DNA-nanotube-induced alignment of membrane proteins for NMR structure determination

Shawn M. Douglas*†‡, James J. Chou*§, and William M. Shih*‡§

Departments of *Biological Chemistry and Molecular Pharmacology and †Genetics, Harvard Medical School, Boston, MA 02115; and ‡Department of Cancer Biology, Dana–Farber Cancer Institute, Boston, MA 02115

Communicated by Tom A. Rapoport, Harvard Medical School, Boston, MA, January 31, 2007 (received for review December 15, 2006)

Membrane proteins are encoded by 20–35% of genes but represent <1% of known protein structures to date. Thus, improved methods for membrane-protein structure determination are of critical importance. Residual dipolar couplings (RDCs), commonly measured for biological macromolecules weakly aligned by liquid-crystalline media, are important global angular restraints for NMR structure determination. For α-helical membrane proteins >15 kDa in size, Nuclear-Overhauser effect-derived distance restraints are difficult to obtain, and RDCs could serve as the main reliable source of NMR structural information. In many of these cases, RDCs would enable full structure determination that otherwise would be impossible. However, none of the existing liquid-crystalline media used to align water-soluble proteins are compatible with the detergents required to solubilize membrane proteins. We report the design and construction of a detergent-resistant liquid crystal of 0.8-μm-long DNA-nanotubes that can be used to induce weak alignment of membrane proteins. The nanotubes are heterodimers of 0.4-μm-long six-helix bundles each self-assembled from a 7.3-kb scaffold strand and >170 short oligonucleotide staple strands. We show that the DNA-nanotube liquid crystal enables the accurate measurement of backbone N_H and $C_\alpha H_\alpha$ RDCs for the detergent-reconstituted ζ-ζ transmembrane domain of the T cell receptor. The measured RDCs validate the high-resolution structure of this transmembrane dimer. We anticipate that this medium will extend the advantages of weak alignment to NMR structure determination of a broad range of detergent-solubilized membrane proteins.

dipolar couplings | nanotechnology | liquid crystal | scaffolded origami

The general applicability of solution NMR spectroscopy to structural characterization of intact α-helical membrane proteins has been demonstrated by a number of recent studies. Examples include the structure determination of the 15-kDa Mistic protein (1) and the 30-kDa pentameric phospholamban (2) as well as the complete assignment of backbone resonances and secondary structures of the 44-kDa trimeric diacylglycerol kinase (3) and the 68-kDa tetrameric KcsA potassium channel (4). Despite such progress, full-scale structure determination of α-helical membrane proteins remains challenging and rare. Because of the large fraction of methyl-bearing residues in membrane proteins and of the added molecular weight of detergent micelles, the low chemical-shift dispersion of α-helical proteins is obscured by resonance overlap and line broadening, making assignment of side-chain methyl resonances extremely difficult. Without side-chain chemical shifts, it is impossible to obtain a sufficient number of long-range Nuclear-Overhauser effect-derived distance restraints for folding secondary segments into the correct tertiary structure.

Dipolar couplings between pairs of nuclei measured from weakly aligned proteins in solution provide global orientation restraints important for molecular structure determination and validation by NMR (5–8). Development of alignment media for accurate residual dipolar coupling (RDC) measurements from α-helical membrane proteins would enhance significantly the capability of solution NMR in structure determination of this important class of targets. The most effective method for weak alignment involves mixing the protein of interest with large particles that form stable liquid crystals at low concentration ($\approx$1.5–5% wt/vol) (8). Tjandra and Bax (9) first demonstrated this method using 1,2-dimyristoyl-sn-glycero-3-phosphocholine/1,2-dicaproyl-sn-glycero-3-phosphocholine (DMPC/DHPC)-bicelle liquid crystals. Subsequently, a number of different liquid crystals, including filamentous phage particles (10, 11), ternary mixtures of cetylpyridinium Cl/Br, hexanol, and sodium Cl/Br (12, 13), binary mixtures of polyethylene glycol and hexanol (14), and cellulose crystallites (15), have been found to be suitable for aligning water-soluble proteins. However, none could be applied to membrane proteins because of incompatibility with the zwitterionic or anionic detergents typically used to solubilize membrane proteins for structural study. Ma and Opella (16) introduced lanthanide ions to an "EF-hand" attached to an HIV-1 membrane protein embedded in lipid micelles that enabled measurement of RDCs, but this method requires modification of proteins that may cause problems for expression and reconstitution of membrane proteins. The only other method currently available for weak alignment of membrane proteins involves the use of strained (radially or axially compressed) polyacrylamide gels (17–19). But dissolving protein–micelle complexes to high concentration in gels is notoriously difficult because of the inhomogeneous pore size of randomly cross-linked gel matrices (20), thus the measured RDCs are of limited accuracy.

Liquid crystals of Pf1 phage, which are rod-like particles that are 7 nm in diameter and 2 μm in length (21), are often the first material tested for alignment of soluble proteins. The structural rigidity and negative-charge surface density (22) of Pf1 phage particles allow them to form a stable and useful liquid crystal at low concentrations (10); however, Pf1 is not compatible with detergents. At 20 mg·ml^{-1}, Pf1 is not liquid-crystalline in modest concentrations of detergent, e.g., ~50 mM DPC (J.J.C., unpublished results). Structural DNA nanotechnology can be used for the design and construction of objects of various shapes (23, 24) that, by virtue of being built from DNA, generally are resistant to detergents. In this article, we describe the design and characterization of a DNA structure mimicking the shape and size of

Author contributions: S.M.D., J.J.C., and W.M.S. designed research, performed research, analyzed data, and wrote the paper.

Conflict of interest statement: S.M.D. and W.M.S. declare competing financial interests. A provisional patent entitled "Nucleic-acid-nanotube liquid crystals and use for NMR structure determination of detergent-solubilized membrane proteins" was filed on April 21, 2006 on behalf of the Dana–Farber Cancer Institute by Edwards Angell Palmer & Dodge LLP, listing S.M.D. and W.M.S. as coinventors. A planned update to the patent filing will add J.J.C. of Harvard Medical School as a coinventor.

Abbreviations: RDC, residual dipolar coupling; SVD, singular value decomposition.

See Commentary on page 6502.

§To whom correspondence may be addressed. E-mail: william_shih@dfci.harvard.edu or james_chou@hms.harvard.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0700930104/DC1.

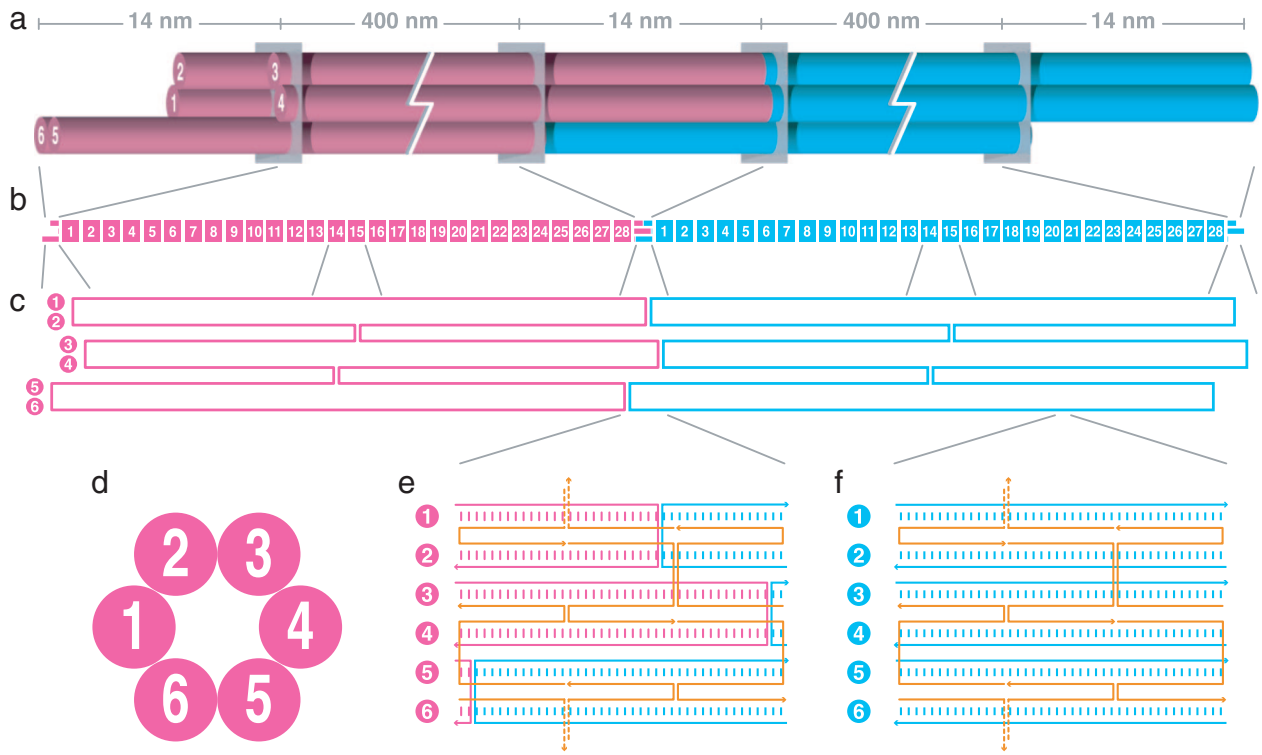

Fig. 1. DNA-nanotube design. (*a*) 3D cartoon view. The front and rear monomers are folded separately and then combined to form heterodimers. (*b*) Segment diagram. Each monomer consists of 28 segments of 42-bp length as well as a head and tail segment on each end. (*c*) Scaffold-only schematic view. Each monomer consists of a modified M13-bacteriophage single-stranded DNA genome of 7,308 bases in length serving as a ''scaffold'' and 168 DNA strands of 42-bp length acting as ''staples'' that constrict the scaffold to the target structure. (*d*) Cross-sectional view. (*e*) Scaffold-plus-staples schematic view of the heterodimer junction. There are two 42-bp staple strands (orange) per 14-bp subsegment. (*f*) Scaffold-plus-staples schematic view of a typical 42-bp segment. The crossover pattern of six staple strands repeats for every 42-bp segment along the length of the nanotube.

a filamentous phage particle. We used a liquid crystal of these DNA structures to align the ζ-ζ transmembrane domain of the T cell receptor, thereby enabling the accurate measurement of RDCs that are in excellent agreement with the previously determined structure. Our results clearly demonstrate the feasibility of a DNA-nanotechnology-based approach for resolving a major bottleneck in membrane-protein NMR spectroscopy.

Results and Discussion

To achieve a Pf1-like DNA structure, a six-helix bundle DNA-nanotube architecture (25) was adopted. This design resembles a parallel array of six double helices for which every set of three adjacent helices frames a dihedral angle of 120° (Fig. 1 *a* and *d*). Adjacent double helices are held together by Holliday-junction crossovers that occur every 42 bp (Fig. 1*f*). Toward building nanotubes of 0.8-μm uniform length, an assembly strategy was conceived to adapt the scaffolded DNA origami method (26) for the self-assembly of 0.4-μm monomers of two types (''front'' and ''rear''). For each monomer, a 7,308-base, M13-derived single-stranded circle of DNA is used as a ''scaffold,'' and 168 single strands of DNA of length 42 bases, programmed with complementarity to three separate 14-base regions of the scaffold, are used as ''staples'' (Fig. 1*f*). The staples self-assemble with the scaffold into the shape of six parallel double helices curled into a tube.

A monomer can be conceptualized as a series of 28 pseudorepeat segments, each consisting of six parallel double helices that are 42 bp long, flanked by jagged overhangs on either end of the object (Fig. 1*b*). Each segment can be conceptualized as a series of three subsegments, for which every double helix is 14 bp long (Fig. 1*f*). Six of the twelve strands of a subsegment are provided by the scaffold strand, three are provided by one staple strand, and three by another staple strand. Adjacent subsegments are related by 120-degree screw pseudosymmetry. The scaffold generally does not cross over between helices, except for four times in the middle of each monomer to produce a ''seam,'' and three times on each monomer end (Fig. 1*c*). The inclusion of a seam in the design allows for the linkage of monomers in a head-to-tail fashion instead of in a head-to-head fashion, as is evident from consideration of the polarity of the scaffold strand within each double helix [see supporting information (SI) Fig. 4*e*]. Three extra staple strands block the head of the front monomer, and four extra staple strands block the tail of the rear monomer (SI Fig. 4 *a* and *d*). To facilitate heterodimerization, three extra staple strands with unpaired bases decorate the tail of the front monomer, and three extra staple strands with unpaired bases decorate the head of the rear monomer (SI Fig. 4 *b* and *c*).

A computer program (see *SI Appendix 1*) was written to generate staple strand sequences (see *SI Appendix 2*) given the sequence of the scaffold. The use of two cyclic permutations of the scaffold sequence as input to the program generates independent sets of staple-strand sequences for folding two different monomer nanotubes. Therefore, copies of the same scaffold molecule can be used to generate two chemically distinct species. Front and rear monomers were folded in separate chambers by heat denaturation, followed by cooling for renaturation (see *SI Methods*). The majority of DNA objects migrate as a single band in agarose-gel electrophoresis (Fig. 2*a*). This population presumably represents well formed nanotube monomers, whereas slower migrating species apparent on the gel presumably represent misfolded or multimerized structures. This mixture was analyzed by

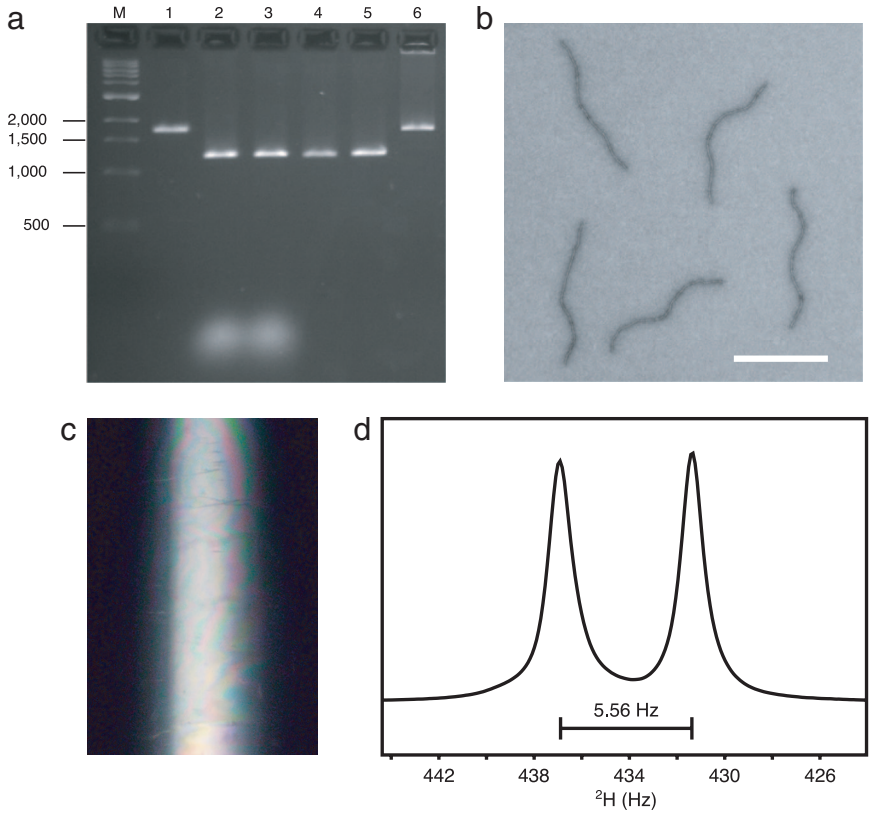

Fig. 2. Characterization of DNA-nanotubes. (*a*) Gel-shift analysis of folding and heterodimerization of DNA-nanotubes. Lane M, marker lane with DNA size standards (number of base pairs indicated at the left); lane 1, M13-derived (p7308) ssDNA scaffold; lanes 2 and 3, front and rear folded monomers (scaffold plus staples); lanes 4 and 5, front and rear monomers after PEG fractionation; lane 6, heterodimers, after 2-h incubation of mixed monomers at 37°C. Samples were electrophoresed in a 2% agarose gel containing 11 mM $MgCl_2$, 0.5 $\mu g \cdot ml^{-1}$ ethidium bromide, 45 mM Tris base, 45 mM boric acid, and 1 mM EDTA (pH 8.0). (*b*) Negative-stain electron micrographs of DNA-nanotube heterodimers. (Scale bar, 500 nm.) Samples were stained with 0.7% uranyl formate and imaged on a Tecnai G2 Spirit BioTWIN microscope. (*c*) Photograph of birefringence exhibited between crossed polarizers by DNA-nanotube dimers at 28 $mg \cdot ml^{-1}$ in a glass NMR tube. (*d*) D ^2H spectrum of a 10 mM NaH_2PO_4, 10 mM $MgCl_2$, 90% H_2O/10% D_2O, 100 mM DPC sample containing 28 $mg \cdot ml^{-1}$ DNA-nanotube heterodimers. The liquid crystal was aligned in an 11.4 T magnetic field and yielded ^2H quadrupolar splitting of 5.56 Hz.

using negative-stain electron microscopy, and the results are consistent with a large fraction of intact nanotubes of length 402 ± 6 nm (SI Fig. 5 *a*, *c*, and *e*). This measured length is in good agreement with the predicted length of 400 nm for 28 segments that are 42 bp long, assuming a rise of 0.34 nm per base pair.

The front and rear monomers were mixed to self-assemble heterodimers (Fig. 1 *a–c* and *e*). The joining of the tail of the front monomer to the head of the rear monomer should generate a 42-bp pseudorepeat segment (Fig. 1*e*). In this segment, all six staple strands bridge the two scaffolds, although by varying numbers of base pairs. In total, a net 62 bp must be broken to sever the linkage between successfully heterodimerized monomers. Agarose-gel electrophoresis of heterodimers assembled from the two monomers indicates that the majority of DNA objects migrate as a single band, although some misfolded objects are evident, as are a small population of monomeric nanotubes (Fig. 2*a*). Negative-stain electron microscopy revealed nanotubes of length 813 ± 9 nm (SI Fig. 5 *b*, *d*, and *f*). This measured length agrees well with the predicted length of 814 nm for 57 segments that are 42 bp long.

A low-salt, aqueous suspension of DNA-nanotube heterodimers at a concentration of 28 $mg \cdot ml^{-1}$ forms a stable liquid crystal, as indicated by strong birefringence observed through crossed polarizers (Fig. 2*c*). When the liquid crystal is aligned in an 11.4 T magnetic field in the presence of 100 mM dodecylphosphocholine (DPC) and 10% D_2O, the weakly oriented HDO yields ^2H quadrupolar splitting of 5.56 Hz (Fig. 2d). The nanotubes were tested for weak alignment of the transmembrane (TM) domain (residue 7–39) of the ζ-ζ chain of the T cell receptor complex reconstituted in mixed DPC/SDS detergent micelles (SI Fig. 6). The measured ^1H-^{15}N and $^1H_\alpha$-$^{13}C_\alpha$ RDCs agree very well with the known NMR structure of the ζ-ζ TM domain (27), with a correlation coefficient of the singular value decomposition (SVD) fit, R_{SVD}, of 0.98, or a free quality factor, Q_{free}, of 16% (Fig. 3*a*). The magnitude of the alignment tensor, D_a, is 9.9 Hz (normalized to D_{NH}), which is ideal for RDC measurement and structure calculation. In addition, the axis of C_2 rotational symmetry of ζ-ζ is parallel to the largest principal axis, A_{zz}, of the alignment tensor (Fig. 3*b*). This result is expected from the rotational averaging of the dimeric complex around its C_2 axis in the aligned medium.

The DNA-nanotube liquid crystal also was used to enable measurement of RDCs for a truncated version of BM2 (residues 26–109), which includes a 20-residue membrane anchor and a soluble coiled-coil tetramerization domain, reconstituted in 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rao-(1-glycerol)] (LMPG) detergent micelles (SI Fig. 7). BM2 is a pH-gated, proton-selective channel from influenza B virus. It is better known as the functional homolog of the well characterized M2 proton channel from influenza A virus. Like M2, BM2 plays a central role in equilibrating pH between compartments and is required in the life cycle of the virus (28, 29). The structure of BM2 is not known.

This alignment medium should be broadly useful for providing global structural restraints in solution NMR studies of mem-

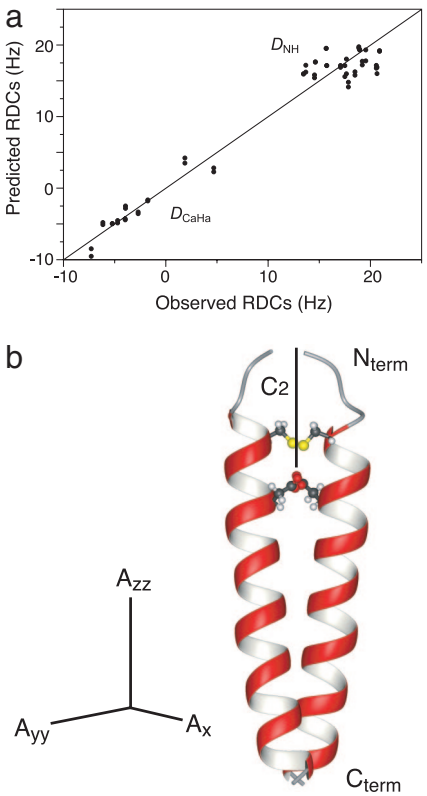

Fig. 3. Analysis of RDCs measured for detergent-reconstituted transmembrane domain of the ζ chain (residues 7–39) of the T cell receptor complex, weakly aligned in 28 mg·ml^{-1} DNA-nanotube. (*a*) Correlation between observed backbone RDCs (normalized to D_{NH}) and RDCs predicted for the known NMR structure of ζ-ζ TM domain (PDB ID code 2HAC) by using an alignment tensor obtained from the SVD fit. The correlation coefficient R_{SVD} is 0.98, and Q_{free} is 16%. (*b*) Principal axes of the alignment tensor relative to 2HAC: A_{zz}, −19.8 Hz; A_{yy}, 15.2 Hz; A_{xx}, 4.6 Hz; D_{aNH}, −9.9 Hz; and rhombicity R, 0.357. The solid line in the ribbon structure of the ζ-ζ TM domain represents the axis of C_2 rotational symmetry.

brane proteins. Because a large number of helical membrane proteins of great biomedical interest are between 20 and 30 kDa, well below the current size limitation of solution NMR spectroscopy, new experimental systems for obtaining NMR structural information in the presence of detergents are of fundamental importance. DNA nanotechnology, which affords versatile molecular design and subnanometer-scale precision, has been pursued as a route toward building host lattices to position guest macromolecules for crystallographic structural studies (25, 30–32). The present study uses solution NMR instead of crystallographic methods to validate the potential of DNA nanotechnology for imposing order on target macromolecules to acquire atomic-resolution structural information.

Methods

DNA-nanotubes were designed and generated as described in *SI Methods* and *SI Appendices 1* and *2*.

Negative Stain Electron Microscopy. After the dimerization step, DNA-nanotube dimers were diluted to 1 nM concentration and prepared for imaging by negative stain with 0.7% uranyl formate (Pfaltz & Bauer, Waterbury, CT) as described (33). Gilder fine bar grids (Ted Pella, Redding, OR), 400 mesh, 3.05 mm o.d., were used. Imaging was performed on a Tecnai G2 Spirit BioTWIN (FEI, Hillsboro, OR).

NMR Sample Preparation. The uniform 15N-, 13C-labeled TM domain (residues 7–39) of the ζ chain of the T cell receptor complex was expressed and purified as described (27). The TM dimer was reconstituted in mixed detergent micelles consisting of a 5:1 molar ratio of DPC/SDS. The aligned sample was made by manual mixing of 250 μl of a 0.5 mM protein solution with 250 μl of 28 mg/ml DNA-nanotube liquid crystal and subsequently concentrated back to ≈250 μl in a YM-50 Centricon (Millipore, Billerica, MA). The final sample, which consisted of ≈28 mg/ml nanotubes, 0.5 mM protein (monomer concentration), 150 mM DPC, 30 mM SDS, 20 mM sodium phosphate (pH 7.0), and 5% D_2O, was loaded to a microcell (Shigemi, Allison Park, PA) for NMR measurement.

NMR Measurements. All NMR experiments were performed on spectrometers (Bruker, Billerica, MA) equipped with cryogenic TXI probes at 30°C. The RDCs were obtained by subtracting J of the unaligned sample from $J + D$ of the aligned sample. The sign of dipolar couplings follows the convention that $|^1J_{NH} + {}^1D_{NH}| < 90$ Hz when $^1D_{NH}$ is positive. The ^{1}H-^{15}N RDCs were obtained from $^1J_{NH}/2$ and $(^1J_{NH} + {}^1D_{NH})/2$, which were measured at 600 MHz (^{1}H frequency) by interleaving a regular gradient-enhanced HSQC and a gradient-selected TROSY, both acquired with 80 ms of ^{15}N evolution. The $^1H_\alpha$-$^{13}C_\alpha$ RDCs ($^1D_{C_\alpha H_\alpha}$) were measured at 500 MHz (^{1}H frequency) by using a 2D CACONH quantitative $^1J_{C_\alpha H_\alpha}$ experiment with interleaved spectra recorded at $^1J_{C_\alpha H_\alpha}$ modulation times of 1.83, 3.63, and 7.12 ms. This experiment was modified from the 3D CBCACONH quantitative J_{CH} experiment (34) used primarily for measuring protein side-chain $^1H_\beta$-$^{13}C_\beta$ RDCs. The CACONH was optimized for measuring the backbone $^1H_\alpha$-$^{13}C_\alpha$ RDCs only. The frequency-labeled dimensions in this experiment are ^1H_N (direct) and ^{15}N (indirect). Because of the limited sensitivity of the CACONH experiment, $^1D_{C_\alpha H_\alpha}$ couplings were extracted from those resonances in the reference spectra ($^1J_{C_\alpha H_\alpha}$ modulation time of 3.63 ms) with S/N ratio >15. Moreover, overlapped resonances were not used. Because the ζ-ζ TM domain is a homodimer obeying 2-fold rotational symmetry, the same RDCs are assigned to both subunits.

We thank Payal Pallavi, Eva Volf, and Xingping Su for assistance with preparation of phage DNA; Matthew Call for preparation of the ζ-ζ NMR sample; and J. Markson for helpful comments on the manuscript. This work was supported by the Claudia Adams Barr Program in Innovative Basic Cancer Research (W.M.S.) and by National Institutes of Health Grant AI067438 (to J.J.C.). S.M.D. is supported by the Harvard Biophysics Program, and J.J.C. is supported by the Pew Scholars Program in the Biomedical Sciences.

1. Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S (2005) *Science* 307:1317–1321.
2. Oxenoid K, Chou JJ (2005) *Proc Natl Acad Sci USA* 102:10870–10875.
3. Oxenoid K, Kim HJ, Jacob J, Sonnichsen FD, Sanders CR (2004) *J Am Chem Soc* 126:5048–5049.
4. Chill JH, Louis JM, Miller C, Bax A (2006) *Protein Sci* 15:684–698.
5. Saupe A, Englert G (1963) *Phys Rev Lett* 11:462–464.
6. Tolman J, Flanagan J, Kennedy M, Prestegard J (1995) *Proc Natl Acad Sci USA* 92:9279–9283.
7. Tjandra N, Omichinski JG, Gronenborn AM, Clore GM, Bax A (1997) *Nat Struct Biol* 4:732–738.
8. Bax A, Kontaxis G, Tjandra N (2001) *Methods Enzymol* 339:127–174.
9. Tjandra N, Bax A (1997) *Science* 278:1111–1114.
10. Hansen MR, Mueller L, Pardi A (1998) *Nat Struct Biol* 5:1065–1074.
11. Clore GM, Starich MR, Gronenborn AM (1998) *J Am Chem Soc* 120:10571–10572.
12. Prosser RS, Losonczi JA, Shiyanovskaya IV (1998) *J Am Chem Soc* 120:11010–11011.
13. Barrientos LG, Dolan C, Gronenborn AM (2000) *J Biomol NMR* 16:329–337.
14. Ruckert M, Otting G (2000) *J Am Chem Soc* 122:7793–7797.
15. Fleming K, Gray D, Prasannan S, Matthews S (2000) *J Am Chem Soc* 122:5224–5225.
16. Ma C, Opella SJ (2000) *J Magn Reson* 146:381–384.

17. Tycko R, Blanco FJ, Ishii Y (2000) *J Am Chem Soc* 122:9340–9341.
18. Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S (2000) *J Biomol NMR* 18:303–309.
19. Chou JJ, Gaemers S, Howder B, Louis JM, Bax A (2001) *J Biomol NMR* 21:377–382.
20. Zhang XZ, Yang YY, Chung TS, Ma KX (2001) *Langmuir* 17:6094–6099.
21. Marvin DA (1998) *Curr Opin Struct Biol* 8:150–158.
22. Zimmermann K, Hagedorn H, Heuck CC, Hinrichsen M, Ludwig H (1986) *J Biol Chem* 261:1653–1655.
23. Seeman NC (2003) *Nature* 421:427–431.
24. Seeman NC (2005) *Methods Mol Biol* 303:143–166.
25. Mathieu F, Liao S, Kopatsch J, Wang T, Mao C, Seeman NC (2005) *Nano Lett* 5:661–665.
26. Rothemund PW (2006) *Nature* 440:297–302.
27. Call ME, Schnell JR, Xu C, Lutz RA, Chou JJ, Wucherpfennig KW (2006) *Cell* 127:355–368.
28. Mould JA, Paterson RG, Takeda M, Ohigashi Y, Venkataraman P, Lamb RA, Pinto LH (2003) *Dev Cell* 5:175–184.
29. Paterson RG, Takeda M, Ohigashi Y, Pinto LH, Lamb RA (2003) *Virology* 306:7–17.
30. Seeman NC (1982) *J Theor Biol* 99:237–247.
31. Paukstelis PJ, Nowakowski J, Birktoft JJ, Seeman NC (2004) *Chem Biol* 11:1119–1126.
32. Malo J, Mitchell JC, Venien-Bryan C, Harris JR, Wille H, Sherratt DJ, Turberfield AJ (2005) *Angew Chem Int Ed Engl* 44:3057–3061.
33. Ohi M, Cheng Y, Walz T (2004) *Biol Proc Online* 6:23–34.
34. Chou JJ, Bax A (2001) *J Am Chem Soc* 123:3844–3845.