Journal of Biomolecular NMR, **22:** 1–8, 2002.
KLUWER/ESCOM
© 2002 *Kluwer Academic Publishers. Printed in the Netherlands.*

Perspective article

Residual dipolar couplings: Synergy between NMR and structural genomics

Hashim M. Al-Hashimi* & Dinshaw J. Patel
Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, U.S.A.

Received 5 November 2001; Accepted 16 November 2001

Key words: dynamics and function, long-range constraints, multi-domain proteins, proteomics, RNA tertiary structure

Abstract

Structural genomics is on a quest for the structure and function of a significant fraction of gene products. Current efforts are focusing on structure determination of single-domain proteins, which can readily be targeted by X-ray crystallography, NMR spectroscopy and computational homology modeling. However, comprehensive association of gene products with functions also requires systematic determination of more complex protein structures and other biomolecules participating in cellular processes such as nucleic acids, and characterization of biomolecular interactions and dynamics relevant to function. Such NMR investigations are becoming more feasible, not only due to recent advances in NMR methodology, but also because structural genomics is providing valuable structural information and new experimental and computational tools. The measurement of residual dipolar couplings in partially oriented systems and other new NMR methods will play an important role in this synergistic relationship between NMR and structural genomics. Both an expansion in the domain of NMR application, and important contributions to future structural genomics efforts can be anticipated.

Structural genomics

For many years, biomolecular structures have been determined by X-ray crystallography and NMR spectroscopy to provide an atomic level basis for recognition, catalysis and regulation underlying a known biomolecular function. The birth of structural genomics and the quest for the structure and function of all gene products marks almost a reversal of this tradition, where biomolecular structures will be determined *ad hoc* for assigning functions. Protein structures have evolved to become signature markers for their functions, and it is estimated that 66% of proteins having similar folds also share similar functions despite often having little sequence homology (Koppensteiner et al., 2000). In addition to assigning known functions to new proteins through structural homology, biomolecular structures can guide the *de novo* discovery of new functions, and provide a rational framework for finding and optimizing lead compounds in drug discovery (Norin and Sundstrom, 2001).

To meet the goal of structural genomics, a number of programs have been established for large-scale protein structure determination by X-ray and NMR. Unlike sequencing efforts in genomics, structural genomics consortia do not plan to *experimentally* determine structures for every identified amino-acid sequence; it is estimated that the human genome alone encodes for 30 000 proteins (Sali, 1998). Many proteins resist characterization by X-ray and NMR, and it remains difficult to fully automate the art of protein cloning, expression, and purification. Rather, the goal is to rapidly determine anywhere between 10 000–20 000 carefully selected non-redundant protein structures, and such an expansion in 'structural space' is expected to pave the way for predicting protein structures

*To whom correspondence should be addressed. E-mail: hashimi@sbnmr1.mskcc.org

based on sequence using comparative homology technologies. Such computational approaches generate a template structure for a target sequence on the basis of its alignment to one or more related protein structures (Koehl and Levitt, 1999; Marti-Renom et al., 2000). For >50% sequence identity, modeled structures tend to have an all atom root mean square distance (RMSD) of 1 Å for main chain atoms (compared to a target experimental structure), which is comparable in accuracy to well-determined NMR structures, while structures based on 30%–50% sequence identity tend to have 90% of the main chain within 1.5 Å, with primary deviations arising in side chain alignments (Sanchez and Sali, 1998). An expansion in protein structural coverage would increase access to protein structures computationally allowing useful inferences to be drawn about gene product function at unprecedented rates and cost efficiencies. Predicting protein structures from sequence is thus no longer a distant objective, but an achievable goal that is fueling structural genomics activity (reviewed in Baker and Sali, 2001).

To determine protein structures efficiently and cost effectively, structural genomics consortia are focusing on well-folded single-domain proteins. Such proteins have a high ‘structure determination success rate’, because they tend to have a high propensity to yield diffracting crystals, and because they frequently fall within a size limit that permits NMR structure determination within a competitive time frame. In addition to determining protein structures that may resist crystallization, NMR can be used to screen proteins for ‘foldedness’ and propensity for crystallization, and to characterize intermolecular interactions in validating functional assignments (reviewed in Montelione et al., 2000; Prestegard et al., 2001). New methodologies are being used to enhance the speed of NMR structure determination. Efforts are primarily focusing on reducing the time expenditure associated with assigning resonances and side chain NOEs, by employing automated resonance assignment programs (Montelione et al., 2000; Mumenthaler et al., 1997), various deuterium isotopic labeling strategies for spectral simplification (Venters et al., 1995; reviewed in Gardner and Kay, 1999) and by relying on other sources of long-range structural constraints, including the measurement of residual dipolar couplings (RDCs) between backbone nuclei (reviewed in Prestegard et al., 2001; Bax et al., 2001). Other procedures are also being devised for establishing structural homology between a target protein and protein structures in the PDB without the need for structure determination, but through direct comparison of measured and predicted NMR parameters, such as chemical shifts and RDCs (reviewed in Bax et al., 2001; Tolman, 2001). These improvements in cost and speed efficiencies of NMR applications, and other developments in automation technologies to assist protein cloning, expression, purification, and sample conditioning, will benefit a variety of biomolecular NMR investigations conducted at the individual laboratory level. Thus, a positive synergistic relationship is evolving between NMR and structural genomics.

Structural genomics will also have other consequences on biomolecular NMR investigation. Structure determination of single-domain proteins – frequent targets of NMR structural studies – are comprehensively being pursed by large-scale structural genomics consortia operating with the benefits of economies of scale, as well as by rapid and cost effective computational homology routines. Hence, in parallel to the demands for high throughput NMR structure determination, structural genomics is creating the impetus to expand NMR applications towards larger and more complex biomolecular systems. Indeed, the majority of eukaryotic genes do not encode for aqueous single-domain proteins, but multi-domain, membrane, and ‘unstructured’ proteins, and gene products include other biopolymers such as RNA. Structural studies of these biomolecules is generally more challenging, both for NMR due to large molecular weights and broad resonances, and X-ray due to conformational flexibility and difficulties in generating well diffracting crystals. Moreover, as protein structures are determined and functions and therapeutic agents sought, studies of molecular interaction and dynamics will become increasingly important. Being a multi-faceted spectroscopic technique for examining structure, interaction, and dynamics that does not require crystallization, NMR can play an important role in these investigations. The purpose of this article is to not only discuss how recent NMR methods, and in particular the measurement of residual dipolar couplings may enhance the feasibility of these NMR investigations, but to also point out how this methodology may be tailored in the future to exploit some of the advances borne out of first generation structural genomics programs.

Residual dipolar couplings

Over the last five years, advances in NMR methodology have tremendously expanded its domain of applicability (reviewed in Ferentz and Wagner, 2000). Among the many developments, we focus on the measurement of RDCs, which provide unique long-range orientational constraints that are important for a variety of biomolecular NMR applications (reviewed in Prestegard et al., 2000; Bax et al., 2001). Dipolar couplings between pairs of nuclei depend on their distance of separation (r) and the angle (θ) between the internuclear vector and the applied magnetic field, as given by $(3\cos^2\theta - 1)/2r^3$. While the latter angular dependence bears long-range structural information that is highly complementary to traditional short-range distance constraints derived from NOEs, dipolar couplings are not detectable under solution conditions because the time average of the angular term is zero when molecular tumbling is isotrpic (random). Attempts to impart weak levels of molecular alignment in a manner that preserves the benefits of high-resolution NMR spectra date back to the 1960s, with the use of ordered liquid crystalline media to induce alignment of organic compounds (Saupe, 1968), followed by direct magnetic field alignment of organic molecules, and biomolecules having large magnetic susceptibility anisotropies (Bothner-By, 1995; Prestegard et al., 1999). A major breakthrough was the demonstration that a dilute solution of phospholipid bilayers called ‘bicelles’ (Ram and Prestegard, 1988; Sanders and Prosser, 1998; Vold and Prosser, 1996), which forms an ordered liquid crystalline phase at temperatures >25 °C, could be used to induce a tunable level of molecular alignment for a variety of biomolecular solutes (Tjandra and Bax, 1997). Even for small departures from isotropic tumbling, on the order of one molecule in a thousand being on average completely aligned, dipolar couplings (D) can be accurately measured as new contributions to splittings of resonances (J), in an aligned state ($J + D$) relative to an isotropic state (J). The introduction of bicelles as an ordering medium subsequently led to the development of an array of ordering media having diverse stability profiles and chemical compositions. While there has been some skepticism about the physiological relevance of these ordering media, it should be noted that 20–30% of cellular interiors are typically occupied by macromolecules (Ellis, 2001). Ordering media that are often composed of lipids, nucleic acids, and carbohydrates (reviewed in Prestegard and Kishore, 2001) may therefore bring solute biomolecules into heterogeneously crowded environments that actually more closely mimic the cellular interior.

The unique angular information in dipolar data can be included in traditional refinement protocols (reviewed in Prestegard et al., 2000; Bax et al., 2001) to constrain the orientation of internuclear vectors into a cone of possible solutions, defined relative to a common order tensor frame. More decisive constraints can be extracted from dipolar couplings provided knowledge is available about the local structure of individual molecular fragments in a biomolecule, and from the measurement of RDCs in two or more different ordering media. Here, the measurement of more than five independent RDCs per molecular fragment can be used to uniquely determine relative fragment orientations in the biomolecule. This unique property of RDCs that enables exploitation of previous structural information, and the ease with which RDC can be measured between readily assignable backbone nuclei, may effectively be combined with an endeavor that seeks to determine structures for ‘building block’ components in biomolecules – structural genomics is a case in point.

Quaternary organizations in multi domain/subunit macromolecules

Most genes in eukaryotes encode for multi-domain proteins rather than their single domain counterparts (domains have on average 153 residues) (Orengo et al., 1999). These domains often fold autonomously, carry out distinct functions, and are frequently observed in different protein contexts. Oligomers are a similar type of biomolecules that carry multiple copies of a given domain. In all these cases, the spatial organization of domains is critical for coordinating and sometimes enhancing function. However, for many years now, determining relative domain orientations has been challenging for both NMR spectroscopists and X-ray crystallographers. Traditional NMR short-range distance constraints derived from NOEs are not effective in constraining the geometric disposition of remotely positioned domains, and many multi-domain proteins resist crystallization or are susceptible to crystal packing forces. This, and the difficulties in characterizing domain orientations computationally have led to the exclusion of multi-domain proteins from the majority of first generation

structural genomics consortia. However, as demonstrated in a number of recent reports, domain orientations can now readily be established using long-range orientational constraints derived from measurements of RDCs (reviewed in Kay, 2001; Prestegard and Kishore, 2001; Tolman, 2001). Indeed, the expanding repertoire of multi-domain protein structures determined using RDC-NMR is highlighting important differences between domain orientations determined in solution and their solid-state counterparts, presumably due to the effects of crystal packing forces. Structure determination of multi-domain proteins under solution conditions may therefore become an increasingly important consideration in the future.

Studies of multi-domain protein systems is an important area of future collaboration between RDC-NMR and structural genomics. While RDC-NMR can readily be applied to determine relative domain orientations when having knowledge about the average backbone fold of individual domains, individual domain structures that are typically less susceptible to crystal packing forces need not come from NMR. In fact, in the first RDC application of its kind, coordinates from the X-ray structure of a homologous protein were used in the RDC determination of relative domain orientations (Fischer et al., 1999). Through experimental and computational approaches, structural genomics will expand the number of protein domain structures whose organization in the context of multi-domain protein needs to be determined. With knowledge about individual domain structures in hand, measurement of as few as five independent RDCs per domain between readily assignable backbone nuclei, and preferably in two different ordering media (Ramirez and Bax, 1998; Al-Hashimi et al., 2000) can allow rapid determination of relative domain orientation. Minimizing requirements for assigning side chain resonances and their associated NOEs would also increase the size limit of NMR application. Importantly, by exploiting the ramifications of point group symmetry on order tensor orientations, RDC-NMR can also be used to determine the relative orientation of monomeric units in homooligomers in a manner that circumvents traditional limitations associated with ambiguous intra- and inter-monomer NOEs (Drohat et al., 1999; Prestegard et al., 2000; Bewley and Clore, 2000; Al-Hashimi et al., 2001b). This may also prove useful in solution studies of solubilized membrane proteins that often have a high propensity for oligomerization, for which direct field alignment and RDC measurements have recently been demonstrated using advantageously bound lanthanide ions (Veglia and Opella, 2000). Several different types of ordering media have now been introduced for achieving partial solute alignment (reviewed in Prestegard and Kishore, 2001) and some of these may also be compatible with detergent-solubilized membrane proteins.

Transverse relaxation optimized spectroscopy (TROSY) (Pervushin et al., 1997) may allow extension of the latter RDC applications to biomolecules with large molecular weights ($>$100 kDa), thereby providing constraints on quaternary organizations of multi-protein or ribonucleoprotein assemblies (Prestegard, 1998). TROSY relies on reducing contributions to transverse relaxation through relaxation interference between the same dipolar interactions that gives rise to RDCs, and chemical shift anisotropy (CSA). In proteins, TROSY experiments can optimally be applied to backbone amide nitrogen/proton pairs, and this TROSY element has been incorporated into many NMR experiments for measuring RDCs between various backbone nuclei (for examples see Yang et al., 1999; Permi and Annila, 2000). Developments in enzymatic (Yamazaki et al., 1998) and chemical (Xu et al., 1999) ligation approaches for segmental labeling are also providing new avenues for overcoming spectral overcrowding in large molecular systems. Together, these developments may allow NMR investigation into the organization of entire protein and/or nucleic acids sub-units in macromolecular assemblies, and also shed light on concerted reorientations typically associated with the function of such macromolecules. Having previous knowledge about individual component structures in the assembly will be important for these RDC studies, and structural genomics may again be an important source for this information (Chou et al., 2000a).

Because RDCs measured between covalently bonded nuclei do not provide distance constraints, an important limitation in many of the above applications will be the definition of the translational disposition of molecular components. Such distance constraints may also be necessary to overcome the orientational degeneracy that can arise when relying solely on RDCs measured in single ordering medium. The measurement of well-selected NOEs as well as other distance-dependent relaxation phenomena will therefore be important, as will be considerations about the covalent geometry of sites linking molecular components. However, another important obstacle will also be the

ability to attain backbone resonance assignments in such large and complex molecular systems.

From structure to assignments: A reversal in trends?

Although structural genomics will determine a large number of protein structures by X-ray and comparative homology routines, many of these proteins will nevertheless undergo NMR investigation. Beside the latter applications involving the determination of component structures in the context of physiologically relevant macromolecules, NMR can also be used to investigate molecular interactions and dynamics. However, unlike *de novo* structure determination, many of these studies do not require side-chain or even complete backbone resonance assignments. In addition to RDC studies which can require as little as the assignment of five pairs of backbone nuclei per molecular fragment, perturbations in backbone amide $^{15}N/^{1}H$ chemical shifts, especially for residues clustered at the protein surface, are primarily used to probe protein interactions by NMR. Current resonance assignment strategies based on through bond correlations do not fully exploit knowledge about a three-dimensional structure and require ^{13}C labeling (especially for NMR experiments involving ^{13}C-^{13}C coherence transfers) as well as acquisition of multiple triple resonance NMR experiments. Successful acquisition and interpretation of the latter experiments also defines the size limit of many NMR applications. Can knowledge about a molecular structure be exploited to facilitate backbone resonance assignments?

The anticipated expansion in the number of protein structures having NMR resonance assignments should energize the development of computational approaches for predicting chemical shifts based on a biomolecular structure (reviewed in Case, 1999). However, RDCs may also be used in the resonance assignment step when *a priori* structural information is available. Simply put, correctly assigned resonances and corresponding RDCs should agree with a given structure, and orientational degeneracies can be reduced through acquisition of multiple RDC data in different ordering media. Recent reports use RDCs in the resonance assignments (Zweckstetter and Bax, 2001), and in one case, resonance assignment and structure determination are pursued in concert with heavy reliance on RDCs and without the requirements for ^{13}C labeling (Tian et al., 2001). More recently, it was shown that sequential assignments in an RNA A-form stem could be achieved by maximizing agreement between RDCs and expected A-form geometry for two helical stems, using resonances belonging to Watson-Crick base pairs as a basis set for exploring all possible commutations (Al-Hashimi et al., 2001a). Such approaches may allow more rapid resonance assignments, reduce requirements for ^{13}C labeling, and possibly allow applications to be extended to larger systems.

Rapid determination of RNA tertiary architecture

Although first generation structural genomics projects are targeting proteins, RNA also exhibits significant structural and functional diversity, as demonstrated by the increasing repertoire of RNA structures, both free and in complex with protein target proteins (reviewed in Hermann and Patel, 1999; Williamson, 2000). Secondary structures can readily be predicted for RNA using various thermodynamic and covariation phylogenetic analyses, but unlike for proteins, tertiary interactions in RNA are not numerous and often involve backbone interactions that are almost impossible to predict *a priori*. Tertiary structures are ultimately critical for understanding RNA function, and expanding the database of RNA tertiary architectures would facilitate future development of computational approaches for predicting RNA folds in analogy to the progress made in protein structure prediction. For these reasons, and because of the fundamental importance of RNA in all aspects of genetic processing, an RNA structural genomics program may be a worthwhile endeavor (reviewed in Doudna, 2000). NMR can be uniquely applied to investigate flexible RNA molecules, which often fail to yield diffracting crystals, or can be affected by crystal packing forces. Although the paucity of protons and NOEs in extended nucleic acids have made application of traditional NOE-based NMR difficult in the past, this has changed with the advent of modern NMR methodology and isotopic labeling of samples (reviewed in Mollova and Pardi, 2000; Zidek et al., 2001).

From the onset of RDC measurements in biomolecules, it was recognized that derivable orientational constraints could have a tremendous impact on determining nucleic acid tertiary folds (Prestegard et al., 1999). Today, applications involving structure determination of bent and oligomeric DNA structures and

RNA tertiary folds, both free and in complex with protein targets, have lived up to these expectations (reviewed in Zidek et al., 2001; Zhou et al., 1999; Tolman, 2001). However, even before RDCs made their impact on nucleic acid structure determination, the discovery of scalar couplings across hydrogen bonds replaced previously indirect methods for characterizing these structurally critical interactions with direct experimental observation (Dingley and Grzesiek, 1998; Pervushin et al., 1998; reviewed in Majumdar and Patel, 2001). The measurement of RDCs and direct detection of hydrogen bonds provide complementary information that can dramatically enhance the speed of RNA structure determination by NMR. In particular, core elements of RNA tertiary architecture can often be defined by specifying helical stem orientations. These secondary structures can readily be determined using phylogenetic analysis, confirmed using *trans*-hydrogen bond NMR spectroscopy, and modeled using idealized geometries. RDCs can then be used to rapidly determine relative stem orientations in either free RNA or protein-RNA complexes (Mollova et al., 2000; Al-Hashimi et al., 2002), and this would facilitate subsequent interpretations of NMR constraints for abounding junctions, loops, turns, and bulges – motifs that often host critical functional groups. Although ^{13}C, ^{15}N, isotopic labeling is particularly expensive for nucleic acids, the sensitivity enhancements awarded by recently introduced cryogenic probes may allow for general NMR applications using only fractionally labeled nucleic acid samples (Chou et al., 2000b; Phan, 2000).

Annotating structures with functional dynamics

Although a quest for protein function through large-scale structure determination seems largely justified, proteins and other biopolymers are not rigid, and characterizing internal mobility is also important for understanding function, and for meeting the objectives of structural genomics. The field of 'functional dynamics' is arguably in its infancy, but much has been learnt to date (reviewed in Kay, 1998; Ishima and Torchia, 2000). To mention a few examples, two proteins can have almost identical structures but different affinities towards a target ligand as well as differences in dynamics; motions can modulate binding affinities through contributions to molecular entropy. The functionally active biomolecular conformation need not be the lowest energy ground state, but one or more excited states accessible through internal fluctuations. This is the basis for many induced fit interactions, where biomolecules undergo significant conformational change upon complex formation. By regulating access to reaction centers, internal motions can also affect kinetic rates of biomolecular reactions. It is also recognized that characterizing internal motions is important for simulations of protein-protein (Robert and Janin, 1998) and protein-ligand (Gohlke et al., 2000) interactions and virtual screening. Especially for 'unstructured' biomolecules that only assume coherent structures upon complex formation with appropriate targets, dynamics cannot be neglected (reviewed in Dyson and Wright, 1998).

Over the last several years, there has been tremendous progress in the NMR characterization of backbone and side chain dynamics over a wide range of timescales. In particular, slow motions that occur at rates (10^3–10^6 s^{-1}) comparable to that of many fundamental biological processes including catalysis, allostery and complex formation, have become more amenable to spectroscopic and atomic characterization. The measurement of chemical exchange contributions to relaxation, which can probe slow motions and particularly their timescales (reviewed in Palmer et al., 2001), have now been complemented by the measurement of RDCs, which are sensitive to a wide range of motional timescales (<ms) and which can potentially provide information about amplitudes and directions of motions (Tolman et al., 1997). New approaches are being devised to simultaneously extract structural and dynamic information from RDCs (Tolman et al., 2001), some of which exploit measurements of independent sets of RDCs in different ordering media (Meiler et al., 2001), while others make use of comparisons with spin relaxation measurements (Wang et al., 2001). In many cases, collective motions of molecular fragments, whose structure is stabilized by local interactions and known *a priori*, are explored. Returning to the example of determining domain orientations in proteins, or relative stem orientations in RNA, the same RDC data can be analyzed to provide information about inter-domain/stem motions with almost no additional effort. Structural genomics and various structure prediction tools will again be an important source of independent structural information, which is required in this case to allow analysis of RDCs in terms of motions. Nevertheless, there is also great promise in completely decoupling structural and dynamic contributions to RDCs without having

any structural information and through measurements of RDCs in different ordering media (Meiler et al., 2001). Such applications will be critical for more in depth analysis of motions and particularly for enabling structural and dynamic studies of 'unstructured' proteins by NMR (Shortle and Ackerman, 2001).

Although progress in NMR methodology is clearly allowing for more detailed characterization of the manifolds of biomolecular dynamics, correlating this information to function is seldom straightforward. Functional genomics will increasingly make available valuable information about binding specificities and affinities that will greatly aid interpretation of dynamics in terms of functional parameters. Structural genomics will further provide a unique opportunity to systematically study dynamics in the context of proteins having related structures and functions. Such a comprehensive approach may be fundamental for formulating useful dynamic descriptors that relate to biomolecular function.

A brighter future for NMR and structural genomics

Regardless of its degree of success, structural genomics will dramatically change future modes of investigation in structural biology. Traditionally reductionist studies are being replaced by a more comprehensive and systematic structural biology that appropriately reflects the vastness and complexity of cellular processes. Although the initial stage of this expansion in 'breadth' is focusing on structures of individual protein folds, it will inevitably expand to include molecular assemblies, membrane, 'unstructured' and post-translationally modified proteins as well as other biomolecules. Much will also be learnt from the pace of progress in this first generation structural genomic effort, which may guide emphasis towards other aspects of biomolecules that go beyond structure, and that include stability, interaction, the chemical environment, dynamics, and perhaps new phenomena that are yet to be discovered. Being a multi-faceted approach for studying biomolecules that is not limited by requirements for crystallization, NMR is poised to make unique contributions to these endeavors. While we have reviewed how certain NMR applications may be expanded in the future through new modes of RDC-NMR investigations that exploit current developments being made in structural genomics, these and other NMR approaches may also become critical for structural genomics to effectively meet its objective of assigning functions to all gene products.

Acknowledgements

We thank Dr Ananya Majumdar and Dr Joel Tolman for a critical reading of the manuscript. This research was supported by NIH Grants GM54777 and CA46778.

References

Al-Hashimi, H.M., Gorin, A., Majumdar, A. and Patel, D.J. (2001a), submitted.
Al-Hashimi, H.M., Majumdar, A., Gorin, A., Kettani, A., Skripkin, E. and Patel, D. J. (2001b) *J. Am. Chem. Soc.*, **123,** 633–640.
Al-Hashimi, H.M., Gosser, Y., Gorin, A., Hu, W., Majumdar, A. and Patel, D.J. (2002) *J. Mol Biol.*, **315**, 95–102.
Al-Hashimi, H.M., Valafar, H., Terrell, M., Zartler, E.R., Eidsness, M.K. and Prestegard, J.H. (2000) *J. Magn. Reson.*, **143,** 402–406.
Baker, D. and Sali, A. (2001) *Science*, **294,** 93–96.
Bax, A., Kontaxis, G. and Tjandra, N. (2001) *Meth. Enzymol.*, **339**, 127–174.
Bewley, C.A. and Clore, G.M. (2000) *J. Am. Chem. Soc.*, **122**, 6009–6016.
Bothner-By, A.A. (1995) In *Encyclopedia of Nuclear Magnetic Resonance.* Grant, D.M. and Harris, R.K. (Eds.), Wiley, Chichester, pp. 2932–2938.
Case, D.A. (1999) *J. Biomol. NMR*, **15**, 95–102.
Chou, J.J., Li, S.P. and Bax, A. (2000a) *J. Biomol. NMR*, **18**, 217–227.
Chou, S.H., Tseng, Y.Y. and Chu, B.Y. (2000b) *J. Biomol. NMR*, **17**, 1–16.
Dingley, A.J. and Grzesiek, S. (1998) *J. Am. Chem. Soc.*, **120**, 8293–8297.
Doudna, J.A. (2000) *Nat Struct Biol.*, **7** (Suppl), 954-956.
Drohat, A.C., Tjandra, N., Baldisseri, D.M. and Weber, D.J. (1999) *Protein Sci.*, **8**, 800–809.
Dyson, H.J. and Wright, P.E. (1998) *Nat. Struct. Biol.*, **5**, 499–503.
Ellis, R.J. (2001) *Trends Biochem Sci.*, **26**, 597–604.
Ferentz, A.E. and Wagner, G. (2000) *Q. Rev. Biophys.*, **33**, 29–65.
Fischer, M.W.F., Losonczi, J.A., Weaver, J.L. and Prestegard, J.H. (1999) *Biochemistry*, **38**, 9013–9022.
Gardner, K.H. and Kay, L.E. 1999. In *Biological Magnetic Resonance*, Vol. 17, Krishna, N.R. and Berliner, L.J. (Eds.), Plenum, New York, NY, pp. 27–69.
Gohlke, H., Hendlich, M. and Klebe, G. (2000) *J. Mol. Biol.*, **295**, 337–356.
Hermann, T. and Patel, D.J. (1999) *J Mol. Biol.*, **294**, 829–849.
Ishima, R. and Torchia, D.A. (2000) *Nat. Struct. Biol.*, **7**, 740–743.
Kay, L.E. (1998) *Nat. Struct. Biol.*, **5**, 513–517.
Kay, L.E. (2001) *Meth. Enzymol.*, **339**, 174–203.
Koehl, P. and Levitt, M. (1999) *Nat Struct Biol.*, **6**, 108–111.
Koppensteiner, W.A., Lackner, P., Wiederstein, M. and Sippl, M.J. (2000) *J. Mol. Biol.*, **296**, 1139–1152.
Majumdar, A. and Patel, D.J. (2001) *Accounts Chem. Res.* (in press).

Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 291–325.

Meiler, J., Prompers, J.J., Peti, W., Griesinger, C. and Bruschweiler, R. (2001) *J. Am. Chem. Soc.*, **123**, 6098–6107.

Mollova, E.T. and Pardi, A. (2000) *Curr. Opin. Struct. Biol.*, **10**, 298–302.

Mollova, E.T., Hansen, M.R. and Pardi, A. (2000) *J. Am. Chem. Soc.*, **122**, 11561–11562.

Montelione, G.T., Zheng, D., Huang, Y.J., Gunsalus, K.C. and Szyperski, T. (2000) *Nat. Struct. Biol.*, **7** (Suppl.), 982–985.

Mumenthaler, C., Guntert, P., Braun, W. and Wuthrich, K. (1997) *J. Biomol. NMR*, **10**, 351–362.

Norin, M. and Sundstrom, M. (2001) *Curr. Opin. Drug. Discov. Devel.*, **4**, 284–290.

Orengo, C.A., Pearl, F.M.G., Bray, J.E., Todd, A.E., Martin, A.C., Lo Conte, L. and Thornton, J.M. (1999) *Nucl. Acids Res.*, **27**, 275–279.

Palmer, A.G., 3rd, Kroenke, C.D. and Loria, J.P. (2001) *Meth. Enzymol.*, **339**, 204–238.

Permi, P. and Annila, A. (2000) *J. Biomol. NMR*, **16**, 221–227.

Pervushin, K., Ono, A., Fernandez, C., Szyperski, T., Kainosho, M. and Wuthrich, K. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 14147–14151.

Pervushin, K., Riek, R., Wider, G. and Wuthrich, K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12366–12371.

Phan, A.T. (2000) *J. Biomol. NMR*, **16**, 175–178.

Prestegard, J.H. (1998) *Nat. Struct. Biol.*, **5**, 517–522.

Prestegard, J.H. and Kishore, A.I. (2001) *Curr. Opin. Chem. Biol.*, **5**, 584–590.

Prestegard, J.H., Al-Hashimi, H.M. and Tolman, J.R. (2000) *Q. Rev. Biophys.*, **33**, 371–424.

Prestegard, J.H., Tolman, J.R., Al-Hashimi, H.M. and Andrec, M. (1999) In *Biological Magnetic Resonance*, Vol. 17, Krishna, N.R. and Berliner, L.J. (Eds.), Plenum, New York, NY, pp. 311–355.

Prestegard, J.H., Valafar, H., Glushka, J. and Tian, F. (2001) *Biochemistry*, **40**, 8677–8685.

Ram, P. and Prestegard, J.H. (1988) *Biochim. Biophys. Acta*, **940**, 289–294.

Ramirez, B.E. and Bax, A. (1998) *J. Am. Chem. Soc.*, **120**, 9106–9107.

Robert, C.H. and Janin, J. (1998) *J. Mol. Biol.*, **283**, 1037–1047.

Sali, A. (1998) *Nat. Struct. Biol.*, **5**, 1029–1032.

Sanchez, R. and Sali, A. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 13597–13602.

Sanders, C.R. and Prosser, R.S. (1998) *Struct. Fold. Des.*, **6**, 1227–1234.

Saupe, A. (1968) *Angew. Chem., Int. Ed. Engl.*, **7**, 97–112.

Shortle, D. and Ackerman, M. S. (2001) *Science*, **293**, 487–489.

Tian, F., Valafar, H. and Prestegard, J.H. (2001) *J. Am. Chem. Soc.*, in press.

Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.

Tolman, J.R. (2001) *Curr. Opin. Struct. Biol.*, **11**, 532–539.

Tolman, J.R., Al-Hashimi, H.M., Kay, L.E. and Prestegard, J.H. (2001) *J. Am. Chem. Soc.*, **123**, 1416–1424.

Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1997) *Nat. Struct. Biol.*, **4**, 292–297.

Veglia, G. and Opella, S.J. (2000) *J. Am. Chem. Soc.*, **122**, 11733–11734.

Venters, R.A., Huang, C.C., Farmer, B.T., 2nd, Trolard, R., Spicer, L.D. and Fierke, C.A. (1995) *J. Biomol. NMR*, **5**, 339–344.

Vold, R.R. and Prosser, R.S. (1996) *J. Magn. Reson. Ser.*, **B113**, 267–271.

Wang, L.C., Pang, Y.X., Holder, T., Brender, J.R., Kurochkin, A.V. and Zuiderweg, E.R.P. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 7684–7689.

Williamson, J.R. (2000) *Nat. Struct. Biol.*, **7**, 834–837.

Xu, R., Ayers, B., Cowburn, D. and Muir, T.W. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 388–393.

Yamazaki, T., Otomo, T., Oda, N., Kyogoku, Y., Uegaki, K., Ito, N., Ishino, Y. and Nakamura, H. (1998) *J. Am. Chem. Soc.*, **120**, 5591–5592.

Yang, D.W., Venters, R.A., Mueller, G.A., Choy, W.Y. and Kay, L.E. (1999) *J. Biomol. NMR*, **14**, 333–343.

Zhou, H., Vermeulen, A., Jucker, F.M. and Pardi, A. (1999) *Biopolymers*, **52**, 168–180.

Zidek, L., Stefl, R. and Sklenar, V. (2001) *Curr. Opin. Struct. Biol.*, **11**, 275–281.

Zweckstetter, M. and Bax, A. (2001) *J Am Chem Soc.*, **123**, 9490–9491.