the presence of diamagnetic contamination or significant noise, such a plot still gives a slope of ∼1 for low PRE values, but as the true PRE increases, the slope of the line decreases, eventually becoming negative, yielding apparent PRE values of ∼0 for very large true PREs. One method to check for whether a PRE value is inaccurate due to diamagnetic contamination or another problem (e.g. a low S/N ratio) is to plot the PRE *versus* $\log(I_{dia}/I_{para})$. Most peaks should follow a straight line with a positive slope, but outliers far below this line deserve further scrutiny.

Some studies have taken an approach that we will term a 'single-time-point' method, analyzing the ratio of I_{para} to I_{dia}, sometimes inappropriately referring to this ratio as the PRE. Because such a ratio depends on ^1H–Γ_2 and ^1H–Γ_1, it is influenced by the delays for coherence transfers, the delays between experiments, and the type of data processing (Clore & Iwahara, 2009). Thus, it alone is not physically meaningful, and it is more suitable for qualitative analysis, although quantitative analysis can be performed with certain caveats. Under conditions of full magnetization recovery between experiments, the ratio of peak intensities in the diamagnetic and paramagnetic spectrum are given by:

$$I_{para}(0) = I_{dia}(0)\frac{R_{2,dia}}{R_{2,dia} + \Gamma_2}\exp(-\Gamma_2\tau), \tag{62}$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

where τ is the overall ^{1}H transverse period for the given pulse sequence (e.g. ~10 ms). In addition to requiring a long inter-scan delay ($T_{rep} \gg {}^1\text{H}$–T_1), such an approach requires identical experimental conditions (sample concentration, number of scans, etc.) for the diamagnetic and paramagnetic samples as well as a Lorentzian line shape in the ^{1}H dimension (Iwahara *et al.* 2007). If the ^{1}H–R_2 relaxation rates for the diamagnetic sample are not already known, they will need to be measured separately. Additionally, if an inter-scan delay long enough for full-spin relaxation is not used, one must also take into account the faster longitudinal relaxation in the paramagnetic sample due to the longitudinal PRE (^{1}H–Γ_1):

$$I_{para}(0) = I_{dia}(0)\frac{R_{2,dia}}{R_{2,dia}+\Gamma_2}\frac{1-\exp\{-(R_{1,dia}+\Gamma_1)T_{rep}\}}{1-\exp(-R_{1,dia}T_{rep})}\exp(-\Gamma_2\tau). \tag{63}$$

In the absence of internal dynamics, ^{1}H–Γ_1 can be calculated directly from ^{1}H–Γ_2, using a formula based on Eqs. (44), (46), and (47):

$$\Gamma_1 = \Gamma_2\left(\frac{2}{3}\omega_I^2\tau_c^2+\frac{7}{6}\right)^{-1}, \tag{64}$$

where τ_c is in units of seconds and ω_I is the ^{1}H frequency in radians·s^{-1}. To avoid having to use Eq. (64), one must use a long repetition delay (>4 × $T_{1,dia}$), especially for large and/or deuterated molecules (Iwahara *et al.* 2007). Thus, under these circumstances, performing a single-time-point measurement offers no appreciable time savings over directly measuring the diamagnetic and paramagnetic transverse relaxation rates using multiple time points. Additionally, the requirement of a Lorentzian line shape for the single-time-point experiment limits data processing to the use of either an exponential window function or no window function at all for the direct ^{1}H dimension. The use of broad Lorentzian lines can lead to increased peak overlap and can be especially problematic for larger proteins. This restriction does not apply to multiple-time-point experiments.

When conducting PRE experiments, the use of concentrations of paramagnetic protein that are too high can lead to PREs from random collisions between molecules, similar to solvent PREs. Relatively low concentrations (~0·3 mM) of sample can be used to avoid this undesirable effect, which can be detected by performing experiments with different concentrations, as the solvent PRE is directly proportional to the concentration of the paramagnetic species. For intramolecular PRE measurements, one can perform a control intermolecular experiment, where instead of having the paramagnetic spin label attached to the isotopically labeled species, one can make a sample with a mix of isotopically labeled (but not spin labeled) protein and paramagnetically tagged (but not isotopically labeled) protein. Any PRE measured in such an experiment would be from intermolecular effects.

Two-time-point PRE measurements can be performed within 1–2 days for a well-behaved small-medium-sized protein. For example, after 20 h each of measurement time for the paramagnetic and diamagnetic control, samples of 0·3 mM perdeuterated CaM–4Ca^{2+} measured using a two-time-point (20 ms delay) TROSY 2D ^{15}N–^1H_N experiment at 600 MHz on a cryogenically cooled probe gave errors in Γ_2 (calculated from S/N using Eq. (60)) of 1–2% for $\Gamma_2 \approx 20$ s^{-1} and 2–10% for $\Gamma_2 \approx 50$ s^{-1} (Anthis *et al.* 2011).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

3.3 Data analysis

3.3.1 Fitting PRE data

Once PRE data are acquired, further analysis is needed to gain insight into macromolecular structure; this often involves back-calculating PRE values from atomic coordinates (either to test the agreement between experimental PRE values and a known structure or as part of a routine to calculate a new structure or ensemble from PRE data). Although it is tempting to interpret PREs directly as distances, this approach is problematic due to the flexibility of most common PRE spin labels (and the inherent flexibility of the macromolecules under study). As the PRE is a population-weighted $\langle r^{-6} \rangle$ average of all states in exchange on a fast timescale, the PRE generally does not correspond to a single distance, and the contribution of a particular state can be very different than its fractional population (hence the utility of PRE for detecting minor states). Historically, this has made rigorous treatment of PRE data complicated, but computational advances have made such studies feasible.

One solution has been the incorporation of an ensemble average approach (Iwahara *et al.* 2004b), which has been implemented in the macromolecular structure determination software package Xplor-NIH (Schwieters *et al.* 2003, 2006). Here, the paramagnetic group is represented by an ensemble of multiple conformers, and PRE values are the sum of the PREs for the individual conformers. Thus, using the 'prePot' pseudoenergy term in Xplor-NIH, one can refine directly against the experimental PRE values. Scripts and detailed instructions for these procedures are available online at http://spin.niddk.nih.gov/clore.

The first step in back-calculating PRE data in Xplor-NIH is the preparation of a suitable .pdb file, beginning with the *in silico* attachment of a spin label. The process is streamlined for common MTSL and EDTA tags. To prepare the .pdb file, one first changes the name of the residue where the tag is to be attached to 'CYSP' (for MTSL) or 'CED' (for EDTA). A script can then be run that fills in the missing atoms and regularizes their geometry (using the regularize module in Xplor-NIH). If the tag is not already in the Xplor-NIH library, one must generate parameter files and then attach the tag to the structure *in silico*. In order to use the SBMF PRE method, multiple conformations (at least 3) of the tag need to be generated.

Tag conformations are then refined by minimizing the target function for the PRE restraints, E_{PRE}, defined as:

$$E_{\mathrm{PRE}} = k_{\mathrm{PRE}} \sum_i w_i \left(\Gamma_2^{\mathrm{obs}}(i) - \Gamma_2^{\mathrm{calc}}(i) \right)^2, \tag{65}$$

where k_{PRE} is a force constant, w_i is a weighting factor, $\Gamma_2^{\mathrm{obs}}(i)$ is the observed PRE, and $\Gamma_2^{\mathrm{calc}}(i)$ is the back-calculated PRE. The paramagnetic group is represented by an ensemble of N different conformers, while the remainder of the molecule is represented by a single structure. The individual members of the ensemble are treated independently and are allowed to overlap by excluding their interactions from the non-bonded energy term. This approach provides equal weights of $1/N$ to each member of the ensemble, but such an approach can account for uneven distributions if a larger N (>3) is used, since members of the ensemble can overlap in the simulation. The ensembleSimulation (Clore & Schwieters, 2004b) module in Xplor-NIH (Schwieters *et al.* 2003, 2006) can also be used to explicitly specify or optimize different weights for each member of the ensemble (Deshmukh *et al.* 2013).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

To calculate the PRE value, the ensemble-averaged distance $\langle r^{-6} \rangle$ for the distance between the observed nucleus (n) and the paramagnetic center (e) must be calculated as follows (Iwahara *et al.* 2004b):

$$\langle r^{-6} \rangle = \frac{1}{N n_{\mathrm{p}}} \sum_{e}^{N} \sum_{n}^{n_{\mathrm{p}}} r_{en}^{-6}, \tag{66}$$

where N is the number of ensembles used to represent the paramagnetic group and n_{p} is the number of equivalent protons (e.g. 1 for NH or CH, 2 for NH_2 or CH_2, and 3 for CH_3). Γ_2 can then be directly calculated from the SB (Eq. (44)) or SBMF (Eq. (49)) equations; the SBMF equation is more accurate but requires additional computational time). Use of the SBMF equation requires the determination of an order parameter (S^2) for each electron–nucleus vector, as defined by Eq. (51). In Xplor-NIH, order parameters are calculated on the fly from the ensemble of paramagnetic conformers, operating under the assumption that the fluctuations in the position of the measured nucleus are small compared with those of the paramagnetic center. Order parameters can be calculated using Eq. (52) together with the following equations (Iwahara *et al.* 2004b):

$$S^2_{\mathrm{PRE,angular}} = \frac{1}{N^2 n_{\mathrm{p}}^2} \sum_{e,d}^{N} \sum_{n.m}^{n_{\mathrm{p}}} \left\{ \frac{3}{2} \left(\frac{\vec{r}_{en} \bullet \vec{r}_{dm}}{r_{en} r_{dm}} \right)^2 - \frac{1}{2} \right\}, \tag{67}$$

$$S^2_{\mathrm{PRE,radial}} = \frac{\left(\sum_{e}^{N} \sum_{n}^{n_{\mathrm{p}}} r_{en}^{-3} \right)^2}{N n_{\mathrm{p}} \sum_{e}^{N} \sum_{n}^{n_{\mathrm{p}}} r_{en}^{-6}}, \tag{68}$$

where r_{en} and r_{dm} are vectors between the paramagnetic center (e, d) and the observed nucleus (n, m) in two different conformers. The values calculated in Eqs. (67) and (68) are the sum of contributions for each pair of conformers in the ensemble.

To fit the PRE values to the structure of a protein, a simulated annealing routine is performed with the protein molecule held rigid and the tags allowed to move to fit the data. Adjacent side chains can also be allowed to move to resolve steric overlap with the tag. At each step, Xplor-NIH will recalculate the order parameter for each PRE vector from the position of the different members of the tag ensemble. Fitting also requires the specification of correlation times, which can be explicitly provided or fit along with the tag positions in Xplor-NIH. If using the SB mode (i.e. Eq. (44)), only a single overall correlation time (τ_{c}) is required, which depends on both the macromolecule rotational correlation time (τ_{r}) and the electron spin relaxation time (τ_{s}) (Eq. (45)). If $\tau_{\mathrm{s}} \gg \tau_{\mathrm{r}}$ (e.g. nitroxide spin labels (Bernini *et al.* 2009)), then the value of τ_{s} can be ignored, and τ_{c} is just simply τ_{r}. However, if τ_{s} is of the order of, or shorter than, τ_{r} (e.g. paramagnetic metal ions), then the value of τ_{s} must be taken into account. Under rigid SB assumptions, τ_{c} can be calculated directly from the ratio of transverse and longitudinal PRE values (Iwahara *et al.* 2003):

$$\tau_{\mathrm{c}} = \frac{1}{\omega_{\mathrm{I}}} \sqrt{\frac{3}{2} \frac{\Gamma_2}{\Gamma_1} - \frac{7}{6}}, \tag{69}$$

where τ_{c} is in units of seconds and ω_{I} is the ^{1}H frequency in rad·s^{-1}. Equation (69) is only formally true for the SB approximation; however, even under SBMF conditions it offers a useful estimate of τ_{c} (yielding maximum errors in the effective $\Gamma_2^{-1/6}$-calculated distance r of 4% for an order parameter of $S^2 = 0{\cdot}5$, with smaller errors for larger order parameters) (Iwahara *et al.* 2004b). To avoid these errors, however, and in the absence of additional data, τ_{c} can instead

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

be optimized within a pre-defined range (Iwahara *et al.* 2004b). If one is using the SBMF mode (i.e. Eq. (49)), which is more appropriate unless a rigid paramagnetic center is used, one must also take into account (in addition to τ_c) the total correlation time τ_t that incorporates τ_r, τ_s, and the internal correlation time for the PRE vector (τ_i) (Eq. (G)). If $\tau_r \gg \tau_s$ and $\tau_r \gg \tau_i$, one can use the approximation $\tau_t \approx 0$, which assumes rapid motion of the PRE spin label. This in effect makes the order parameter S^2 a scaling factor for the back-calculation of PRE data. This approximation does not hold true for shorter values of τ_s. The prePot potential in Xplor-NIH also offers the option of optimizing τ_c and/or τ_t (or τ_i) as part of the simulated annealing routine. Doing so then allows calculations to be performed with no prior knowledge of correlation times.

With these factors in place, Xplor-NIH can perform a simulated annealing calculation, refining directly against the experimental PREs, minimizing a target function including a pseudoenergy term calculated from the difference between back-calculated and experimental PRE values as well as terms to maintain proper geometry and avoid steric clashes. The goodness of fit of the PREs can be assessed by the Q-factor, which is analogous to the crystallographic R-factor and decreases with improved agreement between the observed (Γ_2^{obs}) and back-calculated (Γ_2^{calc}) rates (Iwahara *et al.* 2004b):

$$Q = \sqrt{\frac{\sum_i \left(\Gamma_2^{obs}(i) - \Gamma_2^{calc}(i)\right)^2}{\sum_i \Gamma_2^{obs}(i)^2}}. \tag{70}$$

Alternatively, one can also use a reduced χ^2 statistic:

$$\chi^2 = \frac{1}{M}\sum_N \left(\frac{\Gamma_2^{obs} - \Gamma_2^{calc}}{\sigma\left(\Gamma_2^{obs}\right)}\right)^2, \tag{71}$$

where the degrees of freedom (M) are given by:

$$M = N - 3C, \tag{72}$$

where N is the number of observed PREs and C the number of spin label conformers in the fitting ensemble (C is multiplied by three for the three translational degrees of freedom involved in fitting each nitroxide position). Other measures of the goodness of fit include the root mean-square-deviation (RMSD) between the observed and back-calculated PREs and the correlation coefficient from a linear regression fit of a scatter plot of the back-calculated *versus* observed PREs.

The representation of the paramagnetic label as an ensemble of multiple conformers is essential for fitting PRE data arising from flexible labels (e.g. MTSL). The use of a three-conformer ensemble improved the fit of PRE data from SRY/DNA complexes incorporating dT-EDTA-Mn^{2+} without overfitting the data – as indicated by a reduction in the free Q-factor calculated from data excluded from structure calculations and by the improvement of the fit of RDCs to the calculated coordinates (Iwahara *et al.* 2004b). More rigid spin labels, however, do not require an ensemble representation. For example, it was demonstrated that, whereas multiple tag conformations were required to fit PRE data in CaM conjugated to MTSL, a single conformation was sufficient when CaM was conjugated to a pyridyl analog of MTSL (Fawzi *et al.* 2011a).

A scatter plot of the back-calculated PREs *versus* the experimental PREs is revealing of how good the fit is. However, for understanding the PREs in terms of structural features, a plot of PRE *versus* residue number (plotting both the back-calculated PREs and the experimental

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

PREs) can be more informative. The key feature that indicates that a minor state might be present is the measurement of significantly higher PRE values than those back-calculated from the structure. However, before attempting to identify and characterize minor populations, one must first ensure that the experimental PRE data cannot be reasonably explained by a single structure. If these efforts fail to sufficiently explain the measured PREs, one can then attempt to calculate an ensemble of new structures to explain the data.

3.3.2 Ensemble simulated annealing calculations

If the PRE data collected on a system cannot be explained by known structures, the calculation of new structures or ensembles of structures may be necessary to satisfy the experimental data. The calculation of a molecular ensemble from average (in this case PRE) data is an inverse problem that is ill-posed and underdetermined for large ensembles and highly flexible systems. Thus, various approaches have been proposed (see below). Here, we describe the calculation of minor state structures using ensemble rigid body simulated annealing calculations using the ensembleSimulation module of Xplor-NIH (Clore & Schwieters, 2004b).

In performing a rigid body refinement, the internal structure of all or part of a macromolecule is constrained to behave as a rigid body that is only able to translate and rotate as a cohesive unit (Clore, 2000; Clore & Bewley, 2002). For example, CaM comprises two domains connected by a flexible linker; thus in PRE-driven ensemble simulated annealing calculations, each domain was treated as a rigid body, and residues in the linker were given full torsional degrees of freedom (Anthis *et al.* 2011). Side-chain residues at interfaces can also be allowed to move to prevent steric overlap. Rigid body refinement is appropriate only when the structure of the domains held rigid are not expected to change appreciably from the starting structure, a condition that can be validated by checking the agreement between experimentally measured RDCs and those back-calculated from the starting structure. The need for rigid body refinement arises from the sparse and long-range nature of the PRE, compared with conventional (NOE-based) NMR structural restraints. If sufficient additional data are available, full atomic structure calculations can be performed, although they will be computationally more expensive and the results may be ill defined compared with the results of a rigid-body refinement.

Additional points to consider are how many individual sets of PRE data (each from a molecule with a single paramagnetic tag) need to be included in the structure calculation and where the paramagnetic tags should be placed. Unless detailed information about the minor state(s) is known beforehand, the goal will be to have enough paramagnetic tags so that each position on the surface of the molecule would be within a distance that would generate a measurable PRE from at least one possible minor state conformation. Because some tag positions may disrupt the structure of the protein or have other undesirable effects, it is often necessary to test many more potential tag positions than might ultimately be required. In the case of CaM, which comprises two 8-kDa globular domains, we found that reproducible results required at least one tag on each of its two domains. Regardless of how many data sets are ultimately used, they can all be included in a single simulated annealing calculation.

Before calculating new ensemble structures to fit PRE data, the conformations of the spin labels must first be set using the fitting procedure described in Section 3.3.1, incorporating only the intramolecular or intradomain PRE data (i.e. data that comes primarily from the major state). These tag positions can then be maintained in the same conformations for the subsequent rigid body

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

fitting. When performing simulated annealing calculations involving ensembles of minor states, one must determine an appropriate number of minor states and an appropriate population to fit the experimental data without overfitting. This can be accomplished by performing a grid search, i.e. performing multiple calculations while independently varying the number of minor states and the population attributed to each minor state. Increasing the number of minor states or the population attributed to each state should improve the fit (i.e. reduce the Q-factor) up to a certain threshold. The number N of minor states above which the Q-factor exhibits only marginal reduction is a good working value to use for calculations, as larger values of N will offer little improvement to the fit and will require proportionally more computational time. The minimum population and value of N needed to fit the data can be more rigorously determined by also calculating Q_{free} (calculated the same as the Q-factor (Eq. (70)), but using a small percentage (e.g. 5%) excluded from the data set used for refinement). The value of N needed to fit the data varies from system to system, and a value under 10 is generally sufficient, even for fairly heterogeneous minor states. Likewise, the population above which the Q-factor exhibits only marginal reduction offers a good estimate of the true minor state population. For example, $N = 8$ was sufficient for CaM, in the presence or absence of calcium. A minor state of 10% was necessary to fit the calcium-loaded CaM data, but only 5% was needed for apo CaM (Anthis *et al.* 2011). New updates to Xplor-NIH allow users to directly fit the populations of ensemble members (Deshmukh *et al.* 2013); however, this approach is applicable primarily to ensembles of more highly populated states because the populations of individual members of a minor state ensemble derived from PRE data can be ill-defined due to the dependence of the PRE on both the distance from the paramagnetic center and the population of each ensemble member.

Ensemble calculations can be computationally expensive, although the time required can be minimized through parallelization. The ensembleSimulation module in Xplor-NIH offers two types of parallelization (Clore & Schwieters, 2004b), which can be used simultaneously. As with conventional structural determination in Xplor-NIH, individual ensembles can be calculated independently on different processors, as individual structures can. An additional level of parallelization can be added when using ensembleSimulation, as each member of the ensemble can also be calculated semi-independently on different processors.

The results of an ensemble simulated annealing calculation can be difficult to visualize, given that each ensemble could have tens of members, and the calculation might be repeated 100 or more times. Trying to visualize all individual ensemble members, especially at atomic resolution, can be both overwhelming and uninformative. Other visualization schemes, such as showing just the center of mass for each ensemble member, involve losing most detailed structural information. Reweighted atomic probability density maps, however, allow the consolidation of the full ensemble into a single representation, preserving many structural features and highlighting the major contributing members to the ensemble – a feature that may be lost in other representations (Schwieters & Clore, 2002).

For visualizing NMR ensembles, instead of using an evenly weighted map, it is advantageous to calculate a probability density map that has been reweighted so that the maximum contribution of each atom is equal, allowing visualization of even the highly flexible regions of the molecule (Schwieters & Clore, 2002). This is accomplished by defining the density at each point $\mathbf{q}$ as (Schwieters & Clore, 2002):

$$\rho_{\text{reweight}}(\mathbf{q}) = \frac{1}{N_s} \sum_{l=1}^{N_a} w_l \sum_{k=1}^{N_s} \rho_a(\mathbf{q}, \mathbf{q}_{kl}), \tag{73}$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

where N_s is the number of member structures in the ensemble, N_a the number of atoms in each structure, $\mathbf{q}_{kl}$ the position of atom l in each ensemble member, and w_l a weighting factor. $\rho_a(\mathbf{q}, \mathbf{q}_{lk})$ is a quartic function, ranging from 0 to 1, that represents the electron density for each atom (Schwieters & Clore, 2002):

$$\rho_a(\mathbf{q}, \mathbf{q}_{kl}) = \begin{cases} \frac{105}{32\pi a^7}\left(d(\mathbf{q}, \mathbf{q}_{kl})^2 - a^2\right) & d(\mathbf{q}, \mathbf{q}_{kl}) < a, \\ 0 & d(\mathbf{q}, \mathbf{q}_{kl}) \geq a, \end{cases} \tag{74}$$

where $d(\mathbf{q}, \mathbf{q}_{lk})$ is the distance between point $\mathbf{q}$ and the center of the atom $\mathbf{q}_{kl}$, and a (set to 1 Å in Xplor-NIH) is defined so that $\rho_a(\mathbf{q}, \mathbf{q}_{lk}) = 0{\cdot}5$ at $d(\mathbf{q}, \mathbf{q}_{lk}) = a \times (1 - 1/\sqrt{2})^{1/2}$. The weighting factor for each atom is defined as (Schwieters & Clore, 2002):

$$\frac{1}{w_l} = \max \sum_{k=l}^{N_s} \rho_a(\mathbf{q}, \mathbf{q}_{kl}). \tag{75}$$

These maps can be generated in Xplor-NIH and then visualized in a molecular graphics program such as VMD-Xplor (Schwieters & Clore, 2001) or PyMOL (http://www.pymol.org). Showing this map at different contour levels gives a good indication of where the minor states are located, rather than looking at a large jumble of atoms, bonds, or ribbons. For the clearest results, it is often advantageous to generate multiple different maps in order to visualize the ensembles from multiple perspectives. For example, first, all structures could be aligned to rigid body A, and density maps could be calculated and visualized for rigid body B. Then, all structures could be aligned to rigid body B, and density maps could be calculated and visualized for rigid body A. The calculation and visualization of minor state ensembles allow structural insight into these important, transient species that have previously eluded structural characterization. The extent of insight possible will depend upon the heterogeneity of these ensembles, but if the structures are consistent from one ensemble to the next, it will be possible to analyze specific molecular contacts with a high degree of certainty. For larger or more heterogeneous ensembles, one might be restricted to identifying binding interfaces or other large-scale structural features. Regardless, the insights gained through these calculations shed substantial light on these elusive states and can provide a stepping stone for further biochemical study and manipulation of these minor states.

The PRE-driven ensemble calculation approach has been used to visualize transient interactions between the N-terminal domain of enzyme I (EIN) and the phosphocarrier protein HPr (Tang *et al.* 2006, 2008a), between cytochrome *c* and cytochrome *c* peroxidase (Volkov *et al.* 2006, 2010), between the N- and C-terminal domains of apo MBP (Tang *et al.* 2007), between subunits of the HIV-1 protease precursor protein (Tang *et al.* 2008a), and between the N- and C-terminal domains of the calcium-sensing protein CaM (Anthis *et al.* 2011). Because these interdomain or intermolecular contacts only exist in a minor proportion (<10%) of the total population, they are invisible to conventional structural methods and could only be visualized by long-range paramagnetic NMR measurements. Thus, PRE can unlock fundamental insights into macromolecular-binding events, and many of these studies highlight the subtle interplay between induced fit and conformational selection in these binding events (Clore, 2013).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

3.4 Other approaches to treating PRE data

3.4.1 Maximum occurrence (MO), ensemble selection, and MD

In the previous sections, we discussed analyzing PRE data by fitting to known structures using simulated annealing calculations and characterizing minor states by using PRE-driven ensemble simulated annealing calculations. Various other methods are also available for calculating representative structural ensembles of a dynamic system (based on PRE data or on a variety of other NMR and non-NMR [e.g. SAXS] measurements), and these can broadly be placed in two categories (Fisher & Stultz, 2011). The first category is the calculation of ensembles via simulated annealing calculations (Clore & Gronenborn, 1991; Clore & Schwieters, 2004a, b, 2006), as described above. Another broad class of methods involves first the generation of a large number of structures and then selection amongst these of the most appropriate ensemble to fit the experimental data. Examples include ENSEMBLE (Choy & Forman-Kay, 2001), the ensemble optimization method (EOM) (Bernado *et al.* 2007), ASTEROIDS (Nodet *et al.* 2009), minimum ensemble selection (MES) (Pelikan *et al.* 2009), MO (Bertini *et al.* 2010), and sparse ensemble selection (SES) (Berlin *et al.* 2013).

Some of these methods use molecular dynamics (MD) simulations to generate the initial set of individual conformers from which the final ensemble is selected. However, MD can be incorporated into the analysis of paramagnetic data in various ways. The simulated annealing approach discussed previously involves MD calculations with a simplified force field. Paramagnetic data and other NMR observables can also be incorporated directly into equilibrium MD simulations that also include explicit solvent and electrostatic force fields. Additionally, MD can be used to generate ensembles that can be analyzed for the fit of the experimental data. This approach was used to characterize CaM bound to the IQ-recognition motif from the voltage-gated calcium channel $Ca_v1{\cdot}2$; the CaM/IQ complex had been crystallized in three different conformations, but in order to fit the PCS and paramagnetic RDC data on this complex in solution, all three crystal conformations plus four additional conformations calculated by MD had to be included (Russo *et al.* 2013).

Many different ensemble selection methods have been used for dynamic systems, but one that has been used extensively with paramagnetic data is the MO or maximum allowed probability (MAP) method described by Bertini and colleagues (Bertini *et al.* 2007, 2010, 2012a; Fragai *et al.* 2013) based on prior theoretical work (Gardner *et al.* 2005; Longinetti *et al.* 2006). This method has been used for paramagnetic PCS and RDC data (in combination with SAXS data) (Bertini *et al.* 2010) as well as PRE data (Bertini *et al.* 2012b). A similar approach for PRE data has also been taken by Volkov *et al.* (2010). The MO approach involves first generating individual structures, as with other ensemble selection methods, and then determining a maximum population ('occurrence') that can be ascribed to each structure without violating the data. The result is a statistically rigorous view of the minor conformations observed by paramagnetic data, although without producing a tangible ensemble that can be used for further analysis.

The MO approach first requires generating a pool of random structures that comprehensively sample the rotational and translational conformational space accessible to the system. For a protein with two domains connected by a flexible linker (e.g. CaM), ~50 000 individual structures were required (Bertini *et al.* 2010). For the transient interaction between two globular proteins (e.g. cytochrome *c* and cytochrome *c* peroxidase) ~1 000 000 individual structures were required to sample the interaction surface in 1-Å increments (Volkov *et al.* 2010). A random structure generation protocol can be used to generate individual structures that fulfill basic criteria, such as lack of steric clashes between proteins or domains (Grishaev *et al.* 2012).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

The MO procedure (Bertini *et al.* 2012a) is then performed for some subset of these (thousands to millions) individual structures. First, a structure is randomly selected and given a population. Then, an ensemble of a pre-determined size is selected to best fit the experimental data. A target function (e.g. Eq. (65)) is calculated, and this procedure is repeated for a range of different populations of the selected structure. The MO of this structure is defined as the maximum population that does not exceed the pre-set threshold value for the target function. The entire procedure is then repeated for a certain subset of the random structures available. The MO of each structure corresponds to an upper limit of the structure's population, but as the quantity of experimental data increases, the MO for each conformation should converge to its true population (Fragai *et al.* 2013). The different types of data that have been used in MO studies are highly complementary, being sensitive to either distance from a paramagnetic center (PRE), relative orientation (RDC), both of these (PCS), or overall molecular shape (SAXS). Additionally, the incorporation of multiple paramagnetic tag positions or the use of multiple paramagnetic metals with different *g*-tensors can also help to further refine the MO values (Fragai *et al.* 2013).

The MO method can be computationally expensive, although it lends itself well to parallelization. For this reason, the MaxOcc web portal (http://py-enmr.cerm.unifi.it/access/index/maxocc) has been set up to allow MO calculations on the e-NMR grid infrastructure (Bertini *et al.* 2012a). Another difficulty of MO analysis is finding a reasonable way to visualize the data. To simplify the presentation, the center of mass of each protein or domain can be represented as a dot, although this removes any orientational information. Instead of a dot, a set of axes can be used to represent the orientation of the molecule or domain, although information on interfaces and molecular contacts are still lacking. The MO can be indicated on such a representation through color coding. Also, unlike the ensemble simulated annealing approach, the MO approach does not generate concrete structural ensembles that can be used for further analysis.

The MO approach has been used extensively to characterize CaM (Bertini *et al.* 2010, 2012b; Dasgupta *et al.* 2011; Luchinat *et al.* 2012). These studies have incorporated PCS, RDC, PRE, and/or SAXS data from CaM, primarily using an N60D mutant that selectively binds lanthanide ions in one of the two N-terminal Ca^{2+}-binding sites of CaM. These studies indicate that the two domains of CaM are capable of sampling a wide range of conformational space with respect to one another. Another study demonstrated that even when bound to a peptide from myelin basic protein, CaM was still highly dynamic (Nagulapalli *et al.* 2012). This flexible complex differed from the more rigid complex that Ca^{2+}-loaded CaM forms with many peptide targets, but similar to that formed between apo CaM and a peptide from myosin light-chain kinase (Anthis *et al.* 2011). The MO approach was also used with SAXS, RDC, and PCS restraints to characterize matrix metalloproteinase-1 (MMP-1), revealing that MMP-1 adopts a more open structure in solution than that observed by X-ray crystallography (Cerofolini *et al.* 2013). A similar approach was also used to map the encounter complex between cytochrome *c* and cytochrome *c* peroxidase, although PRE-driven ensemble simulated annealing calculations were used to further refine the model of the encounter complex (Volkov *et al.* 2010), demonstrating that the encounter complex adopts conformations nearby in space to the major species observed by X-ray crystallography (Volkov *et al.* 2006).

3.4.2 PRE studies on intrinsically disordered systems

PRE has proven to be a powerful method for studying unfolded and intrinsically disordered proteins, which, due to their dynamic nature, often defy characterization by other methods (Clore &

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Iwahara, 2009; Clore *et al.* 2007). Many studies on denatured systems have revealed evidence of residual structure in the unfolded state; examples include denatured staphylococcal nuclease (Gillespie & Shortle, 1997a, b), the unfolded form of the N-terminal SH3 domain of drk (in slow-exchange equilibrium with the folded form) (Choy & Forman-Kay, 2001; Marsh & Forman-Kay, 2009; Marsh *et al.* 2007), acid-denatured apomyoglobin (Felitsky *et al.* 2008; Lietzow *et al.* 2002; Silvestre-Ryan *et al.* 2013), denatured acyl coenzyme A-binding protein (Kristjansdottir *et al.* 2005; Ozenne *et al.* 2014), and the γ-subunit of cGMP phosphodiesterase (Song *et al.* 2008). (However, PRE data on other systems have indicated that native-like populations constitute a very small or non-existent population in the disordered state (Cliff *et al.* 2009; Vise *et al.* 2007)). Studies on the intrinsically unstructured and aggregation-prone protein α-synuclein have revealed a more compact average hydrodynamic radius than expected for a truly random coil peptide as well as transient long-range interactions that appear to inhibit oligomerization and aggregation (Bertoncini *et al.* 2005; Dedmon *et al.* 2005). PRE can also be used to identify interactions that could be productive for aggregation, as observed for the protein tau (Peterson *et al.* 2008).

Since disordered systems cannot be described by a single structure, ensemble approaches are usually appropriate (Fisher & Stultz, 2011), including some of the methods described in Sections 3.3.2 and 3.4.1), although a very large and heterogeneous ensemble may be required to describe an unstructured system. Other alternative treatments of the data are also possible. For example, it has been demonstrated that the calculation of a single average structure can provide high-resolution information about residual molecular contacts in acid-unfolded apomyoglobin (Silvestre-Ryan *et al.* 2013), although it should be recognized that this does not provide an accurate representation of the reality in solution.

PRE data can be used to characterize flexible parts of otherwise more ordered systems. For example, recently the pH gating ability of loop 6 of the β-barrel outer membrane protein G was studied by PRE (Zhuang *et al.* 2013). PRE data have been used for studying systems involving structured domains linked by a flexible linker, including CaM (Anthis *et al.* 2011). The nature of the linker in this case was explored in further detail, as the extent of association between the two domains of CaM was determined by interdomain PRE measurements for linkers of varying length, and the results were fit to a simple random-coil model of the linker describing the effective concentration of the ends of the linker as a function of linker length (Anthis & Clore, 2013).

3.4.3 Solvent PRE

In addition to the measurement of PREs arising from a spin label or paramagnetic ion bound directly to a macromolecule (intramolecular PRE) or from a tag or ion bound to another macromolecule (intermolecular PRE), PREs arising from small co-solute molecules (solvent PRE) can also be measured. Such co-solute molecules include molecular oxygen (O_2 (Hernandez *et al.* 2002; Sakakura *et al.* 2005)), nitroxides (e.g. TEMPOL (Fesik *et al.* 1991; Petros *et al.* 1990)) and chelated metal ions (e.g. Gd-diethylenetriamine pentacetic acid-bismethylamide [Gd-DTPA-BMA or 'Omniscan'] (Iwahara *et al.* 2006; Pintacuda & Otting, 2002)). These particular co-solutes are neutrally charged, making them particularly appropriate as probes for solvent accessibility, although care must be taken with TEMPOL, which can preferentially interact with hydrophobic surfaces. Solvent PRE measurements can be used as structural restraints as a measure of solvent accessibility or proximity to the surface of a macromolecule, and they can also

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

be used to detect intermolecular interactions (i.e. decreased PRE values near the interaction surface) (Iwahara *et al.* 2006; Sakakura *et al.* 2005).

Analyzing solvent PRE data is complicated by the need to take into account the diffusion rates and correlation times of many co-solute molecules in addition to those of the macromolecule. Different theoretical frameworks have been proposed (Bernini *et al.* 2009; Clore & Iwahara, 2009), including an inner-sphere model and an outer sphere model. In the inner-sphere model, the solvent PRE is assumed to arise from interactions between the paramagnetic protein and the macromolecular surface (Bloembergen, 1957; Solomon, 1955; Solomon & Bloembergen, 1956), whereas in the outer-sphere model, the solvent PRE is assumed to arise from paramagnetic centers that behave as diffusive solvent molecules (Freed, 1978; Hwang & Freed, 1975). Hwang and Freed have developed a theoretical framework to describe solvent PRE values by such an outer-sphere model, although due to its complexity it is difficult to apply to macromolecular structures.

Otting and LeMaster independently proposed an empirical method for back-calculating solvent PRE values from macromolecular structures (Hernandez *et al.* 2002; Pintacuda & Otting, 2002). The macromolecule is imbedded in a tightly spaced cubic lattice *in silico*, and the solvent PRE for a given nucleus in the macromolecule is the sum of the r^{-6} distances to each of the lattice points (excluding those that overlap with the macromolecule), weighted by a single empirical factor, a, which takes into account the concentration of the paramagnetic co-solute as well as its dynamics. This approach has been applied to solvent PRE data from O_2 on ^{1}H nuclei in rubredoxin (Hernandez *et al.* 2002), from Gd–DTPA–BMA on ^{1}H nuclei in ubiquitin (Pintacuda & Otting, 2002), and from Gd–DTPA–BMA on ^{1}H nuclei in specific and non-specific protein/DNA complexes (Iwahara *et al.* 2006).

Some of these same solvent PRE co-solvents can be used in imaging studies as contrast agents. Conventionally, a non-toxic paramagnetic molecule (generally chelated Gd^{3+}) is injected, and contrast is provided by the enhancement to longitudinal relaxation, measured as an increase in signal intensity where greater concentrations of the contrast agent are present (Caravan *et al.* 1999). Contrast agents have also been developed that take advantage of CEST and, more recently, paramagnetic contrast agents have been developed for paraCEST (chelated lanthanide ions, especially Eu^{3+}) to generate even more dramatic chemical shift differences by hyperfine shifts for CEST imaging (Hancu *et al.* 2010; Vinogradov *et al.* 2013).

3.5 PRE in practice: application to CaM

CaM is the prototypical calcium-sensing protein in eukaryotes (Yamniuk & Vogel, 2004), comprising two globular domains, each with two Ca^{2+}-binding sites, connected by a short linker. CaM is a highly dynamic and pliable protein that binds to a wide range of target sequences, exerting different effects on various signaling pathways. A common mechanism is the release of target self-inhibition by binding to a regulatory peptide sequence and sequestering it to relieve inhibition. In such complexes, CaM forms a closed, compact structure, with its two domains in close proximity on opposite sides of a helical target peptide (Ikura *et al.* 1992; Meador *et al.* 1992) (Fig. 9). In the absence of target peptide, however, CaM forms a much more open structure. Although X-ray crystallography studies have consistently showed peptide-free Ca^{2+}-loaded CaM (CaM-4Ca^{2+}) in a highly extended 'dumbbell' conformation (Babu *et al.* 1985), studies using a range of techniques, including NMR, have shown that the central linker is highly flexible (Barbato *et al.* 1992) and that in solution CaM forms structures intermediate between the fully

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

extended and compact extremes (Trewhella, 1992). However, there was previously no direct evidence for contacts between the two domains, and it was unclear whether compact states like those in the peptide-bound structure were accessible in solution, meaning that such compact states would have to occupy a low proportion of the total population.

In order to detect and characterize sparsely populated compact states in CaM, we collected PRE data on CaM with and without peptide (and with and without calcium; Anthis *et al.* 2011). Using paramagnetic nitroxide spin labels on each domain, we found that Ca^{2+}-loaded peptide-bound CaM PRE data were fully consistent with a previously published crystal structure. However, even in the absence of peptide, significant interdomain PREs were observed that were inconsistent with the 'dumbbell' structure (which predicts no significant interdomain PREs) and which could not be explained by a combination of known structures. CaM also displayed interdomain PREs in the absence of Ca^{2+}, although these PREs were much smaller than for CaM-Ca^{2+}, and they were largely unaffected by the addition of target peptide. Because these PREs in peptide-free CaM could not be explained by pre-existing structures, we performed new calculations to characterize the minor states of CaM with close interdomain contacts.

PRE-driven simulated annealing calculations were performed in Xplor-NIH with the two domains of CaM treated as rigid bodies and the residues of the central linker given torsional degrees of freedom. The major species was represented by CaM in the 'dumbbell' conformation. Simulations were performed while varying both the size (N) and population of the minor state ensemble, and the goodness of fit was ascertained by a decrease in the Q-factor. Increasing the number N of molecules in the minor state improved the fit substantially up to $N = 8$, and then leveled off. Likewise, increasing the population improved the fit, but only to ~10% for CaM-4Ca^{2+} and ~5% for apo CaM. The improvement in fit can also be seen qualitatively in plots comparing the experimental PRE values with the values back-calculated from the ensembles (Fig. 9).

These simulations were performed using the SBMF mode of prePot in Xplor-NIH (Iwahara *et al.* 2004b), with PRE order parameters calculated on the fly from a five-member ensemble of spin label conformers. This calculation was performed independently for each member of the minor state ensemble, because the motion of the spin label is much faster (ps) than the interdomain motion of CaM (ns). Each simulated annealing calculation was run 100 times, and the results were represented as reweighted atomic probability density maps (Schwieters & Clore, 2002) (Fig. 9), which were calculated in Xplor-NIH and visualized using PyMOL (http://www.pymol.org).

From the probability density maps, it is clear that the two domains of CaM-4Ca^{2+} sample a wide range of conformational space, but they preferentially sample conformations similar to the compact peptide-bound state. However, such states are not accessible to apo CaM, even in the presence of target peptide. Thus, loading with Ca^{2+} primes CaM to bind to its peptide targets in part by making the compact peptide-bound conformations accessible even in the absence of target peptide. However, because these closed states occupied by CaM-4Ca^{2+} are similar, but not identical, to the conformations found in the closed peptide-bound form, it is clear that CaM target binding cannot proceed through a purely conformational selection mechanism, especially considering that the peptide-binding faces of the two domains of CaM would be inaccessible to the peptide in such a fully closed conformation. Thus, CaM target binding involves an interplay between induced fit and conformational selection, a phenomenon observed by PRE in other systems as well (Clore, 2013; Tang *et al.* 2007).

We found that in addition to this intramolecular interaction between the two domains of CaM, the N- and C-terminal domains of CaM form an intermolecular interaction that is similar in nature

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

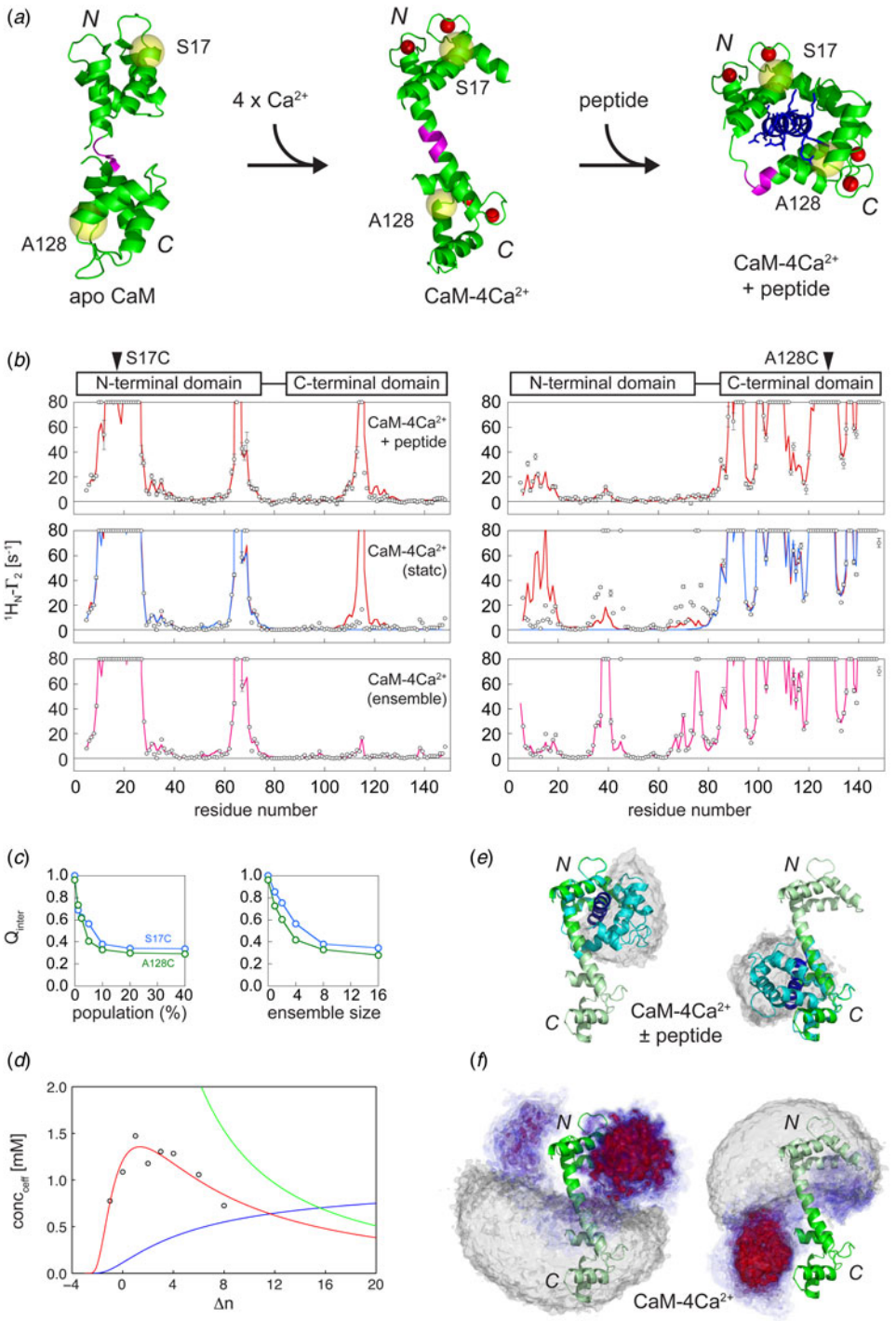


Fig. 9. Visualizing sparsely populated compact 'dark' states in Ca^{2+} activation of CaM. (a) Structures of CaM in various physiological states. Ca^{2+} ions are shown as red spheres, and spin-labeling sites S17C and A128C are indicated by yellow spheres. The flexible linker (residues 77–81) is shown in magenta. (b) Experimental PRE profiles for Ca^{2+}-loaded CaM (with a spin label at S17C or A128C), in the presence or absence of peptide, are shown as circles (error bar, 1 s.d.). PREs too large (>80 s^{-1}) to be accurately

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

to the intramolecular interaction (as judged by the similarity between the intramolecular and intermolecular PREs) (Anthis *et al.* 2011). The nature of the attraction between the two domains is electrostatic, as increasing the ionic strength of the solution decreases the magnitude of interdomain PREs (Anthis *et al.* 2011). Further studies demonstrated that the formation of this intramolecular interdomain interaction is dependent upon the length of the flexible linker between the two domains, and the extent of interaction is directly proportional to the effective concentration of the two linkers (Fig. 9d). Because the two domains are constrained to be close in space by the linker, they interact much more tightly than in the absence of a linker, and the extent of interdomain association decreases as the linker is lengthened and the two domains become less constrained in space. However, decreasing the length of the linker from the wild type length also reduces interdomain association, as the two domains become highly constrained and are not able to form their interaction due to steric factors. We were able to explain these results with a simple model of two points at the ends of a random-coil linker, and further studies on the affinity of CaM for a target peptide showed that affinity is affected by linker length in the same manner as interdomain association (Anthis & Clore, 2013). Thus, the ability of CaM–4Ca^{2+} to form these compact structures in the absence of target peptide is directly related to its biological activity.

4. Characterizing large transient assemblies by DEST and lifetime line broadening

4.1 Theory

Characterizing large macromolecular assemblies by NMR is limited by the fact that as molecular weight increases, the rate of tumbling decreases, and the rate of transverse relaxation and

measured are plotted at the top of the charts. PRE profiles back-calculated from the structures of CaM–4Ca^{2+}–peptide (red) and CaM–4Ca^{2+} (blue) are shown as solid lines. Only for the CaM–4Ca^{2+}–peptide complex are the experimental interdomain PREs correctly predicted by the corresponding known X-ray structure. The bottom panel shows the resulting fits for a minor state population of 10% represented by an eight-member ensemble (magenta lines). (c) Dependence of the interdomain PRE Q-factor for S17C (blue) and A128C (green) as a function of minor state population for CaM–4Ca^{2+}. (d) Modeling interdomain association in CaM–4Ca^{2+} (measured by PREs) as a local concentration effect. Empirical values of effective molarity (circles) were calculated from interdomain PRE data for CaM–4Ca^{2+} with different linker lengths (Anthis & Clore, 2013). These data were fit to an equation describing the theoretical values of the effective concentration of two points at the end of a random coil linker, where the interaction only occurs when the two points are separated by a defined distance. The best-fit line is shown in red (units of mM) and is a product of a pre-exponential factor (green) and exponential factor (blue). (e) The complex of CaM–4Ca^{2+} with an MLCK peptide is shown with CaM in cyan and the peptide in blue overlaid on the CaM–4Ca^{2+} dumbbell structure (green), best-fitted to either the N-terminal (left panel) or C-terminal (right panel) domains. The gray probability density maps represent 26 additional peptide-bound structures overlaid in the same manner. (f) Atomic probability density maps showing the conformational space sampled by the minor species ensemble for CaM–4Ca^{2+}. The minor state atomic probability maps are derived from PRE-driven simulated annealing calculations using an eight-member ensemble with a population of 10%, and plotted at multiple contour levels ranging from 0·1 (transparent blue) to 0·5 (opaque red). The gray probability density maps, plotted at a single contour level of 0·1 of maximum, show the conformational space consistent with interdomain PRE values $\leq 2\ s^{-1}$ and represent the major state ensemble characterized by no interdomain contacts and an occupancy of ~90%. In the left-hand panel, all ensemble members are best fitted to the N-terminal domain (dark green) and the probability density maps are shown for the C-terminal domain. In the right-hand panel, all ensembles are best-fitted to the C-terminal domain (dark green) and the probability density maps are shown for the N-terminal domain. Adapted from Anthis *et al.* (Anthis & Clore, 2013; Anthis *et al.* 2011).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

therefore linewidth increases. Above a certain threshold, resonance linewidths become too broad and fall below the noise of the spectrum, making detection difficult or impossible. Additional difficulties arise if the large assembly is transient (and thus sparsely populated under equilibrium conditions) and if it is heterogeneous (e.g. intermediates in protein misfolding and aggregation). However, under conditions where a large 'dark' state is in equilibrium with a smaller visible species, DEST and lifetime line broadening (R_{llb}) can be used to visualize states that are otherwise invisible to conventional structural biology (Fawzi *et al.* 2010b, 2011b, 2012; Libich *et al.* 2013). Whereas the characteristic that makes minor states visible by PRE is distance (i.e. close proximity between a spin label and an atom in the minor state), the property that makes dark states visible by DEST and R_{llb} is size (one of the very factors that makes such states invisible in the first place). Thus, DEST and R_{llb} are used to visualize molecules (i.e. small- to medium-sized proteins or peptides) in exchange with a much larger macromolecular assembly. As long as the visible species exchanges with the dark state on a timescale of ~500 μs–1 s, the footprint of the invisible species will be apparent as a broadening of the saturation profile and an increase in the apparent observed transverse relaxation rate (R_2^{obs}) of the major species. To date, these methods have been used to study the aggregation of the Aβ peptide into protofibrils (Fawzi *et al.* 2010b, 2011b, 2014) and its interaction with the chaperonin GroEL (Libich *et al.* 2013).

4.1.1 Lifetime line broadening

Here, we consider a process, such as protein aggregation or the interaction of a protein with a larger macromolecular complex, where the system is in exchange between two states:

$$\mathrm{A} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}^{\mathrm{app}}}{\rightleftharpoons}} \mathrm{B}, \tag{76}$$

where state A is the visible species and state B is the invisible dark state. Exchange from states A to B is described by the pseudo-first-order rate constant $k_{\mathrm{on}}^{\mathrm{app}}$, and exchange from states B to A is described by the first-order rate constant k_{off}. Thus, at equilibrium:

$$k_{\mathrm{on}}^{\mathrm{app}}[\mathrm{A}] = k_{\mathrm{off}}[\mathrm{B}]. \tag{77}$$

The rate constant $k_{\mathrm{on}}^{\mathrm{app}}$ is a pseudo-first-order rate constant because although it describes a reaction that is first-order with respect to A, the actual reaction could be of higher order.

In such an exchanging system, the observed transverse relaxation rate (R_2^{obs}) of state A can be enhanced by broadening due to differences in chemical shift between states A and B (R_{ex}) and by broadening due to differences in transverse relaxation rate between the two states (lifetime line broadening, R_{llb}). Overall, the enhancement ($\Delta R_2^{\mathrm{obs}}$) to the transverse relaxation rate of state A is given by:

$$R_2^{\mathrm{obs}} = R_{2,\mathrm{A}}^0 + \Delta R_2^{\mathrm{obs}}, \tag{78}$$

where $R_{2,\mathrm{A}}^0$ is the intrinsic relaxation of state A in the absence of exchange. The relative contributions of R_{ex} and R_{llb} to $\Delta R_2^{\mathrm{obs}}$ depend on the exchange rate and on the magnitude of chemical shift difference ($\Delta\omega = |\omega_{\mathrm{A}} - \omega_{\mathrm{B}}|$) and transverse relaxation rate difference ($\Delta R_2^0 = R_{2,\mathrm{B}}^0 - R_{2,\mathrm{A}}^0$) between the states (Fig. 5).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

If the rate of transverse relaxation in state B (the 'dark' state; $R^0_{2,\mathrm{B}} = R_2^{\mathrm{dark}}$) is very rapid (due a large size and therefore slow tumbling), then any transverse magnetization that is transferred from A to B will effectively be destroyed. If the exchange is slow on the transverse relaxation timescale ($k_{\mathrm{ex}} \ll \Delta R^0_2$), lifetime line broadening will be equal to the apparent first-order rate constant ($R_{\mathrm{llb}} = k_{\mathrm{on}}^{\mathrm{app}}$). If, however, exchange is fast on the transverse relaxation timescale ($k_{\mathrm{ex}} \gg \Delta R^0_2$), R_2^{obs} will be a population-weighted average of the two relaxation rates:

$$R_2^{\mathrm{obs}} = p_{\mathrm{A}} R^0_{2,\mathrm{A}} + p_{\mathrm{B}} R^0_{2,\mathrm{B}} \tag{79}$$

(providing $k_{\mathrm{ex}} \ll 1/\tau_{\mathrm{r}}$, where τ_{r} is the rotational correlation time) and R_{llb} is given by:

$$R_{\mathrm{llb}} = R_2^{\mathrm{obs}} - R^0_{2,\mathrm{A}} = p_{\mathrm{B}} \Delta R^0_2. \tag{80}$$

For cases where the dark state B is very large and has a very rapid rate of transverse relaxation (e.g. large macromolecular assemblies and aggregated species), lifetime line broadening can only be measured in the slow exchange regime, unless p_{B} is very small, because as R_2^{obs} approaches values of $\sim$100 s^{-1} or larger, it becomes difficult or impossible to measure. For smaller assemblies with more moderate $R^0_{2,\mathrm{B}}$, lifetime line broadening can be measured in the intermediate and fast exchange regimes as well.

If states A and B differ both in chemical shift and transverse relaxation rate, the manner by which R_{ex} and R_{llb} add up to $\Delta R_2^{\mathrm{obs}}$ depends on the value of the exchange rate (Fig. 5). For all exchange rates, the $\Delta R_2^{\mathrm{obs}}$ will be at least the magnitude of the greater of R_{ex} and R_{llb}, and $\Delta R_2^{\mathrm{obs}}$ will be at most the sum of the two. In the slow exchange regime ($k_{\mathrm{ex}} \ll \Delta R^0_2$ and $k_{\mathrm{ex}} \ll \Delta\omega$), $\Delta R_2^{\mathrm{obs}} = R_{\mathrm{ex}} = R_{\mathrm{llb}} = k_{\mathrm{on}}^{\mathrm{app}}$. In the fast exchange regime ($k_{\mathrm{ex}} \gg \Delta R^0_2$ and $k_{\mathrm{ex}} \gg \Delta\omega$), $\Delta R_2^{\mathrm{obs}} = R_{\mathrm{ex}} + R_{\mathrm{llb}}$ (and in the very fast regime, $\Delta R_2^{\mathrm{obs}} = R_{\mathrm{llb}}$, as R_{ex} approaches zero). In the intermediate exchange regime, $\Delta R_2^{\mathrm{obs}}$ will occupy a value intermediate between these two extremes (Fig. 5).

Values of $\Delta R_2^{\mathrm{obs}}$, R_{ex}, and R_{llb} can be calculated by solving the McConnell equations (Eqs. (16)–(20)) and fitting the resulting time series to a single-exponential decay to derive a single apparent transverse relaxation rate (R_2^{obs}). Subtracting $R^0_{2,\mathrm{A}}$ from R_2^{obs} yields $\Delta R_2^{\mathrm{obs}}$. The same procedure can be used to calculate R_{ex} and R_{llb} by setting $\Delta R^0_2 = 0$ (to calculate R_{ex}) and by setting $\Delta\omega = 0$ (to calculate R_{llb}). Because R_{llb} does not depend on chemical shift differences, it can be calculated by a simplified procedure only requiring the McConnell equations to be solved once. If all chemical shifts are set to zero (or if the $-i\boldsymbol{\Omega}$ matrix in Eq. (16) is omitted), R_{llb} can be calculated by the following expression:

$$R_{\mathrm{llb}} = -\frac{1}{T_R} \ln\left(\frac{M_{x,\mathrm{A}}(T_{\mathrm{R}})}{M_{x,\mathrm{A}}(0)}\right) - R^0_{2,\mathrm{A}}, \tag{81}$$

where $M_{x,\mathrm{A}}(0)$ is the initial value of x transverse magnetization on state A and $M_{x,\mathrm{A}}(T_{\mathrm{R}})$ is the magnetization at time T_{R} (the specific value of T_{R} is unimportant, but it can be set to an experimentally realistic value such as 0·1 s).

4.1.2 DEST

If a weak CW RF field is applied to a rapidly tumbling macromolecule, NMR signals at the same frequency as the CW RF radiation will be saturated, signals sufficiently far off resonance will be

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

unaffected, and those near the RF frequency will be partially saturated. The width of this saturation profile increases with the transverse relaxation rate (R_2) of the nucleus and the strength of the CW RF field. For proteins or peptides that tumble rapidly and exhibit slow R_2 rates – and are thus visible by NMR – this profile is narrow. Large macromolecular complexes and assemblies (≥1 MDa), on the other hand, tumble slowly and exhibit large R_2 rates that make direct observation by solution NMR difficult or impossible due to the dramatically broadened linewidths; these large systems exhibit much broader saturation profiles. However, if such a dark state is in exchange with a visible species, it may be possible to indirectly study it via DEST, even if the dark state is also sparsely populated and/or has an amorphous or heterogeneous structure.

For a visible species in exchange with an invisible species, the overall R_2 rate for a nucleus signal will have contributions from both states. For the type of very large species that one might study by DEST (e.g. MW > 1 MDa), the transverse relaxation rate can become incredibly fast (e.g. >1000 s^{-1}). Thus, transverse magnetization in the invisible species will quickly disappear, whereas more slowly relaxing longitudinal magnetization will remain intact. Thus, by just partially tilting the bulk magnetization of the sample toward the transverse plane (i.e. only partially saturating the magnetization) by the application of an off-resonance CW RF field, the overall signal will be attenuated but not destroyed. Under these conditions, nuclei in the invisible state will relax more quickly than nuclei in the visible state, but their magnetization will not disappear completely. Thus, due to the large difference in R_2 between the visible and invisible, the otherwise invisible large species can be detected by DEST (Fawzi *et al.* 2012).

The data from a DEST experiment can be modeled using a homogeneous form of the McConnell equations describing a system in equilibrium (Helgstrand *et al.* 2000). If the spins can be treated as isolated spins free of spin diffusion (e.g. ^{15}N), the time dependence of the magnetization in this two-site exchanging system is given by (Helgstrand *et al.* 2000):

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M} = -(\mathbf{E} + \mathbf{R} + \mathbf{K} + \mathbf{\Omega} + \mathbf{S})\mathbf{M}, \tag{82}$$

where **M** is the vector describing the transverse magnetization of each state, **R** is the matrix describing the transverse relaxation rates of each state, **K** is the matrix describing the kinetic rates of exchange between each state, **Ω** is the matrix describing the chemical shift offset of the saturation pulse relative to the chemical shift of each state, **S** is the matrix describing the field strength of the saturation pulse, and **E** is the matrix describing the equilibrium magnetization. The exact equations will be dictated by the kinetic model describing the system, and here we consider a system in exchange between two states (Eq. (76)), where state A is the visible species and state B is the invisible dark state:

$$\mathbf{M} = \begin{bmatrix} E/2 \\ M_{x,\mathrm{A}} \\ M_{y,\mathrm{A}} \\ M_{z,\mathrm{A}} \\ M_{x,\mathrm{B}} \\ M_{y,\mathrm{B}} \\ M_{z,\mathrm{B}} \end{bmatrix}, \tag{83}$$

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

$$\mathbf{E}+\mathbf{R}+\mathbf{K}+\Omega+\mathbf{S}$$
$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R^0_{2,\mathrm{A}}+k^{\mathrm{app}}_{\mathrm{on}} & \Omega_{\mathrm{A}} & -\omega_y & -k_{off} & 0 & 0 \\ 0 & -\Omega_{\mathrm{A}} & R^0_{2,\mathrm{A}}+k^{\mathrm{app}}_{\mathrm{on}} & \omega_x & 0 & -k_{off} & 0 \\ -2R_{1,\mathrm{A}}M_{\mathrm{eq,A}} & \omega_y & -\omega_x & R_{1,\mathrm{A}}+k^{\mathrm{app}}_{\mathrm{on}} & 0 & 0 & -k_{off} \\ 0 & -k^{\mathrm{app}}_{\mathrm{on}} & 0 & 0 & R^0_{2,\mathrm{B}}+k_{off} & \Omega_{\mathrm{B}} & -\omega_y \\ 0 & 0 & -k^{\mathrm{app}}_{\mathrm{on}} & 0 & -\Omega_{\mathrm{B}} & R^0_{2,\mathrm{B}}+k_{off} & \omega_x \\ -2R_{1,\mathrm{B}}M_{\mathrm{eq,B}} & 0 & 0 & -k^{\mathrm{app}}_{\mathrm{on}} & \omega_y & -\omega_x & R_{1,\mathrm{B}}+k_{off} \end{bmatrix}, \tag{84}$$

where E is unity, M is the magnetization of the A or B spin along each axis, and M_{eq} is the equilibrium magnetization of each state. The longitudinal relaxation rates are given by R_1, and the intrinsic transverse relaxation rates (in the absence of exchange) are given by R^0_2. ω is the strength of the CW RF B_1 saturation field applied to each axis (ω [rad/s] = $2\pi \times \nu$ [Hz]), and Ω_{A} and Ω_{B} are the frequency offsets between the resonance frequency of the nucleus and the carrier frequency (in rad s^{-1}). Equations describing three-state exchange (Fawzi *et al.* 2011b) and two-state-exchange involving a multi-spin system (Fawzi *et al.* 2010b) have been described elsewhere.

A DEST (or CEST) profile can be calculated by solving Eq. (82) for a series of CW RF frequency offsets. For each offset, the signal intensity (I_{sat}) is given by solving these equations for the time that the CW RF (ω_x) field is applied. The DEST profile is plotted as $I_{\mathrm{sat}}/I_{\mathrm{ref}}$ *versus* carrier offset, where, I_{ref} is the intensity in the absence of the saturation CW RF field, calculated as for I but with $\omega_x = 0$. Calculations are performed by starting with magnetization on the z-axis ($M_{z,\mathrm{A}}[t=0] = 1 \times p_{\mathrm{A}}$; all other components equal zero, including $M_{z,\mathrm{B}}[t=0]$, since the dark state magnetization would not survive the INEPT transfer if $R^0_{2,\mathrm{B}}$ is very large). Longitudinal magnetization decays to the equilibrium values of $M_{\mathrm{eq,A}}$ and $M_{\mathrm{eq,B}}$. For ^{15}N DEST, a value of $M_{\mathrm{eq,A}} = M_{z,\mathrm{A}}[t=0]/10$ can be used, because $\gamma_{\mathrm{H}}/\gamma_{\mathrm{N}} \approx 10$. A calculation is performed with magnetization beginning on $+z$, then a calculation is performed with magnetization beginning on $-z$ (i.e. $M_{z,\mathrm{A}}[t=0] = 1 \times p_{\mathrm{A}}$, then $M_{z,\mathrm{A}}[t=0] = -1 \times p_{\mathrm{A}}$). The values from the second iteration (beginning on $-z$) are subtracted from the first (beginning on $+z$) because this difference yields a single-exponential curve that decays to zero in the absence of a saturation field). To properly simulate experimental data, the inhomogeneity of the B_1 saturation field should also be taken into account, which can be accomplished by repeating the calculations with a series of slightly different values of ω_x (e.g. a Gaussian distribution about the desired value of ω_x if the standard deviation of the field strength is known (Vallurupalli *et al.* 2012) or just ~10 fields strengths evenly spaced over a range of $\omega_x \pm 10\%$ (Fawzi *et al.* 2014)).

A DEST (or CEST) profile can also be approximated by treating it as an off-resonance $R_{1\rho}$ experiment (Palmer, 2014):

$$I_{\mathrm{sat}}/I_{\mathrm{ref}} = \cos^2\theta \exp(-T_{\mathrm{sat}}R_{1\rho}), \tag{85}$$

where T_{sat} is the total saturation time, the angle θ is given by Eq. (34), and $R_{1\rho}$ is given by Eq. (40) (as long as $p_{\mathrm{A}} \gg p_{\mathrm{B}}$) for a DEST experiment (where $R^0_{2,\mathrm{A}}$ and $R^0_{2,\mathrm{B}}$ differ) or by simpler expressions (e.g. Eq. (39)) for CEST (as long as $R^0_{2,\mathrm{A}} = R^0_{2,\mathrm{B}}$). Palmer has demonstrated good agreement between this analytical approximation, especially when Eq. (40) is used, compared to numerical solutions of the McConnell equations (Palmer, 2014).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

4.2 Data acquisition

A system undergoing slow exchange between an NMR-visible species and a slowly tumbling invisible dark state can be well characterized by employing both lifetime line broadening measurements and DEST. These measurements have been performed on backbone amide ^{15}N nuclei (Conicella & Fawzi, 2014; Fawzi *et al.* 2011b, 2012; Libich *et al.* 2013), methyl ^{13}C nuclei (Fawzi *et al.* 2014) and $^{1}H_N$ nuclei (Fawzi *et al.* 2010b) (although the latter lack single-residue resolution due to rapid ^{1}H–^{1}H spin diffusion in the dark state). Measurements of both R_{llb} and DEST for the same nucleus allow simultaneous fitting, which is necessary for precisely fitting the various parameters. For systems in the intermediate exchange regime, additional observables, including CPMG relaxation dispersion and exchange-induced chemical shifts, can also be useful (Libich *et al.* 2013).

4.2.1 Measuring lifetime line broadening

Measuring lifetime line broadening requires being able to measure the R_2 values for the major state alone and in the presence of the dark state. To characterize the exchange of Aβ monomers with protofibrils, this was accomplished by acquiring data at both low concentrations of Aβ (no dark state present) and high concentrations (dark protofibrils present; Fawzi *et al.* 2010b, 2011b, 2014). To characterize the interaction of Aβ with GroEL, data were measured in the absence and presence of GroEL (Libich *et al.* 2013). Transverse relaxation can be measured using conventional pulse sequences (Cavanagh *et al.* 2007) under conditions that eliminate any R_{ex} contributions, using either a CPMG experiment with a large value of ν_{CP} or a $R_{1\rho}$ experiment with high spin-lock field strength. Under these conditions:

$$R_{\mathrm{llb}} = \Delta R_2^{\mathrm{obs}} = R_2^{\mathrm{vis+dark}} - R_2^{\mathrm{vis}}, \tag{86}$$

where R_2^{vis} is R_2^{obs} in the absence of the dark state and $R_2^{\mathrm{vis+dark}}$ is in the presence of the dark state. R_2 is measured by varying, in an interleaved manner, the length of a delay during which the magnetization of the observed nucleus is on the *xy*-plane (generally at least ~8 different times, ranging from $0\times$ to $\sim 1{\cdot}3 \times T_2$), and fitting the intensity of each peak to an exponential decay function. Because it can be more straightforward to measure R_{llb}, it is usually advisable to attempt to measure R_{llb} before attempting to make more complex DEST measurements. Detection of lifetime broadening of $R_{\mathrm{llb}} \geq 1\ \mathrm{s}^{-1}$ is a strong indicator that DEST measurements will also be successful.

4.2.2 Acquiring DEST data

Most DEST experiments to date have been performed on ^{15}N, which is advantageous because it provides one probe per residue, and it is not subject to spin diffusion – which complicates analysis of ^{1}H DEST measurements – or spin–spin coupling – which complicates ^{13}C measurements in universally labeled samples. The setup and analysis of a DEST experiment, which has recently been described in greater detail (Fawzi *et al.* 2012), will be summarized here.

^{15}N DEST can be measured with a 2D HSQC-based experiment using the pulse sequence in Fig. 10. Pulse sequences for measuring DEST are available online at http://spin.niddk.nih.gov/clore/Software/software.html. A series of interleaved 2D experiments are performed, varying the frequency of the ^{15}N CW saturation field. The optimal duration of the CW period is $T_1 = 1/R_1$. Experiments are performed at multiple saturation field strengths (e.g. 250 and 500 Hz) to assist in

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

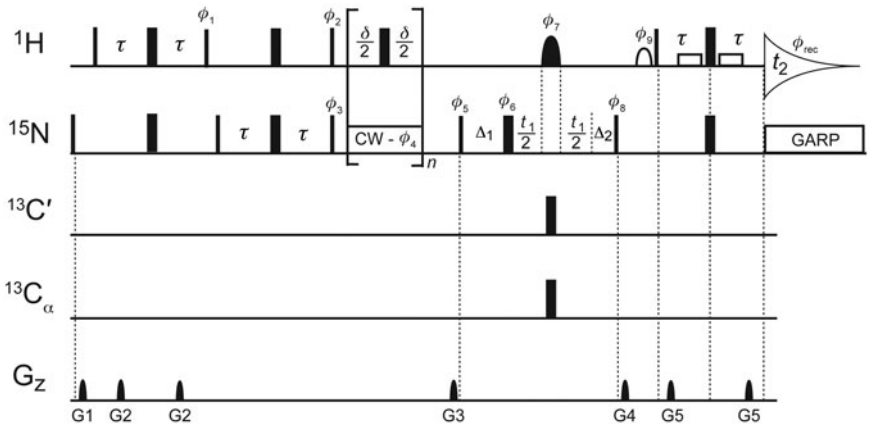


Fig. 10. Pulse sequence for the 2D ^{15}N–DEST experiment. Full details are given in Fawzi *et al.* (Fawzi *et al.* 2011b). During saturation experiments, the ^{15}N CW saturation field is applied at a range of offset frequencies and at two different RF field strengths. During this period, a 180° ^{1}H pulse train is applied (with 100 ms spacing between pulses) in order to eliminate longitudinal cross-correlated relaxation interference between the ^{15}N–^{1}H dipolar interaction and the ^{15}N chemical shift anisotropy. Thus, for a saturation time of 0·9 s, for example, the saturation block would be repeated a total of $n = 9$ times. The optimal time for application of ^{15}N saturation is the ^{15}N–T_1 longitudinal relaxation time. In the reference experiment, the ^{15}N saturation pulse is turned off (but the 180° ^{1}H pulse train is still applied).

optimizing parameter values and discriminating between models. A reference experiment also needs to be performed with the saturation field turned off (i.e. strength of 0 Hz). The intensity of each peak is measured for each of these 2D experiments and plotted against the frequency offset of the CW saturation radiation. The spectrometer time required to acquire DEST data varies, but about 1 week was required to acquire the complete set of DEST and R_2 data on ∼150 μM ^{15}N-labeled Aβ using a cryogenically cooled probe (including the measurement of DEST at two different saturation field strengths; Fawzi *et al.* 2011b).

In order for DEST to be measured, the visible species, which gives rise to a 2D ^{1}H–^{15}N spectrum with well-resolved peaks and high S/N ratio, must be in exchange with a high-molecular-weight species (i.e. the 'dark state'). Additionally, both k_{on}^{app} and the rate constant for the transition of the dark state to the light state (k_{off}) must be of the order of, or faster than, the R_1 rate of the nucleus undergoing saturation (e.g. ≥1 s^{-1} for backbone ^{15}N). The lower limit on k_{off} arises because longer residency periods in the dark state would cause too great a loss of signal before a significant number of transitions from the dark to visible state could be measured. The reason for the lower limit on k_{on}^{app} is that, due to the extremely fast R_2 relaxation rate in the dark state, the creation of ^{15}N$_{\tilde{x}}$ magnetization must take place in the light state due to the inefficiency of the INEPT magnetization transfer in the dark state. Thus, it is also necessary to have a significant number of transitions from the light to the dark state within a time of $T_1 = 1/R_1$. Under conditions where $R_2^{dark} \gg k_{off}$, there is also an upper limit of $k_{on}^{app} \leq 100$ s^{-1} as larger values will cause too great a loss of NMR signal and make measurement of DEST or R_2 difficult or impossible. Additionally, a minimum population of the dark state (∼0·5% or higher) is required; leading to a rough lower limit of $k_{on}^{app}/k_{off} > 0{\cdot}005$.

4.3 Data analysis

DEST and $R_{1\rho}$ data are fit simultaneously, as they offer complementary information on the same type of exchange processes. DEST data are analyzed by fitting the DEST profiles (intensity *versus*

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

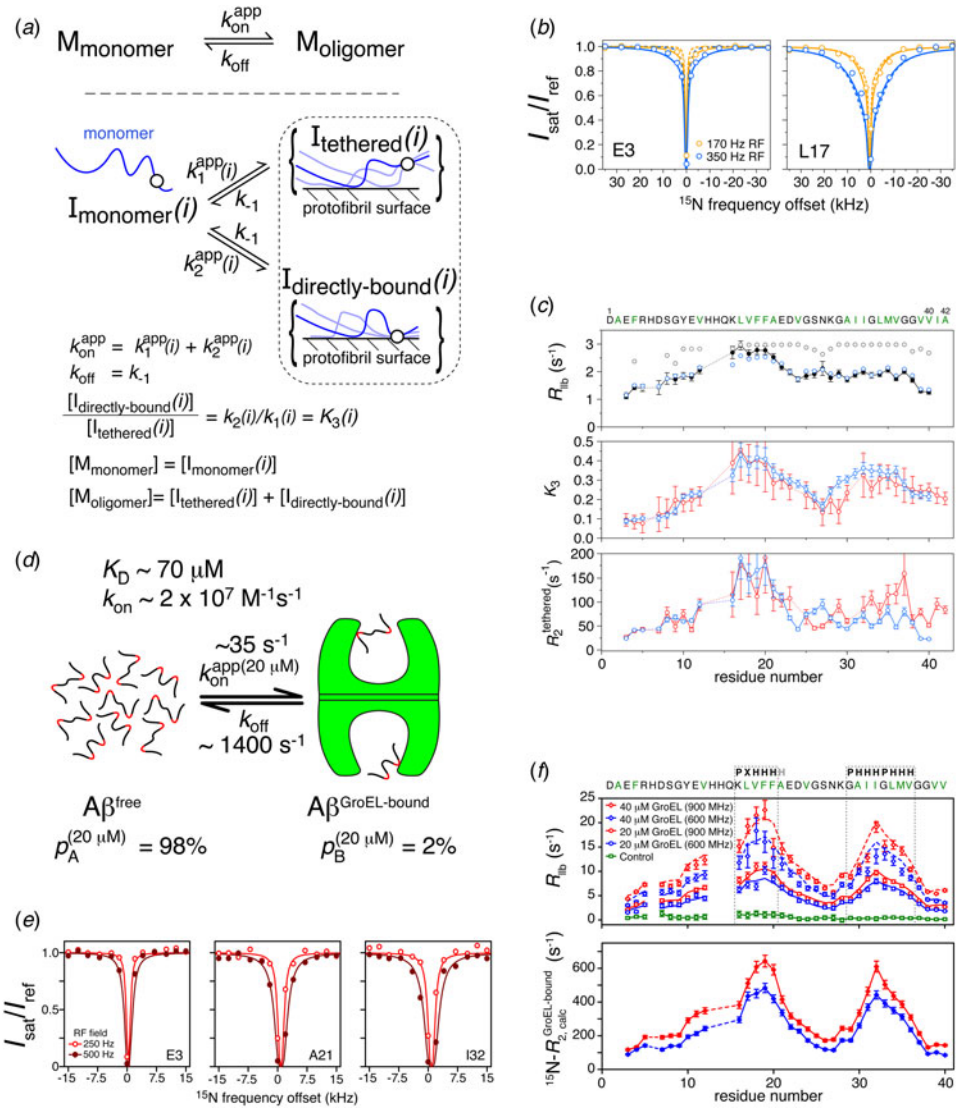


Fig. 11. Characterizing the exchange of Aβ peptides on the surface of large macromolecular assemblies by ^{15}N–DEST and ^{15}N lifetime line broadening ($R_{1\rho}$). (a) Pseudo-two-state kinetic scheme for exchange of Aβ monomers with the surface of amyloid protofibrils. The 'dark' state comprises an ensemble of states where each residue can either be tethered or in direct contact with the surface of the oligomer. (b) ^{15}N–DEST profiles for Glu3 and Leu17 at two different RF fields (170 Hz, orange; 350 Hz, blue). Experimental data are shown as circles, and lines represent the best fit for the simple two-state model (dashed) and the tethered/direct contact model (solid). (c) Residue-by-residue characterization of Aβ in exchange with the protofibril surface. The top panel shows ^{15}N–$R_{1\rho}$, with experimental data shown as black filled-in circles and calculated data shown as open circles for the simple two-state model (gray) or the tethered/direct contact model (blue). Also shown are profiles for the residue-specific partition coefficient K_3 (given by the ratio of direct-contact to tethered states) and ^{15}N–R_2 values for the dark tethered states derived from the fits to the experimental $R_{1\rho}$ and DEST data. Data are shown for Aβ(1–40) (blue) and Aβ(1–42) (red). (d) Kinetic scheme for binding of Aβ(1–40) to GroEL. Values are listed for the populations of free (p_A) and GroEL-bound (p_B) Aβ in the presence of 20 μM (in subunits) GroEL. (e) ^{15}N–DEST profiles for Glu3, Ala21, and Ile32 at two different RF fields (250 Hz, light; 500 Hz, dark). Experimental data are shown as circles, and lines represent the best fit from the simple two-state

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

CW RF offset) to a kinetic model based on the McConnell equations (McConnell, 1958). The simplest model for fitting DEST data is a two-state model of exchange between a visible light state and an invisible dark state (Eqs. (82)–(84)). For a two-state model, the adjustable parameters are the global on- and off-rates (from which the relative visible and dark state populations are calculated), and R_2^{dark} values (one for each observed nucleus). In addition to the DEST profiles and R_{1lb} values, R_2^{visible} values need to be provided.

DEST and R_{1lb} data can be fit using the program DESTfit (Fawzi *et al.* 2012), which is available online at http://spin.niddk.nih.gov/clore/Software/software.html and has been described in detail previously (Fawzi *et al.* 2012). Once DEST and R_{1lb} data have been acquired and processed, the first step is to select the model to be used to fit the data. One should begin with a two-state model, only attempting fits with more complicated models if necessary. For a two-state model, the output includes the DEST profiles (one per nucleus) and plots of R_{1lb} *versus* residue and R_2^{dark} *versus* residue; both experimental and back-calculated data are plotted together to help qualitatively judge the quality of the fit (Fig. 11) (Fawzi *et al.* 2012). By fitting DEST data to a kinetic model, one is able to gain great insight into the system under study, including the number of states and exchange events involved, and the kinetic parameters describing their exchange. Additionally, the fitted parameters of the dark state R_2 values provide a direct estimate of the size of the dark state as well as serve to locate the sites of interaction. Thus, DEST provides unique insight into systems where a visible state is in exchange with a larger invisible state on an intermediate-to-slow timescale.

4.4 Combined approaches involving DEST

The information gained from DEST is complementary to other techniques used to characterize dark states on the same or different timescales (Fig. 1). Acquiring additional data using such complementary methods can help refine the values of fitting parameters and help characterize concurrent exchange events taking place on a timescale close to the faster exchange limits of DEST. For processes on the intermediate exchange regime, the measurement of relaxation dispersion and exchange-induced chemical shifts (ω_{ex}) can be very helpful (Libich *et al.* 2013). Although DEST and lifetime line broadening are optimal for describing exchange events on a timescale of ~1 ms–1 s, combining these observables with CPMG relaxation dispersion and exchange-induced chemical shifts enables study of exchange events of the order of 500 μs or faster.

ω_{ex} is the observed difference in chemical shift of a nucleus in the presence of the dark state *versus* in the visible state alone (Eqs. (13) and (15)), and it is greater for larger differences in chemical shift between the two states ($\Delta\omega$) and for faster exchange rates (k_{ex}). Extreme care must be taken when measuring ω_{ex} because the values for ω_{ex} in these types of systems can be very small (<1 Hz) (Libich *et al.* 2013). Both ω_{ex} (which is measured directly) and R_{ex} (which can be measured indirectly by CPMG relaxation dispersion) arise from chemical shift differences between

model. (f) Residue-by-residue characterization of Aβ in exchange with GroEL. The top panel shows ^{15}N–R_{1lb} for 50 μM Aβ in the presence of 20 (circles) and 40 (diamonds) μM (in subunits) GroEL at spectrometer frequencies of 600 (blue) and 900 (red) MHz. Data are also shown for a control sample at 600 MHz (green) containing 50 μM Aβ40, 2·9 μM acid-denatured Rubisco, and 20 μM GroEL. The bottom panel shows back-calculated ^{15}N–R_{1lb} values for GroEL-bound Aβ at 600 (blue) and 900 (red) MHz spectrometer frequencies. Adapted from Fawzi *et al.* (2011b) and Libich *et al.* (2013).

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

the visible and dark states, and they are most useful as probes of dynamics near $k_{ex} \approx \Delta\omega$. R_{ex} has a maximum value at $k_{ex} = \Delta\omega$. ω_{ex} is most sensitive to k_{ex} near $k_{ex} \approx \Delta\omega$, and it varies in value from zero in the slow exchange regime ($k_{ex} \ll \Delta\omega$) up to the population-weighted average chemical shift of the different states in the fast exchange regime ($k_{ex} \gg \Delta\omega$). Thus, CPMG relaxation dispersion and ω_{ex} provide useful insight into exchange over timescales of ~500 μs–5 ms. A modification of the DESTfit program allows simultaneous fitting of CPMG relaxation dispersion and exchange-induced chemical shifts in addition to DEST and $R_{1\rho}$ data (Libich *et al.* 2013).

4.5 DEST in practice: application to Aβ

A hallmark of Alzheimer's disease is the formation of extracellular plaques consisting of the Aβ peptide in fibrillar amyloid cross-β-strand structures (Goedert & Spillantini, 2006). However, smaller, more transient oligomers and protofibrillar forms of Aβ have more recently been linked with the presence of Alzheimer's disease, and intermediates in the Aβ aggregation pathway have been hypothesized to constitute the actual toxic species in Alzheimer's disease (Walsh & Selkoe, 2007). Despite much interest, little is known about the structures of these transient intermediates, how they form, and their role in Alzheimer's disease. These protofibrils are much too large (1·8–85 MDa) (Fawzi *et al.* 2011b) to be directly characterized by conventional NMR and are not amenable to study by X-ray crystallography. Lower resolution methods, such as EM and AFM, offer some insight into their size and global structure, but provide little insight into their dynamic nature (Fawzi *et al.* 2011b). DEST and $R_{1\rho}$, however, have recently helped shed light on the details of these protofibrillar forms of Aβ (Fawzi *et al.* 2010b, 2011b).

In the course of conducting NMR studies on transient interactions of Aβ, concentration-dependent ^{1}H–R_2 values were observed that were largely uniform across the peptide. It was later shown that this was due to $R_{1\rho}$ – a loss of signal due to the exchange of the visible monomeric Aβ peptide with an invisible dark state on a slow ($k_{ex} \ll R_2^{dark} - R_2^{visible}$) timescale (Fawzi *et al.* 2010b). Fawzi *et al.* (2010b) then showed that by applying ^{1}H off-resonance saturation to Aβ displaying $R_{1\rho}$, the broad saturation profile of the dark state became apparent in the overall saturation transfer profile of the visible species. This indicated that Aβ exchanges with a large species on a slow timescale, offering strong evidence that Aβ forms transient oligomeric intermediates (Fawzi *et al.* 2010b). These findings were further backed up by the visualization of such species by EM (Fawzi *et al.* 2010b). Although these initial studies were highly illuminating, quantitative analysis of ^{1}H saturation transfer was difficult due to spin diffusion, limiting the conclusions to broader insights, and making atomic-resolution insights not possible. A major advance, however, was the development of ^{15}N DEST, which removed the problem of spin diffusion, allowing atomic-resolution (or at least residue-resolution) insights to be gained. Using DEST, the authors were able to measure the kinetics of protofibril formation and show that two hydrophobic segments of the peptide (residues ~16–25 and ~28–37) were more tightly associated with the protofibrils (Fig. 11), and hence likely play a more central role in protofibril formation (Fawzi *et al.* 2011b).

Lifetime broadening and DEST data were acquired on samples of Aβ(1–40) and Aβ(1–42) over a range of concentrations. The data could not be adequately fit with a simple two-state model, so more complex models had to be incorporated. The best fit was given by a pseudo-two-state model, whereby the Aβ peptide is in a two-state exchange between a visible state and a dark state, but in the dark state, each residue exists in either a direct contact or tethered state (i.e. two-site global exchange for the Aβ peptide, but three-site exchange on the level of

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

each individual residue) (Fawzi *et al.* 2011b). Using this kinetic model, the data were fit to a global overall rate of association with the protofibril surface (in either the tethered or direct-contact form), giving a concentration-dependent pseudo-first-order rate constant of $k_{on}^{app} = 3{\cdot}0 \pm 0{\cdot}7$ s^{-1} (for measurements on a sample comprising a total concentration of 270 μM Aβ(1–40) of which ~50% is in the protofibrillar form), consistent with the maximum R_{1lb} of 3·0 s^{-1} observed under these conditions. These data were also fit to global parameters for the dissociation rate constant from the dark state ($k_{off} = 51 \pm 12\ s^{-1}$) and the ^{15}N-$R_2$ rate in the direct-contact dark state (^{15}N-$R_2^{contact} = 18\,800 \pm 700\ s^{-1}$) (Fawzi *et al.* 2011b). Residue-specific parameters for the ^{15}N-R_2 rate in the tethered state ($R_2^{tethered}$) and the partition coefficient between the direct-contact and tethered states (K_3) were also fit. A larger value of K_3 indicates that the residue spends a higher fraction of time in the dark state in direct contact with the protofibrils surface rather than more loosely tethered (via other residues in direct contact with the protofibril surface), and a larger value of $R_2^{tethered}$ indicates a reduced average distance from the tethered residue to the site of direct contact. As a whole, these results provide an indirect picture of the otherwise invisible protofibrils of Aβ (Fawzi *et al.* 2011b), showing that insight can be gained by solution NMR even on a system that is so large that it approaches the Redfield limit (Redfield, 1957) where $T_2 \gg \tau_r$ no longer holds true (for Aβ 1–40 protofibrils (Fawzi *et al.* 2011b), ^{15}N-$T_2^{contact} = 1/^{15}$N-$R_2^{contact} = 53\ \mu s$; $\tau_r \approx 40\ \mu s$ for a spherical 85-MDa protein assembly at 10°C). Studies on Aβ(1–40), (1–42), and (1–43) have revealed broadly similar results, although faster association rates are observed for the longer peptides, consistent with their more rapid overall aggregation kinetics (Conicella & Fawzi, 2014; Fawzi *et al.* 2011b). Additional lifetime line broadening and DEST studies were performed on the methyl ^{13}C resonances of Aβ (1–40), revealing a strong correlation between backbone ^{15}N-R_{1lb} and methyl ^{13}C R_{1lb} (Fawzi *et al.* 2014). These data were best fit to a pseudo-two-state model with the same kinetic parameters as the ^{15}N data. The best-fit value of $1/^{13}$C-$R_2^{contact} = 1230 \pm 30\ s^{-1}$ was significantly slower than ^{15}N–$R_2^{contact}$, consistent with the rapid rotation of methyl groups.

In another study, the authors characterized the transient interaction of the Aβ peptide with the chaperonin GroEL (Libich *et al.* 2013). The chaperonins constitute a subclass of chaperone proteins, which facilitate protein folding *in vivo* (Horwich & Fenton, 2009; Thirumalai & Lorimer, 2001). GroEL comprises two seven-member rings of 56-kDa subunits; substrates are known to bind to the inner cavity of GroEL, although the structures and dynamics of these complexes are largely unknown, due to the invisibility of the GroEL-associated substrate to conventional structural biology techniques. Although GroEL (780 kDa) is much smaller than the Aβ protofibrils previously characterized by DEST (up to 85 MDa), it is still large and slowly tumbling enough ($R_2 > 1000\ s^{-1}$) that the GroEL-associated Aβ is invisible to direct observation by NMR but also susceptible to off-resonance saturation at offsets greater than 1 kHz that could be transferred to the visible monomeric species (Libich *et al.* 2013) (Fig. 11).

Experiments were performed on dilute (50 μM) Aβ in the absence or presence of GroEL at concentrations of 20 and 40 μM (in individual subunits). The addition of GroEL caused a large increase in ^{15}N-R_2, which varied on a residue-by-residue basis from 1 to 23 s^{-1}. This ΔR_2^{obs} could be attributed almost fully to lifetime line broadening (*versus* chemical exchange broadening) because ΔR_2^{obs} did not correlate with exchange-induced chemical shifts. DEST experiments were performed, and saturation transfer was observed in Aβ in the presence of GroEL (50% saturation at up to ~2 kHz off resonance for a saturation field strength of 500 Hz), although this effect was much smaller than seen with concentrated Aβ in exchange between monomeric and protofibrillar Aβ (Fawzi *et al.* 2011b). CPMG relaxation dispersion experiments were also

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

performed on Aβ, and dispersion was seen in the presence of GroEL, but not in its absence. These data were fit together (along with lifetime broadening data and exchange-induced chemical shifts) to a two-state exchange model characterizing the interaction of Aβ with GroEL, revealing both the kinetics of this process and the sites in Aβ most tightly associated with GroEL (Libich *et al.* 2013). Based on these fits, in the presence of 20 μM GroEL (in subunits), 2·3% of Aβ peptides were bound to GroEL, and it associated with a pseudo-first-order rate constant of $k_{on}^{app} = 35\ s^{-1}$ and dissociated with a first-order rate constant of $k_{off} = 1440\ s^{-1}$ (Libich *et al.* 2013). The degree of direct contact with GroEL on a residue-by-residue basis could be determined by the fitted R_2^{dark} for each residue. The tightest association was observed for Aβ residues 16–21 and 30–36 (Libich *et al.* 2013; Fig. 11), which are largely hydrophobic and correspond to consensus GroEL-binding sequences (Stan *et al.* 2006). Additionally, ^{15}N, $^{1}H_N$, $^{13}C_\alpha$, and $^{13}C_\beta$ chemical shifts of GroEL-bound Aβ were obtained through the fitting procedure, offering direct structural insight into the otherwise invisible dark bound state. Interestingly, the small magnitude of these chemical shifts (<1 ppm) and the fact that for all residues $R_2^{dark} \ll 1000\ s^{-1}$ (the value expected for a rigid ^{1}H-^{15}N bond in an assembly the size of GroEL), indicate that in the dark state the Aβ peptide remains predominately only partially associated with GroEL (i.e. only binding with one hydrophobic sequence at a time; Libich *et al.* 2013).

5. Conclusion

Intermediate states in a range of biological processes, including protein folding, protein–protein interactions, enzyme catalysis, and protein aggregation, have defied high-resolution characterization. Various factors contribute to the invisibility of these states, including their transient nature, low population at equilibrium, and – in some cases, including protein aggregation – heterogeneity and large molecular weight, characteristics which are anathema to high-resolution structural biology by NMR, X-ray crystallography, or other methods. However, over recent years, developments in NMR methodology have opened these previously inaccessible states up to high-resolution characterization. These NMR methods are applicable over a wide range of timescales (Fig. 1) from the very fast (ps, ns, and μs) processes characterized by PRE (Clore & Iwahara, 2009) and PCS (Bertini *et al.* 2010) (e.g. protein–protein encounter complexes), to the intermediate (μs–ms) processes characterized by $R_{1\rho}$ (Palmer & Massi, 2006) and CPMG relaxation dispersion (Korzhnev & Kay, 2008; Palmer *et al.* 2001) (e.g. protein folding and enzyme catalysis), intermediate-slow (μs–s) processes characterized by DEST (Fawzi *et al.* 2012) and lifetime line broadening (Fawzi *et al.* 2010b) (e.g. protein association aggregation), slow (ms) processes characterized by CEST (Vallurupalli *et al.* 2012) (protein folding and association), to the very slow (ms, s, and longer) processes characterized by hydrogen exchange (Englander *et al.* 2007) (e.g. protein folding and aggregation).

The key to these methods is that they specifically and disproportionately amplify some aspect of the invisible state. Other methods, either using NMR (e.g. RDCs) or other techniques (e.g. SAXS/SANS) can provide information about minor states as part of the overall ensemble of structures being characterized. However, these average methods only report upon each structure proportionally to its population, limiting their utility for sparsely populated states. As long as the dark state is in chemical exchange with a visible species, though, and if one can take advantage of some unique aspect of the dark state – e.g. a shorter distance in a transient state than in the visible species (PRE) or a much larger molecular weight in the dark state than in the visible state (DEST or lifetime broadening) – it is possible to amplify the information from the dark state and make it

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

visible. Underlying these methods is the enhancement of relaxation rates due to some aspect of the exchange with the invisible species, allowing the information from the invisible species to be viewed through the lens of the visible species.

An advantage of NMR is that from a single experiment, a great deal of information at the atomic or single-residue level can be acquired on an otherwise invisible state. Other methods for measuring distances, for example, such as FRET (Hillisch *et al.* 2001) and DEER (Altenbach *et al.* 2008; Klug & Feix, 2008) also magnify the effects of minor states, but only a single distance can be measured from each double fluorescent- or spin-labeled sample. However, a single NMR experiment offers hundreds to thousands of individual probes of distance and dynamics. Another advantage of NMR, though, which is true of any solution method, is that it can observe molecules under native or native-like conditions. Additionally, recent developments have also made it possible to perform some biomolecular NMR experiments on live cells (Dedmon *et al.* 2002; Inomata *et al.* 2009; Selenko *et al.* 2006), allowing measurements potentially in the most physiological environment possible. Thus, although much insight into the invisible states underlying biomolecular processes has already been gained from these NMR methods, the field may have only just begun to explore the potential of what it is capable.

6. Acknowledgements

We thank Nicolas Fawzi, Dennis Torchia, Vitali Tugarinov, Vincenzo Venditti and Jinfa Ying for useful discussions and other assistance. This work was supported by the Intramural Program of the NIH, NIDDK, and the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.).

7. References

Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. (2008). High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 7439–7444.

Altenbach, C., Oh, K. J., Trabanino, R. J., Hideg, K. & Hubbell, W. L. (2001). Estimation of inter-residue distances in spin labeled proteins at physiological temperatures: experimental strategies and practical limitations. *Biochemistry* **40**, 15471–15482.

Anthis, N. J. & Clore, G. M. (2013). The length of the calmodulin linker determines the extent of transient interdomain association and target affinity. *Journal of American Chemical Society* **135**, 9648–9651.

Anthis, N. J., Doucleff, M. & Clore, G. M. (2011). Transient, sparsely populated compact states of apo and calcium-loaded calmodulin probed by paramagnetic relaxation enhancement: interplay of conformational selection and induced fit. *Journal of American Chemical Society* **133**, 18966–18974.

Babu, Y. S., Sack, J. S., Greenhough, T. J., Bugg, C. E., Means, A. R. & Cook, W. J. (1985). Three-dimensional structure of calmodulin. *Nature* **315**, 37–40.

Bai, Y. & Englander, S. W. (1996). Future directions in folding: the multi-state nature of protein structure. *Proteins* **24**, 145–151.

Bai, Y. W. (2006). Protein folding pathways studied by pulsed-and native-state hydrogen exchange. *Chemical Reviews* **106**, 1757–1768.

Bai, Y. W., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein-folding intermediates – native-state hydrogen-exchange. *Science* **269**, 192–197.

Baker, M. L., Hryc, C. F., Zhang, Q., Wu, W., Jakana, J., Haase-Pettingell, C., Afonine, P. V., Adams, P. D., King, J. A., Jiang, W. & Chiu, W. (2013). Validated near-atomic resolution structure of bacteriophage epsilon15 derived from cryo-EM and modeling. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 12301–12306.

Balasubramaniam, D. & Komives, E. A. (2013). Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochimica et Biophysica Acta* **1834**, 1202–1209.

Baldwin, A. J. & Kay, L. E. (2009). NMR spectroscopy brings invisible protein states into focus. *Nature Chemical Biology* **5**, 808–814.

Baldwin, A. J. & Kay, L. E. (2013). An R(1rho) expression for a spin in chemical exchange between two sites with

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

unequal transverse relaxation rates. *Journal of Biomolecular NMR* **55**, 211–218.

Ban, D., Gossert, A. D., Giller, K., Becker, S., Griesinger, C. & Lee, D. (2012). Exceeding the limit of dynamics studies on biomolecules using high spin-lock field strengths with a cryogenically cooled probe-head. *Journal of Magnetic Resonance* **221**, 1–4.

Ban, D., Mazur, A., Carneiro, M. G., Sabo, T. M., Giller, K., Koharudin, L. M., Becker, S., Gronenborn, A. M., Griesinger, C. & Lee, D. (2013). Enhanced accuracy of kinetic information from CT-CPMG experiments by transverse rotating-frame spectroscopy. *Journal of Biomolecular NMR* **57**, 73–82.

Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W. & Bax, A. (1992). Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **31**, 5269–5278.

Bartesaghi, A., Merk, A., Borgnia, M. J., Milne, J. L. & Subramaniam, S. (2013). Prefusion structure of trimeric HIV-1 envelope glycoprotein determined by cryo-electron microscopy. *Nature Structural and Molecular Biology* **20**, 1352–1357.

Bashir, Q., Volkov, A. N., Ullmann, G. M. & Ubbink, M. (2009). Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome c and cytochrome c peroxidase. *Journal of American Chemical Society* **132**, 241–247.

Battiste, J. L. & Wagner, G. (2000). Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355–5365.

Baum, J., Dobson, C. M., Evans, P. A. & Hanley, C. (1989). Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig alpha-lactalbumin. *Biochemistry* **28**, 7–13.

Benetis, N. & Kowalewski, J. (1985). Nuclear-spin relaxation in paramagnetic systems (S = 1) in the slow-motion regime for the electron-spin .4. Motional anisotropy andnNoncoinciding dipole-dipole and zero-field Splitting tensors. *Journal of Magnetic Resonance* **65**, 13–33.

Berlin, K., Castaneda, C. A., Schneidman-Duhovny, D., Sali, A., Nava-Tudela, A. & Fushman, D. (2013). Recovering a representative conformational ensemble from underdetermined macromolecular structural data. *Journal of American Chemical Society* **135**, 16595–16609.

Bermejo, G. A., Strub, M.-P., Ho, C. & Tjandra, N. (2009). Determination of the solution-bound conformation of an amino acid binding protein by NMR paramagnetic relaxation enhancement: use of a single flexible paramagnetic probe with improved estimation of its sampling space. *Journal of American Chemical Society* **131**, 9532–9537.

Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *Journal of American Chemical Society* **129**, 5656–5664.

Bernini, A., Venditti, V., Spiga, O. & Niccolai, N. (2009). Probing protein surface accessibility with solvent and paramagnetic molecules. *Progress in Nuclear Magnetic Resonance Spectroscopy* **54**, 278–289.

Bertini, I., Ferella, L., Luchinat, C., Parigi, G., Petoukhov, M. V., Ravera, E., Rosato, A. & Svergun, D. I. (2012a). MaxOcc: a web portal for maximum occurrence analysis. *Journal of Biomolecular NMR* **53**, 271–280.

Bertini, I., Giachetti, A., Luchinat, C., Parigi, G., Petoukhov, M. V., Pierattelli, R., Ravera, E. & Svergun, D. I. (2010). Conformational space of flexible biological macromolecules from average data. *Journal of American Chemical Society* **132**, 13553–13558.

Bertini, I., Gupta, Y. K., Luchinat, C., Parigi, G., Peana, M., Sgheri, L. & Yuan, J. (2007). Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. *Journal of the American Chemical Society* **129**, 12786–12794.

Bertini, I., Luchinat, C., Nagulapalli, M., Parigi, G. & Ravera, E. (2012b). Paramagnetic relaxation enhancement for the characterization of the conformational heterogeneity in two-domain proteins. *Physical Chemistry and Chemical Physics* **14**, 9149–9156.

Bertini, I., Luchinat, C. & Parigi, G. (2001a). *Solution NMR of Paramagnetic Molecules: Applications to Metallobiomolecules and Models*. Elsevier Science, Amsterdam.

Bertini, I., Luchinat, C. & Parigi, G. (2002). Magnetic susceptibility in paramagnetic NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* **40**, 249–273.

Bertini, I., Luchinat, C., Parigi, G. & Pierattelli, R. (2005). NMR spectroscopy of paramagnetic metalloproteins. *Chembiochem* **6**, 1536–1549.

Bertini, I., Luchinat, C. & Piccioli, M. (2001b). Paramagnetic probes in metalloproteins. *Methods in Enzymology* **339**, 314–340.

Bertoncini, C. W., Jung, Y. S., Fernandez, C. O., Hoyer, W., Griesinger, C., Jovin, T. M. & Zweckstetter, M. (2005). Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 1430–1435.

Bloch, F. (1946). Nuclear induction. *Physics Reviews* **70**, 460–474.

Bloembergen, N. (1957). Proton relaxation times in paramagnetic solutions. *Journal of Chemical Physics* **27**, 572–573.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Bloembergen, N. & Morgan, L. O. (1961). Proton relaxation times in paramagnetic solutions effects of electron spin relaxation. *Journal of Chemical Physics* **34**, 842–850.

Bodner, C. R., Dobson, C. M. & Bax, A. (2009). Multiple tight phospholipid-binding modes of alpha-synuclein revealed by solution NMR spectroscopy. *Journal of Molecular Biology* **390**, 775–790.

Bodner, C. R., Maltsev, A. S., Dobson, C. M. & Bax, A. (2010). Differential phospholipid binding of alpha-synuclein variants implicated in Parkinson's disease revealed by solution NMR spectroscopy. *Biochemistry* **49**, 862–871.

Boehr, D. D., Dyson, H. J. & Wright, P. E. (2008). Conformational relaxation following hydride transfer plays a limiting role in dihydrofolate reductase catalysis. *Biochemistry* **47**, 9227–9233.

Boehr, D. D., Mcelheny, D., Dyson, H. J. & Wright, P. E. (2006). The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642.

Boehr, D. D., Mcelheny, D., Dyson, H. J. & Wright, P. E. (2010). Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1373–1378.

Bouvignies, G. & Kay, L. E. (2012a). A 2D 13C-CEST experiment for studying slowly exchanging protein systems using methyl probes: an application to protein folding. *Journal of Biomolecular NMR* **53**, 303–310.

Bouvignies, G. & Kay, L. E. (2012b). Measurement of proton chemical shifts in invisible states of slowly exchanging protein systems by chemical exchange saturation transfer. *Journal of Physical Chemistry B* **116**, 14311–14317.

Bouvignies, G., Korzhnev, D. M., Neudecker, P., Hansen, D. F., Cordes, M. H. & Kay, L. E. (2010). A simple method for measuring signs of (1)H (N) chemical shift differences between ground and excited protein states. *Journal of Biomolecular NMR* **47**, 135–141.

Bouvignies, G., Vallurupalli, P., Hansen, D. F., Correia, B. E., Lange, O., Bah, A., Vernon, R. M., Dahlquist, F. W., Baker, D. & Kay, L. E. (2011). Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* **477**, 111–114.

Bouvignies, G., Vallurupalli, P. & Kay, L. E. (2014). Visualizing side chains of invisible protein conformers by solution NMR. *Journal of Molecular Biology* **426**, 763–774.

Bridges, M. D., Hideg, K. & Hubbell, W. L. (2010). Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery EPR. *Applied Magnetic Resonance* **37**, 363–390.

Bruschweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M. & Ernst, R. R. (1992). Influence of rapid intramolecular motion on NMR cross-relaxation rates – a molecular-dynamics study of antamanide in solution. *Journal of American Chemical Society* **114**, 2289–2302.

Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein-folding – a synthesis. *Proteins* **21**, 167–195.

Bryngelson, J. D. & Wolynes, P. G. (1987). Spin-glasses and the statistical-mechanics of protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7524–7528.

Burnley, B. T., Afonine, P. V., Adams, P. D. & Gros, P. (2012). Modelling dynamics in protein crystal structures by ensemble refinement. *eLife* **1**, e00311.

Caravan, P., Ellison, J. J., Mcmurry, T. J. & Lauffer, R. B. (1999). Gadolinium(III) chelates as MRI contrast agents: structure, dynamics, and applications. *Chemical Reviews* **99**, 2293–2352.

Card, P. B., Erbel, P. J. A. & Gardner, K. H. (2005). Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization. *Journal of Molecular Biology* **353**, 664–677.

Carr, H. Y. & Purcell, E. M. (1954). Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physics Reviews* **94**, 630–638.

Carver, J. P. & Richards, R. E. (1972). General 2-site solution for chemical exchange produced dependence of T2 upon Carr-Purcell pulse separation. *Journal of Magnetic Resonance* **6**, 89–105.

Cavalli, A., Salvatella, X., Dobson, C. M. & Vendruscolo, M. (2007). Protein structure determination from NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 9615–9620.

Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Skelton, N. J. & Rance, M. (2007). *Protein NMR Spectroscopy: Principles and Practice*. Elsevier, Amsterdam.

Cayley, P. J., Albrand, J. P., Feeney, J., Roberts, G. C., Piper, E. A. & Burgen, A. S. (1979). Nuclear magnetic resonance studies of the binding of trimethoprim to dihydrofolate reductase. *Biochemistry* **18**, 3886–3895.

Cerofolini, L., Fields, G. B., Fragai, M., Geraldes, C. F., Luchinat, C., Parigi, G., Ravera, E., Svergun, D. I. & Teixeira, J. M. (2013). Examination of matrix metalloproteinase-1 in solution: a preference for the pro-collagenolysis state. *Journal of Biological Chemistry* **288**, 30659–30671.

Chamberlain, A. K., Handel, T. M. & Marqusee, S. (1996). Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nature Structural Biology* **3**, 782–787.

Chamberlain, A. K. & Marqusee, S. (2000). Comparison of equilibrium and kinetic approaches for determining protein folding mechanisms. *Advances in Protein Chemistry* **53**, 283–328.

Cheng, H. & Markley, J. L. (1995). NMR spectroscopic studies of paramagnetic proteins: iron-sulfur proteins.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Annual Review of Biophysics and Biomolecular Structure **24**, 209–237.

Chevelkov, V., Xue, Y., Rao, D. K., Forman-Kay, J. D. & Skrynnikov, N. R. (2010). 15N H/D-SOLEXSY experiment for accurate measurement of amide solvent exchange rates: application to denatured drkN SH3. *Journal of Biomolecular NMR* **46**, 227–244.

Chou, J. J., Case, D. A. & Bax, A. (2003). Insights into the mobility of methyl-bearing side chains in proteins from (3)J(CC) and (3)J(CN) couplings. *Journal of American Chemical Society* **125**, 8959–8966.

Choy, W. Y. & Forman-Kay, J. D. (2001). Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *Journal of Molecular Biology* **308**, 1011–1032.

Chu, R., Pei, W., Takei, J. & Bai, Y. (2002). Relationship between the native-state hydrogen exchange and folding pathways of a four-helix bundle protein. *Biochemistry* **41**, 7998–8003.

Cliff, M. J., Craven, C. J., Marston, J. P., Hounslow, A. M., Clarke, A. R. & Waltho, J. P. (2009). The denatured state of N-PGK Is compact and predominantly disordered. *Journal of Molecular Biology* **385**, 266–277.

Clore, G. M. (2000). Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9021–9025.

Clore, G. M. (2008). Visualizing lowly-populated regions of the free energy landscape of macromolecular complexes by paramagnetic relaxation enhancement. *Molecular Biosystems* **4**, 1058–1069.

Clore, G. M. (2011). Exploring sparsely populated states of macromolecules by diamagnetic and paramagnetic NMR relaxation. *Protein Science* **20**, 229–246.

Clore, G. M. (2013). Interplay between conformational selection and induced fit in multidomain protein-ligand binding probed by paramagnetic relaxation enhancement. *Biophysical Chemistry* **186**, 3–12.

Clore, G. M. & Bewley, C. A. (2002). Using conjoined rigid body/torsion angle simulated annealing to determine the relative orientation of covalently linked protein domains from dipolar couplings. *Journal of Magnetic Resonance* **154**, 329–335.

Clore, G. M., Driscoll, P. C., Wingfield, P. T. & Gronenborn, A. M. (1990a). Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear 15N-1H NMR spectroscopy. *Biochemistry* **29**, 7387–7401.

Clore, G. M. & Gronenborn, A. M. (1982). Theory and applications of the transferred nuclear overhauser effect to the study of the conformations of small ligands bound to proteins. *Journal of Magnetic Resonance* **48**, 402–417.

Clore, G. M. & Gronenborn, A. M. (1983). Theory of the time-dependent transferred nuclear overhauser effect – applications to structural-analysis of ligand protein complexes in solution. *Journal of Magnetic Resonance* **53**, 423–442.

Clore, G. M. & Gronenborn, A. M. (1991). Structures of larger proteins in solution – 3-dimensional and 4-dimensional heteronuclear nmr-spectroscopy. *Science* **252**, 1390–1399.

Clore, G. M. & Gronenborn, A. M. (1998). New methods of structure refinement for macromolecular structure determination by NMR. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5891–5898.

Clore, G. M. & Iwahara, J. (2009). Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chemical Reviews* **109**, 4108–4139.

Clore, G. M. & Schwieters, C. D. (2004a). Amplitudes of protein backbone dynamics and correlated motions in a small alpha/beta protein: correspondence of dipolar coupling and heteronuclear relaxation measurements. *Biochemistry* **43**, 10678–10691.

Clore, G. M. & Schwieters, C. D. (2004b). How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation? *Journal of American Chemical Society* **126**, 2923–2938.

Clore, G. M. & Schwieters, C. D. (2006). Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: a unified picture of high probability, fast atomic motions in proteins. *Journal of Molecular Biology* **355**, 879–886.

Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C. & Gronenborn, A. M. (1990b). Deviations from the simple two-parameter model-free approach to the interpretation of N-15 nuclear magnetic-relaxation of proteins. *Journal of American Chemical Society* **112**, 4989–4991.

Clore, G. M., Tang, C. & Iwahara, J. (2007). Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Current Opinion in Structural Biology* **17**, 603–616.

Cole, R. & Loria, J. P. (2002). Evidence for flexibility in the function of ribonuclease A. *Biochemistry* **41**, 6072–6081.

Columbus, L., Kalai, T., Jeko, J., Hideg, K. & Hubbell, W. L. (2001). Molecular motion of spin labeled side chains in alpha-helices: analysis by variation of side chain structure. *Biochemistry* **40**, 3828–3846.

Conicella, A. E. & Fawzi, N. L. (2014). The C-terminal threonine of Abeta43 nucleates toxic aggregation via

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

structural and dynamical changes in monomers and protofibrils. *Biochemistry* **20**, 3095–3105.

Crisma, M., Deschamps, J. R., George, C., Flippen-Anderson, J. L., Kaptein, B., Broxterman, Q. B., Moretto, A., Oancea, S., Jost, M., Formaggio, F. & Toniolo, C. (2005). A topographically and conformationally constrained, spin-labeled, alpha-amino acid: crystallographic characterization in peptides. *Journal of Peptide Research* **65**, 564–579.

Dasgupta, S., Hu, X., Keizers, P. H. J., Liu, W.-M., Luchinat, C., Nagulapalli, M., Overhand, M., Parigi, G., Sgheri, L. & Ubbink, M. (2011). Narrowing the conformational space sampled by two-domain proteins with paramagnetic probes in both domains. *Journal of Biomolecular NMR* **51**, 253–263.

Davis, D. G., Perlman, M. E. & London, R. E. (1994). Direct measurements of the dissociation-rate constant for inhibitor-enzyme complexes via the T1 rho and T2 (CPMG) methods. *Journal of Magnetic Resonance B* **104**, 266–275.

Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M. & Dobson, C. M. (2005). Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *Journal of American Chemical Society* **127**, 476–477.

Dedmon, M. M., Patel, C. N., Young, G. B. & Pielak, G. J. (2002). FlgM gains structure in living cells. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12681–12684.

Deshmukh, L., Schwieters, C. D., Grishaev, A., Ghirlando, R., Baber, J. L. & Clore, G. M. (2013). Structure and dynamics of full-length HIV-1 capsid protein in solution. *Journal of American Chemical Society* **135**, 16133–16147.

Deverell, C., Morgan, R. E. & Strange, J. H. (1970). Studies of chemical exchange by nuclear magnetic relaxation in rotating frame. *Molecular Physics* **18**, 553–559.

Dharmasiri, K. & Smith, D. L. (1996). Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins. *Analytical Chemistry* **68**, 2340–2344.

Donaldson, L. W., Skrynnikov, N. R., Choy, W. Y., Muhandiram, D. R., Sarkar, B., Forman-Kay, J. D. & Kay, L. E. (2001). Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. *Journal of American Chemical Society* **123**, 9843–9847.

Dvoretsky, A., Gaponenko, V. & Rosevear, P. R. (2002). Derivation of structural restraints using a thiol-reactive chelator. *FEBS Letters* **528**, 189–192.

Ebright, Y. W., Chen, Y., Pendergrast, P. S. & Ebright, R. H. (1992). 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* **31**, 10664–10670.

Ehrhardt, M. R., Urbauer, J. L. & Wand, A. J. (1995). The energetics and dynamics of molecular recognition by calmodulin. *Biochemistry* **34**, 2731–2738.

Eichmuller, C. & Skrynnikov, N. R. (2007). Observation of microsecond time-scale protein dynamics in the presence of Ln^{3+} ions: application to the N-terminal domain of cardiac troponin C. *Journal of Biomolecular NMR* **37**, 79–95.

Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E. & Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **438**, 117–121.

Englander, J. J., Del Mar, C., Li, W., Englander, S. W., Kim, J. S., Stranz, D. D., Hamuro, Y. & Woods, V. L. Jr. (2003). Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 7057–7062.

Englander, S. W. (2006). Hydrogen exchange and mass spectrometry: a historical perspective. *Journal of American Society for Mass Spectrometry* **17**, 1481–1489.

Englander, S. W. & Kallenbach, N. R. (1983). Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Quarterly Reviews of Biophysics* **16**, 521–655.

Englander, S. W., Mayne, L., Bai, Y. & Sosnick, T. R. (1997). Hydrogen exchange: the modern legacy of Linderstrom-Lang. *Protein Science* **6**, 1101–1109.

Englander, S. W., Mayne, L. & Krishna, M. M. G. (2007). Protein folding and misfolding: mechanism and principles. *Quarterly Reviews of Biophysics* **40**, 287–326.

Ermacora, M. R., Delfino, J. M., Cuenoud, B., Schepartz, A. & Fox, R. O. (1992). Conformation-dependent cleavage of staphylococcal nuclease with a disulfide-linked iron chelate. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 6383–6387.

Eyles, S. J. & Kaltashov, I. A. (2004). Methods to study protein dynamics and folding by mass spectrometry. *Methods* **34**, 88–99.

Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. (1994). A heteronuclear correlation experiment for simultaneous determination of 15N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *Journal of Biomolecular NMR* **4**, 727–734.

Fawzi, N. L., Doucleff, M., Suh, J. Y. & Clore, G. M. (2010a). Mechanistic details of a protein-protein association pathway revealed by paramagnetic relaxation enhancement titration measurements. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1379–1384.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Fawzi, N. L., Fleissner, M. R., Anthis, N. J., Kalai, T., Hideg, K., Hubbell, W. L. & Clore, G. M. (2011a). A rigid disulfide-linked nitroxide side chain simplifies the quantitative analysis of PRE data. *Journal of Biomolecular NMR* **51**, 105–114.

Fawzi, N. L., Libich, D. S., Ying, J., Tugarinov, V. & Clore, G. M. (2014). Characterizing methyl-bearing sidechain contacts and dynamics mediating amyloid β protofibril interactions using $^{13}C_{methyl}$-DEST and lifetime line broadening. *Angewandte Chimie International Edition* **53**, 10345–10349.

Fawzi, N. L., Ying, J., Ghirlando, R., Torchia, D. A. & Clore, G. M. (2011b). Atomic-resolution dynamics on the surface of amyloid β protofibrils probed by solution NMR. *Nature* **480**, 268–272.

Fawzi, N. L., Ying, J., Torchia, D. A. & Clore, G. M. (2010b). Kinetics of amyloid β monomer-to-oligomer exchange by NMR relaxation. *Journal of American Chemical Society* **132**, 9948–9951.

Fawzi, N. L., Ying, J., Torchia, D. A. & Clore, G. M. (2012). Probing exchange kinetics and atomic resolution dynamics in high-molecular-weight complexes using dark-state exchange saturation transfer NMR spectroscopy. *Nature Protocols* **7**, 1523–1533.

Felitsky, D. J., Lietzow, M. A., Dyson, H. J. & Wright, P. E. (2008). Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6278–6283.

Ferreiro, D. U., Cho, S. S., Komives, E. A. & Wolynes, P. G. (2005). The energy landscape of modular repeat proteins: topology determines folding mechanism in the ankyrin family. *Journal of Molecular Biology* **354**, 679–692.

Fesik, S. W., Gemmecker, G., Olejniczak, E. T. & Petros, A. M. (1991). Identification of solvent-exposed regions of enzyme-bound ligands by nuclear-magnetic-resonance. *Journal of American Chemical Society* **113**, 7080–7081.

Fisher, C. K. & Stultz, C. M. (2011). Constructing ensembles for intrinsically disordered proteins. *Current Opinion in Structural Biology* **21**, 426–431.

Fitzkee, N. C., Torchia, D. A. & Bax, A. (2011). Measuring rapid hydrogen exchange in the homodimeric 36 kDa HIV-1 integrase catalytic core domain. *Protein Science* **20**, 500–512.

Fleissner, M. R., Bridges, M. D., Brooks, E. K., Cascio, D., Kalai, T., Hideg, K. & Hubbell, W. L. (2011). Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16241–16246.

Fleissner, M. R., Brustad, E. M., Kalai, T., Altenbach, C., Cascio, D., Peters, F. B., Hideg, K., Peuker, S., Schultz, P. G. & Hubbell, W. L. (2009). Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 21637–21642.

Flippen-Anderson, J. L., George, C., Valle, G., Valente, E., Bianco, A., Formaggio, F., Crisma, M. & Toniolo, C. (1996). Crystallographic characterization of geometry and conformation of TOAC, a nitroxide spin-labelled C-alpha, C-alpha-disubstituted glycine, in simple derivatives and model peptides. *International Journal of Peptide and Protein Research* **47**, 231–238.

Forge, V., Wijesinha, R. T., Balbach, J., Brew, K., Robinson, C. V., Redfield, C. & Dobson, C. M. (1999). Rapid collapse and slow structural reorganisation during the refolding of bovine alpha-lactalbumin. *Journal of Molecular Biology* **288**, 673–688.

Forsén, S. & Hoffman, R. A. (1963). Study of moderately rapid chemical exchange reactions by means of nuclear magnetic double resonance. *Journal of Chemical Physics* **39**, 2892–2901.

Fragai, M., Luchinat, C., Parigi, G. & Ravera, E. (2013). Conformational freedom of metalloproteins revealed by paramagnetism-assisted NMR. *Coordination Chemistry Reviews* **257**, 2652–2667.

Fraser, J. S., Clarkson, M. W., Degnan, S. C., Erion, R., Kern, D. & Alber, T. (2009). Hidden alternative structures of proline isomerase essential for catalysis. *Nature* **462**, 669–673.

Freed, J. H. (1978). Dynamic effects of pair correlation-functions on spin relaxation by translational diffusion in liquids .2. Finite jumps and independent T1 processes. *Journal of Chemical Physics* **68**, 4034–4037.

Friedman, J. I., Mcmahon, M. T., Stivers, J. T. & Van Zijl, P. C. (2010). Indirect detection of labile solute proton spectra via the water signal using frequency-labeled exchange (FLEX) transfer. *Journal of American Chemical Society* **132**, 1813–1815.

Fuentes, E. J. & Wand, A. J. (1998a). Local dynamics and stability of apocytochrome b562 examined by hydrogen exchange. *Biochemistry* **37**, 3687–3698.

Fuentes, E. J. & Wand, A. J. (1998b). Local stability and dynamics of apocytochrome b562 examined by the dependence of hydrogen exchange on hydrostatic pressure. *Biochemistry* **37**, 9877–9883.

Gaponenko, V., Howarth, J. W., Columbus, L., Gasmi-Seabrook, G., Yuan, J., Hubbell, W. L. & Rosevear, P. R. (2000). Protein global fold determination using site-directed spin and isotope labeling. *Protein Science* **9**, 302–309.

Gardner, R. J., Longinetti, M. & Sgheri, L. (2005). Reconstruction of orientations of a moving protein domain from paramagnetic data. *Inverse Problems* **21**, 879.

Gillespie, J. R. & Shortle, D. (1997a). Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *Journal of Molecular Biology* **268**, 158–169.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Gillespie, J. R. & Shortle, D. (1997b). Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *Journal of Molecular Biology* **268**, 170–184.

Goedert, M. & Spillantini, M. G. (2006). A century of Alzheimer's disease. *Science* **314**, 777–781.

Gottstein, D., Reckel, S., Dotsch, V. & Guntert, P. (2012). Requirements on paramagnetic relaxation enhancement data for membrane protein structure determination by NMR. *Structure* **20**, 1019–1027.

Griffith, O. H. & Mcconnell, H. M. (1966). A nitroxide-maleimide spin label. *Proceedings of the National Academy of Sciences of the United States of America* **55**, 8–11.

Grishaev, A., Anthis, N. J. & Clore, G. M. (2012). Contrast-matched small-angle X-ray scattering from a heavy-atom labeled protein in structure determination: application to a lead-substituted calmodulin–peptide complex. *Journal of American Chemical Society* **134**, 14686–14689.

Gross, J. D., Moerke, N. J., Von Der Haar, T., Lugovskoy, A. A., Sachs, A. B., Mccarthy, J. E. G. & Wagner, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4 G and eIF4E. *Cell* **115**, 739–750.

Gueron, M. (1975). Nuclear-relaxation in macromolecules by paramagnetic-ions – novel mechanism. *Journal of Magnetic Resonance* **19**, 58–66.

Gupta, R. K. & Redfield, A. G. (1970). Double nuclear magnetic resonance observation of electron exchange between ferri- and ferrocytochrome c. *Science* **169**, 1204–1206.

Hancu, I., Dixon, W. T., Woods, M., Vinogradov, E., Sherry, A. D. & Lenkinski, R. E. (2010). CEST and PARACEST MR contrast agents. *Acta Radiology* **51**, 910–923.

Hansen, A. L., Lundstrom, P., Velyvis, A. & Kay, L. E. (2012). Quantifying millisecond exchange dynamics in proteins by CPMG relaxation dispersion NMR using side-chain 1H probes. *Journal of American Chemical Society* **134**, 3178–3189.

Hansen, A. L., Nikolova, E. N., Casiano-Negroni, A. & Al-Hashimi, H. M. (2009). Extending the range of microsecond-to-millisecond chemical exchange detected in labeled and unlabeled nucleic acids by selective carbon R1(rho) NMR spectroscopy. *Journal of American Chemical Society* **131**, 3818–3819.

Hansen, D. F., Vallurupalli, P. & Kay, L. E. (2008a). Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *Journal of Biomolecular NMR* **41**, 113–120.

Hansen, D. F., Vallurupalli, P., Lundstrom, P., Neudecker, P. & Kay, L. E. (2008b). Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: how well can we do? *Journal of American Chemical Society* **130**, 2667–2675.

Hanson, P., Millhauser, G., Formaggio, F., Crisma, M. & Toniolo, C. (1996). ESR characterization of hexameric, helical peptides using double TOAC spin labeling. *Journal of American Chemical Society* **118**, 7618–7625.

Hass, M. A., Keizers, P. H., Blok, A., Hiruma, Y. & Ubbink, M. (2010). Validation of a lanthanide tag for the analysis of protein dynamics by paramagnetic NMR spectroscopy. *Journal of American Chemical Society* **132**, 9952–9953.

Haussinger, D., Huang, J. R. & Grzesiek, S. (2009). DOTA-M8: an extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. *Journal of American Chemical Society* **131**, 14761–14767.

Helgstrand, M., Hard, T. & Allard, P. (2000). Simulations of NMR pulse sequences during equilibrium and non-equilibrium chemical exchange. *Journal of Biomolecular NMR* **18**, 49–63.

Hernandez, G., Teng, C. L., Bryant, R. G. & Lemaster, D. M. (2002). O_2 penetration and proton burial depth in proteins: applicability to fold family recognition. *Journal of American Chemical Society* **124**, 4463–4472.

Hillisch, A., Lorenz, M. & Diekmann, S. (2001). Recent advances in FRET: distance determination in protein–DNA complexes. *Current Opinion in Structural Biology* **11**, 201–207.

Hitchens, T. K. & Bryant, R. G. (1998). Pressure dependence of amide hydrogen–deuterium exchange rates for individual sites in T4 lysozyme. *Biochemistry* **37**, 5878–5887.

Horwich, A. L. & Fenton, W. A. (2009). Chaperonin-mediated protein folding: using a central cavity to kinetically assist polypeptide chain folding. *Quarterly Reviews of Biophysics* **42**, 83–116.

Hu, K., Vogeli, B. & Clore, G. M. (2007). Spin-state selective carbon-detected HNCO with TROSY optimization in all dimensions and double echo-antiecho sensitivity enhancement in both indirect dimensions. *Journal of American Chemical Society* **129**, 5484–5491.

Huang, F., Pei, Y. Y., Zuo, H. H., Chen, J. L., Yang, Y. & Su, X. C. (2013). Bioconjugation of proteins with a paramagnetic NMR and fluorescent tag. *Chemistry* **19**, 17141–17149.

Hubbell, W. L., Cafiso, D. S. & Altenbach, C. (2000). Identifying conformational changes with site-directed spin labeling. *Nature Structural Biology* **7**, 735–739.

Hubbell, W. L., Lopez, C. J., Altenbach, C. & Yang, Z. (2013). Technological advances in site-directed spin labeling of proteins. *Current Opinion in Structural Biology* **23**, 725–733.

Hughson, F. M., Wright, P. E. & Baldwin, R. L. (1990). Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249**, 1544–1548.

Hwang, L. P. & Freed, J. H. (1975). Dynamic effects of pair correlation-functions on spin relaxation by

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

translational diffusion in liquids. *Journal of Chemical Physics* **63**, 4017–4025.

Igumenova, T. I., Frederick, K. K. & Wand, A. J. (2006). Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chemical Reviews* **106**, 1672–1699.

Ikegami, T., Verdier, L., Sakhaii, P., Grimme, S., Pescatore, B., Saxena, K., Fiebig, K. M. & Griesinger, C. (2004). Novel techniques for weak alignment of proteins in solution using chemical tags coordinating lanthanide ions. *Journal of Biomolecular NMR* **29**, 339–349.

Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B. & Bax, A. (1992). Solution structure of a calmodulin–target peptide complex by multidimensional NMR. *Science* **256**, 632–638.

Inomata, K., Ohno, A., Tochio, H., Isogai, S., Tenno, T., Nakase, I., Takeuchi, T., Futaki, S., Ito, Y., Hiroaki, H. & Shirakawa, M. (2009). High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458**, 106–109.

Ishima, R., Freedberg, D. I., Wang, Y. X., Louis, J. M. & Torchia, D. A. (1999). Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. *Structure* **7**, 1047–1055.

Ishima, R. & Torchia, D. A. (1999). Estimating the time scale of chemical exchange of proteins from measurements of transverse relaxation rates in solution. *Journal of Biomolecular NMR* **14**, 369–372.

Ishima, R. & Torchia, D. A. (2000). Protein dynamics from NMR. *Nature Structural Biology* **7**, 740–743.

Ishima, R. & Torchia, D. A. (2003). Extending the range of amide proton relaxation dispersion experiments in proteins using a constant-time relaxation-compensated CPMG approach. *Journal of BiomolecularNMR* **25**, 243–248.

Iwahara, J., Anderson, D. E., Murphy, E. C. & Clore, G. M. (2003). EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. *Journal of American Chemical Society* **125**, 6634–6635.

Iwahara, J. & Clore, G. M. (2006a). Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **440**, 1227–1230.

Iwahara, J. & Clore, G. M. (2006b). Direct observation of enhanced translocation of a homeodomain between DNA cognate sites by NMR exchange spectroscopy. *Journal of American Chemical Society* **128**, 404–405.

Iwahara, J. & Clore, G. M. (2010). Structure-independent analysis of the breadth of the positional distribution of disordered groups in macromolecules from order parameters for long, variable-length vectors using NMR paramagnetic relaxation enhancement. *Journal of American Chemical Society* **132**, 13346–13356.

Iwahara, J., Schwieters, C. D. & Clore, G. M. (2004a). Characterization of nonspecific protein-DNA interactions by H-1 paramagnetic relaxation enhancement. *Journal of American Chemical Society* **126**, 12800–12808.

Iwahara, J., Schwieters, C. D. & Clore, G. M. (2004b). Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *Journal of American Chemical Society* **126**, 5879–5896.

Iwahara, J., Tang, C. & Clore, G. M. (2007). Practical aspects of (1)H transverse paramagnetic relaxation enhancement measurements on macromolecules. *Journal of Magnetic Resonance* **184**, 185–195.

Iwahara, J., Zweckstetter, M. & Clore, G. M. (2006). NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15062–15067.

Jain, N. U., Venot, A., Umemoto, K., Leffler, H. & Prestegard, J. H. (2001). Distance mapping of protein-binding sites using spin-labeled oligosaccharide ligands. *Protein Science* **10**, 2393–2400.

Jen, J. (1978). Chemical exchange and NMR T2 relaxation – multisite case. *Journal of Magnetic Resonance* **30**, 111–128.

Jennings, P. A. & Wright, P. E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* **262**, 892–896.

Jensen, M. R., Lauritzen, C., Dahl, S. W., Pedersen, J. & Led, J. J. (2004). Binding ability of a HHP-tagged protein towards Ni^{2+} studied by paramagnetic NMR relaxation: the possibility of obtaining long-range structure information. *Journal of Biomolecular. NMR* **29**, 175–185.

Jeschke, G. (2013). Conformational dynamics and distribution of nitroxide spin labels. *Progress in Nuclear Magnetic Resonance Spectroscopy* **72**, 42–60.

Johnson, P. E., Brun, E., Mackenzie, L. F., Withers, S. G. & Mcintosh, L. P. (1999). The cellulose-binding domains from Cellulomonas fimi beta-1,4-glucanase CenC bind nitroxide spin-labeled cellooligosaccharides in multiple orientations. *Journal of Molecular Biology* **287**, 609–625.

Kamerlin, S. C. L. & Warshel, A. (2010). Reply to Karplus: conformational dynamics have no role in the chemical step. *Proceedings of the National Academy of Sciences of the United States of America* **107**, E72.

Karplus, M. (2010). Role of conformation transitions in adenylate kinase. *Proceedings of the National Academy of Sciences of the United States of America* **107**, E71.

Kateb, F., Pelupessy, P. & Bodenhausen, G. (2007). Measuring fast hydrogen exchange rates by NMR spectroscopy. *Journal of Magnetic Resonance* **184**, 108–113.

Kay, L. E., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

sensitivity. *Journal of American Chemical Society* **114**, 10663–10665.

Kay, L. E., Torchia, D. A. & Bax, A. (1989). Backbone dynamics of proteins as studied by N-15 inverse detected heteronuclear NMR-spectroscopy – application to Staphylococcal nuclease. *Biochemistry* **28**, 8972–8979.

Keizers, P. H. & Ubbink, M. (2011). Paramagnetic tagging for protein structure and dynamics analysis. *Progress in Nuclear Magnetic Resonance Spectroscopy* **58**, 88–96.

Keizers, P. H. J., Desreux, J. F., Overhand, M. & Ubbink, M. (2007). Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. *Journal of American Chemical Society* **129**, 9292.

Keizers, P. H. J., Saragliadis, A., Hiruma, Y., Overhand, M. & Ubbink, M. (2008). Design, synthesis, and evaluation of a lanthanide chelating protein probe: CLaNP-5 yields predictable paramagnetic effects independent of environment. *Journal of American Chemical Society* **130**, 14802–14812.

Kempf, J. G., Jung, J.-Y., Ragain, C., Sampson, N. S. & Loria, J. P. (2007). Dynamic requirements for a functional protein hinge. *Journal of Molecular Biology* **368**, 131–149.

Klug, C. S. & Feix, J. B. (2008). Methods and applications of site-directed spin labeling EPR spectroscopy. *Methods in Cell Biology* **84**, 617–658.

Konermann, L., Tong, X. & Pan, Y. (2008). Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *Journal of Mass Spectrometry* **43**, 1021–1036.

Korzhnev, D. M. & Kay, L. E. (2008). Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. *Accounts of Chemical Research* **41**, 442–451.

Korzhnev, D. M., Kloiber, K., Kanelis, V., Tugarinov, V. & Kay, L. E. (2004a). Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *Journal of American Chemical Society* **126**, 3964–3973.

Korzhnev, D. M., Neudecker, P., Zarrine-Afsar, A., Davidson, A. R. & Kay, L. E. (2006). Abp1p and fyn SH3 domains fold through similar low-populated intermediate states. *Biochemistry* **45**, 10175–10183.

Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R. & Kay, L. E. (2010). A transient and low-populated protein-folding intermediate at atomic resolution. *Science* **329**, 1312–1316.

Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Di Nardo, A. A., Davidson, A. R., Dobson, C. M. & Kay, L. E. (2004b). Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586–590.

Kosen, P. A. (1989). Spin labeling of proteins. *Methods in Enzymology* **177**, 86–121.

Kosen, P. A., Scheek, R. M., Naderi, H., Basus, V. J., Manogaran, S., Schmidt, P. G., Oppenheimer, N. J. & Kuntz, I. D. (1986). Two-dimensional 1H NMR of three spin-labeled derivatives of bovine pancreatic trypsin inhibitor. *Biochemistry* **25**, 2356–2364.

Krishna, M. M., Hoang, L., Lin, Y. & Englander, S. W. (2004a). Hydrogen exchange methods to study protein folding. *Methods* **34**, 51–64.

Krishna, M. M., Lin, Y., Mayne, L. & Englander, S. W. (2003). Intimate view of a kinetic protein folding intermediate: residue-resolved structure, interactions, stability, folding and unfolding rates, homogeneity. *Journal of Molecular Biology* **334**, 501–513.

Krishna, M. M. G., Hoang, L., Lin, Y. & Englander, S. W. (2004b). Hydrogen exchange methods to study protein folding. *Methods* **34**, 51–64.

Kristjansdottir, S., Lindorff-Larsen, K., Fieber, W., Dobson, C. M., Vendruscolo, M. & Poulsen, F. M. (2005). Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. *Journal of Molecular Biology* **347**, 1053–1062.

Lange, O. F., Lakomek, N.-A., Fares, C., Schroeder, G. F., Walter, K. F. A., Becker, S., Meiler, J., Grubmueller, H., Griesinger, C. & De Groot, B. L. (2008). Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **320**, 1471–1475.

Li, P., Martins, I. R. & Rosen, M. K. (2011). The feasibility of parameterizing four-state equilibria using relaxation dispersion measurements. *Journal of Biomolecular NMR* **51**, 57–70.

Liang, B. Y., Bushweller, J. H. & Tamm, L. K. (2006). Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy. *Journal of American Chemical Society* **128**, 4389–4397.

Libich, D. S., Fawzi, N. L., Ying, J. & Clore, G. M. (2013). Probing the transient dark state of substrate binding to GroEL by relaxation-based solution NMR. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11361–11366.

Lietzow, M. A., Jamin, M., Dyson, H. J. & Wright, P. E. (2002). Mapping long-range contacts in a highly unfolded protein. *Journal of Molecular Biology* **322**, 655–662.

Lindfors, H. E., De Koning, P. E., Drijfhout, J. W., Venezia, B. & Ubbink, M. (2008). Mobility of TOAC spin-labelled peptides binding to the Src SH3 domain studied by paramagnetic NMR. *Journal of Biomolecular NMR* **41**, 157–167.

Lindfors, H. E., Venkata, B. S., Drijfhout, J. W. & Ubbink, M. (2011). Linker length dependent binding of a focal adhesion kinase derived peptide to the Src SH3-SH2 domains. *FEBS Letters* **585**, 601–605.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Lipari, G. & Szabo, A. (1982a). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules .1. Theory and range of validity. *Journal of American Chemical Society* **104**, 4546–4559.

Lipari, G. & Szabo, A. (1982b). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules .2. Analysis of experimental results. *Journal of American Chemical Society* **104**, 4559–4570.

Liu, W. M., Keizers, P. H., Hass, M. A., Blok, A., Timmer, M., Sarris, A. J., Overhand, M. & Ubbink, M. (2012). A pH-sensitive, colorful, lanthanide-chelating paramagnetic NMR probe. *Journal of American Chemical Society* **134**, 17306–17313.

Liu, Z., Gong, Z., Guo, D. C., Zhang, W. P. & Tang, C. (2014). Subtle dynamics of holo glutamine binding protein revealed with a rigid paramagnetic probe. *Biochemistry* **53**, 1403–1409.

Long, D., Bouvignies, G. & Kay, L. E. (2014). Measuring hydrogen exchange rates in invisible protein excited states. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8820–8825.

Long, D., Marshall, C. B., Bouvignies, G., Mazhab-Jafari, M. T., Smith, M. J., Ikura, M. & Kay, L. E. (2013). A comparative CEST NMR study of slow conformational dynamics of small GTPases complexed with GTP and GTP analogues. *Angewandte Chemie International Edition English* **52**, 10771–10774.

Longinetti, M., Luchinat, C., Parigi, G. & Sgheri, L. (2006). Efficient determination of the most favoured orientations of protein domains from paramagnetic NMR data. *Inverse Problems* **22**, 1485.

Loria, J. P., Berlow, R. B. & Watt, E. D. (2008). Characterization of enzyme motions by solution NMR relaxation dispersion. *Accounts of Chemical Research* **41**, 214–221.

Luchinat, C., Nagulapalli, M., Parigi, G. & Sgheri, L. (2012). Maximum occurrence analysis of protein conformations for different distributions of paramagnetic metal ions within flexible two-domain proteins. *Journal of Magnetic Resonance* **215**, 85–93.

Lundstrom, P. & Akke, M. (2005). Microsecond protein dynamics measured by 13Calpha rotating-frame spin relaxation. *Chembiochem* **6**, 1685–1692.

Lundstrom, P., Hansen, D. F., Vallurupalli, P. & Kay, L. E. (2009a). Accurate measurement of alpha proton chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy. *Journal of American Chemical Society* **131**, 1915–1926.

Lundstrom, P., Lin, H. & Kay, L. E. (2009b). Measuring C-13(beta) chemical shifts of invisible excited states in proteins by relaxation dispersion NMR spectroscopy. *Journal of Biomolecular NMR* **44**, 139–155.

Lundstrom, P., Vallurupalli, P., Religa, T. L., Dahlquist, F. W. & Kay, L. E. (2007). A single-quantum methyl C-13-relaxation dispersion experiment with improved sensitivity. *Journal of BiomolecularNMR* **38**, 79–88.

Lyumkis, D., Julien, J. P., De Val, N., Cupo, A., Potter, C. S., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Carragher, B., Wilson, I. A. & Ward, A. B. (2013). Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1484–1490.

Macnaughtan, M. A., Kamar, M., Alvarez-Manilla, G., Venot, A., Glushka, J., Pierce, J. M. & Prestegard, J. H. (2007). NMR structural characterization of substrates bound to N-acetylglucosaminyltransferase V. *Journal of Molecular Biology* **366**, 1266–1281.

Mahoney, N. M., Rastogi, V. K., Cahill, S. M., Girvin, M. E. & Almo, S. C. (2000). Binding orientation of proline-rich peptides in solution: polarity of the profilin-ligand interaction. *Journal of American Chemical Society* **122**, 7851–7852.

Maity, H., Maity, M., Krishna, M. M. G., Mayne, L. & Englander, S. W. (2005). Protein folding: the stepwise assembly of foldon units. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4741–4746.

Mal, T. K., Ikura, M. & Kay, L. E. (2002). The ATCUN domain as a probe of intermolecular interactions: application to calmodulin-peptide complexes. *Journal of American Chemical Society* **124**, 14002–14003.

Marsh, J. A. & Forman-Kay, J. D. (2009). Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. *Journal of Molecular Biology* **391**, 359–374.

Marsh, J. A., Neale, C., Jack, F. E., Choy, W.-Y., Lee, A. Y., Crowhurst, K. A. & Forman-Kay, J. D. (2007). Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. *Journal of Molecular Biology* **367**, 1494–1510.

Mayer, M. & Meyer, B. (1999). Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie International Edition English* **38**, 1784–1788.

Mcconnell, H. M. (1958). Reaction rates by nuclear magnetic resonance. *Journal of Chemical Physics* **28**, 430–431.

Meador, W. E., Means, A. R. & Quiocho, F. A. (1992). Target enzyme recognition by calmodulin: 2·4 A structure of a calmodulin–peptide complex. *Science* **257**, 1251–1255.

Meiboom, S. (1961). Nuclear magnetic resonance study of proton transfer in water. *Journal of Chemical Physics* **34**, 375.

Meiboom, S. & Gill, D. (1958). Modified spin-echo method for measuring nuclear relaxation times. *Reviews of Scientific Instruments* **29**, 688–691.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Milov, A. D., Tsvetkov, Y. D., Gorbunova, E. Y., Mustaeva, L. G., Ovchinnikova, T. V. & Raap, J. (2002). Self-aggregation properties of spin-labeled zervamicin IIA as studied by PELDOR spectroscopy. *Biopolymers* **64**, 328–336.

Mittermaier, A. & Kay, L. E. (2006). New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224–228.

Mittermaier, A. K. & Kay, L. E. (2009). Observing biological dynamics at atomic resolution using NMR. *Trends in Biochemical Sciences* **34**, 601–611.

Miyashita, O., Onuchic, J. N. & Wolynes, P. G. (2003). Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12570–12575.

Miyashita, O., Wolynes, P. G. & Onuchic, J. N. (2005). Simple energy landscape model for the kinetics of functional transitions in proteins. *Journal of Physical Chemistry B* **109**, 1959–1969.

Monaco, V., Formaggio, F., Crisma, M., Toniolo, C., Hanson, P., Millhauser, G., George, C., Deschamps, J. R. & Flippen-Anderson, J. L. (1999a). Determining the occurrence of a 3(10)-helix and an alpha-helix in two different segments of a lipopeptaibol antibiotic using TOAC, a nitroxide spin-labeled C-alpha-tetrasubstituted alpha-amino acid. *Bioorganic and Medicinal Chemistry* **7**, 119–131.

Monaco, V., Formaggio, F., Crisma, M., Toniolo, C., Hanson, P. & Millhauser, G. L. (1999b). Orientation and immersion depth of a helical lipopeptaibol in membranes using TOAC as an ESR probe. *Biopolymers* **50**, 239–253.

Montelione, G. T., Winkler, M. E., Burton, L. E., Rinderknecht, E., Sporn, M. B. & Wagner, G. (1989). Sequence-specific 1H-NMR assignments and identification of two small antiparallel beta-sheets in the solution structure of recombinant human transforming growth factor alpha. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 1519–1523.

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

Nagulapalli, M., Parigi, G., Yuan, J., Gsponer, J., Deraos, G., Bamm, V. V., Harauz, G., Matsoukas, J., De Planque, M. R., Gerothanassis, I. P., Babu, M. M., Luchinat, C. & Tzakos, A. G. (2012). Recognition pliability is coupled to structural heterogeneity: a calmodulin intrinsically disordered binding region complex. *Structure* **20**, 522–533.

Neudecker, P., Lundstrom, P. & Kay, L. E. (2009). Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding. *Biophysical Journal* **96**, 2045–2054.

Neudecker, P., Robustelli, P., Cavalli, A., Walsh, P., Lundstrom, P., Zarrine-Afsar, A., Sharpe, S., Vendruscolo, M. & Kay, L. E. (2012). Structure of an intermediate state in protein folding and aggregation. *Science* **336**, 362–366.

Neudecker, P., Zarrine-Afsar, A., Davidson, A. R. & Kay, L. E. (2007). Phi-Value analysis of a three-state protein folding pathway by NMR relaxation dispersion spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 15717–15722.

Nodet, G., Salmon, L., Ozenne, V., Meier, S., Jensen, M. R. & Blackledge, M. (2009). Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. *Journal of American Chemical Society* **131**, 17908–17918.

Oganesyan, V. S. (2011). A general approach for prediction of motional EPR spectra from Molecular Dynamics (MD) simulations: application to spin labelled protein. *Physical Chemistry and Chemical Physics* **13**, 4724–4737.

Ogawa, S. & Mcconnel, H. M. (1967). Spin-label study of hemoglobin conformations in solution. *Proceedings of the National Academy of Sciences of the United States of America* **58**, 19-&.

Ohgushi, M. & Wada, A. (1983). 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Letters* **164**, 21–24.

Olejniczak, E. T., Dobson, C. M., Karplus, M. & Levy, R. M. (1984). Motional averaging of proton nuclear Overhauser effects in proteins – predictions from a molecular-dynamics simulation of lysozyme. *Journal of American Chemical Society* **106**, 1923–1930.

Ollerenshaw, J. E., Lidar, D. A. & Kay, L. E. (2003). Magnetic resonance realization of decoherence-free quantum computation. *Physics Reviews Letters* **91**, 217904.

Onuchic, J. N., Lutheyschulten, Z. & Wolynes, P. G. (1997). Theory of protein folding: the energy landscape perspective. *Annual Review of Physical Chemistry* **48**, 545–600.

Otten, R., Villali, J., Kern, D. & Mulder, F. A. (2010). Probing microsecond time scale dynamics in proteins by methyl ^{1}H-Carr-Purcell-Meiboom-Gill relaxation dispersion NMR measurements. Application to activation of the signaling protein NtrC(r). *Journal of American Chemical Society* **132**, 17004–17014.

Otting, G. (2010). Protein NMR using paramagnetic ions. *Annual Review of Biophysics* **39**, 387–405.

Ozenne, V., Noel, J. K., Heidarsson, P. O., Brander, S., Poulsen, F. M., Jensen, M. R., Kragelund, B. B., Blackledge, M. & Danielsson, J. (2014). Exploring the minimally frustrated energy landscape of unfolded ACBP. *Journal of Molecular Biology* **426**, 722–734.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

Palmer, A. G. III (2014). Chemical exchange in biomacromolecules: past, present, and future. *Journal of Magnetic Resonance* **241**, 3–17.

Palmer, A. G. III, Grey, M. J. & Wang, C. (2005). Solution NMR spin relaxation methods for characterizing chemical exchange in high-molecular-weight systems. *Methods in Enzymology* **394**, 430–465.

Palmer, A. G. III, Kroenke, C. D. & Loria, J. P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods in Enzymology* **339**, 204–238.

Palmer, A. G. III & Massi, F. (2006). Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chemical Reviews* **106**, 1700–1719.

Pelikan, M., Hura, G. L. & Hammel, M. (2009). Structure and flexibility within proteins as identified through small angle X-ray scattering. *General Physiology and Biophysics* **28**, 174–189.

Peng, J. W. & Wagner, G. (1994). Investigation of protein motions via relaxation measurements. *Methods in Enzymology* **239**, 563–596.

Peterson, D. W., Zhou, H., Dahlquist, F. W. & Lew, J. (2008). A soluble oligomer of tau associated with fiber formation analyzed by NMR. *Biochemistry* **47**, 7393–7404.

Petros, A. M., Mueller, L. & Kopple, K. D. (1990). NMR identification of protein surfaces using paramagnetic probes. *Biochemistry* **29**, 10041–10048.

Pintacuda, G., Kaikkonen, A. & Otting, G. (2004a). Modulation of the distance dependence of paramagnetic relaxation enhancements by CSA × DSA cross-correlation. *Journal of Magnetic Resonance* **171**, 233–243.

Pintacuda, G., Moshref, A., Leonchiks, A., Sharipo, A. & Otting, G. (2004b). Site-specific labelling with a metal chelator for protein-structure refinement. *Journal of Biomolecular NMR* **29**, 351–361.

Pintacuda, G. & Otting, G. (2002). Identification of protein surfaces by NMR measurements with a paramagnetic Gd(III) chelate. *Journal of American Chemical Society* **124**, 372–373.

Pisliakov, A. V., Cao, J., Kamerlin, S. C. L. & Warshel, A. (2009). Enzyme millisecond conformational dynamics do not catalyze the chemical step. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 17359–17364.

Post, C. B. (2003). Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Current Opinion in Structural Biology* **13**, 581–588.

Ramos, A. & Varani, G. (1998). A new method to detect long-range protein–RNA contacts: NMR detection of electron–proton relaxation induced by nitroxide spin-labeled RNA. *Journal of American Chemical Society* **120**, 10992–10993.

Rao, B. D. (1989). Nuclear magnetic resonance line-shape analysis and determination of exchange rates. *Methods in Enzymology* **176**, 279–311.

Redfield, A. G. (1957). On the theory of relaxation processes. *Ibm Journal of Research and Development* **1**, 19–31.

Redfield, C. (2004). Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods* **34**, 121–132.

Roder, H., Elove, G. A. & Englander, S. W. (1988). Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* **335**, 700–704.

Roosild, T. P., Greenwald, J., Vega, M., Castronovo, S., Riek, R. & Choe, S. (2005). NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* **307**, 1317–1321.

Rueda, M., Ferrer-Costa, C., Meyer, T., Perez, A., Camps, J., Hospital, A., Gelpi, J. L. & Orozco, M. (2007). A consensus view of protein dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 796–801.

Rule, G. S. & Hitchens, T. K. (2006). *Fundamentals of Protein NMR Spectroscopy*. Springer, Dordrecht, The Netherlands.

Rumpel, S., Becker, S. & Zweckstetter, M. (2008). High-resolution structure determination of the CylR2 homodimer using paramagnetic relaxation enhancement and structure-based prediction of molecular alignment. *Journal of Biomolecular NMR* **40**, 1–13.

Ruschak, A. M. & Kay, L. E. (2010). Methyl groups as probes of supra-molecular structure, dynamics and function. *Journal of Biomolecular NMR* **46**, 75–87.

Russo, L., Maestre-Martinez, M., Wolff, S., Becker, S. & Griesinger, C. (2013). Interdomain dynamics explored by paramagnetic NMR. *Journal of American Chemical Society* **135**, 17111–17120.

Sahu, D., Clore, G. M. & Iwahara, J. (2007). TROSY-based z-exchange spectroscopy: application to the determination of the activation energy for intermolecular protein translocation between specific sites on different DNA molecules. *Journal of American Chemical Society* **129**, 13232–13237.

Sakakura, M., Noba, S., Luchette, P. A., Shimada, I. & Prosser, R. S. (2005). An NMR method for the determination of protein-binding interfaces using dioxygen-induced spin-lattice relaxation enhancement. *Journal of American Chemical Society* **127**, 5826–5832.

Schanda, P. & Brutscher, B. (2005). Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *Journal of American Chemical Society* **127**, 8014–8015.

Schmidt, P. G. & Kuntz, I. D. (1984). Distance measurements in spin-labeled lysozyme. *Biochemistry* **23**, 4261–4266.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

SCHWARTZ, A. L. & CUTNELL, J. D. (1983). One- and two-dimensional NMR studies of exchanging amide protons in glutathione. *Journal of Magnetic Resonance* **53**, 398–411.

SCHWIETERS, C. D. & CLORE, G. M. (2001). The VMD-XPLOR visualization package for NMR structure refinement. *Journal of Magnetic Resonance* **149**, 239–244.

SCHWIETERS, C. D. & CLORE, G. M. (2002). Reweighted atomic densities to represent ensembles of NMR structures. *Journal of Biomolecular NMR* **23**, 221–225.

SCHWIETERS, C. D., KUSZEWSKI, J. J. & CLORE, G. M. (2006). Using Xplor-NIH for NMR molecular structure determination. *Progress in Nuclear Magnetic Resonance Spectroscopy* **48**, 47–62.

SCHWIETERS, C. D., KUSZEWSKI, J. J., TJANDRA, N. & CLORE, G. M. (2003). The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance* **160**, 65–73.

SEEHOLZER, S. H. & WAND, A. J. (1989). Structural characterization of the interactions between calmodulin and skeletal muscle myosin light chain kinase: effect of peptide (576–594)G binding on the Ca2+-binding domains. *Biochemistry* **28**, 4011–4020.

SEKHAR, A. & KAY, L. E. (2013). NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 12867–12874.

SELENKO, P., SERBER, Z., GADEA, B., RUDERMAN, J. & WAGNER, G. (2006). Quantitative NMR analysis of the protein G B1 domain in *Xenopus laevis* egg extracts and intact oocytes. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11904–11909.

SENGUPTA, I., NADAUD, P. S., HELMUS, J. J., SCHWIETERS, C. D. & JARONIEC, C. P. (2012). Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy. *Nature Chemistry* **4**, 410–417.

SENGUPTA, I., NADAUD, P. S. & JARONIEC, C. P. (2013). Protein structure determination with paramagnetic solid-state NMR spectroscopy. *Accounts of Chemical Research* **46**, 2117–2126.

SHEN, Y. & BAX, A. (2013). Protein backbone and side-chain torsion angles predicted from NMR chemical shifts using artificial neural networks. *Journal of Biomolecular NMR* **56**, 227–241.

SHEN, Y., LANGE, O., DELAGLIO, F., ROSSI, P., ARAMINI, J. M., LIU, G., ELETSKY, A., WU, Y., SINGARAPU, K. K., LEMAK, A., IGNATCHENKO, A., ARROWSMITH, C. H., SZYPERSKI, T., MONTELIONE, G. T., BAKER, D. & BAX, A. (2008). Consistent blind protein structure generation from NMR chemical shift data. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4685–4690.

SHEN, Y., VERNON, R., BAKER, D. & BAX, A. (2009). *De novo* protein structure generation from incomplete chemical shift assignments. *Journal of Biomolecular NMR* **43**, 63–78.

SHENKAREV, Z. O., PARAMONOV, A. S., BALASHOVA, T. A., YAKIMENKO, Z. A., BARU, M. B., MUSTAEVA, L. G., RAAP, J., OVCHINNIKOVA, T. V. & ARSENIEV, A. S. (2004). High stability of the hinge region in the membrane-active peptide helix of zervamicin: paramagnetic relaxation enhancement studies. *Biochemical and Biophysical Research Communications* **325**, 1099–1105.

SHI, L. & KAY, L. E. (2014). Tracing an allosteric pathway regulating the activity of the HslV protease. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 2140–2145.

SILVESTRE-RYAN, J., BERTONCINI, C. W., FENWICK, R. B., ESTEBAN-MARTIN, S. & SALVATELLA, X. (2013). Average conformations determined from PRE data provide high-resolution maps of transient tertiary interactions in disordered proteins. *Biophysical Journal* **104**, 1740–1751.

SKRYNNIKOV, N. R., DAHLQUIST, F. W. & KAY, L. E. (2002). Reconstructing NMR spectra of "invisible" excited protein states using HSQC and HMQC experiments. *Journal of American Chemical Society* **124**, 12352–12360.

SOLOMON, I. (1955). Relaxation processes in a system of two spins. *Physics Reviews* **99**, 559–565.

SOLOMON, I. & BLOEMBERGEN, N. (1956). Nuclear magnetic interactions in the HF molecule. *Journal of Chemical Physics* **25**, 261–266.

SONG, J., GUO, L.-W., MURADOV, H., ARTEMYEV, N. O., RUOHO, A. E. & MARKLEY, J. L. (2008). Intrinsically disordered gamma-subunit of cGMP phosphodiesterase encodes functionally relevant transient secondary and tertiary structure. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1505–1510.

SPERA, S., IKURA, M. & BAX, A. (1991). Measurement of the exchange rates of rapidly exchanging amide protons: application to the study of calmodulin and its complex with a myosin light chain kinase fragment. *Journal of Biomolecular NMR* **1**, 155–165.

SPRANGERS, R. & KAY, L. E. (2007). Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **445**, 618–622.

STAN, G., BROOKS, B. R., LORIMER, G. H. & THIRUMALAI, D. (2006). Residues in substrate proteins that interact with GroEL in the capture process are buried in the native state. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4433–4438.

SUGASE, K., DYSON, H. J. & WRIGHT, P. E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025.

SWIFT, T. J. & CONNICK, R. E. (1962). NMR-relaxation mechanisms of O_{17} in aqueous solutions of paramagnetic cations and lifetime of water molecules in first coordination sphere. *Journal of Chemical Physics* **37**, 307–320.

TAKAYAMA, Y. & CLORE, G. M. (2011). Intra- and intermolecular translocation of the bi-domain transcription

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

factor Oct1 characterized by liquid crystal and paramagnetic NMR. *Proceedings of the National Academy of Sciences of the United States of America* **108**, E169–E176.

Takayama, Y. & Clore, G. M. (2012). Interplay between minor and major groove-binding transcription factors Sox2 and Oct1 in translocation on DNA studied by paramagnetic and diamagnetic NMR. *Journal of Biological Chemistry* **287**, 14349–14363.

Tang, C., Ghirlando, R. & Clore, G. M. (2008a). Visualization of transient ultra-weak protein self-association in solution using paramagnetic relaxation enhancement. *Journal of American Chemical Society* **130**, 4048–4056.

Tang, C., Iwahara, J. & Clore, G. M. (2006). Visualization of transient encounter complexes in protein–protein association. *Nature* **444**, 383–386.

Tang, C., Louis, J. M., Aniana, A., Suh, J. Y. & Clore, G. M. (2008b). Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. *Nature* **455**, 693–696.

Tang, C., Schwieters, C. D. & Clore, G. M. (2007). Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **449**, 1078–1082.

Thirumalai, D. & Lorimer, G. H. (2001). Chaperonin-mediated protein folding. *Annual Review of Biophysics and Biomolecular Structure* **30**, 245–269.

Tjandra, N., Marquardt, J. & Clore, G. M. (2000). Direct refinement against proton-proton dipolar couplings in NMR structure determination of macromolecules. *Journal of Magnetic Resonance* **142**, 393–396.

Todd, A. P., Cong, J. P., Levinthal, F., Levinthal, C. & Hubbell, W. L. (1989). Site-directed mutagenesis of colicin-E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation. *Proteins* **6**, 294–305.

Toledo Warshaviak, D., Khramtsov, V. V., Cascio, D., Altenbach, C. & Hubbell, W. L. (2013). Structure and dynamics of an imidazoline nitroxide side chain with strongly hindered internal motion in proteins. *Journal of Magnetic Resonance* **232**, 53–61.

Trewhella, J. (1992). The solution structures of calmodulin and its complexes with synthetic peptides based on target enzyme binding domains. *Cell Calcium* **13**, 377–390.

Trott, O. & Palmer, A. G. (2004). Theoretical study of R-1p rotating-frame and R-2 free-precession relaxation in the presence of n-site chemical exchange. *Journal of Magnetic Resonance* **170**, 104–112.

Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E. & Kay, L. E. (2003). Cross-correlated relaxation enhanced H-1-C-13 NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *Journal of American Chemical Society* **125**, 10420–10428.

Tugarinov, V., Kanelis, V. & Kay, L. E. (2006). Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature Protocols* **1**, 749–754.

Ubbink, M., Worrall, J. A. R., Canters, G. W., Groenen, E. J. J. & Huber, M. (2002). Paramagnetic resonance of biological metal centers. *Annual Review of Biophysics and Biomolecular Structure* **31**, 393–422.

Udgaonkar, J. B. & Baldwin, R. L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature* **335**, 694–699.

Ueda, T., Kato, A., Ogawa, Y., Torizawa, T., Kuramitsu, S., Iwai, S., Terasawa, H. & Shimada, I. (2004). NMR study of repair mechanism of DNA photolyase by FAD-induced paramagnetic relaxation enhancement. *Journal of Biological Chemistry* **279**, 52574–52579.

Vallurupalli, P., Bouvignies, G. & Kay, L. E. (2011). Increasing the exchange time-scale that can be probed by CPMG relaxation dispersion NMR. *Journal of Physical Chemistry B* **115**, 14891–14900.

Vallurupalli, P., Bouvignies, G. & Kay, L. E. (2012). Studying "invisible" excited protein states in slow exchange with a major state conformation. *Journal of American Chemical Society* **134**, 8148–8161.

Vallurupalli, P., Bouvignies, G. & Kay, L. E. (2013). A computational study of the effects of (13) C-(13) C scalar couplings on (13) C CEST NMR spectra: towards studies on a uniformly (13) C-labeled protein. *Chembiochem* **14**, 1709–1713.

Vallurupalli, P., Hansen, D. F. & Kay, L. E. (2008a). Probing structure in invisible protein states with anisotropic NMR chemical shifts. *Journal of American Chemical Society* **130**, 2734–2735.

Vallurupalli, P., Hansen, D. F. & Kay, L. E. (2008b). Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 11766–11771.

Vallurupalli, P., Hansen, D. F., Stollar, E., Meirovitch, E. & Kay, L. E. (2007). Measurement of bond vector orientations in invisible excited states of proteins. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18473–18477.

Vallurupalli, P. & Kay, L. E. (2006). Complementarity of ensemble and single-molecule measures of protein motion: a relaxation dispersion NMR study of an enzyme complex. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11910–11915.

Vallurupalli, P. & Kay, L. E. (2013). Probing slow chemical exchange at carbonyl sites in proteins by chemical exchange saturation transfer NMR spectroscopy. *Angewandte Chemie International Edition English* **52**, 4156–4159.

Van Heel, M., Gowen, B., Matadeen, R., Orlova, E. V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M. & Patwardhan, A. (2000). Single-particle

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

electron cryo-microscopy: towards atomic resolution. *Quarterly Reviews of Biophysics* **33**, 307–369.

Van Zijl, P. C. & Yadav, N. N. (2011). Chemical exchange saturation transfer (CEST): what is in a name and what isn't? *Magnetic Resonance in Medicines* **65**, 927–948.

Varani, L., Gunderson, S. I., Mattaj, I. W., Kay, L. E., Neuhaus, D. & Varani, G. (2000). The NMR structure of the 38 kDa U1A protein – PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein. *Nature Structural Biology* **7**, 329–335.

Venditti, V., Fawzi, N. L. & Clore, G. M. (2011). Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear Overhauser enhancement spectroscopy. *Journal of Biomolecular NMR* **51**, 319–328.

Venditti, V., Fawzi, N. L. & Clore, G. M. (2012). An efficient protocol for incorporation of an unnatural amino acid in perdeuterated recombinant proteins using glucose-based media. *Journal of Biomolecular NMR* **52**, 191–195.

Villali, J. & Kern, D. (2010). Choreographing an enzyme's dance. *Current Opinion in Structural Biology* **14**, 636–643.

Vinogradov, E., Sherry, A. D. & Lenkinski, R. E. (2013). CEST: from basic principles to applications, challenges and opportunities. *Journal of Magnetic Resonance* **229**, 155–172.

Vise, P., Baral, B., Stancik, A., Lowry, D. F. & Daughdrill, G. W. (2007). Identifying long-range structure in the intrinsically unstructured transactivation domain of p53. *Proteins* **67**, 526–530.

Vlasie, M. D., Comuzzi, C., Van Den Nieuwendijk, A. M. C. H., Prudencio, M., Overhand, M. & Ubbink, M. (2007). Long-range-distance NMR effects in a protein labeled with a lanthanide-DOTA chelate. *Chemistry* **13**, 1715–1723.

Volkov, A. N., Ubbink, M. & Van Nuland, N. A. J. (2010). Mapping the encounter state of a transient protein complex by PRE NMR spectroscopy. *Journal of Biomolecular NMR* **48**, 225–236.

Volkov, A. N., Worrall, J. A. R., Holtzmann, E. & Ubbink, M. (2006). Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 18945–18950.

Wagner, G. & Wuthrich, K. (1982). Amide protein exchange and surface conformation of the basic pancreatic trypsin inhibitor in solution. Studies with two-dimensional nuclear magnetic resonance. *Journal of Molecular Biology* **160**, 343–361.

Walsh, D. M. & Selkoe, D. J. (2007). A beta oligomers – a decade of discovery. *Journal of Neurochemistry* **101**, 1172–1184.

Wand, A. J. & Englander, S. W. (1996). Protein complexes studied by NMR spectroscopy. *Current Opinion in Biotechnology* **7**, 403–408.

Wang, L., Brock, A., Herberich, B. & Schultz, P. G. (2001). Expanding the genetic code of Escherichia coli. *Science* **292**, 498–500.

Wang, X., Srisailam, S., Yee, A. A., Lemak, A., Arrowsmith, C., Prestegard, J. H. & Tian, F. (2007). Domain-domain motions in proteins from time-modulated pseudocontact shifts. *Journal of Biomolecular NMR* **39**, 53–61.

Ward, K. M., Aletras, A. H. & Balaban, R. S. (2000). A new class of contrast agents for MRI based on proton chemical exchange dependent saturation transfer (CEST). *Journal of Magnetic Resonance* **143**, 79–87.

Watt, E. D., Shimada, H., Kovrigin, E. L. & Loria, J. P. (2007). The mechanism of rate-limiting motions in enzyme function. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 11981–11986.

Weininger, U., Blissing, A. T., Hennig, J., Ahlner, A., Liu, Z., Vogel, H. J., Akke, M. & Lundstrom, P. (2013). Protein conformational exchange measured by 1H R1rho relaxation dispersion of methyl groups. *Journal of Biomolecular NMR* **57**, 47–55.

Wickramasinghe, N. P., Kotecha, M., Samoson, A., Past, J. & Ishii, Y. (2007). Sensitivity enhancement in (13)C solid-state NMR of protein microcrystals by use of paramagnetic metal ions for optimizing (1)H T(1) relaxation. *Journal of Magnetic Resonance* **184**, 350–356.

Wider, G., Neri, D. & Wuthrich, K. (1991). Studies of slow conformational equilibria in macromolecules by exchange of heteronuclear longitudinal 2-spin-order in a 2D difference correlation experiment. *Journal of Biomolecular NMR* **1**, 93–98.

Woessner, D. E. (1962). Spin relaxation processes in a two-proton system undergoing anisotropic reorientation. *Journal of Chemical Physics* **36**, 1–4.

Wohnert, J., Franz, K. J., Nitz, M., Imperiali, B. & Schwalbe, H. (2003). Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. *Journal of American Chemical Society* **125**, 13338–13339.

Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z. & Kern, D. (2004). Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nature Structural and Molecular Biology* **11**, 945–949.

Wolynes, P. G. (2005). Recent successes of the energy landscape theory of protein folding and function. *Quarterly Review of Biophysics* **38**, 405–410.

Wright, P. E. & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology* **19**, 31–38.

Xia, J., Deng, N. J. & Levy, R. M. (2013). NMR relaxation in proteins with fast internal motions and slow conformational exchange: model-free framework and Markov

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press

state simulations. *Journal of Physical Chemistry B* **117**, 6625–6634.

YAMNIUK, A. P. & VOGEL, H. J. (2004). Calmodulin's flexibility allows for promiscuity in its interactions with target proteins and peptides. *Molecular Biotechnology* **27**, 33–57.

YAN, S., GAWLAK, G., SMITH, J., SILVER, L., KOIDE, A. & KOIDE, S. (2004). Conformational heterogeneity of an equilibrium folding intermediate quantified and mapped by scanning mutagenesis. *Journal of Molecular Biology* **338**, 811–825.

YAN, S., KENNEDY, S. D. & KOIDE, S. (2002). Thermodynamic and kinetic exploration of the energy landscape of Borrelia burgdorferi OspA by native-state hydrogen exchange. *Journal of Molecular Biology* **323**, 363–375.

ZHANG, Z. & SMITH, D. L. (1993). Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Science* **2**, 522–531.

ZHOU, Z. H. (2008). Towards atomic resolution structural determination by single-particle cryo-electron microscopy. *Current Opinion in Structural Biology* **18**, 218–228.

ZHUANG, T., CHISHOLM, C., CHEN, M. & TAMM, L. K. (2013). NMR-based conformational ensembles explain pH-gated opening and closing of OmpG channel. *Journal of American Chemical Society* **135**, 15101–15113.

https://doi.org/10.1017/S0033583514000122 Published online by Cambridge University Press