Journal of Biomolecular NMR, 8 (1996) 1–14
ESCOM

J-Bio NMR 356

Macromolecular structural elucidation with solid-state NMR-derived orientational constraints

R.R. Ketchem, K.-C. Lee*, S. Huo** and T.A. Cross***

Center for Interdisciplinary Magnetic Resonance at the National High Magnetic Field Laboratory, Institute of Molecular Biophysics and Department of Chemistry, Florida State University, Tallahassee, FL 32306, U.S.A.

Received 20 December 1995
Accepted 25 March 1996

Keywords: Gramicidin A; Ion channels; ^{2}H NMR; Side-chain conformation; Membrane protein structure; Oriented bilayers

Summary

The complete structure determination of a polypeptide in a lipid bilayer environment is demonstrated built solely upon orientational constraints derived from solid-state NMR observations. Such constraints are obtained from isotopically labeled samples uniformly aligned with respect to the B_0 field. Each observation constrains the molecular frame with respect to B_0 and the bilayer normal, which are arranged to be parallel. These constraints are not only very precise (a few tenths of a degree), but also very accurate. This is clearly demonstrated as the backbone structure is assembled sequentially and the i to i+6 hydrogen bonds in this structure of the gramicidin channel are shown on average to be within 0.5 Å of ideal geometry. Similarly, the side chains are assembled independently and in a radial direction from the backbone. The lack of considerable atomic overlap between side chains also demonstrates the accuracy of the constraints. Through this complete structure, solid-state NMR is demonstrated as an approach for determining three-dimensional macromolecular structure.

Introduction

The characterization of protein structure in anisotropic environments is one of the current frontiers in the field of structural biology. Solid-state NMR spectroscopy (SS-NMR) of such biological systems has been developed in recent years to respond to this challenge. Both distance (Gullion and Schaefer, 1989; Griffiths and Griffin, 1993) and orientational (Cross and Opella, 1983; Shon et al., 1991; Ketchem et al., 1993) constraints have been demonstrated as quantitative structural constraints in protein and polypeptide structures, as recently reviewed for membrane-bound systems (Cross, 1994; Cross and Opella, 1994). While numerous examples of these constraints have been reported, no complete structures have heretofore been achieved. The method for analyzing orientational constraints to achieve a complete atomic-resolution backbone and side-chain description of a polypeptide, gramicidin A, in a hydrated lipid bilayer is presented here. The complete structure of this channel in a lipid environment has not been achieved previously.

A new approach is needed to attack protein structures in anisotropic environments. The formation of three-dimensional crystals of membrane proteins has proven to be exceptionally difficult and consequently few X-ray crystal structures have been achieved for such proteins (Michel and Deisenhofer, 1990; Weiss et al., 1991; Schiffer et al., 1992; Kuhlbrandt et al., 1994). When it has been possible to prepare two-dimensional crystals, these have been used to characterize membrane proteins through electron diffraction (Henderson et al., 1990; Wang and Kuhlbrandt, 1991; Jap et al., 1991; Kuhlbrandt et al., 1994). Solution NMR has also had limited success because the detergents, or lyso-lipid molecules that are used to solubilize membrane proteins, add considerably to the protein molecular weight and increase the global correlation time. Therefore, solution NMR has been restricted to small membrane proteins such as glucagon (Braun et al., 1983),

*Present address: Department of Chemistry and Biochemistry, University of Maryland, P.O. Box 223, College Park, MD 20742, U.S.A.
**Present address: MB-2, The Scripps Research Institute, 10666 N. Torrey Pines Road, La Jolla, CA 92037, U.S.A.
***To whom correspondence should be addressed.

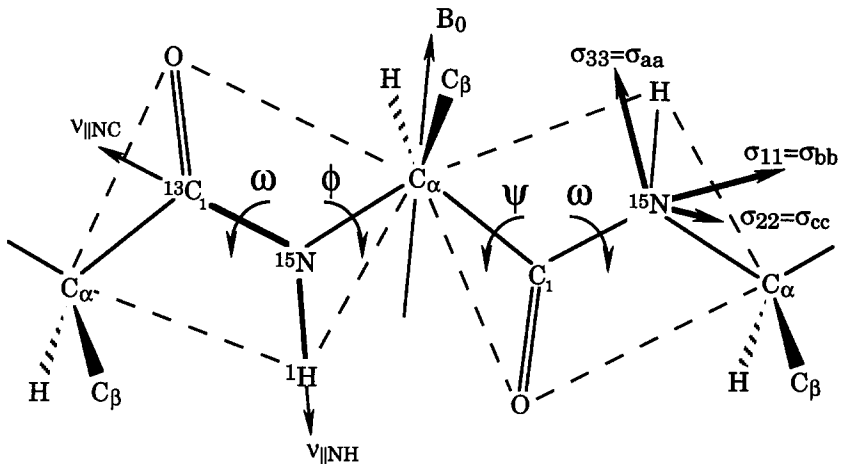

Fig. 1. Linked peptide planes. The peptide plane can be defined by two bond vectors, N-H and N-C^1. The orientation of these vectors with respect to the external magnetic field determined from dipolar interactions defines the orientation of the peptide plane. Once the orientations of the individual peptide planes have been determined with respect to the external magnetic field, the orientations of the planes with respect to each other is determined by joining the planes through the shared C$^\alpha$ carbon which has a well-defined covalent geometry. This allows for the determination of the (ϕ,ψ) torsion angles. The ^{15}N chemical shift interaction presented here in both its frequency-dependent (σ_{11}, σ_{22}, σ_{33}) and frequency-independent (σ_{aa}, σ_{bb}, σ_{cc}) labels (Teng et al., 1992) is needed as a filter to eliminate structural ambiguities.

calcitonin (Morelli et al., 1992) and the filamentous viral coat proteins (McDonnell et al., 1993). SSNMR has been shown to provide not only precise, but accurate structural constraints. These constraints have been difficult to obtain in quantity, but several research groups are addressing this problem. We have previously shown how (ϕ,ψ) torsion angles can be determined from peptide plane dipolar and chemical shift interactions (Brenneman and Cross, 1990; Teng et al., 1991). Here, we describe how to combine the structural constraints from numerous sites and how to minimize conformational ambiguities, resulting in a complete structural solution.

Nuclear spin interactions, such as chemical shift, dipolar and quadrupolar interactions have a well-defined orientation dependence with respect to the magnetic field direction in anisotropic environments. In unoriented samples this leads to powder-pattern lineshapes, whereas in oriented samples the observed sharp resonance can be used to constrain the orientation of the nuclear spin interaction tensor with respect to the magnetic field. If the tensor orientation is known with respect to the molecular frame, then the resonance can be used to constrain the molecular frame with respect to the magnetic field direction. For heteronuclear dipolar and carbon-bound ^{2}H quadrupolar interactions the unique axis of these nearly axially symmetric spin interactions is aligned with the internuclear vector, for instance the dipolar interaction between ^{15}N of one amino acid and the $^{13}C_1$ carbon of an adjacent amino acid has the ^{15}N-^{13}C heteronuclear dipolar interaction axis aligned with this peptide bond direction. Axially asymmetric interaction tensors, such as ^{15}N and ^{13}C chemical shifts have been extensively characterized, both by means of single-crystal studies (Stark et al., 1983; Harbison et al., 1984) and by relative alignment of spin interaction tensors (Hartzell et al., 1987; Valentine et al., 1987; Teng et al., 1992). From the determination of numerous ^{15}N chemical shift tensor orientations in the backbone of gramicidin A, it appears that it is not necessary to experimentally characterize the tensors for each site of interest (Mai et al., 1993). The variability in tensor orientation is less than 2°, a tolerable error for the structural constraints.

The interpretation of orientational constraints not only requires a knowledge of the tensor orientation, but also of the motional averaging of the tensor. The use of powder-pattern spectra to measure the tensor-element magnitudes and to characterize the local and global dynamics for a nuclear site is well-known. Once the global dynamics are characterized, as they have been for gramicidin in hydrated lipid bilayers (Lee et al., 1993), these parameters can be applied to all interactions and all sites within the channel. This leaves the local dynamics which may vary from site to site. However, such dynamics within a lipid bilayer are difficult to characterize independently of the global motions. By lowering the temperature of a bilayer preparation below the gel-to-liquid-crystalline phase transition, the global motion can be stopped, but a phase change of the lipids induces distortions in the bilayer. Consequently, methods have been developed to fast freeze thin films of the bilayers and in this way powder-pattern spectra can be obtained with little heterogeneous broadening of the discontinuities (Lazo et al., 1992; Evans et al., 1993). Spectra below 200 K yield static tensor-element magnitudes. The temperature dependence of the powder-pattern spectra up to the phase transition has been used to characterize details of local librational dynamics that include a description of the motional axis and amplitude (Hu et al., 1995; Lazo et al., 1995). The detailed tensor

characterizations not only permit accurate interpretation of orientational constraints, but they also generate detailed dynamic characterizations that can lead to dynamic-function correlations (Hu and Cross, 1995; North and Cross, 1995).

The (ϕ,ψ) torsion angles have been determined for independent pairs of peptide planes by utilizing two dipolar orientational constraints, the ^{15}N-^{1}H and ^{15}N-$^{13}C^{1}$ couplings, and the ^{15}N chemical shift from each peptide plane. Initially, the peptide plane is assumed to be planar ($\omega = 180°$). The orientations of the N-H and N-C^{1} bonds define the orientation of the peptide plane with respect to the magnetic field, and since the sample is uniformly aligned, the orientation of peptide planes is also defined with respect to a unique molecular axis (Fig. 1). The initial structure is built by linking individual peptide planes together, utilizing the tetrahedral geometry about the C^{α} carbon which joins the planes and utilizing the C^{α}-^{2}H orientation. In doing so, the relative orientation of the peptide planes is defined, i.e. the (ϕ,ψ) torsion angles about the C^{α} carbon are determined, and simultaneously the planes are independently defined with respect to the magnetic field direction and to the unique molecular axis.

While the torsion angles are determined the solution set is not unique. The orientation dependence of the dipolar interaction is given by $\Delta\nu_{obs} = \nu_{\parallel}(3\cos^2\theta - 1)$, where $\Delta\nu_{obs}$ is the observed dipolar splitting, $\nu_{\parallel}$ is the interaction magnitude and θ is the angle of the internuclear vector with respect to the magnetic field direction. The sign of $\Delta\nu_{obs}$ is only known when $|\Delta\nu_{obs}/\nu_{\parallel}|$ is greater than 1. This sign ambiguity will lead to different orientational solutions, as will the undefined sign of $\cos\theta$. For helices in lipid bilayers, aligned such that the bilayer normal is parallel to the magnetic field, one of the two dipolar observations will have a positive sign. Furthermore, many possible combinations of the N-H and N-C^{1} bond orientations are not consistent with the known covalent geometry (i.e. bond angles) of the peptide plane and the combination of many diplane possibilities are not consistent with the covalent geometry about the C^{α} carbon. Moreover, the observed ^{15}N chemical shifts for the two planes and the central C^{α}-^{2}H quadrupolar splitting can be used as filters to reduce the solution set. In this way a unique, or nearly unique orientation for each peptide plane can be obtained (see Fig. 2), albeit with a remaining sign ambiguity in the orientation of the peptide plane normal with respect to the magnetic field (Brenneman and Cross, 1990; Teng et al., 1991; Brenneman et al., unpublished results). In this report, as multiple residues are added to the backbone structure, the remaining (ϕ,ψ) ambiguities are analyzed. In addition, the torsion angles for each of the side chains are determined from orientational constraints.

To illustrate this method, the structure of the cation-channel-forming polypeptide, gramicidin A, is solved. This polypeptide has an alternating stereochemical sequence of amino acids with both termini blocked: HCO-Val[1]-Gly[2]-Ala[3]-D-Leu[4]-Ala[5]-D-Val[6]-Val[7]-D-Val[8]-Trp[9]-D-Leu[10]-Trp[11]-D-Leu[12]-Trp[13]-D-Leu[14]-Trp[15]-$NHCH_2CH_2OH$. The backbone structure has been previously reported (Ketchem et al., 1993), but a generalized method for solving the structure has not been reported and neither has the problem of the structural ambiguities been adequately addressed.

Materials and Methods

The approach for structure determination described here relies upon orientational constraints obtained from local nuclear spin interactions throughout the polypeptide. The method requires samples that are uniformly aligned with respect to the magnetic field and for which the local nuclear spin interactions can be resolved. Samples containing isotopically labeled gramicidin A (gA) in hydrated dimyristoyl phosphatidylcholine (DMPC) bilayers were prepared by codissolving gA with DMPC (1:8 molar ratio) in a 95%-benzene/5%-ethanol solvent system (Mai et al., 1993). This solution is spread on 25 thin glass slips, obtained from Marienfeld Laboratory Glassware (Bad Mergentheim, Germany), measuring 5 mm × 20 mm with a thickness in the range of 0.06–0.08 mm and the slips are vacuum-dried. They are then stacked into a square glass tube and approximately 50%-by-weight water is added. The tube is fully sealed and incubated at 40 °C for a minimum of two weeks, resulting in a sample containing gA in fully hydrated and oriented lipid bilayers. The orientation of the gA channel axis is parallel to the normal of the glass slips and the lipid bilayers (Smith and Cornell, 1986; Fields et al., 1988). The orientational mosaic spread can be assessed from the resonance lineshapes and has been shown to be as little as 0.3° (Cross et al., 1992). This provides a well-oriented sample so that precise orientational constraints can be obtained.

Isotopically labeled amino acids were obtained from Cambridge Isotope Labs; the Fmoc blocking chemistry and peptide synthesis were performed as described previously (Fields et al., 1989). The labeled leucine amino acids were obtained as DL racemic mixtures. The D-amino acid was partially resolved by a combined chemical and enzymatic procedure (Teng et al., 1991). The resulting impure peptides were purified to more than 98% purity using a semipreparative HPLC protocol, designed to resolve peptides having a strict alternating pattern of D and L residues (Fields et al., 1989). By using isotopically labeled peptides, specific-site local interactions can be resolved without assignment ambiguity.

The ^{2}H NMR data were acquired using a spectrometer built around a Chemagnetics data acquisition system and an Oxford Instruments 400/89 magnet. Spectra were obtained using a quadrupole echo pulse sequence with a

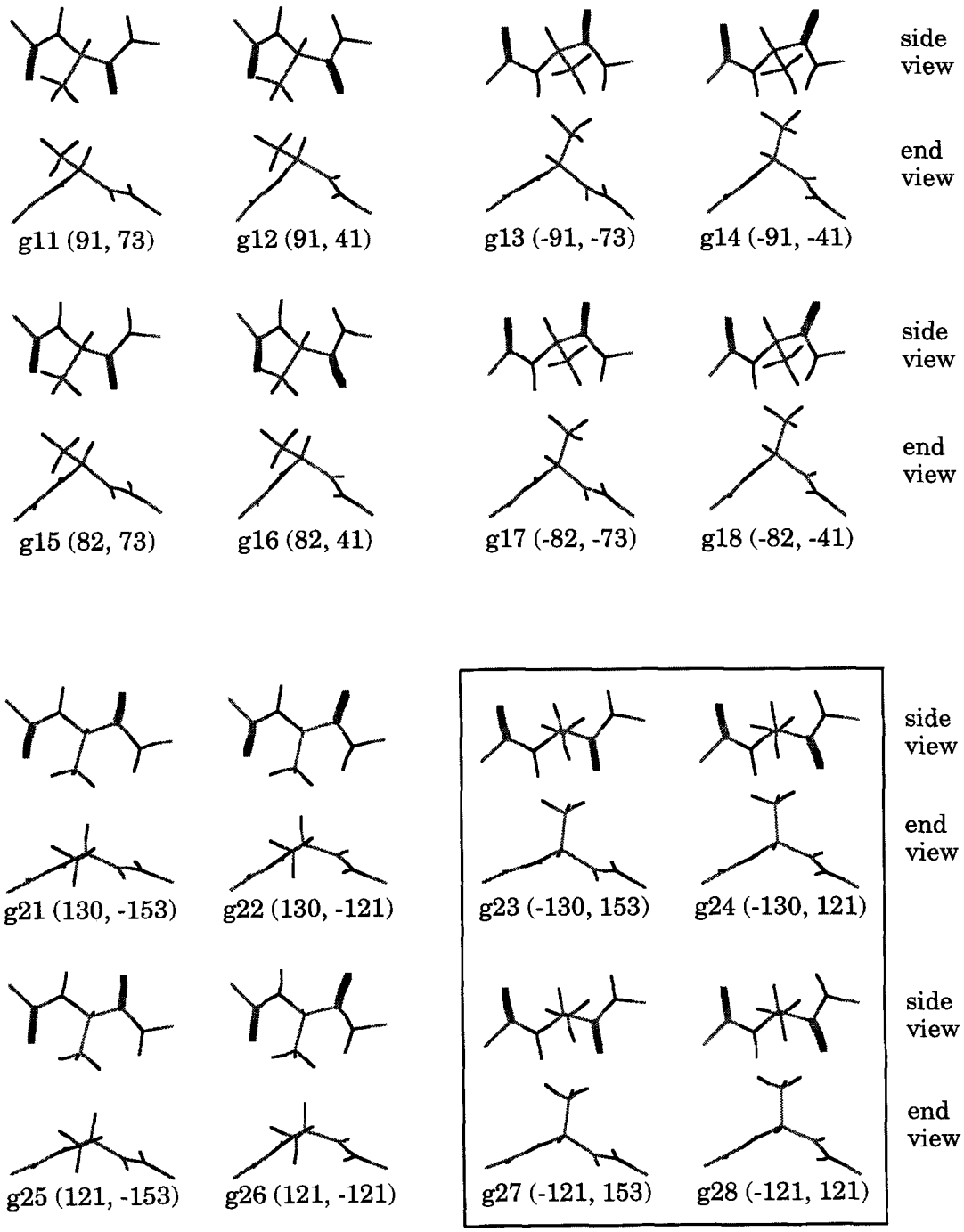

Fig. 2. All sixteen possible solutions for the diplane centered on the Ala[3] C^{α} carbon consistent with the ^{15}N-^{1}H and ^{15}N-$^{13}C'$ dipolar splittings and the ^{15}N chemical shifts for both planes. The carbonyls are displayed as thickened bonds. Twelve of the solutions are inconsistent with the observed C^{α}-^{2}H quadrupolar splitting (195.4 kHz) for the central C^{α} carbon. These invalid solutions have predicted C^{α}-^{2}H quadrupolar splittings that deviate by 56 to 170 kHz from the observed value. The remaining diplanes, g23, g24, g27 and g28, have a predicted C^{α}-^{2}H quadrupolar splitting value of 206 kHz, only 11 kHz from the observed value. The remaining solutions all have similar torsion angles that fall within the β-strand region of the Ramachandran diagram. Consequently, the folding motif is uniquely defined for this one site, despite the four torsional solutions.

30-μs echo delay and 2.8-μs 90° pulse widths. Typical recycle delays for the hydrated samples were 0.7 s. Prior to the interpretation of the signals from oriented samples the nuclear spin interaction tensors were characterized. For the ^{2}H NMR data presented in this report the unique axis of the static carbon-bound ^{2}H quadrupolar interaction is assumed to lie along the internuclear vector. The motional averaging of this interaction was characterized

TABLE 1
POSSIBLE Ala^3 DIPLANE DIRECTION COSINE COMBINATIONS

Bond vector combination	Plane 1		Plane 2	
	^{15}N-^{1}H	^{15}N-^{13}C	^{15}N-^{1}H	^{15}N-^{13}C
g1	0.9871	−0.3970	0.9160	−0.7405
g2	0.9871	−0.3970	−0.9160	0.7405

from unoriented preparations of gA in hydrated lipid bilayers prepared as described above for oriented samples, except that the organic solution was not spread on glass plates. Hydration to 40% of the sample weight for specific-site-deuterated samples was done with deuterium-depleted water.

The spectral data were processed using Felix from Biosym Technologies (San Diego, CA). The individual (ϕ,ψ) torsion angles were calculated with the program CNFCS (CoNFormation with reduction by Chemical Shift), developed in our lab, essentially as previously described (Brenneman and Cross, 1990; Teng et al., 1991), except that the ^{15}N chemical shift has now been introduced as a direct constraint on the possible peptide plane orientations before (ϕ,ψ) solutions were calculated. Diplane coordinates were calculated by using the diplane direction cosine solutions along with the (ϕ,ψ) solutions through the program COORDS, developed in our lab.* The diplanes were visualized and paired using Insight II from Biosym Technologies. All computer work was performed on a Silicon Graphics Indigo 2 Extreme.

Results

The potential solution set for the peptide plane orientations on either side of the Ala^3 C^α carbon and the Ala^3 torsion angles are shown in Table 1 to illustrate the structural ambiguities and their resolution. For the Gly^2-Ala^3 peptide plane the N-H orientation is such that $\Delta\nu_{obs}$ is positive and $\cos\theta$ is assumed to be positive. The consequence of this one-time assumption is that either the structure of the top or bottom monomer will be determined and all subsequent diplanes are chosen such that they are solutions for the same monomer. Since both monomers have identical torsion angles, this assumption does not affect the structural characteristics, but only the symmetry-related orientation of the structure to the bilayer normal. The $\Delta\nu_{obs}$ value for the N-C^1 bond can be either positive or negative, but only $\cos\theta = -0.3970$ is consistent with the peptide plane bond angles, the N-H dipolar result and the ^{15}N chemical shift. Consequently, a unique set of vector orientations is achieved for this peptide plane. For the Ala^3-Leu^4 peptide plane, the $\Delta\nu_{obs}$ value for the N-H dipolar interaction is positive and both signs for $\cos\theta$ are possible. For the N-C^1 dipolar interaction both signs of $\Delta\nu_{obs}$ are possible, but only $\cos\theta = \pm 0.7405$ is consistent with the various peptide plane constraints, as for the Gly^2-Ala^3 plane. The solution sets derive from just two possible combinations of bond vector orientations for the two peptide planes, g1 and g2 (Table 1). For each of these bond vector combinations eight possible torsion angle solutions can be found that are consistent with the C^α carbon tetrahedral geometry (Table 2). These result from two ambiguities that arise when the planes are combined for the torsion angle solutions. Firstly, with a unique plane orientation the sign of the normal to this plane is not defined, hence a sign ambiguity exists for each plane. Secondly, the orientation in the XY plane is not defined and so for each pair of planes there is another (ϕ,ψ) combination that can be defined within the constraints of the C^α carbon tetrahedral geometry. Therefore, for each combination g1 or g2 there are eight possible torsion angle pairs, resulting in a total of sixteen possible solutions for each residue (Fig. 2).

Additional data are needed to further reduce the solution set and these new data need to have the primary interaction vector out of the peptide linkage plane. ^2H data, such as those shown in Fig. 3, that include the C^α-^2H-labeled leucine sites in gramicidin, provide the necessary data. For the Ala^3 site, spectra have previously been published showing a C^α-^2H quadrupolar splitting of 195.4 kHz (Lee et al., 1993). These data eliminate 12 of the 16 possible (ϕ,ψ) torsion angle solutions. Four solutions remain, g23, g24, g27 and g28, with a unique set of direction cosines for the diplane. Furthermore, the variation in the (ϕ,ψ) torsion angles is very limited among these four solutions. These conformers predict C^α-^2H orientations that

TABLE 2
POSSIBLE Ala^3 TORSION ANGLE SOLUTIONS

Bond vector combination	Torsion angle solution	ϕ (°)	ψ (°)
g1	g11	91	73
	g12	91	41
	g13	−91	−73
	g14	−91	−41
	g15	82	73
	g16	82	41
	g17	−82	−73
	g18	−82	−41
g2	g21	130	−153
	g22	130	−121
	g23	−130	153
	g24	−130	121
	g25	121	−153
	g26	121	−121
	g27	−121	153
	g28	−121	121

*The programs CNFCS and COORDS, developed in our lab, are available from the authors upon request.

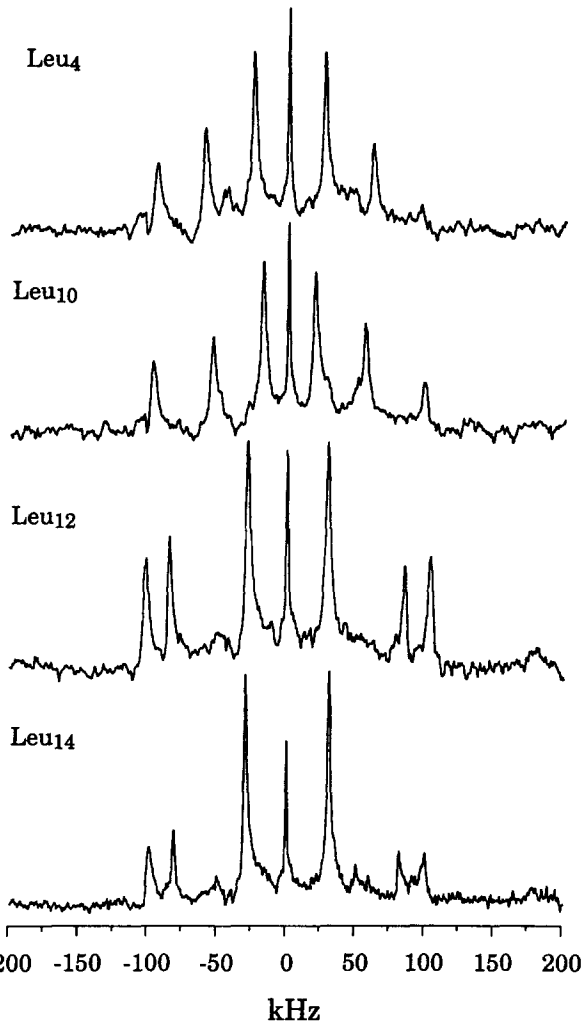

Fig. 3. ^{2}H NMR spectra of oriented d_3-(α,β_1,β_2)-Leu^{4,10,12,14}-labeled gramicidin A at 36 °C in hydrated DMPC. Spectra have been obtained at 61.55 MHz and are the result of approximately 15 000 acquisitions. The samples were oriented with the bilayer normal parallel to the magnetic field.

are within 4° of those calculated from the observed data. Furthermore, all of these solutions represent β-strand-type structures with the carbonyl groups alternating between parallel and antiparallel to the channel axis and, unlike the invalid solutions, they have the C^α-C^β axis approximately perpendicular to the helical axis and bilayer normal. This orientation minimizes the steric hindrance of the side chains with the backbone.

Because the channel is a relatively uniform structure, there are the same number of solutions for each C^α carbon. But, not all of the possible (ϕ,ψ) solutions for a single C^α carbon pair with all of the possible (ϕ,ψ) solutions for the next C^α carbon. Each diplane associated with residue i shares peptide planes with the diplanes of residues i−1 and i+1. Only the diplanes that share peptide planes that are identically aligned with the channel axis can be combined (Fig. 4). For this reason, 2^{n+1} structures are formed, and not 4^n, where n is the number of C^α carbons. The possible backbone conformations represent simple permutations of the sign of the individual peptide plane normals with respect to the channel axis. A unique set of bond orientations has been established for each plane and a unique fold for the polypeptide is established, since a change in the bilayer normal orientation does not affect the helical parameters. The sign permutations result in four base backbone conformations using the g23, g24, g27 and g28 (ϕ,ψ) solutions: (i) alternating in and out of the carbonyl groups; (ii) all out; (iii) all in; and (iv) alternating out and in (Fig. 4). The first structure is built using g23 for odd-numbered or L- residues and g28 for even-numbered or D-residues and is termed g2328. The second structure is built using g24 for all residues and is termed g2424. The third structure is built using g27 for all residues and is termed g2727. The fourth structure is built using g28 for odd-numbered or L- residues and g23 for even-numbered or D-residues and is termed g2823. These four backbone structures have the same structural motif. They are all right-handed helices with approximately 6.6 residues per turn and an average helical pitch of 4.8 Å. Their structural repeat is a dipeptide with β-strand torsion angles and the side chains are directed radially. The hydrogen bonding pattern for these four structures is identical. Since they were built from the same data, they all match the observed N-H and N-C^1 bond orientations and show a reasonable fit to the observed ^{15}N chemical shifts and C^α-^{2}H quadrupolar splittings. Consequently, the difference between these structures is subtle.

The existence of four possible initial backbone conformations does not pose a problem for the determination of a single final structure. A refinement procedure (Ketchem, 1995) has recently been developed that has the ability to search the local conformational space of the peptide plane orientations covering the possible peptide plane orientation permutations with respect to the channel axis. Refinement of the four base structures consistently leads to a single backbone conformation that exhibits a unique pattern of peptide plane orientations (Ketchem et al., in preparation).

Side-chain structures are assembled in much the same way by assuming a fixed covalent geometry for the structural units connected by bonds about which there is torsional flexibility. For the leucine side chains these bonds are the C^α-C^β and C^β-C^γ axes that correspond to the χ_1 and χ_2 axes, respectively, and hence two torsion angles define the structural solution for these side chains. The C-Me bonds exhibit rapid three-state jumps and therefore, methyl-based spin interactions are averaged to the C-Me bond vector. Consequently, no structural information is available on the χ_3 torsion angles of leucine or the χ_2 torsion angles for valine. The analysis of the structural data is aided by the prior knowledge of the backbone structure; in other words, the orientations of both the C^α-^{2}H and C^α-C^β bond vectors are uniquely defined by the

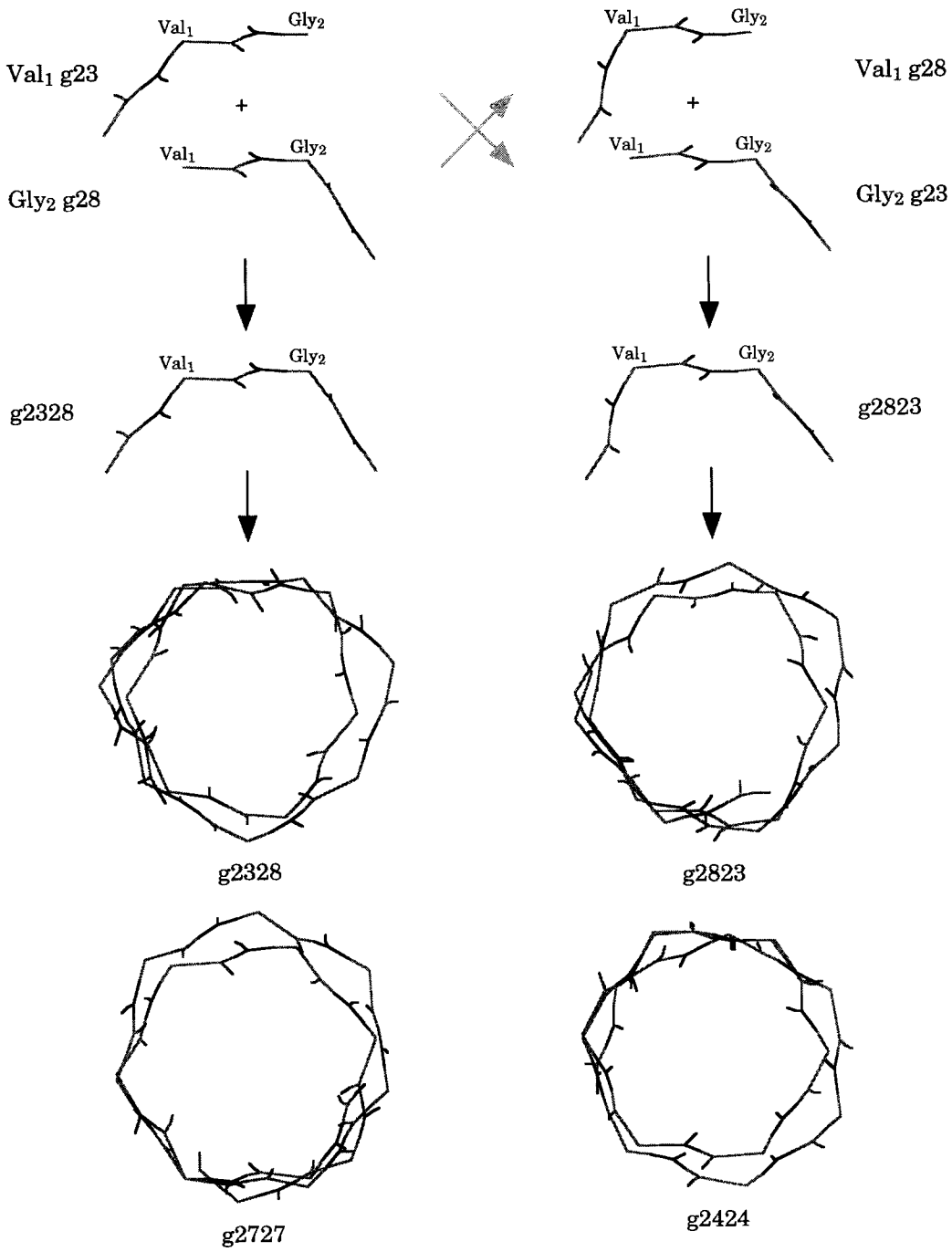

Fig. 4. Overlapping shared peptide planes. Although four (ϕ,ψ) solutions exist about each C^{α} carbon, not all diplane combinations are possible; only those shared planes that have identical orientations can be overlapped to build a structure, as shown by the solid arrows. The shaded arrows indicate invalid plane overlap. The resulting assemblies form the initial backbone structures of which four base initial structures are shown.

backbone structure. The existence of four base backbone structures does not generate ambiguity in these bond orientations. Uncertainty in the data analysis increases as dynamic flexibility increases near the ends of the side chains and therefore, the structure will be elucidated by first constraining χ_1 and then χ_2.

To characterize the structure and dynamics of leucine side chains, two different isotopic labeling patterns have been used for ^2H NMR spectroscopy: a d_3-sample in which the α- and β-protons have been deuterated, and a d_7-sample in which the γ- and δ-protons have been deuterated. The d_3-sample has predictable characteristics for the α-deuteron, while the two β-deuterons will be subject to the additional motional flexibility about the χ_1 torsion axis. The powder-pattern spectra shown in Fig. 5A reflect the superposition of signals from the α- and β-sites, and yet there is no indication of two dynamic states in these spectra. To avoid the dynamic averaging due to global

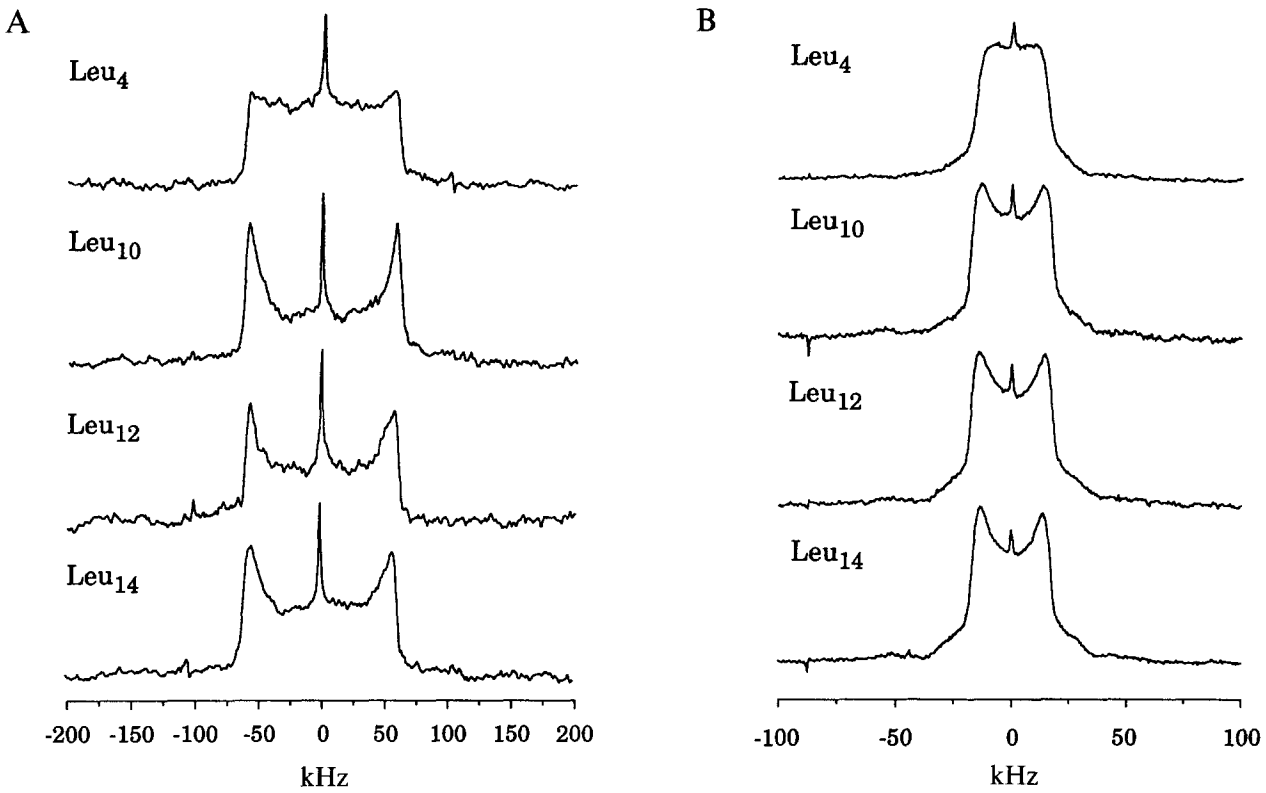

Fig. 5. ^{2}H NMR powder-pattern spectra of labeled gramicidin A in hydrated DMPC bilayers recorded at 5 °C (below the gel-to-liquid-crystalline phase-transition temperature) with about 60 000 acquisitions. (A) Leucine d_3-(α,β)-labeled gramicidin A. (B) Leucine d_7-(γ,δ)-labeled gramicidin A.

rotational motions (Lee et al., 1993), these spectra have been obtained in fully hydrated lipid bilayers at 5 °C, well below the gel-to-liquid-crystalline phase-transition temperature. Since the C^{α}-^{2}H and C^{β}-^{2}H powder patterns are not distinguishable, additional large amplitude dynamics are not implicated about the χ_1 axis and average quadrupole coupling constants (QCC) can be estimated from simulations of the spectra for each of the leucine sites.

Similar powder-pattern spectra have been obtained for d_7-labeled leucine in gramicidin, as shown in Fig. 5B. Here the motionally averaged methyl deuterons dominate the spectra, because six of the seven deuterons are methyls and because the frequency space is restricted for the methyls to 1/3 of that for the single γ-deuteron. These spectra have slightly distorted Pake powder patterns, especially Leu^4, suggesting motions in addition to the three-state jump of the methyl group between equally populated states. However, if this results from multiple χ_2 rotameric states the population of the minor states must be very low, because the tensor averaging is only slight. For the analysis presented here each of the leucines are confined to a single rotameric state based on the powder-pattern results.

To relate the various Electric Field Gradient (EFG) tensors of the deuteron in the laboratory frame obtained from data such as those shown in Fig. 3, to the molecular frame of the leucine side chain, the tensors need to be transformed to their Principal Axis System (PAS) frame. The Euler transformations (Rose, 1957) have been previously described for the valine side chains, for which χ_1 is the only definable side-chain torsion angle. Here this approach has been extended for the longer leucine side chains, so that both χ_1 and χ_2 can be determined (Fig. 6): β_2 is the angle of the C^{α}-C^{β} bond with respect to Z_{lab} and γ_2 is the angle between ^{2}H-C^{α}-C^{β} and Z_{lab}-C^{α}-C^{β} planes; α_3 is the angle between ^{2}H-C^{α}-C^{β} and C^{α}-C^{β}-C^{γ} planes and differs from χ_1 by a constant equal to the angle between the N-C^{α}-C^{β} and ^{2}H-C^{α}-C^{β} planes; α_4 is the angle between C^{α}-C^{β}-C^{γ} and C^{β}-C^{γ}-C^{δ_1} planes and is the χ_2 angle. Furthermore, perfect tetrahedral geometry was assumed for all side-chain carbon sites.

Because the absolute orientation of C^{α}-^{2}H is known, the assignment of one of the quadrupolar splittings (approximately 200 kHz in each leucine spectrum) in Fig. 3 is straightforward. The other two splittings from C^{β_1}-^{2}H and C^{β_2}-^{2}H cannot be assigned without additional information. The ^{2}H NMR spectra for oriented d_7-(γ,δ_3^1,δ_3^2)-leucine-labeled gramicidin is shown in Fig. 7, as well as the room temperature powder-pattern spectra to aid in resonance assignment. As described previously, C^{γ}-^{2}H has low spectral intensity in the powder-pattern spectra, and consequently the C^{γ}-^{2}H assignment can be readily made in the spectra of oriented samples with the exception of Leu^4. Furthermore, it is clear that some powder-pattern intensity is observed in the spectra of the oriented samples; it arises from a sample that has seeped from between

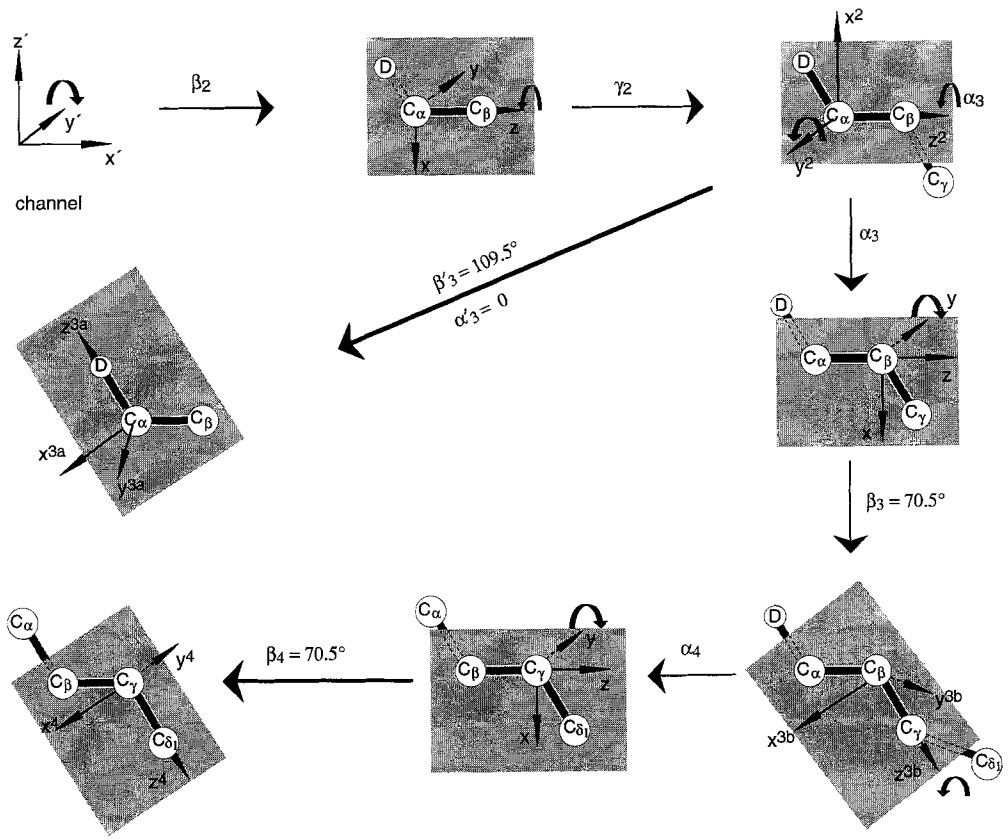

Fig. 6. Euler transformation angles from the channel axis to the Principal Axis System (PAS). The rotations follow the convention of Rose (1957). The rotations, $\Omega_2 = (0,\beta_2,\gamma_2)$, bring the z-axis along the C^{α}-C^{β} bond and the xz plane into the D-C^{α}-C^{β} plane. $\Omega_3 = (0,109.5,0)$ rotates the axes to the PAS of C^{α}–D, while the rotations, $\Omega_3 = (\alpha_3,70.5,0)$, take the xz plane into the C^{α}-C^{β}-C^{γ} plane and the z-axis along the C^{β}-C^{γ} bond. The last set of rotations, $\Omega_4 = (\alpha_4,70.5,0)$, first brings the xz plane to the C^{β}-C^{γ}-C^{δ} plane and then aligns the z-axis with the C^{γ}-C^{δ} bond. No further rotation is needed, since the D^{δ} PAS is averaged along the C^{γ}-C^{δ} bond by fast methyl rotation. The unknown angles are β_2, γ_2, α_3, and α_4. The angles, α_3 and α_4 are related to χ_1 and χ_2, respectively. The shaded areas in the figure represent the xz plane.

the glass plates. Such intensity can be ignored in the structural analysis of the quadrupole splittings. The observed quadrupole splitting from oriented samples occurs at the averaged value of $\nu_{\|}$ in the powder-pattern spectra, as opposed to the static value of $\nu_{\|}$. The best-fit values of averaged $\nu_{\|}$ are presented in Table 3.

In addition to finding solutions for χ_1 and χ_2, the β_2 angle is allowed to have some flexibility about the value range determined from the backbone analysis, which varied from 81° to 86° for the four leucines. This recognizes the potential error in the backbone characterization. Furthermore, the β_2 and γ_2 variables are dependent variables, as shown previously in the analysis of the valine sites, and therefore a total of three variables, χ_1, χ_2 and β_2 need to be solved for. In the conformational space defined by a β_2 range from 70° to 92° and by the full range of χ_1, the quadrupole splittings were calculated and compared by the rmsd to the observed data for C^{α}-^{2}H, C^{β_1}-^{2}H and C^{β_2}-^{2}H. While the C^{α}-^{2}H assignment has been made, the C^{β}-^{2}H assignments were not known and neither was the sign of the quadrupole splittings. These ambiguities were allowed to vary in searching for the rmsd minimum. For each of the leucine sites there are six minima of less than or equal to 1 kHz (only four minima for Leu[4]), such

TABLE 3
THE EFFECTIVE QUADRUPOLE COUPLING CONSTANT AND QUADRUPOLE SPLITTING ($\Delta\nu_q$) FOR THE CALCULATION OF THE LEUCINE CONFORMATIONS

	D^{α}		D^{β_1}		D^{β_2}		D^{γ}		$D_3^{\delta_1}$		$D_3^{\delta_2}$	
	QCC_{eff}	$\Delta\nu_q$	QCC_{eff}	$\Delta\nu_q$	QCC_{eff}	$\Delta\nu_q$	QCC_{eff}	$\Delta\nu_q$	QCC_{eff}	$\Delta\nu_q$	QCC_{eff}	$\Delta\nu_q$
Leu[4]	160	191	153	121.0	153	51.2	138	8.3[a]	130	29.0	130	−8.3
Leu[10]	160	196	153	37.5	156	−110.5	156	−49.1	128	−11.5	128	−31.4
Leu[12]	160	207	153	170.6	153	48.5	115	−76.3	108	43.2	108	6.4
Leu[14]	160	199	153	163.2	153	60.4	115	−79.1	101	38.4	101	5.5

[a] The quadrupole splitting overlapped with the $D_3^{\delta_2}$ splitting.

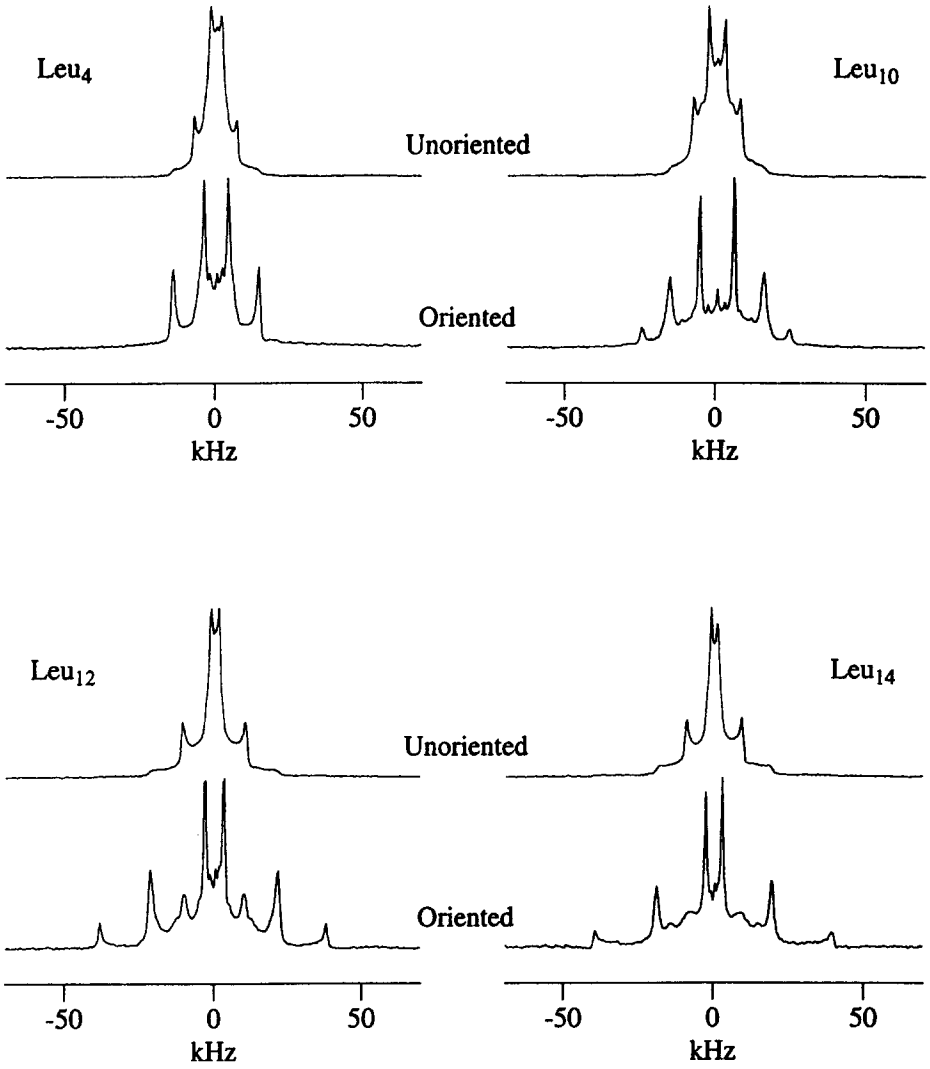

Fig. 7. ^{2}H NMR spectra of d_7-(γ,δ_3^1,δ_3^2)-Leu^{4,10,12,14} gramicidin A from oriented (bilayer normal and channel axis parallel to B_0) and unoriented sample preparations at 36 °C (above the gel-to-liquid-crystalline phase-transition temperature). The powder and oriented spectra were obtained with approximately 20000 and 10000 acquisitions, respectively, and are plotted together to facilitate the assignments. Note that some of the minor peaks from the oriented samples are the result of a small fraction of unoriented sample.

as those shown in Fig. 8 for Leu^{12}. Of these six minima only one is consistent with a χ_1 rotameric state and low rmsd values in the χ_1/χ_2 conformational map which is calculated with both the d_3- and d_7-sample data. This latter map is calculated based on a specific β_2 angle. When $\beta_2 = 87°$ is used, which is also consistent with two of the minima in the β_2/χ_1 map, the rmsd values were substantially greater. The β_2 angle of 84° is only consistent with the eclipsed χ_1 conformational state, and therefore a unique value ($\beta_2 = 79°$ for Leu^{12}) is achieved and it is similar to the value achieved from the backbone data (82°). The unique value of β_2 leads to a unique χ_1 angle of −173° (the angle of 20° is just 20° from an eclipsed conformation, therefore this option is not considered to be viable). The corresponding χ_2 angle of −61° is considered more probable than 146°, which is 34° from the 180° rotameric state.

The Leu^4 analysis was missing the C$^\gamma$-^{2}H assignment because this splitting was unresolved from the other splittings. There are eight χ_2 possible solutions, as shown in Fig. 9, but six are readily eliminated since the predicted C$^\gamma$-^{2}H splitting is inconsistent with any observed spectral intensity. In other words, the C$^\gamma$-^{2}H splitting is assumed to overlap with the observed intensity in the d_3-sample spectrum of Fig. 7. This yields two possible χ_2 values, 160° and −87°. The value of 160° is somewhat closer to a rotameric state and is considered the most probable. Table 3 shows the motionally averaged QCC values used in these rmsd calculations and the resultant sign for the observed quadrupole splittings. Table 4 presents the torsional solutions from the rmsd analysis for each of the leucine side chains, as well as the previously published torsion angles for the valines (Lee et al., 1995) and tryptophans (Hu and Cross, 1995).

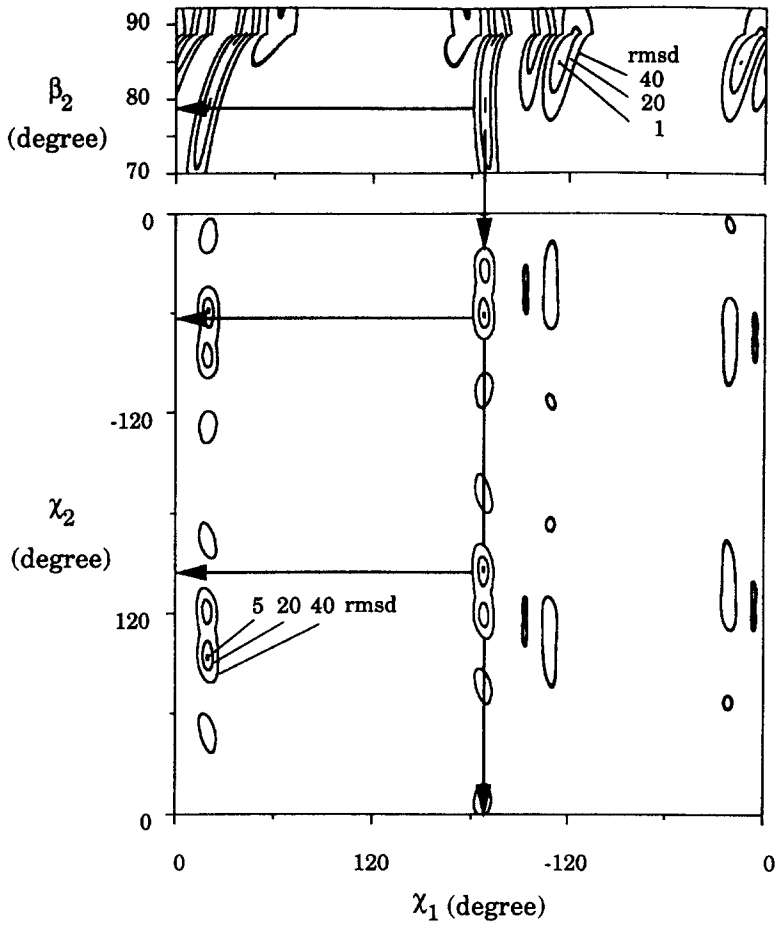

Fig. 8. The rms deviation between observed data and predicted values as a function of β_2, χ_1 and χ_2 for the Leu[12] gramicidin side chain. In the top figure, β_2 and χ_1 are varied and only the data from d_3-leucine is considered. In the bottom figure the data from d_3- and d_7-leucine are analyzed as a function of χ_1 and χ_2, using the fixed value of β_2 defined above.

Discussion

For the backbone structure it has been shown that the orientational ambiguities inherent in these structural constraints do not lead to multiple secondary structures as the structural elements are assembled. This is because the constraints for each structural unit are independent and uniquely defined. Unlike distance constraints in both solution and solid-state NMR, which constrain one part of the macromolecule to another part, the orientational constraints constrain the macromolecule to the laboratory frame of reference. In other words, distance constraints are relative constraints, while orientational constraints are absolute constraints. Through reasonable covalent-geometry assumptions; that the peptide linkage is planar and that the stereochemistry for each C^{α} carbon site is known, a single structural motif results in which all of the hydrogen bonds are identified, and the residues per turn and the helix sense determined.

This amply illustrates the high-resolution nature of these structural constraints. Despite concerns such as the tensor-element magnitudes, the tensor orientation for the chemical shift interaction, and the degree of motional averaging of these tensors, the error in these constraints are small and do not appear to accumulate. In this initial backbone structure where i to i+6 hydrogen bonding is

TABLE 4
THE SIDE-CHAIN CONFORMATIONS DETERMINED BY ^2H AND ^{15}N EXPERIMENTAL DATA

	β_2 (°)		γ_2 (°)		χ_1 (°)	χ_2 (°)
	^{15}N	^2H	^{15}N	^2H	^2H	^2H
Val[1]	101	106	190	198	177	–
Leu[4]	86	77	16	22	−154	160[a]
						−87
Val[6]	88	88	18	16	55	–
Val[7]	93	103	189	199	−145	–
Val[8]	88	90	15	17	55	–
Trp[9]	100	–	189	–	−72	−97
Leu[10]	84	81	16	18	−58	−12[a]
						121
Trp[11]	102	–	190	–	−70	−81
Leu[12]	82	79	16	16	−173	−61
Trp[13]	101	–	191	–	−63	−90
Leu[14]	81	82	15	17	−173	−63
Trp[15]	98	–	191	–	−58	−96

[a]The probable solution.

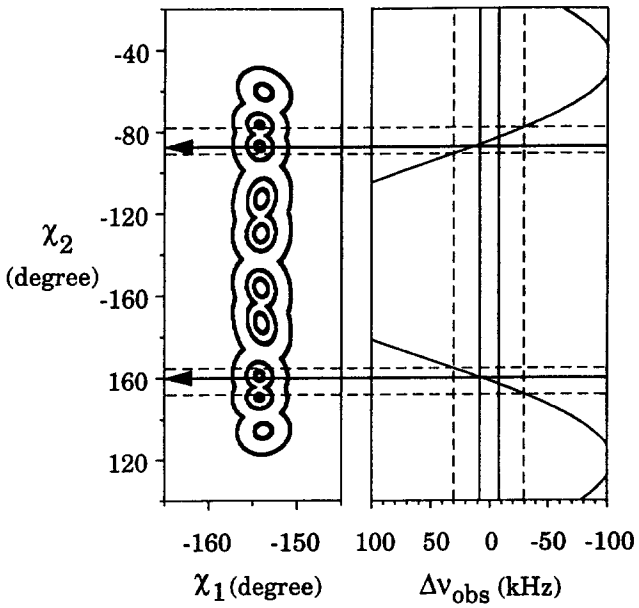

Fig. 9. An expanded region of the χ_1/χ_2 rmsd contour plot for the Leu[4] gramicidin side-chain conformational space. Plotted along the same χ_2 scale are the calculated quadrupole splittings for the γ-deuteron which was not resolved from the other resonances.

shown, despite errors associated with each peptide plane orientation the rmsd from ideal β-strand hydrogen-bond distances is approximately 0.5 Å. If each peptide plane was only defined relative to the adjacent plane, structural distortions would accumulate, but instead each plane is defined relative to the laboratory frame of reference, and consequently errors are just as likely to cancel as to add. The result is a three-dimensional structure that defines a unique molecular fold.

While this result demonstrates the quality of the orientational constraints, the dipolar, chemical shift and quadrupolar data have not been used uniformly. The dipolar interactions have been used to define the peptide plane orientations, while the chemical shifts and C^{α}-^2H quadrupolar interactions have been used as structural filters to accept or reject specific structural possibilities. A refinement protocol has recently been devised which takes an initial structure and refines it against all of the experimental constraints, as well as the CHARMm force field (Ketchem, 1995; Ketchem et al., 1996). Not only does this refinement protocol fully utilize the experimental constraints, but all of the covalent-geometry assumptions can be relaxed, including the imposed planarity of the peptide linkage.

Experimental refinement is being undertaken as well as computational refinement. The tensor characterizations used in the determination of this initial structure were obtained using dry powders (Mai et al., 1993). Methods for the experimental observation of tensors in fast-frozen hydrated lipid bilayers have been developed in this laboratory (Lazo et al., 1993,1995) and have the potential to more accurately define the structure. These tensors are observed over a temperature range from 120 K to room temperature for characterizing the local dynamics of the structural unit. This information could be used to motionally average the static nuclear spin tensors for all of the interactions in the structural unit, and in doing so a better determination of the molecular orientation would be achieved.

Dynamics are shown to increase significantly with distance from the polypeptide backbone, and yet it is shown that 'most probable', if not unique torsion angle solutions can be achieved in these aliphatic side chains with just ^2H labels. One of the most challenging tasks in using the data from oriented samples is the separation of dynamic and structural influences on the observed spectra. The use of powder patterns as a function of temperature to resolve different dynamic modes is often the key to accurate spectral interpretation. More detailed dynamic characterizations are possible in the side chains as well as the backbone with fast-frozen samples that yield a model for the molecular motions that can be used for the interpretation of relaxation data (North and Cross, 1993,1995).

The primary conclusions from the backbone analysis are further substantiated by the side-chain analysis. Conformational ambiguities do not accumulate as the structural units of the side chains are assembled and neither do the orientational errors accumulate. The assembly of a complete initial structure is shown in Fig. 10, with the exception of the ethanolamine blocking group at the C-terminus. The four representative backbone structures are superimposed by minimizing the all-atom rmsd without changing the channel axis orientation. Since all backbone structures have identical side-chain orientations, the side chains are shown only for the g2328 backbone structure. The structure of the four tryptophan side chains (Hu et al., 1993,1995) and the four valines (Lee et al., 1995) have been solved. The glycine and two alanine side chains are defined by the backbone structure. The structure in Fig. 10 has been determined entirely from solid-state NMR derived orientational constraints and the amino acid sequence. No assumptions about the secondary structure were used and the constraints were not fit to a preconceived model. Energetic constraints have only been used to avoid eclipsed side-chain conformational states and to assume planar peptide linkages. This initial structure defines the molecular fold and many of the important structural features, such as the helix sense, the residues per turn, and the orientation of the indole rings at the bilayer surface. While all of this is clear from the initial structure, refinement is needed to resolve van der Waals contacts between side chains, which is not severe in most cases but is significant in some, to improve hydrogen-bonding geometry and to relax the ideal covalent geometry used throughout the initial structure. It is the refinement of these four structures against all of the structural

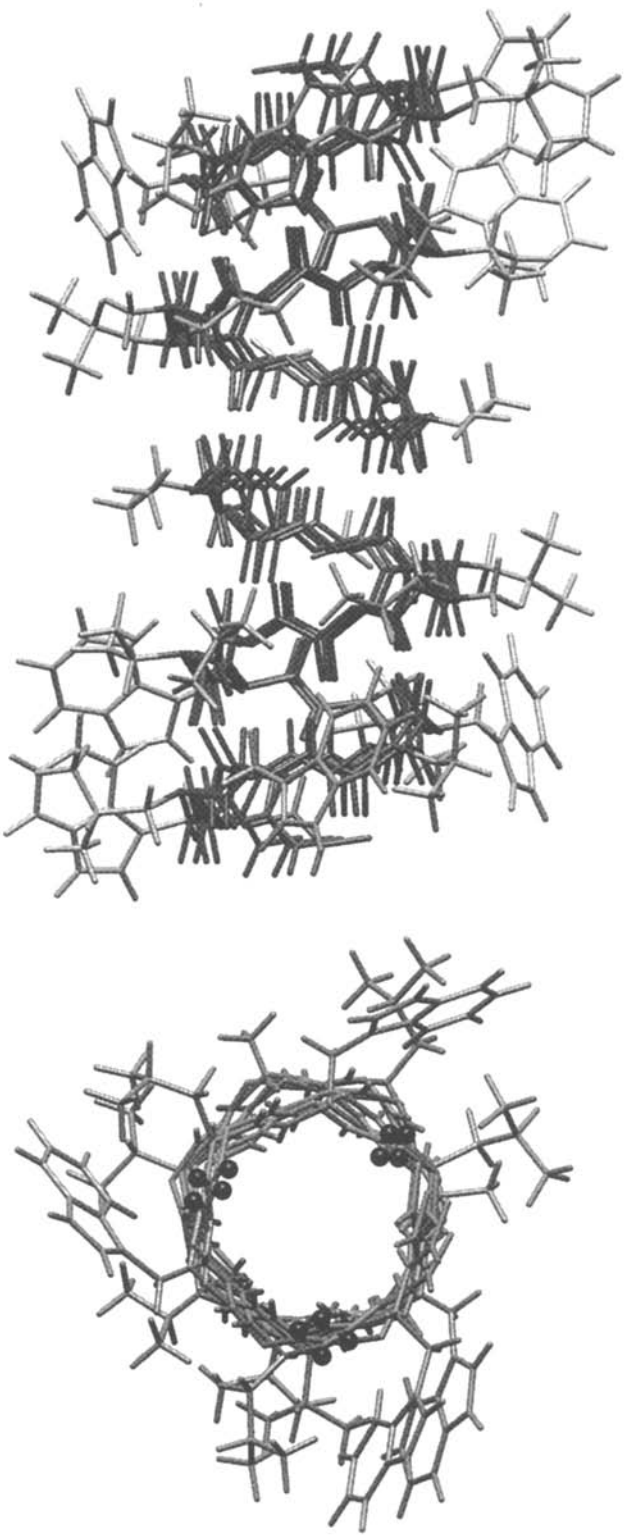

Fig. 10. A complete initial structure determined solely from solid-state NMR-derived orientational constraints and the amino acid sequence. The data were not fit to a model nor were any assumptions made concerning the secondary structure. Shown here are the four superimposed base initial structures (g2328, g2424, g2727 and g2823) which span the complete conformational space arising from the one remaining ambiguity per amino acid: the sign of the direction cosine for the peptide plane normal. This ambiguity gives rise to 2^{n+1} backbone structures, all of which define the same secondary structure, the same helix sense, the same hydrogen-bonding pattern, nearly the same residues per turn, and nearly the same helical pitch. The carbonyl oxygens for Leu^{10}, Leu^{12} and Leu^{14} are shown as dark spheres for the four base structures in the end view at the bottom. For each site two of the structures have the carbonyls oriented in towards the channel and two away from the channel axis. The side view at the top shows the symmetric dimer that is essential for spanning the lipid bilayer. The side-chain orientations have been independently solved and are placed directly on the backbone of just the g2328 structure.

constraints as well as the CHARMm force field that has led to a unique torsion angle solution set (Ketchem et al., in preparation).

Solid-state NMR has been demonstrated here as a new approach for determining complete three-dimensional structures. Not only is a folding motif uniquely defined, but the structure is oriented relative to the anisotropic environment of the lipid bilayer. This is one of the few protein or polypeptide structures that has been solved in a lipid environment, much less a lipid bilayer environment. To make this approach more generally applicable, spectra with a larger number of constraints and a mechanism for assigning multiple resonances will be needed. For now it is shown that these constraints can be used to determine a unique molecular fold of both the polypeptide backbone and the side chains. Furthermore, these constraints hold the potential for very high resolution structural details through refinement against all of the experimental constraints.

Acknowledgements

Many thanks to Scott Prosser, Regitze Vold and Antoinette Killian for bringing to our attention the availability of the thin glass slips used in the sample preparation. The authors are indebted to the staff of the FSU NMR facility: J. Vaughn, R. Rosanske and T. Gedris for their skillful maintenance and service of the NMR spectrometers and to H. Henricks and U. Goli in the Bioanalytical Synthesis and Services Facility for their expertise and maintenance of the ABI 430A peptide synthesizer and HPLC equipment. This work has been supported by NSF grant (DMB 9317111) and the National High Magnetic Field Laboratory.

References

Braun, W., Wider, G., Lee, K.H. and Wüthrich, K. (1983) *J. Mol. Biol.*, **169**, 921–948.

Brenneman, M.T. and Cross, T.A. (1990) *J. Chem. Phys.*, **92**, 1483–1494.

Brenneman, M., Quine, J. and Cross, T.A., unpublished results.

Cross, T.A. and Opella, S.J. (1983) *J. Am. Chem. Soc.*, **105**, 306–308.

Cross, T.A., Ketchem, R.R., Hu, W., Lee, K.-C., Lazo, N.D. and North, C.L. (1992) *Bull. Magn. Reson.*, **14**, 96–101.

Cross, T.A. (1994) *Annu. Rep. NMR Spectr.*, **29**, 123–167.

Cross, T.A. and Opella, S.J. (1994) *Curr. Opin. Struct. Biol.*, **4**, 574–581.

Evans, J.N.S., Appleyard, R.J. and Shuttleworth, W.A. (1993) *J. Am. Chem. Soc.*, **115**, 1588.

Fields, G.B., Fields, C.G., Petefish, J., Van Wart, H.E. and Cross, T.A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1384–1388.

Fields, C.G., Fields, G.B., Noble, R.L. and Cross, T.A. (1989) *Int. J. Pept. Protein Res.*, **33**, 298–303.

Griffiths, J.M. and Griffin, R.G. (1993) *Anal. Chim. Acta*, **283**, 1081–1101.

Gullion, T. and Schaefer, J. (1989) *J. Magn. Reson.*, **81**, 196–200.

Harbison, P.K., Jelinski, L.W., Stark, R.E., Torchia, D.A., Herzfeld, J. and Griffin, R.G. (1984) *J. Magn. Reson.*, **60**, 79–82.

Hartzell, C.J., Whitfield, M., Oas, T.G. and Drobny, G.P. (1987) *J. Am. Chem. Soc.*, **109**, 5966–5969.

Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H. (1990) *J. Mol. Biol.*, **213**, 899–929.

Hu, W., Lee, K.C. and Cross, T.A. (1993) *Biochemistry*, **32**, 7035–7047.

Hu, W. and Cross, T.A. (1995) *Biochemistry*, **34**, 14147–14155.

Hu, W., Lazo, N.D. and Cross, T.A. (1995) *Biochemistry*, **34**, 14138–14146.

Jap, B.K., Walian, P.J. and Gehring, K. (1991) *Nature*, **350**, 167–170.

Ketchem, R.R., Hu, W. and Cross, T.A. (1993) *Science*, **261**, 1457–1460.

Ketchem, R.R. (1995) *Structural Determination and Refinement of the Gramicidin A Transmembrane Channel as Studied by Solid-State Nuclear Magnetic Resonance Spectroscopy*, Ph.D. Thesis, The Florida State University, Tallahassee, FL.

Ketchem, R.R., Roux, B. and Cross, T.A. (1996) In *Membrane Structure and Dynamics* (Eds, Merz, K.M. and Roux, B.), Birkhauser, Boston, MA, pp. 299–322.

Kuhlbrandt, W., Wang, D. and Fujiyoshi, Y. (1994) *Nature*, **367**, 614–621.

Lazo, N.D., Hu, W. and Cross, T.A. (1992) *J. Chem. Soc. Ser. Chem. Commun.*, 1529–1531.

Lazo, N.D., Hu, W., Lee, K.C. and Cross, T.A. (1993) *Biochem. Biophys. Res. Commun.*, **197**, 904–909.

Lazo, N.D., Hu, W. and Cross, T.A. (1995) *J. Magn. Reson.*, **B107**, 43–50.

Lee, K.C., Hu, W. and Cross, T.A. (1993) *Biophys. J.*, **65**, 1162–1167.

Lee, K.C., Huo, S. and Cross, T.A. (1995) *Biochemistry*, **34**, 857–867.

Mai, W., Hu, W., Wang, C. and Cross, T.A. (1993) *Protein Sci.*, **2**, 532–542.

McDonnell, P.A., Shon, K., Kim, Y. and Opella, S.J. (1993) *J. Mol. Biol.*, **233**, 447–463.

Michel, H. and Deisenhofer, J. (1990) In *Current Topics in Membranes and Transport*, Vol. 36, Academic Press, New York, NY, pp. 53–69.

Morelli, M.A.C., Pastore, A. and Motta, A. (1992) *J. Biomol. NMR*, **2**, 335.

North, C.L. and Cross, T.A. (1993) *J. Magn. Reson.*, **B101**, 35–43.

North, C.L. and Cross, T.A. (1995) *Biochemistry*, **34**, 5883–5895.

Rose, M.E. (1957) *Elementary Theory of Angular Momentum*, Wiley, New York, NY.

Schiffer, M., Chang, C.-H. and Stevens, F.J. (1992) *Protein Eng.*, **5**, 213–214.

Shon, K., Kim, Y., Colnago, L.A. and Opella, S.J. (1991) *Science*, **252**, 1303–1305.

Smith, R. and Cornell, B.A. (1986) *Biophys. J.*, **49**, 117–118.

Stark, R.E., Jelinski, L.W., Ruben, D.J., Torchia, D.A. and Griffin, R.G. (1983) *J. Magn. Reson.*, **55**, 266–273.

Teng, Q., Nicholson, L.K. and Cross, T.A. (1991) *J. Mol. Biol.*, **218**, 607–619.

Teng, Q., Iqbal, M. and Cross, T.A. (1992) *J. Am. Chem. Soc.*, **114**, 5312–5321.

Valentine, K.G., Rockwell, A.L., Gierasch, L.M. and Opella, S.J. (1987) *J. Magn. Reson.*, **73**, 519–523.

Wang, D.N. and Kuhlbrandt, W. (1991) *J. Mol. Biol.*, **217**, 691–699.

Weiss, M.S., Abele, U., Weckesser, J., Welte, W., Schiltz, E. and Schulz, G.E. (1991) *Science*, **254**, 1627–1630.