Review

Recent excitements in protein NMR: Large proteins and biologically relevant dynamics

Sai Chaitanya Chiliveri and Mandar V Deshmukh*

CSIR-Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007, India

**Corresponding author (Email: mvdesh@ccmb.res.in)*

The advent of Transverse Relaxation Optimized SpectroscopY (TROSY) and perdeuteration allowed biomolecular NMR spectroscopists to overcome the size limitation barrier (~20 kDa) in *de novo* structure determination of proteins. The utility of these techniques was immediately demonstrated on large proteins and protein complexes (e.g. GroEL-GroES, ClpP protease, Hsp90-p53, 20S proteasome, etc.). Further, recent methodological developments such as Residual Dipolar Couplings and Paramagnetic Relaxation Enhancement allowed accurate measurement of long-range structural restraints. Additionally, Carr-Purcell-Meiboom-Gill (CPMG), rotating frame relaxation experiments ($R_1\rho$) and saturation transfer experiments (CEST and DEST) created never-before accessibility to the μs–ms timescale dynamic parameters that led to the deeper understanding of biological processes. Meanwhile, the excitement in the field continued with a series of developments in the fast data acquisition methods allowing rapid structural studies on less stable proteins. This review aims to discuss important developments in the field of biomolecular NMR spectroscopy in the recent past, i.e., in the post TROSY era. These developments not only gave access to the structural studies of large protein assemblies, but also revolutionized tools in the arsenal of today's biomolecular NMR and point to a bright future of biomolecular NMR spectroscopy.

[Chiliveri SC and Deshmukh MV 2016 Recent excitements in protein NMR: Large proteins and biologically relevant dynamics. *J. Biosci.* **41** 787–803]

1. Introduction to deuteration and TROSY

Solution NMR spectroscopy is an ideal choice for proteins that are difficult to crystallize and warrant information about dynamics. In a conventional protein structure determination project, the first step would be the use of triple-resonance backbone and side-chain chemical shift assignment experiments. Later the data derived from several heteronuclear edited NOESY spectrum allow assignments of distances leading to the solution structure. These methods have been quite promising for gaining structural information on proteins up to molecular weight of ~20 kDa. Until recently, the study of large macromolecules by conventional biomolecular NMR spectroscopy posed two problems, namely, larger number of residues proportionately increase resonances in the spectra, leading to severe resonance overlap, and increase in the molecular weight linearly enhances the transverse relaxation rates (R_2) of the molecule, leading to broad signals and poor spectral quality. These problems severely limit NMR studies of large molecules, which is directly reflected in a fewer number of NMR-derived protein structures with molecular weights above 20 kDa. While the issue of resonance overlap can be overcome either by selective labelling of amino acids or by increasing the dimensionality of NMR data collection. However, increase in R_2 poses a more severe technical challenge. As the R_2 of ^{13}C and ^{1}H increases with molecular weight, less magnetization survives through pulse sequence delays, leading to poor sensitivity of the spectrum. Hence, standard triple-resonance experiments are ineffective for larger proteins. Dense networks of hydrogen atoms are the major sources of dipolar relaxation and their replacement with deuterium substantially increases spectral resolution. Significant improvement in resolution and

Keywords. Biomolecular NMR; CEST; CPMG; DEST; methyl TROSY; PRE; RDCs; relaxation dispersion; TROSY

Supplementary materials pertaining to this article are available on the Journal of Biosciences Website.

DOI: 10.1007/s12038-016-9640-y

Published online: 14 October 2016

sensitivity can be obtained by decreasing the relaxation rates of ^{13}C and ^{1}H by replacing the aliphatic protons by deuterium (^{2}H). Overall, 16-fold less effective dipolar relaxation can be observed by replacing ^{1}H with ^{2}H (equation SE1).

Deuteration was used in the late 1960s to reduce the complexity of ^{1}H NMR spectra (Crespi *et al.* 1968; Markley *et al.* 1968). Deuteration was first demonstrated for triple resonance in 4D HN(COCA)NH experiment by Bax and co-workers in the year 1993 (Grzesiek *et al.* 1993). Two types of labelling schemes are followed for deuteration, (i) site-independent deuteration (uniform or random deuteration), and (ii) site-specific deuteration (Gardner and Kay 1998). In both cases, protons can be re-introduced at the amide ($^{1}H^{N}$) position by dissolving the protein in buffer containing H_2O, which allows back exchange of labile protons. Replacement of protons by deuterium causes isotopic effects on nearby nuclei. These effects are cumulative, and for sidechain carbons that are surrounded by many ^{2}H's, a shift of up to 1 ppm can be observed. For backbone atoms, shifts in the range of -0.3 ppm and -0.5 ppm were observed for ^{15}N and $^{13}C\alpha$ atoms, respectively (Venters *et al.* 1996; Gardner *et al.* 1997). However, deuteration alone cannot extend the size limit beyond 50 kDa.

The magnetic field fluctuations by the field generated from dipole-dipole (DD) coupled spin and chemical shift anisotropy (CSA) of the individual spins drive spin relaxation. TROSY (Transverse Relaxation Optimized SpectroscopY) implements the interference of DD and CSA, which results in decreasing the transverse relaxation rate of the spin, leading to shaper linewidths (Pervushin *et al.* 1997; Pervushin 2000). Consider a ^{1}H-^{15}N spin pair, when the ^{1}H is in β spin state, the contributions of the CSA and the ^{1}H magnetic moment on the ^{15}N nucleus always have opposite sign, irrespective of the orientation of the molecule. In contrast, when ^{1}H is in α spin state, both CSA and DD (^{1}H magnetic moment) on the ^{15}N nucleus always have the same sign. While DD interaction is field independent, CSA increases as the square of the magnetic field. Thus, the optimal TROSY effect can be obtained by choosing the appropriate field strength, where the cancellation of CSA and dipolar fields for the nitrogen spin in a proton-nitrogen group would be maximum. For N-H coupled spin pair, an optimum TROSY effect can be observed at 23.5 T (1000 MHz) (Wüthrich 1998). The significance of TROSY with deuteration effect is illustrated in figure 1. In ^{1}H-^{15}N Hetereonuclear Single Quantum Coherence (HSQC), resonances arising from ^{15}N RDE-4 are broad due to the dense network of background protons, while the combination of deuteration with TROSY resulted in sharper linewidths.

In a heteronuclear, ^{1}H-^{15}N spin systems, the NMR signal is split into two components, separated by scalar coupling constant ($^{1}J_{NH}$). Hence, in a 2D ^{1}H-^{15}N coupled-HSQC, one can observe a single peak splitting into four components (two in each dimension), represented as four red cross-peaks in figure 2. A glance at these four cross-peaks clearly

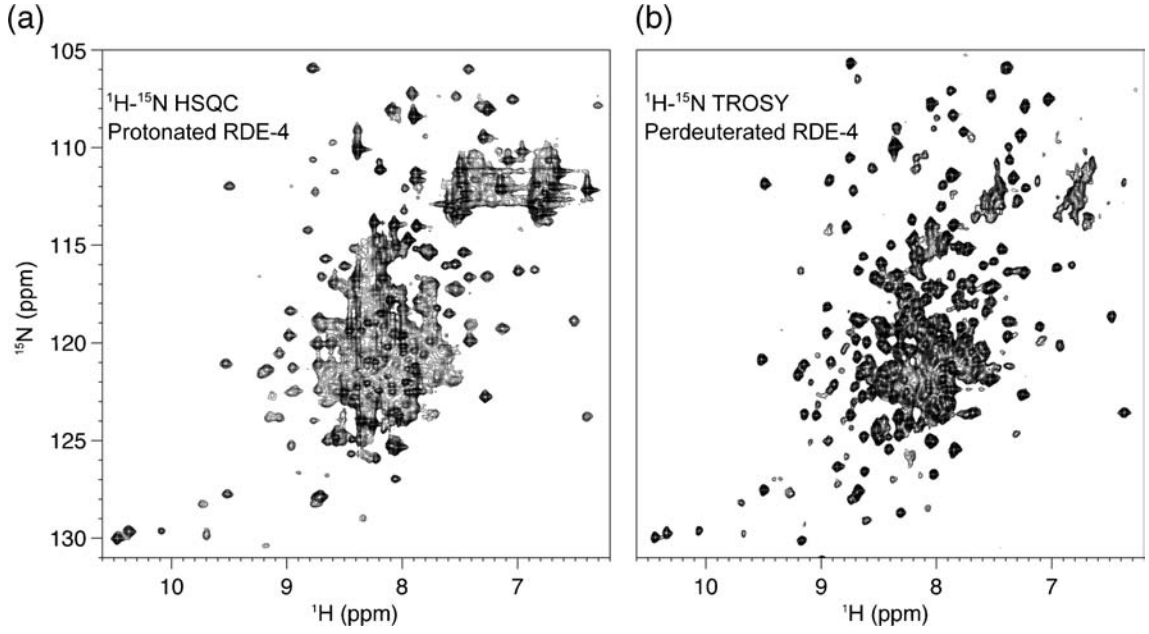

Figure 1. ^{1}H-^{15}N chemical shift correlation of the amide fingerprint region in RDE-4 homodimer (~88 kDa) using HSQC and TROSY methods. (**a**) ^{1}H-^{15}N HSQC of protonated sample exhibits severe broadening for the majority of resonances due to faster relaxation. (**b**) ^{1}H-^{15}N TROSY of RDE-4 with perdeuteration resulting in sharper resonances. Many new cross-peaks can be noticed. The crowding of the central region is deceptive and occurs due to low counter levels to represent broad peaks in other areas. The data was collected under identical conditions at RT using a 600 MHz NMR spectrometer equipped with a cryoprobe.

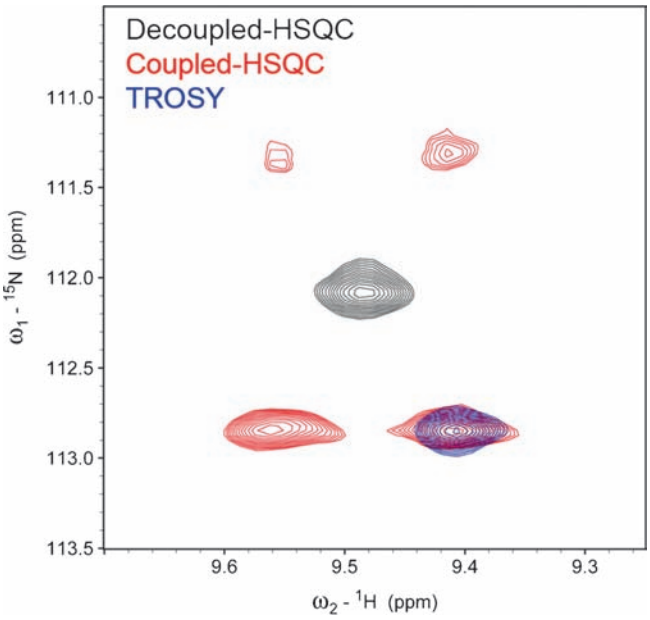

Figure 2. Excerpts from ^{1}H-^{15}N chemical shift correlation spectra. A single amide N-H spin pair is represented under three different experimental conditions: decoupled HSQC (black), coupled HSQC (red) and TROSY (blue). For Large MW proteins, the slowest relaxing TROSY component is preferred due to its narrow linewidth.

indicates that the right bottom cross-peak is sharper than the remaining three. For smaller proteins (<20 kDa) it is advantageous to average all four components despite different relaxation rates of multiple components to achieve enhanced sensitivity and to reduce spectral crowding (black cross-peaks in figure 2). However, for large proteins (>20 kDa), averaging differential transverse relaxation rates result in severe line broadening for the averaged peak. Using TROSY approach, the signal possessing narrow linewidth is selected by employing appropriate phase cycling under the evolution of the coupling (blue cross-peak in figure 2). Development of TROSY and its better versions, i.e., clean-TROSY which utilizes spin-state selective excitation, led to the understanding of several large protein complexes and their structural assemblies that were not previously amenable to solution NMR or X-ray crystallography (Schulte-Herbrüggen and Sørensen 2000). Lastly, the utility of TROSY based selection was extended to many triple resonance experiments commonly used for the chemical shift assignment of backbone and side-chain resonances (Loria *et al.* 1999; Salzmann *et al.* 1999).

2. Residual dipolar couplings

Structure determination by NMR heavily relies on dipolar relaxation, which manifests the Nuclear Overhauser Enhancement. Nonetheless, the NOESY effect rapidly succumbs to the larger proton network in high molecular weight proteins. Moreover, internal motions in the macromolecule allow averaging of NOESY cross-peaks, reflecting only an average conformation of the protein structure. These inherent limitations of NOE provide local information (measurement of inter-proton distances only up to <5Å) and do not carry long-range information. The following problems may be difficult to address without long-range information.

1. If there are few or no NOE contacts between the two non-parallel helices, what would be the angle between them?
2. Is a long structured α helix straight or bent or possess a kink?
3. What is the relative orientation between domains in a multi-domain protein?

Some of these issues can be tackled by measurement of dipolar coupling between two nuclei under tailored alignment. Unlike scalar couplings, which are mediated through bonds, dipolar coupling allows nuclear spins to interact through space. Dipolar coupling of an inter-atomic vector can be used to calculate its orientation with the applied magnetic field. As the orientational restraints are distance independent the above-mentioned problems can be addressed with dipolar couplings (equation SE2).

Due to isotropic tumbling of molecules, dipolar couplings are averaged to zero in solution. Residual dipolar couplings (RDCs) can be created by using a weak alignment either by inherent high magnetic susceptibility of proteins, which allows them to align in an external magnetic field (Tolman *et al.* 1995; Tjandra *et al.* 1997), or by inducing anisotropy in protein solution using external alignment media (Tjandra and Bax 1997). RDCs induced by the former method are very weak and need invariantly higher magnetic fields, whereas the latter is tunable and can induce at least 10-fold higher RDCs.

Over time, numerous alignment media were introduced (Prestegard *et al.* 2004). Of them, bicelles and filamentous Pf1 phages are widely used for biomolecules. Dimyristoyl phosphatidylcholine (DMPC), dihexanoyl phosphatidylcholine (DHPC) and hexadecyl (cetyl) trimethyl ammonium bromide (CTAB) at molar ratios of 3:1:0.1 form bicelles, and weak alignment can be achieved over narrow concentration and temperature in the magnetic field (Losonczi and Prestegard 1998; Ottiger and Bax 1998). Macromolecules attain a degree of alignment due to steric and electrostatic interactions with the bicelle phase. The degree of alignment can be adjusted by the concentration of bicelles. The Pf1 filamentous bacteriophage has been widely accepted for alignment due to its stability over wide range of concentrations, pH, temperature, and salt conditions (Hansen *et al.* 1998). The full solution of the traceless alignment tensor can only be achieved by five independent RDCs per bond vector in different alignment media or availability of five-bond

vectors per residue in a single alignment media (Kramer *et al.* 2004).

RDCs are measured as the difference in J-couplings under isotropic and anisotropic conditions and can be directly measured as J-couplings from heteronuclear edited spectra or using spin-state separation, such as IPAP, S^3E, which allow resolution of crowded spectra or by the use of E.COSY experiments to resolve small couplings between a pair of spins through a larger couplings between another pair of spins. A large variation in the experimental techniques was reviewed by Prestegard *et al.* (2004).

RDCs proved to be a highly valuable tool to biomolecular NMR spectroscopy as it directly provided orientation of all possible bond angles in the principle alignment frame. RDCs have been a huge success and have provided answers to many inconsistent structures (Skrynnikov *et al.* 2000; Deshmukh *et al.* 2006; Mackereth *et al.* 2011). However, RDCs also suffer from molecular averaging and hence, RDCs through external alignment alone are often not sufficient for solution structure determination of large and multi domain proteins.

3. Methyl TROSY

Although deuteration of aliphatic carbons resulted in enhanced sensitivity and spectral resolution, concomitant reduction in NOE restraints due to the lack of side chain protons, poses serious challenges for structure calculations. Towards this, a strategy involving measurement of NOEs emanating from methyl protons of Ile (δ1), Leu and Val (henceforth, ILV) was developed by Lewis E. Kay's group (Tugarinov *et al.* 2003). The idea of Methyl-TROSY revolves around the reintroduction of protons at non-exchangeable sites in a highly deuterated sample. The specific methyl protonation strategy for ILV residues appears to be advantageous as the methyl groups have relatively longer relaxation times allowing retention of excited states for efficient NOE transfers. Hence, methyl spectra of high molecular weight systems yield good resolution. Secondly, distances derived from methyl proton based NOESY experiments provide valuable structural restraints as methyl group of ILVs are often confined to the hydrophobic core of proteins. Thirdly, effective use of labelling patterns preserves many of the important features of perdeuteration while maintaining a critical number of protons at ILV side chain positions for further structural and dynamics studies. Lastly, improved sensitivity and reduced spectral overlap of ILV methyl labelling offers an apt replacement of the complete side chain chemical shift assignment for high molecular weight proteins, which is often tedious and may fail due to issues involving poor sensitivity and resolution. Following section briefly describes the basic concept of the Methyl-TROSY experiment.

Large ^1H-^1H and ^1H-^{13}C dipolar interactions dominate relaxation mechanism in CH_3 spin systems. In comparison to HMQC, conventional ^1H-^{13}C HSQC has additional pulses that transform multiple quantum coherence terms into single quantum, however, it leads to mixing of the slow and fast relaxing components in CH_3 spin system, yielding poor sensitivity. HMQC pulse sequence (when coupled in ^1H dimension during ^{13}C chemical shift evolution) does not mix the fast and slow mixing terms, thus acts as TROSY for methyl groups of large molecular weight proteins.

For small molecules with τ_c in sub-nanosecond range, ^1H-coupled ^{13}C HSQC results in quartet peaks with an intensity ratio of 4:1:1:4 (figure 3a). In contrast, for higher molecular weight proteins, the outer components appear as broad and the inner lines appear sharp. The effect is more pronounced with increasing molecular weights of the protein (figure 3a). The average of all four components would result in a low intensity signal upon decoupling with significant loss of the chemical shift information. ^1H-coupled ^{13}C HMQC displays triplet peaks with intensity ratios of 1:2:1 (figure 3b). For higher molecular weight proteins, outer components become significantly broad with respect to the central component. Hence, ^1H-coupled ^{13}C HMQC would result in better quality of the spectrum, with improvement in the signal to noise ratio of ~3 as was observed in comparison to ^1H-coupled ^{13}C HSQC experiment (Tugarinov *et al.* 2003, 2004).

Methyl TROSY allows sidechain chemical shift assignment of large proteins and has been used in many systems till date (Tugarinov and Kay 2003; Sprangers and Kay 2007; Tzeng and Kalodimos 2013; Shi and Kay 2014). In fact, emerging areas like proton-less protein NMR, where the chemical shifts of only heavy residues are assigned, depend on methyl TROSY for the chemical shift assignment of selective ILV protons so as to allow key distance measurement by NOESY for the determination of *de novo* structures. One of the major deterrents in the high molecular structure determination using this strategy is its inability to gain structural refinements in regions lacking ILVs. Additionally, ILV based distances are also limited to <5Å thus limiting its utility in multi-domain systems. A solution to this problem came from a complementary technique, the electron spin resonance, that depends on the presence of at least one unpaired electron.

4. Paramagnetic NMR

Paramagnetism arises from unpaired electrons. As the magnetic moment of electrons is ~1000 times more than protons, the presence of an unpaired electron near the nucleus causes pronounced paramagnetic effects; leading to dramatic outcome of NMR spectra. As, NMR restraints are sparse for higher molecular weight proteins due to deuteration and signal overlap, Residual Dipolar Couplings (RDCs) and

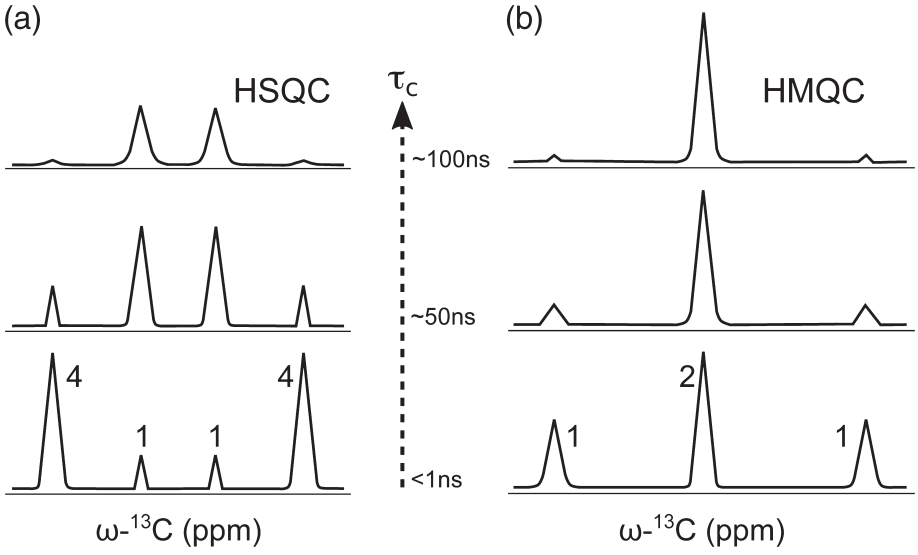

Figure 3. Effect of methyl TROSY. Comparison of 1D ^{13}C strip from ^{1}H-^{13}C correlation spectra of (**a**) ^{1}H coupled ^{1}H-^{13}C HSQC and (**b**) ^{1}H-^{13}C HMQC as a function of rotational correlation time (τ_c). For small molecules (τ_c<1 ns), quartet of 4:1:1:4 intensity ratio is observed in ^{1}H-^{13}C HSQC experiment, whereas, HMQC results in a triplet with 1:2:1 intensity ratio. In ^{1}H-^{13}C HSQC, the split in the slow relaxing components (inner lines) leads to spectral complication and line broadening. In ^{1}H-^{13}C HMQC, the fast relaxing (outer resonances) components broaden beyond detection with increasing molecular weights and the single slow relaxing (inner line) component remains as sharp resonance. Hence, for higher molecular weight proteins ^{1}H-^{13}C HMQC provides better quality spectra with an enhanced signal to noise ratio of ~3 times in comparison to ^{1}H-^{13}C HSQC experiment.

paramagnetic relaxation enhancement (PRE) together offer crucial information about orientation and long-range distances, respectively (Battiste and Wagner 2000; Pintacuda *et al.* 2007; Marius Clore and Iwahara 2009; Madl *et al.* 2010; Göbl *et al.* 2014).

In general, paramagnetic tags are categorized into two types, depending on its magnetic susceptibility tensor, χ (three principal axes: χ_x, χ_y and χ_z). For those paramagnetic centers, where $\chi_x = \chi_y = \chi_z$, the χ tensor is said to be isotropic. For anisotropic χ tensors, at least one of the principal axes is different from others. Paramagnetic centers with very little anisotropy cause extensive line broadening (paramagnetic relaxation enhancement, PRE), whereas the paramagnetic centers with high anisotropy cause extensive shifts (pseudocontact shifts, PCSs) in the resonances as well as weak alignment of the molecules in a strong magnetic field (residual dipolar couplings, RDCs). Also, electronic relaxation times, τ_s, are observed in the range of ns-μs and ps-ns, for isotropic and anisotropic paramagnetic centers, respectively.

4.1 *Paramagnetic relaxation enhancement*

PRE arises from the dipolar interactions between the nucleus and the paramagnetic electron center. Unlike NOEs, the effect of PRE is large and distance restraints of up to 30Å could be obtained. The presence of unpaired electron enhances the relaxation rates of the nearby nuclei (figure 4). Solomon-Bloembergen (SB) describes the longitudinal (Γ_1) and transverse (Γ_2) PRE rates as provided in equation SE3-5(Bloembergen and Morgan 1961; Bertini *et al.* 2001):

As seen from equations (equation SE3 and SE4), relaxation rates are directly proportion to the square of magnetic moment of the nuclei. Hence, in comparison to ^{1}H, PRE effects are 16 and 100 times less for ^{13}C and ^{15}N nuclei, respectively. Paramagnetic centers that induce PRE are nitroxyl radicals (MTSL), Mn^{2+}, Gd^{3+}, Cu^{2+}, etc.

4.2 *Pseudocontact shifts*

Paramagnetic centers with anisotropic magnetic susceptibility tensor exhibit PCSs. Examples include Co^{2+} and all the lanthanide ions except La^{3+}, Lu^{3+}, and Gd^{3+}. The paramagnetic shifts arise from both through-bond (Fermi contact shifts, δ_c) and through-space (PCS, δ_{PCS}) interactions (figure 4). As the δ_c is through-bond, its magnitude drops significantly with increasing covalent bonds, and the effect is local. Hence, contact shifts are not really helpful in obtaining structural information over large distances. PCS arise due to the dipolar coupling between the time-averaged, anisotropic unpaired electron and the magnetic moment of the nucleus. In contrast to contact shifts, PCS are long-range and depend on the polar coordinates r, θ, and ϕ of the nuclear spin with

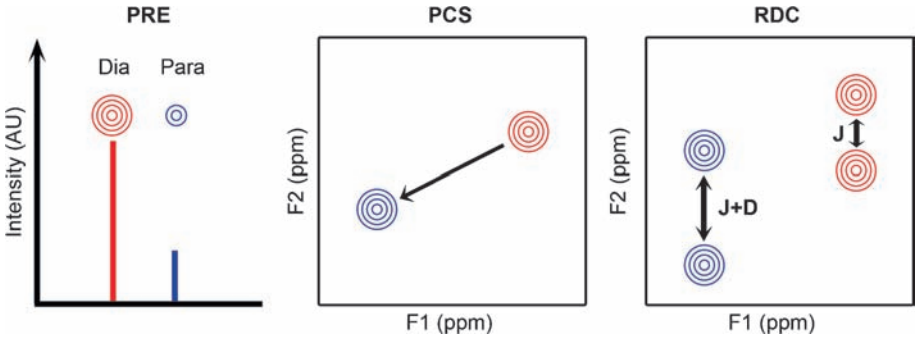

Figure 4. Restraints observed from paramagnetic NMR. Left panel: Reduction in the peak intensity is observed for the resonances that are close by isotropic paramagnetic center. With anisotropic paramagnetic centers, both PCS (middle panel) and RDCs (right panel) are observed. Peaks with red and blue represent resonances in diamagnetic and paramagnetic conditions, respectively. Figure adapted and modified with permission from Clore and Iwahara (2009).

respect to the $\Delta\chi$ tensor of the anisotropic electron center (equation SE6 and SE7) (Bertini *et al.* 2002):

Unlike PRE, PCS exhibits r^{-3} distance dependency and therefore, for lanthanide metal ions, PCS effects can be seen up to 60Å from the paramagnetic center, making PCS as the important structural restraints for large proteins and protein complexes (Ubbink *et al.* 1998; Allegrozzi *et al.* 2000).

4.3 *Residual dipolar couplings (internal alignment)*

Paramagnetic centers with non-vanishing $\Delta\chi$ tensor or anisotropic g-tensor, exhibit partial alignment in the presence of high magnetic field strength, giving rise to residual dipolar coupling (RDCs) (figure 4). In contrast to the external alignment media, RDCs induced by lanthanides in multi-domain or protein complex systems, provides direct information about the alignment of the observed domain with respect to the tagged domain (Bertini *et al.* 2004; Xu *et al.* 2009). The RDC induced on a spin system A-B (D_{AB}) is given by equation SE8 (Bertini *et al.* 2002):

Biologically important relative orientations of individual domains in a multi-domain system are difficult to obtain unless the domains sample a single conformation with close spatial proximity. In reality, a large number of multi-domain proteins undergo conformation sampling of domain orientations that proves to be advantageous to exhibit multiple functions. RDCs obtained from external alignment media orients individual domains, and thus the partial orientation observed for one domain is not transferred to the other, but it is mainly induced by the interaction of domains with the alignment media. In another case, RDCs induced by internal alignment (RDCs induced by lanthanide/paramagnetic centers) are valuable source of information for the investigation of the relative orientation of the individual domains. For an instance, consider a multi-domain protein with domains A and B, where the paramagnetic ion is tagged to domain A (figure 5). The left-hand panel of figure 5 represents a rigid system, while the right-hand panel displays a flexible system. When both the domains have a fixed orientation with respect to each other, a wide dispersion of RDCs will be observed for both domains. In case of a flexible system, wide dispersion of RDCs will be seen only for domain A, as it is internally aligned with paramagnetic ion, whereas, the observed RDC distribution would be narrow for domain B, as it experiences random averaging due to motions leading to almost isotropic orientation with respect to domain A. This approach was applied to probe the orientation of Calmodulin domains in the presence and absence of its interacting partner, α-Synuclein (Bertini *et al.* 2007).

4.4 *Methods for inducing paramagnetic spin labelling*

A wide variety of spin labels are available for paramagnetic NMR studies (see reviews: (Keizers and Ubbink 2011; Hubbell *et al.* 2013; Jeschke 2013)). Paramagnetic spin labels are broadly classified into two types: (i) spin labels containing nitroxide group, and (ii) chelators for a paramagnetic cation (e.g. Mn^{2+}, Cu^{2+}, and Gd^{3+}).

Nitroxide groups are the most widely used spin labels due to their relatively stable radicals at physiological conditions (Griffith and McConnell 1966). Site-specific spin labelling is the common method employed for spin labelling, where a paramagnetic tag is conjugated to a surface exposed cysteine residues in a protein of interest (Todd *et al.* 1989; Hubbell *et al.* 2000). For acquiring specificity, native cysteines (if any) present in the protein must be mutated and a single surface exposed cysteine should be introduced by site-directed mutagenesis (figure 6).

In the case of metalloproteins, paramagnetic ions can be exchanged against natural cation for inducing paramagnetic

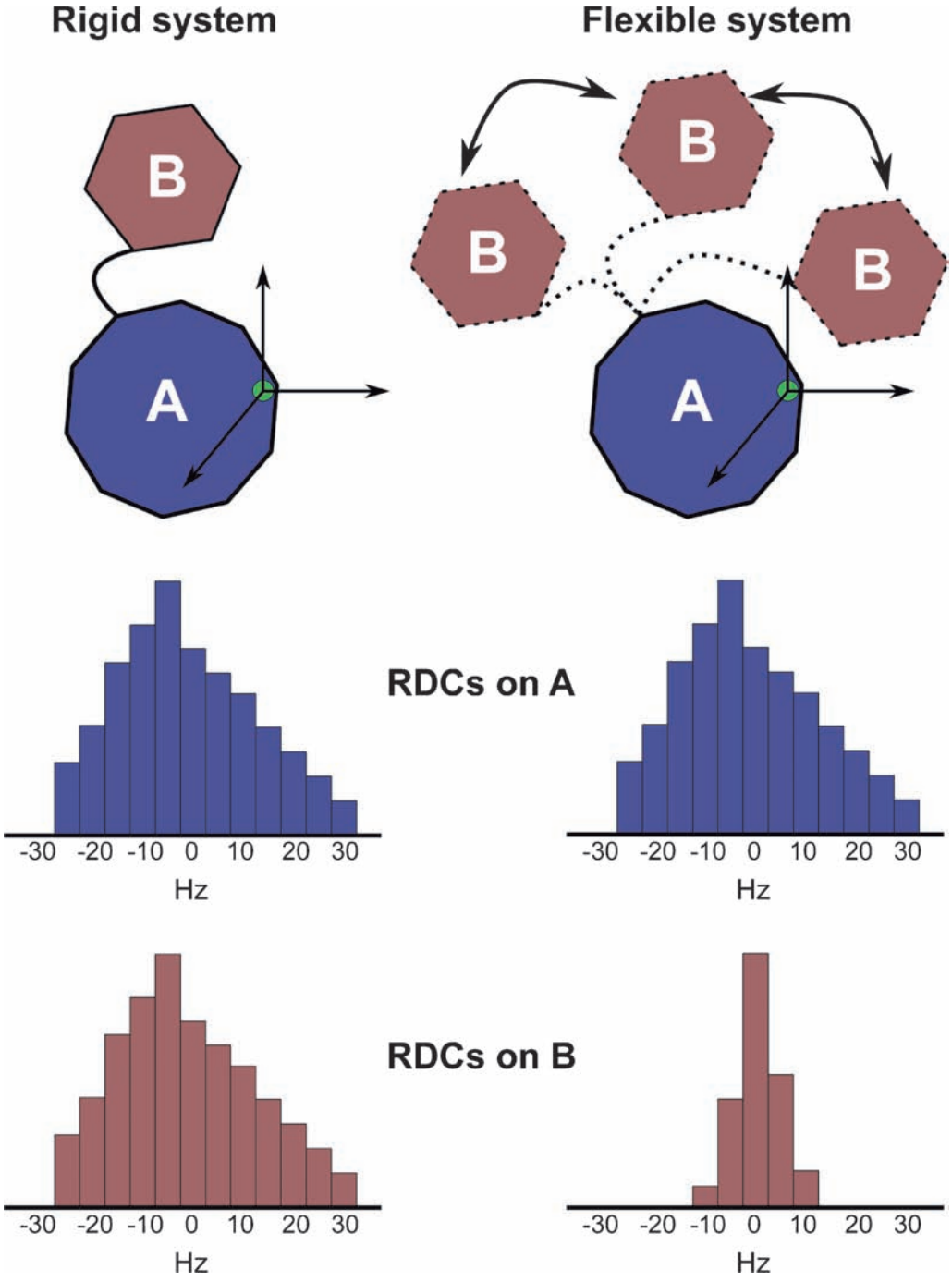

Figure 5. Lanthanide induced RDCs in multi-domain system. A multi-domain protein is shown in left and right panels as rigid and flexible systems, respectively. The green circle with axes represent internal alignment on domain A. RDCs induced on domain A in both conditions is similar. RDCs induced on domain B under flexible system are low. RDC values portrayed in the figure are simulated. Figure inspired from Bertini *et al.* (2008).

effects. For non-native metalloproteins, short metal-binding sequences were genetically encoded to bind lanthanides or paramagnetic cations (Donaldson *et al.* 2001; Wöhnert *et al.* 2003; Jensen *et al.* 2004). For surface immobilization or protein attachment, metal chelators like ethylenediamine tetra-acetic acid (EDTA), ethylene glycol tetraacetic acid (EGTA), diethylenetriamine pentaacetic acid (DTPA) and tetraazacyclododecane tetraacetic acid (DOTA) were

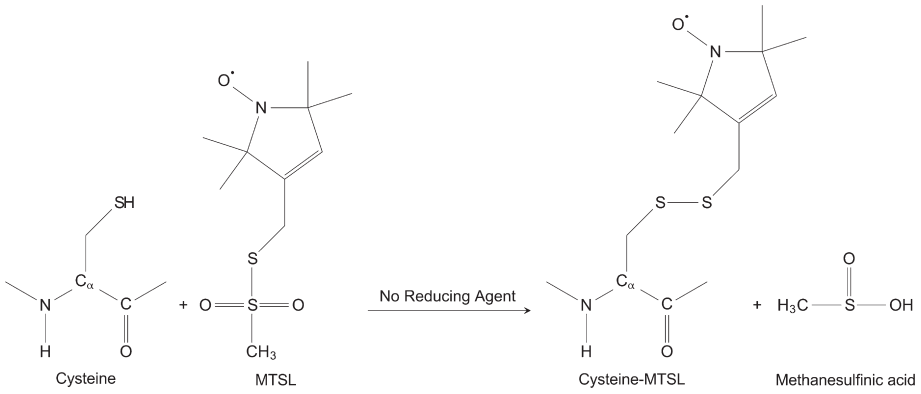

Figure 6. Standard strategy for MTSL spin labeling on the single-surface exposed cysteine. In the absence of reducing agents, the reactive thiol group of MTSL makes a disulphide bond with cysteine. The diamagnetic condition can be achieved by the addition of molar excess ascorbic acid.

modified with cysteine or reactive amines and used as paramagnetic tags (Lauffer and Brady 1985; Arano *et al.* 1996; Lewis and Shively 1998; Tei *et al.* 2008). Due to the orientational dependence of PCS, more rigid tags were introduced, where a metal binding tag conjugates to two juxtapositioned cysteines (Keizers *et al.* 2007, 2008).

4.5 *Case study: Open and closed conformation of U2AF65*

Multi-domain proteins often drive many cellular functions. These proteins are composed of independent structural units connected by flexible linkers. The flexibility induced by the linker may provide multiple conformations, which further provides dynamic interaction of multiple domains in recognizing different ligands. We describe here a case study where PRE derived data has significantly helped in understanding the biological mechanism in U2AF65, where the tandem RNA recognition motif (RRM1-RRM2) domains of U2 snRNP auxiliary factor 65 (U2AF65) recognize the 3′ splice-site associated polypyrimidine (Py) tracts of the target pre-mRNA (Ito *et al.* 1999; Sickmier *et al.* 2006). However, there exists significant diversity in the sequence, length and functional strength of Py tract, reflecting dynamic range of events in splice site acceptor site usage during alternative splicing.

Using PRE studies, Sattler and co-workers have characterized the recognition of sequence variation in Py tract by U2AF65 (Mackereth *et al.* 2011; Hennig *et al.* 2015). Structural studies have shown that U2AF65 RRM1-RRM2 (henceforth, RRM12) adopts two distinct three-dimensional arrangements correlated to the presence or absence of a high-affinity RNA (figure 7). In the presence of RNA, the beta sheet from both RRMs is aligned parallel forming an extended basic-RNA binding surface leading to an 'open' conformation for RRM12 (figure 7). However, in the absence of RNA, RRM12 majorly adopts 'closed' conformation in which the RNA binding surface β2 from RRM1 is occluded by α helices from RRM2. Mackereth *et al.* performed PRE studies with different sequences of RNA to examine whether a dynamic equilibrium between the open and closed forms of RRM12 provide a rationale for regulating strength of Py tract binding. Binding of U4 with RRM12 resulted in chemical shift changes in only RRM2, as the beta sheet in RRM1 is occluded (Mackereth *et al.* 2011). Thus, RRM12 with U4 RNA primarily adopt a closed conformation (figure 7). The ligands U4A4, U4A8U4, and U9 exhibited gradual increase in affinity for RRM12 (figure 7). Interestingly regardless of the overall affinity, all the RNA ligands show similar binding to RRM2, whereas increase in the affinity to particular Py tract was observed with increased perturbation in RRM1 resonances. Moreover, PRE data analysis with each Py tract shows a gradual shift in the population from close to open conformation with increasing affinity for Py tract (figure 7). Thus, the data strongly favors that the population shift from the closed to open conformation is coupled to the functional U2AF65 activity during spliceosome assembly.

5. Newer methods in relaxation and dynamics

In addition to the developments in the determination of high-resolution solution structures of large protein assemblies and multi-domain proteins, biomolecular NMR methodologies focusing on the study of protein dynamics have also experienced a surge in developments during last two decades.

The static structures represent the lowest energy state of the protein and mostly fail to address multiple functions or

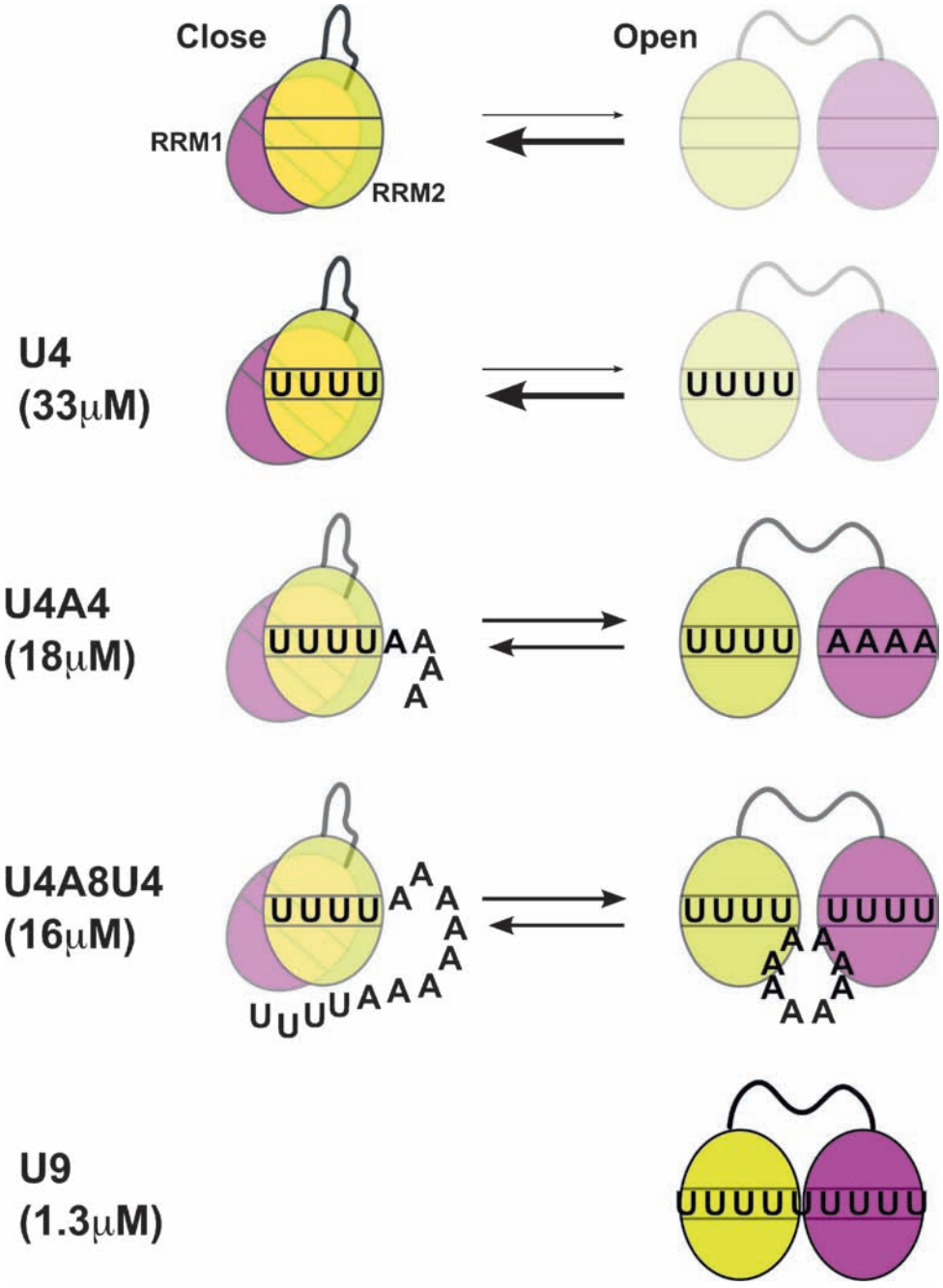

Figure 7. RRM12 binding with Py tract RNA. RRM12 primarily adopts a closed conformation in the absence of RNA, however, a minor population of open confirmation is also observed. RRM1 and RRM2 are represented in pink and yellow, respectively. The gradient from close to open conformation is observed in the presence of higher affinity RNA (U4<U4A4<U4A8U4<U9). Figure adapted and modified with permission from Mackereth *et al.* (2011).

mechanisms exhibited by them. A multi-dimensional energy landscape, which provides relative probabilities of conformational states (thermodynamics) and the energy barrier between them (kinetics), offers a complete description of proteins. Solution NMR spectroscopy is ideally suited to study functionally relevant motions exhibited by biological macromolecules (i.e., proteins and nucleic acids) at a wide range of timescales at an atomic resolution.

From the energy landscape, it is clear that in addition to timescale fluctuations, protein dynamics are characterized by the amplitude and directionality of the fluctuations (figure 8a). The large energy barrier (>kT, k is Boltzmann

constant) between two distinct states (A and B) is often characterized by dynamics on slow timescale (μs to ms). A large ensemble of similar structures is separated by small energy barriers (<kT), which results in local small-amplitude picosecond to nanosecond fluctuations. NMR experimental methods that can detect their respective dynamic processes at each timescale are shown in figure 8b.

5.1 *Picosecond–Nanosecond dynamics*

Oscillations within the energy landscape occur as 'fast' timescale dynamics. Loop motions and side chain fluctuations occur at nanosecond timescale. Notably, bond vibrations occur at femtosecond timescale. Motions on ps-ns timescales, that are comparable to or faster than overall rotational correlation times (τ_c) of the biomolecules, are accessible to spin relaxation resulting from modulation of dipole-dipole (DD) and chemical shift anisotropy (CSA). Generalized squared order parameter (S^2), which defines the order of the nuclei and internal motions (τ_e) exhibited by the nucleus can be extracted from spin-lattice relaxation rates (R_1), spin-spin relaxation rates (R_2) and heteronuclear NOE (Farrow *et al.* 1994). S^2 ranges from 0 (isotropic rotation) to 1 (completely rigid) and is commonly measured for backbone amide (^{15}N) atoms and side-chain methyl (^{13}C) atoms. For an IS spin system, the spin-lattice relaxation rates for spin S (e.g., ^{13}C or ^{15}N) due to dipole-dipole interactions with the I (^1H) spin and to CSA of the S spin are given by equation SE9-SE11 (Cavanagh *et al.* 2007).

The analysis of the relaxation data is often performed within the framework of Model-Free Formulation (Lipari and Szabo 1982a; Lipari and Szabo 1982b) and its extended version, Extended Model-Free Formalism (Clore *et al.* 1990) or Reduced Spectral Density Approach (Farrow *et al.* 1995, Ishima and Nagayama 1995, Peng and Wagner 1995, Lefèvre *et al.* 1996). In a recent study, Jeremko and co-workers have extensively analysed the existing R_1, R_2 and hetNOE data available in the literature for the globular proteins (Jaremko et al. 2015c). The analysis of the multiple field relaxation datasets suggested that the model-free analysis is sufficient in describing slow internal motions. The simulations performed in the study suggests that the careful interpretation of the model-free analysis allows access to both, sub-τ_c and supra-τ_c internal motions.

5.2 *Microsecond–Second dynamics*

Studies from biophysical techniques such as florescence spectroscopy, X-ray crystallography, are restricted to highly populated and homogeneous samples and evade the attention of many biological processes involving low populated intermediate steps that are short lived. These states, referred as dark or invisible states, are noticed during

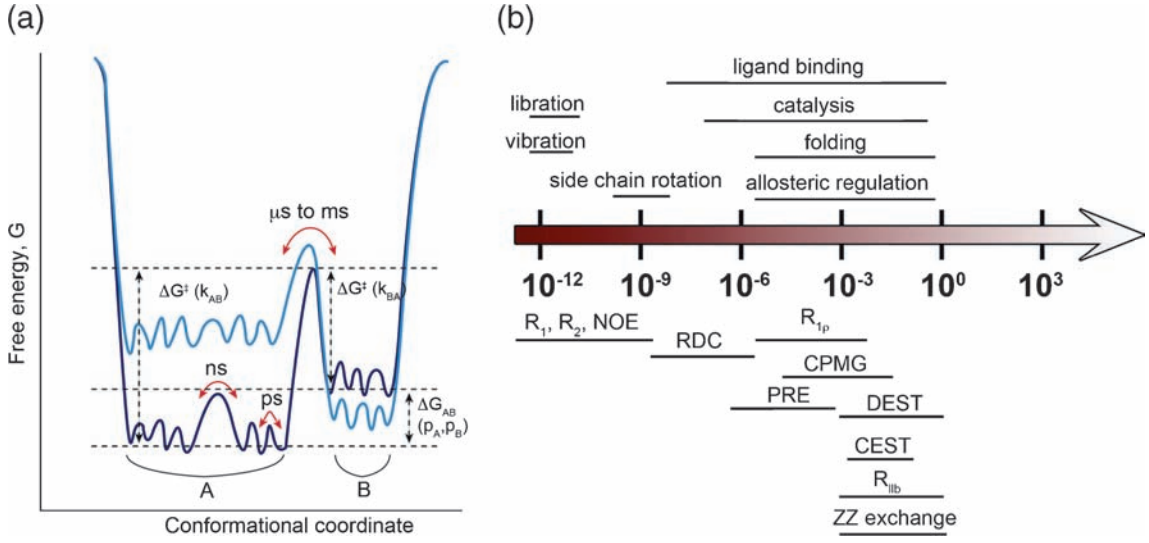

Figure 8. Energy landscape and NMR methods for dynamics. (**a**) Free energy landscape is represented along conformational coordinate. Population, p_A and p_B are the Boltzmann distributions based on their difference in free energy (ΔG_{AB}). Any change in the system, such as temperature or pressure, alters the energy barrier (from dark blue to cyan). Figure adapted with permission from Henzler-Wildman and Kern (2007). (**b**) Faster to slower timescale fluctuations in protein motions are represented with a gradient from red to white. Text above the timescale shows different motions at a particular timescale, and text below the timescale shows different NMR methods to study dynamics of macromolecules.

protein folding, signal transduction, conformational selection, enzyme catalysis, protein aggregation, etc. In recent years, several NMR experiments were developed to access invisible states that occur over a wide range of timescales (figure 8b). These are paramagnetic relaxation enhancement for fast timescale (ns timescale); relaxation dispersion experiments (i.e., rotating frame R_1 ($R_1\rho$) and CPMG (Carr-Purcell-Meiboom-Gill)), CEST (Chemical Exchange Saturation Transfer) for the intermediate timescale (i.e., motions that occur at ms-μs timescale); DEST (Dark state Excitation Saturation Transfer), lifetime line broadening (R_{llb}) for slow timescale (ms-s timescale); and H exchange ZZ exchange for very slow timescale (motions occurring at sec or higher timescale). The process can be further explained by considering a molecule undergoing conformation exchange from state A to state B as follows:

$$A \rightleftharpoons B$$

A molecule exhibiting chemical exchange between states 'A' and 'B' is shown above. PRE, R_{llb} and DEST experiments are utilized to study conformers that exhibit significantly different intrinsic transverse relaxation rates ($R^0_{2,B} >> R^0_{2,A}$); $R_1\rho$, CEST, CPMG and H-exchange experiments suits for systems that undergo chemical exchange with identical rates for both conformers.

5.3 *Relaxation dispersion: CPMG (Carr-Purcell-Meiboom-Gill) and $R_1\rho$*

In the conformational exchange model, the molecule exists in either of the states by inter-converting from one state to the other (figure 9a). The chemical shift of the molecule in conformation A and B is given as ω_A and ω_B, respectively. For instance, if a spin existed in a particular state for time T, and when a refocusing pulse (180°) is applied exactly at the center of the time interval, then the total frequency evolution of the spin is refocused. In a low-frequency CPMG (figure 9a), 180° pulses are applied at regular long intervals of time where the molecule does not stay in a single state. Hence, the spins are not refocused. However, when the refocusing pulses are applied at regular short intervals, as in high-frequency CPMG (figure 9a), for most of the time intervals, molecules exist in one conformation, hence more frequency refocusing occurs. The sum of all the signals from low-frequency CPMG leads to broader linewidths (high $R_{2,eff}$), while signals from high-frequency CPMG lead to sharp linewidths (low $R_{2,eff}$). R_{ex} rate, population of the major and minor state, and chemical shift information of the minor state can be extracted from the relaxation dispersion profile shown in figure 9b.

The $R_1\rho$ experiment is performed by locking transverse magnetization in the rotating frame by applying spin-lock field (equation SE12 and SE13). While CPMG experiments can characterize motions in the range of ~500 μs-10 ms, $R_1\rho$ can be used to decipher motions at much faster rates (~50 μs–2 ms). CPMG and $R_1\rho$ based relaxation dispersion studies have helped in understanding some of the very important biological processes, e.g., Wright and co-workers showed that the ligand guides the dihydrofolate reductase through the reaction pathway along a preferred kinetic path by allowing several short-lived intermediate conformations (Boehr *et al.* 2006). In another example, Kay and co-workers have explored the kinetics and energetics of the low populated folding intermediates in Fyn SH3 domain and its mutants which could decide the fate of the folding event and which were previously inaccessible to any structural biology tool (Korzhnev *et al.* 2004). Al-Hashimi and co-workers have used off-resonance ^{13}C and ^{15}N $R_1\rho$ relaxation dispersion experiments to identify the extensive presence of the Hoogsteen base pairs with a wide range of energies in the canonical DNA duplexes (Alvey *et al.* 2014). An example elaborating the strength of relaxation dispersion studies is given below.

5.4 *Application of relaxation dispersion to catabolite activator protein*

In the therapeutic treatment of human diseases, the discovery of small molecules has found widespread applications by inhibiting protein binding or enzymatic activity (Wells and McClendon 2007; Lee and Craik 2009). In comparison to orthosteric regulation, allostery offers several advantages, in regard to selectivity (Hardy and Wells 2004; Maksay 2011). The majority of the allosteric inhibitors were serendipitously discovered and in many cases the mechanism of inhibitory action remain elusive due to the indistinguishable structural differences in the active sites, in the absence and presence of the inhibitor (Arkin and Wells 2004; Tzeng and Kalodimos 2011). The catabolite activator protein is a gene regulatory protein (transcriptional activator) and exists as a homodimer in solution. It consists of a ligand binding domain (CBD, cyclic AMP binding domain) at its N-terminus and a DNA binding domain (DBD) at its C-terminus. CAP is activated by the binding of cAMP to CBD, which elicits an allosteric response to DBD to bind DNA (Passner *et al.* 2000). Contrarily, DNA binding activity of a double mutant $CAP^{T127L,S128I}$ (CAP*) is inhibited by cGMP binding to CBD (Youn *et al.* 2006).

In both wild-type CAP and CAP*, DBD adopts inactive conformation (Tzeng and Kalodimos 2013). However, ITC studies showed that CAP binds to DNA (K_d~0.5 μM). Relaxation dispersion experiments provide structural and

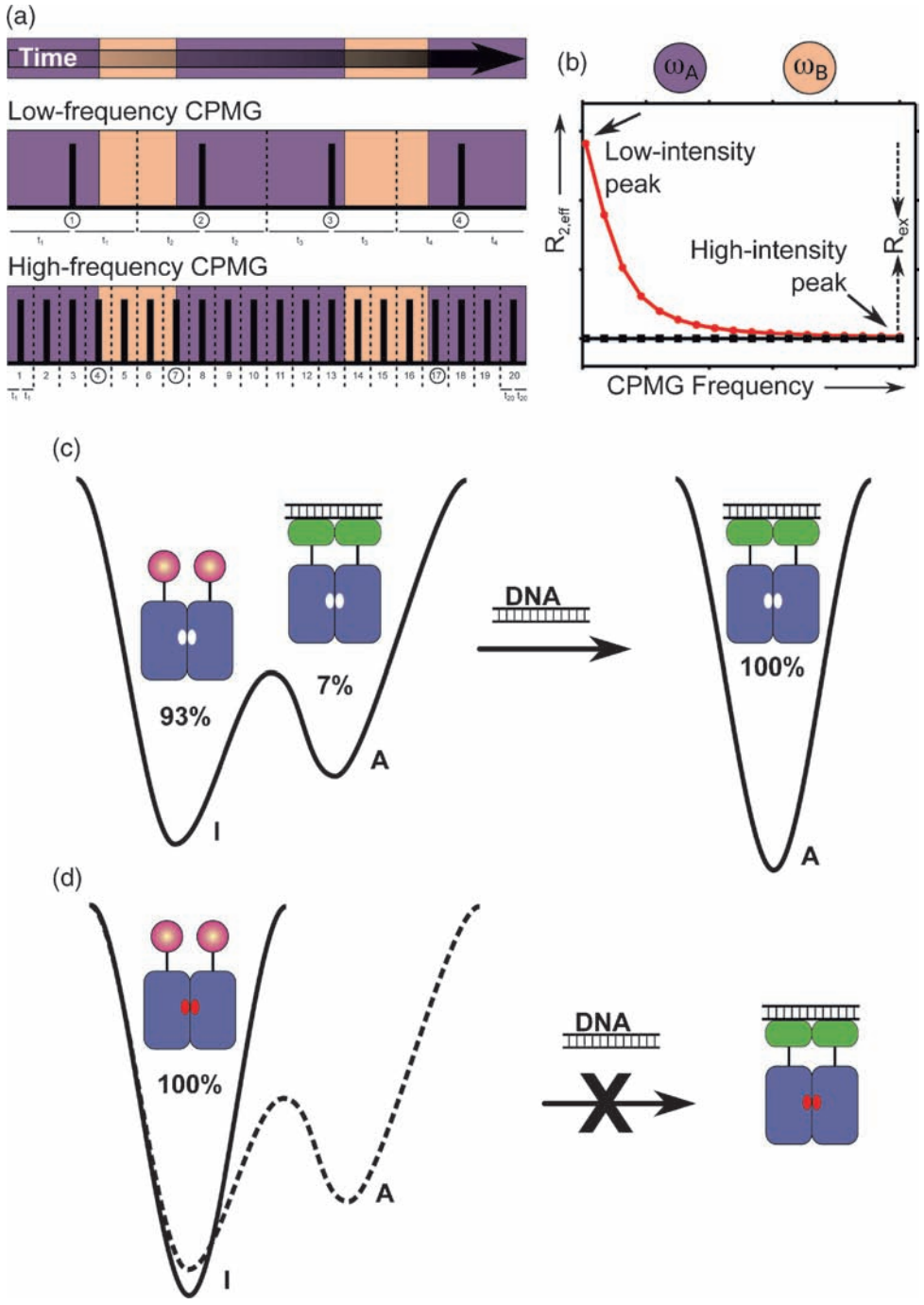

Figure 9. CPMG experiment and energy landscape modulation of CAP* by cGMP. (**a**) As a function of time, a molecule inter-converts between two conformational states, A (purple) and B (salmon). Low- and high-frequency CPMG pulses are applied to modulate the relaxation rates of the molecule. The circles at the bottom of pulses denote non-refocused time intervals. (**b**) Relaxation dispersion profile (red) of molecule can be obtained by plotting $R_{2,eff}$ along different CPMG frequencies, which can be used to extract exchange rates between the two states. The black-coloured line represents relaxation dispersion profile for a residue undergoing no chemical shift exchange. (**c**) Energy landscape of CAP* showing the two states, inactive (I) and active (A), and their fractional populations (93% and 7%, respectively). DNA binding selects the active conformation in a population shift mechanism. (**d**) Energy landscape of CAP*-cGMP2 showing that depletion of the active conformation in CAP*-cGMP2 results in DNA binding inhibition. Panel (**a**) and (**b**) are adapted and modified with permission from Baldwin and Kay (2009). Panel (**c**) and (**d**) are adapted and modified with permission from Tzeng and Kalodimos (2013).

dynamic information about the conformations states of poorly populated states. Hence, relaxation dispersion studies were conducted on CAP to elucidate its binding with DNA. Results indicated the existence of a conformational state that undergoes transition with the ground state on the μs-ms timescale. R_{ex} rates of up to 15 s^{-1} were detected for many residues in DBD. Results suggested the presence of excited state (active state) with a population of 7%, which exchanges at a rate of ~200 s^{-1} with the ground state (inactive state) (figure 9c). Although the presence of cGMP did not induce chemical shift changes in DBD, it has lost its affinity for DNA. The loss of DNA binding by CAP by the inhibitory action of cGMP without affecting DBD resonances (or DBD conformational state) is quite unexpected. To probe the allosteric effect, relaxation dispersion experiments were also conducted on CAP-cGMP. Results showed the absence of μs–ms dynamics, which were previously present in the absence of cGMP. These results clearly demonstrated that DNA binding in CAP was driven by the 7% excited state conformation, which is depleted upon binding of cGMP to CBD, thereby inducing allosteric effect (figure 9d). Collectively, the relaxation dispersion experiments performed by Kalodimos and co-workers showed that DNA binds to the excited state (7% population) of CAP and thus shift the equilibrium from inactive state to active state through a population-shift mechanism. These experiments also provided a direct evidence for allosteric mechanism wherein inhibitors can directly destabilizing the lowly populated excited state without affecting the ground state of the molecules.

5.5 *DEST and CEST*

Using saturation transfer, invisible states can be detected by selectively saturating invisible (or minor) states by off-resonance radio frequency (RF); further, the saturation can be transferred to visible major state by the chemical exchange. NMR saturation experiments were developed in the early 1960s and found their application in protein–ligand interactions by late 1990s (Forsen and Hoffman 1963; Mayer and Meyer 1999).

CEST experiments provide information similar to CPMG, although for processes on slower timescales, which is a limiting factor for CPMG. CEST can be used to visualize a minor state if the exchange timescale is ~2–100 ms with its major state. CEST experiments can be carried out on systems provided the chemical shifts of major and minor states are significantly different, as well as, the intrinsic R_2 rates of conformers undergoing exchange are identical (Vallurupalli *et al.* 2012).

DEST experiments are used to characterize visible states that are in chemical exchange with the dark states, which are invisible due to its extremely large size (>1 MDa) (Fawzi *et al.* 2012; Libich *et al.* 2013). The resonances from large and slowly tumbling invisible states characterized with rapid transverse relaxation rates can be saturated with off-resonance RF that does not excite the narrow signals from rapidly tumbling visible states. The saturation can be transferred to the visible state if the rate of conversion from dark state to invisible state is in the order of or faster than its longitudinal relaxation rates and can be manifested by the loss of cross-peak intensity for visible species. Exchanges on an intermediate-to-slow timescales (0.5–1000 ms) can be characterized using DEST. The pulse sequence for DEST is similar to ^{15}N-R_1 longitudinal relaxation time measurement, except that the variable T_1 relaxation delay is replaced by a weak, continuous wave saturation pulse at variable radio-frequency offsets. The template of pulse sequence used for CEST is identical to the one used for DEST, except for a couple of points: (i) The continuous wave RF B_1 field is applied far (in kHz) off-resonance in DEST, whereas RF field is applied off-resonance by few Hz in CEST, and (ii) Low (~10-30 Hz) B_1 field strengths are applied in CEST, while higher field strengths (>100 Hz) are used in DEST. CEST provides the kinetics and thermodynamics of the exchange process, in addition to the chemical shift of the excited states. For a molecule undergoing 10–100 ms exchange rate, a field strength of 5–10 Hz is best suited to observe the minor state using CEST approach. Contrarily, DEST involves much larger field strengths, as the outcome does not dependent on resolving chemical shift difference between the visible and invisible states.

Using DEST approach, the exchange rates of 3 s^{-1} was calculated for conversion of amyloid-β40 (Aβ40) monomers to Aβ40 protofibrils (Fawzi *et al.* 2011). An exchange rate of 150 s^{-1} was obtained for free and peptide-bound SH3 domain using CEST (Vallurupalli *et al.* 2012). Conventionally, PREs have been used to study the process of conformational exchange in the fast exchange regime (k_{ex}>>10000 s^{-1}) (Sekhar *et al.* 2016). However, the recent work by Kay and co-workers demonstrated the utility of the ^1H CEST-based method for obtaining PREs for proteins exhibiting slow exchange (~300 s^{-1}) (Sekhar *et al*. 2016).

6. Solution NMR of challenging membrane proteins

Membrane proteins account about one-third of the proteins expressed by most organisms and perform several vital cellular functions, such as energy generation, communication between cells, transport nutrients, etc. (Wallin and Heijne 1998). Despite advances in structural biology, of ~120,000 structures (July 2016) deposited in PDB, membrane proteins account for less than 2%, due to their complexity with their expression in soluble form as well as sample preparation. Additionally, the enhanced τ_c due to membrane association

induces broader lines making the NMR data incomprehensible.

E. coli outer membrane protein A (OmpA) was the first membrane structure determined by NMR in the year 2001 (Arora *et al.* 2001). With advances in NMR-based structural biology, significant progress has been made in the recent years in understanding structure and function of integral membrane proteins. For example, the structural studies on Helical transport protein (TSPO) which mediates the import of cholesterol and porphyrins into mitochondria, in the presence of ligand (PK11195) depicted a ligand-induced modulation of the TSPO structure, thereby suggesting a molecular mechanism for the stimulation of cholesterol transport into mitochondria (Jaremko *et al.* 2014, 2015b). Additionally, dynamic properties of the side-chain methyl groups of TSPO and its physiologically important mutant TSPO-A147T revealed an identical binding mode for WT and A147T mutant (Jaremko et al. 2015a). A recent review on this subject by Liang and Tamm covers recent advances in the field of structures of the membrane protein (Liang and Tamm 2016).

7. Other notable developments

The other areas of development in biomolecular NMR in last two decades are fast data acquisition techniques such as Non-uniform sampling, G-Fourier Transform, projection-reconstruction and so-FAST or BEST pulsing for conventional experiments. The biggest advantage of these techniques is the significant reduction in generating NMR data. These methods are already routinely used as they provide significant advantages over the conventional data collection methods. A structure determination project often requires over two-three weeks of data collection with conventional methods. The fast data acquisition methods can acquire all necessary experimental data in about 4–5 days (Hyberts *et al.* 2012). Additionally, a combination of fast data acquisition methods may yield all data for structure determination in just 3–4 h, which is a significant advancement (A. Haribabu, personal communication). In addition to the reduction in spectrometer usage, these developments have huge potential in understanding structural details of proteins that are inherently less stable. The discussion on fast data acquisition methods is out of the scope of this review as the methodological development is highly technical. Interested readers may obtain more information on these and other techniques in an excellent literature on this topic (Kupče and Freeman 2004; Schanda and Brutscher 2005; Szyperski and Atreya 2006; Hyberts *et al.* 2014). Other notable emerging techniques are proton-less NMR, direct detection of heteronuclei, dual detection, etc.

8. Conclusions

Developments in biomolecular NMR Spectroscopy can be broadly classified into three segments. Initial developments heavily relied on ^{1}H based protein structure determination. Although this approach was successful in deriving some of the early structures, the spectral complexity deterred the process of its generalized use, particularly for proteins above 8–10 kDa in size. A significant improvement was achieved with the development and use of heteronuclear edited experiments involving ^{1}H, ^{13}C and ^{15}N to obtain complete sequential chemical shift and NOE assignments for calculating high-resolution three-dimensional structure. While these developments allowed quicker structure determination, it was limited to globular proteins of size up to 20 kDa, limiting structure determination of larger proteins. Moreover, these methods are limited in long-range restraints generation, which are vital for structure and orientation determination for multi-domain systems. These problems can be greatly annihilated by the use of techniques like TROSY, PRE, RDCs, and PCS. Moreover, labelling strategies involving perdeuteration and methyl-specific labelling of ILVs allows generation of key structural restraints for large proteins. Significant developments for probing protein motions, in timescales that are significant to biological processes but are not accessible by any other technique, have occurred in last two decades. These include use of CPMG, relaxation dispersion, CEST, and DEST methods. Together, these methodological improvements continue to excite the field of biomolecular NMR and assure the bright future of the field in aiding solution structural biology.

Acknowledgements

This work was supported by Council of Scientific and Industrial Research (CSIR) network projects GenCODE (BSC0123) and SplenDID (BSC0104). CSC acknowledges a PhD fellowship from the CSIR, India.

References

Allegrozzi M, Bertini I, Janik MBL, Lee YM, Liu G and Luchinat C 2000 Lanthanide-induced pseudocontact shifts for solution structure refinements of macromolecules in shells up to 40 Å from the metal ion. *J. Am. Chem. Soc.* **122** 4154–4161

Alvey HS, Gottardo FL, Nikolova EN and Al-Hashimi HM 2014 Widespread transient Hoogsteen base pairs in canonical duplex DNA with variable energetics. *Nat. Commun.* **5** 4786

Arano Y, Uezono T, Akizawa H, Ono M, Wakisaka K, Nakayama M, Sakahara H, Konishi J, *et al.* 1996 Reassessment of diethylenetriaminepentaacetic acid (DTPA) as a chelating agent for indium-111 labeling of polypeptides using a newly synthesized monoreactive DTPA derivative. *J. Med. Chem.* **39** 3451–3460

Arkin MR and Wells JA 2004 Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **3** 301–317

Arora A, Abildgaard F, Bushweller JH and Tamm LK 2001 Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nat. Struct. Biol.* **8** 334–338

Baldwin AJ and Kay LE 2009 NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.* **5** 808–814

Battiste JL and 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

Bertini I, Janik MBL, Lee YM, Luchinat C and Rosato A 2001 Magnetic susceptibility tensor anisotropies for a lanthanide ion series in a fixed protein matrix. *J. Am. Chem. Soc.* **123** 4181–4188

Bertini I, Luchinat C and Parigi G 2002 Magnetic susceptibility in paramagnetic NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **40** 249–273

Bertini I, Del Bianco C, Gelis I, Katsaros N, Luchinat C, Parigi G, Peana M, Provenzani A, *et al*. 2004 Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. *Proc. Natl. Acad. Sci. U. S. A.* **101** 6841–6846

Bertini I, Gupta YK, Luchinat C, Parigi G, Peana M, Sgheri L and Yuan J 2007 Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. *J. Am. Chem. Soc.* **129** 12786–12794

Bertini I, Luchinat C, Parigi G and Pierattelli R 2008 Perspectives in paramagnetic NMR of metalloproteins. *Dalton Trans.* 3782–3790

Bloembergen N and Morgan LO 1961 Proton relaxation times in paramagnetic solutions. Effects of electron spin relaxation. *J. Chem. Phys.* **34** 842

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

Cavanagh J, Fairbrother W, Palmer AG, Rance M and Skelton N 2007 Protein NMR spectroscopy. Principles and Practice.

Clore GM and 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. *Chem. Rev.* **109** 4108–4139

Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC and Gronenborn AM 1990 Deviations from the simple two-parameter model-free approach to the interpretation of nitorgen-15 nuclear magnetic-relaxation of proteins. *J. Am. Chem. Soc.* **112** 4989–4991

Crespi HL, Rosenberg RM and Katz JJ 1968 Proton magnetic resonance of proteins fully deuterated except for ^{1}H-leucine side chains. *Science.* **161** 795–796

Deshmukh MV, John M, Coles M, Peters J, Baumeister W and Kessler H 2006 Inter-domain orientation and motions in VAT-N explored by residual dipolar couplings and ^{15}N backbone relaxation. *Magn. Reson. Chem.* **44** S89–S100

Donaldson LW, Skrynnikov NR, Choy WY, Muhandiram DR, Sarkar B, Forman-Kay JD and Kay LE 2001 Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. *J. Am. Chem. Soc.* **123** 9843–9847

Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, *et al*. 1994 Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry.* **33** 5984–6003

Farrow NA, Zhang OW, Szabo A, Torchia DA and Kay LE 1995 Spectral density-function mapping using ^{15}N relaxation data exclusively. *J. Biomol. NMR.* **6** 153–162

Fawzi NL, Ying J, Ghirlando R, Torchia DA and Clore GM 2011 Atomic-resolution dynamics on the surface of amyloid-β protofibrils probed by solution NMR. *Nature.* **480** 268–272

Fawzi NL, Ying J, Torchia DA and Clore GM 2012 Probing exchange kinetics and atomic resolution dynamics in high-molecular-weight complexes using dark-state exchange saturation transfer NMR spectroscopy. *Nat. Protoc.* **7** 1523–1533

Forsen S and Hoffman RA 1963 Study of moderately rapid chemical exchange reactions by means of Nuclear Magnetic Double Resonance. *J. Chem. Phys.* **39** 2892

Gardner KH and Kay LE 1998 The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27** 357–406

Gardner KH, Rosen MK and Kay LE 1997 Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. *Biochemistry.* **36** 1389–1401

Göbl C, Madl T, Simon B and Sattler M 2014 NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **80** 26–63

Griffith OH and McConnell HM 1966 A nitroxide-maleimide spin label. *Proc. Natl. Acad. Sci. U. S. A.* **55** 8–11

Grzesiek S, Anglister J, Ren H and Bax A 1993 Carbon-13 line narrowing by deuterium decoupling in deuterium/carbon-13/nitrogen-15 enriched proteins. Application to triple resonance 4D J connectivity of sequential amides. *J. Am. Chem. Soc.* **115** 4369–4370

Hansen MR, Mueller L and Pardi A 1998 Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5** 1065–1074

Hardy JA and Wells JA 2004 Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* **14** 706–715

Haribabu A personal communication

Hennig J, Warner LR, Simon B, Geerlof A, Mackereth CD and Sattler M 2015 Structural analysis of protein-RNA complexes in solution using NMR paramagnetic relaxation enhancements. *Methods Enzymol.* **558** 333–362

Henzler-Wildman K and Kern D 2007 Dynamic personalities of proteins. *Nature.* **450** 964–972

Hubbell WL, Cafiso DS and Altenbach C 2000 Identifying conformational changes with site-directed spin labeling. *Nat. Struct. Biol.* **7** 735–739

Hubbell WL, López CJ, Altenbach C and Yang Z 2013 Technological advances in site-directed spin labeling of proteins. *Curr. Opin. Struct. Biol.* **23** 725–733

Hyberts SG, Milbradt AG, Wagner AB, Arthanari H and Wagner G 2012 Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J. Biomol. NMR.* **52** 315–327

Hyberts SG, Arthanari H, Robson SA and Wagner G 2014 Perspectives in magnetic resonance: NMR in the post-FFT era. *J. Magn. Reson.* **241** 60–73

Ishima R and Nagayama K 1995 Protein backbone dynamics revealed by quasi spectral density function analysis of amide N-15 nuclei. *Biochemistry.* **34** 3162–3171

Ito T, Muto Y, Green MR and Yokoyama S 1999 Solution structures of the first end second RNA-binding domains of human U2 small nuclear ribonucleoprotein particle auxiliary factor (U2AF65). *EMBO J.* **18** 4523–4534

Jaremko Ł, Jaremko M, Giller K, Becker S and Zweckstetter M 2014 Structure of the mitochondrial translocator protein in complex with a diagnostic ligand. *Science.* **343** 1363–1366

Jaremko M, Jaremko Ł, Giller K, Becker S and Zweckstetter M 2015a Structural integrity of the A147T polymorph of mammalian TSPO. *ChemBioChem.* **16** 1483–1489

Jaremko M, Jaremko Ł, Jaipuria G, Becker S and Zweckstetter M 2015b Structure of the mammalian TSPO/PBR protein. *Biochem. Soc. Trans.* **43** 566–571

Jaremko Ł, Jaremko M, Nowakowski M and Ejchart A 2015c The quest for simplicity: remarks on the free-approach models. *J. Phys. Chem. B.* **119** 11978–11987

Jensen MR, Lauritzen C, Dahl SW, Pedersen J and Led JJ 2004 Binding ability of a HHP-tagged protein towards Ni^{2+} studied by paramagnetic NMR relaxation: The possibility of obtaining long-range structure information. *J. Biomol. NMR.* **29** 175–185

Jeschke G 2013 Conformational dynamics and distribution of nitroxide spin labels. *Prog. Nucl. Magn. Reson. Spectrosc.* **72** 42–60

Keizers PHJ and Ubbink M 2011 Paramagnetic tagging for protein structure and dynamics analysis. *Prog. Nucl. Magn. Reson. Spectrosc.* **58** 88–96

Keizers PHJ, Desreux JF, Overhand M and Ubbink M 2007 Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. *J. Am. Chem. Soc.* **129** 9292–9293

Keizers PHJ, Saragliadis A, Hiruma Y, Overhand M and Ubbink M 2008 Design, synthesis, and evaluation of a lanthanide chelating protein probe: CLaNP-5 yields predictable paramagnetic effects independent of environment. *J. Am. Chem. Soc.* **130** 14802–14812

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

Kramer F, Deshmukh MV, Kessler H and Glaser SJ 2004 Residual dipolar coupling constants: an elementary derivation of key equations. *Concepts Magn. Reson. Part A.* **21** 10–21

Kupče E and Freeman R 2004 Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy. *J. Am. Chem. Soc.* **126** 6429–6440

Lauffer RB and Brady TJ 1985 Preparation and water relaxation properties of proteins labeled with paramagnetic metal chelates. *Magn. Reson. Imaging.* **3** 11–16

Lee GM and Craik CS 2009 Trapping moving targets with small molecules. *Science.* **324** 213–215

Lefèvre JF, Dayie KT, Peng JW and Wagner G 1996 Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N-H spectral density functions. *Biochemistry.* **35** 2674–2686

Lewis MR and Shively JE 1998 Maleimidocysteineamido-DOTA derivatives: New reagents for radiometal chelate conjugation to antibody sulfhydryl groups undergo pH-dependent cleavage reactions. *Bioconjug. Chem.* **9** 72–86

Liang B and Tamm LK 2016 NMR as a tool to investigate the structure, dynamics and function of membrane proteins. *Nat. Struct. Mol. Biol.* **23** 468–474

Libich DS, Fawzi NL, Ying J and Clore GM 2013 Probing the transient dark state of substrate binding to GroEL by relaxation-based solution NMR. *Proc. Natl. Acad. Sci.* **110** 11361–11366

Lipari G and Szabo A 1982a Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules I. Theory and range of validity. *J. Am. Chem. Soc.* **104** 4546–4559

Lipari G and Szabo A 1982b Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules II. Analysis of experimental results. *J. Am. Chem. Soc.* **104** 4559–4570

Loria JP, Rance M and Palmer AG 1999 Transverse-relaxation-optimized (TROSY) gradient-enhanced triple-resonance NMR spectroscopy. *J. Magn. Reson.* **141** 180–184

Losonczi JA and Prestegard JH 1998 Improved dilute bicelle solutions for high-resolution NMR of biological macromolecules. *J. Biomol. NMR.* **12** 447–451

Mackereth CD, Madl T, Bonnal S, Simon B, Zanier K, Gasch A, Rybin V, Valcárcel J, *et al*. 2011 Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature.* **475** 408–411

Madl T, Felli IC, Bertini I and Sattler M 2010 Structural analysis of protein interfaces from ^{13}C direct-detected paramagnetic relaxation enhancements. *J. Am. Chem. Soc.* **132** 7285–7287

Maksay G 2011 Allostery in pharmacology: Thermodynamics, evolution and design. *Prog. Biophys. Mol. Biol.* **106** 463–473

Markley JL, Putter I and Jardetzky O 1968 High-resolution nuclear magnetic resonance spectra of selectively deuterated staphylococcal nuclease. *Science.* **161** 1249–1251

Mayer M and Meyer B 1999 Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* **38** 1784–1788

Ottiger M and Bax A 1998 Characterization of magnetically oriented phospholipid micelles for measurement of dipolar couplings in macromolecules. *J. Biomol. NMR.* **12** 361–372

Passner JM, Schultz SC and Steitz TA 2000 Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-cAMP at 2.1 Å resolution. *J. Mol. Biol.* **304** 847–859

Peng JW and Wagner G 1995 Frequency spectrum of NH bonds in eglin c from spectral density mapping at multiple fields. *Biochemistry.* **34** 16733–16752

Pervushin K 2000 Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Q. Rev. Biophys.* **33** 161–197

Pervushin K, Riek R, Wider G and Wüthrich K 1997 Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **94** 12366–12371

Pintacuda G, John M, Su XC and Otting G 2007 NMR structure determination of protein-ligand complexes by lanthanide labeling. *Acc. Chem. Res.* **40** 206–212

Prestegard J, Bougault C and Kishore A 2004 Residual Dipolar Couplings in structure determination of biomolecules. *Chem. Rev.* **104** 3519–3540

Salzmann M, Wider G, Pervushin K, Senn H and Wüthrich K 1999 TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *J. Am. Chem. Soc.* **121** 844–848

Schanda P and Brutscher B 2005 Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* **127** 8014–8015

Schulte-Herbrüggen T and Sørensen OW 2000 Clean TROSY: compensation for relaxation-induced artifacts. *J. Magn. Reson.* **144** 123–128

Sekhar A, Rosenzweig R, Bouvignies G and Kay LE 2016 Hsp70 biases the folding pathways of client proteins. *Proc. Natl. Acad. Sci. U. S. A.* **113** E2794–E2801

Shi L and Kay LE 2014 Tracing an allosteric pathway regulating the activity of the HslV protease. *Proc. Natl. Acad. Sci. U. S. A.* **111** 2140–2145

Sickmier EA, Frato KE, Shen H, Paranawithana SR, Green MR and Kielkopf CL 2006 Structural basis for polypyrimidine tract recognition by the essential pre-mRNA aplicing factor U2AF65. *Mol. Cell.* **23** 49–59

Skrynnikov NR, Goto NK, Yang D, Choy WY, Tolman JR, Mueller GA and Kay LE 2000 Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin. *J. Mol. Biol.* **295** 1265–1273

Sprangers R and Kay LE 2007 Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature.* **445** 618–622

Szyperski T and Atreya HS 2006 Principles and applications of GFT projection NMR spectroscopy. *Magn. Reson. Chem.* **44** S51–S60

Tei L, Baranyai Z, Botta M, Piscopo L, Aime S and Giovenzana GB 2008 Synthesis and solution thermodynamic study of rigidified and functionalised EGTA derivatives. *Org. Biomol. Chem.* **6** 2361–2368

Tjandra N and Bax A 1997 Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science.* **278** 1111–1114

Tjandra N, Omichinski JG, Gronenborn AM, Clore GM and Bax A 1997 Use of dipolar ^{1}H-^{15}N and ^{1}H-^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.* **4** 732–738

Todd AP, Cong J, Levinthal F, Levinthal C and Hubbell WL 1989 Site-directed mutagenesis of colicin E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation. *Proteins Struct. Funct. Genet.* **6** 294–305

Tolman JR, Flanagan JM, Kennedy MA and Prestegard JH 1995 Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc. Natl. Acad. Sci. U. S. A.* **92** 9279–9283

Tugarinov V and Kay LE 2003 Ile, Leu, and Val Methyl Assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc.* **125** 13868–13878

Tugarinov V, Hwang PM, Ollerenshaw JE and Kay LE 2003 Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125** 10420–10428

Tugarinov V, Hwang PM and Kay LE 2004 Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu. Rev. Biochem.* **73** 107–146

Tzeng SR and Kalodimos CG 2011 Protein dynamics and allostery: An NMR view. *Curr. Opin. Struct. Biol.* **21** 62–67

Tzeng SR and Kalodimos CG 2013 Allosteric inhibition through suppression of transient conformational states. *Nat. Chem. Biol.* **9** 462–465

Ubbink M, Ejdebäck M, Karlsson BG and Bendall DS 1998 The structure of the complex of plastocyanin and cytochrome *f*, determined by paramagnetic NMR and restrained rigid-body molecular dynamics. *Structure.* **6** 323–335

Vallurupalli P, Bouvignies G and Kay LE 2012 Studying "invisible" excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.* **134** 8148–8161

Venters RA, Farmer BT, Fierke CA and Spicer LD 1996 Characterizing the use of perdeuteration in NMR studies of large proteins: ^{13}C, ^{15}N and ^{1}H assignments of human carbonic anhydrase II. *J. Mol. Biol.* **264** 1101–1116

Wallin E and Heijne G 1998 Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7** 1029–1038

Wells JA and McClendon CL 2007 Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature.* **450** 1001–1009

Wöhnert J, Franz KJ, Nitz M, Imperiali B and Schwalbe H 2003 Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. *J. Am. Chem. Soc.* **125** 13338–13339

Wüthrich K 1998 The second decade into the third millenium. *Nat. Struct. Biol.* **5** 492–495

Xu X, Keizers PHJ, Reinle W, Hannemann F, Bernhardt R and Ubbink M 2009 Intermolecular dynamics studied by paramagnetic tagging. *J. Biomol. NMR.* **43** 247–254

Youn H, Kerby RL, Conrad M and Roberts GP 2006 Study of highly constitutively active mutants suggests how cAMP activates cAMP receptor protein. *J. Biol. Chem.* **281** 1119–1127

MS received 20 May 2016; accepted 15 September 2016

Corresponding editor: DURGADAS P KASBEKAR