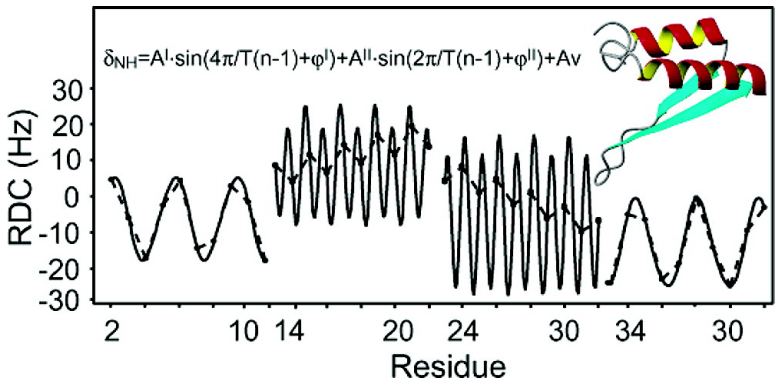
Article

Theoretical Analysis of Residual Dipolar Coupling Patterns in Regular Secondary Structures of Proteins

Alessandro Mascioni, and Gianluigi Veglia

J. Am. Chem. Soc., **2003**, 125 (41), 12520-12526• DOI: 10.1021/ja0354824 • Publication Date (Web): 18 September 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 5 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 09/18/2003

Theoretical Analysis of Residual Dipolar Coupling Patterns in Regular Secondary Structures of Proteins

Alessandro Mascioni and Gianluigi Veglia*

Contribution from the Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455

Received April 5, 2003; E-mail: veglia@chem.umn.edu.

Abstract: A new approach to the interpretation of residual dipolar couplings for the regular secondary structures of proteins is presented. This paper deals with the analysis of the steric and chiral requirements of protein secondary structures and establishes a quantitative correlation between structure periodicity and the experimental values of the backbone residual dipolar couplings. Building on the recent interpretation of the periodicity of residual dipolar couplings in α-helices (i.e., "dipolar waves"), a general parametric equation for fitting the residual dipolar couplings of any regular secondary structure is derived. This equation interprets the modulation of the residual dipolar couplings' periodicity in terms of the secondary structure orientation with respect to an arbitrary reference frame, laying the groundwork for using backbone residual dipolar couplings as a fast tool for determining protein folding by NMR spectroscopy.

Introduction

The natural[1] or induced[2] anisotropy of nuclear spin interactions permits the measurements of residual dipolar couplings (RDCs), providing long-range structural information on the relative orientations of secondary structure elements in biomacromolecules. Since their introduction, RDCs have found several different applications in structural biology. In particular, RDCs have been used for structure refinement as harmonic constraints,[3,4] the validation of NOE-based NMR and X-ray structures,[5] de-novo structure calculation in conjunction with pseudocontact shifts,[6] protein domain orientations and dynamics,[7,8] protein−protein, and protein−ligand binding studies[9−11] and, finally, the identification of folds using homology models or molecular fragment replacement.[12]

Dipolar couplings (full or residual) are diagnostic indicators of the intrinsic periodicity found in protein structures. The periodic patterns of dipolar couplings were first observed in helical membrane proteins embedded in oriented lipid bilayers studied by solid-state NMR[13−15] and, in a recent report, Mesleh et al.[16] showed that these distinctive patterns can also be found in weakly aligned samples of α-helical proteins. These patterns, named "dipolar waves", are reminiscent of helical patterns obtained by EPR spin-label experiments[17] and by NMR paramagnetic quenching using O_2[18] and can be fitted by simple sinusoids of periodicity 3.6. The periodicity of dipolar waves in both solution and solid-state NMR spectra is related to the secondary structure type, whereas the amplitude and the average of the sinusoids can be directly linked to the orientation of the helical domains with respect to the magnetic field.[16,19] This observation opens up the possibility of exploiting backbone residual (or full) dipolar couplings for mapping both proteins' secondary and tertiary structures through nonlinear fitting of RDC data as a function of the residue number. Given the availability of several new methods for weakly aligning both soluble[20] and membrane proteins,[21−23] this approach would provide a valuable tool for rapidly determining protein domain orientations.

Because RDCs are very sensitive probes for both local geometries and domain orientations, a correct fitting procedure for the experimental data is crucial to extract both secondary and tertiary information about protein structures. In the present article, we present a parametric equation that describes the oscillation of the residual dipolar couplings associated with any

(1) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283.
(2) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(3) Tjandra, N.; Omichinski, A. M.; Gronenborn, M. A.; G., C. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732−738.
(4) Tjandra, N.; Marquardt, J.; G., C. M. *J. Magn. Res.* **2000**, *142*, 393−396.
(5) Skrynnikov, N. R.; Goto, N. K.; Yang, D.; Choy, W. Y.; Tolman, J. R.; Mueller, G. A.; Lewis, E. K. *J. Mol. Biol.* **2000**, *295*, 1265−1273.
(6) Hus, J. C.; Marion, D.; Blackledge, M. *J. Mol. Biol.* **2000**, *295*, 927−936.
(7) Prestegard, J. H.; Al-Hashimi, H. M.; Tolman, J. R. *Q. Rev. Biophys.* **2000**, *33*, 371−424.
(8) Fischer, M. W. F.; Losonczi, J. A.; Weaver, J. L.; Prestegard, J. H. *Biochemistry* **1999**, *38*, 9013−9022.
(9) Clore, M. G.; Gronenborn, M. A. *J. Magn. Res.* **2002**, *154*, 329−335.
(10) McCoy, M. A.; Wyss, D. F. *J. Am. Chem. Soc.* **2002**, *124*, 2104−2105.
(11) Koenig, B. W.; Kontaxis, G.; Mitchell, D. C.; Louis, J. M.; Litman, B. J.; Bax, A. *J. Mol. Biol.* **2002**, *322*, 441−461.
(12) Delaglio, F.; Kontaxis, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 2142−2143.
(13) Marassi, F. M.; Opella, S. J. *J. Magn. Res.* **2000**, *144*, 150−155.
(14) Wang, J.; Denny, J.; Tian, C.; Kim, S.; Mo, Y.; Kovacs, F. A.; Song, Z.; Nishimura, K.; Gan, Z.; Fu, R.; Quine, J. R.; Cross, T. A. *J. Magn. Reson.* **2000**, *144*, 162−167.
(15) Marassi, F. M. *Biophys. J.* **2001**, *80*, 994−1003.
(16) Mesleh, M. F.; Veglia, G.; DeSilva, T. M.; Marassi, F. M.; Opella, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 4206−4207.
(17) Altenbach, C.; Marti, T.; Khorana, H. G.; Hubbel, W. J. *Science* **1990**, *148*, 1088−1092.
(18) Luchette, P. A.; Prosser, R. S.; Sanders, C. R. *J. Am. Chem. Soc.* **2002**, *124*, 1778−1781.
(19) Mesleh, F. M.; Lee, S.; Veglia, G.; Thiriot, D. S.; Marassi, F. M.; Opella, S. J. *J. Am. Chem. Soc.* **2003**, *125*, 8928−8935.
(20) Bax, A.; Kontaxis, G.; Tjandra, N. *Methods Enzymol.* **2001**, *339*, 127−174.

10.1021/ja0354824 CCC: $25.00

periodic secondary structure. Rigorous theoretical treatment was used to derive this general formula that describes the value of the residual dipolar couplings for each residue of the polypeptide chain as a function of residue number. By using this approach, it is possible to obtain a precise and quantitative fit for experimental dipolar patterns. Although our mathematical model can be applied to polypeptide (or polymer) conformations of any periodicity, we show the fitting for α-helices, β-strands, 3_{10}-helices, and π-helices.

Theoretical Basis

The regular structures of proteins can be described in terms of repeating monomeric-unit conformations by specifying Cartesian or polar coordinates with respect to an external reference frame. Eyring transformation matrixes[24] have proved to be a useful method to describe the structural parameters of regular secondary structures of macromolecules.[25−28] Because the value of the residual dipolar couplings depend only on the orientation of the coupled nuclei with respect to the direction of the static magnetic field, RDCs can be interpreted in a similar way, using rotation matrixes in polar coordinates.

If we assume that a periodic element of a secondary structure (an α-helix, 3_{10}-helix, π-helix, or β-strand) is initially oriented with its principal axis aligned along the *Z*-axis of an arbitrary Principal Axis System (PAS), which we will in turn assume to be parallel to the direction of the static magnetic field, then the components of the first amide bond vector, $\vec{r}^{\,1}_{NH}$, can be described as

$$\vec{r}^{\,1}_{NH} = \begin{pmatrix} d_{NH}\cdot\sin(\delta)\cdot\cos(\rho) \\ d_{NH}\cdot\sin(\delta)\cdot\sin(\rho) \\ d_{NH}\cdot\cos(\delta) \end{pmatrix}$$

where ρ is the angle that the projection of the $\vec{r}^{\,1}_{NH}$ vector on the X−Y plane makes with the *X* axis, d_{NH} is the proton-nitrogen bond vector (∼1.07 Å[29]), and δ is the angle that the $\vec{r}^{\,1}_{NH}$ vector makes with the chain axis. Also, ρ represents the angle of rotation of the entire chain around its axis. Figure 1 depicts these parameters for an ideal α-helix.

The components of the n^{th} vector ($\vec{r}^{\,n}_{NH}$) can be described by rotating the $\vec{r}^{\,1}_{NH}$ bond around the *Z* axis of an angle $\alpha = 2\pi(n-1)/T$, where *T* is the periodicity of the secondary structure under analysis. This operation can be accomplished by the matrix transformation

$$\vec{r}^{\,n}_{NH} = \mathscr{R}_Z(\alpha)\cdot\vec{r}^{\,1}_{NH}$$

where $\mathscr{R}_Z$ is a right-handed rotation matrix around the *Z* axis. For an ideal α-helix oriented at θ and ϕ degrees with respect to the PAS, the orientation of each $\vec{r}_{NH}$ bond vector can be determined by applying additional rotations of θ and ϕ degrees with respect to the *Y*- and *Z*-axis, respectively. This is accomplished by the following matrix transformation

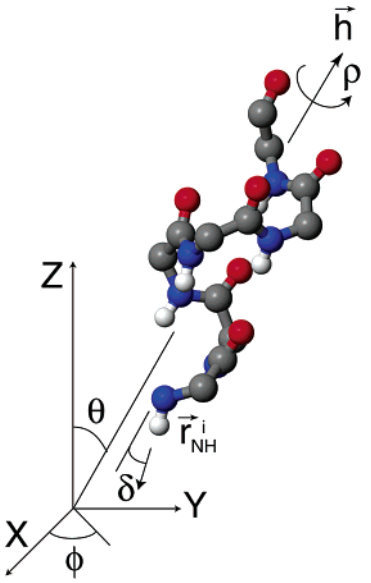


Figure 1. Schematic representation of the geometric parameters which are used to define the dipolar coupling as a function of the secondary structure element orientation. The components of each $\vec{r}_{NH}$ bond in the polypeptide chain are determined by rotating the first amide bond around the helix axis n−1 times the repeat angle (100° for α-helices ∼180° for β-strands).

$$\vec{r}^{\,n}_{NH}* = \mathscr{R}_Z(\phi)\cdot\mathscr{R}_Y(\theta)\cdot\vec{r}^{\,n}_{NH}$$

where $\mathscr{R}_Y$ and $\mathscr{R}_Z$ are the right-handed rotation matrixes that rotate the $\vec{r}^{\,n}_{NH}$ vector by θ and ϕ around the *Y*- and the *Z*-axis, respectively. The symbol * indicates that the vector components are now referred to an orientation of the secondary structure element at θ and ϕ angles with respect to the PAS.

Starting from the component of the first $\vec{r}_{NH}$ bond vector, the overall transformation can be described by the equation

$$\vec{r}^{\,n}_{NH}* = \mathscr{R}_Z(\phi)\cdot\mathscr{R}_Y(\theta)\cdot\mathscr{R}_Z(\alpha)\cdot\vec{r}^{\,1}_{NH} \qquad (1)$$

Therefore, eq 1 provides the components of each $\vec{r}_{NH}$ vector with respect to the PAS as a function of the components of the first amide bond. Using this approach, all the amide bond vectors within the same secondary structure element can be univocally described.

The magnitude of the residual dipolar coupling as a function of the components of the $\vec{r}_{NH}$ bond vector can be expressed in Cartesian coordinates as[20]

$$\delta_{NH} = D_a\left(\frac{2z^2 - x^2 - y^2}{d^2_{NH}}\right) + \frac{3}{2}D_aR\left(\frac{x^2 - y^2}{d^2_{NH}}\right) \qquad (2)$$

where *x*, *y*, and *z* are the components of the bond vector, d_{NH} is the proton-nitrogen bond distance, and D_a and *R* are the axial component and the rhombicity of the alignment tensor, respectively. Substituting the components of the $\vec{r}^{\,n}_{NH}*$ vector from eq 1 into eq 2, it is possible to write the residual dipolar couplings as a function of both residue number and structure orientation

$$\delta_{NH} = k^{I}\cos^2(\alpha+\rho) + k^{II}\sin^2(\alpha+\rho) + k^{III}\sin(\alpha+\rho)\cos(\alpha+\rho) + k^{IV}\cos(\alpha+\rho) + k^{V}\sin(\alpha+\rho) + k^{VI} \qquad (3)$$

(21) Veglia, G.; Opella, S. J. *J. Am. Chem. Soc.* **2000**, *122*, 11 733−11 734.
(22) Ma, C.; Opella, S. J. *J. Magn. Reson.* **2000**, *146*, 381−384.
(23) Chou, J. J.; Gaemers, S.; Howder, B.; Louis, J. M.; Bax, A. *J. Biomol. NMR* **2001**, *21*, 377−382.
(24) Eyring, H. *Phys. Rev.* **1932**, *39*, 746.
(25) Damiani, A.; DeSantis, P. *J. Chem. Phys.* **1968**, *48*, 4071−4075.
(26) Shimanouchi, T.; Mizushima, J. *J. Chem. Phys.* **1955**, *23*, 82.
(27) Miyazawa, T. *J. Polym. Sci.* **1961**, *55*, 215.
(28) DeSantis, P.; Giglio, E.; Liquori, M. A.; Ripamonti, A. *Nature* **1965**, *206*.
(29) Song, X. J.; Rienstra, C. M.; McDermott, A. E. *Magn. Reson. Chem.* **2001**, *39*, S30−S36.

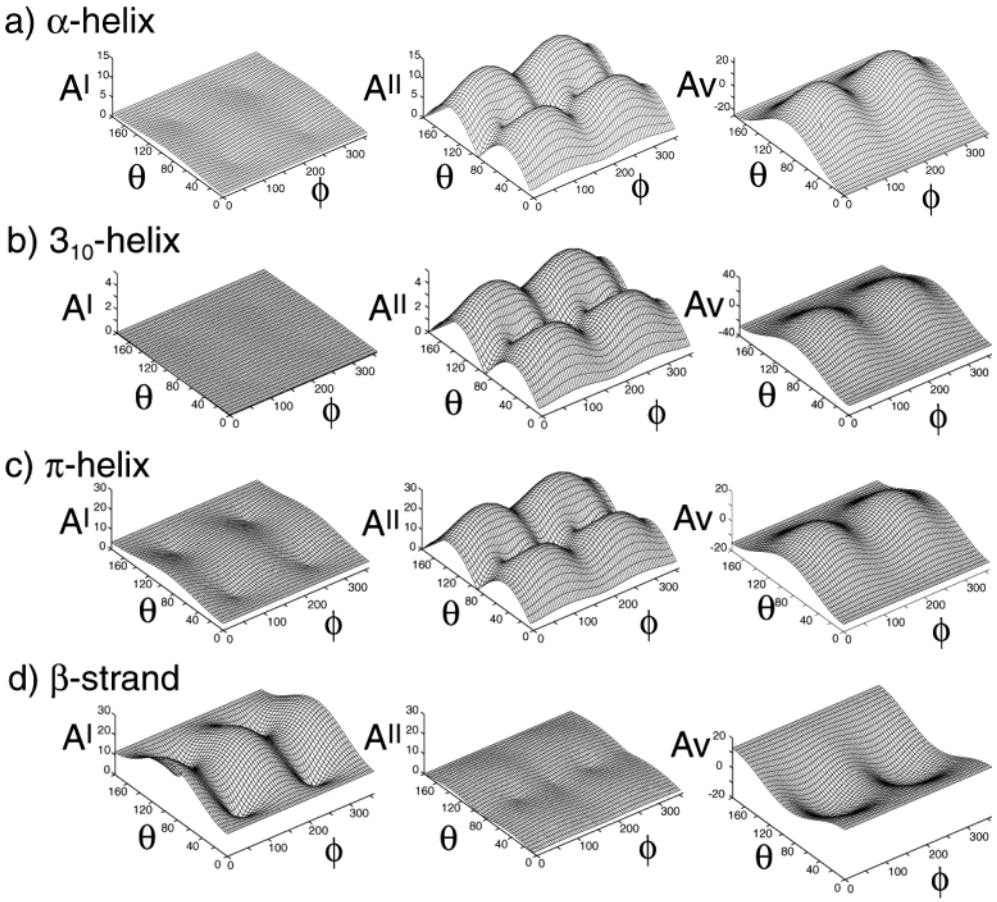


Figure 2. Absolute values of the amplitudes A^{I} and A^{II} and averages Av of the RDC pattern as plotted from eq 5 and as functions of the angles θ and ϕ for: an ideal α-helix (a), a 3_{10}-helix (b), a π-helix (c), and for an antiparallel β-strand (d).

Table 1. Mathematical Expressions of the Coefficients k^{I} to k^{VI} of Eq 3

$k^{I} = 2 \cdot D_a \sin^2\delta \cdot \sin^2\theta - D_a \sin^2\delta \cdot \cos^2\theta + 3/2 \cdot D_a R \cdot \sin^2\delta \cdot \cos^2\theta \cdot \cos 2\phi$
$k^{II} = -D_a \sin^2\delta - 3/2 \cdot D_a R \cdot \sin^2\delta \cdot \cos 2\phi$
$k^{III} = -6 D_a R \cdot \sin^2\delta \cdot \sin\phi \cdot \cos\phi \cdot \cos\theta$
$k^{IV} = -6 D_a \sin\delta \cdot \cos\delta \cdot \sin\theta \cdot \cos\theta + 3 D_a R \cdot \sin\delta \cdot \cos\delta \cdot \sin\theta \cdot \cos\theta \cdot \cos 2\phi$
$k^{V} = -6 D_a R \cdot \sin\delta \cdot \cos\delta \cdot \sin\phi \cdot \cos\phi \cdot \sin\theta$
$k^{VI} = 2 \cdot D_a \cos^2\delta \cdot \cos^2\theta - D_a \cos^2\delta \cdot \sin^2\theta + 3/2 \cdot D_a R \cdot \cos^2\delta \cdot \sin^2\theta \cdot \cos 2\phi$

where the residue number n and the periodicity T of the secondary structure are incorporated into the angle $\alpha = 2\pi(n - 1)/T$. The coefficients for eq 3, k^{I}, k^{II}, k^{III}, k^{IV}, k^{V}, and k^{VI} are functions of the angles δ, θ, and ϕ (Table 1).

Using simple trigonometric transformations, eq 3 can be rearranged as follows

$$\delta_{NH} = \left(\frac{k^{I} - k^{II}}{2}\right)\cos[2(\alpha + \rho)] + \frac{k^{III}}{2}\sin(2(\alpha + \rho)) + k^{IV}\cos(\alpha + \rho) + k^{V}\sin(\alpha + \rho) + k^{VI} + \frac{k^{I} + k^{II}}{2} \quad (4)$$

This equation can be further simplified and transformed into a linear combination of two phase-shifted sinusoids

$$\delta_{NH} = A^{I} \cdot \sin\left[\frac{4\pi}{T}(n - 1) + \varphi^{I}\right] + A^{II} \cdot \sin\left[\frac{2\pi}{T}(n - 1) + \varphi^{II}\right] + Av \quad (5)$$

where A^{I} and A^{II} are the amplitudes, and φ^{I} and φ^{II} the phases

Table 2. Mathematical Expressions of the Amplitudes and Phases of the Two Sinusoids of Eq 5

$$Av = \frac{k^{I} + k^{II} + 2^{VI}}{2}$$

$$A^{I} = \frac{k^{I} - k^{II}}{2\sin\varphi^{I}}$$

$$A^{II} = \frac{k^{IV}}{\sin\varphi^{II}}$$

$$\varphi^{I} = 2\rho + \arctan\left(\frac{k^{I} - k^{II}}{k^{III}}\right)$$

$$\varphi^{II} = \rho + \arctan\left(\frac{k^{II}}{k^{V}}\right)$$

of the two sinusoids. Av is an additional term, independent of the residue number and the periodicity (n and T). The values of the amplitudes, the phases, and Av expressed as functions of the coefficients of eq 3 are reported in Table 2.

This new parametric equation describes the theoretical behavior of residual dipolar couplings for any periodic regular secondary structure of periodicity T and directly links the modulation of their periodicity to their three-dimensional orientations. In particular, this new mathematical treatment makes it possible to fit β-strand structures, which together with α-helices are the most common motifs in proteins. The qualitative and quantitative agreement of our results with simulated and experimental residual dipolar couplings for the different secondary structure elements is reported in the following section.

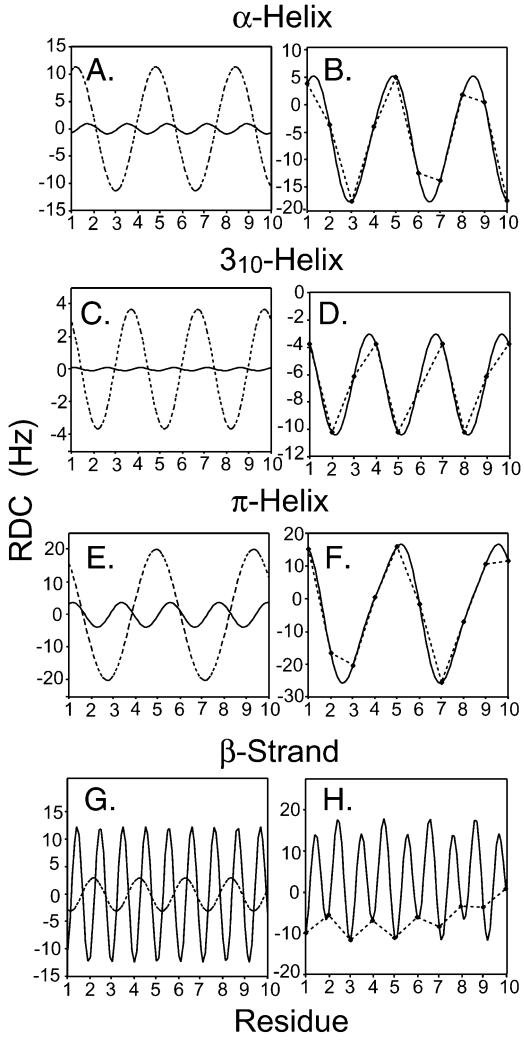


Figure 3. (a) First sinusoid term (solid line) and second sinusoid term (dashed line) of eq 5 plotted as a continuous function of the residue number for an α-helix (a) and (b), a 3_{10}-helix (c) and (d), a π-helix (e) and (f), and a β-strand (g) and (h) all oriented at $\theta = \phi = 45°$. The periodicity T and the angle δ are respectively: α-helix $T = 3.6$ residues/turn and $\delta = 15.8°$; 3_{10} helix $T = 3.0$ residues/turn, $\delta = 4.9°$; π-helix: $T = 4.4$ residues/turn, $\delta = 33.4°$; and β-strand: $T = 2.07$ residues/turn, $\delta = 94.0°$. In (b), (d), (f), and (h), eq 5 is plotted (solid line) and the predicted RDC values (black dots) are reported corresponding to a sampling "rate" of 1 coupling per residue.

Results

Application of Eq 5 to α-Helices. Using arbitrary values of D_a and R ($D_a = -13.6$ Hz and $R = 0.56$), we calculated the dependence of the absolute values of the amplitudes A^I and A^{II}, and $A\nu$ for an ideal α-helix as a function of θ and ϕ in the ranges of $0 \le \theta \le 180°$ and $0 \le \phi \le 360°$ (see Figure 2a). From the surface plots of Figure 2a, it is apparent that the first term A^I of eq 5 is negligible for a broad range of θ and ϕ, whereas the remaining terms A^{II} and $A\nu$ are rather significant. This is also apparent from Figure 3a, where the first term of eq 5 is plotted as a function of residue number. The term A^I (solid line) has a relatively low weight in the equation, while the second term, A^{II} (broken line) is the dominating term. The fitting of the residual dipolar couplings with and without the first term leads to similar results. This is illustrated in Figure 3b, where eq 5 is plotted as a function of the residue number for a helix oriented at $\theta = 45°$ and $\phi = 45°$. In fact, since for an ideal α-helix the angle is $\delta \approx 15.8°$, $\sin^2\delta$ is a relatively small number, and A^I is approximately equal to 0. Therefore, eq 5 for ideal α-helices becomes

$$\delta_{HN} = A^{II}\cdot\sin\left[\frac{2\pi}{T}(n-1) + \varphi^{II}\right] + A\nu \qquad (6)$$

Note that in this case, the term $A\nu$ is approximately equal to k^{VI} and corresponds to the average values of the residual dipolar couplings. This equation is identical to the "dipolar wave" sinusoid proposed by Opella and co-workers for fitting residual or full dipolar couplings of polypeptides in an ideal helical conformation.[16,19] Because the value of periodicity T for an ideal helix is 3.6 residues/turn, the equation depends only on three parameters: the average ($A\nu$), the amplitude (A), and the phase (φ). These three parameters are functions of the orientation of the helix with respect to the PAS (θ and ϕ) and the angle δ (see Tables 1 and 2).

For orientations parallel ($\theta = 0°$) or perpendicular ($\theta = 90°$) with respect to the PAS, the first term of eq 5 becomes significant and the curve deviates from a perfect sinusoidal. This is clearly illustrated in Figures 4a, where A^I is plotted as a function of the θ angle for different values of ϕ. For $\theta = 90^o$ and $\phi = 0^o$, A^I reaches its maximum value. Figure 4, parts b and c, shows the oscillations of the two terms of eq 5 as functions of the residue number for different values of θ, with ϕ fixed at 45°. When the helix is oriented perpendicularly to the magnetic field ($\theta = 90°$), A^I and A^{II} have the same order of magnitude and the RDC pattern is not represented by a simple sinusoid (top of Figure 4, parts b and c). When $\theta = 0°$, the second term of eq 5 cancels out ($A^{II} = 0$) and the RDC pattern follows the behavior indicated by the first term (bottom of Figure 4, parts b and c). Here, the dipolar pattern is best fitted by eq 5.

The sinusoidal fitting of the RDC pattern not only provides information about the orientation of the helix with respect to the magnetic field, it also indicates the phase of rotation of the helix about its axis. The latter information is obtained by extrapolating the value of the angle ρ from the mathematical expression of the parameter φ reported in Table 2. Changes in the RDC pattern and the helical wave for different values of the angle ρ are shown in Figure 5. This type of analysis is crucial for residue assignments in both weakly and strongly oriented helical proteins.[30]

Application of Eq 5 to 3_{10}-Helices and π-Helices. The analysis carried out for ideal α-helices is also valid for both 3_{10}-helices and π-helices. However, unlike the case for ideal α-helices, for 3_{10}-helices the term in A^I is almost zero for all the values of θ and ϕ. Here, eq 6 represents a good approximation of the dipolar pattern. On the other hand, the fitting of dipolar patterns for π-helices, which are not common in nature, requires eq 5. In fact, as with α-helices, the term A^I becomes significant, particularly for orientations close to parallel or perpendicular.

Application of Eq 5 to β-Strands. β-strands in extended conformation are constituted by a flat polypeptide chain with backbone torsion angles of: $\varphi = 180°$, $\psi = 180°$, $\omega = 180°$.[31] In reality, natural occurring β-strands present either a right-hand twist (antiparallel) or a left-hand twist (parallel) around the strand axis, with the antiparallel conformation being the

(30) Marassi, F. M.; Opella, S. J. *Protein Sci.* **2003**, *12*, 403−411.
(31) IUPAC−IUB *Biochemistry* **1970**, *9*, 3471−3479.

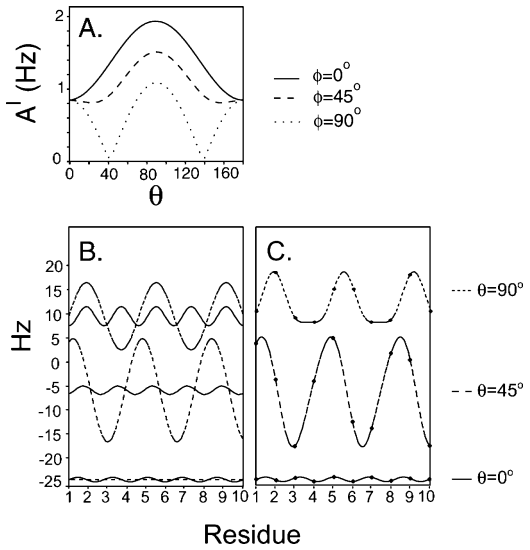


Figure 4. (a) Value of the amplitude A^I as a function of θ for different values of ϕ. (b) and (c) Theoretical patterns of RDCs of a α-helix for three tilt angles $\theta = 0°$, 45° and 90° (and $\phi = 45°$) as predicted by eq 5; (b) plots of the first term (solid lines) and the second term (dashed lines) of eq 5 scaled by the average Av; (c) plot of eq 5 (dashed lines) and predicted theoretical RDCs (black squares).

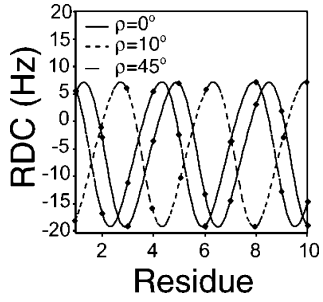


Figure 5. Phase shifts of the helical waves for an ideal α-helix as a function of the angle ρ.

preferred one.[32] The torsion angles are: $\varphi = -139°$, $\psi = 135°$, and $\omega = -178°$ for the antiparallel, and $\varphi = -119°$, $\psi = 113°$, and $\omega = 180°$ for the parallel conformation.[31] However, in naturally occurring globular proteins, the distribution of the torsion angles around these values is quite broad.[32] These deviations are due to the stabilization energy derived from hydrogen bond formation between the strands, and result in periodicities (T) slightly higher (antiparallel) or lower (parallel) than the theoretical 2.0.[32−34] In addition, naturally occurring β-strands often deviate from linearity, presenting remarkable curvatures of the strand axis. All of these factors make the analysis of the periodic dipolar pattern for these structures rather challenging. Nonetheless, by employing eq 5 it is possible to interpret the periodic behavior of β-strands in terms of both local secondary structure and orientation with respect to a reference frame.

For example, we analyzed the case of an antiparallel β-strand polypeptide with $\delta = 94°$ and $T = 2.03$ residues/turn. As with the α-helices, we used arbitrary values of D_a and R to simulate the residual dipolar couplings. The surface plots generated from eq 5 for the absolute values of the amplitudes A^I and A^{II}, and Av as functions of the angles θ and ϕ in the range $0° \leq \theta \leq 180°$ and $0° \leq \phi \leq 360°$ are reported in Figure 2d. The antiparallel β-strand polypeptide behaves differently from the α-helices, with the first term of eq 5 becoming more significant (Figure 2a). This is apparent in Figure 3G, where each of the two terms of eq 5 are plotted as a function of residue number for an antiparallel β-strand with an orientation of θ and ϕ at 45°. Figure 3H shows that when plotted as a continuous function, eq 5 describes the periodic pattern of this antiparallel β-strand.

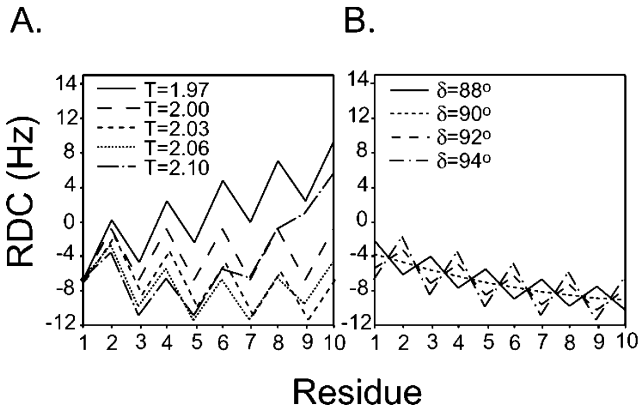


Figure 6. Changes in the RDC pattern for different values of the periodicity T (a), and the angle δ (b). The polypeptide chain is oriented at angles $\theta = 45°$, $\phi = 45°$ and $\rho = 10°$. In (a) $\delta = 94°$. In (b) $T = 2.03$ residues/turn.

Nonetheless, particular care must be taken when analyzing β-strands. We found that when simulating the values of residual dipolar couplings for different β-strand periodicities at a fixed orientation (θ and ϕ) and keeping δ and ρ constants, the pattern of the residual dipolar couplings is affected significantly by the value of the periodicity T, whereas it is only slightly affected by changing δ. This is illustrated in Figure 6, parts A and B, where the predicted residual dipolar couplings of a β-strand oriented at angles $\theta = 45°$, $\phi = 45°$ and $\rho = 10°$ are reported.

Since naturally occurring β-strands exhibit substantial deviations from the ideal β-strand conformation with T and δ values unknown a priori, the fitting of experimental data sets with eq 5 requires the optimization of all six parameters: A^I, A^{II}, T, φ^I, φ^{II}, and Av. This may lead to degenerate solutions with different combinations of parameters that are still good fits of the experimental data.

Comparison of Back-Calculated Residual Dipolar Couplings versus Values Predicted by Eq 5 for α-Helices, 3_{10}-Helices, π-Helices and β-Strands. To show the quantitative fitting of the dipolar couplings, we have back-calculated RDC patterns starting from ideal α-helix, 3_{10}-helix, π-helix, and β-strand built with MOLMOL.[35] In Figure 7, the results of the back-calculations of RDCs are compared with the data predicted by eq 5. The agreement for α-helices, 3_{10}-helices, and π-helices is shown in Figure 7, parts A−C.

Figure 7D shows an identical calculation performed using an ideal antiparallel β-strand structure ($\varphi = -139°$, $\psi = 135°$, $\omega = -178°$).[31] The dashed line interpolates the back-calculated RDC from a PDB file generated with MOLMOL oriented with angles $\theta = 325°$, $\phi = 330°$, and $\rho = 175°$. The optimal fitting (solid lines) for the β-strand is observed for a value $T = 2.035$. Deviations of some of the back-calculated points with respect to the theoretical values are attributable to local deviations from the ideality of the structure used to generate the RDCs.

(32) Chothia, C. *J. Mol. Biol.* **1973**, *75*, 295−302.
(33) Arnott, S.; Dover; S. D. *J. Mol. Biol.* **1967**, *30*, 209.
(34) Schellman, G. N.; Schellman, C. *Proteins* **1964**, *2*, 1.
(35) Koradi, R.; Billeter, M.; Wuthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.

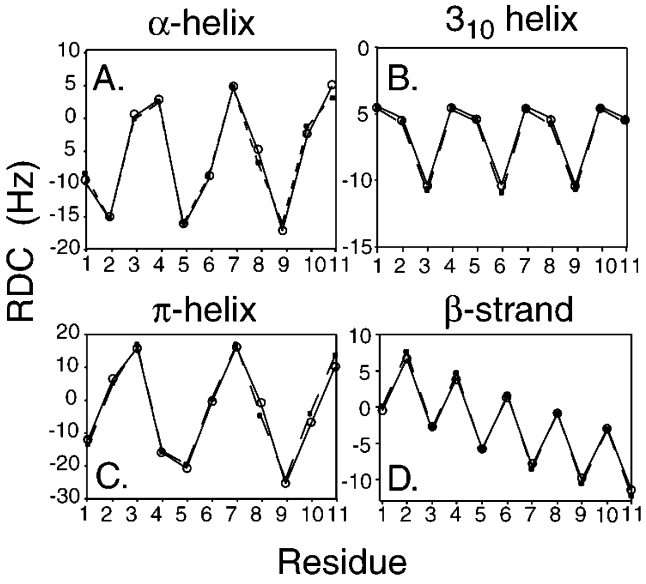


Figure 7. Back-calculated (black squares) and predicted (open circles) RDCs for α-Helix (a), 3_{10}-helix (b), π-helix (c), and a β-strand (d). Black squares represent RDC values back calculated from: (a) an ideal α-helix oriented at $\theta = 45°$, $\phi = 45°$, $T = 3.6$ residues/turn, $\delta = 15.8°$ and $\rho = 45°$; (b) a 3_{10}-helix oriented at $\theta = 45°$, $\phi = 45°$, $T = 3.0$ residues/turn, $\delta = 4.9°$ and $\rho = -31°$; (c) a π-helix oriented at $\theta = 45°$, $\phi = 45°$, $T = 4.4$ residues/turn, $\delta = 33.4°$ and $\rho = 25°$; and (d) a theoretical antiparallel β-strand ($\varphi = -139°$, $\psi = +135°$, $\omega = -180°$ [31]) oriented at $\theta = 45°$, $\phi = 45°$, $\rho = 45°$. In this case, the best fitting has been obtained with $T = 2.034$ residues/turn and $\delta = 95.6°$. Open circles represent RDCs calculated from eq 5. RMSDs between the back calculate and predicted RDC values are (a) 0.300 Hz, (b) 0.119 Hz, (c) 0.646 Hz, and (d) 0.490 Hz.

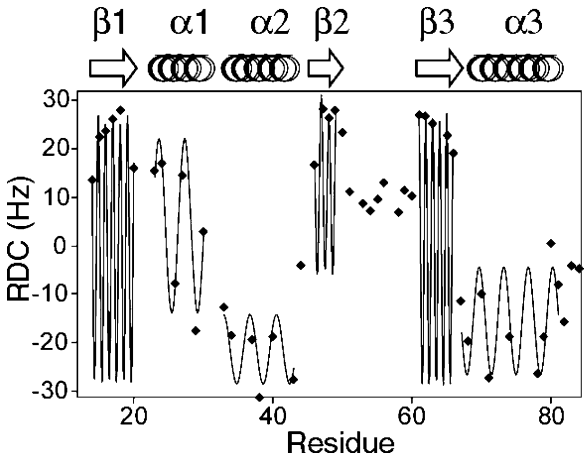


Figure 8. Fitting of the experimental dipolar couplings for of C-terminal KH Domain of heterogeneous nuclear ribonucleoprotein K (KH3) using eq 5.

Application of Eq 5 to Fitting the Dipolar Patterns of the C-Terminal KH Domain of Heterogeneous Nuclear Ribonucleoprotein K (KH3). As an example, we applied eq 5 to fit the dipolar pattern of the C-terminal KH domain of Hnrnp K (KH3), whose structure was recently solved by Tjandra and co-workers.[36] Because this protein adopts a $\beta\alpha\alpha\beta\beta\alpha$ fold with both α-helices and β-strands present in the structure, it serves as an ideal test for our approach. Figure 8 shows the experimental residual dipolar couplings for KH3 obtained in a liquid crystalline medium.[36] The fitting obtained using eq 5 shows that the dipolar patterns for the different β-strands (β1, β2, and β3) have higher frequencies (or low periodicities), while the patterns for the α-helices (α1, α2, and α3) show the characteristic 3.6 periodicity. The accuracy of the fitting was estimated using the root-mean-square deviation (rmsd) of the experimental RDCs from the ones calculated using eq 5. For β1, β2, and β3 domains, we obtained rmsd values of 1.34, 0.14, and 0.43 Hz, respectively, whereas for the three α-helical domain (α1, α2, and α3), the obtained rmsd values were 2.72, 3.23, and 3.76 Hz, respectively.

Discussion

Using a rigorous approach, we have derived a general formula for the interpretation of the distinct periodic pattern presented by the dipolar couplings for regular structures of proteins. The novelty introduced by this approach is the extension of the concept of "dipolar waves" to β-strands, which are widely distributed among protein folds. Equation 5 is a linear combination of two sinusoids, whose amplitudes reflect (a) the different types of secondary structures, and (b) the different orientations of the secondary structure elements with respect to the static magnetic field.

For ideal α-helices, our theoretical approach confirms the previous observations made by Mesleh et al.[16] In this case, using eq 5 we show that the behavior of the RDC pattern can be accurately explained by a sinusoid of periodicity $1/T$ (second term in eq 5). Indeed, eq 5 is also capable of fitting dipolar patterns of β-strand conformations. Unlike α-helices, where one can provide an estimate of the type of secondary structure by simple visual inspection of the dipolar pattern, the pattern of β-strands is not easily recognizable. In this case, the dipolar patterns can be fitted using the linear combination of two sinusoids as described by eq 5, with frequencies $2/T$ and $1/T$, respectively. The difficulties in interpreting the residual dipolar coupling for β-strands using simple "dipolar waves" can be understood by analyzing the dipolar oscillations using the theorem of discretely sampled data.[37] In fact, for a complete representation of a sine wave, it is necessary to have two sample points per cycle (Nyquist critical frequency, $f_c = 1/2\Delta$, where Δ is the sampling interval). For the backbone residual dipolar couplings, Δ is equal to 1 and the Nyquist critical frequency is 0.5. Because an ideal α-helix has $T = 3.6$, the "sampling" of the residual dipolar coupling is within the bandwidth defined by the Nyquist frequency and the pattern is easily recognizable. On the contrary, for β-strands the periodicity is lower than the Nyquist frequency and the oscillating dipolar pattern does not appear as a wave of periodicity $T \approx 2$ (see Figure 3H).

The fitting of dipolar patterns for β-strands also poses some challenges. In fact, naturally occurring β-strands are generally short (6−10 residues[32]) and often present deviations from linearity, making the analysis of their periodicity more cumbersome. Nonetheless, the application of eq 5 to the fitting of experimental data from KH3 shows a remarkable agreement between the theoretical and experimental points, making this method very well suited for determining the relative three-dimensional orientation of protein secondary structure elements. Although eq 5 is not applicable to bent structures in its present form, we are currently working to implement our approach to accurately describe curved and kinked helices and strands whose axes deviate from linearity. However, it should be mentioned

(36) Baber, J. L.; Libutti, D.; Levens, D.; Tjandra, N. *J. Mol. Biol.* **1999**, *289*, 949−962.

(37) Marple, S. L. *Digital Signal Processing*; Englewood Cliffs: N. J., 1987.

that while curved helices are widely distributed in globular proteins, ideal α-helices are expected to be more common in membrane protein structures.[38]

The parametric eq 5 is also able to fit other less common helical conformations, namely 3_{10}-helices and π-helices. For 3_{10}-helices, the first term of eq 5 is close to zero and the fittings of the RDC patterns result in an almost pure sinusoidal curve (Figure 2b). For π-helices on the other hand, the A^1 term in eq 5 is not negligible (Figure 2c).

Because this method is sensitive to both different secondary structure elements and slight distortions of the bond vectors, it can also be used as a tool to establish the ideality of the secondary structure elements. Though, we should point out that additional complications in the analysis of RDCs arise from the inevitable errors in the experimental measurements as well as the "structural noise" introduced by the refinement process for structure calculations.

In summary, the parametric eq 5 extends the concept of pattern recognition to any secondary structure, making it possible to "assemble" secondary structure domains into tertiary folds. A major advantage of using dipolar pattern analysis over RDC constraints for individual bond vectors is the obvious reduction of the ambiguity intrinsic in the vector orientation that characterizes residual dipolar couplings.[19] While the value of a residual dipolar coupling for an individual bond vector is compatible with infinite positions of the vector on the surface of "taco shaped" cones,[39] the steric and chiral requirements of the secondary structure restrict the number of possible orientations for each individual vector within the cone itself. This effect is due to the δ angles that $\vec{r}_{NH}$ bond vectors have in the secondary structures (α-helices, β-strands etc.), which cause the $\vec{r}_{NH}$ vectors to be nonparallel. In fact, although parallel bond vectors would have identical values of the residual dipolar couplings (i.e., identical orientations with respect to the PAS), slight deviations introduce dipolar coupling oscillations. Therefore, only specific orientations will simultaneously satisfy the residual dipolar couplings and the geometry imposed by the secondary structure. From this point of view, the use of dipolar patterns to determine protein topology is conceptually similar to the approach proposed by Prestegard and co-workers,[40] e.g., proteins are considered to be an ensemble of rigid domains that can be assembled through the use of dipolar couplings.

Conclusions

In this work, we derived a parametric equation for a quantitative interpretation of dipolar couplings in proteins weakly aligned in magnetic fields. This theoretical analysis represents an exact mathematical solution for fitting RDC patterns for any periodic secondary structure. Such accurate fitting of the RDC pattern is crucial for extracting both precise local secondary structure parameters and the orientation of the structural fragments with respect to an external reference frame. This analysis can be easily extended to strongly aligned systems such as membrane proteins aligned in oriented lipids, providing a valuable tool for recognizing secondary structure elements and for directly linking their dipolar patterns to the three-dimensional structure. In addition, this method can be used to identify local distortions or deviation of secondary structures from ideality.

Finally, this model is independent of the nature of the nuclei, and in principle, can be applied to any repeating polymer chain with a regular structure.

Acknowledgment. The authors would like to thank Dr. Claudio Anselmi and Professor Pasquale De Santis at the Department of Chemistry of the University of Rome "La Sapienza", as well as Professor Antonio Palleschi at the Department of Chemistry of the University of Rome "Tor Vergata" for helpful discussion. Also, many thanks to Becky Eggimann and Professor Ilja Siepmann for his helpful comments on the manuscript and on the mathematical treatment. This work was supported in part by grants to G.V. (NIH GM64742; AHA 0160465Z).

JA0354824

(38) Kim, S.; Cross, T. A. *Biophys. J.* **2002**, *83*, 2084−2095.
(39) Ramirez, B. E.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 9106−9107.
(40) Al-Hashimi, H. M.; Valafar, H.; Terrell, M.; Zartler, E. R.; Eidsness, M. K.; Prestegard, J. H. *J. Magn. Res.* **2000**, *143*, 402−406.