Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 191 (2008) 24–30
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Continuity conditions and torsion angles from ssNMR orientational restraints

S. Achuthan [a,*], T. Asbury [c], J. Hu [f], R. Bertram [a,d], T.A. Cross [b,e], J.R. Quine [a,b]

[a] Department of Mathematics, Florida State University, Tallahassee, FL 32306-4510, USA
[b] National High Magnetic Field Laboratory, Tallahassee, FL 32310-3706, USA
[c] Department of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina, Charleston, SC 32306, USA
[d] Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, USA
[e] Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL 32306, USA
[f] NIDDK, NIH, Bethesda, MD 20892, USA

Received 27 June 2007; revised 16 November 2007
Available online 3 December 2007

Abstract

The backbone torsion angle pair (ϕ, ψ) at each amino acid of a polypeptide is a descriptor of its conformation. One can use chemical shift and dipolar coupling data from solid-state NMR PISEMA experiments to directly calculate the torsion angles for the membrane-spanning portion of a protein. However, degeneracies inherent in the data give rise to multiple potential torsion angles between two adjacent peptide planes (a diplane). The molecular backbone structure can be determined by gluing together the consecutive diplanes, as in the PIPATH algorithm [T. Asbury, J.R. Quine, S. Achuthan, J. Hu, M.S. Chapman, T.A. Cross, R. Bertram, PIPATH: an optimized alogrithm for generating α-helical structures from PISEMA data, J. Magn. Reson. 183 (2006) 87–95.]. The multiplicities in torsion angles translate to multiplicities in diplane orientations. In this paper, we show that adjacent diplanes can be glued together to form a permissible structure only if they satisfy *continuity conditions*, described quantitatively here. These restrict the number of potential torsion angle pairs. We rewrite the torsion angle formulas from [J.R. Quine, M.T. Brenneman, T.A. Cross, Protein structural analysis from solid-state NMR-drived orientational constraints, Biophys. J. 72 (1997) 2342–2348.] so that they automatically satisfy the continuity conditions. The reformulated torsion angle formulas have been applied recently in the PIPATH algorithm [T. Asbury, J.R. Quine, S. Achuthan, J. Hu, M.S. Chapman, T.A. Cross, R. Bertram, PIPATH: an optimized alogrithm for generating α-helical structures from PISEMA data, J. Magn. Reson. 183 (2006) 87–95.] and will be helpful in other applications in which diplane gluing is used to construct a protein backbone model.
Published by Elsevier Inc.

Keywords: Solid state NMR; Torsion angles; Peptide planes; Diplanes; Degeneracies; PISEMA; PIPATH

1. Introduction

Solid state nuclear magnetic resonance (ssNMR) spectroscopy has the unique capability to characterize membrane protein structure in a liquid crystalline lipid bilayer environment ([1–7]) yet there are degeneracies in the interpretation of the data that complicate the structural analysis. Here, we show that the number of degeneracies is reduced in the case of permissible protein structures. The structure of a protein model can be built up using a recursive approach [10] with residue specific assignments. Another approach would be to combine segments of the protein structure that have been determined independently. We refer to this process as the *gluing method*. The gluing method is especially useful for solving the structure of proteins from partially assigned ssNMR data. The shotgun approach [18] for determining α-helical membrane protein

* Corresponding author. Present address: Neuroscience Center of Excellence, School of Medicine, Louisiana State University Health Sciences Center, New Orleans, LA-70112, USA. Fax: +1 504 599 0891.
E-mail address: sachut@lsuhsc.edu (S. Achuthan).

1090-7807/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.jmr.2007.11.018

structures is an example of the gluing method. Another example is the PIPATH algorithm [25] which builds an atomic model of a protein by gluing together the diplanes. The PIPATH algorithm was used in the determination of the atomic structure of the amantadine-blocked transmembrane domain of the M2 proton channel from Influenza A virus [29].

Orientational restraints [10,16] from ssNMR form the basis of the recursive and gluing methods. These restraints give rise to degeneracies [11–13,17] in peptide plane orientations. That is, multiple peptide plane orientations are consistent with a given data point. Each degeneracy is 2-fold and can be expressed in the form of ± 1. Unique orientations are determined from the signs of the associated degeneracies. In this sense, final resolution of a protein structure computed from ssNMR orientational restraints is equivalent to solving for a sequence of sign degeneracies.

Torsion angle formulas from Cross et al. [22] illustrate how to form a diplane from two PISEMA data points. Overlapping diplanes can then be glued together to form the structure of the full protein. However, due to the degeneracies, some combinations of diplanes obtained through the torsion angle formulas are not permissible. These were described in [13,18,23]. The primary goal of this paper is to rewrite the torsion angle formulas from Cross et al. [22] so that they incorporate a rule, which we call the continuity condition, that ensures the proper gluing of overlapping diplanes. This makes it much simpler to build structures using gluing, and is a key element of the PIPATH algorithm [25] for building structures using uniformly labelled two-dimensional ssNMR data. The torsion angle formulas can provide an initial model, but further refinement of the model is needed [20] using both stereochemical restraints and the NMR data.

2. Theory

2.1. Oriented structures and degeneracies

PISEMA, Polarization Inversion Spin Exchange at Magic Angle [8,9], is a 2-dimensional ssNMR experiment that correlates the ^{15}N–^{1}H dipolar coupling and ^{15}N anisotropic chemical shift for each of the labelled residues. A single PISEMA data point can be analyzed to calculate the orientation of the associated peptide plane with respect to $\boldsymbol{B}_0$ [10], but with degeneracies [12,13]. Each degeneracy is a sign degeneracy and can be represented by ± 1. We refer to a structure together with an external vector as an *oriented structure*.

2.1.1. Degeneracies associated with a single oriented peptide plane

The orientation of a peptide plane with respect to $\boldsymbol{B}_0$ corresponding to a single PISEMA data point can be specified by calculating $\boldsymbol{B}_0$ in the peptide plane frame [10]. The $\boldsymbol{B}_0$ coordinates can be expressed as a function of four sign degeneracies $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$ [12,17]. That is, the orientation is determined by the choice of ± 1 for each ε. We work with degeneracies that have been expressed mathematically in the principal axis frame (PAF) of the ^{15}N–^{1}H dipolar coupling interaction [17]. The definition of all the sign degeneracies were given in [17]. For the sake of brevity we do not repeat them here.

The set of possible PISEMA data points corresponding to all $\boldsymbol{B}_0$ orientations is a powder pattern [12,17,20] (Fig. 1). Each data point can have 4, 8 or 12 sign degeneracies depending on its location in the powder pattern [20]. The ε_1 sign degeneracy for the dipolar coupling is resolved in regions A, B and D of the powder pattern. It can also often be resolved in regions C and E if the data is part of a PISA (Polar Index Slant Angle) wheel reflecting an α-helical structure. For simplicity, we assume that ε_1 is resolved everywhere. If we further assume that the PISEMA data point lies in the PISEMA ellipse where the degeneracy is 4-fold (i.e. region A or B in Fig. 1), then $\varepsilon_3 = -\varepsilon_2$ [17]. This assumption holds for transmembrane helices with a tilt angle to the bilayer normal of less than 40°. Thus, the sign degeneracies ε_2 and ε_4 determine the orientations of a peptide plane. For the remainder of this article we assume that this is really the case. Table 1 depicts equivalent ways ([12,13,17,18]) of expressing the four degeneracies associated with a single peptide plane.

2.1.2. Degeneracies associated with an oriented diplane

Two adjacent peptide planes of a protein, corresponding to two consecutive PISEMA data points, constitute a

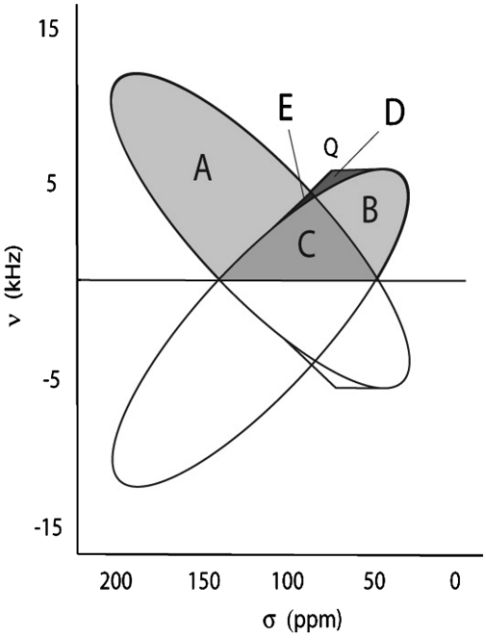

Fig. 1. A typical PISEMA powder pattern bounded by primary and reflected ellipses (PISEMA ellipses) and a small extra-elliptical triangle (PISEMA triangle) near Q. The experimental data $(\sigma, \frac{|\nu|}{2})$ fall within the shaded regions A–E. The number of degeneracies associated with each region of the powder pattern is as follows: regions A and B (4-fold degeneracy), regions C and D (8-fold degeneracy) and region E (12-fold degeneracy). We assume that the angle between the N–H bond vector and σ_{33} is $\alpha = 0^\circ$ and $\beta = 17^\circ$ [17].

Table 1
Equivalent notations for the degeneracies associated with a single peptide plane

Polar angles	Dipolar coupling frame	$(\varepsilon_2, \varepsilon_4)$
(α, β)	(u, v, y)	$(1, 1)$
$(\alpha, \pi - \beta)$	$(u, v, -y)$	$(1, -1)$
$(\pi + \alpha, \beta)$	$(-u, -v, y)$	$(-1, 1)$
$(\pi + \alpha, \pi - \beta)$	$(-u, -v, -y)$	$(-1, -1)$

The degeneracies have been expressed in polar angles (α, β) [13,14] where α is the angle between the NH bond and the projection of $\boldsymbol{B}_0$ on the peptide plane, and β is the angle between the normal to the peptide plane and the direction of $\boldsymbol{B}_0$ (Column 1). They have also been expressed in terms of the coordinates (u, v, y) of $\boldsymbol{B}_0$ [18] (Column 2) or in terms of ε_2 and ε_4 in the principal axis frame of the dipolar coupling interaction (Column 3) [17]. Note ε_2 is the sign of u and ε_4 is the sign of y.

diplane [15,22]. Apart from the two degeneracies that arise from the data for each peptide plane of a diplane, there is an additional degeneracy capturing the chirality of the C_α bond and $\boldsymbol{B}_0$ for the α-carbon connecting the peptide planes [11,22]. Thus, there are altogether five degeneracies associated with the PISEMA data for a diplane, and $2^5 (= 32)$ possible oriented diplanes, consistent with two PISEMA data points given the assumptions described above.

Let $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$ be unit vectors in the direction of the $C_\alpha C$, CN and NC_α bonds of a peptide plane (Fig. 2). Also, let $\overrightarrow{NH}$ be a unit vector in the direction of the NH bond. Then, for a diplane, we associate the following sequence of sign degeneracies:

$$\mathbf{D} = (D_1, D_2, D_3, D_4, D_5), \tag{1}$$

where

$$\begin{aligned} D_1 &= \mathrm{sign}(\boldsymbol{B}_0 \cdot \overrightarrow{\mathrm{NH}}), \\ D_2 &= \mathrm{sign}(\boldsymbol{B}_0 \cdot (\mathbf{b} \times \mathbf{c})), \\ D_3 &= \mathrm{sign}(\boldsymbol{B}_0 \cdot (\mathbf{c} \times \mathbf{a}')), \\ D_4 &= \mathrm{sign}(\boldsymbol{B}_0 \cdot \overrightarrow{\mathrm{NH}}'), \\ D_5 &= \mathrm{sign}(\boldsymbol{B}_0 \cdot (\mathbf{b}' \times \mathbf{c}')). \end{aligned} \tag{2}$$

The primed superscripts for the bond vectors in (2) represent the second peptide plane of a diplane. The sign degeneracies D_1 and D_4 are the ε_2 degeneracies for the first and second peptide planes of a diplane, respectively. Similarly, D_2 and D_5 are the ε_4 degeneracies [17].

Since $\boldsymbol{B}_0 \cdot (\boldsymbol{a}' \times \boldsymbol{b}')$ is used in computing the ψ torsion angle, and $\boldsymbol{B}_0 \cdot (\boldsymbol{b}' \times \boldsymbol{c}')$ is not used for ϕ or ψ, we redefine the degeneracy D_5 in terms of $\boldsymbol{B}_0 \cdot (\boldsymbol{a}' \times \boldsymbol{b}')$. Thus,

$$D_5 = \mathrm{sign}(\boldsymbol{B}_0 \cdot (\mathbf{a}' \times \mathbf{b}')). \tag{3}$$

We next show that if two diplanes share a peptide plane, the degeneracies associated with the two diplanes are not independent.

2.2. The continuity conditions

Three consecutive peptide planes of a protein can be formed by gluing together two oriented diplanes. Let **D** and **D′** denote the degeneracy sequences of consecutive oriented diplanes that share an oriented peptide plane (Fig. 3). Since the two diplanes have a common peptide plane, there is an overlap between the degeneracy sequences. This can be expressed as:

$$\begin{aligned} D_1' &= D_4, \\ D_2' &= -D_5. \end{aligned} \tag{4}$$

We refer to these as the *continuity conditions*. The continuity conditions reduce the number of possible orientations of the two adjacent diplanes to $32 \times 8 (= 256)$ from a total of $32 \times 32 (= 1024)$ possibilities [11]. Note that the negative sign in front of the second equation in (4) is a result of using protein backbone vectors $(\mathbf{a}', \mathbf{b}')$.

Next, suppose that we have a data set of n PISEMA data points corresponding to a protein containing n residues. Since each data point is consistent with multiple orientations for a peptide plane, there are totally $2^5 \times (2^3)^{n-2} = 2^{3n-1}$ protein structures that satisfy the continuity conditions for this data set. Here, as before, we have

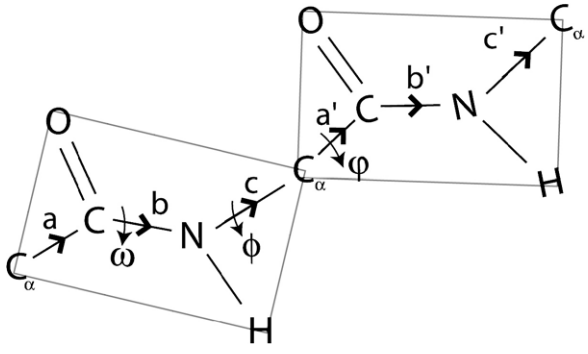

Fig. 2. Torsion angles and bond vectors of a diplane. The bond vectors of the second peptide plane have primed superscripts.

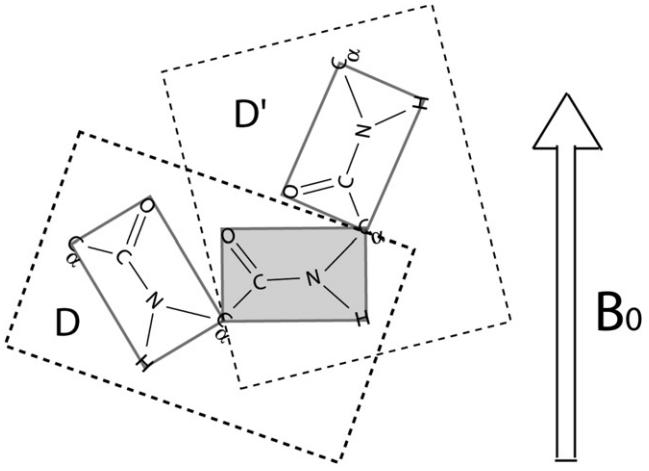

Fig. 3. An oriented structure consisting of three peptide planes obtained by gluing two oriented diplanes with sign degeneracy sequences **D** and **D′**. The diplanes share a peptide plane (shaded gray). All the orientations are with respect to $\boldsymbol{B}_0$.

assumed that each data point lies in either region A or region B of the powder pattern (Fig. 1). For a point in region C or D, the total number of structures increases by a factor of two. For each point that lies in region E of the powder pattern, the total number increases by a factor of three.

3. Discussion

3.1. Torsion angles

The (ϕ,ψ) torsion angles (see Fig. 2) between the two peptide planes of a diplane are internal angles that define the structure of the diplane. The torsion angles do not provide information on the orientation of the diplane relative to $\boldsymbol{B}_0$. However, these angles can be determined from two PISEMA data points, which themselves contain this orientation information [21–24]. Because of degeneracies, there emerge many possible torsion angles. When gluing together two diplanes, one must ensure that the degeneracies are chosen so that the diplanes satisfy the continuity conditions. These conditions would be satisfied automatically if the torsion angles retained orientation information relative to $\boldsymbol{B}_0$, but this is not the case. We next discuss how our continuity conditions are incorporated into the torsion angle formulas.

3.1.1. Torsion angle formulas

We use $\mathrm{tor}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)$ to represent the torsion angle formed from the vectors $\boldsymbol{v}_1,\boldsymbol{v}_2$ and $\boldsymbol{v}_3$. The following formula [22] is used for computing the torsion angle if $\boldsymbol{v}_1,\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ are vectors of length 1:

$$\mathrm{tor}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)=\arg\left(-\boldsymbol{v}_1\cdot\boldsymbol{v}_3+(\boldsymbol{v}_1\cdot\boldsymbol{v}_2)(\boldsymbol{v}_2\cdot\boldsymbol{v}_3)+\mathrm{i}\boldsymbol{v}_1\cdot(\boldsymbol{v}_2\times\boldsymbol{v}_3)\right). \tag{5}$$

From this it follows that

$$\mathrm{tor}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)=\mathrm{tor}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{w})+\mathrm{tor}(-\boldsymbol{w},\boldsymbol{v}_2,\boldsymbol{v}_3), \tag{6}$$

where $\boldsymbol{w}$ is any vector. Expression (6) is especially useful in ssNMR when we take $\boldsymbol{w}=\boldsymbol{B}_0$, since in this case we can only measure orientation angles of bond vectors with respect to the external magnetic field vector. Thus, (6) provides an expression for the backbone torsion angles in terms of torsion angles involving $\boldsymbol{B}_0$.

The triple product in (5) can be written in terms of the Gramian determinant. Let

$$g(x,y,z)=\begin{vmatrix}1 & x & y\\ x & 1 & z\\ y & z & 1\end{vmatrix}=1-x^2-y^2-z^2+2xyz. \tag{7}$$

Then

$$g(\boldsymbol{v}_1\cdot\boldsymbol{v}_2,\boldsymbol{v}_2\cdot\boldsymbol{v}_3,\boldsymbol{v}_3\cdot\boldsymbol{v}_1)=\det[(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)^t(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)]$$

is the *Gramian* determinant [10,22] (where the vectors $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ are taken to be column vectors and the matrix $(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)^t$ is the transpose of the matrix $(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)$. For unit vectors $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$,

$$\boldsymbol{v}_1\cdot(\boldsymbol{v}_2\times\boldsymbol{v}_3)=\varepsilon[g(\boldsymbol{v}_1\cdot\boldsymbol{v}_2,\boldsymbol{v}_2\cdot\boldsymbol{v}_3,\boldsymbol{v}_3\cdot\boldsymbol{v}_1)]^{1/2}, \tag{8}$$

where ε is the chirality of the cross product. Written in this way, the scalar triple products can be computed in terms of known dot products.

3.1.2. Torsion angles for an oriented diplane

We now derive the torsion angle formulas for a diplane as a function of the diplane's degeneracies.

The ϕ and ψ torsion angles for a diplane are given by:

$$\phi=\mathrm{tor}(\boldsymbol{b},\boldsymbol{c},\boldsymbol{a}')\qquad\psi=\mathrm{tor}(\boldsymbol{c},\boldsymbol{a}',\boldsymbol{b}'). \tag{9}$$

We assume the ω torsion angle of each peptide plane of the diplane is 180°. Using (6),

$$\phi=\phi_1+\phi_2\qquad\psi=\psi_1+\psi_2, \tag{10}$$

where

$$\phi_1=\mathrm{tor}(\boldsymbol{b},\boldsymbol{c},\boldsymbol{B}_0)\quad\phi_2=\mathrm{tor}(-\boldsymbol{B}_0,\boldsymbol{c},\boldsymbol{a}'), \tag{11}$$

$$\psi_1=\mathrm{tor}(\boldsymbol{c},\boldsymbol{a}',\boldsymbol{B}_0)\quad\psi_2=\mathrm{tor}(-\boldsymbol{B}_0,\boldsymbol{a}',\boldsymbol{b}'). \tag{12}$$

Using (5),

$$\begin{aligned}\phi_1&=\arg(-\boldsymbol{B}_0\cdot\boldsymbol{b}+(\boldsymbol{B}_0\cdot\boldsymbol{c})(\boldsymbol{b}\cdot\boldsymbol{c})+i\boldsymbol{B}_0\cdot(\boldsymbol{b}\times\boldsymbol{c})),\\ \phi_2&=\arg(\boldsymbol{B}_0\cdot\boldsymbol{a}'-(\boldsymbol{B}_0\cdot\boldsymbol{c})(\boldsymbol{c}\cdot\boldsymbol{a}')-i\boldsymbol{B}_0\cdot(\boldsymbol{c}\times\boldsymbol{a}')),\\ \psi_1&=\arg(-\boldsymbol{B}_0\cdot\boldsymbol{c}+(\boldsymbol{B}_0\cdot\boldsymbol{a}')(\boldsymbol{c}\cdot\boldsymbol{a}')+i\boldsymbol{B}_0\cdot(\boldsymbol{c}\times\boldsymbol{a}')),\\ \psi_2&=\arg(\boldsymbol{B}_0\cdot\boldsymbol{b}'-(\boldsymbol{B}_0\cdot\boldsymbol{a}')(\boldsymbol{a}'\cdot\boldsymbol{b}')-i\boldsymbol{B}_0\cdot(\boldsymbol{a}'\times\boldsymbol{b}')).\end{aligned} \tag{13}$$

The bond orientation cosines and the scalar triple products involving backbone bond vectors of a diplane with $\boldsymbol{B}_0$ in (13) can be written in terms of the degeneracies using the peptide plane geometry and (8) as:

$$\begin{aligned}&\boldsymbol{B}_0\cdot\boldsymbol{b}=D_1\mu_b\quad\text{and}\quad\boldsymbol{B}_0\cdot\boldsymbol{c}=D_1\mu_c,\\ &\boldsymbol{B}_0\cdot\boldsymbol{a}'=D_4\mu_{a'}\quad\text{and}\quad\boldsymbol{B}_0\cdot\boldsymbol{b}'=D_4\mu_{b'},\\ &\boldsymbol{B}_0\cdot(\boldsymbol{b}\times\boldsymbol{c})=D_2\quad g(\boldsymbol{B}_0\cdot\boldsymbol{b},\boldsymbol{b}\cdot\boldsymbol{c},\boldsymbol{c}\cdot\boldsymbol{B}_0)^{1/2},\\ &\boldsymbol{B}_0\cdot(\boldsymbol{c}\times\boldsymbol{a}')=D_3\quad g(\boldsymbol{B}_0\cdot\boldsymbol{c},\boldsymbol{c}\cdot\boldsymbol{a}',\boldsymbol{a}'\cdot\boldsymbol{B}_0)^{1/2},\\ &\boldsymbol{B}_0\cdot(\boldsymbol{a}'\times\boldsymbol{b}')=D_5\quad g(\boldsymbol{B}_0\cdot\boldsymbol{a}',\boldsymbol{a}'\cdot\boldsymbol{b}',\boldsymbol{b}'\cdot\boldsymbol{B}_0)^{1/2},\end{aligned} \tag{14}$$

where

$$\begin{aligned}\mu_a&=p\cos 123^\circ-q\sin 123^\circ,\\ \mu_b&=p\cos 58^\circ-q\sin 58^\circ,\\ \mu_c&=p\cos 117^\circ-q\sin 117^\circ,\\ p&=\sqrt{\frac{2v+v_{||}}{3v_{||}}},\\ q&=\frac{Bp+\sqrt{C(\sigma-\sigma_{22})+fp^2}}{C}\end{aligned} \tag{15}$$

with

$$
\begin{aligned}
f &= B^2 - AC,\\
A &= \sigma_{11}\sin^2\beta + \sigma_{33}\cos^2\beta - \sigma_{22},\\
B &= (\sigma_{11} - \sigma_{33})\sin\beta\cos\beta,\\
C &= \sigma_{11}\cos^2\beta + \sigma_{33}\sin^2\beta - \sigma_{22}.
\end{aligned}
\tag{16}
$$

The equations in (15) and (16) were derived in [17] assuming ideal peptide plane geometry [26]. Here, $(\sigma_{11}, \sigma_{22}, \sigma_{33})$ are the principal values of the ^{15}N chemical shift tensor written in the order of increasing magnitude, β is the angle between the NH bond and the principal axis vector $\boldsymbol{\sigma}_{33}(\approx 17^\circ$[17]), and $\nu_{||}$ is the NH dipolar coupling constant. In (14), μ_a, μ_b and μ_c are functions of the PISEMA data point (σ, ν) as well as the peptide plane geometry. Similar functions corresponding to the next PISEMA data point $(\sigma\prime, \nu\prime)$ are denoted as $\mu_{a'}, \mu_{b'}$ and $\mu_{c'}$. The negative sign in (14) is due to the fact that the direction of $\boldsymbol{b}\times\boldsymbol{c}$ is opposite to the direction of $\boldsymbol{a}\times\boldsymbol{b}$.

There are 32 torsion angle combinations corresponding to the 5 degeneracies $(D_1, D_2, D_3, D_4, D_5)$:

$$
\begin{aligned}
\phi_1 &= \arg(-D_1\mu_b + D_1\mu_c(\boldsymbol{b}\cdot\boldsymbol{c}) + iD_2[g(\mu_b, \mu_c, \boldsymbol{b}\cdot\boldsymbol{c})]^{1/2}),\\
\phi_2 &= \arg(D_4\mu_{a'} - D_1\mu_c(\boldsymbol{c}\cdot\boldsymbol{a}') - iD_3[g(D_1\mu_c, D_4\mu_{a'}, \boldsymbol{c}\cdot\boldsymbol{a}')]^{1/2}),\\
\psi_1 &= \arg(-D_1\mu_c + D_4\mu_{a'}(\boldsymbol{c}\cdot\boldsymbol{a}') + iD_3[g(D_1\mu_c, D_4\mu_{a'}, \boldsymbol{c}\cdot\boldsymbol{a}')]^{1/2}),\\
\psi_2 &= \arg(D_4\mu_{b'} - D_4\mu_{a'}(\boldsymbol{a}'\cdot\boldsymbol{b}') - iD_5[g(\mu_{a'}, \mu_{b'}, \boldsymbol{a}'\cdot\boldsymbol{b}')]^{1/2})
\end{aligned}
\tag{17}
$$

with $\boldsymbol{b}\cdot\boldsymbol{c} = \cos 59^\circ \approx .51, \boldsymbol{c}\cdot\boldsymbol{a}' = \cos 70^\circ \approx .34$ and $\boldsymbol{a}'\cdot\boldsymbol{b}' = \cos 65^\circ \approx .42$ from peptide plane geometry [19,24].

Although (17) implies that there are 32 possible torsion angle combinations for the diplane, only 16 are actually distinct. Because torsion angles alone cannot describe oriented structures, ssNMR experimental data and the torsion angle formulas are unchanged by replacing $\boldsymbol{B}_0$ with $-\boldsymbol{B}_0$. Thus, the same torsion angles are produced if the signs of all the degeneracies are flipped.

3.2. Application of continuity conditions

As an application of the continuity conditions, the tables in Fig. 4 show the possible torsion angles between two adjacent diplanes. If the torsion angles for the first diplane are taken from the second column in table A then by the continuity conditions the choice for the second diplane's torsion angles must be from the first row in table B (matching shade of gray). However, if the choice for the torsion angles is from the third column in table A then the choice for the torsion angle in table B must be from the fourth row (matching shade of gray).

As an example of the use of continuity conditions in computing the torsion angles between adjacent diplanes, consider two adjacent diplanes of a protein corresponding to three consecutive PISEMA data points each of which lies in regions A or B of the powder pattern. For the purpose of illustration, let us focus on a diplane corresponding to the residues ALA2 - ALA3 - ALA4 of a protein. Let the

TABLE A : First Diplane

Peptide Plane 1 (D_1, D_2)	Peptide Plane 2 (D_4, D_5) (1,1)	(1,-1)	(-1,1)	(-1,-1)	D_3
(1,1)	$\gamma_1\ \delta_1$ / $-\gamma_2\ -\delta_2$	$\gamma_1\ \delta_2$ / $-\gamma_2\ -\delta_1$	$\gamma_3\ -\delta_3$ / $\gamma_4\ -\delta_4$	$\gamma_3\ \delta_4$ / $\gamma_4\ \delta_3$	1 / -1
(1,-1)	$\gamma_2\ \delta_1$ / $-\gamma_1\ -\delta_2$	$\gamma_2\ \delta_2$ / $-\gamma_1\ -\delta_1$	$-\gamma_4\ -\delta_3$ / $-\gamma_3\ -\delta_4$	$-\gamma_4\ \delta_4$ / $-\gamma_3\ \delta_3$	1 / -1
(-1,1)	$-\gamma_3\ \delta_3$ / $-\gamma_4\ \delta_4$	$-\gamma_3\ -\delta_4$ / $-\gamma_4\ -\delta_3$	$-\gamma_1\ -\delta_1$ / $\gamma_2\ \delta_2$	$-\gamma_1\ -\delta_2$ / $\gamma_2\ \delta_1$	1 / -1
(-1,-1)	$\gamma_4\ \delta_3$ / $\gamma_3\ \delta_4$	$\gamma_4\ -\delta_4$ / $\gamma_3\ -\delta_3$	$-\gamma_2\ -\delta_1$ / $\gamma_1\ \delta_2$	$-\gamma_2\ -\delta_2$ / $\gamma_1\ \delta_1$	1 / -1

TABLE B : Second Diplane

Peptide Plane 2 (D'_1, D'_2)	Peptide Plane 3 (D'_4, D'_5) (1,1)	(1,-1)	(-1,1)	(-1,-1)	D'_3
(1,1)	$\alpha_1\ \beta_1$ / $-\alpha_2\ -\beta_2$	$\alpha_1\ \beta_2$ / $-\alpha_2\ -\beta_1$	$\alpha_3\ -\beta_3$ / $\alpha_4\ -\beta_4$	$\alpha_3\ \beta_4$ / $\alpha_4\ \beta_3$	1 / -1
(1,-1)	$\alpha_2\ \beta_1$ / $-\alpha_1\ -\beta_2$	$\alpha_2\ \beta_2$ / $-\alpha_1\ -\beta_1$	$-\alpha_4\ -\beta_3$ / $-\alpha_3\ -\beta_4$	$-\alpha_4\ \beta_4$ / $-\alpha_3\ \beta_3$	1 / -1
(-1,1)	$-\alpha_3\ \beta_3$ / $-\alpha_4\ \beta_4$	$-\alpha_3\ -\beta_4$ / $-\alpha_4\ -\beta_3$	$-\alpha_1\ -\beta_1$ / $\alpha_2\ \beta_2$	$-\alpha_1\ -\beta_2$ / $\alpha_2\ \beta_1$	1 / -1
(-1,-1)	$\alpha_4\ \beta_3$ / $\alpha_3\ \beta_4$	$\alpha_4\ -\beta_4$ / $\alpha_3\ -\beta_3$	$-\alpha_2\ -\beta_1$ / $\alpha_1\ \beta_2$	$-\alpha_2\ -\beta_2$ / $\alpha_1\ \beta_1$	1 / -1

Fig. 4. Backbone torsion angles obtained from the combination of degeneracies corresponding to two adjacent diplanes of a protein. Table A depicts all possible torsion angles (γ, δ) for the first diplane. Table B shows all possible torsion angles (α, β) for the second diplane. Out of 32 possible torsion angles in each table, 16 are unique. Suppose the choice of torsion angles of the first diplane is from one of the columns of table A (highlighted by different shades of gray). Then, the continuity conditions restrict the choice of the torsion angles for the second diplane to a row of table B highlighted by the matching shade of gray. The arrow indicates one such possibility, i.e., if the torsion angles for a diplane are from the second column in table A then the torsion angles for the immediately following diplane must be from the first row of table B.

diplane following it correspond to the residues ALA3 - ALA4 - ALA5. Suppose that the PISEMA data points corresponding to the residues ALA2 - ALA3, ALA3 - ALA4 and ALA4 - ALA5 are (171.79 ppm, 7.37 kHz), (142.72 ppm, 9.1 kHz) and (45.27 ppm, 5.15 kHz), respectively. Also, $\sigma_{11} = 30$ ppm, $\sigma_{22} = 60$ ppm, $\sigma_{33} = 205$ ppm and $\nu_{||} = 11.335$ kHz. Now, suppose we choose the sign degeneracy sequence associated with the first diplane, i.e., $\mathbf{D}$, to be $(-1, 1, -1, -1, 1)$, then from (17), the (ϕ, ψ) angles are $(-52.64^\circ, -50.25^\circ)$. There are 15 additional choices for $\mathbf{D}$ that would produce unique torsion angles. The sign degeneracy sequence of the second diplane, $\mathbf{D}'$, is restricted by the continuity conditions (4). Because of this, the choice of $\mathbf{D} = (-1, 1, -1, -1, 1)$ limits $\mathbf{D}'$ to eight possibilities i.e., $\mathbf{D}' = (-1, -1, \pm 1, \pm 1, \pm 1)$, all of which satisfy the continuity conditions for adjacent

diplanes. If $\mathbf{D}' = (-1, -1, 1, 1, 1)$ is used, then the (ϕ, ψ) angles of the second diplane come out as (-75.7°, -89.5°).

4. Conclusions

In this paper, the continuity conditions for adjacent diplanes of a protein have been quantified. The degeneracies associated with a diplane have been incorporated into equations for torsion angles, and the continuity conditions were rewritten so that they could be used with the torsion angle formulas. These formulas are in a convenient form for algorithms such as PIPATH [25] which derive plausible protein structures directly from PISEMA data. The formulas in fact form an essential aspect of the PIPATH algorithm in the process of deriving protein structures from uniformly labelled two-dimensional ssNMR data. The formulas make the whole process of protein structure building quite simple. It is important to note that the torsion angle formulas provide only an initial model of a protein. Further refinement of the model [20] using both stereochemical restraints and the NMR data needs to be carried out to obtain final feasible model structure/ structures of a protein molecule. Our approach to computing the torsion angles is very general and can be applied effectively to determine the backbone torsion angles of any secondary protein structure. In particular, we do not require the secondary structure to be an α-helix.

To find the structure of a protein by ssNMR experiments, such as PISEMA, the data must first be assigned [25]. The problem here is to match each resonance peak in the experimental data with a residue in the amino acid sequence of the protein or peptide under investigation. In [27], the assignment problem and the protein structure were solved simultaneously when the secondary structure of the protein was known. In [25], α-helical assignments are obtained before solving the protein structure. Often, only partially assigned data are available. In these situations, using the spectral frequencies as well as sign degeneracies the orientation of peptide planes corresponding to the assigned resonances can be determined. These lead to solving the isolated segments of a protein. To combine various isolated segments of a protein along with unassigned resonances, diplanes must be glued. An important implication of the present work is that diplanes not satisfying the continuity conditions cannot be glued together, and thus are not adjacent to each other.

Wang and Donald showed [30] that the protein backbone of secondary structures within a large protein can be built in a sequential approach (a.k.a. iterative technique) using residual dipolar coupling (RDC) data obtained from two aligning media. In particular, they computed the N–H bond vector for each specifically labelled residue of a protein. Since they obtained two data points for each residue, they were able to minimize a function that splits the difference between the 2 aligning media. This vector established the reference frame for further computations. They then computed all the subsequent N–H vectors by using the iterative technique. In the gluing approach to protein structure modelling the molecular backbone structure of a protein is obtained by gluing together the consecutive diplanes, as in the PIPATH algorithm. The PIPATH algorithm allows for building of peptides in arbitrary residue order as opposed to a sequential order. The gluing of adjacent diplanes introduces ''continuity conditions'' that are automatically satisfied when using the sequential technique.

To build an initial model of a protein structure it is necessary to compute the backbone torsion angles. We derive torsion angle formulas that incorporate the continuity conditions. This is important because the formulas make the process of building structures using gluing method simpler. Moreover, they form a key element of the PIPATH algorithm for building structures using uniformly labeled two-dimensional ssNMR data.

The full potential of ssNMR derived structure is only achieved when all of the sign degeneracies are completely resolved. However, many degeneracies may lead to minor modifications of the structure, such that a moderate resolution structure is achievable without resolving all of the sign degeneracies [28]. Furthermore, because of the absolute nature of these orientational restraints, errors from one peptide plane (e.g. degeneracy) do not propagate throughout the structure [23]. For a protein comprising n residues, the continuity conditions significantly reduce the number of sign degeneracies that need to be resolved. This number is typically reduced further by incorporating additional structural information. This can be achieved, for example, by studying the dipolar interactions of ^{15}N–^{1}H, ^{15}N–^{13}C and $^{13}C_{\alpha}$-^{1}H for each peptide plane of the diplane [15] or by studying the dipolar interaction of ^{15}N–^{1}H, ^{15}N–^{13}C and chemical shift of ^{15}N. The $^{13}C_{\alpha}$-^{2}H quadrupolar splittings can as well be used as effective filters to reduce the degeneracies due to the central alpha-carbon between the peptide planes of a diplane [23].

For helical proteins, the symmetry properties of PISA wheels can be utilized to resolve the peptide plane sign degeneracies [13,18]. The (ϕ, ψ) torsion angles can be calculated once the orientation of the corresponding peptide planes are predicted from the helical tilt and rotation angles. In this case, the position of each data point on the helical wheel determines the peptide plane orientation. The continuity conditions between adjacent diplanes of a protein are satisfied automatically here. However, when PISA wheel patterns are not clear, the peptide plane degeneracies reappear. In such situations, we should glue together only those diplanes that satisfy the continuity conditions.

In [23], the ^{15}N–^{1}H, ^{15}N-^{13}C dipolar couplings and ^{15}N chemical shifts for peptide planes of gramicidin A were used to compute the possible orientations for each diplane of the protein. As a result of measuring the C_{α}-^{2}H quadrupolar couplings for the C_{α} between adjacent peptide planes, the number of possible orientations for each diplane were reduced from 16 to 4. The model structure of gramicidin A was then built by combining diplanes that share peptide

planes that are identically aligned with an external vector. This guaranteed that the continuity conditions between diplanes were satisfied. Furthermore, it should be noted that the remaining degeneracies were all consistent with a single fold of the polypeptide and a single set of hydrogen bonds [28]. The continuity conditions were explained qualitatively in figures 2 and 4 of [23]. In the current work we have derived the continuity conditions in terms of the sign degeneracies associated with each diplane. This kind of quantification of the continuity conditions is quite essential when building structures using the computer algorithms such as PIPATH [25].

Acknowledgments

This work was supported by American Heart Association Grant 0415075B (T.A and R.B.) and National Institutes of Health Grant P01-GM64676 (J.R.Q., R.B. and T.A.C.) and National Institutes of Health Grant ROI-AI23007 (T.A.C.).

References

[1] R. Fu, T.A. Cross, Solid state NMR investigation of protein and polypeptide structure, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 235–268.
[2] F.M. Marassi, S.J. Opella, NMR structural studies of membrane proteins, Curr.Opin.Struct.Biol. 8 (1998) 640–648.
[3] R.R. Ketchem, W. Hu, T.A. Cross, High resolution conformation of gramicidin A in a lipid bilayer by solid state NMR, Science 261 (1993) 1457–1460.
[4] J. Wang, S. Kim, F. Kovacs, T.A. Cross, Structure of the transmembrane region of the M2 protein H^+ channel, Protein Sci. 10 (2001) 2241–2250.
[5] Y. Kim, K. Valentine, S.J. Opella, S.L. Schendel, W.A. Cramer, Solid-state NMR studies of the membrane-bound closed state of the colicin E1 channel domain in lipid bilayers, Protein Sci. 7 (1998) 342–348.
[6] F.M. Marassi, J.J. Gesell, A.P. Valente, Y. Kim, M. Oblatt-Montal, M. Montal, S.J. Opella, Dilute spin-exchange assignment of solid-state NMR spectra of oriented proteins: acetylcholine M2 in bilayers, J. Biomol. NMR 14 (1999) 141–148.
[7] S.H. Park, A.A. Mrse, A.A. Nevzorov, M.F. Mesleh, M. Oblatt-Montal, M. Montal, S.J. Opella, Three-dimensional structure of the channel-forming trans-membrane domain of virus protein "u" (Vpu) from HIV-1, J. Mol. Biol. 333 (2003) 409–424.
[8] C.H. Wu, A. Ramamoorthy, S.J. Opella, High-resolution heteronuclear dipolar solid-state NMR spectroscopy, J. Magn. Reson. 109 (1994) 270–282.
[9] A. Ramamoorthy, Y. Wei, D. Lee, PISEMA solid state NMR spectroscopy, Ann. Reports on NMR Spectrosc. 52 (2004) 1–52.
[10] J.R. Quine, T.A. Cross, Protein structure in anisotropic environments: unique structural fold from orientational constraints, Concepts Magn. Reson. 12 (2000) 71–82.
[11] T.A. Cross, Solid state nuclear magentic resonance approach for determining the structure of gramicidin A without model fitting, Biophys. J. 49 (1986) 124–126.
[12] J. Denny, J. Wang, T.A. Cross, J.R. Quine, PISEMA powder patterns and PISA wheels, J. Magn. Reson. 152 (2001) 217–226.
[13] R. Tycko, P.L. Stewart, S.J. Opella, Peptide plane orientations determined by fundamental and overtone ^{14}N NMR, J. Am. Chem. Soc. 108 (1986) 5419–5425.
[14] F.M. Marassi, S.J. Opella, Using PISA pies to resolve ambiguities in angular constraints from PISEMA spectra of aligned proteins, J. Biomol. NMR 23 (2002) 239–242.
[15] M.T. Brenneman, T.A. Cross, A method for the analytic determination of polypeptide structure using solid state nuclear magnetic resonance: The "metric method", J. Chem. Phys. 92 (2) (1990) 1483–1494.
[16] T.A. Cross, J.R. Quine, Protein structure in anisotropic environments: development of orientational constraints, Concepts Magn. Reson. 12 (2000) 55–70.
[17] J.R. Quine, S. Achuthan, T. Asbury, R. Bertram, M. Chapman, J. Hu, T.A. Cross, Intensity and mosaic spread analysis from PISEMA tensors in solid-state NMR, J. Magn. Reson. 179 (2006) 190–198.
[18] F.M. Marassi, S.J. Opella, Simultaneous assignment and structure determination of a membrane protein from NMR orientational restraints, Protein Sci. 12 (2003) 403–411.
[19] R.A. Engh, R. Huber, Accurate bond and angle parameters for X-ray protein structure refinement, Acta Crystallogr. A47 (1991) 392–400.
[20] R. Bertram, T. Asbury, F. Fabiola, J.R. Quine, T.A. Cross, M.S. Chapman, Atomic refinement with correlated solid-state NMR restraints, J. Magn. Reson. 163 (2003) 300–309.
[21] T.A. Cross, S.J. Opella, Protein structure by solid state nuclear magnetic resonance: Residues 40 to 45 of bacteriophage fd coat protein, J. Mol. Biol. 182 (1985) 367–381.
[22] J.R. Quine, M.T. Brenneman, T.A. Cross, Protein structural analysis from solid-state NMR-derived orientational constraints, Biophys. J. 72 (1997) 2342–2348.
[23] R.R. Ketchem, K.C. Lee, S. Huo, T.A. Cross, Macromolecular structural elucidation with solid-state NMR derived orientational constraints, J. Biomol. NMR 8 (1996) 1–14.
[24] Z. Song, F.A. Kovacs, J. Wang, J.K. Denny, S.C. Sekhar, J.R. Quine, T.A. Cross, Transmembrane domain of M2 protein from influenza A virus studied by solid-state ^{15}N polarization inversion spin exchange at magic angle NMR, Biophys. J. 79 (2000) 767–775.
[25] T. Asbury, J.R. Quine, S. Achuthan, J. Hu, M.S. Chapman, T.A. Cross, R. Bertram, PIPATH: An optimized alogrithm for generating α-helical structures from PISEMA data, J. Magn. Reson. 183 (2006) 87–95.
[26] W. Mai, W. Hu, C. Wang, T.A. Cross, Orientational constraints as three-dimensional structure constraints from chemical shift anisotropy-the polypeptide backbone of gramicidin-A in a lipid bilayer, Protein Sci. 2 (1993) 532–542.
[27] A.A. Nevzorov, S.J. Opella, Structural fitting of PISEMA spectra of aligned proteins, J. Magn. Reson. 163 (2003) 300–309.
[28] R.R. Ketchem, B. Roux, T.A. Cross, High resolution polypeptide structure in a lamellar phase lipid environment from solid state nmr derived orientational constraints, Structure 5 (1997) 1655–1669.
[29] J. Hu, T. Asbury, S. Achuthan, R. Bertram, J.R. Quine, T.A. Cross, Backbone structure of the Amantadine-blocked trans-membrane domain M2 proton channel from Influenza A Virus, Biophys. J. 92 (2007) 4335–4343.
[30] L. Wang, B.R. Donald, Exact solutions for internuclear vectors and backbone dihedral angles from NH residual dipolar couplings in two media, and their application in a systematic search algorithm for determining protein backbone structure, J. Biomol. NMR 29 (2004) 223–242.