J|A|C|S
ARTICLES
Published on Web 08/24/2006

Structure Determination of a Membrane Protein with Two Trans-membrane Helices in Aligned Phospholipid Bicelles by Solid-State NMR Spectroscopy

Anna A. De Angelis, Stanley C. Howell, Alexander A. Nevzorov, and Stanley J. Opella*

Contribution from the Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, 0307, La Jolla, California 92093-0307

Received June 1, 2006; E-mail: sopella@ucsd.edu

Abstract: The structure of the membrane protein MerFt was determined in magnetically aligned phospholipid bicelles by solid-state NMR spectroscopy. With two trans-membrane helices and a 10-residue inter-helical loop, this truncated construct of the mercury transport membrane protein MerF has sufficient structural complexity to demonstrate the feasibility of determining the structures of polytopic membrane proteins in their native phospholipid bilayer environment under physiological conditions. PISEMA, SAMMY, and other double-resonance experiments were applied to uniformly and selectively ^{15}N-labeled samples to resolve and assign the backbone amide resonances and to measure the associated ^{15}N chemical shift and ^{1}H−^{15}N heteronuclear dipolar coupling frequencies as orientation constraints for structure calculations. ^{1}H/^{13}C/^{15}N triple-resonance experiments were applied to selectively ^{13}C′- and ^{15}N-labeled samples to complete the resonance assignments, especially for residues in the nonhelical regions of the protein. A single resonance is observed for each labeled site in one- and two-dimensional spectra. Therefore, each residue has a unique conformation, and all protein molecules in the sample have the same three-dimensional structure and are oriented identically in planar phospholipid bilayers. Combined with the absence of significant intensity near the isotropic resonance frequency, this demonstrates that the entire protein, including the loop and terminal regions, has a well-defined, stable structure in phospholipid bilayers.

Introduction

Because the native phospholipid bilayer environment for membrane proteins is both asymmetric and liquid crystalline, daunting technical problems are encountered in the preparation of samples suitable for structure determination by conventional methods. On the one hand, the samples are not well enough ordered for X-ray crystallography, and on the other, they are too ordered for solution NMR spectroscopy. As a result, nearly all of the reported structures of membrane proteins[1] (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) have been determined in micelle, cubic phase, or other isotropic chemical environments.[2] Solid-state NMR spectroscopy is beginning to fill this gap in technology, as demonstrated by the deposition into the Protein Data Bank[3] of the structures of several membrane proteins determined in phospholipid bilayers. The initial examples have been relatively small polypeptides with the equivalent of one trans-membrane helix.[4−8] Advances in sample preparation, instrumentation, experimental methods, and calculations enable the structures of larger membrane proteins to be determined with this approach. This is demonstrated in this article by the structure determination of MerFt; with two trans-membrane helices and a 10-residue inter-helical loop, this truncated construct of the mercury transport membrane protein MerF has sufficient structural complexity to demonstrate the feasibility of determining the structures of polytopic membrane proteins in fully hydrated phospholipid bilayers under physiological conditions.

Early experiments demonstrated that solid-state NMR spectroscopy is capable of determining molecular structures with atomic resolution.[9] Protein structure determination by solid-state NMR of aligned samples is based on the narrowing of resonances that can be achieved for immobile molecules with multiple-pulse[10] and double-resonance[11] experiments, the spectroscopic simplifications associated with sample alignment,[12] and the ability to measure orientation constraints for individual atomic sites with separated local field spectroscopy.[13] Solid-state NMR of aligned samples was originally demonstrated with

(1) Raman, P.; Cherezov, V.; Caffrey, M. *Cell Mol. Life Sci.* **2006**, *63*, 36−51.
(2) Rosenbusch, J. P. *J. Struct. Biol.* **2001**, *136*, 144−157.
(3) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(4) Ketchem, R. R.; Hu, W.; Cross, T. A. *Science* **1993**, *261*, 1457−1460.
(5) Opella, S. J.; Marassi, F. M.; Gessel, J. J.; Valente, A. P.; Kim, Y.; Oblatt-Montal, M.; Montal, M. *Nat. Struct. Biol.* **1999**, *6*, 374−379.
(6) Wang, J.; Kim, S.; Kovacs, F.; Cross, T. A. *Protein Sci.* **2001**, *10*, 2241−2250.
(7) Marassi, F. M.; Opella, S. J. *Protein Sci.* **2003**, *12*, 403−411.
(8) Park, S. H.; Mrse, A. A.; Nevzorov, A. A.; Mesleh, M. F.; Oblatt-Montal, M.; Montal, M.; Opella, S. J. *J. Mol. Biol.* **2001**, *333*, 409−424.
(9) Pake, G. E. *J. Chem. Phys.* **1948**, *16*, 327−336.
(10) Waugh, J. S.; Huber, L. M.; Haeberlen, U. *Phys. Rev. Lett.* **1968**, *20*, 180−182.
(11) Pines, A.; Gibby, M. G.; Waugh, J. S. *J. Chem. Phys.* **1973**, *59*, 569−590.
(12) Opella, S. J.; Waugh, J. S. *J. Chem. Phys.* **1977**, *66*, 4919−4924.
(13) Waugh, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1394−1397.

10.1021/ja063640w CCC: $33.50

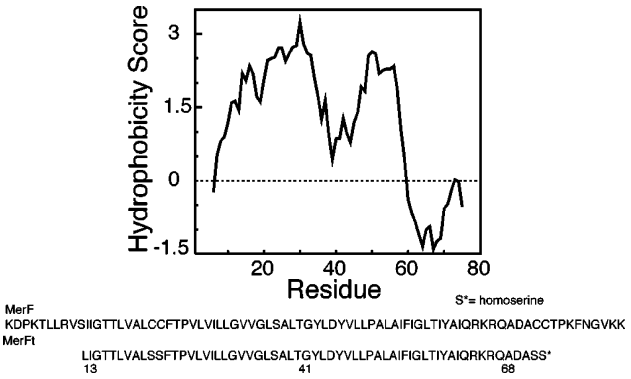

Figure 1. Kyte−Doolittle hydropathy plot[71] and the amino acid sequences[26,29] of MerF and MerFt.

uniaxially oriented polymer fibers[12] and subsequently applied to proteins in magnetically aligned crystals[14] and filamentous bacteriophage particles,[15,16] mechanically aligned phospholipid bilayers,[17] and magnetically aligned bicelles.[18,19] Magic angle spinning solid-state NMR is a complementary approach for structural characterization of membrane proteins.[20−24]

In the bacterial mercury detoxification system,[25] Hg(II) is transported across the inner cell membrane into the cytoplasm by a membrane protein, such as MerF.[26] The amino acid sequence of the truncated construct MerFt used in the experimental studies is compared to that of full-length MerF in Figure 1. The accompanying hydropathy plot suggests that MerF has two hydrophobic trans-membrane helices, although the N-terminal portion of the plot is somewhat atypical. Wild-type MerF has two pairs of vicinal Cys residues (replaced by serines in MerFt), which is unusual, and contrasts with the −CXXC− sequences typically found in proteins that bind mercury and other heavy metals.[27] The vicinal arrangements of the cysteines have profound consequences for their metal-binding properties and the selectivity for Hg(II) over other metals.[28] With both metal transfer and transport activities, the structural biology of MerF is too elaborate to be explained simply by Hg(II) binding to pairs of cysteines. Therefore, it is essential to determine its three-dimensional structure in phospholipid bilayers, where it executes its functions. We previously reported the structure of MerFt in micelles determined by solution NMR[29] and have used this protein in a variety of NMR studies building toward its structure determination in a phospholipid bilayer environment. We have also previously reported the structure of MerP,[30] the periplasmic protein that binds Hg(II) and then transfers it to MerF. With further structural studies of these proteins alone and together, the potential exists to describe the mechanism of mercury detoxification in molecular detail.

Bicelles are nearly ideal samples for experimental studies of membrane proteins[19] since they consist principally of long-chain phospholipids arranged in planar bilayers. An added fraction of short-chain phospholipids interrupts the long-range order and enables them to be aligned by a magnetic field.[31−33] The membrane proteins are immobilized (on the 10^4 Hz time scale of the NMR experiments) by their interactions within the bilayer portion of the bicelles and are aligned along with the phospholipids by the magnetic field, fulfilling the two prerequisites for solid-state NMR of aligned samples. The samples are fully hydrated with about 28% (w/v) lipids in water and are stable for many months under physiological conditions of temperature and pH.

Experimental Methods

Samples for NMR Experiments. We have previously described the expression and purification of isotopically labeled MerFt for NMR studies.[29] Briefly, the gene for the 60-residue sequence of MerFt shown in Figure 1 was inserted into a pET31(+)b vector encoding a 125-residue N-terminal KetoSteroid Isomerase (KSI), which resulted in the formation of inclusion bodies upon expression in bacteria and a C-terminal 6X His-tag, which assisted protein isolation. Expression of the fusion protein was induced after 7 h of bacterial growth by the addition of IPTG. The collected cells were disrupted by sonication; inclusion bodies were isolated by centrifugation; after solubilization in guanidine hydrochloride, the fusion protein was isolated using Ni−NTA affinity chromatography. The polypeptide corresponding to MerFt was cleaved from the fusion protein with CNBr and purified by size-exclusion chromatography in the presence of SDS, which was subsequently removed by exhaustive dialysis. We estimate that the actual yield through the isolation and purification steps is 85% of theoretical. The overall yield of purified protein is ∼15 mg/L of minimal M9 growth media. The precipitated protein was stored as a lyophilized powder prior to reconstitution in lipids.

The two-dimensional HSQC spectrum of uniformly ^{15}N-labeled MerFt in SDS micelles has been assigned,[29] and this enabled the isotopic labeling of each protein sample to be verified by comparison of the relative intensities of resonances of their solution NMR spectra. Samples of MerFt in 500 mM SDS, 10 mM sodium phosphate, and 1 mM sodium azide at pH 5.0 were used for solution NMR spectroscopy.

We have previously described the preparation and properties of the protein-containing bicelle samples used for solid-state NMR spectroscopy.[19,34] Non-hydrolyzable ether-linked analogues of DMPC and DHPC were used to ensure long-term stability. Samples of MerFt in bicelles were prepared by solubilizing the purified polypeptide in an aqueous solution containing the short-chain lipid (DHPC) and then adding this solution to a dispersion of the long-chain lipid (DMPC) in water. The lipids 1,2-*O*-dihexyl-*sn*-glycero-3-phosphocholine (6-O-PC) and 1,2-*O*-ditetradecyl-*sn*-glycero-3-phosphocholine

(14) Oldfield, E.; Rothgeb, T. M. *J. Am. Chem. Soc.* **1980**, *102*, 3635−3637.
(15) Cross, T. A.; Opella, S. J. *J. Am. Chem. Soc.* **1983**, *105*, 306−308.
(16) Zeri, A. C.; Mesleh, M. F.; Nevzorov, A. A.; Opella, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6458−6463.
(17) Nicholson, L. K.; Moll, F.; Mixon, T. E.; LoGrasso, P. V.; Lay, J. C.; Cross, T. A. *Biochemistry* **1987**, *26*, 6621−6626.
(18) Losonczi, J. A.; Tian, F.; Prestegard, J. H. *Biochemistry* **2000**, *39*, 3804−3816.
(19) De Angelis, A. A.; Nevzorov, A. A.; Park, S. H.; Howell, S. C.; Mrse, A. A.; Opella, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 15340−15341.
(20) Eilers, M.; Ying, W.; Reeves, P. J.; Khorana, H. G.; Smith, S. O. *Methods Enzymol.* **2002**, *343*, 212−222.
(21) Rosay, M.; Lansing, J. C.; Haddad, K. C.; Bachovchin, W. W.; Herzfeld, J.; Temkin, R. J.; Griffin, R. G. *J. Am. Chem. Soc.* **2003**, *125*, 13626−13627.
(22) Andronesi, O. C.; Becker, S.; Seidel, K.; Heise, H.; Young, H. S.; Baldus, M. *J. Am. Chem. Soc.* **2005**, *127*, 12965−12974.
(23) Sharpe, S.; Yau, W. M.; Tycko, R. *Biochemistry* **2006**, *45*, 918−933.
(24) Yamamoto, K.; Tuzi, S.; Saito, H.; Kawamura, I.; Naito, A. *Biochim. Biophys. Acta* **2006**, *1758*, 181−189.
(25) Barkay, T.; Miller, S. M.; Summers, A. O. *FEMS Microbiol. Rev.* **2003**, *27*, 355−384.
(26) Wilson, J. R.; Leang, C.; Morby, A. P.; Hobman, J. L.; Brown, N. L. *FEBS Lett.* **2000**, *472*, 78−82.
(27) Opella, S. J.; DeSilva, T. M.; Veglia, G. *Curr. Opin. Chem. Biol.* **2002**, *6*, 217−223.
(28) DeSilva, T. M.; Veglia, G.; Porcelli, F.; Prantner, A. M.; Opella, S. J. *Biopolymers* **2002**, *64*, 189−197.
(29) Howell, S. C.; Mesleh, M. F.; Opella, S. J. *Biochemistry* **2005**, *44*, 5196−5206.
(30) Steele, R. A.; Opella, S. J. *Biochemistry* **1997**, *36*, 6885−6895.
(31) Sanders, C. R.; Hare, B. J.; Howard, K. P.; Prestegard, J. H. *Prog. NMR Spectrosc.* **1994**, *26*, 421−444.
(32) Gaemers, S.; Bax, A. *J. Am. Chem. Soc.* **2001**, *123*, 12343−12352.
(33) Harroun, T. A.; Koslowsky, M.; Nieh, M. P.; de Lannoy, C. F.; Raghunathan, V. A.; Katsaras, J. *Langmuir* **2005**, *21*, 5356−5361.
(34) De Angelis, A. A.; Jones, D. H.; Grant, C. V.; Park, S. H.; Mesleh, M. F.; Opella, S. J. *Methods Enzymol.* **2005**, *394*, 350−382.

(14-O-PC) were obtained in chloroform solutions from Avanti Polar Lipids (www.avantilipids.com). Volumes corresponding to 46.5 mg of 14-O-PC and to 9.5 mg of 6-O-PC were transferred separately into glass tubes.

The labeled protein was dissolved in 1 mL of trifluoroethanol/chloroform 4/1 (v/v) and sonicated for 5 min in a bath sonicator. The organic solvents were evaporated under a stream of nitrogen gas from the protein and lipid samples, which were then placed under high vacuum overnight. An aqueous solution of 6-O-PC, obtained by adding 100 μL of water to the short-chain lipid, was added to the dry protein film. This solubilizes the protein in micelles. A dispersion of 14-O-PC in water was prepared by adding 80 μL of water to the long-chain lipid, followed by extensive vortexing and three freeze/heat cycles (liquid nitrogen/45 °C). The MerFt/6-O-PC micelle solution was added to the 14-O-PC dispersion, previously warmed above 40 °C, vortexed, and subjected to several freeze/heat cycles, and then allowed to equilibrate at room temperature. The opaque dispersion of 14-O-PC becomes transparent upon mixing, indicating that bicelles are formed. The bicelle samples for $q = 3.2$ and a lipid concentration, cL = 28% (w/v), were ~300 mM in 14-O-PC in a 200 μL volume, with pH 5.0, and contained between 4 and 12 mg of MerFt to give final protein concentrations between about 3 and 9 mM. A short, flat-bottomed tube with 5 mm outer diameter (New Era Enterprises (http://newera-spectro.com)) was filled with 160 μL of the bicelle solution, using a precooled glass pipet, at 4 °C. The same volume, which fit entirely within the confines of the solenoid coil, was used for each sample to ensure that susceptibility effects would not influence the observed chemical shift frequencies. The sample tube was sealed with a tight-fitting rubber cap and then pierced with a thin syringe to remove any residual air and create a tight seal. Bicelle samples for ^{2}H NMR measurements were prepared by including a small amount of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine lipid deuterated in the fatty acid chains (DMPC-d_{54} (Avanti Polar Lipids (www.avantilipids.com))), with [DMPC-d_{54}]/[14-O-PC] = 1/50. Parallel bicelles were prepared by adding $YbCl_3 \cdot 6H_2O$ (Sigma (www.sigmaaldrich.com)) directly to the samples in the NMR tube. The tube was cooled to 4 °C, the rubber cap quickly removed, and 5 μL of freshly prepared 100 mM $YbCl_3 \cdot 6H_2O$ in water added to the sample. The ~3 mM lanthanide concentration, which corresponds to ~1 mol % of the long-chain lipid, was chosen to be slightly above the minimum needed to induce parallel bicelle alignment, as verified by ^{31}P NMR. The sample used to obtain the spectrum in Figure 2A was prepared by depositing on glass plates ~3 mg of MerFt and 75 mg of lipids co-dissolved in trifluoroethanol/chloroform solution, drying completely under vacuum, and then hydrating to 95% relative humidity in a chamber saturated with ammonium phosphate (pH 7.0) at 42 °C.[8]

NMR Instrumentation. The double-resonance ^{1}H/^{15}N solid-state NMR experiments were performed on spectrometers with ^{1}H resonance frequencies of 700 and 750 MHz configured with Bruker (www.bruker-.com) Avance consoles and Magnex (www.magnex.com) magnets and room temperature shim coils. The home-built lumped-element probes have 5 mm inner diameter solenoid coils double-tuned to the ^{1}H and ^{15}N resonance frequencies. The bicelle order parameter was measured by ^{2}H NMR on the 700 MHz spectrometer using a single-channel home-built probe with a 5 mm inner diameter solenoid coil tuned to the deuterium resonance frequency of 107.47 MHz.

The triple-resonance ^{1}H/^{13}C/^{15}N solid-state NMR experiments were performed on spectrometers with ^{1}H resonance frequencies of 500 and 750 MHz. The 500 MHz spectrometer was configured with a Varian (www.varianinc.com) Inova console and a wide-bore Magnex 500/89 magnet and room temperature shim coils. A Varian triple-tuned probe (T3 HCN-NS) with a 5 mm inner diameter solenoid coil was used. On the 750 MHz spectrometer, a Bruker triple-tuned probe (PH BIOPE 750SB N/C(P)/H) with a 5 mm inner diameter solenoid coil and Bruker room temperature shim coils were utilized. Magnet drift was compensated for by a linear ramp of current to the Z_0 coil of the room temperature shims since the spectrometers were operated without field/frequency locks.

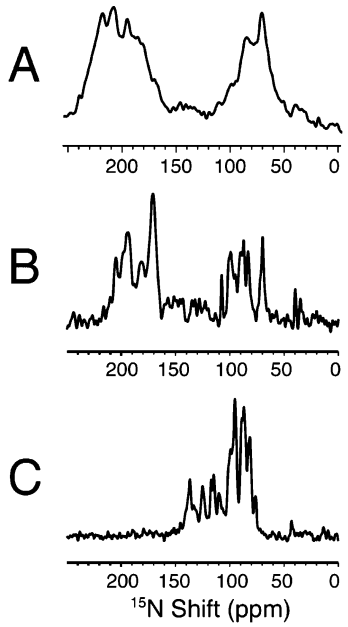

Figure 2. One-dimensional ^{15}N NMR spectra of aligned samples of uniformly ^{15}N-labeled MerFt. (A) Mechanically aligned in 14-O-PC bilayers on glass plates. (B) Magnetically aligned in 14-O-PC/6-O-PC, $q = 3.2$, parallel bicelles. (C) Magnetically aligned in perpendicular bicelles.

NMR Experiments. All samples were equilibrated in the magnetic field at 40 °C for a period >30 min prior to the initiation of NMR measurements. The common parameters for both the one-dimensional ^{1}H−^{15}N cross-polarization (CP) and two-dimensional separated local field experiments (PISEMA[35] and SAMMY[36]) were 1 ms cross-polarization mix time and radio frequency field strengths of 50−55 kHz for ^{1}H irradiation. CP-MOIST[37] cross-polarization was used to compensate for radio frequency power mismatches, and SPINAL-16[38] modulation was used to extend the bandwidth of the ^{1}H irradiation used for heteronuclear decoupling during data acquisition. The recycle delays were typically between 6 and 10 s to avoid sample heating; between 1000 and 8000 scans were co-added with 5 or 10 ms acquisition times in the one-dimensional experiments; between 96 and 512 scans were co-added with a 5 ms acquisition time for the two-dimensional experiments. The two-dimensional PISEMA and SAMMY spectra of uniformly ^{15}N-labeled samples resulted from a total of 128 t_1 increments and 512 t_2 complex points. The spectra of selectively ^{15}N-labeled samples resulted from between 32 and 72 t_1 increments, depending upon the resolution needed for each sample. The same ^{1}H B_1 field strength was utilized during cross-polarization, Lee−Goldburg irradiation, and heteronuclear decoupling during data acquisition. The frequency jumps used to satisfy the Lee−Goldburg off-resonance condition in PISEMA experiments were between ±35.4 and 39.3 kHz. The SAMMY pulse sequence was applied as previously described,[36] with the delays during the t_1 contact time chosen as $\delta_1 = \pi/4$ and $\delta_2 = \pi/2$ to compensate for the finite pulse length, and only a single phase alternation of the ^{15}N irradiation during each increment in order to minimize phase transients. The best ^{1}H carrier frequency for experiments on perpendicular bicelle samples was determined empirically to be ~9 ppm for resonances from residues in the trans-membrane helices. The most favorable ^{1}H carrier frequencies for residues in the loop and

(35) Wu, C. H.; Ramamoorthy, A.; Opella, S. J. *J. Magn. Reson. A* **1994**, *A109*, 270−272.
(36) Nevzorov, A. A.; Opella, S. J. *J. Magn. Reson.* **2003**, *164*, 182−186.
(37) Levitt, M. H.; Suter, D.; Ernst, R. R. *J. Chem. Phys.* **1986**, *84*, 4243−4255.
(38) Fung, B. M.; Khitrin, A. K.; Ermolaev, K. *J. Magn. Reson.* **2000**, *142*, 97−101.

N-terminal regions of the protein were generally somewhat different, and this necessitated the acquisition of multiple spectra in order to obtain optimal results for all of the resonances. The experimental data sets were zero filled in both the t_1 and t_2 dimensions to form a 2048 × 1024 real matrix. The theoretical scaling factor of 0.816 for heteronuclear dipolar couplings measured in PISEMA experiments was applied in the indirect dimension.[35] The scaling factor for SAMMY spectra was 0.61, corresponding to the theoretical 0.67 scaling factor[36] adjusted with an empirical correction. Phase-shifted sine bell multiplication was followed by exponential multiplication corresponding to 50−100 Hz line broadening in both dimensions prior to Fourier transformation. The NMR data were processed and displayed using the programs NMRPipe/NMRDraw[39] and Sparky.[40] The triple-resonance experiments on selectively $^{13}C'$,^{15}N-labeled samples were performed with a ^1H decoupling field of 60 kHz and a ^{13}C decoupling field of 42 kHz on the 500 MHz Varian spectrometer, and with a ^1H decoupling field of 50 kHz and a ^{13}C decoupling field of 52 kHz on the 750 MHz Bruker spectrometer.

The ^1H chemical shift frequencies were referenced to DSS by assigning the observed internal ^1H_2O resonance to 4.5 ppm at 40 °C. The ^{15}N chemical shift frequencies were referenced to liquid ammonia by assigning the observed external ^{15}N resonance of solid ammonium sulfate to 26.8 ppm.[41] The ^{13}C chemical shift frequencies were referenced to TMS by assigning the observed external low-field resonance of adamantane to 38.4 ppm.

The alignments of the phospholipids in perpendicular and parallel bicelles were verified by ^{31}P NMR spectroscopy. The protein-containing bicelles formed a stable, aligned liquid crystalline phase above 37 °C. The bicelle order parameter was determined experimentally by measuring the ^2H quadrupolar splitting of the methylene deuterons in the fatty acyl chains in one-dimensional ^2H NMR spectra of the protein-containing bicelles.[42] The splitting measured from an aligned bicelle sample was compared to that for a full powder pattern obtained from a liposome sample containing only the long-chain phospholipid.

Results and Discussion

One-Dimensional NMR Spectra of MerFt. The two principal types of aligned samples used in solid-state NMR studies of membrane proteins are mechanically aligned phospholipid bilayers on glass plates and magnetically aligned bicelles in aqueous solution. Although, cylindrical lipid bilayers on aluminum oxide nanopores have recently been proposed as an alternative.[43] The one-dimensional ^{15}N NMR spectrum of uniformly ^{15}N-labeled MerFt in Figure 2A is typical of that for a mechanically aligned bilayer sample with nearly all of the resonance intensity segregated into two bands of overlapping resonances. The clear separation between signals from residues in the trans-membrane helices (170−220 ppm) and those from residues in the terminal and loop regions (70−140 ppm) indicates that the sample is well aligned with the bilayer normals parallel to the magnetic field.

Comparison of the spectra in Figure 2A and B demonstrates the improvement in spectral resolution available by using magnetically aligned bicelles.[19] In addition to yielding narrower line widths, which accounts for the increase in fine structure in the spectrum in Figure 2B, magnetically aligned samples improve sensitivity by eliminating the glass plates from the active sample volume. The tightly sealed sample tubes ensure that the protein-containing phospholipid bilayers are maintained at full hydration throughout the course of experiments, and the use of cylindrical sample tubes means that they can be placed in solenoid coils with the attendant improvements in spectroscopic performance over the flat coils used with samples aligned on glass plates. A potential complication is that the protein-containing bicelles align with their bilayer normals perpendicular rather than parallel to the direction of the magnetic field; this means that the proteins must undergo rotational diffusion about their direction of alignment in order to average the partial powder patterns to the narrow single line resonances essential for high-resolution spectroscopy.[35] This turns out not to be a problem for membrane proteins with one, two, three, or seven trans-membrane helices, examples of which have been shown to yield narrow single-line resonances in bicelles aligned with their normals perpendicular to the magnetic field.[19,44,45] An additional advantage of using bicelles is that it is possible to change the direction of alignment 90° by adding relatively small amounts of lanthanide ions to the sample.[46] The spectrum in Figure 2B was obtained on a "flipped" bicelle sample to enable direct comparison with the mechanically aligned sample on glass plates in Figure 2A, which also has its bilayer normal parallel to the magnetic field. The large differences in chemical shift frequencies (>100 ppm) for the same amide sites in samples of MerFt in parallel (Figure 2B) and perpendicular (Figure 2C) bicelles are a consequence of the 90° difference in orientations of the ^{15}N chemical shift tensors. The reduction in the span of the spectrum in Figure 2B compared to that in Figure 2A is an effect of the bicelle order parameter[31] $S \sim 0.8$.

The high spectral resolution available from membrane proteins in magnetically aligned bicelle samples can be more fully appreciated in the one-dimensional spectra of selectively ^{15}N-labeled samples of MerFt shown in Figure 3. The ∼2 ppm line widths of resolved resonances are narrower than those observed in spectra of mechanically aligned membrane proteins in phospholipid bilayers on glass plates or single crystals of model peptides.

The spectra in Figure 3 contain single line resonances from each ^{15}N-labeled site in the protein. In spectra obtained from samples of parallel bicelles (top row), the span of the ^{15}N chemical shift frequencies is twice that observed in the spectra of the perpendicular bicelles (bottom row) (vide infra the "Bicelle Equation"). As a result, spectra obtained on parallel bicelle samples generally have better resolution in the ^{15}N chemical shift dimension since the resonance line widths are similar in both types of samples. We have not observed significant broadening of ^{15}N resonance line widths in the parallel bicelle samples containing added lanthanide ions, which is a potential concern. Nonetheless, we perform most experiments on perpendicular bicelles because the corresponding reduction in the span of the ^1H chemical shift frequencies reduces the impact of the bandwidth limitations of the pulse sequences in high-field spectrometers. This provides the best

(39) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(40) Goddard, T. D.; Kneller, D. G. *SPARKY 3*; University of California, San Francisco.
(41) Wishart, D. S.; Bigam, C. G.; Yiao, J.; Abilgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J.; Sykes, B. *J. Biomol. NMR* **1995**, *6*, 135−140.
(42) Prosser, R. S.; Hwang, J. S.; Vold, R. R. *Biophys. J.* **1998**, *74*, 2405−2418.
(43) Chekmenev, E. Y.; Hu, J.; Gor'kov, P. L.; Brey, W. W.; Cross, T. A.; Ruuge, A.; Smirnov, A. I. *J. Magn. Reson.* **2005**, *173*, 322−327.
(44) Nevzorov, A. A.; De Angelis, A. A.; Park, S. H.; Opella, S. J. In *NMR Spectroscopy of Biological Solids*; Ramamoorthy, A., Ed.; CRC Press: Boca Raton, FL, 2005; pp 177−190.
(45) Park, S. H.; Mrse, A. A.; Nevzorov, A. A.; De Angelis, A. A.; Opella, S. J. *J. Magn. Reson.* **2005**, *178*, 162−165.
(46) Prosser, R. S.; Hunt, S. A.; Di Natale, J. A.; Vold, R. R. *J. Am. Chem. Soc.* **1996**, *118*, 269−270.

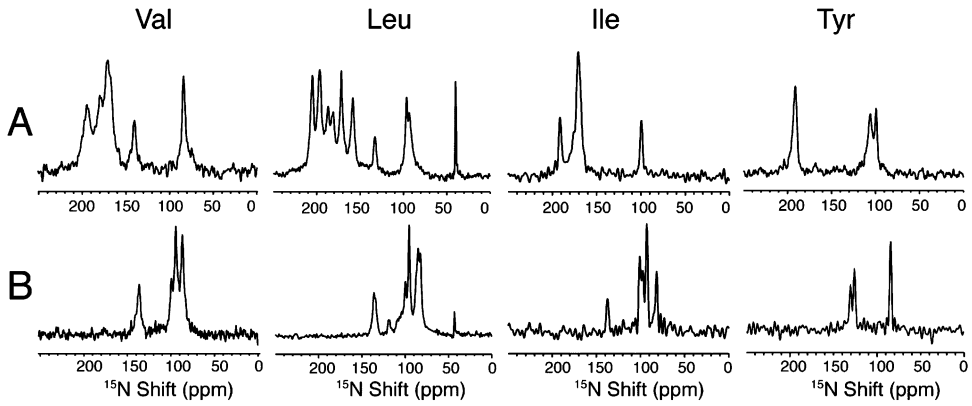

Figure 3. One-dimensional ^{15}N NMR spectra of magnetically aligned samples of selectively ^{15}N-labeled MerFt in 14-O-PC/6-O-PC, $q = 3.2$, bicelles. (A) Parallel bicelles. (B) Perpendicular bicelles. The samples in panel A were "flipped" from the perpendicular to the parallel alignment by addition of 3 mM $YbCl_3$. Spectra of valine (6 sites), leucine (13 sites), isoleucine (6 sites), and tyrosine (3 sites) labeled MerFt samples are shown left to right. The narrow resonance near 40 ppm in the leucine spectra is assigned to the N-terminal amino group.

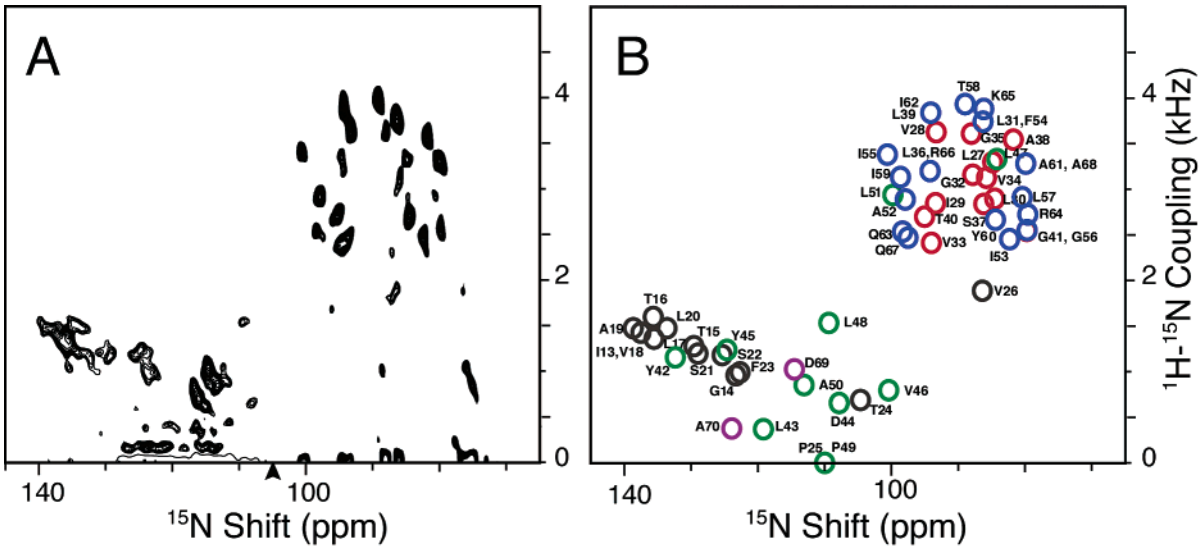

Figure 4. Two-dimensional ^{15}N chemical shift/^{1}H−^{15}N dipolar coupling NMR spectrum of uniformly ^{15}N-labeled MerFt aligned in 14-O-PC/6-O-PC, $q = 3.2$, perpendicular bicelles. (A) Composite spectrum consisting of a PISEMA spectrum (^{15}N shift < 105 ppm) and a SAMMY spectrum (^{15}N shift > 105 ppm), as marked by the arrow on the chemical shift axis. (B) Representation of the two-dimensional spectrum with the centers of the circles corresponding to the values in Table 1, which were obtained from the best available measurement. The diameter of the circles corresponds to the estimated experimental uncertainty in the measurements. The resonance assignments are noted. Colors are used to denote resonances with segments of the protein: N-terminus, residues 13−26 (gray); helix 1, residues 27−41 (red); loop, residues 42−51 (green); helix 2, residues 52−68 (blue); C-terminus, residues 69−70 (purple).

balance between resolution in the ^{15}N chemical shift and ^{1}H−^{15}N heteronuclear dipolar coupling frequency dimensions in two-dimensional spectra; nevertheless, with currently available instrumentation, it is often necessary to repeat crucial spectra using several ^{1}H carrier frequencies and pulse sequences in order to obtain optimal resolution at all combinations of frequencies and to ensure accurate measurements of individual dipolar couplings. In this context, comparisons between spectra obtained on perpendicular and parallel bicelles are useful in confirming results. In addition, the comparison provides a method for measuring the isotropic chemical shift frequencies for individual resonances, which enables the influence of orientation on the observed resonance frequencies to be precisely determined.

Two-Dimensional NMR Spectra of MerFt. The separated local field spectrum of uniformly ^{15}N-labeled MerFt in perpendicular bicelles in Figure 4A is a composite of spectra obtained from experiments optimized for resonances from residues in the trans-membrane helices (70−105 ppm) and separately for resonances from residues in the loop and terminal regions of the protein (105−140 ppm). Each amide resonance is characterized by frequencies from the ^{15}N chemical shift and ^{1}H−^{15}N heteronuclear dipolar coupling interactions. In addition to spectral resolution, the angular dependencies of these spectral frequencies result in the mapping of the protein secondary and tertiary structure onto the spectrum. It is possible to calculate solid-state NMR spectra from the coordinates of proteins with known structures.[47,48] However, our goal is to determine the three-dimensional structures of proteins *from* their experimental solid-state NMR spectra, which requires the resolution and assignment of the resonances so that their frequencies can serve as orientation constraints for all residues.

For regular secondary structure elements, such as α-helix and β-sheet, the relationship between the protein structure and spectral appearance is particularly direct and can be readily observed in many different multidimensional spectra of aligned

(47) Opella, S. J.; Stewart, P. L. *Methods Enzymol.* **1989**, *176*, 242−275.
(48) Bak, M.; Schultz, R.; Vosegaard, T.; Nielsen, N. C. *J. Magn. Reson.* **2002**, *154*, 28−45.

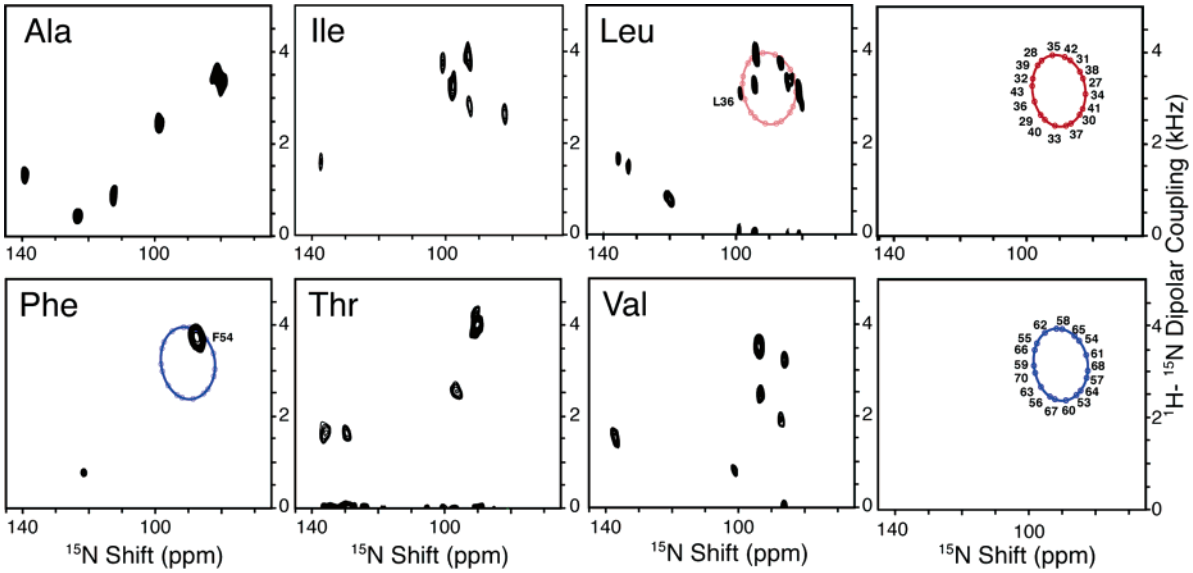

Figure 5. Two-dimensional ^{15}N chemical shift/^{1}H−^{15}N dipolar coupling NMR spectra of selectively ^{15}N-labeled MerFt samples aligned in 14-O-PC/6-O-PC, $q = 3.2$, perpendicular bicelles. The spectra are composites derived from PISEMA and SAMMY experiments, which were individually optimized at various ^{1}H carrier frequencies, typically between 8 and 9.5 ppm, and resulting from between 32 and 72 t_1 increments, depending upon the resolution needed to assign for each sample. Also shown are indexed, ideal PISA wheels calculated for an α-helix with uniform dihedral angles ($\Phi = -61°$, $\Psi = -45°$), helix tilt of 18°, and bicelle order parameter $S = 0.8$. The wheels are superimposed on the spectra of Ile- (2 sites) and Leu-labeled (13 sites) MerFt. Comparison with ideal PISA wheels gives an immediate visual estimate of the helix tilt, which here appears to be the same for both α-helices. Assignment of the only helical Phe residue, F54, by comparison with the PISA wheel is straightforward and enables indexing of the second helix. For more complex cases, like the Leu spectrum, indexing of the PISA wheels can be carried out by least-squares fitting of the patterns of the selectively labeled samples, assigned by residue type. The indexed wheels give an estimate of the phase of the helices.

proteins.[49] The "wheel-like" pattern of resonances that occurs between about 70 and 105 ppm in the two-dimensional PISEMA spectrum (Figure 4A) of uniformly ^{15}N-labeled MerFt in perpendicular bicelles is a signature indicator of the presence of tilted trans-membrane helices in the protein. PISA (polarity index slant angle) wheels correspond to helical wheel projections and serve as indices of secondary structure and topology.[50,51] Consequently, the principal features of the wheel-like patterns of resonances observed in experimental spectra provide information about the slant angle (tilt) and polarity (rotation angle) of a helix in a phospholipid bilayer. When the helix axis is parallel to the direction of sample alignment, all amide sites have the same orientation relative to the field, and the resonances overlap. Tilting the helix varies the orientations of the amide nitrogen chemical shift tensors and N−H bond vectors. This results in dispersions in both the chemical shift and dipolar coupling frequencies and is the reason that it is possible to resolve many resonances from residues in the trans-membrane helices in two-dimensional spectra. For example, nearly all of the resonances from two similar trans-membrane helices in MerFt are resolved in the spectrum in Figure 4A. PISA wheels are complemented by Dipolar waves,[52,53] which reflect the periodic wave-like variations of the magnitudes of the heteronuclear dipolar couplings as a function of residue number without chemical shift effects, which can cause distortions in the wheel-like patterns due to local variations of the amide nitrogen chemical shift tensors. Because they are based on structural regularity, the most fundamental characteristic of protein secondary structure, Dipolar waves provide accurate determinations of the number of residues in a helix as well as its tilt and rotation angles, and they are particularly effective at identifying kinks or curvature in otherwise ideal helices.[54]

NMR spectra of selectively isotopically labeled samples were used to identify resonances arising from individual types of amino acids. Combined with the nearly complete segregation of the spectrum into regions associated with the trans-membrane helices and the loop and terminal residue, this serves as the first step in the resonance assignment process. The next step was to analyze the PISA wheels since the primary sequence and hydropathy plot provide guidance as to the residues that constitute the trans-membrane helices. As sequential assignments are made, the dipolar couplings are plotted on trial Dipolar waves as checks on the results. A mis-assigned resonance from a helical residue almost invariably stands out by virtue of its falling far off the sinusoid of a Dipolar wave.

Many sequential assignments result from indexing the amino acid sequence of the helical wheel to the PISA wheel observed in the spectrum of a uniformly ^{15}N-labeled sample; this can be accomplished with one independent resonance assignment or by matching the unique pattern of resonances simulated for a helix with only one type of labeled amino acid. Figure 5 illustrates how the overlapping PISA wheels from the two trans-membrane helices in MerFt can be recognized by comparisons among the spectra of several selectively ^{15}N-labeled samples and a uniformly ^{15}N-labeled sample. The indexing by the patterns of resonances around the wheels also determines the rotation angles of the helices. The use of multiple ^{15}N-labeled samples ensures that partially overlapped resonances are rec-

(49) Marassi, F. M. *Biophys. J.* **2001**, *80*, 994−1003.
(50) Marassi, F. M.; Opella, S. J. *J. Magn. Reson.* **2000**, *144*, 150−155.
(51) Wang, J.; Denny, J.; Tian, C.; Kim, S.; Mo, Y.; Kovacs, F.; Song, Z.; Nishimura, K.; Gan, Z.; Fu, R.; Quine, J. R.; Cross, T. A. *J. Magn. Reson.* **2000**, *144*, 162−167.
(52) Mesleh, M. F.; Veglia, G.; DeSilva, T. M.; Marassi, F. M.; Opella, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 4206−4207.
(53) Mesleh, M. F.; Opella, S. J. *J. Magn. Reson.* **2003**, *163*, 288−299.
(54) Mesleh, M. F.; Lee, S.; Veglia, G.; Thiriot, D. S.; Marassi, F. M.; Opella, S. J. *J. Am. Chem. Soc.* **2003**, *125*, 8928−8935.

ognized, provides a reliable source of residue-type assignments, and confirms the indexing of the wheels. In applications of "shotgun" NMR[7] (named for the spread of labeled sites in the protein sequence and on the PISA wheel target) to small membrane proteins consisting of few helices and short loop and terminal segments, the assignment of resonances and structure determination can be performed in parallel. The structure determination of more complex proteins, such as MerFt, requires that each resonance be assigned to a specific residue prior to the analysis of Dipolar waves and the calculation of three-dimensional structures. The tilt, polarity, and ideality of the helices characterized by the analysis of PISA wheels and Dipolar waves are separately determined by the structure calculations that utilize the same experimental resonance frequencies as input.

Assignment of the Two-Dimensional Spectrum of MerFt. The resonance assignments noted in Figure 4B and listed in Table 1 were obtained largely from comparisons among spectra obtained from uniformly ^{15}N, selectively ^{15}N, and selectively ^{13}C′,^{15}N-labeled samples. The principles of the shotgun assignment method could be applied to the helical residues, even though the protein has two trans-membrane helices because their resonances are nearly completely resolved and occur in well defined, albeit overlapping, PISA wheels. Selectively labeled samples were prepared through the biosynthetic incorporation of ^{15}N-labeled Ala, Ile, Leu, Phe, Pro, Thr, Tyr, Val, (Gly/Ser), (Gln/Asp), and (Asp/Gln). We were unable to label the Gly, Gln, or Asp residues with high selectivity by incorporation of the labeled amino acids from the bacterial growth media; therefore, the resulting proteins are referred to as combinations with the other principal labeled amino acid identified in the solution NMR spectra of samples in micelles. The difference in the extent of the labeling between the two types of residues was sufficient to assign Gln and Asp by comparison of intensities of the Gln/Asp and Asp/Gln samples. Minor isotopic scrambling to other types of residues did not interfere with the assignment of resonances. In addition, a uniformly ^{15}N-labeled sample with Arg "unlabeled" was useful in confirming some assignments. Spectra obtained on longer constructs of MerF were also used to confirm a few of the MerFt assignments. Comparisons of spectra obtained in parallel and perpendicular bicelles (Figure 3) contributed to the assignment process by identifying a number of resonances that are partially overlapped in the spectra of perpendicular bicelles. Finally, some assignments were based on the observation of splittings in the ^{15}N amide resonances in proteins labeled selectively with ^{13}C′ in one type of amino acid and ^{15}N in another.

A single resonance is observed for each labeled site in the experimental spectra in Figures 4 and 5. Therefore, each residue has a unique conformation, and all protein molecules in the sample have the same three-dimensional structure and are oriented identically in the planar phospholipid bilayer portions of the bicelles. Combined with the absence of significant intensity near the isotropic resonance frequency (~120 ppm) in the one- and two-dimensional spectra, this demonstrates that the entire protein, including the loop and terminal regions, has a well-defined, stable structure in phospholipid bilayers. Most of the frequencies listed in Table 1 were measured from the spectrum of uniformly ^{15}N-labeled MerFt in Figure 4A. However, some values were obtained from separately optimized spectra. For example, all dipolar couplings with magnitudes <500 Hz were measured in both perpendicular and parallel bicelles, and the values listed in Table 1 and used in the structure calculations are actually half the experimental values measured in parallel bicelles where more reliable measurements could be made. The two Pro residues in MerFt were not distinguishable and displayed partially overlapping peaks in the ^{15}N spectrum; therefore, we chose the average value of their chemical shifts for both sites (Table 1). Whenever resonances were spread across the full ^{15}N and ^{1}H chemical shift ranges in spectra of

Table 1. ^{1}H−^{15}N Dipolar Coupling and ^{15}N Chemical Shift Frequencies of MerFt Assignment of the Two-Dimensional Spectrum of MerFt

residue	^{15}N chemical shift (ppm)	^{1}H–^{15}N dipolar coupling (Hz)
I13	137.5	1428
G14	123.3	962
T15	129.6	1286
T16	135.6	1609
L17	135.5	1367
V18	137.5	1428
A19	138.7	1479
L20	133.6	1483
S21	129	1198
S22	125.4	1183
*F*23	122.7	994
T24	104.7	690
P25	110	0
V26	86.5	1890
L27	84.9	3312
V28	93.4	3638
I29	93.5	2858
L30	84.7	2903
L31	86.3	3754
G32	88	3168
V33	94.1	2420
V34	85.9	3132
G35	88.1	3621
L36	94.4	3206
S37	86.3	2849
A38	81.8	3557
L39	94.2	3853
T40	95.2	2709
G41	79.8	2556
Y42	132.4	1161
L43	119.2	367
D44	107.9	659
Y45	124.7	1239
V46	100.5	793
L47	84.3	3344
L48	109.4	1537
P49	110	0
A50	113.1	852
L51	99.8	2948
A52	98	2901
I53	82.4	2457
F54	86.3	3754
I55	100.6	3395
G56	79.8	2556
L57	80.5	2920
T58	89.1	3946
I59	98.7	3146
Y60	84.5	2672
A61	80	3290
I62	94.2	3853
Q63	98.4	2540
R64	79.7	2735
K65	86.3	3893
R66	94.4	3206
Q67	97.6	2477
A68	80	3290
D69	114.5	1025
A70	123.9	380

selectively ^{15}N-labeled samples, the measurements were repeated using both PISEMA[35] and SAMMY[36] at several different ^{1}H resonance offset frequencies to ensure that the same scaling factor was effective for all of the dipolar coupling measurements. Variations of peak positions for the same resonances in different samples and between selectively and uniformly ^{15}N-labeled samples were less than the line widths and thus were included in the experimental uncertainties used in the structure calculations.

The segregation of signals from residues in trans-membrane helices and those in irregularly structured loop or terminal segments into two distinct spectral regions is sufficient to assign some resonances in the spectra of selectively labeled samples. For example, MerFt contains only two Phe residues, one in the N-terminal segment (Phe 23) and one in a trans-membrane helix (Phe 54). There are two resonances in the spectrum of selectively ^{15}N Phe-labeled MerFt in Figure 5. Since resonances from residues in trans-membrane helices have chemical shift frequencies between 70 and 110 ppm in ^{15}N NMR spectra of proteins in perpendicular bicelles, the resonance at 80 ppm can be assigned to Phe 54 and that at 122 ppm to Phe 23 with considerable confidence. Notably, the identification of the resonance of Phe 54 also serves to determine the rotation angle of the second trans-membrane helix because it falls on the PISA wheel pattern observed in the spectrum of uniformly ^{15}N-labeled MerFt. The relatively well-defined wheel-like pattern of resonances in the PISEMA spectrum in Figure 4A indicates that both trans-membrane helices have similar tilt angles. Resonances from the Ile, Leu, and Val residues could be readily identified by comparison of the spectra obtained from the selectively and uniformly ^{15}N-labeled samples. Guided by the amino acid sequence, two ideal PISA wheels (with the same tilt angles) were generated by simultaneous least-squares fitting to the resonances from the Phe, Tyr, Ile, Leu, and Val residues in the trans-membrane helices. These resonances also define the phase of the PISA wheel and hence the rotation angles of the corresponding helical wheels for the trans-membrane helices. The PISA wheels with the best fits to the resonances are tilted 18° from the bilayer normal, and this was confirmed by the quantitative Dipolar wave analysis and the calculated three-dimensional structure of the protein.

The assignment of resonances to specific residues in the loop and terminal segments of MerFt was facilitated by the elimination of many possibilities due to the prior assignment of the helical region of the spectrum. Some resonances could be assigned in the N-terminal region of the protein because there was only a single possibility for the type of amino acid after the resonances from the helical residues were assigned, for example, Ile 13 and Phe 23. Comparisons with constructs of MerF containing different N-terminal sequences helped discriminate among possibilities when there was more than one residue of a type in the loop or terminal segments. For example, once all of the Val residues in the trans-membrane helices were assigned, the Val resonance at 137 ppm could be assigned to Val 18 because it was significantly shifted in a construct of MerF with a different N-terminal sequence. By elimination, this enabled the remaining Val resonance to be assigned to Val 46 in the inter-helical loop region of the protein.

Although, the majority of resonances could be assigned in a straightforward manner using double-resonance experiments on

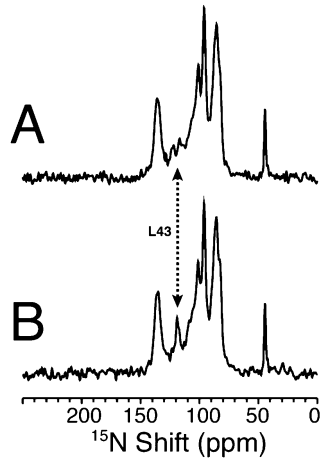

Figure 6. One-dimensional ^{15}N NMR spectra of selectively ^{13}C′ Tyr- and ^{15}N Leu-labeled MerFt aligned in perpendicular bicelles. (A) With only ^{1}H decoupling. (B) With both ^{1}H and ^{13}C decoupling.

selectively and uniformly ^{15}N-labeled samples, a few assignments required the use of ^{13}C′ and ^{15}N double-labeled samples. Double selective labeling can be used to identify ^{15}N-labeled amino acid Y in an XY pair when amino acid X is ^{13}C′-labeled.[55] Whenever the XY pair appears in the primary sequence, the one-dimensional ^{15}N spectrum of a ^{13}C′ X-labeled and ^{15}N Y-labeled sample has a doublet due to the ^{13}C′−^{15}N dipolar coupling. This splitting, which is less than 1.1 kHz and varies with the orientation of the XY peptide plane in the magnetic field, is affected by the application of ^{13}C decoupling during acquisition of the ^{15}N signals. As shown in Figure 6, the ^{15}N resonance of Leu 43 can be unambiguously assigned by comparing the spectra of (^{13}C′ Tyr, ^{15}N Leu)-labeled MerFt acquired without (Figure 6A) and with (Figure 6B) ^{13}C decoupling during data acquisition. The spectrum obtained with ^{13}C decoupling was indistinguishable from that obtained from a sample labeled only with ^{15}N Leu (Figure 3). The spectra of several of the double-labeled samples acquired without ^{13}C decoupling showed only a broadening or other distortion of the corresponding ^{15}N resonance. This can be attributed to a combination of an orientation of the peptide plane that yields a small C′−N coupling and only partial labeling with ^{13}C, such that there is a central component consisting of a ^{15}N resonance flanked by the spectral components from the ^{15}N bonded to ^{13}C. However, as long as there is a discernible difference between spectra obtained without and with ^{13}C decoupling during data acquisition, the resonance assignments can be made with this approach. Samples of MerFt were prepared selectively ^{13}C′,^{15}N double-labeled with (^{13}C′ Tyr, ^{15}N Leu), (^{13}C′ Val, ^{15}N Leu), (^{13}C′ Leu,^{15}N Leu), (^{13}C′ Gly, ^{15}N Tyr), and (^{13}C′ Pro, ^{15}N Ala).

The assignment of the Tyr resonances combined elements of both the double- and triple-resonance methods described above. In Figure 7, the two-dimensional spectrum of ^{15}N Tyr-labeled MerFt is compared to the one-dimensional ^{15}N spectra of (^{13}C′ Gly, ^{15}N Tyr)-labeled MerFt obtained in the absence and presence of ^{13}C decoupling. The resonance from Tyr 60 can be assigned by inspection since it is the only Tyr residue in a trans-membrane helix. There are two additional resonances in the spectrum, corresponding to residues 42 and 45 in the inter-helical loop. Since there is only one GlyTyr pair in MerFt,

(55) Cross, T. A.; Opella, S. J. *J. Mol. Biol.* **1985**, *182*, 367−381.

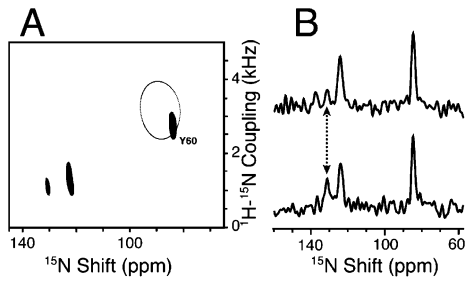

Figure 7. The assignment of the three tyrosine resonances in the spectrum of MerFt. (A) Two-dimensional PISEMA spectrum of selectively ^{15}N Tyr-labeled MerFt in perpendicular bicelles. An ideal PISA wheel calculated for an α-helix with uniform dihedral angles ($\Phi = -61°$, $\Psi = -45°$), helix tilt of 18°, and bicelle order parameter $S = 0.8$ is superimposed on the spectrum. (B) One-dimensional ^{15}N NMR spectra of selectively ^{13}C′ Gly- and ^{15}N Tyr-labeled MerFt in perpendicular bicelles. With only ^{1}H decoupling (top) and with both ^{1}H and ^{13}C decoupling (bottom).

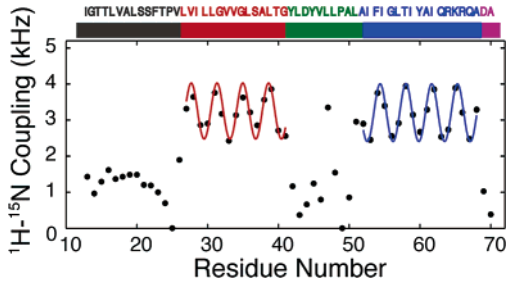

Figure 8. Plot of experimental dipolar coupling frequencies as a function of residue number with Dipolar wave fits. The frequencies are divided into four groups that are color-coded by residue number in the primary sequence of the peptide corresponding to the secondary structural domains of the protein: N-terminus, residues 13−26 (gray); first helix, residues 27−41 (red wave); loop, residues 42−51 (green); second helix, residues 52−68 (blue wave); and C-terminus (purple).

Tyr 42 and 45 were distinguished, as indicated, on the basis of the one-dimensional spectra of (^{13}C′ Gly, ^{15}N Tyr) MerFt acquired without and with ^{13}C decoupling.

Structural Analysis Using Dipolar Couplings. The experimental dipolar coupling frequencies listed in Table 1 are plotted as a function of residue number in Figure 8. For clarity of presentation and to focus on the principal structural features of the protein, the amino acid sequence of the protein is divided into five sections: residues 13−26 are the N-terminal segment, 27−41 are the first trans-membrane helix, 42−51 are the inter-helical loop, 52−67 are the second trans-membrane helix, and the remainder constitute the C-terminus. The Dipolar waves that denote and characterize the two trans-membrane helices are shown in Figure 8, where sinusoids with a 3.6 residues per turn period were fit to the experimental dipolar coupling frequencies. The fitting of the experimental data to the Dipolar waves was accomplished using rigid-body rotations of an ideal α-helix by setting $\Phi = -61°$, $\Psi = -45°$ for all residues in the structural fitting algorithm. The Dipolar waves show that the helical segments of MerFt in phospholipid bilayers are nearly ideal, and that they have very similar tilt angles. The RMSDs for the fits of the experimental data to Dipolar waves are less than the experimental errors for the measurements of the dipolar coupling frequencies (<200 Hz).

Although this type of analysis is straightforward for secondary structure elements, where the regularities of the structural and spectroscopic parameters are coincident, it also provides valuable qualitative insights into the properties of other regions of the protein. Notably, the irregular pattern of dipolar couplings found for the residues between the two helices indicates that the loop has a unique, complex conformation. By contrast, there is a much simpler pattern for the N-terminal residues, although the distinct nonzero dipolar couplings for these residues demonstrate that they too have unique conformations up to the N-terminus.

Calculation of the Three-Dimensional Structure of MerFt. The experimental data in Table 1 provide input for calculation of the three-dimensional structure of MerFt using the previously described structural fitting algorithm,[56] which was modified for application to membrane proteins in perpendicular bicelles. Rapid uniaxial rotational diffusion about the bilayer normal results in narrow resonances and scales their frequencies relative to their isotropic values.[57] By taking into account the bicelle order parameter S arising from the fast restricted "wobble" motion of bicelles, the following linear equation (the "Bicelle Equation") that interconverts the heteronuclear dipolar coupling and chemical shift frequencies observed in spectra of bilayer samples aligned on glass plates with their normals parallel to the field with those in spectra of perpendicular bicelles can be written

$$\nu_{\text{bicelle}} = \left(1 + \frac{S}{2}\right)\sigma_{\text{iso}} - \frac{S}{2}\nu_{\text{bilayer}}$$

where σ_{iso} is the isotropic chemical shift for a given resonance; for the dipolar couplings, the equivalent parameter is zero. The Ramachandran angles Φ and Ψ are determined by linking the solutions for the peptide plane orientations of adjacent residues obtained from the calculated 0° alignment frequencies (ν_{bilayer}) as previously described.[56] This is accomplished by minimizing the difference between the experimental and calculated frequencies using a closed-form expression, which describes the relationship between the ^{15}N chemical shift and ^{1}H−^{15}N heteronuclear dipolar coupling frequencies for the neighboring amide sites in terms of peptide plane rotations given by the angles Φ and Ψ.[57]

Polytopic membrane proteins present a more complex situation for structure calculations than the earlier monotopic examples mainly because they have a larger number of residues in inter-helical loops and terminal regions for which there are no compact regions on the Ramachandran plot, where a limited range of solutions for the torsion angles Φ and Ψ can be sought. Because the mapping of the spectra onto a three-dimensional backbone structure is not unique, there are multiple structures consistent with a given set of frequencies. The most general way to deal with the structural degeneracies is through the incorporation of additional experimental frequency measurements into the structure calculations, and a host of three-dimensional and triple-resonance experiments are being developed for this purpose. However, it is possible to determine protein structures using only data from ^{1}H/^{15}N double-resonance experiments by inclusion of several structural constraints derived from analysis of previously determined protein structures. The integration of experimental and bioinformatics constraints in the structure calculations is illustrated by the flowchart in Figure 9.

(56) Nevzorov, A. A.; Opella, S. J. *J. Magn. Reson.* **2003**, *160*, 33−39.
(57) Nevzorov, A. A., Opella, S. J., Webb, G. A., Eds. *Modern Magnetic Resonance Part III*; Springer: Berlin, 2006; pp 1−10.

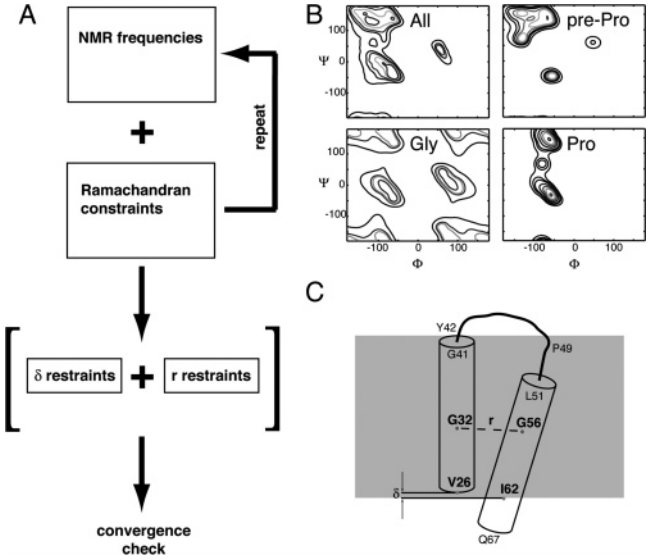

Figure 9. (A) Flowchart for the structure calculations. (B) Ramachandran plots (RAMA) of prolines, pre-prolines, glycines, and all other residues used to restrict the space of the torsion angles for nonhelical residues. (C) Definitions of parameters for the helix−helix packing constraints described in the text.

The two types of bioinformatics constraints incorporated into the structure calculations are based on Ramachandran plots derived from high-resolution crystal structures of globular proteins and on the helix−helix packing arrangements observed in membrane protein crystal structures. The recently compiled individual Ramachandran (RAMA) plots for prolines, glycines, pre-prolines, and the general type, including all other residues, are a substantial improvement over previously available plots.[58] The contour plots were derived from high-resolution protein structures excluding the secondary structure regions (i.e., α-helix and β-sheet). The Ramachandran plots shown in Figure 9B were highly effective in limiting the torsion angle space (combinations of Φ and Ψ) available for the loop residues. For the helical regions, the structural solutions were sought within predefined limits relative to the ideal values, $\Phi = -61°$, $\Psi = -45°$, as described previously.[59,60] The N−H bond length used in all calculations, including Dipolar waves, was 1.05 Å. The principal values of the ^{15}N chemical shift tensor were $\sigma_{11} = 64$ ppm, $\sigma_{22} = 77$ ppm, and $\sigma_{33} = 222$ ppm, for nonglycine residues; $\sigma_{11} = 41$ ppm, $\sigma_{22} = 64$ ppm, and $\sigma_{33} = 215$ ppm for glycine residues. The proline residues are a special case because their ^{15}N chemical shift tensor is significantly different from that of an amide nitrogen in a peptide bond, and they do not have a N−H bond to provide a second orientationally dependent frequency from a dipolar coupling. The ^{15}N Pro chemical shift frequencies were included in the fitting as restraints using the principal values $\sigma_{11} = 38$ ppm, $\sigma_{22} = 127$ ppm, and $\sigma_{33} = 231$ ppm.[61] The angle between the N−H(R) bond vector and the σ_{33} component of the ^{15}N chemical shift tensor was 18.5° for all nonproline residues and 20° for proline residues.

Helix packing in membrane proteins can be parametrized by δ and r, which are defined in Figure 9C. The parameter δ describes the relative displacement of the two trans-membrane helices in the z-direction and reflects hydrophobic matching.[62] Specifically, the α-carbon of the first hydrophobic residue of the first helix is allowed to deviate from the α-carbon of the last hydrophobic residue of the second helix by no more than $\pm\delta$ in the vertical dimension. Since the distance between the α-carbons of any two adjacent residues is 3.8 Å, which can be regarded as a rather large displacement, we selected $\delta = \pm 1.5$ Å as the maximum allowed deviation.

Additional support for this choice comes from the analysis of the 35 trans-membrane helices in bacteriorhodopsin[63] (PDB 1C3W, 7 TM helices), halorhodopsin[64] (1E12, 7TM helices), a photosynthetic reaction center[65] (1OGV, 10TM helices), KcsA[66] (1K4C, 2TM helices), and Ca ATPase[67] (1SU4, 9 TM helices). These examples were chosen because their crystallographic z-axes are coincident with the normal to the membrane plane, enabling the z-positions of the C_α atoms to be directly examined. We found that the C_α atoms of the highly hydrophobic residues (Ile, Val, Leu, Phe, Trp) located near the ends of the trans-membrane helices could be simultaneously confined between two bounding planes which are roughly 3 Å apart and which lie in the membrane interface region. Moving the planes up or down by as little as 1 Å would cause hydrophilic residues to appear. Moreover, the corresponding bounding planes on the other side of the membrane could be found by a simple translation in the z-direction by 21−23 Å, which corresponds to an average depth of the hydrophobic membrane interior. Among the few exceptions were lysines whose side chains may snorkel. These findings support our postulate about the existence of a well-defined membrane boundary layer of essentially hydrophobic amino acids that determines the position of the trans-membrane helices due to hydrophobic matching. In the case of monotopic membrane proteins, this defines their tilt angles.[62] Along with the Dipolar waves, which define the helical regions precisely, this enables the application of the δ restraint to the trans-membrane helices of MerFt. As can be seen from the Dipolar waves in Figure 8, suitable candidates are C_α of Val 26 (at the beginning of the first helix) and Ile 62 (the last hydrophobic residue of the second helix), which are highly hydrophobic residues and, therefore, are assumed to be confined within the 3 Å bounding layer at the membrane interface consistent with our choice of $\delta = \pm 1.5$ Å in the structure calculations.

We also impose distance restraints on the helix packing by accepting only those solutions for which the distance between the α-carbons of the middle residues of the two helices are separated by a distance between 8 and 15 Å.[68,69]

The Dipolar waves in Figure 8 show that both helices are tilted by the same angle relative to the bilayer normal; this means that the middle residue of the first helix (Gly 32) lies away

(58) Lovell, S. C.; Davis, I. W.; Arendall, W. B., III; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins* **2003**, *50*, 437−450.
(59) Thiriot, D. S.; Nevzorov, A. A.; Zagyanskiy, L.; Wu, C. H.; Opella, S. J. *J. Mol. Biol.* **2004**, *341*, 869−879.
(60) Hildebrand, P. W.; Preisssner, R.; Frömmel, C. *FEBS Lett.* **2004**, *559*, 145−151.
(61) Waddell, K. W.; Checkmenev, E. Y.; Wittebort, R. J. *J. Am. Chem. Soc.* **2005**, *127*, 9030−9035.
(62) Park, S. H.; Opella, S. J. *J. Mol. Biol.* **2005**, *350*, 310−318.
(63) Luecke, H.; Schobert, B.; Richter, H. T.; Cartailler, J. P.; Lanyi, J. K. *J. Mol. Biol.* **1999**, *291*, 899−911.
(64) Kolbe, M.; Besir, H.; Essen, L. O.; Oesterhelt, D. *Science* **2000**, *288*, 1390−1396.
(65) Katona, G.; Andreasson, U.; Landau, E. M.; Andreasson, L.-E.; Neutze, R. *J. Mol. Biol.* **2003**, *331*, 681−692.
(66) Zhou, Y.; Morais-Cabral, J. H.; Kaufman, A.; MacKinnon, R. *Nature* **2001**, *414*, 43−48.
(67) Toyoshima, C.; Nakasako, M.; Nomura, H.; Ogawa, H. *Nature* **2000**, *405*, 647−655.
(68) Reddy, B. V. B.; Blundell, T. L. *J. Mol. Biol.* **1993**, *233*, 464−479.
(69) Bowie, J. U. *J. Mol. Biol.* **1997**, *272*, 780−789.

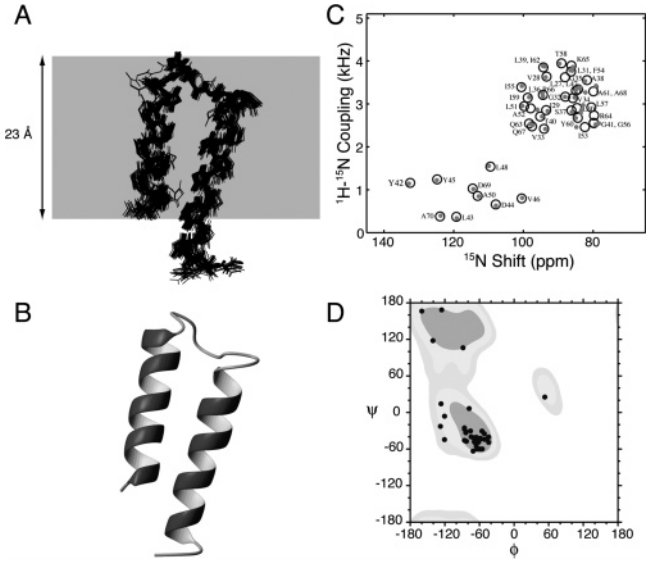

Figure 10. Three-dimensional backbone structure of residues 27−68 of MerFt in phospholipid bilayers. (A) Wire frame representation. (B) Ribbon diagram. (C) Back-calculated spectrum (dots) from one of the structures shown in part A superimposed on the representation of the experimental data. (D) Ramachandran plot with all of the backbone angles marked with dots.

from the cytoplasmic membrane surface at approximately the same distance as the middle residue of the second helix (Gly 56) since they both reside 6 residues away from that surface. Therefore, in the structure calculations, residues Gly 32 and Gly 56 were used to check against the loose helix−helix distance restraints in the calculations. Only solutions for which the distance between the α-carbons of these residues fell between 8 and 15 Å were accepted.

The calculated three-dimensional backbone structure of residues 27−68 of MerFt is illustrated in Figure 10A and B (deposited in the PDB as 2H3O). The positions of the resonances for these residues in the two-dimensional spectrum are shown as circles that encompass the experimental uncertainties in the measurements of the chemical shift and dipolar coupling frequencies. The "dots" in Figure 10B correspond to the spectrum back-calculated from a structure in Figure 10A. To address the issue of experimental uncertainty, the protein structure was calculated from data that were artificially randomized within ±1.5 ppm for the chemical shifts and ±100 Hz for the dipolar couplings relative to the values listed in Table 1. Several thousand structures that satisfy both the NMR data and the Ramachandran restraints were calculated in this way. The application of the δ-restraint on Val 26 and Ile 62 selects roughly 1 out of every 10 structures, and the further application of the weak distance restraints on residues Gly 32 and Gly 56 results in a close set of about 20 structures having RMSD of ∼1.5 Å for residues 27−68 of MerFt. To include all possible torsion angle solutions consistent with the experimental data, as the structural fitting progressed from residue to residue, the starting Φ and Ψ values for the loop residues were randomized within ±180°. Significantly, the resulting calculated structures were not highly sensitive to minor changes of parameters or small deviations of the spectral frequencies, within the limits of the experimental measurements. For example, imposing the helix packing constraints with other pairs of residues, such as Leu 27 and Ala 61 (for δ) or Val 33 and Ile 55 (for r), yields structures that are essentially indistinguishable from those superimposed in Figure 10A.

Structure of MerFt. The results of the structure calculations shown in Figure 10 are consistent with the analysis of the PISA wheels and Dipolar waves (Figure 8). Although the same experimental data are used with these methods of analysis, they provide separate views of the helical portions of the protein structure since information about neither the tilt nor the rotation of the helices is input explicitly into the structure calculations.

The protein structure shown in Figure 10B is dominated by two nearly parallel trans-membrane helices. The first helix starts at residue 27, although residue 26 appears to be in a transition into the helix following Pro 25, and ends at residue 41, and the second helix spans residues 52−68. Both helices have the same tilt angle relative to the membrane normal. The first helix has 15 or possibly 16 residues; the second helix is slightly longer with 17 residues and contains hydrophilic residues at the end (QRQRQA) that extend beyond the hydrophobic portion of the bilayer. The inter-helical loop consists of 10 residues (42−51) and appears to reside in the interfacial region of the phospholipid bilayer.

Notably, the NMR spectra contain a resonance from each labeled nitrogen site, including those in residues 13−24, which indicates that the N-terminal region of MerFt has a unique conformation in the bilayer environment, even though it is truncated compared to that of the full-length wild-type protein. The frequencies of the ^{15}N chemical shifts and especially the ^{1}H−^{15}N dipolar couplings indicate that the residues 13−24 have an extended structure that is approximately parallel to the surface of the bilayer. No single converged solution emerged for these residues using the approach outlined in Figure 9; however, since spectra of longer constructs of the protein that include all of the N-terminal residues show significant differences for these same resonances, the structure of residues 13−24 most relevant for understanding the mechanism of mercury transport across membranes will be that determined for the full-length protein. There are many similarities between the structure of MerFt determined in phospholipid bilayers by solid-state NMR and that determined in micelles by solution NMR;[29] however, they differ most in the dynamics of residues 13−24, which are mobile and unstructured in micelles. With the implications of the N-terminal truncation of the polypeptide in mind, we postpone a detailed discussion of structural differences resulting from the lipid environments, spectroscopic approaches, and time scales of NMR experiments until the studies of the full-length protein are completed.

Conclusions

It is possible to isotopically label polytopic membrane proteins by expression in bacteria and reconstitute them in phospholipid bilayer samples suitable for high-resolution NMR spectroscopy. The observation of well-defined resonances in both one- and two-dimensional spectra from all labeled sites, none of which are at the isotropic chemical shift frequency, is more than a spectroscopic curiosity. It shows that every residue in the polypeptide, including those in the loop and terminal region, has a unique conformation on the 10^4 Hz time scale of the NMR experiments. It also shows that every polypeptide in the sample has the same three-dimensional structure, and that they are all aligned in exactly the same way in phospholipid bilayers. While

the structures of the helices and their tilt angles relative to the membrane normal can be determined with great confidence, their relative positions clearly depend on both the uniqueness of the structure of the inter-helical loop and the hydrophobic and distance constraints imposed after the structural fitting. We found that only a single close set of loop structures is simultaneously consistent with the assigned data points, RAMA restraints, and the weak distance restraints. The added hydrophobic and distance[68,69] criteria are independent of the NMR data and are applied as weak restraints. Throughout the calculations, we allowed for experimental error by slightly randomizing the experimental data within ± 1.5 ppm in the chemical shift dimension and ± 100 Hz in the dipolar dimension. These values are quite conservative compared to the precision and reproducibility of the experimental measurements. Calculations yield three-dimensional backbone structures with a precision corresponding to RMSD values of <2.0 Å. Additional spectroscopic measurements and further refinement of the calculation procedures will almost certainly yield protein structures with higher precision and improve the accuracy of determining tertiary arrangements by solid-state NMR of aligned samples. The resonance line widths and other spectroscopic features indicate that the structures of membrane proteins in phospholipids bilayers are at least as well defined as those of globular proteins in crystals or solution.

Membrane proteins with between one[19] and seven[70] transmembrane helices from viruses and humans yield similar spectroscopic results, indicating that the approach is not limited to bitopic proteins of bacterial origin. There is no fundamental size limitation to solid-state NMR spectroscopy since the averaging of the spin interactions required for line narrowing is accomplished with radio frequency irradiations rather than rapid isotropic molecular reorientation. With continued development, this approach to structure determination should be applicable to substantially larger polypeptides, although proteins with several hundred residues will require more sophisticated experiments to deal with the overlap of resonances. Three-dimensional experiments will provide opportunities for resolution of overlapping resonances and the measurement of additional orientation-dependent frequencies. Triple-resonance experiments will not only improve the structure determination of the backbone through measurements that include the Cα and C′ sites but also provide a mechanism for characterizing the side chains so that they can be included in complete three-dimensional structures of the proteins.

Solid-state NMR can be used to characterize the details of backbone and side chain dynamics. Thus, by combining structure determination with studies of dynamics, solid-state NMR spectroscopy has the potential to provide a comprehensive description of polytopic membrane proteins in their native phospholipid bilayer environment under physiological conditions of pH and temperature. In the case of MerF, this will provide insight into the mechanism of transporting Hg(II) across the membrane from the periplasm into the cytoplasm of the bacterial cell where it is detoxified by reduction to Hg(0).

Acknowledgment. We thank S.H. Park for preparing the mechanically aligned bilayer sample on glass plates used to obtain the spectrum in Figure 2A, and C.H. Wu, C.V. Grant, and A.A. Mrse for assistance with the instrumentation and helpful discussions. We also thank I.W. Davis (Duke University) for advice on the new Ramachandran plots. This research was supported by grants from the National Institutes of Health RO1GM066978 and P41EB002031, which supports the Biomedical Technology Resource for NMR Molecular Imaging of Proteins at the University of California, San Diego; A.A.D. was supported by postdoctoral fellowship F32GM65833, and S.C.H. by Training Grant DK54441.

JA063640W

(70) Park, S. H.; Prytulla, S.; De Angelis, A. A.; Brown, J. M.; Kiefer, H.; Opella, S. J. *J. Am. Chem. Soc.* **2006**, *128*, 7402−7403.
(71) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. *Nucleic Acids Res.* **2003**, *31*, 3784−3788.