Tilted, uninterrupted, monomeric HIV-1 gp41 transmembrane helix from residual dipolar couplings

Sai Chaitanya Chiliveri,[1] John M. Louis,[1] Rodolfo Ghirlando,[2] James L. Baber,[1] Ad Bax[1,*]

[1]Laboratory of Chemical Physics and [2]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Supporting Information

EXPERIMENTAL PROCEDURES

Protein expression, purification and reconstitution in bicelles

HIV-1 gp41 spanning residues 677-716 from clade D (GenBank: U43386.1) was cloned in fusion with the immunoglobulin binding domain B1 (GB1; Figure S1A) in pJ414 vector (ATUM). The plasmid was transformed into BL-21 (DE3) competent cells, grown and induced for expression with a final concentration of 1 mM isopropyl β-D-1-thiogalactopyranoside at an optical density (600 nm) of ~0.6 for 4 hr at 37 °C. Cells were harvested by centrifugation at 5 K rpm for 30 min at 4 °C.

The cell pellet derived from 0.5 L of culture was suspended in 70 ml of buffer containing 50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 10 mM DTT and 0.1 mg/ml lysozyme, and sonicated. The lysate was spun at 20,000 g for 30 min at 4 °C. Supernatant was discarded and the pellet was briefly sonicated in 35 ml of buffer-A (50 mM Tris (pH 8.0), 6 M GnHCl, 0.2 M NaCl and 1% Triton-X 100). After centrifugation at 20,000 g for 30 min at 18 °C, the supernatant was subjected to Ni-NTA affinity chromatography (GE Biosciences), equilibrated and washed in buffer-A. The column was further washed with buffer-B [50 mM Tris-HCl (pH 8.0), 8 M urea, 0.2 M NaCl, 0.02 M imidazole] and the bound protein was eluted in buffer-B containing 0.4 M imidazole and 10 mM DPC. Peak fractions were pooled (up to 20 mg), concentrated to ~1.5 mL and an equal volume of 10% SDS was added. The solution was heated at 60 °C for 12 min, cooled to room temperature and the fusion protein was fractionated on a size exclusion Superdex-200 column (1.6 x 60 cm, GE Bioscience) equilibrated in 8M urea, 50 mM Tris-HCl (pH 8.0), 0.2 M NaCl and 20 mM SDS. Peak fractions were pooled, adjusted to a final concentration of 10% formic acid and subjected to reverse phase high performance liquid chromatography (RP-HPLC, Zorbax SB-C3 column). Pure GB1-TM fractions were pooled, lyophilized, re-dissolved in 10% formic acid, heated for 2 hr at 80 °C to promote cleavage between Asp-Pro,[1] and again subjected to the same HPLC column to fractionate the TM peptide from the other cleavage products.

An aliquot of the lyophilized TM peptide was dissolved in 25 mM MES (pH 6.5) buffer containing 8 M urea and 150 mM DMPC/DHPC bicelles (q=0.5) and dialyzed against 25 mM MES buffer (pH 6.5) to remove urea. The loss of DHPC was monitored by 1D ^{1}H NMR and compensated after dialysis to achieve a q of 0.4. We did not observe any change in TROSY-HSQC spectrum when dissolving lyophilized TM directly in DMPC/DHPC bicelles (q=0.4), and this procedure therefore was used for part of the samples prepared.

Isotope labeling was carried out by growing the cells in M9/D$_2$O medium supplemented with ^{15}N NH$_4$Cl and ^{13}C-d7 D-Glucose.

Backbone chemical shift assignments

Backbone chemical shift assignment experiments were carried out on 0.5 mM ^{13}C/^{15}N/^{2}H TM in a buffer containing 25 mM MES (pH 6.5), 150 mM DMPC/DHPC (q=0.4) and 8% D$_2$O. H^N, ^{15}N, ^{13}C$^\alpha$ and ^{13}C´ chemical shifts were obtained from TROSY-based HNCO and HNCA spectra collected on a 900MHz Bruker Avance II spectrometer equipped with a z-gradient TXI cryogenic probe. Further, a 3D ^{1}H^N-^{15}N-^{1}H^N NOESY HMQC (τ_m=250 ms) spectrum was collected for validating assignments. Data were processed with NMRPipe[2] and analyzed using CCPNMR.[3] Dihedral angles were obtained for ^{2}H isotope corrected backbone chemical shifts using TALOS-N.[4] Unless mentioned, all NMR spectra were collected at 318 K.

Backbone relaxation experiments

^{15}N spin-lattice (R$_1$), spin-spin (R$_{1\rho}$) relaxation, and ^{15}N-{^{1}H} NOE were collected on samples containing 500 μM ^{13}C/^{15}N/^{2}H TM in 25 mM MES (pH 6.5), 150 mM DMPC/DHPC (q=0.4) at 900 MHz using TROSY-based heteronuclear experiments.[5] The R$_1$ spectra were collected with 8 relaxation delays of 0, 80, 160, 320, 640, 1280, 1760 and 2240 ms. The R$_{1\rho}$ spectra were collected using a spin lock RF field strength of 2 kHz and relaxation delays of 0, 3, 9, 12, 27, 33, 40 and 50 ms. R$_2$ rates were extracted from R$_{1\rho}$ after correcting for ^{15}N offset effects.[6] The ^{15}N-{^{1}H} NOE experiment was collected in an interleaved fashion, where alternate free induction decays were collected with and without 6 s of proton saturation.

Paramagnetic NMR

A chemically synthesized single cysteine variant of TM (TM R707C) was further purified using RP-HPLC. TM R707C in [50 mM Tris (pH 7.5), 60% isopropanol and 20% acetonitrile] was incubated overnight at room temperature with 10-fold excess MTSL [(1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) Methanethiosulfonate]. Unreacted or excess MTSL was further removed by performing RP-HPLC. Labeling to completion of TM R707C with MTSL was verified by ESI-MS. For homogenous sample preparation, 75% isopropanol dissolved ^{15}N, ^{2}H TM (100 µM) and ^{1}H TM R707C-MTSL (200 µM) were mixed and lyophilized. The TM mixture was dissolved in a buffer solution (25 mM MES (pH 6.5), 100 mM DMPC/DHPC (q=0.4)). Proton R_2 rates for the paramagnetic and diamagnetic samples were measured using seven relaxation delays (0, 3, 6, 10, 16, 25, and 36 ms) at 900 MHz as described by Anthis et al.[7] The corresponding diamagnetic sample was prepared by quenching the MTSL label with 10-fold excess of ascorbic acid.

Residual dipolar couplings

Alignment of 0.1 mM ^{15}N, ^{2}H TM in 25 mM MES (pH 6.5), 100 mM DMPC/DHPC (q=0.4) was obtained in acrylamide (AA, 4.77 % w/v), 0.13 % bis(acrylamide) and positively charged (3-acrylamidopropyl)-trimethylammonium chloride (ATAC, 1.4 % w/v). Gels were radially compressed from 5.4 mm to 4.2 mm diameter by means of a funnel, used for entry of the sample into the NMR tube.[8] Alignment in neutral polyacrylamide gel (4.87 % w/v AA and 0.13 % bisacrylamide) was also obtained by radially compressing it from 5.4 mm to 4.2 mm diameter.

Synthesis and purification of 4R,4S-DOTA-M8 loaded with thulium and lutetium was described elsewhere,[9] and the tags were kindly provided by Dr. Haussinger (Biozentrum, Basel). HPLC purified ^{15}N, ^{2}H TM R707C in 50 mM MES (pH 6.5), 60% isopropanol, 20% acetonitrile was incubated overnight at room temperature with 5-fold molar excess 4R,4S-DOTA-M8, which was already charged with either thulium or lutetium. Unreacted DOTA-M8 was further removed by RP-HPLC. Complete charging of the lanthanide tag to the TM peptide was confirmed by mass spectrometry. TM R707C (40-70 µM) with lanthanide tag was then resuspended in ~100 mM DMPC/DHPC (q=0.4) and 25 mM MES (pH 6.5).

Amide RDCs were measured using transverse relaxation optimized spectroscopy (ARTSY) from the difference of couplings under anisotropic and isotropic conditions.[10] RDCs for samples aligned by gels and lanthanide tag were measured at 900 MHz and 800 MHz, respectively.

Structure calculation

Structures were calculated with the Xplor-NIH software, using a simulated annealing protocol.[11] Convergence of NMR-derived structures when starting from randomized initial models and using RDC restraints is a well known problem. Based on the α-helical chemical shifts we chose to start from either an idealized helical backbone model with backbone torsion angles set to φ=−60° and ψ=−40° for residues 679-709, or set to φ=−50° and ψ=−50°, or set to the starting angles of PDB entry 5JYN of the previously reported structure. As only chemical shifts and backbone RDCs were available as experimental input restraints, initial structure calculations were carried out for a model where all residues had been changed to Ala, thereby preventing steric clashing of the sidechains and greatly improving convergence. As input restraints, we used TALOS-N[4] derived backbone torsion angles, with a flat bottom potential spanning a width of 40° in both φ and ψ, with the force constant of 200 kcal mol^{-1} rad^{-2}. Structures yielded improved cross validation statistics when additionally employing a potential-of-mean-force (PMF) for hydrogen bonding[12] (see below).

For the backbone calculations on the poly-Ala model of gp41-TM, starting from the helical model (φ,ψ=−60°,−40° or −50°,−50°, or the torsion angles of structure 5JYN) at 1000 K, the temperature was ramped down to 5K in 995 steps with 0.4 ps simulated annealing duration per time step. Harmonic potentials with no flat bottom were used for the RDCs, with force constants ramped from 0.05-0.2 kcal mol^{-1} Hz^{-2} for the gel-derived RDCs, and five-fold weaker for the RDCs measured by paramagnetic alignment. The final force

constant of 0.2 kcal mol^{-1} Hz^{-2} was chosen such as to yield a final RMSD between experimental and structure-fitted RDCs (~0.5 Hz for polyacrylamide RDCs; ~1.6 Hz for paramagnetic RDCs) that is comparable to the experimental error in the measured couplings. A total of 60 structures was calculated and the 15 lowest energy structures were selected for representation.

Quality of the structures (Q$_{free}$) was evaluated by eliminating all RDCs corresponding to a given bond vector, one residue at a time, and repeating the structural calculation from the modified set of input RDC restraints and calculating:

$$Q = \mathrm{rms}(D_{obs}\text{-}D_{pred}) \, / \, \mathrm{rms}(D_{pred})$$

for the lowest energy structure, where D_{obs} and D_{pred} are the observed and predicted value of the omitted RDC, respectively, and rms is the root mean square function. Because the orientation of ^{15}N-^{1}H bond vectors in a nearly straight α-helix deviates strongly from random, it is preferred to replace the denominator in the above expression by the rms value predicted for a uniform spherical distribution:[13]

$$\mathrm{rms}(D_{pred}) = \{D_a^2[4 + 3Rh^2]/5\}^{1/2}$$

where D_a and Rh are the axial and rhombic components of the alignment tensor, and this was the procedure used in our analysis.

Incorporating the HBDB potential[12] resulted in better quality structures, as evidenced by comparing Q$_{free}$ values of the structures calculated in the presence and absence of HBDB (Table S4). An ensemble of the poly-Ala backbone models of gp41-TM is shown in Figure S4A, and the corresponding statistics are listed in Table S2. Structure calculations on the actual sequence of gp41-TM was carried out by fixing the backbone atoms of the lowest energy poly-Ala structures and side chain energies were minimized by ramping the torsionDB potential[14] from 0.002-2 kcal mol^{-1} Hz^{-2}. The structures with added sidechains are shown in Figure S4B. The Ramachandran map of the structures shows a narrow distribution of backbone torsion angles (centered around ϕ=-62°, ψ=-43°; Figure S4C,D) that is indistinguishable for different helical starting structures.

Analytical ultracentrifugation

Sedimentation equilibrium experiments were carried out on a Beckman Optima XL-A at 20 °C following standard protocols.[15] Samples of 60 mM DMPC/DHPC (q = 0.4) bicelles in 50 mM NaCl and 50 mM MES (pH 6.5) were reconstituted in solutions containing 80 – 100% (v/v) D$_2$O. The samples were analyzed at 50,000 rpm by monitoring the absorbance at 230 nm, and sedimentation equilibrium data were analyzed in SEDPHAT 13.0a[15] to obtain the buoyant molar mass. In this manner, the density for neutral buoyancy was found at 92.5% D$_2$O, consistent with observations made by sedimentation velocity experiments carried out at 50,000 rpm and 20 °C on a Beckman Coulter ProteomeLab XL-I using the interference optical detection system. GB1-TM samples were reconstituted in 60 mM DMPC/DHPC (q=0.4), 50 mM NaCl, 50 mM MES (pH 6.5) and 92.5% (v/v) D$_2$O, and studied at rotor speeds of 10,000, 20,000 and 35,000 rpm. The ^{15}N and ^{2}H isotopically labeled form of GB1-TM Clade D was studied at 19 μM, whereas GB1-TM Clade C was studied at 15 and 30 μM. The radial distribution of the protein was monitored using the absorbance optics at 280 nm. Multi-speed equilibrium data were processed in GUSSI 1.3.0 [16] and analyzed in terms of a single ideal species in SEDPHAT 13.0a to obtain the molar mass. The partial specific volumes of GB1-TM Clade C and Clade D were calculated based on the amino acid composition in SEDNTERP,[17] and corrected for isotopic labeling (Clade D) and deuterium exchange (Clade C). The density of 50 mM NaCl, 50 mM MES (pH 6.5) in H$_2$O was measured at 20 °C on an Anton Paar DMA 5000 density meter, and corrected for the presence of 92.5% (v/v) D$_2$O.

EPR DEER

TM R707C-MTSL (50-500 µM) was dissolved in a D_2O buffer containing 25 mM MES (pH 6.5), and 50-200 mM DMPC/DHPC (fully protonated, q=0.4), and subsequently 10%(w/v) perdeuterated glycerol was added. Samples were flash frozen in liquid N_2 and data were collected at 50 K. Four-pulse DEER experiments[18] were collected at Q-band (33.8 GHz) on a Bruker E-580 spectrometer equipped with a 150 W traveling wave tube amplifier and a model ER5107D2 resonator. All experiments employed 8 ns pump (ELDOR) π pulses, 12 ns $\pi/2$ and 24 ns π observe pulses, and a 95 MHz frequency difference between pump and observe pulses. The pump frequency was placed at the maximum of the echo-detected field-swept spectrum. The 400 ns half-echo periods of the first echo were incremented 8 times in 16 ns increments to average the effect of ^2H modulation. DeerAnalysis2015 was used to analyze all data.[19] The typical exponential function with a dimension of 3.0 was not quite adequate to fit the backgrounds observed in Figure S6A-D, presumably because the lipids concentrate the monomers in a spatially non-homogeneous manner. A 2nd order polynomial function starting at T=1.5µs yielded reasonable background fits as judged by the resulting background-corrected dipolar evolution curves (not shown). An exponential fit of dimension 3.0 provided an adequate background fit for the three samples with the highest lipid-to-peptide ratios (Figure S6E-G), indicating that these samples better approximate a spatially homogeneous MTSL distribution.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis

It is fairly well established that α-helical TM peptides can exhibit anomalous migration on SDS-PAGE.[20-22] Nevertheless, to allow comparison with the results reported by Dev et al., we also carried out such measurements. Our data show that TM separates from bicelles during SDS-PAGE, with the bicelles migrating faster than TM (Figure S2A). Depending on the SDS concentration in the running buffer, TM migrates either as a ~10 kD band in 20 mM SDS or as a 15 kD band in 3.5 mM SDS (Figure S2B,C). This anomalous migration is typical of helical proteins bound to SDS, as referenced above. Importantly, GB1-TM migrates somewhat slower than a theoretical monomer (14 kDa) but exhibits the same migration at both SDS concentrations (Figure S2B) and is not compatible with TM being a trimer. In this respect, it is important to note that the NMR spectrum of TM and the TM resonances of GB1-TM in its NMR spectrum are essentially indistinguishable (Figure S5A), demonstrating that the oligomerization state of TM in bicelles is the same for TM and GB1-TM. Anomalous migration of TM is also highlighted by the comparison of crosslinked TM (R707C) dimer (~15 kD) with that of the corresponding monomer (~10 kD) under the same conditions (Figure S2C).

(A)
MGSSHHHHHHHHHGSSGIEGRGGSMQYKLILNGKTLKGET
TTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTE
GSSGDPGN^{677}WLWYIRIFIIIVGSLIGLRIVFAVLSLVNR
VRQGYSPLS^{716}

(B)
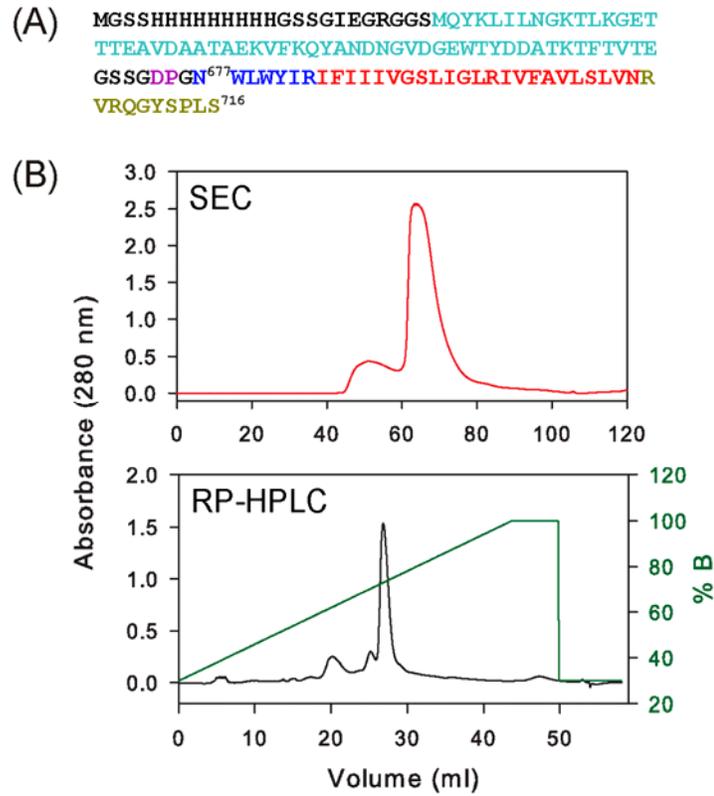
Figure S1. Purification of HIV-1 clade D gp41 TM. (A) GB1-TM primary amino acid sequence with GB1 in cyan, DP cleavage site in magenta, MPER in blue, TM in red and CT in olive. (B) The top panel shows the size exclusion chromatography elution profile of GB1-TM on Superdex 200 column (1.6 x 60 cm) at a flow-rate of 1.5 ml/min in 8M urea, 50 mM Tris (pH 8.0), 0.2 M NaCl and 20 mM SDS at room temperature. The bottom panel shows the RP-HPLC profile of TM subsequent to acid cleavage, fractionated using a linear gradient of 5% isopropanol/0.1 % TFA/water to 75 % isopropanol/25 % acetonitrile/0.1 % TFA/water.
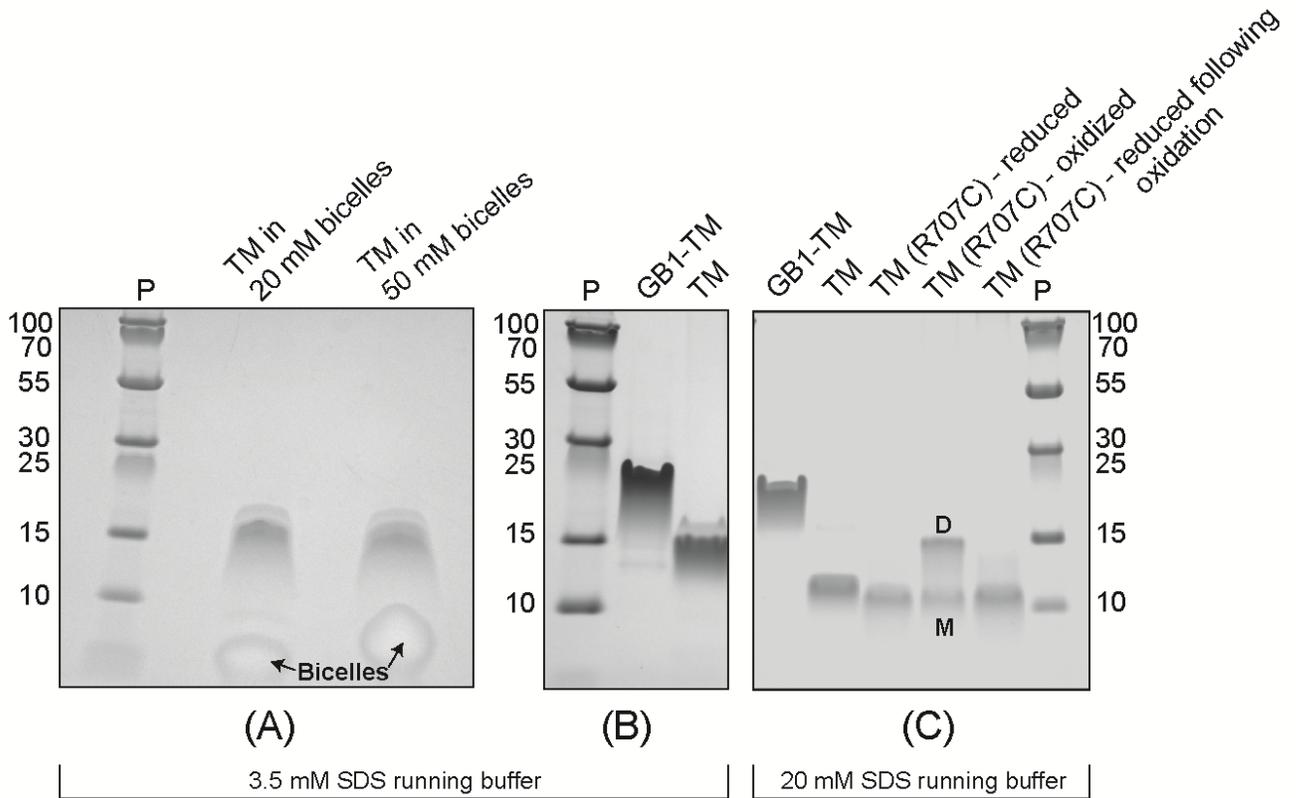
Figure S2. SDS-PAGE of TM and its analogues on 10-20% Tris Tricine gels. Samples were dissolved either in bicelles (A) or in buffer-A: 50 mM Tris-HCl (pH 8), 8 M urea, 200 mM NaCl and 20 mM SDS (B and C) and mixed with SDS-PAGE sample buffer to give a final concentration of 70 mM SDS prior to loading. (A) Migration of TM and bicelles (indicated by arrows). (B) Migration of GB1-TM relative to TM. (C) Migration of mostly oxidized TM (R707C) relative to its reduced form, GB1-TM, and TM. TM (R707C) dissolved in buffer-A was allowed to oxidize overnight at ambient temperature prior to addition of SDS-PAGE sample buffer (without 2-mercaptoethanol). Markings D (above) and M (below) the band denote the crosslinked dimer and monomer form of TM(R707C), respectively. Mass spectrometry of the oxidized sample showed masses of 9494 and 4748 (theoretical monomer mass = 4748), corresponding to dimer and monomer, respectively. P denotes Pageruler plus prestained marker (ThermoFisher). Bands were visualized using PageBlue protein staining solution (ThermoFisher).
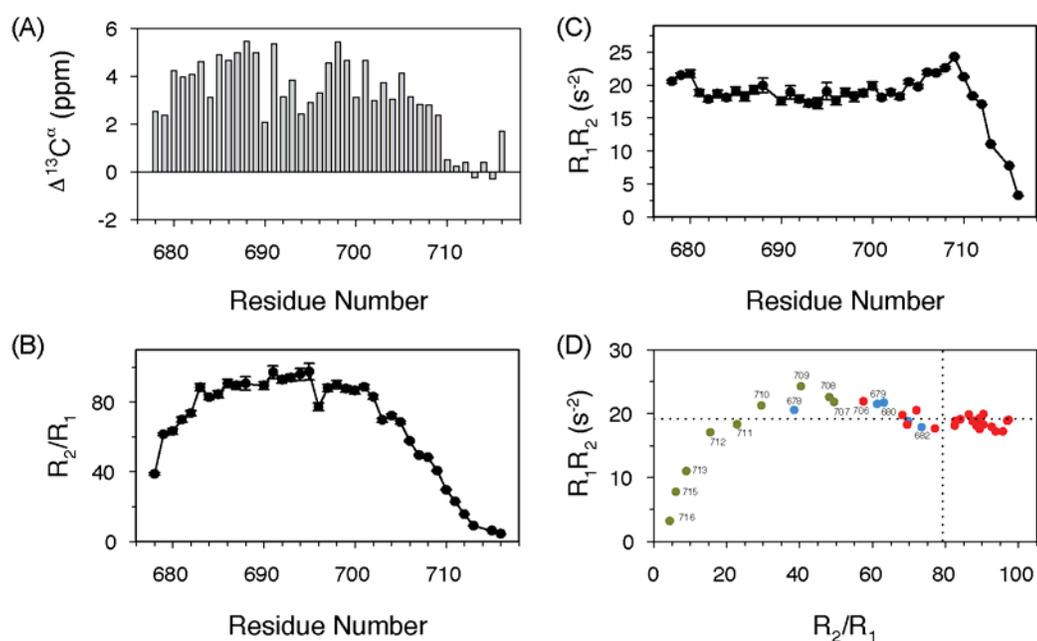
Figure S3. Analysis of secondary shifts and backbone dynamics of gp41-TM. (A) Differences in $^{13}C^{\alpha}$ chemical shifts from the neighbor-corrected random coil values.[23] Large positive values for residues 678 to 709 are consistent with $\alpha-$helical structure. Plots of (B) R_2/R_1 and (C) R_1R_2 as a function of residue number. (D) Plot of R_1R_2 versus R_2/R_1, with residues from MPER, TM and CT color coded as blue, red and olive, respectively. Dotted lines represent the 10% trimmed mean for residues with ^{15}N-{^{1}H} NOE values above 0.65. The highly homogeneous values of R_1R_2 for residues Y681-V705 indicate the absence of significant exchange broadening or motional anisotropy for this region of the peptide, whereas depressed values for the C-terminal residues are consistent with their relaxation being increasingly dominated by the fast internal motion of this dynamic tail.[24]
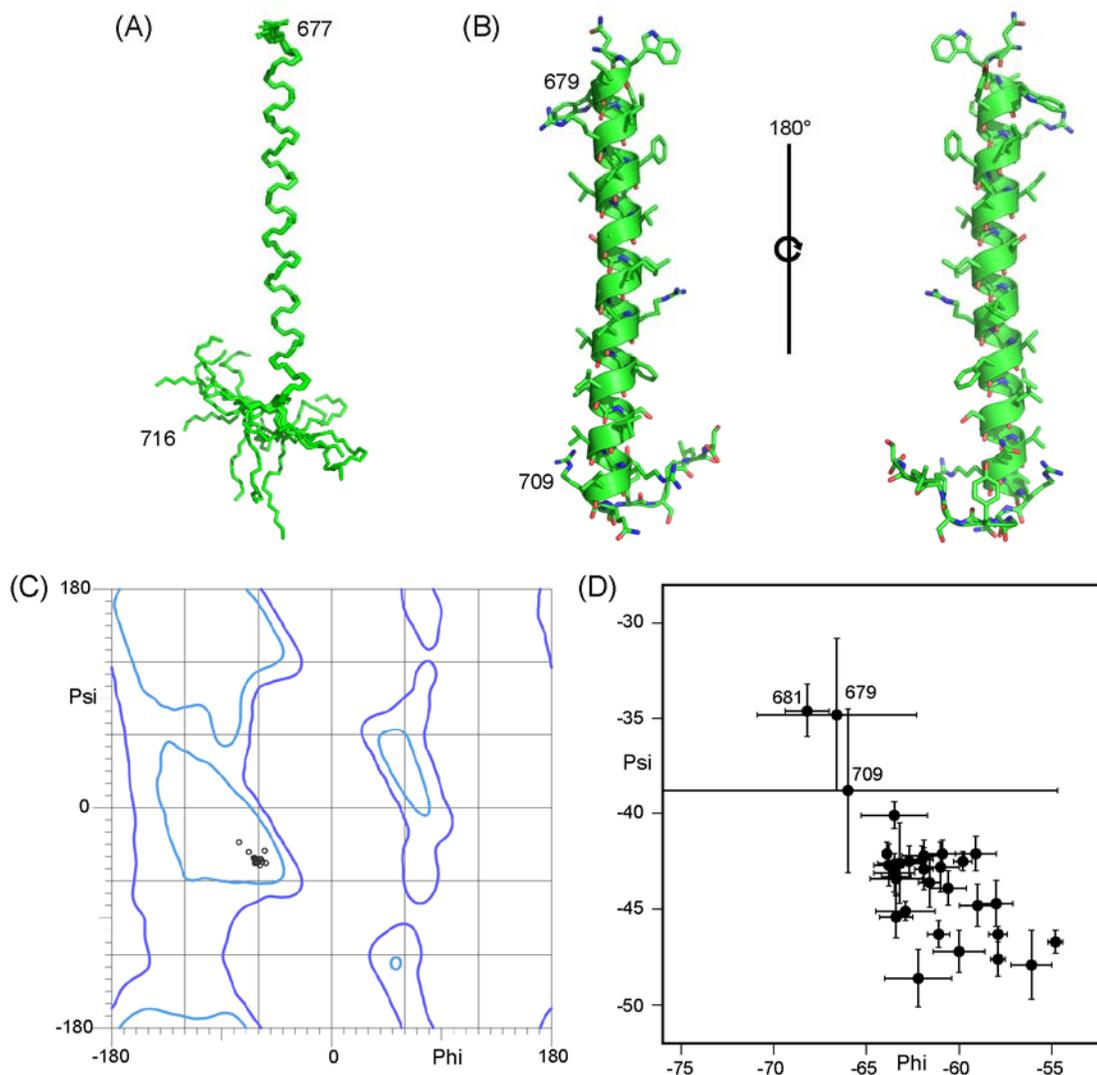
Figure S4. Backbone structure of HIV-1 clade D gp41 TM in q=0.4 phospholipid bicelles. (A) Backbone ensemble of 15 lowest energy structures in stick representation. (B) Lowest energy TM in ribbon representation, with sidechains shown in stick representation. (C) Ramachandran map distribution for residues 679-709 of the lowest energy model, derived using Molprobity.[25] (D) Zoomed region of the Ramachandran plot, displaying the average ϕ/ψ angles (residues 679-709) and their standard deviations for the 15 lowest energy structures.
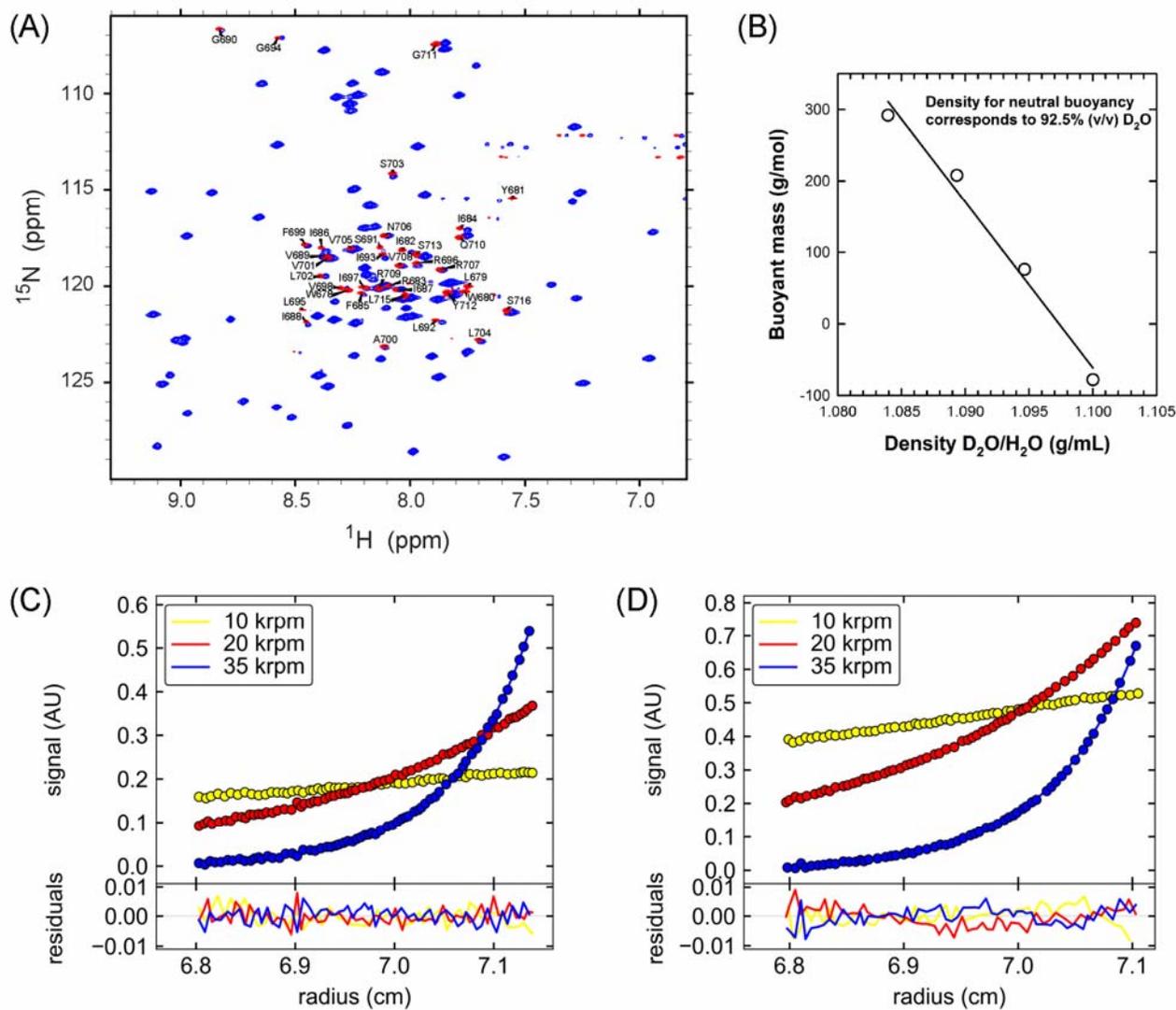
Figure S5. Comparison of experimental data for GB1-TM and TM. (A) ^1H-^{15}N TROSY-HSQC superposition of GB1-TM clade D (blue) with TM clade D (red). (B) Density matching of 60 mM DMPC/DHPC (q=0.4) bicelles in 50 mM MES buffer (pH 6.5) and 50 mM NaCl, against varying concentrations of D_2O (80-100%). Neutral buoyancy was obtained at 92.5% D_2O. (C, D) Sedimentation equilibrium absorbance profiles of (C) 15 µM and (D) 30 µM GB1-TM (Clade C), each containing 60 mM DMPC/DHPC (q=0.4) bicelles and solvent conditions of (B), at 10000 (yellow), 20000 (red) and 35000 (blue) rpm rotor speeds. Profiles are essentially the same as those obtained for clade D (Figure 4B, main text). Global fitting of the data to a single species resulted in masses of 16.5 kD and 17.5 kD, for 15 µM and 30 µM, respectively. Residuals are shown below the profiles.

(A) 200μM TM
60mM DMPC/DHPC (q=0.4)
lipid:peptide=300

(B) 500μM TM
250mM DMPC/DHPC (q=0.4)
lipid:peptide=500

(C) 50μM TM
25mM DMPC/DHPC (q=0.4)
lipid:peptide=500

(D) 50μM TM
25mM DMPC/DHPC (q=0.5)
lipid:peptide=500

(E) 100μM TM
100mM DMPC/DHPC (q=0.4)
lipid:peptide=1000

(F) 50μM TM
200mM DMPC/DHPC (q=0.4)
lipid:peptide=4000

(G) 100μM TM
100mM DHPC
lipid:peptide=1000
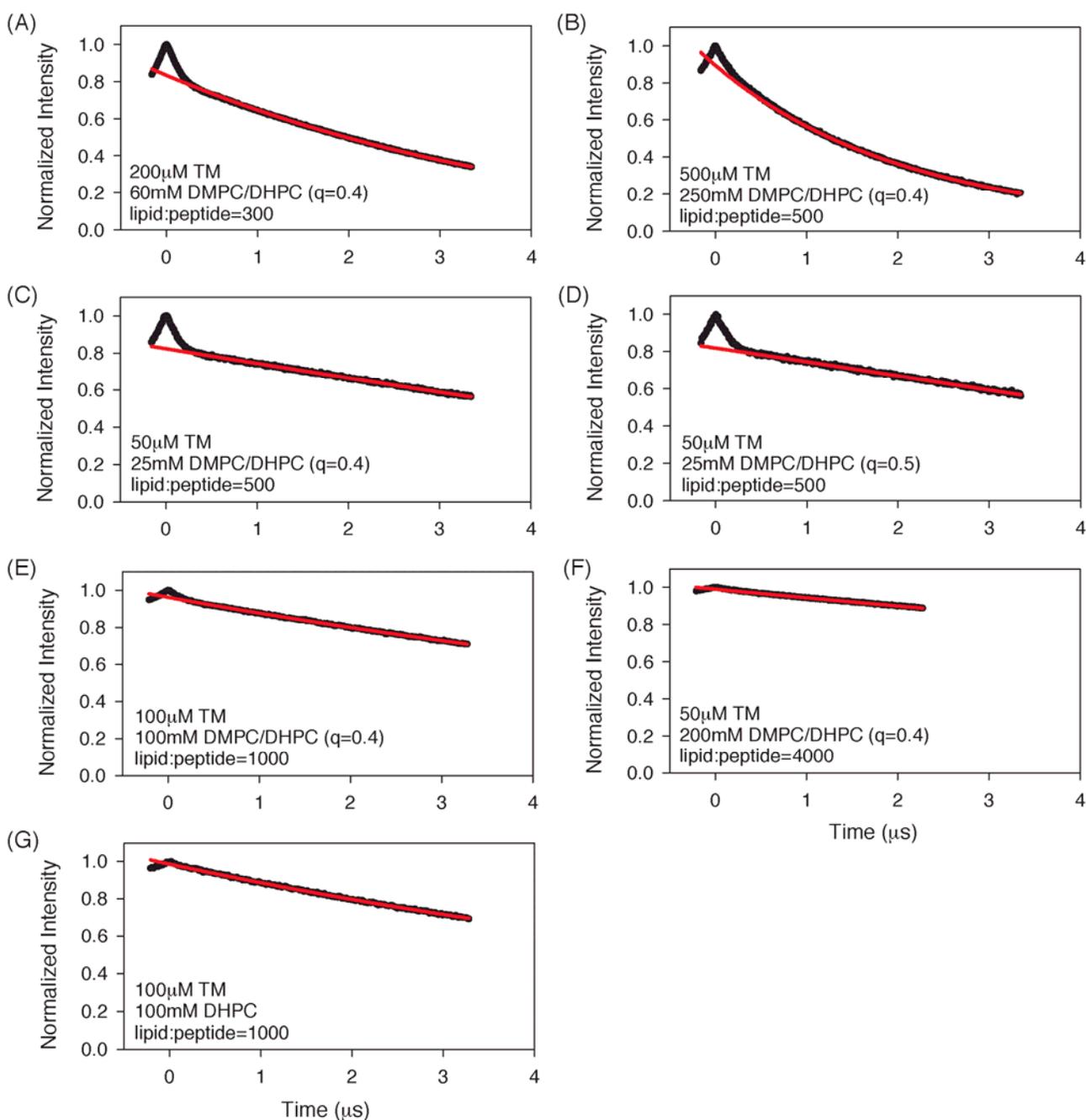
Figure S6. DEER time domain data of TM with a nitroxide label at residue R707C. Sample conditions are marked in the panels, with each sample also containing 10% (w/v) glycerol. Red lines in the panels represent the exponential background fit. The initial, small amplitude decay at short echo times (<0.5 μs) results from random intra-particle dipolar couplings and is attributed to the small fraction of particles that have more than one peptide in the same bicelle. Modulation depths of ~0.43 and ~0.68 would be measured for a dimer or a trimer, respectively, on our spectrometer using the parameters stated and for ~100% MTSL labeling efficiency. Corresponding echo-detected field swept spectra are shown in Figure S7, and demonstrate that the absence of modulation is not caused by excessive dipolar coupling.
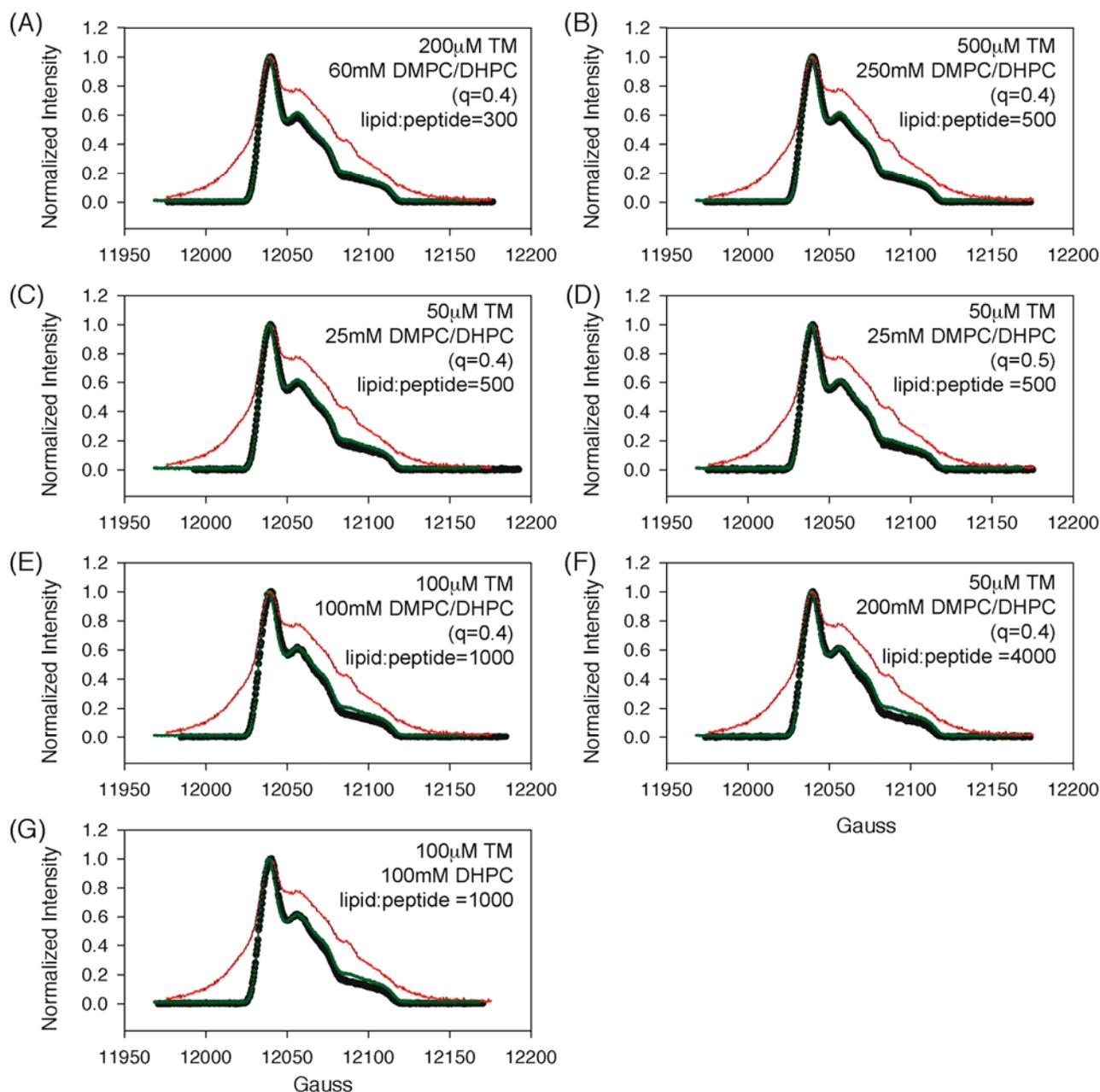
Figure S7. Normalized echo-detected Q-band field swept spectra of TM samples for which DEER time domain data are presented in Figure S6. For reference, analogous spectra for N-Cys CoreS (red), a homo trimer with spin labels closer than 1.5 nm and extensive dipolar broadening, which potentially can result in a depressed modulation depth,[26] and a protonated doubly MTSL labeled protein A sample (green) which shows the modulation depth expected for a two-MTSL sample, are superimposed on the TM spectra (black). A Hahn echo with a 12 ns $\pi/2$ pulse, a 24 ns π pulse, a 400 ns half-echo period, and a 80-ns echo integration window was employed to record these spectra. As the spectrometer frequency is variable between the samples, field values were adjusted by small, constant offsets to facilitate comparison. All spectra were recorded at a temperature of 50 K.
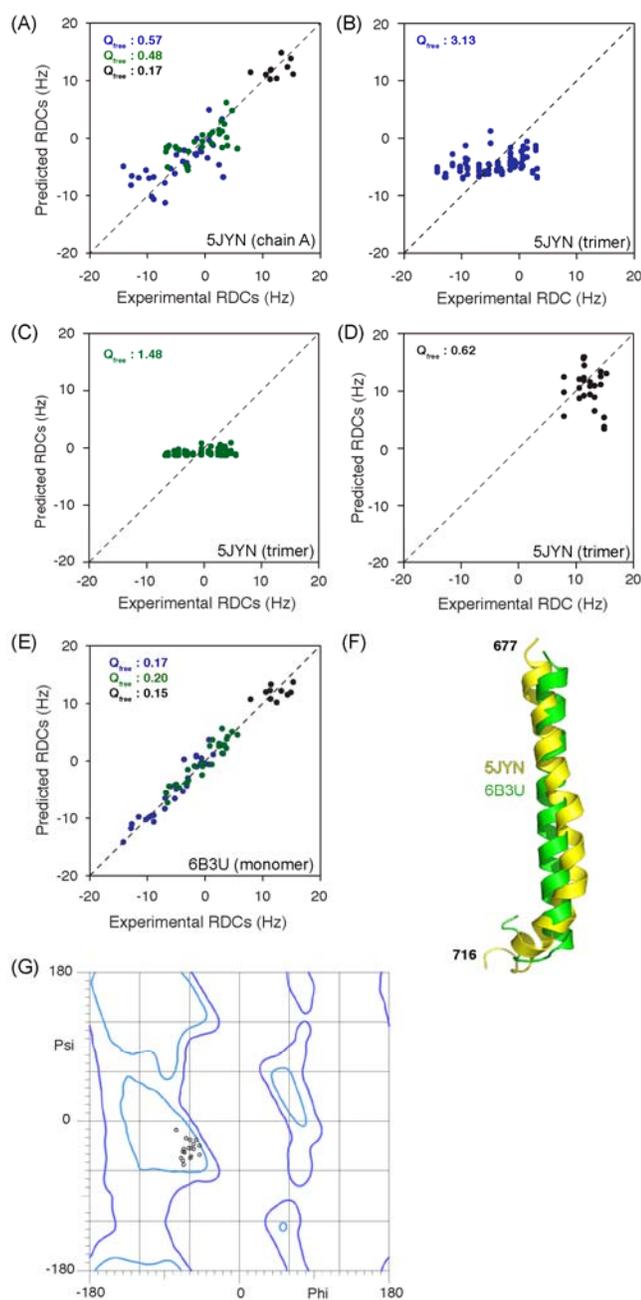
Figure S8. Comparison of structures and RDC fits of gp41-TM. (A) Plot of experimental RDCs against predicted values obtained from a singular value decomposition (SVD) fit of experimental RDCs to a single chain of the previously reported structure (PDB: 5JYN, chain A). Green, blue, and black symbols represent RDCs from neutral gel, positively charged gel, and thulium lanthanide alignment, respectively. (B-D) Fits to the trimer. Lack of perfect C_3 symmetry in 5JYN results in 3 predicted RDCs for each peptide residue. (E) Cross-validated comparison of experimental and predicted ^{1}H-^{15}N RDCs from the current study. Predicted data for each residue were obtained from a structure calculation that included the three RDC data sets of all other residues, but not of the residue whose RDCs are predicted and compared to experimental values. (F) Superposition of the RDC-derived TM structure (green) with the prior TM structure (yellow; PDB entry 5JYN, chain A). A backbone RMSD of 3.6 Å was obtained over 31 residues (679-709). (G) Ramachandran plot for residues 679 to 709 (PDB ID: 5JYN) derived by Molprobity.[25]

Supplementary Table S1. H^N, N, C′, C^α chemical shifts for HIV-1 clade D gp41 TM.

ResID	ResName	H^N (ppm)	N(ppm)	C′(ppm)	C^α(ppm)
677	Asn	-	-	175.96	-
678	Trp	8.27	120.22	177.69	59.89
679	Leu	7.75	120.02	178.76	58.05
680	Trp	7.76	120.29	177.92	61.47
681	Tyr	7.55	115.45	178.12	61.21
682	Ile	8.03	118.14	177.29	64.84
683	Arg	8.09	119.99	177.81	60.64
684	Ile	7.78	117.01	177.4	64.16
685	Phe	8.21	120.36	176.73	62.01
686	Ile	8.39	118.03	177.79	65.22
687	Ile	8.06	120.18	178.82	65.70
688	Ile	8.45	121.86	177.29	66.15
689	Val	8.37	118.46	178.05	67.36
690	Gly	8.83	106.66	175.19	47.47
691	Ser	8.13	117.99	175.42	63.60
692	Leu	7.89	121.8	178.82	58.31
693	Ile	8.12	118.34	178.09	65.19
694	Gly	8.58	107.13	174.88	47.67
695	Leu	8.46	121.25	178.15	58.14
696	Arg	7.97	118.79	179.52	59.08
697	Ile	8.2	120.05	177.37	65.59
698	Val	8.31	120.06	177.72	67.73
699	Phe	8.46	117.86	178.1	61.82
700	Ala	8.11	123.09	180.47	55.30
701	Val	8.36	118.55	177.69	66.95
702	Leu	8.39	119.5	178.88	58.27
703	Ser	8.07	114.14	176.78	61.82
704	Leu	7.7	122.76	178.59	58.21
705	Val	8.26	118.05	177.44	66.56
706	Asn	8.12	117.45	177.02	56.20
707	Arg	7.86	119.2	178.85	58.94
708	Val	8.04	118.88	178.26	65.17
709	Arg	8.14	120.12	177.35	58.41
710	Gln	7.79	117.49	176.71	56.84
711	Gly	7.88	107.51	173.79	45.49
712	Tyr	7.83	120.26	175.17	57.93
713	Ser	7.97	118.35	-	55.83
714	Pro	-	-	176.4	63.61
715	Leu	8.03	120.58	176.34	55.11
716	Ser	7.58	121.28	-	60.06

Supplementary Table S2. Structural Statistics for 20 lowest energy structures of gp41 TM.[a]

RMSD for NMR distance, orientation and dihedral angle restraints (no. of restraints)	
Dihedral angle restraints (68)	0°
Total RDC restraints (70)	0.5 Hz
NH RDCs from positively charged gel (30)	0.4 Hz
NH RDCs from neutral charged gel (29)	0.5 Hz
NH RDCs from DOTA-M8 Thulium (11)	1.6 Hz
RDC Cross validation	
Q_{free} positively charged gel	16.5%
Q_{free} neutral gel	20.4%
Q_{free} DOTA-M8 Thulium	15.2%
Structure Statistics[b]	
Backbone RMSD[b]	0.2 Å
Ramachandran Statistics[b]	
Most Favored	100%
Additionally allowed	0%

[a] Structures were deposited in PDB under accession ID of 6B3U
[b] Over residues 679-709

Supplementary Table S3. Alignment tensor parameters and normalized scalar product (NSP) between different alignment tensors.

Medium	D_a (Hz)	R	α	β	γ	$10^4 A_{zz}$	$10^4 A_{yy}$	$10^4 A_{xx}$	NSP[a]
Positive	-9.7	0.32	93.9	69.3	63.3	-8.96	6.63	2.33	NA
Neutral	-6.1	0.60	89.4	73.9	44.0	-5.65	5.37	0.28	0.77
Thulium	6.7	0.08	148.9	39.2	115.9	6.20	-3.47	-2.73	-0.35
Thulium[b]	-13.5	0.08	-53.6	49.0	39.3	-12.50	7.00	5.50	-0.15

[a] Normalized scalar product for the different alignment tensors relative to positively charged polyacrylamide charged gel.
[b] This second solution provides an essentially indistinguishable quality of the fit, but corresponds to an alignment strength substantially larger than expected for the DOTA-M8-Tm tag at 318K.

Supplementary Table S4. Quality (Q_{free}) of structures derived in the absence and presence of RDC restraints and HBDB potential.

Media	Q_{free}		
	With RDC; without HBDB	With HBDB; without RDC	With RDC; with HBDB
Positive	36.0%	24.5%	16.5%
Neutral	41.0%	30.0%	20.4%
Thulium	14.8%	12.5%	15.2%

Supplementary Table S5. Relaxation parameters for gp41 TM collected at 900 MHz ^{1}H frequency.

Residue	NOE	R_1 (s^{-1})	R_2 (s^{-1})	S^2
678	0.65 (±0.06)	0.73 (±0.01)	28.14 (±0.24)	0.67 (±0.01)
679	0.65 (±0.04)	0.59 (±0.01)	36.29 (±0.28)	0.88 (±0.01)
680	0.61 (±0.07)	0.59 (±0.01)	37.01 (±0.60)	0.88 (±0.01)
681	0.79 (±0.06)	0.52 (±0.01)	36.22 (±0.44)	
682	0.74 (±0.05)	0.49 (±0.01)	36.19 (±0.42)	0.88 (±0.01)
683	0.73 (±0.05)	0.46 (±0.01)	40.56 (±0.57)	0.96 (±0.01)
684	0.76 (±0.04)	0.47 (±0.01)	38.64 (±0.45)	0.95 (±0.01)
685	0.75 (±0.05)	0.48 (±0.01)	40.06 (±0.64)	0.96 (±0.01)
686	0.82 (±0.05)	0.45 (±0.01)	40.62 (±0.64)	0.98 (±0.01)
687	0.76 (±0.05)	0.46 (±0.01)	41.49 (±0.63)	0.96 (±0.01)
688	0.76 (±0.08)	0.47 (±0.02)	42.42 (±1.07)	1.00 (±0.01)
690	0.82 (±0.05)	0.44 (±0.01)	39.60 (±0.55)	0.97 (±0.01)
691	0.74 (±0.09)	0.44 (±0.01)	42.76 (±1.0)	1.00 (±0.01)
692	0.80 (±0.05)	0.44 (±0.01)	40.68 (±0.53)	0.98 (±0.01)
693	0.78 (±0.05)	0.43 (±0.01)	40.14 (±0.55)	0.98 (±0.01)
694	0.90 (±0.08)	0.42 (±0.01)	40.61 (±0.79)	1.00 (±0.01)
695	0.87 (±0.11)	0.44 (±0.02)	43.03 (±1.29)	1.00 (±0.01)
696	0.83 (±0.06)	0.48 (±0.01)	36.87 (±0.62)	
697	0.84 (±0.05)	0.46 (±0.01)	40.74 (±0.60)	0.98 (±0.01)
698	0.77 (±0.06)	0.45 (±0.01)	40.41 (±0.72)	0.97 (±0.01)
699	0.79 (±0.04)	0.46 (±0.01)	40.51 (±0.50)	0.97 (±0.01)
700	0.80 (±0.05)	0.48 (±0.01)	41.40 (±0.60)	0.96 (±0.01)
701	0.83 (±0.04)	0.45 (±0.01)	40.01 (±0.37)	0.97 (±0.01)
702	0.83 (±0.04)	0.48 (±0.01)	39.51 (±0.4)	0.96 (±0.01)
703	0.76 (±0.04)	0.51 (±0.01)	35.60 (±0.36)	
704	0.69 (±0.03)	0.53 (±0.01)	38.40 (±0.32)	0.92 (±0.01)
705	0.67 (±0.03)	0.54 (±0.01)	36.70 (±0.30)	0.88 (±0.01)
706	0.69 (±0.03)	0.62 (±0.01)	35.49 (±0.26)	0.85 (±0.01)
707	0.67 (±0.03)	0.66 (±0.01)	32.84 (±0.18)	0.79 (±0.01)
708	0.69 (±0.03)	0.68 (±0.01)	32.94 (±0.18)	0.79 (±0.01)
709	0.71 (±0.03)	0.78 (±0.01)	31.29 (±0.16)	0.75 (±0.01)
710	0.53 (±0.03)	0.85 (±0.01)	25.06 (±0.12)	0.59 (±0.01)
711	0.50 (±0.03)	0.90 (±0.01)	20.44 (±0.10)	0.48 (±0.01)
712	0.50 (±0.02)	1.05 (±0.01)	16.25 (±0.04)	0.37 (±0.01)
713	0.33 (±0.02)	1.11 (±0.01)	9.94 (±0.04)	0.21 (±0.01)
715	0.10 (±0.01)	1.12 (±0.01)	6.88 (±0.03)	0.14 (±0.01)
716	-0.30 (±0.01)	0.86 (±0.01)	3.75 (±0.02)	

Supplementary Table S6. $^1D_{NH}$ RDCs obtained on gp41 TM in positively and neutral polyacrylamide gel and with thulium paramagnetic label.

Residue	Positive (Hz)	Neutral (Hz)	Thulium (Hz)
678	-1.7	2.3	
679	-14.2	-6.6	
680	-3.7	2.4	
681	0.4		7.8
682	-12.8	-6.7	11.4
683	-8.96	-4.5	11.3
684	-1.52	1.3	13.2
685	-3.9	1.2	11.3
686	-12.8	-6.8	8.3
687	-6.0	-1.9	10.6
688	-0.8	2.5	14.5
690	-10.3	-6.3	14.3
691	0.7	3.7	14.9
692	-0.7	3.7	12.4
693	-9.7	-5.0	
694	-5.3	-0.5	
695	0.7	3.0	
696	-3.0	3.0	
697	-9.0	-3.1	
698	-1.5	0.8	
699	3.1	5.6	
700	-7.0	-3.0	
701	-7.0	-2.9	
702	2.4	3.9	
703	-1.5	2.8	
704	-11.6	-5.2	
705	-3.2	-0.5	
706	2.9	4.7	
707	-4.7	-0.5	
708	-9.2	-3.4	
709	1.4	3.4	
710[a]	1.4	3.3	
711[a]	-12.2	-3.8	
712[a]	-1.0	-0.2	
713[a]	-2.4	-0.9	3.9
715[a]	1.9	1.5	1.6
716[a]	1.0	0.9	0.1

[a] Values for these residues are significantly impacted by internal motion, based on ^{15}N relaxation data, and were not used in structure calculations or validations.

Supplementary Table S7. Proton R_2 rates of gp41 TM under paramagnetic and diamagnetic conditions.

Residue	$R_{2, para}$ (s^{-1})	$R_{2, dia}$ (s^{-1})	ΔR_2 (s^{-1})
678	60.5 (±2.5)	46.37 (±1.48)	14.13 (±2.95)
679	37.7 (±0.7)	32.09 (±0.54)	5.65 (±0.89)
680	71.0 (±3.0)	61.51 (±2.64)	9.44 (±3.97)
681	52.1 (±1.5)	50.65 (±1.42)	1.45 (±2.10)
682	40.9 (±0.9)	36.31 (±0.79)	4.61 (±1.19)
683	38.3 (±0.8)	29.54 (±0.56)	8.80 (±0.97)
684	40.3 (±0.8)	32.61 (±0.7)	7.6 (±1.0)
685	37.0 (±0.7)	30.4 (±0.7)	6.6 (±1.0)
686	34.1 (±0.6)	28.3 (±0.6)	5.9 (±0.8)
687	38.0 (±0.8)	33.5 (±0.8)	4.6 (±1.2)
688	46.0 (±1.3)	41.3 (±1.6)	4.6 (±2.0)
689	38.1 (±0.6)	34.3 (±0.7)	3.8 (±0.9)
690	35.1 (±0.7)	28.5 (±0.7)	6.6 (±1.0)
691	52.7 (±2.4)	46.7 (±2.2)	6.1 (±3.2)
692	45.3 (±1.0)	40.9 (±1.1)	4.4 (±1.5)
693	40.2 (±0.8)	32.3 (±0.7)	7.9 (±1.1)
694	35.4 (±0.8)	29.6 (±0.7)	5.9 (±1.1)
695	66.86 (±2.3)	67.5 (±2.8)	-0.6 (±3.7)
696	41.8 (±1.0)	35.7 (±0.8)	6.1 (±1.3)
697	41.7 (±0.8)	33.7 (±0.6)	8.0 (±1.0)
698	47.9 (±1.2)	42.5 (±1.3)	5.4 (±1.8)
699	38.4 (±0.8)	30.6 (±0.6)	7.8 (±1.0)
700	48.2 (±1.1)	45.5 (±1.1)	2.7 (±1.5)
701	34.2 (±0.5)	29.2 (±0.4)	5.0 (±0.6)
702	30.3 (±0.5)	25.2 (±0.4)	5.2 (±0.7)
703	47.4 (±1.2)	35.1 (±0.8)	12.2 (±1.4)
704	32.8 (±0.5)	27.2 (±0.4)	5.6 (±0.6)
705	33.9 (±0.5)	26.7 (±0.3)	7.2 (±0.6)
706	37.1 (±0.6)	26.2 (±0.3)	10.9 (±0.7)
707	31.3 (±0.4)	25.7 (±0.3)	5.6 (±0.5)
708	29.0 (±0.3)	24.1 (±0.2)	4.8 (±0.4)
709	30.5 (±0.4)	22.3 (±0.2)	8.2 (±0.5)
710	37.5 (±0.6)	31.6 (±0.4)	5.9 (±0.7)
711	34.2 (±0.5)	29.5 (±0.4)	4.6 (±0.6)
712	22.1 (±0.2)	18.2 (±0.1)	3.9 (±0.2)
713	28.2 (±0.4)	25.3 (±0.3)	2.9 (±0.5)
715	12.1 (±0.1)	9.4 (±0.1)	2.7 (±0.1)
716	8.5 (±0.1)	7.5 (±0.1)	1.0 (±0.1)

REFERENCES

1.	Piszkiew, D.; Landon, M.; Smith, E.L., *Biochem. Biophys. Res. Commun.* **1970,** 40, 1173-1178.
2.	Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A., *J. Biomol. NMR* **1995,** 6, 277-293.
3.	Vranken, W.F.; Boucher, W.; Stevens, T.J.; Fogh, R.H.; Pajon, A.; Llinas, P.; Ulrich, E.L.; Markley, J.L.; Ionides, J.; Laue, E.D., *Proteins* **2005,** 59, 687-696.
4.	Shen, Y.; Bax, A., *J. Biomol. NMR* **2013,** 56, 227-241.
5.	Lakomek, N.A.; Ying, J.F.; Bax, A., *J. Biomol. NMR.* **2012,** 53, 209-221.
6.	Massi, F.; Johnson, E.; Wang, C.Y.; Rance, M.; Palmer, A.G., *J. Am. Chem. Soc.* **2004,** 126, 2247-2256.
7.	Anthis, N.J.; Doucleff, M.; Clore, G.M., *J. Am. Chem. Soc.* **2011,** 133, 18966-18974.
8.	Chou, J.J.; Gaemers, S.; Howder, B.; Louis, J.M.; Bax, A., *J. Biomol. NMR* **2001,** 21, 377-382.
9.	Haussinger, D.; Huang, J.R.; Grzesiek, S., *J. Am. Chem. Soc.* **2009,** 131, 14761-14767.
10.	Fitzkee, N.C.; Bax, A., *J. Biomol. NMR* **2010,** 48, 65-70.
11.	Schwieters, C.D.; Kuszewski, J.J.; Tjandra, N.; Clore, G.M., *J. Magn. Reson.* **2003,** 160, 65-73.
12.	Grishaev, A.; Bax, A., *J. Am. Chem. Soc.* **2004,** 126, 7281-7292.
13.	Clore, G.M.; Garrett, D.S., *J. Am. Chem. Soc.* **1999,** 121, 9008-9012.
14.	Bermejo, G.A.; Clore, G.M.; Schwieters, C.D., *Protein Sci.* **2012,** 21, 1824-1836.
15.	Zhao, H.; Brautigam, C.A.; Ghirlando, R.; Schuck, P., *Current protocols in protein science,* Wiley, New York, NY, **2013,** Chapter 20, Unit20.12.
16.	Brautigam, C.A., *Methods Enzymol.* **2015,** 562, 109-133.
17.	Cole, J.L.; Lary, J.W.; Moody, T.P.; Laue, T.M., *Methods Cell Biol.* **2008,** 84, 143-179.
18.	Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H.W., *J. Magn. Reson.* **2000,** 142, 331-340.
19.	Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C.R.; Hilger, D.; Jung, H., *Appl. Magn. Reson.* **2006,** 30, 473-498.
20.	Walkenhorst, W.F.; Merzlyakov, M.; Hristova, K.; Wimley, W.C., *Biochim. Biophys. Acta* **2009,** 1788, 1321-1331.
21.	Rath, A.; Glibowicka, M.; Nadeau, V.G.; Chen, G.; Deber, C.M., *Proc. Natl. Acad. Sci. U. S. A.* **2009,** 106, 1760-1765.
22.	Tatulian, S.A.; Tamm, L.K., *Biochemistry* **2000,** 39, 496-507.
23.	Kjaergaard, M.; Poulsen, F.M., *J. Biomol. NMR* **2011,** 50, 157-165.
24.	Kneller, J.M.; Lu, M.; Bracken, C., *J. Am. Chem. Soc.* **2002,** 124, 1852-1853.
25.	Chen, V.B.; Arendall, W.B., III; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C., *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2010,** 66, 12-21.
26.	Louis, J.M.; Baber, J.L.; Clore, G.M., *Biochemistry* **2015,** 54, 6796-6805.