Residual Dipolar Couplings in Structure Determination of Biomolecules

J. H. Prestegard,*,† C. M. Bougault,‡ and A. I. Kishore†

Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30602, and Institut de Biologie Structurale, 41 rue Jules Horowitz, 38027 Grenoble Cedex 01, France

Received March 3, 2004

Contents

1. Introduction

The use of residual dipolar couplings (RDCs) in the analysis of biomolecular structure and dynamics has expanded rapidly since its potential as a source of structural information on proteins was demonstrated in the mid 1990s.[1,2] Of course, this work on proteins rested on applications to smaller biomolecular systems that occurred much earlier,[3] and even these early applications benefited from prior research on organic molecules in partially ordered liquid crystals.[4] However, in the 1990s, the existence of efficient means of introducing magnetically active isotopic labels (^{13}C and ^{15}N) and the availability of triple resonance strategies for selective manipulation and assignment of NMR resonances made widespread application to large biomolecules possible. It was fortuitous that the ^{13}C and ^{15}N labels introduced had small magnetogyric ratios, allowing simple dipolar interactions with directly bonded protons to dominate RDC observations. Prior work had focused on systems with couplings coming from the much larger $^1H-^1H$ dipolar and ^2H quadrupolar interactions. While large interactions and the resultant increased size of observable couplings may have seemed an advantage, these large interactions also lead to complex spectra and broader lines. In the case of $^1H-^1H$ interactions, additional splittings of resonances from protons at long distances arose, and in both cases broader lines resulted from enhanced spin relaxation processes.

Since the recognition of the potential of RDCs in protein structure determination, applications have spread to nucleic acid structure, carbohydrate structure, protein−ligand interactions, protein domain relationships, high-throughput strategies for structural genomics, and studies of motional amplitudes in flexible assemblies. Related pieces of data coming from interactions with paramagnetic sites and chemical shift anisotropy (CSA) offsets have also come onto the scene. Each new application demands parallel improvements in sample preparation, data acquisition, and data analysis methods.

The development of RDC applications has been reviewed periodically since their introduction to the structural biology field,[5−13] and the reader is referred to these reviews for a more complete description of the history and the underlying theory. Here, we will provide a brief introduction to RDCs and related data as they are used today. Advances that have been made in alignment techniques, data acquisition techniques, and analysis methods will be reviewed. In the course of this review, we will provide examples of applications that use these methods. Applications, per se, have become too numerous to attempt a comprehensive review.

* To whom correspondence should be addressed. James H. Prestegard, Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Road, Athens, GA, 30602. Phone: (706) 542-6281. Fax: (706) 542−4412. E-mail: jpresteg@ccrc.uga.edu.
† University of Georgia.
‡ Institut de Biologie Structurale.

10.1021/cr030419i CCC: $48.50
Published on Web 07/07/2004

James H. Prestegard was born and raised in Minnesota, and received his bachelor's degree from the University of Minnesota in 1966. After earning his Ph.D. in chemistry from Caltech in 1971 under the direction of Professor Sunney Chan, he moved directly to a faculty position in the Chemistry Department at Yale University where he remained for 27 years. He moved to The University of Georgia in 1998 where he is a Professor in the Department of Biochemistry and Molecular Biology as well as in the Department of Chemistry. His laboratories are in the Complex Carbohydrate Research Center where he holds the title of Georgia Research Alliance Eminent Scholar in Nuclear Magnetic Resonance Spectroscopy. His fundamental studies involving membrane-associating molecules dissolved in magnetically oriented phospholipid micelles fueled the now widespread interest in structural applications of residual dipolar couplings. Current applications of residual dipolar couplings in his laboratory now include determining the relative orientation of remote domains of macromolecular systems, the backbone structures of protein targets in structural genomics projects, and the geometry of carbohydrate binding in protein complexes.

Catherine M. Bougault was born in 1967 in Bourges, France. She received a national Agrégation teaching degree in Chemistry (1990), her B.Sc. degree (1991) from the Ecole Normale Supérieure de Lyon, France, and her Ph.D. degree (1994) in inorganic synthetic chemistry from the Joseph Fourier University in Grenoble, France. As a postdoctoral Lavoisier fellow with Prof. Gerd La Mar at the University of California, Davis, her interests moved to the NMR study of metalloproteins in their paramagnetic states. She became Assistant Professor in Chemistry in Joseph Fourier University in 1995 and recently (2001–2003) spent a sabbatical in Athens, Georgia, where she worked, with Prof. James H. Prestegard, on the hyperthermostability of a small metalloprotein, rubredoxin.

Anita I. Kishore was born in Carbondale, Illinois, earned a B.S. in chemistry from Yale University, and received an M.Sc. in chemistry from the University of Melbourne under the direction of Prof. Frances Separovic. She is currently a Ph.D. student in chemistry at the University of Georgia with Prof. Prestegard. Her dissertation work is on developing NMR techniques to study membrane-associated biomolecules.

2. Origin of Residual Dipolar Couplings and Complementary Observables

2.1. Residual Dipolar Couplings (RDCs)

RDCs arise when molecular systems containing proximate pairs of magnetic nuclei are partially ordered in magnetic fields. The underlying mechanism is the same as the through-space dipole−dipole coupling that dominates solids NMR spectra. For a pair of spin $^1/_2$ nuclei in a magnetic field, the distance and angle dependence are shown in eq 1 below, where r is the distance between a specific pair of nuclei, $\gamma_{i,j}$ are the magnetogyric ratios for the nuclei, μ_0 is the permittivity of space, h is Planck's constant, and θ is the angle between the considered internuclear vector and the magnetic field. When all parameters are given in SI units, the resulting D_{ij} is given in Hertz. Many measurements of RDCs are made between pairs of bonded nuclei, so that r is fixed; RDCs have, thus, been used primarily to provide angular information.

$$D_{ij} = -\frac{\mu_0 \gamma_i \gamma_j h}{(2\pi r)^3}\left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle \qquad (1)$$

The formula in eq 1 differs from the dipolar coupling expression commonly used in solids NMR applications in some important ways. First, the brackets around the angular term denote averaging over the fast molecular motion that occurs in solution or liquid crystal media. If motion allows vectors to sample directions uniformly in space, the expression reduces to zero; hence, the requirement that partial alignment be used to produce measurable values. Rapid averaging also means that, under partial alignment, interactions result in a single uniform splitting of resonances rather than a solids-like powder pattern. Second, the expression differs by a factor of 2 from that used in the solids NMR area. This is done so that values correspond directly to the increment in splitting of the doublets that would be seen in through-bond-coupled spectra of an isolated pair of spin $^1/_2$ nuclei. And third, the expression assumes that all couplings are first order (coupled spins are effectively "unlike" spins).

The effects of RDC are illustrated in the energy level diagram of Figure 1 for a pair of bonded ^{1}H and ^{15}N spins. The ^{15}N (first spin) has a negative magnetogyric ratio, and the ^{1}H (second spin) has a positive magnetogyric ratio. The common spin operator for dipolar and scalar parts ($2H_zN_z$) produces the perturbations shown to the right of the initial energy

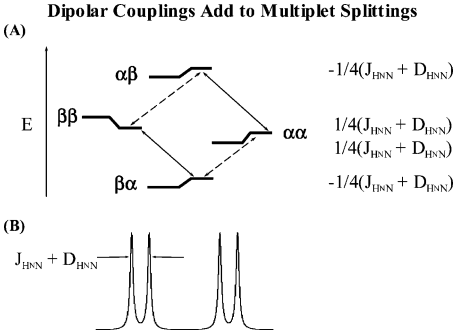

Figure 1. (A) Energy level diagram for a $^1H-^{15}N$ spin system. The dashed arrows are ^{15}N transitions, and the solid arrows are ^1H transitions. The effects of scalar and dipolar couplings, assuming a negative $J + D$ value, are denoted to the right of the diagram. (B) The expected ^{15}N and ^1H doublets are shown at the bottom.

levels as dictated by just a Zeeman interaction with an external magnetic field (energies not to scale). Because of the identical spin operators, RDCs add to scalar couplings (J_{HNN}) to produce splittings of $J_{HNN} + D_{HNN}$. As drawn in the diagram, the sum of J_{HNN} and D_{HNN} would have to be negative. The actual sum could, of course, be either positive or negative, depending on the nature of molecular orientational averaging. Two resulting doublets (one at the ^{15}N frequency and one at the ^1H frequency) are shown at the bottom. Splittings are usually measured under isotropic (J_{HNN}) and aligned ($J_{HNN} + D_{HNN}$) conditions to isolate the RDC contribution.

The need for partial alignment actually introduces additional unknowns into the above expression, making analysis of RDCs in terms of angular constraints less straightforward than the simplicity of eq 1 would imply; at a minimum, there must be parameters specifying the level and direction of order. The order also need not be axially symmetric, making it necessary to specify the orientation of a complete ordering frame relative to a molecule fixed coordinate system, as well as the extent and asymmetry of order. Specifying order results in the appearance of five independent variables as opposed to the single angular variable explicitly occurring in eq 1.

Following a practice set in the liquid crystal literature, these five variables can be introduced as elements of a 3×3 order matrix (S_{kl}, eq 2).[14] $D_{\max\ ij}$ is the coupling for a pair of nuclei at a 1.0 Å separation with their internuclear vector along the magnetic field, and the $\cos(\theta_{k,l})$ are the direction cosines relating the internuclear vector to the axes (x,y,z) of the order matrix frame. Because the matrix is traceless and symmetric, only five elements are independent. These five elements can also be viewed as the three Euler angles needed to transform molecular representations from the initial molecular frame to the principal frame in which the order matrix is diagonal, plus a principal order parameter (S'_{zz}) and an asymmetry parameter ($\eta = (S'_{xx} - S'_{yy})/S'_{zz}$) for this new frame. The elements of order matrices are dimensionless and of magnitude 10^{-3} for systems discussed here, reflecting the fact that motional averaging reduces observable dipolar couplings by approximately 3 orders of magnitude from their static counterparts ($^1H-^{15}N$ RDCs observed are on the order of ±25 Hz as opposed to the maximum coupling of 24 350 Hz that would be observed for a static directly bonded pair at 1.02 Å oriented in the direction of the magnetic field).

$$D_{ij} = \frac{D_{\max\ ij}}{r^3}\sum_{k,l} S_{kl}\cos(\theta_k)\cos(\theta_l) \qquad (2)$$

Some alternate practices for specifying molecular order have also arisen. One stems from an assumption that the principal alignment frame may be known. Once in the principal alignment frame, all S'_{ij} for $i \neq j$ are zero and eq 2 can be simplified. The remaining S'_{ii} are replaced with new variables such as an axial alignment parameter, D_a, and a rhombicity parameter, R. These new parameters are simply related to order parameters as follows: $S'_{zz} = 2D_{a\ ij}/D_{\max\ ij}$, and $\eta = 3/2\ R$.

$$D_{ij} = \frac{D_{a\ ij}}{r^3}\left[(3\cos^2\theta - 1) + \frac{3}{2}R\sin^2\theta\cos(2\phi)\right] \qquad (3)$$

For procedures that generate structural solutions by minimizing error functions through rotation of a molecular fragment toward a principal alignment frame, writing trial RDCs in the form of eq 3 is convenient. It is also convenient in the case of magnetic alignment of a molecule containing a paramagnetic center, where approximate directions for axes of a susceptibility tensor may be known from local coordination geometry. In this latter case, $D_{a\ ij}$ and R are related to the magnitude and rhombicity of a susceptibility tensor.[15] Here $D_{a\ ij} = (D_{\max\ ij}/2)$-$(B^2\Delta\chi/(15kT\mu_0))$ and $R = \delta\chi/\Delta\chi$, where $\Delta\chi$ is the axial anisotropy of magnetic susceptibility and $\delta\chi/\Delta\chi$ is the rhombicity. Note that a dependence on magnetic field squared enters (B^2). This points to the potential importance of an alignment mechanism dependent on inherent anisotropic magnetic susceptibilities of molecules of interest as higher field magnets become available.

2.2. Chemical Shift Anisotropy (CSA)

There are other partial alignment effects on NMR spectra that can complement RDC information in the sense that they display similar angular dependencies. Among these effects are resonance offsets that arise from anisotropies in chemical shielding tensors.[16−19] The chemical shielding of nuclei in most molecular groups varies with orientation in a magnetic field due to the anisotropy of the group's electronic distribution. Normally in solution NMR, only the isotropic average of the resonance position is seen, δ_{iso}. However, offsets to the isotropic chemical shift under partial alignment can be significant for nuclei in chemical groups with particularly large CSAs, ^{13}C in a carbonyl group, for example. These offsets can be used to place constraints on geometries of molecular models.

The chemical shift offset, δ^{csa}, can be expressed as shown in eq 4, where the δ_{kk} are the principal elements of the anisotropic part of the chemical shift tensor, θ_{ik} and θ_{jk} are the angles between the prin-

cipal axes of the shift tensor and an arbitrary molecular frame, and S_{ij} are the elements of the order tensor in the molecular frame. The values of δ_{kk} are assumed to be well defined in a frame oriented in an individual molecular group. In practice, the values and directions of their principal axes have been taken from suitable model compounds examined in the solid state. As an alternate procedure, accurate ab initio calculations of chemical shift tensors seem to offer promising results.[20,21]

$$\delta^{csa} = \delta_{align} - \delta_{iso} = \frac{2}{3}\sum_{i=x,y,z}\sum_{j=x,y,z}\sum_{k=x,y,z} S_{ij}\cos(\theta_{ik})\cos(\theta_{jk})\delta_{kk} \quad (4)$$

The similarity of eq 4 to the RDC eq 3 can be recognized when eq 4 is re-expressed as the sum of two pseudo-dipolar couplings in the principal alignment frame (eq 5).[16] This equation uses the same alignment parameters as eq 3. It presumes that one knows the angles between the principal alignment frame and the principal axes of the shift tensor for a group of interest (or that one can find them through a simulated annealing procedure).

$$\delta^{csa} = \frac{1}{3}\Big[D_a(2\delta_{xx} + \delta_{yy})\Big\{(3\cos^2\theta_1 - 1) + \frac{3}{2}R\sin^2\theta_1\cos(2\phi_1)\Big\} + D_a(2\delta_{xx} + \delta_{yy})\Big\{(3\cos^2\theta_2 - 1) + \frac{3}{2}R\sin^2\theta_2\cos(2\phi_2)\Big\}\Big] \quad (5)$$

Applications of CSA offsets as structural restraints have largely been confined to solid-state NMR where offsets are large and more easily measured than in weakly ordered solution NMR. The anisotropic contribution to the chemical shift in the case of weakly ordered systems requires high resolution and precise chemical shift referencing. Chemical shift offsets have been measured for carbonyl carbons in both glycolipids and amphipathic peptides strongly oriented by incorporation into phospholipid bicelles[22,23] and in both low[24] and high molecular weight proteins[16] weakly oriented by collisional interactions with alignment media. Phosphorus chemical shift offsets in oligonucleotides[17] and nitrogen shift offsets in proteins[18,25] have been measured in a variety of alignment media. One of the primary difficulties in the use of CSA offsets is that both an isotropic reference spectrum and an aligned spectrum must be collected under conditions in which variations in environment (such as temperature) introduce no additional chemical shift perturbations; differences in peak positions between aligned and reference spectra can then be associated solely with CSA offsets. An interesting approach to solving this problem is to compare spectra of the same sample taken under static and spinning conditions, where spinning at the magic angle destroys the ordering of most liquid crystalline media and allows the dipolar interaction to average to zero.[26]

2.3. Pseudocontact Shifts in Paramagnetic Systems

A similar set of equations results for resonance offsets seen in systems carrying a magnetically anisotropic paramagnetic center. These systems exhibit not only field-induced alignment, but also chemical shift offsets that depend on pseudocontact (or dipolar) shifts, δ^{pc}.[27,28] First-order contributions to pseudocontact shifts do not actually depend on molecular alignment but come from a net electron spin moment that depends on a molecule fixed susceptibility tensor. We include a discussion here because of the complementary nature of the data. Pseudocontact shifts are distinct from through-bond contact shifts that occur over short distances. Contact shifts can also be useful, but they depend on electronic structure theories for interpretation.[15,29,30] The form of the equation for pseudocontact shifts (eq 6) also bears a marked similarity to eq 3 with axial and rhombic susceptibilities being substituted for axial alignment and rhombicity parameters. The electron-to-nucleus distance (r) also appears in place of an internuclear distance. There is a subtle difference in that the susceptibilities used here are those for the paramagnetic center and not the total (paramagnetic plus diamagnetic) susceptibilities that would be used in an expression for RDCs resulting from magnetic alignment. The expression also lacks a dependence on nuclear properties and gives resonance offsets in dimensionless units appropriate for measurement in parts per million. Measurement of the offset requires comparison to a diamagnetic analogue having the same geometry and charge distribution. While preparation of a suitable reference sample can be a challenge, there are illustrations of successful measurement,[28,31−33] and when successful, the data provide both angular and long-range distance constraints.

$$\delta^{pc} = \frac{1}{12\pi r^3}\Big[\Delta\chi(3\cos^2\theta - 1) + \frac{3}{2}\delta\chi\sin^2\theta\cos(2\phi)\Big] \quad (6)$$

2.4. Cross-Correlated Relaxation

More recently, cross-correlated relaxation interference has also provided angular information in a form similar to RDCs. In particular, interference between dipole−dipole interactions and Curie interactions has been used to some extent.[32,33] While observation of this interference, like observation of pseudocontact shifts, does not require partial alignment, we include a discussion here because of the complementary nature of the data. Cross-correlated relaxation arising from interference between a contribution from a paramagnetic center-nuclear dipole interaction (the Curie contribution dominating paramagnetic at higher magnetic fields) and a contribution from a nuclear dipole−nuclear dipole interaction can be expressed as in eq 7. Here i is usually a proton and j is an ^{15}N or ^{13}C site in a paramagnetic protein, r_i is the distance between the paramagnetic center and the j spin, ω_i is the proton resonance frequency, τ_r is the rotational correlation time of the protein, and θ is

the angle between the ^1H-X vector and the electron−proton vector. The effects are very similar to the effects of CSA and dipole−dipole cross-correlation exploited in TROSY experiments, and, in fact, cross-correlated relaxation can be measured from differences in line widths of spin coupled multiplets. Analysis of dipole−dipole/CSA interference can give similar angular information.[34] The angular dependence for both effects is again of similar functional form to RDCs but complementary in that the vectors involved are different in each case.

$$\eta^{CCR} = \kappa \frac{(3\cos^2\theta - 1)}{r_i^3}\left(4\tau_r + \frac{3\tau_r}{1+\omega_i^2\tau_r^2}\right) \quad (7)$$

Recently authors have used RDC, pseudocontact shift, cross-correlated relaxation and paramagnetic relaxation enhancement data together to develop a novel strategy for fast NMR resonance assignment in ^{15}N HSQC spectra of proteins, whose structures are known from X-ray. This strategy yields the magnetic susceptibility tensor at the same time.[35] As applications progress, there are certain to be new sources to exploit for orientational and angular information.

3. Alignment of Samples

Partial alignment of samples is the key to the observation of RDCs, as well as other observables that come from anisotropies in spin interactions. The need for this alignment is embodied in the dependence on order parameters that average to zero any time molecules sample an isotropic distribution of orientations in a time short compared to the reciprocal of the interaction. Early applications used high magnetic fields and inherent anisotropies in magnetic susceptibilities of the molecules of interest to directly induce nonisotropic distributions,[1] or they used the field-induced order of liquid crystalline media to indirectly induce nonisotropic distributions through collisional interactions of the molecules of interest.[22] These early alignment methods are depicted in Figure 2 for the case of field-induced order of myoglobin (Figure 2A) and collision-induced order of a two-domain fragment of barley lectin in a bicelle medium (Figure 2B). A number of different media for aligning samples exist, and the most commonly used media are described here.

3.1. Bicelles

Bicelles were the first medium used to collect RDCs in biomolecules. The actual physical nature of the bicelle medium depicted in Figure 2B has come into question recently.[36] The original model for bicelles composed of mixtures of dimyristoylphosphatidylcholine and dihexanoylphosphatidylcholine (3:1) is one of lipid bilayer disks 30−40 nm in diameter, having the shorter chain lipid concentrated at the edges. The model is actually supported by a variety of data on closely related systems including low angle X-ray scattering data,[37] data on the anisotropic diffusion of solvent about the disks,[38] NMR data on the orientation and distribution of the two lipids

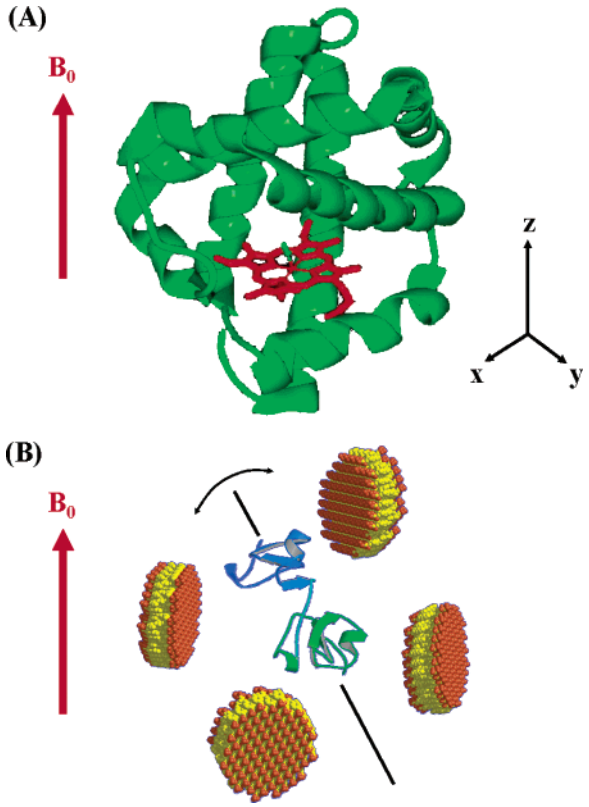

Figure 2. Partial alignment of biomolecules: (A) myoglobin oriented by interaction of its anisotropic susceptibility tensor (approximate axes shown) with the magnetic field (B_0); (B) a two-domain construct from barley lectin oriented by collisions with bicelles. The protein's long axis tends to align parallel to the bicelle surfaces.

making up the particles,[39,40] as well as electron microscopy, fluorescence, and light scattering data on particle size.[39] Nevertheless, these data are often on samples prepared at higher concentrations than those typically used for high resolution biomolecular studies, or they are on somewhat different lipid compositions, or interpretation relies on model-dependent theories to relate less direct observations to the question at hand. Recently, an alternate "Swiss cheese" model has been proposed in which bilayers are punctuated by holes lined with the shorter lipid.[36] Support for this model depends on relating less direct diffusional data for a lipid-soluble probe to the structural model. Debate about the actual structures that exist may continue for some time,[41,42] but both models share a basic bilayer structure oriented with the normal perpendicular to the magnetic field. Both are, therefore, capable of imparting order though collisional interactions of soluble, nonspherical molecules with the bilayer surfaces.

3.2. Bacteriophage

The options for aligning soluble molecules for the measurement of RDCs have increased dramatically in the years following initial applications of bicelles. Use of bacteriophage followed quickly and proved particularly useful for the study of nucleic acid systems as well as protein systems.[43−46] The filamentous phage used in these studies are up to 1000 nm in length and of order 10 nm in diameter. They are

Table 1. Alignment Media Commonly Used to Measure Residual Dipolar Couplings

medium	molecular species	charge	temp range (°C)	features and limitations	ref
ester-linked phospholipid bicelles	DMPC/DHPC	neutral	27−45	+ easy preparation − expensive, susceptible to hydrolysis	2 183
ether-linked phospholipid bicelles	DIODPC/CHAPSO DIODPC/DIOHPC	neutral	10−55	low pH	184 185
phospholipid bicelles doped with charged lipids	DMPC/DHPC/ CTAB/SDS	CTAB, positive SDS, negative	27−40		69
poly(ethylene) glycol ether bilayers	C_nE_m/*n*-alcohol	neutral	0−60	+ easy preparation, inexpensive + highly compatible with biomolecules − kinetics of alignment esp. with dissolved biomolecules unknown	186
poly(ethylene) glycol ether bilayers doped with charged lipids	C_nE_m/*n*-alcohol/CTAB/SDS	positive, negative	0−60		109
bacteriophage	rod-shaped viruses	negative	5−60	+ easy preparation, sample recovery − only suitable for negatively charged biomolecules	43, 44 187, 188
purple membranes	cooperative anisotropic membranes	charged	<70		189, 190
stretched or strained polyacrylamide gels	polyacrylamide gels	neutral	5−45	+ easy sample recovery + can accommodate larger MW (esp. membrane) proteins − difficult to align homogeneously − strong steric interactions cause broad lines	47 48 89 191
charged polyacrylamide gels	acrylamide/acrylate	charged	5−45	+ decreased line broadening − delicate and easily ruptured	50 51
immobilized media	gel- or polymer-stabilized purple membranes or phage	neutral		+ fixed director orientation	48, 192, 193
lanthanide ions/ Ln-binding tags	align by anisotropy of susceptibility			+ no compatibility problems − very small degree of alignment	54 56, 57
Helfrich phases	CPyBr/*n*-hexanol/NaBr	neutral	0−70	+ stable, wide temperature range − very sensitive to salt, buffer, pH	194 195

covered by a coat protein and are highly negatively charged. Recovery of sample by high-speed centrifugation of phage particles is possible.

3.3. Polyacrylamide Gels

Another distinctly different alignment medium is based on anisotropically compressed polyacrylamide gels[47,48] in which alignment is independent of the magnetic field direction. Orientation is based on collisional factors much like the initial bicelle systems, but the collisional barriers are strands of polyacrylamide instead of bilayer surfaces. Strands are given a preferred direction by casting a gel of a diameter larger or smaller than the diameter of an NMR tube and then compressing or stretching the gel inside the tube. Suitable devices for facilitating this process have been described.[49] The time required to diffuse a sample into the gel can be limiting, but dialysis after observation also presents a convenient option for sample recovery. Adding a charge to the gel allows somewhat better line shape for signals observed and provides more options for varying alignment properties.[50,51]

3.4. Other Media

There is a wide variety of other media in use. Many of them derive from media described above. There are, for example, a number of bicelle-like media that use different combinations of amphiphilic molecules. Some, such as those based on alkyl-poly(ethylene glycol)/alcohol mixtures, are proving easy to use and low in cost. In Table 1 we attempt to summarize the current options for alignment media.

3.5. Practical and Theoretical Considerations

The identification of suitable media for a particular application is not necessarily trivial. It is not simply sufficient that media do not perturb molecular structures; they must also induce a proper level of align-

ment. Alignment must be sufficient to give measurable RDCs but not so large as to introduce spectral complexity. A principal order parameter of 10^{-3} will, for example, give a maximum $^{15}N-^{1}H$ splitting for a directly bonded amide pair of about 25 Hz. This is large compared to typical line widths, particularly in the ^{15}N dimension of a $^{1}H-^{15}N$ HSQC spectrum, but it is not so large that long-range interactions begin to broaden resonances or cause second-order distortions of one bond couplings. Alignment is also weak enough that INEPT transfers optimized for one-bond scalar couplings in TROSY- or HSQC-based experiments do not begin to fail due to mismatch of transfer delays to total couplings. Simple adjustment of media concentration is sometimes enough to scale alignment to a proper level. However, the concentration range over which cooperative alignment occurs at all is small for some media.

Several factors beyond simple concentration of the orienting medium must also be taken into account when attempting to predict the level of order. For example, the overall charge and charge distribution of a protein must be considered when attempting to orient it in an electrically charged medium; a positively charged protein will, for example, interact strongly with negatively charged filamentous phage, leading to broad lines and poor resolution. A highly asymmetric charge distribution (large quadrupole moment) will also lead to greatly enhanced RDCs. In some cases, problems with strong charge-induced association or orientation can be alleviated by raising ionic strength. However, this solution can be problematic with less salt-tolerant high-sensitivity cryogenic probes. In addition to electrostatic considerations, hydrophobic patches in biomolecules can associate with amphiphilic alignment media. It is clear that a number of media are necessary if one is to find media that are compatible with the biomolecule, spectroscopically useful, and convenient to prepare.

One solution that may offer promise for the future is the use of media-free, field-induced orientation of biomolecules. Besides the initial application to myoglobin,[1] there have been several examples of induced alignment that take advantage of inherent metal binding sites in proteins or nucleic acids.[35,52−55] Paramagnetic ions in appropriate sites have a large anisotropic magnetic susceptibility that results in the molecule containing the site having a preferred alignment in a magnetic field. RDC contributions to splittings of $^{15}N-^{1}H$ doublets in the original work on cyanometmyoglobin, which has a highly anisotropic Fe(III) center, were slightly less than 3 Hz at 750 MHz. However, field-induced alignment goes up as field squared, suggesting that contributions would be more than 4 Hz at 900 MHz. Paramagnetic ions can also be substituted for nonparamagnetic native ions. Certain lanthanides have proven useful in this respect. Proteins can be engineered to have metal-binding tags that bind lanthanides with very high affinities.[56,57] RDC contributions to splittings for some of these systems have exceeded 10 Hz at 800 MHz. Contributions of this magnitude can easily be measured.

One should also not forget that diamagnetic anisotropies can be large if anisotropic entities such as aromatic rings are arranged to coherently add to total anisotropic susceptibilities.[58] A 16-base pair DNA double helix, for example, has about the same anisotropic susceptibility as the myoglobin molecule discussed above. Use of natural orienting properties minimizes concern about media-induced anomalies.

In addition to finding a single compatible medium, it is very useful to have the option of using several alignment media in the course of a study. Dipolar couplings suffer from the multivalued nature of the dipolar coupling function. Even after transforming to a principal alignment frame, two different θ values (in a 0−180° range) can give the same value for the $(3\cos^2\theta - 1)/2$ RDC function. The direction of a single interaction vector in a molecular fragment can, therefore, be defined no better than being on the surface of two opposing cones having these values of θ as half angles. Similarly, equations for RDCs expressed in terms of order parameters have five independent parameters that must be determined before the orientation of a molecular fragment in a principal alignment frame can be determined. This limitation can be partially addressed by collecting data for several (five or more) different interaction vectors in a molecular fragment of known geometry ($^{15}N-^{1}H$, $^{13}C-^{1}H$, $^{13}C-^{15}N$, etc. for a residue in a protein).

The degeneracy problem does not unfortunately stop at this point. Principal order parameters (dependent on $(3\cos^2\theta_{x,y,z} - 1)$) are insensitive to inversion of axis directions. It becomes clear that an order frame can equally well be described by any one of four possible combinations of positive and negative x, y, and z axis directions that yield a right-handed Cartesian coordinate system. One way to lift this degeneracy and determine a unique orientation of a molecular fragment is to use at least two alignment media with different alignment tensors.[59]

A second reason for using multiple media is related to the feasibility of collecting sufficient RDCs in a fragment of known geometry. In some cases, it is not possible to collect the minimum five RDCs required to define all independent members of an order matrix. Even if there is an adequate number of interaction vectors, they can be accidentally collinear and provide redundant information.[60] A second aligning medium can aid in gathering additional RDCs. It is difficult to see how this would work for a single isolated fragment since new axis directions and principal order parameters are introduced. However, the fact that all fragments of a rigid molecule must share the same alignment frame and have identical order parameters means that new alignment variables are introduced only once. Extreme examples of this strategy exist where structures have been determined by collecting only $^{15}N-^{1}H$ data but doing so in many alignment media.[61,62] Use of multiple media also facilitates the analysis of internal motion properties of biomolecules. This analysis introduces additional parameters that must be evaluated in any RDC-based analysis.[9,51,63−67]

A third reason for using multiple alignment media is for the elimination of the occasional concern about media-induced distortion of molecular geometries. Significant distortions as a direct result of an interaction between the medium and the molecule of interest are unlikely because the energy of interaction required to induce order of one part in 10^3 is extremely small. However, it is possible to preferentially orient populations of minor states that are in rapid equilibrium with major states. An average RDC is observed in these cases, and it can be dominated by properties of the minor species.[68] Collecting RDCs in multiple media can provide some protection against improper interpretation of measurements skewed by a minority of conformations. It is unlikely that preferential orientation of the same minority species will occur in all media, and derived conformations will be inconsistent if significant problems of this type occur.

To achieve most of the above advantages, each aligning medium must orient the biomolecule in a significantly different way so that the alignment tensors are not coincident. Simple solutions such as increasing the concentration of a liquid crystal from a lower to a higher percentage, for example, will not work; this will simply scale the dipolar couplings to larger values and will not, in general, provide new orientational restraints for use in structural determination. Use of media with perpendicularly oriented directors also does not yield new information. In alignment media that orient with their directors 90 degrees apart, for example, perpendicular and parallel to the magnetic field, elements of the alignment tensors will simply be scaled by a factor of $-1/2$. Alignment media that orient the biomolecule through primarily similar steric interactions also tend to yield similarly oriented alignment tensors. For example, aligning a protein in either poly(ethylene glycol) ether or phospholipid bicelles may give alignment tensors differing primarily by a scaling factor. However, by changing the nature of the liquid crystalline interactions with the biomolecule, for example, by changing the electrostatics of the environment, unique alignment tensors can be generated. This is easily accomplished by doping the bicelles or poly(ethylene glycol) ether with a charged lipid[69] or by combining RDC data sets from phage and bicelles.[70] In any event, effective use of multiple alignment media in a study of biomolecular structure or dynamics benefits greatly from the availability of a large number of alignment media.

4. RDC Data Acquisition

Initial attempts at data acquisition concentrated on the measurement of one bond $^1H-^{15}N$ and $^1H-^{13}C$ RDCs ($^1D_{HN-N}$ and $^1D_{C-H}$) in proteins and nucleic acids.[1,71−73] The choice of these particular couplings was dictated by practicality. The H_N-N and $H-C$ internuclear distances are rather short (1.02 and 1.08 Å) and reasonably invariant. Thus, they exhibit substantial dipolar interactions dominated by directly bonded spins, and interactions could be interpreted in terms of angular constraints without concern about distance variations. Splittings could be seen in high sensitivity HSQC spectra with minimal modifications of pulse sequences. Also, H_N-N pairs are well distributed throughout proteins providing global structural information.

4.1. One-Bond H_N–N and C–H RDCs

Early studies on field-aligned biomolecules, where one-bond RDCs were expected to be on the order of only a couple of Hertz, focused on the design of experiments achieving precision of measurement on the order of a couple of tenths of Hz. These experiments, based on modified HSQC schemes, were essentially divided into two categories: frequency-resolved methods, where the coupling is observed as a difference in positions of the two lines of a doublet in a frequency domain, and intensity-based methods, where the coupling is encoded in the signal intensity.[74] Most couplings were measured from frequency or intensity modulation in the indirect heteronuclear dimension because less efficient spin relaxation gives narrower lines, and artifacts from $^1H-^1H$ couplings, and cross-correlated relaxation effects could be minimized. The introduction of alignment media and tunable degrees of order reduced concern about precision of the measurement but did not fundamentally alter these two approaches. As a result, a large number of experimental schemes are available for the measurement of $^1D_{HN-N}$ and $^1D_{H-C}$. These experiments have been extensively discussed in previous reviews for small and medium-size proteins,[6−8,11] for RNAs,[75] and for oligosaccharides.[76,77] In practice, the selection of a particular sequence depends on the field at which the data are collected, the degree of induced alignment, the level of isotopic labeling, and the size of the molecular system studied.

The use of higher degrees of alignment ($^1D_{HN-N}$ of up to 40 Hz can be obtained in bicelles) is not without negative consequences. In protonated samples, for example, where the proton density is quite high, such conditions cause severe ^1H line broadening resulting from unresolved $^1H-^1H$ dipolar interactions. Reduction of signal intensity as well as line shape distortions can also be observed in the indirect dimension due to long-range heteronuclear $^1H-^{15}N$ or $^1H-^{13}C$ couplings. Some more recent developments have been directed at reducing these effects. To reduce distortions of the ^1H line shape, vander Kooi et al.[78] suggest the use of band-selective ^1H homonuclear WURST decoupling during acquisition. To solve the problem of line splittings or line broadening in the indirect dimension, Feher et al.[79] and Pham et al.[80] propose the inclusion of a G-BIRD sequence in the middle of the t_1 interval, respectively, in modified sensitivity-enhanced HSQC and *J*-modulated HSQC sequences.

In the case of small-to-medium size biomolecules, intensity-based and frequency-resolved 2D HSQC schemes have both been used. In general, intensity-based methods have the advantage of offering high precision for the measurement of RDCs, but they may suffer from considerable loss of accuracy if systematic errors, such as relaxation-induced errors, pulse imperfections, or presence of passive couplings, are not carefully considered.[81−83] In frequency-resolved experiments, phase distortion can complicate measure-

ments, and accuracy is highly dependent on digital and inherent resolution.[84] Some efforts at comparison of methods have been made. For example, values of $^1D_{HN-N}$ were extracted, for a galectin-3C (142 amino acids) sample, from phase-encoded (intensity-based) HSQC and coupling-enhanced (frequency-resolved) HSQC data sets. A root-mean square deviation between sets of 1.1 Hz for isotropic and 2.1 Hz for aligned samples suggests that both measurements are precise and in good agreement with one another; however, it was possible to collect the intensity-based data in a somewhat shorter time.[85] Agreement of measured couplings with couplings back-calculated from a crystal structure was substantially less than precision would indicate. This is very likely due in part to "structural noise" associated with improper modeling of local peptide geometry in terms of uniform bond lengths and bond angles.

As the size of the biomolecule increases, the number of cross-peaks in even the simplest 2D-spectra becomes large, and the appearance of pairs of lines for each site in frequency-resolved approaches becomes undesirable. In the last five years, several techniques have been used to achieve spin-state separation/selection, in the indirect dimension. In this context, IPAP, spin-state (S^3E, S^3CT), and α/β-selection elements have been included in sensitivity-enhanced pulse schemes. Key features of corresponding elements are presented in Figure 3.

In the IPAP sequences,[86−90] an antiphase doublet is generated by the inclusion of the element presented in Figure 3 at the beginning of the t_1 evolution. Addition/subtraction of the obtained spectrum to the in-phase doublet generated in the absence of this element separates upfield and downfield components into two spectra from which frequency positions can be measured, each with half the number of peaks. In the S^3-type (refs 78, 91, and 92 and references therein) and α/β-type experiments,[92,93] spin-state selective excitation is achieved through manipulation of pulse phases during the INEPT transfer, and the reverse INEPT phase, as to avoid spin-state mixing. All these sequences are prone to relaxation-induced artifacts resulting in spurious peaks at the position of the canceled resonance and are sensitive to J-mismatch. Elegant modifications have been proposed such as the double in-phase single antiphase-(DIPSAP) experiment which corrects for the latter problem.[94] A last alternative to circumvent the problem of overlap in the J-coupled spectra has been presented very recently. In the J-evolved heteronuclear transverse relaxation optimized spectroscopy (JE-TROSY),[95] J-coupling of the sharpest cross-peak multiplet component selected in a TROSY experiment is allowed to evolve during the INEPT transfer in addition to during the heteronuclear chemical shift evolution time. As a result, the value of the $^1H-^{15}N$ or $^1H-^{13}C$ coupling can be determined from the frequency displacement from the zero frequency in a third dimension. A disadvantage results from the additional dimension; nevertheless, the acquisition time is comparable to that of other J-modulated experiments.[96]

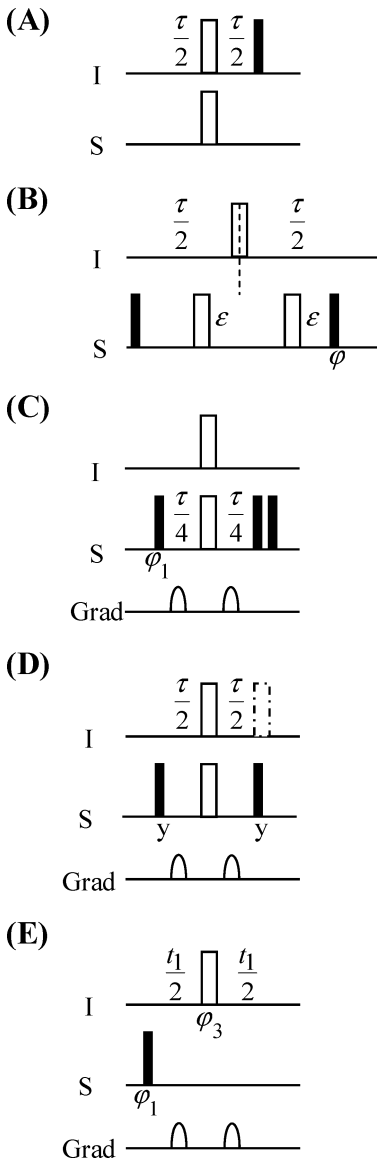

Figure 3. Pulse sequence elements used for the separation of doublet components in frequency resolved HSQC/TROSY RDC measurements. (A) Element added to collect the antiphase component of the IPAP at the beginning of the t_1 period; $= 1/2J_{IS}$.[86] (B) DIPSAP element. Three datasets need to be recorded with (i) $\epsilon = \tau/4$, $\varphi = x$; (ii) $\epsilon = 0$, $\varphi = x$; (iii) $\epsilon = \tau/8$, $\varphi = y$; $\tau = 1/J_{IS}$.[94] (C) S^3E[205] and (D) S^3CT[224] elements inserted just before the t_1 period with $= 1/2J_{IS}$. (E) α/β filter element with $\varphi_1 = x, -x$ and $\varphi_3 = x, x, -x, -x$.[198]

For medium-size to large ^{15}N- or $^{15}N,^2H$-labeled biomolecules, combining TROSY elements[97] with the measurement of RDCs is very appealing. Differential cross-correlated relaxation of the downfield and upfield components of a $^1H-^{15}N$ doublet means that precision of measurement will be limited by the broadest component. It would be ideal to measure coupling simply from the sharpest (TROSY) components. Because coupling inherently mixes fast and slow relaxing line components, however, it is difficult to retain the full TROSY advantage. Nevertheless, much progress has been made.[84,91] RDCs measured in a TROSY-HSQC pair of experiments clearly surpass those measured in the IPAP-HSQC. The precision at which $^1D_{HN-N}$ can be measured with the

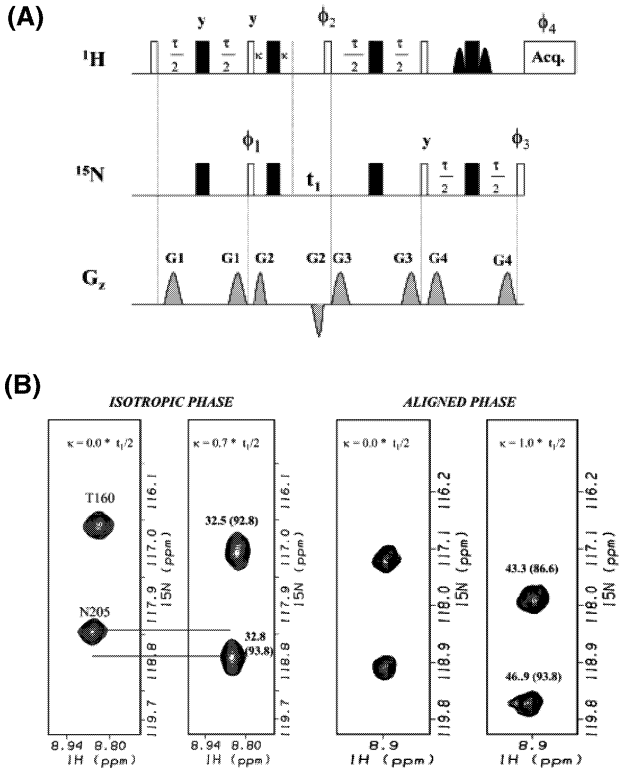

Figure 4. Example of $^1H_N-^{15}N$ RDC measurement in a large protein using a mixed HSQC-TROSY pulse sequence (CE-TROSY):[98] (A) pulse sequence in which the period, $\kappa = k\tau$, introduces variable amounts of coupling evolution into a TROSY sequence. (B) Two representative cross-peaks showing a difference in displacement for isotropic and aligned conditions. The numbers in parentheses are offsets adjusted by κ to directly reflect the couplings. The sample is 0.7 mM in the 53-kDa ^{15}N,(50% randomly deuterated)-labeled mannose binding protein trimer in 6% (w/v) bicelle solution.

TROSY-HSQC has been estimated to be about 1 Hz on a 30 kDa protein, precision that is fully sufficient for use in structure calculation.[84] A very similar 2D coupling-enhanced (CE)-TROSY-HSQC has recently allowed the measurement of sufficient $^1D_{HN-N}$ in a ^{15}N,^{2}H-labeled 53-kDa homomultimeric trimer from mannose-binding protein.[98] A representative dataset and the corresponding sequence are shown in Figure 4. When resonance overlap in the 2D HSQC spectrum becomes problematic, ^{15}N,^{13}C,^{2}H-labeled samples are required and a 3D TROSY-HNCO pulse scheme may become the method of choice, as in the case of a 723-residue enzyme, malate synthase.[99] In both of the latter examples, $^1D_{HN-N}$ couplings are extracted from the frequency displacement of the $^1H-^{15}N$ cross-peak in a reference TROSY experiment and the J-scaled TROSY sequence. Despite ^{15}N relaxation rates on the order of 10−15 ms at 800 MHz for the fast relaxing components, these procedures allow measurement of couplings within 2−3 Hz for large proteins in aligned media.

4.2. Other Protein Backbone RDCs

Despite success with measurement and interpretation of one-bond couplings, there is a clear need to measure a larger variety of couplings. In fact, five independent measurements are required to define order and orientation of small fragments, and additional parameters must be defined if dynamics are to be studied. Thus, a variety of experimental schemes has been designed for the collection of $^1D_{N-C'}$, $^1D_{N-C\alpha}$, $^1D_{C\alpha-C'}$, $^1D_{C\alpha-C\beta}$, $^2D_{N-C\alpha}$, $^2D_{HN-C'}$, $^2D_{HN-C\alpha}$, and $^3D_{HN-C\alpha}$ RDCs in ^{15}N,^{13}C or ^{15}N,^{13}C,^{2}H-labeled proteins. These experiments take advantage of the methodology developed for the measurement of $^1D_{HN-N}$ and $^1D_{H-C}$ couplings, as well as the triple-resonance experiments developed for assignment purposes. They have recently been extensively reviewed by Bax and coworkers.[7] Nevertheless, one principal advancement in recent years deserves special attention, namely, the collection of multiple couplings in a single experiment. This is motivated not only by the need to reduce spectrometer time requirements but also by the need to work within limits imposed by instability of biomolecules or alignment media.

Multiple couplings were first collected in a 30-kDa perdeuterated protein, MAP 30.[100] At 700−800 MHz, the ^{15}N line-narrowing obtained by constructive use of relaxation interference between $^1H-^{15}N$ dipolar coupling and ^{15}N CSA (TROSY) allows the accurate simultaneous measurement of $^1D_{HN-N}$, $^1D_{N-C'}$, and $^2D_{HN-C'}$. $^1D_{HN-N}$ splittings are further separated in two different spectra by addition/subtraction of the in-phase and antiphase $^{15}N-^1H$ doublets collected in the indirect dimension of a ^{1}H,^{15}N HSQC spectrum. In each subspectrum, the absence of $^{13}C'$ and ^{1}H decoupling during the ^{15}N evolution yields E.COSY-type patterns, from which the $^1D_{N-C'}$ and $^2D_{HN-C'}$ can be extracted, respectively, from the indirect and direct dimensions. Since that time, several strategies including the IPAP strategy, spin-state selective strategies, and J-modulated strategies have been used for the concomitant measurement of many other couplings. Adding to the collection of experiments previously reported by Griesinger et al.,[101] many of these experiments, the couplings measured, their experimental basis, and their respective references have been collected in Table 2. As the size of the studied protein increases drastically, some of these approaches are hampered by resolution and sensitivity issues and by the severe attenuation of the anti-TROSY component.[97] For large ^{15}N,^{13}C,^{2}H-labeled proteins, Yang et al.[102] proposed three modified TROSY-HNCO pulse schemes for the collection of $^1D_{HN-N}$ and $^2D_{HN-C'}$, $^1D_{N-C'}$ and $^2D_{HN-C'}$, as well as $^1D_{C'-C\alpha}$ and $^3D_{HN-C\alpha}$ RDCs. These pulse sequences take advantage of the E.COSY strategy[103] as well as accordion style spectroscopy[104] (except in the last case). When collected on a 1 mM sample of maltose-binding protein (42 kDa), they yielded a set of five RDCs for 275 well-resolved resonances with reasonable agreement between repeated experiments and values calculated from the crystal structure. Incorporated into a structure refinement procedure, this dataset improved the precision of the family of structures from 5.5 to 2.2 Å, while the rmsd with respect to the X-ray structure reduced from 5.1 to 3.3 Å.[105]

While RDCs have been most frequently used in combination with other types of NMR data to characterize biomolecular structures, there have also been

Table 2. Pulse Sequences Used for the Collection of Residual Dipolar Coupling Data

biomolecule	labeling	dipolar coupling	method	principle	ref
small/medium size protein	^{15}N	H_N-N	J-HSQC	J-modulation	71
			phase-encoded HSQC	peak volume in two spectra	81,196
			SCE-HSQC	line position from two spectra, normalization	98,197
			IPAP-HSQC	line position from two spectra	84,86,89
			E.COSY-HSQC	E.COSY extraction	87
			S^{3}E-HSQC	line position from two spectra	78
			S^{3}CT-HSQC	line position from two spectra	91
			α/β-HSQC	line position from two spectra	93,198
	^{13}C	$C_\alpha-H_\alpha$	CT-J-HSQC	J-modulation	82
	^{15}N,^{13}C		HNCO	E.COSY type spectra	199
			(HACACO)NH	J-modulation	96
			(HACACO)NH	J-modulation	96
			IPAP-(HA)CANH	line position from two spectra	107
			HCCH-COSY	line position from two spectra	90
	^{15}N,^{13}C	N−C′, N−C_α	TROSY-HNCO	J-modulation	200
			J-correlated HNC	peak volume in two spectra	201
			SE-HSQC	line position from two spectra	202
		C′−C_α	HN-(α/β-COCA-J)	line position from two spectra	203
	^{13}C	C′−C_α	CT-HSQC	line position	204
		H_N-C_α, H_N-H_α	CT-J-HSQC	J-modulation	82
			S^{3}E-HSQC	line position from two spectra	205
	^{15}N, (10−15% ^{13}C)	H_N-C_α, H_N-H_α	soft-HNCA-E.COSY	E.COSY extraction	70
	^{15}N,^{13}C	N−C′, H_N-C′	semi-CT-HSQC	line position difference, normalization	206
			DIPSAP J-HNCO	line position in three spectra	111
			S^{3}E/IPAP-HNCO	E.COSY extraction	100
		H_N-N, $C_\alpha-C'$ H_N-N, $C_\alpha-H_\alpha$	IPAP-HNCO	line position from two spectra for H_N-N, from one spectrum $C_\alpha-X$	207
		$H_\alpha-N$, $C_\alpha-H_\alpha$	E.COSY HNCA	E.COSY extraction	208
		side-chain CH_2: C−H	CT-J-HSQC	J-modulation	115
			CB(CA)CONH	peak volume from three spectra	117
	^{15}N,^{13}C	side-chain CH_2: C−H, H−H	SPITZE-HSQC	line positions from four spectra	118
		side-chain CH_3: C−H	CT-J-HSQC	J-modulation	115
			IPAP-CT-HSQC	line positions from two spectra	116
	^{15}N,^{13}C, 50% ^{2}H-fract lab	side-chain CH_3: C−H	filtered-CT-HSQC	H,H coupling as antiphase splitting	114,209
		side-chain CH_3: C−C	CT-HSQC	J-modulation	122,209
	^{15}N,^{13}C	side-chain CH_3: H−H	DiM	H,H coupling as antiphase split	119
	^{15}N,^{13}C, 50% ^{2}H-fract lab	side-chain CH_3: H−H	filtered-CT-HSQC	line separation from two spectra	120
	no label	H,H	COSY	ACME amplitude-constrained multiplet evaluation	132
			CT-COSY	intensity modulation	133,210
			signed COSY		138
			MOCCA-SIAM	ACME amplitude-constrained multiplet evaluation	211
		N−H, H_H-H_α	JHH-NOESY	E.COSY extraction	139
		H_HH_α	HNHA	J-modulation	2
	^{15}N,^{13}C	H_HH_α	HNCA-E.COSY	E.COSY extraction	134
	^{15}N,^{13}C,^{2}H	H−H	SS-HMQC	peak volume	126
			COSY-HMQC	peak volume	127
large size protein	^{15}N,^{2}H	H_N, N	JE-TROSY	J resolved spectroscopy in the third dimension	95
			SCE-HSQC	line position from two spectra, normalization	98
	^{15}N,^{13}C,^{2}H	H_N, N	TROSY-HNCO	line position from two spectra	84,102
		H_N-C_α, H_N-H_α	TROSY-HNCO	line position from two spectra	102,212
		N−C′, H_N-C′	TROSY-HNCO	line position from two spectra	102,212
	^{15}N,^{13}C,^{2}H	side-chain CH_3: C−C	^{13}C−^{13}C-TOCSY	H,H coupling as splitting	123
RNA/DNA	^{15}N,^{13}C	C−H	J-modulated HSQC	intensity-modulation	213
			TROSY-HSQC	line position from two spectra in ^{1}H dimension	214
		N9−C, $H8_N$−N9 purine N1−C, $H6_N$−N1 pyri	S^{3}E-HC[N]	line position from two spectra	112
			MQ-HCN	E.COSY extraction	215
		N1−C, $H1_N$-N9 purine N3−C, $H3_N$-C pyrim	S^{3}E-HN[C]	line position from two spectra	112
		$H_{2'}-H_{1'}$, $H_{2'}-C_{1'/2'}$, $H_{1'}-C_{1'/2'}$	CT-HMQC	E.COSY extraction from C1′−C2′ and H1′−C2′ planes	113
	no label	H,H	CT-COSY	intensity modulation	216
			selective-CT-COSY	peak volume	128
		H−P	CT-NOESY	intensity modulation	217
	^{19}F	H−F	E.COSY	E.COSY extraction	218,219
	^{15}N,^{13}C	through H bond H−N	HNN	E.COSY extraction	220
polysaccharide	no label	C−H	CT-CE-HSQC	C−H splittings	129
			HMBC	line fitting	221
		H−H	CT-COSY	intensity modulation	133
			E.COSY	E.COSY extraction	222
	^{13}C	H−H, C−C	CT-HSQC COSY	intensity modulation	223

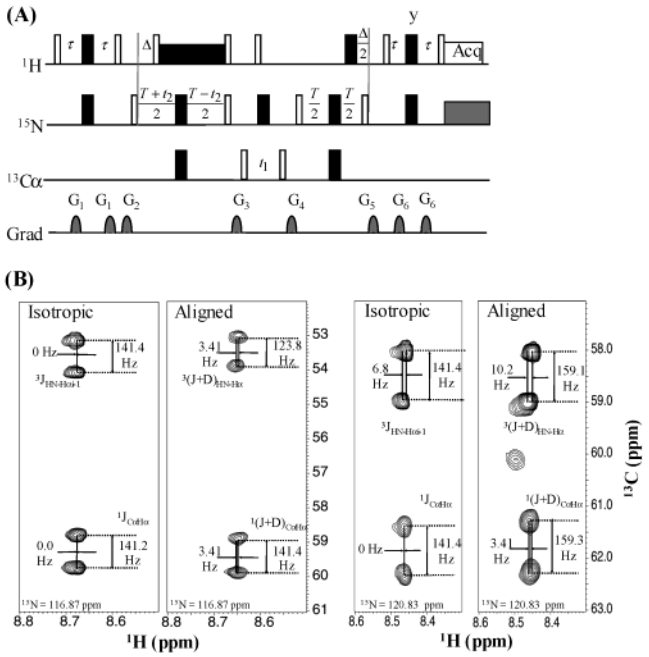

Figure 5. HNCA-E.COSY experiment for the simultaneous measurement of ^{1}H−^{13}C, and ^{1}H$_N$−^{1}H$_\alpha$ couplings in ^{15}N-labeled, low percentage ^{13}C-labeled proteins.[70] (A) Pulse sequence generating a 3D spectrum with E.COSY multiplets in the ^{1}H−^{13}C planes. (B) Example segments at discrete nitrogen frequencies showing intraresidue and interresidue multiplets for two residues. The protein, a 77-residue conserved hypothetical protein from *Pyrococcus furiosus* labeled to 90% in ^{15}N and 16% in ^{13}C is at 0.5 mM in 30 mM phosphate buffer, 50 mM KCl, 4% (w/v) C12E5/hexanol/CTAB (27:31:1). Data were collected at 600 MHz with a cryogenic probe. Note that no ^{1}H$_N$−^{1}Hα splittings are observed for interresidue multiplets under isotropic conditions.

efforts to use RDCs as primary sources of structural information as a more efficient route to backbone structures for proteins targeted by structural genomics projects.[106] This provides additional motivation for simultaneous collection of many types of couplings. Our own application imposed additional restrictions on the types of experiments that could be used by targeting simultaneous resonance assignment and structure determination and by emphasizing ^{15}N as opposed to ^{15}N/^{13}C enrichment. ^{13}C connectivities were not excluded from the experiments, but only those experiments with sufficient sensitivity to work at natural (or low) levels of ^{13}C abundance were chosen. With recent improvements in the sensitivity of NMR instruments and the advent of cryogenic-probe technology, the most sensitive triple resonance experiments such as HNCO and HNCA become feasible. In this context, a set of three experiments has been designed for the simultaneous backbone assignment and fold determination of small to medium-size proteins. It includes a 2D IPAP or phase-encoded HSQC, a 2D or 3D IPAP-HNCO, and a 3D soft-HNCA-E.COSY experiment, in which ^{1}D$_{HN-N}$, ^{1}D$_{N-C'}$ and ^{2}D$_{HN-C'}$, and ^{1}D$_{N-C\alpha}$ and ^{2}D$_{HN-C\alpha}$ are respectively collected.[70,107,108]

A representative dataset of the 3D HNCA experiment, including the isotropic and aligned sample in a mixture of DMPC/DHPC/CTAB bicelles, is presented in Figure 5. The small protein studied contains 77 amino acids and constituted one of the targets of the South East Collaboratory for Structural Genomics (SECSG).[109] Several sections of ^{1}H$_N$−^{13}C$_\alpha$ planes at particular ^{15}N chemical shifts are shown. These establish connections of a given ^{1}H$_N$ resonance to both intra- (i) and inter- (i-1) ^{13}C$_\alpha$ resonances. The intraresidue E.COSY-type multiplets show ^{1}H−^{13}C$_\alpha$ splittings in the vertical dimension and ^{1}H$_N$−^{1}H$_\alpha$ splittings in the horizontal dimension for both aligned and isotropic conditions. The interresidue E.COSY-type multiplets show ^{1}H$_N$−^{13}C$_\alpha$ splittings in the vertical dimension and ^{1}H$_N$−^{1}H$_\alpha$ splittings in the horizontal dimension only for the aligned condition as there is no significant interresidue four-bond ^{1}H$_N$-^{1}H$_\alpha$ scalar coupling. In addition to orientational information from RDC values, the HNCA-E.COSY experiment provides key information to establish interresidue connectivities. As in the normal HNCA experiment, matching the intraresidue C$_\alpha$ shift seen in one segment to the interresidue C$_\alpha$ shift seen in another segment provides connectivity. However, the HNCA-E.COSY as used for RDC measurement has additional advantages. The H$_{Ni}$ to C$_{\alpha(i-1)}$ pair is readily identified since the ^{1}H$_N$ upfield and downfield peaks present no frequency shift in the direct dimension. Also, accidental degeneracy in establishing sequential connectivities is minimized by insisting on matching of both doublet components in the aligned spectrum; because RDC values are different for different (^{1}H$_N$,^{13}C$_\alpha$) pairs, this adds an additional filter.

Analysis of the data can be automated to a considerable extent. In the case illustrated, this started with processing and automatic peak-picking, followed by a manual screening within the NMRPipe/NMRDraw suite of programs.[110] Peak positions were refined using Pipp and the nlinLS procedure within NMRPipe. It was followed by the simultaneous assignment of resonances to a particular (^{1}H$_N$,^{15}N) pair and measurement of couplings, which appeared as a frequency displacement in the direct and/or indirect dimension of the corresponding spectra. This step was automated with the help of Tcl/Tk scripts in NMRPipe. Matching of C$_\alpha$ chemical shift as well as splitting was also pursued automatically within a Tcl/Tk script.

For many protein samples, the time requirement for recording a full 3D spectrum with a signal-to-noise ratio sufficient for the precise determination of RDCs can be prohibitive. In fact, datasets containing three or more RDCs per residue are limited (BMRB and PDB databases). In this context, reduced dimensionality approaches may offer a good compromise, at least for small to medium-size proteins. With this idea in mind, Bersch et al.[111] recently demonstrated the concomitant measurement of ^{1}D$_{HN-N}$ and ^{2}D$_{HN-C'}$ in ^{15}N,^{13}C-labeled MerAa, a 68-amino acid protein. The pulse sequence, a 2D J-HNCO(H), is based on a regular HNCO experiment and uses the same incremented time delay t_1 to encode for the time evolution of the two spins ^{15}N and ^{13}C′. An adapted quadrature detection scheme, the absence of ^{1}H decoupling during the ^{15}N and ^{13}C′ frequency editing, and the use of a J-mismatched compensated DIPSAP filter (Figure 3) lead to the generation of two sub-

spectra α and β, where the resonance is displaced in the indirect dimension by the sum and the difference of the two targeted couplings. In this example, $^1D_{HN-N}$ and $^2D_{HN-C'}$ could be measured with a precision of 1.0 and 0.5 Hz, respectively, for 52 out of the 68 amino acids of the protein in only 45 min.

In DNA and RNA samples, measurements of multiple RDCs are sparse, partly due to the still onerous labeling task and to the lower spectral resolution. A suite of 2D spin-state selective experiments, including the S^3E excitation scheme, have been proposed for the simultaneous measurement of one- and two-bond RDCs in pyrimidines and purines.[112] In addition, there are some 3D experiments yielding as many as five couplings.[113] References to these experiments have been included in Table 2.

4.3. One-Bond C–C and C–H RDCs in Protein Side-Chains

One- to three-bond backbone RDCs provide invaluable orientational information for structure validation, for de novo backbone fold determination, and for the characterization of protein−protein, protein−DNA, or protein−ligand interactions. However, side-chain conformations are responsible for the critical details at the protein surface and can be responsible for the recognition between a protein and its partner. At the active site or the core of a protein, accurate positioning of methylene or methyl groups often defines hydrophobic side-chain contacts. Gaining orientational information on these fragments could obviously complement the distance restraints obtained through NOEs and help in structure refinement. In addition, orientational restraints can facilitate stereospecific assignments and provide an insight into χ1 torsion angles.[114]

Three RDCs are commonly used to provide side-chain orientational information: $^1D_{H-C}$, $^2D_{H-H}$, and $^1D_{C-C}$. Some of the simplest and most convenient approaches to measuring $^1D_{C-H}$ couplings are to use a J-modulated constant-time HSQC,[115] an IPAP-CT-HSQC,[116] or a 3D CB-(CA)CONH.[117] For methylene groups, Carlomagno et al.[118] have proposed a sequence that separates four of the eight normally unresolved components of a CH_2 coupled-HSQC multiplet. The carefully designed sequence, SPIn sTate selective ZEro overlap HSQC (SPITZE-HSQC), yielded the measurement of $H_{\alpha1}-C_\alpha$, $H_{\alpha2}-C_\alpha$, and $H_{\alpha1}-H_{\alpha2}$ scalar and RDCs for all six glycines present in a $^{15}N,^{13}C$ ubiquitin sample. One cautionary note is that values measured need to be interpreted carefully because of possible second-order effects.

In the case of methyl groups, the alignment introduces splittings in the proton spectrum which are the result of nonaveraged dipolar interactions between methyl protons. If the reorientation of the methyl group around its C_3 symmetry axis is fast, the $^1D_{H-H}$ coupling is, except for a scaling factor of −3/4, formally equivalent to the dipolar splitting between two weakly coupled protons. A 2D pulse sequence was designed by Kaikkonen and Otting[119] to measure the separation of the two outermost lines of the triplet by creating antiphase magnetization which suppresses the central resonance of the triplet. The observed relationship of $^1D_{H-C}$ and $^2D_{H-H}$ experimental couplings, $^2D_{H-H} \cong 2.3\ ^1D_{H-C}$, confirms that measurement of $^1D_{H-C}$ and $^2D_{H-H}$ couplings are in fact redundant. As the size of the protein increases, a high level of deuteration, 50% or so, is used to slow the relaxation of heteronuclei. Methyl groups are then present as three isotopomers, CH_3, CH_2D, CHD_2, which show slightly different chemical shifts. Filters can be introduced in the pulse sequence of interest to select one of the isotopomers, CH_2D, for example, and $^2D_{H-H}$ can be measured quite easily. Using an α/β spin-state selection filter and a CT-HSQC based experiment Sibille et al.[120] measured 65 methyl D_{H-H} couplings in SiR−FP18, a 18-kDa protein. More importantly, the authors could pursue a significant number of stereospecific assignments for valine, threonine, and isoleucine methyl groups by reproducing these measurements in two different alignment media and fitting the experimental values with RDCs calculated using the analysis program MODULE.[121]

Alternatively, orientation of the methyl C_3 symmetry axis can be determined by the measurement of $^1D_{C-C}$ couplings. In high molecular weight, fractionally deuterated proteins, these couplings can be measured from a modified J-modulated constant-time HSQC.[122] Very recently, an approach using a ^{13}C-start and ^{13}C-observe experiment in combination with broadband homonuclear cross-polarization has been presented for uniformly deuterated proteins of molecular weight above 40 kDa.[123] To date, this approach is still limited by spectral overlap and the multiplicity of many ^{13}C-correlation cross-peaks, but it may offer advantages for uniformly deuterated proteins up to 100 kDa.

4.4. Proton–Proton RDCs

Most of the RDCs presented so far have involved spins separated by one or two bonds; the distance between spins is fixed by bond geometry and the observables relate primarily to orientation. However, in the case of proton pairs, RDCs over much longer ranges can be observed. When Hansen et al.[124] executed a TOCSY-type pulse sequence on a 10-mer DNA duplex in phage they noted the appearance of new cross-peaks. The corresponding long-range interproton dipolar interactions could be detected between proton pairs separated by up to 7.5 Å. This is the result of the $1/r^3$ dependence of the dipolar coupling as compared to the $1/r^6$ dependence of the Nuclear Overhauser effect. Constraints based on measurement of these couplings now involve both distance and angular variables. Their value for structure determination of proteins was quickly recognized; measurement and inclusion of D_{H-H} couplings in the structure refinement of ubiquitin, for example, resulted in a backbone atomic rms shift of 0.43 Å toward the crystal structure.[125] Perdeuteration of proteins can extend the range over which couplings can be measured and also allows the measurement of D_{HN-HN} couplings between remote backbone segments in proteins.[126,127] A more general approach for extending distances over which couplings can be observed is to effectively decouple all

protons outside the spectral regions of interest during the homonuclear coherence transfer. A band-selective ^{1}H homonuclear decoupling pulse element has been designed and implemented in a COSY pulse sequence. Couplings on an unlabeled 12-mer DNA duplex have been detected between two protons up to three base pairs (12 Å) apart.[128]

Another very useful application occurs in oligosaccharides. One-bond H−C dipolar couplings, which are relatively easy to measure in these molecules, frequently do not provide a sufficiently diverse set of RDCs to determine structure because of the nearly parallel orientation of many C−H vectors in pyranose rings. The addition of proton−proton RDCs provides the information needed and allows constraint of the orientation of one ring relative to another, as well as some assessment of conformational flexibility in these molecules.[129,130]

^{1}H−^{1}H couplings are also unique in that many of the experiments used to measure them do not require isotopic labeling of the molecule of interest. Simple COSY experiments and their derivatives can provide convenient means of visualizing couplings.[131] Actual measurement of couplings is more challenging. A recent procedure proposes an amplitude-constrained multiplet evaluation (ACME) to interactively extract the coupling constants.[132] An alternative is to use a constant-time COSY. In the absence of significant differential relaxation, the cross-peak intensity and the autopeak intensity are respectively modulated by a sine and cosine function of the homonuclear coupling constant. To avoid inaccuracy related to the different phase properties of the cross- and autopeaks, data sets are collected for different constant times, and amplitudes of the signals are fit to extract the necessary couplings. Data collected on a trisaccharide are presented in Figure 6, together with subsequent analysis of the observed ratio of cross-peak and autopeak amplitudes.[133] Other variants of these basic procedures are included in the experiments listed in Table 2.

In macromolecules that do have a high level of ^{15}N and/or ^{13}C labeling, the heteronucleus can be advantageous in resolving additional proton−proton couplings. A variety of triple resonance experiments[134−136] have been designed for ^{15}N,^{13}C-labeled proteins that take advantage of the E.COSY approach proposed by Sørensen and co-workers.[135−137] This approach can also be integrated in experiments designed for ^{15}N-only labeled proteins. In the corresponding experiments, the transfer of coherence from the ^{1}H nucleus to the ^{15}N nucleus without nitrogen decoupling provides the necessary E.COSY. After ^{15}N evolution and transfer to an amide proton, a TOCSY[138] or NOESY[139] transfer establishes a connection to a remote proton. This strategy provides both the magnitude and the sign of the D_{H-H} coupling in contrast to approaches based on the HNHA or HSQC/HMQC experiments which only provide magnitudes.[140]

There are of course a number of other types of RDCs that can and have been measured, including those between ^{31}P or ^{19}F and any number of the spins we have already discussed. The experiments used to measure these couplings parallel those discussed above. We will not discuss these experiments in detail, but again we have included appropriate references in Table 2.

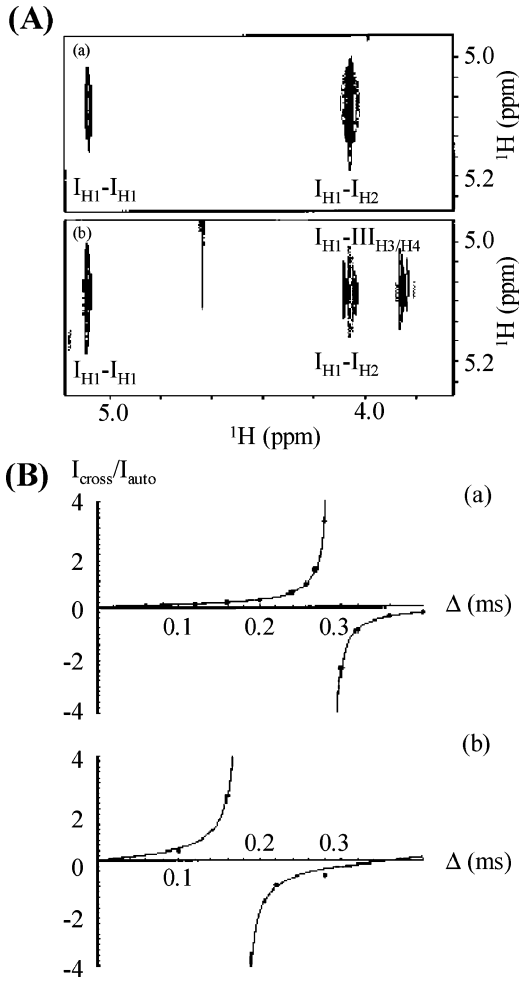

Figure 6. Constant time COSY for the measurement of ^{1}H−^{1}H RDCs. (A) Spectrum of trimannoside in aqueous bicelle media under (a) isotropic, 20 °C and (b) aligned, 36 °C conditions. Note the appearance of long-range couplings between the anomeric proton (H1) of ring I and the transglycosidic protons (H3 and H4) of ring III under aligned conditions. (B) Plots of intensity ratio used to extract couplings ($I_{cross}/I_{auto} = k \tan(\pi(J + D)\Delta)$) for the (a) isotropic and (b) aligned cases.

5. Structural Interpretation of RDCs

The advances in the acquisition of RDC data have clearly enhanced the potential for extracting structural information on a variety of biomolecular systems. It is tempting to think of use in structure determination as a relatively straightforward undertaking. After all, programs for the production of structures based on NOE distance constraints are highly refined and readily available. However, RDCs in most applications offer a fundamentally different type of information, information that is orientational, in addition to distance dependent. This fact, in some ways, makes RDC information an ideal complement to NOEs; RDCs can provide long-range constraints on structures in situations in which few NOEs can be observed. However, this fundamental difference in the nature of the information offered has required the development of a number of new procedures and

new software programs for interpreting RDCs in terms of structural models. We discuss these developments in terms of four classes: (1) a group of programs combining RDC information with other structural constraints (NOE-based, for example) using target, pseudoenergy, or error functions in MD-like algorithms, (2) a group of programs building structures primarily from RDCs by a direct search for fragment geometries compatible with molecular alignment data, (3) a group of programs using RDC data to search for homologous structures that may exist in structural databases or to validate structures produced by computer modeling strategies, and (4) a group of programs using RDCs in combination with existing structures of complex systems (ligands, protein domains, nucleic acid segments, or subunits of macromolecular complexes) to deduce orientational relationships of the components. Each class has its own advantages and special areas of application.

5.1. Additional Constraints in Structure Determination and Refinement

The first class of RDC implementation has most commonly been built on programs initially introduced for the analysis of NOE data. Programs such as XPLOR/CNS,[141,142] DYANA,[143,144] and AMBER[145,146] are good examples of cases in which error or target functions have been added that minimize when RDC data are satisfied. These programs work by searching for a global minimum in a total energy function by simulated annealing, Monte Carlo, or other approaches. These approaches work well when substantial amounts of NOE data are included to minimize problems with the multiple minima inherent in RDC error functions. In cases in which RDC data are more numerous, search efficiencies can also be improved by reducing weighting functions for RDCs during early stages of the search to allow NOE data to dominate initial folds. The optimum protocols may vary with the number and distribution of constraints as well as the size of the system studied.[16] When RDC data are added to structure determination protocols in this way, substantial improvements in the quality of NMR structures can be realized. An interesting observation is that when a corresponding high resolution X-ray structure exists, the NMR structures not only become better defined (according to rmsds of backbone atoms) but also converge toward the X-ray structure.[147]

One interesting application has been in the comparison of protein structures in the presence and absence of bound ligands.[85] Galectin-3 is a carbohydrate binding protein (lectin) with an affinity for lactose or *N*-acetyl-lactosamine. Crystal structures exist for the protein with ligand bound but not for the apoprotein. A set of NOE data existed, and a set of ^{1}H-^{15}N RDC data was acquired for the protein in the absence of ligand. Starting from the crystal structure, a ligand-free structure was generated using a target function and simulated annealing protocol. The two structures, as compared in Figure 7, present an interesting picture of binding site dynamics and the possible exposure of peptide segments involved in apoptotic signaling in the presence of ligand.

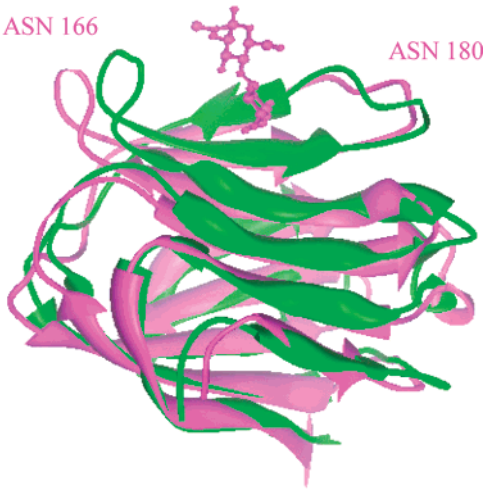

Figure 7. Structure of galectin-3 in the absence (green) and presence (magenta) of ligand. The structure in the absence was refined with $^{1}H-^{15}N$ RDCs and NOEs[85] beginning with the ligand-bound X-ray structure, 1A3K.

RDC constraints can be combined with a variety of constraints other than NOEs using similar pseudoenergy/target function approaches. As discussed above, CSA offsets for resonances from highly anisotropic groups such as carbonyl carbons display geometry dependence similar to RDCs; in fact, these offsets can be written in terms of a pair of RDC-like constraints. A procedure for adding these constraints has been added to the program CNS.[16] A novel addition to DYANA combines RDC data not only with NOE data but also with data from pseudocontact shifts and electron spin-induced relaxation of nuclear sites in systems substituted with paramagnetic metal ions.[144] Paramagnetic relaxation provides distance constraints between a paramagnetic (metal, spin label) center and the nuclear site in much the same way that NOEs provide distance constraints between pairs of nuclear sites, but the distances have a longer range. Pseudocontact shifts display a distance ($1/r^3$) and an angular dependence similar to RDCs, but now the vector that defines that distance is that between the paramagnetic center and the single nuclear site observed, rather than a vector connecting a pair of spin $^1/_2$ nuclei.[55] These new additions are particularly intriguing because of the recent introduction of paramagnetic tags for alignment purposes and the move to higher fields where measurement of these additions will become easier (see section on alignment media).[57]

5.2. Direct Structure Determination

The second class of programs has usually been based on a procedure that calculates the best order tensor solution for a subset of dipolar couplings belonging to a small peptide fragment of assumed geometry. An early program solved a set of equations analogous to eq 2 in section 2.1 to obtain a best order tensor solution using singular value decomposition.[148] The resulting order tensor elements were assembled into a matrix, and the matrix was diagonalized to yield principal order tensor values and a definition of the principal axis directions in the molecular fragment frame. Various fragment geometries can

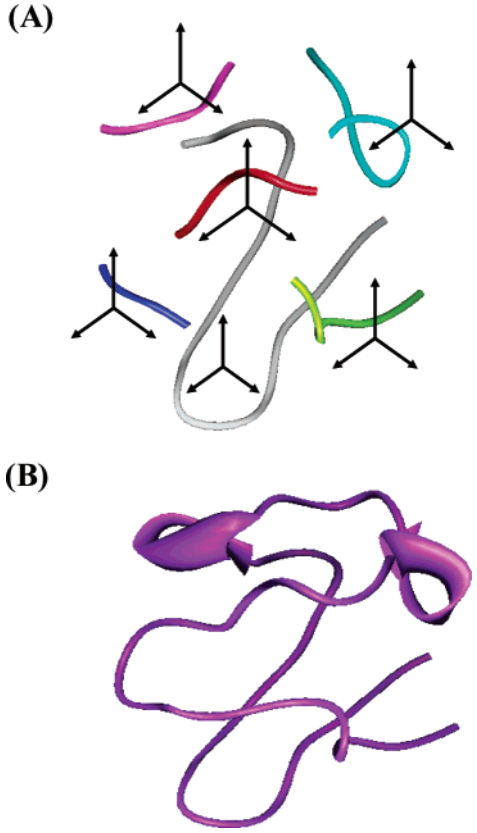

Figure 8. Protein backbone structure determination from RDCs.[70] (A) The local structure and global orientation was determined from RDC data for six fragments of the small protein, rubredoxin. (B) Translation and connection of the fragments resulted in a complete structure, 1RWD.

also be tested by back-calculating RDCs from a best order tensor solution and comparing them with experiment to give a score used in selecting an appropriate model.[149] In an early application,[70] the phi and psi angles connecting a pair of peptide planes were systematically varied using a comparison to a set of six couplings per peptide plane (nine unique to a pair of planes) to select the best geometry. The couplings included $H_{Ni}-N_i$, $C'_{i-1}-N_i$, $C'_{i-1}-H_{Ni}$, $C_{\alpha i}-H_{\alpha i}$, $H_{Ni}-H_{\alpha i}$, and $H_{Ni}-H_{\alpha i-1}$ couplings. Additional peptides were then added along with additional couplings (3−6) to extend the fragments to lengths of 6−16 residues. In the case of the small protein rubredoxin (54 residues) geometries for six fragments, each separated by a single proline, were identified. These fragments were then combined using the fact that fragments must be oriented so that order frames as viewed from each fragment must coincide. Several of the fragments with their alignment frames are shown in Figure 8. In most cases, bonding restrictions imposed by the single connecting proline provided unique translational positioning. RDC data from a second medium helped resolve any remaining 4-fold orientational degeneracy of the individual order frames.[59] The procedure led to a backbone structure that agreed within 1.8 Å with a closely homologous X-ray structure.[70] A unique aspect of the procedure is that assignment of resonances to a specific point in a protein sequence is not a prerequisite to structure determination. Sets of couplings involving particular resonances do need to be connected, but this can be accomplished without many of the experiments normally required for sequential assignment.

This general procedure has been systematized and improved to allow a recent application to a larger protein as a part of a structural genomics project.[109] One of the tenants of structural genomics initiatives is that determination of structures for proteins in poorly populated fold families will allow computational prediction of structures for proteins with homologous sequences.[150] Since sequence identity of homologues may be as low as 30%, side-chain geometry is of little relevance. This raises interest in direct determination of backbone structures, something that RDC-based approaches can accomplish when NOE-based approaches cannot because of the rarity of backbone-to-backbone NOEs in most protein structures.

Similar order tensor-based strategies can be adopted when taking fragments from libraries of existing structures. Delaglio and Bax[151] introduced a procedure that uses a reduced set of structures from the Protein Data Bank to provide this library. Fragments 5−7 residues in length, having a set of four backbone couplings per residue ($H_{Ni}-N_i$, $H_{\alpha i}-C_{\alpha i}$, $N_i-C'_{i-1}$, $H_{Ni}-C'_{i-1}$), are threaded through members of the reduced set incrementing the alignment one residue at a time. A best alignment tensor is calculated, and an rmsd between experimental and back-calculated couplings is used to evaluate the alignment. The procedure has been tested on ubiquitin with significant success. In this case, complete sequential assignments were done ahead of structural analysis and fragment length was selected simply based on structural precision and database limitations. For shorter fragments, this database approach appears to provide an adequate sampling of conformational space but at some compromise in precision. For longer fragments precision improves, but there are concerns about adequate sampling and bias of structures toward conformations that are already well represented in the database.

Levy and Andrec have introduced a similar protocol.[152,153] Their procedure differs in that it employs a filter that requires overlapping segments to select the same structural form. The data set is again quite extensive with a set of RDCs similar to that used by Delaglio and Bax. They too find optimum performance with segments of approximately seven residues. Performance is improved with their procedure; however, there are more stringent requirements for extensive sequential assignment.

A procedure that appears successful with fewer data is built on methods developed for homology and ab initio protein modeling.[154] These methods rely not only on local structural characteristics but also employ pairwise contact potentials to select proper folds. A structural template selected based on sequence homology is also used. Without RDCs, competitions such as the CASP program show that one might expect a structure accurate to 3.5 Å when a template with 30% sequence identity exists. When just H_N-N RDCs are added, accuracy improves dramatically. A related method called ROSETTA

employs many of the pairwise potentials of threading algorithms but works by doing a Monte Carlo search of a fragment library rather than starting from a homology template.[155] This obviously broadens applicability by eliminating the need for a homology template. Test cases using RDC data on proteins from 52 to 128 residues gave structures with backbone atom rmsds ranging from 1 to 6 Å from structures determined by X-ray or NOE methods. In most cases, this represents a very significant improvement over modeling without experimental constraints. Again, these methods require extensive sequential assignment.

5.3. A Tool for Structure Validation or Homology Searches

The above categories clearly merge with our third class of programs, those that simply screen for similarity to proteins with known structure. This type of application was in fact among the earliest application of RDCs to protein structure determination.[156] Assigned RDCs are sequentially placed along a trial structure, a search for a proper alignment tensor is conducted, and an rmsd between back-calculated and experimental couplings is determined. Rmsd values approaching expected experimental error provide a positive hit. Griesinger et al. have expanded on this protocol by introducing a more efficient algorithm in a program called DipoCoup.[61] They have also incorporated modules that accept additional pseudocontact shift information when paramagnetic proteins are examined.

Recently Valafar et al. have demonstrated that it may be possible to omit assignment to specific sites and rely on pattern matching to identify protein folds.[157] An illustration is given in Figure 9 in which a comparison of the distribution of predicted dipolar couplings for a pair of structurally related proteins and a pair of structurally unrelated proteins is shown. In both cases, the width of the distribution is based on order predicted for a bicelle medium using the program PALES.[158] The structurally related pair (Figure 9A) are domains from glycosyltransferases that have less than a 15% sequence identity over 180 residues but have a 2.5 Å backbone rmsd.[159] One can clearly see the similarity in distribution of dipolar coupling values. The structurally unrelated pair (Figure 9B) is one of the glycosyltransferases and a similarly sized protein, ADP ribosylating factor. There is clearly a dramatic difference in pattern. Similarity in distributions can also be detected when the alignment tensor cannot be independently predicted and can be adjusted for a best fit to a distribution. However, recognizing similarity and dissimilarity in these cases depends on small deviations of experimental distributions from the powder pattern shapes. These distributions are calculated from a random distribution of vector directions for a large number of vectors using the best fit alignment tensor. The procedure in this latter case will break down for larger proteins and may not be able to distinguish certain symmetrically related structures. Nevertheless, application to a variety of folds in the 100-amino acid range appears to be quite feasible.

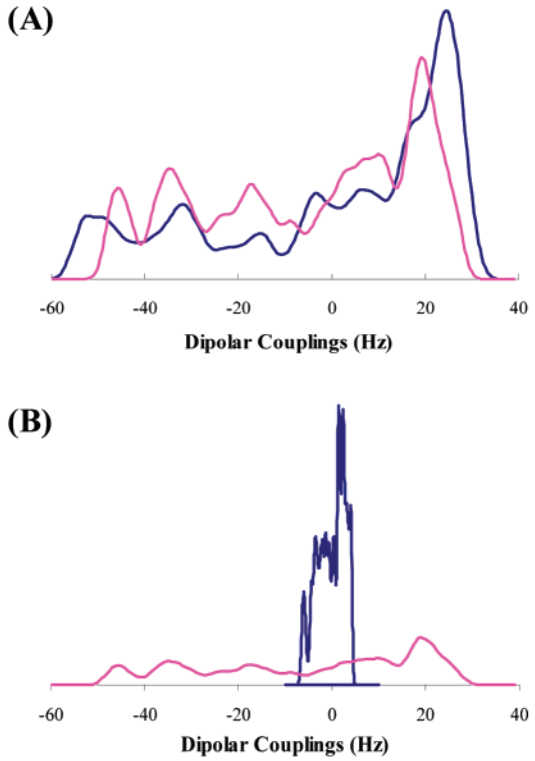

Figure 9. Dipolar coupling distributions calculated for two proteins (A) of similar structure, glycosyltransferase C-terminal domains of 1F0K and 1IIR, which overlay to 2 rmsd, and (B) of dissimilar structure, glycosytransferase C-terminal domain of 1F0K and HUR1, which are structurally unrelated.[157] Alignment tensors were predicted with PALES.[158]

The identification of a structural homologue can be a stepping stone to further structure refinement using some of the methods described above. However, methods of this type are also advantageous in structural genomics applications. Emphasis here is on the structure determination of proteins with novel folds. New proteins are typically screened based on the extent of primary sequence identity to eliminate targets with redundant folds. However, even with very low levels of sequence identity, a significant fraction of structures determined turn out to belong to well-populated fold families. An efficient experimental method for identifying proteins belonging to well-populated fold families will clearly avoid the more time-consuming task of structure determination of proteins with less novel folds. All of the above pattern recognition procedures, whether requiring assignments or not, are quite efficient and usually need only H_N-N couplings from ^{15}N-labeled proteins.

5.4. Orientational Relationships between Components

The fourth class of programs is designed to take advantage of existing structures to extend investigations to the way biomolecular components interact and function. These are important applications and ones in which use of RDCs can be particularly advantageous. NOEs across the interface of interacting pairs of proteins, protein−nucleotide complexes, or protein−ligand complexes can be sparse and difficult to observe. Components of complex systems

may also have dynamic aspects that are difficult to characterize by other means. Programs designed to address these problems often make direct use of the order tensors that were used to back-calculate RDCs in previous examples. As in the case of the protein fragment assembly strategy described above, the direction of axes of the principal order frame must coincide for elements of the same structure, placing restrictions on relative orientations of components; the magnitude of the principal order elements must also be the same if the complex under investigation is rigid. If elements for different components are not the same, they can, under the right circumstances, be interpreted in terms of the extent of motion between components.

Programs useful in analyses of the type described above include the ones that directly calculate order tensors from a structure and RDCs, including the ones employing singular value decomposition to find a best least-squares solution to a set of equations relating RDCs to order and orientation. A new program called REDCAT (REsidual Dipolar Coupling Analysis Tool) based on this algorithm has recently become available.[160] Input is any PDB structure along with a list of atom pairs showing an RDC and the corresponding RDCs. Providing adequate data exist (in principle, more than five independent couplings, in practice more than 15) a set of principal order parameters and a set of Euler angles for transforming the PDB file to the principal alignment frame are returned. There are also a number of tools for assessing the quality of data. When applied to different parts of a complex, the transformed structures can be viewed in a common frame allowing assessment of probable modes of interaction. Again, with data from media giving different molecular orientations, or with the addition of a few distance constraints, it is possible to eliminate the normal 4-fold degeneracy in relative orientations.

A program called MODULE offers some very similar tools.[121] The program is based on a different algorithm for finding an order tensor solution but allows the same transformation of coordinates for pieces of a complex and provides a more extensive graphical interface. There are also programs that are based on rigid body orientational minimization algorithms.[161,162] The latter may provide a convenient approach, particularly when a variety of other data must be integrated with a structure search.

One of the earliest applications to relative domain orientation was in the case of a two-domain fragment from barley lectin.[163] This application included an observation of interdomain motion. Magnitudes of order parameters were significantly larger for one domain than the other. This was interpreted as a tendency for one domain to preferentially interact with the bicelle orientating medium, resulting in a greater degree of order for this domain. The other domain exhibits reduced order due to internal motion relative to the first domain. An assumed simple model for motion led to an estimate of internal fluctuations of ± 30°. An illustration showing the relative domain orientations for this particular problem is shown in Figure 10.

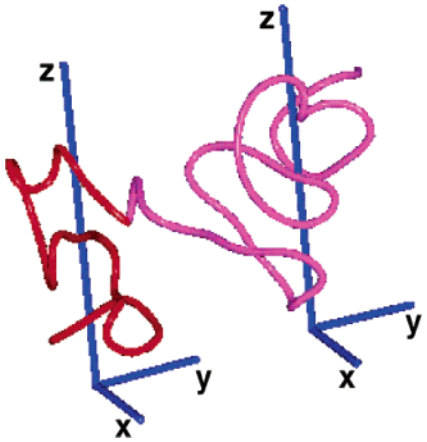

Figure 10. Domain orientations for B and C domains of barley lectin. $^{1}H_{N}-^{15}N$ dipolar couplings were used to solve for alignment frames of each domain independently. Domains were then rotated until directions of frame axes coincided producing the structure depicted.[163] The structure must be viewed as an average structure as substantial interdomain motion is detected.

Some more functionally relevant studies include those on T4 lysozyme and maltose binding protein in which closure of domains around ligand binding pockets has been quantitatively studied and analyzed in terms of hinge motion parameters.[164,165] Other studies include domain orientations of homomultimers[166] and mannose binding protein.[98] Studying domain orientation in homomultimers has been a challenging problem when based on NOEs because of the single set of resonances observed and ambiguities between intra- and intermolecular NOEs. This problem can be avoided to a certain extent in RDC studies because certain symmetry axes must coincide with principal frame orientation axes. This is the case with mannose binding protein which has a 3-fold axis[98] and in certain dimers.[167] One of the largest protein systems tackled using RDCs to constrain domain orientations is hemoglobin.[168] Here an interesting observation was that the structure determined in solution with the assistance of RDCs proved to be a dynamic average between previously existing crystal structures. There are also some notable examples of determining relative domain orientations in nucleic acids and nucleic acid−protein complexes.[45,46,169−171]

Determining the bound orientation of a ligand complexed with a protein presents challenges that are much the same as for protein−protein complexes. One must independently determine the preferred orientation of both the protein and the ligand. Each is treated as a rigid body and the programs discussed above are well-suited to this type of problem. The same issues related to the 4-fold degeneracy of orientation arise, and one must either use multiple orientation media or supplementary distance constraints to remove this degeneracy. In some cases, the ligand is actually a peptide and sources of data are identical to those discussed above, namely, a variety of peptide backbone RDCs from a labeled ligand.[172] For non-peptide ligands, the data source for the ligand may be dominated by $^{1}H-^{1}H$ and $^{1}H-^{13}C$ RDCs. While early work employed ^{13}C-labeled

ligands,[173,174] recent advances in probe sensitivity (cold probes) have made it possible to work at natural abundance.[98]

There can be some added difficulties in working toward bound ligand geometries using RDCs. Sensitivity limitations in ligand−protein interactions are sometimes overcome by working in the fast exchange limit and using an excess of ligand (10−20-fold). In this case, an average RDC would be observed, with RDCs from both bound and free states contributing. This would be akin to transferred NOE studies except there is no inherent emphasis of bound state RDCs; induced orientation of free ligand can be quite substantial. The need to separately measure and remove the free state contributions can significantly reduce accuracy of results. Work to improve this situation continues. Particularly promising are orientation protocols that preferentially orient the protein. This can occur when the protein is paramagnetic, either inherently or by addition of an appropriate tag. One interesting application is peptide binding to fragments of G-protein-containing membranes.[172] These fragments have sufficient anisotropy in susceptibility to orient in the presence of high fields. Without the dense fragment distributions found in liquid crystalline media, there is minimal orientation of free ligand.

Another problem that arises frequently with bound ligands is the lack of sufficient information to determine both orientation and bound conformation from RDCs alone. Determining conformation is critical for flexible ligands, both peptides and oligosaccharides. The oligosaccharide case is challenging because $^{1}H-^{13}C$ one-bond couplings at natural abundance are often the primary source of information, and the number of independent pieces of information for each sugar ring can be minimal (all beta-linked glucose H−C vectors are nearly parallel for example).[175] In these cases, bound conformations are best determined from other sources. In some cases, conformational flexibility about glycosidic linkages is sufficiently restricted to allow assumption of a particular bound conformation.[176] In other cases, RDC data for the determination of orientation can be combined with transferred NOE data on the ligand for determination of bound conformation.[174]

Proper assembly of complexes, whether they are protein−protein, protein−nucleic acid, or protein−ligand complexes, necessitates translational as well as orientational constraints. There is a variety of procedures for identifying contact surfaces in complexes. These range from simple chemical shift perturbation of resonances from proximate residues by added components to various protection experiments (from paramagnetic induced relaxation, for example) on adding a second component.[177,178] Implementing contact constraints in combination with orientational constraints in assembling a complex can also be challenging. Fortunately, there are software packages emerging that address this issue as well.[179]

6. Limitations and Future Directions

The information available from RDC measurements and anisotropies in other NMR observables is obviously varied and quite powerful. It does, of course, have limitations. Chief among these is the absence of translational constraints in most applications. We have mentioned a few contact-sensitive experiments in our discussion of complex assembly above, but the need for more quantitative sources of long-range distance constraints is clear. One obvious source is paramagnetic relaxation enhancement from natural or artificially introduced centers. Under the right circumstances, these can provide very long-range restraints (20−30 Å). Several applications using relaxation enhancement from paramagnetic sites, either spin labels or metal centers, have already been used to provide some complementary distance information.[28,180,181]

Another potential limitation is the effect of internal motion on RDC values. Most of our structural analysis procedures assume a rigid model for the molecule under study. When motion exists, without our knowledge, errors in structure can in principle be introduced. Fortunately, amplitudes of motion have to be rather large before this is a major concern.[9]

RDCs can actually be turned into an advantage in the study of motion in cases in which adequate data exist to simultaneously determine structure. We do not plan to review this type of application here but simply point to the fact that reduction in order parameters by internal motion relates primarily to amplitudes of motion and is sensitive to motion on a wide range of time scales (picoseconds to milliseconds). Thus, RDC measurement complements spin relaxation studies which tend to be more sensitive to faster motions, and for slower motion, returns information on a time scale more readily than amplitudes of motion. The interested readers are referred to recent articles by the Griesinger[67] and Tolman groups and their included references.[66,182]

7. Abbreviations

ACME	amplitude-constrained multiplet evaluation[132]
AMBER	force-field and package of molecular simulation programs[145]
ADP	adenosine 5′-diphosphate
BMRB	Biomolecules Magnetic Resonance Databank
CASP	critical assessment of techniques for protein structure prediction
CHAPSO	3-(chloramidopropyl)dimethylammonio-2-hydroxyl-1-propane sulfonate
C_nE_m	*n* is the number of carbons in a poly(ethylene)glycol headgroup and *m* refers to the number of repeating ethylene oxide units
CNS	crystallography and NMR system[141]
CPyBr	cetylpyyridinium bromide
CSA	chemical shift anisotropy
CT	constant-time
CTAB	cetyltrimethylammonium bromide
COSY	correlation spectroscopy
DNA/RNA	deoxyribonucleic acid/ribonucleic acid
DIOHPC	1,2-di-*O*-hexyl-*sn*-glycero-3-phosphocholine
DIODPC	1,2-di-*O*-dodecyl-*sn*-glycero-3-phosphocholine
DIPSAP	double in-phase single antiphase

DHPC	1,2-dihexanoyl-*sn*-glycero-3-phosphocholine
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DYANA	dynamics algorithm for NMR applications[143]
E.COSY	exclusive COSY
HMQC	heteronuclear multiple quantum correlation
HSQC	heteronuclear single quantum coherence
INEPT	insensitive nucleus enhancement by polarization transfer
IPAP	in-phase antiphase
MD	molecular dynamics
MODULE	program allowing the determination of alignment tensors[121]
NaBr	sodium bromide
NOE(NOESY)	nuclear Overhauser effect(spectroscopy)
NMR	nuclear magnetic resonance
PALES	prediction of alignment from structure[158]
PDB	Protein Data Bank
RDC	residual dipolar coupling
REDCAT	residual dipolar coupling analysis tool[160]
ROSETTA	CASP algorithm[155]
rmsd	root-mean-square deviation
SDS	sodium dodecylsulfate
S^3-CT/S^3-E	spin state selective constant time/exclusive
SPITZE-HSQC	spin state selective zero overlap HSQC
TOCSY	total correlation spectrscopy
TROSY	transverse relaxation optimized spectroscopy
XPLOR	structure determination program[142]

8. Acknowledgments

Research on residual dipolar couplings has been funded by grants from the National Institutes of Health, GM062407 and RR05351.

9. References

(1) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279.
(2) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111.
(3) Bastiaan, E. W.; Maclean, C.; Van Zijl, P. C. M.; Bothner-By, A. A. *Annu. Rep. NMR Spectrosc.* **1987**, *19*, 35.
(4) Emsley, J. W.; Lindon, J. C. *NMR Spectroscopy Using Liquid Crystal Solvents*; Pergamon Press: Oxford, 1975.
(5) Zhou, H.; Vermeulen, A.; Jucker, F. M.; Pardi, A. *Biopolymers* **1999**, *52*, 168.
(6) Prestegard, J. H.; Al-Hashimi, H. M.; Tolman, J. R. *Q. Rev. Biophys.* **2000**, *33*, 371.
(7) Bax, A.; Kontaxis, G.; Tjandra, N. *Methods Enzymol.* **2001**, *339*, 127.
(8) Brunner, E. *Concepts Magn. Res.* **2001**, *13*, 238.
(9) Tolman, J. R. *Curr. Opin. Struct. Biol.* **2001**, *11*, 532.
(10) Al-Hashimi, H. M.; Patel, D. J. *J. Biomol. NMR* **2002**, *22*, 1.
(11) de Alba, E.; Tjandra, N. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 175.
(12) MacDonald, D.; Lu, P. *Curr. Opin. Struct. Biol.* **2002**, *12*, 337.
(13) Bax, A. D. *Protein Sci.* **2003**, *12*, 1.
(14) Saupe, A. *Angew. Chem., Int. Ed. Engl.* **1968**, *7*, 97.
(15) Bertini, I.; Luchinat, C.; Parigi, G. *Concepts Magn. Reson.* **2002**, *14*, 259.
(16) Choy, W. Y.; Tollinger, M.; Mueller, G. A.; Kay, L. E. *J. Biomol. NMR* **2001**, *21*, 31.
(17) Wu, Z. R.; Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **2001**, *123*, 3617.
(18) Lipsitz, R. S.; Tjandra, N. *J. Magn. Reson.* **2003**, *164*, 171.
(19) Wu, Z.; Delaglio, F.; Tjandra, N.; Zhurkin, V. B.; Bax, A. *J. Biomol. NMR* **2003**, *26*, 297.
(20) Case, D. A. *Curr. Opin. Struct. Biol.* **1998**, *8*, 624.
(21) Czernek, J. *J. Phys. Chem. A* **2001**, *105*, 1357.
(22) Losonczi, J. A.; Prestegard, J. H. *Biochemistry* **1998**, *37*, 706.
(23) Sanders, C. R. *Biophys. J.* **1993**, *64*, 171.
(24) Lipsitz, R. S.; Tjandra, N. *J. Am. Chem. Soc.* **2001**, *123*, 11065.
(25) Boyd, J.; Redfield, C. *J. Am. Chem. Soc.* **1999**, *121*, 7441.
(26) Kurita, J.; Shimahara, H.; Utsunomiya-Tate, N.; Tate, S. *J. Magn. Reson.* **2003**, *163*, 163.
(27) Bertini, I.; Janik, M. B. L.; Lee, Y. M.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **2001**, *123*, 4181.
(28) Bertini, I.; Donaire, A.; Jimenez, B.; Luchinat, C.; Parigi, G.; Piccioli, M.; Poggi, L. *J. Biomol. NMR* **2001**, *21*, 85.
(29) Volkman, B. F.; Wilkens, S. J.; Lee, A. L.; Xia, B.; Westler, W. M.; Beger, R.; Markley, J. L. *J. Am. Chem. Soc.* **1999**, *121*, 4677.
(30) Wilkens, S. J.; Xia, B.; Weinhold, F.; Markley, J. L.; Westler, W. M. *J. Am. Chem. Soc.* **1998**, *120*, 4806.
(31) Allegrozzi, M.; Bertini, I.; Janik, M. B. L.; Lee, Y. M.; Lin, G.; Luchinat, C. *J. Am. Chem. Soc.* **2000**, *122*, 4154.
(32) Hus, J. C.; Marion, D.; Blackledge, M. *J. Mol. Biol.* **2000**, *298*, 927.
(33) Bertini, I.; Cavallaro, G.; Cosenza, M.; Kummerle, R.; Luchinat, C.; Piccioli, M.; Poggi, L. *J. Biomol. NMR* **2002**, *23*, 115.
(34) Duchardt, E.; Richter, C.; Ohlenschlager, O.; Gorlach, M.; Wohnert, J.; Schwalbe, H. *J. Am. Chem. Soc.* **2004**, *126*, 1962.
(35) Pintacuda, G.; Keniry, M. A.; Huber, T.; Park, A. Y.; Dixon, N. E.; Otting, G. *J. Am. Chem. Soc.* **2004**, *126*, 2963.
(36) Gaemers, S.; Bax, A. *J. Am. Chem. Soc.* **2001**, *123*, 12343.
(37) Hare, B. J.; Prestegard, J. H.; Engelman, D. M. *Biophys. J.* **1995**, *69*, 1891.
(38) Chung, J.; Prestegard, J. H. *J. Phys. Chem.* **1993**, *97*, 9837.
(39) Glover, K. J.; Whiles, J. A.; Wu, G.; Yu, N. J.; Deems, R.; Struppe, J. O.; Stark, R. E.; Komives, E. A.; Vold, R. R. *Biophys. J.* **2001**, *81*, 2163.
(40) Vold, R. R.; Prosser, R. S. *J. Magn. Reson. B* **1996**, *113*, 267.
(41) Arnold, A.; Labrot, T.; Oda, R.; Dufourc, E. J. *Biophys. J.* **2002**, *83*, 2667.
(42) Nieh, M. P.; Glinka, C. J.; Krueger, S.; Prosser, R. S.; Katsaras, J. *Langmuir* **2001**, *17*, 2629.
(43) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571.
(44) Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065.
(45) McCallum, S. A.; Pardi, A. *J. Mol. Biol.* **2003**, *326*, 1037.
(46) Williams, D. C.; Cai, M.; Clore, G. M. *J. Biol. Chem.* **2004**, *279*, 1449.
(47) Tycko, R.; Blanco, F. J.; Ishii, Y. *J. Am. Chem. Soc.* **2000**, *122*, 9340.
(48) Sass, H. J.; Musco, G.; Stahl, S. J.; Wingfield, P. T.; Grzesiek, S. *J. Biomol. NMR* **2000**, *18*, 303.
(49) Chou, J. J.; Gaemers, S.; Howder, B.; Louis, J. M.; Bax, A. *J. Biomol. NMR* **2001**, *21*, 377.
(50) Meier, S.; Haussinger, D.; Grzesiek, S. *J. Biomol. NMR* **2002**, *24*, 351.
(51) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179.
(52) Beger, R. D.; Marathias, V. M.; Volkman, B. F.; Bolton, P. H. *J. Magn. Reson.* **1998**, *135*, 256.
(53) Contreras, M. A.; Ubach, J.; Millet, O.; Rizo, J.; Pons, M. *J. Am. Chem. Soc.* **1999**, *121*, 8947.
(54) Bertini, I.; Felli, I. C.; Luchinat, C. *J. Biomol. NMR* **2000**, *18*, 347.
(55) Barbieri, R.; Bertini, I.; Cavallaro, G.; Lee, Y. M.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **2002**, *124*, 5581.
(56) Ma, C.; Opella, S. J. *J. Magn. Reson.* **2000**, *146*, 381.
(57) Wohnert, J.; Franz, K. J.; Nitz, M.; Imperiali, B.; Schwalbe, H. *J. Am. Chem. Soc.* **2003**, *125*, 13338.
(58) Kung, H. C.; Wang, K. Y.; Goljer, I.; Bolton, P. H. *J. Magn. Reson. B* **1995**, *109*, 323.
(59) Al-Hashimi, H. M.; Valafar, H.; Terrell, M.; Zartler, E. R.; Eidsness, M. K.; Prestegard, J. H. *J. Magn. Reson.* **2000**, *143*, 402.
(60) Fushman, D.; Ghose, R.; Cowburn, D. *J. Am. Chem. Soc.* **2000**, *122*, 10640.
(61) Meiler, J.; Peti, W.; Griesinger, C. *J. Biomol. NMR* **2000**, *17*, 283.
(62) Hus, J. C.; Bruschweiler, R. *J. Chem. Phys.* **2002**, *117*, 1166.
(63) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Nat. Struct. Biol.* **1997**, *4*, 292.
(64) Al-Hashimi, H. M.; Gosser, Y.; Gorin, A.; Hu, W.; Majumdar, A.; Patel, D. J. *J. Mol. Biol.* **2002**, *315*, 95.
(65) Peti, W.; Meiler, J.; Bruschweiler, R.; Griesinger, C. *J. Am. Chem. Soc.* **2002**, *124*, 5822.
(66) Tolman, J. R. *J. Am. Chem. Soc.* **2002**, *124*, 12020.
(67) Meiler, J.; Peti, W.; Griesinger, C. *J. Am. Chem. Soc.* **2003**, *125*, 8072.
(68) Berthault, P.; Jeannerat, D.; Camerel, F.; Alvarez Salgado, F.; Boulard, Y.; Gabriel, J. C. P.; Desvaux, H. *Carbohydr. Res.* **2003**, *338*, 1771.
(69) Losonczi, J. A.; Prestegard, J. H. *J. Biomol. NMR* **1998**, *12*, 447.
(70) Tian, F.; Valafar, H.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 11791.
(71) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6264.
(72) Bax, A.; Tjandra, N. *J. Biomol. NMR* **1997**, *10*, 289.

(73) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732.
(74) Prestegard, J. H.; Tolman, J. R.; Al-Hashimi, H. M.; Andrec, M. In *Biological Magnetic Resonance: Structure Computation and Dynamics in Protein NMR*; Kluwer Academic/Plenum Publishers: New York, 1999; Vol. 17.
(75) Mollova, E. T.; Pardi, A. *Curr. Opin. Struct. Biol.* **2000**, *10*, 298.
(76) Martin Pastor, M.; Bush, C. A. In *NMR Spectroscopy of Glyco-Conjugates*; Jimenez Barbero and Peters, Eds.; Wiley-VCH: Weinheim, 2002.
(77) Prestegard, J. H.; Glushka, J. In *NMR Spectroscopy of Glyco-Conjugates*; Jimenez-Barbero and Peters, Eds.; Wiley-VCH: Weinheim, 2002.
(78) Vander Kooi, C. W.; Kupce, E.; Zuiderweg, E. R. P.; Pellecchia, M. *J. Biomol. NMR* **1999**, *15*, 335.
(79) Feher, K.; Berger, S.; Kover, K. E. *J. Magn. Reson.* **2003**, *163*, 340.
(80) Pham, T. N.; Liptaj, T.; Barlow, P. N.; Uhrin, D. *Magn. Reson. Chem.* **2002**, *40*, 729.
(81) Tolman, J. R.; Prestegard, J. H. *J. Magn. Reson. B* **1996**, *112*, 245.
(82) Tjandra, N.; Bax, A. *J. Magn. Reson.* **1997**, *124*, 512.
(83) Pham, T. N.; Liptaj, T.; Bromek, K.; Uhrin, D. *J. Magn. Reson.* **2002**, *157*, 200.
(84) Kontaxis, G.; Clore, G. M.; Bax, A. *J. Magn. Reson.* **2000**, *143*, 184.
(85) Umemoto, K.; Leffler, H.; Venot, A.; Valafar, H.; Prestegard, J. H. *Biochemistry* **2003**, *42*, 3688.
(86) Ottiger, M.; Delaglio, F.; Bax, A. *J. Magn. Reson.* **1998**, *131*, 373.
(87) Ding, K.; Gronenborn, A. M. *J. Magn. Reson.* **2003**, *163*, 208.
(88) Giesen, A. W.; Bae, L. C.; Barrett, C. L.; Chyba, J. A.; Chaykovsky, M. M.; Cheng, M. C.; Murray, J. H.; Oliver, E. J.; Sullivan, S. M.; Brown, J. M.; Dahlquist, F. W.; Homans, S. W. *J. Biomol. NMR* **2001**, *19*, 255.
(89) Ishii, Y.; Markus, M. A.; Tycko, R. *J. Biomol. NMR* **2001**, *21*, 141.
(90) Vallurupalli, P.; Moor, P. B. *J. Biomol. NMR* **2002**, *24*, 63.
(91) Lerche, M. H.; Meissner, A.; Poulsen, F. M.; Sorensen, O. W. *J. Magn. Reson.* **1999**, *140*, 259.
(92) Srensen, M. D.; Meissner, A.; Srensen, O. W. *J. Magn. Reson.* **1999**, *137*, 237.
(93) Cordier, F.; Dingley, A. J.; Grzesiek, S. *J. Biomol. NMR* **1999**, *13*, 175.
(94) Brutscher, B. *J. Magn. Reson.* **2001**, *151*, 332.
(95) Luy, B.; Marino, J. P. *J. Magn. Reson.* **2003**, *163*, 92.
(96) Hitchens, T. K.; McCallum, S. A.; Rule, G. S. *J. Magn. Reson.* **1999**, *140*, 281.
(97) Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12366.
(98) Jain, N. U.; Noble, S.; Prestegard, J. H. *J. Biomol. NMR* **2003**, *328*, 451.
(99) Tugarinov, V.; Kay, L. E. *J. Mol. Biol.* **2003**, *327*, 1121.
(100) Wang, Y. X.; Marquardt, J. L.; Wingfield, P.; Stahl, S. J.; Lee-Huang, S.; Torchia, D.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 7385.
(101) Griesinger, C.; Meiler, J.; Peti, W. In *Protein NMR for the Millennium*; Rama Krishna, N., Berliner, L. J., Eds.; Kluwer Academic/Plenum Publishers: New York, 2003; Vol. 20.
(102) Yang, D.; Venters, R. A.; Mueller, G. A.; Choy, W. Y.; Kay, L. E. *J. Biomol. NMR* **1999**, *14*, 333.
(103) Griesinger, C.; Srensen, O. W.; Ernst, R. R. *J. Chem. Phys.* **1986**, *85*, 6837.
(104) Bodenhausen, G.; Ernst, R. R. *J. Magn. Reson.* **1981**, *45*, 367.
(105) Mueller, G. A.; Choy, W. Y.; Yang, D.; Forman-Kay, J. D.; Venters, R. A.; Kay, L. E. *J. Mol. Biol.* **2000**, *300*, 197.
(106) Adams, M. W. W.; Dailey, H. A.; Delucas, L. J.; Luo, M.; Prestegard, J. H.; Rose, J. P.; Wang, B. C. *Acc. Chem. Res.* **2003**, *36*, 191.
(107) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2001**, *123*, 9490.
(108) The present approach was developed on a ^{15}N-only labeled rubredoxin, a small 54-amino acid protein. A similar approach, based on the recording of an IPAP-(HA)CANH experiment, was developed independently by Zweckstetter and Bax.[107]
(109) Mayer, K. L., Valafar, H, Leblond, P. and Prestegard, J. H. *J. Struct. Funct. Genomics* **2004**, submitted.
(110) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(111) Bersch, B.; Rossy, E.; Coves, J.; Brutscher, B. *J. Biomol. NMR* **2003**, *27*, 57.
(112) Zidek, L.; Wu, H.; Feigon, J.; Sklenar, V. *J. Biomol. NMR* **2001**, *21*, 153.
(113) O'Neil-Cabello, E.; Bryce, D. L.; Nikonowicz, E. P.; Bax, A. *J. Am. Chem. Soc.* **2004**, *126*, 66.
(114) Mittermaier, A.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 6892.
(115) Ottiger, M.; Delaglio, F.; Marquardt, J. L.; Tjandra, N.; Bax, A. *J. Magn. Reson.* **1998**, *134*, 365.
(116) Kontaxis, G.; Bax, A. *J. Biomol. NMR* **2001**, *20*, 77.
(117) Chou, J. J.; Bax, A. *J. Am. Chem. Soc.* **2001**, *123*, 3844.
(118) Carlomagno, T.; Peti, W.; Griesinger, C. *J. Biomol. NMR* **2000**, *17*, 99.
(119) Kaikkonen, A.; Otting, G. *J. Am. Chem. Soc.* **2001**, *123*, 1770.
(120) Sibille, N.; Bersch, B.; Covs, J.; Blackledge, M.; Brutscher, B. *J. Am. Chem. Soc.* **2002**, *124*, 14616.
(121) Dosset, P.; Hus, J. C.; Marion, D.; Blackledge, M. *J. Biomol. NMR* **2001**, *20*, 223.
(122) Mittermaier, A.; Kay, L. E. *J. Am. Chem. Soc.* **1999**, *121*, 10608.
(123) Vogeli, B.; Kovacs, H.; Pervushin, K. *J. Am. Chem. Soc.* **2004**, *126*, 2414.
(124) Hansen, M. R.; Rance, M.; Pardi, A. *J. Am. Chem. Soc.* **1998**, *120*, 11210.
(125) Tjandra, N.; Marquardt, J.; Clore, G. M. *J. Magn. Reson.* **2000**, *142*, 393.
(126) Wu, Z.; Bax, A. *J. Am. Chem. Soc.* **2002**, *124*, 9672.
(127) Meier, S.; Haussinger, D.; Jensen, P.; Rogowski, M.; Grzesiek, S. *J. Am. Chem. Soc.* **2003**, *125*, 44.
(128) Boisbouvier, J.; Delaglio, F.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11333.
(129) Tian, F.; Al-Hashimi, H. M.; Craighead, J. L.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 485.
(130) Martin Pastor, M.; Bush, C. A. *J. Biomol. NMR* **2001**, *19*, 125.
(131) Bolon, P. J.; Prestegard, J. H. *J. Am. Chem. Soc.* **1998**, *120*, 9366.
(132) Delaglio, F.; Wu, Z.; Bax, A. *J. Magn. Reson.* **2001**, *149*, 276.
(133) Tian, F.; Bolon, P. J.; Prestegard, J. H. *J. Am. Chem. Soc.* **1999**, *121*, 7712.
(134) Cai, M.; Wang, H.; Olejniczak, E. T.; Meadows, R. P.; Gunasekera, A. H.; Xu, N.; Fesik, S. W. *J. Magn. Reson.* **1999**, *139*, 451.
(135) Permi, P.; Kilpelainen, I.; Heikkinen, S. *Magn. Reson. Chem.* **1999**, *37*, 821.
(136) Lohr, F.; Schmidt, J. M.; Rterjans, H. *J. Am. Chem. Soc.* **1999**, *121*, 11821.
(137) References included in ref 136.
(138) Otting, G.; Ruckert, M.; Levitt, M. H.; Moshref, A. *J. Biomol. NMR* **2000**, *16*, 343.
(139) Peti, W.; Griesinger, C. *J. Am. Chem. Soc.* **2000**, *122*, 3975.
(140) Pellecchia, M.; Vander Kooi, C. W.; Keliikuli, K.; Zuiderweg, E. R. P. *J. Magn. Reson.* **2000**, *143*, 435.
(141) Brunger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr. D* **1998**, *54*, 905.
(142) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65.
(143) Guntert, P.; Mumenthaler, C.; Wuthrich, K. *J. Mol. Biol.* **1997**, *273*, 283.
(144) Banci, L.; Bertini, I.; Cremonini, M. A.; Gori-Savellini, G.; Luchinat, C.; Wuthrich, K.; Guntert, P. *J. Biomol. NMR* **1998**, *12*, 553.
(145) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. R.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.
(146) Tsui, V.; Zhu, L.; Huang, T. H.; Wright, P. E.; Case, D. A. *J. Biomol. NMR* **2000**, *16*, 9.
(147) Clore, G. M.; Starich, M. R.; Bewley, C. A.; Cai, M. L.; Kuszewski, J. *J. Am. Chem. Soc.* **1999**, *121*, 6513.
(148) Losonczi, J. A.; Andrec, M.; Fischer, M. W. F.; Prestegard, J. H. *J. Magn. Reson.* **1999**, *138*, 334.
(149) Hus, J. C.; Marion, D.; Blackledge, M. *J. Am. Chem. Soc.* **2001**, *123*, 1541.
(150) Prestegard, J. H.; Valafar, H.; Glushka, J.; Tian, F. *Biochemistry* **2001**, *40*, 8677.
(151) Delaglio, F.; Kontaxis, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 2142.
(152) Andrec, M.; Du, P.; Levy, R. M. *J. Biomol. NMR* **2001**, *21*, 335.
(153) Andrec, M.; Du, P.; Levy, R. M. *J. Am. Chem. Soc.* **2001**, *123*, 1222.
(154) Haliloglu, T.; Kolinski, A.; Skolnick, J. *Biopolymers* **2003**, *70*, 548.
(155) Rohl, C. A.; Baker, D. *J. Am. Chem. Soc.* **2002**, *124*, 2723.
(156) Annila, A.; Aitio, H.; Thulin, E.; Drakenberg, T. *J. Biomol. NMR* **1999**, *14*, 223.
(157) Valafar, H.; Prestegard, J. H. *Bioinformatics* **2003**, *19*, 1549.
(158) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791.
(159) Unligil, U. M.; Rini, J. M. *Curr. Opin. Struct. Biol.* **2000**, *10*, 510.
(160) Valafar, H.; Prestegard, J. H. *J. Magn. Reson.* **2004**, *167*, 228.
(161) Clore, G. M.; Bewley, C. A. *J. Magn. Reson.* **2002**, *154*, 329.
(162) Clore, G. M.; Schwieters, C. D. *J. Am. Chem. Soc.* **2003**, *125*, 2902.
(163) Fischer, M. W. F.; Losonczi, J. A.; Weaver, J. L.; Prestegard, J. H. *Biochemistry* **1999**, *38*, 9013.
(164) Goto, N. K.; Skrynnikov, N. R.; Dahlquist, F. W.; Kay, L. E. *J. Mol. Biol.* **2001**, *308*, 745.
(165) Evenas, J.; Tugarinov, V.; Skrynnikov, N. R.; Goto, N. K.; Muhandiram, R.; Kay, L. E. *J. Mol. Biol.* **2001**, *309*, 961.
(166) Bewley, C. A.; Clore, G. M. *J. Am. Chem. Soc.* **2000**, *122*, 6009.
(167) Al-Hashimi, H. M.; Tolman, J. R.; Majumdar, A.; Gorin, A.; Patel, D. J. *J. Am. Chem. Soc.* **2001**, *123*, 5806.
(168) Lukin, J. A.; Kontaxis, G.; Simplaceanu, V.; Yuan, Y.; Bax, A.; Ho, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 517.

(169) Bondensgaard, K.; Mollova, E. T.; Pardi, A. *Biochemistry* **2002**, *41*, 11532.
(170) Kikuchi, J.; Iwahara, J.; Kigawa, T.; Murakami, Y.; Okazaki, T.; Yokoyama, S. *J. Biomol. NMR* **2002**, *22*, 333.
(171) Lukavsky, P. J.; Kim, I.; Otto, G. A.; Puglisi, J. D. *Nat. Struct. Biol.* **2003**, *10*, 1033.
(172) Koenig, B. W.; Kontaxis, G.; Mitchell, D. C.; Louis, J. M.; Litman, B. J.; Bax, A. *J. Mol. Biol.* **2002**, *322*, 441.
(173) Bolon, P. J.; Al-Hashimi, H. M.; Prestegard, J. H. *J. Mol. Biol.* **1999**, *293*, 107.
(174) Shimizu, H.; Donohue-Rolfe, A.; Homans, S. W. *J. Am. Chem. Soc.* **1999**, *121*, 5815.
(175) Stevensson, B.; Landersjo, C.; Widmalm, G.; Maliniak, A. *J. Am. Chem. Soc.* **2002**, *124*, 5946.
(176) Azurmendi, H. F.; Martin Pastor, M.; Bush, C. A. *Biopolymers* **2002**, *63*, 89.
(177) Nakanishi, T.; Miyazawa, M.; Sakakura, M.; Terasawa, H.; Takahashi, H.; Shimada, I. *J. Mol. Biol.* **2002**, *318*, 245.
(178) Niccolai, N.; Ciutti, A.; Spiga, O.; Scarselli, M.; Bernini, A.; Bracci, L.; Di Maro, D.; Dalvit, C.; Molinari, H.; Esposito, G.; Temussi, P. A. *J. Biol. Chem.* **2001**, *276*, 42455.
(179) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. *J. Am. Chem. Soc.* **2003**, *125*, 1731.
(180) Donaldson, L. W.; Skrynnikov, N. R.; Choy, W. Y.; Muhandiram, D. R.; Sarkar, B.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 9843.
(181) Mal, T. K.; Ikura, M.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 14002.
(182) Briggman, K. B.; Tolman, J. R. *J. Am. Chem. Soc.* **2003**, *125*, 10164.
(183) Ottiger, M.; Bax, A. *J. Biomol. NMR* **1998**, *12*, 361.
(184) Cavagnero, S.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1999**, *13*, 387.
(185) Ottiger, M.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 187.
(186) Ruckert, M.; Otting, G. *J. Am. Chem. Soc.* **2000**, *122*, 7793.
(187) Hansen, M. R.; Hanson, P.; Pardi, A. *Methods Enzymol.* **2000**, *317*, 220.
(188) Zweckstetter, M.; Bax, A. *J. Biomol. NMR* **2001**, *20*, 365.
(189) Koenig, B. W.; Hu, J.-S.; Ottiger, M.; Bose, S.; Hendler, R. W.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 1385.
(190) Sass, J.; Cordier, F.; Hoffmann, A.; Cousin, A.; Omichinski, J. G.; Lowen, H, Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 2047.
(191) Chou, J. J.; Kaufman, J. D.; Stahl, S. J.; Wingfield, P. T.; Bax, A. *J. Am. Chem. Soc.* **2002**, *124*, 2450.
(192) Riley, S. A.; Giuliani, J. R.; Augustine, M. P. *J. Magn. Reson.* **2002**, *159*, 82.
(193) Trempe, J. F.; Morin, F. G.; Xia, Z.; Marchessault, R. H.; Gehring, K. *J. Biomol. NMR* **2002**, *22*, 83.
(194) Barrientos, L. G.; Dolan, C.; Gronenborn, A. M. *J. Biomol. NMR* **2000**, *16*, 329.
(195) Barrientos, L. G.; Gawrisch, K.; Cheng, N.; Steven, A. C.; Gronenborn, A. M. *Langmuir* **2002**, *18*, 3773.
(196) Cutting, B.; Tolman, J. R.; Nanchen, S.; Bodenhausen, G. *J. Biomol. NMR* **2002**, *23*, 195.
(197) Tolman, J. R.; Prestegard, J. H. *J. Magn. Reson. B* **1996**, *112*, 269.
(198) Andersson, P.; Nordstrand, K.; Sunnerhagen, M.; Liepinsh, E.; Turovskis, I.; Otting, G. *J. Biomol. NMR* **1998**, *11*, 445.
(199) Yang, D. W.; Tolman, J. R.; Goto, N. K.; Kay, L. E. *J. Biomol. NMR* **1998**, *12*, 325.
(200) Chou, J. J.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **2000**, *18*, 101.
(201) Wienk, H. L. J.; Martnez, M. M.; Yalloway, G. N.; Schmidt, J. M.; Prez, C.; Ruterjans, H.; Lohr, F. *J. Biomol. NMR* **2003**, *25*, 133.
(202) Heikkinen, S.; Permi, P.; Kilpelainen, I. *J. Magn. Reson.* **2001**, *148*, 53.
(203) Permi, P.; Sorsa, T.; Kilpelainen, I.; Annila, A. *J. Magn. Reson.* **1999**, *141*, 44.
(204) Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 4690.
(205) Meissner, A.; Duus, J..; Sorensen, O. W. *J. Magn. Reson.* **1997**, *128*, 92.
(206) Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 12334.
(207) de Alba, E.; Suzuki, M.; Tjandra, N. *J. Biomol. NMR* **2001**, *19*, 63.
(208) Permi, P. *J. Biomol. NMR* **2003**, *27*, 341.
(209) Mittermaier, A.; Kay, L. E. *J. Biomol. NMR* **2002**, *23*, 35.
(210) Tian, F.; Fowler, C. A.; Zartler, E. R.; Jenney, F. A.; Adams, M. W.; Prestegard, J. H. *J. Biomol. NMR* **2000**, *18*, 23.
(211) Moglich, A.; Wenzler, M.; Kramer, F.; Glaser, S. J.; Brunner, E. *J. Biomol. NMR* **2002**, *23*, 211.
(212) Permi, P.; Rosevear, P. R.; Annila, A. *J. Biomol. NMR* **2000**, *17*, 43.
(213) Tjandra, N.; Tate, S.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 6190.
(214) Brutscher, B.; Boisbouvier, J.; Pardi, A.; Marion, D.; Simorre, J. P. *J. Am. Chem. Soc.* **1998**, *120*, 11845.
(215) Yan, J.; Corpora, T.; Pradhan, P.; Bushweller, J. H. *J. Biomol. NMR* **2002**, *22*, 9.
(216) Wu, Z.; Bax, A. *J. Magn. Reson.* **2001**, *151*, 242.
(217) Wu, Z.; Tjandra, N.; Bax, A. *J. Biomol. NMR* **2001**, *19*, 367.
(218) Luy, B.; Marino, J. P. *J. Biomol. NMR* **2001**, *20*, 39.
(219) Luy, B.; Barchi, J. J.; Marino, J. P. *J. Magn. Reson.* **2001**, *152*, 179.
(220) Wu, Z.; Ono, A.; Kainosho, M.; Bax, A. *J. Biomol. NMR* **2001**, *19*, 361.
(221) Verdier, L.; Sakhaii, P.; Zweckstetter, M.; Griesinger, C. *J. Magn. Reson.* **2003**, *163*, 353.
(222) Neubauer, H.; Meiler, J.; Peti, W.; Griesinger, C. *Helv. Chim. Acta* **2001**, *84*, 243.
(223) Martin Pastor, M.; Canales-Mayordomo, A.; Jimenez-Barbero, J. *J. Biomol. NMR* **2003**, *26*, 345.
(224) Meissner, A.; Duus, J..; Sorensen, O. W. *J. Biomol. NMR* **1997**, *10*, 89.

CR030419I