

Nuclear magnetic resonance captures the elusive

Steven R Van Doren

Address: Department of Biochemistry, 105 Schweitzer Hall, University of Missouri, Columbia, MO 65211, USA

E-mail: vandorens@missouri.edu

F1000 Biology Reports 2009,1:24 (doi: 10.3410/B1-24)

The electronic version of this article is the complete one and can be found at: <http://F1000.com/Reports/Biology/content/1/24>

Abstract

Innovative nuclear magnetic resonance approaches have emerged as outstanding means of tracking down functionally important properties of biomolecular assemblies that evade detection. These enigmatic attributes include dynamic equilibria with very lightly populated excited states and structures of complexes obscured by slow tumbling in solution.

Introduction and context

Biomolecular nuclear magnetic resonance (NMR) in solution has been known for structures derived from hundreds of nuclear Overhauser effect (NOE)-based distance measurements [1] and studies of equilibrium fluctuations based on relaxation rates [2]. Increasingly, NMR has demonstrated abilities to characterize challenging structural states and dynamically sampled substates that are elusory. Technical impediments to detecting them may arise from severe line broadening associated with high molecular weights, unfolded character, or fleeting populations of higher energy substates. NMR approaches can detect excursions occurring on time scales that are orders of magnitude longer than are accessible by molecular dynamics. This report visits recent milestones of NMR insight into biomolecular complexes and 'invisible' states.

Major recent advances

Large complexes elucidated by NMR

Characterization of biomolecular complexes of high molecular weight in solution relies upon sharpening NMR peaks by perdeuteration and transverse relaxation optimized spectroscopy (TROSY), which exploits slow tumbling with high magnetic field strength [3]. An extreme challenge would be the lack of any tumbling, as prevails in insoluble protein fibrils. Associations in which one partner is effectively immobile have had their interfaces mapped via transferred cross-saturation (TCS) [4]. TCS requires the affinity to be moderate, with

a dissociation rate exceeding 0.1 per second [5]. It detects the free state of the perdeuterated soluble partner at large excess over the insoluble component. The free-state spectra are influenced by the part of its lifetime when it is bound to the insoluble component, where it is exposed to the spectral bleaching spreading to it from the insoluble partner [4]. TCS was applied to locate the collagen-binding groove among loops of the discoidin domain of a receptor tyrosine kinase called DDR2 [6].

Quantitative NMR insight into motion and binding events of oligomeric proteins as large as 1 MDa has been pioneered by exploiting methyl side chains that are distributed throughout protein cores [7,8]. This has relied upon sharpening of methyl peaks by (a) protonating one or both prochiral methyl groups of each Val, Leu, and Ile by means of α -keto acid precursors in otherwise perdeuterated minimal media, and (b) detecting with methyl-TROSY (heteronuclear multiple quantum correlation) spectra [8,9]. Ala methyl peaks should also become suitable probes, with the aid of a new assignment strategy for large proteins [10]. Assignments of methyl peaks of the 670-kDa 20S proteasome, in α_7 heptameric rings sandwiching central β_7 rings, began with a 21-kDa α_{monomer} and were adjusted for the methyl peak positions in the 360-kDa α_7 - α_7 double ring and 670-kDa α_7 - β_7 - β_7 - α_7 proteasome [7]. NOEs between methyl groups of spatially neighboring side chains facilitated assignments in the α_7 rings. ^{13}C and ^2H NMR relaxation detailed mobile interfaces between

subunits of the α_7 rings. Methyl-TROSY relaxation dispersion revealed concerted motions occurring at 1200 to 1600 per second inside the α_7 antechamber en route to the proteolytic chamber [7]. Monitoring changes of methyl peaks elucidated gating of the entrance and provided affinities and binding sites for either the 11S activator [7] or the weak binding of an antimalaria drug between α and β subunits [11].

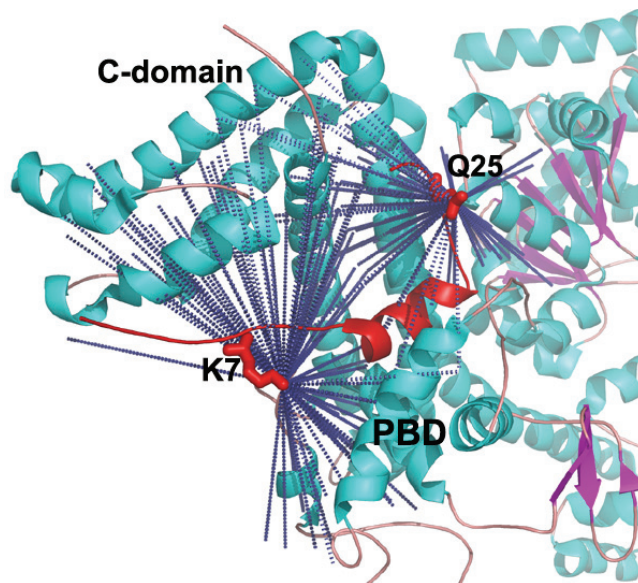
Methyl peak assignments in a fivefold longer chain, built up by interpreting spectra of progressively larger deuterated fragments, provided the foundation for NMR structural models of a SecA complex [12,13]. The SecA component of bacterial protein translocation is a 204-kDa homodimer of 901-residue subunits. It bound a helical signal peptide from a porin in a groove between the preprotein-binding domain (PBD) and the C-terminal domain [13] (Figure 1). Most contacts in the groove are hydrophobic and are surrounded by electrostatic and polar contacts. The insights were facilitated by identifying the NMR peaks of the backbone and side-chain methyl groups of progressively larger deuterated fragments with additional domains [12,13]. To establish this, a single cysteine-tethered nitroxide spin label was placed toward either terminus of the signal peptide. This introduced line broadening to TROSY spectra of SecA methyl groups [7] out to around 25 Å. These paramagnetic relaxation enhancements (PREs) provided long-range distance restraints to dock the signal peptide (helical based on its transferred NOEs) to SecA [13] (Figure 1). Spin labeling of the PBD introduced PREs to the C-terminal domain that indicated the cleft between them to be 90% in the open conformation and 10% in the closed conformation [13] in the free state.

New NMR windows on structural diversity

Rare and 'invisible' high-energy states were discovered by NMR to have populations down to less than 1 part per 10 million using native-state hydrogen exchange (NHX). These fleeting protein states, called partially unfolded forms, are accessible by transitory visits from the fully folded state at equilibrium and were detected initially in a helical protein [14] and an α/β fold [15] and later in a β sandwich [16].

The structural diversity encompassing higher energy substates of low population can now be identified and quantified by multiple means. Elusive manifolds of structural substates have been observed using bold modern applications of a variety of NMR methodologies known to be sensitive to dynamic equilibria. 'Invisible' states populated to at least 0.5% have often been detected via the line broadening accompanying conformational equilibria by means of a Carr-Purcell-

Figure 1. Long-range distances between SecA and signal peptide measured by paramagnetic relaxation enhancement (PRE) of SecA spectra by nitroxide spin labeling of either sequence position marked on the peptide bound



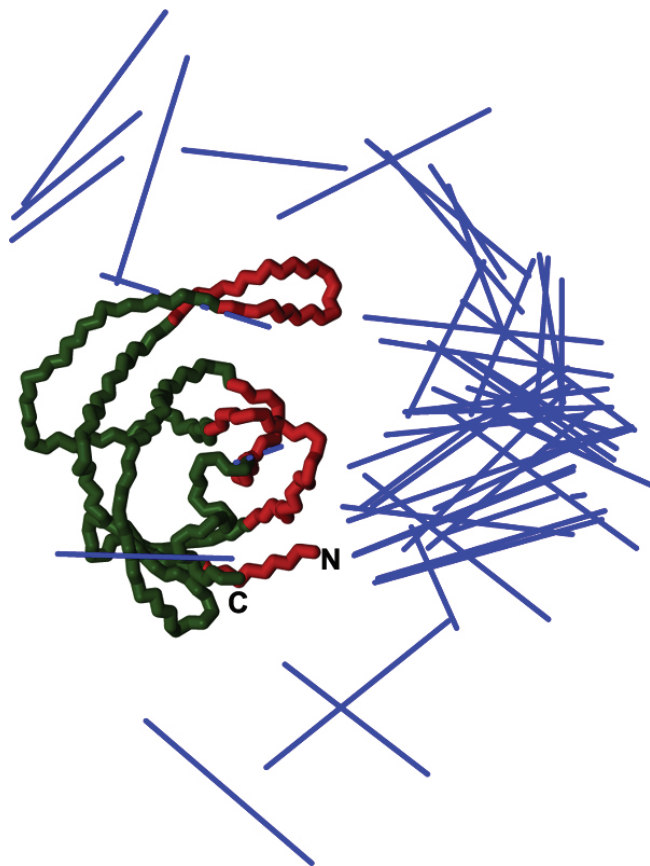
Dashed lines represent PRE-derived distances between methyl groups of SecA (PDB code 2VDA) and the nitroxide spin label linked to a cysteine substitution for Lys7 or Gln25 of the peptide colored red. PBD, preprotein-binding domain.

Meiboom-Gill (CPMG) radiofrequency (RF) pulsing strategy [17]. Sparsely populated structural states have recently been computed from structural restraints of either residual dipolar couplings (RDCs) or PRE-based line broadening that depends on distance from a paramagnetic center. 'Invisible' states have been described in protein-protein encounter complexes, enzymes, and other structures.

Protein encounter complexes revealed

The PRE approach is very responsive to transient close approaches to protons of a usually distant paramagnetic spin label. This results from the steep distance dependence of the line broadening proportional to $\langle r^{-6} \rangle$ over the range of the proton line broadening out to 24 Å or more. For protein-DNA and protein-protein associations driven by electrostatics, fitting of intermolecular PREs by rigid body docking visualized lightly populated (higher energy) encounter complexes differing from static structural coordinates [18,19]. This approach recently resolved the dilemma of how the apparently monomeric precursor to HIV protease activates itself into mature HIV protease dimer where the termini are sequestered in a β sheet remote from the active site [20]. Using a truncated precursor form, large intermolecular PREs were

Figure 2. The predominantly monomeric precursor of HIV protease accompanied by lightly populated dimeric species



The blue lines represent some of the multitude of orientations in the transient encounter complexes that the second precursor subunit may adopt relative to the monomer plotted. The lines connect the C α atoms of Gly27 of the active site and Ile62 on the distant side of the protomer. In the mature dimer, this vector runs horizontally in the middle on the right. The red coloring marks segments for which the proton NMR peaks are broadened by intermolecular paramagnetic relaxation enhancements from nitroxide labeling of half of the molecules in the samples of the truncated precursor. Termini are labeled.

introduced to segments of the dimer interface by individual spin labels. The intermolecular PREs of the truncated precursor of HIV protease were modeled adequately by four-member structural ensembles in which dimers were populated to 3–5% with a similar interface with mature HIV protease but highly variable orientations [20] (Figure 2). Spin labeling near the N terminus suggests that, in a subset of the transient dimers, the N terminus makes transient contacts with the active site. This accounts for how the otherwise mostly monomeric precursor of HIV protease can process the N terminus to generate the mature full length [20].

Another association involving an encounter complex was characterized instead by NMR relaxation dispersion, which reports on equilibrium kinetics, thermodynamics, and indirectly upon structural change. Exchange within milliseconds between states in equilibrium can contribute line broadening (R_{ex}) that depends on RF pulsing rate ν_{CPMG} in CPMG experiments. R_{ex} as a function of ν_{CPMG} and magnetic field strength can be fitted to the rate of interconversion k_{ex} between 'sites' or states, the population of each state that gives their free energy differences, and the chemical shift difference $\Delta\omega$ between sites (the difference in spectral peak positions) [21]. The phosphorylated kinase-inducible activation domain (pKID) of transcription factor CREB (cAMP response element binding) is intrinsically unfolded but is structured when bound to the KIX domain of CREB-binding protein. Titrations suggested this association to involve an encounter complex that is only partly structured and that makes transient hydrophobic contacts with KIX [22]. The coupled binding and folding processes include an intermediate in the equilibrium: free \leftrightarrow encounter \leftrightarrow intermediate \leftrightarrow bound [22]. The line broadening data were fitted well by a model with three sites. The chemical shift changes $\Delta\omega$ suggested the equilibrium to be among encounter complex, intermediate, and bound state [22]. The two helices of pKID behaved a little differently, with αB featuring a dissociation that was fivefold faster. In regard to the intermediate, $\Delta\omega$ values suggested that αA nearly reached the bound structure but αB less so [22].

Dynamic energy landscape of enzyme catalysis

Rate constants k_{ex} of equilibrium conformational fluctuations, from CPMG relaxation dispersion studies, correlated elegantly with enzymatic rate constants of cyclophilin A [23] and dihydrofolate reductase (DHFR) (both chemical transformation and product release) [24]. The insights became even more profound by recognizing the structural implications of $\Delta\omega$ obtained from fitting ^{15}N relaxation dispersions of the conformational fluctuations of intermediates of DHFR. The intermediates are binary or ternary complexes with cofactor, substrate, product, or analogue [24]. $\Delta\omega$ of the fluctuations was consistently highly correlated with chemical shift changes δ accompanying the addition of a ligand to shift a binary complex to the next or previous ternary complex in the catalytic cycle. This provides strong evidence that the higher energy substate of each DHFR intermediate adopts a structure very similar to that of the dominant ground state of the next or previous intermediate in the catalytic cycle [24]. This demonstration of transient excited states (populated to a few percent) being linked to the adjacent intermediate by

protein conformational fluctuations is a landmark in understanding enzyme guidance through the dynamic energy landscape of its catalytic cycle [25].

Structures of sparsely populated conformers

The concept of conformational selection evident in the molecular recognition by DHFR [24] was recently also articulated for ubiquitin's associations with other proteins [26]. The latter case depended upon access to the structures of sparsely populated higher energy states, achieved by innovative use of RDCs. RDCs introduce peak splittings to NMR peaks of solutions aligned very weakly, typically by liquid crystalline components. The RDCs report on the orientation, relative to the small overall alignment, of bond vectors joining J-coupled nuclei [27].

An ensemble of ubiquitin conformers was fitted to many sets of backbone RDCs (arising from independent means of alignment) using a new strategy of structural calculations [26]. The diversity of the ensemble was validated by its consistency with both the conformational restriction of bonds indicated by backbone order parameters derived from the RDCs and principal components analysis [26]. The diversity of the conformations in the ensemble is so great as to encompass all of the conformations ever reported for ubiquitin complexes [26]. The regions of greater mobility and greater spread of conformations correlate very well with the sites of contacts within ubiquitin complexes. These results suggested that all of these conformations are accessed transiently, within microseconds, as sparsely populated higher energy substates that can be selected by the binding partners of ubiquitin when they associate [26].

Kay's group has developed an independent strategy using RDCs that is also a technical tour de force. The solution structure was determined for one particular 'invisible' bound state of an SH3 domain populated to approximately 5% with phosphopeptide [28]. The excited state's backbone dihedral angles were estimated from measurements of nitrogen, proton, and carbon chemical shifts by means of relaxation dispersion methods [21,29–31]. Bond orientations in the excited state were estimated from RDCs and residual carbonyl chemical shift anisotropies introduced by weak alignment, each measured using clever new relaxation dispersion schemes [32–35].

Future directions

An array of approaches, including relaxation dispersion, PREs, RDCs, NHX, and TCS, is available for revealing elusive conformational states. Remarkably good structural models of proteins up to 15 kDa can now be derived from chemical shifts [36]. As capabilities

improve for accurately predicting chemical shifts and using them to calculate structures larger than 15 kDa, the chemical shifts of invisible states accessible from CPMG spectra [31] could make structural models of invisible states more easily accessible.

Abbreviations

CPMG, Carr-Purcell-Meiboom-Gill; CREB, cAMP response element binding; DHFR, dihydrofolate reductase; NHX, native-state hydrogen exchange; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; PBD, preprotein-binding domain; pKID, phosphorylated kinase-inducible activation domain; PRE, paramagnetic relaxation enhancement; RDC, residual dipolar coupling; RF, radiofrequency; TCS, transferred cross-saturation; TROSY, transverse relaxation optimized spectroscopy.

Competing interests

The author declares that he has no competing interests.

Acknowledgments

The author appreciates the comments of Chun Tang, Shaokai Jiang, Rajagopalan Bhaskaran, and Mark Palmer on the manuscript and support from the National Institutes of Health and the American Heart Association. Chun Tang kindly provided the coordinates used to prepare Figure 2.

References

1. Wüthrich K: *NMR of Proteins and Nucleic Acids*. New York: John Wiley & Sons; 1986.
2. Jarymowycz VA, Stone MJ: **Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences**. *Chem Rev* 2006, **106**:1624-71.
3. Riek R, Pervushin K, Wuthrich K: **TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution**. *Trends Biochem Sci* 2000, **25**:462-8.
4. Shimada I: **NMR techniques for identifying the interface of a larger protein-protein complex: cross-saturation and transferred cross-saturation experiments**. *Methods Enzymol* 2005, **394**:483-506.
5. Nishida N, Sumikawa H, Sakakura M, Shimba N, Takahashi H, Terasawa H, Suzuki EI, Shimada I: **Collagen-binding mode of vWF-A3 domain determined by a transferred cross-saturation experiment**. *Nat Struct Biol* 2003, **10**:53-8.
6. Ichikawa O, Osawa M, Nishida N, Goshima N, Nomura N, Shimada I: **Structural basis of the collagen-binding mode of discoidin domain receptor 2**. *EMBO J* 2007, **26**:4168-76.
7. Sprangers R, Kay LE: **Quantitative dynamics and binding studies of the 20S proteasome by NMR**. *Nature* 2007, **445**:618-22.

F1000 Factor 3.2 Recommended

Evaluated by Mitsuhiro Ikura 10 Jan 2003, Jia-huai Wang 13 Jan 2003

F1000 Factor 8.7 Exceptional

Evaluated by Russ Hille 13 Feb 2007, Xiayang Qiu 14 Feb 2007, Antonio Rosato 15 Feb 2007, Dorothee Kern 20 Feb 2007, Gottfried Otting 28 Mar 2007, Josep Rizo 09 Aug 2007

8. Sprangers R, Velyvis A, Kay LE: **Solution NMR of supramolecular complexes: providing new insights into function.** *Nat Methods* 2007, **4**:697-703.
9. Tugarinov V, Kay LE: **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* 2003, **125**:13868-78.
10. Sheppard D, Guo C, Tugarinov V: **4D (1)H-(13)C NMR spectroscopy for assignments of alanine methyls in large and complex protein structures.** *J Am Chem Soc* 2009, **131**:1364-5.
11. Sprangers R, Li X, Mao X, Rubinstein JL, Schimmer AD, Kay LE: **TROSY-based NMR evidence for a novel class of 20S proteasome inhibitors.** *Biochemistry* 2008, **47**:6727-34.
- F1000 Factor 3.2 Recommended
Evaluated by H Jane Dyson 26 Jun 2008, Wolfgang Jahnke 15 Aug 2008
12. Keramisanou D, Biris N, Gelis I, Sianidis G, Karamanou S, Economou A, Kalodimos CG: **Disorder-order folding transitions underlie catalysis in the helicase motor of SecA.** *Nat Struct Mol Biol* 2006, **13**:594-602.
13. Gelis I, Bonvin AM, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, Economou A, Kalodimos CG: **Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR.** *Cell* 2007, **131**:756-69.
- F1000 Factor 9.0 Exceptional
Evaluated by Ian Collinson 20 Dec 2007
14. Bai Y, Sosnick TR, Mayne L, Englander SW: **Protein folding intermediates: native-state hydrogen exchange.** *Science* 1995, **269**:192-7.
15. Chamberlain AK, Handel TM, Marqusee S: **Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH.** *Nat Struct Biol* 1996, **3**:782-7.
16. Liang X, Lee GI, Van Doren SR: **Partially unfolded forms and non-two-state folding of a beta-sandwich: FHA domain from Arabidopsis receptor kinase-associated protein phosphatase.** *J Mol Biol* 2006, **364**:225-40.
17. Mittermaier A, Kay LE: **New tools provide new insights in NMR studies of protein dynamics.** *Science* 2006, **312**:224-8.
18. Iwahara J, Clore GM: **Detecting transient intermediates in macromolecular binding by paramagnetic NMR.** *Nature* 2006, **440**:1227-30.
- F1000 Factor 3.2 Recommended
Evaluated by Janet Morrow 27 Jun 2006, Vincent Hilser 01 Aug 2006
19. Tang C, Iwahara J, Clore GM: **Visualization of transient encounter complexes in protein-protein association.** *Nature* 2006, **444**:383-6.
- F1000 Factor 6.4 Must Read
Evaluated by Gideon Schreiber 24 Nov 2006, Alfred Pingoud 27 Nov 2006
20. Tang C, Louis JM, Aniana A, Suh JY, Clore GM: **Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease.** *Nature* 2008, **455**:693-6.
21. Palmer AG 3rd, Kroenke CD, Loria JP: **Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules.** *Methods Enzymol* 2001, **339**:204-38.
22. Sugase K, Dyson HJ, Wright PE: **Mechanism of coupled folding and binding of an intrinsically disordered protein.** *Nature* 2007, **447**:1021-5.
- F1000 Factor 9.6 Exceptional
Evaluated by Mitsuhiro Ikura 01 Jun 2007, Matthias Buck 02 Jul 2008
23. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D: **Intrinsic dynamics of an enzyme underlies catalysis.** *Nature* 2005, **438**:117-21.
- F1000 Factor 9.9 Exceptional
Evaluated by Kristina Downing 04 Nov 2005, Gottfried Otting 10 Nov 2005, David Cowburn 15 Nov 2005, Arieh Warshel 22 Nov 2005, Laura Itzhaki 20 Dec 2005
24. Boehr DD, McElheny D, Dyson HJ, Wright PE: **The dynamic energy landscape of dihydrofolate reductase catalysis.** *Science* 2006, **313**:1638-42.
- F1000 Factor 8.2 Exceptional
Evaluated by Deyou Zheng 25 Sep 2006, Mitsuhiro Ikura 17 Oct 2006, Andrea Mattevi 18 Oct 2006
25. Vendruscolo M, Dobson CM: **Structural biology: Dynamic visions of enzymatic reactions.** *Science* 2006, **313**:1586-7.
26. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL: **Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution.** *Science* 2008, **320**:1471-5.
- F1000 Factor 10.5 Exceptional
Evaluated by Matthias Buck 16 Jun 2008, Gottfried Otting 26 Jun 2008, Antonio Rosato 26 Jun 2008, Torleif Härd 02 Jul 2008, Moosa Mohammadi 03 Jul 2008
27. Tjandra N, Bax A: **Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium.** *Science* 1997, **278**:1111-4.
28. Vallurupalli P, Hansen DF, Kay LE: **Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy.** *Proc Natl Acad Sci U S A* 2008, **105**:11766-71.
29. Ishima R, Torchia DA: **Extending the range of amide proton relaxation dispersion experiments in proteins using a constant-time relaxation-compensated CPMG approach.** *J Biomol NMR* 2003, **25**:243-8.
30. Ishima R, Baber J, Louis JM, Torchia DA: **Carbonyl carbon transverse relaxation dispersion measurements and ms-micros timescale motion in a protein hydrogen bond network.** *J Biomol NMR* 2004, **29**:187-98.
31. Hansen DF, Vallurupalli P, Lundstrom P, Neudecker P, Kay LE: **Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: how well can we do?** *J Am Chem Soc* 2008, **130**:2667-75.
32. Igumenova TI, Brath U, Akke M, Palmer AG 3rd: **Characterization of chemical exchange using residual dipolar coupling.** *J Am Chem Soc* 2007, **129**:13396-7.
33. Vallurupalli P, Hansen DF, Stollar E, Meirovitch E, Kay LE: **Measurement of bond vector orientations in invisible excited states of proteins.** *Proc Natl Acad Sci U S A* 2007, **104**:18473-7.
34. Hansen DF, Vallurupalli P, Kay LE: **Quantifying two-bond 1HN-13CO and one-bond 1H(alpha)-13C(alpha) dipolar couplings of invisible protein states by spin-state selective relaxation dispersion NMR spectroscopy.** *J Am Chem Soc* 2008, **130**:8397-405.
35. Vallurupalli P, Hansen DF, Kay LE: **Probing structure in invisible protein states with anisotropic NMR chemical shifts.** *J Am Chem Soc* 2008, **130**:2734-5.
36. Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D, Bax A: **Consistent blind protein structure generation from NMR chemical shift data.** *Proc Natl Acad Sci U S A* 2008, **105**:4685-90.