Assembly of a Polytopic Membrane Protein Structure from the Solution Structures of Overlapping Peptide Fragments of Bacteriorhodopsin

Madan Katragadda, James L. Alderfer, and Philip L. Yeagle
Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269, and Department of Biophysics, Roswell Park Cancer Institute, Buffalo, New York 14263 USA

ABSTRACT Three-dimensional structures of only a handful of membrane proteins have been solved, in contrast to the thousands of structures of water-soluble proteins. Difficulties in crystallization have inhibited the determination of the three-dimensional structure of membrane proteins by x-ray crystallography and have spotlighted the critical need for alternative approaches to membrane protein structure. A new approach to the three-dimensional structure of membrane proteins has been developed and tested on the integral membrane protein, bacteriorhodopsin, the crystal structure of which had previously been determined. An overlapping series of 13 peptides, spanning the entire sequence of bacteriorhodopsin, was synthesized, and the structures of these peptides were determined by NMR in dimethylsulfoxide solution. These structures were assembled into a three-dimensional construct by superimposing the overlapping sequences at the ends of each peptide. Onto this construct were written all the distance and angle constraints obtained from the individual solution structures along with a limited number of experimental inter-helical distance constraints, and the construct was subjected to simulated annealing. A three-dimensional structure, determined exclusively by the experimental constraints, emerged that was similar to the crystal structure of this protein. This result suggests an alternative approach to the acquisition of structural information for membrane proteins consisting of helical bundles.

INTRODUCTION

Determination of membrane protein structure has proven a more difficult subject than the determination of soluble protein structure due to the hydrophobic nature of membrane proteins and their resulting insolubility in aqueous medium. Crystals suitable for diffraction studies are difficult to obtain for these proteins. Much effort has been invested in growing both two-dimensional and three-dimensional crystals to enable diffraction studies. However, because progress is still slow on the determination of membrane protein structures, alternative approaches to structure would be helpful.

A growing body of data suggests that solution structures of peptides derived from some classes of proteins retain the secondary structure of the parent protein because of the dominance of short-range interactions that can be captured in peptides. Studies on segments of proteins forming α-helices show that peptides containing these sequences form α-helices in almost every case (Gao et al., 1999; Ramirez-Alvarado et al., 1997; Gegg et al., 1997; Hamada et al., 1995; Callihan and Logan, 1999; Wilce et al., 1999; Jimenez et al., 1999; Fan et al., 1998; Cox et al., 1993; Hunt et al., 1997). Peptides representing segments that are turns in the native protein also show turns as peptides in solution (Chandrasekhar et al., 1991; Ghiara et al., 1994; Blumenstein et al., 1992; Blanco and Serrano, 1994; Goudreau et al., 1994; Adler et al., 1995; Campbell et al., 1995; Wilce et al., 1999; Cox et al., 1993; Katragadda et al., 2000). Both α-helices and turns are dominated by short-range interactions (Yang et al., 1996a). In some cases, the sequence of an entire protein, built around helical bundles, has been incorporated in a series of peptides spanning that sequence, and the individual peptides have reported the secondary structure of the native proteins with fidelity (Blanco and Serrano, 1994; Behrends et al., 1997; Reymond et al., 1997; Dyson et al., 1992; Padmanabhan et al., 1999).

It might therefore be hypothesized that peptides containing the amino acid sequences for turns or for transmembrane helices of membrane proteins built of helical bundles will exhibit the same secondary structures in solution that those sequences adopt in the native protein. This hypothesis could provide access to important structural information for a transmembrane protein that may be available from no other approach.

Bacteriorhodopsin from *Halobacteria* offers an excellent opportunity to examine this hypothesis. The structure of bacteriorhodopsin consists of a bundle of seven transmembrane helices connected by turns. Engelman and co-workers have found that the transmembrane helices of this protein are independently stable folding units (Hunt et al., 1997) and thus can be considered protein domains (Popot and Engelman, 2000). Several x-ray crystal structures are available for this membrane protein (Gouaux, 1998; Grigorieff et al., 1996; Luecke et al., 1999; Pebay-Peyroula et al., 1997). We synthesized a series of overlapping peptides spanning the sequence of bacteriorhodopsin and determined the solution structures of these peptides using NMR in dimethylsulfoxide (DMSO) solution. The solution structures of these peptides compared favorably with the corresponding re-

Received for publication 12 January 2001 and in final form 30 April 2001.
Address reprint requests to Dr. Philip L. Yeagle, University of Connecticut, Department of Molecular and Cell Biology, &-125, Storrs, CT 06269-4331; Tel.: 860-486-4363; Fax: 860-486-4331; E-mail: yeagle@uconnvm.uconn.edu.

0006-3495/01/08/1029/08 $2.00

gions of the x-ray crystal structure. These individual peptide structures, and the constraints obtained for them, were then used to build the three-dimensional structure for the whole protein by a new method. The resulting structure was similar to structures determined from electron and x-ray diffraction data.

MATERIALS AND METHODS

Peptide synthesis

Peptides were synthesized through solid-phase synthesis in the Biotechnology Center of the University of Connecticut. Synthesis was performed on an Applied Biosystems 433A peptide synthesizer, using fastMoc chemistry on HMP resin at room temperature. HBTU was used for activation and the column was washed with NMP. The N- and C-termini were not capped.

NMR spectroscopy

All NMR spectra were recorded on a Bruker AMX-600 spectrometer at 30°C in DMSO. Standard pulse sequences and phase cycling were employed to record double quantum filtered (DQF) COSY, TOCSY, and NOESY (Kumar et al., 1980) (data were collected with 400-ms mixing times). Previous work with other similar-sized peptides at mixing times of 150, 250, and 400 ms showed no evidence of spin diffusion and 400 ms showed the most useful interactions in the NOESY. All spectra were accumulated in a phase-sensitive manner using time-proportional phase incrementation for quadrature detection in F1. All the data for the structure determinations were obtained in DMSO at 25°C or 30°C. Chemical shifts were referenced to the residual protons in the d_6-DMSO. DMSO was chosen as a solvent only because it was one solvent system in which both the loop peptides (which are insoluble in water) and the transmembrane helix peptides were soluble. Some helices and turns were not soluble in detergent micelles. (Another solvent system consisting of $CDCl_3/CD_3OD/D_2O$ has been used successfully for a pair of hydrophobic helices connected by a turn (Dmitriev et al., 1999).) The torsion angle constraints (ϕ) were obtained from the coupling constants measured as described (Ludvigsen et al., 1991) using the program PRONTO (written by M. Kjaer).

Structure refinement

The sequence-specific assignment of the ^{1}H NMR spectrum for each peptide was carried out using standard methods employing FELIX (MSI, San Diego, CA). Assigned NOE cross-peaks were segmented using a statistical segmentation function and characterized as strong, medium, and weak corresponding to upper bounds distance range constraints of 2.7, 3.5, and 5.0 Å, respectively. Lower bounds between nonbonded atoms were set to the sum of their van der Waals radii (~1.8 Å). Pseudo-atom corrections were added to inter-proton distance restraints where necessary (Wüthrich et al., 1983). Structures were obtained using simulated annealing with these experimental distance constraints from the NOESY data and the torsion angle constraints obtained from the COSY data. The Kollman All Atom force field and Kollman charges were used within Sybyl (Tripos, St. Louis, MO). The molecule was heated to 800 K for 1000 fs followed by cooling to 200 K during 1500 fs. Five or ten consecutive cycles were calculated. These calculations were performed on a Silicon Graphics R10000 computer.

Assembly of the peptide fragments

A construct for the whole protein can be assembled from the pieces whose individual structures are determined as described above. Because of the design of the set of peptides, the overlap in the sequences of these peptides can be superimposed. The superposition was done with recognition of the disordering that is typically observed at the ends of the peptide. Taking advantage of the 10-residue overlap, the most disordered residues were not used in the superposition algorithm. The unused residues were discarded from the file. The construct then consists of the structures of all the peptides with the overlapping sequences. A model of helix G (because an NMR structure could not be obtained for this peptide) was temporarily added to the construct to be used as an anchor point for some of the inter-helical distances to adequately pack the transmembrane helical bundle of bacteriorhodopsin. In the final structure, helix G was removed because its structure could not be determined experimentally. Retinal was not included in this structure, because it is bound to helix G. Redundant sequences are then removed from the Protein Data Bank file. To test the limits of this approach, another construct was assembled, by first superimposing the experimental peptide helical structures on the corresponding helices of one of the structures (1brd, 7 Å resolution), removing the crystal structure, and then adding the loop peptide structures by simultaneously superimposing both ends of the loop structure onto the ends of the appropriate transmembrane helices. The ultimate result using this construct, after the simulated annealing (see below), was similar to the result using the first construct, with a somewhat improved match to the crystal structures.

Simulated annealing

On the construct, all the distance range constraints and dihedral angle constraints obtained from the solution structures of each of the peptides were written in a mol2 file with SYBYL (Tripos). To this were added inter-helical distance constraints derived from intermediate-resolution electron diffraction studies (Henderson et al., 1990) to define the packing of the helical bundle. These constraints consist of distances (as range constraints) between the tops of pairs of helices in the transmembrane bundle as well as between the middle and the bottom of pairs of helices of the bundle (see Table 2). Hydrogen bond constraints were added where hydrogen bonds were observed in the original solution structures of the peptides. A simulated annealing protocol was used to fold a structure that was consistent with the available experimental distance constraints. The construct for bacteriorhodopsin was heated to 800°C for 1000 fs and then cooled for 1500 fs to 200°C. The Kollman All Atom force field and Kollman charges were used within Sybyl. The process of simulated annealing, with all the NOE-derived distance constraints on all parts of the molecule, at one point led to a distortion of one of the helical segments (helix C) so that its structure differed from the original structure observed in the solution NMR of the individual peptide. As a refinement process, we selected a few appropriate inter-atomic distances on that distorted segment, measured those inter-atomic distances on the original NMR solution structure, and added those as distance constraints to the construct of the whole molecule. After another simulated annealing, the structure of that particular helix once again was close to the structure observed in the original solution structure of the peptide. These calculations were performed on a Silicon Graphics R10000 computer.

RESULTS AND DISCUSSION

Three-dimensional structures of the set of overlapping peptides spanning the sequence of rhodopsin

A set of peptides was designed for the complete primary sequence of bacteriorhodopsin. Fig. 1 shows the sequences of these peptides. Each peptide was designed to encompass either a transmembrane helix of the protein or a turn. Peptides containing a turn also included some of the helices

FIGURE 1 Sequences of overlapping peptides from bacteriorhodopsin used in this study. These peptides were synthesized by solid-phase peptide synthesis as described previously (Katragadda et al., 2000). The tubes show the approximate extent of the helical regions in the crystal structure, which are labeled A through G. Overlaps in sequences of the peptides are indicated by overlaps of the representations of the peptides in the figure.

(one to two turns) that connect to the turn on both sides. Peptides containing a turn that was shorter (15–17 residues) than the ones used here usually showed in the solution structure only the turn with no information on how that turn connected to the transmembrane helices. Therefore we moved to longer peptides from which one might expect (and usually got) a helix-turn-helix motif. Furthermore, each peptide was designed to overlap each of its neighbors in the series by 10 amino acids. The overlap was necessary because structures of peptides in solution typically show ill-defined termini. To obtain structural information on the entire sequence thus necessitated a design in which the disordered ends of each of the peptides could be ignored. After trying shorter overlaps, we found that a 10-amino-acid overlap is probably minimal. Each of the overlap regions turned out to be helical and to have enough structure to perform a superposition to connect the two structures required a minimum of four to five residues that were well defined and in a helical conformation. Because the last two to four residues of peptides are frequently disordered in solution, the overlap region needs to be 8–13 amino acids long. In most cases, an overlap of 10 residues worked well.

The structures of these peptides could most easily be obtained in solution by NMR. To do so, a common solvent was required. Most of the peptides in this set from bacteriorhodopsin were hydrophobic and not soluble in water. Although some peptides (coding for transmembrane helices) were soluble in detergent micelles (Hunt et al., 1997), peptides coding for loops were not soluble in detergent micelles in a manner that preserved structure as measured by circular dichroism. All peptides in the series were soluble in DMSO except for the peptide corresponding to helix G, which was not soluble in DMSO or in chloroform/methanol solutions. Previous work had shown that helix G was not

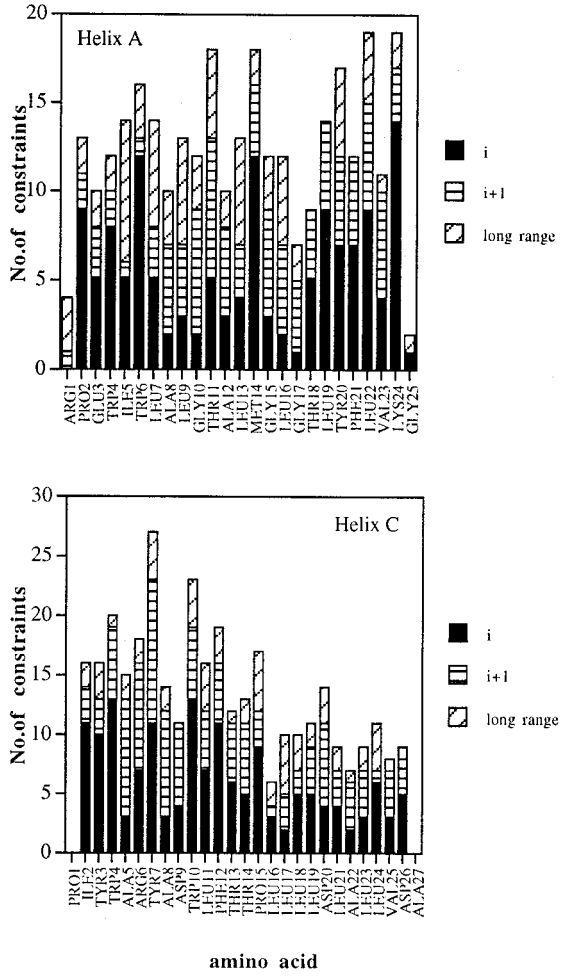
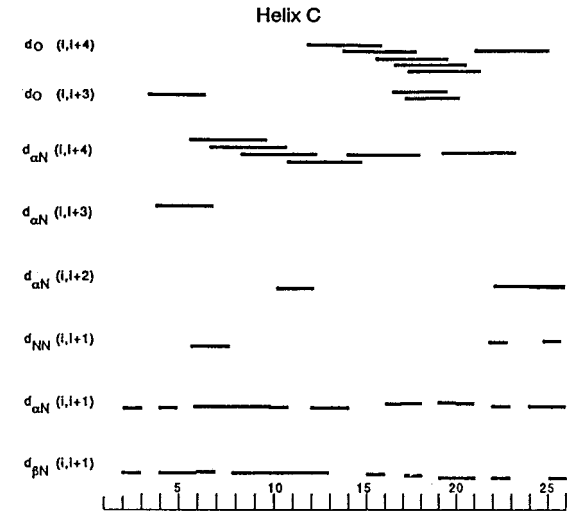


FIGURE 2 Numbers of constraints (per residue) obtained from the NOESY data for the peptides from helix A and from helix C. Black bar, intra-residue constraints; horizontal hatched bar, sequential constraints; diagonal hatched bar, long-range constraints.

stable in detergent micelles either (Hunt et al., 1997) (however, helix G was stable in organic solvent when part of a peptide that included helix F as well (Barsukov et al., 1992)). Previous studies had shown that peptides encoding three loops of bacteriorhodopsin were not only soluble in DMSO but also exhibited the same structure in solution in DMSO as the same sequences adopted in the crystal structure of the intact protein (Katragadda et al., 2000). DMSO was therefore chosen as a solvent that could be used in common for the 12 of the 13 peptides.

The structures of these 12 peptides in solution were determined by two-dimensional homonuclear ^{1}H NMR as described in detail previously (Yeagle et al., 1997, 1995; Katragadda et al., 2000). Details of the structure determination for three of the loops were described previously (Katragadda et al., 2000). Details of two of the helices, helix A and helix C, are presented here as examples, including the number of constraints per residue (Fig. 2), the connectivities



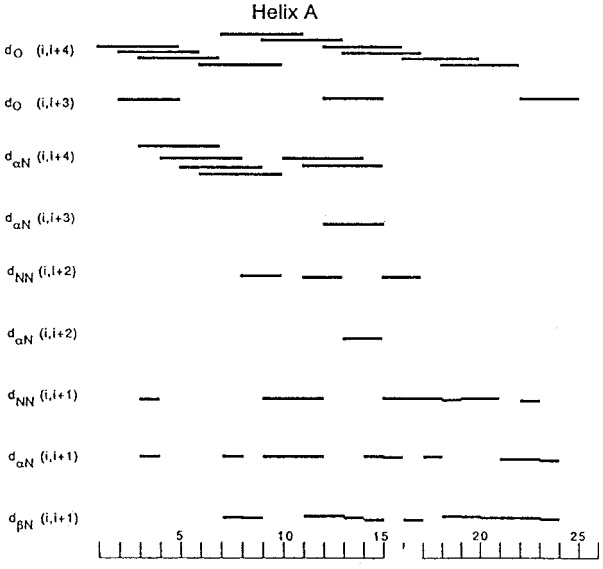


FIGURE 3 Connectivities observed in the NOESY data for the peptides from helix A and from helix C.

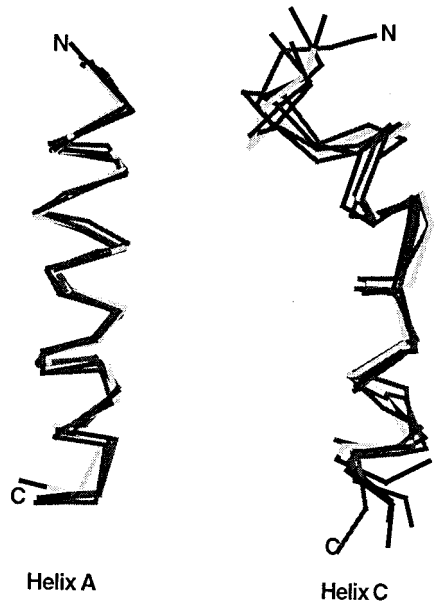


FIGURE 4 Families of structures obtained for the peptides from helix A and from helix C. Six structures are superimposed for each peptide in this representation.

from the NOE data (Fig. 3), and the families of structures superimposed (Fig. 4). These are similar to results from solution NMR of helices from other membrane proteins (Chopra et al., 2000). Table 1 describes the constraints used in the structure determinations for these peptides. These include distance constraints from NOESY data and torsion angle constraints (ϕ) from COSY data. Each of the 12 peptides exhibited one family of structures in solution. Families of structures were calculated from the constraints. Good overlap of the members of the family was observed over most of the sequence. The ends of each of the peptides were disordered as expected for peptide structures. Table 1 shows the average pair-wise rmsd of the families of structures. Table 1 also indicates which residues were in the well-ordered regions of the structures.

Peptides from the loop regions CD, DE, and FG of bacteriorhodopsin formed loops in solution (Katragadda et al., 2000). Fig. 2 shows that the other three peptides corresponding to loops of bacteriorhodopsin also formed loops in solution. Furthermore, peptides corresponding to helices A, B, C, D, E, and F of bacteriorhodopsin formed helices in solution. This is in agreement with previous studies that showed a helical structure in organic solvent and/or detergent micelles for peptides of helix B, helices A + B, and helices F + G (Lomize et al., 1992; Barsukov et al., 1992; Pervushin et al., 1994).

TABLE 1 Summary of data for peptide structures

Peptide	Length	Number of NOE constraints	Number of angle constraints	RMSD of the family	Well ordered residues
Helix A	25	223	5	0.75	2–24
Loop AB	32	263	2	1.7	10–30
Helix B	25	213	4	1.9	3–20
Loop BC	28	269	4	2.5	7–22
Helix C	27	246	5	1.1	6–25
Loop CD	23	156	3	1.2	2–21
Helix D	23	196	5	1.6	3–19
Loop DE	15	168	0	0.8	2–13
Helix E	32	311	2	1.5	2–30
Loop EF	22	192	2	1.2	3–20
Helix F	32	280	6	1.6	10–30
Loop FG	24	163	3	0.9	2–21

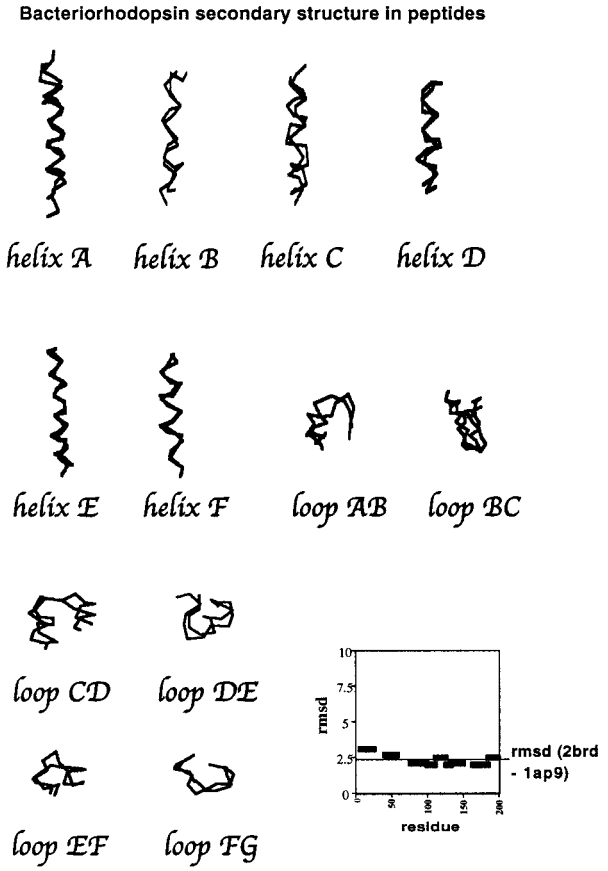


FIGURE 5 α-Carbon maps of the superposition of the backbone atoms of the peptide structures on the corresponding sequences in one crystal structure (2BRD) of bacteriorhodopsin. In each case, one member of the family of peptide structures, randomly chosen, was superimposed on the crystal structure. Similar results were obtained from superposition of these peptide structures on 1AP9. The superpositions were calculated using only the well-ordered portions of the peptide structures, as listed in Table 1. (*Inset*) The rmsd (Å) of the superposition is plotted for each peptide as a function of the sequence of bacteriorhodopsin. The horizontal line represents the average rmsd of superposition of 2BRD on 1AP9.

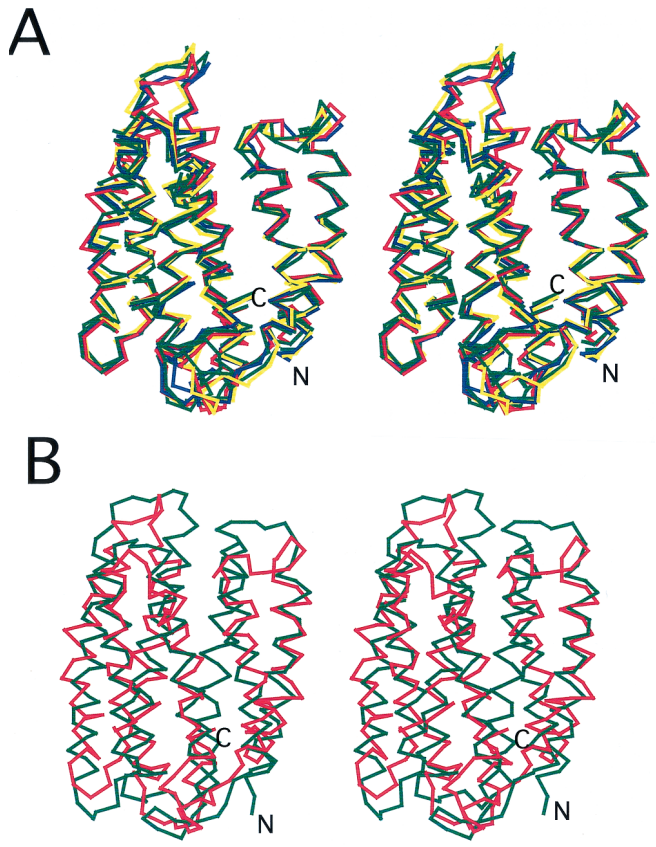


FIGURE 6 Three-dimensional structure of bacteriorhodopsin determined as described in this work. (*A*) Five structures obtained from simulated annealing as described in the text superimposed on each other, with the non-experimental part of helix G removed. (*B*) One of the structures from *A* (red) superimposed on one of the crystal structures (green) of bacteriorhodopsin (2brd).

Comparison of peptide structures to crystal structure of bacteriorhodopsin

Fig. 2 also shows the superposition of the structures of these peptides on the corresponding portions of one of the crystal structures. The peptides corresponding individually to helices A through F from bacteriorhodopsin each form a helix that mimics the crystal structure. The peptides corresponding to all six turns from bacteriorhodopsin form turns with the same residues in both the crystal structure and the peptide. Most of these peptide structures superimpose on the crystal structure with an average backbone rmsd less than 2.5 Å (see *inset*), indicating that the structures in the peptides are similar to the crystal structure (given the limited number of constraints available from NMR from modest-sized peptides in solution) (Katragadda et al., 2000). For comparison, the overall backbone rmsd between two crystal structures of bacteriorhodopsin (2BRD and 1AP9) is ~2.3.

Analysis of the B factors from the crystal structures shows that the electron densities of loops AB, BC, and EF are not well defined in the available crystal structures. Because of these limitations, it is not possible to quantitatively compare the structures of these peptides to the structures of the corresponding loops in the crystal structure. Therefore, there are blank regions in the plot inset in Fig. 2. Nevertheless, the residues that populate the turns in the crystal structures and in the peptides are the same.

These data indicate that most of the secondary structure of bacteriorhodopsin can be accurately captured in this series of overlapping peptides that span the sequence of the protein. This information is useful in determining the start and stop points of helices, whether the helices are straight or bent, and which residues are in the turns that connect the transmembrane helices. Some suggestion of the ability to get such information on secondary structure has been observed previously for another protein. Work on rhodopsin, also consisting of a bundle of seven transmembrane helices,

showed that the stop point of helix 5 in the third cytoplasmic loop determined from the structure of a peptide (Yeagle et al., 1997) was the same as had been identified from spin-label experiments on intact rhodopsin (Yang et al., 1996b) (the stop point of this helix cannot be determined from the recent crystal structure of rhodopsin due to the high B factors (Palczewski et al., 2000)). Likewise, the stop point of helix 7 determined from peptide structure (Yeagle et al., 2000) was the same as identified from spin-label experiments on intact rhodopsin (Altenbach et al., 1999) and the same as in the recent crystal structure of rhodopsin (Palczewski et al., 2000).

Use of peptide structures to determine the three-dimensional structure of bacteriorhodopsin

Exploiting the overlap of adjacent peptides, a continuous construct (residues 1–203) of all the peptides can be made by superimposing the backbone atoms of the overlapping sequences. All available experimental distance constraints were then written on this construct, including 2681 distance range constraints and 41 angle constraints from the NMR data on the individual peptides. As described below, some inter-helical distance constraints were used in this structure determination to help organize the helical bundle, and they were obtained from 1brd, a 7-Å-resolution diffraction study of two-dimensional crystals of bacteriorhodopsin (Henderson et al., 1990). We determined inter-helical distances between pairs of helices at the top, middle, and bottom, and assigned top, for example, to a residue near the top of the helix in our solution structure. These inter-helical distances had to be converted to distances between specific α-carbons of specific residues to be included in the constraint list for simulated annealing. This conversion was performed by taking into account the differences between a center-to-center distance for two helices and the distances between atoms on the two helices that are not at the helical center. The distances are range constraints to account for the uncertainties in this procedure (Table 2). Alternatively, the necessary constraints can be obtained from a variety of measurements, including solid-state NMR experiments (rotational resonance), cysteine scanning (disulfide bond formation), dipolar interactions between spin labels, fluorescence energy transfer measurements, engineering of metal binding sites, or complementary mutagenesis experiments.

Simulated annealing is used to optimize the conformation of the construct with respect to all the experimental constraints simultaneously. The result is shown in Fig. 3 as a family of structures with an average pair-wise backbone rmsd of 1.55. Fig. 3 also shows an overlay of this structure on a previously determined crystal structure of bacteriorhodopsin. The rmsd of simultaneous superposition of all the helices of the current structure on 2brd is 2.9 (the B factors from x-ray and electron diffraction studies are high in the loops, thus limiting quantitative comparisons to the transmembrane helical bundle with much lower B factors). Somewhat poorer superpositions result if 1ap9 or 1c3w are used. Thus, considerable agreement between the three-dimensional structure determined by the method described here and the structures determined by electron and x-ray diffraction is observed.

These results suggest that for transmembrane proteins built around helical bundles, considerable structural information can be obtained from the segmented approach de-

TABLE 2 Inter-helical constraints taken from the electron diffraction studies of bacteriorhodopsin (1BRD)

Atom 1	Atom 2	Distance range
Bottom A(THR17)	Bottom B(TYR57)	3.1–7.1
Top A(GLY31)	Top B(TYR43)	3.4–7.4
Middle A(THR24)	Middle B(ALA51)	3.4–7.4
Middle A(MET20)	Middle C(ASP85)	9.7–13.7
Top A(GLY31)	Top C(ASP96)	12.4–16.4
Bottom C(ILE78)	Bottom A(LEU13)	7.3–11.3
Bottom A(LEU13)	Bottom D(LEU123)	17.7–21.7
Top A(GLY31)	Top D(ILE108)	22–26
Bottom B(SER59)	Bottom C(TRP80)	7.8–11.8
Top B(ILE45)	Top C(LEU95)	6.2–10.2
Middle B(ILE52)	Middle C(PHE88)	5.7–9.7
Top C(LEU95)	Top D(LEU109)	7–11
Middle C(LEU87)	Middle D(ILE117)	7.4–11.4
Bottom C(TRP80)	Bottom D(VAL124)	8.6–12.6
Bottom D(THR121)	Bottom E(TRP137)	4.2–8.2
Middle D(ALA114)	Middle E(ALA144)	3.9–7.9
Top D(THR107)	Top E(VAL151)	6.5–10.5
Top E(PHE156)	Top F(LYS172)	5.7–9.7
Middle E(LEU146)	Middle F(SER183)	7.1–11.1
Bottom E(PHE135)	Bottom F(LEU190)	6–10
Bottom A(GLU9)	FG loop(ASN202)	3.7–7.7
Bottom A(TRP12)	Bottom F(VAL188)	13.6–17.6
Top A(GLY31)	Top F(THR170)	14.6–18.6
Middle A(GLY23)	Middle F(VAL177)	15.4–19.4
Bottom B(SER59)	Bottom D(VAL124)	18.2–22.2
Top B(ILE45)	Top D(LEU109)	14–18
Bottom B(SER59)	Bottom E(TRP137)	23.5–17.5
Top B(LYS41)	Top E(GLY155)	17.2–21.2
Middle C(LEU84)	Middle E(ILE140)	16.7–20.7
Top C(LEU95)	Top E(VAL151)	10.7–14.7
Middle C(TYR83)	Middle F(PRO186)	10.9–14.9
Top C(LEU94)	Top F(ASN176)	11.5–15.5
Top D(LEU111	Top F(ASN176)	15.9–19.9
Bottom D(GLY122)	Bottom F(PRO186)	7.3–11.3
Top C(ASP96)	Top G(LEU223)	7.3–11.3
Middle C(ASP85)	Middle G(ASP212)	8.5–12.5
Bottom C(TYR79)	Bottom G(THR205)	7.4–11.4
Top F(PHE171)	Top G(LEU223)	6.6–10.6
Middle F(SER184)	Middle G(LEU211)	6–10
Bottom F(SER193)	Bottom G(LEU201)	3.7–7.7
Bottom A(TRP12)	Bottom G(LEU206)	3.4–7.4
Middle A(MET20)	Middle G(VAL213)	4.8–8.8
Top A(LEU28)	Top G(LEU223)	11.6–15.6
Middle B(PRO50)	Middle G(LYS216)	4–8
Top B(PHE42)	Top G(LEU223)	5.7–9.7
Bottom B(LEU61)	Bottom G(THR205)	8.9–12.9

Distances are α-carbon to α-carbon (see text).

scribed in this work. These results offer an alternative approach to obtain some useful structural information in the absence of a high-resolution x-ray crystallographic study.

We thank Dr. Arlene Albert for many helpful discussions, J. Yeagle for assistance in data analysis, and Drs. F. Mauri and J. Landin for synthesizing the peptides used in this work.

This work was supported by National Institutes of Health grant EY03328 and in part by CA16056.

REFERENCES

Adler, M., M. H. Sato, D. E. Nitecki, J.-H. Lin, D. R. Light, and J. Morser. 1995. The structure of a 19 residue fragment from the C loop of the fourth epidermal growth factor-like domain of thrombomodulin. *J. Biol. Chem.* 270:23366–23372.

Altenbach, C., K. Cai, H. G. Khorana, and W. L. Hubbell. 1999. Structural features and light-dependent changes in the sequence 306–322 extending from helix VII to the palmitoylation sites in rhodopsin: a site-directed spin-labeling study. *Biochemistry.* 38:7931–7937.

Barsukov, I. L., D. E. Nolde, A. L. Lomize, and A. S. Arseniev. 206. 1992. Three-dimensional structure of proteolytic fragment 163–231 of bacterioopsin determined from nuclear magnetic resonance data in solution. *Eur. J. Biochem.* 665–672.

Behrends, H. W., G. Folkers, and A. G. Beck-Sickinger. 1997. A new approach to secondary structure evaluation: secondary structure prediction of porcine adenylate kinase and yeast guanylate kinase by CD spectroscopy of overlapping synthetic peptide segments. *Biopolymers.* 41:213–231.

Blanco, F. J., and L. Serrano. 1994. Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. *Eur. J. Biochem.* 230:634–649.

Blumenstein, M., G. R. Matsueda, S. Timmons, and J. Hawiger. 1992. A β turn is present in the 392–411 segment of the human fibrinogen γ chain: effects of structural changes in this segment on affinity to antibody 4A5. *Biochemistry.* 31:10692–10698.

Callihan, D. E., and T. M. Logan. 1999. Conformations of peptide fragments from the FK506 binding protein: comparison with the native and urea-unfolded states. *J. Mol. Biol.* 285:2161–2175.

Campbell, A. P., C. McInnes, R. S. Hodges, and B. D. Sykes. 1995. Comparison of NMR structures of the receptor binding domains of *Pseudomonas aeruginosa* Pili strains PAO, KB7, and PAK: implications for receptor binding and synthetic vaccine design. *Biochemistry.* 34: 16255–16268.

Chandrasekhar, K., A. T. Profy, and H. J. Dyson. 1991. Solution conformation preferences of immunogenic peptides derived from the principal neutralizing determinant of the HIV-1 envelope glycoprotein gp120. *Biochemistry.* 30:9187–9194.

Chopra, A., P. L. Yeagle, J. A. Alderfer, and A. Albert. 2000. Solution structure of the sixth transmembrane helix of the G-protein coupled receptor, rhodopsin. *Biochim. Biophys. Acta.* 1463:1–5.

Cox, J. P. L., P. A. Evans, L. C. Packman, D. H. Williams, and D. N. Woolfson. 1993. Dissecting the structure of a partially folded protein. *J. Mol. Biol.* 234:483–492.

Dmitriev, O., P. C. Jones, W. Jiang, and R. H. Fillingame. 1999. Structure of the membrane domain of subunit b of the *Escherichia coli* F0F1 ATP synthase. *J. Biol. Chem.* 274:15598–15604.

Dyson, H. J., G. Merutka, J. P. Waltho, R. A. Lerner, and P. E. Wright. 1992. Folding of peptide fragments comprising the complete sequence of proteins. *J. Mol. Biol.* 226:795–817.

Fan, J. S., H. C. Cheng, and M. Zhang. 1998. A peptide corresponding to residues asp177 to asn208 of human cyclin a forms an alpha helix. *Biochem. Biophys. Res. Commun.* 253:621–627.

Gao, J., Y. Li, and H. Yan. 1999. NMR solution structure of domain 1 of human annexin I shows an autonomous folding unit. *J. Biol. Chem.* 274:2971–2977.

Gegg, C. V., K. E. Bowers, and C. R. Matthews. 1997. Probing minimal independent folding units in dihydrofolate reductase by molecular dissection. *Protein Sci.* 6:1885–1892.

Ghiara, J. B., E. A. Stura, R. L. Stanfield, A. T. Profy, and I. A. Wilson. 1994. Crystal structure of the principal neutralization site of HIV-1. *Science.* 264:82–85.

Gouaux, E. 1998. It's not just a phase: crystallization and x-ray structure determination of bacteriorhodopsin in lipidic cubic phases. *Structure.* 6:5–10.

Goudreau, N., F. Cornille, M. Duchesne, F. Parker, B. Tocqué, C. Garbay, and B. P. Roques. 1994. NMR structure of the N-terminal SH3 domain of GRB2 and its complex with a proline-rich peptide from SOS. *Nat. Struct. Biol.* 1:898–907.

Grigorieff, N., T. A. Ceska, K. H. Downing, J. M. Baldwin, and R. Henderson. 1996. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* 259:393–421.

Hamada, D., Y. Kuroda, T. Tanaka, and Y. Goto. 1995. High helical propensity of the peptide fragments derived from β-lactoglobulin, a predominantly β-sheet protein. *J. Mol. Biol.* 254:737–746.

Henderson, R., J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing. 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* 213: 899–929.

Hunt, J. F., T. N. Earnest, O. Bousche, K. Kalghatgi, K. Reilly, C. Horvath, K. J. Rothschild, and D. M. Engelman. 1997. A biophysical study of integral membrane protein folding. *Biochemistry.* 36:15156–15176.

Jimenez, M. A., J. A. Evangelio, C. Aranda, A. Lopez-Brauet, D. Andreu, M. Rico, R. Lagos, and J. M. Andreu. 1999. Helicity of alpha(404–451) and beta(394–445) tubulin C-terminal recombinant peptides. *Protein Sci.* 8:788–799.

Katragadda, M., J. L. Alderfer, and P. L. Yeagle. 2000. Solution structure of the loops of bacteriorhodopsin closely resemble the crystal structure. *Biochim. Biophys. Acta.* 1466:1–6.

Kumar, A., R. R. Ernst, and K. Wuthrich. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* 95:1–6.

Lomize, A. L., K. V. Pervushin, and A. S. Arseniev. 1992. Spatial structure of (34–65)bacterioopsin polypeptide in SDS micelles determined from nuclear magnetic resonance data. *J. Biomol. NMR.* 2:361–372.

Ludvigsen, S., K. V. Andersen, and F. M. Poulsen. 1991. Accurate measurements of coupling constants from two-dimensional nuclear magnetic resonance spectra of proteins and determination of phi-angles. *J. Mol. Biol.* 217:731–736.

Luecke, H., B. Schobert, H.-T. Richter, J.-P. Cartailler, and J. J. K. Lanyi. 1999. Structure of bacteriorhodopsin at 1.55 angstrom resolution. *J. Mol. Biol.* 291:899.

Padmanabhan, S., M. A. Jimenez, and M. Rico. 1999. Folding propensities of synthetic peptide fragments covering the entire sequence of phage 434 Cro protein. *Protein Sci.* 8:1675–1688.

Palczewski, K., T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, and M. Miyano. 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science.* 289:739–745.

Pebay-Peyroula, E., G. Rummel, J. P. Rosenbusch, and E. M. Landau. 1997. X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science.* 277:1676–1681.

Pervushin, K. V., V. Y. Orekhov, A. I. Popov, L. Y. Musina, and A. S. Arseniev. 1994. Three-dimensional structure of (1–71)bacterioopsin solubilized in methanol/chloroform and SDS micelles determined by 15N–1H heteronuclear NMR spectroscopy. *Eur. J. Biochem.* 219: 571–583.

Popot, J.-L., and D. M. Engelman. 2000. Helical membrane protein folding, stability, and evolution. *Annu. Rev. Biochem.* 69:881–922.

Ramirez-Alvarado, R., L. Serrano, and F. J. Blanco. 1997. Conformational analysis of peptides corresponding to all the secondary structure elements of protein L B1 domain. *Protein Sci.* 6:162–174.

Reymond, M. T., G. Merutka, H. J. Dyson, and P. E. Wright. 1997. Folding propensities of peptide fragments of myoglobin. *Protein Sci.* 6:706–716.

Wilce, J. A., D. Salvatore, J. D. Waade, and D. J. Craik. 1999. H-1 NMR structural studies of a cystine-linked peptide containing residues 71–93 of transthyretin and effects of a ser84 substitution implicated in familial amyloidotic polyneuropathy. *Eur. J. Biochem.* 262:586–594.

Wüthrich, K., M. Billeter, and W. J. Braun. 1983. Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* 169:949–961.

Yang, A.-S., B. Hitz, and B. Honig. 1996a. Free energy determinants of secondary structure formation: β turns and their role in protein folding. *J. Mol. Biol.* 259:873–882.

Yang, K., D. L. Farrens, W. L. Hubbell, and H. G. Khorana. 1996b. Structure and function in rhodopsin: single cysteine substitution mutants in the cytoplasmic interhelical E-F loop region show position-specific effects in transducin activation. *Biochemistry*. 35: 12464–12469.

Yeagle, P. L., J. L. Alderfer, and A. D. Albert. 1995. Structure of the carboxyl terminal domain of bovine rhodopsin. *Nat. Struct. Biol.* 2:832–834.

Yeagle, P. L., J. L. Alderfer, and A. D. Albert. 1997. The first and second cytoplasmic loops of the G-protein receptor, rhodopsin, independently form β-turns. *Biochemistry.* 36:3864–3869.

Yeagle, P. L., C. Danis, G. Choi, J. L. Alderfer, and A. D. Albert. 2000. Three dimensional structure of the seventh transmembrane helical domain of the G-protein receptor, rhodopsin. *Molecular Vision*. www.molvis.org/molvis/v 6/a17/.