CrossMark click for updates

Structure of CrgA, a cell division structural and regulatory protein from *Mycobacterium tuberculosis*, in lipid bilayers

Nabanita Das[a,b], Jian Dai[a,c], Ivan Hung[b], Malini R. Rajagopalan[d], Huan-Xiang Zhou[a,c], and Timothy A. Cross[a,b,e,1]

[a]Institute of Molecular Biophysics, and Departments of [c]Physics and [e]Chemistry and Biochemistry, Florida State University, Tallahassee, FL 32306; [b]National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL 32310; and [d]Biomedical Research Branch, The University of Texas Health Science Center at Tyler, Tyler, TX 75708

Edited by Huilin Li, Stony Brook University, Stony Brook, NY, and accepted by the Editorial Board December 2, 2014 (received for review August 17, 2014)

The 93-residue transmembrane protein CrgA in *Mycobacterium tuberculosis* is a central component of the divisome, a large macromolecular machine responsible for cell division. Through interactions with multiple other components including FtsZ, FtsQ, FtsI (PBPB), PBPA, and CwsA, CrgA facilitates the recruitment of the proteins essential for peptidoglycan synthesis to the divisome and stabilizes the divisome. CrgA is predicted to have two transmembrane helices. Here, the structure of CrgA was determined in a liquid–crystalline lipid bilayer environment by solid-state NMR spectroscopy. Oriented-sample data yielded orientational restraints, whereas magic-angle spinning data yielded interhelical distance restraints. These data define a complete structure for the transmembrane domain and provide rich information on the conformational ensembles of the partially disordered N-terminal region and interhelical loop. The structure of the transmembrane domain was refined using restrained molecular dynamics simulations in an all-atom representation of the same lipid bilayer environment as in the NMR samples. The two transmembrane helices form a left-handed packing arrangement with a crossing angle of 24° at the conserved Gly39 residue. This helix pair exposes other conserved glycine and alanine residues to the fatty acyl environment, which are potential sites for binding CrgA's partners such as CwsA and FtsQ. This approach combining oriented-sample and magic-angle spinning NMR spectroscopy in native-like lipid bilayers with restrained molecular dynamics simulations represents a powerful tool for structural characterization of not only isolated membrane proteins, but their complexes, such as those that form macromolecular machines.

membrane protein structure | transmembrane helix binding motif | intrinsically disordered proteins | solid-state NMR | oriented samples

Better understanding of cell division in *Mycobacterium tuberculosis* (*Mtb*), the causative agent of tuberculosis (TB), will generate new opportunities for pharmaceutical development. CrgA, a transmembrane (TM) protein, is a central component of the *Mtb* divisome (1). CrgA has homologs in other actinomycetes (2, 3), but not in the two bacteria, *Escherichia coli* and *Bacillus subtilis*, with better characterized cell division mechanisms. Conversely, many cell division proteins in the latter organisms, such as FtsA, FtsN, FtsL, and ZipA, appear to have no homologs in *Mtb.* CrgA is localized at the poles and septum, and interacts with multiple cell division proteins, including FtsZ, FtsQ, FtsI (PBPB), PBPA, and CwsA. One function of these interactions is to stabilize the divisome (1, 4). The interaction with CwsA, a protein that is unique to mycobacteria (5), might coordinate elongation at the poles and division at midcell (4). Moreover, CrgA appears to have an important role in peptidoglycan (PG) formation during cell division, by recruiting PG synthases to the divisome (4). Reduced production of CrgA results in elongated cells and reduced growth rate (1), and loss of CrgA impairs PG synthesis (5). In addition to CwsA, the *Mtb* divisome involves other atypical players such as FipA (FhaB), ChiZ, and MtrB (6–8), and thus there is much yet to be learned about the participants in mycobacterial cell division (9). Here, we determined the structure of CrgA in a lipid bilayer environment using solid-state NMR (ssNMR) spectroscopy.

TB is a devastating human disease that kills ∼1.3 million people each year with 8.6 million new cases diagnosed annually worldwide (10). Rising extreme drug-resistant *Mtb* strains do not succumb to the frontline antibiotics, generating a dire need for new drugs (11). Pathways critical for bacterial survival such as DNA replication and cell division include numerous potential drug targets and represent a major focus for structural biology. Also, TB treatment is expensive and significantly toxic and requires an extensive period caused by *Mtb*'s ability to exist in a latent state. Hence, there are additional motivations for characterizing the proteins associated with its survival in active and nonreplicative persistent states.

CrgA was first described from *Streptomyces* as being required for sporulation through coordinating several aspects of its reproductive growth (2, 3). The *Mtb* CrgA consists of 93 residues, with two predicted TM helices (12) (TM1: residues 29–51; and TM2: residues 66–88; Fig. 1*A*). The N-terminal 17 residues are

Significance

Understanding the structure and function of the cell division apparatus of *Mycobacterium tuberculosis* is crucial for advancing drug development against tuberculosis. Here, we report the solid-state NMR structure of a transmembrane protein, CrgA, that is a central component of the *M. tuberculosis* divisome. Small helical membrane protein structures are particularly sensitive to their environment, and consequently, we characterized CrgA in an environment that models well the biophysical properties of the native membrane. To determine the structure, both oriented sample and magic-angle spinning NMR data from liquid–crystalline lipid bilayer preparations were used along with refinement by restrained molecular dynamics simulations in the same lipid environment. The structure suggests how CrgA serves as a platform for binding and recruiting other proteins of the divisome.

Author contributions: H.-X.Z. and T.A.C. designed research; N.D., J.D., I.H., and M.R.R. performed research; N.D., J.D., I.H., H.-X.Z., and T.A.C. analyzed data; and N.D., J.D., M.R.R., H.-X.Z., and T.A.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. H.L. is a guest editor invited by the Editorial Board.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2mmu). Anisotropic chemical shifts and dipolar couplings from oriented-sample NMR, and isotropic chemical shifts from magic-angle spinning NMR, are deposited in the Biological Magnetic Resonance Bank, www.bmrb.wisc.edu (accession no. 19867).

[1]To whom correspondence should be addressed. Email: cross@magnet.fsu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1415908112/-/DCSupplemental.

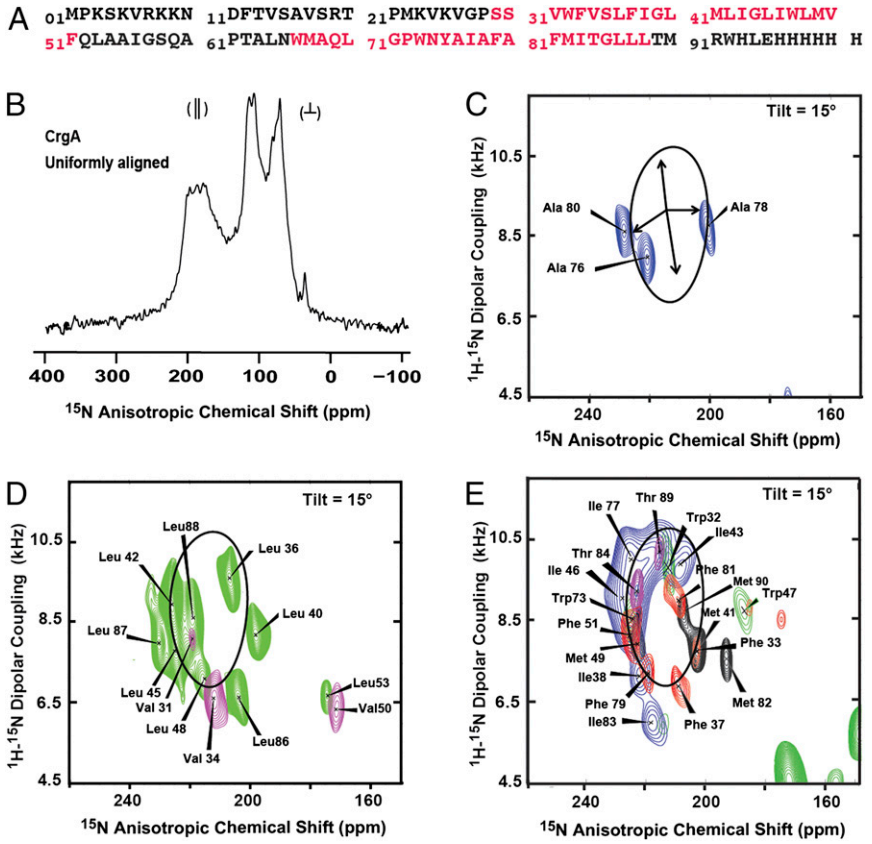

Fig. 1. Amino acid sequence and ssNMR spectra of full-length *Mtb* CrgA membrane protein. (*A*) The sequence of the expressed CrgA with a C-terminal His_6 tag. The predicted transmembrane (TM) helical residues (TMHMM, version 2.0) are indicated by red lettering. (*B*) One-dimensional ^{15}N cross-polarization spectrum of ^{15}N uniformly labeled CrgA with the bilayer normal parallel to the magnetic field. The resonance intensities near 200 ppm are from the TM helix residues, with their backbone amide NH bonds nearly parallel to the bilayer normal. The intensity near 120 ppm is from the highly dynamic sites, and most of the remaining intensity is from structured residues not in the TM domain. (*C*) Two-dimensional PISEMA spectra of ^{15}N Ala-labeled CrgA with a superimposed PISA wheel calculated for an ideal helix [$(\varphi, \psi) = (-60°, -45°)$] with a 15° tilt angle relative to the bilayer normal. The set of four arrows in the middle of the PISA wheel define 90° rotational increments for the backbone nitrogen sites. Resonance assignments based on the PISA analysis are shown. (*D*) Superimposed PISEMA spectra of ^{15}N Leu (green)- and ^{15}N Val (purple)-labeled protein with resonance assignments. (*E*) Overlay of PISEMA spectra for ^{15}N Ile (blue)-, ^{15}N Met (black)-, ^{15}N Thr (purple)-, ^{15}N Phe (red)-, and ^{15}N Trp (green)-labeled CrgA, with sequence-specific assignments shown. All spectra were collected at 60 MHz for ^{15}N in oriented POPC/POPG (4:1 mol/mol) liquid–crystalline phospholipid bilayers, pH 7.0, 13 °C.

predicted to be disordered by the software PONDR (13); the C terminus is predicted to be just five residues, whereas the loop between the TM helices is predicted to be just 14 residues. The predicted TM1 sequence contains a pair of conserved tryptophan residues (W32 and W47) that appear from the sequence to be positioned for anchoring the helix to the membrane interfacial regions. A second pair of conserved tryptophan residues is at positions 73 and 92. Because the TM2 prediction has W73 eight residues into the helix and W92 four residues beyond the end of the predicted helix, this prediction may not be as accurate. Both predicted helices contain a number of other conserved residues, whereas the loop between the helices is much more variable both in length and in composition (*SI Appendix*, Fig. S1).

Surprisingly, this small membrane protein binds a large number of other proteins, all of which are transmembrane proteins except for FtsZ. In particular, FtsI, with a single TM helix, is a transpeptidase responsible for synthesis of the septal PG (1). A *crgA*-deletion mutant results in the loss of septal and polar localization of FtsI, suggesting the importance of CrgA for PG synthesis through its recruitment of FtsI. CwsA also contains a single TM helix. A *crgA* and *cwsA* double-deletion mutant showed the importance of the corresponding gene products for cell wall synthesis and cell shape maintenance (8).

The CrgA TM helices contain a number of conserved glycine and alanine residues (*SI Appendix*, Fig. S1). Although glycine residues are known to be helix breakers in water-soluble proteins, in TM helices, they may allow local helix bending in the low dielectric membrane environment where intrahelical hydrogen bonds are strengthened for maintaining the overall integrity of the helical structure. In addition, glycine and alanine residues permit close approach of adjacent helical backbones, resulting in backbone–backbone electrostatic and side-chain–side-chain van der Waals interactions that stabilize the tertiary structure. Therefore, glycines may allow helical membrane proteins to sacrifice secondary structural stability for tertiary structural stability (14–16). This is needed because the amino acid composition in the interior of membrane proteins is more hydrophobic than the interior of water-soluble proteins where there are more frequent tertiary hydrogen bonds than in TM domains (17). In addition, conserved glycine residues are rarely found on the fatty-acyl exposed surface of multihelix membrane proteins (16). In such a location, they would expose their hydrophilic backbone atoms to the low dielectric environment of the protein. If present, it is a strong indication that they are exposed for a required function such as binding another protein. Interestingly, *E. coli* FtsQ is thought to localize to the divisome through interactions with other components via its single TM helix (18).

Only a couple of full-length *Mtb* membrane protein structures have been determined. One is an X-ray structure of the mechanosensitive channel of large conductance, and the other is a single TM helix protein, Rv1761 (19, 20). In addition, water-soluble domains of other *Mtb* membrane proteins have been characterized such as those from PknB and FtsX (21–24). Although X-ray crystallographers have focused on large membrane proteins, the majority of the 1,162 ORFs of the *Mtb* genome code for small helical membrane proteins containing one to three TM helices with <40-kDa molecular weight (25). Structure–function studies of these small membrane proteins are essential for understanding *Mtb* cell division and other cellular processes. Small polytopic membrane protein structures are stabilized not just by interactions between their TM helices, but also by interactions with their membrane environment. Consequently, it is necessary to solve their structures in an appropriate membrane mimetic environment, one that possesses many of the restraining influences of the native membrane such as a relatively fixed hydrophobic thickness, a dramatic lateral pressure profile, and a hydrophobic core essentially devoid of water (26, 27).

For the structure determination of CrgA, here we used both oriented-sample (OS) and magic-angle spinning (MAS) ssNMR to characterize the full-length protein in lipid bilayers. All ssNMR spectroscopy was performed on fully hydrated liquid–crystalline lipid bilayer preparations of CrgA. The use of such bilayer preparations for supporting the native-like conformation of the M2 protein from Influenza A has been validated with the comparison of spectra from synthetic bilayers and from cellular membranes where the protein has been inserted by the cellular machinery and never removed from this environment or exposed to a detergent environment (28). Multiple recent membrane protein structures have now been determined by OS ssNMR (29–34), and the first membrane protein structure has been obtained from MAS ssNMR (35). OS ssNMR generates information on the orientations of peptide planes with respect to the bilayer normal, and for a TM helix it yields the tilt angle of the helix relative to the lipid bilayer normal and rotational orientation about the helical axis along the entire length of helix. However, it does not directly provide information on the helix–helix packing interface. The latter information can be ascertained by relatively few distance restraints between the helices, as the degrees of freedom for packing the helices have been minimized by the orientational restraints, to just the relative rotation around the bilayer normal and relative translation in the bilayer plane. The combination of OS and MAS ssNMR thus allows the complete determination of the helical TM domain structure.

Based on the OS and MAS data, we refined the structure using restrained molecular dynamics simulations in an all-atom representation of the same lipid bilayer environment as in the protein samples. The two TM helices both have a tilt angle of 13° but are tilted in nearly opposite directions such that they form a left-handed packing arrangement with a crossing angle of 24° at the conserved Gly39 residue. The two-helix TM domain exposes other conserved glycine and alanine residues that potentially form binding sites for TM helices of CrgA binders. Much of the N-terminal region is disordered, but a nine-residue motif therein appears to form an amphipathic helix. In the interhelical loop, a short segment appears to be disordered while a 12-residue motif appears to form a β-hairpin in the membrane interface. The C terminus comprises just two residues. Overall, the structure suggests how CrgA serves as a platform where other proteins of the divisome assemble.

Results

Sample Characterization. ^{15}N, ^{13}C uniform-labeled and amino acid-specific–labeled full-length CrgA was expressed with a C-terminal His_6 tag through an LE linker (Fig. 1*A*). In 12% (wt/vol) SDS/PAGE gel, the purified protein appeared as a single band at 12 kDa (*SI Appendix*, Fig. S2*A*). The purified protein was then reconstituted in 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC)/1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPG) (4:1 mol/mol) liposomes for structural characterization. As an initial secondary structure characterization, the circular dichroism (CD) spectrum of the reconstituted proteoliposomes was obtained suggesting 49% helix, 12% β-sheet, 17% turn, and 22% random coil [CD Pro analysis (36); *SI Appendix*, Fig. S2*B*].

OS ssNMR. The orientation of full-length CrgA in glass-slide–supported POPC/POPG lipid bilayers (with a gel-to-liquid crystalline phase transition temperature of −2 °C) was first assessed using the one-dimensional (1D) ^{15}N ssNMR chemical shift spectrum of uniform ^{15}N-labeled protein at 13 °C, pH 8.0 (Fig. 1*B*). The spectrum is dominated by the amide backbone resonances. Approximately 10% of the intensity is at the isotropic frequency (110–130 ppm) indicating structural disorder for a significant portion of the amino acid sequence in addition to the His-tag, considering that cross-polarization of the isotropic intensity is often weak. Here, the B_0 field direction was parallel to the bilayer normal, and thus the considerable intensity near the parallel edge of the anisotropic chemical shift tensor (~230–180 ppm) suggests that the TM helices account for most but maybe not all of the 49% α-helical content indicated by CD spectroscopy. The rest of the spectral intensity accounts for the loops and terminal regions.

Fig. 1*C* shows the ^{15}N Ala polarization inversion spin exchange at magic angle (PISEMA) (37) spectrum for three alanine residues (at positions 76, 78, 80) predicted to be in TM2. Similar to helical wheels, there are 3.6 resonances per turn in a polarity index slant angle (PISA) wheel for an α-helix (i.e., 100° separation between adjacent resonances) (38, 39), as shown superimposed on the spectrum. The tilt of the helix controls the position and size of the PISA wheel. This wheel, for a 15° tilt of the helix axis relative to the bilayer normal, was chosen for best alignment with all of the spectral data (not just the alanine data) and provides an initial assessment of helix tilt. The assignments for these alanine residues were based on the fact that residues i and $i + 2$ should be on opposite sides of the wheel (~200° apart) and residues i to $i + 4$ should be near each other (~40° apart). The unique assignment of these residues also provided an initial assessment of the rotational orientation of the helix. The full ^{15}N Ala spectrum shows resolved resonances for all nine alanine residues in CrgA (*SI Appendix*, Fig. S3) including resonances from the disordered regions of the protein and the structured interfacial regions. Although a bacterial two-hybrid assay suggested that CrgA forms a dimer (1), the single resonances observed here for all of the residues indicate a monomer (or symmetric dimer) in these preparations.

Fig. 1*D* shows the superimposed PISEMA spectra of ^{15}N Val- and ^{15}N Leu-labeled CrgA obtained from separate samples. All three valine residues predicted to be in the TM helices are in TM1 (at positions 31, 34, and 50), but residue 50, due to its predicted position in the last turn of TM1 has a potentially deformed orientation for its peptide plane (40), and consequently it was excluded from initial assignment. Residues 31 and 34, corresponding to i and $i + 3$ positions, should be on one side of the wheel (~60° apart). The initial tilt assignment of TM1, again by alignment of the corresponding PISA wheel with all of the data, is ~15°, similar to that of TM2. Via the assignment of the two valine resonances, an initial assessment of the rotational orientation for TM1 was also made. Based on the tilt and rotational characterizations for the two TM helices, the remaining resonance assignments were made.

The leucine residues in TM1 (at positions 36, 40, 42, 45, and 48) are distributed almost uniformly around the helical and PISA wheels, whereas those in TM2 (at positions 86, 87, and 88) are on one-half of its wheel. Fig. 1*E* shows five additional superimposed PISEMA spectra for ^{15}N Phe (residues 33, 37, 51, 79, and 81),

Met (residues 41, 49, 82, and 90), Trp (residues 32, 47, and 73), Ile (residues 38, 43, 46, 77, and 83), and Thr (residues 84 and 89). The vast majority of the resonances for the TM helices (Fig. 1*E*) fall in close proximity to the 15° PISA wheels for both TM helices, with the exception of the resonances from residues in the first or last turns of the helices. All of the helical backbone sites in the TM helices were observed except for Gly44. The anisotropic chemical shifts and dipolar couplings for the TM helices are listed in Table 1. In addition, the conformity of the data to PISA wheels permits the use of typical α-helical restraints associated with hydrogen bonding distances and membrane protein helical torsion angle ($\phi = -60°$ and $\psi = -45°$) restraints (41). Most of the resonances in the N-terminal region, the two-residue C terminus, and interhelical loop were observed, and many indicate significant structural disorder by displaying very little anisotropy in their oriented sample resonance frequencies.

The orientational restraints for the TM helices (Table 1) when plotted as a function of residue number, known as dipolar and chemical shift waves (42, 43), provide additional structural detail for the experimentally characterized helices consisting of residues 31–52 and residues 73–91 (*SI Appendix*, Fig. S4). There appear to be some local perturbations to the helices in addition to the first and last turns, such as the anisotropic chemical shifts observed at Met41 in TM1 and Tyr75 in TM2, where the data lie significantly off the waves. Perturbations primarily associated with the anisotropic chemical shifts may reflect differences in their chemical shift tensors. Perturbations in the first or last turn are likely to reflect the interaction of this helical turn with the membrane interface through amphipathic side chains, anchoring and stabilizing the TM domain in the lipid bilayer (44, 45), and hence the helical backbone may be somewhat perturbed at these sites. However, it is clear that the tilt and rotational orientation of the helices remain constant throughout their lengths.

Most of the remaining PISEMA resonances could not be sequence-specifically assigned except for Ile56 and Trp66 to a structured portion of the interhelical loop and Phe12 to the disordered N-terminal region. The other resonances were amino acid-specifically assigned because of the isotopic labeling strategy. For methionine, there are two residues in addition to those in the TM helices. Met67 is in the interhelical loop, next to Trp66, and could be assigned to a Met resonance similar to that of Trp66. That left Met22 with a resonance indicating a structured portion in the N-terminal region. The rest of the observed resonances were tentatively assigned according to proximity in the spectra (sequentially neighboring residues were assigned to nearby resonances, i.e., those having similar dynamics or conformation; *SI Appendix*, Table S1).

MAS ssNMR. Although OS ssNMR characterizes the helical backbone structure and the orientations of the helices relative to the bilayer normal, distances between the helices are valuable for unique determination of the tertiary structure. We obtained the latter data using ^{13}C-^{13}C correlation [dipolar-assisted rotational resonance (DARR)] spectra. Before distances could be obtained, sequence-specific assignments were made for the isotropic ^{13}C resonances. This was complicated by the uniformity of the helical structures resulting from the hydrophobic amino acid content and the low dielectric environment that strengthens intrahelical hydrogen bonds (41, 46). However, we were able to make unique assignments for segments of the helical backbone and some of the side chains from 3D MAS experiments [including NCACX, NCOCX, and CAN(CO)CX experiments (47, 48)] of uniformly ^{15}N and ^{13}C-labeled CrgA samples (*SI Appendix*, Fig. S5). The resonance assignments for residues 38–43 and 48–50 of TM1 and 75–85 of TM2 are listed in *SI Appendix*, Table S2, representing enough assignments to obtain a set of unique interhelical distances. Fig. 2 *A* and *B* shows the superposition of 2D DARR spectra (49, 50) with different mixing times from a ^{13}C, ^{15}N reverse-labeled sample in which all of the residues were labeled except for Thr, Ile, Phe, Ser, and Trp. These spectra resulted in interhelical cross peaks between Met49 and both Tyr75 and Ala78 as well as a cross peak between Leu42 and Ala80 (distances listed in *SI Appendix*, Table S3). In a second ^{13}C, ^{15}N reverse-labeled sample, all of the amino acid residues except for Ile, Leu, Phe, Tyr, and Ser were isotopically labeled, and another interhelical resonance between Gly39 and Thr84 was observed in the DARR spectra (Fig. 2*C*).

Table 1. CrgA ^{1}H-^{15}N dipolar coupling (DC) and ^{15}N anisotropic chemical shift (ACS) data from the PISEMA spectra

TM1 (residue Val31–Gln52)			TM2 (residue Trp73–Arg91)		
Residues	DC, kHz	ACS, ppm	Residues	DC, kHz	ACS, ppm
Val31	8.07	221.3	Trp73	8.40	224.6
Trp32	9.23	210.3	Asn74	9.57	208.9
Phe33	7.84	203.5	Tyr75	4.28	189.7
Val34	6.5	211.5	Ala76	7.80	220.9
Ser35	9.07	227.5	Ile77	10.07	224.8
Leu36	9.57	206.9	Ala78	8.75	200.8
Phe37	6.96	209.4	Phe79	7.20	219.1
Ile38	7.23	222.0	Ala80	8.70	228.5
Gly39	9.66	220.4	Phe81	9.10	210.0
Leu40	8.20	198.3	Met82	7.80	193.7
Met41	7.90	202.9	Ile83	6.06	218.1
Leu42	8.93	226.0	Thr84	9.17	223.0
Ile43	9.90	208.6	Gly85	9.77	203.3
Gly44	NA	NA	Leu86	6.62	203.8
Leu45	7.59	224.6	Leu87	7.95	230.4
Ile46	9.02	225.3	Leu88	8.57	219.1
Trp47	8.75	188.6	Thr89	10.35	215.8
Leu48	7.09	215.8	Met90	8.79	207.4
Met49	8.07	222.5	Arg91	4.73	177.3
Val50	6.42	171.3			
Phe51	8.23	225.1			
Gln52	6.64	174.3			

CrgA Structure. Based on the OS and MAS ssNMR data described above, we calculated the structure of the CrgA TM domain (34). First, a set of 960 initial conformations was generated by simulated annealing (51), with the dipolar couplings and anisotropic chemical shifts of helical residues restrained to the values observed in the OS ssNMR experiments (Table 1) and the interhelical distances observed in the MAS ssNMR experiments restrained to a conservative upper bound of 8 Å (*SI Appendix*, Table S3). After filtering by TM1 and TM2 tilt angles and clustering, the average pairwise backbone root-mean-square deviation among 10 selected conformations is 0.38 Å (see *SI Appendix*, Table S4, for other statistics of the structure calculation). A representative conformation was then further refined by restrained molecular dynamics simulations, again with the restraints from the ssNMR observables, but now with an all-atom representation of the same lipid bilayer environment used for collecting the ssNMR data. ssNMR parameters calculated on the final structure are in close agreement with the experimental values (*SI Appendix*, Fig. S6).

The refined structure has residues 31–49 forming TM1, two residues shorter on both termini of the helix compared with the prediction (Fig. 1*A*). These terminal residues are likely to be part of the helix, but their peptide plane orientations appear to deviate substantially from those in the helix core (Table 1), and consequently their PISEMA resonances fall outside of the PISA wheel, but still in close proximity. TM2 is formed by residues 73–91, considerably different from the predicted residues 66–88 (Fig. 1*A*). Residues 66, 67, and 68 were labeled, but their resonances did not appear in the vicinity of the PISA wheel, and

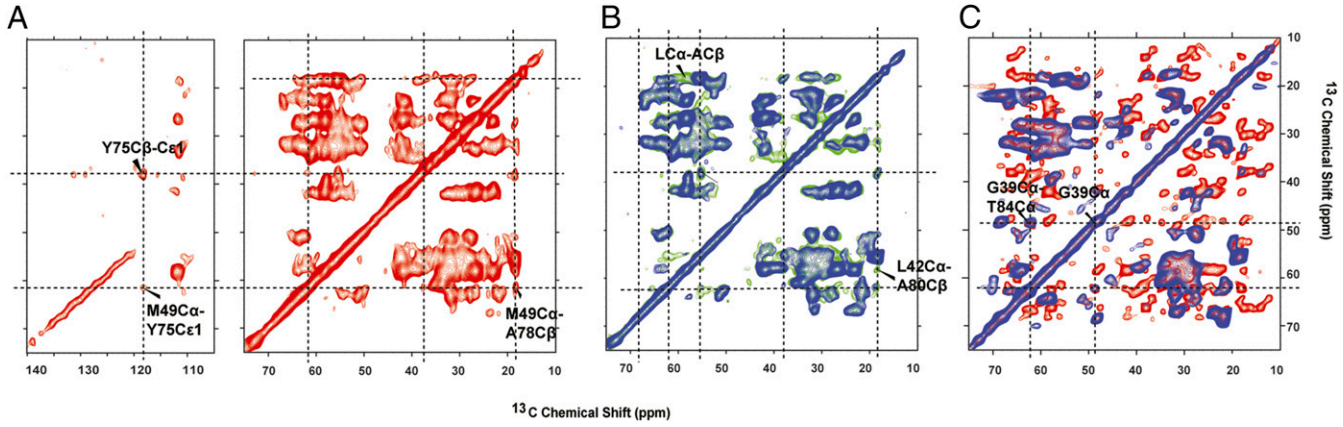

Fig. 2. Two-dimensional ^{13}C-^{13}C DARR-MAS spectra of ^{13}C, ^{15}N reverse-labeled CrgA. (*A* and *B*) DARR spectra of CrgA with all but Thr, Ile, Phe, Ser, and Trp labeled, using mixing times of 1,000 ms (red), 700 ms (green), and 300 ms (blue). Cross peaks corresponding to three interhelical distance restraints are indicated. (*C*) Superimposed DARR spectra of CrgA with unlabeled Ile, Leu, Phe, Tyr, and Ser at 100-ms mixing time (blue) and ^{13}C, ^{15}N uniform-labeled CrgA at 50-ms mixing time (red). The Gly39–Thr84 cross peak represents an interhelical distance restraint. All spectra were collected at 600 MHz proton frequency in liquid–crystalline POPC/POPG liposomes at 13 °C and 12-kHz spinning rate.

therefore these residues are unlikely to be part of TM2. Gly71 appears to be disordered, and the next residue is a proline, a known helix breaker, leaving Trp73 as the starting residues of TM2. It is thus important to experimentally characterize these helices and not depend solely on prediction tools.

The side view of the refined structure of CrgA TM domain shows a pair of antiparallel α-helices with a tilt angle of 13° with respect to the bilayer normal for both TM1 and TM2 (Fig. 3*A*). These tilt values are close to those obtained initially from the PISA wheel and wave analysis of the PISEMA spectra. The four interhelical distances from MAS ssNMR define the packing arrangement of the two helices (Fig. 3*B*), in which the two TM helices tilt in nearly opposite directions, resulting in a left-handed bundle with a crossing angle of 24° and a crossing midpoint near Gly39 and Leu42 of TM1 and Phe81 and Thr84 of TM2, positioned near the center of the membrane (Fig. 3).

The interhelical interface can be well described as "knobs in holes" (52). Close van der Waals interactions extend over the entire length of the helix pair, involving 11 residues from TM1 (at positions 31, 32, 35, 36, 38, 39, 42, 43, 45, 46, and 49) and 9 residues from TM2 (at positions 74, 77, 78, 80, 81, 84, 87, 88, and 91), and contribute significantly to the tertiary structural stability (Fig. 4*A*). Position 39 in the middle of the interface is conserved as a small residue, either glycine or alanine (*SI Appendix*, Fig. S1). One of the primary reasons for the conservation of Gly/Ala in TM domains is for forming stable helical complexes (53, 54). Two more conserved small residues, A78 and A80, also contribute to the interhelical interface. Interestingly, two highly conserved glycines, Gly44 and Gly85, are not part of the interface and can potentially form interaction sites for other TM proteins or as a CrgA dimer (Fig. 5*A*; *Discussion*).

Gly39 CαH potentially forms a hydrogen bond across the interface with the side-chain hydroxyl of Thr84 (Fig. 4*B*), a type of interaction known to stabilize helix–helix packing (33, 55). Besides Thr84, there are two more polar side chains (Ser35 and Thr89) within the helices that form intrahelical hydrogen bonds with neighboring backbone carbonyls. Ser30 is in a helix-capping position and its side chain hydrogen bonds with a backbone amide. In addition, the side chain of Asn74 reaches across the interhelical interface to hydrogen bond with the backbone carbonyl of Leu53. This variety of side-chain–backbone hydrogen bonding is typical of polar residues in TM helices (56).

The structure is defined not only internally but also with respect to its environment (Fig. 5*A*). These helices span nearly 30 Å of the POPC/POPG bilayer hydrophobic thickness. Two tryptophan residues at the helix termini (Trp47 on TM1 and Trp73 on TM2) have their indole N–H oriented toward the membrane interfacial regions and serve as anchors (44, 57); two other tryptophans (Trp32 and Trp92) form an interhelical stacking interaction. Neighboring polar and charged residues (Ser30 and Gln52 for TM1 and Tyr75 and Arg91 for TM2) form hydrogen bonds with lipid phosphate and carbonyl groups (Fig. 5 *B*–*E*). All four tryptophans are absolutely conserved and somewhat less so for Tyr75. Clearly, these residues help to orient the protein in the membrane environment as has been noted for other proteins (45, 58). Interestingly, three consecutive polar residues at the N terminus of TM2 appear to interact with the membrane interface, something that is only possible when the helical tilt angle is small.

The OS ssNMR data outside TM1 and TM2, although sparsely sequence-specifically assigned, contain rich information for structural characterization of the full-length protein. They confirm the disorder of the N-terminal 18 residues predicted by PONDR and further reveal a short disordered segment (Ala68 to Pro72) at

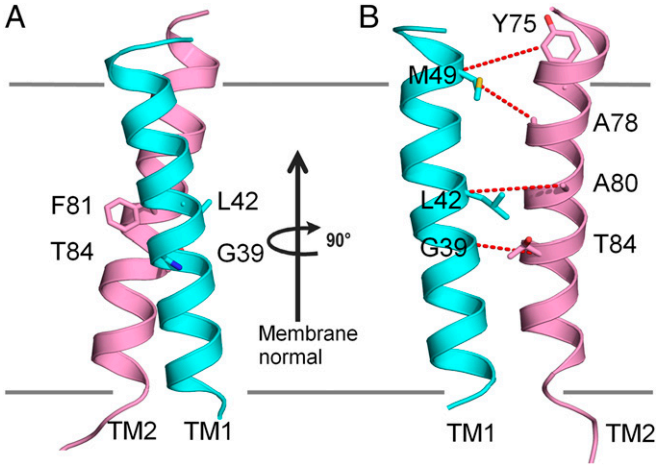

Fig. 3. Structure of CrgA transmembrane domain determined in lipid bilayers. (*A*) TM1 (cyan) and TM2 (pink) both have a tilt angle of 13°, and cross each other at a 24° angle near Gly39 and Leu42 of TM1 and Phe81 and Thr84 of TM2. (*B*) A rotated view showing the four interhelical distance restraints (dashed red lines) determined from the MAS spectra.

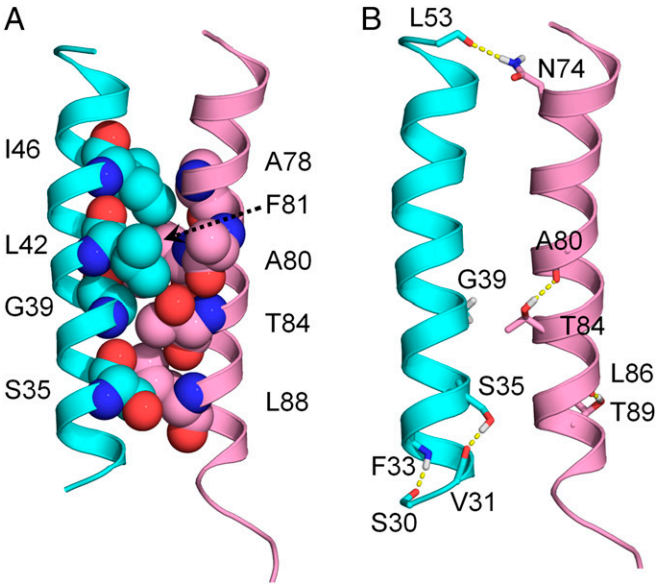

Fig. 4. Interhelical and intrahelical interactions in CrgA. (*A*) Interhelical interface with a "knobs-in-holes" pattern involving complementary large and small side chains on the two helices. (*B*) Interhelical and intrahelical side-chain–backbone hydrogen bonds.

the C terminus of the interhelical loop. The sequences for both of these disordered segments are highly conserved among mycobacterial species (*SI Appendix*, Fig. S1). The structured portions in the N-terminal region (residues 19–27) and in the loop (residues 56–67) have remarkably uniform chemical shifts between 70 and 85 ppm and dipolar couplings between 4.3 and 5.5 kHz, with only a couple of outliers for each parameter. The immediate conclusion is these residues form α-helices or β-strands, with the peptide N–H bonds nearly parallel to the bilayer surface, such as an amphipathic helix or a β-sheet in the plane of the bilayer interface. Furthermore, the CD spectral analysis suggests that there are ∼10 residues of additional helix and ∼10 residues of β-strand, consistent with the preceding OS ssNMR data. Although both segments could form amphipathic helices, the segment in the loop has Pro61 in the middle, which potentially facilitates the formation of a γ-turn. Thereby, this segment can form a β-hairpin, with the terminal residues in close proximity, which is required by the close packing of the TM helices at the extracytoplasmic end (Fig. 3). Although these results are strongly suggestive for the structure of the full-length protein, these segments were not included in the refinement.

Discussion

Recently, general and robust protocols have been published for OS and MAS ssNMR sample preparation (59). Although the ssNMR structural approach demonstrated here for characterizing the structure of CrgA is relatively new, it has general applicability to a large segment of the proteome that has been largely inaccessible by other structural tools. This is because the crystallization of small helical membrane proteins is very difficult (60, 61) and the structures characterized in detergent micelles are often significantly distorted by the single hydrophilic surface of the micelle (27). Consequently, it is important to characterize these relatively small membrane proteins in lipid bilayer environments. As demonstrated here, the combination of OS and MAS ssNMR provides an ideal approach.

CrgA has the responsibility of coordinating the assembly of multiple membrane proteins for cell division (1). CrgA binds to FtsZ, the protein responsible for initiating the formation of the Z-ring when cells divide. Because FtsZ is a water-soluble protein, its interaction with CrgA, documented by bacterial two-hybrid assays, pull-down assays, and colocalization of fluorescent labeled proteins in vivo, must occur through the cytoplasmic N-terminal region of CrgA, which as shown here is largely disordered and consequently has certain advantages for binding other proteins (62). Two penicillin-binding proteins, the transpeptidases FtsI (PBPB) and PBPA, are single-TM helix membrane proteins responsible for synthesis of the septal PG. Based on *crgA* deletion experiments, this gene product appears to be responsible for the recruitment of FtsI to the divisome. Bacterial two-hybrid assays have demonstrated that the interaction between CrgA and FtsI occurs through the latter's cytoplasmic domain (1). Similar assays have shown that CrgA also binds PBPA, but this protein has a minimalistic cytoplasmic N terminus of five or six residues that is unlikely to be responsible for the binding and its TM helix has no obvious helix binding motif. Therefore, we anticipate that PBPA binds to the interhelical loop of CrgA, which appears to form a β-hairpin followed by a disordered segment. The structure of CrgA determined here serves as an important platform for assembling these partner proteins (Fig. 6*A*).

Conserved residues are typically not observed on the lipid-facing surface of membrane proteins, where hydrophobic residues can replace each other without significantly affecting the tertiary stability of the protein in the lipid bilayer. Glycines on this surface, however, expose hydrophilic backbone atoms to the hydrophobic interstices of the bilayer, thereby reducing tertiary stability (16). Hence conserved glycine residues on the lipid-facing surface may indicate a functionally important role, such as stabilizing protein–protein interactions. For CrgA, although the highly conserved Gly39 residue is used in the intraprotein helix packing, the strictly conserved Gly44 and Gly85 residues are exposed to lipid acyl chains (Fig. 5*A*). Their positions strongly suggest that these are sites for the binding of additional proteins.

In addition to FtsZ, FtsI, and PBPA, CrgA interacts with FtsQ and CwsA, which are also associated with *Mtb* cell division and are TM helical proteins (1). FtsQ and CwsA both have helix-binding motifs in their single TM helices. The FtsQ TM helix has an AxxGxGxGxA sequence that translates into two helix-binding

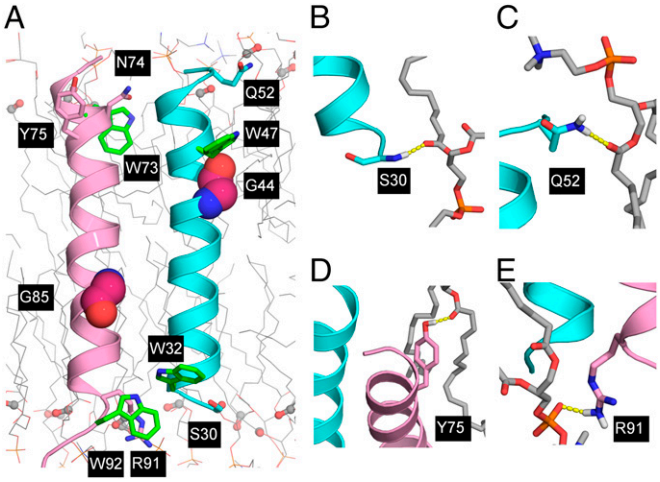

Fig. 5. Interactions of CrgA with its environment. (*A*) Positioning of the transmembrane helices in a lipid bilayer. Conserved Trp side chains are shown as green sticks for carbon atoms; neighboring polar and charged side chains are shown as cyan (TM1) or pink (TM2) sticks for carbon atoms. Conserved outward-facing Gly residues are shown as magenta spheres for carbons. Lipid carbonyls are shown as ball-and-stick. In all cases, oxygen atoms are in red and nitrogen atoms are in blue. (*B–E*) Hydrogen bonds of Ser30, Gln52, Tyr75, and Arg91 with lipid phosphate and carbonyl groups.

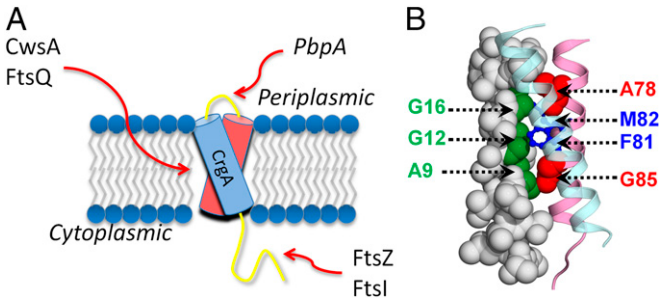

Fig. 6. The structure of *Mtb* CrgA as a platform for assembling other interacting proteins. (*A*) Schematic view of the likely binding sites for five cell division proteins known to interact with CrgA. FtsZ and FtsI interact with the cytoplasmic N terminus of CrgA. CwsA and FtsQ have binding motifs in their transmembrane (TM) helices and are therefore likely to interact with the CrgA TM helices. PBPA has no such binding motif in its TM helix and a minimalistic N-terminal sequence and therefore is hypothesized to interact with the CrgA interhelical loop. (*B*) A structural model of the CrgA–FtsQ complex. TM1 and TM2 of CrgA are shown as cyan and pink ribbons, respectively. FtsQ is shown as spheres. Residues involved in binding are shown as red spheres (Thr78 and Gly85) and blue ball-and-stick (Phe81 and Met82) on the CrgA side and as green spheres (Gly16, Gly12, and Ala9) on the FtsQ side.

motifs on opposite sides of the FtsQ helix, an AxxGxxxG motif on one side and a GxxxA motif on the other side. Based on the CrgA TM domain structure, modeling suggests how the AxxGxxxG motif may take advantage of two exposed small residues in CrgA TM2, Ala78 and Gly85, in generating an extensive van der Waals binding interface between these two proteins (Fig. 6*B*). So we propose that the binding of CrgA to FtsQ is through the TM domain. This leaves CrgA G44 still exposed to the lipid interstices, and it could be a site for dimerization or binding CwsA. This small protein has a single TM helix, WxxAGxAAAxxAGGAxAxSxxR, that is rich in small residues. This sequence has a plethora of surfaces for protein interactions to the extent that large hydrophobic residues that provide the majority of the van der Waals interactions for helix–helix binding are scarce. A *cwsA* and *crgA* double-deletion mutant showed that these proteins are required for cell wall PG synthesis and cell morphology (5).

Fig. 6*A* summarizes the potential binding sites on CrgA for its important function in the divisome. CwsA and FtsQ may bind through the TM helices of CrgA, such as illustrated by the model in Fig. 6*B* for the FtsQ–CrgA complex. FtsZ and FtsI may bind to the cytoplasmic N-terminal region and PBPA may bind to the extracytoplasmic loop of CrgA. Although it is not clear how many of these proteins bind at the same time, there is potential through these multiple binding sites for CrgA to play the central role in recruitment of the PG synthesis machinery during cell division.

Methods

Methods used in this study are briefly summarized below. Full descriptions are given in *SI Appendix*.

CrgA Protein Expression, Purification, and Reconstitution. The protocol for preparing the protein samples was similar to one described previously (1). For ^{13}C, ^{15}N uniformly labeled CrgA protein preparation, 1 L of M9 media was supplemented with 1 g of $^{15}NH_4Cl$ and 2 g of ^{13}C-glucose. Media preparation for other types of labeling is found in *SI Appendix*. Details of the procedure for CrgA protein purification and of the buffer preparations were previously published (1, 59). Full-length CrgA was reconstituted in POPC/POPG (4:1 mol/mol) liposomes at pH 8.0 as previously described (59).

Circular Dichroism Spectroscopy. Initial secondary structural analysis of CrgA in POPC/POPG (4:1 mol/mol) liposome was performed by CD spectroscopy at 25 °C. The secondary structure content was analyzed and calculated using the CDPro program (63) and the spectrum was plotted as molar ellipticity ($[\theta]_m$) (in degree·square centimeters per decimole) (*SI Appendix*, Fig. S2*B*).

ssNMR Spectroscopy. Glass plate-supported OS preparation was described previously (59). OS ssNMR spectroscopy was performed on oriented lipid bilayer samples containing full-length CrgA with various labeling schemes (^{15}N uniform labeling, ^{15}N amino acid-specific labeling, and ^{15}N reverse labeling). For MAS ssNMR spectroscopy, 2D DARR (49, 50) experiments and 3D NCACX, NCOCX (64), and CAN(CO)CX (48) experiments were performed on ^{13}C, ^{15}N uniform-labeled and two reverse-labeled ^{13}C, ^{15}N CrgA samples.

Structure Calculation. Based on the OS and MAS ssNMR data (Table 1 and *SI Appendix*, Table S3), we calculated the structure of CrgA TM domain, largely following a previous protocol (34).

ACKNOWLEDGMENTS. We thank Dr. Y. Tian and Dr. F. Marassi from Sanford–Burnham Medical Research Institute for providing the Xplor-NIH script. We also thank P. L. Gor'kov for his design of the OS sample holder and the sample transfer base. This work was supported, in part, by NIH Grants AI073891, AI074805, and AI023007. The NMR experiments were performed at the National High Magnetic Field Laboratory supported by a National Science Foundation Cooperative Agreement (DMR-1157490) with the State of Florida.

1. Plocinski P, et al. (2011) Characterization of CrgA, a new partner of the *Mycobacterium tuberculosis* peptidoglycan polymerization complexes. *J Bacteriol* 193(13):3246–3256.
2. Del Sol R, Pitman A, Herron P, Dyson P (2003) The product of a developmental gene, *crgA*, that coordinates reproductive growth in *Streptomyces* belongs to a novel family of small actinomycete-specific proteins. *J Bacteriol* 185(22):6678–6685.
3. Del Sol R, Mullins JG, Grantcharova N, Flärdh K, Dyson P (2006) Influence of CrgA on assembly of the cell division protein FtsZ during development of *Streptomyces coelicolor*. *J Bacteriol* 188(4):1540–1550.
4. Kieser KJ, Rubin EJ (2014) How sisters grow apart: Mycobacterial growth and division. *Nat Rev Microbiol* 12(8):550–562.
5. Plocinski P, et al. (2012) *Mycobacterium tuberculosis* CwsA interacts with CrgA and Wag31, and the CrgA-CwsA complex is involved in peptidoglycan synthesis and cell shape determination. *J Bacteriol* 194(23):6398–6409.
6. Sureka K, et al. (2010) Novel role of phosphorylation-dependent interaction between FtsZ and FipA in mycobacterial cell division. *PLoS One* 5(1):e8590.
7. Chauhan A, et al. (2006) Interference of *Mycobacterium tuberculosis* cell division by Rv2719c, a cell wall hydrolase. *Mol Microbiol* 62(1):132–147.
8. Plocinska R, et al. (2012) Septal localization of the *Mycobacterium tuberculosis* MtrB sensor kinase promotes MtrA regulon expression. *J Biol Chem* 287(28): 23887–23899.
9. Hett EC, Rubin EJ (2008) Bacterial growth and cell division: A mycobacterial perspective. *Microbiol Mol Biol Rev* 72(1):126–156.
10. World Health Organization (2013) *Global Tuberculosis Report 2013* (World Health Organization, Geneva), pp 1–3.
11. Amaral L, Molnar J (2012) Potential therapy of multidrug-resistant and extremely drug-resistant tuberculosis with thioridazine. *In Vivo* 26(2):231–236.
12. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol* 305(3):567–580.
13. Li X, Romero P, Rani M, Dunker AK, Obradovic Z (1999) Predicting protein disorder for N-, C-, and internal regions. *Genome Inform Ser Workshop Genome Inform* 10:30–40.
14. Senes A, Engel DE, DeGrado WF (2004) Folding of helical membrane proteins: The role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol* 14(4):465–479.
15. MacKenzie KR, Prestegard JH, Engelman DM (1997) A transmembrane helix dimer: Structure and implications. *Science* 276(5309):131–133.
16. Dong H, Sharma M, Zhou HX, Cross TA (2012) Glycines: Role in α-helical membrane protein structures and a potential indicator of native conformation. *Biochemistry* 51(24):4779–4789.
17. Zhou HX, Cross TA (2013) Modeling the membrane environment has implications for membrane protein structure and function: Influenza A M2 protein. *Protein Sci* 22(4): 381–394.
18. Scheffers DJ, et al. (2007) Contribution of the FtsQ transmembrane segment to localization to the cell division site. *J Bacteriol* 189(20):7273–7280.
19. Chang G, Spencer RH, Lee AT, Barclay MT, Rees DC (1998) Structure of the MscL homolog from *Mycobacterium tuberculosis*: A gated mechanosensitive ion channel. *Science* 282(5397):2220–2226.
20. Page RC, Lee S, Moore JD, Opella SJ, Cross TA (2009) Backbone structure of a small helical integral membrane protein: A unique structural characterization. *Protein Sci* 18(1):134–146.
21. Mieczkowski C, Iavarone AT, Alber T (2008) Auto-activation mechanism of the *Mycobacterium tuberculosis* PknB receptor Ser/Thr kinase. *EMBO J* 27(23):3186–3197.
22. Mavrici D, et al. (2014) *Mycobacterium tuberculosis* FtsX extracellular domain activates the peptidoglycan hydrolase, RipC. *Proc Natl Acad Sci USA* 111(22):8037–8042.

23. Yao Y, Barghava N, Kim J, Niederweis M, Marassi FM (2012) Molecular structure and peptidoglycan recognition of *Mycobacterium tuberculosis* ArfA (Rv0899). *J Mol Biol* 416(2):208–220.
24. Wells RM, et al. (2013) Discovery of a siderophore export system essential for virulence of *Mycobacterium tuberculosis*. *PLoS Pathog* 9(1):e1003120.
25. Korepanova A, et al. (2005) Cloning and expression of multiple integral membrane proteins from *Mycobacterium tuberculosis* in *Escherichia coli*. *Protein Sci* 14(1): 148–158.
26. Zhou HX, Cross TA (2013) Influences of membrane mimetic environments on membrane protein structures. *Annu Rev Biophys* 42:361–392.
27. Cross TA, Murray DT, Watts A (2013) Helical membrane protein conformations and their environment. *Eur Biophys J* 42(10):731–755.
28. Miao Y, et al. (2012) M2 proton channel structural validation from full-length protein samples in synthetic bilayers and *E. coli* membranes. *Angew Chem Int Ed Engl* 51(33): 8383–8386.
29. Park SH, et al. (2012) Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* 491(7426):779–783.
30. Verardi R, Shi L, Traaseth NJ, Walsh N, Veglia G (2011) Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. *Proc Natl Acad Sci USA* 108(22):9101–9106.
31. Lu GJ, Tian Y, Vora N, Marassi FM, Opella SJ (2013) The structure of the mercury transporter MerF in phospholipid bilayers: A large conformational rearrangement results from N-terminal truncation. *J Am Chem Soc* 135(25):9299–9302.
32. Ahuja S, et al. (2013) A model of the membrane-bound cytochrome b5-cytochrome P450 complex from NMR and mutagenesis data. *J Biol Chem* 288(30):22080–22095.
33. Jean-Francois FL, et al. (2014) Binding of MgtR, a Salmonella transmembrane regulatory peptide, to MgtC, a *Mycobacterium tuberculosis* virulence factor: A structural study. *J Mol Biol* 426(2):436–446.
34. Sharma M, et al. (2010) Insight into the mechanism of the influenza A proton channel from a structure in a lipid bilayer. *Science* 330(6003):509–512.
35. Wang S, et al. (2013) Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat Methods* 10(10):1007–1012.
36. Sreerama N, Woody RW (2004) On the analysis of membrane protein circular dichroism spectra. *Protein Sci* 13(1):100–112.
37. Ramamoorthy A, Opella SJ (1995) Two-dimensional chemical shift/heteronuclear dipolar coupling spectra obtained with polarization inversion spin exchange at the magic angle and magic-angle sample spinning (PISEMAMAS). *Solid State Nucl Magn Reson* 4(6):387–392.
38. Wang J, et al. (2000) Imaging membrane protein helical wheels. *J Magn Reson* 144(1): 162–167.
39. Marassi FM, Opella SJ (2000) A solid-state NMR index of helical membrane protein structure and topology. *J Magn Reson* 144(1):150–155.
40. Murray DT, Hung I, Cross TA (2014) Assignment of oriented sample NMR resonances from a three transmembrane helix protein. *J Magn Reson* 240:34–44.
41. Page RC, Kim S, Cross TA (2008) Transmembrane helix uniformity examined by spectral mapping of torsion angles. *Structure* 16(5):787–797.
42. Mesleh MF, et al. (2003) Dipolar waves map the structure and topology of helices in membrane proteins. *J Am Chem Soc* 125(29):8928–8935.
43. Kovacs FA, Denny JK, Song Z, Quine JR, Cross TA (2000) Helix tilt of the M2 transmembrane peptide from influenza A virus: An intrinsic property. *J Mol Biol* 295(1): 117–125.
44. Braun P, von Heijne G (1999) The aromatic residues Trp and Phe have different effects on the positioning of a transmembrane helix in the microsomal membrane. *Biochemistry* 38(30):9778–9782.
45. von Heijne G (2006) Membrane-protein topology. *Nat Rev Mol Cell Biol* 7(12): 909–918.
46. Kim S, Cross TA (2002) Uniformity, ideality, and hydrogen bonds in transmembrane alpha-helices. *Biophys J* 83(4):2084–2095.
47. Igumenova TI, et al. (2004) Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J Am Chem Soc* 126(21):6720–6727.
48. Franks WT, Kloepper KD, Wylie BJ, Rienstra CM (2007) Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *J Biomol NMR* 39(2):107–131.
49. Takegoshi K, Nakamura S, Terao T (2001) C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett* 344(5-6):631–637.
50. Takegoshi K, Nakamura S, Terao T (2003) C-13-H-1 dipolar-driven C-13-C-13 recoupling without C-13 rf irradiation in nuclear magnetic resonance of rotating solids. *J Chem Phys* 118(5):2325–2341.
51. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160(1):65–73.
52. Walther D, Eisenhaber F, Argos P (1996) Principles of helix-helix packing in proteins: The helical lattice superposition model. *J Mol Biol* 255(3):536–553.
53. Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, Engelman DM (1992) Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry* 31(51): 12719–12725.
54. Javadpour MM, Eilers M, Groesbeek M, Smith SO (1999) Helix packing in polytopic membrane proteins: Role of glycine in transmembrane helix association. *Biophys J* 77(3):1609–1618.
55. Senes A, Ubarretxena-Belandia I, Engelman DM (2001) The Cα—H···O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci USA* 98(16):9056–9061.
56. Vijayakumar M, Qian H, Zhou HX (1999) Hydrogen bonds between short polar side chains and peptide backbone: Prevalence in proteins and effects on helix-forming propensities. *Proteins* 34(4):497–507.
57. Killian JA, von Heijne G (2000) How proteins adapt to a membrane-water interface. *Trends Biochem Sci* 25(9):429–434.
58. de Planque MR, et al. (2003) Interfacial anchor properties of tryptophan residues in transmembrane peptides can dominate over hydrophobic matching effects in peptide-lipid interactions. *Biochemistry* 42(18):5341–5348.
59. Das N, Murray DT, Cross TA (2013) Lipid bilayer preparations of membrane proteins for oriented and magic-angle spinning solid-state NMR samples. *Nat Protoc* 8(11): 2256–2270.
60. Höfer N, Aragão D, Caffrey M (2010) Crystallizing transmembrane peptides in lipidic mesophases. *Biophys J* 99(3):L23–L25.
61. Separovic F, Killian JA, Cotten M, Busath DD, Cross TA (2011) Modeling the membrane environment for membrane proteins. *Biophys J* 100(8):2073–2074, author reply 2075.
62. Oldfield CJ, Dunker AK (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem* 83:553–584.
63. Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287(2):252–260.
64. Igumenova TI, Wand AJ, McDermott AE (2004) Assignment of the backbone resonances for microcrystalline ubiquitin. *J Am Chem Soc* 126(16):5323–5331.

Structure of CrgA, a Cell Division Structural and Regulatory Protein from *Mycobacterium tuberculosis*, in Lipid Bilayers

Nabanita Das[a,b], Jian Dai[a,c], Ivan Hung[b], Malini Rajagopalan[d], Huan-Xiang Zhou[a,c], and Timothy A. Cross[a,b,e]

[a.] Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306

[b.] National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL 32310

[c.] Department of Physics, Florida State University, Tallahassee, FL 32306

[d.] The University of Texas Health Science Center at Tyler, Biomedical Research Branch, Tyler, TX 75708

[e.] Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL 32306

Supporting Information:

Supporting Text

Supporting Figures

Supporting Tables

Supporting Text

CrgA Protein Expression. The protocol for preparing the protein samples was similar to one described previously (1). Briefly, the *Mtb* (H37Rv strain) *crgA* gene was cloned into pET29b vector with a non-cleavable C-terminal His_6 tag. The plasmid was transformed into BL21(DE3)RP codon plus *E. coli* strains for large-scale protein expression. A single colony was picked to inoculate two 50 ml precultures of LB media with 50 µg/ml kanamycin antibiotic in each. Each preculture was grown overnight at 37 °C in a shaker incubator and transferred to two 1 L of LB media with the same kanamycin antibiotic (50 µg/ml). The 2 L of LB media were grown first at 37 ºC until the $O.D_{600}$ reached 1.0-2.0. Cells were harvested by centrifugation, washed twice with M9 media and transferred to 1 L of M9 media with stable isotopes (see below for media preparation). The M9 media was incubated for 30 minutes at 37 ºC in a shaker incubator followed by adding 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to induce. CrgA membrane protein overexpression. After the $O.D_{600}$ reached 4.0-5.0 the cells were harvested by centrifugation at 5000-6000 g for 10 minutes at room temperature and the cell pellet was resuspended in 20 ml of T80 buffer (40 mM Tris, pH 7.5 and 500 mM NaCl) and stored at -80 ºC.

For ^{13}C, ^{15}N reverse labeled protein production the cell culture conditions were similar. The cell pellet from 2 L LB media was transferred to 0.7 L of M9 media with 50 µg/ml kanamycin antibiotic at 37 °C. Cells were grown for 30 minutes in a shaker incubator and then the remaining 0.3 L of the M9 medium containing all the desired natural abundance amino acids was added to the 0.7 L of culture. After 5-7 min cells were induced with IPTG and harvested when the $O.D_{600}$ reached 4.0. The harvested cell pellet was resuspended in 20 ml of 40 mM Tris, pH 7.5 and 500 mM NaCl buffer and stored at -80ºC for further protein purification.

M9 Media Preparation. For ^{13}C, ^{15}N uniformly labeled CrgA protein preparation, 1 L M9 media was supplemented with 1 g of $^{15}NH_4Cl$ and 2 g of ^{13}C glucose.

For ^{15}N amino acid specific labeled CrgA protein preparation, 20 amino acids (natural abundance) except the one intended to be ^{15}N labeled was added to the 1 L M9 media. The quantity of the natural abundance amino acids per L of M9 media were as follows: 800 mg of Asp and Glu, 500 mg of Ala, Val, Leu, Ile, and 200 mg of each of the remaining amino acids. The amount of ^{15}N labeled amino acids was 200 mg/L of M9 media.

For ^{15}N reverse labeled protein the protocol is similar to that previously published by Griffin's group (2). 1 L M9 media containing 1 gm/L $^{15}NH_4Cl$ was divided into two portions of 0.7 L and 0.3 L. The 0.3 L portion was supplemented, in the amounts described above, with all 20 amino acids in natural abundance except those amino acids that were intended for labeling.

For ^{13}C, ^{15}N reverse labeled protein 1 L of M9 media was prepared, but here the media was supplemented with 1 g of $^{15}NH_4Cl$ and 2 g of ^{13}C glucose. The 0.3 L portion was again supplemented with all 20 amino acids in natural abundance except those amino acids that were intended for labeling.

CrgA Protein Purification. Details of the procedure for CrgA protein purification and of the buffer preparations were previously published (1, 3). Briefly, the harvested cells from 1 L of M9 media were thawed and mixed with 4 µL of benzonase (Novagene) and 0.25 mg of lysozyme at room temperature. Then 20% (vol/vol) Lysis buffer (20 mM Tris, pH 8.0, 1% TritonX-100, 8 M urea) for 60 ml of total whole cell lysate was added to the mixture. The ratio between T80 and Lysis buffer was 4:1 vol/vol. The solution was mixed for 30 minutes at room temperature and passed through a syringe with a 0.1 mm needle three times. 3% Empigen (vol/vol) was added to the solution and incubated overnight at 4 °C with gentle rotation. The solution was centrifuged at 230,000 g for 30-45 minutes at 8 ºC. This step removed most of the cellular debris in the pellet

while the supernatant contained detergent-solubilized membrane protein. The supernatant was transferred to a 5 ml His TrapTM FF Nickel Column (GE Healthcare) equilibrated with buffer 1 (300 mM NaCl, 20 mM Tris, pH 8.0, 40 mM Imidazole and 0.7% (v/v) Empigen) and purified by AKTA-Xpress FPLC (GE Healthcare). Two consecutive washing steps were performed. Washing step 1 was performed with 20 ml (4 times column volume) of buffer 1. Washing step 2 (also known as the detergent exchange step) was performed with buffer 2 (20 mM Tris, pH 8.0, 60 mM Imidazole and 0.2% (v/v) DPC, 100 mM NaCl). Finally, the protein was eluted with the elution buffer (20 mM Tris pH 8.0, 400 mM imidazole, 0.4% (v/v) DPC, and 100 mM NaCl) at room temperature. The elution fraction absorbance was monitored at 280 nm by Nanodrop. The yield of the purified CrgA membrane protein in DPC detergent was 25-30 mg/L of minimal media. The protein purity was checked by 12% SDS-PAGE gel (Fig. S2a).

CrgA Reconstitution in POPC/POPG Lipid Bilayer. Full-length CrgA was reconstituted in POPC/POPG (4:1 mol/mol) liposomes at pH 8.0 as previously described (3). For all OS samples, the protein to lipid molar ratio was 1:80 (mol/mol) while for MAS samples it was 1:30 (mol/mol). Briefly, a thinly prepared POPC/POPG lipid film was dissolved in 3 ml of 5 mM Tris-HCl (pH 8.0) and mixed with 300 µl of 20% SDS. The mixture was bath sonicated and 8-10 mg of purified CrgA in 0.2% DPC detergent was added. After brief mixing, methyl β-cyclodextrin (MβCD) was added to the solution (1:3 mol/mol detergent:MβCD). After 10 minutes of incubation with gentle rocking at room temperature, the mixture was centrifuged at 223,000 g for 3 hours at 8° C. The MβCD bound detergent was removed by suspending the proteoliposome pellet into 20 ml 5 mM Tris-HCl three times with subsequent centrifugation. The final proteoliposome pellet was harvested for later characterization.

Circular Dichroism Spectroscopy. Initial secondary structural analysis of CrgA in POPC/POPG (4:1 mol/mol) liposome was performed by CD spectroscopy at 25 °C. The reconstituted proteoliposome sample (protein to lipid molar ratio 1:30 mol/mol) was diluted with

5 mM Tris buffer (pH 8.0) and bath sonicated until completely clear. The CD spectrum of a sample containing 20 μM CrgA in 300 μl of total volume was collected on an AVIV 202 CD spectrometer equipped with 0.1 cm quartz cuvette. The recorded spectrum was the average of 3 scans from 260 nm to 190 nm in 0.5 nm increments with 1 sec integration time, with the baseline of a blank containing 5 mM Tris buffer (pH 8.0) and POPC/POPG liposome (4:1) subtracted. The secondary structure content was analyzed and calculated using the CDPro program (4) and the spectrum was plotted as molar ellipticity ($[\theta]_m$ (deg cm^2/ dmol) (Fig. S2b).

OS Glass plate supported Oriented sample preparation. Glass plate supported OS preparation was described previously (3). The final proteoliposome (1:80 molar ratio) pellet was dissolved in 5 mM Tris-HCl with 0.01% sodium azide (w/v). A homogeneous suspension was prepared by bath sonication of the pellet in a final volume of 1.2 ml. Approximately 32 μl of the solution was spread onto each of 35-40 glass slides (5.7 mm × 12.0 mm × 0.06-0.08 mm). Then the glass slides were partially dehydrated by incubating them in 16% relative humidity (RH) at 37 °C. After ~20 minutes, glass slides were removed from the incubator having a small opaque region remaining at the center of each slide. They were then stacked and incubated for 4-5 days at 96% RH and 37 °C. Once the stack was almost transparent, it was inserted into the sample cell with inner dimension 3 mm x 6.4 mm x 20 mm (New Era Enterprises Inc) and incubated again at 96% RH and 37 °C for a day or two. Sample clarity started from the center of the cell to the corners upon rehydration. Once the stack of slides became completely clear the sample cell was sealed by a cap and screw with bees wax and parafilm.

Solid state NMR spectroscopy. OS ssNMR spectroscopy was performed on oriented lipid-bilayer samples containing full-length CrgA with various labeling schemes (^{15}N uniform labeling, ^{15}N amino acid specific labeling, and ^{15}N reverse labeling). 1D cross polarization (CP) and 2D polarization inversion spin exchange at magic angle (PISEMA) (5) spectra were collected. All NMR experiments were performed in a Bruker Avance 600 MHz NMR spectrometer at 13 °C

using a home-built low-E static NMR probe (6). For all spectra ^{15}N chemical shifts were referenced to the ^{15}N signal of an aqueous solution of ^{15}N- labeled ammonium sulfate (5%, pH 3.1) at 26.8 ppm. The typical CP experimental conditions were ^{1}H 90° pulse length of 4 μs, ^{1}H and ^{15}N RF fields of 50 kHz, ^{1}H decoupling RF field of 62.5 kHz, recycle delay of 4 s, CP contact time of 1000 μs with 2000 scans, resulting in a total acquisition time of 2-3 hours per sample. Similar conditions were used for PISEMA, with 4000-5000 scans acquired for each of 28-32 increments in the dipolar coupling dimension. The orientational restraints were interpreted using a motionally averaged chemical shift tensor (σ_{11}=57.3, σ_{22}=81.2, σ_{33}=227.8ppm) and a motionally averaged ^{15}N-^{1}H dipolar interaction was set to 10.735kHz. In addition the relative orientation of σ_{33} and the $\nu_{||}$ component of the dipolar interaction was set at 17° (7).

For magic angle spinning (MAS) ssNMR spectroscopy, 2D dipolar assisted rotational resonance (DARR) (8, 9) experiments and 3D NCACX, NCOCX (10) and CAN(CO)CX (11) experiments were performed on ^{13}C, ^{15}N uniform labeled and two reverse labeled (^{13}C, ^{15}N $CrgA_{TIFSW}$, ^{13}C, ^{15}N $CrgA_{ILFYS}$) CrgA samples at 13 °C. All spectra were collected using a Bruker Avance 600 MHz NMR spectrometer with a 3.2 mm Low-E triple resonance MAS probe (6). ^{13}C spectra were referenced to the carbonyl carbon resonance of Glycine at 178.4 ppm with respect to tetramethylsilane. The sample spinning rate was controlled by a Bruker MAS unit at 12 kHz ± 3 Hz. The DARR experimental conditions were: ^{1}H 90° pulse length of 2.5 μs, (90-110%) linear ramped RF field on ^{1}H channel during CP, 50 kHz RF field on the ^{13}C channel during the cross polarization time of 1 ms, ^{1}H decoupling RF field of 100 kHz using the SPINAL64 decoupling sequence (12). Each DARR spectrum was collected with 128-256 scans for each increment, 1.5 - 2 s recycle delay with 2048/324 (t2/t1) complex data points, resulting in a total experimental time of 1-2 days. DARR mixing times of 30, 50, 100, 300, 500, 700 and 1000 ms were used. For 3D experiments, the CP conditions were similar to the DARR experiments. All 3D experiments had 30 ms DARR mixing time. The number of scans per increment was 256 for NCACX, 368 for

NCOCX and 448 for CANCOCX experiments. More scans were used for CANCOCX and NCOCX than the NCACX because the sensitivities of the 3D experiments follow the order NCACX > NCOCX > CANCOCX. Three spectra were collected using the following complex data points: NCACX experiment, 2048/60/48 (t3/t2/t1) points; NCOCX experiment, 3072/24/24 (t3/t2/t1) points; CAN(CO)CX experiment, 3072/20/42 (t3/t2/t1) points. The total acquisition times for single NCACX and NCOCX experiments were 7-10 days and that for CANCOCX was 12 days. All spectra were processed by Topspin version 2.1. Data analysis and spectral plotting were facilitated by Sparky version 3.114.

Structure Calculation. Using the ssNMR data (Table 1 and Table S3), CrgA TM domain conformations were generated by XPLOR-NIH in torsion-angle space (13). First, two ideal helices were constructed for the TM1 and TM2 sequences. Sixty starting models of the helical pairs were generated with randomized sidechain dihedral angles by MODELLER (14). For each starting model, 16 conformations were generated by simulated annealing in XPLOR-NIH, resulting in a total of 960 conformations. In generating each conformation, the pair of helices was equilibrated at 3,500 K for 10 ps, and simulated annealing was performed from 3,500 K to 20 K with a decrement of 10 K. The simulation was run for 0.2 ps and the force constants of restraint terms mentioned below were gradually increased from their initial values to the final values (see Table S5). This procedure was followed by 500 steps of energy minimization. In order to maintain the α-helical conformation, the backbone dihedral angles for TM1 (residues 31-49) and TM2 (residues 73-87) core regions were restrained during the simulation to flat-bottom harmonic potentials, with the minima of the right and left branches shifted ±15° from the ideal α-helix values (ϕ = -60°, ψ = -45°) (7). Also, 60% reduced force constants were applied to residues 88 to 92 when compared to the TM core regions, with ±30° shifts in the flat-well potential minima. The distance between *i* and *i* + 4 backbone hydrogen bonding atoms of the helices were also restrained to their respective ideal values (O-N at 3.0 Å and O-HN at 2.06 Å).

The 960 energy-minimized conformations were filtered according to the restraints from OS ssNMR on the tilt angles of the TM helices with respect to the *z*-axis. Specifically, conformations were retained only if the TM1 tilt angle was between 12 and 16° and the TM2 tilt angle was between 14 and 18°. The resulting 85 conformations were clustered according to their pairwise root-mean-square-deviations (RMSDs), and a representative conformation in the largest cluster (consisting of 51 conformations) was chosen for further refinement by restrained molecular dynamics simulations.

In restrained molecular dynamics simulations, the POPC/POPG (4:1 mol/mol) lipid bilayer was obtained from CHARMM-GUI (15), which contained 120 POPC, 30 POPG molecules to match the 4:1 mol/mol POPC/POPG ratio in the ssNMR samples and was solvated by 6000 water molecules followed by addition of Na^+ and Cl^- ions to give a salt concentration of 0.15 M. The hydrated lipid bilayer system was equilibrated for 50 ns under constant pressure of 1 atm and constant temperature of 310 K. The initial structure of CrgA TM domain was then embedded in the lipid bilayer. Overlapped lipid and water molecules were removed from the system, giving 109 POPC, 24 POPG and 5930 water molecules in the final system. Energy minimization was performed, followed by molecular dynamics simulations for 20 ns, with the protein backbone atoms fixed. After this, the ideal α helical secondary structure of the CrgA TM helix core regions (TM1 residues 31 to 49 and TM2 residues 73 to 87) were restrained by the hydrogen bond distances and dihedral angle restraints by harmonic potentials with 20.0 kcal/mol/Å and 200.0 kcal/mol/deg force constants, respectively. During this step, the PISEMA restraints were gradually introduced to maintain the ^{15}N anisotropic chemical shifts and ^{1}H-^{15}N dipolar couplings at their experimental values (16, 17). The simulations were continued for 7 ns with the force constants linearly ramped from 0.01 to 0.05 kcal/mol/ppm^2 and from 0.5 to 5.0 kcal/mol/kHz^2. The four interhelical distances were restrained to be under 8.0 Å using a force constant of 100.0 kcal/mol/Å^2. The final snapshot of the simulations was taken as the refined structure of CrgA TM domain.

The simulations used the CHARMM 27 protein force field (18) and CHARMM36 lipid force field (19), running NAMD 2.9 (20). The water model was TIP3P (21). During the simulations, constant temperature and pressure were maintained by Langevin dynamics with 1.0 ps^{-1} damping coefficient and the Nosé-Hoover Langevin piston method, respectively (22). A TCL force script was used to impose the PISEMA restraints. The distance module in the colvars package of NAMD was used to restrain the interhelical distances (23). The extra bonds module was used to restrain the backbone dihedral angles and hydrogen bonds. Long-range electrostatic interactions were treated with the particle-mesh Ewald sum method (24). The time step was 2 fs for the 50 ns equilibration of the lipid system and the 20 ns equilibration of the lipid-protein system, but reduced to 1 fs for the final 5 ns restrained simulations. In every time step bonded and nonbonded interactions were calculated and every second step the long-range electrostatic interactions were updated. Between 10-12 Å, van der Waals interactions were switched off. The nonbonded interaction pairlist was updated every 10 steps with a cutoff 14 Å.

References

1. Plocinski P*, et al.* (2011) Characterization of CrgA, a new partner of the Mycobacterium tuberculosis peptidoglycan polymerization complexes. *J Bacteriol* 193(13):3246-3256.
2. Andreas LB, Eddy MT, Chou JJ, & Griffin RG (2012) Magic-angle-spinning NMR of the drug resistant S31N M2 proton transporter from influenza A. *J Am Chem Soc* 134(17):7215-7218.
3. Das N, Murray DT, & Cross TA (2013) Lipid bilayer preparations of membrane proteins for oriented and magic-angle spinning solid-state NMR samples. *Nat Protoc* 8(11):2256-2270.
4. Sreerama N & Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287(2):252-260.

5. Wu CH, Ramamoorthy A, & Opella SJ (1994) High-resolution heteronuclear dipolar solid-state NMR spectroscopy. *J Magn Reson Ser A* 109(2):270-272.

6. Gor'kov PL*, et al.* (2007) Using low-E resonators to reduce RF heating in biological samples for static solid-state NMR up to 900 MHz. *J Magn Reson* 185(1):77-93.

7. Page RC, Kim S, & Cross TA (2008) Transmembrane helix uniformity examined by spectral mapping of torsion angles. *Structure* 16(5):787-797.

8. Takegoshi K, Nakamura S, & Terao T (2001) C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett* 344(5-6):631-637.

9. Takegoshi K, Nakamura S, & Terao T (2003) C-13-H-1 dipolar-driven C-13-C-13 recoupling without C-13 rf irradiation in nuclear magnetic resonance of rotating solids. *J Chem Phys* 118(5):2325-2341.

10. Igumenova TI, Wand AJ, & McDermott AE (2004) Assignment of the backbone resonances for microcrystalline ubiquitin. *J Am Chem Soc* 126(16):5323-5331.

11. Franks WT, Kloepper KD, Wylie BJ, & Rienstra CM (2007) Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *J Biomol NMR* 39(2):107-131.

12. Fung BM, Khitrin AK, & Ermolaev K (2000) An improved broadband decoupling sequence for liquid crystals and solids. *J Magn Reson* 142(1):97-101.

13. Schwieters CD, Kuszewski JJ, Tjandra N, & Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160(1):65-73.

14. Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234(3):779-815.

15. Jo S, Kim T, Iyer VG, & Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* 29(11):1859-1865.

16. Lee J, Chen J, Brooks CL, 3rd, & Im W (2008) Application of solid-state NMR restraint potentials in membrane protein modeling. *J Magn Reson* 193(1):68-76.

17. Sharma M*, et al.* (2010) Insight into the mechanism of the Influenza A proton channel from a structure in a lipid bilayer. *Science* 330(6003):509-512.

18. Mackerell AD, Jr., Feig M, & Brooks CL, 3rd (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Computl Chem* 25(11):1400-1415.

19. Klauda JB*, et al.* (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114(23):7830-7843.

20. Phillips JC*, et al.* (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781-1802.

21. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926-935.

22. Martyna GJ, Tobias DJ, & Klein ML (1994) Constant-pressure molecular dynamics algorithms. *J Chem Phys* 101:4177-4189.

23. Fiorin G, Klein ML, & Henin J (2013) Using collective variables to drive molecular dynamics simulations. *Mol Phys* 111(22-23):3345-3362.

24. York DM, Darden TA, & Pedersen LG (1993) The effect of long-range electrostatic interactions in simulations of macromolecular crystals - A comparison of the Ewald and trucated list methods. *J Chem Phys* 99:8345-8348.

Supporting Figures

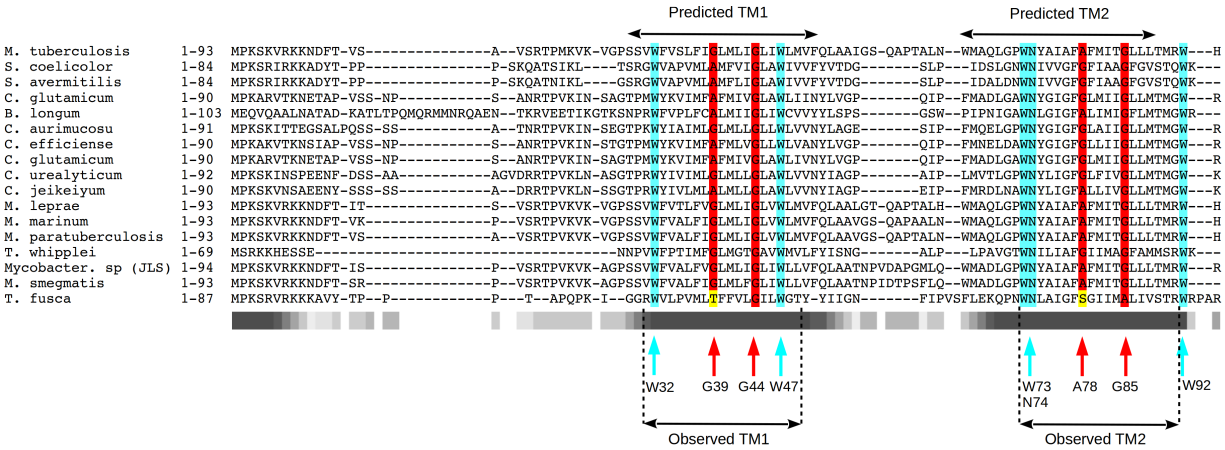

Figure S1. Sequence alignment and conservation of CrgA in 17 species. Conserved tryptophan and asparagine residues are shaded in cyan; conserved glycine and alanine residues are in red; and residues not conforming to the consensus are in yellow. *M. tuberculosis* CrgA transmembrane helical residues predicted by TMHMM are indicated by black arrows; helices determined experimentally are marked by vertical dashes. Conservation scores calculated by T-Coffee are displayed on a gray scale bar.

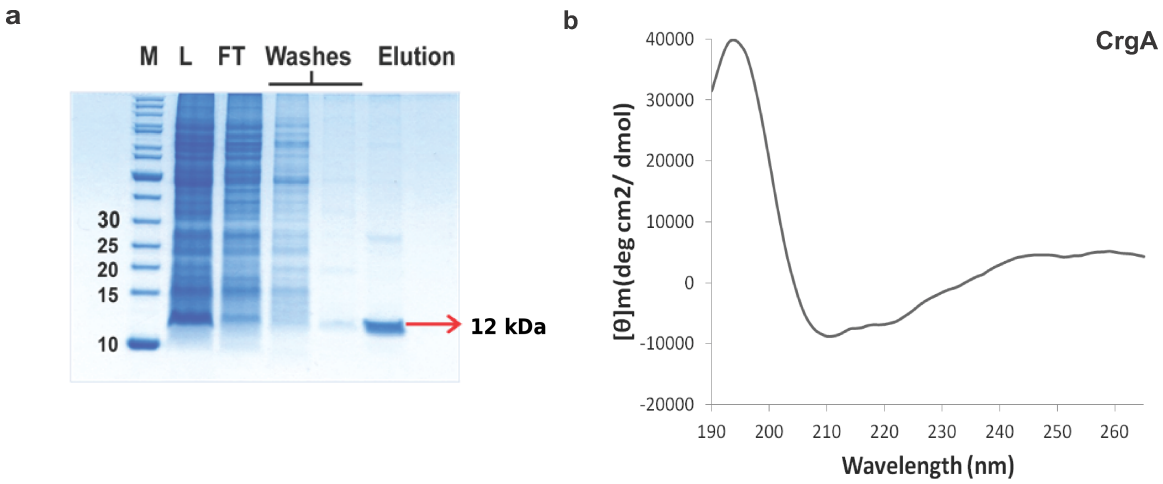

Figure S2. Expression/purification and CD spectroscopy of full-length *Mtb* CrgA. (a) 12% SDS-PAGE gel showing products of the expression and purification steps. M: molecular weight markers; L: whole cell lysate containing inclusion body and membrane fractions; FT: flow through from the nickel column; Washes: two consecutive washing steps; and Elution: protein elution from the nickel column. Molecular weight of the purified protein is indicated by red arrow. (b) Circular dichroism spectrum of CrgA in POPC/POPG (4:1 mol/mol) liposome at pH 7.0 and 25 °C, showing significant α-helix content. Protein to lipid molar ratio is 1:30.

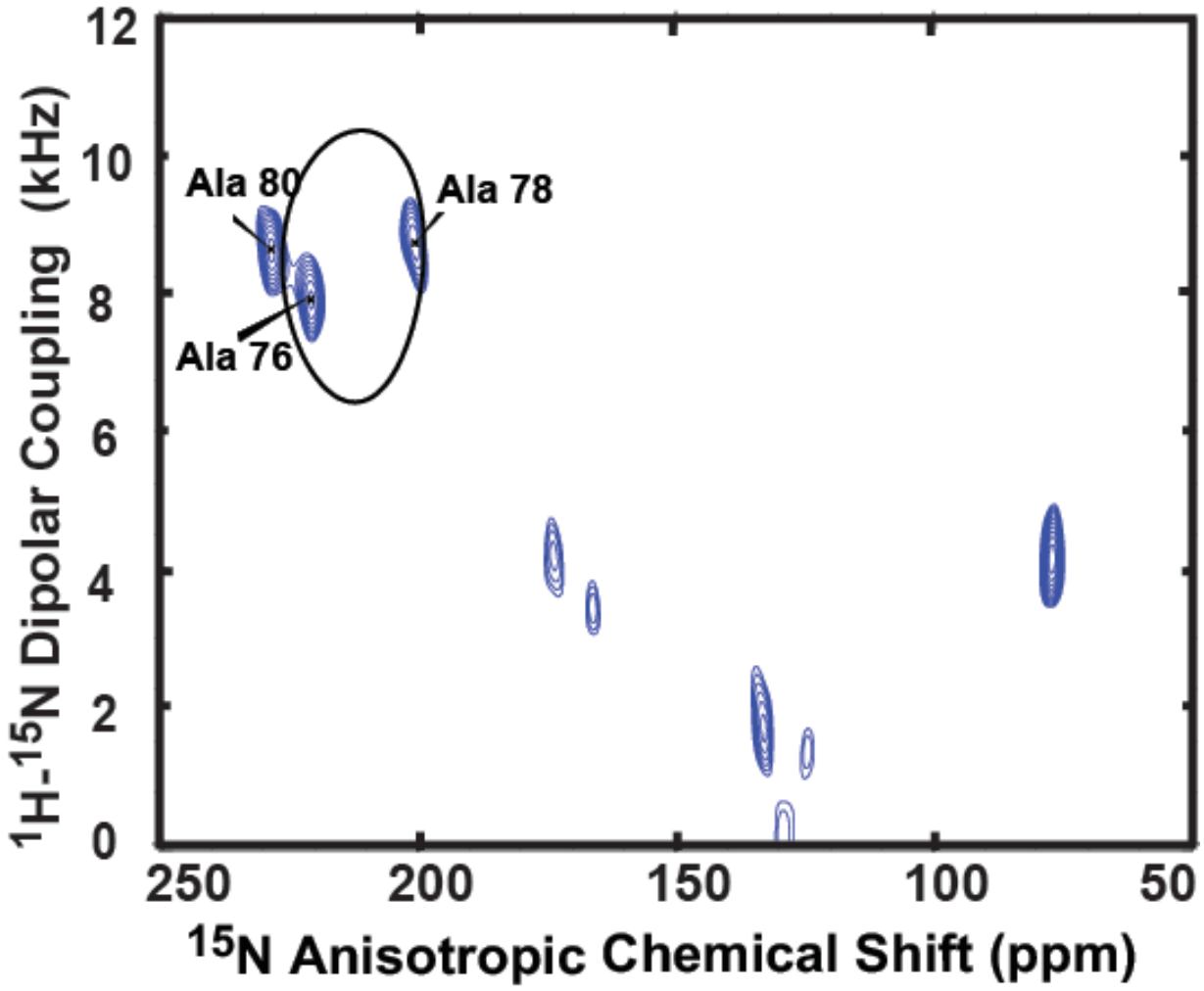

Figure S3. 2D PISEMA spectra of ^{15}N Ala labeled CrgA in POPC/POPG (4:1 mol/mol) lipid bilayers at 13° C, showing resonances of three transmembrane α-helix alanines (residue numbers identified and overlaid to a PISA wheel) and six other alanines (tentative assignments listed in Table S1).

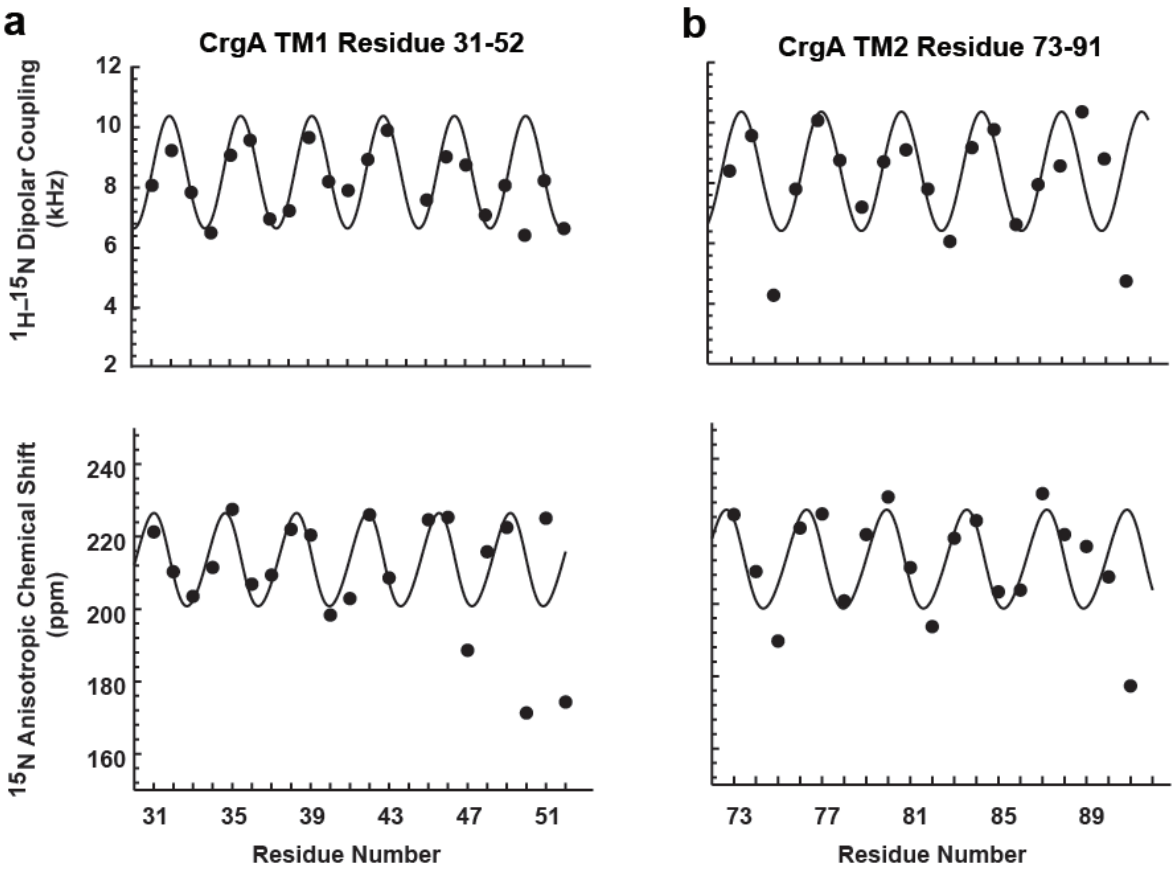

Figure S4 Assessment of *Mtb* CrgA helix uniformity by dipolar and chemical shift waves. (a) ^{1}H-^{15}N dipolar couplings and (b) ^{15}N anisotropic chemical shifts as functions of residue number. Left and right panels show results for TM1 and TM2, respectively; wave patterns are generated by fitting the PISEMA data to predictions of PISA wheels with 15° tilt for TM1 and 16° tilt for TM2.

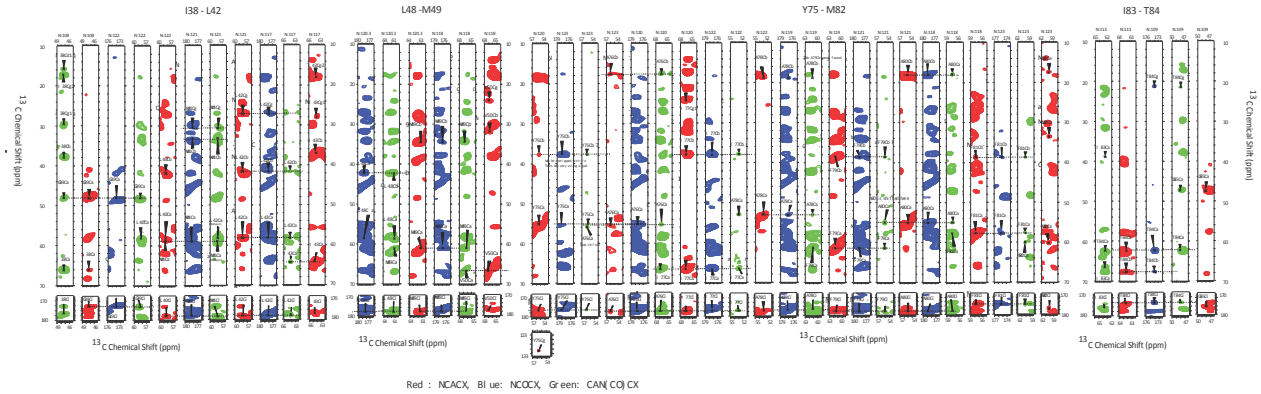

Figure S5. Strip plot display of 3D NCACX, NCOCX and CANCOCX spectra for partial backbone resonance assignment of CrgA in POPC/POPG lipid bilayers. The backbone walking process is shown by dotted lines connecting the resonances in three strip plots. All three types of spectra were collected with 30 ms DARR mixing time on a 600 MHz spectrometer at 13 °C.

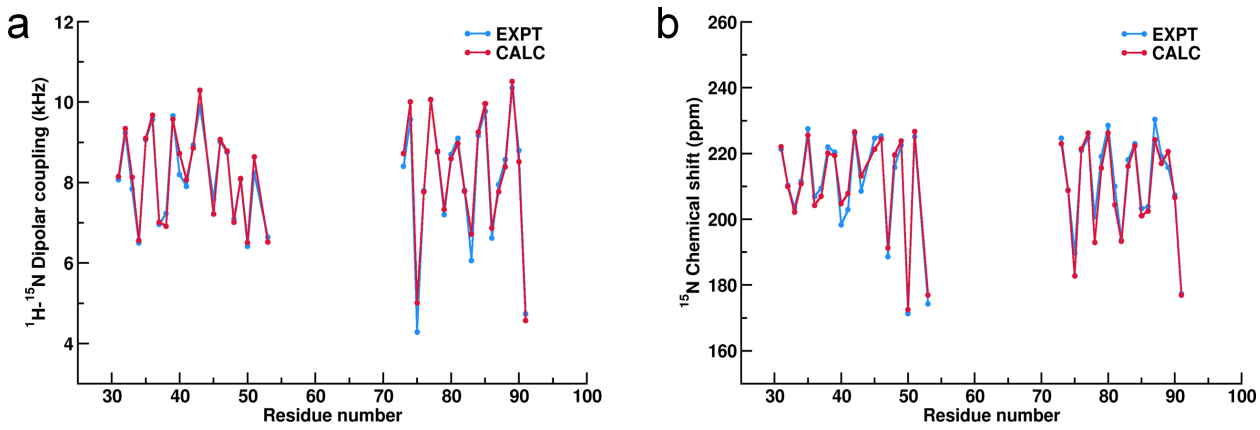

Figure S6. Comparison of experimental dipolar couplings and anisotropic chemical shifts with those calculated on the final structure. (a) Dipolar couplings. Root-mean-square-deviation (RMSD) between observed and calculated results is 0.26 kHz. (b) Anisotropic chemical shifts. RMSD is 3.2 ppm. In addition, all the four interhelical distances are within the upper bound of 8 Å.

Table S1: ^{1}H-^{15}N dipolar coupling (DC in kHz) and ^{15}N anisotropic chemical shift (ACS in ppm) data for the N- and C-terminal regions and the interhelical loop.

Residue	DC	ACS	Residue	DC	ACS
Ser4	0.0	110.0	Leu53	6.6	178.0
Val6	0.0	125.0	Ala54	4.2	173.0
Asn10	1.0	115.0	Ala55	3.6	166.0
Phe12	**0.0**	**120.0**	**Ile56**	**5.8**	**78.0**
Thr13	0.0	105.0	Gly57	5.2	80.0
Val14	0.0	125.0	Ser58	5.0	65.0-90.0
Ser15	1.0	110.0	Ala60	1.3	125.0
Ala16	0.0	130.0	Thr62	5.5	62.0
Val17	0.0	125.0	Ala63	4.3	78.0
Ser18	1.0	110.0	Leu64	5.5	65.0
Arg19	5.2	70.0	Asn65	5.0	80.0
Thr20	5.6	75.0	**Trp66**	**5.0**	**85.0**
Met22	3.9	105.0	Met67	5.4	88.0
Val24	5.4	70.0	Ala68	1.7	133.0
Val26	4.3	75.0	Leu70	0.0	105.0
Gly27	5.1	70.0	Gly71	0.0	110.0
Ser29	4.0	175.0			
Ser30	8.5	225.0	Trp92	6.0	212.0
			Leu94	0.0	105.0

Resonances for Phe12, Ile56 and Trp66 were unambiguously assigned and are highlighted in bold.

Table S2: ^{13}C ,^{15}N chemical shift assignments (ppm) for CrgA residues 38-43, 48-50, and 75-85.

Residue	**N**	**CO**	**Cα**	**Cβ**	**Cγ1/Cγ2**	**Cδ1/Cδ2**	**Cε1/Cε2**
Ile38	121.0	177.2	66.2	38.0	29.9/17.3	16.0	
Gly39	108.0	174.2	47.2				
Leu40	122.0	178.4	58.5	41.8			
Met41	121.0	178.6	59.0	33.3	30.6		
Leu42	121.0	178.4	58.3	41.6	27.3		
Ile43	117.0	177.2	64.2	37.0	27.8/18.1		
Leu48	121.5	178.6	58.5	42.2			
Met49	120.3	177.9	61.3	34.3	32.0		
Val50	118.0	177.0	66.4	31.9			
Tyr75	118.0	177.2	55.4	37.3	129.3	131.2	118.2
Ala76	123.0	177.7	55.7	17.7			
Ile77	120.0	177.3	66.49	38.12	24.2		
Ala78	122.0	177.5	53.0	18.6			
Phe79	119.0	177.6	61.5	38.3			
Ala80	121.0	178.2	55.1	18.0			
Phe81	118.0	175.4	57.5	39.3			
Met82	123.0	177.8	60.2	33.6			
Ile83	119.0	177.6	66.6	39.0			
Thr84	113.0	174.5	62.3	67.8			
Gly85	109.0	176.4	47.6				

Errors are estimated to be 0.2 ppm.

Table S3: Interhelical distances from DARR spectra.

TM1 nucleus	TM2 nucleus	Mixing time & labeling scheme	Distance upper bound (Å)
Gly39Cα	Thr84Cα	50 ms uniform labeling; 100 ms ILFYS reverse labeling	8.0
Leu42Cα	Ala80Cβ	700 & 1000 ms TIFSW reverse labeling	8.0
Met49Cα	Tyr75Cε1	300 ms TIFSW reverse labeling	8.0
Met49Cα	Ala78Cβ	300 ms TIFSW reverse labeling	8.0

Table S4: NMR restraints and statistics for structure calculation

NMR distance and dihedral restraints	
Distance restraints	
Total distances	60
Intra-residue	0
Inter-residue	60
Sequential (|*i* – *j*| = 1)	0
Medium-range (|*i* – *j*| < 4) (PISA wheel based[a])	56
Long-range (|*i* – *j*| > 5) (DARR based)	4
Intermolecular	0
Hydrogen bonds[a]	56
Total dihedral angle restraints (PISA wheel based[a])	78
Phi	39
Psi	39
Total orientational restraints from OS ssNMR	80
^{1}H-^{15}N dipolar couplings	40
^{15}N anisotropic chemical shifts	40
Structure statistics[b]	
Violations (mean and s.d.)	
Distance restraints (Å)	0.22 +/- 0.2
Dihedral angle restraints (º)	0.67 +/- 0.7
Max. dihedral angle violation (º)	3.1
Max. distance restraint violation (Å)	1.1
Deviations from idealized geometry	
Bond lengths (Å)	0.003 +/- 0.0005
Bond angles (º)	0.51 +/- 0.025
Impropers (º)	0.59 +/- 0.042
Average pairwise RMSD (Å)[c]	

Heavy	1.68 +/- 0.09
Backbone	0.38 +/- 0.03

[a]Ideal values of hydrogen bonding distances and torsion angles were used for these PISA wheel based restraints.

[b]Structure statistics were calculated from the 10 conformations out of XPLOR-NIH simulated annealing, after filtering, and with the lowest average pairwise RMSDs.

[c]Average pairwise RMSD was calculated among the 10 conformations using only the TM helix cores (residues 31 to 49 and 73 to 87).

Table S5: Force constants for restraint terms during simulated annealing.

Keyword	Restraint type	Force constant	
		Initial	**Final**
NOEPot1 (kcal/mol/Å^2)	Distance restraints of hydrogen bonds within helices	0.01	10.0
NOEPot2 (kcal/mol/Å^2)	Distance restraints between helices	0.01	40.0
CDIH (kcal/mol/rad^2)	Dihedral angle	200	1000
CSAPot (kcal/mol/ppm^2)	Anisotropy chemical shift	0.05	0.5
RDCPot (kcal/mol/kHz^2)	Dipolar coupling	0.1	2.0
RAMA (kcal/mol)	Knowledge-based dihedral angle database	0.02	2.0
ANGL (kcal/mol/rad^2)	Bond angle	0.4	1.0
IMPR (kcal/mol/rad^2)	Improper dihedral angle	0.1	1.0
VDW (kcal/mol/Å^2)	Atom-atom repulsion	0.004	4.0