

Supplementary Information for

Structure of the membrane proximal external region of HIV-1 envelope glycoprotein

James J. Chou,
email: james_chou@hms.harvard.edu;
Bing Chen,
email: bchen@crystal.harvard.edu;
Stephen C. Harrison,
email: harrison@crystal.harvard.edu

This PDF file includes:

Supplementary text
Figs. S1 to S8
Table S1
References for SI reference citations

SI Methods

Protein expression and purification

Addition of the MPER to our previous gp41^{HIV1D(677-716)} construct led to severe protein aggregation. We therefore tested several HIV-1 gp41 fragments, including 660-716, 660-710, 665-716, 671-716, and 660-705, and found that the 660-710 fragment from HIV-1 clade D isolate 92UG024.2 was very stable in DMPC/DHPC bicelles with $0.5 \leq q \leq 0.6$, showing excellent solution NMR properties. This construct, designated gp41^{HIV1D(660-710)} or MPER-TMD, contains the entire MPER (residues 660-683) and the TMD (residues 684-705). The expression construct of the MPER-TMD was created as a trpLE fusion following a procedure described previously (1). Mutant constructs were generated by standard PCR protocols and confirmed by DNA sequencing. For NMR sample preparation, transformed *E. coli* strain BL21 (DE3) cells were grown in M9 minimal media supplemented with Centrum multivitamins and stable isotopes. Cultures were grown at 37° C to an absorbance of ~0.6 at 600 nm, and cooled to 22° C before induction with 100 μM isopropyl β-D-thiogalactopyranoside at 22° C for overnight. For fully deuterated proteins, bacterial culture in 1 ml LB medium was spun down and the cells were adapted in 99.8% D₂O medium (100 ml, with deuterated glucose) over night. Then collected cells were grown in 99.8% D₂O (4 L) (Sigma Aldrich, St. Louis, MO) with deuterated glucose (Cambridge Isotope Laboratories, Tewksbury, MA). The MPER-TMD protein was extracted, cleaved by cyanogen bromide, purified and lyophilized as described (1). Protein identity was confirmed by MALDI-TOF mass spectrometry and SDS-PAGE.

Reconstitution of the MPER-TMD in bicelles

To reconstitute the MPER-TMD in bicelles, 2 mg of lyophilized protein was mixed with 9 mg 1,2-dimyristoyl-*sn*-Glycero-3-Phosphocholine (DMPC; protonated or deuterated from Avanti Polar Lipids, Alabaster, AL) and dissolved in hexafluoro-isopropanol. The mixture was slowly dried to a thin film under nitrogen stream, followed by overnight lyophilization. The dried thin film was redissolved in 3 ml of 8 M urea containing ~20 mg 1,2-dihexanoyl-*sn*-Glycero-3-Phosphocholine (D6PC; protonated or deuterated from Avanti Polar Lipids), with subsequent addition of ~2-4 mg 1,2-diheptanoyl-*sn*-Glycero-3-Phosphocholine (D7PC; protonated or deuterated from Avanti Polar Lipids). D7PC, with a critical micelle concentration (CMC) of ~1.4 mM, was used to prevent formation of liposomes, because D6PC, which has a CMC of ~15 mM, is often lost very rapidly during dialysis. The mixture was dialyzed twice against a 40 mM MOPS buffer (pH 6.8) (1 L each time) to remove the denaturant, and 15 mg D6PC was added to the sample before the second dialysis to compensate its loss. After the

second dialysis, we estimate that D7PC accounted for less than 5% of the total detergent (D6PC and D7PC). The DMPC:DHPC ratio was monitored by 1D NMR throughout the reconstitution process. If needed, additional D6PC was added to make the final DMPC: DHPC ratio between 0.5 and 0.6. The sample was concentrated in a Centricon (EMD Millipore, Billerica, MA) to ~350 μ l. The final NMR sample contained ~0.8 mM MPER-TMD (monomer), ~50 mM DMPC, ~100 mM D6PC, 40 mM MOPS (pH 6.7), 0.02% NaN_3 and 5% D_2O . For all NOE experiments, the protein was reconstituted using DMPC and DHPC with deuterated acyl chains (Avanti Polar Lipids).

Analysis of the oligomeric state of the MPER-TMD in bicelles by SDS-PAGE and chemical crosslinking

Wild type MPER-TMD was reconstituted in bicelles ($q=0.5$), then mixed with a SDS sample buffer (Invitrogen) without boiling, followed by SDS-PAGE at 200 volts for 30 minutes and Commassie blue staining. The MPER-TMD protein migrated slightly faster than did a marker protein with M.W. of 18 kDa (Fig. S1B), consistent with the theoretical size of a trimer (~18.5 kDa).

As an independent validation, we examined trimerization of the MPER-TMD in bicelles ($q=0.5$) by chemical crosslinking followed by urea-PAGE. Double mutation K665R/D674K was introduced retrospectively using the structure of the MPER-TMD to facilitate crosslinking with dithiobis-succinimidyl propionate (DSP). DSP was dissolved in DMSO at 200 mM, and added to reconstituted MPER-TMD (K665R/D674K) at 0.05 mM (monomer) concentration. Cross-linking was performed using DSP at a final concentration of 0.6 mM for 30 min at room temperature, and the reaction was quenched with 200 mM Tris buffer (pH 7.5) and then used for UREA-PAGE. The mutant MPER-TMD migrated in 8 M urea as a diffuse band at roughly the position expected for a monomer; the DSP-treated mutant migrated predominantly as a sharp band at the position expected for a trimer (M.W. ~20 kDa including DSP) (Fig. S1D).

NMR spectroscopy

Assignment of NMR resonances. All NMR data were collected at 35°C on Bruker spectrometers operating at ^1H frequency of 900 MHz, 800 MHz, 750 MHz, or 600 MHz and equipped with cryogenic probes. NMR data were processed using NMRPipe (2), and spectra were analyzed using NMRPipe (2), Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco), and XEASY (3). Sequence specific assignment of backbone chemical shifts was accomplished using two pairs of TROSY-enhanced triple resonance experiments (4), recorded from a (^{15}N , ^{13}C , 85% ^2H)-

labeled sample. The triple resonance experiments included HNCA, HN(CO)CA, HN(CA)CO, and HNCO. Backbone amide resonance assignments were confirmed with a 3D ^{15}N -edited NOESY-TROSY-HSQC spectrum ($\tau_{\text{NOE}} = 200$ ms), which was recorded on a 750 MHz spectrometer from a (^{15}N , ^2H)-labeled sample. Protein aliphatic and aromatic resonances were assigned using a combination of 2D ^{13}C HSQC, 3D ^{15}N -edited NOESY-TROSY ($\tau_{\text{NOE}} = 120$ ms) and ^{13}C -edited NOESY-HSQC ($\tau_{\text{NOE}} = 150$ ms) recorded on a 900 MHz spectrometer. These experiments were performed with (^{15}N , ^{13}C)-labeled protein samples in bicelles with deuterated DMPC and DHPC acyl chains (Anatrace). For residues 685-702, most assignments of methyl group resonances were taken directly from those of our previous TMD structure (BMRB accession number 30090), as chemical shifts for those residues are essentially the same for the MPER-TMD and TMD. Specific stereo assignment of the methyl groups of valines and leucines were obtained from a 28 ms constant-time ^1H - ^{13}C HSQC spectrum recorded using a 15% ^{13}C -labeled sample (5). Finally, the assignment of the indole amine of W672 and W680 was confirmed using a W672Y mutant.

Backbone chemical shift and TALOS+ analyses. The assigned chemical shift values of backbone ^{15}N , $^{13}\text{C}\alpha$, and $^{13}\text{C}'$ were used as input for the TALOS+ program (6) to predict backbone dihedral angles. Out of 48 residues with assignments, the dihedral angles of 42 residues were considered by TALOS+ as 'GOOD'. Furthermore, the carbon secondary shifts of gp41^{HIV1D(660-710)} and gp41^{HIV1D(677-716)} are compared in Fig. S2D&E, which provided a secondary structure mapping of residues 660-716 of HIV-1 Env.

Assignment of NOE restraints. Intramolecular distance restraints derived from nuclear Overhauser enhancements (NOEs) were obtained from the ^{15}N -edited and ^{13}C -edited NOESY spectra recorded on a 900 MHz spectrometer (see above). Many backbone amide resonances between 682 and 701 were extremely weak in non-deuterated protein. Thus, in the ^{15}N -edited NOESY recorded with the non-deuterated protein, most of the local NOEs could not be analyzed. The spectrum of residues 685-702 is, however, the same as that of the gp41^{HIV1D(677-716)} sample (Fig. S2A&B). We thus applied the NOE restraints assigned for residues 685-702 of the gp41^{HIV1D(677-716)} sample.

Determining the inter-protomer distance restraints faced the challenge of measuring NOEs between structurally equivalent protomers having the same chemical shifts. To solve this problem, we prepared a mixed sample in which half of the protomers were ^{15}N , ^2H -labeled (0.4 mM) and the other half 15% ^{13}C -labeled (0.4 mM). Recording a 3D ^{15}N -edited NOESY-TROSY-HSQC ($\tau_{\text{NOE}} = 200$ ms)

on this sample allowed measurement of exclusive NOEs between the ^{15}N -attached protons of one subunit and aliphatic protons of the neighboring subunits. The non-deuterated peptide was 15% ^{13}C -labeled for recording the ^1H - ^{13}C HSQC spectrum as an internal aliphatic proton chemical shift reference. Examples of residues in the MPER that had inter-protomer NOEs are shown in Fig. S3A&B. In parallel, a control sample with only (^{15}N , ^2H)-labeled MPER-TMD (0.4 mM) was used to record an identical 3D ^{15}N -edited NOESY-TROSY-HSQC (Fig. S3A&B). The ^{15}N , ^2H -labeled MPER-TMD in the mixed and control samples were from the same protein expression batch. Comparison of the mixed and control spectra allowed rigorous confirmation that a particular inter-protomer NOE was due solely to the mixing of protomers with different labeling schemes.

Structure calculation

Structures were calculated with XPLOR-NIH (7). The matching resonances, shown in Fig. S2, between gp41^{HIV1D(660-710)} and gp41^{HIV1D(677-716)} for the TM core (residues 685-702) justified the strategy to determine the MPER structure using the known structure of the TM core as a starting point (1). To implement this strategy, we first modeled secondary structures of the MPER-TMD monomer using the ‘GOOD’ dihedral angles generated by TALOS+ using backbone chemical shift values (6). Second, a trimer was assembled using the previously assigned NOE restraints for residues 685-702 (taken from PDB accession code: 5JYN). Finally, we applied the newly assigned inter-protomer restraints (most of them from residues 660-683) to complete the trimer of the MPER-TMD. For each inter-protomer restraint between two adjacent protomers, three identical distance restraints were assigned respectively to all pairs of neighboring protomers to satisfy the condition of C3 symmetry. The assembled trimer was then refined against the complete set of NOE restraints (both intra-protomer and inter-protomer) and dihedral angles using a simulated annealing (SA) protocol in which the temperature was lowered from 1000 to 200 K in steps of 40 K. The NOE restraints were enforced by flat-well harmonic potentials, with the force constant ramped from 2 to 50 kcal/mol \AA^{-2} during annealing. Backbone dihedral angle restraints were taken from the ‘GOOD’ dihedral angles from TALOS+, all with a flat-well (\pm the corresponding uncertainties from TALOS+) harmonic potential with force constant ramped from 10 to 1000 kcal/mol rad^2 . A total of 150 structures were calculated, and the 15 lowest energy structures were selected as the final structural ensemble (Fig. S4C and Table S1).

Can we account for the observations in Chiliveri et al (8), who concluded that a construct essentially identical to the one that yielded our previous TMD structure was monomeric rather than trimeric? Their spectra were recorded at 45°C rather than 35°, a possible distinction. We believe,

however, that their data are also consistent with a stable trimer, as indeed suggested by the SDS gel in the supplementary material of that paper. Two sets of NMR observations were presented as showing lack of oligomer. With regard to the alignment tensor derived from residual dipolar coupling measurements they performed, those authors state that for a "C3 symmetric structure, alignment in an anisotropic medium must be axially symmetric, with the unique axis of the alignment tensor coinciding with the C3 axis." This will not be the case, however, if there are non-C3-symmetric fluctuations in the trimer. The C-terminal part of the TMD (the "hydrophilic core") indeed shows substantial dynamics, and there is no a priori reason to suppose that those fluctuations will be symmetrical. They also write that they were unable to detect cross-chain PRE. We have found that unless the two chains (in this case, the spin-labeled chain and the ^{15}N -labeled chain) are mixed in the trpLE-fusion form *before* purification, the trimers that form after CNBr cleavage and subsequent isolation are too stable to exchange. According to the Chiliveri et al paper (8), the chains in their experiments were mixed after purification, and we believe that they may have had a mixture of homotrimers.

^{13}C selective inter-protomer NOE experiment

For definitive assignment of inter-molecular NOE cross peaks, we prepared an isotopically mixed sample in which half of the proteins were (^{15}N , ^2H) labeled and the other half (^{13}C , ^1H) labeled. The new NOE experiment was based on the same 3D ^{15}N -edited NOESY-TROSY-HSQC as above, but used the ^1H - ^{13}C J coupling to distinguish inter vs. intra -molecular NOEs. The ^1H evolution (or frequency labeling) period before the NOE mixing was changed to a "mixed constant-time" evolution to introduce ^1H - ^{13}C J evolutions. Two interleaved 3D spectra were recorded. In one case, the total J evolution was 0 ms, making both inter and intra -molecular NOE peaks positive. In the other case, the total J evolution was 8 ms, making intermolecular NOE peaks negative and intra-molecular NOE peaks positive. Collecting the two data sets interleaved with each other allowed us to add the two spectra to see only intra-molecular NOEs or to subtract the two spectra to see only inter-molecular NOEs.

Measurement of inter-protomer PRE

To obtain inter-protomer information independent of NOEs, we performed PRE analysis on a mixed sample containing ~1:1 ratio of (^{15}N , ~70% ^2H) labeled MPER-TMD and ^{14}N MPER-TMD with a Cys mutation at position 660 or 710, reconstituted in bicelles with $q = 0.5$. Cells that expressed the two proteins with different labeling schemes were mixed prior to purification. After bicelle reconstitution, 100 mM MTSL dissolved in DMSO was added to the samples at a final ratio (MTSL to MPER-TMD)

of 5:1 and the mixture incubated at room temperature overnight. The sample was then passed through a PD-10 desalting column to remove free MTSL. DHPC was added to the final sample to adjust the q to ~ 0.5 . For measuring PRE, defined as the ratio of peak intensities of the non-reduced (I) and reduced (I_0) samples, the ^1H - ^{15}N TROSY-HSQC spectrum was recorded before and after the addition of 10 mM sodium ascorbate. A mixed TMD sample, labeled at position 716, was prepared similarly.

Solvent PRE analysis of TM partition

We use a previously published solvent PRE method (9) to determine the transmembrane partition of the MPER-TMD trimer in bicelles. This method is based on the notion that if the bicelle is sufficiently wide ($q > 0.5$), the lateral solvent PRE becomes negligible, thus allowing the use of measurable solvent PRE to probe residue-specific depth immersion of the protein in the bilayer region of the bicelle (Fig. S6). To ensure that our bicelles are wide enough, we reconstituted the MPER-TMD in bicelles with $q = 0.6$ (Fig. S6A) to perform the solvent PRE analysis as described in ref (9). The water-soluble and membrane-inaccessible paramagnetic agent, Gd-DOTA (Sigma), was used to generate solvent PREs. Gd-DOTA (in 200 mM stock solution) was titrated into the bicelle sample to reach final concentrations of 0.0, 1.0, 2.0, 4.0, 6.0, 8.0, 10.0 and 15.0 mM. At each concentration, a 2D ^1H - ^{15}N TROSY-HSQC spectrum was recorded at 600 MHz to measure residue-specific PRE, defined here as the ratio of peak intensity in the presence (I) and absence (I_0) of the paramagnetic agent. Peak intensities were measured at peak local maxima using quadratic interpolation to identify peak centers. For each of the residues, we used *Origin* (OriginLab, Northampton, MA) to fit the PRE titration curve to exponential decay

$$\frac{I}{I_0} = 1 - \text{PRE}_{\text{amp}} \left(1 - e^{-[\text{Gd-DOTA}]/\tau} \right) \quad (1)$$

to derive the residue-specific PRE amplitude (PRE_{amp}) (Fig. S6B). To determine the position of the MPER-TMD trimer relative to the bilayer center, we calculated, for each residue i , the distance (r_z) along the protein symmetry axis, which is parallel to the bilayer normal, from the amide proton to an arbitrary reference point based on the structure of the MPER-TMD trimer. This calculation converted PRE_{amp} vs. (residue number) to PRE_{amp} vs. r_z , which was then analyzed using *the sigmoidal fitting* method (see detailed discussion in ref (9)). Briefly, the trimer structure was moved along the 3-fold axis in increment of 0.3 Å (Fig. S6C) to achieve the best fit to the symmetric sigmoid equation

$$\text{PRE}_{\text{amp}} = \text{PRE}_{\text{amp}}^{\text{min}} + \frac{(\text{PRE}_{\text{amp}}^{\text{max}} - \text{PRE}_{\text{amp}}^{\text{min}})}{1 + e^{(r_z^I - |r_z|)/\text{SLOPE}}}, \quad (2)$$

where $\text{PRE}_{\text{amp}}^{\text{min}}$ and $\text{PRE}_{\text{amp}}^{\text{max}}$ are the limits within which PRE_{amp} can vary, r_z^I is the inflection point (the distance from the bilayer center at which PRE_{amp} is halfway between $\text{PRE}_{\text{amp}}^{\text{min}}$ and $\text{PRE}_{\text{amp}}^{\text{max}}$), and

SLOPE is a parameter which reports the steepness of the curve at the inflection point. The best fit (Fig. S6C) gave an adjusted coefficient of determination (R^2_{adj}) of 0.96, and was used to determine the position of the trimer structure with respect to the bilayer center ($r_z = 0$).

Generation of Env mutant constructs and production of monoclonal antibodies

Full-length Env mutants were generated using the 92UG037.8 gp160, described previously (10), as a template by QuikChange Site-Directed Mutagenesis (Agilent Technologies). All constructs were confirmed by restriction digestion and DNA sequencing. Anti-HIV-1 Env monoclonal antibodies and their Fab fragments were produced as described (10, 11). We have generated expression constructs of antibodies

PG9, PG16, PGT145, 2F5, 4E10 and 10E8 using synthetic genes made by GeneArt Gene Synthesis (Life Technologies, Carlsbad, CA) or GenScript. The DH570 expression constructs were kindly provided by Barton Haynes (Duke University); antibody VRC01 by John Mascola (VRC, NIH); the CHO stable line expressing antibody b6 by Dennis Burton (Scripps); 17b hybridoma by James Robinson (Tulane University); 3791 hybridoma by Susan Zolla-Pazner (New York University). The expression constructs of 10E8 were initially obtained from the NIH AIDS reagent program. Anti-His antibody MA1-21315 was purchased from ThermoFisher Scientific.

Antibody binding to the MPER-TMD in bicelles

The bicelle-reconstituted MPER-TMD trimer was examined for binding to several antibodies, including antigen binding fragment (Fab) of 2F5, 4E10, 10E8 and DH570, as well as a control Fab derived from an anti-His antibody MA1-21315. For binding to 2F5 or 4E10, the (^{15}N , 80% ^2H)-labeled MPER-TMD was reconstituted in bicelles with $q = 0.5$. For binding to the anti-His Fab, the (^{15}N , 80% ^2H)-labeled MPER-TMD with N-terminal 6His-tag reconstituted in the same bicelles was used. In all three cases, we mixed the MPER-TMD at 0.1 mM and Fab at 0.07 mM with a final molar ratio of 1.0:0.7. First, a reference 1D spectrum of the MPER-TMD before Fab addition was recorded at 600 MHz using a ^{15}N -edited TROSY-HSQC experiment. The same spectrum was then recorded at 2, 4, 6, and 8-hour time points after addition of Fab. Spectrum intensities at each time point relative to those of the reference spectrum were measured in the Bruker Topspin3.2 operating system using the multiple-spectra display mode. These intensities directly correlate to the fraction of the MPER-TMD in the NMR samples not bound to antibody. Hence, the fraction of the MPER-TMD in a conformation incompatible with Fab binding (Fraction Unbound), or equivalently, the fraction of the Fab unbound to the MPER-TMD, at various time points can be calculated as $(I - 0.3) / (I_0 - 0.3)$, where I_0 is the

reference intensity set to 1, I is the fraction intensity relative to I_0 at a particular time, and subtraction of 0.3 was made to account for the 30% molar access of the MPER-TMD to the Fab. Linear fitting of the Fraction Unbound vs. Time points was performed using the *Origin* software (OriginLab, Northampton, MA) (Fig. 3G).

Production of pseudoviruses containing mutant Envs

Preparation of HIV-1 Env pseudoviruses containing mutations in the MPER, titration of pseudovirus stocks to determine the 50% tissue culture infectious dose per ml (TCID₅₀/ml) and the neutralization assay using TZM.bl cells were performed as previously described (1, 12).

HIV-1 p24 Antigen ELISA assay and Western blot

Viral stocks were boiled in a buffer containing 0.5% Triton X-100 for 60 min and analyzed for p24 antigen using a HIV-1 p24 antigen ELISA 2.0 kit (ZeptoMetrix Corporation, Buffalo, New York), as described (1). Western blots were carried out following our published protocol (1). Virus lysates were made by directly mixing p24-normalized virus stocks with Laemmli Sample Buffer (Bio-Rad, Hercules, CA) and boiling for 10 min. Lysates of cells expressing Env or its mutants were prepared by resuspending the cells in PBS (phosphate buffered saline) at a density of 2×10^6 cells/ml, followed by treatment with the Sample Buffer and boiling for 10 min.

Flow cytometry

293T cells were transiently transfected with 2 μ g of the 92UG037.8 gp160 expression construct or its MPER mutants. Flow cytometry was carried out as described (1, 10). All data were analyzed by FlowJo (FlowJo, LLC, Ashland, OR).

Cell-cell fusion assay

Cell-cell fusion assay, based on the α -complementation of *E. coli* β -galactosidase, was carried out as described previously (1). Briefly, 293T cells were cotransfected with either HIV-1 Env and the α fragment of β -galactosidase or CD4, CCR5 and the ω fragment of β -galactosidase. Env-expressing cells (2.0×10^6 cells/ml) were mixed with CD4- and CCR5-expressing cells (2.0×10^6 cells/ml). Cell-cell fusion was allowed to proceed at 37°C for 2 hr. Cell-cell fusion activity was quantified using a chemiluminescent assay system, Gal-Screen (Applied Biosystems, Foster City, CA).

Viral infectivity and antibody neutralization assays

Viral infectivity of HIV-1 92UG037.8 Env and the MPER mutants was measured by infecting TZM.bl cells with p24-normalized pseudovirus in growth media containing DEAE-dextran (11 µg/ml).

Luciferase activity of the reporter gene was quantified 48 hours post-infection, as previously described (13). Neutralization by monoclonal antibodies was also determined by the luciferase-based virus neutralization assay in TZM.bl cells (13). The assay measures the reduction in luciferase reporter gene expression in TZM-bl cells following a single round of virus infection. Briefly, 5-fold serial dilutions of antibody samples were performed in duplicate (96-well flat bottom plate) in 10% DMEM growth medium (100 µl/well). Virus was added to each well in a volume of 50 µl, and the plates were incubated for 1 hour at 37°C. TZM.bl cells were then added (1×10^4 /well in 100 µl volume) in 10% DMEM growth medium containing DEAE-Dextran. Following a 48 hour incubation, luminescence was measured using Bright-Glo luciferase reagent (Promega, Madison, WI). Murine leukemia virus (MuLV) was used as a negative control virus for all assays. Antibodies used in this assay include IgG of b6, 3791, 17b, PG9, PG16, PGT145 and VRC01.

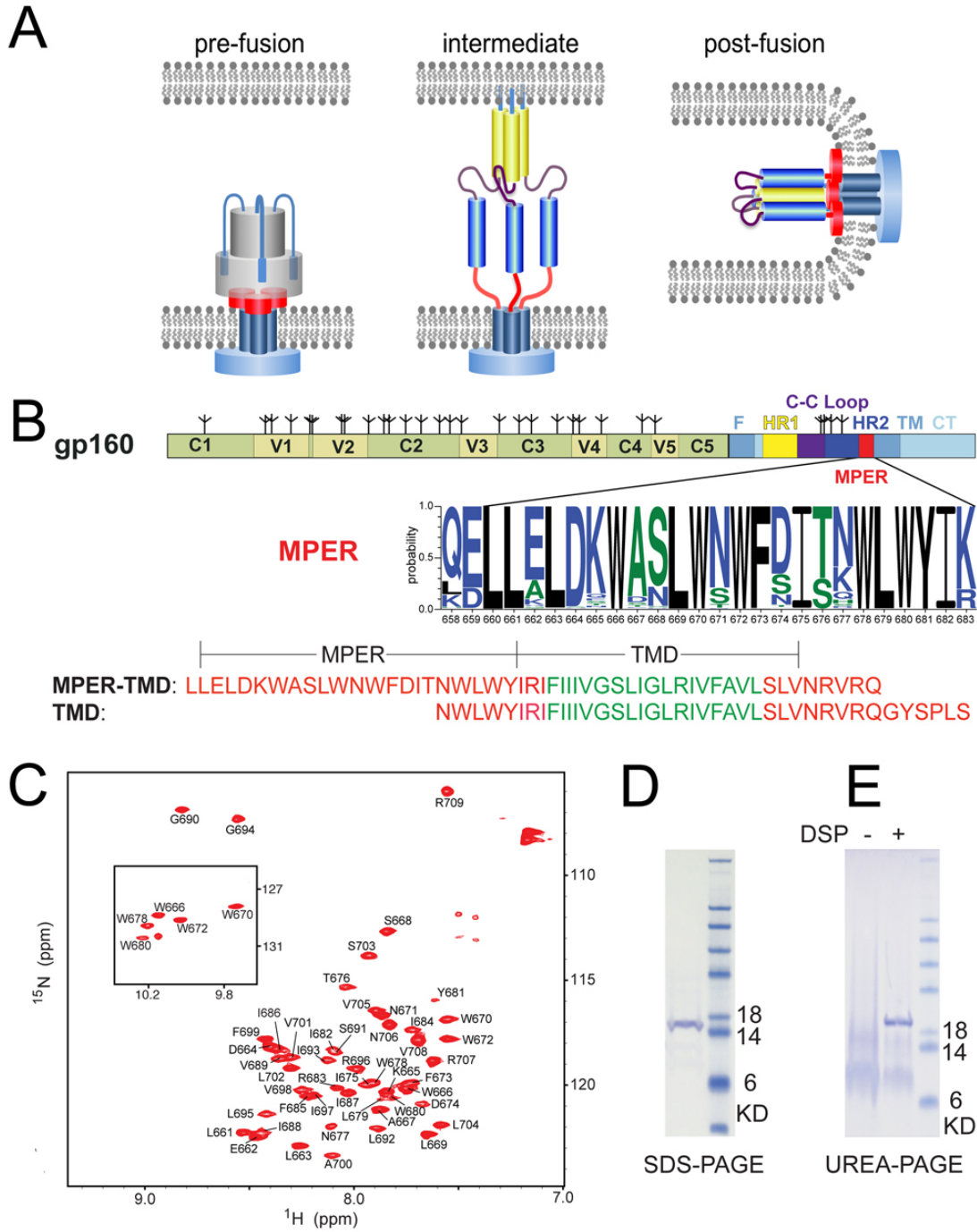


Figure S1. HIV-1 MPER and NMR sample preparation of the MPER-TMD in bicelles. (A)

Schematic of HIV-1 gp41 in the pre-fusion, intermediate, and post-fusion states, with the MPER-TMD highlighted in red. **(B)** Schematic of HIV-1 gp160 with segments of gp120 and gp41 designated as follows: C1-C5, conserved regions 1-5; V1-V5, variable regions 1-5; F, fusion peptide; HR1, heptad repeat 1; C-C loop, the immunodominant loop with a disulfide bond; HR2, heptad repeat 2; MPER, membrane proximal external region; TM, transmembrane anchor; CT, cytoplasmic tail, and tree-like symbols, glycans. Sequence logos showing variations in the MPER (residues 660-683) among 5132 HIV-1 sequences with small residues in green, hydrophobic in black, and strongly polar or charged in blue (generated using ANALYZEALIGN of the Los Alamos National Laboratory). Amino acid sequences of the gp41^{HIV1D(660-710)} (or MPER-TMD) construct used in the current study and the gp41^{HIV1D(677-716)} (or TMD) construct in earlier study(1) showing the overlapped TM region (residues 683-704) of HIV-1 clade D isolate 92UG024.2. The TMD core which has essentially the same NMR chemical shifts in both constructs (residues 685-702) is shown in green; the MPER in red and the rest near the cytoplasmic tail in light blue. **(C)** Assigned NMR spectra of the MPER-TMD in bicelles. The ¹H-¹⁵N TROSY-HSQC spectrum recorded at ¹H frequency of 750 MHz using (¹⁵N, ²H)-labeled protein. **(D)** SDS-PAGE of the MPER-TMD reconstituted in bicelles showing a homogeneous band at ~17 kDa position, consistent with the M.W. of a MPER-TMD trimer (~18.5 kDa). **(E)** Urea-PAGE analysis of crosslinking of an MPER-TMD mutant, K665R/D674K, in bicelles. We used standard, SDS-containing Laemmli buffers with a final urea concentration of 4.7M in the separating gel, 6.8M in the stacking gel, and 8M in the running buffer. The loading buffer was 9.6M in urea, 3 % in SDS, and 5% in glycerol before mixing (25 μL plus 5 μL sample) with protein sample and boiling. The uncrosslinked sample ran as a diffuse band near the position expected for a monomer (lane 1), whereas the mutant crosslinked by DSP ran as a sharp band at the position expected for a trimer (~20 kDa including DSP) (lane 2). The mutations were designed retrospectively from examination of the structure, to enhance crosslinking efficiency (see SI Methods).

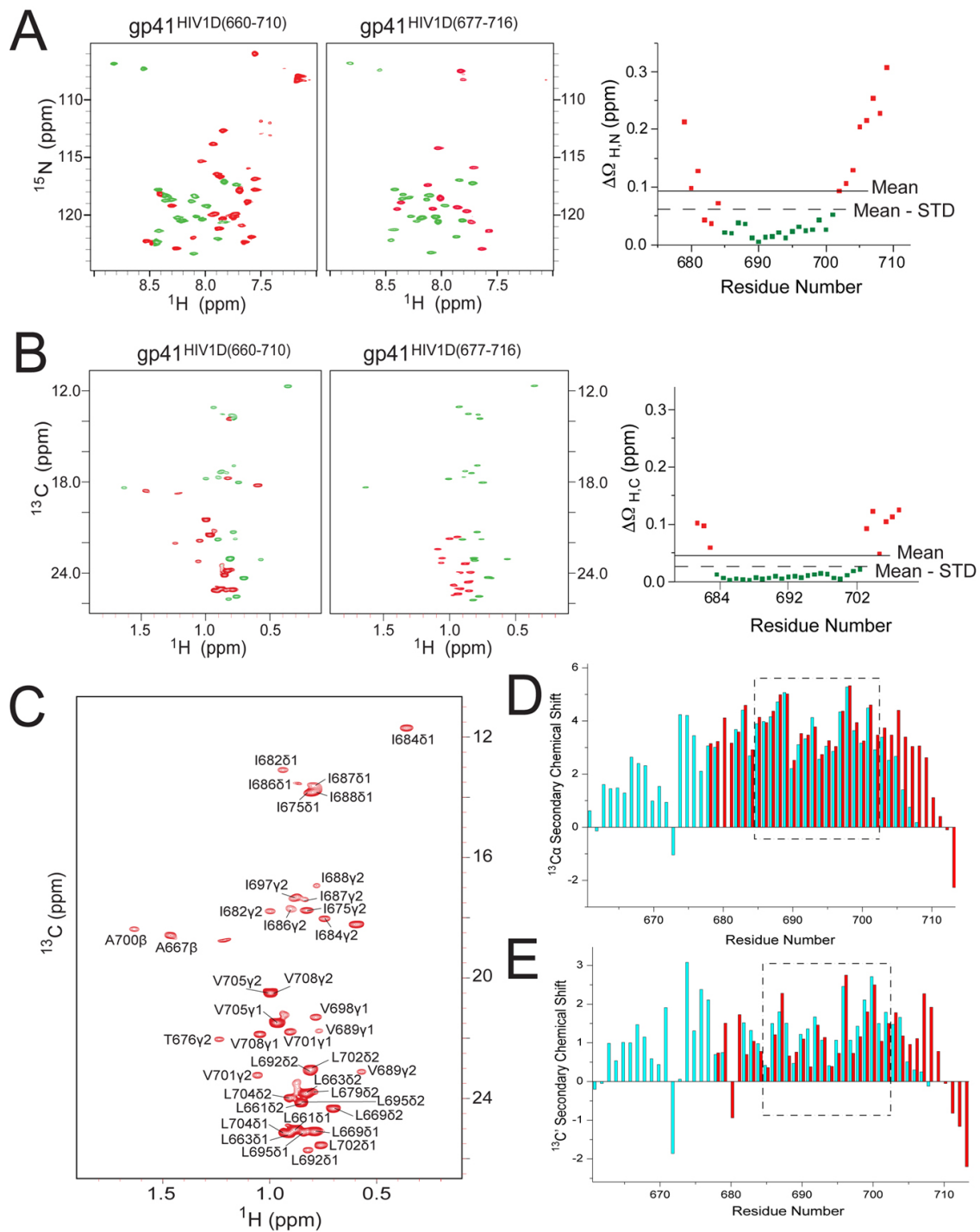


Figure S2. NMR resonances of the MPER-TMD of HIV-1 gp41 in bicelles. (A) ^1H - ^{15}N TROSY-HSQC spectra of the MPER-TMD (left) and TMD (right) displayed in the color scheme of (Fig. S1B), showing that the TMD cores of the two constructs have the same resonance shifts (green). The spectra were recorded with MPER-TMD and TMD samples, reconstituted in DMPC/DHPC bicelles ($q = 0.5$) in the same way as described in the SI Methods. The right-most panel shows the normalized chemical shift difference, $\Delta\Omega_{\text{H,N}}$ (Supplementary Methods), between the two spectra for each of the overlapping residues. The solid and dashed lines represent the mean $\Delta\Omega_{\text{H,N}}$ and mean $\Delta\Omega_{\text{H,N}} - \text{standard deviation}$ of $\Delta\Omega_{\text{H,N}}$, respectively. The error of $\Delta\Omega_{\text{H,N}}$, estimated based on peak line-width and signal/noise ratio, are all less than 0.0015 ppm and are thus not visible in the plot. (B) The methyl group chemical shift differences between the ^1H - ^{13}C HSQC spectra of the MPER-TMD (left) and TMD (right) constructs, displayed in the same manner as in (A). The residue number labels in the $\Delta\Omega_{\text{H,C}}$ plot are approximations, as there is not a methyl group for every residue. (C) Assigned methyl group resonances of the MPER-TMD in bicelles. The ^1H - ^{13}C HSQC (28 ms ^{13}C constant time) recorded at ^1H frequency of 800 MHz using (^{15}N , ^{13}C)-labeled protein, showing the assignments of the methyl group resonances. The proteins were reconstituted in DMPC/DHPC bicelles ($q = 0.5$), in which the acyl chains of DMPC and DHPC were deuterated. (D and E) Comparison of the $^{13}\text{C}\alpha$ and $^{13}\text{C}'$ secondary chemical shifts, respectively, between the gp41^{HIV1D(660-710)} (cyan) and the gp41^{HIV1D(677-716)} (red) reconstituted in bicelles. The secondary chemical shift values were generated using the program TALOS+ (6). The boxed region is the core region of the TMD including residues 685-702.

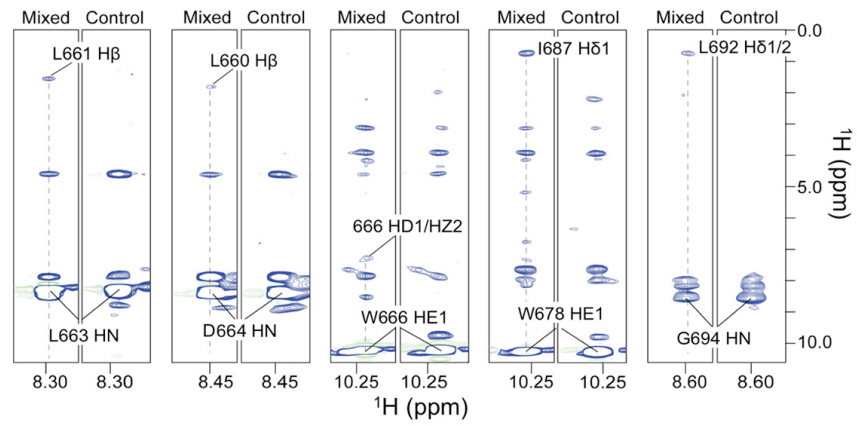
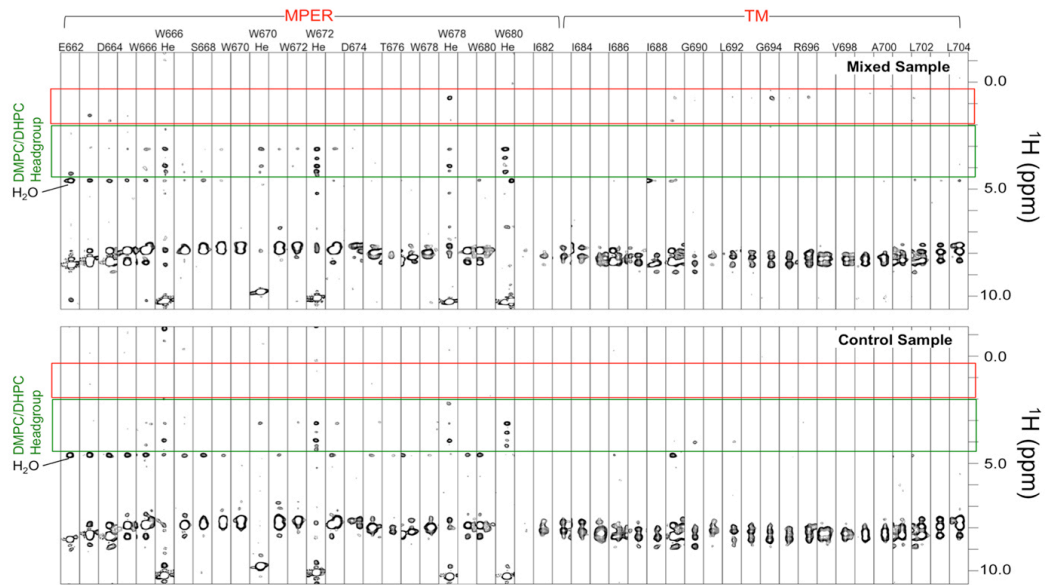
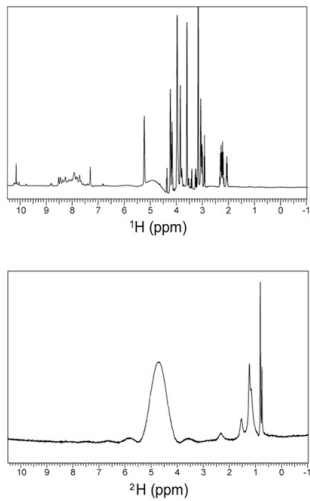
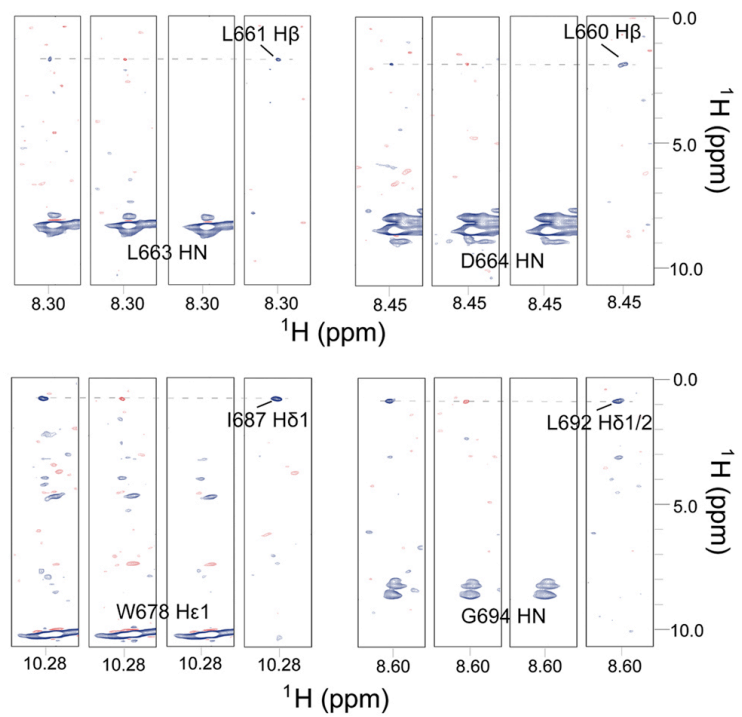
A**B****C****D**

Figure S3. Inter-protomer NOEs for the MPER. (A) Example strips from 3D ^{15}N -edited NOESY-TROSY-HSQC ($\tau_{\text{NOE}} = 200$ ms) spectra recorded using an isotopically mixed sample (left strips) and a control sample (right strips). The mixed sample consists of 0.4 mM (^{15}N , ^2H)-labeled MPER-TMD and 0.4 mM (15% ^{13}C)-labeled MPER-TMD, premixed at the cell level before cell lysis and protein purification. The control sample contains only 0.4 mM (^{15}N , ^2H)-labeled MPER-TMD. The spectra were recorded at ^1H frequency of 800 MHz. The inter-protomer NOE crosspeaks between amides and aliphatic protons due to isotope mixing are labeled. (B) Essentially complete NOE strips from the mixed and control samples. In panels A and B, the NOEs must have ^1H chemical shift > 0.3 ppm, but those with ^1H chemical shift between 2 and 4.4 ppm (green box) might have come from the protons of the DMPC/DHPC headgroups, which were not deuterated in our sample (only the acyl chains were perdeuterated). Although some appeared to be inter-chain NOEs, we chose not to use them, leaving a chemical shift region between 0.3 and 2 ppm (red box), within which most of the inter-chain NOEs were identified. The ^{15}N , ^2H labeled protein used for the mixed and control samples were from the same batch of expressing cells. Had the NOEs assigned as intermolecular been from site-specific ^1H contamination, they would also have appeared in the spectrum of the control sample. (C) 1D NMR spectra of the sample used for recording the reference NOE spectrum. In this sample, the MPER-TMD is (^{15}N , ^2H) labeled, and the acyl chains of DMPC and DHPC are perdeuterated. The left spectrum is a 1D ^1H spectrum, showing essentially no signals < 2 ppm. The right spectrum is a 1D ^2H spectrum, showing that the methylene and methyl groups in the sample are ^2H labeled. (D) Mixed NOE experiment with $J(^{13}\text{C}-^1\text{H})$ modulation. See Methods. Strips from 3D ^{15}N -edited NOESY-TROSY-HSQC ($\tau_{\text{NOE}} = 200$ ms) spectra recorded using an isotopically mixed sample that contained 0.8 mM MPER-TMD, as a 50:50 mixture of (^{15}N , ^2H)-labeled MPER-TMD and (^{13}C , ^1H)-labeled MPER-TMD. Two interleaved 3D spectra were recorded with ^1H - ^{13}C J evolution introduced before the NOE mixing ($J^{\text{evol1}} = 0$ ms, $J^{\text{evol2}} = 8$ ms). For each NOE pair, four panels are shown: Spectrum1 (with positive inter-NOEs, blue), Spectrum2 (with negative inter-NOEs, red), Summation of spectra 1 & 2 (inter-NOE peaks canceled), and Subtraction of spectra 1 & 2 (inter-NOE peaks enhanced and other peaks canceled). Upper left: NOEs between L663^{HN} and L661^{H β} ; upper right: NOEs between D664^{HN} and L660^{H β} ; lower left: NOEs between W678^{H ϵ 1} and I687^{H δ 1}; lower right: NOEs between G694^{HN} and L692^{H δ 1/2}.

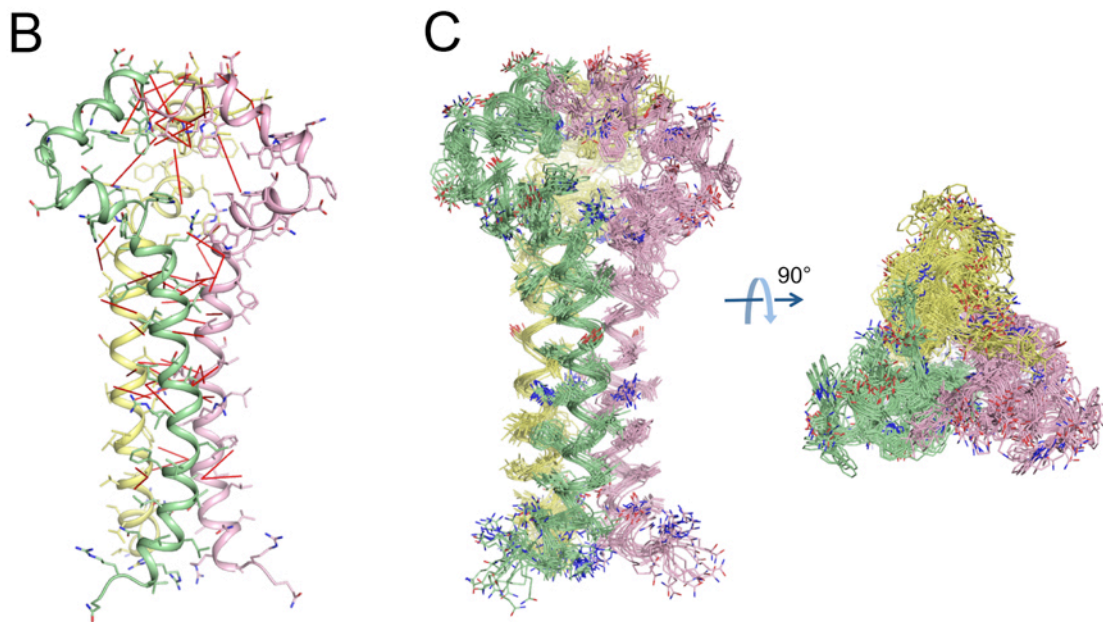
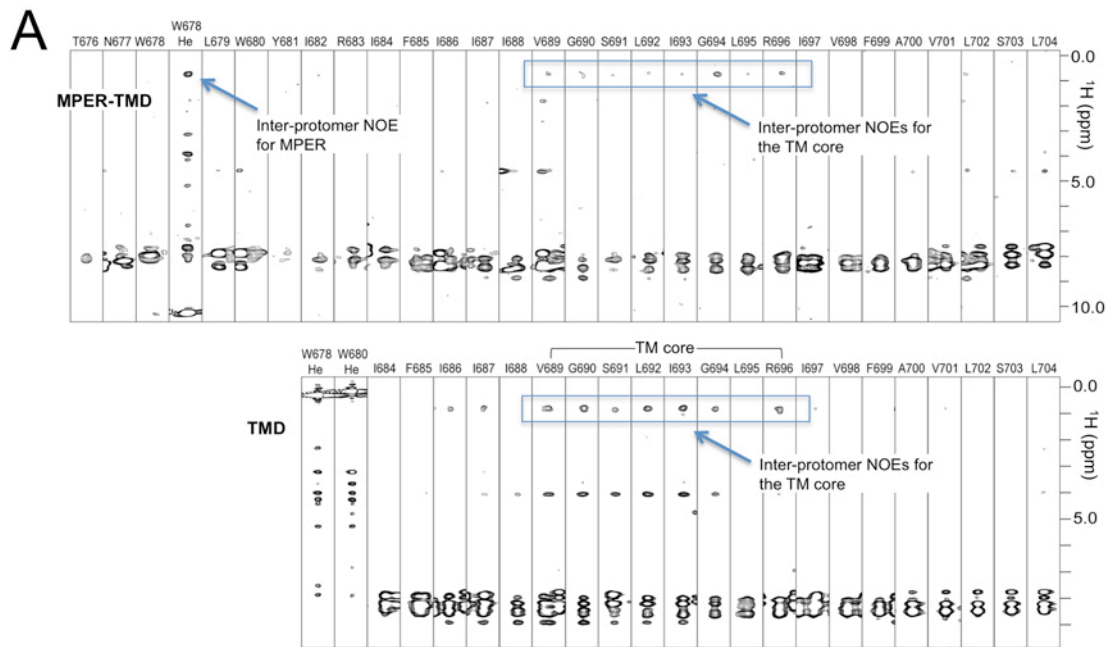


Figure S4. MPER-TMD and TMD alone. (A) Comparison of inter-protomer, TM-core NOEs for the MPER-TMD and the TMD samples. The NOE spectrum of the TMD sample was taken from Figure S5 of Dev J, *et al.* (2016). *Science* 353:172-175 (1). Reprinted with permission of AAAS. (B) Ribbon representation of the trimer structure showing NOE-derived intermonomer restraints (red lines). (C) Ensemble of 15 lowest energy structures calculated using NMR-derived structural restraints (Table S1). Structures are shown as thin ribbon representation of the backbones and stick representation of the side chains.

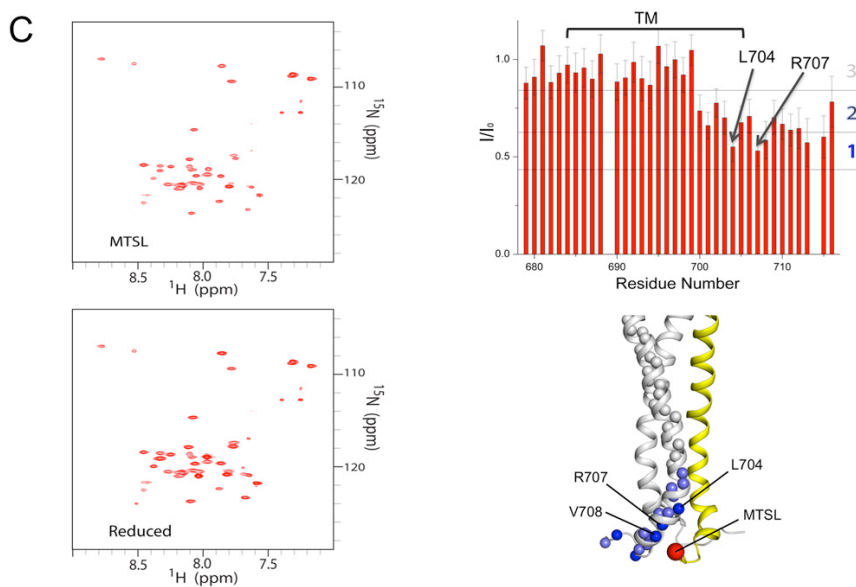
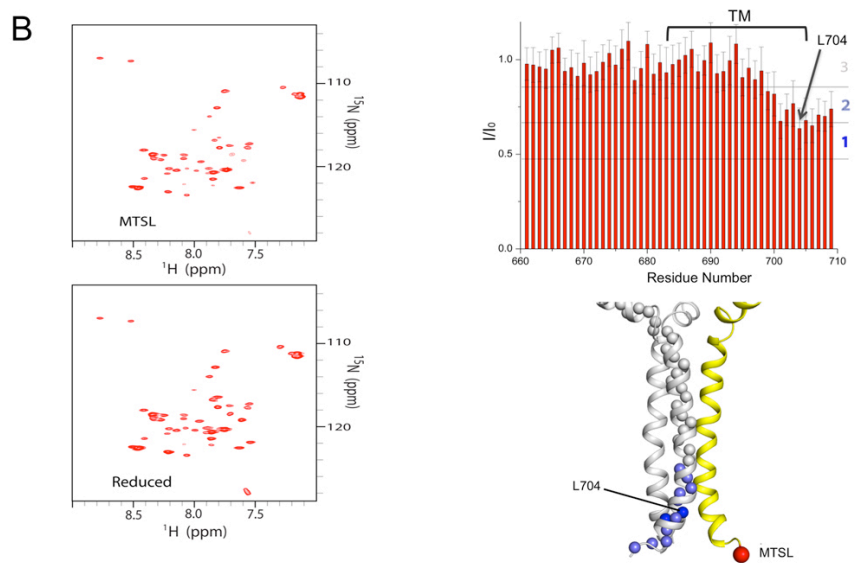
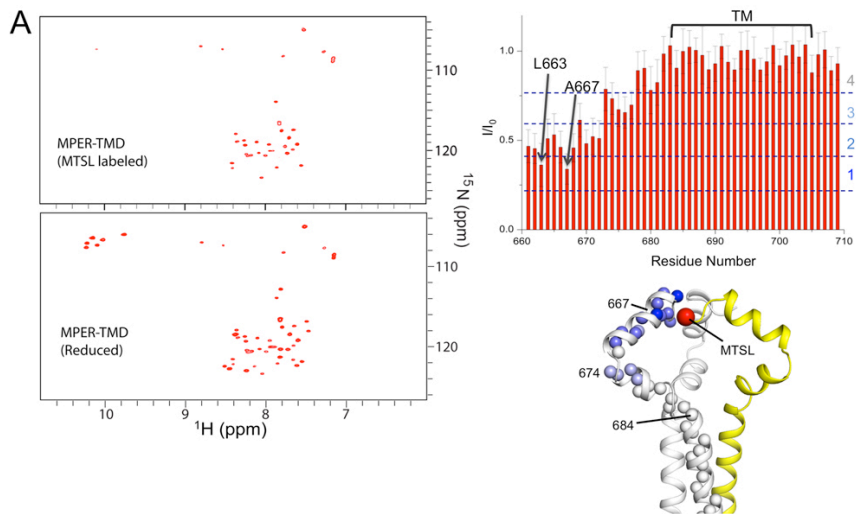


Figure S5. Mixed PRE data for the MPER-TMD and the TMD: both form homogeneous oligomers. The samples consist of ~1:1 ratio of (^{15}N , ~70% ^2H) labeled protein and unlabeled protein with MTSL near either the N- or the C-terminus of the construct, reconstituted in bicelles with $q \approx 0.5$.

(A) Mixed PRE data for the MPER-TMD with MTSL labeled at position 660. *Left:* ^1H - ^{15}N TROSY-HSQC spectra of the sample before (top) and after (bottom) the addition of 10 mM sodium ascorbate. *Top Right:* Residue-specific PRE, defined as the ratio of intensity before (I) and after (I_0) reducing with ascorbate. Regions between the dashed lines are defined as different PRE ranges. *Bottom Right:* A structural view showing the positions of backbone amide protons (spheres) that show strong PRE. Ribbon representations of the ^{15}N and MTSL labeled strands are shown in gray and yellow, respectively. The MTSL position is indicated by the red sphere. The amide protons are colored according to the PRE ranges defined in the top right panel. **(B)** Mixed PRE data for the MPER-TMD with MTSL labeled at position 710, presented in the same arrangement as in A. **(C)** Mixed PRE data for the TMD (residues 677-716) with MTSL labeled at position 716, presented in the same arrangement as in A.

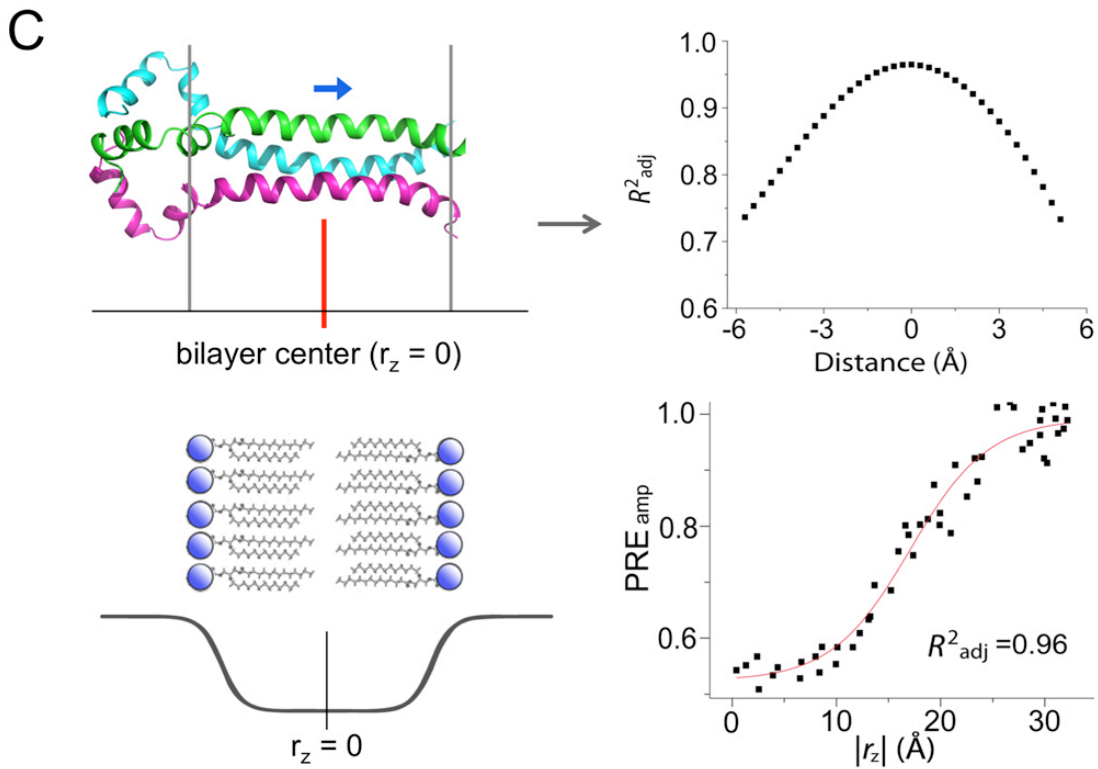
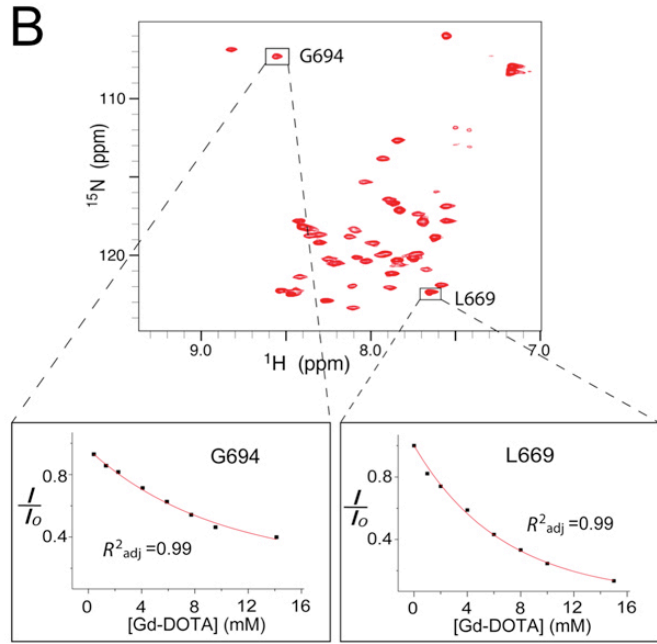
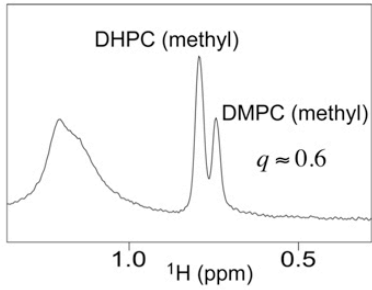
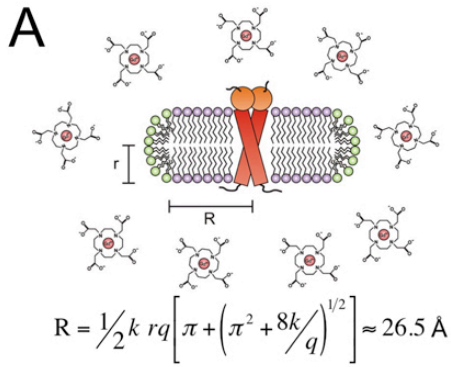
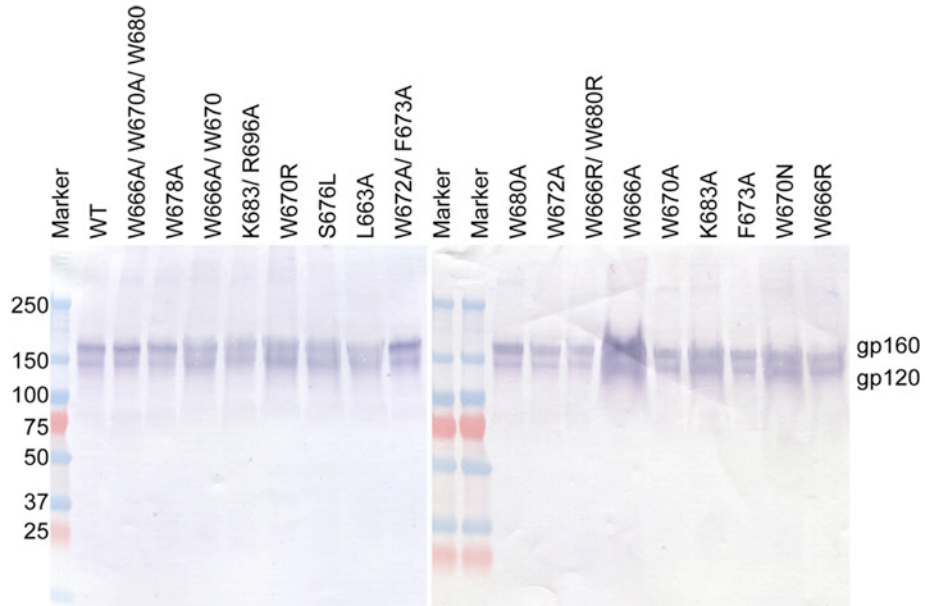
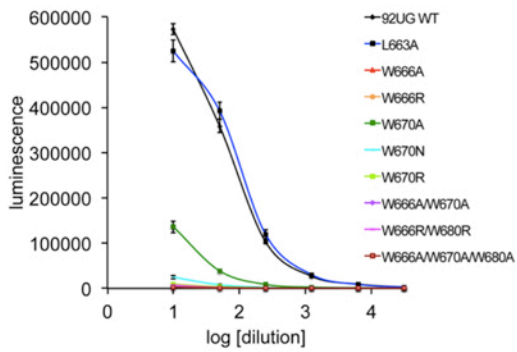


Figure S6. Solvent PRE analysis for characterizing membrane partition of the MPER-TMD. (A) Diagram illustrating titration of the bicelle-reconstituted MPER-TMD with Gd-DOTA (upper panel). The radius of the bilayer region of the bicelle (R) is given by the equation (14, 15) below the drawing, where r is the radius of the DHPC rim (20 Å), q is the molar ratio of DMPC to DHPC (0.6), and k is the ratio of the head group area of DMPC to that of DHPC (0.6). Lower panel is the 1D ^1H spectrum of the bicelle sample recorded at 600 MHz, showing that the peak intensity ratio of the DMPC terminal methyl groups to that of the DHPC terminal methyl groups is 0.6 (or $q = 0.6$). (B) PRE vs. [Gd-DOTA] for L669 and G694, measured with a series of ^1H - ^{15}N TROSY-HSQC spectra recorded at various Gd-DOTA concentrations. The data were fitted to the exponential decay function (Eq. 1 in Methods) to determine the PRE amplitude (PREamp). The PRE is defined as the ratio of peak intensity in the presence (I) and absence (I_0) of Gd-DOTA. The PREamp vs. residue number is shown in Fig. 2A. (C) Assignment of the bilayer center to the MPER-TMD by data fitting. Left: Illustration for sliding the MPER-TMD along the 3-fold axis (or bilayer normal) to yield best fit to the symmetric sigmoidal function (Eq. 2 in SI Methods). The $r_z = 0$ of the sigmoidal function corresponds to the bilayer center. Right upper: The adjusted coefficient of determination (R^2_{adj}) from data fitting versus deviation from the true bilayer center. The plot shows that R^2_{adj} is a reliable indicator of the protein position with an error of about ± 0.3 Å. Right lower: The best fit of the PREamp vs. r_z data to the symmetric sigmoidal function. The same fit is shown in Fig. 2B.

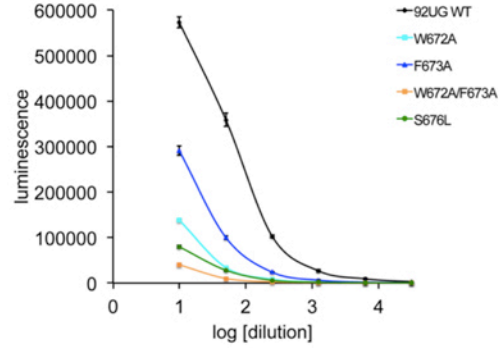
A



B



C



D

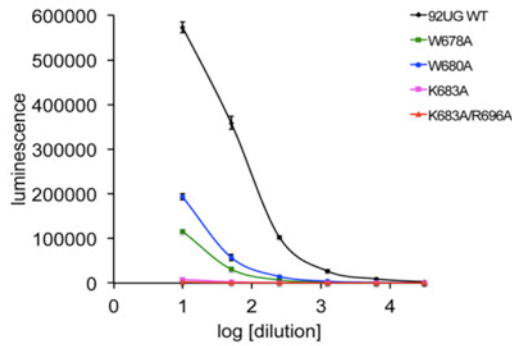
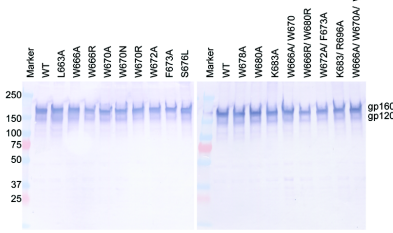


Figure S7. Viral infectivity of Env mutants. (A) Incorporation of Env mutants into pseudoviruses analyzed by western blot. Env samples prepared from p24-normalized pseudoviruses containing either HIV-1 gp92UG037.8 Env or each of its MPER mutants were detected by an anti-V3 antibody 3791. (B-D) Viral infectivity. Pseudoviruses containing Env MPER mutants were normalized by p24-antigen, titrated 10 times using 5-fold dilution series, and tested for viral infectivity in TZM.bl cells. The Env mutants are grouped in (B) for those containing mutations in the hydrophobic core; in (C) containing mutations in the turn; and in (D) containing mutations in the kink. The experiment was carried out in quadruplicate. Error bars indicate the standard deviation calculated by Excel.

A



C

Mutation	Mean Fluorescence Intensity (MFI)		
	2G12	VRC01	PG16
WT	1254	820	648
L663A	1482	960	737
W666A	578	395	351
W666R	1044	794	526
W670A	1214	918	526
W670N	1039	776	400
W670R	505	347	222
W672A	820	805	454
F673A	1124	893	574
T676L	930	641	383
W678A	1099	841	497
W680A	1011	755	496
K683A	711	517	243
W666A/W670A	736	494	288
W666R/W680R	867	649	371
W672A/F673A	1374	1169	454
K683A/R696A	930	715	301
W666A/W670A/W680A	910	646	347
293T	21	10	10

B

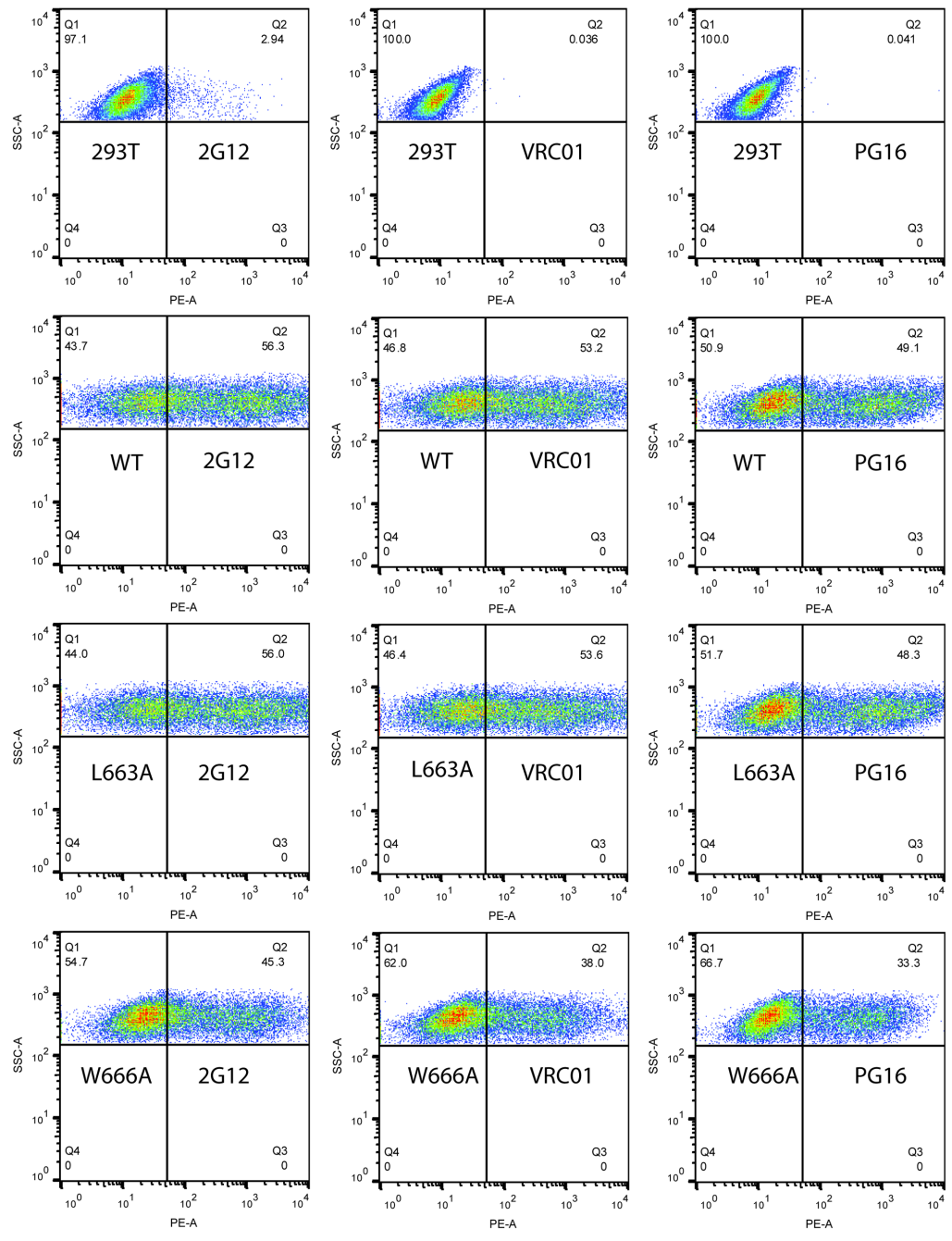


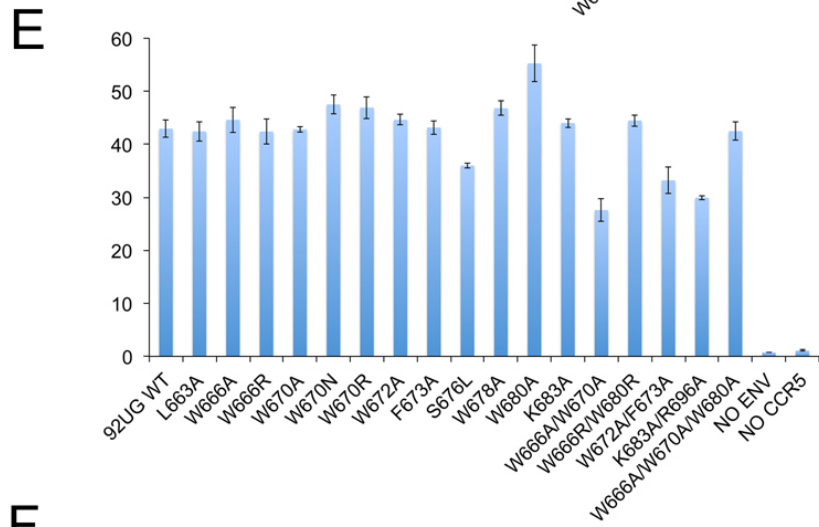
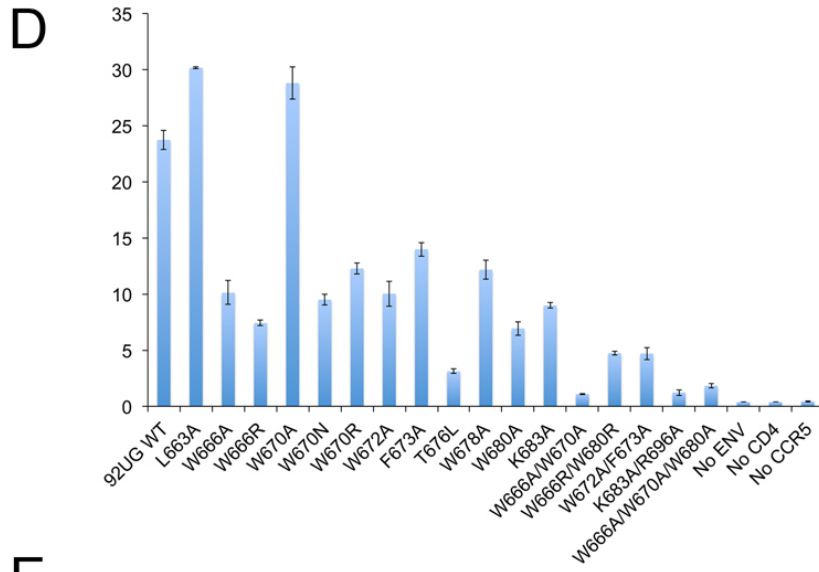
Figure S8. Expression and cell-cell fusion of HIV-1 Env and its MPER mutants.

(A-C)

(A) Expression and processing of Env mutants expressed in 293T cells. Env samples prepared from 293T cells transiently transfected with 1 µg of the HIV-1 gp92UG037.8 gp160 expression plasmid or each of MPER mutants were detected by an anti-V3 antibody 3791.

(B) Cell-surface expression of Env mutants detected by flow cytometry. Representative dot plots for negative control 293T cells, cells expressing the wildtype Env (WT), or mutants L663A and W666A, measured by flow cytometry using monoclonal antibodies 2G12, VRC01 and PG16.

(C) Summary of mean fluorescence intensity (MFI) from data shown in (B) for the wildtype and 17 MPER mutants, as well as for the 293T cell control.



F

Mutation	Structural element	Viral infectivity ^a	Cell-cell fusion ^b (low Env level)	Cell-cell fusion (high Env level)
WT	n/a	100.0±2.1	100.0±3.6	100.0±3.9
L663A	hydrophobic core	91.5±4.2	127.1±0.4	98.8±4.3
W666A		0.3±0.1	42.6±1.9	103.9±5.8
W666R		0.4±0.1	31.2±0.3	98.7±5.4
W670A		23.7±2.3	121.3±7.4	99.7±1.2
W670N		4.3±0.7	40.0±0.8	110.6±4.5
W670R		1.6±0.1	51.7±1.0	109.3±5.3
W666A/W670A		0.8±0.1	4.6±0.0	64.3±3.2
W666R/W680R		1.2±0.0	19.9±0.1	103.6±2.6
W666A/W670A/W680A		0.6±0.0	7.7±0.1	98.9±4.0
W672A		turn	24.2±0.6	42.3±2.0
F673A	50.8±1.9		58.9±1.5	100.5±3.0
W672A/F673A	6.9±0.3		19.8±0.5	77.3±4.4
S676L	kink	13.9±0.5	13.1±0.1	83.8±1.0
W678A		20.2±0.5	51.3±1.8	109.0±3.5
W680A		33.7±1.1	29.2±0.7	128.7±10.5
K683A		1.4±0.2	37.8±0.4	102.5±2.0
K683A/R696A		0.3±0.0	5.2±0.1	69.7±0.7

Figure S8. Expression and cell-cell fusion of HIV-1 Env and its MPER mutants.

(D-F)

(D) Cell-cell fusion of HIV-1 Env and its MPER mutants at a low expression level. 293T cells transfected with 75 ng of the 92UG037.8 Env expression plasmid or each of its TMD mutants were mixed with CD4- and CCR5-expressing cells. Cell-cell fusion led to reconstitution of active β -galactosidase; fusion activity was quantified by a chemiluminescent assay. No Env, no CD4 and no CCR5 were negative controls. The experiment was carried out in triplicate and repeated at least twice with similar results. Error bars indicate the standard deviation calculated by the Excel STDEV function. Data points are means \pm standard deviations from triplicate measurements. The data are also summarized in Fig. S8F.

(E) Cell-cell fusion of HIV-1 Env and its MPER mutants at a high expression level. 293T cells transfected with 10 μ g of the 92UG037.8 Env expression plasmid or each of its MPER mutants were fused with CD4- and CCR5-expressing cells. No Env, no CCR5 were negative controls. The experiment was carried out in triplicate and repeated at least twice with similar results. Error bars indicate the standard deviation calculated by the Excel STDEV function. Data points are means \pm standard deviations from triplicate measurements.

(F) Functional analysis of mutations in the HIV-1 TMD. ^a Viral infectivity was determined using p24-normalized pseudoviruses; titration data for each virus with 5-fold dilution series are also shown in Fig. S7B-D. ^b Cell-cell fusion assay was performed at both high and low Env expression levels, as displayed in panels **D** and **E**.

Table S1. NMR and refinement statistics

NMR distance and dihedral constraints^a	MPER (660-683)	TM (684-710)
Distance constraints from NOE ^b	282	432
Short-range intramolecular ($ i-j \leq 4$)	237	390
Long-range intramolecular ($ i-j \geq 5$)	21	0
Intermolecular	24	42
Total dihedral angle restraints ^c	252	
ϕ (TALOS)	126	
ψ (TALOS)	126	
Structure statistics^d		
Violations (mean \pm s.d.)		
Distance constraints (\AA)	0.155 \pm 0.008	
Dihedral angle constraints ($^\circ$)	1.281 \pm 0.095	
Deviations from idealized geometry		
Bond lengths (\AA)	0.032 \pm 0.000	
Bond angles ($^\circ$)	1.641 \pm 0.019	
Improper ($^\circ$)	0.931 \pm 0.039	
Average pairwise r.m.s. deviation (\AA) ^e		
Heavy	1.714	
Backbone	1.192	

^a The numbers of constraints are summed over all three subunits.

^b For the MPER, all NOE restraints were obtained in the current study. For the TM, the NOE restraints for residues 685 – 701 (total 204) were taken from our earlier NMR study of the TMD, and the rest were obtained in the current study.

^c Backbone ϕ and ψ restraints and their respective uncertainties were obtained from the “GOOD” dihedrals generated by the TALOS+ program (6) based on the backbone chemical shift values.

^d Statistics are calculated and averaged over an ensemble of the 15 lowest energy structures out of 150 calculated structures.

^e The precision of the atomic coordinates is defined as the average r.m.s. difference between the 15 final structures and their mean coordinates.

References

1. Dev J, *et al.* (2016) Structural basis for membrane anchoring of HIV-1 envelope spike. *Science* 353(6295):172-175.
2. Delaglio F, *et al.* (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6:277-293.
3. Bartels C, Xia TH, Billeter M, Guntert P, & Wuthrich K (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *Journal of biomolecular NMR* 6(1):1-10.
4. Salzmann M, Wider G, Pervushin K, & Wuthrich K (1999) Improved sensitivity and coherence selection for [15N,1H]-TROSY elements in triple resonance experiments. *Journal of biomolecular NMR* 15(2):181-184.
5. Szyperski T, Neri D, Leiting B, Otting G, & Wuthrich K (1992) Support of 1H NMR assignments in proteins by biosynthetically directed fractional 13C-labeling. *Journal of biomolecular NMR* 2(4):323-334.
6. Shen Y, Delaglio F, Cornilescu G, & Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR* 44(4):213-223.
7. Schwieters CD, Kuszewski JJ, Tjandra N, & Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *Journal of magnetic resonance* 160(1):65-73.
8. Chiliveri SC, Louis JM, Ghirlando R, Baber JL, & Bax A (2018) Tilted, Uninterrupted, Monomeric HIV-1 gp41 Transmembrane Helix from Residual Dipolar Couplings. *J Am Chem Soc* 140(1):34-37.
9. Piai A, Fu Q, Dev J, & Chou JJ (2017) Optimal Bicelle Size q for Solution NMR Studies of the Protein Transmembrane Partition. *Chemistry* 23(6):1361-1367.
10. Chen J, *et al.* (2015) HIV-1 ENVELOPE. Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. *Science* 349(6244):191-195.
11. Kovacs JM, *et al.* (2012) HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120. *Proc Natl Acad Sci U S A* 109(30):12111-12116.
12. Li M, *et al.* (2005) Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol* 79(16):10108-10125.
13. Sarzotti-Kelsoe M, *et al.* (2014) Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J Immunol Methods* 409:131-146.
14. Sanders CR, 2nd & Schwonek JP (1992) Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry* 31(37):8898-8905.
15. Glover KJ, *et al.* (2001) Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophysical journal* 81(4):2163-2171.