NMR Structure Determination of a Membrane Protein with Two Transmembrane Helices in Micelles: MerF of the Bacterial Mercury Detoxification System[†,‡]

Stanley C. Howell, Michael F. Mesleh, and Stanley J. Opella*

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0307

Received September 3, 2004; Revised Manuscript Received February 8, 2005

ABSTRACT: The three-dimensional backbone structure of a membrane protein with two transmembrane helices in micelles was determined using solution NMR methods that rely on the measurement of backbone $^{1}H-^{15}N$ residual dipolar couplings (RDCs) from samples of two different constructs that align differently in stressed polyacrylamide gels. Dipolar wave fitting to the $^{1}H-^{15}N$ RDCs determines the helical boundaries based on periodicity and was utilized in the generation of supplemental dihedral restraints for the helical segments. The $^{1}H-^{15}N$ RDCs and supplemental dihedral restraints enable the determination of the structure of the helix−loop−helix core domain of the mercury transport membrane protein MerF with a backbone RMSD of 0.58 Å. Moreover, the fold of this polypeptide demonstrates that the two vicinal pairs of cysteine residues, shown to be involved in the transport of Hg(II) across the membrane, are exposed to the cytoplasm. This finding differs from earlier structural and mechanistic models that were based primarily on the somewhat atypical hydropathy plot for MerF and related transport proteins.

NMR spectroscopy is well suited for determining the structures of helical membrane proteins; difficult to obtain crystals are not necessary, and there is sufficient flexibility in the experimental methods for studies of proteins in the principal types of lipid assemblies (*1*). Solid-state NMR of aligned bilayer samples (*2*−*7*) as well as solution NMR of micelle samples (*8*−*16*) has been used to determine the three-dimensional structures of membrane peptides, proteins, or their domains with the equivalent of one transmembrane helix. There have been NMR structural studies of helices associated with the surfaces of micelles as well (*17*−*19*). However, many of the solution NMR studies of larger membrane proteins containing more than one transmembrane helix in micelles have encountered spectroscopic complexities in the forms of missing, doubled, or broadened resonances that are generally attributed to the variable effects of internal dynamics of loop regions as well as the helices (*20*−*23*). NMR studies of membrane proteins in mixed organic solvents are problematic in terms of spectroscopy and uncertainties about the effect of environment on the structures (*24*−*27*), and indeed for subunit c of the F_1F_0 ATPase, the most thoroughly studied example, SDS[1] micelles have been shown to provide more relevant structural insights than solvents based on TFE or chloroform and methanol (*28*). With the increasing interest in structure determination of membrane proteins, a wide range of lipids have been screened (*29*, *30*) for their potential to form small, homogeneous protein-containing micelles; however, the best characterized and most reliable choices for both maintaining native, functional protein structures and obtaining well-resolved solution NMR spectra remain SDS (*31*) and DPC (*32*).

Although high-quality solution NMR data, including resonance assignments, have been obtained for proteins with two and three transmembrane helices in both SDS and DPC micelles (*20*, *33*, *34*), there have been no three-dimensional structures reported, largely because of the difficulties encountered in resolving and assigning a sufficient number of "long-range" NOEs to determine their folds. Previously, we have shown that dipolar waves (*34*−*36*), which describe the periodic variation of the magnitudes of dipolar couplings in the backbone of a protein as a function of residue number, accurately identify the residues in helices, any deviations (kinks or curvature) from ideality, and the relative orientations of helices in weakly aligned micelle samples of membrane proteins (*36*, *37*).

The results presented in this paper demonstrate that it is possible to determine the three-dimensional backbone structure of a helical membrane protein by measuring residual dipolar couplings for two different alignments of protein-containing micelles. The 60-residue polypeptide corresponding to the core of the mercury transport membrane protein, MerF, contains two transmembrane helices and an interhe-

[†] This research was supported by Grants RO1EB002169 and RO1GM066978 from the National Institutes of Health and utilized the Biomedical Technology Resource for NMR Molecular Imaging of Proteins supported by Grant P41EB002031. S.C.H. is supported by training grant DK54441.

[‡] Structure coordinates have been deposited in the Protein Data Bank as entry 1WAZ.

* To whom correspondence should be addressed. Phone: (858) 822-4820. Fax: (858) 822-4821. E-mail: sopella@ucsd.edu.

[1] Abbreviations: APS, ammonium persulfate; DHPC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HMQC-NOESY, heteronuclear multiple-quantum coherence nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single-quantum coherence; IPTG, isopropyl-*β*-thiogalactoside; KSI, ketosteroid isomerase; NMR, nuclear magnetic resonance; RDC, residual dipolar coupling; RMSD, root mean square deviation; SDS, sodium dodecyl sulfate; TEMED, *N*,*N*,*N*′,*N*′-tetramethylethylenediamine; TFE, trifluoroethanol.

10.1021/bi048095v CCC: $30.25
Published on Web 03/11/2005

lical loop, representative of the principal structural features of helical membrane proteins identified in 20−30% of the proteins encoded in a typical genome (*38*). Dipolar wave analysis contributes to and complements the calculation of the three-dimensional structure with RDCs as the principal input for the program X-PLOR-NIH (*39*); some refinement of the structure results from the inclusion of chemical shift derived dihedral angle constraints.

Bacteria that survive in mercury-polluted sediment contain a transmissible plasmid with an operon that encodes the sequences of several proteins whose functions constitute a mercury detoxification system that enzymatically transforms Hg(II) to Hg(0) (*40*). Hg(II) is transported from the periplasm, where it binds to the protein MerP (*41*, *42*), across the membrane into the cytoplasm by a mercury transport membrane protein (*43*). Genes for three different mercury transport membrane proteins have been identified in various strains of mercury-resistant bacteria (*44*). MerF has 81 residues and two transmembrane helices and has 20% sequence identity with the N-terminal two-thirds of MerT (116 residues and three transmembrane helices); MerC has 143 residues and four transmembrane helices (*45*), and similar levels of sequence similarities with MerF and MerT. All of the proteins that handle Hg(II) contain one or more pairs of cysteine residues. In MerF, both pairs of cysteine residues are vicinal, and it has been hypothesized that they participate in a conduit for Hg(II) from the pair of Cys residues on MerP (*42*) to the pairs of Cys residues on the homologous N-terminal metal binding domains of MerA (*48*), the cytoplasmic enzyme that reduces Hg(II) to Hg(0) (*45*). This model will require some modification in order to accommodate the placement of the pairs of cysteine residues based on the structural results described in this paper. One pair of cysteines that is essential for function is located immediately before the first transmembrane helix and a second pair that affects function directly after the second transmembrane helix (*45*−*47*). This places both pairs of cysteine residues on the cytoplasmic side of the membrane, which is markedly different than the earlier models based on hydropathy plots (*45*).

With only 81 residues, two transmembrane helices, one interhelical loop, and the essential biological function of transporting Hg(II) across membranes, MerF is an attractive target for structure determination. On the basis of results obtained in our earlier studies (*33*, *34*, *49*), two protein constructs were designed to address the principal sources of experimental difficulties encountered in NMR experiments. The wild-type sequence contains four cysteines and requires large amounts of a reducing agent in the samples to prevent aggregation and maintain the native structure; however, since the reducing agent is oxidized over time, the spectral quality degrades during signal averaging. Additionally, the spectral overlap due to the narrow resonances from the mobile, unstructured residues near the N- and C-termini is problematic. A truncated 60-residue version of MerF, designated $MerF_t$, was cloned, expressed, and purified. This polypeptide encompasses residues 13−72 of the full-length protein and includes all of the residues in the two transmembrane helices and interhelical loop and a few from the N- and C-terminal regions. All four cysteine residues are replaced by isosteric serines in this construct. As a result, samples of $MerF_t$-containing micelles are stable at significantly higher protein concentrations and for longer periods of time than those containing wild-type MerF. Similarly, $MerF_m$, an 80-residue construct containing the sequence of the full-length MerF, except for the four Cys to Ser mutations and the N-terminal Met residue, retains the favorable solubility and stability properties of $MerF_t$. Valuable spectroscopic and structural comparisons can be made between these two constructs. Moreover, the micelles containing the 60-residue $MerF_t$ and the 80-residue $MerF_m$ polypeptides align differently in stressed polyacrylamide gels. This is significant, since it enables the measurement of backbone residual dipolar couplings at two different alignments, which is essential for resolving orientational ambiguities encountered in the analysis of dipolar waves fit to the helical segments and the constraints used for the calculation of the three-dimensional structures (*50*, *51*). Furthermore, spectroscopic comparisons among the truncated polypeptide ($MerF_t$) and the full-length polypeptides ($MerF_m$ and wild-type MerF) demonstrate that the three-dimensional structure of the 46 residues that constitute the helix−loop−helix core domain is unaffected by the presence of the terminal residues or cysteine substitutions.

EXPERIMENTAL PROCEDURES

Materials. Enzymes were purchased from New England Biolabs (www.neb.com) unless otherwise noted, and the oligonucleotides were synthesized by Integrated DNA Technologies (www.idtdna.com). Plasmid DNA of pET-31(+)b and bacterial strains NovaBlue and BL21(λDE3)pLysS were purchased from Novagen (www.emdbiosciences.com). Bacterial strain C43(λDE3) was obtained from Avidis (www.avidis.fr). [$^{15}N_2$]Ammonium sulfate, ^{15}N-labeled amino acids, [$^{13}C_6$]glucose, and sodium dodecyl-d_{25} sulfate were obtained from Cambridge Isotope Laboratories (www.isotope.com).

Cloning, Expression, and Purification of MerF Constructs. The coding sequences of $MerF_t$ and $MerF_m$ were constructed independent of the MerF parental cDNA to enable additional genetic optimization, simultaneous insertion of the targeted mutations, and incorporation of restriction sites for vector preparation. The $MerF_t$ sequence was constructed from mutual priming of long oligonucleotide fragments and the $MerF_m$ sequence by overlap extension of the $MerF_t$ coding sequence. Both sequences were restricted at engineered *Alw*NI and *Xho*I restriction sites and inserted into the pET31-(+)b vector, which contains a 125-residue N-terminal ketosteroid isomerase (KSI) fusion for direction of protein expression to inclusion bodies and a C-terminal 6× His tag for affinity purification. Plasmids were transformed into NovaBlue cells and verified by DNA sequencing on an ABI3100 by the Center for AIDS Research Molecular Biology Core at the University of California, San Diego. Amino acid sequences of the protein constructs are given in Table 1.

Verified plasmids were transformed for protein expression into C43(λDE3) cells for uniform isotopic labeling and into BL21(λDE3)pLysS for selective isotopic labeling. Expression was carried out by inducing cells in M9 minimal medium containing $(^{15}NH_4)_2SO_4$ for 7 h by the addition of IPTG. Cells were disrupted by sonication. The inclusion bodies were isolated by centrifugation and then solubilized in 6 M guanidine hydrochloride and purified using Ni^{2+} affinity

Table 1: Amino Acid Sequence of MerF Constructs

```
MerF    MKDPKTLLRVSIIGTTLVALCCFTPVLVILLGVVGLSALTGYLDYVLLPALAIFIGLTIYAIQRKRQADACCTPKFNGVKK
MerFt             IGTTLVALSSFTPVLVILLGVVGLSALTGYLDYVLLPALAIFIGLTIYAIQRKRQADASS*
MerFm    KDPKTLLRVSIIGTTLVALSSFTPVLVILLGVVGLSALTGYLDYVLLPALAIFIGLTIYAIQRKRQADASSTPKFNGVKKS*
                  10        20        30        40        50        60        70        80
                                                                          S* - homoserine
```

chromatography. The polypeptides were cleaved from the N-terminal KSI fusion and C-terminal His tag by CNBr in 70% formic acid and then purified by size-exclusion chromatography on a Sephacryl S-200 column (100 mM Na_2HPO_4, 4 mM SDS, 1 mM EDTA, 1 mM NaN_3, pH = 7.4). In the case of wild-type MerF, the buffer also contained 20 mM DTT. Detergent was removed from the protein-containing fractions by exhaustive dialysis, and the precipitated protein was pelleted by centrifugation and lyophilized prior to further use.

NMR Sample Preparation. All NMR samples were prepared by solubilizing the lyophilized protein in the SDS-containing NMR buffer (90% H_2O/10% D_2O, 500 mM SDS-d_{25}, 10 mM sodium phosphate, 1 mM sodium azide, pH 6.0). Samples of wild-type MerF also contained 40 mM DTT. The protein-containing micelles were incubated at 40 °C for 30 min with intermittent bath sonication and microcentrifuged at 13000 rpm for 15 min to remove any undissolved protein or other debris. Selectively ^{15}N-labeled samples were prepared in the same manner. Hydrogen−deuterium fractionation experiments (*52*) were performed by lyophilizing the NMR sample, resolubilizing the protein-containing micelles in the same volume with increasing ratios of ^2H_2O to ^1H_2O, and then incubating and bath sonicating the sample for 1 h before putting it back into the NMR tube. Samples were made with ^2H_2O concentrations of 10%, 25%, 50%, 75%, 90%, 95%, and 99%, and the pH was adjusted for all samples taking into account the uncorrected pH meter readings, 0.4 unit lower for 100% ^2H_2O. Small "isotropic" bicelles (q = 0.25) were prepared using 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) and 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) with a total lipid content of 7.5% (w/v) (*53*, *54*). Protein dissolved in trifluoroethanol was combined with DHPC in chloroform. The solvent was removed, and the resulting film of protein and lipid was redissolved in 90% ^1H_2O/10% ^2H_2O, 1 mM PO_4 buffer, and 1 mM NaN_3, pH 6.0, and bath sonicated. The resulting micelle solution was added to DMPC and subjected to repeated freeze−thaw cycles.

Protein-containing micelles were weakly aligned in stressed polyacrylamide gels (*55*). Separate samples were aligned in a 6% polyacrylamide gel using the method of compression to induce alignment in the isotropic samples (*56*, *57*) and in a 5% polyacrylamide gel using the method of Bax and coworkers (*58*). For the first method, a solution containing 30% (w/v) acrylamide and 0.8% bisacrylamide was diluted with deionized H_2O and polymerized by addition of TEMED and APS in a long glass tube with a 3.5 mm inner diameter. Following polymerization, the gel was squeezed out of the glass tube and washed in deionized water overnight to remove any unpolymerized acrylamide and unreacted TEMED/APS. The gel was then cut to 2.6 cm in length and dried overnight at 42 °C. The dried gel was then added to a solution of protein-containing micelles in a Shigemi (www.geocities.com/~shigemi/) NMR tube with an inner diameter of 4.2 mm, and the plunger was set to restrict the expansion of the length of the polyacrylamide gel to 1.8 cm. For the second alignment method, the gel was cast with a 5 mm diameter, cut to 2 cm in length, and dried as described above. After the solution of protein-containing micelles was soaked into the gel overnight, the gel was squeezed into a standard 4.2 mm inner diameter NMR tube that was sealed with epoxy. Induction of weak alignment in the samples was confirmed in a light microscope by optical birefringence prior to initiating NMR experiments.

NMR Spectroscopy. The NMR spectra shown in the figures were obtained on a Bruker (www.bruker-biospin.com) DMX 600 MHz spectrometer using a triple-resonance $^1H/^{13}C/^{15}N$ probe equipped with three-axis pulsed field gradients. The fractionation and dynamics experiments were performed on a Varian (www.varianinc.com) Inova 500 MHz spectrometer using a double-resonance $^1H/^{15}N$ probe equipped with a *z*-axis pulsed field gradient. All NMR experiments were performed at 60 °C using a 1.5 s recycle delay. The chemical shifts are referenced to the ^1H_2O resonance set to its expected position of 4.3935 ppm at 60 °C (*59*). The standard fHSQC experiment was used for isotropic samples with 1024 points in t_2 and 256 in t_1 (*60*). $^1H-^{15}N$ heteronuclear NOE measurements were made using difference experiments with and without 3.0 s of saturation of the ^1H resonances between scans (*61*). Initial backbone resonance assignments were made using a standard HNCA experiment with constant time evolution for ^{15}N, and solvent suppression was accomplished with a water flip-back pulse after the original $^1H-^{15}N$ magnetization transfer (*62*). However, due to the extensive overlap among the C_α resonances, which is typical of helical membrane proteins in micelles, spectra from several selectively ^{15}N-labeled protein samples were necessary to resolve assignment ambiguities. Two-dimensional HMQC-NOESY experiments performed with a 150 ms mixing time were also helpful in dealing with the resonance overlap encountered in two-dimensional spectra of uniformly ^{15}N-labeled samples (*37*, *63*).

NMR data were processed using the NMRPipe software package (*64*). Spectra were assigned, peak intensities measured, and peak volumes integrated using the program Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). With all of the ^1H, ^{15}N, and $^{13}C_\alpha$ backbone resonances assigned, the chemical shift index was then used to verify the residues in the helical segments.

For the $^1H/^2H$ fractionation experiments, the same sample was used for all measurements in order to quantitatively compare the peak intensities. Individual fractionation factors (χ) were obtained for each residue by plotting the normalized peak intensities as a function of mole fraction of ^1H_2O and fitting the expression

$$(yC)^{-1} = \chi(1 - X)/X + 1$$

where y is the peak intensity, C is a normalization constant for the intensity of the peak with 10% ^2H_2O, and X is the mole fraction of $^1H/^2H$ in the sample. The value of χ is equivalent to the slope of the line describing this relationship (*52*). Fractionation factors were fitted by linear regression in MATLAB (www.mathworks.com).

Weak alignment of the sample was confirmed by the observation of a splitting of the ^2H signal in the lock channel of the NMR spectrometer, and the spectral quality was monitored with two-dimensional spectra obtained using a version of the $^1H-^{15}N$ HSQC experiment modified for suppression of the NH_2 signals from the acrylamide (*57*). Measurements of the $^1H-^{15}N$ residual dipolar couplings were made using a sensitivity-enhanced version of the IPAP experiment (*65*). The contribution to the $^1H-^{15}N$ splitting from the isotropic scalar coupling was determined by performing the same experiment on an isotropic micelle sample and subtracting the value of the isotropic *J*-coupling obtained from that measured for the weakly aligned gel sample.

Analysis of Dipolar Couplings. The analysis of the experimentally measured RDCs was performed using custom scripts in MATLAB. Helical regions were identified by applying a sliding window fitting algorithm to the experimental RDCs as described previously (*36*). For each window of five residues, a simple sinusoid of periodicity 3.6 was best fitted to the experimental data, and the RMSD between that sinusoid and the experimental data was calculated. When the RMSDs are plotted as a function of residue number, the regions having low scores are identified as helices and are fitted as single sinusoids. Helices have RMSDs below the experimental error of ~1.5 Hz; higher scores are generally interpreted as deviations from ideality such as kinks and curvature (*34*, *36*). Simple fitting of an expression relating the orientation of the helix to the amplitude, average value, and phase of the sinusoid is an initial step toward structure determination and precisely determines the relative orientations of helices to within four degenerate solutions (*36*). Data from at least two different alignments of the protein are essential for resolving among the orientational ambiguities and obtaining unique structural solutions (*50*, *51*).

Two helical regions were identified on the basis of the periodicity of their dipolar waves in both orientations of both constructs. The first segment is from residues L27 to L39 and the second from residue A50 to residue R66 with both regions having RMSD values of less than 1.3 Hz, slightly lower than the experimental error. The inclusion of data from one or more additional residues on either end of the helices significantly reduces the quality of the fits. Also, the preceding residue is a proline in both cases (P25, P49). Values of D_a and R were determined by directly fitting the expression describing the sinusoidal oscillations for $^1H-^{15}N$ couplings for helices to the experimental data, and the values that optimized the fit to the data from *both* helices of the same sample were selected. For the compressed gel sample of $MerF_t$ the values of D_a and R were -9.15 Hz and 0.25, respectively. Residual dipolar couplings measured from $MerF_m$ gave similar results, but the protein was aligned differently within the pores of the gel ($D_a = -8.21$ Hz; $R = 0.14$).

Three-dimensional protein structures were calculated from the experimental data using a basic simulated annealing protocol in the program X-PLOR-NIH (*39*). An extended template with favorable covalent geometry was annealed against dihedral angle restraints having a force constant of 300 kcal/K in the regions of the protein found to be α-helical based on the periodicity of their dipolar waves. First, the protein was subjected to 10 ps of dynamics at 3000 K. An error bound of 1° was applied to dihedral angle restraints during dynamics to promote formation of the helical secondary structural elements and maintained through slow cooling to minimize helical distortion not present in the initial fitting. To facilitate the transition to the termination of helices, the first set and last set of ϕ and ψ dihedrals were given an error bound of 5°. The harmonic restraints for the $^1H-^{15}N$ residual dipolar couplings were introduced with the force constant linearly increased from 0.01 to 7 kcal Hz^{-1} K^{-1} as the temperature was lowered to 25 K in steps of 12.5 K over 120 ps with a fixed error bound of 1.05 Hz per coupling, as described in the rdcpot module of X-PLOR-NIH (*39*). A supplemental radius of gyration term (R_{gyr}), which reflects the packing tendencies of secondary structure elements, is frequently applied in order to improve the quality of structures calculated for soluble proteins from NMR data (*66*). Similarly, a R_{gyr} potential was applied in the calculations of the structure of the $MerF_t$ at a force constant of 50 kcal mol^{-1} $Å^{-2}$; additional optimization of R_{gyr} was performed to improve convergence by varying the residues influenced by this term and the applied radius as described below (*67*). Data from multiple alignments were incorporated using a separate pseudoatom for each alignment and weighting each dataset equally. Supplemental dihedral angle restraints were generated from ^{15}N and ^{13}C chemical shifts using the TALOS program (*68*) and implemented with the dihedral angles derived from the dipolar wave analysis at a reduced force constant of 30 kcal/K and an increased error bound of 50°. The inclusion of these restraints improved the convergence of the structure calculations, especially for the nonhelical residues. Side chain conformations were restrained using a database potential of mean force to maintain realistic geometries (*69*). The accepted structures have violations of the RDC values that are less than 1.05 Hz and deviations of the helical dihedral angles of less than 10° from those determined by dipolar wave fitting. Optimization of the R_{gyr} term was performed by arraying both the residues directly affected by this potential and then the magnitude of the radius in order to minimize structural distortion away from the experimental data. R_{gyr} values, between 8 and 30 Å, were tested, and the statistical information on the population of structures generated for each value is included in the Supporting Information. Results from the R_{gyr} optimization were spatially profiled using the DEFINE STRUCTURE program (*70*), which has previously been utilized in database analysis of membrane proteins of known structure (*71*). Parameters derived from the DEFINE STRUCTURE program are based upon analysis of the Cα distance matrix generated from the Cartesian coordinates of the calculated structures and are listed in the Supporting Information.

RESULTS

Two-dimensional $^1H-^{15}N$ HSQC spectra of uniformly ^{15}N-labeled polypeptides corresponding to three different constructs of MerF are shown in Figure 1. Sample optimization involved the comparison of spectra obtained in several

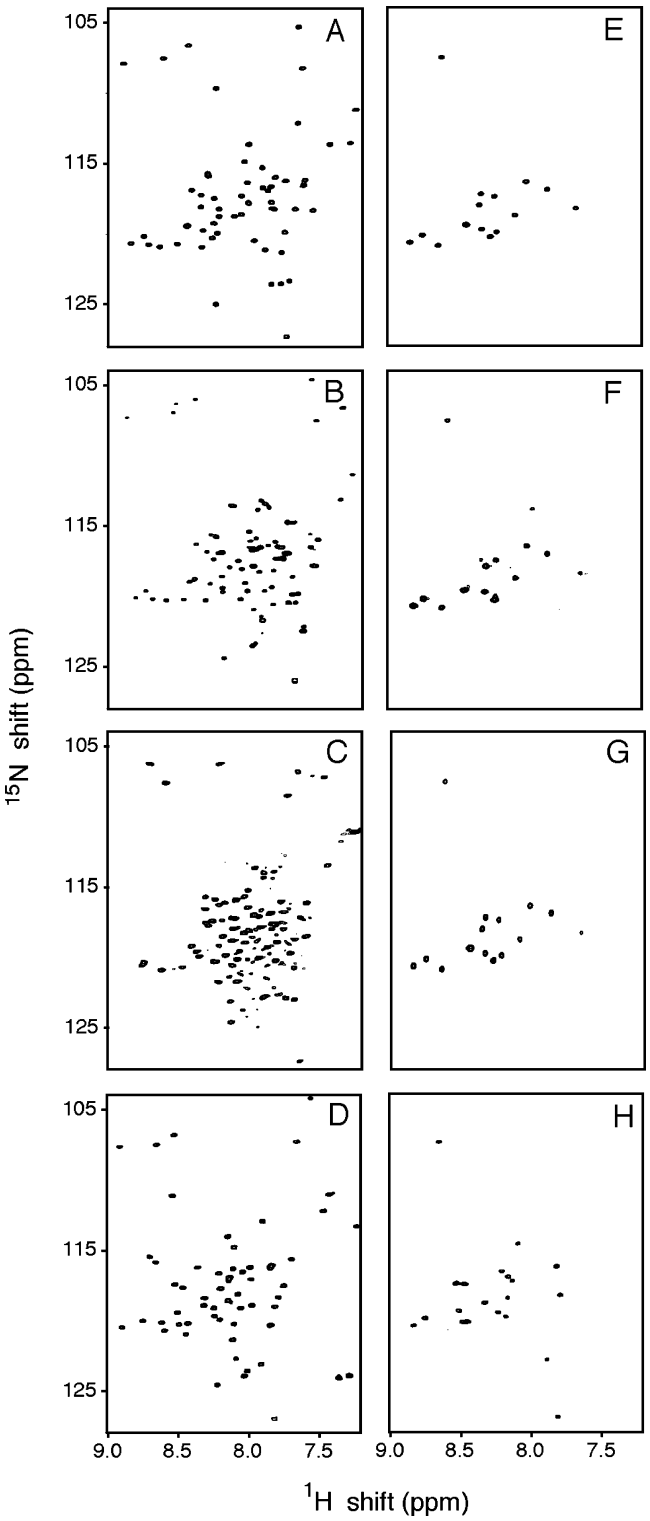

FIGURE 1: Two-dimensional ^{1}H/^{15}N HSQC spectra of uniformly ^{15}N-labeled samples of polypeptides with sequences derived from MerF. (A) $MerF_t$ in micelles in H_2O. (B) $MerF_m$ in micelles in H_2O. (C) Wild-type MerF in micelles in H_2O. (D) $MerF_t$ in isotropic bicelles in H_2O. (E) $MerF_t$ in micelles in D_2O. (F) $MerF_m$ in micelles in D_2O. (G) Wild-type MerF in micelles in $^{2}H_2O$. (H) $MerF_t$ in isotropic bicelles in $^{2}H_2O$. The micelle samples contained 500 mM SDS, the bicelle sample contained 7.5% DHPC/DMPC ($q = 0.25$), and the protein concentrations were 0.7 mM. The temperature for micelle samples was 60 °C; the temperature for bicelle samples was 42 °C ($td_1 = 1024$, $td_2 = 256$, $d_1 = 1.5$ s).

different micelle-forming lipids over a range of protein, lipid, and salt concentrations, lipid to protein ratios, and temperatures. The best-resolved and most reproducible spectra were obtained in SDS micelles at 60 °C. These and other experimental conditions are typical of those we have previously used in solution NMR studies of helical membrane proteins in micelles (*9*, *17*, *33*, *37*, *72*). Under these conditions, the wild-type protein binds mercury, and as demonstrated by the NMR data all three constructs adopt a unique folded structure with no evidence of internal mobility, except for residues near the N- and C-termini. NMR spectra of the truncated 60-residue $MerF_t$ (Figure 1A) and the full-length 80-residue $MerF_m$ (Figure 1B) and MerF (Figure 1C) polypeptides have many similar features. However, the additional 20 resonances present in the spectra of the full-length polypeptides overlap in the central region because they are from mobile and unstructured residues near the N- and C-termini and have similar ^{1}H chemical shift frequencies. Significantly, nearly identical spectra were obtained from all three proteins in $^{2}H_2O$ solutions (Figure 1E−G) where the observable resonances are from residues in the transmembrane helices. The quantitative comparisons of the ^{1}H and ^{15}N chemical shift frequencies for the resonances of $MerF_t$ and $MerF_m$ in Figure 2C demonstrate that residues 26−66, which constitute the helix−loop−helix core domain of MerF, are not substantially affected by the presence of the residues at the N- and C-termini. There are significant chemical shift differences only near the truncation sites that form the N- and C-termini of $MerF_t$.

In the context of membrane protein structure determination, another highly relevant comparison is between the spectra of $MerF_m$ in micelles (Figure 1A) and in small isotropic bicelles (Figure 1D). Although the ^{1}H resonance line widths observed in the bicelle sample are noticeably broader (20 vs 10 Hz), reflecting the slower overall reorientation of the protein when associated with bicelle-forming lipids, there are striking similarities between the patterns of resonances. This is especially true for the spectra obtained in $^{2}H_2O$ solutions (Figure 1E vs Figure 1H). These data indicate that the helix−loop−helix core domain of MerF has similar properties in both micelle and bicelle samples. Solid-state NMR results (*73*) on samples of the same polypeptides in large aligned bicelles indicate that the core domain also has similar properties in planar phospholipid bilayers.

Further evidence that residues 26−66 have similar properties in the truncated $MerF_t$ and full-length $MerF_m$ constructs comes from comparisons of their dynamics. Relative resonance intensities (peak heights) are readily measured surrogates for line widths in HSQC spectra and reflect local backbone dynamics, since large-amplitude backbone motions that are rapid compared to the overall reorientation rate of the protein reduce the line widths and increase the intensities of the resonances. The plots of peak intensities as a function of residue number are similar for the core helix−loop−helix residues in both constructs (Figure 2D,E). Significantly, the only resonances with substantially greater intensities (narrower line widths) are from residues in the N- and C-terminal regions of the polypeptides. The heteronuclear ^{1}H−^{15}N NOE measurements, which are highly reliable indicators of local backbone motions in helical membrane proteins in micelles (*74*), confirm that residues in the helices and the interhelical loop have the same rotational correlation time and that additional backbone motions are present only in residues near the N- and C-termini (Figure 2H).

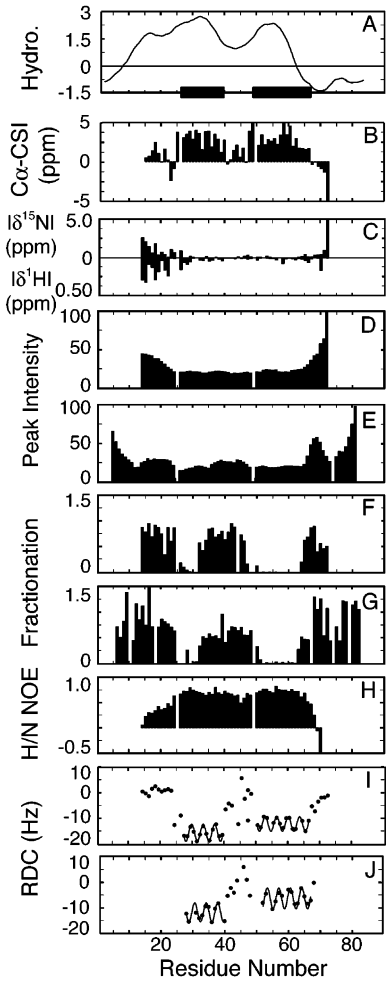

FIGURE 2: (A) Kyte−Doolittle hydropathy plot of MerF. (B) Cα chemical shift index of $MerF_t$. (C) Comparison of the chemical shifts of the two constructs reveals that the only differences are at the N- and C-termini, with the most significant ones being at the N-terminus for residues 13−26. (D, E) Peak intensities compared for these constructs show that the differences at the N-terminus are related to the change in dynamics from one construct to the other. (F, G) Hydrogen/deuterium fractionation factors identify the two α-helical regions of the protein that are resistant to exchange with the solvent in $MerF_t$ and $MerF_m$, respectively. (H) The values of the $^1H/^{15}N$ NOEs show that the N- and C-termini of the protein are dynamic relative to the helices while the interhelical loop is not. (I, J) Values of the $^1H-^{15}N$ RDCs for both $MerF_t$ and $MerF_m$ show the characteristic α-helical periodicity between residues 27 and 39 and between residues 50 and 66, with nonzero values in the interhelical loop.

Deuterium/hydrogen fractionation is an effective method for monitoring hydrogen exchange at amide sites in membrane proteins (*52*), because it can provide valid comparisons over the wide range of exchange rates encountered for residues in transmembrane helices and exposed loops, turns, and terminal regions. The fractionation values determined for $MerF_t$ and $MerF_m$ are plotted as a function of residue number in panels F and G of Figure 2, respectively. The residues in the two helical segments are readily identified in these plots because they undergo negligible hydrogen exchange under these conditions; in contrast, the loop and terminal regions undergo substantial exchange.

Not only do the $^1H-^{15}N$ backbone RDCs provide a robust index of secondary structure and topology, but they are also the source of orientation constraints as input for calculations of three-dimensional structures. The regular secondary structure of the protein is mapped onto the spectroscopic measurements through the dipolar waves. The signature periodicity of 3.6 of the sine waves fit to the data identifies the residues in a α-helix and provides a fine delineation of the boundaries of the helix based on periodicity. In contrast, a traditional hydropathy plot (Figure 2A) provides only a coarse prediction of the transmembrane helices because the helical boundaries are strongly influenced by the choice of threshold and the size of the window (*75*). Moreover, the shape of the plot for the first transmembrane helix is distorted for MerF. In contrast, dipolar wave fitting to the experimental RDCs (Figure 2I,J) identifies the boundaries of the helical segments within one residue. Additionally, the dipolar waves show that both of the helices in MerF are straight and nearly ideal, and their magnitudes and average values can be used to determine the relative alignment of the helices (*34*−*36*).

Through fitting of a sinusoid to the measured residual dipolar couplings as a function of residue number, the orientations of the structural elements can be parametrized. Importantly, rather than a trace of infinite possible orientations for a single dipolar coupling, the degeneracy of an ensemble of couplings for each individual structural element is reduced to four orientations. Thus, for a polypeptide with two helices, there is an expected 16-fold degeneracy for the absolute orientations. By focusing on the relative orientations of the structural elements, the 16 degenerate solutions are reduced to four nonsuperimposable orientations of the two helices as shown in Figure 3A−D. A further reduction in the number of orientations consistent with the experimental data can be achieved by comparing the results obtained from samples where the proteins have different alignments. This was accomplished by comparing the RDCs for the residues in the core domains of $MerF_t$ and $MerF_m$. Since the resonances from these residues have the same isotropic chemical shifts (Figure 2C), dynamics (Figure 2D,E), and secondary structure (Figures 2F,G and 2I,J) in both constructs, the differences in the measured residual dipolar couplings reflect only variations in the alignments of the proteins in the samples. The angles between the helices noted in Figure 3 are all consistent with the experimental data. The models of helices shown in Figure 3D,H are the only pair that is the same within experimental error, are compatible with both helices passing through the bilayer, are consistent with the selective broadening of resonances preceding the first helix by paramagnetic reagents (*76*−*78*) covalently linked to Cys 71 after the second helix, and are in agreement with the results of solid-state NMR experiments on bicelle samples (*73*) and the protein structure calculated (Figures 3I and 4) using X-PLOR-NIH. This demonstrates the consistency of the analysis of the data and the unique determination of helix topology.

The structure calculations of the helix−loop−helix core domain (residues 27−66) based on the RDCs converge to an ensemble of structures with a backbone RMSD of 1.56 Å. To facilitate the formation of the protein fold, a radius of gyration potential was applied to the helical segments across a wide range of values and was optimized in order to minimize the tendency to distort geometry and violate experimental constraints. Included in the Supporting Information is a summary of the R_{gyr} array and statistics of the observed populations. The inclusion of the TALOS-derived

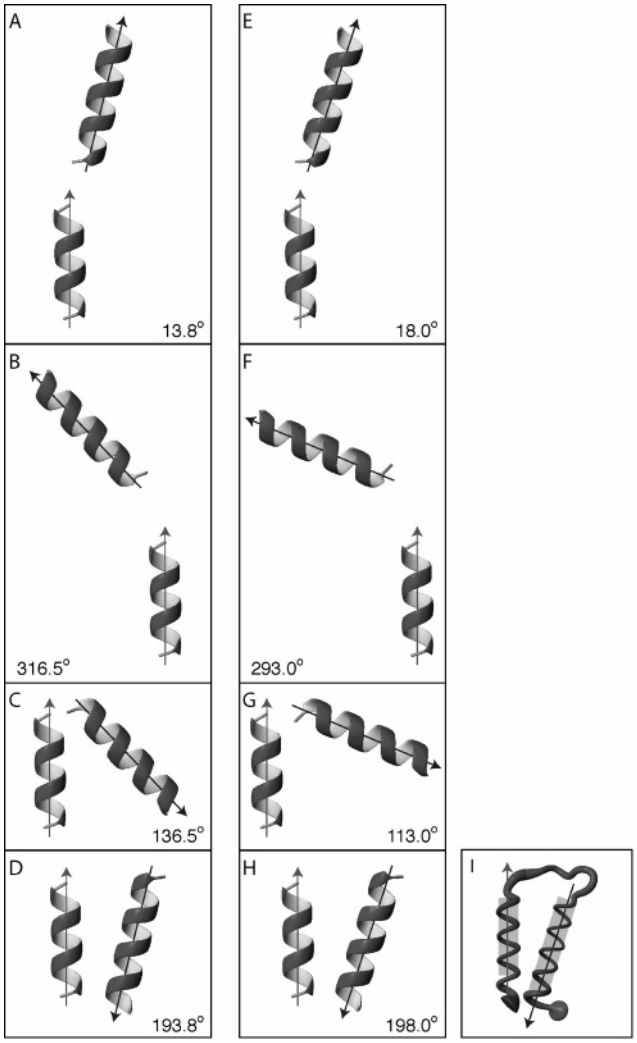

FIGURE 3: Four-fold orientational degeneracy arising from dipolar wave fitting of residual dipolar couplings for $MerF_t$ (left column) and $MerF_m$ (right column). (I) Structure of $MerF_t$.

dihedral constraints in the calculations does not significantly alter the average structure of the protein; however, they do effect a reduction of the backbone RMSD of the ensemble to 0.58 Å. In the final structure calculations, 15 out of 100 possible structures were accepted, and the ensemble of structures has an RMSD of less than 1.05 Hz to the experimental RDCs. The ensemble of accepted structures is shown as a variable radius spline in Figure 4A.

DISCUSSION

Many soluble, globular proteins have large amounts of helical secondary structure; for example, hemoglobin is about 70% α-helix. Thus, the most distinctive feature of a helical membrane protein is the requirement of lipids for it to adopt a stable, functional conformation. The lipid composition of biological membranes is diverse, and many natural and synthetic lipids are available from commercial suppliers. For structural studies of membrane proteins, the most significant types of lipid assemblies are micelles, small isotropic bicelles, large magnetically aligned bicelles, and bilayers (*1*). Large bicelles and bilayers are the most attractive systems for structure determination because the proteins reside in a planar, liquid-crystalline, phospholipid environment with chemical and topological properties essentially identical to

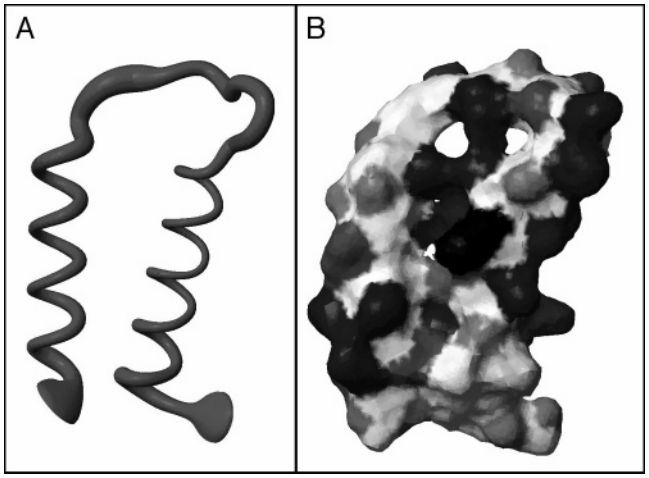

FIGURE 4: (A) Spline representation of an ensemble of structures calculated using the NMR data (residues 24−69). Structure of the helix−loop−helix core of MerF in micelles. The family of the structure has an RMSD of 0.58. (B) Space-filling surface displaying the hydrophobicity based upon the whole-residue partitioning coefficients for the structure in (A) transitioning from black (negative coefficient, hydrophobic) to white (positive coefficient, hydrophilic).

those found in biological membranes; however, they can only be studied by solid-state NMR spectroscopy (*1*, *73*). Micelles and small bicelles are complementary experimental systems that reorient rapidly enough to be studied by solution NMR spectroscopy. The genuine differences found in the structures and dynamics of membrane proteins in various lipid assemblies (*34*) enable the effects of lipids on protein structure and function to be examined. As a practical matter, difficulties resulting from aggregation or solubility are often encountered in the preparation of samples of membrane proteins, and flexibility in the choices of lipids and conditions is needed to overcome these problems. The experimental NMR spectra in Figure 1 demonstrate that lipid and polypeptide components can be selected and manipulated as part of the sample optimization process to yield well-resolved HSQC spectra. The ranges of ^1H and ^{15}N isotropic chemical shift frequencies and the line widths and shapes reflect the properties of native, folded proteins. The chemical shift frequencies observed in HSQC spectra are highly sensitive to local structure, and even small conformational changes can result in substantial variations of the chemical shifts. Therefore, the spectra in Figure 1 demonstrate that the helix−loop−helix core domain of MerF has the same structure in truncated and full-length constructs and in micelles and small bicelles.

The results summarized in Figure 2 provide a more detailed view of the structure and dynamics of MerF in micelles. By itself, the hydropathy plot in Figure 2A is somewhat ambiguous, since it can be viewed as being consistent with MerF having two or possibly three transmembrane helices, as well as a substantial range of lengths for the first transmembrane helix, and this contributed to differences among earlier models of the protein (*34*, *45*, *49*). Both the Cα chemical shift index (*79*) and, more definitively, the dipolar wave fits to the $^1H-^{15}N$ RDCs support the definition of the two α-helical segments of MerF containing residues 27−39 and 50−66. Further support for the N-terminal boundaries of the helices comes from the presence of proline residues at positions 26 and 49. The fractionation

factors indicate that residues in these two regions of the protein are highly resistant to hydrogen exchange with solvent water; indeed, as shown in Figure 1, amide resonances from these residues remain in samples that have been exposed to high levels of ^2H_2O for many weeks, which provides strong qualitative support for the locations and lengths of the helices. The dipolar waves for both the truncated (Figure 2I) and full-length (Figure 2J) polypeptides show the 3.6 residues per turn periodicity characteristic of an α-helix for these same residues. Other fitting parameters indicate that both helices are straight and nearly ideal (*36*).

Both the resonance intensities (line widths) and $^1H-^{15}N$ heteronuclear NOEs show that there are local backbone motions at the C-terminus of the truncated protein (Figures 2D and H) and both the N- and C-termini of the full-length protein (Figure 2E) in micelles. Significantly, none of the residues in the helix−loop−helix core domain of MerF shows evidence of internal motions in micelles. Therefore, residues 40−49 that form the interhelical loop adopt a single stable conformation on the time scales of the solution NMR experiments. This conclusion is reinforced by the large magnitudes and irregular variation in the amplitudes of the RDCs for these residues (Figure 2I,J), especially when compared to the near-zero RDCs observed for the residues near the termini of the proteins in the same samples.

Approximately 10 residues at the N-terminus and 6 residues at the C-terminus of the truncated $MerF_t$ are unstructured and flexible, as indicated by their relative intensities (line widths), $^1H/^2H$ fractionation factors, $^1H-^{15}N$ heteronuclear NOEs, and the observation of near-zero values for their RDCs. The corresponding residues show little additional evidence of structure in the full-length $MerF_m$ polypeptide. However, these N- and C-terminal residues adopt unique rigid conformations in bilayer environments, since solid-state NMR spectra of these proteins show little or no evidence of isotropic resonance intensity indicative of internal motions (*73*). The dynamics of residues in terminal and loop regions is emerging as one of the largest areas of differences between membrane proteins in micelle and bilayer environments (*34*); it is a principal reason why the demonstration that the residues in the interhelical loop and the helices do not display evidence of internal dynamics and adopt unique structures by the data in Figure 2 is crucial for the structure determination of $MerF_t$ in micelles.

The whole-residue hydrophobicity based upon partitioning constants (*80*) is displayed on the surface of an accepted protein structure in Figure 4B. The gray scale gradient reflects the variation from hydrophobic (black) to hydrophilic side chains (white). The impact of the relatively large hydrophobic core on the protein structure is evident, both from the prevalence of hydrophobic residues in the helical segments and from their exclusion from the terminal segments. Additional views of the structure are contained in the Supporting Information.

The structure determined for the central helix−loop−helix domain of MerF demonstrates that converged structures can be calculated with remarkably high structural precision from sparse NMR data, in this example, principally two sets of $^1H-^{15}N$ residual dipolar couplings from samples with nondegenerate alignments. Two additional terms were applied in the calculations to facilitate the convergence to a final family of protein structures. The predictions derived from chemical shift analysis made by the TALOS program have been shown to be reliable in many cases; however, the finite database being searched in the prediction as well as the composition of the databases (typically small, soluble proteins) raises concerns about the application of TALOS predictions to a membrane protein. Therefore, the TALOS-derived dihedrals were applied with a significantly smaller force constant than the dipolar wave-derived dihedrals to the helical segments (30 vs 300 kcal deg^{-1} K^{-1}) and without being evaluated in the acceptance criteria of the structures. In the accepted protein structures, the dihedrals display significant variance from the initial TALOS predictions. The second term applied to aid in the convergence of the MerF structures was a radius of gyration potential. DEFINE STRUCTURE analysis of the R_{gyr} optimization failed to associate the helices for the second population of acceptable structures at $R_{gyr} = 15.5$ Å, and the closest helical contact (based upon the Cα distance matrix) for the $R_{gyr} = 15.5$ Å population was ∼20 Å. The ensemble of structures generated with this R_{gyr}, while satisfying the experimental restraints within the acceptance criteria, exhibited a general lack of convergence with respect to relative helical tilts and phases and gave the largest deviation from the relative helical tilts compared to the dipolar wave analysis of the helical segments. The results of the first population at $R_{gyr} = 9.43$ Å gave an axial helical separation of 7.1 ± 1.1 Å. The analysis of structures calculated at $R_{gyr} = 9.5$ Å also gave a mean axial distribution of 7.1 Å as did other members of this subpopulation of R_{gyr} values that yielded accepted structures, which suggests that the calculation is relatively insensitive to minor variation in the R_{gyr} term. Previous analyses of the crystal structures of membrane proteins have determined two major distributions of axial separation, one centered at 7.3 Å and another at 10.8 Å (*71*). Thus, the structure of MerFt determined by NMR has an axial separation between helices that is consistent with previously determined structures of membrane proteins.

On the basis of the periodicity of the α-helix, $MerF_t$ has two helical segments, encompassing residues 27−39 (13 residues) and residues 50−66 (17 residues), in micelles. Both helical segments of MerF have proline residues, which are associated with helix distortions (*81*), at their N-terminal ends. Since prolines are imino acids, they do not have a backbone amide site for the measurement of $^1H-^{15}N$ dipolar couplings. However, in situations where a proline is located in the middle of a helix, dipolar waves corresponding to distinct helical segments can be fit to residues on both sides, generally exhibiting differences in the observed phasing or amplitude of the dipolar oscillations that are characteristic of a kink (*36*). Attempts to extend the dipolar wave fittings beyond the proline residues in both helices of $MerF_t$, in particular at T24, V26, and L48, found evidence of substantial distortions. Although some evidence of periodicity could be found in the segment before helix 1, residues G14−F23 have dynamic properties, fractionation factors, and near-zero residual dipolar couplings that are inconsistent with the presence of a stable helix. A short stretch of residues before helix 2, Y45−L47 or V46−L48, also displays some evidence of periodicity; however, such small stretches are generally artifacts resulting from overparametrization, and we do not consider them to be part of the helix as defined by fits to dipolar waves.

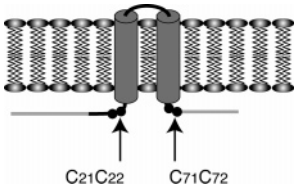

FIGURE 5: Model of MerF in a membrane bilayer based on the structure of MerFt in micelles shown in Figure 4. The locations of the two pairs of vicinal cysteines on the cytoplasmic side of the membrane are identified with arrows.

The program DEFINE STRUCTURE generally finds somewhat longer transmembrane helices. The discrepancies between the helical lengths predicted by dipolar wave fitting and DEFINE STRUCTURE arise from the high sensitivity of the wave fits to deviations in the periodicity of helices (*36*). Helices with between 13 and 36 residues are found in the other membrane proteins with their structures determined (*82*), and the average length of helices that appear to span the hydrophobic portion of bilayers is 17.3 residues (*83*). The analysis of the $MerF_t$ structure with DEFINE STRUCTURE assigns lengths of 16 residues, spanning 26.94 Å, and 18 residues, spanning 27.42 Å, to the two helices. The distances are based upon the Cα matrix. Using either the more conservative values determined by fitting dipolar waves to the RDCs (13 residues and 17 residues) or those from a structural analysis (16 residues and 18 residues), the lengths of the helices of $MerF_t$ are typical of those found in the structures of other membrane proteins. The structure in Figure 4 and the model in Figure 5 show the two helices spanning the bilayer. Solid-state NMR spectra of $MerF_t$ in aligned bicelles demonstrate that the two helices have similar tilt angles in bilayers of dimyristoylphosphatidylcholine (*73*).

The selective reintroduction of the cysteine residues, the determination of the structures of more members of this family, and the extension of structural studies to the side chains should provide an opportunity to describe the structural basis for the transport function of MerF and its role in the detoxification of mercury by bacteria. Little is known about the oligomeric state of these proteins *in vivo* or *in vitro*, and as this information becomes available, it too will influence the understanding of the structural biology of the helical segments and their roles in the transport of mercury across the membrane.

Genetic studies on a related protein, MerT, have shown that the two pairs of vicinal cysteine residues play key roles in the transport process (*84*). The first set of vicinal cysteines has been shown to be essential for functionality; the residues corresponding to this first cysteine pair some evidence of minor truncation effects between the two constructs tested, as indicated by differences in chemical shifts, RDCs, and dynamic properties for these residues in $MerF_t$ and $MerF_m$. The second set of vicinal cysteines appears to be located in a region of the protein that undergoes additional backbone motions, which may have functional significance. The structure of MerF in micelles determined by NMR spectroscopy shows that both pairs of cysteine residues are located on the cytoplasmic side of the membrane. The structural model shown in Figure 5 is distinctly different than that derived principally from the hydrophobicity plot that located the first pair near the periplasm with only the second pair exposed to the cytoplasm (*45*). Thus, the structure of the helix−loop−helix core domain of MerF illustrated in Figure 4 not only demonstrates that NMR can be used to determine the three-dimensional structures of polytopic membrane proteins but also sets the stage for studying the structural biology of heavy metal transport across membranes.

ACKNOWLEDGMENT

We thank Robert Koneckny as a part of the W. M. Keck Laboratory for Integrated Biology Computer Facility at UCSD for computational assistance.

SUPPORTING INFORMATION AVAILABLE

An example spectrum with assignments is given in Figure S1. Correlation plots of $^1H-^{15}N$ RDC measurements for degenerate alignments of $MerF_t$ in compressed and stretched polyacrylamide gels and nondegenerate alignments of $MerF_t$ and $MerF_m$ are given in Figure S2. Correlation plots of experimental $^1H-^{15}N$ RDC values to back-calculated $^1H-^{15}N$ RDC values from each of the proposed models for $MerF_t$ and $MerF_m$ are shown in Figure S3. Additional views of the hydrophobicity surface given in Figure 4C are included in Figure S4 in color scale to illustrate the interface between the helices and indicate the sites of charged residues. Figure S5 gives alternative representations of the determined structures, including the conventional overlay of the full ensemble of structures. Tables contain the experimental data of residue-specific ^1H, ^{15}N, and $^{13}C_\alpha$ chemical shift assignments (Table 1) and $^1H-^{15}N$ RDC (Table 2) and additional information on the structure calculation including general model statistics (Table 3), values of the derived dihedrals from the dipolar wave fitting and TALOS prediction (Tables 4 and 5, respectively), experimental and simulated RDCs (Table 6), radius of gyration convergence statistics (Table 7), and DEFINE STRUCTURE model analysis (Table 8). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Opella, S. J., and Marassi, F. M. (2004) Structure determination of membrane proteins by NMR spectroscopy, *Chem. Rev. 104*, 3587−3606.
2. Ketchem, R. R., Hu, W., and Cross, T. A. (1993) High-resolution conformation of gramicidin a in a lipid bilayer by solid-state NMR, *Science 261*, 1457−1460.
3. Opella, S. J., Marassi, F. M., Gesell, J. J., Valente, A. P., Kim, Y., Oblatt-Montal, M., and Montal, M. (1999) Structures of the m2 channel-lining segments from nicotinic acetylcholine and nmda receptors by NMR spectroscopy, *Nat. Struct. Biol. 6*, 374−379.
4. Wang, J., Kim, S., Kovacs, F., and Cross, T. A. (2001) Structure of the transmembrane region of the m2 protein h(+) channel, *Protein Sci. 10*, 2241−2250.
5. Valentine, K. G., Liu, S. F., Marassi, F. M., Veglia, G., Opella, S. J., Ding, F. X., Wang, S. H., Arshava, B., Becker, J. M., and Naider, F. (2001) Structure and topology of a peptide segment of the 6th transmembrane domain of the *Saccharomyces cerevisae* alpha-factor receptor in phospholipid bilayers, *Biopolymers 59*, 243−256.
6. Marassi, F. M., and Opella, S. J. (2003) Simultaneous assignment and structure determination of a membrane protein from NMR orientational restraints, *Protein Sci. 12*, 403−411.
7. Park, S. H., Mrse, A. A., Nevzorov, A. A., Mesleh, M. F., Oblatt-Montal, M., Montal, M., and Opella, S. J. (2003) Three-dimensional structure of the channel-forming trans-membrane domain of virus protein "u" (vpu) from hiv-1, *J. Mol. Biol. 333*, 409−424.

8. Williams, K. A., Farrow, N. A., Deber, C. M., and Kay, L. E. (1996) Structure and dynamics of bacteriophage ike major coat protein in mpg micelles by solution NMR, *Biochemistry 35*, 5145−5157.
9. Almeida, F. C., and Opella, S. J. (1997) Fd coat protein structure in membrane environments: Structural dynamics of the loop between the hydrophobic trans-membrane helix and the amphipathic in-plane helix, *J. Mol. Biol. 270*, 481−495.
10. MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) A transmembrane helix dimer: Structure and implications, *Science 276*, 131−133.
11. Papavoine, C. H., Christiaans, B. E., Folmer, R. H., Konings, R. N., and Hilbers, C. W. (1998) Solution structure of the m13 major coat protein in detergent micelles: A basis for a model of phage assembly involving specific residues, *J. Mol. Biol. 282*, 401−419.
12. Townsley, L. E., Tucker, W. A., Sham, S., and Hinton, J. F. (2001) Structures of gramicidins a, b, and c incorporated into sodium dodecyl sulfate micelles, *Biochemistry 40*, 11676−11686.
13. Sorgen, P. L., Cahill, S. M., Krueger-Koplin, R. D., Krueger-Koplin, S. T., Schenck, C. C., and Girvin, M. E. (2002) Structure of the *Rhodobacter sphaeroides* light-harvesting 1 beta subunit in detergent micelles, *Biochemistry 41*, 31−41.
14. Zamoon, J., Mascioni, A., Thomas, D. D., and Veglia, G. (2003) NMR solution structure and topological orientation of monomeric phospholamban in dodecylphosphocholine micelles, *Biophys. J. 85*, 2589−2598.
15. Yushmanov, V. E., Mandal, P. K., Liu, Z., Tang, P., and Xu, Y. (2003) NMR structure and backbone dynamics of the extended second transmembrane domain of the human neuronal glycine receptor alpha1 subunit, *Biochemistry 42*, 3989−3995.
16. Yushmanov, V. E., Xu, Y., Tang, P., Mandal, P. K., and Liu, Z. (2003) NMR structure and dynamics of the second transmembrane domain of the neuronal acetylcholine receptor beta 2 subunit, *Biochemistry 42*, 13058−13065.
17. Gesell, J., Zasloff, M., and Opella, S. J. (1997) Two-dimensional ^{1}H NMR experiments show that the 23-residue magainin antibiotic peptide is an alpha-helix in dodecylphosphocholine micelles, sodium dodecyl sulfate micelles, and trifluoroethanol/water solution, *J. Biomol. NMR 9*, 127−135.
18. Chou, J. J., Kaufman, J. D., Stahl, S. J., Wingfield, P. T., and Bax, A. (2002) Micelle-induced curvature in a water-insoluble hiv-1 env peptide revealed by NMR dipolar coupling measurement in stretched polyacrylamide gel, *J. Am. Chem. Soc. 124*, 2450−2451.
19. Okon, M., Frank, P. G., Marcel, Y. L., and Cushley, R. J. (2002) Heteronuclear NMR studies of human serum apolipoprotein a-i. Part i. Secondary structure in lipid-mimetic solution, *FEBS Lett. 517* (1−3), 139−143.
20. Oxenoid, K., Kim, H. J., Jacob, J., Sonnichsen, F. D., and Sanders, C. R. (2004) NMR assignments for a helical 40 kDa membrane protein, *J. Am. Chem. Soc. 126*, 5048−5049.
21. Schubert, M., Kolbe, M., Kessler, B., Oesterhelt, D., and Schmieder, (2002) Heteronuclear multidimensional NMR spectroscopy of solubilized membrane proteins: Resonance assignment of native bacteriorhodopsin, *ChemBioChem 3*, 1019−1023.
22. Klein-Seetharaman, J., Reeves, P. J., Loewen, M. C., Getmanova, E. V., Chung, J., Schwalbe, H., Wright, P. E., and Khorana, H. G. (2002) Solution NMR spectroscopy of [alpha-^{15}N]lysine-labeled rhodopsin: The single peak observed in both conventional and trosy-type hsqc spectra is ascribed to lys-339 in the carboxyl-terminal peptide sequence, *Proc. Natl. Acad. Sci. U.S.A. 99*, 3452−3457.
23. Klein-Seetharaman, J., Yanamala, N. V., Javeed, F., Reeves, P. J., Getmanova, E. V., Loewen, M. C., Schwalbe, H., and Khorana, H. G. (2004) Differential dynamics in the G protein-coupled receptor rhodopsin revealed by solution NMR, *Proc. Natl. Acad. Sci. U.S.A. 101*, 3409−3413 (Epub 2004 Feb 3427).
24. Norwood, T. J., Crawford, D. A., Steventon, M. E., Driscoll, P. C., and Campbell, I. D. (1992) Heteronuclear ^{1}H−^{15}N nuclear magnetic resonance studies of the c subunit of the *Escherichia coli* F_1F_0 ATP synthase: Assignment and secondary structure, *Biochemistry 31*, 6285−6290.
25. Girvin, M. E., Rastogi, V. K., Abildgaard, F., Markley, J. L., and Fillingame, R. H. (1998) Solution structure of the transmembrane H^+-transporting subunit c of the F_1F_0 ATP synthase, *Biochemistry 37*, 8817−8824.
26. Schwaiger, M., Lebendiker, M., Yerushalmi, H., Coles, M., Groger, A., Schwarz, C., Schuldiner, S., and Kessler, H. (1998) NMR investigation of the multidrug transporter emre, an integral membrane protein, *Eur. J. Biochem. 254*, 610−619.
27. Dmitriev, O. Y., Altendorf, K., and Fillingame, R. H. (2004) Subunit a of the *E. coli* ATP synthase: Reconstitution and high-resolution NMR with protein purified in a mixed polarity solvent, *FEBS Lett. 556*, 35−38.
28. Matthey, U., Kaim, G., Braun, D., Wuthrich, K., and Dimroth (1999) NMR studies of subunit c of the ATP synthase from *Propionigenium modestum* in dodecylsulphate micelles, *Eur. J. Biochem. 261*, 459−467.
29. Vinogradova, O., Sonnichsen, F., and Sanders, C. R., II (1998) On choosing a detergent for solution NMR studies of membrane proteins, *J. Biomol. NMR 11*, 381−386.
30. Krueger-Koplin, R. D., Sorgen, P. L., Krueger-Koplin, S. T., Rivera-Torres, I. O., Cahill, S. M., Hicks, D. B., Grinius, L., Krulwich, T. A., and Girvin, M. E. (2004) An evaluation of detergents for NMR structural studies of membrane proteins, *J. Biomol. NMR 28*, 43−57.
31. Cross, T. A., and Opella, S. J. (1979) NMR of fd coat protein, *J. Supramol. Struct. 11*, 139−145.
32. Lauterwein, J., Brown, L. R., and Wuthrich, K. (1980) High-resolution ^{1}H NMR studies of monomeric melittin in aqueous solution, *Biochim. Biophys. Acta 622*, 219−230.
33. Opella, S. J. (1997) NMR and membrane proteins, *Nat. Struct. Biol. 4* (Suppl.), 845−848.
34. Mesleh, M. F., Lee, S., Veglia, G., Thiriot, D. S., Marassi, F. M., and Opella, S. J. (2003) Dipolar waves map the structure and topology of helices in membrane proteins, *J. Am. Chem. Soc. 125*, 8928−8935.
35. Mesleh, M. F., Veglia, G., DeSilva, T. M., Marassi, F. M., and Opella, S. J. (2002) Dipolar waves as NMR maps of protein structure, *J. Am. Chem. Soc. 124*, 4206−4207.
36. Mesleh, M. F., and Opella, S. J. (2003) Dipolar waves as NMR maps of helices in proteins, *J. Magn. Reson. 163*, 288−299.
37. Lee, S., Mesleh, M. F., and Opella, S. J. (2003) Structure and dynamics of a membrane protein in micelles from three solution NMR experiments, *J. Biomol. NMR 26*, 327−334.
38. Wallin, E., and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, *Protein Sci. 7*, 1029−1038.
39. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The X-PLOR-NIH NMR molecular structure determination package, *J. Magn. Reson. 160*, 65−73.
40. Barkay, T., Miller, S. M., and Summers, A. O. (2003) Bacterial mercury resistance from atoms to ecosystems, *FEMS Microbiol. Rev. 27*, 355−384.
41. Brown, N. L., Camakaris, J., Lee, B. T., Williams, T., Morby, A. P., Parkhill, J., and Rouch, D. A. (1991) Bacterial resistances to mercury and copper, *J. Cell. Biochem. 46*, 106−114.
42. Steele, R. A., and Opella, S. J. (1997) Structures of the reduced and mercury-bound forms of merp, the periplasmic protein from the bacterial mercury detoxification system, *Biochemistry 36*, 6885−6895.
43. Lund, P. A., and Brown, N. L. (1987) Role of the mert and merp gene products of transposon tn501 in the induction and expression of resistance to mercuric ions, *Gene 52* (2−3), 207−214.
44. Hobman, J., Kholodii, G., Nikiforov, V., Ritchie, D. A., Strike, P., and Yurieva, O. (1994) The sequence of the mer operon of pmer327/419 and transposon ends of pmer327/419, 330 and 05, *Gene 146*, 73−78.
45. Wilson, J. R., Leang, C., Morby, A. P., Hobman, J. L., and Brown, N. L. (2000) Merf is a mercury transport protein: Different structures but a common mechanism for mercuric ion transporters?, *FEBS Lett. 472*, 78−82.
46. Morby, A. P., Hobman, J. L., and Brown, N. L. (1995) The role of cysteine residues in the transport of mercuric ions by the tn501 mert and merp mercury-resistance proteins, *Mol. Microbiol. 17*, 25−35.
47. Sahlman, L., Hagglof, E. M., and Powlowski, J. (1999) Roles of the four cysteine residues in the function of the integral inner membrane Hg^{2+}-binding protein, merc, *Biochem. Biophys. Res. Commun. 255*, 307−311.
48. Rossy, E., Champier, L., Bersch, B., Brutscher, B., Blackledge, M., and Coves, J. (2004) Biophysical characterization of the merp-like amino-terminal extension of the mercuric reductase from ralstonia metallidurans ch34, *J. Biol. Inorg. Chem. 9*, 49−58.
49. Veglia, G., and Opella, S. J. (2000) Lanthanide ion binding to adventitious sites aligns membrane proteins in micelles for solution NMR spectroscopy, *J. Am. Chem. Soc. 122*, 11733−11734.

50. Ramirez, B. E., and Bax, A. (1998) Modulation of the alignment tensor of macromolecules dissolved in a dilute liquid crystalline medium, *J. Am. Chem. Soc. 120*, 9106−9107.
51. Hus, J. C., Marion, D., and Blackledge, M. (2001) Determination of protein backbone structure using only residual dipolar couplings, *J. Am. Chem. Soc. 123*, 1541−1542.
52. Veglia, G., Zeri, A. C., Ma, C., and Opella, S. J. (2002) Deuterium/hydrogen exchange factors measured by solution nuclear magnetic resonance spectroscopy as indicators of the structure and topology of membrane proteins, *Biophys. J. 82*, 2176−2183.
53. Glover, K. J., Whiles, J. A., Wu, G., Yu, N., Deems, R., Struppe, J. O., Stark, R. E., Komives, E. A., and Vold, R. R. (2001) Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules, *Biophys. J. 81*, 2163−2171.
54. Vold, R. R., Prosser, R. S., and Deese, A. J. (1997) Isotropic solutions of phospholipid bicelles: A new membrane mimetic for high-resolution NMR studies of polypeptides, *J. Biomol. NMR 9*, 329−335.
55. Jones, D. H., and Opella, S. J. (2004) Weak alignment of membrane proteins in stressed polyacrylamide gels, *J. Magn. Reson. 171*, 258−269.
56. Sass, H. J., Musco, G., Stahl, S. J., Wingfield, P. T., and Grzesiek, S. (2000) Solution NMR of proteins within polyacrylamide gels: Diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes, *J. Biomol. NMR 18*, 303−309.
57. Ishii, Y., Markus, M. A., and Tycko, R. (2001) Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel, *J. Biomol. NMR 21*, 141−151.
58. Chou, J. J., Gaemers, S., Howder, B., Louis, J. M., and Bax, A. (2001) A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles, *J. Biomol. NMR 21*, 377−382.
59. Cavanaugh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996) *Protein NMR Spectroscopy Principles and Practice*, Academic Press, New York.
60. Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. (1995) Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (fhsqc) detection scheme that avoids water saturation, *J. Magn. Reson. B 108*, 94−98.
61. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry 33*, 5984−6003.
62. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients, *Prog. Nucl. Magn. Reson. Spectrosc. 34*, 93−158.
63. Shon, K., and Opella, S. J. (1989) Detection of proton homonuclear NOE between amide sites in proteins with proton/nitrogen-15 heteronuclear correlation spectroscopy, *J. Magn. Reson. 82*, 193−197.
64. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on unix pipes, *J. Biomol. NMR 6*, 277−293.
65. Ding, K. Y., and Gronenborn, A. M. (2003) Sensitivity-enhanced 2D IPAP, TROSY-anti-TROSY, and E. Cosy experiments: alternatives for measuring dipolar N-15-H-1(N) couplings, *J. Magn. Reson. 163*, 208−214.
66. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1999) Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration, *J. Am. Chem. Soc. 121*, 2337−2338.
67. Skolnick, J., Kolinski, A., and Ortiz, A. R. (1997) Monsster: A method for folding globular proteins with a small number of distance restraints, *J. Mol. Biol. 265*, 217−241.
68. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
69. Kuszewski, J., and Clore, G. M. (2000) Sources of and solutions to problems in the refinement of protein NMR structures against torsion angle potentials of mean force, *J. Magn. Reson. 146*, 249−254.
70. Richards, F. M., and Kundrot, C. E. (1988) Identification of structural motifs from protein coordinate data: Secondary structure and first level supersecondary structure, *Proteins 3*, 71−84.
71. Eilers, M., Shekar, S. C., Shieh, T., Smith, S. O., and Fleming, P. J. (2000) Internal packing of helical membrane proteins, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5796−5801.
72. McDonnell, P. A., and Opella, S. J. (1993) Effect of detergent concentration on multidimensional solution NMR spectra of membrane proteins in micelles, *J. Magn. Reson. 102*, 120−125.
73. De Angelis, A., Nevzorov, A., Park, S. H., Howell, S. C., Mrse, A. A., and Opella, S. J. (2004) High-resolution NMR spectroscopy of membrane proteins in "aligned" bicelles, *J. Am. Chem. Soc. 126*, 15340−15341.
74. Bogusky, M. J., P., Schiksnis, R. A., Leo, G. C., and Opella, S. J. (1987) Protein backbone dynamics by solid state and solution ^{15}N NMR spectroscopy, *J. Magn. Reson. 72*, 186−190.
75. Moller, S., Croning, M. D., and Apweiler, R. (2001) Evaluation of methods for the prediction of membrane spanning regions, *Bioinformatics 17*, 646−653.
76. Gaponenko, V., Howarth, J. W., Columbus, L., Gasmi-Seabrook, G., Yuan, J., Hubbell, W. L., and Rosevear, P. R. (2000) Protein global fold determination using site-directed spin and isotope labeling, *Protein Sci. 9*, 302−309.
77. Ikegami, T., Verdier, L., Sakhaii, P., Grimme, S., Pescatore, B., Saxena, K., Fiebig, K. M., and Griesinger, C. (2004) Novel techniques for weak alignment of proteins in solution using chemical tags coordinating lanthanide ions, *J. Biol. NMR 29*, 339−349.
78. Pintacuda, G., Moshref, A., Leonchiks, A., Sharipo, A., and Otting, G. (2004) Site-specific labelling with a metal chelator for protein-structure refinement, *J. Biol. NMR 29*, 351−361.
79. Wishart, D. S., and Sykes, B. D. (1994) The ^{13}C chemical-shift index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data, *J. Biomol. NMR 4*, 171−180.
80. Wimley, W. C., and White, S. H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces, *Nat. Struct. Biol. 3*, 842−848.
81. Cordes, F. S., Bright, J. N., and Sansom, M. S. (2002) Proline-induced distortions of transmembrane helices, *J. Mol. Biol. 323*, 951−960.
82. Bowie, J. U. (1997) Helix packing in membrane proteins, *J. Mol. Biol. 272*, 780−789.
83. Hildebrand, P. W., Preissner, R., and Frömmel, C. (2004) Structural features of transmembrane helices, *FEBS Lett. 559*, 145−151.
84. Morby, A. P., Hobman, J. L., and Brown, N. L. (1995) The role of cysteine residues in the transport of mercuric ions by the tn501 mert and merp mercury-resistance proteins, *Mol. Microbiol. 17*, 25−35.

BI048095V