ELSEVIER

Progress in Nuclear Magnetic Resonance Spectroscopy 46 (2005) 23–61

PROGRESS IN NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

www.elsevier.com/locate/pnmrs

Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings

Martin Blackledge

Institut de Biologie Structurale Jean-Pierre Ebel, U.J.F.–C.N.R.S.–C.E.A., 41 rue Jules Horowitz, 38027 Grenoble Cedex, France

Received 18 October 2004
Available online 24 January 2005

Keywords: Dipolar couplings; RDC; Structure; Dynamics; Protein; Nucleic acid

Contents

E-mail address: martin.blackledge@ibs.fr.

doi:10.1016/j.pnmrs.2004.11.002

1. Introduction

Nuclear magnetic resonance (NMR) spectroscopy is now established as the method of choice for the study of the structure and dynamics of proteins and nucleic acids in solution [1]. The success of this technique has relied primarily on the ability to measure a large number of cross-relaxation effects (nOe) between protons close together in space ($<$5–6 Å), providing approximate distance restraints mapping inter-proton proximity throughout the molecule of interest [2]. In combination with simple modelling of covalent and through-space interactions, these short-range constraints, correlating structural information from parts of the molecule which are far apart in primary sequence but close in three-dimensional coordinate space, allow accurate reconstruction of the secondary and tertiary molecular geometry of globular proteins [3]. The strength of nOe-based structure determination is however undermined by some significant limitations. For instance, the measurement of insufficient distances, particularly at interfacial or hinge regions, can result in ill-defined relative positioning of distant regions of extended or modular macromolecules. More fundamental barriers originate from the fact that nOes between potential partners should eventually be unambiguously identified in order to define the correct fold of the molecule, a task that becomes increasingly difficult as molecular size increases, and spectra become more complex. Finally, and possibly most importantly, nOes become difficult to measure in larger protonated molecules due to prohibitive transverse relaxation effects [4]. The development of alternative approaches to structure determination of larger biomolecules is therefore one of the major challenges for solution state NMR spectroscopy.

The first part of this review concerns the use of dipolar couplings measured in weakly aligned macromolecules for the study of the structure in the solution state. Dipolar couplings are normally averaged to zero in isotropic solution by the complete averaging of orientations experienced by the interaction tensor [5]. In order to measure this coupling it is necessary to reintroduce preferential alignment, or order, into the sample. Forty years ago Saupe and Englert [6] demonstrated that anisotropic interactions could be measured in the NMR spectrum of benzene by dissolving the small solute molecule in the nematic-phase of an organic solvent (*p*-azoxyanisole), thereby inducing residual order in the molecular orientation. Following this, the field of liquid crystal NMR developed around the measurement of molecular structural and dynamic parameters in highly ordered systems from non-averaged anisotropic interactions [7]. Smaller molecules were later oriented with respect to the magnetic field using their own natural magnetic susceptibility anisotropy, either diamagnetic or paramagnetic in origin, and non-averaged dipolar and quadrupolar effects were measured [8–10]. Finally proteins were aligned relative to the field; initially in a protein containing a native paramagnetic centre (cyanometmyoglobin) with a significant magnetic susceptibility anisotropy [11], and then using the protein Ubiquitin, which is diamagnetic and has a much

smaller magnetic susceptibility anisotropy [12]. The field dependent dipolar couplings in the latter case were approximately 10 times smaller (less than 0.5 Hz), but the correlation of these numbers with expected values from known structure in both cases demonstrated the enormous potential of such measurements should they become routine and easily accessible. A further means of aligning molecules derives from the large magnetic susceptibility anisotropy of nucleotides, which, in stacked conformation, can also induce sufficient order to align the molecule sufficiently to give quite significant effects [13,14].

Despite these advances, the field of application of this potentially very valuable source of structural information remained restricted to molecules where high susceptibility anisotropies could be found or induced. The generalisation of the approach occurred following the remarkable and elegant demonstration [15] that simple dissolution of a protein in a dilute liquid crystal solution made up of phospholipid bicelles [16] would allow the measurement of large (tens of Hertz), easily measurable couplings, while retaining the high quality spectra necessary for high resolution protein NMR. Importantly, the degree of order was apparently 'tuneable' to allow induction of the ideal level of alignment, and the precision with which the parameters could be measured was very high. This was surely a key turning point for the field of solution state NMR. Very rapidly additional solvent systems were developed to provide the necessary partial alignment (e.g. [17–23]).

Since this initial demonstration, residual dipolar couplings or RDCs, as the non-averaged dipolar coupling has now come to be known, have evolved into one of the most important sources of structural and dynamic information available for biomolecules in solution [24–28]. The routine application of dipolar couplings to structure refinement, fold determination and quaternary structure determination of proteins and nucleic acids testifies to the ability of these measurements to provide relatively easy access to biomolecular conformation. The development of novel approaches to de novo structure determination based only on RDCs also demonstrates the precise geometric definition available when multiple RDCs are combined in local, rigid structural units.

NMR spectroscopy also occupies a unique position among biophysical techniques due to its particular sensitivity to the effects of intra- and intermolecular motions [29–32]. Rapid movements, up to the range of the characteristic rotational correlation time of the molecule (around 10 ns for medium size proteins in aqueous solution at room temperature), can be studied in a relatively straightforward manner using spin relaxation measurements [33]. These experimental methods, combined with an intuitively simple analysis in terms of dynamic amplitudes and frequencies [34,35], are now part of the standard set of experiments applied in many protein NMR studies [36]. While chemical shift exchange occurring on slower time-scales can be detected using relaxation dispersion experiments [37,38], routine measurement of slower backbone motions has so far remained elusive. These time-scales are of particular interest because a large number of functionally important biological processes are suggested to occur in this time range.

It is in terms of molecular dynamics that a second, equally important aspect of residual dipolar couplings becomes apparent [39]. RDCs are averages over all inter-spin orientations of the dipolar interaction that are sampled up to the time-scale defined by the inverse of the alignment-induced coupling, and consequently report on averages over relatively long time-scales (up to and beyond the millisecond range) [40]. Measurement of these parameters therefore provides direct access to key information for understanding protein motions on time-scales that are highly complementary to the dynamic picture derived from spin relaxation measurements in solution, where motional sampling is restricted to the faster (picosecond–nanosecond) time-scale. This relevance is particularly evident for ^{15}N–^{1}H couplings because rapid reorientation of the same inter-nuclear dipole–dipole interaction also dominates experimental ^{15}N relaxation rates. In this case, comparison of the motional averaging on the two time-scales can in principle deliver otherwise inaccessible information on the slower motional sampling of the interatomic vector. In the second part of this article, a review is presented of recent approaches to the interpretation of RDCs to extract information about the dynamic behaviour of biomolecules in solution.

This review is therefore a synopsis of recent developments for exploiting RDCs measured in partially aligned samples for the characterisation of biomolecular structure and dynamics. The emphasis is oriented towards the geometric nature of non-averaged dipolar couplings, covering diverse examples of the application from recent literature. In the interests of space, and to focus the review on analytical aspects, no attempt will be made to follow the numerous recent experimental developments to facilitate the routine measurement of RDCs. Interested readers are referred to recent reviews that cover this field extensively [27,28].

2. Analytical description of dipolar couplings in terms of molecular structure and dynamics

Here, we consider two non-equivalent spins (i, j) present in a partially aligned molecule, that is consequently experiencing restricted orientational sampling. The dipolar contribution to the observed splitting between i and j derives from the secular part of the magnetic dipole–dipole interaction between the spins. In the high field limit we

can write the effective Hamiltonian

$$H_{ij}^{\mathrm{D}}(t) = -\frac{\gamma_i\gamma_j\mu_0 h}{8\pi^3 r_{ij}^3(t)} I_{iz} I_{jz} \frac{(3\cos^2\theta_{ij}(t) - 1)}{2} \tag{1}$$

r_{ij} is the distance between the two nuclei, γ_i and γ_j are the gyromagnetic ratios of the two spins, and I_{mz} are angular momentum spin operators, h is Planck's constant, μ_0 the permittivity of free space, and θ_{ij} is the angle made between the (ij) inter-nuclear vector and the static magnetic field.

2.1. Incomplete averaging of the dipolar interaction

Note that the dipolar Hamiltonian depends on the orientation θ of the inter-nuclear vector between the coupled spins, relative to the magnetic field, following a second order Legendre polynomial dependence In solution state NMR, the measured dipolar coupling, is described by the time and ensemble average of this dipolar Hamiltonian over all sampled orientations

$$D_{ij} = -\frac{\gamma_i\gamma_j\mu_0 h}{8\pi^3} \left\langle \frac{(3\cos^2\theta_{ij}(t) - 1)}{2r_{ij}^3(t)} \right\rangle \tag{2}$$

This averaging, denoted by the angular brackets, reduces the measured coupling to zero when all orientations are equally probable, but to a non-zero value when there is an anisotropic distribution of orientations relative to the static field direction.

The simple averaging described in Eq. (2) contains information about the mean laboratory frame orientation of the inter-nuclear vector with respect to the magnetic field. In the case of a macromolecule, this average can therefore be described as a convolution of the restricted motion of the solute molecules, defined by the average over all orientations of the molecule relative to the magnetic field (ξ_x, ξ_y and ξ_z), and the orientation of the inter-nuclear vector with respect to the axes of the macromolecular frame (ζ_x, ζ_y and ζ_z) (see Fig. 1). The term $\cos\theta_{ij}$ can then be expressed as follows:

$$\cos\theta_{ij} = \begin{pmatrix}\cos\zeta_x\\ \cos\zeta_y\\ \cos\zeta_z\end{pmatrix}\begin{pmatrix}\cos\xi_x\\ \cos\xi_y\\ \cos\xi_z\end{pmatrix}$$
$$= \cos\zeta_x\cos\xi_x + \cos\zeta_y\cos\xi_y + \cos\zeta_z\cos\xi_z \tag{3}$$

Assuming that we can consider the molecule, or molecular domain of a molecule, to be rigid, we can then express Eq. (2) in the following form

$$\left\langle \frac{3\cos^2\theta_{ij} - 1}{2} \right\rangle = \frac{3}{2}\langle(\cos\zeta_x\cos\xi_x + \cos\zeta_y\cos\xi_y + \cos\zeta_z\cos\xi_z)^2\rangle - \frac{1}{2}, \tag{4}$$

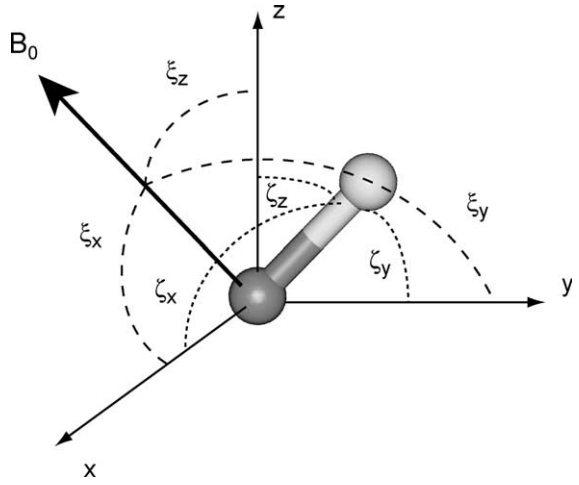

Fig. 1. Orientation of the B_0 field and the inter-nuclear vector in the molecular frame. The angles represent the orientation of the molecular frame (diagonal here) relative to the magnetic field (ξ_x, ξ_y and ξ_z), and the orientation of the inter-nuclear vector with respect to the axes of the macromolecular frame (ζ_x, ζ_y and ζ_z).

which we can recast in the following way using $C_i = \cos\zeta_i$ and $c_i = \cos\xi_i$:

$$\left\langle \frac{3\cos^2\theta_{ij} - 1}{2} \right\rangle = \frac{3}{2}[\langle c_x^2\rangle C_x^2 + \langle c_y^2\rangle C_y^2 + \langle c_z^2\rangle C_z^2 + 2\langle c_x c_y\rangle C_x C_y + 2\langle c_x c_z\rangle C_x C_z + 2\langle c_y c_z\rangle C_y C_z] - \frac{1}{2} \tag{5}$$

Note here that the only averaging that is occurring concerns the molecule relative to the magnetic field, since the two spins are assumed to be rigid within the molecular frame. This is clearly an approximation for spins present in a biomolecule in aqueous solution. We will come back to this point below, and correct this approximation to take into consideration local motion of the dipolar interaction relative to the molecular frame.

We can then describe the preferential orientational averaging of the molecule in terms of a symmetric order matrix (or alignment tensor) **A** whose units are dimensionless, and whose trace is zero [41]:

$$A_{kl} = \frac{3}{2}\langle\cos\xi_k\cos\xi_l\rangle - \frac{1}{2}\delta_{kl} \tag{6}$$

The measured RDC can then be described as:

$$D_{ij} = -\frac{\gamma_i\gamma_j\mu_0 h}{8\pi^3 r_{ij,\mathrm{eff}}^3} \sum_{kl} A_{kl}\cos\zeta_k\cos\zeta_l \tag{7}$$

Note that an effective inter-nuclear distance $r_{ij,\mathrm{eff}}$ has been introduced to account for averaging of this term. In a general molecule-fixed axis system, the tensor **A** has all elements non-zero. It is convenient to use a specific molecular frame, the principal axis system (PAS) in which the off-diagonal elements of the Saupe matrix are zero and only the diagonal terms, A_{xx}, A_{yy} and A_{zz} remain. The orientation of the PAS or

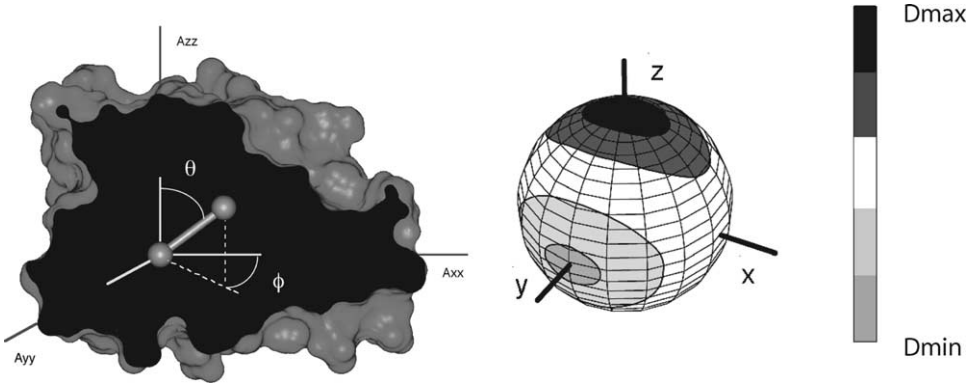

Fig. 2. Schematic representation of the dependence of RDC values on the orientation of the inter-dipolar vector and the alignment tensor. The available orientations are plotted on the surface of the sphere that is shaded as a function of the range of couplings. The range of RDC values is shown as a discrete distribution for reasons of clarity, this distribution should of course be continuous. In the example shown the alignment tensor is maximally rhombic.

alignment tensor with respect to the coordinate frame of the molecule can in return be defined simply via a three-dimensional Euler rotation $R(\alpha,\beta,\gamma)$.

The measured coupling can then be described in terms of the orientation of the inter-spin vector (the polar angles $\{\theta,\phi\}$ are often used) in the eigenframe of the alignment tensor, with eigenvalues A_{xx}, A_{yy} and A_{zz}

$$D_{ij}(\theta,\phi) = -\frac{\gamma_\iota\gamma_j\mu_0 h}{8\pi^3 r_{ij,\mathrm{eff}}^3}[A_{zz}\cos^2\theta + A_{xx}\sin^2\theta\cos^2\phi + A_{yy}\sin^2\theta\sin^2\phi] \quad (8)$$

or

$$D_{ij}(\theta,\phi) = -\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\left[A_\mathrm{a}(3\cos^2\theta - 1) + \frac{3}{2}A_\mathrm{r}\sin^2\theta\cos 2\phi\right] \quad (9)$$

By convention $|A_{xx}| \le |A_{yy}| \le |A_{zz}|$. $A_\mathrm{a} = A_{zz}/2$ is the axial component of the alignment tensor and $A_\mathrm{r} = (1/3)(A_{xx} - A_{yy})$ is the rhombic component. It is useful to note that the maximum measurable coupling is given when an inter-spin vector is aligned with the z-axis of the alignment tensor. For any known structure or sub-structure, we therefore have five parameters that determine the alignment tensor or Saupe matrix; (A_a, A_r, α, β, γ) using this formalism or A_{kl} from Eq. (6). These parameters can be determined directly via singular value decomposition [42] or by some alternative form of optimization, from a set of (≥ 5) RDCs. Variations on the formalisms given in Eqs. (7) and (9) are found in the literature, and are of course entirely equivalent. The 3×3 matrix **A** is known as the Saupe matrix and is presented as **S** or **A** as shown here.

2.2. Structural dependence of residual dipolar couplings

The orientations of an interaction vector consistent with a single measured residual dipolar coupling for a known alignment tensor are depicted in Fig. 2 on the surface of a sphere that is shaded as a function of the value associated with each direction. There is clearly a strong angular degeneracy—while extreme values can be easily identified as aligning the inter-nuclear vector along the limiting axes of the alignment tensor, a measured coupling with an intermediary value, and therefore the general case, can be aligned either along the $\pm y$-axis, in the $\pm x \pm z$ plane, or along a continuum of orientations in between. At first glance it would therefore appear that the structural importance of RDCs is somewhat limited, leading to a large number of potential solutions for the orientation of each inter-dipolar vector. It is clearly important to raise this degeneracy if RDCs are to be used to provide less ambiguous structural information. This can be achieved either by measuring more couplings in structures of known conformation, or by measuring RDCs in the presence of liquid crystals that orient the molecule differently, and thereby give rise to a different alignment tensor. These two approaches will be described in more detail below.

2.2.1. Raising orientational degeneracy: introduction of structural coherence

The orientational degeneracy continuum for a single coupling can be raised by measuring multiple couplings in a structural domain of known conformation such that the relative orientation of the different inter-nuclear vectors acts as a supplementary restraint as shown in Fig. 3.

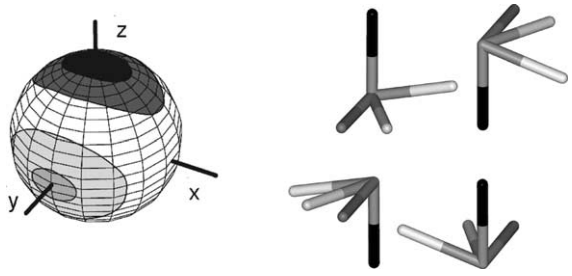

Fig. 3. Schematic representation of the structural information available when multiple couplings can be measured, whose relative orientation is known. A model chiral motif is shown here. This reduces the number of degenerate orientations to 4, related to each other via π rotations about the axes of the alignment tensor.

Sketched on the right-hand side of the figure, are the equivalent orientations for an imaginary sub-structure consisting of differently oriented vectors. There are now four equivalent orientations of the differently valued couplings (shaded in the figure), which are in agreement with measure values. This four-fold degeneracy is inherent to the orientation of any three-dimensional structure relative to a molecular alignment tensor, and derives from simple symmetry operations (180° rotations around A_{xx}, A_{yy} and A_{zz}).

2.2.2. Raising orientational degeneracy: multiple alignment systems

Further gain in orientational definition can be achieved by measuring RDCs in a molecule using samples which have different alignment tensors. As described above, this will require that we have access to liquid crystal media whose interactions with the solute molecules are different. The effect of the orientational definition for a single vector is sketched in Fig. 4a. The continuum of solutions in

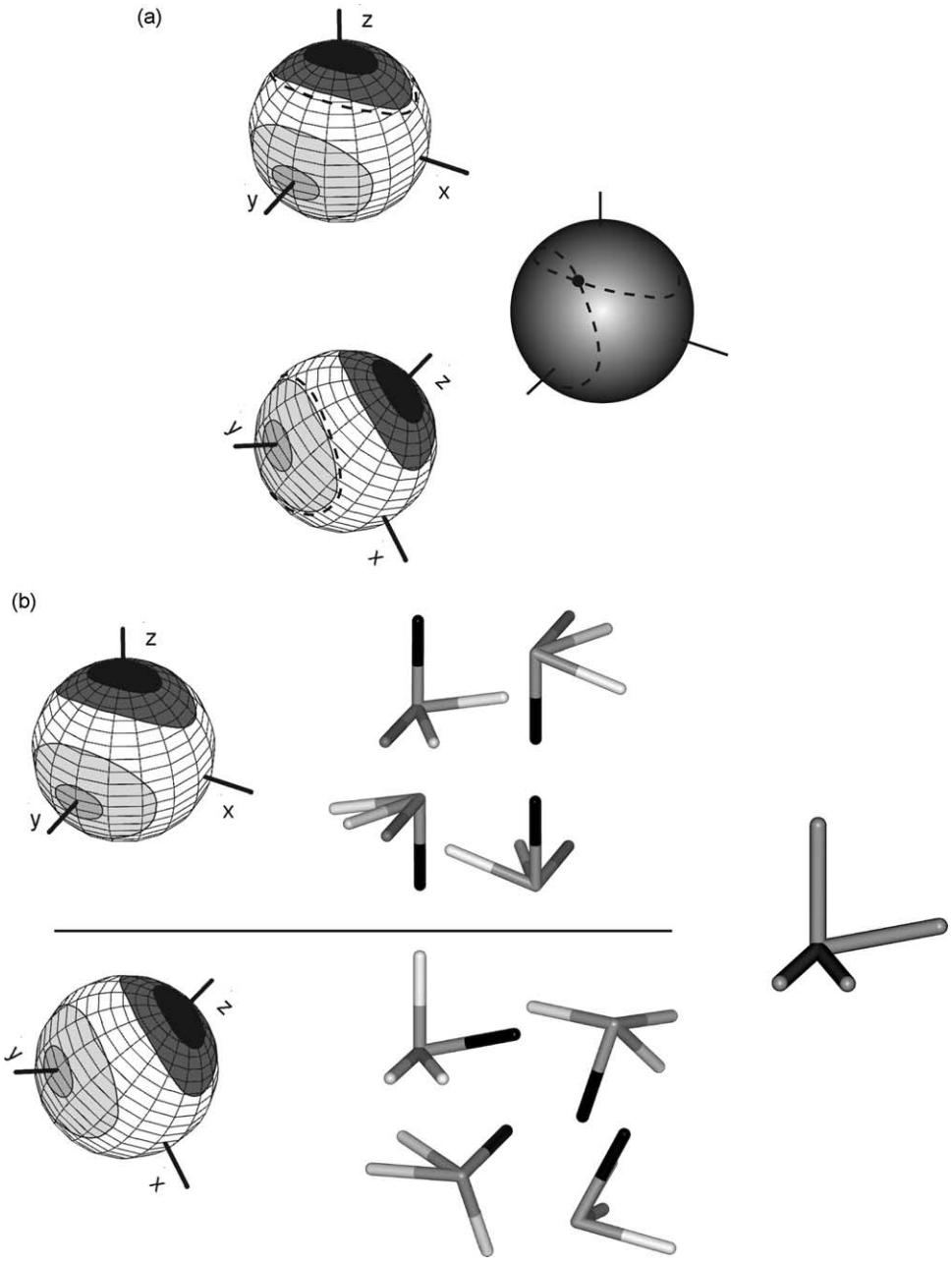

Fig. 4. (a) Schematic representation of the structural information available when two different alignment media are used to orient the molecule. The degenerate orientations giving rise to a single RDC value in the presence of one tensor (dashed line) can be resolved to fewer orientations (maximum 8) in the presence of a second, differently aligning medium (RDCs values shown on dashed line). The intersection points of the two sets of curves provides the solutions, as sketched on the right. (b) The orientation of a chiral motif in these conditions is determined uniquely. Four solutions exist for each tensor, but only one of these is common and therefore real.

the presence of a single tensor is mapped as a distorted cone of RDC-isocontour on the surface of the sphere. If we add data for the same coupling present in a sample with a second, different alignment tensor, the distorted cones intersect [43] to provide a maximum of eight equivalent solutions.

The gain in orientational precision when we have access to two alignment media is even clearer for the previously described case, where a three-dimensional, or chiral, motif was found to have four equivalent orientations in the presence of a single alignment medium (Fig. 4b). Four equivalent orientations also exist for the sub-structure in the presence of the second alignment tensor, related again to the correct orientation by π rotations about the axes of the alignment tensor. Only one of the four solutions (the correct orientation) is common to both media [44]. The structure is therefore uniquely oriented.

2.3. Dynamic averaging of residual dipolar couplings

In Eq. (3), it was assumed that the inter-spin vector has a fixed orientation with respect to the alignment tensor. In reality, in the presence of local internal motion the measured coupling is represented by a more complex relationship incorporating conformational averaging over both time and ensemble

$$\langle D_{ij}\rangle = \sum_n p_n D_{ij,n} \tag{10}$$

where p_n represents the respective weighting of the n conformational sub-states, whose individual dipolar couplings $D_{ij,n}$ are defined using Eq. (9):

$$D_{ij,n} = -\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,n}^3}\left[A_{a,n}(3\cos^2\theta_n - 1) + \frac{3}{2}A_{r,n}\sin^2\theta_n\cos 2\phi_n\right]. \tag{11}$$

The subscript n refers to the value of the relevant parameter when the molecule adopts conformation n. If we can assume that the motion is strictly local and exerts negligible influence on the molecular alignment tensor, and that the inter-nuclear distance averaging is again incorporated into an effective distance $r_{ij,\mathrm{eff}}$, Eq. (10) can be simplified by restricting the conformational averaging to the terms concerning the orientation of the inter-nuclear vector

$$\langle D_{ij}(\theta,\phi)\rangle = -\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\left[A_a\langle 3\cos^2\theta - 1\rangle + \frac{3}{2}A_r\langle\sin^2\theta\cos 2\phi\rangle\right] \tag{12}$$

where the angular brackets indicate averaging over all sampled conformations. This is a common approximation when small amplitude, local dynamics are studied, but is of course unjustified in the case of large-scale motions that would affect the induced alignment of the molecule. The terms in Eq. (12) can be recast in terms of averaged spherical harmonics, containing both structural and dynamic information:

$$\langle D_{ij}(\theta,\phi)\rangle = -\sqrt{\frac{16\pi}{5}}\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\left[A_a\langle Y_{20}(\theta,\phi)\rangle + \sqrt{\frac{3}{8}}A_r(\langle Y_{22}(\theta,\phi)\rangle + \langle Y_{2-2}(\theta,\phi)\rangle)\right] \tag{13}$$

The measured value clearly depends both on the mean orientation of the vector with respect to the alignment frame, and the local dynamic averaging with respect to this mean. It is often convenient to reorient the motional trajectory of each site into a common frame, so that each site can be equivalently analysed. Following the logic presented previously [45–47], this can be achieved using the rotational properties of spherical harmonics to reorient this average into a frame defined by the common local geometry (for example the peptide plane). If we apply an Euler rotation $R(\alpha,\ \beta,\ \gamma)$, the averaged spherical harmonics are then transformed as follows

$$R(\alpha,\beta,\gamma)Y_{2,M}(\theta,\phi) = \sum_{M'=-2}^{+2} \mathrm{e}^{-\mathrm{i}\alpha M'} d^{(2)}_{M'M}(\beta)\mathrm{e}^{-\mathrm{i}\gamma M} Y_{2,M'}(\theta,\phi)$$

where $d^{(2)}_{M'M}$ are the Wigner rotation matrices. Eq. (13) can then be written

$$\langle D_{ij}(\theta,\phi)\rangle = -\sqrt{\frac{16\pi}{5}}\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3} \times\left[A_a\left(\sum_{M'=-2}^{+2}\mathrm{e}^{-\mathrm{i}\alpha M'}d^{(2)}_{M'0}(\beta)\langle Y_{2,M'}\rangle\right) + \sqrt{\frac{3}{8}}A_r\left(\sum_{M'=-2}^{+2}\mathrm{e}^{-\mathrm{i}\alpha M'}d^{(2)}_{M'2}(\beta)\mathrm{e}^{-\mathrm{i}2\gamma}\langle Y_{2,M'}\rangle + \sum_{M'=-2}^{+2}\mathrm{e}^{-\mathrm{i}\alpha M'}d^{(2)}_{M'-2}(\beta)\mathrm{e}^{+\mathrm{i}2\gamma}\langle Y_{2,M'}\rangle\right)\right] \tag{14}$$

The Euler rotation $R(\alpha,\ \beta,\ \gamma)$ here defines the transformation from the alignment tensor frame to the local frame. $R(\alpha,\ \beta,\ \gamma)$ represents a different rotation for each vector, depending on the mean orientation of the vector with respect to the alignment tensor, and allows the conformational averaging, defined only by the spherical harmonics to be considered equivalently for each local frame. We will encounter a number of examples where this formalism has been used to simplify analysis of local dynamic disorder from RDC.

It is instructive to describe the effects of local dynamics in the presence of specific geometric models. A number of analytical motional models have been proposed for the interpretation of spin relaxation measurements in proteins [34,48], possibly the simplest and most intuitive mode for

backbone vector motion describes free diffusion in a cone [49]. Anisotropic peptide plane motions around the $^{\alpha}C$–$^{\alpha}C$ axis have also been shown, from molecular dynamics simulations and spin relaxation measurements, to be both valid and useful in characterising fundamental peptide chain motions in proteins [50–54].

Analytical expressions for dynamically averaged RDCs assuming three basic motional models relevant for peptide plane motions are briefly presented in the following sections.

2.3.1. Axially symmetric motion: diffusion in a cone

Dynamics of a vector under this motional mode [49] are characterized by a symmetric distribution of the vector around an average position. Using Eqs. (13) and (14), we can reorient the distribution from the alignment frame into a local frame such that the average position of the vector of interest is taken as the direction of the z' axis of this frame. Due to axially symmetric motion

$$\langle e^{\pm i\phi'}\rangle = \langle e^{\pm 2i\phi'}\rangle = 0 \tag{15}$$

and consequently

$$\langle Y_{2\pm 2}(\theta',\phi')\rangle = \langle Y_{2\pm 1}(\theta',\phi')\rangle = 0 \tag{16}$$

Eq. (14) is then simplified to

$$\langle D_{ij}(\theta',\phi')\rangle = -\sqrt{\frac{16\pi}{5}}\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\langle Y_{2,0}(\theta',\phi')\rangle \times\left[A_a d_{00}^{(2)}(\beta)+\sqrt{\frac{3}{8}}A_r(d_{02}^{(2)}(\beta)e^{-i2\gamma} +d_{0-2}^{(2)}(\beta)e^{+i2\gamma})\right] \tag{17}$$

which becomes

$$\langle D_{ij}(\theta',\phi')\rangle = -\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\langle 3\cos^2\theta'-1\rangle\left[A_a\left(\frac{3\cos^2\beta-1}{2}\right)+\frac{3}{4}A_r\sin^2\beta\cos 2\gamma\right] \tag{18}$$

where $\beta=-\theta_{av}$ and $\gamma=-\phi_{av}$ with θ_{av}, and ϕ_{av} being the polar angles that define the geometry of the symmetry axis of the movement, equivalent to the mean inter-nuclear vector with respect to the alignment tensor. As a consequence, the motionally averaged RDC, $\langle D_{jk}(\theta',\phi')\rangle$, is dependent only on the amplitude of the motional excursions of the polar angle θ' in the local frame, and the orientation of the z-axis of this frame with respect to the alignment frame. In the presence of axially symmetric motion, the value is therefore simply linearly scaled with respect to the static value $D_{ij,\mathrm{static}}(-\theta_{\mathrm{static}},-\phi_{\mathrm{static}})$ or $D_{ij,\mathrm{static}}(\beta,\gamma)$:

$$\langle D_{ij}(\theta,\phi)\rangle = S_{\mathrm{axial}}D_{ij,\mathrm{static}}(\beta,\gamma). \tag{19}$$

Here, we introduce S_{axial} as the scaling factor that depends on the amplitude of the motion but not on the position of the vector with respect to the alignment tensor. This order parameter can be compared to the generalised order parameters, S^2, derived from NMR relaxation studies, that defines the plateau value of the internal angular autocorrelation function [34,35]. Expressed in terms of averaged spherical harmonics the generalised order parameter is equal to:

$$S_{\mathrm{rdc}}^2 = \frac{4\pi}{5}\sum_{M=-2}^{2}\langle Y_{2M}(\theta',\phi')\rangle\langle Y_{2M}^*(\theta',\phi')\rangle \tag{20}$$

Under an axially symmetric motional regime this simplifies to

$$S_{\mathrm{rdc,axial}}^2 = \frac{4\pi}{5}\langle Y_{2,0}(\theta',\phi')\rangle^2 = \left\langle\frac{3\cos^2\theta'-1}{2}\right\rangle^2 \tag{21}$$

2.3.2. Anisotropic peptide motions: Gaussian axial fluctuation

The Gaussian axial fluctuation (GAF) motional model describes the peptide movement as a distribution of conformations centred around a known orientation, with the ith peptide plane undergoing reorientations about the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis (Fig. 5) [50]. Eq. (14) can then be expanded using the known averaging properties of trigonometric functions under the influence of GAF-like motions

$$\langle\cos(m\theta)\rangle = \int_{-\infty}^{+\infty}p(\theta)\cos(m\theta)d\theta = \exp\left(\frac{-m^2\sigma^2}{2}\right) \tag{22}$$

where

$$p(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-\theta^2}{2\sigma^2}\right) \tag{23}$$

The rotation $R(\alpha,\beta,\gamma)$ is again defined to transform from the alignment tensor principal axis system into the local

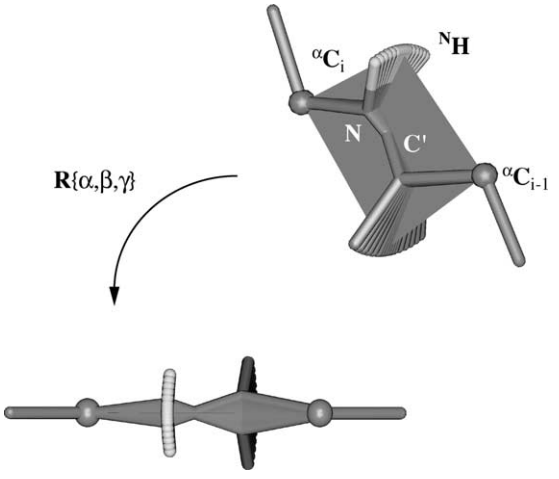

Fig. 5. Representation of anisotropic motion of the peptide plane about the mean $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis. GAF motion samples this range as a Gaussian probability distribution of width σ. The rotation $R(\alpha,\beta,\gamma)$ from the alignment tensor principal axis system into the local peptide plane frame (assumed diagonal in this figure) is indicated.

peptide plane frame. If we constrain the average vector orientation to lie along the z'-axis and the rotation axis to lie along the y' axis (in this case $\phi' = 0$ for the entire motion), a simplification of peptide plane motion (named *ortho*-GAF) can be derived as follows [55]

$$\langle D_{\mathrm{NH}}\rangle_{o-\mathrm{GAF}} = -\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\left[\frac{A_\mathrm{a}}{4}\{s_1(3\cos^2\beta - 1) + 3s_2\sin^2\beta\cos 2\alpha\} + \frac{3}{8}A_\mathrm{r}\{s_1\sin^2\beta\cos 2\gamma + 2s_2\left(\cos^4\frac{\beta}{2}\cos 2\delta_1 + \sin^4\frac{\beta}{2}\cos 2\delta_2\right)\}\right] \tag{24}$$

where

$$\delta_1 = \alpha + \gamma; \quad \delta_2 = \alpha - \gamma; \quad s_1 = 1 + 3\mathrm{e}^{-2\sigma^2}; \quad s_2 = 1 - \mathrm{e}^{-2\sigma^2} \tag{25}$$

By inspection this reduces to Eq. (9) when $\sigma = 0$, with $\beta = \theta$ and $\gamma = \phi$. This simplification has been shown to be relevant for the averaging of $^1D_{\mathrm{NH}}$ couplings in the case of peptide plane reorientations about the $^{\alpha}\mathrm{C}_{i-1}$–$^{\alpha}\mathrm{C}_i$ vector (where the vectors N–^{N}H and $^{\alpha}\mathrm{C}_{i-1}$–$^{\alpha}\mathrm{C}_i$ are not orthogonal, but are bisected by an angle of $\sim 100°$) [55]. An analytical expression for S^2 can be derived in the presence of the *ortho*-GAF motional model with an amplitude of σ [50]:

$$S^2 = 1 - \frac{3}{4}(1 - \mathrm{e}^{-4\sigma^2}) \tag{26}$$

2.3.3. Anisotropic peptide motions: two-site jump model

This model describes the situation where the site of interest experiences jumps between two well-defined, energetically degenerate, conformational states. Analogously to the previously described *ortho*-GAF model, it is convenient to describe the jumps as rotations around the $\mathrm{C}_{i-1}^{\alpha} - \mathrm{C}_{\mathrm{i}}^{\alpha}$ axis with an amplitude of 2θ. The orthogonal approximation has again been used in order to simplify the analytical development. In the peptide plane frame, the angular probability under a two-site jump motional model can be written

$$p(x) = \frac{1}{2}(\delta(x + \theta) + \delta(x - \theta)) \tag{27}$$

where θ represents the half-amplitude of the angular jump.

The equation of a dynamically averaged RDC under this motional regime becomes

$$\langle D_{\mathrm{NH}}\rangle_{\mathrm{jump}} = -\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\left[\frac{A_\mathrm{a}}{4}\{s_1'(3\cos^2\beta - 1) + 3s_2'\sin^2\beta\cos 2\alpha\} + \frac{3}{8}A_\mathrm{r}\{s_1'\sin^2\beta\cos 2\gamma + 2s_2'\left(\cos^4\frac{\beta}{2}\cos 2\delta_1 + \sin^4\frac{\beta}{2}\cos 2\delta_2\right)\}\right] \tag{28}$$

with

$$\delta_1 = \alpha + \gamma; \quad \delta_2 = \alpha - \gamma; \quad s_1' = 2(3\cos^2\theta - 1); \quad s_2' = 2\sin^2\theta$$

Note that the form of Eq. (28) developed for two-site jump model is equivalent to that developed for *ortho*-GAF with only terms s_1' and s_2' differing.

The generalised order parameter of a site experiencing a two-site jump of 2θ amplitude can be expressed as:

$$S^2 = \frac{3\cos^2 2\theta + 1}{4} \tag{29}$$

3. Partial alignment of macromolecules

3.1. Inducing order in the sample

3.1.1. Natural alignment

Almost all molecules in solution will align naturally in the magnetic field due to the anisotropy of the molecular magnetic susceptibility, although the net alignment of diamagnetic molecules is normally extremely small, in the range of 10^{-5}. If repetitive, similarly oriented structural motifs with significant magnetic susceptibility are present this alignment can be much larger, indeed some of the earliest measurements of residual dipolar couplings in partially aligned biomolecules were made using extended nucleic acid molecules, where susceptibilities of coplanar aromatic rings align additively [13,14]. There are distinct advantages to studying systems containing a natural paramagnetic centre with anisotropic magnetic susceptibility. In these systems, additional distance and orientation dependent interactions can be measured between the observed spins and the paramagnetic spin (e.g. pseudo-contact dipolar shifts [56,57], Curie-dipole cross-correlated relaxation [58,59]) to provide complementary conformational information. In combination, these parameters, measured for NH sites in a 128 amino acid cytochrome c', have been shown to be sufficient to determine the overall fold of the protein [60]. For a detailed overview of NMR phenomena that become measurable in the presence of a paramagnetic centre, and the application of their measurement to derive structural information, the reader is referred to a recent review from Bertini et al. [61].

The main drawback to paramagnetically induced alignment is that until very recently such methods had proven difficult to apply generally. A number of successful studies have however recently been carried out in which paramagnetic ions have been attached to biomolecules and apparently sufficient alignment of the protein achieved to derive structural and dynamic information from their measurement [62–65]. It is therefore to be expected that

these powerful methods will become increasingly popular for the study of larger biomolecules in solution.

3.1.2. Alignment by dissolution in a dilute liquid crystal

More generally, order can be induced by dissolving a molecule in a liquid crystal solvent. Initial demonstration of the feasibility of inducing order was achieved using circular bicelles [15], formed from a mixture of phospholipid such as dimyristoyl-phosphatidylcholine (DMPC) and a detergent-like lipid such as dihexanoyl-phosphatidylcholine (DHPC). In aqueous solution, this mixture aligns spontaneously in the magnetic field with the liquid crystalline director perpendicular to the magnetic field [16]. The identification of alternative alignment media for biological macromolecules rapidly followed, these were for the most part empirically developed to satisfy a certain number of basic criteria. Besides ordering the solute to a controllable extent, the medium should be stable over a range of experimental conditions (pH, temperature) and inert with respect to the structural, and hopefully functional, integrity of the solute molecule. Thus, popularly used alignment media currently range from elements of cellular material, such as filamentous bacteriophage [17,18] purple membrane fragments [19,20] to lyotropic alcohol-based mixtures [23] or mechanically stressed acrylamide gels [21,22].

Some important aspects pertaining to induced alignment in proteins are that the degree of order can be tuned as a function of concentration to achieve a suitable compromise between measurable couplings and spectral quality, although this will eventually be reduced as the degree of order increases, because of effective peak broadening (actually additional splitting) due to the presence of multiple significant stray couplings between protons. The nature of the interaction between the solute and the liquid crystal also plays an important role—as we will see later the ability to measure couplings in media that align the molecule differently (giving rise to an alignment tensor whose orientation is defined by different angles $R(\alpha, \beta, \gamma)$ relative to the molecular frame) is of fundamental importance [43]. Simply changing the orientation of the director of the liquid crystalline phase with respect to the field will not achieve this end, as this will only scale the measured couplings [43,66,67]. In order to change the effective alignment tensor it is necessary to change the nature of the interaction between the liquid crystal and the solute molecule. Alignment forces inducing order can be broadly classified into steric and non-steric in nature (see Section 3.3), although it is clear that different types of repulsion will often be present to varying extent and that the exact interaction between solute and liquid crystal particles remains difficult to control a priori.

3.2. Determination of the alignment tensor

One additional point, which will need to be addressed at some stage in any analysis designed to exploit RDCs in proteins, is that the magnitude and orientation of the alignment tensor will not be known a priori. If a sufficient number of independent RDCs have been measured within a motif of known structure it is possible to estimate these parameters, as described above, using singular value decomposition (SVD) or some alternative least-squares optimization procedure. Potential artefacts associated with these technique have been addressed in a recent study, identifying for example under-estimation of the tensor eigenvalues in the case of local error in structural coordinates [68].

The structure will not always be known, specifically if the object of the study is precisely to determine molecular conformation. In this case, the eigenvalues of the Saupe matrix, A_{xx}, A_{yy} and A_{zz} can be estimated from a histogram of the distribution of measured values. This is the Pake pattern, representing dipolar coupling values corresponding to all possible orientations of the interaction with respect to the B_0 field. One can either extract the alignment tensor eigenvalues from the limits of the distribution [69] or perform a least-squares fitting analysis of the entire shape to derive best-fitting values of these parameters [70,71]. These approaches of course require sufficient sampling of orientational space in order to be precise, so that it is often useful to scale all couplings measured between different heteronuclei in order to complete the directions available in angular space. The limitations incurred due to incomplete angular sampling of experimental parameters that depend on the orientation of an interaction vector relative to a second rank tensor, and the potential artefacts encountered in such procedures have been rigorously formalised by Fushman et al. [72].

In the case of planar structural motifs, only three independent dipolar couplings can be theoretically measured, all other couplings in the plane can then be expressed geometrically in terms of trigonometric combinations of these three and their relative orientation. This allows rapid validation that four or more measured couplings measured in the same plane are consistent with a single conformation [73], but also allows the creation of 'extended' histograms, filled with interpolated couplings derived from the original experimental data. It has recently been suggested that this simulation of angular sampling can provide an improved basis for determining the principal components of the alignment tensor using histogram-based approaches [74]. As we will see in Section 4.4, alignment tensor parameters can also be introduced into structure refinement algorithms as global parameters that can be determined simultaneously with the conformation.

3.3. Prediction of alignment properties

If the structure of the molecule is known, the expected alignment tensor may be estimated from the physical properties of the molecule combined with an appropriate description of the mechanism of molecular alignment. Two distinct interactions are expected to contribute to molecular

ordering; short-range repulsive forces that depend essentially on the size and shape of the solute, and long-range interactions depending on the electrostatic properties of both solute and liquid crystal. The field of liquid crystal NMR has addressed this question thoroughly over the last 40 years and no attempt will be made to review the field. It is nevertheless instructive to identify the principal approaches that have been used to describe the mechanism of solute alignment in liquid crystal phases. In order to model short-range interactions, a number of approaches have been proposed to describe the mean molecular field relevant for the solute particle [75]. For example, the anisotropy of the solvent–solute interaction has been parameterised in terms of model-free, direction-dependent interaction coefficients [76], or by defining an explicit ordering potential dependent on the van der Waals dimensions of the solute in each direction [77,78], or alternatively by considering the interaction of the solute with the solvent in terms of a second rank surface tensor and the molecule as an assembly of van der Waals spheres [79].

All of these approaches were developed and tested against experimental data from small organic molecules dissolved in thermotropic liquid crystal solvents. While the molecular alignment forces might be expected to be the same, it was not clear how well alignment could be predicted for the case of more complex biological macromolecules dissolved in aqueous solutions of ordered bicelles or biopolymers. Nevertheless the first application of a short-range, repulsive interaction model applied to the prediction of RDCs from partially aligned proteins reproduced experimental data remarkably well [80], at least for the case of apparently neutral alignment media. In their study, Zweckstetter and Bax used a numerical approach (PALES: prediction of alignment from structure) to efficiently span the relevant orientational space relative to a planar obstacle, ruling out all obstructed orientations, and averaging the Saupe matrices for the non-excluded orientations. The authors have in this case attempted a quantitative analysis so that absolute values of the experimental couplings can be predicted, by calibrating the bicelle concentration and invoking an order parameter (0.8) to account for incomplete bicellar alignment relative to the field. Although the PALES approach is computationally expensive, and therefore probably not suitable for direct incorporation into structure calculation algorithms, its utility was immediately demonstrated for the study of the dimeric form of cyanovirin-N. Experimental RDC were matched to values predicted from shape analysis of docked combinations of the two symmetric monomers, resulting in identification of an antiparallel dimer that could be distinguished from other possible solutions [81].

Non-atomistic models have also been proposed, with the aim of reducing the complexity of the numerical simulation required for the PALES approach. In the first of these, an analytical solution to the steric alignment of an axially symmetric particle in the presence of a planar obstacle was developed [82]. The authors predicted the orientation distribution function of a pure ellipsoid, defined by semi-axes derived from the gyration tensor, and, despite the simplicity of the structural model, demonstrated good agreement with experimental data from eight different proteins. Azurmendi and Bush [83] have also proposed a method based on transforming the information about shape asymmetry present in the molecular inertia tensor into effective alignment features (tracking alignment from the moment of inertia tensor: TRAMITE). The efficiency of this approach was demonstrated by reproducing RDCs from a highly flexible oligosaccharide. A related algorithm, this time based on the hydrodynamic shape, has been used to derive an empirical relationship between diagonal elements of the order matrix and a shape factor, δ, defined by the gyration tensor [84]. In comparison with numerical simulations performed using the atomistic approach PALES, the authors observe better reproduction of simulated data than using the analytical expression of Fernandes et al. [82] for higher rhombicity molecules.

In the case of a non-neutral liquid crystal medium, long-range electrostatic interactions between solute and solvent can be equally important for the effective alignment of the solute molecule, this became immediately apparent with the first studies of charged bicelles [43], and bacteriophage [17,18] (negatively charged), that resulted in entirely different alignment properties to those exhibited by the same molecule dissolved in a neutral liquid crystal. In view of the importance of increasing the number of differently aligning liquid crystal media for each molecule of interest, there is significant interest in the correct prediction of electrostatic alignment. The first successful modelling of macromolecular alignment in dilute filamentous phage was performed by Ferrarini [85], who described electrostatic effects in terms of Coulomb interactions between charges on the solute surface, described at an atomistic level, and a mean electrostatic potential, calculated according to Poisson–Boltzmann theory. This was combined with a hard-body repulsion description of the steric alignment contribution. The simplification of the virus as a uniformly charged rod did not prevent Ferrarini from reproducing both alignment characteristics and the effects of ionic strength and virus concentration and dimension. Independently, Zweckstetter et al. [86] extended their prediction of steric obstruction to incorporate an electrostatic contribution. For each non-excluded orientation from PALES, an electrostatic energy is calculated, again solving the Poisson–Boltzmann equation, treating the biomolecule in terms of a surface charge distribution immersed in an electrostatic potential. Again, the ability to reproduce experimental behaviour as a function of ionic strength is demonstrated, and the relative importance of contributions from net charge, dipolar and quadrupolar moments are analysed. The authors identify potentially limiting approximations implicit in these approaches. The most important are no doubt the continuous description of the electrostatic field exerted by a complex

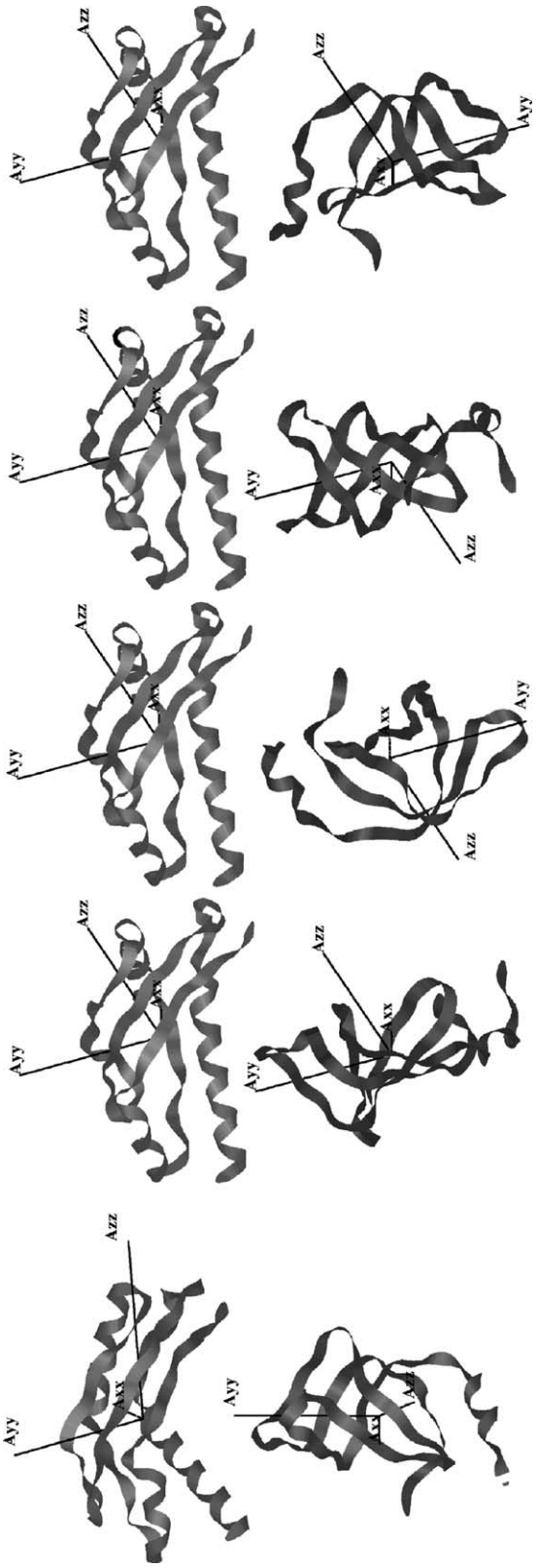

Fig. 6. Schematic representation of the use of RDCs to orient two partners in a two-domain complex. On the left-hand side the alignment tensors determined for the individual domains (top and bottom) are shown in the molecular reference frame of each partner. On the four panels to the right of this the two domains have been placed in a common frame, and the four equivalent orientations of the lower domain are shown relative to the fixed, upper domain. This figure was produced using the program Module [111].

alignment system, without taking into account local surface charge inhomogeneity, the supposed invariance of the electrostatic potential in the presence of the charged solute molecule, and possibly most importantly, conformational disorder of charged sidechains on the protein surface. Nevertheless the successful reproduction of alignment behaviour demonstrated in these two studies is an important step in our understanding of induced ordering effects on biomolecules dissolved in dilute charged liquid crystals, and it is to be expected that progress in this direction in the future will lead to rational selection of liquid crystal media for the alignment of proteins.

4. Use of RDCs in structure determination of proteins and nucleic acids

In this section, recent methodological developments for the use of dipolar couplings in the structure determination of biomolecules are reviewed. The thread of this part of the review will be guided by the idea of structural coherence, introduced above, and the combination of interactions of known relative orientation to characterise molecular alignment properties and determine structure. We will thus initially describe the most obvious, and possibly most frequently exploited application of RDCs in solution NMR, that is their use to determination the relative orientation of domains of known structure in multi-domain biomolecules and the determination of ligand–receptor geometry. This will be followed by applications to fold validation and comparative analysis using conformational databases. A description will then be given of structure determination combining RDCs with locally defined sub-structures, from peptide fragments to individual planes, before reviewing some aspects of restrained molecular dynamics approaches to structure refinement using RDCs as orientational restraints.

4.1. Determination of the relative orientations in biomolecules

It has become clear that the structural coherence present in sub-units of known conformation can be intelligently combined with RDC data to determine the overall geometry of molecular systems. In the explanatory case of a two-domain complex shown in Fig. 6, the structures of the component parts of the molecule are assumed to be known, but not their relative orientation in the complex. If sufficient couplings have been measured in each of the two domains, the alignment tensors for the two domains can be determined independently. This number can theoretically be as few as 5; as this is the number of unknown parameters required to define the alignment tensor/Saupe matrices of each of the domains (see above). In reality, in the presence of noise and non-ideal angular sampling, more couplings will be required in order to achieve sufficient accuracy.

Assuming that the two domains are in a stable complex they experience the same alignment tensor, and once the tensors have been determined with respect to the individual molecular reference frames, we can simply align the molecules such that the component axes are parallel. The four degenerate orientations of one of the domains are shown relative to the second, fixed domain in Fig. 6. It is of course necessary to check the validity of the assumption that the different domains experience the same tensor. As we do not know the relative orientation of the tensors a priori, this essentially means comparing the magnitude of the rhombic and axial components (A_a and A_r) of the alignment tensor. The observation of different effective alignment characteristics of domains within a macromolecule or complex may be evidence of differential dynamics between the component parts (vide infra).

4.1.1. Studying quaternary geometry of multi-domain molecules using RDCs

Despite the inherent four-fold orientational degeneracy of one oriented partner with respect to another, the ability to determine domain orientation is a very powerful complement to classical structure determination and forms the basis of many studies of the molecular architecture of multi-domain systems (e.g. [46,87–90] for recent review see [91]). This complementarity between NMR and, for example, X-ray crystallography, allows characterisation of molecular properties that are otherwise difficult to extract in solution. For example, Skrynnikov and co-workers [46] used backbone RDCs measured from the 379 amino acid maltodextrin binding domain in complex with β-cyclodextrin, in combination with different X-ray structures of samples of the enzyme crystallized in the presence or absence of ligands, to determine the relative orientation of the two domains of this protein in solution. The relative orientation of the different domains could be characterised using three angles (Θ, Φ, ω), while the parameters defining the common alignment tensor (A_a, A_r, α, β, γ) were simultaneously optimised in an eight-parameter non-linear least-squares optimisation procedure. The conformation in solution was found to be significantly (11°) more closed than in the crystal lattice. The same rigid-body approach was applied to the study of the relative orientation of the two domains of T4 lysozyme where in this case the solution conformation was found to be 17° more open than in the crystal structure [92]. In this case, the important observation was made that very similar relative orientations were determined using RDC measured in four differently composed alignment media, indicating that the average conformation is apparently common to the four alignment conditions and its parameterisation robust with respect to experimental conditions.

A series of studies performed by the Pardi group have exploited this capacity of RDCs to define long-range order in RNA molecules, where alternative methods are not yet available in solution. In the first example, a similar rigid-body approach to that presented above was used to determine the relative orientation of two rigid sub-units of $tRNA^{Val}$ from *E. coli* [89]. The local structure of the two sub-units were modelled either on the basis of the structure of these stem regions in the X-ray structure of yeast tRNA or using canonical helical geometry, and 24 RDCs measured from the molecule aligned in Pf1 phage were used to determine, and subsequently align the two domains. The four-fold relative orientational degeneracy could be raised using nOe and geometric considerations of the linker regions. A second example from the same laboratory allowed the determination of the long-range orientation of the three helical stem regions of the hammerhead ribozyme from a total of 41 N–H and C–H RDCs [93]. The four-fold degeneracy could be solved in this case by selecting the only combination of the 16 possible solutions that could be accommodated by the central junction region common to all three stem fragments. The final overall geometry in agreement with the RDCs was found to be significantly different to the geometry found in the X-ray structure, in particular the angle between the main axes of stems I and II was found to increase by nearly 120°.

These examples demonstrate the power of relatively accessible experimental RDC data to provide unique information on long-range solution conformation in large RNA molecules [94]. It is also worth mentioning that the low density of protons in nucleic acid molecules allows the induction of higher levels of alignment than are possible in protonated proteins. Increased molecular alignment allows larger RDCs to be measured, leading to higher levels of confidence in the structural analysis.

4.1.2. Constructing molecular assemblies from RDCs and translational restraints

Dipolar couplings are also complementary to many readily available sources of structural constraint currently used to build models of molecular assemblies in solution, whether these are experimental, such as intermolecular nOe or chemical shift perturbation measured at interacting surfaces, or predicted from existing structural information, for example electrostatic or hydrophobic surface calculations. The study of macromolecular complexes by NMR in solution is often hampered by the fact that inter-domain nOe may be sparse, possibly allowing identification of the situation of the interacting surfaces, but precluding a more precise definition of the overall structure of the complex (for review see [95]). Similarly, while mapping of chemical shift perturbations to define the interaction surfaces of the two partners can broadly locate interaction sites, construction of a structural model is often impossible using only this information. In common applications, for example, the study of molecular complexes where individual component structures are known, dipolar couplings can provide the necessary additional orientational information for the construction of a molecular model. Using these data it is possible to determine the four equivalent relative orientations of individual domains, and subsequently use distance

restraints to identify the correct solution in agreement with all data, as well as positioning the two partners to better characterize the interaction surface. A number of approaches have thus been developed to achieve this end, combining RDCs with rigid, or semi-rigid-body molecular mechanics or dynamic methods using distance restraints derived from intermolecular nOe [96], or using chemical shift mapping [97,98] to determine the tertiary architecture of macromolecular complexes.

One macromolecular complex that has been studied in great detail by solution NMR is the 40 kDa phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. This system has thus served as a test-system for the development of novel methodologies for the determination of protein–protein complex geometry using RDC. The solution structure was initially determined following a Herculean effort by the group of Clore and Gronenborn [88], using 3500 h of measurement time and 14 distinct mixtures of differently isotopically labelled proteins (essentially for spectral filtering purposes) to finally extract 5474 experimental NMR restraints, of which over 3000 were nOe, 110 of which were intermolecular. A total of 244 N–^{N}H RDC distributed evenly over the two proteins were also incorporated into the calculation. In this case, the conformation of the individual domains were determined entirely using NMR data, however, Clore [96] later demonstrated that the structure of the complex could also be determined by driving the two rigid partners from positions far apart in Cartesian space into the conformation native to the complex using rigid-body molecular dynamics with only RDC and intermolecular nOe data. In this case, X-ray crystallographic structures of the two proteins were available and were used in the modelling procedure. The X-ray structures were shown to be close to the conformations found in the complex, by comparison with the individual backbone RDC data sets, but a relaxation period was nevertheless allowed for the side-chains to accommodate the protein interface. Clore noted that the significant number of intermolecular distances are sufficient to determine the overall geometry of the two partners, even in the absence of RDCs, but that when this number is artificially reduced, mimicking more difficult experimental systems, the correct orientation of the two partners relies heavily on the N–^{N}H RDC. Further applications of this approach to studies of HPr in complexation with the glucose specific enzyme IIA^{Glc} [99] and as a complex with cytoplasmic A domain of the mannitol transporter II^{Mannitol} [100] verified the rapid applicability of this approach to the characterisation of protein–protein complexes in solution.

Unambiguous assignment of interprotein distances between two relatively large systems is still a difficult and time-consuming procedure, and a number of groups have attempted to reduce the necessary information required to construct reliable protein–protein complexes. McCoy and Wyss [97] have proposed a method using only chemical shift perturbations (induced at surface residues of the two proteins due to the proximity of the partner in the complex) and residual dipolar couplings. RDCs were used to align the partners and chemical shifts were calculated in different translational positions using the program SHIFTS, developed by Case [101]. The principal perturbation influencing chemical shifts at the surface is assumed to be due to the presence of aromatic rings in the partner structure, that are therefore implicitly assumed to be correctly oriented and static. The authors nevertheless reproduce a complex with quaternary geometry resembling that of the complete NMR determination (2.5 Å rmsd from superposition of the C$^{\alpha}$ atoms), and note that this was not possible in the absence of RDC data.

Clore and Schwieters [98] have taken a similar approach to reducing the necessary data set required for docking the three complexes described above, again removing the nOe constraints entirely, and this time transforming ^{1}H/^{15}N chemical shift mapping on the surface of the different proteins into empirical ambiguous restraints between these sites, again in combination with residual dipolar couplings. The translation between the proteins is again performed using rigid-body/torsion angle dynamics, using a molecular mechanics force field incorporating an additional term minimizing the radius of gyration of the complex to avoid potential expansion at the protein–protein interface. In the case of the enzyme I/HPr complex, unambiguous and correct positioning was again possible using only chemical shift mapping and RDC data, while for enzyme IIA^{Glc}/HPr two distinct solutions were found to agree equally well with the experimental target function. These could be distinguished by comparing, amongst other criteria, the expected alignment tensors for the two candidates with the experimentally determined tensor from the RDC data.

Finally it is worth noting that Bonvin and co-workers [102] developed a protein–protein docking approach that also uses rigid-body energy minimization, followed by semi-rigid simulated annealing in torsion angle space and refinement with explicit solvent. Ambiguous interaction restraints identified from, for example, chemical shift-mapping, are used to drive the component parts of the enzyme I/HPr and enzyme IIA^{Glc}/HPr complexes mentioned above into their correct conformations. Notably in this case only chemical shift perturbation data were used and no RDCs, although the approach is clearly highly compatible with orientational restraints.

4.1.3. Geometry of protein–ligand complexes from RDCs

Surely one of the most promising aspects of solution state NMR is the ability to characterise protein–ligand interactions in solution, that, despite their great importance for drug discovery, can be difficult to study using other structural biology tools. In this respect, RDCs are again very powerful experimental restraints, allowing the determination of the relative position of the interacting partners [103–105]. Under conditions of fast exchange it is possible to report on the bound conformation of a ligand. In this case,

the measured RDCs originate in the bound state but are detected on the free ligand with the spectral properties characteristic of the small molecule, allowing precise characterisation of the bound state.

The Prestegard group demonstrated that the particular averaging properties of symmetric oligomers with a three-fold or higher symmetry axis imposes colinearity on the axes of the inherently axially symmetric alignment tensor and the symmetry axis of the oligomer [106]. They applied this logic to the determination of the relative orientation of α-methyl mannose (AMM), for which RDCs were measured, and for which an axially symmetric alignment tensor was indeed found, and the 53 kDa mannose-binding protein trimer for which data were not collected, but for which the alignment tensor was deduced from the oligomeric symmetry [105]. In rapidly exchanging systems, observed couplings will be a population weighted average of the bound and free forms of the aligned peptide, and the contribution of the latter may not be negligible, especially under conditions of ligand excess. Alignment of the free form of the peptide was therefore calibrated, in the absence of the protein, and the bound-form RDCs derived from the known association constant of the interaction. This study revealed inconsistencies between the crystal-derived ligand-binding geometry and the RDC-derived orientations. In a follow-up study [107], the same group have attempted to identify the source of this disagreement by directly studying the relative geometry of the three components of the trimer in solution. In this case, the backbone resonances were assigned, allowing the orientation of the alignment tensor to be determined, thereby verifying the relative orientation of the monomers in the trimer. Subsequent analysis of C–H RDCs measured at natural abundance in the trimannoside ligand resulted in reproduction of the observed differences in protein–ligand orientation to those found in the crystal structure.

Under certain circumstances the natural alignment of the protein can obviate the need to employ an alignment medium. Apart from the obvious advantage of simplifying the experimental system, this can also be useful as it diminishes the problem caused by averaging of aligned, free and differently aligned, complexed RDCs from the ligand, because the free ligand experiences nearly negligible alignment. In a particularly elegant example, the orientation of an analogue of the C-terminal fragment of the α-sub-unit of G-protein bound to photo-activated rhodopsin was characterised using RDCs [108]. Rhodopsin is the major constituent of disk-shaped vesicles from rod outer segments of bovine retinas that align spontaneously in the magnetic field. Importantly rhodopsin undergoes fast (>10 kHz) axial rotation within the membrane, reducing the measured couplings by a further factor of $(3\cos^2\alpha - 1)/2$ where α is the angle between the membrane normal and the magnetic field. In a careful analysis of the expected behaviour of the observed couplings, the authors firstly show that the transferred RDCs are simply scaled between the bound-value and the observed value as long as k_{off}, the 'off-rate', is much greater than the bound-state coupling, and observe that the net alignment tensor measured on the peptide should be axially symmetric. The extracted RDCs, in combination with distance restraints derived from transferred nOe measurements, were then used to precisely determine the bound ligand conformation.

The precision with which RDCs can define subtle changes in local structure was also illustrated in the study of a 23 peptide fragment from HIV-I envelope protein GP41 bound to micelles and bicelles and aligned in compressed polyacrylamide gel [109]. In both cases, a helical structure is adopted by the peptide as determined from RDCs and chemical shift information however the curvature of the helix was found to increase significantly when bound to the micelle, corresponding to the natural form of the smaller spherical particle. The authors note that such small differences would be difficult to pick up using classical structural restraints.

4.1.4. Tools for rigid-body modelling using RDCs

The interpretation of RDCs for the determination of relative domain orientation requires tools specifically developed for the manipulation of sub-structures within a common reference calculation frame. Currently available graphical molecular modelling programs are not yet adapted to handling this kind of specific analysis. A number of programs have therefore been developed by research groups active within the field to facilitate this kind of analysis (see for example SVD [42], Dipocoup [110], PALES [80]). As an example, Module [111] has been developed for the determination of the alignment characteristics of macromolecules or sub-structures of macromolecules in solution from RDCs and their subsequent analysis in terms of a common alignment frame. The graphical interface allows the user to define modules from the primary sequence, or sequences—these can be contiguous (for example α helices) or non-contiguous (for example β sheets)—and to check the quality of the fit of the three-dimensional structure to these data from correlation diagrams. Tensor parameters are calculated using least-squares minimization of the target function over all couplings associated with a given domain

$$\chi^2 = \sum_n \{D_{ij}^{\exp} - D_{ij}^{\text{calc}}\}^2/\sigma_{ij}^2 \tag{30}$$

where σ_{ij} is the uncertainty in the experimentally measured coupling. The minimization algorithm searches the $\{A_a, A_r, \alpha, \beta, \gamma\}$ parametric space using simulated annealing [112], fuzzy logic [113], and Levenberg–Marquardt minimization [114]. This approach is entirely equivalent to using singular value decomposition (SVD) to find a best-fit solution to simultaneous equations in the form of Eq. (7) to extract values for A_{ij}, the elements of the Saupe matrix [41]. The directions of tensor axes are superimposed on the structural motifs and correlation plots presented for each different coupling type, as well as local χ^2 values for each module. For multiple modules, the program determines the relative

orientation of individual structured domains and provides graphical user-driven rigid-body modelling of the different modules relative to the common tensorial frame after aligning the domains such that all tensors are collinear (equivalent to applying the appropriate Euler rotation $\{\alpha, \beta, \gamma\}$, to each module). Translational freedom in the common frame, and equivalent π rotations about the diagonalized (x, y, z) axes, can be used to position the different modules in this frame to find the correct model in best agreement with experimentally measured couplings. Expected covalent interactions can be incorporated and used to propose the most likely structure, while distance restraints can be used in combination with the orientational data to find an optimal geometry of the oriented structural motifs with respect to the common frame by minimizing the function

$$E_{\text{rstrnt}} = \sum_n \{d_{ij} - d_{ij}^{\text{rstrnt}}\}^2 \tag{31}$$

where d_{ij}^{rstrnt} are the experimentally measured or covalence distances. The program will also calculate, and localise the steric contacts due to atom-overlap between modules.

Recently, a new program REDCAT (residual dipolar coupling analysis tool) has become available [115]. This performs similar analyses, determining a set of alignment matrices for different regions of input structural coordinates from corresponding experimental RDCs, and allowing reorientation into a common alignment frame and subsequent combination with distance constraints for translational positioning. In this case, data from multiple alignment media can be handled, raising the four-fold orientational degeneracy of chiral motifs.

4.2. Combination of RDCs with database folds: structure validation and beyond

The increasing volume of experimentally determined structures present in databases has greatly improved the possibility of modelling structures simply on the basis of primary sequence similarity. Such models are however difficult to verify experimentally using solution NMR data prior to nOe-based structure determination, and were consequently difficult to exploit with confidence. It is therefore something of a revolution for the solution state structural biologist to have access to coherent structural information from throughout the molecule immediately following backbone resonance assignment, and one of the first and most obvious applications of RDCs was to use this information to validate hypothetical models of the protein of interest. This common tool was used in the NMR study of the 18 kDa flavodoxin-like sub-unit of sulphite reductase from *E. coli* (SiR-FP18). At the stage where ^{13}C, ^{15}N and ^{1}H backbone resonance assignment had been accomplished, three RDCs ($^{1}D_{NH}$, $^{1}D_{Ca-C'}$, $^{2}D_{HN-C'}$) were measured from throughout the protein backbone and compared to a homology-built structure, derived from primary sequence comparison. Individual analyses were performed using RDCs from the five different secondary structural elements (so that the local geometry could be assumed to be relatively well defined, and tensors be determined from sufficient couplings). These correspond to the central β-sheet treated as a single structural entity, and the four α-helices (Fig. 7). The magnitude of the alignment tensors for the individual domains were found to be similar, and the orientation of the tensor axes found to be co-axial relative to the molecular frame, providing good evidence that the organisation of the topology in the homology based model of SiR-FP18 is similar to that found in solution [116]. This is a typical example of the power of RDCs, being able to provide primary structural information on the fold of the molecule immediately following backbone resonance assignment. Following such a validation approach it would of course be preferable to refine the structure further against the RDCs using a restrained molecular dynamics calculation. Some of the more practical aspects involved in this kind of procedure are described in further detail in Section 4.4.

4.2.1. Structural homology from systematic database comparisons with RDCs

An obvious development of this kind of analysis is a more direct combination of automated or semi-automated primary sequence-based structural homology approaches with the comparative analysis of three-dimensional structure available from residual dipolar couplings. A number of groups have proposed the direct comparison of measured sets of RDCs with expected couplings from structural motifs present in databases, demonstrating that RDCs can be used to rapidly identify a complete or partial homologous fold from such comparative methods [110,117,118], simply on the basis of ordered coherence of sufficiently large populations of inter-nuclear vectors in orientational space. An evident extrapolation to this kind of technique would be to by-pass the assignment step altogether, and to rely simply on the sampled RDCs, translated into sampled directions (assuming the alignment tensor can be estimated), to recognise the correct fold from a database of folds. Some progress has recently been made in this direction by Valafar and Prestegard, who have been able to distinguish between cases where structures of unassigned proteins are related or unrelated, essentially by comparing Pake pattern profiles, using only N–^{N}H RDCs [119].

A more general incorporation of RDC analysis into the automatic prediction and validation of the tertiary fold in backbone-assigned proteins would be a major step forward in improving the efficiency of ab initio structure prediction in solution. RDCs can no doubt make a considerable contribution to this field, as they provide precisely the kind of long-range structural order that is the most difficult to confidently estimate using structure prediction algorithms. Not surprisingly two groups specialising in the field of structure prediction have successfully combined expertise in this field with the long-range structure available from RDC data to improve structures calculated in this way [120–122].

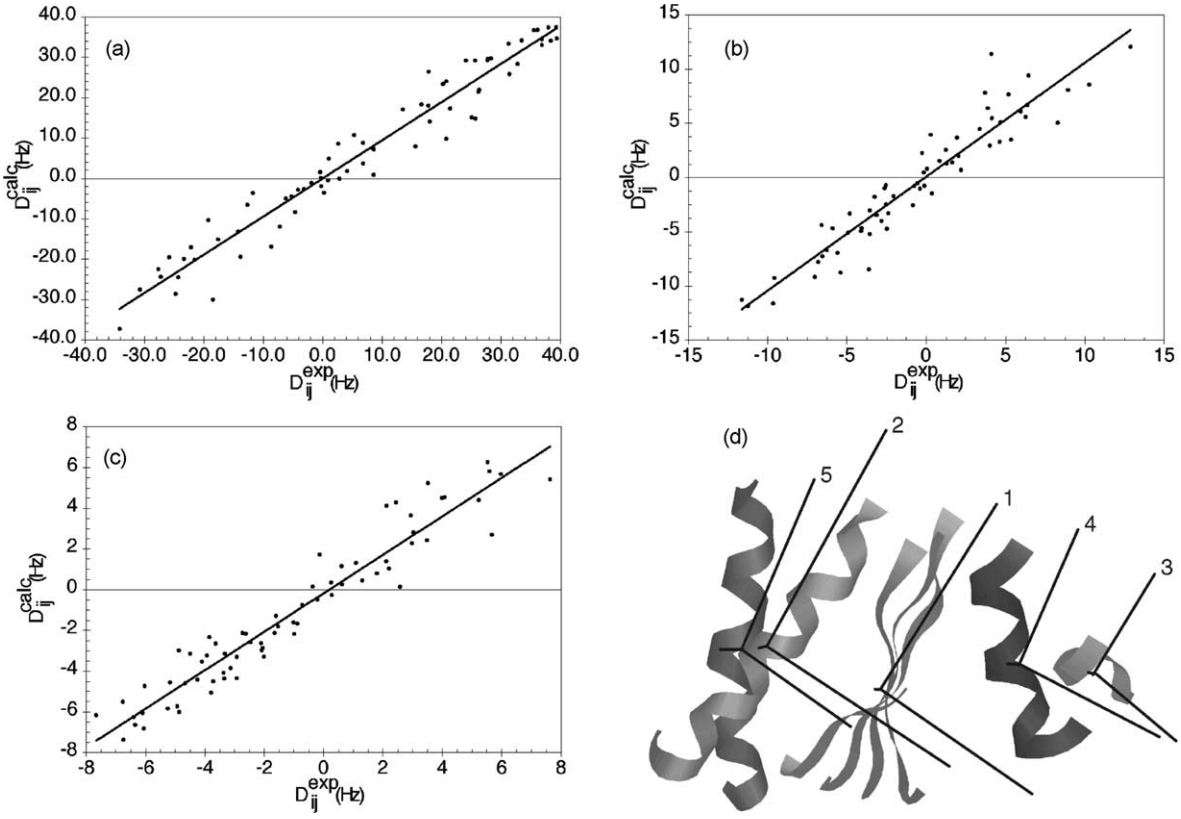

Fig. 7. Model validation using RDCs. The five different secondary structure elements were fitted as a single entity in the orientation of the homology modelled structure of SiR protein. Correlation plots between D_{ij}^{exp} and $D_{ij}^{back\ cal}$ for (a) $^1D_{NH}$, (b) $^1D_{C'HN}$, and (c) $^1D_{C'C\alpha}$. (d) Relative orientation of the five different secondary structure elements in the modelled structure of SiR protein and their associated alignment tensors. Numbers from 1 to 5 correspond to the central β-sheet, and the helices I, II, III, and IV, respectively. The length of the axes shown corresponds to the eigenvalues of the alignment tensors. The axial and rhombic components of the different elements are closely reproduced (data not shown). Analysis and figure using the program Module [111].

RDCs have thus been included in the program Rosetta [123], a structure prediction algorithm that extensively samples fold-space by randomly combining fragments using a Monte-Carlo simulated annealing search, and then selecting probable structures on the basis of similarity to pair-wise long-range interactions found in structure databases of natively folded proteins. The target function was modified to include a term measuring the ability of the models to reproduce experimental residual dipolar couplings. This incorporation of long-range structural information was found to significantly improve the ability of the program to predict the overall folds of backbone-assigned proteins [120]. Despite these examples, the scope of this potentially productive combination is as yet largely untapped and it is to be expected that the overlap between the fields will develop considerably in the future, with major developments occurring in environments where expertise in both ab initio structure prediction and partially aligned solution state NMR are in close contact.

The fact that structure is coded in RDCs and therefore represents primary structural information has also recently been exploited by Hus et al. [124] who demonstrated that in the case where the structure of a protein is already known it is possible to solve the 'inverse' problem and use this structural coherence to assign the backbone resonances of the protein using only three RDCs per amino acid together with chemical shift information.

It is important to note that local structure can also be identified using RDCs; for example it is often possible to recognise α-helical elements from the cyclic dependence of residual dipolar couplings along the primary sequence. This idea has been formalised and the cyclic dependence termed 'dipolar waves' [125] and further extended to β-sheet motifs [126]. The apparent precision with which local conformational information can be defined, combined with the evident long-range structural order inherent to RDCs, suggests a slightly different approach than previously discussed, whereby long-range order is used to define relative alignment, while local structure is extracted from sequential RDCs.

4.2.2. Sequential construction of proteins using molecular fragment replacement

Using this logic, Delaglio et al. [127] extrapolated database comparison with RDCs to an elegant conclusion by demonstrating that it is possible to unambiguously identify sequential segments from a structural database that are in best agreement with a seven-residue continuous strand of

experimental RDCs and chemical shift data. A sliding window of this length is shifted along the primary chain of the protein, superimposing the common regions of the best-fitting peptide conformers, while respecting the common alignment frames, to define the fold of the molecule (molecular fragment replacement). The method by-passes the problem of sequence-based identification of common regions in template structures by relying on the extent of conformational sampling of short peptide strands in available structural databases, and on the ability of multiple RDC measurements to recognise the correct strand conformation from the database sub-structures. Note that this approach used data from two different alignment media, and that the structural definition available under these circumstances is much improved compared to a single medium as will be discussed below. Andrec et al. [128] developed a conceptually similar approach, using longer windows and exploiting only data measured in one alignment medium. False positives due to the remaining degeneracy of orientational solutions were excluded by applying criteria based on similarity of overlapping regions.

More recently, a study by Zweckstetter and co-workers [129] demonstrated that assignment and structure can be simultaneously and automatically achieved using only unassigned chemical shift lists and associated sets of at least three backbone dipolar couplings per peptide unit, measured in a single alignment medium. The method applies standard methods of backbone assignment in an initial step to determine a minimum set of unambiguous RDCs and chemical shifts, and then applies molecular fragment replacement, comparing this information to database fragments, to select peptides of determined structure. A compact structure is then assembled using the ab initio structure calculation program RosettaNMR (vide supra) [120]. The best-fitting structures are then used in further rounds (typically four to eight are required) of assignment and structure calculation, until convergence is achieved. This approach was successfully applied to eight proteins to produce accurate models of known folds.

4.3. RDCs and ab initio structure determination

We have seen that the orientational degeneracy inherent to a single measured coupling can be raised by measuring different directions in a rigid body. As demonstrated above there remain four equivalent orientations of even a chiral motif with respect to a single alignment frame, implying that ab initio calculation of protein structure will be very difficult unless we can find a means to exclude the three incorrect orientations. An early study demonstrated that only very little distance information is actually necessary to unambiguously place secondary structural elements and select between the equivalent orientations to define protein fold, using for example sparse, but assigned, nOe data [130]. In a later study the same group [131] used RDC from two alignment media, combined with chemical shift information and scalar coupling data to construct strands of a small protein that were then combined using connectivity requirements and steric hindrance to find the fold. This study is all the more impressive as it was performed using ^{15}N-only labelled protein, and incorporated simultaneous resonance assignment into the analysis. It has also been shown that single alignment media can in theory be used to determine the correct conformation of the peptide chain as long as a good estimate of the backbone dihedral angles can be used to select among the potential solutions [132]. Similarly it has been shown that fewer RDCs can be used in combination with sparse nOe data to unambiguously determine the tertiary protein fold in terms of positioning of oriented secondary structural elements [133], or that global folds can be derived using a limited number of long-range ^{N}H–^{N}H nOes combined with three RDCs per peptide plane measured in two alignment media [134].

We have seen in Section 2.2.2 that orientational degeneracy can be raised by measuring RDCs in a second, differently aligning liquid crystal medium. In particular, the four-fold degeneracy inherent to the orientation of a chiral motif in the presence of a single medium can be raised to reveal an unambiguous solution. This implies that structure calculation using only residual dipolar couplings can be achieved if the following criteria can be satisfied:

(1) The biopolymer of interest can be divided into sub-structures whose local geometry is sufficiently accurately known.
(2) Sufficient RDCs can be measured for these sub-structures to determine their relative orientation.
(3) These couplings can be measured in two significantly differently alignment media.

It becomes immediately apparent when we attempt to apply this logic to the protein backbone that there are no individual chiral motifs whose local three-dimensional structure is known a priori, that is before the primary sequence has been structured to any extent, along the peptide chain. In fact, the protein backbone can be described as a chain of planar motifs, connected by tetrahedral junctions comprising the apex of plane i, plane $i+1$ and the $^{\alpha}$C–$^{\alpha}$H, $^{\alpha}$C–$^{\beta}$C plane. We are therefore obliged to consider the peptide plane as the only sub-unit of known structure for which numerous couplings can be measured. The orientation of a planar element has two-fold degeneracy when RDCs from two different alignment media are available. This introduces an orientational ambiguity between the correct alignment and the mirror image that must be solved for each planar element in the protein (Fig. 8a). In this case, it is necessary to develop approaches that can raise the two-fold degeneracy relevant to the orientation of the peptide plane from such data sets.

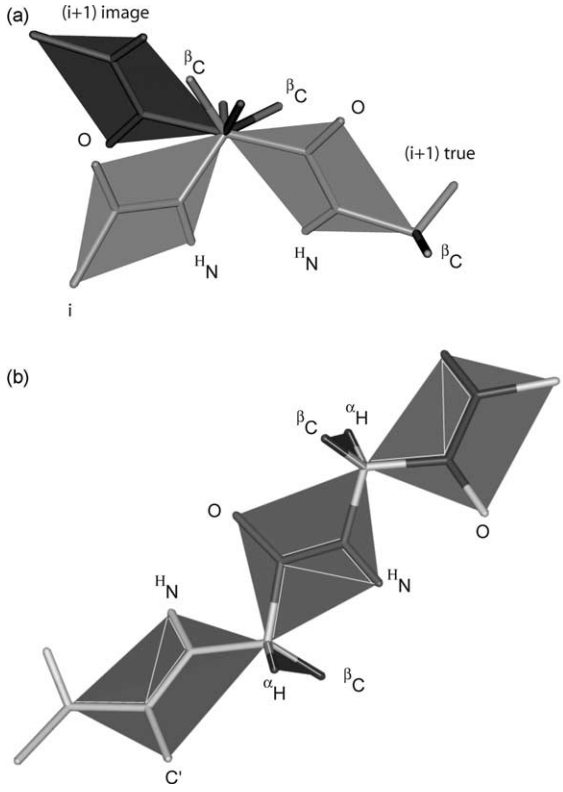

Fig. 8. Meccano (molecular engineering calculations using coherent association of non-averaged orientations). (a) Illustration of the two solutions to the orientation of a peptide plane in the presence of two different alignment tensors. The correct orientation can be distinguished from the image from $^{\alpha}$C–$^{\beta}$C and $^{\alpha}$C–$^{\alpha}$H couplings and expected covalence angles at the $^{\alpha}$C tetrahedral junction. (b) Using RDCs measured between different spins along the protein backbone (N–^{N}H, C′–^{N}H, C′–N, $^{\alpha}$C–C′, $^{\alpha}$C–$^{\beta}$C and $^{\alpha}$C–$^{\alpha}$H) in two different alignment media, the conformation of the peptide chain can then be defined uniquely as a function of the orientation of peptide planes and tetrahedral junctions.

4.3.1. Construction of the protein backbone from RDCs obtained in two alignment media

An approach has thus been proposed [135] to solve the backbone structure of a protein using only residual dipolar couplings obtained for nuclei throughout the peptide chain. The method has been named meccano (molecular engineering calculations using coherent association of non-averaged orientations), and feasibility was initially demonstrated using the dataset from Ubiquitin in bicelles and charged bicelles measured by the Bax group [136,137]. Two sets of 63 N–^{N}H, 61 C′–^{N}H, 61 and 63 C′–N and 59 and 54 $^{\alpha}$C–C′ couplings defining the peptide plane orientations, in addition to two sets of 62 $^{\alpha}$C–$^{\alpha}$H and one set of 39 $^{\alpha}$C–$^{\beta}$C couplings were used in the calculation. These are essentially the same data that were used in the first application of the molecular fragment replacement approach described earlier.

As mentioned earlier, in an ab initio calculation the alignment tensors will not be known a priori. The first step of the meccano approach is therefore designed to calibrate both tensors, in the absence of any structural information concerning the fold of the molecule. This is achieved using a least-squares-based search algorithm to determine the alignment tensors, described by seven parameters in the calculation frame (A_a^1, A_r^1, A_a^2, A_r^2, α, β, γ), where α, β, γ describe the orientation of $\mathbf{A}^2$ with respect to $\mathbf{A}^1$, taken to be diagonal in the calculation frame. Simultaneously, the orientation of each peptide plane $(\theta_1, \theta_2, \theta_3)_i$ is determined with respect to the calculation frame. The number of parameters to be determined is then $(3N+7)$ describing each plane orientation and the common reference frame, while the maximum number of independent RDCs available from N planes is $6N$, assuming we only include peptide planes with complete datasets in this stage of the calculation. This algorithm reliably finds the global minimum of the target function

$$\chi^2 = \sum_n \{D_{ij}^{\exp} - D_{ij}^{\text{calc}}\}^2/\sigma_{ij}^2 \tag{32}$$

over all measured couplings, requiring no a priori estimation of the alignment tensors. The second stage of the algorithm then constructs the molecule in this reference frame.

It turns out to be straightforward to construct the folded peptide chain from the known orientation of individual peptide planes (plane i is defined here as $^{\alpha}C_{i-1}$, C'_{i-1}, N_i, $^{\alpha}C_i$) and α-carbon junctions between the planes. Importantly for the success of this approach, the correct plane orientation can in theory be identified by the tetrahedral geometry requirements at the junctions connecting peptide planes, although for the general experimental case this is not always sufficient. The combination of RDCs measured in the peptide plane and tetrahedral junctions ($^{\alpha}C_{i-1}$–$^{\alpha}H_{i-1}$, $^{\alpha}C_{i-1}$–$^{\beta}C_{i-1}$) however allows unambiguous positioning. When no peptide plane orientation is available for plane $(i+1)$, ϕ_i/ψ_i values are optimized to reproduce ($^{\alpha}$C–$^{\alpha}$H, $^{\alpha}$C–$^{\beta}$C) from (i) and $(i+1)$ and peptide plane data from $(i+2)$. This allows the unambiguous continuation of the peptide chain (Fig. 8b). The final structure is therefore defined as a combination of oriented peptide planes and α-carbon junctions, and compares closely with the structure determined using 2727 nOes, in combination with RDCs, ^{3}J couplings and hydrogen bonding restraints (raw structure 2.0 Å backbone rmsd, and 1.0 Å following restrained molecular dynamics refinement to regularize non-bonding and covalent terms; Fig. 9).

The ability to determine protein structure ab initio simply on the basis of residual dipolar couplings measured along the protein backbone combined with rudimentary covalent considerations concerning the local peptide plane conformation is a promising development for the application of this kind of restraint. The dataset used is however extensive, and apparently highly accurate, raising the question as to whether this kind of analysis can be generalised to larger, and more difficult systems. We should remember that a great deal of free information was not included in this calculation, for example non-bonding terms were ignored, and secondary structural information, available from chemical shift analysis, was not taken into account. The future exploitation

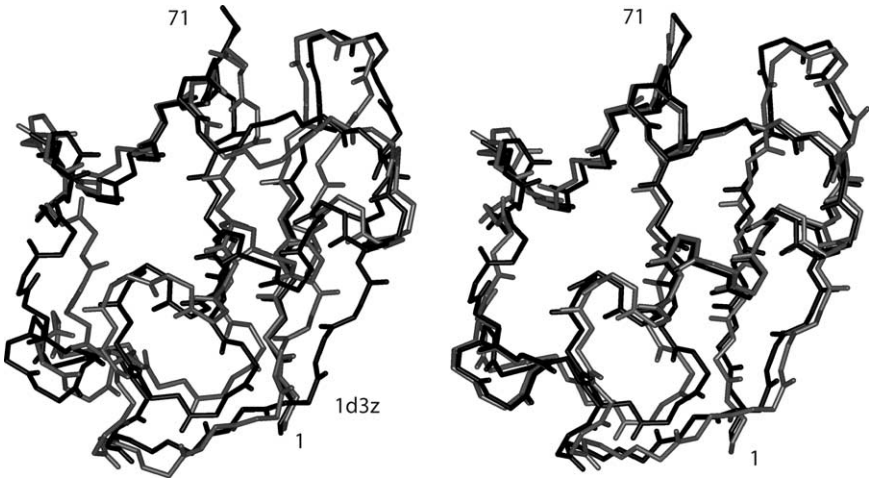

Fig. 9. Results of meccano calculation of the protein Ubiquitin. Left: the coordinates of the structure in comparison to the NMR structure determined using a full nOe/J-coupling/RDC restraint set (pdb code 1d3z). The rmsd at this stage is 2.0 Å between the two structures. Right: following RDC-based refinement/regularisation using the MD program Sculptor the two structures are within 1.0 Å rmsd.

of RDCs for ab initio calculation of protein structure will no doubt combine all of the available structural information to produce molecular models. It is worth noting that the approach described here is conceptually similar to work carried out using solid state NMR methods on molecules embedded in bilayers [138–140].

4.3.2. Focussed structure determination using RDCs

The meccano method has also been applied to determine the backbone conformation of a local site of interest in a somewhat larger molecule, Methionine sulphoxide reductase from *E. chrysanthemi* ($MsrA^{E.\ chmi}$, 221 amino acids). Extensive dipolar couplings were measured in ^{13}C, ^{15}N and ^{2}H labelled $MsrA^{E.\ chmi}$ in phage and alcohol-based alignment media. Using the fold validation approach presented earlier ($N–^{N}H$, $C'–^{N}H$, $^{\alpha}C–C'$ and $^{\alpha}C–^{\beta}C$) RDCs were compared to expected values from the crystal structure of the closest homologue ($MsrA^{Eco}$ 75% identity primary sequence). This revealed satisfactory agreement with experimental RDCs throughout the protein, both in secondary structural elements and in loop regions, excepting the continuous segment from Pro196 to Leu202, where large local inconsistency was systematically found for each individual coupling type, in both alignment media.

The disagreement could in principle be explained by differential dynamics in this region, producing time averaged RDC values that are in disagreement with a single conformational model, or by a different local conformation in $MsrA^{E.\ chmi}$ compared to the $MsrA^{E.\ coli}$ crystal structure. No evidence for rapid time-scale mobility from ^{15}N relaxation measurements was found, leading the authors to calculate the local conformation of this peptide segment using the meccano approach, to determine whether a single conformation could be found satisfying the available data and expected covalent geometry [141].

The remainder of the protein structure was assumed to be identical to $MsrA^{E.\ coli}$, as suggested from the comparative RDC analysis. Only the structure of the P196-L202 peptide was determined de novo. In this case, the two alignment tensors are known from the analysis of the secondary structural core of the molecule. Although the RDC dataset is incomplete over the whole sequence (90% of all possible RDCs), nearly all RDCs (47/48) were available for the peptide region, comprising six peptide planes and seven $^{\alpha}C$ junctions.

The length between the fragment termini P196–$^{\alpha}C$ and L202–$^{\alpha}C$ over the final ensemble (14.7 ± 0.4 Å) closely reproduces the equivalent length in $MsrA^{E.\ coli}$ (14.9 Å). In order to dock the meccano-peptide, the equivalent peptide in $MsrA^{E.\ coli}$ was removed from of the X-ray model. The procedure for positioning the fragment must however then respect the stringent orientational degrees of freedom available for the two structural domains relative to the two common alignment tensor frames. As described above, under these conditions a single relative alignment of the two domains exists, imposing an unambiguous orientation of the peptide fragment relative to the crystal structure. The translational docking must then be performed so that this relative alignment is respected, while best satisfying the expected covalence. This procedure is illustrated with respect to one of the alignment tensors using the program Module (Fig. 10a). The fragment could be directly accommodated into the crystal conformation with no serious violation of expected covalence. The ability to replace the oriented peptide in the molecular scaffold provided an independent validation of the probability that the conformation is realistic.

The local differences between the meccano conformation of $MsrA^{E.\ chmi}$ and the X-ray crystal $MsrA^{E.\ coli}$ structure in this region were found to be significant; with the $^{\alpha}C–^{\beta}C$ vector of Cys200 pointing away from the core in the X-ray structure, and into the core in the solution model (Fig. 10b). Interestingly, the more distantly related bovine enzyme ($MsrA^{bov}$, 51% identical to $MsrA^{E.\ chmi}$), closely reproduced

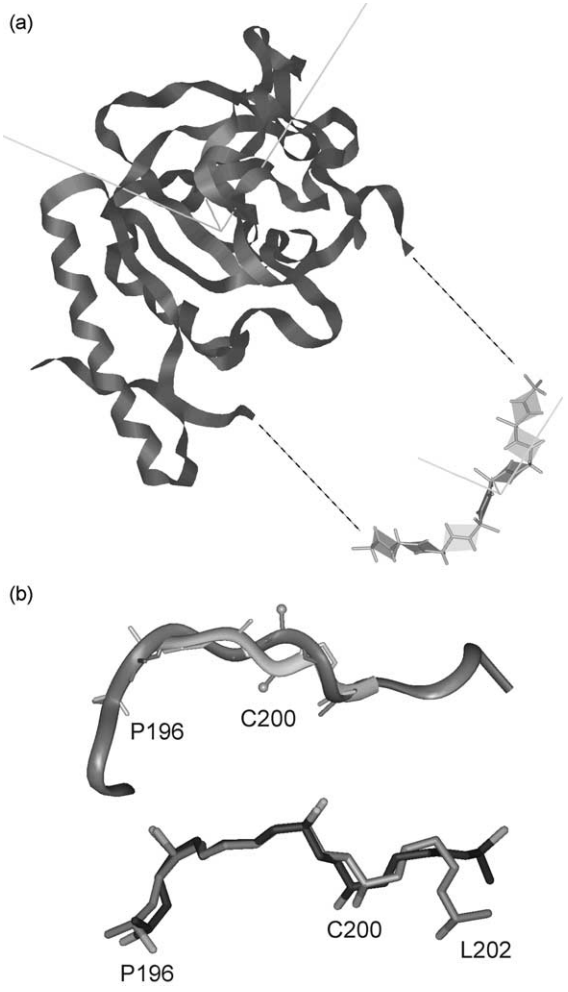

Fig. 10. (a) Representation of the positioning of the P196-L202 meccano-peptide relative to the crystal structure of MsrA$^{E.\ coli}$ using the program Module to place the fragment with respect to the alignment tensor in phage. Only transverse degrees of freedom are available in the common coordinate system. The peptide can be easily accommodated at the C′-K195 and N-G203 positions without significantly violating known covalence. (b) Top: comparison of the meccano model (light grey) and X-ray crystal structures: orientation of the Cys200 $^{\alpha}$C–$^{\beta}$C vector differs significantly in the MsrA$^{E.\ coli}$ conformation (dark grey), such that the Cys200 side-chain point into the protein interior in the MsrA$^{E.\ chmi}$ model and into the solvent in MsrA$^{E.\ coli}$. Bottom: comparison of the peptide backbone conformation in MsrA$^{E.\ chmi}$ determined using meccano and the equivalent conformation in MsrA^{bov} (dark grey).

the MsrA$^{E.\ chmi}$ conformation between residues Pro196 and Cys200 (0.73 Å backbone rmsd over this strand), further validating the model.

This study demonstrates that a combination of RDC-based techniques, initially for fold validation using a primary sequence homologue, followed by a focussed structure determination of the site of interest de novo, can provide precise local structure in localised regions of interest of highly deuterated molecules, using unambiguously assigned structural data requiring only backbone resonance assignment (no sidechain assignment was available for MsrA). Investigation time and effort are therefore greatly economized in comparison to established nOe-based structure calculation techniques.

4.4. Structure refinement using RDCs in a hybrid energy function

The definition of an experimental RDC residual term in a hybrid molecular dynamics energy function has also been popularly used in the refinement stage of structure determination protocols to drive an existing structure into a conformation in agreement with measured couplings (e.g. [60,142–144]). Nevertheless the directional ambiguity of RDCs, resulting, for example, in equivalent energetic minima at 180° rotations about each axis of the alignment tensor, makes the energetic landscape particularly complex. Despite the obvious complementarity of RDCs and ^{1}H–^{1}H nOe it is therefore not yet clear that structure calculation protocols developed for predominantly short-range restraints are the most appropriate for use with orientational restraints. Meiler et al. [145] and independently Skrynnikov and Kay [70] have provided solutions that overcome this problem by transforming measured dipolar couplings into angular restraints between vectors (these are correlated because related to the same tensor). These projection angle restraints can be numerous, because of the number of possible cross-correlations, but provide a much simpler energy landscape.

A number of structure calculation packages now contain an RDC term (for example CNS [146], XPLOR-NIH [147], DYANA [148], AMBER [143,149]), incorporating the constraint as an explicit target potential in addition to the classical potential energy function of the particular force field:

$$E_{\mathrm{RDC}} = k_{\mathrm{RDC}}(D_{ij}^{\mathrm{calc}} - D_{ij}^{\mathrm{exp}})^2/\sigma_{ij}^2. \tag{33}$$

In order to facilitate reorientation of complete structural motifs, and to penalize local violations of planarity, this force field often has higher force constants than normally used to constrain local planar or valence angle geometry.

An obvious application of RDC-driven restrained molecular dynamic calculations is the refinement of possibly low-resolution structural models derived from sequence-homology and subsequently validated using RDCs as described above. Chou et al. [150] have developed a specific approach to tackle this problem, constraining backbone torsion angles relatively close to their starting values and allowing the dipolar coupling restraints to refine the bond vector orientations during a low-temperature refinement approach. Using this approach the same authors have been able to identify helical domains adopting different average orientations in solution than in the crystal structure of Ca^{2+} ligated calmodulin [151].

As noted above, a specific problem that must normally be addressed in RDC-based structure calculation is the determination of tensor eigenvalues (A_a and A_r) and axis

orientations (defined by the Euler angles α, β, γ). These are normally unknown a priori and should be calibrated during any proposed calculation strategy. One approach that is commonly used is to repeat calculations with different combinations of A_a and A_r and to select the ensemble giving the lowest target function [152]. It has also been shown that the linearity of the alignment matrix allows the magnitude to be eliminated from the target function [153].

To illustrate the potential problems of alignment tensor determination, a brief description will be given of a study using the program Sculptor (structure calculation using long-range, paramagnetic, tensorial and orientational restraints) [60], that was specifically written to allow maximal flexibility in the development of conformational search algorithms using long-range structural restraints. Tensor parameters are treated as independent pseudo-molecules, read with the coordinates of the molecule of interest. Eigenvalues and eigenvectors are treated separately, such that one two-point molecule represents A_a, one A_r and a third, 3 point orthogonal axis system represents the orientation (α, β, γ) of the tensor, so that each term could be manipulated independently. Two independent alignment tensor functions are available that can be used separately for restraints from different alignment media, or to refine two regions independently to test the similarity of tensors for the different regions.

4.4.1. Long-range order in extended nucleic acid structures

Nucleic acids frequently form extended structures, thus their NMR-derived solution ensembles often exhibit poorly defined long-range order. This is a more serious problem than for proteins, where the density of protons is higher, and the more globular structure often supplies sufficient inter-proton nOes to define the protein fold. There is clearly a great deal of potential for exploiting the complementarity of RDCs and more local structural information [94,154]. Vermeulen et al. [155] have thus shown from simulation that RDCs should drastically improve local structure and accurately reproduce the overall helical bend in a DNA duplex. The orientational information available from RDCs in nucleic acids was illustrated experimentally by MacDonald et al. [156] who determined the structure of an A-tract DNA bend consistent with biochemical solution data but different from known crystal structures. Tjandra et al. [157] have used a large number of RDCs to determine the structure of the Dickerson dodecamer of DNA, and find a highly regular conformation with negligible bending. The importance of careful treatment of the alignment tensor parameters was also demonstrated in this study.

This aspect is also illustrated in a study of the theophylline-binding RNA, whose NMR structure was calculated using nOes and ^{3}J couplings [158]. The ensemble is shown using three different superposition modes in Fig. 11a, demonstrating the lack of overall long-range order present in the constraint set. Structure refinement using RDCs normally assumes that molecular alignment can be characterized by a single tensor and that the molecule is essentially rigid. However, if the long-range structural disorder present in an NMR ensemble actually represents the real conformational averaging present in solution, then the interpretation of the RDC data as representative of one single conformation under the influence of a common alignment tensor will no longer be valid. To address the validity of this assumption in this case, the authors analyzed distinct and distant stem, and stem-loop domains of the RNA molecule separately, so that local structure and alignment tensors experienced by each region could be independently determined. To achieve this, a restrained molecular dynamics calculation was performed using two independent tensors for these two regions of the molecule. During this calculation, that included nOe and RDC data, the tensors were allowed to float freely for the two regions simultaneously (Fig. 11b). The reasoning behind this is that local deformations in structure, due for example to non-uniform distribution of inter-proton nOes, may effect the eigenvalues extracted for the nOe-based structure. Simultaneous refinement should allow a more self-consistent solution to be attained.

Over the ensemble of initial structures both tensors converged to very similar eigenvalues. While this similarity does not allow us to exclude similar amplitude domain motion (vide infra—Section 5.1) the rest of the analysis assumes that this is not the case, and attempts to determine whether a single conformation can be found in agreement with all experimental restraints. The whole structure was therefore refined to be in agreement with a single alignment, by applying a semi-rigid-body molecular dynamics protocol in the presence of this tensor (Fig. 12). In this protocol, the local conformation in the two stem regions of the molecule were tethered to the conformations that had been refined in the first step, using artificial distance restraints. In this way, the angular coherence resulting from the first step of the calculation was retained and used to reorient the domains. The dispersion of the overall fold of the molecule falls from (3.5 ± 1.2) to (1.5 ± 0.2) Å between the initial and final ensembles, both of which were in equal agreement with the nOe data, underlining the importance of the long-range orientational restraints [159].

A conceptually similar study was performed by McCallum and Pardi [160] in which several approaches to discrete sampling of A_a and A_r were compared to identify optimal alignment parameters. The authors compared the alignment tensors for distant regions of the molecule, and again the similarity of these parameters persuaded the authors to go ahead and refine in the presence of a single tensor. Recently Puglisi et al. have determined the structure of HCV IRES domain II, a 25 kDa RNA using a combination of residual dipolar couplings, nOe and dihedral angle restraints. In this case three molecules were studied [161]. Two shorter forms of the molecule, domains IIa

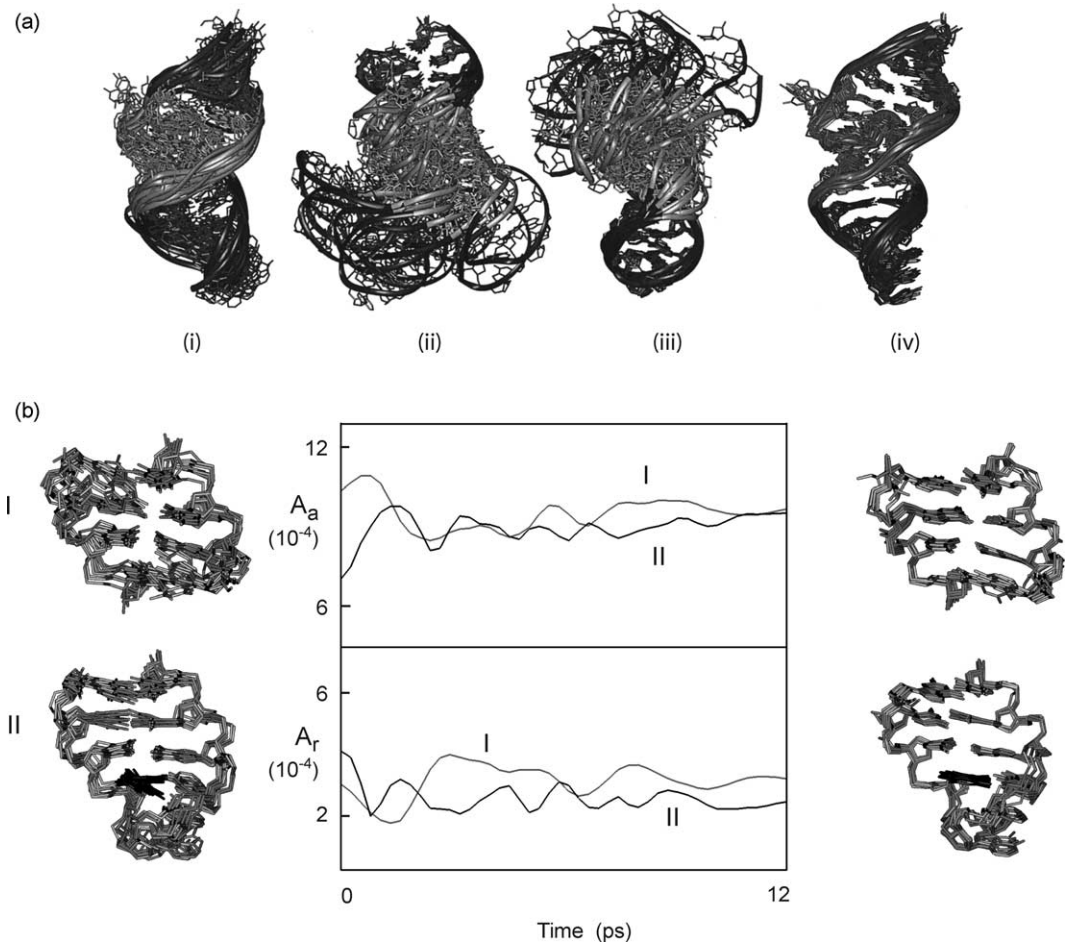

Fig. 11. (a) Calculation of the structure of the theophylline-binding RNA aptamer using ^{13}C–^{1}H residual dipolar couplings and restrained molecular dynamics. The panels (i–iii) represent the lowest target-function conformations from the nOe/J-coupling ensemble: (i) superposed using all the nucleic acids; (ii) superposed using the 3′–5′ stem I region; and (iii) superposed using the stem II—loop region. (iv) The structural ensemble represents the nOe/J-coupling/RDC ensemble superposed on all nucleic acids. (b) Sampling of the alignment tensor elements associated with stem region I and stem-loop region II. A_a and A_r and (α,β,γ) are allowed to vary independently during the initial refinement stage over the 12 ps restrained molecular dynamics. This allows optimal alignment tensors to be simultaneously determined independently, before going on, in a second step, to refine the overall long-range order with respect to a common tensor if these are similar for the two distant regions. The two structural elements are shown before (left) and after (right) the refinement procedure, (top—stem region, bottom stem II and tetraloop). Note that in addition to overall refinement of the local conformation, one of the bases no longer satisfies two orientations.

(55 nucleotides) and IIb (34 nucleotides), were designed and their conformations determined in the isolated forms. The authors then developed a hybrid calculation that used the conformational information measured for the isolated forms, including alignment tensors valid for these forms, in combination with limited RDC data available from the complete domain II, to align the two domains into an overall fold. Although there is no physical reason to retain the tensors determined for the isolated domains for the full domain, this has the advantage of constraining the local structures determined in the first steps, in a similar way to the semi-rigid-body approach of the theophylline-binding RNA described above. The authors note that on the basis of spectral characteristics measured in the different forms, the local structure is probably very similar in the isolated and complete forms of domain II, and indeed show identical chemical shifts in a segmentally labelled 60 nucleotide segment of the entire (100 kDa) HCV IRES.

4.4.2. Cross-validation of refined structural models using RDCs

The use of dipolar couplings as experimental restraints strongly compromises their utility as quality control for the resulting structure. In addition, the highly non-linear energy landscape of Eq. (33) with respect to vector orientation presents a risk of over-fitting data in this structure refinement stage. To address these issues, Clore and Garrett [162] have proposed a cross-validation approach whereby only a certain percentage of the RDC data (work data set) are used in the refinement stage and the non-used data (free data set) is reserved to validate the resulting structure. This kind of protocol reproduces data processing techniques used routinely for biomolecular structure determination by X-ray diffraction. Using this approach, the same authors demonstrated that they could find optimal force constants for use in refinement calculations of protein backbone (k_{RDC} in Eq. (33) above) for different backbone RDCs.

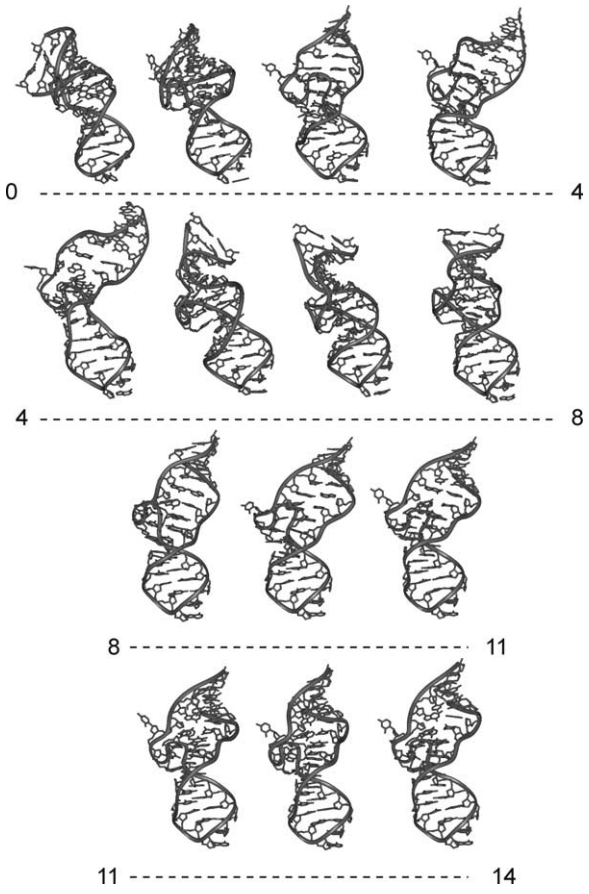

Fig. 12. Simultaneous refinement of long-range order and local structure of the theophylline-binding RNA. Coordinate snapshots from the semi-rigid-body restrained molecular dynamics. Structures were superimposed on the stem IIa region (nucleotides 11–20, the lower part of the molecule in this orientation). The numbers refer to picoseconds in the restrained molecular dynamics simulation.

4.4.3. Allowing for conformational variability in refinement: time and ensemble averaging of RDCs

As indicated in Eq. (10), RDCs depend on the population weighted average of all sampled orientations of the interaction tensor up to the time-scale of dipolar coupling (up to and beyond the millisecond range). This implies that the static structural model normally imposed on NMR-based structure calculations may not be appropriate in the presence of the kind of dynamic averaging that is expected to be present in biomolecules at room temperature [163]. Two approaches, time and ensemble averaged restraints, both previously applied to nOe-based structure determination [164–166], have recently been developed to allow for sampling of multiple conformations to fulfil the RDC restraints, by Hess and Scheek [167], and then by Clore and Schwieters [168]. Rather than requiring that a single conformation is in agreement with the constraint presented in Eq. (33), the average of conformations sampled either during a restrained molecular dynamics simulation or over n copies of the same molecule are used to constrain the sampled conformations. Eq. (33) then becomes

$$E_{\mathrm{RDC}} = k_{\mathrm{RDC}}[(\langle D_n^{\mathrm{calc}}\rangle - D_n^{\mathrm{exp}})^2/\sigma_{ij}^2] \tag{34}$$

with $\langle D_n^{\mathrm{calc}}\rangle$ representing the time or ensemble average RDC over all sampled conformations, where necessary including the possibility that the different conformations are associated with different tensors. This approach clearly presents advantages over structure calculations that impose a static model, although it should be remembered that the averaging of RDCs, giving rise to a single averaged value, differs fundamentally from the averaging effects seen for example when distinct sets of cross-relaxation rates are measured between different pairs of spins due to jumps between different conformations. It may be that a conformational equilibrium between two distinct states, giving rise to two sets of contradictory nOe peaks, is more appropriate to this kind of multiple copy analysis than the same conformers producing a single set of averaged parameters, where unambiguous identification of the participating conformers may be more difficult. In their study, Clore and Schwieters have also suggested that ensemble averaging can be used as a diagnostic tool to study the extent of internal mobility in proteins. The results of this study are presented in more detail in Section 5.3.

5. Using RDCs for the study of molecular dynamics of proteins and nucleic acids

Since the earliest work on weak alignment of proteins, numerous approaches have been developed to extract information on domain dynamics in biomolecules. In an early study of cyanometmyoglobin, a haemoprotein whose anisotropic magnetic susceptibility leads to significant alignment in the magnetic field, dipolar contributions to measured couplings were extracted from data measured at three static field strengths [169]. These were found to disagree with expected values calculated from the magnetic susceptibility tensor predicted from the structure. The authors found that these departures from predicted values could be significantly reduced if collective motions in some helical regions of the protein were allowed. An optimisation procedure using either a diffusive motional model or a simple reorientation about an axis allowed the parameterisation of best-fitting amplitudes for these motions to account for the difference in calculated and experimental dipolar contributions. While the authors cannot be sure that dynamics are the true cause of the observed discrepancy [170], they nevertheless demonstrated the enormous potential of dipolar couplings to characterise the geometric nature of slower motions in proteins; information that had until then been very difficult to access.

5.1. Domain dynamics in biomolecules

As mentioned above, differential domain dynamics can be inferred when two structurally distinct regions of the molecule or complex exhibit different alignment characteristics, specifically with respect to the parameters A_a and A_r describing the extent and anisotropy of the preferential alignment of each domain. An early example again came from the Prestegard group [171], in a study of the two-domain Barley Lectin protein, in which very different values of A_a and A_r were evident from the two domains. In this case, the authors concluded that one of the two domains (B) was interacting partially with the CTAB-doped bicelles, while the other (domain C) was undergoing less restricted dynamics with respect to this domain. The amplitude of the dynamic sampling of domain C relative to domain B could again be estimated, by modelling the difference in the alignment tensor amplitudes via a simple diffusion-in-a-cone model, to be of the order of 40°.

It is clear that the observation of distinct alignment properties in different domains of a macromolecular complex should be interpreted with some caution. In a detailed structural characterisation, Braddock et al. [172] have recently determined the solution conformation of two K-homology domains, KH3 and KH4, of the FUSE-binding protein (FBP) in complex with single stranded DNA. Although the two domains are of equal size and are joined by a highly flexible 30-amino acid linker, they exhibit highly contrasting alignment characteristics when dissolved in dilute phage fd, the level of overall alignment differing by a factor of two. In this case, the authors evoke the possibility that the very different alignment characteristics derive from the different electrostatic charge distribution in the two parts of the complex (the associated DNA strands effectively differing by four negative charges). This assumes that the two domains have sufficient orientational independence to allow the surface charge distribution of each domain to dominate their individual alignment characteristics [173].

A number of more recent studies have also postulated significant domain dynamics in multi-domain systems based on the detection of differential apparent alignment forces in distinct structural domains. Jacobs et al. [174] found levels of alignment differing by a factor of nearly four in the C-terminal catalytic and the N-terminal WW domains when aligned in both Pf1 phage and a lyotropic alcohol mixture. The two domains are very different in size, and are linked by a highly flexible 12-amino acid peptide. The smaller WW domain (less than 30% of the mass and the surface) is considered to act as a flexible tail of the C-terminal domain that is three-times bigger. In the presence of a model peptide mimicking the substrate that apparently simultaneously binds both C and N-terminal domains, the complex becomes much more rigid, and the level of alignment of the WW domain increases by a factor of 3–4 in the two alignment media.

Small amplitude reorientational dynamics of molecular domains can also be studied in some detail using RDC, essentially due to the precision available if many couplings can be measured in each domain. This has been illustrated in the study of maltodextrin binding protein referred to in Section 3, where Skrynnikov and co-workers [46] also compared static and dynamic models of the relative orientation of the two domains of the protein. In this case, motional models allowing exchange between two conformers derived from crystallographic models, differential dynamics of one of the two domains relative to the other or a distribution of relative orientations implying similar amplitude excursions for the two domains were compared. While none of these scenarios reproduces the data significantly better than a rigid-body approach, the authors do note that in the latter case, shared motion of this kind can be very difficult to detect due to the similar averaging effects on the two domains, which give rise to a modified, but common, alignment tensor. Finally, a series of site-directed mutants in the linker region between the two domains were engineered, and RDCs measured in each of these aligned in phage solution [175]. The different conformations were found to be related by a simple rotation about the invariant closure axis, and the free energy of the conformations, measured by fluorescence spectroscopy, was found to correlate linearly with this angle. The authors have thus combined protein engineering and state-of-the-art partial alignment technology to map the pathway between the open and closed state of this enzyme.

A similar approach was applied by the same group to the study of inter-domain dynamics in the two-domain T4 lysozyme using RDC data obtained from different alignment media. In this case, a two state exchange was suggested between conformers representing the ‘open’ (thought to be required for ligand binding) and the ‘closed’ forms of the enzyme [92]. Calculations of the expected sterically aligned tensor based on the overall molecular shape were compared to the effective averaged tensors in order to determine optimal populations of the two exchanging conformers. The authors again note that one cannot distinguish between the dynamic exchange model and an average static model on the basis of statistical analysis of the available data, underlining the need to exercise caution when evoking this kind of inter-domain motion from RDC data.

More recently, Ho and Bax and co-workers [176] have studied the quaternary conformation of the tetrameric human carbonmonoxy-haemoglobin in solution using RDCs measured in two different alignment systems. Comparisons with known crystal conformations allow the authors to propose a dynamic exchange equilibrium in solution between the quaternary structure present in two distinct states. This study again underlines the complementarity of RDC-based analysis in combination with high resolution X-ray crystallography, to characterise conformational transitions in large proteins.

Bertini et al. [177] combined pseudo-contact paramagnetic chemical shifts and residual dipolar coupling data in a study of the amplitude of inter-domain motions in lanthanide bound calmodulin. These sources of dynamic information are highly complementary due to the dependence of the dipolar shifts on the distance between the observed spin and the paramagnetic centre. As expected for this two-domain system, no single conformation can explain the observables, and the authors proceed to place limits on the amplitude of orientational sampling experienced by the two domains, finding that higher amplitudes cones (semi-angle greater than 40°) could be spanned than have previously been suggested from relaxation data [178], reporting on motions with pico-nanosecond characteristic time constants. The authors are also able to comment on the relative populations of known conformational states and on the possible modes of sampling the accessible cones.

The difficulties that can be encountered when probing motions of similar-sized domains are illustrated in the study, by Al-Hashimi, Patel and co-workers who investigated the dynamic characteristics of the transactivation response element (TAR) RNA from human immunodeficiency virus HIV-1 [179]. Similar to the RNA aptamer discussed above, this molecule can be 'dissected' into two helical regions, one capped with a tetraloop (stem II), joined to stem I by a bulge region having no identifiable secondary structure, in this case comprising 3 nucleotides on one side of the strand. Differential alignment tensors from the two stem regions (assumed to adopt canonical A-form helices) were detected from C–H and C–C RDC values measured in phage-aligned RNA. The alignment tensor of stem II was found to have an amplitude more than twice that of stem I. The authors showed that these differences could be interpreted either in terms of a simple motional model based on diffusion of stem I, in a cone of half-amplitude 45° relative to the other, non-diffusing domain (stem II and tetraloop), or in terms of an increased reorientation of stem I relative to a more restricted stem II. This model results in larger angle motions of stem II. In a further study, Al-Hashimi and co-workers [180] dispensed with the alignment medium and studied the natural alignment properties of the same molecule at three magnetic field strengths, and compared these to the calculated magnetic susceptibility tensor derived from summed base-group susceptibilities. In this case, the experimentally extracted alignment tensors from the two stem domains are found to be similar, and in agreement with predicted susceptibility assuming independent domains, but in strong disagreement with predicted susceptibility values for a single static conformation of the whole molecule. This is interpreted as further evidence for large-scale (half-angle around 90°) domain motions of both stem regions around an average conformation represented by the structure that was previously refined with respect to phage-aligned RDC. This research is ongoing, and a further contribution from these authors has recently demonstrated, using the same methodology, that on binding to HIV-1 TAR RNA, argininamide inhibits these global motions essentially completely, stabilising the relative conformation of the two domains into a single static envelope [181].

Quaternary geometry and domain dynamics in proteins can also be studied using heteronuclear relaxation rates measured in the different domains [182]. In combination with RDC-derived orientational information, this can be highly informative due to the different time-scales relevant for relaxation and the averaging of RDCs. For example, Ulmer et al. [183] have used RDC and ^{15}N relaxation, combined with chemical shift perturbation, to characterise domain orientation in the related SH(32) fragment from the Src family member Fyn. In this case, although the two domains of the peptide-bound FynSH32 were found to be coupled, they still exhibit inter-domain flexibility on both nanosecond and longer time-scales. In a recent study, Fushman and co-workers [184,185] have analysed the structural and dynamic properties of poly-Ubiquitin chains in solution using residual dipolar couplings and ^{15}N relaxation. In particular relaxation data collected on di-Ubiquitin at neutral and acidic pH, and RDC data collected at neutral pH, point to a large amplitude ($>110°$ rotation) conformational switch when going from neutral to acidic conditions. In this case, the relative orientation of the two domains determined assuming common rotational diffusion properties from relaxation and common alignment properties from RDC were found to be slightly different, suggesting that differential slower time-scale dynamics that are not present on the fast time-scale to which relaxation is sensitive may provide additional averaging mechanisms for the measured RDCs. In contrast, a ^{15}N relaxation study of MBP, a more compact two-domain protein, closely confirmed the relative orientation determined from the RDCs, suggesting that additional slower time-scale motional averaging modes are not significant in this case [186]. It has also been noted that in the case of steric alignment, dependent on the shape of the molecule, one would expect a strong correlation between R_2/R_1 (dependent on the rotational diffusion tensor, in turn dependent on the shape of the molecule) and measured RDCs. This comparison was suggested as a direct means of detecting chemical shift exchange contributions to R_2 [187].

5.2. RDCs for the study of local backbone dynamics in proteins

Rapid backbone dynamics in proteins are now routinely studied using ^{15}N and ^{13}C spin relaxation [29–36] Because of the overall tumbling of the protein, that essentially reduces the angular correlation function to zero within a small number of molecular rotations, spin relaxation is no longer sensitive to time-scales beyond this limit (5–10 ns for proteins in aqueous solution). Measurement of the averages present in residual dipolar couplings however provides direct access to key information for understanding protein motions in the sub-micro to milli-second range, that is

highly complementary to the shorter time-scale dynamics derived from spin relaxation. This is particularly evident for example for ^{15}N–^{1}H couplings, whose rapid reorientation also dominates experimental spin relaxation rates, where comparison of the dynamic characteristics exhibited over the two time-scales can in principle deliver invaluable information on the slower motional sampling of the same interatomic vector. Perhaps not surprisingly, this prospect has excited considerable recent interest in the field. In view of this intense activity, this section will be presented in more detail than other topics in this review.

5.2.1. Local alignment tensor analysis and generalised order parameters

A number of analogous approaches to those described above for the elucidation of domain motions have been applied to the study of local dynamics in proteins, by determining the differential effective alignment properties of local structural elements. The problem here is again similar to that encountered in structural analysis of RDCs: that is the identification of structural motifs whose conformation is sufficiently well-defined to allow the order parameter to be determined and whose local geometry, even when experiencing motions relative to the molecular frame, can be considered unchanging over the averaging time-scale.

As we have seen above the most obvious integral structural unit present on the main chain of a protein that meets these criteria is the peptide plane. Tolman et al. [188] addressed the question of the determination of local dynamics by considering not only the peptide plane, but extending the fragment to include the three atoms attached to the second $^{\alpha}$C atom. This allows the inclusion of $^{\alpha}$C–C′, $^{\alpha}$C–$^{\beta}$C and $^{\alpha}$C–$^{\alpha}$H RDCs, giving a maximum of eight RDC values for each structural unit, thereby allowing determination of the five parameters defining the local alignment characteristics from these measurements, in combination with the relative conformation of the eight inter-nuclear vectors. As the authors point out, confident modelling of the conformation of the local fragment requires quite precise knowledge of the ϕ angle connecting the two planes to determine the orientation of the tetrahedral geometry relative to the peptide plane. Despite this limitation, local alignment tensors could be extracted along the whole peptide chain of the protein Ubiquitin, and compared, at least qualitatively, to rapid motional amplitudes determined at N–^{N}H and $^{\alpha}$C–$^{\alpha}$H sites from spin relaxation measurements. In this analysis, the authors have introduced a parameter that reflects the absolute degree of order of the fragment of interest, the generalised degree of order (GDO):

$$\vartheta = \sqrt{\frac{2}{3}\sum_{ij} A_{ij}^2} \tag{35}$$

This parameter is essentially the Euclidean norm of the five-dimensional vector that can alternatively be used to express the traceless Saupe matrix **A**, in terms of averages of combinations of Wigner matrices D^2_{i0} ($i = -2, -1, 0, 1, 2$), and, at least in the case of axially symmetric motion, is independent of the mean orientation of the fragment of interest relative to the field. The authors note that in the case of an axially symmetric alignment tensor ϑ is simply given by A_{zz} and that local motional averaging can in favourable cases be separated from overall alignment averaging if one defines a fragment specific internal GDO, ϑ_{int} and considers the ratio of this relative to the global parameter, $\vartheta_{\mathrm{int}}/\vartheta_{\mathrm{global}}$, as a measure of local dynamics.

A careful simulation of expected behaviour in the presence of axially symmetric and anisotropic motional averaging was performed in this study, as well as a detailed consideration of the effects of structural error on the extracted parameters. In particular, it was noted that in the presence of anisotropic motional averaging ϑ reflects a combination of directionality and motional amplitude. Finally, the authors also point the way to future studies, by suggesting that a way to more confident characterisation of dynamic modes could come from measurements in more than one alignment medium. The Prestegard group [189] have also applied the GDO approach to conformational analysis of a flexible oligosaccharide (again a trimannoside). In this case $^{1}D_{\mathrm{CH}}$, $^{2}D_{\mathrm{CH}}$ and D_{HH} couplings were measured in each ring of the molecule in both bicelle and phage media, and the resulting general order parameters could be interpreted in terms of different degrees of motion about the linkages between the central ring (III) and the two peripheral rings attached to this (I and II).

Wang et al. [190] have also looked at local dynamics along the peptide chain by comparing alignment characteristics of structural units, but in this case 5-amino acid strands were used to fit five N–^{N}H RDC values. Reduced eigenvalues for the local alignment tensors determined within the sliding window were interpreted as evidence for dynamic averaging effects.

The use of the GDO approach for detecting motional averaging in local units of proteins where multiple RDC values can be measured is a convenient, and relatively robust method of expressing local mobility in terms of a single parameter. The method is model-free, in that no specific geometric mode is supposed, although as noted above different averaging effects can be extracted if equal amplitude anisotropic motions are present at different mean orientations in the protein.

5.2.2. Model-free approaches to the characterisation of the dynamics of inter-nuclear bond vectors

A number of recent contributions to the study of conformational averaging using residual dipolar couplings have also used ‘model-free’ approaches to characterize motional properties, this time of individual inter-spin vectors [47,191–193]. As shown in Eq. (14), the direction and amplitude of angular sampling of the inter-nuclear vector can be described purely in terms of average second

order spherical harmonics $\langle Y_{2,M'}\rangle$. Clearly if these terms could be determined for a single site, it would be possible to describe both the average structure, and the motional averaging for each individual vector. The drawback here is that one must have access to at least five independent sources of data for each spin-pair in order to determine the five quantities $\langle Y_{2M'}\rangle$, $\{M' = -2, -1, 0, 1, 2\}$. Data from the different alignment media can be related using the following relationship, derived from Eq. (14)

$$\begin{aligned}\langle D_{ij}^{(i)}(\theta,\phi)\rangle = &-\sqrt{\frac{16\pi}{5}}\frac{\gamma_i\gamma_j\mu_0 h}{16\pi^3 r_{ij,\mathrm{eff}}^3}\\ &\times\left[A_{\mathrm{a}}^{(i)}\left(\sum_{M'=-2}^{+2}\mathrm{e}^{-\mathrm{i}\alpha^{(i)}M'}d_{M'0}^{(2)}(\beta^{(i)})\langle Y_{2,M'}\rangle\right)\right.\\ &+\sqrt{\frac{3}{8}}A_{\mathrm{r}}^{(i)}\left(\sum_{M'=-2}^{+2}\mathrm{e}^{-\mathrm{i}\alpha^{(i)}M'}d_{M'2}^{(2)}(\beta^{(i)})\mathrm{e}^{-\mathrm{i}2\gamma^{(i)}}\langle Y_{2,M'}\rangle\right.\\ &\left.\left.+\sum_{M'=-2}^{+2}\mathrm{e}^{-\mathrm{i}\alpha^{(i)}M'}d_{M'-2}^{(2)}(\beta^{(i)})\mathrm{e}^{+\mathrm{i}2\gamma^{(i)}}\langle Y_{2,M'}\rangle\right)\right]\end{aligned} \tag{36}$$

where the superscript (i) indicates data from different alignment media. The principal axis systems of the alignment tensors, with amplitude and rhombicity $(A_{\mathrm{a}}^{(i)}, A_{\mathrm{r}}^{(i)})$, are related by three-dimensional rotations $(\alpha^{(i)}, \beta^{(i)}, \gamma^{(i)})$ applied here as Wigner rotation matrices, such that the invariant averaged spherical harmonics are conserved and can therefore be determined in a given alignment frame, as long as five or more sufficiently independent datasets are available.

Brüschweiler and Griesinger and co-workers [47] have followed this approach, initially applied to simulated $^1D_{\mathrm{NH}}$ data that were extracted by collecting all of the orientations sampled during a 10 ns molecular dynamics trajectory of solvated Ubiquitin and transforming these into different RDCs by imposing 11 different alignment tensors. These simulated data were used to extract the averaged spherical harmonics by expressing N (the number of alignment media) versions of Eq. (36) as a series of simultaneous equations and solving for $\langle Y_{2,M'}\rangle$. It is clear from Eq. (36), that both $(A_{\mathrm{a}}^{(i)}, A_{\mathrm{r}}^{(i)})$ and $(\alpha^{(i)}, \beta^{(i)}, \gamma^{(i)})$ must be accurately known if the extracted dynamically averaged parameters are to have any physical meaning. The authors recognise this and initially investigate the effects of motional averaging on interpretation of RDCs as static parameters in the case of three possible scenarios, and then apply their approach in the context of the following scenarios. (I) This assumes that the 'true' alignment tensors are known from independent sources, for example, analytical calculations of alignment or knowledge of the magnetic susceptibility. Not surprisingly this results in identification of the most severe averaging effects when a static model is used. (II) The motionally averaged data are used to determine the best-fitting alignment tensor to the average structure. In comparison to scenario (I), some of the true motion is found to be absorbed in this process. This is because the values of $(A_{\mathrm{a}}^{(i)}, A_{\mathrm{r}}^{(i)})$ are already scaled (by up to 0.89), and the values of $\langle Y_{2,M'}\rangle$ do not therefore need to be scaled in order to reproduce the dipolar couplings. Scenario (III) attempts to take local motional averaging into account, by scaling the simulated RDC by its associated order parameter S_{LS} where the subscript denotes the equivalence of this parameter to that routinely extracted from spin-relaxation data using the Lipari–Szabo model-free analysis (this would be the source of this parameter experimentally). In this case, the correct values for $(A_{\mathrm{a}}^{(i)}, A_{\mathrm{r}}^{(i)})$ are reproduced, the extracted local orientation of the average vector amplitude agrees very closely with the average calculated directly from the trajectory (generally within 2°), and the local motional amplitudes and anisotropies are apparently correctly extracted. Notably the authors demonstrate that scenario (II), where the tensors are determined without accounting for local motional averaging using S_{LS} results in real difficulties in finding an average vector orientation in agreement with all datasets, but that for well-behaved sites a common solution can be found if a common scaling factor affecting data from all alignment media, and therefore describing axially symmetric motion, is taken into consideration. The consequences of this for structure calculations are considered.

In a second article [191], this approach was applied to experimental data collected from 11 different alignment media, again for the protein Ubiquitin. The list of alignment media used to align this protein is immediately impressive, and although it has been recently pointed out that these tend to fall into two related families [168], there remains a high degree of independent orientational information present in the entire data set. Much emphasis in this study is again given to the determination of the alignment tensor that is not influenced by internal motional averaging. The authors essentially use scenario (II) introduced above, but then correct extracted motional amplitudes via a universal scaling parameter S_{overall}, that is retrospectively determined by the condition that the generalised order parameter derived from the data

$$S_{\mathrm{rdc}}^2 = \frac{4\pi}{5}\sum_{M=-2}^{2}\langle Y_{2M}(\theta,\phi)\rangle\langle Y_{2M}^*(\theta,\phi)\rangle \tag{37}$$

should not exceed 1 (then $S_{\mathrm{overall}} = 1/(S_{\mathrm{rdc}})^{\mathrm{max}}$. For the uninitiated, the five averaged spherical harmonics do not lend themselves to an intuitively obvious interpretation in terms of the nature and amplitude of local structure and dynamics. For this reason, the authors have made an additional effort to transform this information into more tangible attributes, namely $\{\theta_{\mathrm{eff}}, \phi_{\mathrm{eff}}\}$ the effective or mean orientation of the vector, S_{axial} reflecting the extent of axially symmetric motion

$$S_{\mathrm{axial}} = \sqrt{\frac{4\pi}{5}}\langle Y_{2,0}(\theta',\phi')\rangle \tag{38}$$

η the extent of non-axially symmetric motion (the anisotropy of the motion)

$$\eta = \sqrt{\sum_{M=-2,2} \langle Y_{2M}(\theta,\phi)\rangle\langle Y_{2-M}(\theta,\phi)\rangle / \sum_{M=-2}^{2} \langle Y_{2M}(\theta,\phi)\rangle\langle Y_{2-M}(\theta,\phi)\rangle} \tag{39}$$

and $\bar{\phi}'$ the direction of this anisotropy in the local peptide frame. A number of related approaches to extracting these parameters from simultaneous analysis of all 11 data sets are presented, and their distribution along the peptide chain is discussed in detail. The extracted effective or mean orientations of the N–^{N}H vectors very closely reproduce those present in the RDC-refined NMR structure, demonstrating both the accuracy and the precision of the approach. $S_{\rm rdc}$ has an average of 0.78, in comparison to the relaxation derived order parameter $S_{\rm LS}$ of 0.89, while the axial component $S_{\rm axial}$, determined via a separate analysis applying a local scaling factor to the raw spherical harmonics, is found to have an average of 0.85, and is directly compared to GDO values determined by Tolman et al. (vide supra). In contrast to the previous study based on MD simulation, the anisotropic contribution to the movement is found to be much more significant, possibly due to the fact that the range of sampled motional time-scales is increased by 6–7 orders of magnitude for the experimental data compared to the simulation. Interestingly, however, the direction of the anisotropic motions in secondary structural elements correlates quite well with those extracted from the MD simulation. The exact amount of motion derived using this method remains susceptible to scaling artefacts introduced in the estimation of $S_{\rm overall}$ (parameterised at 0.79 in this study), and $S_{\rm axial}$, the consequences of which appear to be that quantitative analysis of dynamic amplitudes may be difficult. Nevertheless approximately twice as much motion was found to be present than has been observed on the relaxation-derived time-scales ($\leq$ns). Recently, the same authors, wary of the comparatively high amplitude motions extracted from this analysis, both for axially symmetric and anisotropic components, have extended their experimental dataset to include 30 different alignment conditions for the protein Ubiquitin. The increase in definition of the extracted motionally averaged spherical harmonics results in somewhat lower amplitude motions, and concomitantly higher order parameters [194].

It is possibly worth noting that while the structural information derived from the analysis is precise, with $\{\theta_{\rm eff}, \phi_{\rm eff}\}$ accurately reproducing known $\{\theta,\phi\}_i$ values from NMR and crystal structures, in its current form the method requires an already existent three-dimensional structure, presumably of quite high resolution, in order to determine the alignment tensors at the beginning of the fitting procedure. One might speculate whether more precise dynamic information could be extracted if $\{\theta_{\rm eff}, \phi_{\rm eff}\}$ were assumed to be known accurately from the available structure throughout the analysis, thereby reducing the complexity of the optimization surface. We will encounter a geometric-model-dependent approach that follows this logic to some extent (Section 5.2.5).

Interestingly, although the authors identify the α-helix in Ubiquitin as having high $S_{\rm axial}$ and relatively low η values, the direction of this anisotropy does not vary significantly along the helix, and in a further study they are able to model all of this motion as a single coherent reorientation of the helix with half-angle amplitudes of 10 and 20° about two axes orthogonal to the main helical axis [195].

Tolman has independently developed a conceptually similar, but analytically quite different approach to the retrieval of structural and dynamic information from RDCs measured in several alignment media [192]. In this case, the simultaneous equations represented in Eq. (36) are replaced by a single matrix relationship

$$\mathbf{D} = -\frac{\gamma_i\gamma_j\mu_0 h}{8\pi^3 r_{jk}^3}\mathbf{BA} \tag{40}$$

where $\mathbf{D}$ is a matrix comprising the RDC measurements, with dimensions $M\times N$, with M the number of different media, and N the number of different spin-pairs. The matrix $\mathbf{A}$ comprises the information describing the alignment tensors of the molecule in the different media and is dimensioned $5\times M$. Each column contains five elements taking the form

$$A_{mn}(j) = \frac{3}{2}\langle \cos\xi_m \cos\xi_n\rangle_j - \frac{1}{2}\delta_{mn} \tag{41}$$

where the subscript j refers to the alignment medium. Matrix $\mathbf{B}$ containing the desired information concerning the structure and dynamics of the interaction vector. This is also defined in terms of a traceless second rank Cartesian tensor, in analogy to the order tensor formalism, in this case the rows contain five terms pertaining to each interaction vector (i):

$$B_{mn}(i) = \frac{3}{2}\langle \cos\zeta_m^i \cos\zeta_n^i\rangle - \frac{1}{2}\delta_{mn} \tag{42}$$

As we saw earlier in Eq. (3) (ξ_x, ξ_y and ξ_z) define the orientation of the alignment frame relative to the magnetic field and (ζ_x, ζ_y and ζ_z) describe the orientation of the internuclear vector with respect to the axes of the macromolecular frame. The five elements $B_{mn}(i)$ code identical information to that stored in the five averaged spherical harmonics in the Griesinger/Brüschweiler approach described above, and can be interpreted in terms of a local order tensor, whose principal axis system is oriented according to a three-dimensional Euler rotation (θ_i, ϕ_i, ψ_i). The five elements of the tensor then encode the direction of the mean vector $\{\theta_i, \phi_i\}$ from the orientational information and the amplitude and asymmetry of any motion from the principal values. Finally, the directionality of the anisotropic motion is defined by the third angle $\{\psi_i\}$. In order to extract this information from matrix $\mathbf{B}$ by solving Eq. (40), some definition of matrix $\mathbf{A}$ is required, as noted by

the author, this will not generally be available. This is the identical problem encountered in the previous study—that the alignment tensors must be characterised before the local structure and dynamics can be investigated, and that these alignment tensors should be unaffected by local dynamic motions. To overcome this problem, the author then demonstrates, in an analysis called DIDC (direct interpretation of dipolar couplings) using logic similar to that presented by Moltke and Grzesiek [153] that the information contained in matrix **A** can be eliminated from the analysis completely in the presence of complete RDC data (matrix **D** is rank 5). This relies on the observation that under these conditions one can express **B** almost completely in terms of the data in matrix **D**, such that

$$\mathbf{B} = \mathbf{U}_D\mathbf{\Lambda} \tag{43}$$

$\mathbf{U}_D$ is determined from singular value decomposition of **D** and the 5×5 matrix $\mathbf{\Lambda}$, apparently containing all the information that cannot be determined from RDC data alone, is estimated by minimising the variation between the effective resulting generalized order parameters. The author demonstrates the method using simulations performed using complete data sets, indicating that the extraction of a generalised order parameter is much more robust than the two parameters describing the anisotropic amplitude and direction, which are apparently generally less reliably characterised. A more practical two-step procedure is then introduced for use with incomplete data sets, including a refinement-step to create a best-fitting, structurally refined model, or set of oriented vectors, before using this as a fixed model with respect to which generalised order parameters are extracted. The author thus demonstrates that local order parameters can be determined in the presence of as few as three different alignment data sets assuming that structural coordinates are available, and applies this approach to Ubiquitin. The results of this initial study indicate higher levels of site-to-site variation in motional amplitudes than reflected by ^{15}N relaxation.

A subsequent application of the complete DIDC approach was later presented using data taken from the multiple sets of RDC utilised in the previous study of Ubiquitin by Meiler et al., thereby removing the dependence on an initial structural model. In this case, the generalised order parameters present lower residue to residue dispersion than in either the Meiler et al. approach or the previous application of this method, although again the order parameters were scaled to be equal to, or lower than the order parameters derived from ^{15}N relaxation [193]. The author notes that the similar profile of the order parameters derived by the two methods does, however, provide some justification for using the minimization procedure applied to the matrix $\mathbf{\Lambda}$ that necessarily reduces site-to-site variation.

One identifiable drawback to the application of approaches requiring the measurement of dipolar couplings in multiple alignment media is the increased potential for interaction between the protein and one of the alignment media surfaces. This problem has been addressed by Hus et al. [196,197], who have developed a self-consistency analysis approach, based on a principal component analysis applied to a weighted covariance matrix C, of RDCs measured in M alignment systems:

$$C_{ij} = \frac{1}{M-1}\sum_{k=1}^{M}(D_i^{(k)} - \langle D_i^{(k)}\rangle)(D_j^{(k)} - \langle D_j^{(k)}\rangle)/\sigma_k^2 \tag{44}$$

The five largest eigenmodes then encode the direction of each vector (ij), and the appearance of additional eigenmodes provides evidence for inconsistencies, due either to structural differences in the presence of the different alignment media, or potentially due to dynamics. Application of this approach to 10 of the data sets used to determine local dynamic amplitudes by Meiler et al. reveals that the secondary structural elements in Ubiquitin behave essentially homogeneously over the different alignment media.

5.2.3. Position and dynamics of amide protons from multiple RDC data sets

The extent of backbone dynamics present over the longer time-scales relevant for averaging of residual dipolar couplings can also be addressed by investigating the possibility of interpreting the data in terms of a single conformation, and evoking the presence of dynamic interconversion between conformers only when this approach fails. In this respect, the recent presentation by Ulmer et al. of the detailed analysis of multiple different RDCs in five differently aligning media contributes important information regarding the study of local dynamics using RDCs [198]. The aim of this study was to use an extensive data set (up to 20 couplings per peptide unit, comprising $^1D_{\mathrm{NH}}$, $^1D_{\mathrm{C\alpha H\alpha}}$, $^1D_{\mathrm{C'C\alpha}}$ and $^1D_{\mathrm{C'N}}$ couplings) to investigate local structure as precisely as possible, including the possible presence of site specific variations of the peptide plane conformation, normally assumed to be uniform in proteins. The authors begin their study by comparing these measured couplings to an existing high-resolution crystal structure of the same protein [199]. Hydrogen electron density is normally too weak for the position of these atoms to be accurately derived from X-ray crystallographic structures of proteins and it is therefore necessary to build hydrogens onto the crystal structure assuming known local geometry. As the authors were primarily interested in the local conformational geometry of the amide proton, they initially compared only the RDC involving heavy atoms ($^1D_{\mathrm{C'C\alpha}}$ and $^1D_{\mathrm{CN}}$). While these couplings compared well with those expected from the structure when using SVD analysis, the residual between calculated and experimental RDC was still found to exceed levels expected on the basis of experimental error. Postulating that these local disagreements may be due to differences between the solid and liquid state conformations, the authors then used a restrained molecular

dynamics calculation to refine the crystal structure with respect to these couplings. This procedure resulted in very close agreement between measured and predicted RDCs from 46 of the 55 amino acids. The authors then demonstrate that, within the estimated experimental error, all $^1D_{NH}$ (and separately all $^1D_{C\alpha H\alpha}$) couplings measured in the 46 retained sites can be in agreement with a single vector orientation in the presence of these five tensors.

Ulmer et al. proceed further to precisely investigate the local configuration of the peptide plane throughout the protein, in particular they are able to parameterise the direction of the N–^{N}H vector in the peptide plane, and observe evidence for slight pyramidalisation (such that the $^\alpha$C–C′–N–^{N}H dihedral angle is adjusted to $-2°$). Finally, the authors investigate whether anisotropic peptide plane motions around the $^\alpha C_{i-1} - {}^\alpha C_i$ axis, that have been previously identified from molecular dynamics simulations and spin relaxation measurements [51–53], can be accommodated by the $^1D_{NH}$ values when compared to the structure refined using the remaining couplings. For the retained sites, the results are found to be consistent with the presence of small amplitude fluctuations of this geometric mode having 10–15° higher amplitude than in-peptide-plane fluctuations.

5.2.4. Ensemble averaging and dynamic motions in protein backbones

Clore and Schwieters [168] have proposed a related approach to characterising the degree of backbone motions present in proteins and their relevance for the averaging of RDCs, again attempting to find conformational sampling conditions that are in agreement with all measured couplings simultaneously As described in Section 4.4.3, a more elaborate restrained molecular dynamics refinement scheme is devised in this case, specifically ensemble averaging, in which each RDC restraint is calculated as the average over all N conformers, thereby allowing the possible presence of interconverting structures if required. One obvious advantage of this approach is that when different conformations are present in the ensemble, the simulated structures immediately deliver a molecular model of a dynamic mode to account for a component of the motional averaging of RDCs.

Ensembles with $N=1,2,4$ and 8 conformers were compared, and cross-validation used to assess the accuracy of the observations. The results demonstrate that refinement against all 11 N–^{N}H data sets can result in a single conformation that is in agreement with the experimental data, at least within the limits of the estimated experimental error. This does not of course preclude the existence of local dynamics, but possibly more interestingly, the authors also show that allowing more conformations to contribute to the data does locally improve the reproduction of the experimental values. The authors then measure the difference between the vector orientations (θ) in the interconverting conformers, and interpret these amplitudes in terms of two (or more) site jumps between vectors, calculating order parameters for this anisotropic motion on the basis of Eq. (29). In a cross-validation analysis, pairs of conformations are found to provide slightly better global agreement with a set of unused $^\alpha$C–$^\alpha$H couplings. Order parameters are found to be close to unity for most sites, although for a few residues (six of the sites considered for $N=2$ averaging), jumps of over 40° are still required to optimally fit the data. It may be relevant to note the use of an additional restraining force on the $^\alpha$C atomic positions of each partner of the $N=2$ ensembles; this is applied to avoid too much divergence between the two copies, but also partially restrains the observed conformational averaging to peptide plane jumps. The authors recognise the potential that this restraint may have for reducing the extracted amplitudes of the actual motion, and show that when it is completely removed a slight drop of around 0.1 in the effective S^2 is observed. In the absence of this constraint, the axis of the motion can of course shift, and the method becomes more sensitive to other dynamic modes.

5.2.5. Use of geometric models to fit motional amplitude and direction

The utility of employing a specific geometric model to describe the intramolecular motion has also recently been investigated [55]. This approach has the advantage of reducing the number of independent measurements required for the characterisation of local motional amplitudes, thereby alleviating the practically stringent requirement of finding multiple (>5) different alignment media that do not interact too strongly with the protein. The disadvantage is that this approach relies on both the relevance of the dynamic model employed, and on the validity of the structural model used to describe the effective mean conformation. One additional aspect of using known geometric motional models to analyse RDCs is the insight gained into the expected behaviour of experimental systems if these motions were present. Some of the points raised in the previous sections can then be addressed.

It is accepted that protein backbone dynamics are complex processes and that no single geometric model will be universal. Nevertheless extensive molecular simulations of peptides [51] and small proteins [52] have shown that anisotropic motions, in the form of Gaussian axial fluctuations (GAF) around the average position, can be useful for characterising fundamental backbone motions in proteins. Brüschweiler and collaborators have shown that rigid peptide plane fragment motion can be described by a three-dimensional GAF (3D-GAF) motion comprising independent rotations of amplitude $\sigma_{j,k,l}$ around three orthogonal axes (j, k, l) fixed in the peptide plane. The main axis of rotation, l, lies along the axis $^\alpha C_{i-1}$–$^\alpha C_i$, and a reduction of the 3D-GAF model comprising this major component alone (1D-GAF) was indeed initially proposed to interpret order parameters derived from spin relaxation [50]. The other two axes, j and k, lie in the peptide plane and perpendicular to it, respectively. ^{15}N and ^{13}C relaxation

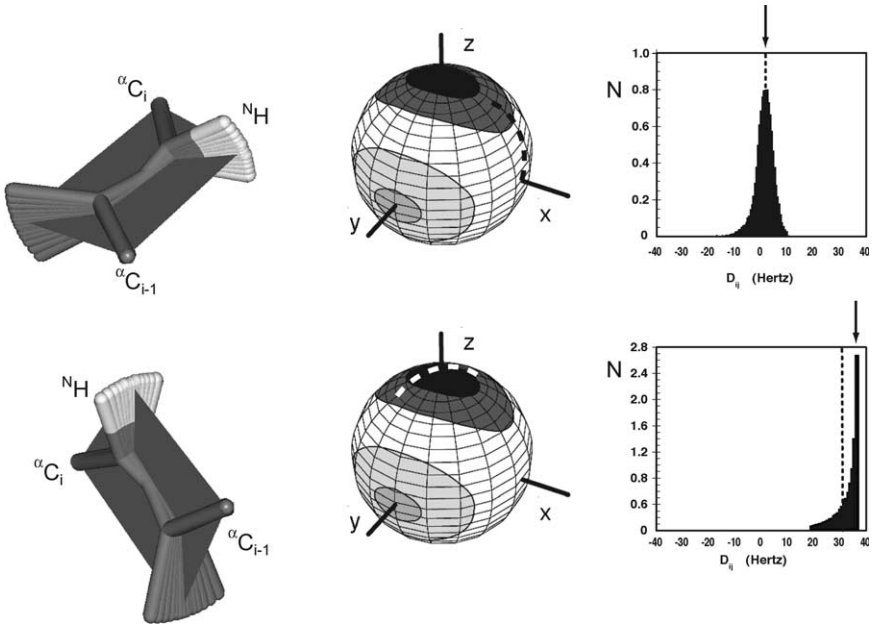

Fig. 13. Illustration of the dependence of an anisotropic component of peptide motion (in this case 1D-GAF motion around the $^{\alpha}C_{i-1} - {}^{\alpha}C_i$ axis) on peptide plane orientation. The amplitude and nature of the motion are identical in both cases, but the averaging has very different effects, because of the differential sampling of the RDC values. The dashed line represents the average RDC, while the arrow indicates the static value that would exist for the average orientation in the absence of motion. Simulations of 50,000 peptide plane conformations from a Gaussian distribution centred around the static plane orientation were performed. The relative population of N–^{N}H RDCs accessed in the presence of this motion (assuming an amplitude of $\sigma = 15°$ for the 1D-GAF motion) is shown in the form of a histogram, where N represents the number of conformations with the relevant RDC value (x-axis). The difference of these values is denoted Δ_{dyn} in Fig. 14. The alignment tensor is defined by ($A_a = -16.0\times10^{-4}$, $A_r = 10.0\times10^{-4}$, $\theta = -59.85°$, $\phi = 97.60°$ and $\psi = 47.85°$).

rates have been fitted to this, or similar dynamic models [51–54]. For the case of Ubiquitin, average angular amplitudes of $\sigma_l = 16.6°$ and $\sigma_j = \sigma_k = 6.7°$ were derived for those residues displaying a Gaussian axial fluctuation behaviour in the MD simulation.

Following these leads, a recent study investigated the effects of anisotropic peptide plane motions on RDCs with the aim of incorporating this kind of dynamic motion into the analysis of experimental data. Using numerical simulations of Gaussian axial fluctuations (GAF) motions, the authors observe highly heterogeneous averaging effects, depending on the orientation of the peptide plane, and therefore both the direction of the N–^{N}H vector and the rotation axis $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ with respect to the alignment tensor (Fig. 13). Equal amplitude anisotropic motions are shown to result in very differently shaped distributions of sampled RDCs resulting in different scaling effects for the effective RDCs from different sites in the molecule. Anisotropic motional averaging is also shown to increase the absolute value of the averaged RDC compared to the static value for certain plane orientations and amplitudes. This contrasts sharply with the case of cylindrically symmetric motion of uniform amplitude, where RDC values would be scaled identically throughout the molecule, independent of the peptide plane orientation. The implications of this observation for quantitative alignment tensor determination are clear, in the presence of average amplitude GAF motions of, for example, $\sigma = 15°$, values of A_a and A_r would already be scaled by 0.91 (Fig. 14). If in addition a common librational axially symmetric component were present, the total alignment tensor amplitude would then be scaled by the product of these two order parameters if dynamics were ignored in the tensor description.

The fact that anisotropic dynamics average differently depending on the orientation of the peptide plane signifies that the effects of these dynamics are encoded with respect

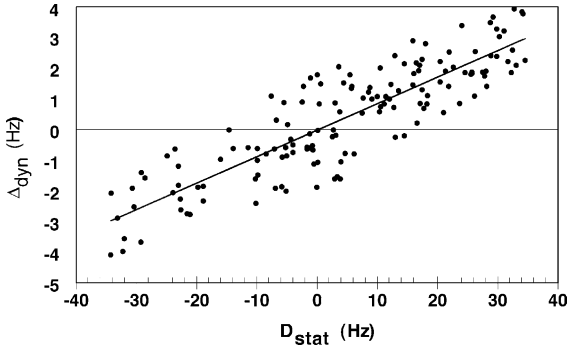

Fig. 14. Effects of anisotropic peptide plane motion on average N–^{N}H RDCs. Distribution of values of $\Delta_{dyn} = (D_{stat} - D_{av})$ with respect to D_{stat} calculated for 144 peptide planes distributed throughout a model protein (sulphite reductase) assuming $\sigma = 15°$ 1D-GAF motion. Simulations were performed with $A_r = 2A_a/3$ and $A_a = -16.0\times10^{-4}$. Linear regression is plotted as a solid line. No motion would result in flat line at $\Delta_{dyn} = 0$, while cylindrically symmetric motion would give a linear relationship between Δ_{dyn} and D_{stat}.

to structure, and allows this component to be characterised using a structure-based analysis. The analytical expression derived in Eq. (24), describes the averaging of RDCs under the influence of GAF motions around the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis, in terms of peptide plane orientation defined by the angles (α, β, γ) and the motional amplitude (σ_i). This means that if (α, β, γ) are known (from a high-resolution structural model for example) anisotropic N–^{N}H reorientation amplitudes can be described using the single parameter σ. As a first application, this expression allowed the incorporation of an average peptide plane reorientation amplitude (σ_{av}) into alignment tensor analyses. The presence of anisotropic reorientational dynamics in proteins was investigated by optimising (σ_{av}) as an additional parameter in the alignment tensor analysis of single-medium N–^{N}H RDC data sets. In this case, six parameters are simultaneously extracted (A_a, A_r, θ, ϕ, ψ, σ_{av}), rather than five, as is the case for the static model. The significance of the improved fitting of the alignment tensor using this oversimplified model of a common or mean amplitude for this motional mode, can then be gauged by comparing the static and dynamic models.

Values of σ_{av} were determined for secondary structural elements from single experimental N–^{N}H RDC data sets from five different proteins, for which high-resolution structural models were available, yielding statistically significant improvement over the static description and detecting similar amplitudes ranging from 14.4 to 17.0° for the different proteins. A higher value of $\sigma_{av} = 20°$ from loop regions was found using two independent sets of N–^{N}H RDCs in the protein Lysozyme, for which a very high-resolution (0.9 Å) structure is available. Comparison of fitting behaviour over 13 structures from Lysozyme of crystal diffraction resolution ranging from 0.9 to 2.1 Å indicates a small spread of motional amplitudes, demonstrating that the method is robust up to this level of resolution. By simulation, the authors also demonstrate that randomly distributed structural error cannot give rise to the observed anisotropic motional amplitudes.

The authors have also examined the sensitivity of other backbone RDCs present in the peptide plane to anisotropic motion, and identified the $^{\alpha}$C–C′ vector as the least affected commonly measured ^{1}D coupling. A combined description of $^{\alpha}$C–C′ and N–^{N}H RDCs under the influence of GAF motions was therefore developed leading to similar values of motional amplitudes compared to those extracted from N–^{N}H RDC values, and, importantly, a more significant improvement with respect to the static model than that extracted using only the dynamically averaged N–^{N}H coupling. This allowed the extraction of similar average motional amplitudes from four of the five recently published data sets measured from protein G. An interesting detail of this study was the confirmation that the commonly used N–^{N}H bond length of 1.041 Å, determined from calibration compared to RDCs measured between heavy atoms [136] is in agreement with a more common static value of 1.02 Å if the vector is undergoing GAF-like internal motions of $\sigma_{av} = 15°$.

Finally, it was demonstrated that the 1D-GAF description closely reproduces data simulated using all three components of the more complex, but also physically more realistic 3D-GAF motion, containing smaller amplitude rotations about axes orthogonal to the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis. A convolution of the analytical 1D-GAF motion, and an additional axially symmetric component $S^{ax}_{C',-^{\alpha}C} = S^{ax}_{N-^{N}H}$ was defined such that:

$$\langle D_{N-^{N}H}\rangle_{3DGAF} \approx S^{ax}_{N-^{N}H}\langle D_{N-^{N}H}\rangle_{o-GAF} \tag{45}$$

where $\langle D_{N-^{N}H}\rangle_{o-GAF}$ is given in Eq. (24), and

$$\langle D_{C'-^{\alpha}C}\rangle_{3DGAF} \approx -\frac{\gamma_t\gamma_j\mu_0 h}{16\pi^3 r^3_{C'-^{\alpha}C}} S^{ax}_{C'-^{\alpha}C} \times \left[A_a(3\cos^2\theta - 1) + \frac{3}{2}A_r\sin^2\theta\cos 2\phi\right] \tag{46}$$

Despite its inherent simplifications, this study therefore provided direct evidence for the existence of GAF-like motions on the protein backbone, and extracted a more accurate definition of the alignment tensor amplitude that is less affected by anisotropic motions. As we have seen, this is a potentially important advance towards the extraction of quantitative dynamic amplitudes.

In a follow-up study, the same authors then explored the possibility of using the *ortho*-GAF description to study the anisotropic component of local peptide dynamics along the protein backbone [200]. In this case, RDC data from m alignment media ($D^i_{m,exp}$) were combined to extract local motional amplitudes (σ_i) at each site i throughout the protein, by minimizing the target function:

$$\chi^2_i = \sum_m (\{\langle D^i_{m,calc}\rangle_{GAF} - D^i_{m,exp}\}^2/\delta^{i2}_m) \tag{47}$$

$D^i_{m,calc}$ are calculated from the known peptide orientations $(\alpha, \beta, \gamma)^i$ present in the high-resolution crystal structure, and the alignment tensor parameters, $(A_a, A_r, \theta, \phi, \psi)^m$ using Eq. (24). The alignment tensor parameters were determined using $^{\alpha}$C–C′ RDCs following the logic from the previous article, that assumes that this vector is the least affected by GAF-like motions. In this case, the success of the technique relies on the different dependence of the same anisotropic motion in the presence of differently oriented tensors (Fig. 15).

The extracted σ values were transformed into an anisotropic order parameter S^2_{aniso} (Eq. (25)), and compared with dynamic amplitudes derived from ^{15}N relaxation [201] (Fig. 16). This comparison should allow the identification of sites undergoing additional longer time-scale motions. In general, the pattern of order parameters along the sequence is reproduced by the two methods. Motions of similar amplitude were found in the central α-helix of the protein, with higher order parameters characteristic of a compact, relatively rigid structural domain that retains a degree of

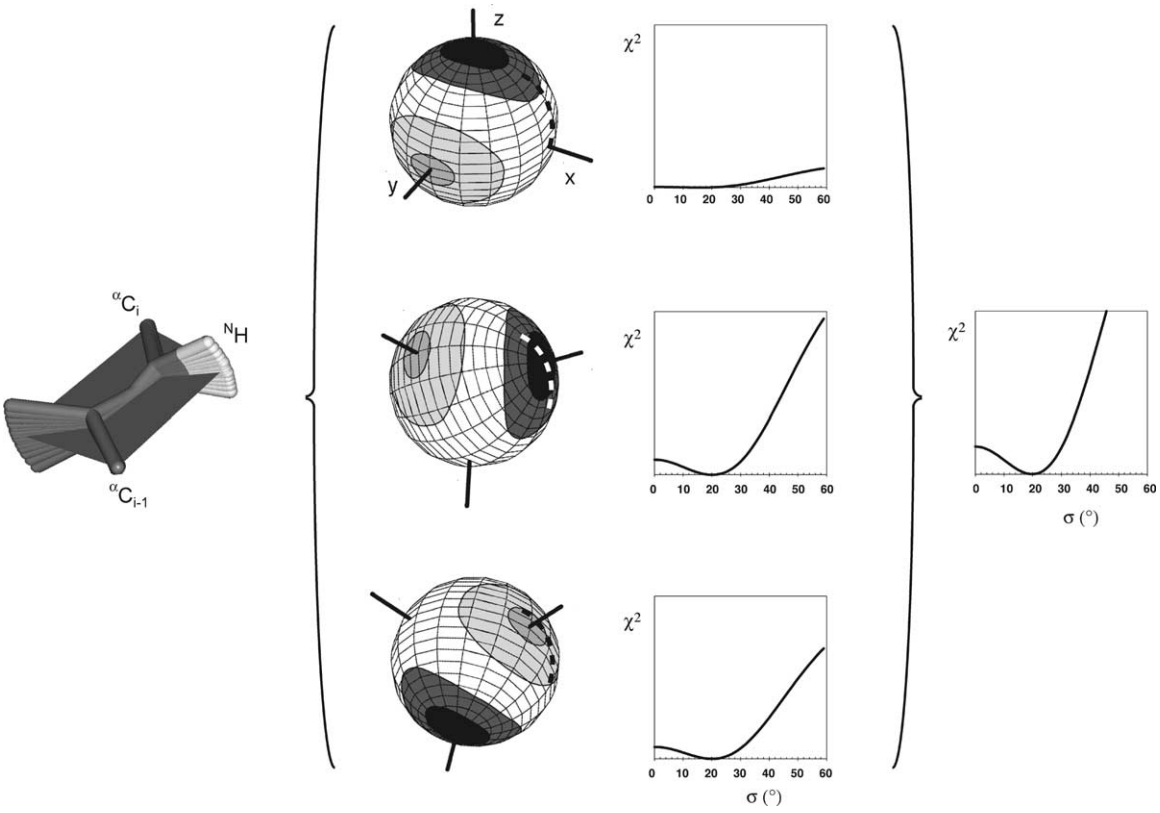

Fig. 15. Determination of the amplitude of anisotropic motions using the different dependence of the same anisotropic motion in the presence of three differently oriented tensors. The right-hand side of the figure shows the χ^2 function with respect to the amplitude of the GAF motion σ, optimised by minimising the function in Eq. (47). Each individual contribution to the total χ^2 function is shown. Note that while the function is not necessarily well defined in each case, the presence of differently oriented tensor axes improves the definition.

structural integrity similar to that observed on the faster time-scale. Some dynamic regions correspond to positions found from ^{15}N relaxation, essentially loop regions, although for certain sites the RDC-derived amplitude is significantly greater, corresponding to sites experiencing additional, longer time-scale motions that are invisible to ^{15}N relaxation. No additional scaling was applied to either data set. It is also of interest that this measure of conformational disorder in solution derived from RDC is in qualitative agreement with the distribution of *B*-factors

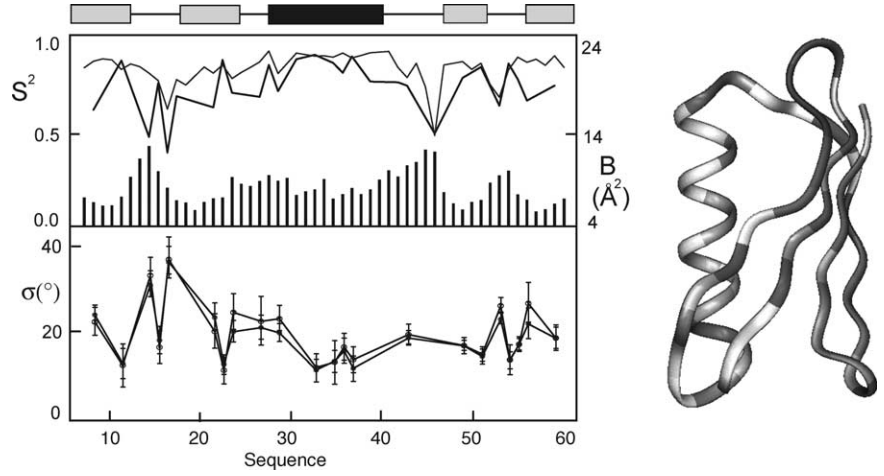

Fig. 16. Dynamic parameters from RDCs measured in protein G. Order parameters derived from GAF analysis of RDC (thick line) compared to S^2 derived from ^{15}N relaxation (thin line) and *B*-factors from the crystal study (histogram). Position of the α-helix is indicated in black, and β-sheets in grey. Bottom: GAF amplitudes derived using all five data sets (open circles) and the two most independent (filled circles). Errors are estimated from noise-based Monte-Carlo simulations. Right: ribbon diagram showing the distribution of anisotropic motions throughout the protein: the darker the shading the lower the order parameters. White indicates that the parameter is not reported.

from the X-ray study, and with sites localised from an extensive study of multiple RDC from this protein G as having the largest disagreement with the crystal conformation [202]. This analysis was also applied to data from Lysozyme, where order parameters derived using structures of differing atomic resolution (0.9, 1.5 and 2.1 Å) give very similar S^2_{aniso} values, demonstrating the robustness of the model selection criteria with respect to structural error.

Analysis carried out using only data from the two most independent alignment tensors results in similar amplitude motions compared to fitting using all five data sets, implying that the method is relatively robust in this respect, and that the dependence on finding five differently aligning media can be avoided. It is important that care is taken to avoid aliasing of local differences between the known structural model and the actual conformation in solution, into fictitious motional amplitudes. The authors are careful to reject sites where the dynamic model is not statistically valid, or where the target function with respect to σ is ill-defined. Extensive numerical simulations of RDC dynamic averaging under different regimes have also been performed to gauge the extent of such artefacts [203], showing that, perhaps not unexpectedly, the anisotropic motional models are more geometrically stringent than the axially symmetric model, reducing the possibility of aliasing structural noise as artificial dynamic amplitudes. The effect of internal motion on the accurate determination of alignment tensors has been demonstrated by simulating RDCs averaged by GAF motions, of pervasive amplitude 20°, in the presence of three independent alignment tensors from throughout a model peptide chain. Not surprisingly extraction of accurate amplitudes can be achieved if the correct, non-averaged eigenvalues are used to describe the tensors. However, if the tensors are derived by fitting to the structure assuming no motion is present, the eigenvalues are scaled, and local motional amplitudes, determined from the data with respect to these tensors, exhibit an average shift of 0.14 (0.85 compared to the expected 0.71) in the extracted S^2 [203].

Recently, the same data from protein G were analysed using the ensemble averaging approach presented earlier [204]. The authors demonstrate a significant improvement in data fitting using a combination of two conformers, extracting jump amplitudes that are very similarly distributed over the protein sequence compared to the GAF fitting procedure described above. The absolute amplitude of these jumps is however lower than that extracted using the *ortho*-GAF method. As described above, the absolute values of these amplitudes may be influenced by the different treatment of global alignment tensors in the two approaches.

5.2.6. RDCs as probes of structure and dynamics: complementarity or contradiction?

There appear therefore to be conflicting reports emanating from the RDC studies described in the previous sections. Depending on the type of analysis, the same data deliver contrasting pictures of the nature and amplitude of backbone mobility in proteins. It is therefore worthwhile to compare the kind of information derived from the different approaches, to attempt to develop a coherent description of the dynamic behaviour of protein backbones. As mentioned above, as applied, both model-free and restrained molecular dynamics-based methods suffer from an inherent incapacity to quantitatively determine dynamic amplitudes, due to the difficulty in estimating the alignment tensors as they would exist in the absence of internal dynamics. In the model-free approaches, all order parameters are therefore effectively scaled so that all are less than either 1 [191], or less than the relaxation-derived order parameter [193]. This calibration could be prone to unnecessarily low order parameters if, for example, the highest order parameters were affected by noise, with the remaining values then being scaled too severely. In the refinement-based approaches, the alignment tensor is optimized with the structure, so that in this case common components of the motion, both axially symmetric and anisotropic, can be absorbed into the tensor eigenvalues. These components, both of which may be significant, could then be effectively removed from the remaining analysis, as they would already be accounted for. When comparing motional amplitudes derived from the different approaches it is perhaps also useful to note that by far the largest contribution to the motions described by Peti et al. comes from the axially symmetric component, and that both the GAF-averaging and the ensemble averaging methods are in theory insensitive to this kind of movement (as applied ensemble averaging is mainly sensitive to motion about the ${}^{\alpha}\text{C}_{i-1}-{}^{\alpha}\text{C}_i$ axis).

A more subtle aspect to the interpretation of dynamic amplitudes using restrained molecular dynamics calculations may also occur due to the inherent capacity of force-field based calculations to absorb intramolecular motion, such that refinement of a structure removes a proportion of the apparent dynamics. The extent of this effect, and the efficiency of such a technique to determine true dynamic amplitudes, certainly depends on the very complex energy surface created by the particular refinement protocol and the weighting of local conformational force field constants (peptide plane planarity, relative vector angles in the peptide) relative to the RDC target function.

The GAF-averaging approach [55] unambiguously identified the presence of common anisotropic motions about the ${}^{\alpha}\text{C}_{i-1} - {}^{\alpha}\text{C}_i$ axis, and characterised global averages of σ in secondary structural elements of a number of proteins. These average amplitudes are comparable with those derived from rapid (ps–ns) motions, although this average value certainly does not exclude the possible presence of individual sites with higher amplitude motion. The local GAF analysis [200] relies on the availability of a high-resolution structure in order to extract accurate dynamic amplitudes, and is therefore less elegant than the model-free approaches. Because of this dependence on structure, and the sensitivity to a defined anisotropic mode,

the method is possibly most comparable to the ensemble averaging technique. The obvious difference is that while the latter simultaneously refines vector orientation, tensor eigenvalues and dynamic amplitude, the GAF approach assumes that the mean vector orientations are correctly known, and that tensor eigenvalues can be accurately determined, with the disadvantage that any small errors in the structure may artificially influence the extracted amplitudes. It is worth noting, however, that because the alignment tensors are determined in the absence of GAF-like motions (using the $^{\alpha}$C–C′ RDCs or using N–^{N}H RDCs with a common $\langle\sigma\rangle$) the GAF-averaging approach can in theory provide an absolute measure of the amplitude of this component of anisotropic motion. In this respect, it is interesting that the RDC-derived order parameters are systematically lower than their relaxation-derived counterparts, and possibly even more interesting that certain sites apparently undergo additional large amplitude slower motions.

The results from the different techniques applied to the extraction of motional amplitudes are therefore not necessarily contradictory, more a reflection of the different methodologies employed, whether structure-based or dynamics-based, highlighting the fact that residual dipolar couplings are sensitive to both components, and that this dual dependence, expressed in Eq. (10), is extremely difficult to unambiguously unravel.

5.3. RDCs for the study of local sidechain dynamics in proteins

A number of recent contributions have also demonstrated the use of RDCs to characterise dynamic modes and extract motional amplitudes of protein sidechains. Bertini et al. [205] demonstrated that lanthanide-induced alignment of Calbindin can deliver information about the dynamic sampling of peripheral $^{15}NH_2$ moieties in asparagine and glutamine residues. Chou and Bax [206] measured $^{\beta}$C–$^{\beta}$H dipolar couplings which, in combination with $^{\alpha}$C–$^{\alpha}$H and $^{\alpha}$C–C′ measurements allows a method of identification of rotameric states that is independent of the alignment tensor. Results from the protein Calmodulin are compatible with nearly ideal staggered rotamer conformations for approximately half of the non-Ala, non-Gly residues, indicating that for these sites motional jumps are not necessary to explain the data. The authors also calculate the effects of 15° amplitude GAF motions on the effective couplings and find that this causes a mean correction of the order of 10%, similar to that found in the backbone analyses presented above. Mittermaier and Kay [207] have directly studied dynamics about the χ_1 dihedral angle in proteins, using data from two differently aligning media for protein L. In this case, the local structure is refined using backbone RDCs, and the resultant $^{\alpha}$C–$^{\beta}$C orientation used to model $^{\beta}$C–$^{\beta}$H dipolar couplings in terms of a static model, GAF motions about the mean orientation of the $^{\alpha}$C–$^{\beta}$C vector and a jump-model between rotamers. In this case, the majority of residues can be accounted for by a static model, and the authors conclude that small amplitude fluctuations are not necessary to reproduce the experimental data. More recently, a study of aromatic sidechain dynamics in the SMN Tudor domain relied on the symmetric properties of different motional averaging regimes to identify the presence of ring-flips in the binding pocket of this protein [208].

6. Conclusions

In this review, we have tried to give a flavour of some of the multiple aspects of biomolecular structure and dynamics that can be probed using dipolar couplings measured in partially aligned systems. Readers may feel that a number of related aspects of the field have been neglected, for example paramagnetic effects have only been briefly touched on, and no mention has been made of non-averaged chemical shift effects. The use of RDCs for the study of the structure and dynamics of partially folded, or unfolded proteins, where conformational information is otherwise very difficult to extract, has recently motivated considerable interest. Again in the interests of space this very active field has not been described here. A large number of biologically important applications of RDC for structure refinement have also not been presented, essentially because the examples presented here are thought to illustrate the methodology sufficiently well. This is evidence of the now-routine application of these still relatively novel techniques, and testifies to the rapidity with which technology is currently transferred from development to application in the field of biomolecular NMR. It is clear that RDC-based analysis of biomolecular conformation will play a major role in future development in this field.

Acknowledgements

The author thanks Dr Pau Bernado for helpful discussion and for reading the manuscript. This work was supported by the Commisariat à l'Energie Atomique, the CNRS and the European Union (contract HPRI-CT-2001-50035).

References

[1] K. Wüthrich, NMR of Proteins and Nucleic Acids, Wiley, New York, 1986.
[2] G.M. Clore, A.M. Gronenborn, Nat. Struct. Biol. 4 (1997) 849.
[3] D. Case, P.E. Wright, in: G.M. Clore, A.M. Gronenborn (Eds.), NMR of Proteins, Macmillan, London, 1993.
[4] R.A. Venters, R. Thompson, J. Cavanagh, J. Mol. Struct. 602 (2002) 275.
[5] A. Abragam, The Principles of Nuclear Magnetism, Oxford University Press, Oxford, 1961.
[6] A. Saupe, G. Englert, Phys. Rev. Lett. 11 (1963) 462.

[7] J.W. Emsley, J.C. Lindon, NMR Spectroscopy using Liquid Crystal Solvents, Pergamon Press, Oxford, 1975.
[8] J.A.B. Lohman, C. Maclean, Chem. Phys. 35 (1978) 269.
[9] C. Gayathri, A.A. Bothner-by, P.C.M. van Zijl, C. Maclean, Chem. Phys. Lett. 87 (1982) 192.
[10] E.W. Bastiaan, C. Maclean, P.C.M. van Zijl, A.A. Bothner-By, Annu. Rev. NMR Spectrosc. 278 (1987) 1111.
[11] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, J.H. Prestegard, Proc. Natl Acad. Sci. USA 92 (1995) 9279.
[12] N. Tjandra, S. Grzesiek, A. Bax, J. Am. Chem. Soc. 118 (1996) 6264.
[13] H.C. Kung, K.Y. Wang, I. Goljer, P.H. Bolton, J. Magn. Reson. B 109 (1995) 323.
[14] N. Tjandra, J. Omichinski, A.M. Gronenborn, G.M. Clore, A. Bax, Nat. Struct. Biol. 4 (1997) 732.
[15] N. Tjandra, A. Bax, Science 278 (1997) 1111.
[16] C.R. Sanders, B.J. Hare, K.P. Howard, J.H. Prestegard, Prog. NMR Spectrosc. 26 (1994) 421.
[17] M.R. Hansen, L. Müller, A. Pardi, Nat. Struct. Biol. 5 (1998) 1065.
[18] G.M. Clore, M.R. Starich, A.M. Gronenborn, J. Am. Chem. Soc. 120 (1998) 10571.
[19] H.-J. Sass, F. Cordier, A. Hoffman, A. Cousin, J.G. Omichinski, H. Lowen, S. Grzesiek, J. Am. Chem. Soc. 121 (1999) 2047.
[20] B. Koenig, J.-S. Hu, M. Ottiger, S. Bose, R. Hendler, A. Bax, J. Am. Chem. Soc. 121 (1999) 1385.
[21] R. Tycko, F. Blanco, Y. Ishii, J. Am. Chem. Soc. 122 (2000) 9340.
[22] J. Sass, G. Musco, S.J. Stahl, P.T. Wingfield, S. Grzesiek, J. Biomol. NMR 18 (2000) 303.
[23] M. Rückert, G. Otting, J. Am. Chem. Soc. 122 (2000) 7793.
[24] J.H. Prestegard, Nat. Struct. Biol. 5 (1998) 517.
[25] A. Bax, G. Kontaxis, N. Tjandra, Methods Enzymol. 339 (2001) 127.
[26] E. de Alba, N. Tjandra, Prog. NMR Spectrosc. 40 (2002) 175.
[27] A. Bax, Protein Sci. 12 (2003) 1.
[28] J.H. Prestegard, C. Bougault, A. Kishore, Chem. Rev. 104 (2004) 3519.
[29] L.E. Kay, Nat. Struct. Biol. 5 (1998) S513.
[30] R. Ishima, D. Torchia, Nat. Struct. Biol. 7 (2000) 740.
[31] A. Palmer, Annu. Rev. Biophys. Biomol. Struct. 30 (2001) 129.
[32] R. Brüschweiler, Curr. Opin. Struct. Biol. 13 (2003) 175.
[33] B. Kieffer, A. Atkinson, Prog. NMR Spectrosc. 44 (2004) 141.
[34] G. Lipari, A. Szabo, J. Am. Chem. Soc. 104 (1982) 4546.
[35] G. Lipari, A. Szabo, J. Am. Chem. Soc. 104 (1982) 4559.
[36] N.A. Farrow, R. Muhandiram, A.U. Singer, S. Pascal, C.M. Kay, G. Gish, S.E. Shoelson, T. Pawson, J.D. Forman-Kay, L.E. Kay, Biochemistry 33 (1994) 5984.
[37] A.G. Palmer, C.D. Kroenke, J.P. Loria, Methods Enzymol. 339 (2001) 204.
[38] M. Akke, Curr. Opin. Struct. Biol. 12 (2002) 642.
[39] J.R. Tolman, H.M. Hashimi, Annu. Rep. NMR Spectrosc. 51 (2003) 105.
[40] J.R. Tolman, Curr. Opin. Struct. Biol. 11 (2001) 532.
[41] A. Saupe, Angew. Chem., Int. Ed. (Engl.) 7 (1968) 97.
[42] J.A. Losonczi, M. Andrec, M.W.F. Fischer, J.H. Prestegard, J. Magn. Reson. 138 (1999) 334.
[43] B.E. Ramirez, A. Bax, J. Am. Chem. Soc. 120 (1998) 9106.
[44] H.M. Al-Hashimi, H. Valafar, M. Terrell, M. Zartler, M. Eidsness, J.H. Prestegard, J. Magn. Reson. 143 (2000) 402.
[45] J.H. Prestegard, H.M. Al-Hashimi, J.R. Tolman, Q. Rev. Biophys. 33 (2000) 371.
[46] N. Skrynnikov, N. Goto, D. Yang, W.Y. Choy, J.R. Tolman, G.A. Mueller, L.E. Kay, J. Mol. Biol. 295 (2000) 1265.
[47] J. Meiler, J. Prompers, C. Griesinger, R. Brüschweiler, J. Am. Chem. Soc. 123 (2001) 6098.
[48] V.A. Daragan, K.H. Mayo, Prog. NMR Spectrosc. 130 (1997) 329.
[49] D.E. Woessner, J. Chem. Phys. 37 (1962) 647.
[50] R. Brüschweiler, P.E. Wright, J. Am. Chem. Soc. 116 (1994) 8426.
[51] T. Bremi, R. Brüschweiler, J. Am. Chem. Soc. 119 (1997) 6672.
[52] S.F. Lienin, T. Bremi, B. Brutscher, R. Brüschweiler, R.R. Ernst, J. Am. Chem. Soc. 120 (1998) 9870.
[53] M.W.F. Fischer, L. Zeng, Y. Pang, W. Hu, A. Majumdar, E.R.P. Zuiderweg, J. Am. Chem. Soc. 119 (1997) 12629.
[54] Y. Pang, E.R.P. Zuiderweg, J. Am. Chem. Soc. 122 (2000) 4841.
[55] P. Bernado, M. Blackledge, J. Am. Chem. Soc. 126 (2004) 4907.
[56] M. Gochin, H. Roder, Protein Sci. 4 (1995) 296.
[57] L. Banci, I. Bertini, J.G. Huber, C. Luchinat, A. Rosato, J. Am. Chem. Soc. 120 (1998) 12903.
[58] J. Boisbouvier, P. Gans, M. Blackledge, B. Brutscher, D. Marion, J. Am. Chem. Soc. 121 (1999) (1999) 7700.
[59] I. Bertini, G. Cavallero, M. Cosenza, R. Kummerle, C. Luchinat, M. Piccoli, L. Poggi, J. Biomol. NMR 23 (2002) 115.
[60] J.-C. Hus, D. Marion, M. Blackledge, J. Mol. Biol. 298 (2000) 927.
[61] I. Bertini, C. Luchinat, G. Parigi, Prog. NMR Spectrosc. 40 (2002) 249.
[62] J. Feeney, B. Birdsall, A.F. Bradbury, R.R. Biekofsky, P.M. Bayley, J. Biomol. NMR 21 (2001) 41.
[63] V. Gaponenko, A.S. Altieri, J. Li, R.A. Byrd, J. Biomol. NMR 24 (2002) 143.
[64] J. Wöhnert, K. Franz, M. Nitz, B. Imperialiand, H. Schwalbe, J. Am. Chem. Soc. 125 (2003) 13338.
[65] T. Ikegami, L. Verdier, P. Sakhaii, S. Grimme, B. Pescatore, K. Saxena, K.M. Fiebig, C. Griesinger, J. Biomol. NMR 29 (2004) 339.
[66] R.S. Prosser, J.S. Hwuang, R.R. Vold, Biophys. J. 74 (1998) 2405.
[67] P. Bernado, R. Barbieri, E. Padros, C. Luchinat, M. Pons, J. Am. Chem. Soc. 124 (2002) 374.
[68] M. Zweckstetter, A. Bax, J. Biomol. NMR 23 (2002) 127.
[69] G.M. Clore, A.M. Gronenborn, A. Bax, J. Magn. Reson. 133 (1998) 216.
[70] N. Skrynnikov, L.E. Kay, J. Biomol. NMR 18 (2000) 239.
[71] J.J. Warren, P.B. Moore, J. Magn. Reson. 149 (2001) 271.
[72] D. Fushman, R. Ghose, D. Cowburn, J. Am. Chem. Soc. 122 (2000) 10640.
[73] L. Zidek, P. Padrta, J. Chmelik, V. Sklenar, J. Magn. Reson. 162 (2003) 385.
[74] D. Bryce, A. Bax, J. Biomol. NMR 28 (2004) 273.
[75] W. Maier, A. Saupe, Z. Naturforsch., A 14 (1959) 882.
[76] J.W. Emsley, G.R. Luckhurst, Mol. Phys. 41 (1980) 19.
[77] A.J. Van der Est, M.Y. Kok, E.E. Burnell, Mol. Phys. 60 (1987) 397.
[78] C.T. Yim, D.F.R. Gilson, Can. J. Chem. 68 (1990) 875.
[79] A. Ferrarini, G. Moro, P. Nordio, G.R. Luckhurst, Mol. Phys. 77 (1992) 1.
[80] M. Zweckstetter, A. Bax, J. Am. Chem. Soc. 122 (2000) 3791.
[81] C. Bewley, G.M. Clore, J. Am. Chem. Soc. 122 (2000) 6009.
[82] M.X. Fernandes, P. Bernado, M. Pons, J. Garcia de al Torre, J. Am. Chem. Soc. 123 (2001) 12037.
[83] H.F. Azurmendi, C.A. Bush, J. Am. Chem. Soc. 124 (2002) 2426.
[84] A. Almond, J.B. Axelsen, J. Am. Chem. Soc. 124 (2002) 9986.
[85] A. Ferrarini, J. Phys. Chem. 107 (2003) 7923.
[86] M. Zweckstetter, G. Hummer, A. Bax, Biophys. J. 86 (2004) 3444.
[87] M. Cai, Y. Huang, R. Zheng, S.Q. Wei, R. Ghuirlando, M.S. Lee, R. Craigie, A.M. Gronenborn, G.M. Clore, Nat. Struct. Biol. 5 (1998) 903.
[88] D.S. Garrett, Y.-J. Seol, A. Peterkovsky, A.M. Gronenborn, G.M. Clore, Nat. Struct. Biol. 6 (1999) 166.
[89] E. Mollova, M.R. Hansen, A. Pardi, J. Am. Chem. Soc. 22 (2000) 11561.
[90] K. Huang, J.M. Louis, L. Donaldson, F.-L. Lim, A. Sharrocks, G.M. Clore, EMBO J. 19 (2000) 2615.
[91] N. Skrynnikov, C. R. Phys. 5 (2004) 359.
[92] N.K. Goto, N.R. Skrynnikov, F.W. Dahlquist, L.E. Kay, J. Mol. Biol. 308 (2001) 745.
[93] K. Bondensgaard, E.T. Mollova, A. Pardi, Biochemistry 41 (2002) 11532.
[94] D. MacDonald, P. Lu, Curr. Opin. Struct. Biol. 12 (2002) 337.

[95] E.R.P. Zuiderweg, Biochemistry 41 (2002) 1.
[96] G.M. Clore, Proc. Natl Acad. Sci. 97 (2000) 9021.
[97] M.A. McCoy, D.F. Wyss, J. Am. Chem. Soc. 124 (2002) 11758.
[98] G.M. Clore, C.D. Schwieters, J. Am. Chem. Soc. 125 (2003) 2902.
[99] G.S. Wang, J.M. Louis, M. Sondej, Y.J. Seok, A. Peterkofsky, G.M. Clore, EMBO J. 19 (2000) 5635.
[100] G. Cornilescu, B.R. Lee, C.C. Cornilescu, G.S. Wang, A. Peterkofsky, G.M. Clore, J. Biol. Chem. 277 (2002) 42289.
[101] K. Osapay, D.A. Case, J. Am. Chem. Soc. 113 (1991) 9436–9444.
[102] C. Dominguez, R. Boelens, A.M.J.J. Bonvin, J. Am. Chem. Soc. 125 (2003) 1731.
[103] E.T. Olejniczak, R.P. Meadows, H. Wang, M. Cai, D.G. Netteshein, S.W. Fesik, J. Am. Chem. Soc. 121 (1999) 9249.
[104] H. Shimizu, A. Donohue-Rolfe, S.W. Homans, J. Am. Chem. Soc. 125 (1999) 1731.
[105] P.J. Bolon, H.M. Al-Hashimi, J.H. Prestegard, J. Mol. Biol. 293 (1999) 107.
[106] H.M. Al-Hashimi, P.J. Bolon, J.H. Prestegard, J. Magn. Reson. 142 (2000) 153.
[107] N.U. Jain, S. Noble, J.H. Prestegard, J. Mol. Biol. 328 (2003) 441.
[108] B.W. Koenig, G. Kontaxis, D.C. Mitchell, J.M. Louis, B.J. Litman, A. Bax, J. Mol. Biol. 322 (2002) 441.
[109] J. Chou, J.D. Kaufman, S.J. Stahl, P.T. Wingfield, A. Bax, J. Am. Chem. Soc. 124 (2002) 2450.
[110] J. Meiler, W. Peti, C. Griesinger, J. Biomol. NMR 17 (2000) 283.
[111] P. Dosset, J.-C. Hus, D. Marion, M. Blackledge, J. Biomol. NMR 20 (2001) 223.
[112] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, J. Chem. Phys. 21 (1953) 1087.
[113] C.T. Leondes, Fuzzy Logic and Expert Systems Applications, Academic Press, San Diego, CA, 1997.
[114] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C, The Art of Scientific Computing, Cambridge University Press, Cambridge, UK, 1988.
[115] H. Valafar, J.H. Prestegard, J. Magn. Reson. 167 (2004) 228.
[116] L. Champier, N. Sibille, B. Bersch, B. Brutscher, M. Blackledge, J. Coves, Biochemistry 41 (2002) 3770.
[117] A. Annila, H. Aitio, E. Thulin, T.J. Drakenberg, J. Biomol. NMR 14 (1999) 223.
[118] M. Andrec, P. Du, R.M. Levy, J. Am. Chem. Soc. 123 (2001) 1222.
[119] H. Valafar, J.H. Prestegard, Bioinformatics 19 (2003) 1549.
[120] C.A. Rohl, D. Baker, J. Am. Chem. Soc. 124 (2002) 2723.
[121] J. Meiler, D. Baker, Proc. Natl Acad. Sci. USA 100 (2003) 15404.
[122] T. Haliloglu, A. Kolinski, J. Skolnick, Biopolymers 70 (2003) 548.
[123] K.T. Simons, C. Kooperberg, E. Huang, D. Baker, J. Mol. Biol. 268 (1997) 209.
[124] J.-C. Hus, J. Prompers, R. Brüschweiler, J. Magn. Reson. 157 (2002) 119.
[125] M.F. Mesleh, G. Veglia, T.M. DeSilva, F.M. Marassi, S.J. Opella, J. Am. Chem. Soc. 124 (2002) 4206.
[126] A. Mascioni, G. Veglia, J. Am. Chem. Soc. 125 (2002) 12520.
[127] F. Delaglio, G. Kontaxis, A. Bax, J. Am. Chem. Soc. 122 (2000) 2142.
[128] M. Andrec, P. Du, R.M. Levy, J. Biomol. NMR 21 (2001) 335.
[129] Y.-S. Jung, M. Sharma, M. Zweckstetter, Angew. Chem., Int. Ed. (Engl.) 43 (2004) 3479.
[130] C.A. Fowler, F. Tian, H.M. Al-Hashimi, J.H. Prestegard, J. Mol. Biol. 304 (2000) 447.
[131] F. Tian, H. Valafar, J.H. Prestegard, J. Am. Chem. Soc. 123 (2001) 11791.
[132] W.J. Wedermeyer, C.A. Rohl, H.A. Scheraga, J. Biomol. NMR 22 (2002) 137.
[133] L. Wang, B.R. Donald, J. Biomol. NMR 29 (2004) 223.
[134] A.W. Giesen, S.W. Homans, J.M. Brown, J. Biomol. NMR 25 (2003) 63.
[135] J.-C. Hus, D. Marion, M. Blackledge, J. Am. Chem. Soc. 123 (2001) 1541.
[136] G. Cornilescu, J.L. Marquardt, M. Ottiger, A. Bax, J. Am. Chem. Soc. 120 (1998) 6836.
[137] M. Ottiger, A. Bax, J. Am. Chem. Soc. 120 (1998) 12334.
[138] R.R. Ketchem, W. Hu, T.A. Cross, Science 261 (1993) 1457.
[139] T.A. Cross, S.J. Opella, Curr. Opin. Struct. Biol. 4 (1994) 574.
[140] S.J. Opella, F.M. Marassi, J.J. Gesell, A.P. Valente, Y. Kim, M. Oblatt-Montal, M. Montal, Nat. Struct. Biol. 6 (1999) 374.
[141] S. Béraud, B. Bersch, B. Brutscher, P. Gans, F. Barras, M. Blackledge, J. Am. Chem. Soc. 124 (2002) 13709.
[142] G.M. Clore, A.M. Gronenborn, Proc. Natl Acad. Sci. USA 95 (1998) 5891.
[143] V. Tsui, L. Zhu, T.-H. Huang, P.E. Wright, D.A. Case, J. Biomol. NMR 16 (2000) 9.
[144] G.M. Clore, M.R. Starich, A.M. Gronenborn, J. Am. Chem. Soc. 120 (1998) 10571.
[145] J. Meiler, N. Blomberg, M. Nilges, C. Griesinger, J. Biomol. NMR 16 (2000) 245.
[146] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Acta Crystallogr. D 54 (1998) 905.
[147] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, G.M. Clore, J. Magn. Reson. 160 (2003) 65.
[148] P. Güntert, C. Mumenthaler, K. Wüthrich, J. Mol. Biol. 273 (1997) 283.
[149] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.R. Ross, T.E. Cheatham, S. Debolt, D. Ferguson, G. Seibel, P. Kollman, Comput. Phys. Commun. 91 (1995) 1.
[150] J.J. Chou, S. Li, A. Bax, J. Biomol. NMR 18 (2000) 217.
[151] J.J. Chou, S. Li, C.B. Klee, A. Bax, Nat. Struct. Biol. 8 (2001) 990.
[152] G.M. Clore, A.M. Gronenborn, N. Tjandra, J. Magn. Reson. 131 (1998) 159.
[153] S. Moltke, S. Grzesiek, J. Biomol. NMR 15 (1999) 77.
[154] E. Mollova, A. Pardi, Curr. Opin. Struct. Biol. 10 (2000) 298.
[155] A. Vermeulen, H. Zhou, A. Pardi, J. Am. Chem. Soc. 122 (2000) 9638.
[156] D. MacDonald, K. Herbert, X.L. Zhang, T. Polgruto, P. Lu, J. Mol. Biol. 306 (2001) 1081.
[157] N. Tjandra, S. Tate, A. Ono, M. Kainosho, A. Bax, J. Am. Chem. Soc. 122 (2000) 6190.
[158] G.R. Zimmerman, R.D. Jenison, C.L. Wick, J.-P. Simorre, A. Pardi, Nat. Struct. Biol. 4 (1997) 644.
[159] N. Sibille, A. Pardi, J.-P. Simorre, M. Blackledge, J. Am. Chem. Soc. 123 (2001) 12135.
[160] S.A. McCallum, A. Pardi, J. Mol. Biol. 326 (2003) 1037.
[161] P. Lukavsky, I. Kim, G. Otto, J.D. Puglisi, Nat. Struct. Biol. 10 (2003) 1033.
[162] G.M. Clore, D.S. Garrett, J. Am. Chem. Soc. 121 (1999) 9008.
[163] C.L. Brooks III, M. Karplus, B.M. Pettitt, Proteins, A Theoretical Persepctive of Dynamics, Structures and Thermodynamics, Wiley, New York, 1987.
[164] A. Torda, R.M. Scheek, W.F. van Gunsteren, J. Mol. Biol. 214 (1989) 223.
[165] A.M.J.J. Bonvin, A.T. Brunger, J. Mol. Biol. 250 (1995) 80.
[166] R.M. Scheek, A. Torda, J. Kemmink, W.F. van Gunsteren, in: J. Hoch (Ed.), Computional Aspects of the Study of Biological Macromolecules by NMR, Plenum Press, New York, 1991.
[167] A. Hess, R.M. Scheek, J. Magn. Reson. 164 (2003) 19.
[168] G.M. Clore, C.D. Schwieters, J. Am. Chem. Soc. 126 (2004) 2923.
[169] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, J.H. Prestegard, Nat. Struct. Biol. 4 (1997) 292.
[170] A. Bax, N. Tjandra, Nat. Struct. Biol. 4 (1997) 254.
[171] M.W.F. Fischer, J.A. Losonczi, J.L. Weaver, J.H. Prestegard, Biochemistry 38 (1999) 9013.
[172] D. Braddock, J.M. Louis, J.L. Baber, D. Levens, G.M. Clore, Nature 415 (2002) 1051.

[173] D. Braddock, M. Cai, J.L. Baber, Y. Huang, G.M. Clore, J. Am. Chem. Soc. 123 (2001) 8643.
[174] D.M. Jacobs, K. Saxena, M. Vogtherr, P. Bernadó, M. Pons, K.M. Fiebig, J. Biol. Chem. 278 (2003) 26174.
[175] O. Millet, R.P. Hudson, L.E. Kay, Proc. Natl Acad. Sci. USA 100 (2003) 12700.
[176] J.A. Lukin, G. Kontaxis, V. Simplaceanu, Y. Yuan, A. Bax, C. Ho, Proc. Natl Acad. Sci. USA 100 (2003) 517.
[177] I. Bertini, C. Del Bianco, I. Gelis, N. Katsaros, C. Luchinat, G. Parigi, M. Peana, A. Provenzano, M.A. Zoruddu, Proc. Natl Acad. Sci. USA 101 (2004) 6841.
[178] J.L. Baber, A. Szabo, N. Tjandra, J. Am. Chem. Soc. 123 (2001) 3953.
[179] H.M. Al-Hashimi, Y. Gosser, A. Gorin, W.D. Hu, A. Majumdar, D.J. Patel, J. Mol. Biol. 315 (2002) 95.
[180] Q. Zhang, R. Throolin, S.W. Pitt, A. Serganov, H.M. Al-Hashimi, J. Am. Chem. Soc. 125 (2003) 10530.
[181] S.W. Pitt, A. Majumdar, A. Serganov, D.J. Patel, H.M. Al-Hashimi, J. Mol. Biol. 338 (2004) 7.
[182] R. Brüschweiler, X. Liao, P.E. Wright, Science 268 (1995) 886.
[183] T. Ulmer, J. Werner, I.D. Campbell, Structure 10 (2002) 901.
[184] R. Varadan, O. Walker, C. Pickart, D. Fushman, J. Mol. Biol. 324 (2002) 637.
[185] D. Fushman, R. Varadan, M. Assfalg, O. Walker, Prog. NMR Spectosc. 44 (2004) 189.
[186] P.M. Hwang, N.R. Skrynnikov, L.E. Kay, J. Biomol. NMR 20 (2001) 83.
[187] E. de Alba, J.L. Baber, N. Tjandra, J. Am. Chem. Soc. 121 (1999) 4282.
[188] J.R. Tolman, H.M. Al-Hashimi, L.E. Kay, J.H. Prestegard, J. Am. Chem. Soc. 123 (2001) 1416.
[189] F. Tian, H.M. Al-Hashimi, J.L. Craighead, J.H. Prestegard, J. Am. Chem. Soc. 123 (2001) 485.
[190] L. Wang, Y. Pang, T. Holder, J.R. Brender, A.V. Kurochkin, E.R.P. Zuiderweg, Proc. Natl Acad. Sci. USA 98 (2001) 7684.
[191] W. Peti, J. Meiler, R. Brüschweiler, C. Griesinger, J. Am. Chem. Soc. 124 (2002) 5822.
[192] J.R. Tolman, J. Am. Chem. Soc. 124 (2002) 12020.
[193] K.B. Briggman, J.R. Tolman, J. Am. Chem. Soc. 125 (2003) 10164.
[194] C. Griesinger, Oral contribution, 45th Experimental NMR Conference, Asilomar (2004).
[195] J. Meiler, W. Peti, C. Griesinger, J. Am. Chem. Soc. 125 (2003) 8072.
[196] J.-C. Hus, R. Brüschweiler, J. Biomol. NMR 24 (2002) 123.
[197] J.-C. Hus, W. Peti, C. Griesinger, R. Brüschweiler, J. Am. Chem. Soc. 125 (2003) 5596.
[198] T.S. Ulmer, B.E. Ramirez, F. Delaglio, A. Bax, J. Am. Chem. Soc. 125 (2003) 9179.
[199] J.P. Derrick, D.B. Wigley, J. Mol. Biol. 243 (1994) 906.
[200] P. Bernado, M. Blackledge, J. Am. Chem. Soc. 126 (2004) 7760.
[201] J.B. Hall, D. Fushman, J. Biomol. NMR 27 (2003) 261.
[202] S. Meier, D. Häussinger, P. Jensen, M. Rogowski, S. Grzesiek, J. Am. Chem. Soc. 125 (2003) 44.
[203] P. Bernado, G. Bouvignies, M. Blackedge, J. Magn. Reson. (2005) in press.
[204] G.M. Clore, C.D. Schwieters, Biochemistry 43 (2004) 10678.
[205] I. Bertini, I.C. Felli, C. Luchinat, J. Biomol. NMR 18 (2000) 347.
[206] J.J. Chou, A. Bax, J. Am. Chem. Soc. 123 (2001) 3844.
[207] A. Mittermaier, L.E. Kay, J. Am. Chem. Soc. 123 (2001) 6892.
[208] R. Sprangers, M.R. Groves, I. Sinning, M. Sattler, J. Mol. Biol. 327 (2003) 507.