JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Cite This: J. Am. Chem. Soc. 2017, 139, 18432−18435
Communication
pubs.acs.org/JACS

Stability and Water Accessibility of the Trimeric Membrane Anchors of the HIV-1 Envelope Spikes

Alessandro Piai,[‡] Jyoti Dev,[‡] Qingshan Fu, and James J. Chou*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, United States

Supporting Information

ABSTRACT: HIV-1 envelope spike (Env) is a type I membrane protein that mediates viral entry. Recent studies showed that its transmembrane domain (TMD) forms a trimer in lipid bilayer whose structure has several peculiar features that remain difficult to explain. One is the presence of an arginine R696 in the middle of the TM helix. Additionally, the N- and C-terminal halves of the TM helix form trimeric cores of opposite nature (hydrophobic and hydrophilic, respectively). Here we determined the membrane partition and solvent accessibility of the TMD in bicelles that mimic a lipid bilayer. Solvent paramagnetic relaxation enhancement analysis showed that the R696 is indeed positioned close to the center of the bilayer, but, surprisingly, can exchange rapidly with water as indicated by hydrogen−deuterium exchange measurements. The solvent accessibility of R696 is likely mediated by the hydrophilic core, which also showed fast water exchange. In contrast, the N-terminal hydrophobic core showed extremely slow solvent exchange, suggesting the trimer formed by this region is extraordinarily stable. Our data explain how R696 is accommodated in the middle of the membrane while reporting the overall stability of the Env TMD trimer in lipid bilayer.

The HIV-1 envelope spike (Env) is a glycoprotein used by the HIV-1 virus to target and infiltrate specific host cells.[1,2] The mature Env spikes [trimeric $(gp120/gp41)_3$] are the sole antigens on the virion surface.[3] Conformational changes in gp120 when triggered by binding to receptor (CD4) and coreceptor (e.g., CCR5 or CXCR4) lead to a cascade of refolding events in gp41, and ultimately to membrane fusion.[2,4−6]

A plethora of structures of the ectodomain (ECD) of the gp120/gp41 complex have been determined,[4−11] but relatively little is known about the transmembrane (TM) and membrane proximal regions of gp41 due to challenges of preserving their native-like folding in membrane mimetic media. Mutagenesis studies showed that the transmembrane domain (TMD) of gp41 is not merely a membrane anchor, but plays critical roles in membrane fusion and viral infectivity.[12−16] Moreover, recent studies showed that truncations in the cytoplasmic tail of gp41 could alter the antigenic surface of the Env ECD on the opposite side of the membrane,[17] suggesting that the ECD, the TMD, and the membrane proximal regions are conformationally coupled, and thus the conformational stability of the TMD is an important consideration for immunogen design in B-cell based HIV-1 vaccine development.

In an earlier study, we determined the NMR structure of the gp41 TMD in bicelles[18] using a gp41 fragment (residues 677−716) from a clade D HIV-1 isolate 92UG024.2, designated $gp41^{HIV1D(677-716)}$. The TMD forms a well-structured trimer, helical almost all the way from the N- to the C-terminal end. The structure shows two unusual features. One is the presence of an arginine (R696), three in the trimer, in the middle of the TM helices (Figure S1a), implying three unbalanced charges in the hydrophobic core of the membrane. R696 is highly conserved with occasional lysine substitution. Another feature is that the N- and C- terminal halves of the trimer are assembled differently. The region N-terminal to R696 (686−695) resembles a coiled-coil motif that forms a hydrophobic core (Figure S1b), whereas the region C-terminal to R696 (697−712) assembles a core that is largely hydrophilic (Figure S1c).

The above features of the gp41 TMD are difficult to comprehend. In particular, the three lipid-facing arginines in the middle of the TMD have no nearby acidic residues to form hydrogen bond with. In this study, we performed extensive analysis of the membrane insertion and solvent accessibility of the gp41 TMD in bicelles using the same protein construct, denoted $gp41^{TMD}$ for simplicity, used for structure determination.[18] Our data provide an explanation for how the membrane-embedded arginines are tolerated in the gp41 TMD and revealed regions of the TMD with very different trimer stability.

Isotropic bicelles are versatile systems for investigating protein immersion in membrane.[19] We first investigated the membrane partition of the gp41 TMD to determine the true position of the R696 in the membrane using a solvent paramagnetic relaxation enhancement (PRE) method developed previously for bicelles.[20] This approach is based on the notion that if the bicelle is sufficiently wide, 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 (Figure S2a). The protein $gp41^{TMD}$ was reconstituted in bicelles with DMPC/DHPC molar ratio (q) of 0.5 (Figure S2b), and titration of the water-soluble paramagnetic probe Gd-DOTA outside the bicelles provided residue-specific PRE amplitudes (PRE_{amp}) (Figure 1a, S2c; Table S2). To determine the position of the TMD trimer relative to the bilayer center, we calculated, for

Received: September 1, 2017
Published: December 1, 2017

DOI: 10.1021/jacs.7b09352
J. Am. Chem. Soc. 2017, 139, 18432−18435

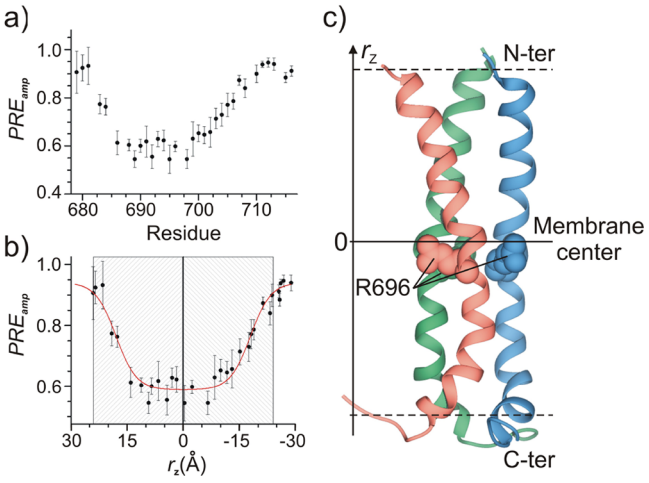


Figure 1. Membrane partition of gp41^{TMD}. (a) Residue-specific PRE_{amp}. (b) PRE_{amp} vs r_z best fitted to the symmetric sigmoidal equation (eq S2), where r_z = 0 corresponds to the bilayer center. The gray-striped box represents the estimated thickness of the bilayer. (c) Position of the gp41^{TMD} structure relative to the center (solid line) and boundaries (dashed lines) of the lipid bilayer.

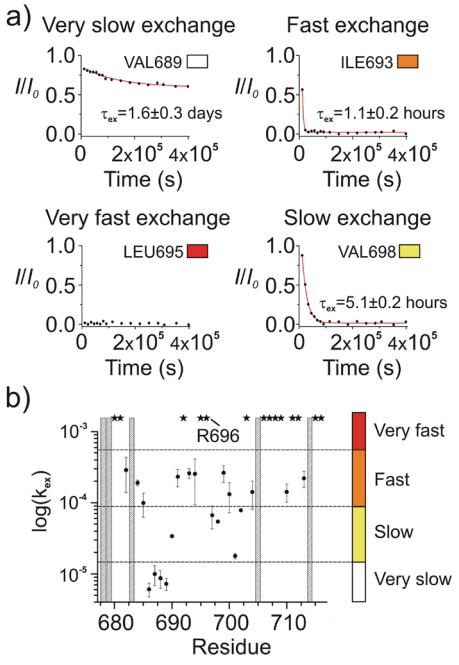


Figure 2. Solvent accessibility of gp41^{TMD}. (a) Signal decay over time for four residues representative of the defined exchange regimes. (b) Residue-specific k_{ex} reported in logarithmic scale. Amide protons with exchange too fast to be measured are marked with stars. Residues that could not be analyzed (overlapping peaks and P714) are marked with gray bars. The color spectrum on the right represents the four different exchange regimes: very fast (red), fast (orange), slow (yellow), and very slow (white: very slow).

each residue *i*, the distance (r_z) along the 3-fold symmetry axis, which is parallel to the bilayer normal, from the amide proton to an arbitrary reference point based on the NMR structure of the 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[20] (Figure S3) to position the TMD structure relative to the center of the bilayer (Figure 1b). Moreover, the sigmoidal fit in Figure 1b shows that the PRE_{amp} reaches the maximal value at about 24 Å away from the bilayer center on either side, indicating that the bilayer thickness around the protein is ∼48 Å. The membrane partition of the gp41 TMD shows that the R696 is indeed deep within the hydrophobic core of the membrane (Figure 1c; Table S3). This result was confirmed by performing a complementary PRE titration in which the lipophilic 16-Doxyl-stearic acid (16-DSA) was used as paramagnetic probe (Figure S4), which also indicated that the lateral solvent PRE is negligible in the experimental condition used.

Having three basic residues in the middle of the membrane suggests that the gp41 TMD trimer might be energetically unstable. Indeed, the N$^{\epsilon}$ chemical shift of R696 is 85 ppm (Table S4), indicating that its side-chain is charged.[21] We thus investigated the trimer stability by performing the hydrogen−deuterium (H-D) exchange experiment. We first prepared a sample of (^{15}N, ^{2}H)-labeled gp41^{TMD} reconstituted in bicelles (q = 0.5) at pH 6.0, and a reference 2D ^{1}H-^{15}N TROSY-HSQC spectrum was recorded (Figure S5). The sample was then flash frozen and lyophilized. To initiate H-D exchange, the dried sample was dissolved in 99.9% D_2O (pD ∼ 6.4). The exchange was monitored as the loss of NMR signal in a series of 2D ^{1}H-^{15}N TROSY-HSQC spectra recorded at different time points (Figure 2a) as exchangeable amide protons were replaced by deuterium.

Our measurements show that H-D exchange among the TMD residues is extremely heterogeneous, e.g., the exchange time constant ($\tau_{ex} = 1/k_{ex}$) of V689 is 1.6 ± 0.3 days whereas that of I693 is 1.1 ± 0.2 h (Figure 2a; Table S5). We note that the k_{ex}, although primarily reporting local water accessibility, could also be affected by the presence/absence of labile protons in the residue side-chain. For this reason, the k_{ex} rates were treated semiquantitatively and grouped into four time regimes of solvent exchange: very fast (τ_{ex} < 1 h), fast (1 ≤ τ_{ex} < 3 h), slow (3 h ≤ τ_{ex} < 1 day), and very slow (τ_{ex} ≥ 1 day). The plot of exchange rate (k_{ex}) vs (residue number) (Figure 2b) shows that on average the N-terminal hydrophobic core has much slower H-D exchange than the C-terminal hydrophilic core (Figure 3a). To our surprise, the R696 amide exchanged rapidly

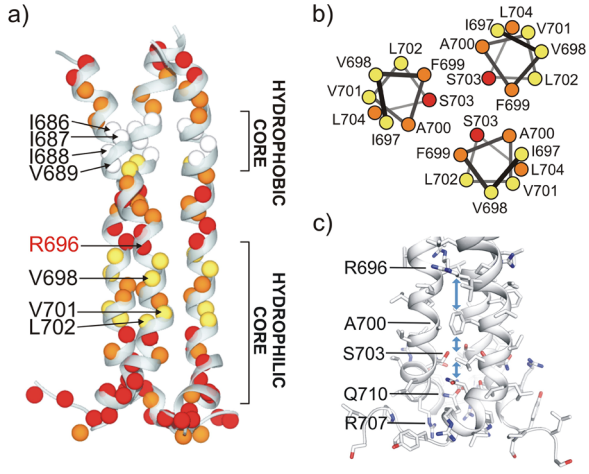


Figure 3. Water accessibility of gp41^{TMD}. (a) Four exchange regimes mapped onto the gp41^{TMD} structure with colors defined as in Figure 2b. (b) Helical wheel representation of residues 697−704 colored according to their exchange regime. (c) Residues lining the hydrophilic core that could mediate water passage to the R696.

DOI: 10.1021/jacs.7b09352
J. Am. Chem. Soc. 2017, 139, 18432−18435

despite being in the most buried region of the bicelles (Figure 1c). The exchange of this residue was too fast to be measured, i.e., the peak vanished by the time of sample preparation and NMR experiment setup and acquisition (~1 h). Because it is unlikely for water to penetrate a lipid bilayer, the fast access to D_2O by the R696 ought to be mediated by the TMD structure. Incidentally, the core-facing amides of the C-terminal hydrophilic core (e.g., F699, A700 and S703) also showed fast exchange whereas the corresponding lipid-facing amides (e.g., V698, V701 and L702) showed slow exchange (Figure 3b). These measurements strongly suggest that D_2O can readily diffuse through the hydrophilic core spanning from residues R696 to R709, and that this hydrophilic core serves as a channel that mediates the fast water access by the R696. In stark contrast, just a few residues away, residues 686−689 in the N-terminal hydrophobic core exchanged very slowly (Figure 2b, 3a), e.g., the peaks were still present even after 4.6 days (Figure S5). The result implies that water diffusion along the trimer core stops at around R696, as it cannot pass through the hydrophobic core formed by residues 686−689. The H-D exchange measurements clearly show that the TMD of HIV-1 gp41 is very different from most of the known oligomeric TMDs in that multiple regions with very different trimer stability and water accessibility exist within the same TMD.

The large heterogeneity in H-D exchange could be accompanied by variations in the dynamic properties. Therefore, we studied the protein backbone dynamics by measuring ^{15}N R_1 and R_2 relaxation rates (Figure S6). As expected, both hydrophobic and hydrophilic cores are well structured (low R_1 and high R_2), with the protein becoming progressively more dynamic on both ends. Though R_1 and R_2 are excellent probes for ns−ps time scale motions, they can also provide qualitative information on slower time scale motions (ms−μs) through their product (R_1R_2). It has been shown that the R_1R_2 product tends to be reduced by ns−ps motions but increased by ms−μs motions associated with chemical shift exchange (R_{ex}).[22] For the well-structured region of the TMD (residues 682−710), the R_1R_2 values of residues 682−700 are quite consistent at ~10 (Figure 4a), indicating the lack of R_{ex} in this region of the trimer. After A700, however, there is a sharp increase in R_1R_2 for residues 702−710 (Figure 4a). We note that this sudden change of R_1R_2 aligns with the kink at L702 that resulted in ~30° tilt of the C-terminal segment (residues 702−710) with respect to the core region of the TM helix (residues 685−700) (Figure 4b). The larger R_1R_2 of residues 702−710 suggest that the C-terminal helical segments undergo substantial conformational exchange.

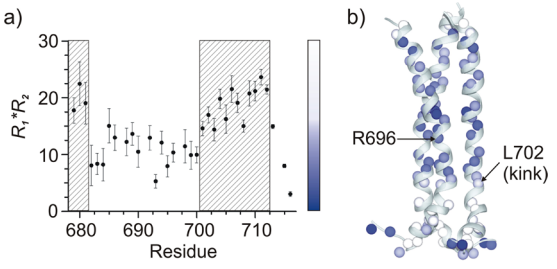


Figure 4. Presence of ms−μs dynamics in $gp41^{TMD}$. (a) Residue-specific R_1R_2. The regions in the striped box are characterized by greater R_{ex} than the rest of the TMD. (**b**) Mapping R_1R_2 onto the $gp41^{TMD}$ structure according to the color spectrum in panel a.

We have shown that the gp41 TMD comprises multiple regions with very different assembly stability and water accessibility. The diverse properties of the TMD afford an explanation for how the conserved basic residue at position 696 (R/K) can be accommodated in the middle of a lipid bilayer while preserving the integrity of the trimer complex.

Due to the unusual position of the R696, the membrane partition of the gp41 TMD has been controversial.[16,18,23] Our solvent PRE analysis now confirms that R696 resides in the middle of the membrane (Figure 1c). The solvent PRE, however, provided no information about the water accessibility of the protein, which we probed using the H-D exchange experiment. We found, unexpectedly, that the R696, although in the middle of the bilayer, showed fast H-D exchange, indicating that water from the bulk solvent can rapidly access the R696. This is consistent with a previous observation that the R696 side-chain H^{ε} showed a water cross-peak in a NOESY spectrum[18] (Figure S7), though it was not understood how water could possibly reach the middle of the lipid bilayer. The H-D exchange data for the C-terminal hydrophilic core (residues 697−709) provided the missing piece of the puzzle. That is, the pore-lining amides of this region all showed reasonably fast exchange (τ_{ex} < 3 h), suggesting that water can diffuse through the hydrophilic core to reach the R696. In contrast, the N-terminal hydrophobic core showed very slow H-D exchange (τ_{ex} > 1 day), which prevents the TMD from becoming a water channel. The R696 is thus strategically located at the interface between the two regions with completely different water accessibilities. To date, the function of the R696 remains unknown, but different studies have suggested that it is important for efficient membrane fusion.[13,15] Our data revealed properties of the gp41 TMD that enables the membrane-embedded basic residue to access water and provided a rationale for the relatively high content of polar residues in the C-terminal core of the TMD.

The very slow H-D exchange of the N-terminal hydrophobic core (residues 686−689) (τ_{ex} > 1 day) suggests that this trimeric region of the gp41 TMD essentially does not dissociate. The extreme stability could ensure that the gp41 remains trimeric in specific stages of the fusion process;[24] it could also play a role in stabilizing the prefusion state of the trimeric HIV-1 Env.

In conclusion, our NMR measurements provided direct evidence that the R696 of HIV-1 gp41 TMD can access water from the bulk solvent via the C-terminal hydrophilic core, thus allowing the arginine to hydrate despite being in the middle of a lipid bilayer. The coexistence of the N-terminal hydrophobic and C-terminal hydrophilic cores is consistent with water permeation through only the C-terminal half of the TMD. This peculiar feature of the gp41 TMD should also apply to HIV-2 and SIV based on structural homology.[23] Although the function of the highly conserved R696 remains elusive, our results show that the bipolar nature of the gp41 TMD allows the natural placement of the arginine in an otherwise unfavorable lipid environment.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b09352.

DOI: 10.1021/jacs.7b09352
J. Am. Chem. Soc. 2017, 139, 18432−18435

Description of the NMR experiment setup and the protocols for the sample preparation and data analysis (PDF)

AUTHOR INFORMATION

Corresponding Author

*james_chou@hms.harvard.edu

ORCID

James J. Chou: 0000-0002-4442-0344

Author Contributions

‡These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Prof. Bing Chen and Prof. Stephen C. Harrison for insightful discussion. This study was supported by NIH grant AI127193 to J.J.C.

REFERENCES

(1) Harrison, S. C. *Adv. Virus Res.* **2005**, *64*, 231−259.
(2) Chan, D. C.; Kim, P. S. *Cell* **1998**, *93*, 681−4.
(3) Wei, X.; Decker, J. M.; Wang, S.; Hui, H.; Kappes, J. C.; Wu, X.; Salazar-Gonzalez, J. F.; Salazar, M. G.; Kilby, J. M.; Saag, M. S.; Komarova, N. L.; Nowak, M. A.; Hahn, B. H.; Kwong, P. D.; Shaw, G. M. *Nature* **2003**, *422*, 307−12.
(4) Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. *Cell* **1997**, *89*, 263−273.
(5) Weissenhorn, W.; Dessen, A.; Harrison, S. C.; Skehel, J. J.; Wiley, D. C. *Nature* **1997**, *387*, 426−430.
(6) Pancera, M.; Zhou, T.; Druz, A.; Georgiev, I. S.; Soto, C.; Gorman, J.; Huang, J.; Acharya, P.; Chuang, G. Y.; Ofek, G.; Stewart-Jones, G. B.; Stuckey, J.; Bailer, R. T.; Joyce, M. G.; Louder, M. K.; Tumba, N.; Yang, Y.; Zhang, B.; Cohen, M. S.; Haynes, B. F.; Mascola, J. R.; Morris, L.; Munro, J. B.; Blanchard, S. C.; Mothes, W.; Connors, M.; Kwong, P. D. *Nature* **2014**, *514*, 455−61.
(7) Chen, B.; Vogan, E. M.; Gong, H.; Skehel, J. J.; Wiley, D. C.; Harrison, S. C. *Nature* **2005**, *433*, 834−41.
(8) Kwong, P. D.; Wyatt, R.; Robinson, J.; Sweet, R. W.; Sodroski, J.; Hendrickson, W. A. *Nature* **1998**, *393*, 648−659.
(9) Huang, C. C.; Tang, M.; Zhang, M. Y.; Majeed, S.; Montabana, E.; Stanfield, R. L.; Dimitrov, D. S.; Korber, B.; Sodroski, J.; Wilson, I. A.; Wyatt, R.; Kwong, P. D. *Science* **2005**, *310*, 1025−1028.
(10) Caffrey, M.; Cai, M.; Kaufman, J.; Stahl, S. J.; Wingfield, P. T.; Covell, D. G.; Gronenborn, A. M.; Clore, G. M. *EMBO J.* **1998**, *17*, 4572−4584.
(11) Lyumkis, D.; Julien, J. P.; de Val, N.; Cupo, A.; Potter, C. S.; Klasse, P. J.; Burton, D. R.; Sanders, R. W.; Moore, J. P.; Carragher, B.; Wilson, I. A.; Ward, A. B. *Science* **2013**, *342*, 1484−90.
(12) Helseth, E.; Olshevsky, U.; Gabuzda, D.; Ardman, B.; Haseltine, W.; Sodroski, J. *J. Virol.* **1990**, *64*, 6314−6318.
(13) Owens, R. J.; Burke, C.; Rose, J. K. *J. Virol.* **1994**, *68*, 570−574.
(14) Shang, L.; Yue, L.; Hunter, E. *J. Virol* **2008**, *82*, 5417−28.
(15) Long, Y.; Meng, F.; Kondo, N.; Iwamoto, A.; Matsuda, Z. *Protein Cell* **2011**, *2*, 369−76.
(16) Rotem, E.; Reuven, E. M.; Klug, Y. A.; Shai, Y. *Biochemistry* **2016**, *55*, 1049−57.
(17) Chen, J.; Kovacs, J. M.; Peng, H.; Rits-Volloch, S.; Lu, J.; Park, D.; Zablowsky, E.; Seaman, M. S.; Chen, B. *Science* **2015**, *349*, 191−5.
(18) 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. *Science* **2016**, *353*, 172−5.
(19) Schmidt, T.; Situ, A. J.; Ulmer, T. S. *J. Phys. Chem. Lett.* **2016**, *7*, 4420−4426.
(20) Piai, A.; Fu, Q.; Dev, J.; Chou, J. J. *Chem. - Eur. J.* **2017**, *23*, 1361−1367.
(21) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno, B. E.; McIntosh, L. P. *Protein Sci.* **2015**, *24*, 752−761.
(22) Kneller, J. M.; Lu, M.; Bracken, C. *J. Am. Chem. Soc.* **2002**, *124*, 1852−3.
(23) Chen, B.; Chou, J. J. *FEBS J.* **2017**, *284*, 1171−1177.
(24) Roche, J.; Louis, J. M.; Grishaev, A.; Ying, J.; Bax, A. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 3425−30.

DOI: 10.1021/jacs.7b09352
J. Am. Chem. Soc. 2017, 139, 18432−18435

Supporting Information

Stability and Water Accessibility of the Trimeric Membrane Anchors of the HIV-1 Envelope Spikes

Alessandro Piai, Jyoti Dev, Qingshan Fu, and James J. Chou*

Department of Biological Chemistry and Molecular Pharmacology
Harvard Medical School
Boston, Massachusetts 02115, United States

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} (obtained in the identical sample condition of that used in the present work, e.g. q = 0.5 DMPC/DHPC bicelles) 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:

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

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 $r_Z = 0$).

Lipophilic PRE analysis

Complementary to the solvent PRE data, a set of lipophilic PRE data was also acquired. Using the identical approach to that employed for the solvent PRE, the bicelle-reconstituted gp41^{TMD} was titrated

with the membrane-embedded paramagnetic agent 16-Doxyl-stearic acid (16-DSA) at various known concentrations. The titrant stock solution (24 mM 16-DSA) was prepared in the same buffer as that of the protein sample, i.e., using q = 0.5 DMPC/DHPC bicelles solution to solubilize the paramagnetic agent to prevent changes in the sample bicelle q during titration. The 16-DSA 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 spectrum at each of the following 16-DSA concentrations: 0 (reference), 0.6, 1.2, 1.8, 2.4, 3.0, 3.6 and 4.2 mM. The residue-specific PRE_{amp} was determined by fitting the peak intensity decay as a function of [16-DSA] to the exponential decay equation Eq. S1.

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/τ_{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.

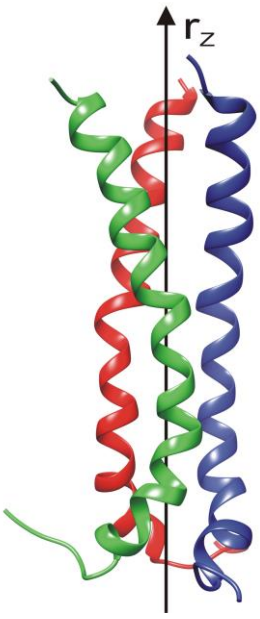
Type of experiment	Spectral widths and chemical shift evolution times		# of scans	Inter-scan delay (s)	Duration of the experiment
Gd-DOTA titration					
2D ^{1}H-^{15}N TROSY-HSQC	1300 Hz (^{15}N) 77.0 ms	9600 Hz ($^{1}H^{N}$) 106.5 ms	32	3.5	6 hours 40 min
16-DSA titration					
2D ^{1}H-^{15}N TROSY-HSQC	1600 Hz (^{15}N) 62.5 ms	8500 Hz ($^{1}H^{N}$) 61.0 ms	24	3.5	5 hours
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)
25.4±0.5	TRP 678
24.1±0.5	LEU 679
23.5±0.5	TRP 680
21.5±0.5	TYR 681
19.7±0.5	ILE 682
19.1±0.5	ARG 683
17.7±0.5	ILE 684
15.5±0.5	PHE 685
14.0±0.5	ILE 686
13.0±0.5	ILE 687
11.2±0.5	ILE 688
9.3±0.5	VAL 689
8.5±0.5	GLY 690
6.7±0.5	SER 691
4.3±0.5	LEU 692
3.0±0.5	ILE 693
1.8±0.5	GLY 694
0.0	*Membrane Center*
-0.4±0.5	LEU 695
-2.4±0.5	ARG 696
-4.3±0.5	ILE 697
-6.6±0.5	VAL 698
-8.4±0.5	PHE 699
-10.0±0.5	ALA 700
-11.7±0.5	VAL 701
-13.1±0.5	LEU 702
-15.2±0.5	SER 703
-17.6±0.5	LEU 704
-18.2±0.5	VAL 705
-18.9±0.5	ASN 706
-21.3±0.5	ARG 707
-23.2±0.5	VAL 708
-22.8±0.5	ARG 709
-23.8±0.5	GLN 710
-26.1±0.5	GLY 711
-26.8±0.5	TYR 712
-28.9±0.5	SER 713
-25.8±0.5	LEU 715
-25.6±0.5	SER 716

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

Table S4. The arginine N^{ε} chemical shifts of the HIV-1 $gp41^{TMD}$.

Residue	δN^{ε} (ppm)
ARG 683	84.45±0.05
ARG 696	85.03±0.05
ARG 707	84.62±0.05
ARG 709	84.05±0.05

The average N^{ε} chemical shifts of the protonated (charged) and unprotonated (uncharged) arginine are 84.8 and 90.7 ppm, respectively[10].

Table S5. 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).

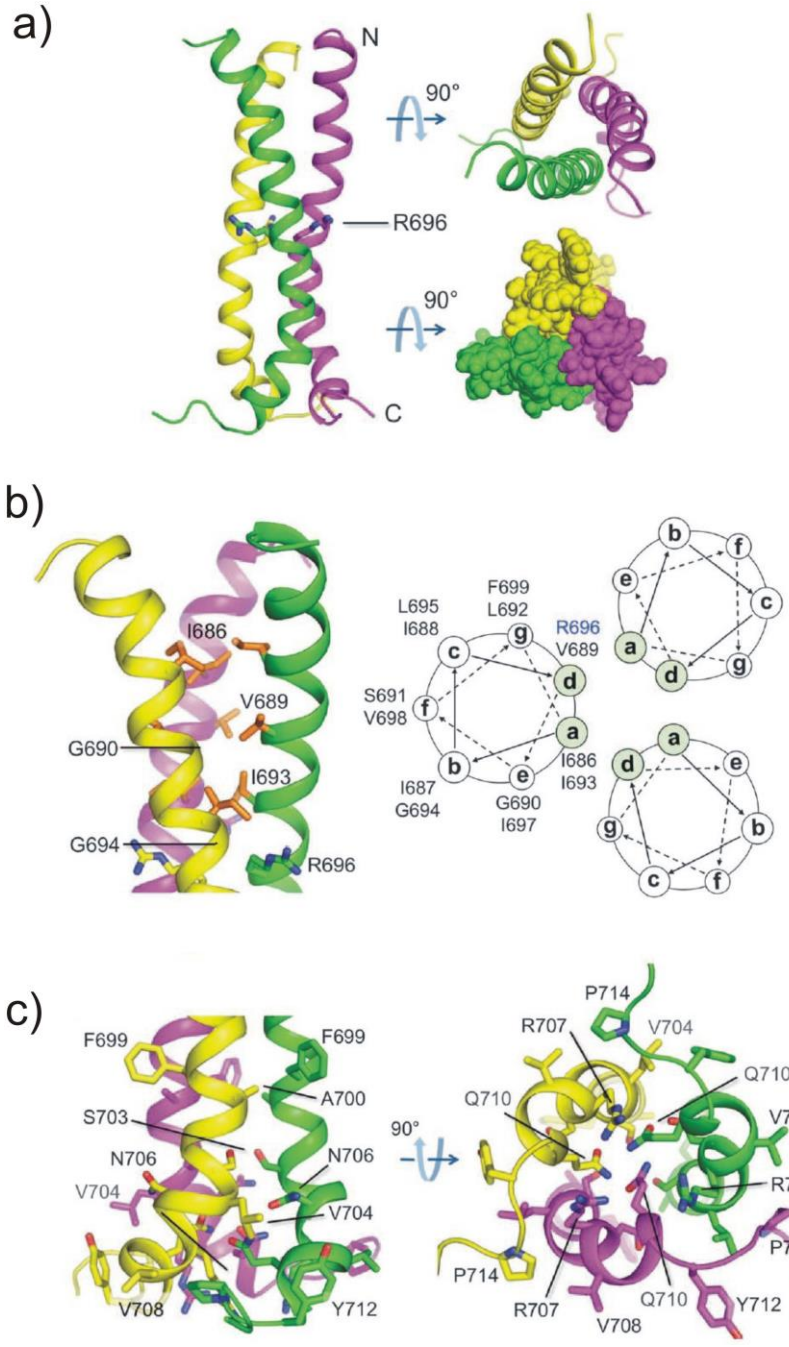


Figure S1. The NMR structure of the HIV-1 gp41^{TMD} in bicelles. (**a**) Ribbon representation of the trimer showing the side-chains of the R696. The sphere representation of the top view (lower right) shows that the trimer has no ion-permeable holes. (**b**) The hydrophobic core of the N-terminal half 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].

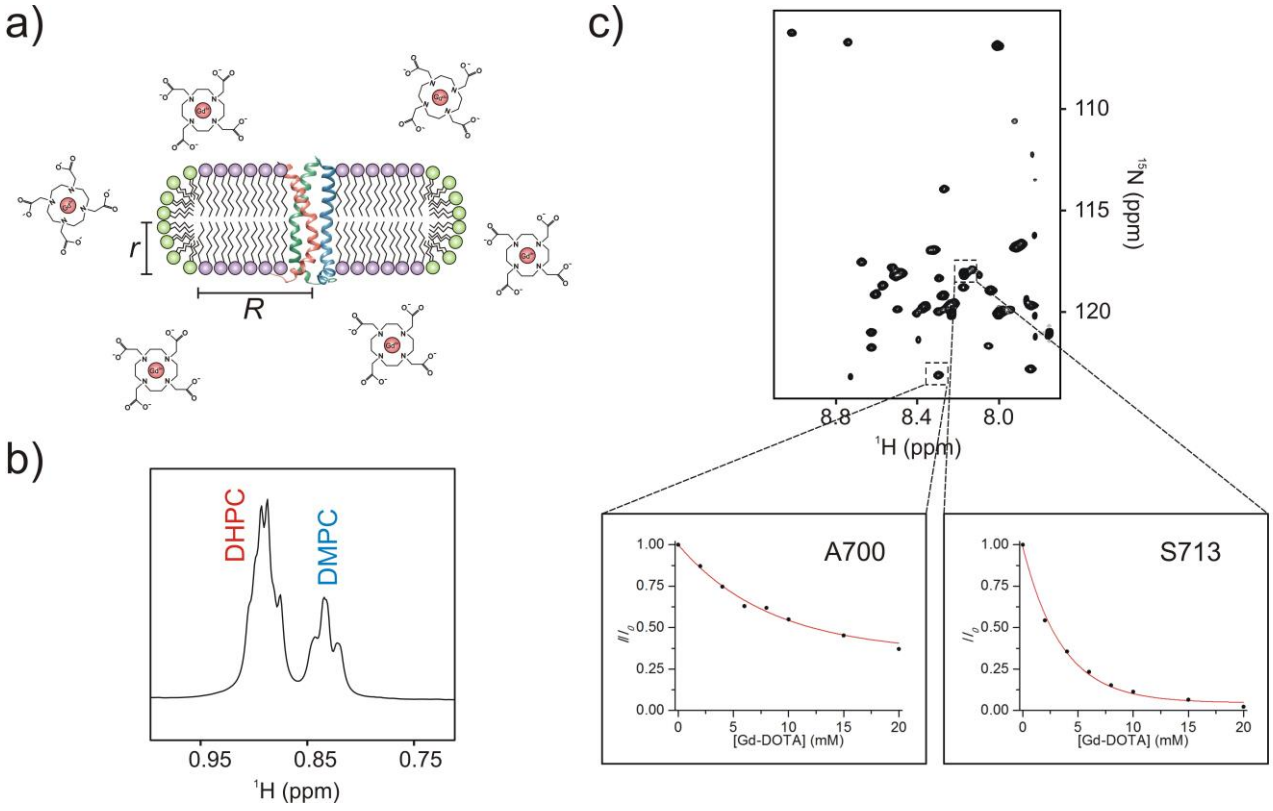


Figure S2. Solvent PRE analysis for determining the membrane partition of the HIV-1 gp41^{TMD} in bicelles. (**a**) Schematic illustration of a bicelle-reconstituted gp41^{TMD} surrounded by Gd-DOTA molecules. The radius of the planar region of the bicelle (R) is given by the equation $R = 1/2\, krq\left[\pi + (\pi^2 + 8k/q)^{1/2}\right]$, 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[11-13]. (**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. S1) yielded PRE_{amp} for the two residues.

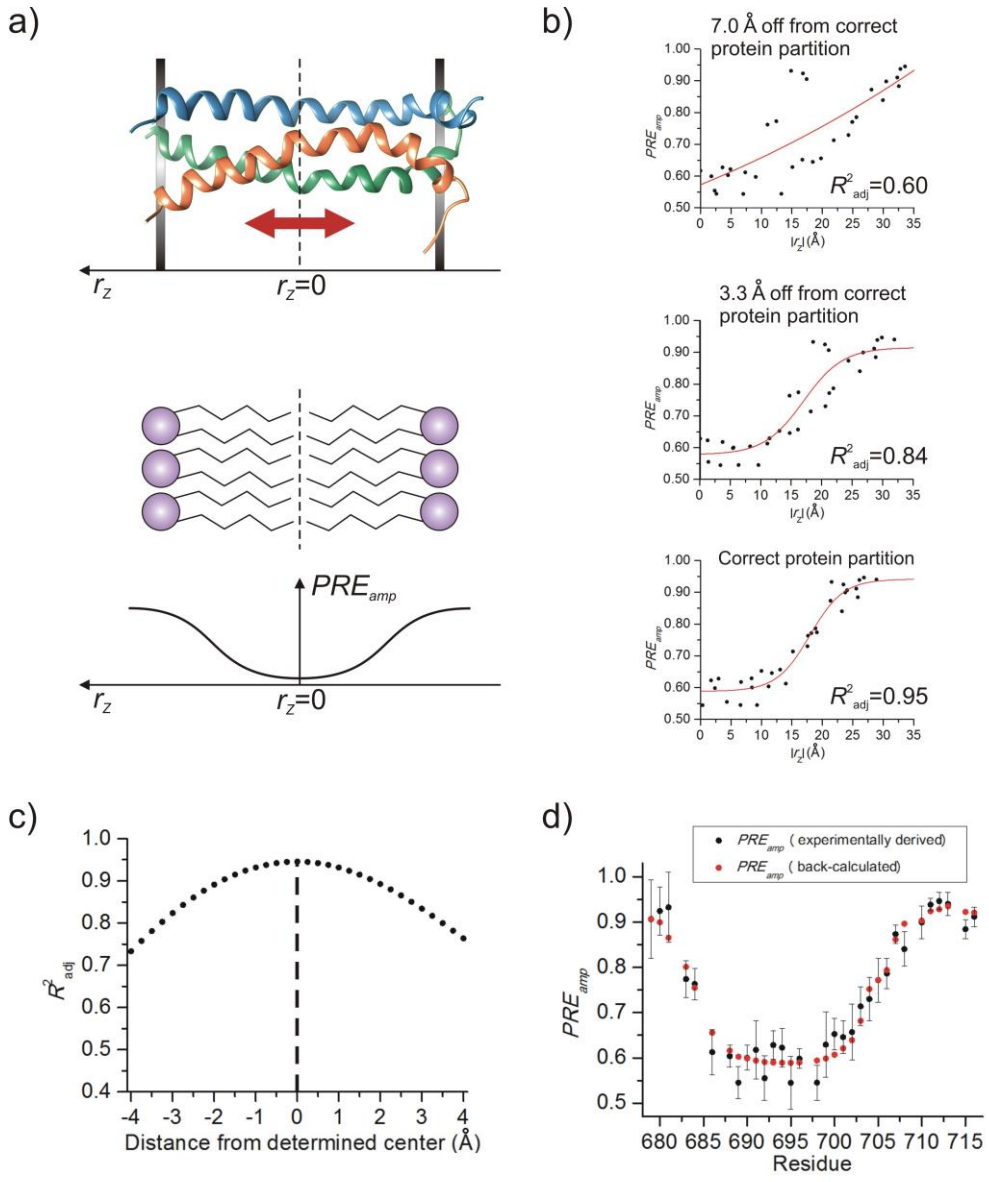


Figure S3. Position of the HIV-1 gp41^{TMD} relative to the bilayer center determined by data fitting. (**a**) Illustration showing the sliding of the gp41^{TMD} structure along the bilayer normal (r_Z axis) to yield the best fit to the symmetric sigmoidal function (Eq. S2). The r_Z = 0 of the sigmoidal 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}.

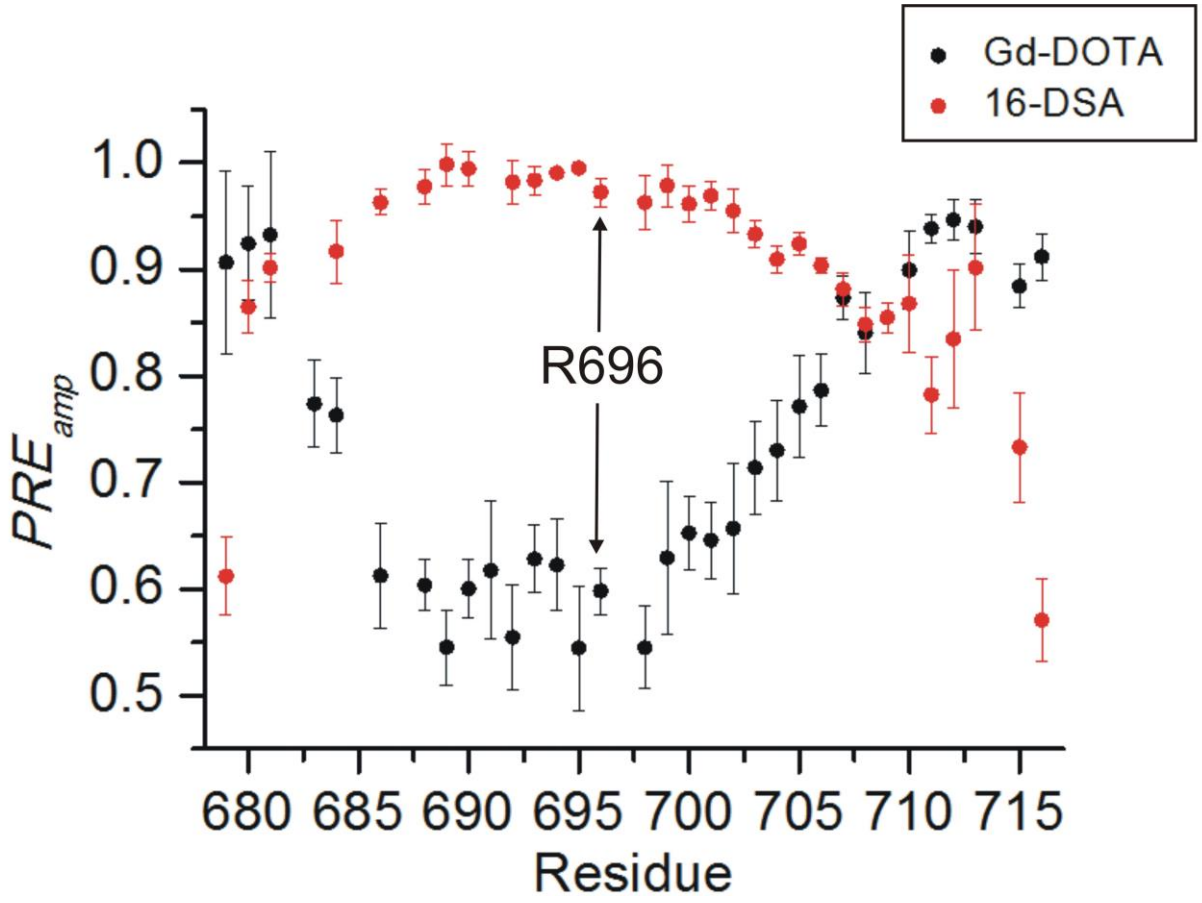


Figure S4. Residue-specific PRE_{amp} determined from the paramagnetic titrations. PRE_{amp} profiles derived from 16-DSA titration (red) and Gd-DOTA titration (black) are compared. The reciprocal profiles of the two data sets are consistent with the expected confinements of the two paramagnetic probes, i.e., the center of the bicelle (16-DSA) and the solvent (Gd-DOTA). Both titrations indicate that R696, highlighted in the figure, is partitioned in the core of the lipid bilayer.

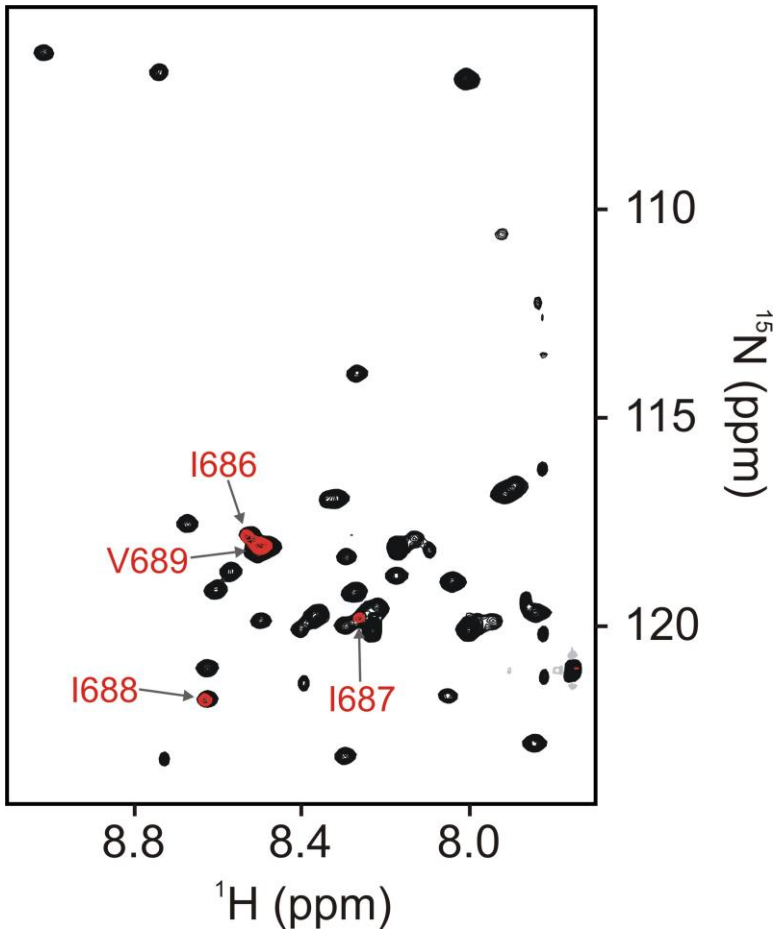


Figure S5. Example NMR spectrum showing the H-D exchange of the HIV-1 gp41^{TMD}. The 2D ^{1}H-^{15}N TROSY-HSQC spectra recorded before and after 4.6 days of H-D exchange are shown in black and red, respectively. The residues forming the hydrophobic core of the protein are marked in the overlaid spectra.

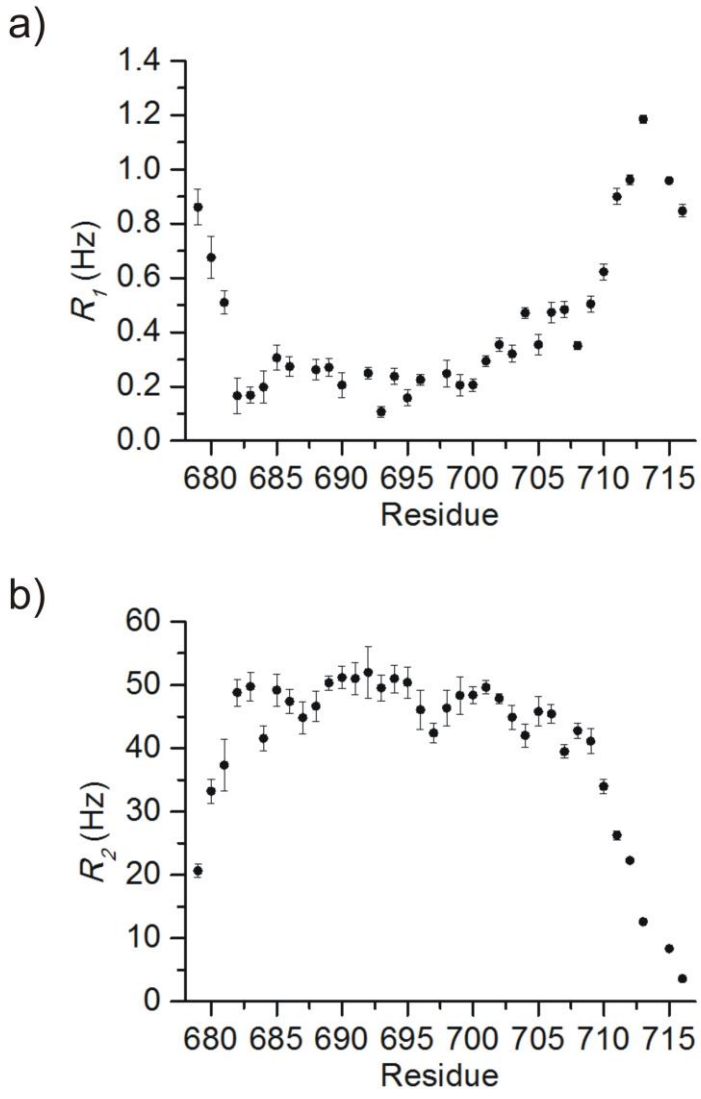


Figure S6. Backbone dynamics of the HIV-1 gp41^{TMD} in bicelles. (**a**) Residue-specific ^{15}N R_1 relaxation rates. (**b**) Residue-specific ^{15}N R_2 relaxation rates. The measurements were carried out at 14.1 T and 303 K on the gp41^{TMD} reconstituted in bicelles with $q = 0.5$.

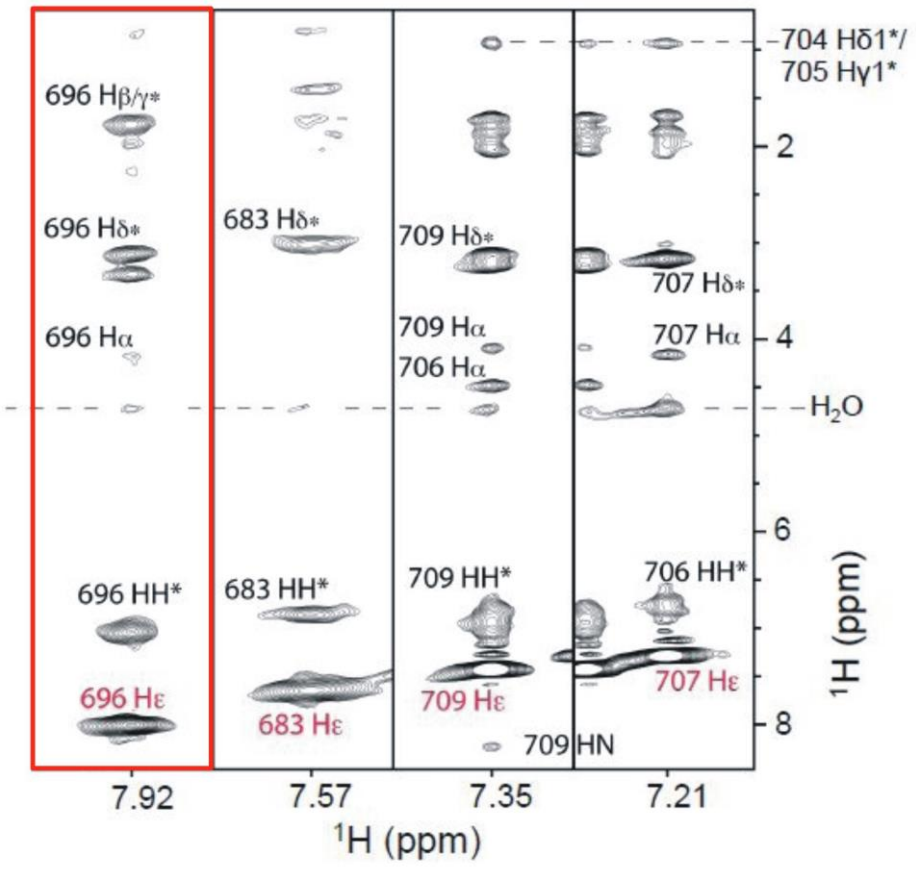


Figure S7. The NOE strips for the four arginine sidechain H^{ε} of the HIV-1 gp41^{TMD}, taken from the 3D ^{15}N-edited NOESY-TROSY-HSQC spectrum. The spectrum was recorded with NOE mixing time of 60 ms at 14.1 T. The NOE strip of the R696 is highlighted in a red box to indicate that the R696 H^{ε} shows a water cross-peak. 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., *Science* **2016,** *353*, 172-5.

(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, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., *Proteins* **2005,** *59*, 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., *Nucleic Acids Res* **2008,** *36*, D402-8.

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

(6) Piai, A.; Fu, Q.; Dev, J.; Chou, J. J., *Chem - Eur J* **2017,** *23*, 1361-1367.

(7) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A., *Biochemistry* **1992,** *31*, 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., *Biochemistry* **1994,** *33*, 5984-6003.

(9) Peng, J. W.; Wagner, G., *Methods Enzymol* **1994,** *239*, 563-96.

(10) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno, B. E.; McIntosh, L. P., *Protein Sci* **2015,** *24*, 752-61.

(11) Glover, K. J.; Whiles, J. A.; Wu, G.; Yu, N.; Deems, R.; Struppe, J. O.; Stark, R. E.; Komives, E. A.; Vold, R. R., *Biophys J* **2001,** *81*, 2163-71.

(12) Sanders, C. R., 2nd; Schwonek, J. P., *Biochemistry* **1992,** *31*, 8898-905.

(13) Wu, H.; Su, K.; Guan, X.; Sublette, M. E.; Stark, R. E., *Biochim Biophys Acta* **2010,** *1798*, 482-8.