Journal of Biomolecular NMR **29:** 517–524, 2004.

Alignment of chain-like molecules

Martti Louhivuori[a], Kai Fredriksson[b], Kimmo Pääkkönen[c], Perttu Permi[c] & Arto Annila[a,b,c,*]
Departments of [a]Physical Sciences, [b]Biosciences, [c]NMR Laboratory, Structural Biology and Biophysics Program, Institute of Biotechnology, University of Helsinki, Finland

Received 5 January 2004; Accepted 30 March 2004

Key words: characteristic ratio, denatured protein, persistence length, random-flight chain, residual dipolar couplings, valence chain

Abstract

The steric obstruction model, that describes the enhanced alignment of folded proteins by anisotropic medium, is extended to account for the residual dipolar couplings of chain-like polypeptides. The average alignment of each chain segment is calculated from an ensemble of conformations represented by a spatial probability distribution. The segmental alignment depends on chain length, flexibility and segment's position in the chain. Residual dipolar couplings in turn depend on internuclear vector directions within each fragment. The results of calculations and simulations explain salient features of the experimental data. With this insight residual dipolar couplings can be interpreted to assess the degree of denaturation, local structures and spatial organization of weakly structured proteins.

Introduction

Residual dipolar couplings (RDCs) are particularly informative NMR parameters for studies of biological macromolecules in solution (Tjandra and Bax, 1997; Mueller et al., 2000; Prestegard et al., 2000; Bax et al., 2001; De Alba and Tjandra, 2002). Most notably RDCs relate to orientations of chemical bonds with respect to a molecular coordinate system. RDCs have been used to determine and refine three-dimensional structures, to build models of macromolecular complexes from known sub-units (Clore, 2000; Fischer et al., 1999; Mattinen et al., 2002), to construct three-dimensional protein models from protein fragments deposited in the PDB (Delaglio et al., 2000; Andrec et al., 2001), to study spatial arrangements of modular proteins (Fischer et al., 1999; Chou et al., 2001), and to recognize (Annila et al., 1999) and classify protein folds (Valafar and Prestegard, 2003). For diamagnetic proteins RDCs are intrinsically small (Tjandra et al., 1996) but become readily observable and applicable when anisotropic molecular tumbling is enhanced by liquid crystal particles (Tjandra and Bax, 1997; Meier et al., 2002) or axial polymer matrices (Tycko et al., 2000). The enhanced alignment due to the obstruction is a well-understood phenomenon for structured macromolecules (Zweckstetter and Bax, 2000; Bax et al., 2001; Fernandes et al., 2001), i.e., for folded proteins.

Recently RDCs were also recorded from denatured proteins (Shortle and Ackerman, 2001; Ackerman and Shortle, 2002; Ohnishi and Shortle, 2003). However, the first interpretations were based on an incorrect implicit assumption that only a molecule with a defined three-dimensional architecture may give rise to non-vanishing RDCs. It was subsequently shown through an analytical calculation that the mere structure of a random-flight chain is sufficient to yield non-zero RDCs (Louhivuori et al., 2003). To progress in interpreting RDCs acquired from denatured proteins we have formulated a valence chain model to account for steric hindrance between residues and simulated chains of varying flexibility.

*To whom correspondence should be addressed. E-mail: arto.annila@helsinki.fi

Alignment of a chain

A denatured protein, i.e., an unstructured and flexible molecule, obviously differs from the single-conformation rigid-body case as there are numerous conformations present at any given time in the solution. Here we limit the analysis to conformations that are in fast exchange with each other. This is the prerequisite to acquire NMR observables, because otherwise resonance lines would be expected to broaden beyond detection. We assume that the polypeptide experiences a purely steric obstruction effect in a liquid crystal or an axial matrix and that the obstruction itself does not induce conformational changes.

Consider a rigid segment denoted here with i that holds two nuclei A and B which will give rise to residual dipolar coupling D^i_{AB}

$$D^i_{AB} = D^{\max}_{AB}\overline{\langle P^i_2(\cos\theta)\rangle}. \quad (1)$$

The first term $D^{\max}_{AB}$ is a constant that includes the gyromagnetic ratios γ_A and γ_B, the length of the internuclear vector r_{AB}, the Planck constant h and the magnetic permiability μ_o. The second term expresses averaging of Legendre polynomial $P_2 = {}^1/_2(3\cos^2\theta - 1)$ where θ is the angle between $\mathbf{r}_{AB}$ and the external magnetic field $\mathbf{B}_o$. The brackets $\langle\rangle$ denote the averaging due to the molecular tumbling in the solution and the overbar $\overline{\ }$ stands for the average over the ensemble of conformations.

To calculate weighted averages over the conformations we introduce a spatial probability density to describe the ensemble *about* the segment i. We denote the spatial distribution of the segments n preceding i with W_n and likewise the spatial spread of the segments m succeeding i by W_m. Obviously the segment i itself aligns as well. It is a rigid body associated with a probability which is unity within i and zero outside (Fernandes et al., 2001). The total distribution W_i is a weighted sum of the probabilities. Next, we note that when the rigid segment i of a finite length l moves or reorients itself, also W_n and W_m translate or rotate in space but do not change otherwise. However, in the presence of an obstruction, e.g. a plane-barrier, W_i is not invariant anymore as the particular conformations that collide with the barrier are excluded from the solution. Consequently the concentration diminishes towards the barrier placed at a position h_z (Figure 1). Most importantly the probability density is a function of the orientation θ, i.e. $W^{obs}_i(x, y, z, h_z, \theta)$. The tiny orientation-dependent concentration effect is the source of non-vanishing RDCs even for the random-flight chain. The effect is larger for stiffer chains.

Above we have implicitly assumed that $\mathbf{r}_{AB}$ is collinear with $\mathbf{i}$, which gives the orientation of the segment, and that $\mathbf{B}_o$ is also parallel with $\mathbf{n}$, the director of the anisotropic medium. This specific geometry is convenient but of course not a general condition. For each case the explicit relationship between an RDC value and P_2 can be determined from the covalent geometry of the fragment, and the type of liquid crystal in question dictates whether $\mathbf{n} \parallel \mathbf{B}_o$ or $\mathbf{n} \perp \mathbf{B}_o$. Therefore we proceed to calculate the average alignment for each segment rather than values of RDCs for particular internuclear vectors. The normalized integration over the space where the conformations roam gives the alignment for each segment i

$$\overline{\langle P^i_2\rangle} = \int P^i_2(\cos\theta)\, W^{obs}_i(x, y, z, h_z, \theta)\, d\Omega. \quad (2)$$

Note that the distribution, W^{obs}_i, that incorporates the obstruction inflicted upon W^{free}_i by the barrier, is used in the calculation (Equation 2), instead of the intact free distribution of conformations, W^{free}_i. Therefore it may at first seem as if RDCs would report from perturbed conformations, which is not the case. Instead, it should be understood that RDCs arise from the 'examination' of W^{free}_i by the barrier as conformational changes are explicitly excluded in the definition of W^{obs}_i and only differences in the concentrations of conformations are allowed.

Valence chain model

The steric hindrance between residues in polypeptides limits the main chain Φ and Ψ torsions to allowed regions in the Ramachandran plot. To take into account the restricted degrees of freedom by a simple model the polypeptide chain is customarily described as a valence chain where a subsequent segment is constrained to a cone by the valence angle. For an average amino acid heteropolymer the valence angle is about 36 degrees (Cantor and Schimmel, 1998). Owing to the free rotations about azimuth angles the original direction of i is lost after a certain distance referred to as the persistence length λ. It varies in the polypeptide depending on the residue types but spans approximately five residues i.e., $\lambda \approx 20$ Å. A polyglycine is the softest chain with $\lambda = 6$ Å and a polyproline chain is very stiff with $\lambda = 220$ Å (Cantor and Schimmel, 1998).

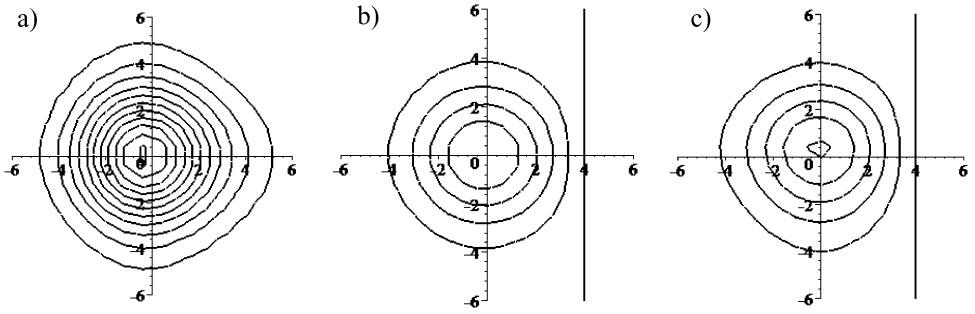

Figure 1. Two-dimensional projections of an eleven segment random-flight chain distribution (a), in the vicinity of a plane-barrier when the segment *i* points towards (b) and along (c) the plane-barrier at the distance equivalent to +4 segments. The distribution is viewed from the frame of the third segment *i* that is placed in the origin. The contours relate to concentration. Note the minute differences between (b) and (c) due to the different orientations of *i*.

In the presence of a plane-barrier at a distance h_z with its normal **n** along z-axis || $\mathbf{B}_o$ we define the valence chain model by the normalized distribution W_i^{obs} (Chadrasekhar, 1943; Louhivuori et al., 2003) as

$$W_i^{obs}(z, h_z, \theta, n, N) = \frac{n}{N+1} W_n^{obs} + \frac{m}{N+1} W_m^{obs}$$
$$= \frac{n}{N+1}\left\{W_n^{free} - W_n^{bar}\right\} + \frac{m}{N+1}\left\{W_m^{free} - W_m^{bar}\right\},$$
$$W_n^{free} = \sqrt{\frac{2}{\pi C_n n l^2}} \exp\left[-\left(z + \frac{l_n^e}{2}\cos\theta\right)^2 / 2C_n n l^2\right],$$
$$W_n^{bar} = \sqrt{\frac{2}{\pi C_n n l^2}} \exp\left[-\left(2h_z - z + \frac{l_n^e}{2}\cos\theta\right)^2 / \right. \quad (3)$$
$$\left. 2C_n n l^2\right],$$

$$C_n = 1 + \frac{2}{n}\sum_{j=1}^{n}\prod_{k=j}^{n}\alpha_k, \alpha_k = \cos\phi_k,$$
$$l_n^e = (C_n + 1)l/2, N = n + m.$$

The value n is the number of segments in the half-chain before the locus i and $m = N - n$ is the number of segments succeeding i. Thus the total number of segments including i equals $N + 1$. The varying flexibility of heteropolymers is introduced for the two half-chains by the characteristic ratios C_n and C_m, which are computed from the locus dependent valence angles ϕ_k analogously to the sum of geometric series for homopolymers (Cantor and Schimmel, 1998). The root mean square length $\langle r_o^2\rangle = C_n n l^2$ is a measure of the spatial extent of the half-chain. We use l_e to express the effective stiffness at the position i that displaces the half-chain distributions from each other. The definitions of C_n and C_m reduce to unity and the effective length l_e reduces to l for the random-flight chain (Louhivuori et al., 2003).

The integration is performed only along the z-axis because the dimensions along x- and y-axis are in our model not obstructed and not correlated to the distribution along z. The position of i is integrated from $-L + h_z$ to h_z, where L is the distance between two obstructions, and h_z is integrated to L from the position where it just makes a contact with l_e. The θ angle is integrated from 0 to 2π using the standard sampling $d\Omega = \sin\theta d\theta$. The integration of exponential distribution functions extends over the total volume to yield directly $\langle P_2\rangle$. For the statistical element the integration is confined to the restricted volume near the barrier (Fernandes et al., 2001) where the concentration is proportional to $l_e(1 - |\cos\theta|)$.

Results from homopolymer valence chain models

We have computed alignments of valence chain segments using Equations 1–3. First we have examined the effect of chain length and flexibility by calculating alignments for various homopolymers. The flexibilities were chosen to match the average of a heteropolymer chain $\alpha_k = 0.627$, polyglycine $\alpha_k = 1.195$ and polyproline chains $\alpha_k = 0.186$ obtained from the reported persistence lengths (Cantor and Schimmel, 1998). We also show results for the random flight-chain, i.e. $\alpha_k = 0$. For comparison we have simulated the corresponding alignments using the PALES program (Zweckstetter and Bax, 2000). An ensemble of 16384 valence chain conformations were generated by a Python script and submitted for the simulations. The average RDC for every segment in the chain was converted to the average alignment $\langle P_2\rangle$ dividing by $D_{AB}^{\max}$ as $\mathbf{r}_{AB}$ was chosen parallel to the segment.

First, for all chains the $\langle P_2\rangle$ values increase from the chain termini to the central residues (Figure 2). This reflects the fact that the locus-dependent distribu-

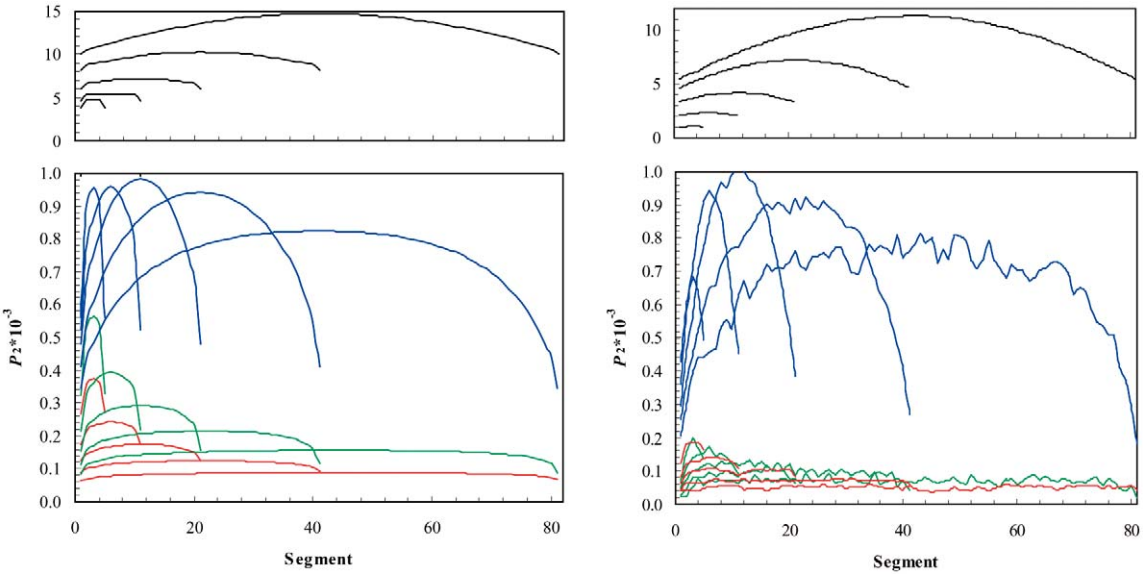

Figure 2. Calculated (left) and simulated (right) alignments for valence chains comprising $N = 5$, 11, 21, 41 and 81 segments. The valence angles correspond to an average polypeptide (blue), polyglycine (green), polyproline (black in separate windows above) and random-flight chain (red). The bicell concentration was 5% w/v that was taken to correspond to a 760 Å separation of the planes in the calculation. The 'noise' in the simulated data is due to the insufficient sampling of conformational space.

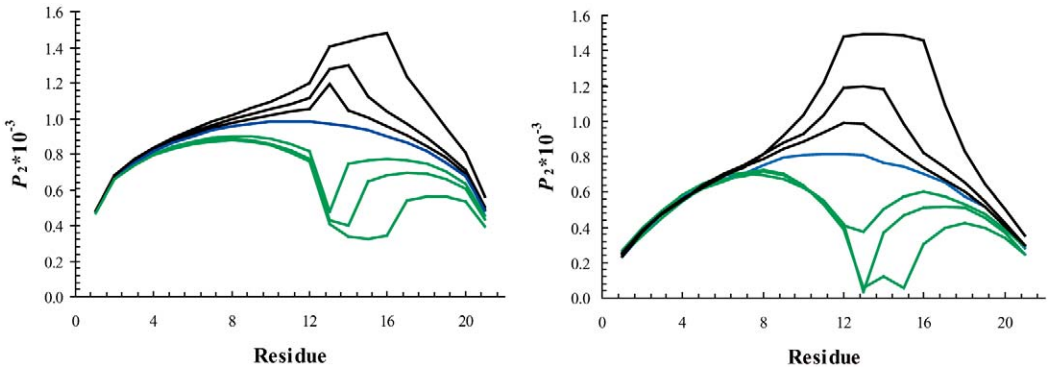

Figure 3. Alignment of 21-mer chains obtained from calculations (left) and valence chain simulations (right). The homopolymer (blue) has the flexibility of an average polypeptide. Variation in flexibility was introduced by incorporating proline (black) and glycine (green) residues at the positions 13, 13 and 14, and 13–16. The distance between the planar obstructions was 760 Å and the bicell concentration was 5%.

tion is more elongated at the middle of the chain than at the termini. Second, stiffness leads to an increase in the alignment as expected. Third, the maximum alignment is obtained when the valence chain ensemble has reached its maximally elongated shape. Thereafter as the chain becomes longer and longer the distribution becomes more and more spherical and the values decrease. The alignment of average polypeptide does not display substantial chain length dependence about the broad maximum around a 20-mer. The homoproline valence chain is so stiff that not even 80 residues are enough to reach the maximum. In fact the homoproline P_2 is dominated by the alignment of the effective length l_e as shown by the calculation (Figure 2). The homoglycine valence chain is so soft that it takes less than five residues to reach the maximum alignment.

There is a qualitative agreement between the alignments obtained from the calculations and simulations as a function of the chain length, residue position, and flexibility. Especially for the most relevant case, i.e. the average chain, the correspondence is good. However, the following reasons could account for the discrepancies between the calculated and simulated values. The exponential distribution function used in the calculation does not represent accurately the bi-

nomial nature of the short chains. The tails of the exponentials reach further than the actual chain. Furthermore, to model the flexibility at the locus i we displaced the half-chain distributions by l_e instead of attempting to describe the true course of the chain. This is a practical but inadequate approximation in particular for stiff chains. Clearly, in the limit of a rigid object the exponential distribution function does not give the step function at the surface of the object (Fernandes et al., 2001). Finally, it should be noted that in the calculations the chain segments were implicitly taken as infinitely thin whereas in the simulations the segments had a finite thickness.

Results from heteropolymer valence chain models

To inspect the varying flexibility of a 21-residue polypeptide we replaced one, two and four consequent residues at the positions 13–16 with prolines and with glycines. The valence chain calculations were carried out as before. Ensembles of 16384 chain conformations were submitted to PALES. The simulated RDC values relate to P_2 by a constant that includes $D^{\max}$ and the orientation of the bond that was set parallel to the segment (Figure 3).

The homoproline chain has more than tenfold higher C_∞ than any other amino acid homopolymer. Indeed the stiffening of the chain at the proline site manifests as an increased alignment both in the calculations and simulations. Nonetheless, the increase remains perhaps surprisingly small. The reason is that e.g. the characteristic ratio for four consecutive prolines, i.e. C_4, differs only by a factor of two from that of the average polypeptide. In longer chains the relative difference is smaller. As a result the spatial rms-dimensions of the ensemble, which has a square-root dependence on C_n, does not increase that much compared with the average polymer and the alignments increase only a little elsewhere in the chain. Analogous reasoning accounts for the decreased alignment at the position of glycines. It takes only two to three glycines to lower the alignment nearly to the level of a homoglycine (Figure 2). However, our simple valence chain model does not take into account excluded volume effects. The differences between the calculated and simulated values stem from the use of exponential functions as before and from the failure to describe the actual course of the chain.

Results from all-atom models

To relate the valence chain results to more realistic models we carried out all-atom simulations with the corresponding 21-mers. For all-atom polypeptide models it is instructive to visualize the ensemble of superimposed conformations by computing probability density maps using VMD-Xplor (Schwieters and Clore, 2002) (Figure 4).

Isodensity surfaces are coarse because the spatial sampling by only 50 conformations is sparse. Nonetheless the maps serve to illustrate the locus-dependent distributions. The probability density map for the most central residue shows an elongated form which accounts for the larger alignment in the middle of the chain whereas the more spherical map for the terminal residue is consistent with the smaller alignment. It is also conceivable that the fragmental alignments may exhibit rhombicity that is not included in our chain model.

An ensemble of 16384 21-mer all-atom conformations was generated using CYANA (Günthert et al., 1997) and relaxed to remove steric conflicts. The family of conformations was then submitted to PALES and the average RDCs were calculated. We simulated one-bond RDCs for NH, HACA, NCO and COCA (Figure 5). The values relate to $\langle P_2 \rangle$ by a factor that consists of $D^{\max}$ and the cosine squared of the average orientation of a particular bond with respect to the fragmental alignment. The couplings from the homopolymer were normalized to D_{NH} by the gyromagnetic ratios and bond lengths.

The RDC data from the all-atom simulations of the glutamate homopolymer follow mostly curves similar to those calculated and simulated for the valence chains but contain noteworthy details. First, it is evident that the various bonds probe various average orientations. A mere inspection of the values reveals that the HACA bonds are mostly perpendicular whereas the COCA bonds are mostly parallel with respect to the average local course of the chain. The NCO bonds are closest to the magic angle where RDCs vanish. Second, the various couplings probe the alignment *within* a residue. This is easiest to notice when comparing e.g. the values at the chain termini. The NH values at the N-terminus are smaller than at the C-terminus because within the first residue the amide pair is closer to the chain terminus than in the last residue where it is followed by the alpha carbon, side chain and carboxyl group. Analogously, the COCA values are larger in the N-terminal than in the C-

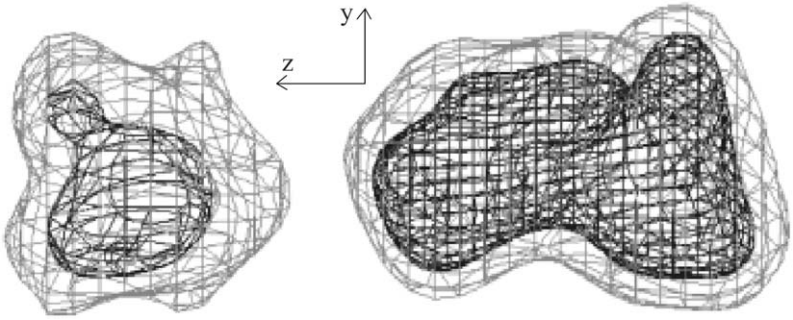

Figure 4. Probability density map of 50 conformations of 21-mer homoglutamate superimposed on the backbone atoms of residue number 1 (left) and 11 (right).

terminal residue. The corresponding HACA values are nearly equal because alpha carbons are in the center of a residue.

The heteropolymer containing prolines displays increased HACA couplings consistent with the increase in the alignment expected on the basis of the valence chain models (Figure 3). This stiffness, which results in a smaller dispersion of conformations, may be the very property why proline-rich sequences appear in molecular recognition motifs (Zarrinpar et al., 2003). However, the COCA and NCO couplings display large variation that must arise from a zig-zag course of the chain. This is not easily realized from the HACA data alone. Obviously an analysis of local structures would benefit from a number of RDCs acquired per residue (Permi et al., 2000).

Perhaps the most flagrant dissimilarity between the all-atom and valence chain models is that the RDCs from the glycine-rich region do not show diminished alignment. The COCA data follows nearly the homopolymer values and the other RDCs show even increased amplitudes. To explain this we reason that the particular conformations with tight turns would diminish the alignment most. However, even if glycines allow for tight turns, the half-chains may exclude each other. Consequently, these conformations are not abundant in the all-atom ensemble and the alignment does not drop. The variation in the RDC values compared to homopolymer may reflect that glycines allow for average bond orientations not present in the glutamate homopolymer. The similar values for the two HACA bonds of glycine residues indicate that they are symmetric about the average local course of the chain.

From the all-atom simulations we infer that the alignment varies rather gradually over the polypeptide chain even in the presence of most rigid and most flexible fragments due to prolines and glycines. Therefore any abrupt variation in the average internuclear directions for adjacent residues is the primary source of sudden deviations in RDCs.

Discussion

In this study it is shown that the spatial distribution of the conformational ensemble and directions of the internuclear vectors in molecular fragments dictate the values of the residual dipolar couplings. This implies that it might be possibile to interpret RDCs from an unstructured protein in terms of an ensemble shape and a local structure. However, a mere inspection of the functional form of RDC (Equation 2) reveals that the interpretation is an underdetermined inverse problem. In practice it is difficult to acquire abundantly non-redundant RDCs per a rigid fragment i.e., the peptide plane or C_α-center to determine accurately the average fragmental alignment frames.

Fortunately, even in the case of an unstructured polypeptide, the inherent difficulties in the inverse problem are largely compensated by the properties of the polypeptide chain itself. The fact that there are about three degrees of freedom per a polypeptide residue is often quoted to result in an astronomical number of degrees of freedom for the entire molecule (Levinthal, 1969). To us the small number of degrees of freedom per residue implies that submolecular fragments are coupled to each other. Presumably the resulting stiffness is a valuable property for a polypeptide to adopt well-defined protein folds. For the interpretation of RDCs the stiffness implies that the consecutive fragmental alignment frames do not change abruptly. The concept of semi-continuous frame is supported by the examples presented above.

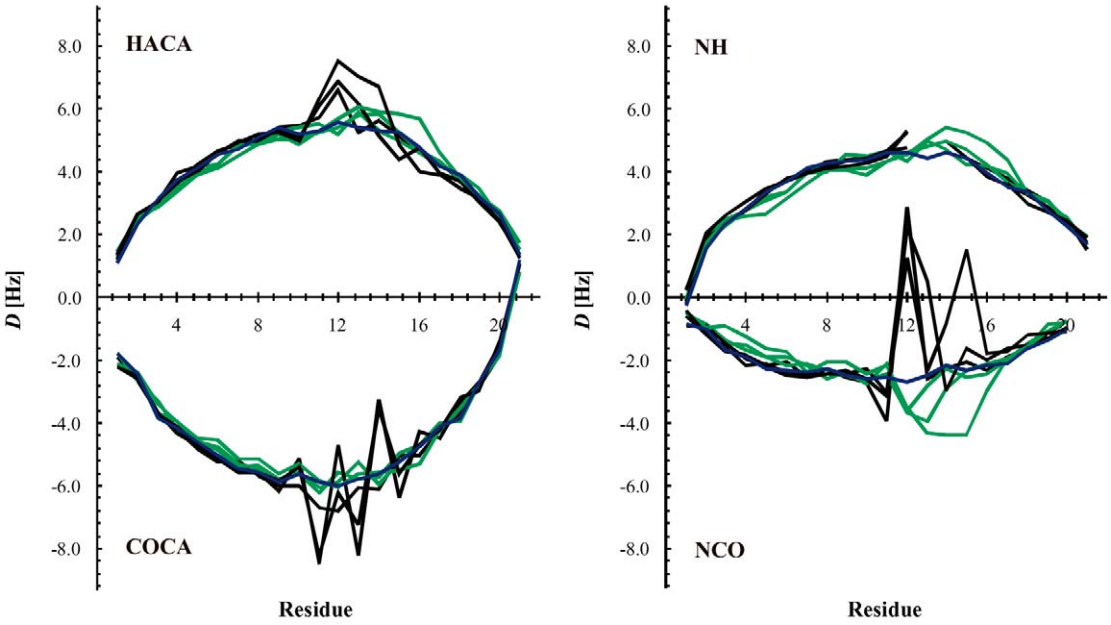

Figure 5. RDCs of 21-residue polymers obtained from all-atom simulations; homoglutamate (blue) and heteropolymers containing proline residues (black) and glycine residues (green). The HACA and COCA (left) and NH and NCO (right) values were scaled by the absolute value of gyromagnetic ratios and bond lengths to correspond the amide nuclear pair. The bicell concentration was 5%.

Therefore any sudden change in the value of a particular RDC in any adjacent residues is mostly due to changes in internuclear vector orientations just as it is in the case of a folded protein. This leads to a gratifying conclusion. To a first approximation RDCs from neighboring residues, with appropriate weights, i.e., α_k, can be used to estimate the size of the alignment and to construct average fragmental alignment frames to which the average bond directions can be related to.

The ensemble shape, *viewed from a locus*, is a surprisingly useful notion to calculate ensemble averages although it is obviously a very poor description for any individual chain. The most important characteristic of the ensemble is its elongated form. The magnitude of alignment is determined by the overall size and shape of the ensemble about the locus. It is conceivable that computer simulations will serve to create families of conformations and to refine them to match measured RDCs. Such a search for a representative ensemble would certainly benefit from additional data, e.g. from diffusion (Choy et al., 2002), relaxation (Schwalbe et al., 1997), scattering (Garcia et al., 2001) or fluorescence resonance energy transfer measurements, to constrain the distribution.

In summary, residual dipolar couplings provide valuable information from denatured proteins or weakly structured polypeptide chains, e.g., conformational ensembles of amyloidogenic peptides, with the detail of an amino acid and beyond. The interpretation of RDCs is based on the following guidelines. The overall shape of the conformational ensemble, perceived at the locus where an RDC is measured, largely determines the magnitude of the alignment. The alignments of neighboring fragments vary smoothly. Variation in RDCs is mainly due to deviations in the internuclear vector directions and to lesser extent due to variation in the chain flexibility. Abrupt changes in RDCs between adjacent corresponding internuclear vectors should be attributed to differences in internuclear vector orientations rather than to substantially different fragmental alignments. However, more detailed data than published to date (Ackerman and Shortle, 2002; Ohnishi and Shortle, 2003) and preferably from well-characterized polymers will be needed to evaluate the details of the presented results. For example, the diminishing alignment at the chain termini, which is inherent in the presented analytical and simulated models that do not allow for conformational changes remains to be challenged.

Acknowledgement

This work has been supported by a research grant from the Academy of Finland.

References

Ackerman, M.S. and Shortle, D. (2002) *Biochemistry*, **41**, 13791–13797.

Andrec, M., Du, P. and Levy, R.M. (2001) *J. Biomol. NMR*, **4**, 335–347.

Annila, A., Aitio, H., Thulin, E. and Drakenberg, T. (1999) *J. Biomol. NMR*, **14**, 223–230.

Bax, A., Kontaxis, G. and Tjandra, N. (2001) *Meth. Enzymol.*, **339**, 127–174.

Cantor, C.R. and Schimmel, P.R. (1998) *Biophysical Chemistry*, Freeman, New York.

Chadrasekhar, S. (1943) *Rev. Mod. Phys.*, **15**, 1–89.

Chou, J.J., Li, S., Klee, C.B. and Bax A. (2001) *Nat. Struct. Biol.*, **11**, 990–997.

Choy, W.Y., Mulder, F.A., Crowhurst, K.A., Muhandiram, D.R., Millett, I.S., Doniach, S., Forman-Kay, J.D. and Kay, L.E. (2002) *J. Mol. Biol.*, **316**, 101–112.

Clore, M. (2000) *Proc. Nat. Acad. Sci. USA*, **97**, 9021–9025.

De Alba, E. and Tjandra, N. (2002) *Prog. NMR Spectrosc.*, **40**, 175–197.

Delaglio, F., Kontaxis, G. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 2142–2143.

Fernandes, M.X., Bernado, P., Pons, M. and Garcia de la Torre, J. (2001) *J. Am. Chem. Soc*., **123**, 12037–12047.

Fischer, M.W., Losonczi, J.A., Weaver, J.L. and Prestegard, J.H. (1999) *Biochemistry*, **38**, 9013–9022.

Garcia, P., Serrano, L., Durand, D., Rico, M. and Bruix, M. (2001) *Protein Sci.*, **6**, 1100–1112.

Günthert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.

Levinthal, C. (1969) *Univ. Illinois Bull.* 67, **41**, 22–24.

Louhivuori, M., Pääkkönen, K., Fredriksson, K., Permi, P., Lounila, J. and Annila, A. (2003) *J. Am. Chem. Soc.*, **125**, 15647–15650.

Mattinen, M.-L., Pääkkönen, K., Ikonen, T., Craven, J., Drakenberg, T., Serimaa, R., Waltho, J. and Annila, A. (2002) *Biophys. J.*, **83**, 1177–1183.

Meier, S., Haussinger, D. and Grzesiek, S. (2002) *J. Biomol. NMR*, **24**, 351–356.

Mueller, G.A., Choy, W.Y., Yang, D., Forman-Kay, J.D., Venters, R.A. and Kay, L.E. (2000) *J. Mol. Biol.*, **300**, 197–212.

Ohnishi, S. and Shortle, D. (2003) *Proteins*, **50**, 546–51.

Permi, P., Rosevear, P.R. and Annila, A. (2000) *J. Biomol. NMR*, **17**, 43–54.

Schwalbe, H., Fiebig, K.M., Buck, M., Jones, J.A., Grimshaw, S.B., Spencer, A., Glaser, S.J., Smith, L.J. and Dobson, C.M. (1997) *Biochemistry*, **36**, 8977–8991.

Schwieters, C.D. and Clore, G.M. (2002) *J. Biomol. NMR*, **23**, 221–225.

Shortle, D. (2002) *Adv. Protein Chem.*, **62**, 1–23.

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

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

Tjandra, N., Grzesiek, S. and Bax, A. (1996) *J. Am. Chem. Soc.*, **118**, 6264–6272.

Tycko, R., Blanco, F.J. and Ishii, Y. (2000) *J. Am. Chem. Soc.*, **122**, 9340–9341.

Valafar, H. and Prestegard, J.H. (2003) *Bioinformatics*, **19**, 1549–1555.

Zarrinpar, A., Bhattacharyya, R.P. and Lim, W.A. (2003) *Sci. STKE*, **179**, re8.

Zweckstetter, M. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 3791–3792.