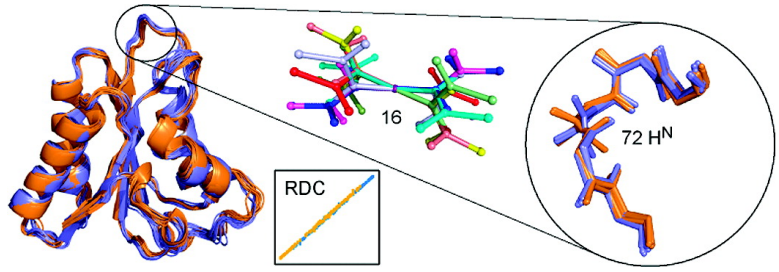
Article

16-Fold Degeneracy of Peptide Plane Orientations from Residual Dipolar Couplings: Analytical Treatment and Implications for Protein Structure Determination

Jean-Christophe Hus, Loi#c Salmon, Guillaume Bouvignies, Johannes Lotze, Martin Blackledge, and Rafael Bru#schweiler

J. Am. Chem. Soc., **2008**, 130 (47), 15927-15937 • DOI: 10.1021/ja804274s • Publication Date (Web): 30 October 2008



More About This Article

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

- Supporting Information
- 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 10/30/2008

16-Fold Degeneracy of Peptide Plane Orientations from Residual Dipolar Couplings: Analytical Treatment and Implications for Protein Structure Determination

Jean-Christophe Hus,[#,†,‡] Loïc Salmon,[‡] Guillaume Bouvignies,[‡] Johannes Lotze,[‡] Martin Blackledge,[*,‡] and Rafael Brüschweiler[*,§]

Clark University, Worcester, Massachusetts 01610, Institut de Biologie Structurale Jean-Pierre Ebel, 38027 Grenoble, France, and Chemical Sciences Laboratory, Department of Chemistry and Biochemistry, and National High Magnetic Field Laboratory, Florida State University, Tallahassee, Florida 32306

Received June 5, 2008; E-mail: martin.blackledge@ibs.fr; bruschweiler@magnet.fsu.edu

Abstract: Residual dipolar couplings (RDCs) measured for internally rigid molecular fragments provide important information about the relative orientations of these fragments. Dependent on the symmetry of the alignment tensor and the symmetry of the molecular fragment, however, there generally exist more than one solution for the fragment orientation consistent with the measured RDCs. Analytical solutions are presented that describe the complete set of orientations of internally rigid fragments that are consistent with multiple dipolar couplings measured in a single alignment medium that is rhombic. For the first time, it is shown that, for a planar fragment such as the peptide plane, there generally exist 16 different solutions with their analytical expressions presented explicitly. The presence of these solutions is shown to be highly relevant for standard structure determination protocols using RDCs to refine molecular structures. In particular, when using standard protein structure refinement with RDCs that were measured in a single alignment medium as constraints, it is found that often more than one of the peptide plane solutions is physically viable; i.e., despite being consistent with measured RDCs, the local backbone structure can be incorrect. On the basis of experimental and simulated examples, it is rationalized why protein structures that are refined against RDCs measured in a single medium can have lower resolution (precision) than one would expect on the basis of the experimental accuracy of the RDCs. Conditions are discussed under which the correct solution can be identified.

1. Introduction

Protein structure determination by NMR spectroscopy is primarily based on nuclear Overhauser distance constraints (NOEs) where the achievable resolution is often limited by the resolution of local structure. Residual dipolar couplings (RDCs) provide complementary information on protein structure.[1−3] They can be used for structural refinement[4]and, in favorable cases, even allow the determination of a protein structure with few or no NOEs.[5−8] When considering RDCs alone, multiple solutions exist for the overall orientation of a molecule or molecular fragment for a given set of RDCs. This is due to the mathematical form of the magnetic dipole−dipole interaction and its properties under anisotropic averaging.[9] The exact number of solutions depends on the relative orientations of the internuclear vectors considered and on internal symmetries of the molecule or the molecular fragment under investigation. The situation is similar for dipolar couplings in the solid state,[10−13] where "dipolar waves" have recently been utilized to relate dipolar couplings to secondary structural motifs.[14,15]

Their well-defined tensorial properties make RDCs amenable to analytical mathematical treatment.[16−24] For example, Meiler et al.[17] and Skrynnikov and Kay[24] have determined analytical

† Clark University.
‡ Institut de Biologie Structurale Jean-Pierre Ebel.
§ Florida State University.
# Current address: Biogen Idec, 12 Cambridge Center, Cambridge, MA 02142.

(1) Bax, A.; Kontaxis, G.; Tjandra, N. *Nucl. Magn. Reson. Biol. Macromol., B* **2001**, *339*, 127–174.
(2) Blackledge, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 23–61.
(3) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. *Chem. Rev.* **2004**, *104*, 3519–3540.
(4) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732–738.
(5) Delaglio, F.; Kontaxis, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 2142–2143.
(6) Hus, J. C.; Marion, D.; Blackledge, M. *J. Am. Chem. Soc.* **2001**, *123*, 1541–1542.
(7) Fowler, C. A.; Tian, F.; Al-Hashimi, H. M.; Prestegard, J. H. *J. Mol. Biol.* **2000**, *304*, 447–460.
(8) Mueller, G. A.; Choy, W. Y.; Yang, D. W.; Forman-Kay, J. D.; Venters, R. A.; Kay, L. E. *J. Mol. Biol.* **2000**, *300*, 197–212.
(9) Saupe, A. *Angew. Chem., Int. Ed. Engl.* **1968**, *7*, 97.
(10) Ketchem, R. R.; Hu, W. D.; Cross, T. A. *J. Cell. Biochem.* **1993**, 277–277.
(11) Tycko, R.; Stewart, P. L.; Opella, S. J. *J. Am. Chem. Soc.* **1986**, *108*, 5419–5425.
(12) Opella, S. J.; Stewart, P. L. *Methods Enzymol.* **1989**, *176*, 242–275.
(13) Quine, J. R.; Cross, T. A. *Concepts Magn. Reson.* **2000**, *12*, 71–82.
(14) Mascioni, A.; Veglia, G. *J. Am. Chem. Soc.* **2003**, *125*, 12520–12526.
(15) Mesleh, M. F.; Veglia, G.; DeSilva, T. M.; Marassi, F. M.; Opella, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 4206–4207.
(16) Meiler, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2001**, *123*, 6098–6107.

expressions for lower and upper bounds for the angle between two dipolar vectors. Wedemeyer et al.[22] developed exact solutions for vector orientations in the presence of additional angular constraints, such as bond angle constraints. Wang and Donald[21] derived analytical expressions for vector orientations when RDCs are available in two alignment media.

When the rhombicity is zero, rotation about the symmetry axis of the alignment tensor does not change the RDCs, and therefore the number of orientations that fulfill a given experimental RDC set belonging to a single alignment is infinite. On the other hand, for nonzero rhombicity, it is commonly assumed that, for a rigid fragment with three or more coplanar dipolar vectors, for example, vectors that lie within the same peptide plane, there are generally eight different solutions for the fragment orientations (see, e.g., refs 6 and 8). It is shown here that, in this case, the number of distinct solutions is in fact twice as large, and analytical expressions are provided that describe all solutions and their orientations. Using the sulfite reductase flavodoxin-like domain as an example, it is demonstrated that multiple solutions can be present within a representative conformational ensemble, even in the case where extensive experimental NOE-derived distance constraints are combined with RDCs from a single data set. In view of the prevalence of this commonly encountered combination of experimental constraints, we discuss in detail the implications for standard structure determination protocols.

2. Theory

We consider a set of interatomic vectors and their associated magnetic dipole−dipole interactions of a molecule expressed in the alignment frame that belongs to an alignment medium with magnitude D_a and rhombicity R. For a rigid internuclear vector $\mathbf{r}_i$ of length r_i, with its direction given by the spherical coordinates (θ_i,ϕ_i), the relationship for a residual dipolar coupling D_i measured for this vector is (in units of Hz)

$$D_i = D_a\left(3\cos^2\theta_i - 1 + \frac{3}{2}R\sin^2\theta_i\cos 2\varphi_i\right) = D_a\left(3z_i^2 - 1 + \frac{3}{2}R(x_i^2 - y_i^2)\right) \quad (1)$$

where $D_a = -\mu_0\gamma_k\gamma_l h/(16\pi^3 r_i^3)A_a$, R is the rhombicity with $0 \le R \le {}^2/_3$, (x_i,y_i,z_i) are the Cartesian coordinates of $\mathbf{r}_i/r_i$ (i.e., after normalizing $\mathbf{r}_i$), A_a is the unitless absolute alignment magnitude, h is Planck's constant, μ_0 is the vacuum permeability, and γ_k, γ_l are the gyromagnetic ratios of the two interacting spins.

Enumeration of Solutions. In general, there are an infinite number of vectors that correspond to a particular value of D_i. One way to reduce this degeneracy is to consider a structural motif with multiple dipolar vectors with known relative orientations. Even in this case, multiple distinguishable orientations of the motif exist that are consistent with the RDCs. A convenient way to enumerate the solutions uses the right-hand-side of eq 1, which makes it immediately clear that the sign inversion of either the x, y, or z axis leaves D_i invariant; i.e., if (x,y,z) is a solution, then any of the eight sign combinations $\mathbf{e}_i = (\pm x,\pm y,\pm z)$ is a solution too.

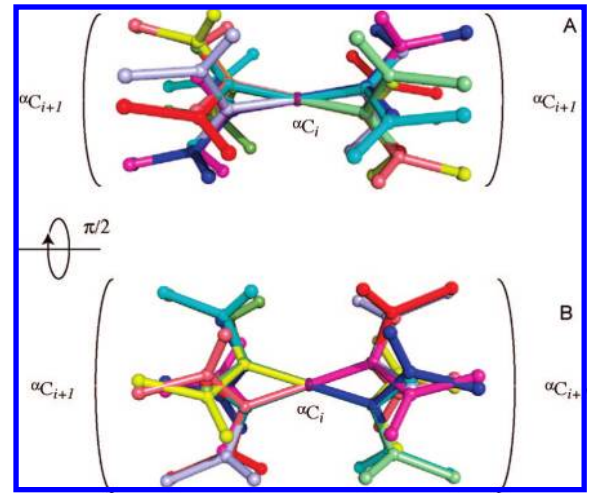


Figure 1. Example of 16 orientations of a peptide plane that are in agreement with a single set of residual dipolar couplings. Cartesian coordinates associated with this example are given in the Supporting Information (Table S1), along with associated RDC values.

For a set of vectors that define a chiral center, however, inversion of the sign with respect to one axis only, for example of all x-components, will change the chirality of the motif and is therefore not allowed. The same applies to the situation where all three axes simultaneously change sign. On the other hand, when two of the three axes change sign, the handedness of the chiral center is conserved. Thus, for a general chiral motif there are only four solutions, (x,y,z), $(x,-y,-z)$, $(-x,y,-z)$, and $(-x,-y,z)$, with the same chirality. (As shown below, for special chiral motives there may be more than four solutions.)

For a structural motif that has no chiral center, which is the case when all vectors lie in the same plane, there is no such restriction, and all eight sign combinations yield allowed solutions. Interestingly, generally there exist *eight additional solutions* that lie in the same planes as the solutions of the first set but whose vector components have orientations that differ from those of the first eight solutions, which is illustrated in Figure 1. These additional solutions flip the planes upside-down (e.g., by a 180° rotation about an axis that lies in the plane). Therefore, in the absence of any in-plane symmetry, the new solutions cannot be aligned with respect to the old solutions by rotation about an axis orthogonal to the plane.

These alternative solutions can be constructed as follows. We consider a plane P depicted in Figure 2A with in-plane vector orientations specified by the three angles (η,ψ,χ) defined in the figure that can be expressed using standard trigonometry:

$$\mathbf{e}(\chi) = (\cos\eta\cos\chi - \sin\eta\sin\psi\sin\chi,\ \sin\eta\cos\chi + \cos\eta\sin\psi\sin\chi,\ \cos\psi\sin\chi) \quad (2)$$

where χ defines the vector orientation within plane P, η is the angle between the intersection of plane P with the xy plane of the alignment frame and the x axis, and ψ is the tilt angle of plane P with respect to the z-axis of the alignment frame. The dipolar couplings can then be expressed as a function of (η,ψ,χ) by insertion of eq 2 into eq 1:

(17) Meiler, J.; Blomberg, N.; Nilges, M.; Griesinger, C. *J. Biomol. NMR* **2000**, *16*, 245–252.
(18) Moltke, S.; Grzesiek, S. *J. Biomol. NMR* **1999**, *15*, 77–82.
(19) Wang, J.; Walsh, J. D.; Kuszewski, J.; Wang, Y. X. *J. Magn. Reson.* **2007**, *189*, 90–103.
(20) Walsh, J. D.; Wang, Y. X. *J. Magn. Reson.* **2005**, *174*, 152–162.
(21) Wang, L. C.; Donald, B. R. *J. Biomol. NMR* **2004**, *29*, 223–242.
(22) Wedemeyer, W. J.; Rohl, C. A.; Scheraga, H. A. *J. Biomol. NMR* **2002**, *22*, 137–151.
(23) Tian, F.; Valafar, H.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 11791–11796.
(24) Skrynnikov, N. R.; Kay, L. E. *J. Biomol. NMR* **2000**, *18*, 239–252.

$$d' = D/D_a = 3\sin^2\chi\cos^2\psi - 1 + \frac{3}{2}R[(\cos^2\chi(1+\sin^2\psi) - \sin^2\psi)\cos 2\eta - \sin(2\chi)\sin(2\eta)\sin\psi] \quad (3a)$$

or

$$d'(\chi) = D/D_a = A\cos 2\chi - B\sin 2\chi + C \quad (3b)$$

where $A = -(3/2)\cos^2\psi + (3/4)R(1+\sin^2\psi)\cos 2\eta$, $B = (3/2)R\sin(2\eta)\sin\psi$, and $C = (3/2)\cos^2\psi - 1 + (3/4)R(1+\sin^2\psi)\cos(2\eta) - (3/2)R\sin^2\psi\cos(2\eta)$

It can be shown that there exists a suitable axis lying in plane P so that a 180° rotation about this axis of the in-plane dipolar vectors yields additional solutions. This is equivalent to demonstrating that a constant offset δ exists that depends on (η,ψ) so that

$$d'(\chi) = d'(-\chi - \delta) \quad (4)$$

Insertion of eq 3 into eq 4 yields

$$A\cos(2\chi) - B\sin(2\chi) + C = A\cos(2\chi + 2\delta) + B\sin(2\chi + 2\delta) + C \quad (5)$$

Equation 5 is indeed fulfilled for arbitrary angles χ if δ fulfills

$$\tan\delta = B/A \quad (6)$$

where the two solutions for $\delta = \arctan(B/A)$ are 180° shifted with respect to each other, as displayed in Figure 2B.

In summary, if for a rhombic alignment tensor ($R \neq 0$) an internally rigid planar structural motif fulfills a given set of RDCs, then there generally exist no fewer than 16 distinct solutions. The 16 solutions cluster into four groups, where the four solutions of each group lie in the same plane, i.e., each group shares the same (parallel or antiparallel) normal axis to the plane. In the alignment frame, each of these axes points toward one of the four upper quadrants. The four solutions of each group cluster into two subgroups, where the solutions in the subgroups are related to each other by a 180° rotation about the normal axis. The subgroups are related to each other by a 180° rotation about the axis that intersects plane P and the *xy* plane of the alignment frame followed by a rotation about the normal by $-\delta$.

It further follows that a chiral motif consisting of dipolar vectors within a plane and dipolar vectors orthogonal to the plane has the same 16 solutions as the planar system without violating the chirality constraint. A systematic description of all solutions in terms of rotations about the three Euler angles is given in the following sections.

Determination of Solutions from Experimental Data. To determine all possible orientations of an internally rigid molecular fragment on the basis of RDC data, we find it convenient to use internal angular coordinates:

$$D_i = 2D_a\sqrt{\frac{4\pi}{5}}\left\{Y_0^2(\theta_i,\varphi_i) + \sqrt{\frac{3}{8}}R(Y_2^2(\theta_i,\varphi_i) + Y_{-2}^2(\theta_i,\varphi_i))\right\} \quad (7)$$

Equation 7 expresses dipolar couplings in terms of normalized spherical harmonics functions of rank 2, $Y_l^2(\theta,\varphi)$, where (θ_i,φ_i) are the polar angles that belong to the corresponding vector $\mathbf{e}_i$ in the alignment tensor frame. For a set of vectors $\mathbf{e}_i$, one can always define a rotation that rotates the vectors from their initial coordinates (θ_i^0,φ_i^0) in a fixed frame of reference to the frame

(25) Zare, R. N. *Angular Momentum*; John Wiley & Sons: New York, 1988.

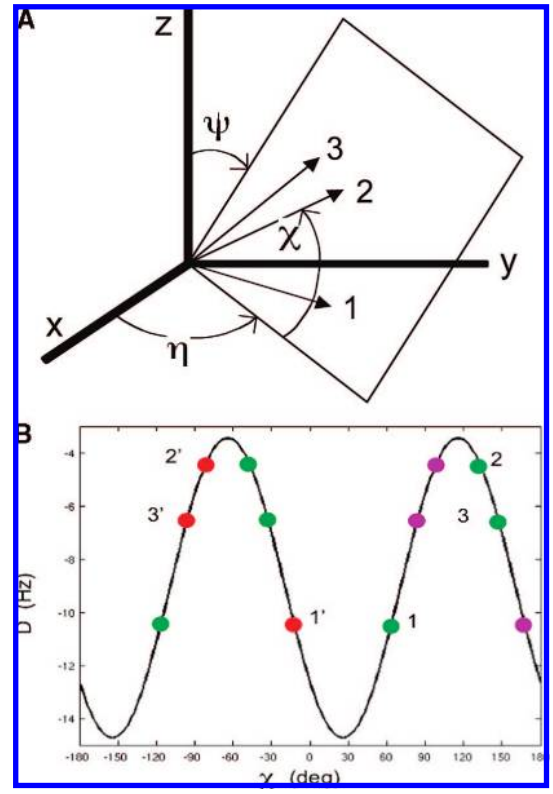


Figure 2. Reverse plane orientations consistent with residual dipolar couplings (RDCs) of three coplanar vectors **1**, **2**, and **3** measured in a single alignment medium with nonzero rhombicity. (A) Definition of angles (η,ψ,χ) that define in-plane vector orientation with respect to alignment frame (x,y,z). (B) For an alignment tensor with $R = 0.35$, the four solutions for vector orientations that lie in the same plane are indicated by four different colors: the red and magenta solutions are related to each other by a $\Delta\chi = 180°$ in-plane rotation. The two sets of solutions (red/magenta vs green) are related to each other by inverting χ followed by a shift δ, i.e., $\chi \rightarrow -\chi - \delta$, where δ is defined by eqs 2−6 as illustrated for vector sets **1**, **2**, **3** and **1′**, **2′**, **3′**.

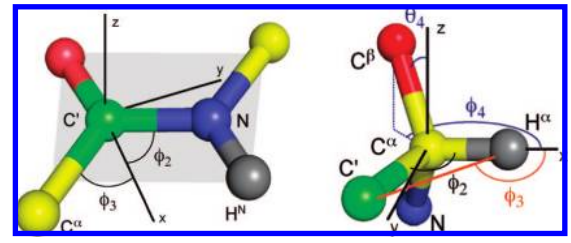


Figure 3. Definition of angles for (A, left) three vectors that lie in the peptide plane and (B, right) four vectors of the chiral amino acid C^α center at the junction between adjacent peptide planes.

of the alignment tensor where their coordinates are (θ_i,φ_i) and their RDC is given by eq 7. Such a rotation can be described by using the Wigner matrices[25] with elements $D^{(2)}_{mn}(\alpha,\beta,\gamma)$ that depend on the Euler angles (α,β,γ). Application of this rotation to eq 7 yields[16]

$$D_i = 2D_a\sqrt{\frac{4\pi}{5}}\left[\sum_{m=-2}^{2} D^2_{m,0}(\alpha,\beta,\gamma)\,Y_m^2(\theta_i^0,\varphi_i^0) + \sqrt{\frac{3}{8}}R\sum_{m=-2}^{2}(D^2_{m,2}(\alpha,\beta,\gamma)\,Y_m^2(\theta_i^0,\varphi_i^0) + D^2_{m,-2}(\alpha,\beta,\gamma)\,Y_m^2(\theta_i^0,\varphi_i^0))\right] \quad (8)$$

After simplification, the following expression for the normalized residual dipolar coupling is obtained:

$$d_i = \frac{D_i}{3D_a} = \sum_{k=1}^{5} b_i^k B_k \tag{9a}$$

where B_k are geometric functions that relate the molecular frame to the alignment frame,

$$\begin{aligned}
B_1 &= \cos^2\beta - \frac{1}{3} + \frac{R}{2}\sin^2\beta\cos 2\gamma \\
B_2 &= \cos 2\alpha\left[-\sin^2\beta + \frac{R}{2}(1+\cos^2\beta)\cos 2\gamma\right] - \\
&\quad R\sin 2\alpha \sin 2\gamma \cos\beta \\
B_3 &= \sin 2\alpha\left[-\sin^2\beta + \frac{R}{2}(1+\cos^2\beta)\cos 2\gamma\right] + \\
&\quad R\cos 2\alpha \sin 2\gamma \cos\beta \\
B_4 &= \sin\beta\left[2\cos\beta\cos\alpha\left(1 - \frac{R}{2}\cos 2\gamma\right) + R\sin\alpha\sin 2\gamma\right] \\
B_5 &= \sin\beta\left[2\cos\beta\sin\alpha\left(1 - \frac{R}{2}\cos 2\gamma\right) - R\cos\alpha\sin 2\gamma\right]
\end{aligned} \tag{9b}$$

with (α,β,γ) corresponding to the rotation angles according to the *zxz* convention (i.e., first rotation by α about the *z*-axis, second rotation by β about the *x*-axis, third rotation by γ about the *z*-axis), and b_i^k represent the dipolar vector orientations in the molecular frame,

$$\begin{aligned}
b_i^1 &= \frac{3\cos^2\theta_i^0 - 1}{2} \\
b_i^2 &= \frac{\sin^2\theta_i^0}{2}\cos 2\varphi_i^0 \\
b_i^3 &= -\frac{\sin^2\theta_i^0}{2}\sin 2\varphi_i^0 \\
b_i^4 &= -\sin\theta_i^0\cos\theta_i^0\cos\varphi_i^0 \\
b_i^5 &= \sin\theta_i^0\cos\theta_i^0\sin\varphi_i^0
\end{aligned} \tag{9c}$$

The linear system of eq 9 can be solved by linear least-squares minimization, e.g., using singular value decomposition (SVD), if there are at least five different dipolar couplings available. This relates to the fact that generally five RDCs are necessary to determine the alignment tensor.[26] Here, we assume that the alignment tensor magnitude D_a and rhombicity R are already known. Therefore, fewer than five RDCs allow establishment of the orientations of a structural motif, as is discussed below.

Equation 9 reflects the fact that, when one knows the coordinates of a set of vectors and their normalized RDCs, one can extract coefficients B_k that contain information about the rotational transformation of these vectors to their correct orientations in the alignment frame. The associated Euler angles and rhombicity R can be extracted from these coefficients as follows:

(a) If $R = 0$:

$$\begin{aligned}
\beta &= \arccos\sqrt{B_1 + \frac{1}{3}} \\
\alpha &= \frac{1}{2}\arctan\frac{B_3}{B_2} = \arctan\frac{B_5}{B_4} \\
\gamma &\quad \text{undefined}
\end{aligned} \tag{10a}$$

which means that there is an infinite number of solutions due to the axial symmetry of the alignment tensor.

(b) If $R \neq 0$:

$$\begin{aligned}
\beta &= \arcsin\sqrt{1 - \frac{\left(B_1 + \frac{4}{3}\right)^2 - B_2^2 - B_3^2}{4 - R^2}} \\
\gamma &= \frac{\pi}{2} - \frac{1}{2}\arccos\frac{2}{R\sin^2\beta}\left(\sin^2\beta + B_1 - \frac{2}{3}\right) \\
\alpha &= \\
&\frac{1}{2}\arctan\frac{\left[-\sin^2\beta + \frac{R}{2}(1+\cos^2\beta)\cos 2\gamma\right]B_3 - R\cos\beta\sin 2\gamma B_2}{\left[-\sin^2\beta + \frac{R}{2}(1+\cos^2\beta)\cos 2\gamma\right]B_2 + R\cos\beta\sin 2\gamma B_3}
\end{aligned} \tag{10b}$$

With the following allowed ranges for the Euler angles, the four-fold degeneracy in orientation, leading to four rotations having the same residual dipolar couplings, is compiled in Table 1:

$$\begin{aligned}
0 \le \alpha \le \pi \\
0 \le \beta \le \frac{\pi}{2} \\
0 \le \gamma \le \pi
\end{aligned} \tag{11}$$

Hence, for the general case of N vectors, eq 10b defines the Euler angles describing the rotation that transforms these vectors from a molecular frame of reference to the alignment frame.

In the presence of symmetries in the molecular fragment under consideration, some of the coefficients B_k can be undetermined, which leads to additional allowed fragment orientations. In the following section, we consider special cases that are relevant for applications to proteins and peptides.

Peptide Plane. The situation where all dipolar vectors lie in the same plane has already been mentioned above. This situation is commonly encountered in practice, with the peptide plane in proteins and nucleic acid bases being the most prominent examples. It is convenient to align the *xy* plane of the coordinate system with the molecular plane so that all in-plane vectors have $\theta_i^0 = \pi/2$. Here and in the following section, ideal geometries and noise-free RDC data are assumed. As a consequence, in eq 9c, coefficients b_i^4 and b_i^5 are zero, and hence B_4 and B_5 are undetermined, which increases the number of possible solutions. Figure 3A shows the angles that define the peptide plane motif. We use a minimum of three vectors, including $N-H^N$, $N^{(i)}-C'^{(i-1)}$, and $C'^{(i-1)}-C^{\alpha(i-1)}$, whose polar coordinates are as given as follows:

$$\begin{aligned}
&\mathbf{1}\text{:}\quad N-H^N,\quad \theta_1^0 = \frac{\pi}{2},\quad \varphi_1^0 = 0 \\
&\mathbf{2}\text{:}\quad C'-N,\quad \theta_2^0 = \frac{\pi}{2},\quad \varphi_2^0 = \varphi_2 \\
&\mathbf{3}\text{:}\quad C'-C^\alpha,\quad \theta_3^0 = \frac{\pi}{2},\quad \varphi_3^0 = \varphi_3
\end{aligned} \tag{12}$$

Table 1. Four-Fold Degeneracy of Fragment Orientation

solution	Euler angles		
1	α	β	γ
2	α	β	$\gamma + \pi$
3	$\alpha + \pi$	$\pi - \beta$	$-\gamma$
4	$\alpha + \pi$	$\pi - \beta$	$-\gamma + \pi$

(26) Losonczi, J. A.; Andrec, M.; Fischer, M. W. F.; Prestegard, J. H. *J. Magn. Reson.* **1999**, *138*, 334–342.

Equation 9a then leads to

$$d_{N-H^N} = -\frac{1}{2}B_1 + \frac{1}{2}B_2$$
$$d_{C'-N} = -\frac{1}{2}B_1 + \frac{\cos 2\varphi_2}{2}B_2 - \frac{\sin 2\varphi_2}{2}B_3$$
$$d_{C'-C^\alpha} = -\frac{1}{2}B_1 + \frac{\cos 2\varphi_3}{2}B_2 - \frac{\sin 2\varphi_3}{2}B_3 \qquad (13)$$

From the first three equations one obtains for B_1, B_2, and B_3,

$$B_1 = 2\frac{\sin 2\varphi_3(d_{C'-N} - d_{N-H^N}) - \sin 2\varphi_2(d_{C'-C^\alpha} - d_{N-H^N})}{\sin 2\varphi_3(\cos 2\varphi_2 - 1) - \sin 2\varphi_2(\cos 2\varphi_3 - 1)} - 2d_{N-H^N}$$
$$B_2 = 2\frac{\sin 2\varphi_3(d_{C'-N} - d_{N-H^N}) - \sin 2\varphi_2(d_{C'-C^\alpha} - d_{N-H^N})}{\sin 2\varphi_3(\cos 2\varphi_2 - 1) - \sin 2\varphi_2(\cos 2\varphi_3 - 1)}$$
$$B_3 = 2\frac{(\cos 2\varphi_3 - 1)(d_{C'-N} - d_{N-H^N}) - (\cos 2\varphi_2 - 1)(d_{C'-C^\alpha} - d_{N-H^N})}{\sin 2\varphi_3(\cos 2\varphi_2 - 1) - \sin 2\varphi_2(\cos 2\varphi_3 - 1)} \qquad (14)$$

which according to eq 10b leads to Euler angles (α,β,γ). Since in eq 9c coefficients $b_i^4 = b_i^5 = 0$, B_4 and B_5 do not impose any restriction. Therefore, for $R \neq 0$, there are additional solutions (α',β',γ'), which are consistent with the first three equations of eqs 9b and 9c:

$$\alpha' = \frac{1}{2}\arctan\frac{\left[-\sin^2\beta + \frac{R}{2}(1+\cos^2\beta)\cos 2\gamma\right]B_3 + R\cos\beta\sin 2\gamma B_2}{\left[-\sin^2\beta + \frac{R}{2}(1+\cos^2\beta)\cos 2\gamma\right]B_2 - R\cos\beta\sin\ 2\gamma B_3}$$
$$\beta' = \pi - \beta$$
$$\gamma' = \pi + \gamma \qquad (15)$$

The (α,β,γ) set and the (α',β',γ') set are both eight-fold degenerate, as listed in Table 2. The complete set of peptide plane orientations satisfying the RDCs is then 16, consistent with the Theory section. In special cases, some of these solutions will coincide such that the effective degeneracy can coalesce to ≤ 8. The two sets of solutions are related to each other by a rotation about an axis that lies in the peptide plane, as discussed above (see also Figures 1 and 2). Figure 4 provides a step-by-step description of Euler angle rotations for the two sets of solutions.

Amino Acid C^α Chiral Center. By using a suitable choice of initial coordinates, eq 9a can be solved analytically in the case of a nonplanar motif. We now demonstrate this result for the C^α chiral center of an amino acid (Figure 3B) using the four one-bond RDCs: $C^\alpha-H^\alpha$, $C^\alpha-C^\beta$, $C^\alpha-C'$, and $C'-H^\alpha$.

Table 2. 16-Fold Degeneracy of Fragment Orientation[a]

solution	Euler angles		
1	α	β	γ
2	α	β	$\gamma + \pi$
3	$\alpha + \pi$	$\pi - \beta$	$-\gamma$
4	$\alpha + \pi$	$\pi - \beta$	$-\gamma + \pi$
5	$\alpha + \pi$	β	γ
6	$\alpha + \pi$	β	$\gamma + \pi$
7	α	$\pi - \beta$	$-\gamma$
8	α	$\pi - \beta$	$-\gamma + \pi$

[a] Solutions 9−16 are identical to 1−8, except for replacing (α,β,γ) by (α',β',γ'), as defined in the text.

We assume the C^α chiral center to have perfect tetrahedral bond geometry, with vector **1** being the $C^\alpha-H^\alpha$ vector along the x axis. The internal coordinates are (Figure 3B)

$$\mathbf{1}: \quad C^\alpha-H^\alpha, \quad \theta_0^1 = \frac{\pi}{2}, \quad \varphi_0^1 = 0$$
$$\mathbf{2}: \quad C^\alpha-C', \quad \theta_0^2 = \frac{\pi}{2}, \quad \varphi_0^2 = \varphi_2$$
$$\mathbf{3}: \quad C'-H^\alpha, \quad \theta_0^3 = \frac{\pi}{2}, \quad \varphi_0^3 = \varphi_3$$
$$\mathbf{4}: \quad C^\alpha-C^\beta, \quad \theta_0^4 = \theta_4, \quad \varphi_0^4 = \varphi_4 \qquad (16)$$

Equation 9 then leads to

$$d_{C^\alpha-H^\alpha} = -\frac{1}{2}B_1 + \frac{1}{2}B_2$$
$$d_{C^\alpha-C'} = -\frac{1}{2}B_1 + \frac{\cos 2\varphi_2}{2}B_2 - \frac{\sin 2\varphi_2}{2}B_3$$
$$d_{C'-H^\alpha} = -\frac{1}{2}B_1 + \frac{\cos 2\varphi_3}{2}B_2 - \frac{\sin 2\varphi_3}{2}B_3$$
$$d_{C^\alpha-C^\beta} = -\frac{3\cos^2\theta_4 - 1}{2}B_1 + \frac{\sin^2\theta_4}{2}\cos 2\varphi_4 B_2 -$$

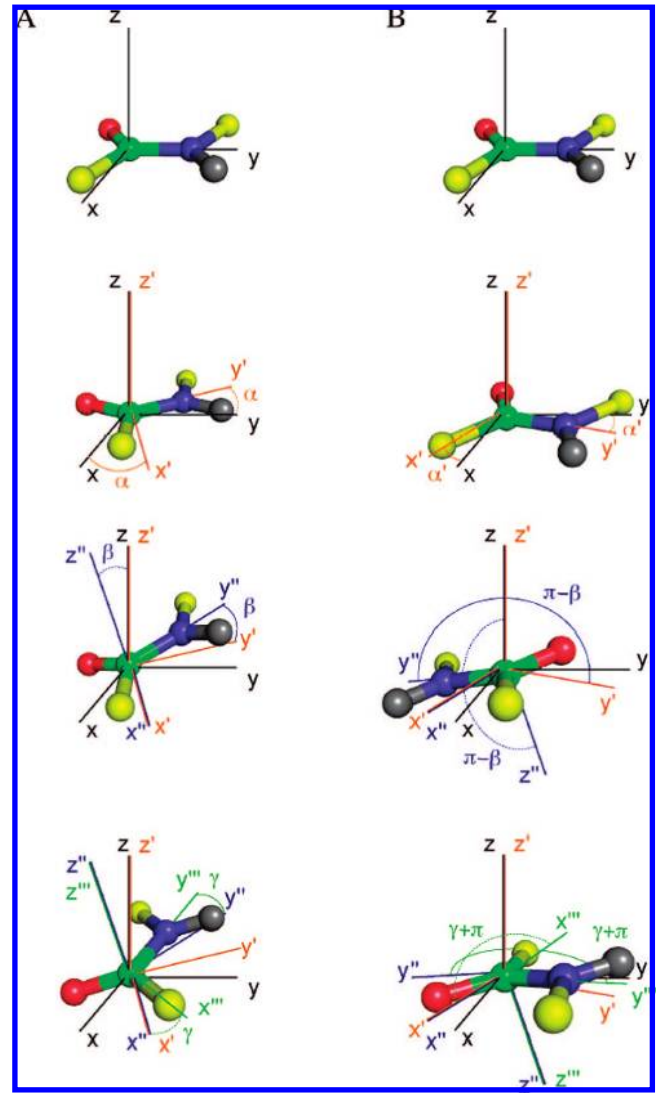


Figure 4. Step-by-step illustration of Euler angle rotations of one vector according to two sets of Euler angles, (α,β,γ) (A, left column) and (α',β',γ') (B, right column).

$$\frac{\sin^2\theta_4}{2}\sin 2\varphi_4 B_3 - \sin\theta_4\cos\theta_4\cos\varphi_4 B_4 + \sin\theta_4\cos\theta_4\sin\varphi_4 B_5$$

One can readily extract B_1, B_2, and B_3 from the $C'-C^\alpha-H^\alpha$ plane using eq 14 and the associated Euler angles (α,β,γ) according to eq 10. Moreover, eq 9b leads to

$$3B_1^2 + B_2^2 + B_3^2 + B_4^2 + B_5^2 = R^2 + \frac{4}{3} \tag{18}$$

with solutions

$$B_4 = \frac{-k_1k_2 \pm \sqrt{k_1^2k_2^2 - (1+k_1^2)\left(3B_1^2 + B_2^2 + B_3^2 - R^2 - \frac{4}{3}\right)}}{1+k_1^2}$$

$$B_5 = k_1B_4 + k_2$$

$$k_1 = \frac{1}{\tan\varphi_4}$$

$$k_2 = \frac{d_{C^\alpha-C^\beta} - \frac{3\cos^2\theta_4 - 1}{2}B_1 + \frac{\sin^2\theta_4}{2}\cos 2\varphi_4 B_2 - \frac{\sin^2\theta_4}{2}\sin 2\varphi_4 B_3}{\sin\theta_4\cos\theta_4\sin\varphi_4} \tag{19}$$

Thus, there are generally two solutions for B_4 and B_5. However, since the Euler angles are known, B_4 and B_5 are also determined by eq 9b and Table 2, which permits identification of the solution that is consistent with the $C'-C^\alpha-H^\alpha$ RDC information. This procedure yields the four sets of unique Euler angles that describe the orientations of the chiral motif that are consistent with the four RDCs. As noticed above, for $\theta_4 = 0$ (vector normal to the plane), the 16 solutions are valid even though it is a chiral motif since $b_4 = 0$, and $b_5 = 0$ and B_4 and B_5 do not enter eq 17.

When the alignment tensor is known, the above formulas allow determination of all fragment orientations consistent with the RDC set. Because these fragments are part of a larger molecular structure to which they are chemically bonded, additional constraints apply such as bond angle and dihedral angle constraints. Moreover, additional information may be available about internuclear distances (from NOEs and trans-hydrogen bond scalar J-couplings) and dihedral angles (from scalar ^3J-couplings and chemical shifts) as discussed in the following section.

3. Application to Biomolecular Structure Determination

By far the most prevalent use of experimental RDCs is found in their application to the refinement of structures whose global architecture is already defined, for example, from distance constraints derived from NOESY experiments. Restrained molecular dynamics calculations are thus routinely used to refine the conformations of proteins and nucleic acids using RDC constraints typically measured in a single alignment medium only. In order to assess the implications of the analytical solutions derived above on commonly applied protocols, we have undertaken a series of calculations using both simulated and experimental data.

Implication of 16 Peptide Plane Orientations for Structural Refinement. Characterization from Simulation. The possible manifestation of the 16 RDC-consistent orientations available for any single planar element using data from one alignment medium has been simulated using RDC-restrained molecular dynamics calculations performed using the program Sculptor[27] under the following conditions:

(i) Peptide plane RDCs ($N-H^N$, $C'-C^\alpha$, $N^{(i)}-C'^{(i-1)}$, and $H^N-C'^{(i-1)}$) were simulated from the known structure of protein GB3[28] using a maximally rhombic alignment tensor. Distances were also simulated between all H^N and H^α in the protein, and all distances less than 5 Å were then used as simulated NOE restraints with lower and upper limits of 0 and 6 Å, respectively. The structure was then subjected to a restrained molecular dynamics simulation with only the NOE restraints applied, followed by refinement using both simulated NOEs and simulated RDCs. Both simulations used Cartesian dynamics with a time step of 0.3 fs, with a sampling period of 8 ps at 1000 K, followed by slow cooling to 100 K over 12 ps and subsequent energy minimization. The alignment tensor eigenvalues and eigenvectors were allowed to evolve during the whole period, resulting in accurate reproduction of the original tensor.

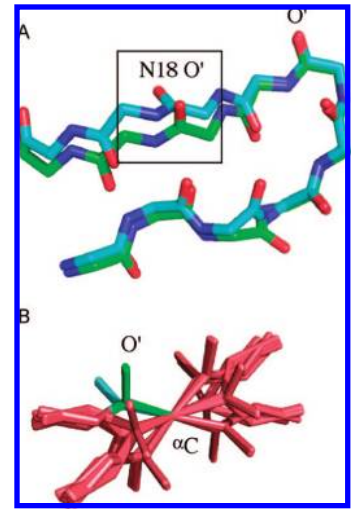


Figure 5. RDC-consistent orientations of peptide plane units produced from RDC-restrained molecular dynamics calculations. (A) Two different RDC-consistent orientations of peptide unit N18-G19 in protein GB3. Data were simulated from a known tensor and subsequently used as restraints, in addition to loose distance restraints between backbone protons (see text). The two structures agree equally well with RDC and distance data as well as tetrahedral geometry. (B) Peptide plane data from peptide plane N18-G19 were used as restraints using the same protocol as in panel A with CNS-Sculptor to determine the possible orientations of the isolated plane. All 16 solutions predicted from eqs 16–18 are shown here. Two of these (the green and the cyan conformations) correspond to the two shown in panel A. Note that four solutions have very similar C^α positions at either end of the plane, illustrating the possibility of finding false minima.

Two low-energy structures are shown in Figure 5 that agree equally well with the simulated RDCs. Two planes are seen to have peptide plane orientations that fulfill the RDCs and do not violate tetrahedral geometry nor the loose distance restraints. In order to check whether all these orientations are RDC-consistent as described by eqs 10a,b the RDCs from the N18/G19 plane were used as restraints to determine the orientation of a single, free peptide plane using the known alignment tensor from the simulation. The results of 100 repeated calculations for which RDC energy violations were equivalent and negligible are shown in Figure 5B. The complete set of 16 peptide plane orientations is visibly present, with two of these corresponding to the two orientations shown in Figure 5A. Clearly both

(27) Hus, J. C.; Marion, D.; Blackledge, M. *J. Mol. Biol.* **2000**, *298*, 927–936.
(28) Bouvignies, G.; Markwick, P.; Brüschweiler, R.; Blackledge, M. *J. Am. Chem. Soc.* **2006**, *128*, 15100–15101.

Table 3. Alignment Tensors Used To Simulate RDCs from SiR-FP18 (1−3) and Experimental Tensors from Bacteriophage and Steric Alignment Media (A, B)

	D_a (Hz)	R ($\times 10^{-4}$)	α (°)	β (°)	γ (°)
tensor 1	23	0.65	45.0	0.0	45.0
tensor 2	23	0.65	0.0	45.0	45.0
tensor 3	23	0.65	60.0	60.0	60.0
tensor A	−21.5	0.63	84.6	107.0	117.4
tensor B	−24.2	0.52	−149.8	153.1	−130.2

covalent and nonbonded terms in the force field, as well as additional experimental restraints, will impose additional restrictions on the physical viability of any of these solutions. In particular, orientations that invert the direction of the peptide chain are highly unlikely and are excluded (e.g., on the basis of the position of the subsequent amino acid). We note, however, that four peptide plane orientations have similarly positioned C^{α} atoms, underlining the real possibility that wrong orientations can be found for peptide planes.

(ii) In the second example, experimental distance constraints have been taken[29] from the sulfite reductase flavodoxin-like domain SiR-FP18 (comprising 1114 NOESY-derived distance restraints, 119 hydrogen bond distance restraints based on trans-hydrogen bond $^{3}J_{NC}$ couplings, and accurate dihedral angle restraints from ^{13}C chemical shifts), together with simulated peptide plane RDC data from imposed rhombic alignment tensors (tensors 1 and 2) using a known structure determined using the NOESY dihedral angle and hydrogen-bond distance restraints from randomized initial coordinates.[29] RDCs were also simulated from a third, differently oriented alignment tensor (tensor 3) for purposes of comparison. The use of the third tensor allows for straightforward identification of incorrectly oriented planes. These three tensors are linearly independent, as gauged by the scalar product of the tensor elements (Table 3). The same restrained molecular dynamics protocol was applied (see simulation i) to refine the distance-only derived structures. Five calculations were performed with the following active RDCs: (A) no RDCs used in the structure refinement; (B) $N-H^N$ RDCs simulated from tensor 1 used in the structure refinement; (C) four RDCs from the peptide plane ($N-H^N$, $C'-C^{\alpha}$, $N^{(i)}-C'^{(i-1)}$, and $H^N-C'^{(i-1)}$) simulated from tensor 1 used in the structure refinement; (D) four RDCs from the peptide plane ($N-H^N$, $C'-C^{\alpha}$, $N^{(i)}-C'^{(i-1)}$, and $H^N-C'^{(i-1)}$) simulated from tensor 1 and tensor 2 used in the structure refinement; and (E) $N-H^N$ RDCs simulated from tensor 1 and tensor 2 used in the structure refinement. Structural ensembles were selected on the basis of their agreement with distance and RDC restraints.

The final ensemble of RDC-refined structures from calculations B−E are all in close agreement with the respective active RDCs from tensors 1 and 2. These lowest energy (combined experimental energy term associated with RDCs and distance restraints) structures were compared with the RDCs from tensor 3 by fitting the vector orientations of the four RDCs ($N-H^N$, $C'-C^{\alpha}$, $N^{(i)}-C'^{(i-1)}$, and $H^N-C'^{(i-1)}$) to an optimal alignment tensor. This provides a simple means to identify correctly and incorrectly oriented planes. The correlation of simulated and calculated RDCs to the best-fitting tensors is shown in Figure 6 for typical members of the optimal ensembles for calculations A−E.

Importantly, it can be seen that there is no visible improvement between refinement against a single vector from a single alignment medium with respect to using no RDCs. This has previously been noted[30] but is strikingly clear from the correlations in Figure 6A,B. We note here that, using a single set of $N-H^N$ RDCs, it is not always possible to determine the alignment tensor to any accuracy. Alignment tensors are normally determined simultaneously during structure refinement, either using discrete sampling of different combinations of the axial and rhombic components of the tensor or allowing the tensor eigenvalues to float. The latter approach is used here, and while we find that the eigenvalues are sometimes correctly chosen, they can also be wrong, and the target function of the "active" constraints does not reflect this. For example, one of the best structures (on the basis of experimental and physical violations) has an axial tensor component of 31 Hz rather than the value of 23 Hz that was used to simulate the data (this is the worst correlation of the ten structures in the ensemble and is shown in Figure 6B).

In order to investigate this effect in more detail, we have calculated the backbone root-mean-square deviation (rmsd) over ensembles calculated with no RDC information and compared this to the results for ensembles calculated using one full $N-H^N$ RDC set, one full set of four RDCs per peptide plane, and two full sets of peptide plane RDCs. The overall rmsd falls slightly (0.61 compared to 0.68 Å) when $N-H^N$ RDCs are introduced. This information is presented in the Supporting Information on a per-residue basis (Figure S2). In this specific case, local and global structures of SiR FP18 are well defined by the experimental NOE and dihedral angle restraints, as demonstrated by the already low rmsd, and under these conditions, refinement against a single set of $N-H^N$ RDCs provides only minor improvement and does not necessarily better orient the $N-H^N$ vectors compared to the RDC-free refinement. In the light of the high degeneracy of vector orientations that are consistent with a given RDC value, it should not surprise that refinement against a single set of couplings provides negligible improvement. Exceptions are conceivable, however, for example, when the mutual orientations of helical elements are poorly determined due to the sparsity of long-range NOEs. In such cases, the combination of $N-H^N$ vectors from multiple sites may provide a better determination of the relative orientations of the helices.

Beyond a single vector, two vectors (results not shown) and four vectors from a single alignment medium define a planar element (Figure 6C) that reproduces RDCs from tensor 3 for the vast majority of sites and results in more precisely defined backbone conformations (rmsd 0.39 Å). However, the presence of alternative peptide plane orientations was also detected by large local violations of the third set of RDCs (Figure 6C). These are visible as clear outliers in the panels, especially for the $N-H^N$ and $H^N-C'^{(i-1)}$ RDCs. Investigation of these sites reveals that each case corresponds to one of the alternative 15 orientations of the planes described above. Two such cases are shown in Figure 7A,B. For peptide plane D99/Y100, two orientations of the plane can be identified. Both are in agreement with the available distance restraints and tetrahedral geometry at the C^{α} positions. The case of F160/C161 (Figure 7B) illustrates that the orientation of the plane is almost inverted between the two solutions, with the hydrogen−bonding moieties positioned on different sides of the peptide chain. Again no violations of the experimental NOEs are found. The positions of the sites in the protein where two orientations are observed are shown in Figure 7C, illustrating the fact that, while such

(29) Sibille, N.; Blackledge, M.; Brutscher, B.; Coves, J.; Bersch, B. *Biochemistry* **2005**, *44*, 9086–9095.
(30) Grishaev, A.; Bax, A. *Curr. Opin. Struct. Biol.* **2005**, *15*, 563–570.

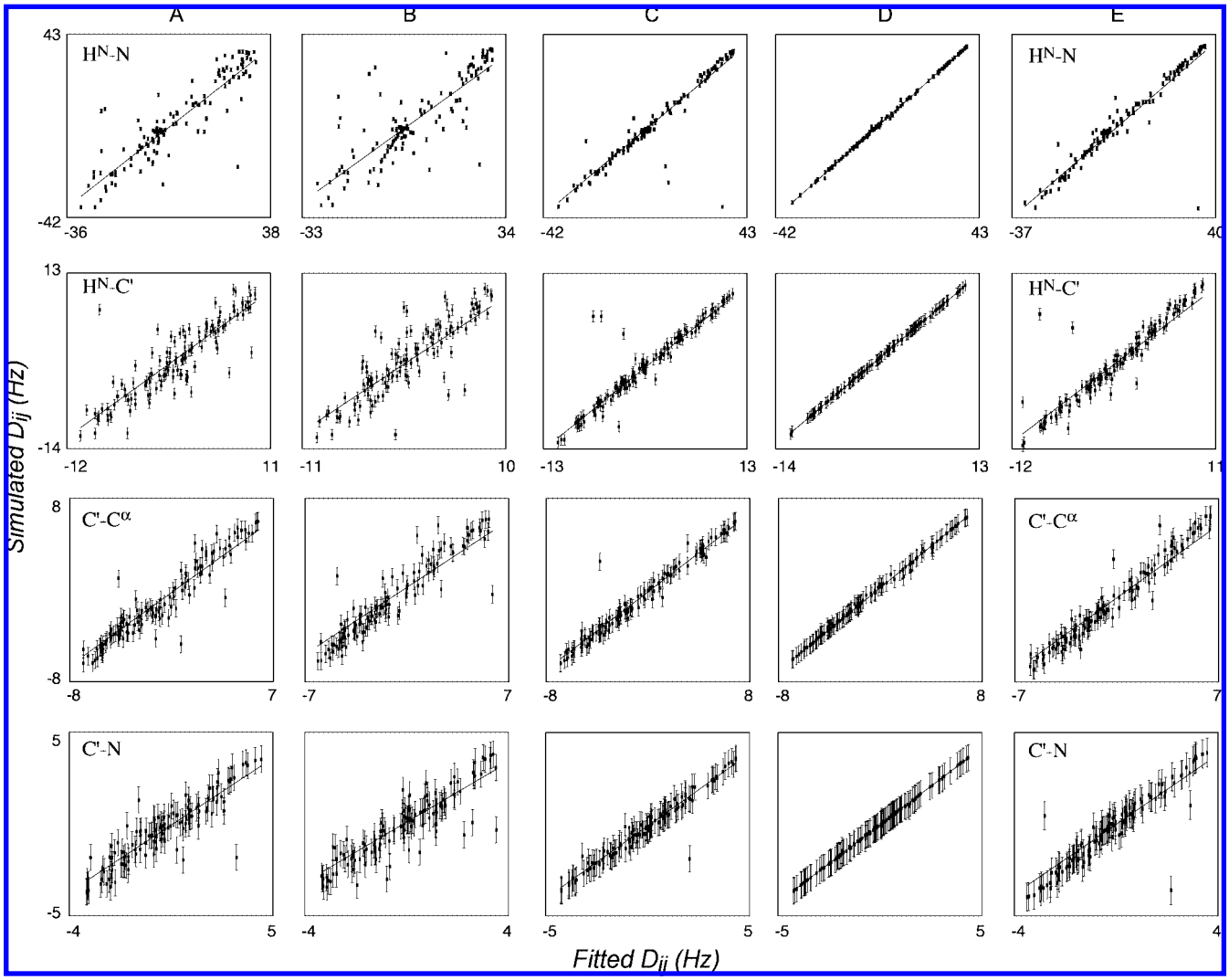


Figure 6. Sampling of RDC-consistent peptide plane orientations using simulated RDCs. Best-fit correlations of RDCs simulated from tensor 3 (Table 3) when fit to optimal alignment tensors determined using structures refined using data from tensors 1 and 2. (A) No RDCs used in the refinement. (B) Only $N-H^N$ RDCs from tensor 1 used in structure refinement. (C) Four peptide plane RDCs from tensor 1 used in structure refinement. (D) Four peptide plane RDCs from tensors 1 and 2 used in structure refinement. (E) Only $N-H^N$ RDCs from tensors 1 and 2 used in structure refinement.

ambiguous peptide plane orientations are more likely to occur in loop regions, due to the generally lower NOE constraint density, they can also occur in regular secondary structural regions. The $C'-C^\alpha$ and $N-C'$ vectors are less sensitive to the presence of incorrect orientations, probably because the overall fold, determined by distance information, essentially defines the position of the C^α atoms, such that the degrees of freedom available to these vectors are more restricted than the vectors involving the H^N atom.

It is also useful to represent the misorientations using a metric that directly relates to the molecular structure. We have therefore calculated the cosines of angles between vectors in each plane and the angle (Ω) between the planes of interest and planes in the target, or known structure. The quantity $\cos(\Omega)$ is shown for each plane for one of the structures present in the ensemble illustrated in Figure 7 (Figure 7D). While for the majority of sites we see that the orientation is correctly determined ($\cos(\Omega) \approx 1$), for other sites the orientation is clearly different, indicating that the plane populates one of the possible equivalent orientations. The cases of orientational degeneracy illustrated in Figure 7A,B are both present and are indicated by asterisks. In the case of plane 160−161, the large angle (>90°) between the correct and incorrect conformations results in a negative value of the cosine.

Not surprisingly, calculations carried out using data from both alignment tensors (Figure 6D) eliminates all of the spurious minima, leaving only the native orientation. The rmsd of the backbone conformation compared to the average structure falls to 0.36 Å, compared to 0.39 Å for a single set of couplings (Figure S2, Supporting Information). Use of RDCs from two differently aligned tensors will indeed reduce the peptide plane orientational degeneracy from 16 to 2, which is the correct plane orientation and its inverse. The structure determination protocol Meccano,[6] developed for the determination of protein backbone conformation using RDCs measured in two different alignment media, exploits the result that the 16-fold degeneracy is reduced to a two-fold degeneracy under these circumstances.

We have also compared the validity of plane orientations refined using only $N-H^N$ RDCs measured in two alignment media (Figure 6E). This again results in isolated occurrences of incorrect orientations of the $N-H^N$ vectors. It is important to note that, in each case, the RDCs used for refinement are essentially in perfect agreement with the structures. Cross-

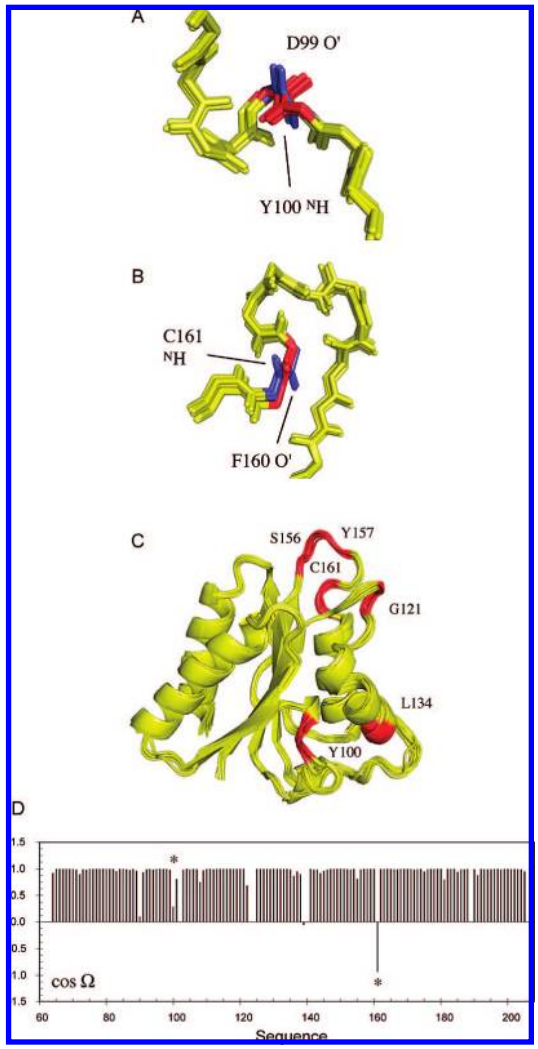


Figure 7. RDC-consistent orientations of peptide plane units produced from RDC-restrained molecular dynamics calculations using experimental distance restraints from sulfite reductase flavodoxin-like domain and simulated RDC data from a single alignment tensor. (A) The 20 structures in the lowest experimental energy ensemble sample two RDC-consistent orientations of peptide unit D99/Y100 that are in agreement with the distance constraints. The blue orientation is the correct orientation. (B) The 20 structures in the lowest experimental energy ensemble sample two RDC-consistent orientations of peptide unit F160/C161 that are in agreement with the distance constraints. The blue orientation again shows the correct orientation. (C) Sites of sulfite reductase flavodoxin-like domain (in red) where two RDC-consistent orientations are observed. (D) Display of cos(Ω) for one of the structures shown in panels A−C, where Ω is the angle between the peptide planes in the structure of interest and the target, or known structure (see text). The positions of peptide planes 99−100 and 160−161 are shown with an asterisk.

validation of the plane orientations using other vectors within the plane does not readily reveal that these orientations are in fact incorrect. It is therefore hard to detect this type of local structural errors in the absence of additional orientational information.

Manifestation in Experimental Ensembles of RDC-Refined Structures. Identification of wrong peptide plane orientations by cross-validation relying on RDCs from different alignments can be directly applied to experimental data. Naturally, demonstration of such effects will also depend on experimental noise in both the "active" data set and the data set used to identify the incorrect equivalent orientations. For this purpose, we have applied the same structure refinement protocols using two high-quality experimental RDC data sets from SiR FP18 in combination with experimental NOESY-derived distances.

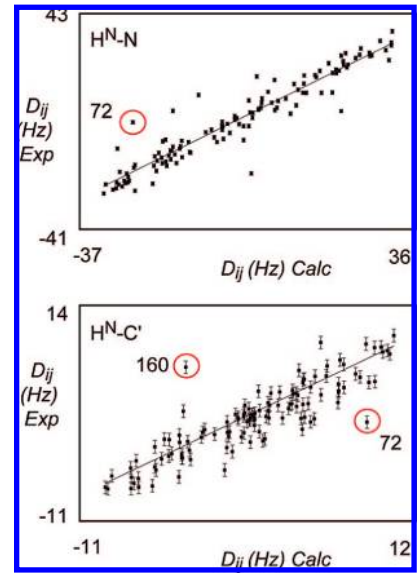


Figure 8. (A) Best-fit correlation of experimental $N-H^N$ RDCs from alignment medium B when fit to optimal alignment tensors determined using structures refined using experimental $N-H^N$, $C'-C^\alpha$, and $C'-H^N$ RDCs data from alignment medium A. Correlation from a typical structure is shown from the ensemble best-fitting the active data set (alignment medium A, bacteriophage). A circled outlier corresponds to one of the sites populating two RDC-consistent orientations with respect to these data (see Figure 9). (B) Best-fit correlation of experimental $C'-H^N$ RDCs from alignment medium A when fit to optimal alignment tensors determined using structures refined using experimental $N-H^N$, $C'-C^\alpha$, and $C'-H^N$ RDCs data from alignment medium B. Correlation from a typical structure is shown from the ensemble best-fitting the active data set (alignment medium B, alcohol-based alignment). Circled outliers correspond to sites populating two RDC-consistent orientations with respect to these data (see Figure 9).

Alignment medium A (bacteriophage) is expected to align on the basis of electrostatic repulsion, while alignment medium B (PEG/hexanol mixture) is expected to align on the basis of steric repulsion. The resulting alignment tensors have previously been shown to be significantly different (Table 3). The structural ensembles resulting from refinement with respect to media A and B show similar behavior to that observed from the simulated data sets: outliers are observed when compared to the "passive" data sets that are not used in the refinement. In Figure 8 we show correlations with respect to $N-H^N$ and $C'-H^N$ RDCs from typical members of the two ensembles.

Investigation of the structural ensembles reveals that the observed outliers indeed correspond to sites where both correct and incorrect RDC-consistent orientations are sampled. Figure 9 illustrates two such sites. In the case of the plane associating amino acids 158−159, two orientations are found from the refinement with respect to data from medium B, while a single orientation is found in the ensemble refined with respect to data from medium A.

The incorrect orientations in the "B" ensemble violate RDCs from medium A, as shown in Figure 8. In the case of the plane associating amino acids 71−72, two orientations are found from the refinement with respect to data from both media (Figure 9). Of course, only the common orientation is correct, while the incorrect orientations violate RDCs from the unused medium in both cases, as shown in Figure 8. In all cases, the structures in the final ensembles reproduce the active RDCs well. We note that these ensembles are very similar with respect to global rmsd measurements (0.52 ± 0.08 and 0.64 ± 0.07 Å), compared to

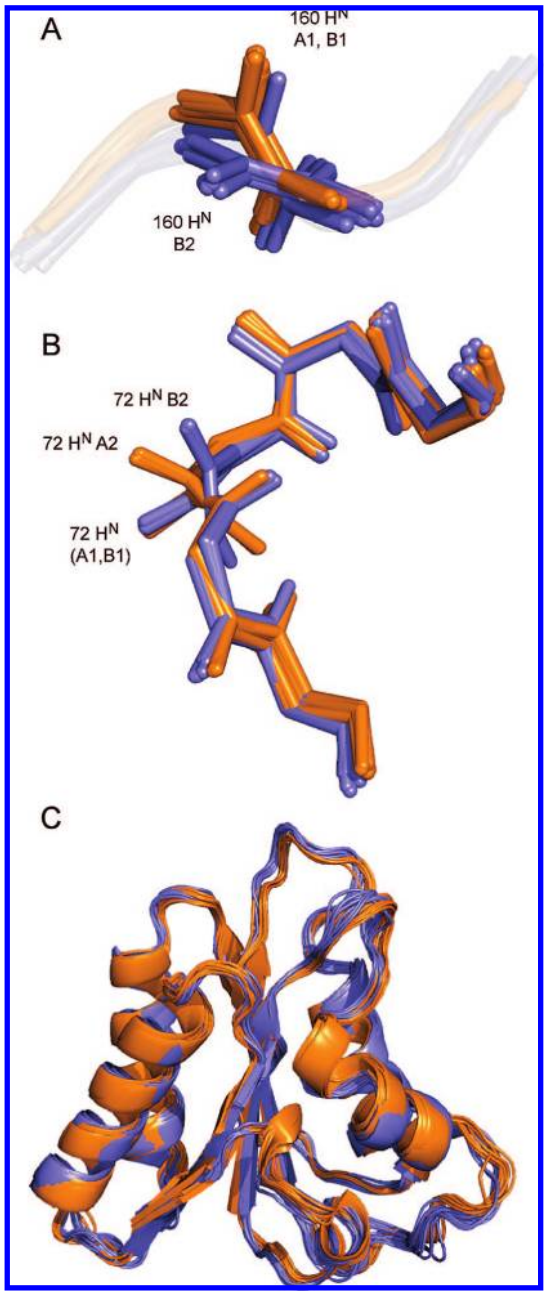


Figure 9. (A) Peptide plane 159−160 shows two orientations in the structural ensemble determined using data from alignment medium B (blue). Two minima, B1 and B2, are indicated. Data from alignment medium A (orange) define a single orientation (A1). (B) Peptide plane 71−72 shows two orientations in the structural ensemble determined using data from alignment medium B (blue: minima B1 and B2). The structural ensemble determined using data from alignment medium A (orange) also show two orientations (A1 and A2). The common orientation is correct. (C) Comparison of the structural ensembles refined using RDCs from media A (orange) and B (blue).

the respective means, and mostly differ in terms of local structure resulting from ambiguous definition of individual peptide plane orientations (Figure 9C).

Implication for Structure Refinement Protocols. Both simulated and experimental data therefore indicate that ambiguous orientations predicted from the analytical descriptions derived above can appear in representative ensembles of structures refined using a single set of RDCs. Their presence is somewhat random, assuming that the correct orientation is equally likely and that other experimental data as well as covalent and noncovalent restraints do not discriminate between the solutions. In this case, the ensemble can contain two or more families of similarly oriented planes. Use of additional RDC restraints from the tetrahedral junction, use of more accurate dihedral angle restraints (e.g., from TALOS), or use of more precise NOE distance information may reduce the possibility of sampling RDC-consistent but wrong peptide plane orientations. In this context it is, however, important to note that the distance information measured from SiR FP18 was relatively complete, including experimentally determined hydrogen-bonding restraints and NOEs extracted from high-resolution NOESY spectra. If one finds numerous sites where multiple orientations are observed under such favorable refinement conditions as illustrated using both noise-free simulated and experimentally measured RDCs, it would be surprising if such cases were not common in standard RDC refinement studies.

A number of additional factors can contribute to the precision of structures refined against RDCs. These include the true geometry of local structural elements that are often constrained to adopt a common optimal geometry. A second source of potential error lies in the supposition that a static model can appropriately describe the experimentally measured parameters. Dynamic averaging, however, can significantly affect measured RDCs[16,31−35] so that refinement against motionally averaged RDCs will be expected to compromise the orientational information present in the final structure. Hence, the analysis presented here primarily applies to amino acids that are part of structured and motionally confined regions of the polypeptide chain.

4. Conclusions

The relationship between a set of RDCs and the associated internuclear vector orientations expressed in an arbitrary molecular frame is determined by the five parameters B_k which are obtained by a linear least-squares fit. Conversion of the B_k parameters to the three Euler angles that define the orientation of the molecular fragment in the alignment frame must take into account that there are multiple solutions due to the symmetry of the alignment tensor and possible symmetries, such as planarity, of the fragment itself. Analytical expressions have been introduced here for the complete set of nonequivalent Euler angles. These relationships, together with bonding constraints, are directly applicable to the reconstruction of biomacromolecular structures from RDCs. Even though this analysis assumes idealized local geometry, it holds interest in methods for the scaffolding of protein backbone structures.

Application of the derived expressions to the description of peptide plane orientations demonstrates that, for nonzero rhombicity, there are generally 16 nonequivalent solutions. Because some of the wrong solutions can have an orientation similar to that of the correct solution, their detection can be difficult,

(31) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A* **1995**, *92*, 9279–9283.
(32) Lakomek, N. A.; Walter, K. F.; Fares, C.; Lange, O. F.; de Groot, B. L.; Grubmuller, H.; Brüschweiler, R.; Munk, A.; Becker, S.; Meiler, J.; Griesinger, C. *J. Biomol. NMR* **2008**, *41*, 139–155.
(33) Bernado, P.; Blackledge, M. *J. Am. Chem. Soc.* **2004**, *126*, 4907–4920.
(34) Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Brüschweiler, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885–13890.
(35) Tolman, J. R.; Ruan, K. *Chem. Rev.* **2006**, *106*, 1720–1736.

directly affecting the validity and resolution of a protein structure determined by standard RDC-assisted refinement. In terms of standard practices for protein structure refinement, we show that the existence of 16 different RDC-consistent peptide plane orientations can result in incorrect local plane orientations, even in the presence of extensive long-range NOE distance information. This has immediate implications for the correctness and resolution of protein and nucleic acid structures that are refined against a single set of RDCs.

Acknowledgment. We dedicate this work to Prof. Richard R. Ernst on the occasion of his 75th birthday. This work was supported by the NIH (grant R01 GM 066041) and NSF (grant 0621482) to R.B. and through EU-NMR JRA3 projects RII3-026145 and ANR NT05-4_42781 to M.B.

Supporting Information Available: Cartesian coordinate representation of 16 distinct solutions for peptide plane orientations; backbone rmsd's of sulfite-reductase flavodoxin-like domain for different RDC refinement protocols. This material is available free of charge via the Internet at http://pubs.acs.org.

JA804274S