Supporting Information

Stability and water accessibility of the trimeric membrane anchors of the HIV-1 envelope spikes

Alessandro Piai, Jyoti Dev, Quingshan Fu, and James J. Chou

Department of Biological Chemistry and Molecular Pharmacology
Harvard Medical School
Boston, Massachusetts 02115, USA.

Methods

Sample preparation

The (^{15}N, ^{2}H)-labeled gp41$^{HIV1D(677-716)}$ (gp41^{TMD}) was expressed and purified as previously described[1]. The lyophilized protein (~1.7 mg) was dissolved in hexafluoro-isopropanal (HFIP) with approximately 16 mg of DMPC and 27 mg of DHPC, followed by drying of the solution under a nitrogen stream to achieve a thin film. The thin film was then dissolved in 3 ml of an 8 M urea solution and dialyzed against a 40 mM MES buffer (pH 6.0) to remove the denaturant. After dialysis, additional DHPC was added to make up for the loss during dialysis and to adjust the DMPC:DHPC ratio (q) to approximately 0.5. Finally, the solution with reconstituted gp41^{TMD} was concentrated using a centricon to 360 μL. The final NMR sample contained ~1 mM HIV-1 gp41^{TMD}, 60 mM DMPC, 120 mM DHPC, 40 mM MES (pH 6.0), and 10% (v/v) D_2O (for the NMR lock). The bicelle q of the NMR sample was determined by integrating the resolved methyl peaks of DMPC and DHPC in the 1D ^{1}H NMR spectrum and adjusted to be exactly 0.5 (Fig. S2b).

NMR data acquisition, processing, and analysis

The NMR experiments were performed at 14.1 T on a Bruker Avance III HD spectrometer operating at 600.13 MHz ^{1}H, 150.90 MHz ^{13}C, and 60.81 MHz ^{15}N frequencies, equipped with a cryogenically cooled probe head. All the measurements were performed at 303 K. The most relevant acquisition parameters of the experiments are reported in Table S1.

The NMR data sets were processed using *nmrPipe*[2] and the resulting NMR spectra were analyzed with *Sparky* (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) and *CcpNMR Analysis*[3]. Peak intensities were measured at peak local maxima using quadratic interpolation to identify peak centers. *Origin* (OriginLab, Northampton, MA) was used to fit the experimental data. The chemical shift assignment of the HIV-1 gp41^{TMD} was taken from the *Biological Magnetic Resonance Bank (BMRB)*[4], entry 30090[1]; the structure of the protein was taken from the *Protein Data Bank* (PDB)[5], entry 5JYN[1].

Solvent PRE analysis

The membrane partition of gp41^{TMD} was determined using a method we previously developed[6]. DMPC/DHPC bicelle with sufficiently large q (≥ 0.5) allows direct use of measurable solvent paramagnetic relaxation enhancement (PRE) to probe residue-specific depth immersion of the protein in the bilayer region of the bicelle (Fig. S2). Therefore, the bicelle-reconstituted gp41^{TMD} was titrated with

the water-soluble and membrane-inaccessible paramagnetic agent Gd-DOTA at various known concentrations. The titrant was taken from a concentrated stock solution (600 mM Gd-DOTA) in the same buffer as that of the protein sample and it was added in small aliquots (few μL per step) to minimize sample dilution. The progress of the titration was monitored by measuring a 2D ^{1}H-^{15}N TROSY-HSQC[7] spectrum at each of the following titrant concentrations: 0 (reference), 2.0, 4.0, 6.0, 8.0, 10.0, 15.0 and 20.0 mM. The residue-specific PRE_{amp}, which is the amplitude of PRE experienced by an amide proton in the protein, was determined by fitting the peak intensity decay as a function of [Gd-DOTA] to the following exponential decay equation:

$$\frac{I}{I_0} = 1\text{-}PRE_{amp}\left(1 - e^{-\frac{[\text{Gd-DOTA}]}{\tau}}\right), \qquad \text{(Eq. S1)}$$

where I and I_0 are the peak intensities in the presence and absence of the paramagnetic agent, respectively, [Gd-DOTA] is the concentration of the paramagnetic agent, τ is the decay constant, and PRE_{amp} is the PRE amplitude.

To determine the position of the gp41^{TMD} relative to the bilayer center, we calculated, for each residue i, the distance (r_z) along the protein symmetry axis (parallel to the bilayer normal), from the amide proton to an arbitrary choice of reference point based on the NMR structure. As such, the PRE_{amp} vs (residue number) plot was converted to PRE_{amp} vs r_z, which was then analyzed using the sigmoidal fitting method (details can be found in Ref. [6]). Briefly, the gp41^{TMD} structure was moved along the bilayer normal (Fig. S3) to achieve the best fit to the symmetric sigmoid equation:

$$PRE_{amp} = PRE_{amp}^{min} + \frac{(\text{PRE}_{\text{amp}}^{\text{max}} - \text{PRE}_{\text{amp}}^{\text{min}})}{1+e^{(r_Z^I - |rz|)/SLOPE}} \qquad \text{(Eq. S2)}$$

where PRE_{amp}^{min} and PRE_{amp}^{max} are the limits within which PRE_{amp} can vary for a particular protein system, r_z^I is the inflection point (the distance from the bilayer center at which PRE_{amp} is halfway between PRE_{amp}^{min} and PRE_{amp}^{max}), and *SLOPE* is a parameter which reports the steepness of the curve at the inflection point. The protein position along the bilayer normal which yielded the best fit to Eq. S2 (or the highest R^2_{adj}) provided the best placement of the gp41^{TMD} with respect to the bilayer center (defined as $r_Z = 0$).

Hydrogen-deuterium (H-D) exchange

The gp41^{TMD}, reconstituted in protonated solvent, was flash-frozen in liquid nitrogen and then thoroughly lyophilized. The dried sample was dissolved in 360 μL of 99.9% D_2O. The progress of the H-D exchange was monitored by measuring a 2D ^{1}H-^{15}N TROSY-HSQC spectrum at uniform time intervals of ~3 hours

up to ~4.6 days. The residue-specific exchange constant, k_{ex} (=$1/\tau_{ex}$), was determined by fitting the fractional peak intensity vs. time to the following exponential decay equation:

$$\frac{I}{I_0} \propto e^{-\frac{t}{\tau_{ex}}} \quad , \qquad \text{(Eq. S3)}$$

where I_0 and I are the peak intensities before and after the H-D exchange, t is the time passed from the beginning of the exchange, and τ_{ex} is the time constant of the decay.

Backbone dynamics

The backbone dynamics of the gp41^{TMD} in bicelles was examined by measuring the ^{15}N R_1 and R_2 relaxation rates. The R_1 and R_2 were measured using the TROSY version of the standard experiments for measuring ^{15}N R_1 and R_2 [7-9]. For the determination of R_1, 9 experiments were acquired with the following relaxation delays: 0 (reference), 10, 50, 100, 200, 300, 600, 800 and 1000 ms. For the determination of R_2, 9 experiments were acquired with the following relaxation delays: 0 (reference), 6.4, 10, 20, 30, 40, 50, 64 and 80 ms. The R_1 and R_2 values were determined by fitting the peak intensity vs. relaxation delay to the exponential decays:

$$\frac{I}{I_0} \propto e^{-R_1 t} \quad , \qquad \text{(Eq. S4)}$$

$$\frac{I}{I_0} \propto e^{-R_2 t} \quad , \qquad \text{(Eq. S5)}$$

where I is the peak intensity at a given relaxation delay, I_0 is the peak intensity in the reference spectrum ($t = 0$), t is the relaxation delay, and R_1 and R_2 are the relaxation rates.

Table S1. NMR acquisition parameters.

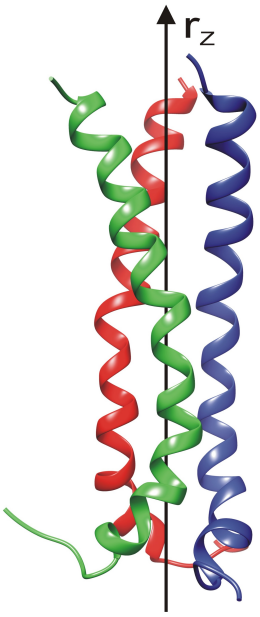
Type of experiment	Spectral widths and chemical shift evolution times		# of scans	Inter-scan delay (s)	Duration of the experiment
Gd-DOTA titration					
2D ^{1}H-^{15}N TROSY-HSQC	1300 Hz (^{15}N) 77.0 ms	9600 Hz ($^{1}H^{N}$) 106.5 ms	32	3.5	6 hours 40 min
H-D exchange					
2D ^{1}H-^{15}N TROSY-HSQC	1600 Hz (^{15}N) 62.5 ms	9600 Hz ($^{1}H^{N}$) 106.5 ms	40	1.2	3 hours 10 min
^{15}N T_1					
2D ^{1}H-^{15}N T_1-TROSY-HSQC	1600 Hz (^{15}N) 62.5 ms	9600 Hz ($^{1}H^{N}$) 106.5 ms	32	1.5	3 hours 10 min
^{15}N T_2					
2D ^{1}H-^{15}N T_2-TROSY-HSQC	1600 Hz (^{15}N) 62.5 ms	9600 Hz ($^{1}H^{N}$) 106.5 ms	32	1.5	3 hours 15 min

Table S2. Residue-specific PRE_{amp} of the HIV-1 gp41^{TMD} in bicelles [*].

Residue	PRE_{amp}	R^2_{adj}
LEU 679	0.906±0.086	0.940
TRP 680	0.924±0.054	0.987
TYR 681	0.932±0.078	0.958
ARG 683	0.774±0.041	0.995
ILE 684	0.763±0.035	0.990
ILE 686	0.613±0.050	0.956
ILE 688	0.604±0.024	0.991
VAL 689	0.545±0.035	0.979
GLY 690	0.600±0.027	0.986
SER 691	0.618±0.065	0.936
LEU 692	0.555±0.049	0.948
ILE 693	0.628±0.032	0.987
GLY 694	0.623±0.043	0.972
LEU 695	0.545±0.058	0.956
ARG 696	0.598±0.022	0.992
VAL 698	0.545±0.039	0.967
PHE 699	0.629±0.072	0.919
ALA 700	0.653±0.035	0.988
VAL 701	0.646±0.036	0.981
LEU 702	0.657±0.062	0.951
SER 703	0.714±0.043	0.981
LEU 704	0.730±0.048	0.975
VAL 705	0.771±0.048	0.989
ASN 706	0.787±0.036	0.989
ARG 707	0.873±0.021	0.997
VAL 708	0.840±0.038	0.988
GLN 710	0.899±0.036	0.991
GLY 711	0.939±0.013	0.999
TYR 712	0.946±0.019	0.997
SER 713	0.940±0.025	0.995
LEU 715	0.884±0.021	0.996
SER 716	0.912±0.022	0.996

[*] Residue-specific PRE_{amp} were determined by fitting I/I_0 vs. [Gd-DOTA] to Eq. S1. The adjusted coefficient of determination (R^2_{adj}) was used to evaluate the quality of the fittings. The R^2_{adj} parameter is a measure of how well the model describes the experimental data.

Table S3. Residue-specific membrane partition of the HIV-1 gp41^{TMD}.



r_z (Å)	Residue (H^N)
25.40	TRP 678
24.07	LEU 679
23.46	TRP 680
21.51	TYR 681
19.66	ILE 682
19.11	ARG 683
17.65	ILE 684
15.51	PHE 685
14.01	ILE 686
13.01	ILE 687
11.20	ILE 688
9.28	VAL 689
8.48	GLY 690
6.67	SER 691
4.34	LEU 692
2.98	ILE 693
1.76	GLY 694
0.00	*Membrane Center*
-0.35	LEU 695
-2.38	ARG 696
-4.29	ILE 697
-6.58	VAL 698
-8.39	PHE 699
-10.02	ALA 700
-11.72	VAL 701
-13.07	LEU 702
-15.16	SER 703
-17.56	LEU 704
-18.19	VAL 705
-18.87	ASN 706
-21.33	ARG 707
-23.22	VAL 708
-22.79	ARG 709
-23.75	GLN 710
-26.08	GLY 711
-26.83	TYR 712
-28.88	SER 713
-25.81	LEU 715
-25.56	SER 716

Data not available for Asn677 (N-terminus) and Pro714

Table S4. Residue-specific k_{ex} of the HIV-1 gp41^{TMD} in bicelles [†].

Residue (H^N)	k_{ex} (Hz)	τ_{ex}	
TRP 680	*	*	
TYR 681	*	*	
ILE 682	$(2.9\pm1.5)\cdot10^{-4}$	1.0±0.5 hours	
ILE 684	$(1.9\pm1.8)\cdot10^{-4}$	1.5±0.1 hours	
PHE 685	$(9.9\pm3.6)\cdot10^{-5}$	2.8±1.0 hours	
ILE 686	$(6.1\pm1.3)\cdot10^{-6}$	1.9±0.4 days	
ILE 687	$(10.0\pm3.1)\cdot10^{-6}$	1.2±0.4 days	
ILE 688	$(8.7\pm2.4)\cdot10^{-6}$	1.3±0.4 days	
VAL 689	$(7.3\pm1.4)\cdot10^{-6}$	1.6±0.3 days	
GLY 690	$(3.4\pm1.3)\cdot10^{-5}$	8.2±0.3 hours	
SER 691	$(2.3\pm0.6)\cdot10^{-4}$	1.2±0.3 hours	
LEU 692	*	*	
ILE 693	$(2.6\pm0.4)\cdot10^{-4}$	1.1±0.2 hours	
GLY 694	$(2.6\pm1.6)\cdot10^{-4}$	1.1±0.7 hours	
LEU 695	*	*	
ARG 696	*	*	
ILE 697	$(6.7\pm2.4)\cdot10^{-5}$	4.2±1.5 hours	
VAL 698	$(5.4\pm0.2)\cdot10^{-5}$	5.1±0.2 hours	
PHE 699	$(2.6\pm0.7)\cdot10^{-4}$	1.1±0.3 hours	
ALA 700	$(1.3\pm0.6)\cdot10^{-4}$	2.1±1.0 hours	
VAL 701	$(1.8\pm0.2)\cdot10^{-5}$	15.6±1.3 hours	
LEU 702	$(7.9\pm0.2)\cdot10^{-5}$	3.5±0.1 hours	
SER 703	*	*	
LEU 704	$(1.4\pm0.6)\cdot10^{-4}$	2.0±0.8 hours	
ASN 706	*	*	
ARG 707	*	*	
VAL 708	*	*	
ARG 709	*	*	
GLN 710	$(1.4\pm0.4)\cdot10^{-4}$	2.0±0.6 hours	
GLY 711	*	*	
TYR 712	*	*	
SER 713	$(2.2\pm0.6)\cdot10^{-4}$	1.3±0.3 hours	
LEU 715	*	*	
SER 716	*	*	

Data not available for Asn677 (N-terminus), Trp678, Leu679, Arg683, Val705, and Pro714.

'*' indicates residues with k_{ex} too fast to be measured.

[†] The colors in the last column represent the four different exchange regimes defined as: very fast (τ_{ex} < 1 hour) (red), fast (1 hour ≤ τ_{ex} < 3 hours) (orange), slow (3 hours ≤ τ_{ex} < 1 day) (yellow), and very slow (τ_{ex} ≥ 1 day) (white).

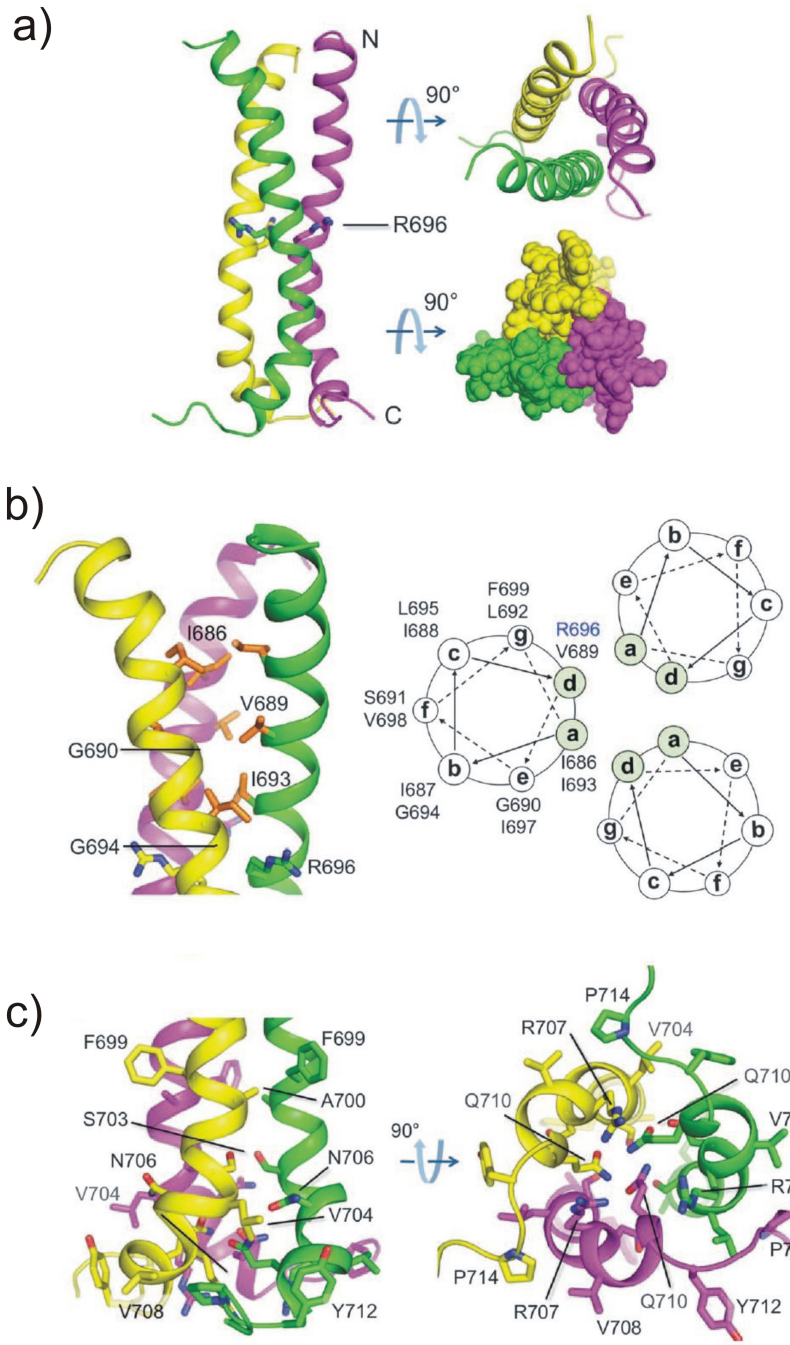


Figure S1. The NMR structure of the HIV-1 gp41^{TMD} in bicelles. (**a**) Ribbon representation of the trimer showing the side-chains of the R696. The sphere representation of the top view (lower right) shows that the trimer has no ion-permeable holes. (**b**) The hydrophobic core of the N-terminal half of the structure with hydrophobic residues (orange) arranged in the coiled-coil pattern (right panel). (**c**) The hydrophilic core in the C-terminal half of the structure, showing an array of polar residues. The figure was taken and adapted from Ref. [1].

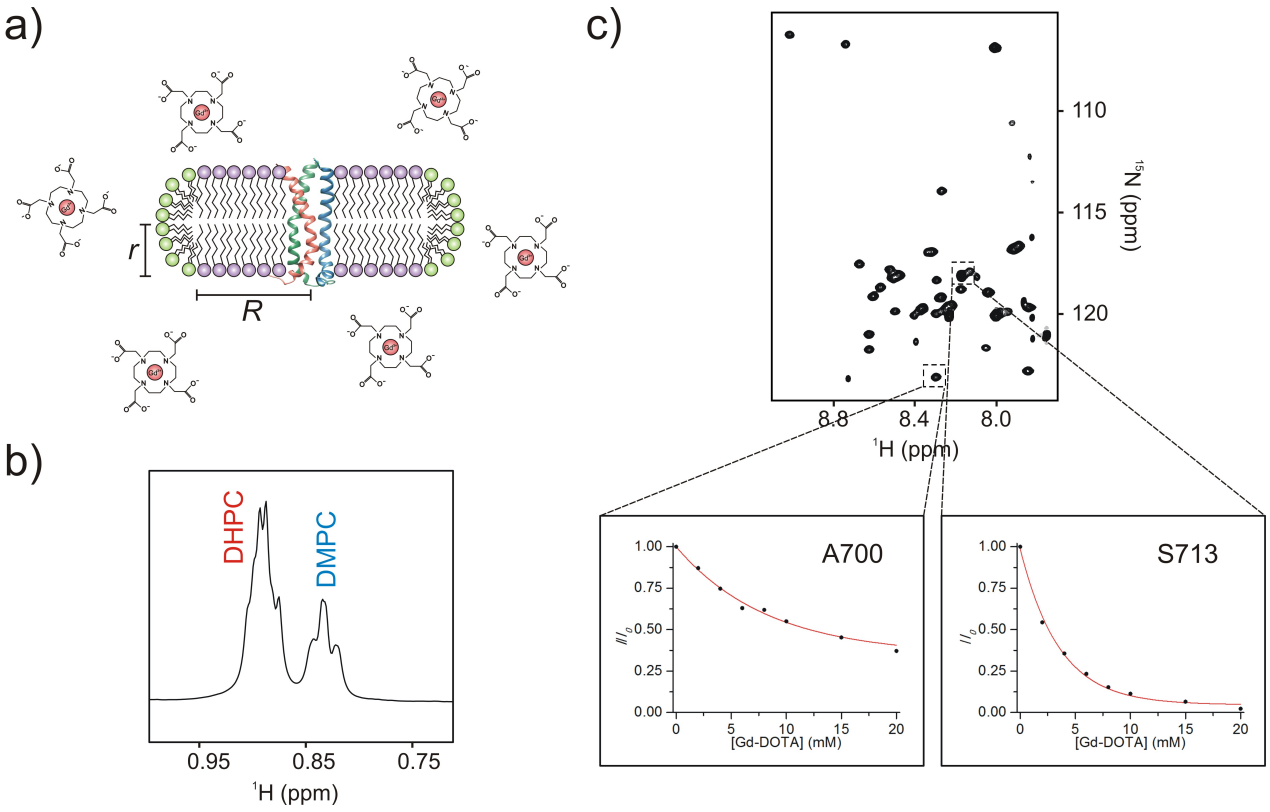


Figure S2. Solvent PRE analysis for determining the membrane partition of the HIV-1 gp41^{TMD} in bicelles. (**a**) Schematic illustration of a bicelle-reconstituted gp41^{TMD} surrounded by Gd-DOTA molecules. The radius of the planar region of the bicelle (R) is given by the equation $R = 1/2\, krq\left[\pi + (\pi^2 + 8k/q)^{1/2}\right]$, where r is the radius of the DHPC rim (20 Å), q is the molar ratio of DMPC to DHPC, and k is the ratio of the head group area of DMPC to that of DHPC[10-12]. (**b**) The zoomed region of the 1D ^{1}H NMR spectrum of the bicelle sample, indicating that the molar ratio of DMPC to DHPC (q) is 0.5. (**c**) The 2D ^{1}H-^{15}N TROSY-HSQC spectrum of the gp41^{TMD} in bicelles with q = 0.5. As examples, I/I_0 vs. [Gd-DOTA] plots of A700 (buried) and S713 (exposed) are shown. The data fitting (red line) to the exponential decay function (Eq. S2) yielded PRE_{amp} for the two residues.

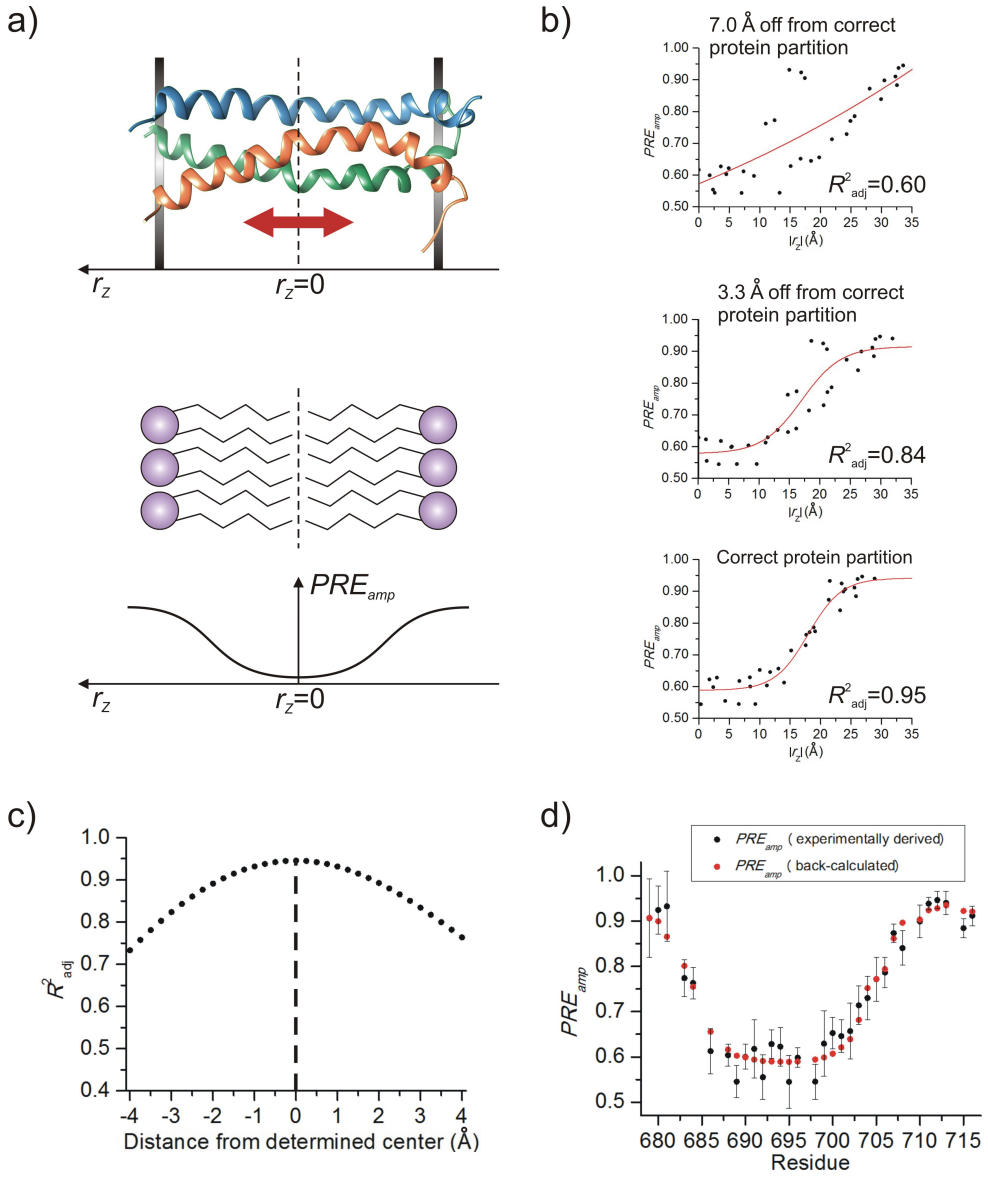


Figure S3. Position of the HIV-1 gp41^{TMD} relative to the bilayer center determined by data fitting. (**a**) Illustration showing the sliding of the gp41^{TMD} structure along the bilayer normal (r_Z axis) to yield the best fit to the symmetric sigmoidal function (Eq. S2). The r_Z = 0 of the sigmoidal function corresponds to the bilayer center. (**b**) The sigmoidal fittings of the PRE_{amp} vs. r_Z data from both sides of the membrane for different protein positions along the bilayer normal. The quality of the fitting (R^2_{adj}) improves as the protein is moved closer to the correct position in the bilayer. (**c**) The R^2_{adj} vs. r_Z plot showing the sensitivity of R^2_{adj} to deviations (± 4 Å) from the r_Z value that yielded the best R^2_{adj}. The plot shows that R^2_{adj} is a reliable indicator of the bilayer center with an approximate error of ± 0.5 Å. (**d**) Comparison between experimentally derived (black) and back-calculated (red) PRE_{amp}.

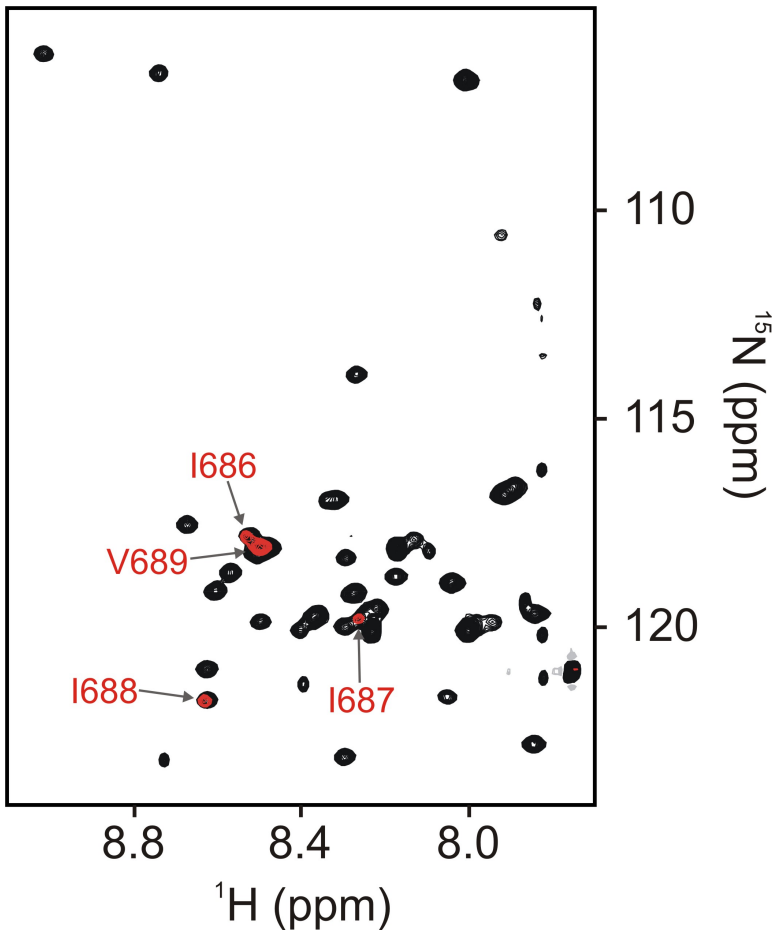


Figure S4. Example NMR spectrum showing the H-D exchange of the HIV-1 gp41^{TMD}. The 2D ^{1}H-^{15}N TROSY-HSQC spectra recorded before and after 4.6 days of H-D exchange are shown in black and red, respectively. The residues forming the hydrophobic core of the protein are marked in the overlaid spectra.

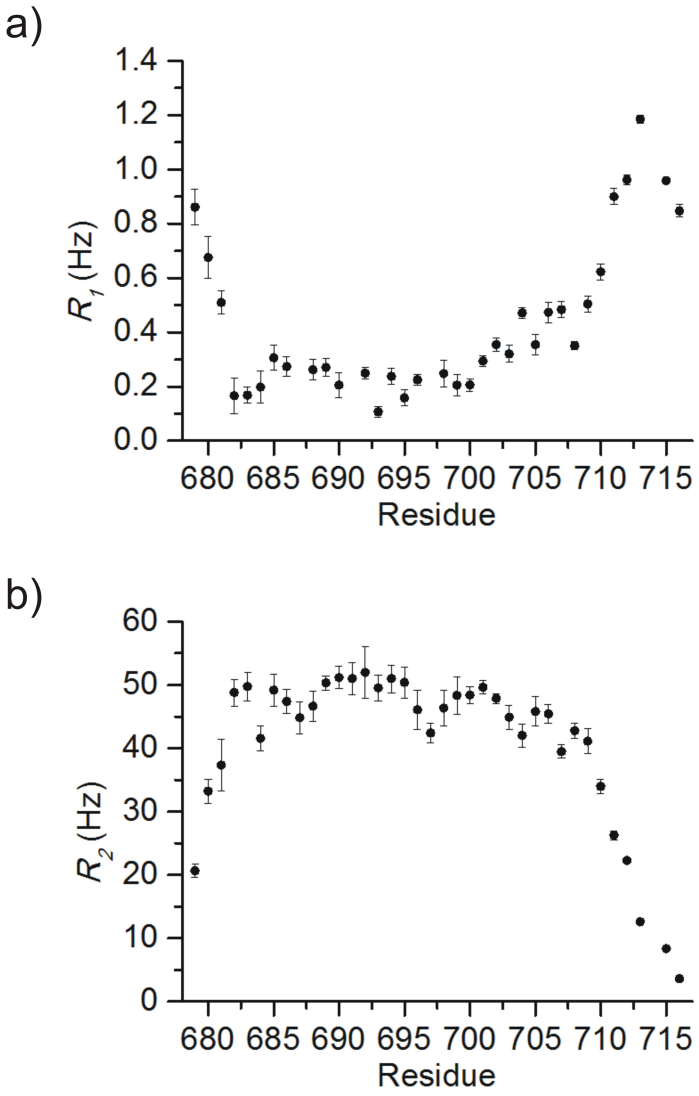


Figure S5. Backbone dynamics of the HIV-1 gp41^{TMD} in bicelles. (**a**) Residue-specific ^{15}N R_1 relaxation rates. (**b**) Residue-specific ^{15}N R_2 relaxation rates. The measurements were carried out at 14.1 T and 303 K on the gp41^{TMD} reconstituted in bicelles with $q = 0.5$.

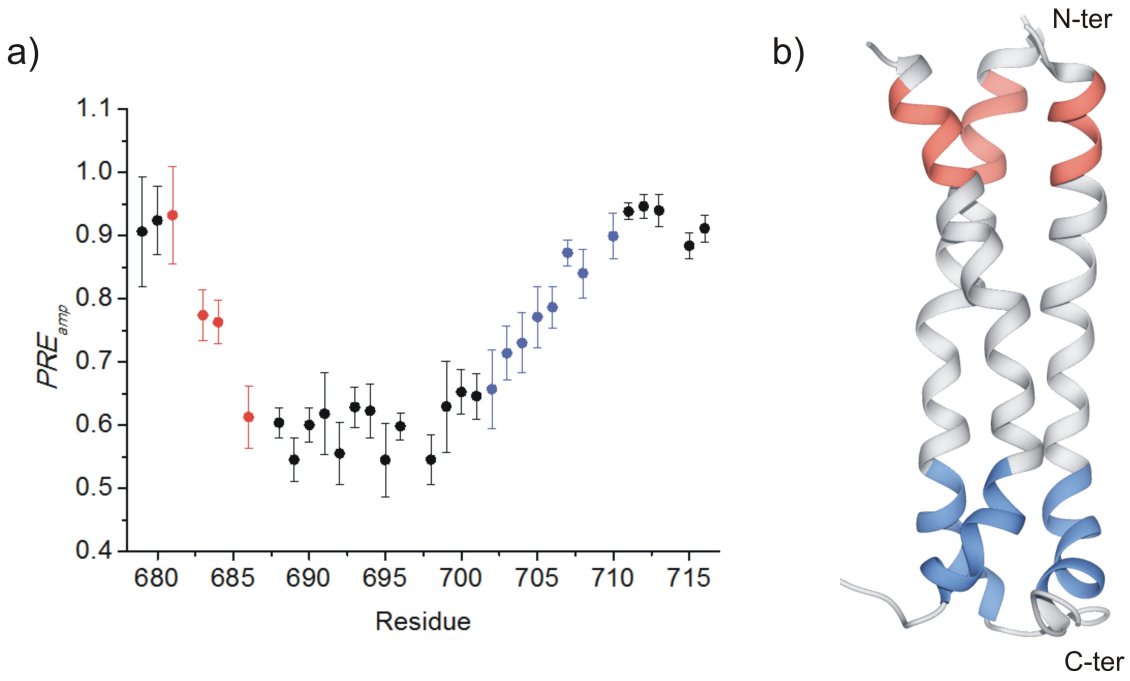


Figure S6. Qualitative agreement between the PRE_{amp} slope and the protein segment tilt angle. (**a**) The PRE_{amp} vs. residue plot in which the PRE-sensitive regions (residues 681-686 and 702-710) are shown in red and blue, respectively. (**b**) The PRE-sensitive regions in (a) were mapped onto the gp41^{TMD} trimer structure, showing the agreement between the PRE_{amp} slope and the helical segment tilt angle (relative to the bilayer normal) for these regions.

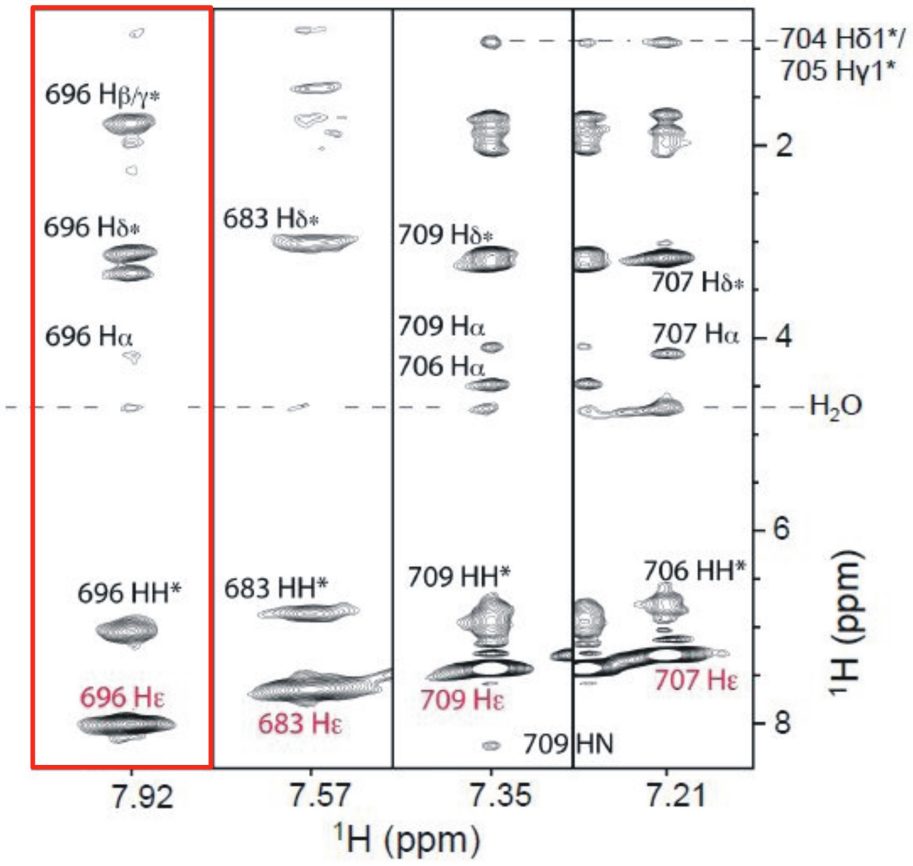


Figure S7. The NOE strips for the four arginine sidechain Hε of the HIV-1 gp41^{TMD}, taken from the 3D ^{15}N-edited NOESY-TROSY-HSQC spectrum. The spectrum was recorded with NOE mixing time of 60 ms at 14.1 T. The NOE strip of the R696 is highlighted in a red box to indicate that the R696 Hε shows a water crosspeak. The figure was taken and adapted from Ref. [1].

References

1. Dev, J.; Park, D.; Fu, Q.; Chen, J.; Ha, H. J.; Ghantous, F.; Herrmann, T.; Chang, W.; Liu, Z.; Frey, G.; Seaman, M. S.; Chen, B.; Chou, J. J., Structural basis for membrane anchoring of HIV-1 envelope spike. *Science* **2016,** *353* (6295), 172-5.

2. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995,** *6*, 277-293.

3. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **2005,** *59* (4), 687-96.

4. Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.; Markley, J. L., BioMagResBank. *Nucleic Acids Res* **2008,** *36* (Database issue), D402-8.

5. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res* **2000,** *28* (1), 235-42.

6. Piai, A.; Fu, Q.; Dev, J.; Chou, J. J., Optimal Bicelle Size q for Solution NMR Studies of the Protein Transmembrane Partition. *Chemistry* **2017,** *23* (6), 1361-1367.

7. Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A., Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **1992,** *31* (23), 5269-78.

8. Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E., Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **1994,** *33* (19), 5984-6003.

9. Peng, J. W.; Wagner, G., Investigation of protein motions via relaxation measurements. *Methods Enzymol* **1994,** *239*, 563-96.

10. Glover, K. J.; Whiles, J. A.; Wu, G.; Yu, N.; Deems, R.; Struppe, J. O.; Stark, R. E.; Komives, E. A.; Vold, R. R., Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophysical journal* **2001,** *81* (4), 2163-71.

11. Sanders, C. R., 2nd; Schwonek, J. P., Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry* **1992,** *31* (37), 8898-905.

12. Wu, H.; Su, K.; Guan, X.; Sublette, M. E.; Stark, R. E., Assessing the size, stability, and utility of isotropically tumbling bicelle systems for structural biology. *Biochim Biophys Acta* **2010,** *1798* (3), 482-8.