Use of a Single Glycine Residue to Determine the Tilt and Orientation of a Transmembrane Helix. A New Structural Label for Infrared Spectroscopy

Jaume Torres, Andreas Kukol, and Isaiah T. Arkin
Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, United Kingdom

ABSTRACT Site-directed dichroism is an emerging technique for the determination of membrane protein structure. However, due to a number of factors, among which is the high natural abundance of ^{13}C, the use of this technique has been restricted to the study of small peptides. We have overcome these problems through the use of a double C-deuterated glycine as a label. The modification of a single residue (Gly) in the transmembrane segment of M2, a protein from the *Influenza* A virus that forms H^+-selective ion channels, has allowed us to determine its helix tilt and rotational orientation. Double C-deuteration shifts the antisymmetric and symmetric stretching vibrations of the CD_2 group in glycine to a transparent region of the infrared spectrum where the dichroic ratio of these bands can be measured. The two dichroisms, along with the helix amide I dichroic ratio, have been used to determine the helix tilt and rotational orientation of M2. The results are entirely consistent with previous site-directed dichroism and solid-state NMR experiments, validating C-deuterated glycine ($GlyCD_2$) as a structural probe that can now be used in the study of polytopic membrane proteins.

INTRODUCTION

Membrane proteins are by far the most biomedically important family of proteins, serving as a target for the vast majority of pharmaceutical agents. Regrettably, the structural determination of membrane proteins is an extremely difficult task using standard methods such as NMR or x-ray diffraction. An alternative technique, based on the application of structural restraints from site-directed infrared dichroism (Arkin et al., 1997) to molecular dynamics protocols, is an emerging method in structural biology to study membrane proteins in their native environment, a lipid bilayer. This method has recently been applied in the determination of structural models of various transmembrane helical homooligomers (Kukol and Arkin, 1999, 2000; Kukol et al., 1999). This technique relies on the measurement of the dichroic ratio of the band due to the $^{13}C{=}O$ carbonyl vibration of a specifically labeled residue.

The use of this technique, however, was until now restricted to such small proteins (25–30 residues per helix) due to the fact that ^{13}C is a relatively abundant isotope (1.11%) which is therefore found also at random positions, thus diluting the label. Additionally, the $^{13}C{=}O$ vibrational absorption band, although shifted relative to the main $^{12}C{=}O$ amide I, is still partially overlapped with the latter, which again limits the study to relatively small α-helical proteins. Finally, at least two samples, with residues labeled at different positions, are needed to determine the helix tilt and rotational orientation of an α-helix (Arkin et al., 1997).

All of these limitations can be overcome with the use of an isotopomer of glycine, $GlyCD_2$, in which both α-protons have been substituted for deuterium. This substitution shifts the symmetric, *ss*, and antisymmetric, *as*, stretching modes of the CH_2 group to shorter wavenumbers (Suzuki et al., 1966), which isolates these bands from the CH_2 and CH_3 stretching spectral region of the lipids and also from bands arising from the protein. Additionally, the natural abundance of deuterium is 0.016%, i.e., 0.000256% for two deuterium atoms, which implies that dilution of the label is vanishingly small, even for proteins that contain thousands of residues, as only glycine can dilute the signal.

Furthermore, a different advantage comes from the fact the symmetric and antisymmetric vibrational modes of the CD_2 group in $GlyCD_2$ are mutually perpendicular, allowing the determination of two dichroic ratios, *ss* and *as*, in a single residue. This is important because according to the theory of site-directed dichroism at least three dichroic ratios have to be known to determine the main structural parameters of the helix (see below). In the case of the label $^{13}C{=}O$, the dichroism of the helix and at least two labels are needed, i.e., two samples, each of them labeled at a different residue. In the case of $GlyCD_2$, *ss* and *as* dichroic ratios can be determined using a single residue.

We have confirmed this experimentally by introducing the label $GlyCD_2$ in the sequence of a synthetic peptide encompassing the transmembrane region of the 97-residue M2 protein from *Influenza* A virus, Ser-22–Leu-46. The result obtained is of utmost importance in structural biology, as it allows the use of site-directed infrared dichroism to the structural determination of polytopic membrane proteins, which probably constitute the majority of drug targets in the body.

MATERIALS AND METHODS

Peptide synthesis and reconstitution

The peptide was made by standard solid-phase F-moc synthesis and reconstituted in DMPC lipids. The peptide was cleaved from the resin and

Received for publication 22 March 2000 and in final form 30 June 2000.

Address reprint requests to Isaiah T. Arkin, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK. Tel.: +44-(0)1223-766-048; Fax: +44-(0)1223-766-002; E-mail: sa232@cam.ac.uk.

0006-3495/00/12/3139/05 $2.00

lyophilized. The peptide contained a C-deuterated glycine ($GlyCD_2$, Cambridge Isotopes Laboratories, Andover, MA) at position 34 and was purified and reconstituted in DMPC liposomes as described previously (Kukol et al., 1999).

FTIR measurements

FTIR spectra were recorded on a Nicolet Magna 560 spectrometer (Madison, WI) purged with N_2 and equipped with an MCT/A detector cooled with liquid nitrogen. Attenuated total reflection (ATR) spectra were measured with a 25-reflection ATR accessory from Graseby Specac (Kent, UK) and a wire grid polarizer (0.25 μM, Graseby Specac). Approximately 200 μl of sample ($\sim$2.5 mg/ml protein and 12.5 mg/ml lipid) were dried onto a trapezoidal ($50 \times 2 \times 20$ mm) internal reflection element (KRS-5 or Ge). A total of 1000 interferograms collected at a resolution of 2 cm^{-1} were averaged for every sample and processed with 1 point zero filling and Happ-Genzel apodization. The dichroism of the amide I bands was calculated integrating between 1670 and 1645 cm^{-1} either the original or the Fourier self-deconvoluted (FSD) (Kauppinen et al., 1982) spectra. The enhancement factor used in FSD was 2.0, well below log (S/N). Integration from either spectra gave indistinguishable results. The area corresponding to the CD_2 vibration bands was measured from the original spectra using a straight baseline that contains points immediately before and after the band. The dichroic ratio was calculated as the ratio between the integrated absorptions of parallel and perpendicular polarized light.

Data analysis

The data were analyzed according to the theory of site-specific dichroism presented in detail elsewhere (Arkin et al., 1997), based on the fact that the measured dichroic ratios, $\mathcal{R}$, of a particular transition dipole moment is a function of the sample fractional order, f, and the spatial orientation of the dipole. This is defined by the parameters shown in Fig. 1: β, the helix tilt, α, which relates the transition dipole moment to the helix director, and ω, the rotational pitch angle.

From each measurement, three different dichroic ratios were obtained. The first is $\mathcal{R}_{Helix}$, the dichroic ratio that corresponds to the $^{12}C{=}O$ transition dipole moments distributed around the helical axis (i.e., $\omega_{n+1} = \omega_n + 100°$ for a standard α-helix). Therefore, this dichroic ratio is dependent on β and f, but independent of ω (Arkin et al., 1997):

$$\mathcal{R}_{Helix}(\beta, f) = \frac{e_z^2\left(f\mathcal{H}_z + \frac{1-f}{3}\right) + e_x^2\left(f\mathcal{H}_x + \frac{1-f}{3}\right)}{e_y^2\left(f\mathcal{H}_y + \frac{1-f}{3}\right)} \quad (1)$$

where $\mathcal{H}_x$, $\mathcal{H}_y$, and $\mathcal{H}_z$ are the x, y, and z components of the rotationally averaged integrated absorption coefficients. The parameter f represents the fractional order, i.e., $f = 1$ if the sample is completely ordered and $f = 0$ if the sample is completely random. Finally, e_x, e_y, and e_z are the electric field components for each axis given by Harrick (1967). The other dichroic ratios observed, due to *ss* and *as* stretching ($\mathcal{R}_{ss}$ and $\mathcal{R}_{as}$) correspond to the CD_2 label. Consequently, they will be dependent on the ω angle of the residue that contains the label:

$$\mathcal{R}_{ss}(\beta, f, \omega) = \frac{e_z^2\left(f\mathcal{H}_{ss\,z}(\omega) + \frac{1-f}{3}\right) + e_x^2\left(f\mathcal{H}_{ss\,x}(\omega) + \frac{1-f}{3}\right)}{e_y^2\left(f\mathcal{H}_{ss\,y}(\omega) + \frac{1-f}{3}\right)} \quad (2)$$

$$\mathcal{R}_{as}(\beta, f, \omega) = \frac{e_z^2\left(f\mathcal{H}_{as\,z}(\omega) + \frac{1-f}{3}\right) + e_x^2\left(f\mathcal{H}_{as\,x}(\omega) + \frac{1-f}{3}\right)}{e_y^2\left(f\mathcal{H}_{as\,y}(\omega) + \frac{1-f}{3}\right)} \quad (3)$$

These three equations are sufficient to obtain the three unknowns β, ω, and f.

The nonlinear equations were solved with Newton's method as implemented in the FindRoot function in Mathematica 3.0 (Wolfram Research, Champaign, IL). The angle α for the peptidic C=O bond is known to be 39° from fiber diffraction studies (Tsuboi, 1962).

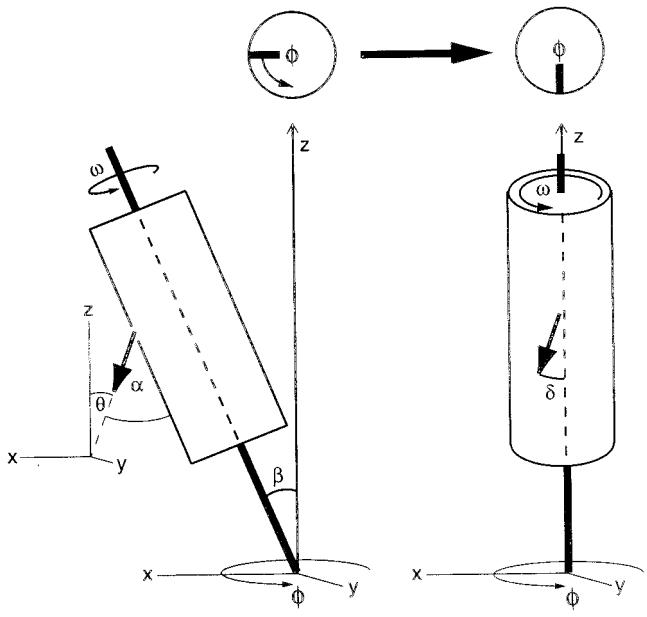

FIGURE 1 Schematic representation of the geometric parameters that define a transition dipole moment orientation (in this case the amide I C=O stretching mode is shown) in a transmembrane helix. The helix tilt β and the rotational orientation ω are derived from the experimental data.

RESULTS AND DISCUSSION

The amide I band for both polarizations is shown (see Fig. 2). The spectra are typical for a predominantly α-helical peptide (Byler and Susi, 1986), having an amide I band with maximum at 1657 cm^{-1}. Neither the original nor the deconvoluted (not shown) spectra show significant intensity around 1640–1630 cm^{-1}, indicating the absence of β structure (Byler and Susi, 1986). The orientation of the *ss* and *as* transition dipole moments (tdp) of the CH_2 (or CD_2) vibrations in glycine are as shown in Fig. 3. The angles between the antisymmetric tdp (as-tdp) or symmetric tdp (ss-tdp) and the helix axis were calculated using a model of glycine where this residue is located in an α-helical environment in which the helix axis coincides with the z axis, using the function CREATE in CHI (CNS, Crystallography and NMR System (CNS Version 0.3) (Brunger et al., 1998)).

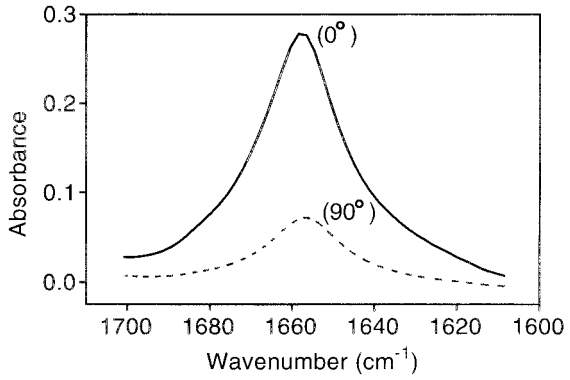

FIGURE 2 Original infrared spectra collected with parallel (0°) or perpendicular polarization (90°) of the amide I band for M2 labeled with $GlyCD_2$ and incorporated in DMPC liposomes. The dichroism was calculated as described in Materials and Methods.

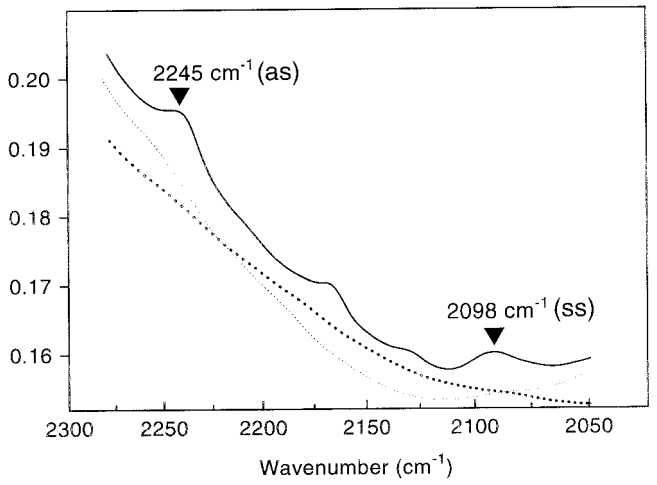

FIGURE 4 Original infrared spectra of native M2 (- - -) and labeled M2 at Gly-34 (—) or $ValC^{13}$ (· · ·) in the absence of lipids after drying from a chloroform/methanol solution (see text). The bands used in the analysis corresponding to symmetric (*ss*) or antisymmetric (*as*) stretching of CD_2 are indicated.

It was assumed that the as-tdp is aligned along with a vector that joins the two C-hydrogens of glycine, and that the ss-tdp is perpendicular to as-tdp and contains the Cα and a point in the *z* axis. The angles between as-tdp and ss-tdp relative to the helix axis were found to be 38° and 84°, respectively. The angle δ, which is defined by the angle between the tdp and the helix axis when the *z* axis, the helix axis, and the residue reside in the same plane, with the residue located in the direction of the tilt, was found to be $-51°$ for as-tdp and 0° for ss-tdp.

The bands that arise from the CH_2 stretching modes were isolated from the CH_2 and CH_3 stretching region of DMPC by substituting the C-protons with deuterons. The precise location of these bands in the infrared spectrum was determined using a sample in which lipids were absent after dissolving the peptide in chloroform/methanol and drying.

Four bands were observed (Fig. 4) in a region consistent with the red-shift expected ($\sim$700 cm^{-1}) if the CH_2 group is treated as a harmonic oscillator. As expected, nonlabeled M2 or M2 labeled with $^{13}C{=}O$ valine do not display any of these bands in these conditions (see figure), nor does any other protein tested (not shown). Four bands have also been observed for crystalline polyglycine II (Dwivedi and Krimm, 1982), where the two high-energy bands were assigned to *as* vibration and the two lower-energy bands to *ss* vibration. Only two of these bands, located at 2242 cm^{-1} and 2098 cm^{-1}, indicated in Fig. 4, were used to obtain the dichroic ratios (see Fig. 5). The other two bands were found to overlap with observed small periodic bands, $\Delta\nu \sim 30$ cm^{-1}, that appear when lipids are present (not shown). For example, in Fig. 5 (*right panel*), the band expected at 2130 cm^{-1} according to Fig. 4 is hidden by a band centered at 2150 cm^{-1}, a band also present in the absence of protein. The origin of these bands is completely unclear to us, but appears to be related to the presence of a bilayer or its interaction with water, because they are not present in dry

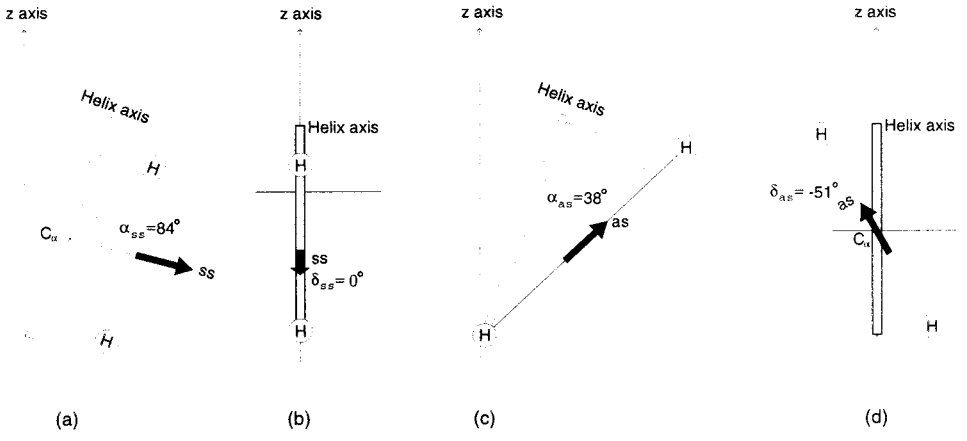

FIGURE 3 Schematic diagrams showing the orientation of symmetric and antisymmetric transition dipole moments for the CH_2 stretching vibration in glycine. (*a*) Orientation of the symmetric transition dipole moment relative to the helix axis. (*b*) Orientation of the symmetric transition dipole moment relative to the helix axis when the *z* axis, the helix axis, and the α carbon of glycine are all in the same plane and the residue is in a proximal position (this defines $\omega = 0$). (*c*) Same as in (*a*) for as-tdp. (*d*) Same as in (*b*) for as-tdp. The δ and α angle are represented.

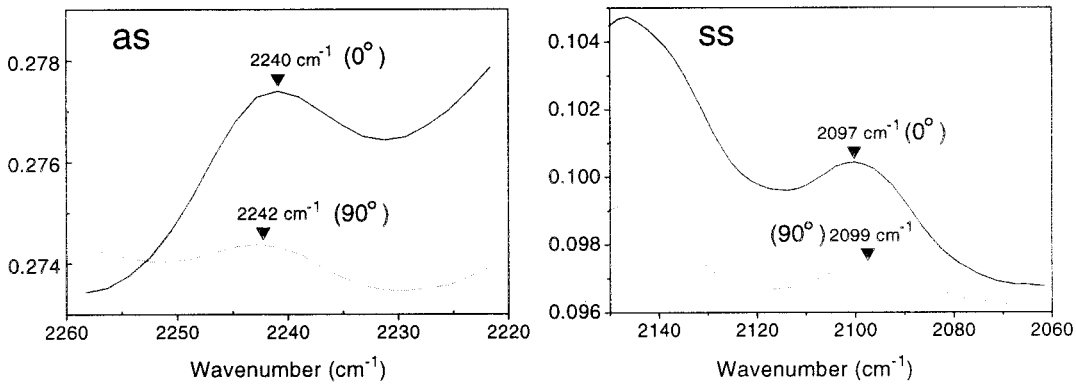

FIGURE 5 Original infrared spectra collected with parallel (0°) or perpendicular polarization (90°) of symmetric (*ss*) or antisymmetric (*as*) stretching bands for M2 labeled with $GlyCD_2$ and incorporated in DMPC liposomes. The position of the maxima are indicated.

lipid spectra or lipid dissolved in an organic solvent. The frequency of the bands used (see *arrows* in Fig. 4) is almost identical to the bands previously assigned to the CD_2 vibration in C-deuterated β-form polyglycine I (2240 and 2118 cm^{-1}) and α-helical polyglycine II (2244 and 2117 cm^{-1}) (Suzuki et al., 1966). Note that for the samples used by these authors, N-deuteration induced a red-shift only of the low-frequency band, to 2108 cm^{-1}. The samples without the label did not show any band in the region shown in Fig. 5.

The results of the five separate measurements are given in Table 1. The average dichroic ratios of the *ss* band was 2.15 ± 1.1 and for the *as* band 5.2 ± 1.2. Analysis of the dichroic ratios according to the theory for site-specific dichroism as detailed in the Methods section yields a rotational pitch angle ω of 146° ± 11° and a helix tilt β of 35° ± 4°, whereas f ranged from 0.6 to 0.8, depending on the sample. These results are almost identical to the values obtained from the model reported previously (Kukol et al., 1999) (ω_{Gly34} = 139° ± 9.9° and β = 31.6° ± 6.2°) in which the labels used were ^{13}C-Ala-29 and ^{13}C-Ala-30. This clearly indicates that the results obtained from both labels ($^{13}C{=}O$ and $GlyCD_2$) are entirely consistent. Furthermore, both the rotational orientation and the β angle are almost identical to those determined from solid-state NMR data (Kovacs and Cross, 1997). The complex relation between the dichroisms (*ss* and *as*) and the rotational orientation of the peptide is illustrated visually in Fig. 6. Obviously, as mentioned in Materials and Methods, both dichroisms (*ss* and *as*) depend on three parameters, ω, β, and f. We have chosen a representation in which β is constant (31°) (Kukol et al., 1999) and the calculated dichroisms are plotted in a plane as a function of f and (ω_{Gly34}). Each one of the lines (contours) represents coordinates (f,ω_{Gly34}) with the same calculated dichroism. In this figure, a broken line in both left and right graphs represents the expected ω_{Gly34} according to Kukol et al. (1999). The observed dichroisms for Gly-34, therefore, should fall in an intersection between the broken line and a contour line.

TABLE 1 Helix dichroic ratios $\mathcal{R}_{Helix}$ and site-specific dichroic ratios of the symmetric $\mathcal{R}_{CD_2ss}$ and asymmetric stretching $\mathcal{R}_{CD_2as}$ for different samples of lipid vesicle reconstituted M2 transmembrane domain, labeled with $GlyCD_2$ at residue 34.

$\mathcal{R}_{Helix}$	$\mathcal{R}_{CD_2ss}$	$\mathcal{R}_{CD_2as}$
3.1	5.7	3.5
3.4	4.0	3.0
2.3	4.5	2.4
2.5	4.3	1.7
2.9	6.7	1.7

Variation of the data reflects the different sample order achieved by drying the lipid vesicles on the ATR internal reflection element.

CONCLUSIONS

The advantage of $GlyCD_2$ over $^{13}C{=}O$ is clear, as $GlyCD_2$ does not have natural abundance (the percentage of natural ^{13}C is 1%). Additionally, the bands that originate from $GlyCD_2$ are located in a region where no other protein bands contribute. Thirdly, and more importantly, the fact that $GlyCD_2$ contributes with two perpendicular stretching modes (only the $C{=}O$ stretching mode is used in $^{13}C{=}O$ labels), allows the determination of the rotational orientation of a transmembrane segment using a single label, which is not dependent on the assumption that $\omega_{n+1} = \omega_n + 100°$, as in labels with a single measurable tdp, i.e., $^{13}C{=}O$. Obviously, because of slight bends and coils of the helix, this assumption is less true as the distance between two labels increases.

The double amount of information to be obtained from this label will be particularly valuable in the study of polytopic membrane proteins, a group that perhaps constitutes the major target for drugs. For these proteins the contribution of individual helices to the amide I, and therefore their respective dichroic ratios, cannot be determined. This increases from two to three the number of labels required to

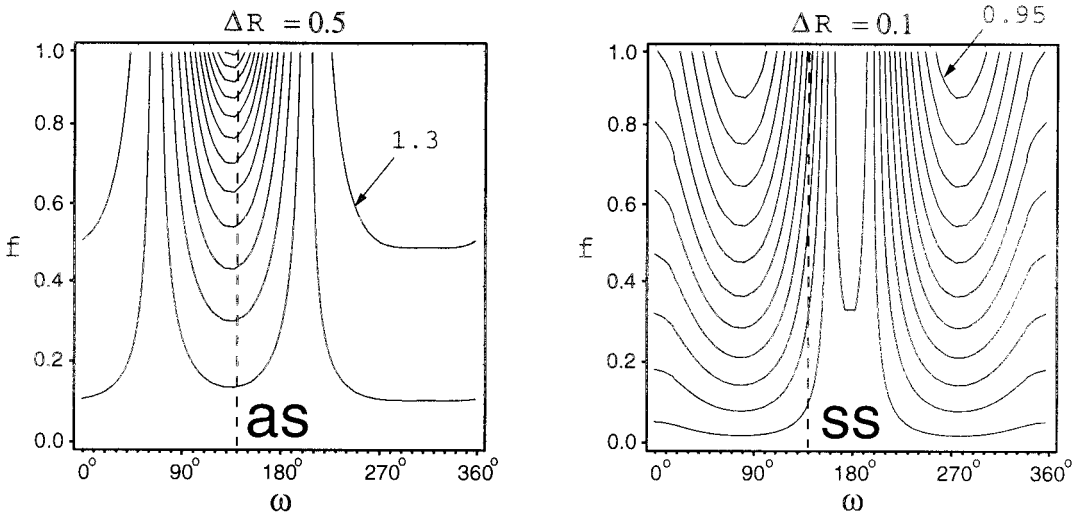

FIGURE 6 Contour diagrams of the distribution of the dichroic ratios $\mathcal{R}$ as a function of ω and f when β is 31° (Kukol et al., 1999) for the antisymmetric (*left*) and symmetric (*right*) stretching. The broken vertical line is located at $\omega = 139°$, the value predicted previously for Gly-34 in M2 (Kukol et al., 1999). The dichroism increases toward the center from 0.8 to 8.3, at increments of 0.5 units for the antisymmetric stretching (*left*) and from 0.85 to 2.15, at increments of 0.1 units for the symmetric stretching (*right*). The dichroism corresponding to the first contour plotted is also shown.

determine helix tilt and rotational orientation for a label producing a single tdp, i.e., $^{13}C{=}O$ stretching. A label like $GlyCD_2$, which produces two tdp, allows keeping at two the number of samples required per helix, with obvious benefits in terms of time and financial resources. We note also that, although glycine is not one of the most commonly found residues in transmembrane helices, it is particularly abundant in small and rigid loops that connect transmembrane segments. Important structural restraints can be obtained in these regions that will help to determine the topology of the protein, although different structural assumptions will be required, as these regions are not α-helical.

This work was supported by a grant from the BBSRC and the Wellcome Trust.

REFERENCES

Arkin, I., K. MacKenzie, and A. Brunger. 1997. Site-directed dichroism as a method for obtaining rotational and orientational constraints for oriented polymers. *J. Am. Chem. Soc.* 119:8973–8190.

Brunger, A., P. Adams, G. Clore, W. Gros, R. Grosse-Kunstleve, J. Jiang, J. Kuszewski, M. Nilges, N. Pannu, R. Read, L. Rice, T. Simonson, and G. Warren. 1998. Crystallography and NMR system: a new software system for macromolecular structure determination. *Acta Crystallogr. D.* 54:905–921.

Byler, D., and H. Susi. 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers.* 25:469–487.

Dwivedi, A., and S. Krimm. 1982. Vibrational analysis of peptides, polypeptides and proteins. XV. Crystalline polyglycine II. *Biopolymers.* 21:2377–2397.

Harrick, N. 1967. Internal Reflection Spectroscopy, 1st Ed. Interscience publishers, New York.

Kauppinen, J., D. Moffatt, H. Mantsch, and D. Cameron. 1982. Fourier self-deconvolution: a method for resolving intrinsically overlapped bands. *Appl Spectrosc.* 35:271–276.

Kovacs, F., and T. Cross. 1997. Transmembrane four-helix bundle of influenza A M2 protein channel: structural implications from helix tilt and orientation. *Biophys. J.* 73:2511–2517.

Kukol, A., P. Adams, L. Rice, A. Brunger, and I. Arkin. 1999. Experimentally based orientational refinement of membrane proteins: a structure for the influenza A M2 H+ channel. *J. Mol. Biol.* 286:951–962.

Kukol, A., and I. Arkin. 1999a. Vpu transmembrane peptide structure obtained by site-specific Fourier transform infrared dichroism and global molecular dynamics searching. *Biophys. J.* 77:1594–1601.

Kukol, A., and I. Arkin. 2000. Structure of the Influenza C virus CM2 protein transmembrane domain obtained by site-specific infrared dichroism and global molecular dynamics searching. *J. Biol. Chem.* 275: 4225–4229.

Suzuki, S., Y. Iwashita, and T. Shimanouchi. 1966. Infrared spectra of isotopic polyglycines. *Biopolymers.* 4:337–350.

Tsuboi, M. 1962. Infrared dichroism and molecular conformation of α-form poly-g-benzyl-L-glutamate. *J. Polym. Sci.* 59:139–153.