nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-16165-0 OPEN

Structural basis of transmembrane coupling of the HIV-1 envelope glycoprotein

Alessandro Piai[1,5], Qingshan Fu[1,5], Yongfei Cai[2,3,5], Fadi Ghantous[4], Tianshu Xiao[2,3], Md Munan Shaik[2,3], Hanqin Peng[2], Sophia Rits-Volloch[2], Wen Chen[1], Michael S. Seaman[4], Bing Chen[2,3]✉ & James J. Chou[1]✉

The prefusion conformation of HIV-1 envelope protein (Env) is recognized by most broadly neutralizing antibodies (bnAbs). Studies showed that alterations of its membrane-related components, including the transmembrane domain (TMD) and cytoplasmic tail (CT), can reshape the antigenic structure of the Env ectodomain. Using nuclear magnetic resonance (NMR) spectroscopy, we determine the structure of an Env segment encompassing the TMD and a large portion of the CT in bicelles. The structure reveals that the CT folds into amphipathic helices that wrap around the C-terminal end of the TMD, thereby forming a support baseplate for the rest of Env. NMR dynamics measurements provide evidences of dynamic coupling across the TMD between the ectodomain and CT. Pseudovirus-based neutralization assays suggest that CT-TMD interaction preferentially affects antigenic structure near the apex of the Env trimer. These results explain why the CT can modulate the Env antigenic properties and may facilitate HIV-1 Env-based vaccine design.

[1] Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Boston, MA 02115, USA. [2] Division of Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA. [3] Department of Pediatrics, Harvard Medical School, 3 Blackfan Street, Boston, MA 02115, USA. [4] Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, USA. [5] These authors contributed equally: Alessandro Piai, Qingshan Fu, Yongfei Cai. ✉email: bchen@crystal.harvard.edu; james_chou@hms.harvard.edu

HIV-1 envelope glycoprotein [Env; trimeric $(gp160)_3$, cleaved to $(gp120/gp41)_3$] catalyzes fusion of viral and target cell membranes leading to viral entry[1,2]. Binding of gp120 to receptor (CD4) and co-receptor (e.g., CCR5 or CXCR4) triggers a cascade of refolding events in gp41 that promote membrane fusion[1,3,4]. The prefusion conformation of Env trimer is the state recognized by most broadly neutralizing antibodies (bnAbs)[5–7], and thus considered a major vaccine target. Several studies demonstrated that alterations of its membrane-related components, including the transmembrane domain (TMD) and cytoplasmic tail (CT), can reshape the antigenic structure of the Env ectodomain exposed outside of viral membrane[5,8,9], suggesting that there are intricate interconnections among them.

The intact HIV-1 Env has been visualized on the surface of virion at modest resolutions by cryo-electron tomography (cryo-ET)[10], revealing the trimeric organization of gp120 and a part of gp41, but leaving the TMD and CT regions completely unresolved. Its ectodomain density is in overall agreement with subsequent high-resolution structures, determined by X-ray crystallography and cryo-electron microscopy (cryo-EM), of a soluble form of the Env trimer stabilized by a disulfide crosslink between gp120 and gp41[11–13]. Other cryo-EM structures of detergent-solubilized Env constructs with or without the CT has been reported[14–16], but the MPER, TMD, and CT regions are all disordered in detergent micelles that probably failed to mimic a real membrane. We have recently determined the NMR structures of the trimeric MPER and TMD reconstituted in bicelles that mimic a lipid bilayer[8,9]. In particular, the MPER matches well the low-resolution cryo-ET density near the membrane of the unliganded viral Env, suggesting the MPER conformation observed by NMR is consistent with the structure of prefusion Env on the virion[9].

HIV-1 Env and related lentiviral fusion proteins have an unusually long CT (~150 residues), which has been implicated in Env cellular trafficking, as well as incorporation into virions[17,18]. The HIV-1 Env CT can be divided into distinct regions based on their biophysical properties (Supplementary Fig. 1a): a loop, commonly known as the Kennedy sequence (KS), followed by three segments predicted to form amphipathic α-helices, named lentivirus lytic peptide 2 (LLP2), LLP3, and LLP1[17]. Earlier studies suggested that the CT forms three membrane-bound amphipathic helices in an extended conformation[19,20]. These structures are very informative about the secondary structures of the CT, but they fall short of explaining how truncation in the CT can influence the antigenic structure of the ectodomain on the opposite side of the membrane[5,21,22]. In this study, we used NMR to obtain high-resolution information of the HIV-1 Env CT folding in the context of the TMD and lipid bilayer. We find that the CT adopts a structure different from the previous model[19] that can explain the physical coupling between the CT and the ectodomain.

Results

Identification of a suitable TMD–CT fragment for structural investigation. To prepare an NMR sample suitable for structural analysis, we designed a protein construct derived from a clade D HIV-1 isolate 92UG024.2 (residues 677–788) that encompasses the TMD and a portion of the CT containing the KS and the LLP2. We first mutated the palmitoylation site C764 to serine (C764S) to avoid non-physiological disulfide formation. We then found that the KS (residues 710–738) was completely unstructured according to the NMR data and that removal of its central region (residues 726–736) did not affect the protein structure (Supplementary Fig. 2). This region was thus deleted to reduce NMR signal overlap. The final NMR construct, designated TMD–CT^{LLP2}, included residues 677–725 and 737–788. TMD–CT^{LLP2} is trimeric in DMPC–DHPC bicelles with $q = 0.5$ (Supplementary Fig. 1b). Comparison of the NMR spectrum of the bicelle-reconstituted TMD–CT^{LLP2} with those of the MPER–TMD (residues 660–710)[9] and the TMD (residues 677–716)[8] used in previous studies showed that the peaks corresponding to the TM core (residues 685–700) are almost superimposable (Supplementary Figs. 1c and 3a), indicating that its structure is essentially identical in all three overlapping constructs. Therefore, the CT did not disrupt the TMD structure.

NMR structure of the TMD–CT^{LLP2} trimer in bicelles. The structure of the trimeric TMD–CT^{LLP2} in bicelles was determined using mainly two types of NMR-derived structural restraints. One is the inter-chain proton–proton distance, derived from nuclear Overhauser effects (NOEs) using previously established protocols[9,23] (see the Methods section and Supplementary Fig. 4). The other is the plane restraint (confinement of an assigned atom in a plane)[24], derived from NMR-based membrane partition analysis of the protein in bicelles[25]. The plane restraints are rarely used in NMR structure determination but useful here for constraining the CT segments that reside mostly on the plane of the bicelle. The residue-specific plane restraints were derived by (1) measuring solvent and lipophilic paramagnetic relaxation enhancement (PRE) amplitudes along the bicelle normal using the paramagnetic probe titration (PPT) method[25], and (2) calibrating the PRE amplitudes against the established structure and membrane partition of the TMD (residues 678–710)[8,26] (see Methods and Supplementary Fig. 5). The final structure of the TMD–CT^{LLP2} trimer was independently validated by inter-chain PRE analyses of multiple site-directed spin-labeling (Supplementary Fig. 6).

The structure of the TMD–CT^{LLP2} trimer in bicelles shows a novel membrane protein fold in which the CT winds around the TMD (Fig. 1a). The LLP2 is not a long continuous helix as previously proposed[19]; instead it forms two helices, designated H1 (residues 741–764) and H2 (residues 769–786), adopting a ring-like structure around the C-terminal region of the TMD trimer. The three H1s make direct contacts with the TMD trimer to form an inner ring around the TMD (Fig. 1a, b). The CT–TMD interactions involve the H1 nonpolar residues (L748 and L755), which form a hydrophobic cluster with TM residues (V701, L704, V705, and V708) of the same chain, and the polar H1 residue D759, which can make a salt bridge with the TM residue R709 of the neighboring chain. These residues thus collectively mediate the specific association between the CT ring and the TMD trimer. Indeed, hydrogen–deuterium (H–D) exchange data confirmed that residues 704–706, which showed fast H–D exchange in the TMD construct[26], are protected in the TMD–CT^{LLP2} sample, as expected (Supplementary Fig. 7). The three H2s fold around the inner ring at its periphery (Fig. 1a, c). This folding appears to be driven mainly by hydrophobic interactions involving H2 residues (V778, I781, and V782) of one chain and H1 residues (I746 and V749) of the neighboring chain. Based on the membrane partition data (Fig. 1d), H1s and H2s all reside in the headgroup region of the lipid bilayer. The palmitoylation site C764 (S764 in this construct) faces the lipid bilayer interior, consistent with its role in anchoring the LLP2 to the membrane.

A model of the MPER–TMD–CT^{LLP2} in lipid bilayer. Since the TM core (residues 685–700) has the same structure in the TMD, MPER–TMD, and TMD–CT^{LLP2} samples (Supplementary Figs. 1c, 3a, and 8), it can guide the creation of a composite model containing all three components. As such, we merged the NMR

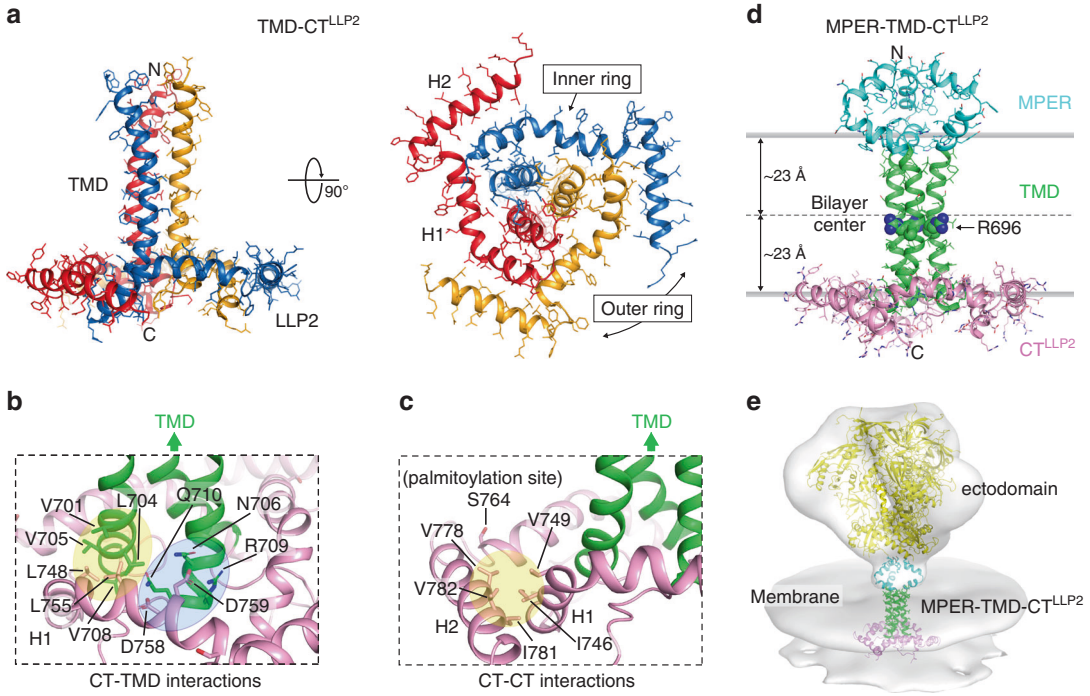


Fig. 1 Structures of TMD-CT^{LLP2} and MPER-TMD-CT^{LLP2} trimers in bicelles. a Ribbon representation of the TMD-CT^{LLP2} average structure from the calculated ensemble. The unstructured KS (residues 711–736) is omitted for clarity. The H1 and H2 helices forming the inner and outer rings of the CT baseplate, respectively, are indicated in the bottom view (right). **b** Close-up view of the residues establishing the CT–TMD interactions. Hydrophobic and hydrophilic interactions are shaded in yellow and light blue, respectively. The TMD and the CT are shown as green and pink ribbons, respectively. **c** Same as (**b**) but for the CT–CT interactions. The palmitoylation site C764 (S764 in our construct) faces the lipid bilayer interior. **d** Ribbon representation of the merged MPER-TMD-CT^{LLP2} model showing the MPER, TMD, and CT^{LLP2} in cyan, green, and pink, respectively. The placement of the structure in the lipid bilayer was determined experimentally using the PPT method[25]. The conserved intramembrane R696 is represented as spheres. **e** Fit of the MPER-TMD-CT^{LLP2} model and the structure of the SOSIP Env trimer (yellow; pdb ID: 5T3Z[61]) into the low-resolution EM density (gray) of the HIV-1 Env trimer on the virion surface by cryo-electron tomography[10] (Env trimer EMDB ID: EMD-5019; viral membrane EMDB ID: EMD-5020).

restraints from the three samples and constructed a model of the MPER–TMD–CT^{LLP2} (residues 660–788) (Fig. 1d, e) in the context of lipid bilayer, as described in the Methods section. The model, which accounts for ~75% of the entire membrane region of the Env, shows that the membrane-related components all form well-structured trimeric assemblies capable of relaying structural perturbation from one component to another across the membrane.

Physical interaction between the CT and the TMD. To investigate the propensity of the CT to oligomerize and interact with the TMD, we split the TMD–CT^{LLP2} construct into two fragments: one consisting only of the H1 and H2 helices (residues 737–788; designated CT^{LLP2}), and the other including the TMD and the KS (residues 677–725; designated TMD–KS) (Fig. 2a). When reconstituted in bicelles, the CT^{LLP2} did not trimerize according to the OG-label analysis[27] (Supplementary Fig. 9a). In contrast, when reconstituted in the presence of the TMD–KS, the CT^{LLP2} trimerized (Supplementary Fig. 9b) and folded around the TMD as shown by the PRE analysis (Fig. 2). In addition, to evaluate the structural impact of weakening the interaction between CT and TMD, we introduced five mutations in the CT H1 including L748S, L755S, D758A, D759A, and S762A. The mutant (designated CT2-tmd, also used in the antigenicity studies below) appeared to have maintained the overall trimeric structure of the wild type in bicelles as the NMR resonances of the TMD and H2 did not change significantly (Fig. 3a). The conformational stability of the mutant was examined by inter-chain PRE analysis, as was done in Supplementary Fig. 6b. Comparison of residue-specific PREs to that measured under identical condition for the TMD–CT^{LLP2} showed that the TMD PRE values of the mutant were reduced by as much as 30% (Fig. 3b). Since the two samples carried the same spin-label at C764, the results indicate that the H1 mutations above indeed had weaken the CT–TMD interaction and probably loosened the CT baseplate. The above independent experiments collectively suggest that the specific TMD–CT interactions drive spontaneous formation of the CT ring around the TMD.

Interaction between the CT and TMD influences Env antigenicity. In an earlier study, we showed that truncation or deletion of the CT diminishes Env binding to trimer-specific bnAbs that target the epitopes near the apex of the Env trimer[5,28–30]. With the new CT structural information, we now probe this effect with CT mutations designed to disrupt CT–TMD or CT–CT interactions shown in Fig. 1. The effect of the mutations on the antigenic structure of the Env ectodomain was evaluated by a pseudovirus-based neutralization assay[31], using bnAbs VRC01 (CD4 binding site)[32], PG9, PG16, and PGT145 (trimer-specific)[33,34], as well as non-neutralizing or strain-specific neutralizing antibodies, including b6 (CD4 binding site)[32], 3791 (V3)[35], and 17b (CD4-induced)[36] (mutant list and results in Supplementary Tables 1 and 2, respectively; epitope mapping is shown in Supplementary Fig. 10). Although virus neutralization is not a direct measure of antibody binding to the Env as the process could be influenced by other factors such as possible CT-matrix protein (MA) interaction and membrane fusion kinetics, earlier studies have shown that loss of neutralization is overall correlated with the loss of Env binding to bnAbs[5,37]. We generated 28 Env mutants using the sequence of a

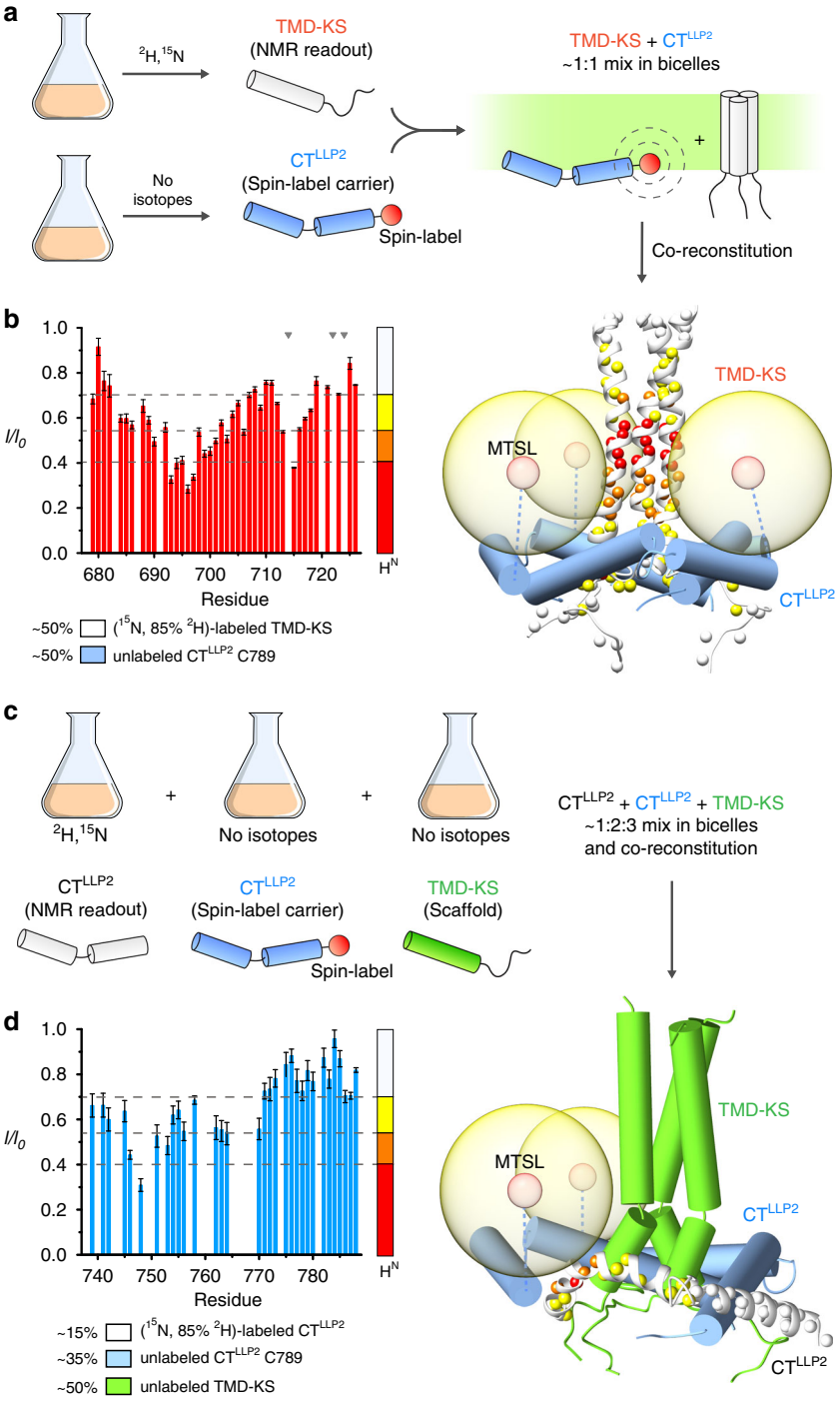


Fig. 2 Independent evidence of CT^{LLP2}–TMD interaction. a Schematic illustration of the sample preparation strategy used for intermolecular PRE analysis. TMD-KS and CT^{LLP2} (C789) are expressed separately: TMD-KS is isotopically-enriched for NMR readout while CT^{LLP2} carries the spin-label (at C789). After purification, the two segments are mixed at ~1:1 molar ratio and co-reconstituted in bicelles. Upon interaction, the TMD-KS is expected to experience PRE generated by the CT^{LLP2} spin-label. **b** Residue-specific PRE (I/I_0) of (^{15}N, 85% ^{2}H)-labeled TMD-KS mixed with MTSL-labeled CT^{LLP2} (left). Error bars represent the uncertainty derived from cross-peaks signal to noise. Missing bars are due to prolines (indicated by gray triangles) or overlapping residues. The horizontal dash lines mark the four PRE regimes used to map the PREs onto the protein structure (right). The TMD-KS and the CT^{LLP2} are shown as white ribbons and blue cylinders, respectively. **c** Labeling scheme for probing intermolecular CT^{LLP2}–CT^{LLP2} interaction. CT^{LLP2} (white) is isotopically-enriched for NMR readout while CT^{LLP2} (blue) carries the spin-label at C789, and TMD-KS (green) serves as scaffold. After purification, the three proteins are mixed at ~1:2:3 molar ratio, respectively, and co-reconstituted in bicelles. **d** Residue-specific PRE (I/I_0) of (^{15}N, 85% ^{2}H)-labeled CT^{LLP2} mixed with MTSL-labeled CT^{LLP2} and scaffold TMD-KS (left). Error bars represent the uncertainty derived from cross-peaks signal to noise. Missing bars are due to overlapping residues. The horizontal dash lines mark the four PRE regimes used to map the PREs onto the protein structure (right). Source data are provided as a Source data file.

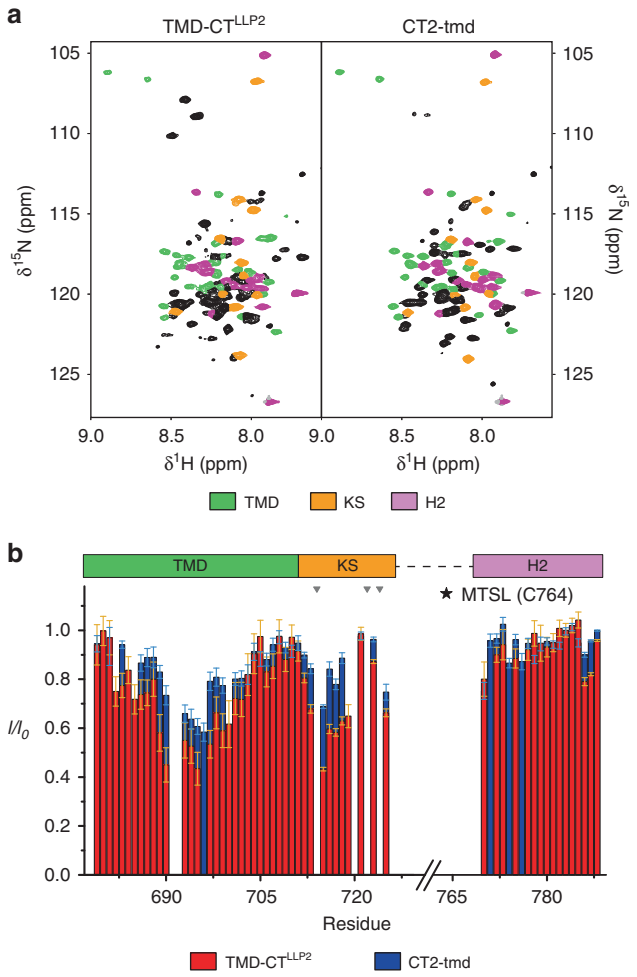


Fig. 3 Destabilization of the TMD–CT interactions loosens the CT baseplate. a Comparison of the 2D ^{1}H–^{15}N TROSY-HSQC spectra of the TMD–CT^{LLP2} (left) and CT2-tmd (right), showing strong agreement among the two constructs of the chemical shift of residues 679–710 (TMD), 711–725 (KS), and 770–788 (H2) highlighted in green, orange, and pink, respectively. Bearing the mutation site, H1 exhibited strong chemical shift changes (black cross-peaks) and thus was not further analyzed. **b** Comparison of residue-specific PRE (I/I_0) of TMD–CT^{LLP2} (red) versus those of CT2-tmd (blue). Each sample was prepared mixing ~1:1 (^{15}N, 85% ^{2}H)-labeled protein (NMR readout) with unlabeled (and thus "NMR-invisible") protein carrying the MTSL spin-label at position S764C, following the strategy summarized in Fig. 2a (see the Methods section for details). Error bars represent the uncertainty derived from cross-peaks signal to noise. Missing bars are due to prolines (indicated by gray triangles) or overlapping residues. The position of the paramagnetic tag is marked by a star. Source data are provided as a Source data file.

clade A isolate 92UG037.8, and we further characterized 11 of them (indicated in Supplementary Table 1) which showed significant phenotype in the antibody neutralization assay. When expressed in 293T cells, these mutants produced comparable levels of Env, with similar extents of cleavage between gp120 and gp41 (Supplementary Fig. 11). In particular, none of these mutations in the CT region have led to dramatic increase in cell-surface levels (Supplementary Fig. 12). At a low Env expression level to mimic the low surface density on HIV-1 virions[38], only the mutant TMD-ct showed a moderate reduction in cell–cell fusion activity, but still at ~50% of that of the wild-type Env (Supplementary Fig. 13a). At a high Env expression level, the fusion activity of all the mutants was essentially indistinguishable from that of the wild type (Supplementary Fig. 13b), suggesting the high cell-surface density of Env can compensate for defects in membrane fusion caused by the CT mutations, as observed for mutations in the TMD and MPER previously[8,9]. When these mutations were introduced into pseudoviruses, there were moderate changes in Env incorporation for mutants V749K and A756N, and in processing for those with large-scale mutations, including the mutant TMD-ct (Supplementary Fig. 14), consistent with the role of the CT in Env incorporation[39]. In addition, the mutant F774N showed substantially increased viral infectivity (165%) while the rest of mutants have impaired infectivity, ranging from 10 to 63% of that of the wild-type Env, suggesting that the CT can both positively and negatively modulate the efficiency of viral entry (Supplementary Fig. 15).

As expected, the wild-type Env is neutralized by VRC01, PG9, PG16, and PGT145, but resistant to b6, 3791, and 17b (Fig. 4a). The single mutants all showed neutralization patterns very similar to that of wild-type Env (Supplementary Fig. 16). Among the four double mutants, the two containing D759R showed minor but noticeable reduction (~20%) in sensitivity to the trimer-specific bnAbs. The triple mutant (L702S-A756N-D759R) showed very significant resistance (~40%) to the trimer-specific bnAbs, especially for PG9 and PGT145 (Fig. 4b). These results suggest the important role of D759 in interacting with either R709 or N706 of the TMD (Fig. 1b), while underlining the nature of the CT structure being resistant to small changes. Larger scale mutagenesis of the CT–TMD interface resulted in mutants consistently less sensitive to the trimer-specific bnAbs, but remained totally resistant to the non-neutralizing antibodies (Fig. 4c and Supplementary Fig. 16). As a negative control, the mutant (A756N-L760S-L763N-C764E), containing mutations on the other side of the H1 helix, did not show any phenotype (Fig. 4d). Finally, hydrophobicity/hydrophilicity swapping of the key TMD residues in the CT–TMD interface generated a mutant (TMD-ct) that, while still sensitive to VRC01, was completely resistant to the trimer-specific bnAbs and also substantially sensitive to b6, 3791, and 17b (Fig. 4e). These results suggest that disrupting the packing interface between the TMD and CT can destabilize the Env ectodomain and shift it to an open conformation[10,40,41], further supporting our previous notion that the CT can modulate the antigenic structure of the Env trimer[5].

To further verify the impact of these mutations, we performed tier phenotyping[42] using seven HIV+ chronic serum samples with the 92UG037.8 (a tier 2 virus) wild-type and the TMD-ct mutant, along with the CT2-tmd and CT3-tmd mutants as representative "intermediate" phenotype viruses (Supplementary Table 3). As expected, the TMD-ct mutant became much more sensitive to the HIV+ chronic sera than the wild type, consistent with an open Env conformation and the tier-1 phenotype[42], while the two intermediate mutants were similar to the WT in the tier phenotype probably due to limited local changes near the trimer apex. We also tested additional bnAbs that target epitopes in CD4bs (binding site), V3-glycan, V1/V2-glycan, gp120/gp41 interface, and the MPER (Supplementary Table 4). While the wild-type and mutant viruses showed no major differences in sensitivity to most of these bnAbs, the mutants were ~10-times more sensitive to the MPER-directed antibody 4E10 than the wild-type virus, suggesting that the mutations in the CT and/or TM regions significantly alter the MPER structure.

TMD dynamic parameters suggest pivotal motion of the TMD trimer. Physical coupling between the CT and the apex of the ectodomain must be mediated via the TMD, but no obvious differences are observed between the NMR structures of the TMDs with and without the CT. We thus examined backbone

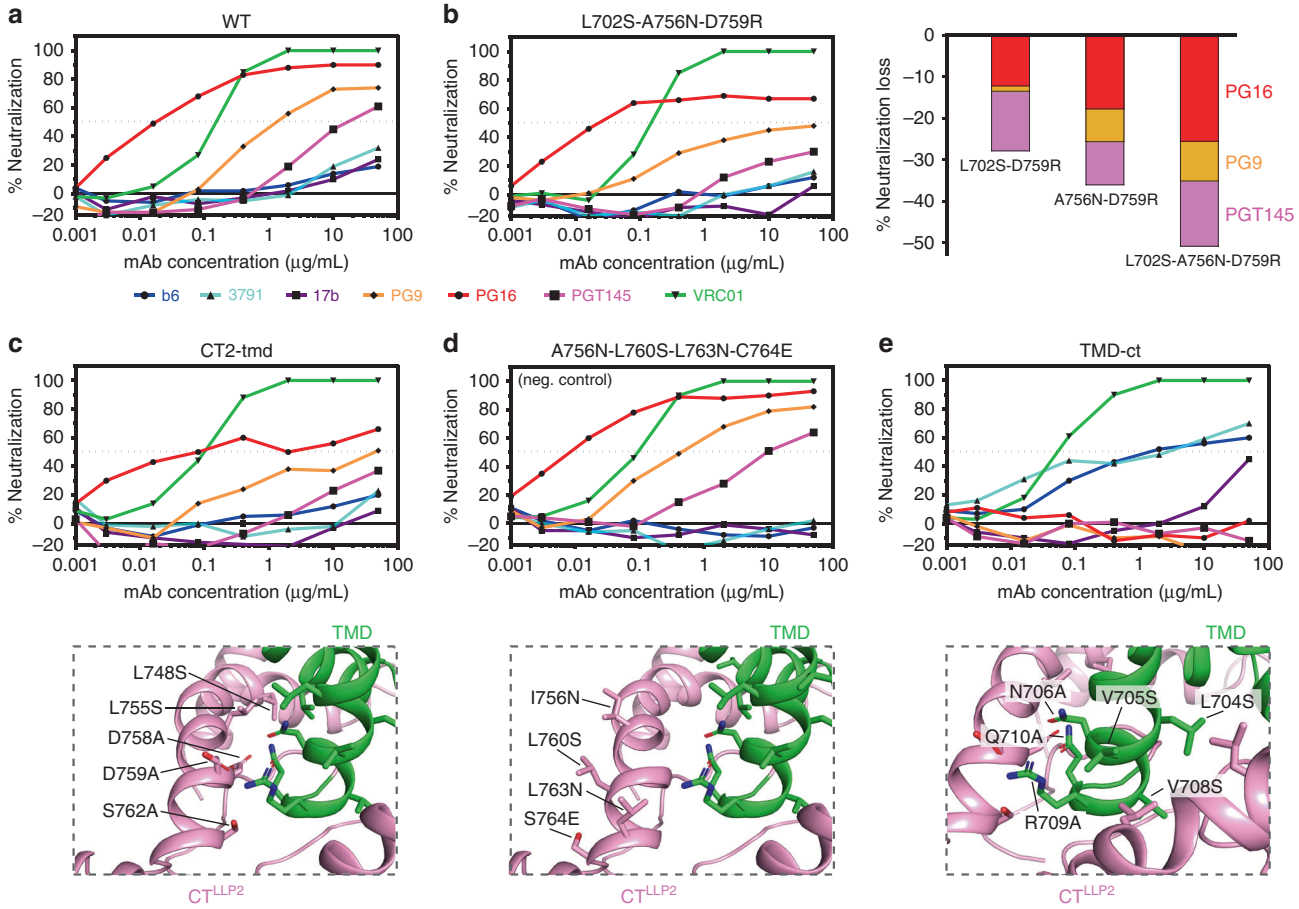


Fig. 4 Effect of mutations in the CT on Env antibody sensitivity. a Antibody neutralization of pseudovirus containing the 92UG037.8 Env determined for non-neutralizing antibodies, including b6 (CD4 binding site; blue), 3791 (V3; cyan), and 17b (CD4-induced; purple), and trimer-specific bnAbs, including PG9 (orange), PG16 (red), and PGT145 (magenta). The CD4 binding site bnAb VRC01, used as a control antibody, is shown in green. **b** Antibody neutralization of pseudovirus containing the L702S-A756N-D759R mutant. Right panel shows the relative decrease of sensitivity to the trimer-specific bnAbs of double and triple mutants bearing the D759R. **c** Antibody neutralization of pseudovirus (upper panel) containing a larger number of mutations in the CT (lower panel). **d** Same as (**c**) but for the mutant with mutations on the opposite side of the CT-TMD interface. **e** Same as (**c**) but for the mutant containing TMD mutations that break the hydrophobic and hydrophilic TMD-CT interactions. Source data are provided as a Source Data file.

dynamics across the TM in the MPER–TMD and TMD–CT^{LLP2} constructs by measuring the product of NMR ^{15}N R_1 and R_2 relaxation rates, R_1R_2, which is a probe for relative ms–μs motion associated with conformational exchange[43]. The MPER–TMD shows significantly higher dynamics in the MPER and C-terminal region of the TMD than the hydrophobic core of the TMD (residues 686–689), suggesting that both ends are unconstrained (Fig. 5a). We then locked the MPER by introducing the L660C and A667C mutations, which are positioned to form inter-chain disulfide in the MPER prefusion structure (Supplementary Fig. 17a). Locking the MPER resulted in substantially reduced motion not only in the MPER, as expected, but also in the C-terminal region of the TMD at the opposite side of the membrane; no change in the TMD hydrophobic core (Fig. 5a). Similar effects are observed in the TMD–CT^{LLP2} where the R_1R_2 values for the TMD C-terminal region are less than that of the unlocked MPER–TMD (Fig. 5a). These data suggest that the three TM helices undergo a "scissor-like" movement around the hydrophobic core (or the hinge), at which the trimer remains tightly associated as shown by H–D exchange analysis (Supplementary Fig. 7 and ref. [26]), and that constraining either N- or C- terminal end of the TM helix with locked MPER or CT, respectively, can affect the opposite end of the TMD. Therefore, disruption or deletion of the CT baseplate can in principle destabilize the MPER structure via the TMD (Fig. 5b). Since the MPER is a control relay that modulates the ectodomain conformation and antigenic properties[9], the CT-MPER coupling is also expected to shift conformational equilibria of the ectodomain and alter its antibody binding profile.

Discussion

We have shown that the LLP2 segment of the Env CT forms amphipathic helices that wrap around the TMD, forming a support baseplate for the TMD and the rest of the Env. The other two LLP segments (LLP3 and LLP1) of the CT can also form amphipathic helices[19] and we suggest that they may fold around the LLP2 ring to further expand the baseplate (Fig. 5b). This model provides a structural explanation for our previous observations that progressive truncations of the CT incrementally reduced sensitivity of the Env ectodomain to PG16-like bnAbs, and that the effect reached maximum after the complete deletion of the LLP2[5], as progressive truncations of the CT would gradually weaken the base support for the TMD. Indeed, our data on the co-refolding of the TMD–KS and CT^{LLP2} fragments show that the interaction between H1 and TMD is critical to the proper formation of the baseplate and possibly to the structural integrity of the entire Env. It is unclear whether the CT baseplate is

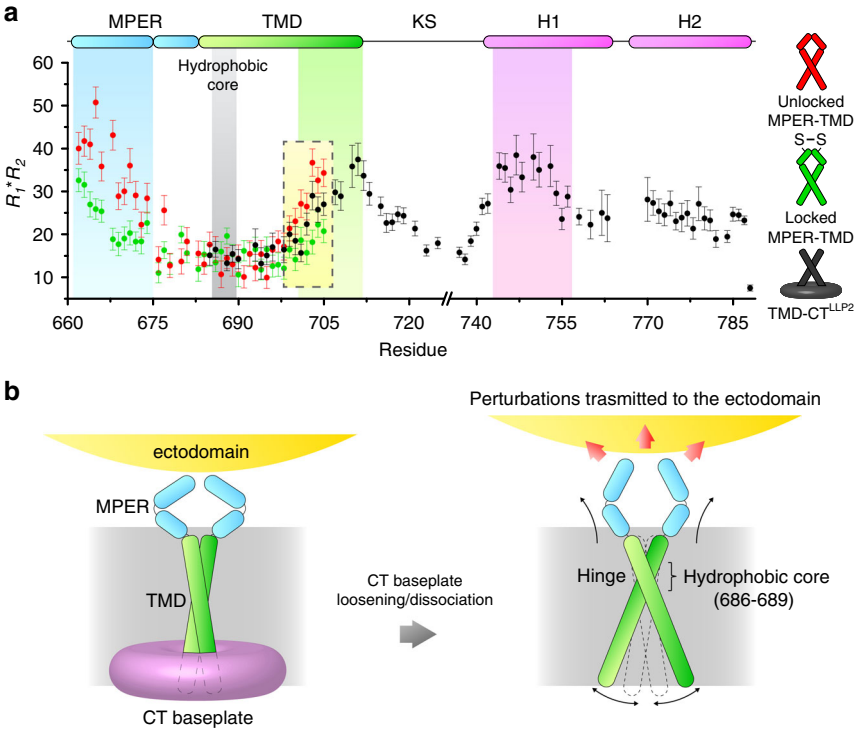


Fig. 5 Conformational coupling between the Env CT and ectodomain. a Backbone conformational dynamics of the unlocked MPER-TMD (red), locked MPER-TMD (green) and TMD-CT^{LLP2} (black) probed by the product of ^{15}N R_1 and R_2 relaxation rates, which reflects ms-μs motion. Error bars represent the uncertainty derived from R_1 and R_2 fitting errors. The regions characterized by higher R_1R_2, i.e., the MPER, the C-terminal part of the TMD, and its interacting CT region (H1), are highlighted in cyan, green, and purple, respectively. The position of the hydrophobic core (or the TM hinge) is shown in gray. The yellow box highlights the decrease of R_1R_2 in the C-terminal part of the TMD due to locking the MPER or CT presence. **b** Proposed mechanism of CT-ectodomain coupling mediated by the TMD. For simplicity, only two chains of the trimer are illustrated. Left: the CT baseplate confines the TMD motion, important for stabilizing the Env trimer in the prefusion state. Right: deletion of the CT baseplate or disruption of CT–TMD interactions allows greater TMD hinge motion, destabilizing the MPER, which in turn modulates the antigenic structure of the ectodomain. Source data are provided as a Source data file.

relevant to viral entry. The new structural information further consolidates the notion that conformational state of the membrane-related components of the Env can influence the antigenic structure of the ectodomain.

Rigorous understanding of the mode of coupling between the Env ectodomain and CT across the TMD would require visualization of the full-length Env at high resolution. Our NMR dynamics study of the MPER–TMD and TMD–CT^{LLP2} fragments in bicelles provided compelling data suggesting physical coupling between the MPER and CT. The observation that locking the MPER with inter-chain disulfides reduced dynamics of the C-terminal end of the TMD ~50 residues away suggests that the proposed cross-membrane coupling of the Env is entirely possible by structural consideration. It was somewhat surprising that TMD C-terminal end showed slightly more dynamics in the TMD–CT^{LLP2} than in the locked MPER–TMD, because we thought CT–TMD interaction may impose a more direct constraint on the TMD motion. Nevertheless, this result is not inconsistent with the proposed model in Fig. 5. Comparing to the unlocked MPER–TMD, the TMD–CT^{LLP2} construct indeed showed reduced dynamics at the C-terminal end of the TMD, although the CT–TMD interaction still may not produce an effect as strong as covalently locking the MPER. The key message from the dynamics data is that the pivotal or the scissor-like motion of the TMD can be modulated by either MPER or CT.

To the best of our knowledge, a structural arrangement of a cytoplasmic baseplate supporting the ectodomain on the other side of membrane has never been observed in any known TM proteins including cellular receptors and viral fusion proteins. Other lentiviral fusion proteins also have long CTs that are predicted to form multiple amphipathic helical segments, such as those from EIAV, MVV, and CAEV[17,18,44] (Supplementary Fig. 18). Whether these viral proteins adopt the unusual membrane-proximal fold seen for the HIV-1 Env CT remains to be investigated. Nevertheless, a trimeric baseplate of the CT together with the TMD appears to pose for specific interactions with the trimeric matrix (MA) protein[45–47], in the context of membrane.

In conclusion, our study provides a structural basis for how the CT of HIV-1 Env is physically coupled to its ectodomain through the TM and membrane-proximal regions and how the CT stabilizes the antigenic structure of the Env trimer on the opposite side of membrane. It is widely believed that the prefusion conformation of a functional HIV-1 Env on the surface of infectious virions is probably the most appropriate vaccine candidate for eliciting effective antibody responses, because it is recognized by most bnAbs and not by other tier-1 neutralizing antibodies[5,48,49]. Structural definition of the prefusion Env trimer at high resolution, which can provide a useful tool to facilitate immunogen selection and design, has limitations due to artificial modifications, extra ligands required for stability or the absence of membrane, compounded by observations with other spectroscopic approaches, such as single-molecule fluorescence resonance energy transfer (smFRET) and double electron–electron resonance (DEER) spectroscopy[50–52]. Our data presented here suggest new strategies to stabilize the native state of Env and help develop more effective Env trimer immunogens for clinical studies.

Methods

Protein expression and purification. The gp41 fragment TMD–CT^{LLP2} (residues 677–788) from a clade D HIV-1 isolate 92UG024.2 was synthesized by GenScript (Piscataway, NJ). The expression construct was created by fusing the TMD–CT^{LLP2} to the C-terminus of the His$_9$-TrpLE expression sequence in pMM-LR6 vector with an added methionine in-between for subsequent cleavage during protein purification, as previously described[8,9,23]. Mutants and additional constructs (e.g., TMD–KS and CT^{LLP2}, Supplementary Tables 5 and 6) were generated by standard PCR protocols and confirmed by DNA sequencing.

Each protein construct was expressed by growing transformed *E. coli* strain BL21 (DE3) cells in LB or M9 minimal media (when isotopic labeling was required). Cultures were grown at 37 °C until they reached an optical density of ~0.6 and were then cooled to 20 °C before induction with 0.1 mM isopropyl β-D-thiogalatopyranoside (IPTG). Protein was expressed at 20 °C for ~18–24 h.

After growth, cells were harvested, suspended in a lysis buffer (50 mM Tris, pH 8.0, and 200 mM NaCl), and lysed by sonication. Inclusion bodies were separated by centrifugation at 25,400 × *g* and suspended in a denaturing buffer (1% Triton X-100, 6 M guanidine hydrochloride, 50 mM Tris, pH 8.0, and 200 mM NaCl). Inclusion bodies were homogenized using a glass tissue grinder, dissolved and centrifuged at 25,400 × *g*. The fusion protein was bound to nickel affinity resin (Sigma-Aldrich), washed with 8 M Urea and dH_2O, and finally eluted with 90% formic acid (FA). The gp41 construct was cleaved from the TrpLE by hydrolyzing the peptide bond at the C-terminus of the methionine using cyanogen bromide (CNBr) (~0.1 g/mL) in 90% FA for 1 h. The reaction mixture was dialyzed (MWCO 3.5 kDa) to remove the excess of CNBr and FA and then lyophilized. The protein powder was dissolved in 90% FA and purified by reverse-phase high-performance liquid chromatography (RP-HPLC) in a Zorbax SB-C3 column (Agilent Technologies, Santa Clara, CA) with a gradient from 95% dH_2O, 5% isopropanol (IPA), 0.1% trifluoroacetic acid (TFA) (buffer A) to 75% IPA, 25% acetonitrile, 0.1% TFA (buffer B) (Supplementary Fig. 2a-b). Fractions containing the pure gp41 construct (confirmed by SDS-PAGE analysis) were collected and lyophilized.

Protein reconstitution. The lyophilized gp41 construct was dissolved in 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) and mixed with 9 mg of 1,2-Dimyristoyl-sn-glycero-3-phosphocholine (DMPC) (from Avanti Polar Lipids, Alabaster, AL or from FB Reagents, Boston, MA, if deuterated) and 27 mg of 1,2-Dihexanoyl-sn-glycero-3-phosphocholine (DHPC) (from Avanti Polar Lipids Alabaster, AL or from FB Reagents Boston, MA, if deuterated). The mixture was dried under a nitrogen stream until a thin film was achieved and lyophilized overnight to completely remove any trace of residual organic solvent. The thin film was then dissolved in 3 mL of 8 M Urea and dialyzed (MWCO 3.5 kDa) against 40 mM MES buffer, pH 6.7 to remove the denaturant. During and after the dialysis, additional DHPC was added to make up for the DHPC lost during the dialysis, adjusting the DMPC:DHPC ratio (q) to ~0.5. The reconstituted protein was then concentrated using a Centricon (EMD Millipore, Billerica, MA) (MWCO 3.5 kDa) to ~300–350 µL. The buffer of the final NMR sample contained 40 mM DMPC, 80 mM DHPC, 40 mM MES, pH 6.7, 1% NaN_3, and 10% (v/v) D_2O (for the NMR lock). Finally, the bicelle q of the NMR sample was quantified by signal integration of the DMPC and DHPC methyl peaks in the 1D ^{1}H NMR spectrum and adjusted to exactly 0.5.

Determination of the protein oligomeric state. To determine the oligomeric state of the bicelle-reconstituted TMD–CT^{LLP2}, TMD–KS, and CT^{LLP2}, standard SDS-PAGE analysis was initially used. The proteins were reconstituted in DMPC/DHPC bicelles at $q = 0.5$, mixed with an SDS sample buffer (Invitrogen) and boiled for 5 min, followed by SDS-PAGE (12% Bis-Tris protein gel) at 200 V for 30 min and Commassie blue staining. The TMD–KS migrated at the apparent MW of ~17 kDa (theoretical MW is 5.8 kDa), indicating that it had remained trimeric despite SDS denaturation as previously shown[8,9] (Supplementary Fig. 9b, lane 2 of the gel). The TMD–CT^{LLP2} and CT^{LLP2}, instead, migrated at the apparent MW of ~12 and ~7 kDa (theoretical MW are 11.8 and 6.1 kDa, respectively), indicating that both had become monomeric upon SDS denaturation (Supplementary Fig. 1b, lane 1 of the gel, and Supplementary Fig. 9, lanes 1 and 2 of the gels, respectively). Therefore, their native oligomeric states were quantified using the non-denaturing method known as OG-label[23,27].

In the OG-label method, each protomer of the oligomer to be studied is non-covalently labeled with a soluble cross-linkable protein (SCP), so that the latter can be cross-linked with Lomant's reagents to read out the sample oligomeric state. The small Ig-fold protein named GB1 (MW = 8.4 kDa) has been proven to serve as the SCP very effectively. A TriNTA molecule is linked via PEG-2-SMCC (succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate) to the N-terminus of GB1, while a His$_6$-tag is added to the C-terminus of the oligomeric protein, so that the TriNTA-GB1 conjugate can strongly attach to the protomer (the binding affinity of TriNTA to His$_6$-tag is 20 ± 10 nM). The GB1s are then cross-linked to report the oligomeric state of the protein, as the local concentration of stoichiometric amount of GB1 to the oligomer allows for more efficient cross-linking than for the free GB1 in solution. Finally, the cross-linked GB1s are released from the oligomer by addition of EDTA and analyzed by SDS-PAGE.

To implement the OG-label method for the TMD–CT^{LLP2} and CT^{LLP2}, a His$_6$-tag was added at the C-terminus of the two proteins. Three samples were prepared containing either (i) TMD–CT^{LLP2}, (ii) CT^{LLP2} or (iii) CT^{LLP2} in presence of TMD–KS in ~1:1 ratio. The His$_6$-tagged proteins were expressed, purified and reconstituted in bicelles ($q = 0.5$) as previously described, except the sample buffer was a phosphate buffer (pH 7.4) for better cross-linking efficiency. To prevent unwanted cross-linking between the membrane protein and GB1, all the free primary amines of TMD–CT^{LLP2}, CT^{LLP2}, or TMD–KS were blocked by reacting with 100-fold molar excess of Sulfo-NHS acetate (Thermo Fisher Scientific) at room temperature for 1.5 h. Excess Sulfo-NHS acetate was removed by dialysis while tightly controlling the bicelle q. After dialysis, the samples were concentrated to 40 µM and mixed with 60 µM TriNTA-GB1 to ensure that all the His$_6$-tags were saturated with TriNTA-GB1. The mixtures were then incubated at room temperature with various concentration of BS3(PEG9) (0.3, 0.9, 3.0, and 3.0 mM, corresponding to lane 4, 5, 6, and 7 in the gels shown in Supplementary Figs. 1b and 9, respectively) for 30 min, followed by a second incubation at room temperature with various concentration of glutaraldehyde (0.6, 0.6, 0.6, and 1.8 mM, corresponding to lane 4, 5, 6, and 7 in the gels shown in Supplementary Figs. 1b and 9, respectively) for 3 min. The cross-linking reactions were quenched with 20 mM Tris buffer (pH 7.5) upon incubation at room temperature for 15 min. As negative control, 0.9 mM BS3(PEG9) and 0.6 mM glutaraldehyde were sequentially added to 60 µM TriNTA-GB1 in the absence of the His$_6$-tagged samples (lane 3 in the gels shown in Supplementary Figs. 1b and 9), indicating that the TriNTA-GB1 alone remains mostly monomeric and only partially dimerizes in the conditions used. The cross-linked GB1s were then released from the samples by adding 50 mM EDTA and examined by SDS-PAGE using 12% Bis-Tris protein gels (Thermo Fisher Scientific) (Supplementary Figs. 1b and 9).

NMR data acquisition and processing. The NMR experiments were performed at a) 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; (b) 17.6 T on a Bruker Avance III spectrometer operating at 749.66 MHz ^{1}H, 188.50 MHz ^{13}C, and 75.96 MHz ^{15}N frequencies; (c) 18.8 T on a Bruker Avance III spectrometer operating at 800.28 MHz ^{1}H, 201.23 MHz ^{13}C, and 81.09 MHz ^{15}N frequencies; (d) 21.1 T on a Bruker Avance III spectrometer operating at 900.17 MHz ^{1}H, 226.35 MHz ^{13}C, and 91.21 MHz ^{15}N frequencies. All the spectrometers were equipped with a cryogenic probe. All the measurements were performed at 308 K if not stated otherwise. The most relevant acquisition parameters of the experiments are reported in Supplementary Table 7.

The NMR data sets were processed with *nmrPipe*[53] and the resulting NMR spectra were analyzed with *Sparky* (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) and *XEASY*[54]. 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. For comparison purpose, the chemical shift assignments of the TMD and MPER–TMD were taken from the Biological Magnetic Resonance Bank (BMRB), entries 30090 and 30503, respectively[8,9]; the TMD and MPER–TMD average structures were taken from the Protein Data Bank (PDB), entries 5JYN and 6E8W, respectively[8,9].

NMR resonance and NOE assignment. Sequence specific assignment of TMD–CT^{LLP2} backbone chemical shifts (BMRB accession code 30678) was accomplished using a set of TROSY-enhanced triple resonance experiments (HNCA, HN(CO)CA, HN(CA)CO, HNCO and HNCACB)[55,56], recorded on a (^{15}N, ^{13}C, 85% ^{2}H)-labeled sample. In addition, an ultra-high-resolution 3D ^{15}N-edited NOESY-TROSY-HSQC ($\tau_{mix} = 180$ ms) spectrum of a (^{15}N, ^{2}H)-labeled sample resulted extremely useful in completing the backbone resonance assignment. Protein aliphatic and aromatic resonances were assigned using a combination of 2D ^{13}C HSQC, 3D ^{15}N-edited NOESY-TROSY-HSQC ($\tau_{mix} = 80$ ms), and ^{13}C-edited NOESY-HSQC ($\tau_{mix} = 150$ ms). These NOESY experiments were performed on a (^{15}N, ^{13}C)-labeled sample reconstituted in bicelles with deuterated DMPC and DHPC acyl chains.

NOE-derived intra-chain distance restraints for the TMD–CT^{LLP2} were obtained from the 3D ^{15}N-edited and ^{13}C-edited NOESY spectra used for aliphatic and aromatic resonance assignment above. Assigning inter-chain distance restraints, however, faced the challenge of measuring NOEs between structurally equivalent protomers having the same chemical shifts. To overcome this problem, we used a mixed sample in which half of the protomers were ^{15}N, ^{2}H-labeled and the other half ^{13}C-labeled. Recording a 3D ^{15}N-edited NOESY-TROSY-HSQC ($\tau_{mix} = 200$ ms) on this sample allowed identification of NOEs exclusively between the ^{15}N-attached protons of one protomer and the aliphatic protons of the neighboring one. The identified NOEs were unambiguously confirmed by performing the 3D J_{CH}-modulated NOE experiment[9,23], in which two interleaved NOESY spectra were recorded with varying the J_{CH} evolution ($J_{CH} = 0$ ms and $J_{CH} = 8$ ms) before the NOE mixing. Subtraction of the two spectra allowed selection of only the inter-chain NOEs. Finally, a control sample containing only (^{15}N, ^{2}H)-labeled protein was used to perform an identical 3D ^{15}N-edited NOESY-TROSY-HSQC experiment as done for the mixed sample above (note: the ^{15}N, ^{2}H-labeled TMD–CT^{LLP2} in the mixed and control samples were from the same protein expression batch). Comparison of the mixed and control spectra provided another confirmation that the detected inter-chain NOEs were due to the mixing of protomers with different labeling schemes and not to residual protonation. This

assignment strategy is summarized in Supplementary Fig. 4, where several examples of unambiguously assigned inter-chain NOEs are shown.

NMR-based membrane partition analyses. The membrane partition of the TMD–CT^{LLP2} was determined using the paramagnetic probe titration (PPT) method[23,25]. As previously shown, DMPC/DHPC bicelle with sufficiently large q (≥0.5) allows direct use of measurable paramagnetic relaxation enhancement (PRE) to probe residue-specific immersion depth of the protein in the bilayer region of the bicelle. Two PPT analyses were performed: titrating the bicelle-reconstituted TMD–CT^{LLP2} with (1) the soluble paramagnetic agent Gd-DOTA, and (2) the lipophilic paramagnetic agent 16-Doxyl-stearic acid (16-DSA). The titrants were taken from concentrated stock solutions (600 mM Gd-DOTA and 24 mM 16-DSA) in the same buffer as that of the protein sample and were added in small aliquots (few µL per step) to minimize sample dilution. The PRE increase was monitored by recording a 2D ^{1}H–^{15}N TROSY-HSQC spectrum at each of the titrant concentrations: 0 (reference), 2.0, 4.0, 6.0, 8.0, 10.0, 15.0, and 20.0 mM for Gd-DOTA; 0 (reference), 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, and 4.2 mM for 16-DSA. The residue-specific PRE$_{amp}$, which is the amplitude of the PRE experienced by an amide proton in the protein, was determined by fitting the peak intensity decay as a function of [paramagnetic probe] to the following exponential decay equation:

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

where I and I_0 are the peak intensities in the presence and absence of the paramagnetic probe, respectively, [paramagnetic probe] is the concentration of the paramagnetic agent (Gd-DOTA or 16-DSA), τ is the decay constant and PRE$_{amp}$ is the PRE amplitude. The residue-specific PRE$_{amp}$ (Supplementary Table 8) were then used to determine the membrane partition of the protein (r_Z) by the sigmoidal fitting method[23,25], in which the position of the TMD–CT^{LLP2} trimer along the bilayer normal was fitted to satisfy Eq. (2):

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

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.

Measurement of inter-chain PREs. Inter-chain PREs were measured for cross validation of the NOE-derived TMD–CT^{LLP2} structure (Supplementary Fig. 6). Three mixed samples were prepared by mixing, at ~1:1 molar ratio, the (^{15}N, 85% ^{2}H)-labeled TMD–CT^{LLP2} with one of the following unlabeled TMD–CT^{LLP2} construct containing a specific Cys mutation/addition: (a) G738C; (b) S764C (reintroducing the native cysteine); (c) C789 (Cys added to the C-terminus). The mutant proteins were prepared as in the "Protein expression and purification" section above. After reconstitution in bicelles under reducing conditions, DTT was removed from the samples by dialysis at low pH (6.2). The pH was then rapidly raised to 7.4 and 100 mM MTSL (in DMSO) was added to a final ratio of 10:1 (MTSL to Cys-mutant TMD–CT^{LLP2}) and allowed to react at room temperature overnight. Excess MTSL was removed by extensive dialysis (pH 6.7). The samples were then concentrated to 360 µL for NMR measurements. During the entire sample preparation, MTSL-containing solutions were shielded from light. As a negative control, a sample containing only (^{15}N, 85% ^{2}H)-labeled TMD–CT^{LLP2} (without Cys) was prepared using the same MTSL-labeling procedure used for the mixed samples, to ensure that free MTSL removal with our protocol was complete (Supplementary Fig. 6d). To quantify the PREs, defined as the ratio of the peak intensities in the paramagnetic (I) and diamagnetic state (I_0), a 2D ^{1}H–^{15}N TROSY-HSQC spectrum was recorded before and after reducing the spin-label by addition of 20 mM sodium ascorbate (pH 6.7).

The same protocol was used to study the interaction between the TMD and the CT (Figs. 2 and 3). Three mixed samples were prepared by mixing: (1) (^{15}N, 85% ^{2}H)-labeled TMD–KS with unlabeled CT^{LLP2} C789 (~1:1); (2) (^{15}N, 85% ^{2}H)-labeled CT^{LLP2} with unlabeled CT^{LLP2} C789 and unlabeled TMD–KS (~1:2:3); (3) (^{15}N, 85% ^{2}H)-labeled CT2-tmd (L748S-L755S-D758A-D759A-S762A) with unlabeled CT2-tmd S764C (~1:1).

Measurement of H–D exchange. Solvent accessibility of the TMD–CT^{LLP2} was examined by performing an H–D exchange experiment at 303 K (Supplementary Fig. 7). The TMD–CT^{LLP2}, reconstituted in protonated solvent (pH 6.0), was flash-frozen in liquid nitrogen and thoroughly lyophilized. The dried sample was then dissolved in 360 µL of 99.9% D_2O (pD ~ 6.4). 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 h up to ~3.25 days. The residue-specific exchange constant, k_{ex} ($=1/\tau_{ex}$) (Supplementary Table 9), was determined by fitting the fractional peak intensity vs. time to the following exponential decay equation:

$$I_{(t)} \propto e^{-\frac{t}{\tau_{ex}}} \quad (3)$$

where I is the peak intensity, t is the time passed from the beginning of the exchange, and τ_{ex} is the time constant of the decay. Finally, k_{ex} values were divided in four different exchange regimes defined as: very fast (τ_{ex} < 1 h), fast (1 h ≤ τ_{ex} < 3 h), slow (3 h ≤ τ_{ex} < 1 day), and very slow (τ_{ex} ≥ 1 day).

Measurement of NMR relaxation rates. Backbone ^{15}N dynamics of the bicelle-reconstituted TMD–CT^{LLP2} (14.1 T, 303 K) was investigated by measuring ^{15}N R_1 and R_2 relaxation rates using the TROSY version of the standard experiments[57] (Supplementary Fig. 17b). For ^{15}N R_1, 8 experiments were acquired with the following relaxation delays: 10, 50, 100, 200, 300, 600, 800, and 1000 ms. For ^{15}N R_2, 8 experiments were acquired with the following relaxation delays: 6.4, 10, 20, 30, 40, 50, 64, and 80 ms. The ^{15}N R_1 and R_2 values were determined by fitting the peak intensity vs. relaxation delay to the exponential decays:

$$I_{(t)} \propto e^{-R_1 t} \quad (4)$$

$$I_{(t)} \propto e^{-R_2 t} \quad (5)$$

where I is the peak intensity at a given relaxation delay, t is the relaxation delay, and R_1 and R_2 are the relaxation rates.

The same type of analysis was performed also on the unlocked and locked MPER–TMD. The unlocked MPER–TMD was prepared by reconstituting in bicelles a MPER–TMD mutant (L660C, A667C) under reducing condition (10 mM DTT). The production of the mutant MPER–TMD followed the same protocol used for the wild-type MPER–TMD[9]. Briefly, a fragment of HIV-1 gp41 (clade D, isolate 92UG024.2) containing residues 660–710 was expressed in *E. coli* strain BL21 (DE3) cells as a trpLE fusion and purified as in the "Protein expression and purification" section above. The locked MPER–TMD was subsequentially obtained upon complete removal of the DTT by dialysis and gradual addition of Glutathione (ox) to the final concentration of 10 mM. The complete locking of the MPER was confirmed by Urea-PAGE and mass spectrometry analyses (Supplementary Fig. 17a).

Plane restraints. The PPT method[23,25] was used to derive a set of plane restraints to aid structure calculation. Provided that the bicelle is sufficiently large ($q \geq 0.5$), the PPT method allows to determine the projection of each amide proton (r_Z) along the protein C_3 symmetry axis, which is also parallel to the bicelle normal and aligned to the Cartesian Z axis for convenience. Therefore, r_Z can be assigned as residue-specific plane restraint if the PRE$_{amp}$ values are calibrated against a known structure with known membrane partition. Since the TMD–CT^{LLP2} and the TMD show remarkable chemical shift and structure similarity for residues 677–710 (Supplementary Figs. 1c, 3a, and 8), this region was used to calibrate the PRE$_{amp}$ values of the TMD–CT^{LLP2} (from both Gd-DOTA and 16-DSA) against the previously established structure and membrane partition of the TMD[8,26] (Supplementary Fig. 5e). Specifically, we generated the PRE$_{amp}$ vs. r_Z plot for residues 679–710 using the residue-specific PRE$_{amp}$ from the TMD–CT^{LLP2} and the known r_Z from the known TMD structure. The data were then fitted to the sigmoidal function (Eq. (2)) to yield parameterized Eq. 2 for Gd-DOTA and 16-DSA, which were then used to determine r_Z for the other residues of the TMD–CT^{LLP2} not used for the calibration (711–788). Out of the calculated r_Z, only those in the sensitive region of the sigmoidal curves were retained. Finally, r_Z derived from Gd-DOTA and 16-DSA data sets were averaged and merged into one single data set, yielding the final set of plane restraints. The plane restraints used for the TMD–CT^{LLP2} structure calculation included r_Z from the published TMD structure (residues 679–-710) and the newly calculated r_Z for residues 711–788, with an uncertainty of ±1 Å.

Using the same procedure, plane restraints were also assigned for the MPER and included in the MPER–TMD–CT^{LLP2} model calculation. In this case, the previously published PRE$_{amp}$ values of the MPER–TMD (from Gd-DOTA titration)[9] were calibrated against residues 695–708 of the TMD (Supplementary Fig. 5g) and used to derive plane restraints for residues 660–694.

The data used to generate the plane restraints for the TMD–CT^{LLP2} and MPER–TMD–CT^{LLP2} (Supplementary Table 10) are summarized in Supplementary Fig. 5.

Structure calculation. NMR structures were calculated using *XPLOR-NIH*[58]. Since the TMD–CT^{LLP2} (677–788) is an extension of the previously studied TMD (677–716) and both constructs exhibit almost identical chemical shift for residues 677–710 (Supplementary Figs. 1c, 3a, and 8), we used the structure of the TMD[8] as a starting point for the new calculation. To implement this strategy, we first performed *TALOS+*[59] analysis using the assigned backbone chemical shift values (^{15}N, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C'$), and then used the "GOOD" dihedral angles generated by TALOS+ to build a secondary structural model of the TMD–CT^{LLP2}. Second, we assembled a trimer using the previously assigned NOE restraints for residues 677–702 (from PDB accession code 5JYN). Finally, we applied the newly assigned inter-chain restraints (most of them from residues 704–785) to complete the trimer of the TMD–CT^{LLP2}. For each inter-chain restraint between two adjacent protomers, three identical distance restraints were assigned respectively to all pairs of neighboring protomers to satisfy the condition of C_3 symmetry. The assembled trimer was then refined against the complete set of NOE restraints (intra- and inter-chain) and dihedral angles using a simulated annealing (SA) protocol in

which the temperature was lowered from 1000 to 200 K in steps of 50 K. The NOE restraints were enforced by flat-well harmonic potentials, with the force constant ramped from 2 to 40 kcal/mol Å^{-2} during annealing. Backbone dihedral angle restraints were enforced by flat-well (± the corresponding uncertainties from TALOS+) harmonic potential with force constant ramped from 10 to 1000 kcal/mol rad^{-2}. The plane restraints were fixed in space and enforced by flat-well (±1 Å) harmonic potentials, with force constant ramped from 1 to 5 kcal/mol Å^{-2} during annealing. A total of 150 structures were calculated, and the 15 lowest energy structures were selected as the final structural ensemble (PDB accession code 6UJU) (Supplementary Fig. 19b and Supplementary Table 11).

The model of the MPER–TMD–CT^{LLP2} was generated in a similar manner. The matching resonances of the TM core (residues 685–700) in the MPER–TMD (660–710), TMD (677–716), and TMD–CT^{LLP2} (677–788) (Supplementary Figs. 1c, 3a, and 8) allowed to merge the three structures at the TM core to generate a MPER–TMD–CT^{LLP2} starting model for further refinement. NOEs and backbone dihedral angle restraints for the MPER and TMD (residues 660–693 and 694–710, respectively) were taken from previously published studies[8,9], while those for the CT^{LLP2} (711–788) were taken from the current study. Plane restraints from the three regions were also applied. The model was refined using identical parameters and potentials as those used for the TMD–CT^{LLP2} (see above). A total of 150 structures were calculated, and the 15 lowest energy structures were selected as the final structural ensemble (PDB accession code 6UJV) (Supplementary Fig. 19c and Supplementary Table 12).

Env mutant constructs and monoclonal antibodies. Full-length Env mutants were generated using the 92UG037.8 gp160[5] 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 previously described[5,6]. Briefly, the intact antibodies or Fab fragments were expressed in 293T cells either by transient transfection or using selected stably transfected clones, or from hybridomas or CHO cells. The antibodies were purified by affinity chromatography using Gamma bind plus resin (GE Healthcare, Pittsburgh, PA), followed by gel-filtration chromatography. Expression constructs of antibodies PG9, PG16, and PGT145 were generated using synthetic genes made by GeneArt Gene Synthesis (Life Technologies, Carlsbad, CA) or GenScript (Supplementary Table 13). The VRC01 expression constructs were kindly provided 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).

Production of pseudoviruses containing mutant Envs. Preparation of HIV-1 Env pseudoviruses of CT mutants, and titration of pseudovirus stocks to determine the 50% tissue culture infectious dose per mL ($TCID_{50}$/mL) were performed as previously described[8,60]. Briefly, each pseudovirus was prepared by transfecting 293T/17 cells (ATCC, Manassas, VA) with 4 µg of Env expression plasmid and 8 µg of an Env-deficient HIV-1 backbone vector (pSG3ΔEnv). Pseudovirus-containing culture supernatant was harvested 24 h after transfection, filtered (0.45 µm), and stored at −80 °C. To determine $TCID_{50}$/mL, a 5-fold serial dilution of virus stock was performed in quadruplicate wells and incubated with TZM.bl cells (NIH AIDS reagent program) in growth media containing DEAE-dextran (Sigma-Aldrich) at a final concentration of 11 µg/mL. After 48 h, the cells were measured for luciferase reporter gene expression, indicating the ability of the pseudovirus to infect cells. $TCID_{50}$/mL was calculated using an Excel macro made available on the Las Alamos National Laboratories website (www.hiv.lanl.gov).

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)[8]. 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 2.0 × 10^6 cells/mL, followed by treatment with the Sample Buffer and boiling for 10 min. Western blot was performed following our published protocol[8]. Briefly, Env samples were resolved in 4–15% Mini-Protean TGX gel (Bio-Rad) and transferred onto PVDF membranes (Millipore, Billerica, MA) by an Iblot2 (Life Technologies). Membranes were blocked with 5% skimmed milk in PBS for 1 h and incubated with anti-V3 loop antibody 3791 for another hour at room temperature. Alkaline phosphatase conjugated anti-human Fab IgG (1:5000) (Sigma-Aldrich, LOT number: A8542) was used as a secondary antibody. Env proteins were visualized using one-step NBT/BCIP substrates (Thermo Scientific, Cambridge, MA).

Flow cytometry. 293T cells were transiently transfected with 2 µg of the 92UG037.8 gp160 expression construct or its CT mutants in six-well plates. Flow cytometry was carried out as previously described[5,8]. Briefly, Env-expressing cells were detached from plates using PBS and washed with ice-cold PBS containing 1% BSA. Hundred and six cells were incubated for 30–40 min on ice with either VRC01 Fab, 2G12 Fab, or PG16 IgG at concentrations of 10 µg/ml in PBS containing 1% BSA. The cells were then washed twice with PBS containing 1% BSA and stained with R-Phycoerythrin AffiniPure F(ab′)2 fragment goat anti-human IgG, F(ab′)2 Fragment specific secondary antibody (Jackson ImmunoResearch Laboratories, West Grove, PA) at 5 µg/ml. All the fluorescently labeled cells were washed twice with PBS containing 1% BSA and analyzed immediately using a BD FACScanto II instrument and program FACSDIVA (BD Biosciences, San Jose, CA). All data were analyzed by *FlowJo* (FlowJo, LLC, Ashland, OR). The flow cytometry gating strategy is shown in Supplementary Fig. 20.

Cell–cell fusion assay. Cell–cell fusion assay was based on the α-complementation of *E. coli* β-galactosidase[8]. 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 h. 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 CT mutants was measured by infecting TZM.bl cells with p24-normalized pseudovirus in growth media containing DEAE-dextran (11 µg/mL). Forty eight hours post-infection, luciferase activity of the reporter gene was quantified[31]. Likewise, neutralizing IC50 and IC80 titers of monoclonal antibodies were determined also by the luciferase-based virus neutralization assay in TZM.bl cells, which measures the reduction in luciferase reporter gene expression in TZM-bl cells following a single round of virus infection (Supplementary Table 2)[31]. 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 h at 37 °C. TZM.bl cells were then added (1.0 × 10^4/well in 100 µL volume) in 10% DMEM growth medium containing DEAE-Dextran. Following a 48 h incubation, luminescence was measured using Bright-Glo luciferase reagent (Promega, Madison, WI). The IC50 and IC80 titers were calculated as the antibody dilution that resulted in a 50% or 80% reduction in relative luminescence units (RLU), respectively, compared with the virus control wells after the subtraction of cell control RLU. Maximum percent inhibition (MPI) indicates the highest percent inhibition of virus infection observed with the tested concentrations of each antibody. 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, VRC01, 3BNC117, PGT121, 10-1074, PGDM1400, PGT151, 8ANC194, and 4E10. The tier phenotyping assay used the following HIV+ chronic serum samples: HIV-018, HIV-019, HIV-021, HIV-023, HIV-024, HIV-025, and HIV-026, which are all chronic clade B HIV+ serum. The assay was described previously[42] and contained in the same protocol described above. The IC50 and IC80 titers were calculated as the serum dilution that caused a 50 and 80% reduction in relative RLU compared with the level in the virus control wells after subtraction of cell control RLU.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The source data underlying Figs. 2–5 and Supplementary Figs. 1, 2, 5, 6, 8, 9, 11, 13-17 are provided as a Source Data file. The atomic structure coordinate and structural constraints have been deposited in the Protein Data Bank (PDB), accession numbers 6UJU and 6UJV. The chemical shift values have been deposited in the Biological Magnetic Resonance Data Bank (BMRB), accession number 30678. Other data that support the findings of this study are available from the corresponding authors upon reasonable request.

Received: 26 November 2019; Accepted: 16 April 2020;
Published online: 08 May 2020

References

1. Harrison, S. C. Mechanism of membrane fusion by viral envelope proteins. *Adv. Virus Res.* **64**, 231–259 (2005).
2. Chen B. Molecular mechanism of HIV-1 entry. *Trends Microbiol.* **27**, 878–891 (2019).
3. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263–273 (1997).
4. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**, 426–430 (1997).
5. Chen, J. et al. HIV-1 ENVELOPE. Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. *Science* **349**, 191–195 (2015).

6. Kovacs, J. M. et al. HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120. *Proc. Natl Acad. Sci. USA* **109**, 12111–12116 (2012).
7. Sanders, R. W. et al. A next-generation cleaved, soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog.* **9**, e1003618 (2013).
8. Dev, J. et al. Structural basis for membrane anchoring of HIV-1 envelope spike. *Science* **353**, 172–175 (2016).
9. Fu, Q. et al. Structure of the membrane proximal external region of HIV-1 envelope glycoprotein. *Proc. Natl Acad. Sci. USA* **115**, E8892–E8899 (2018).
10. Liu, J., Bartesaghi, A., Borgnia, M. J., Sapiro, G. & Subramaniam, S. Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455**, 109–113 (2008).
11. Julien, J. P. et al. Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1477–1483 (2013).
12. Lyumkis, D. et al. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1484–1490 (2013).
13. Pancera, M. et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514**, 455–461 (2014).
14. Lee, J. H., Ozorowski, G. & Ward, A. B. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science* **351**, 1043–1048 (2016).
15. Pan, J., Peng, H., Chen, B. & Harrison, S. C. Cryo-EM structure of full-length HIV-1 Env bound with the fab of antibody PG16. *J. Mol. Biol.* **14**, 1158–1168 (2020).
16. Torrents de la Pena, A. et al. Similarities and differences between native HIV-1 envelope glycoprotein trimers and stabilized soluble trimer mimetics. *PLoS Pathog.* **15**, e1007920 (2019).
17. Postler, T. S. & Desrosiers, R. C. The tale of the long tail: the cytoplasmic domain of HIV-1 gp41. *J. Virol.* **87**, 2–15 (2013).
18. Santos da Silva, E., Mulinge, M. & Perez Bercoff, D. The frantic play of the concealed HIV envelope cytoplasmic tail. *Retrovirology* **10**, 54 (2013).
19. Murphy, R. E., Samal, A. B., Vlach, J. & Saad, J. S. Solution Structure and Membrane Interaction of the Cytoplasmic Tail of HIV-1 gp41 Protein. *Structure* **25**, 1708–1718 e1705 (2017).
20. Chou, J. J., Kaufman, J. D., Stahl, S. J., Wingfield, P. T. & Bax, A. Micelle-induced curvature in a water-insoluble HIV-1 Env peptide revealed by NMR dipolar coupling measurement in stretched polyacrylamide gel. *J. Am. Chem. Soc.* **124**, 2450–2451 (2002).
21. Edwards, T. G. et al. Truncation of the cytoplasmic domain induces exposure of conserved regions in the ectodomain of human immunodeficiency virus type 1 envelope protein. *J. Virol.* **76**, 2683–2691 (2002).
22. Durham, N. D. et al. Neutralization resistance of virological synapse-mediated HIV-1 Infection is regulated by the gp41 cytoplasmic tail. *J. Virol.* **86**, 7484–7495 (2012).
23. Fu Q., Piai A., Chen W., Xia K. & Chou J. J. Structure determination protocol for transmembrane domain oligomers. *Nat. Protoc.* **14**, 2483–2520 (2019).
24. Xu, C. et al. Regulation of T cell receptor activation by dynamic membrane binding of the CD3ε cytoplasmic tyrosine-based motif. *Cell* **135**, 702–713 (2008).
25. Piai, A., Fu, Q., Dev, J. & Chou, J. J. Optimal bicelle size q for solution NMR studies of the protein transmembrane partition. *Chemistry* **23**, 1361–1367 (2017).
26. Piai, A., Dev, J., Fu, Q. & Chou, J. J. Stability and water accessibility of the trimeric membrane anchors of the HIV-1 envelope spikes. *J. Am. Chem. Soc.* **139**, 18432–18435 (2017).
27. Chen, W., Dev, J., Mezhyrova, J., Pan, L., Piai, A. & Chou, J. J. The unusual transmembrane partition of the hexameric channel of the hepatitis C virus. *Structure* **26**, 627–634 e624 (2018).
28. Julien, J. P. et al. Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. *Proc. Natl Acad. Sci. USA* **110**, 4351–4356 (2013).
29. Lee, J. H. et al. A broadly neutralizing antibody targets the dynamic HIV envelope trimer apex via a long, rigidified, and anionic beta-hairpin structure. *Immunity* **46**, 690–702 (2017).
30. Pancera, M. et al. Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nat. Struct. Mol. Biol.* **20**, 804 (2013).
31. Sarzotti-Kelsoe, M. et al. Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J. Immunol. Methods* **409**, 131–146 (2014).
32. Scheid, J. F. et al. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633–1637 (2011).
33. Walker, L. M. et al. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**, 285–289 (2009).
34. Walker, L. M. et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466–470 (2011).
35. Swetnam, J., Shmelkov, E., Zolla-Pazner, S. & Cardozo, T. Comparative magnitude of cross-strain conservation of HIV variable loop neutralization epitopes. *PLoS ONE* **5**, e15994 (2010).
36. Thali, M. et al. Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. *J. Virol.* **67**, 3978–3988 (1993).
37. Han, Q. et al. Difficult-to-neutralize global HIV-1 isolates are neutralized by antibodies targeting open envelope conformations. *Nat. Commun.* **10**, 2898 (2019).
38. Zhu, P. et al. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* **441**, 847–852 (2006).
39. Postler, T. S. & Desrosiers, R. C. The tale of the long tail: the cytoplasmic domain of HIV-1 gp41. *J. Virol.* **87**, 2–15 (2013).
40. Wang, H., Cohen, A. A., Galimidi, R. P., Gristick, H. B., Jensen, G. J. & Bjorkman, P. J. Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proc. Natl Acad. Sci. USA* **113**, E7151–E7158 (2016).
41. Ozorowski, G. et al. Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. *Nature* **547**, 360–363 (2017).
42. Seaman, M. S. et al. Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J. Virol.* **84**, 1439–1452 (2010).
43. Kneller, J. M., Lu, M. & Bracken, C. An effective method for the discrimination of motional anisotropy and chemical exchange. *J. Am. Chem. Soc.* **124**, 1852–1853 (2002).
44. Evans, D. T., Tillman, K. C. & Desrosiers, R. C. Envelope glycoprotein cytoplasmic domains from diverse lentiviruses interact with the prenylated Rab acceptor. *J. Virol.* **76**, 327–337 (2002).
45. Wyma, D. J. et al. Coupling of human immunodeficiency virus type 1 fusion to virion maturation: a novel role of the gp41 cytoplasmic tail. *J. Virol.* **78**, 3429–3435 (2004).
46. Hill, C. P., Worthylake, D., Bancroft, D. P., Christensen, A. M. & Sundquist, W. I. Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. *Proc. Natl Acad. Sci. USA* **93**, 3099–3104 (1996).
47. Alfadhli, A., Barklis, R. L. & Barklis, E. HIV-1 matrix organizes as a hexamer of trimers on membranes containing phosphatidylinositol-(4,5)-bisphosphate. *Virology* **387**, 466–472 (2009).
48. Kovacs, J. M. et al. HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120. *Proc. Natl Acad. Sci. USA* **109**, 12111–12116 (2012).
49. Sanders, R. W. et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog.* **9**, e1003618 (2013).
50. Ma, X. et al. HIV-1 Env trimer opens through an asymmetric intermediate in which individual protomers adopt distinct conformations. *eLife* **7**, e34271 (2018).
51. Lu, M. et al. Associating HIV-1 envelope glycoprotein structures with states on the virus observed by smFRET. *Nature* **568**, 415–419 (2019).
52. Stadtmueller, B. M. et al. DEER spectroscopy measurements reveal multiple conformations of HIV-1 SOSIP envelopes that show similarities with envelopes on native virions. *Immunity* **49**, 235–246 e234 (2018).
53. 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* **6**, 277–293 (1995).
54. Bartels, C., Xia, T. H., Billeter, M., Guntert, P. & Wuthrich, K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **6**, 1–10 (1995).
55. Salzmann, M., Wider, G., Pervushin, K. & Wuthrich, K. Improved sensitivity and coherence selection for [15N,1H]-TROSY elements in triple resonance experiments. *J. Biomol. NMR* **15**, 181–184 (1999).
56. Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. Three-dimensional triple resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson* **89**, 496–514 (1990).
57. 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* **31**, 5269–5278 (1992).
58. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
59. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. biomolecular NMR* **44**, 213–223 (2009).
60. Li, M. et al. 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**, 10108–10125 (2005).
61. Gristick, H. B. et al. Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. *Nat. Struct. Mol. Biol.* **23**, 906–915 (2016).

Acknowledgements

We thank Stephen Harrison for insightful discussion and critical reading of the manuscript. Alexander Daniels and Catharina Saenger kindly helped with protein construct design and testing. Zhijun Liu helped with NMR data collection at NCPSS. This work was supported by NIH grants AI127193 (to B.C. and J.J.C.), GM116898 (to J.J.C.), AI129721 (to B.C.), a Merck Fellowship (to Q.F.), and Collaboration for AIDS Vaccine Discovery (CAVD) grant OPP1169339 (to Dan H. Barouch from the Bill and Melinda Gates Foundation). The NMR data were collected at the MIT-Harvard CMR (supported by NIH grant P41 EB-002026).

Author contributions

A.P., B.C., and J.J.C. conceived the study; A.P. designed protein constructs for structural studies; Q.F. designed protein constructs for NMR dynamics studies; A.P. and Q.F. prepared samples for NMR studies; A.P., Y.C., B.C., and J.J.C. designed mutants for functional studies; F.G., Y.C., T.X., M.M.S., H.P., S.R.-V., and M.S.S. performed functional studies; W.C. and A.P. performed OG-label analysis; A.P. and J.J.C. solved the NMR structure; A.P., B.C., and J.J.C. wrote the paper and all authors contributed to editing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16165-0.

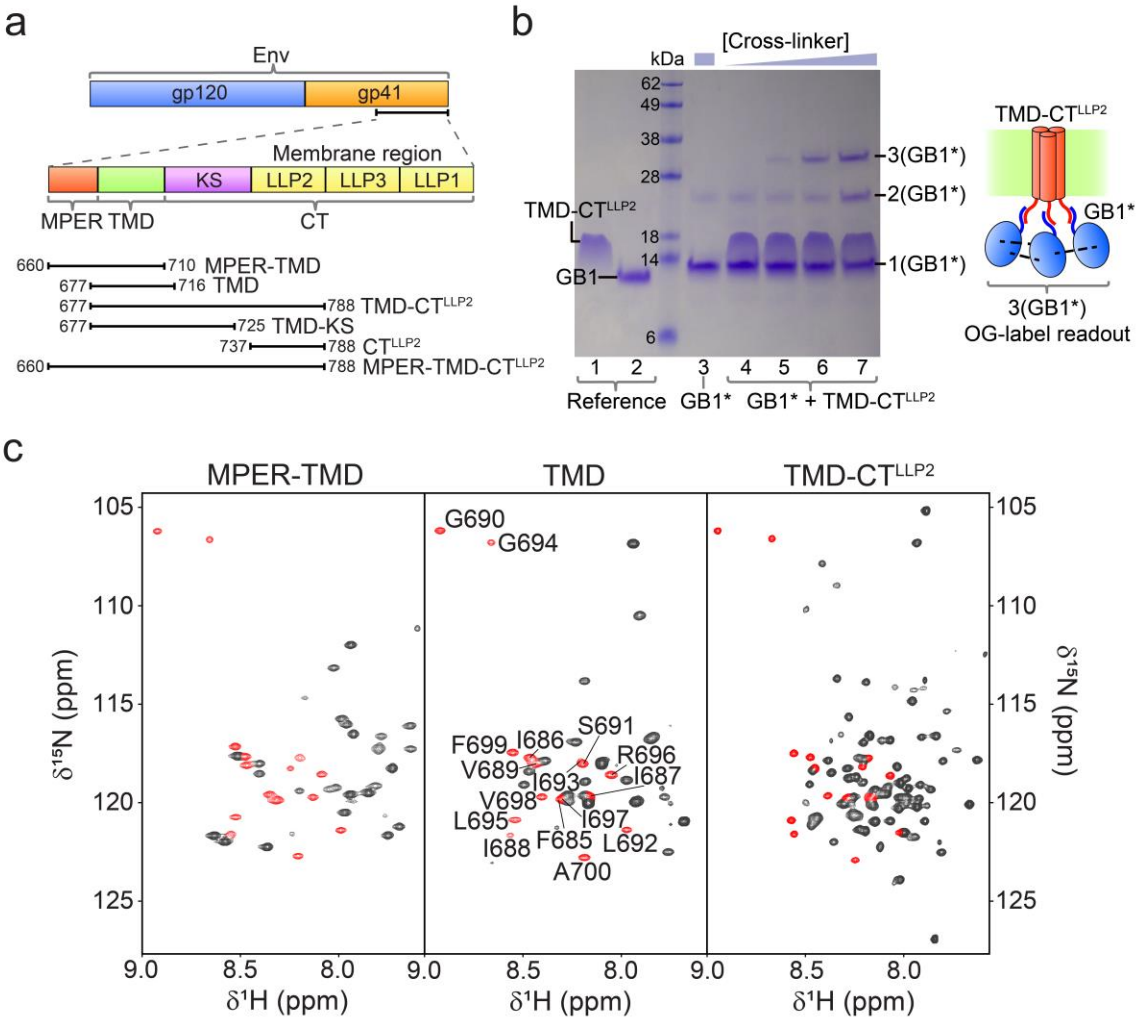
Correspondence and requests for materials should be addressed to B.C. or J.J.C.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

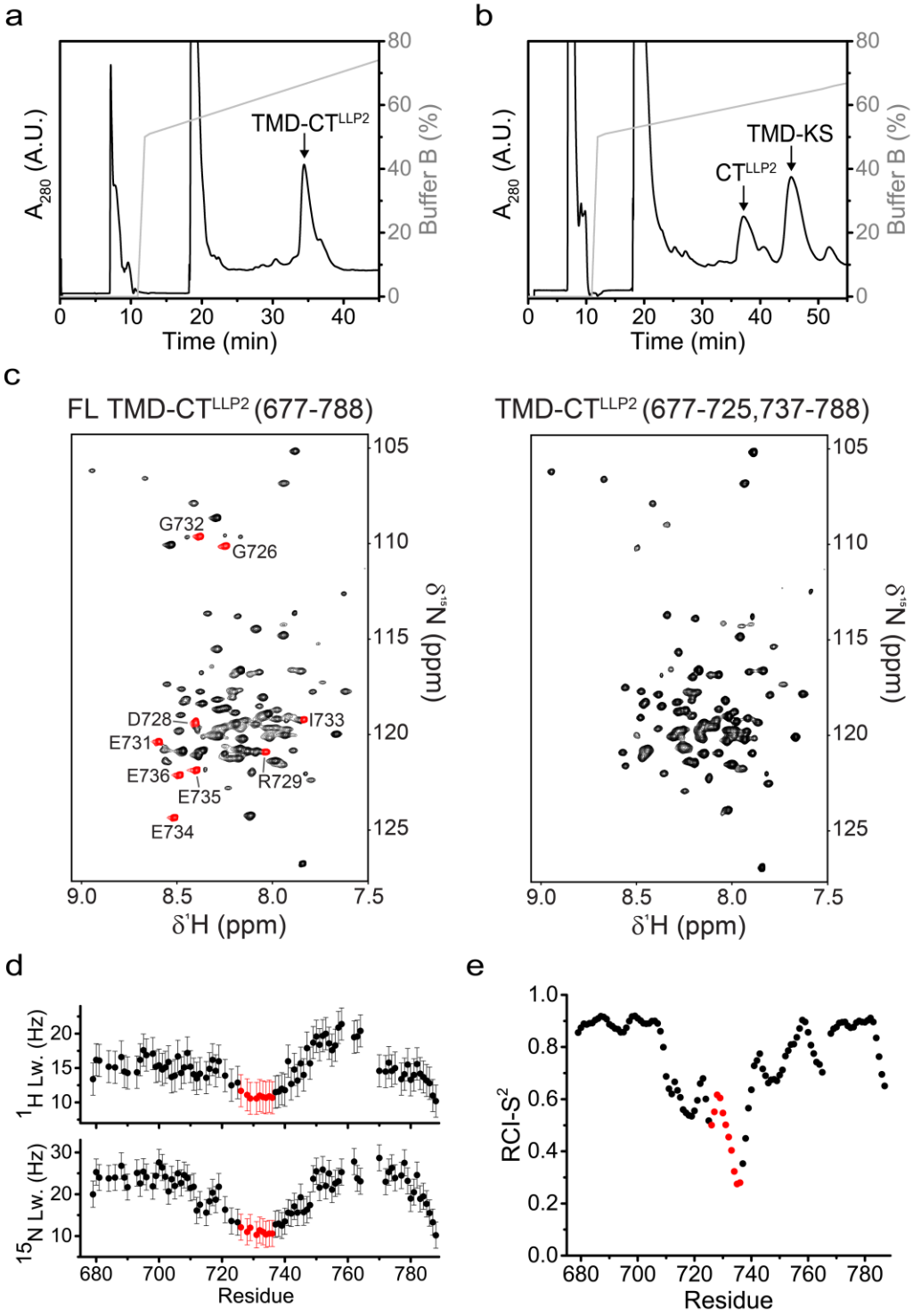
Supplementary Information

Structural basis of transmembrane coupling of the HIV-1 envelope glycoprotein

Alessandro Piai, Qingshan Fu, Yongfei Cai et al.

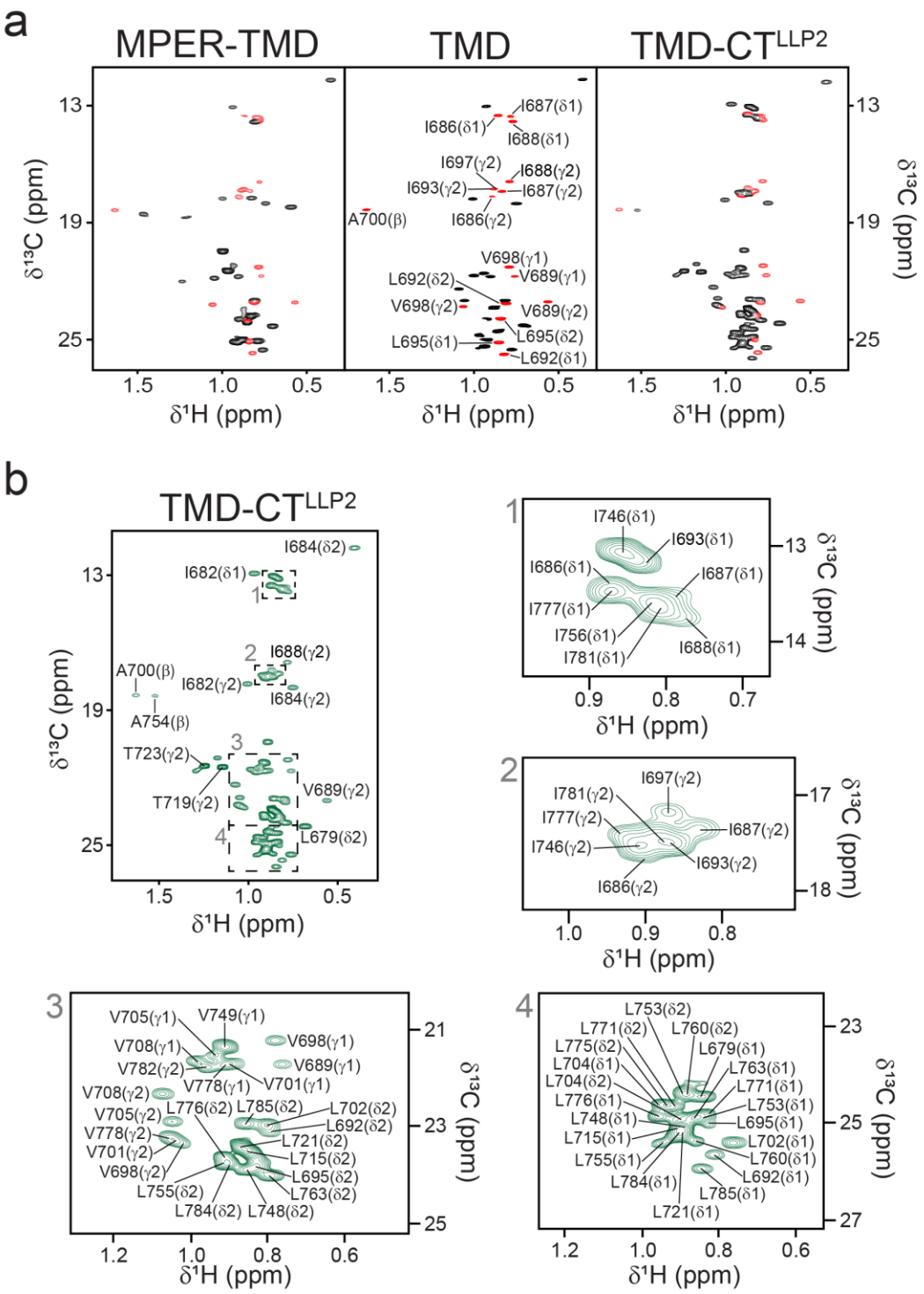


Supplementary Fig. 1 | Biochemical and NMR characterization of TMD-CT^{LLP2} in bicelles. **a,** Designation of the membrane-related components of HIV-1 gp41 (top) and definition of the constructs used in this study (bottom). **b,** Determination of the oligomeric state of TMD-CT^{LLP2} reconstituted in DMPC-DHPC bicelles (q = 0.5) using the OG-label method[1] that uses the GB1 protein as reporter of TMD-CT^{LLP2} oligomerization. The SDS-PAGE gel lanes are: 1 – TMD-CT^{LLP2} (monomeric in SDS); 2 – GB1 (n.b. GB1* denotes GB1 conjugated to a triNTA molecule); 3 – GB1* cross-linked alone; 4-7 – GB1* cross-linked in the presence of TMD-CT^{LLP2} at increasing cross-linker concentration. The experiment was performed once. **c,** Comparison of the 2D ^{1}H-^{15}N TROSY-HSQC spectra of the MPER-TMD, TMD and TMD-CT^{LLP2}, showing strong agreement among the three constructs of the chemical shift of residues 685-700 (red). Source data are provided as a Source Data file.

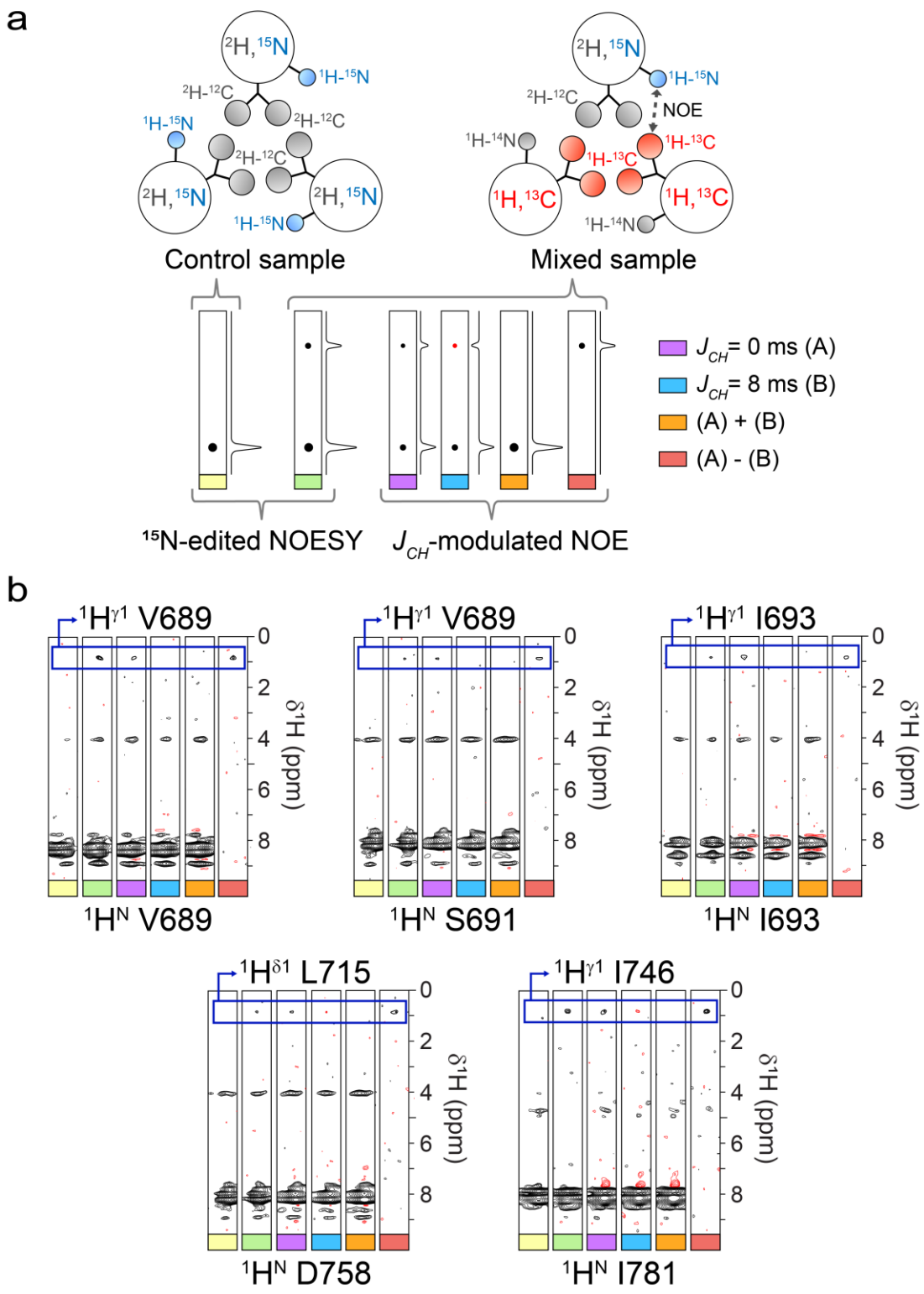


Supplementary Fig. 2 | TMD-CT^{LLP2}, TMD-KS and CT^{LLP2} purification and deletion of residues 726-736. a, RP-HPLC chromatogram of the TMD-CT^{LLP2}. The protein absorbance was measured at 280 and 214 nm (the latter not shown for clarity). The gray line indicates the gradient of the mobile phase composition (% of Buffer B). **b,** Same as (a) but for the TMD-KS and CT^{LLP2}.

The constructs were obtained by introducing in the TMD-CT^{LLP2} an additional Met between the KS and the CT^{LLP}, so that the two fragments were cleaved by CNBr together with the TrpLE. **c,** 2D ^{1}H-^{15}N TROSY-HSQC spectra of the full-length (FL) TMD-CT^{LLP2} (left) and the TMD-CT^{LLP2} with residues 736-736 deleted (right). The cross-peaks of residues 726-736 are shown in red. The two spectra are virtually identical indicating that the deletion did not alter the structure of the protein. **d,** Residue-specific backbone $^{1}H^{N}$ (top) and ^{15}N (bottom) NMR peak linewidth of the FL TMD-CT^{LLP2} obtained from the 2D ^{1}H-^{15}N TROSY-HSQC spectrum in panel (c). Error bars represent the uncertainty derived from cross-peaks signal to noise. The linewidths of residues 726-736, highlighted in red, are significantly smaller than those of the rest of the protein, indicating that the deleted region is highly flexible. **e,** Random coil index (RCI)-based order parameter (S^2) calculated from experimental chemical shifts (^{15}N, $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and $^{13}C'$) using the *MICS* program[2]. The low values computed for residues 726-736 (red dots) indicates the absence of secondary structure in the deleted region. Source data are provided as a Source Data file.

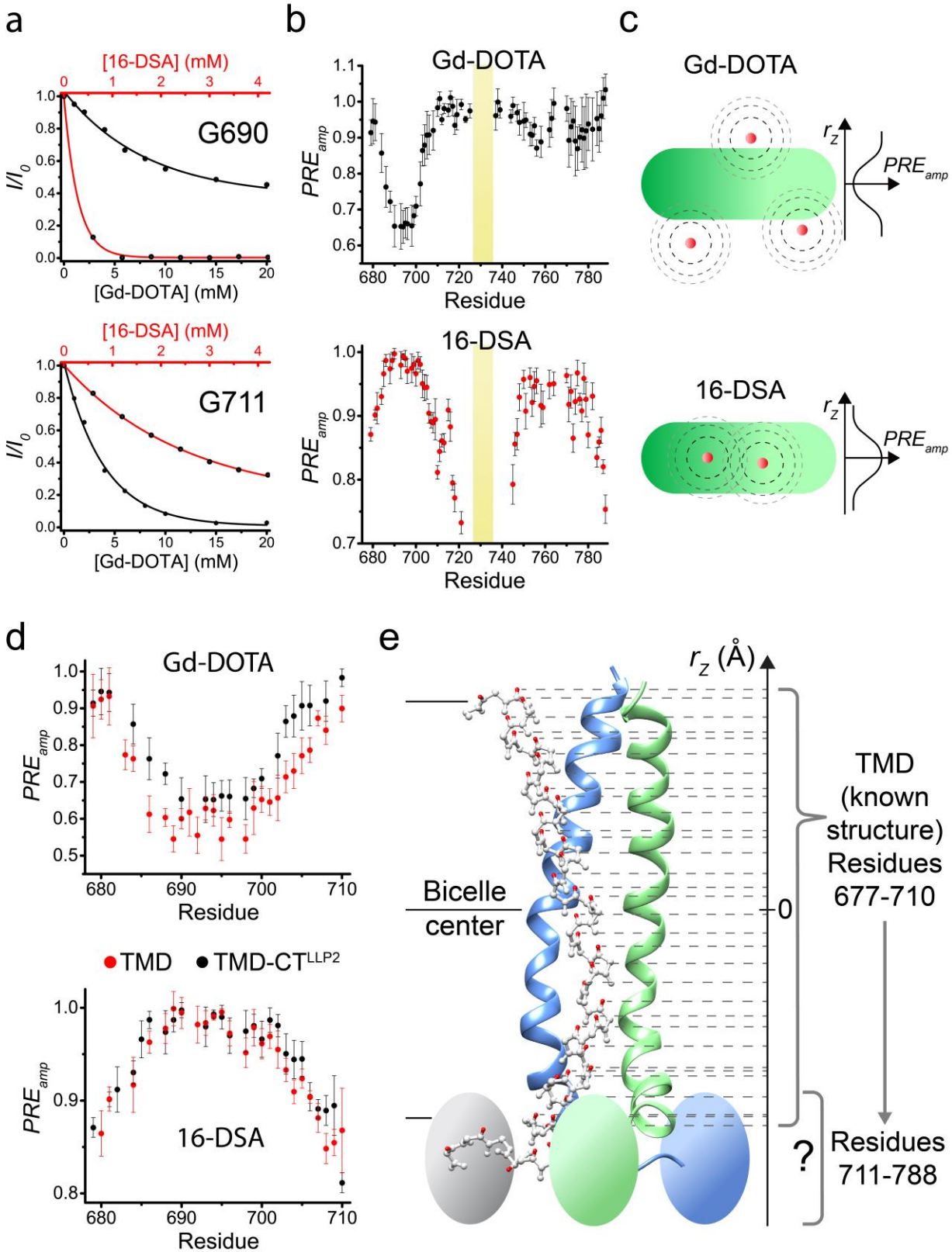


Supplementary Fig. 3 | Methyl group resonances of the MPER-TMD, TMD and TMD-CT^{LLP2} in bicelles. a, Comparison of the methyl group regions of the 2D ^{1}H-^{13}C HSQC spectra of the MPER-TMD, TMD and TMD-CT^{LLP2}, showing strong agreement among the three constructs of the chemical shift of residues 685-700 (red). **b,** Assigned methyl group resonances of the TMD-CT^{LLP2}. The assignment of the most crowded regions is shown in the four expansions.

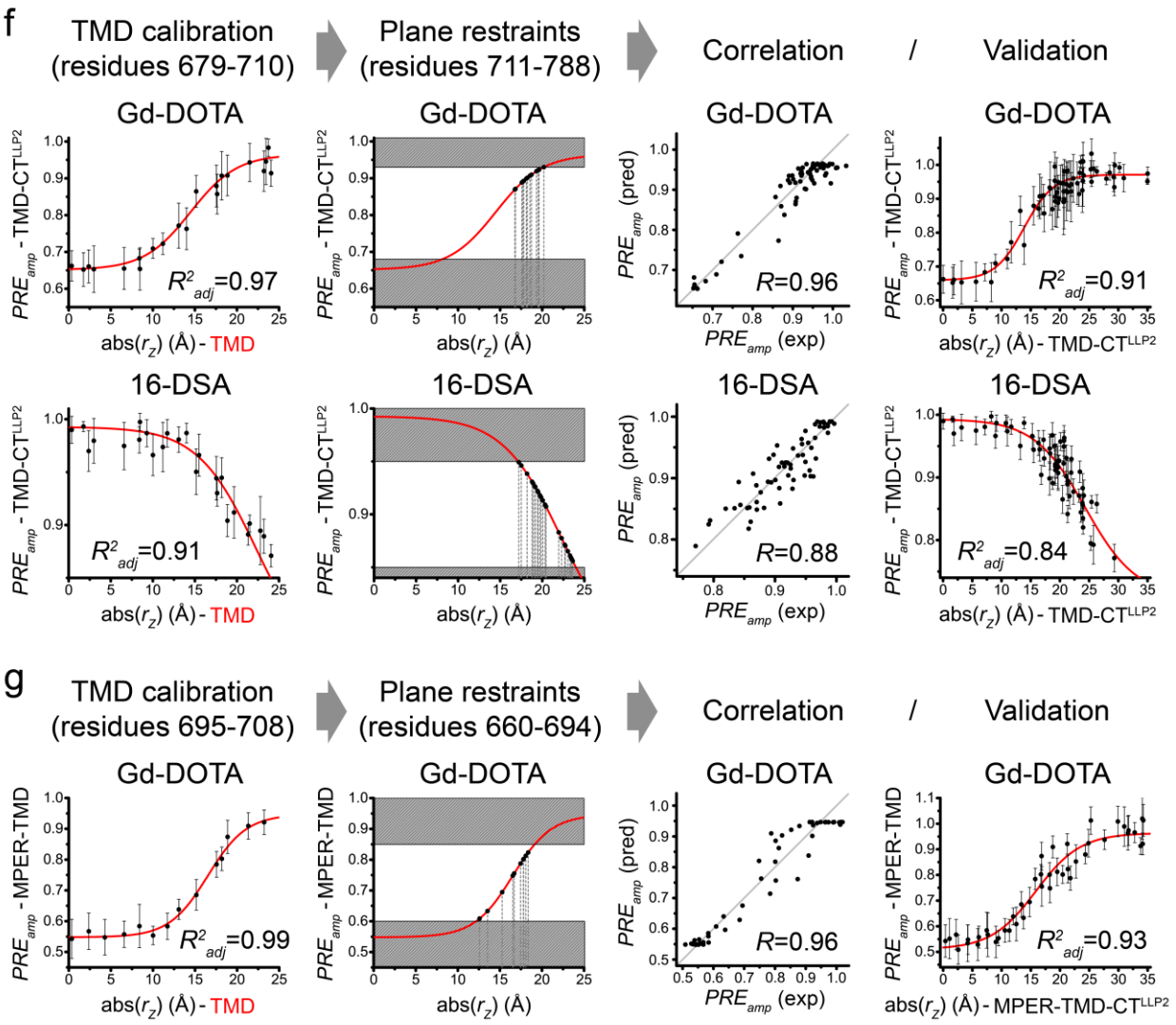


Supplementary Fig. 4 | Identification of unambiguous inter-chain NOEs of the TMD-CT^{LLP2}. **a,** Schematic representation of the samples used to record the NOESY experiments. The control sample (left) contained only ^{15}N, ^{2}H-labeled TMD-CT^{LLP2}, and the mixed sample (right) contained ~1:1 ratio of ^{15}N, ^{2}H-labeled TMD-CT^{LLP2} to ^{13}C-labeled TMD-CT^{LLP2}. ^{1}H-^{15}N and ^{1}H-^{13}C groups

are shown in blue and red, respectively, whereas the undetectable groups (e.g., ^{1}H-^{14}N and ^{2}H-^{12}C) are shown in gray. The mixed sample was used to detect inter-chain NOEs from ^{1}H-^{15}N to ^{1}H-^{13}C. In the bottom, the expected diagonal and cross-peaks in each experiment are schematically illustrated, with positive and negative peaks in black and red, respectively. **b,** NOESY strips showing examples of the unambiguous inter-chain NOEs of the TMD-CT^{LLP2}, identified using the above strategy. The control sample was used to acquire a 3D ^{15}N-edited NOESY-TROSY-HSQC spectrum, while the mixed sample was used to record an identical ^{15}N-edited NOESY-TROSY-HSQC spectrum as well as the 3D J_{CH}-modulated NOE spectrum. A mixing time of 200 ms was employed in all the experiments. Comparison of the strips from the ^{15}N-edited NOESY-TROSY-HSQC spectra of the control and mixed samples allowed the identification of inter-chain NOEs. The identified NOEs were then unambiguously validated by recording the J_{CH}-modulated NOE spectrum, in which two interleaved spectra were acquired with varying the J_{CH} evolution (J_{CH} = 0 ms and J_{CH} = 8 ms) before the NOE mixing. The subtraction of the two spectra allowed selection of only the inter-protomer NOEs. Positive and negative peaks are shown in black and red, respectively. The inter-protomer NOEs shown are: $^{1}H^{N}$ V689 to $^{1}H^{\gamma 1}$ V689; $^{1}H^{N}$ S691 to $^{1}H^{\gamma 1}$ V689; $^{1}H^{N}$ I693 to $^{1}H^{\gamma 1}$ I693; $^{1}H^{N}$ D758 to $^{1}H^{\delta 1}$ L715; $^{1}H^{N}$ I781 to $^{1}H^{\gamma 1}$ I746.

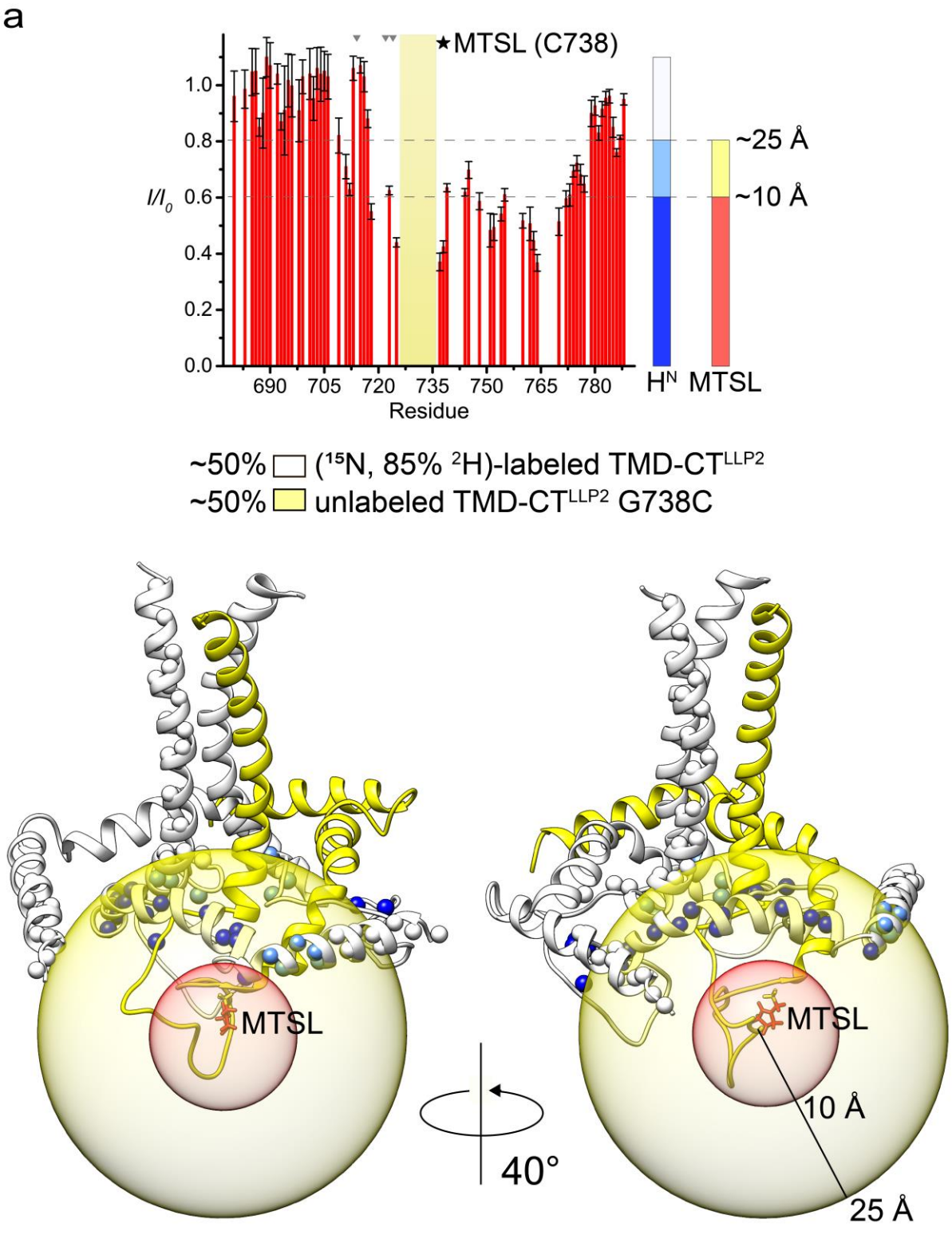


Supplementary Fig. 5 (continue in the next page)

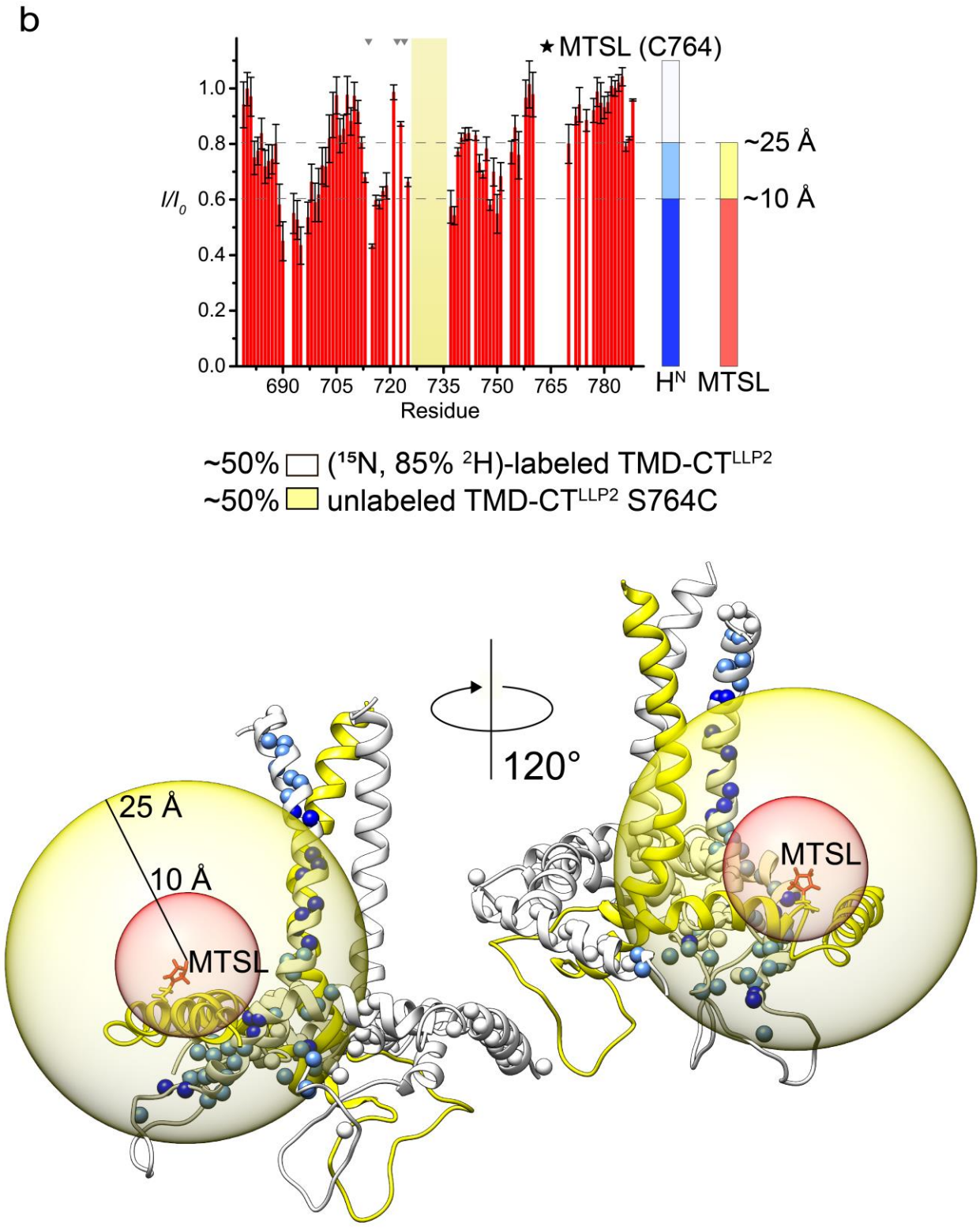


Supplementary Fig. 5 | Derivation of the plane restraints used in the TMD-CT^{LLP2} and MPER-TMD-CT^{LLP2} structure calculation. a, The I/I_0 vs. [Gd-DOTA] (black) and I/I_0 vs. [16-DSA] (red) plots for residues G690 (buried) and G711 (exposed) of the TMD-CT^{LLP2} are shown as examples. Data fitting (black and red lines for Gd-DOTA and 16-DSA, respectively) to the exponential decay function (Eq. 1) yielded the PRE_{amp} of the two residues. Variation in signal decays reflects the different localization of the two residues. **b,** PRE_{amp} vs (residue number) plots for TMD-CT^{LLP2} obtained from Gd-DOTA (black, top) and 16-DSA (red, bottom) titrations. Error bars represent the uncertainty derived from fitting error. Missing values are due to prolines or residues with overlapping peaks. The yellow box marks the deleted residues 726-736. **c,** Schematic illustration of the distribution of Gd-DOTA (top) and 16-DSA (bottom) relative to the bicelle, with the expected reciprocal PRE_{amp} profiles shown on the right. **d,** Comparison of the PRE_{amp} vs (residue number) plots of residues 679-710 between the TMD-CT^{LLP2} (black) and the TMD (red), derived from Gd-DOTA (top) and 16-DSA (bottom) titrations. Error bars represent the uncertainty derived from fitting error. The two data sets collectively indicate that residues 679-710 have the same membrane partition in both constructs. **e,** Schematic illustration of the strategy used to assign plane restraints for residues 711-788 of the TMD-CT^{LLP2}. The close agreement in the TM region

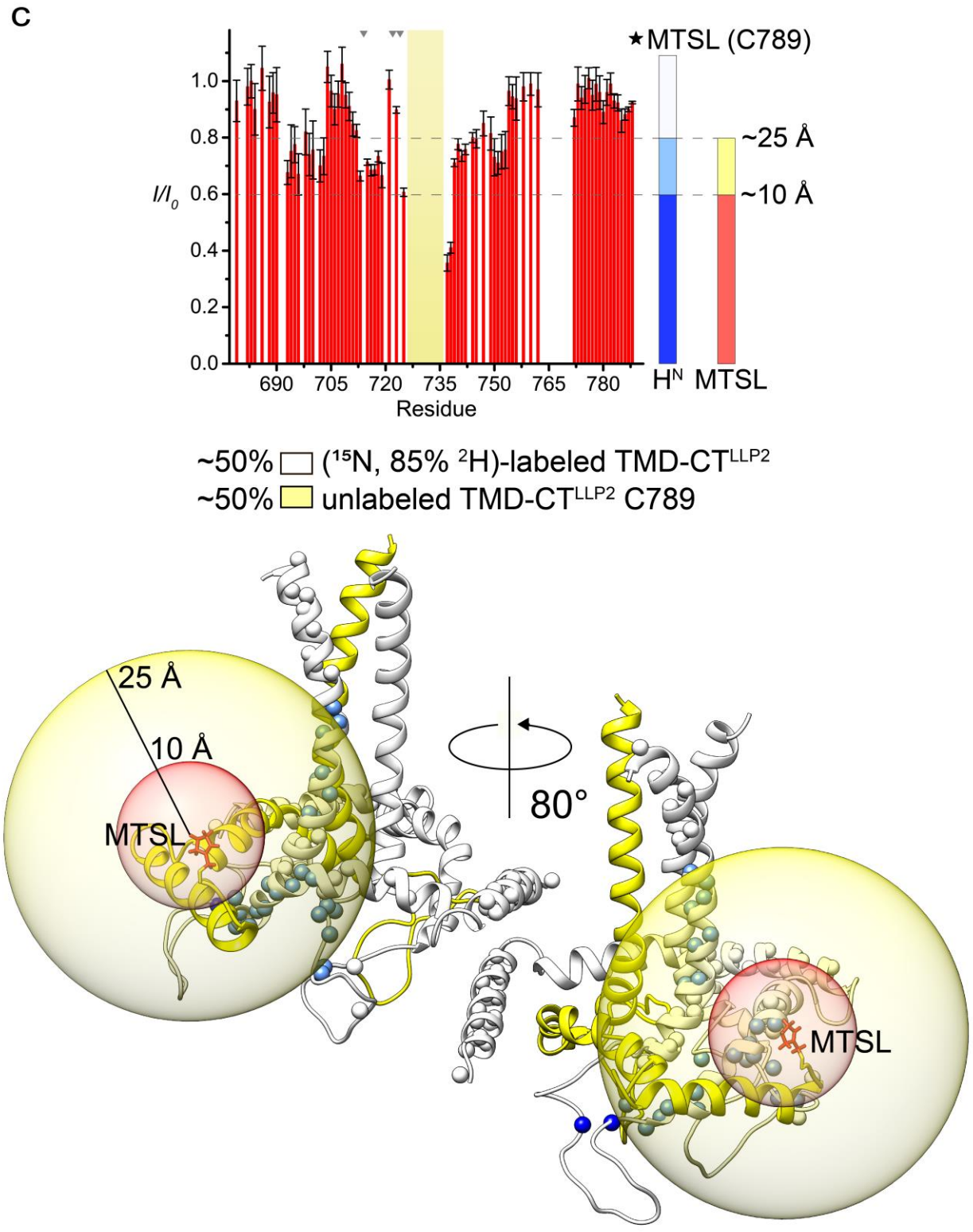
between the TMD-CT^{LLP2} and TMD constructs allowed to calibrate the PRE_{amp} values of the TMD-CT^{LLP2} against the previously established structure and membrane partition of the TMD[3,4] such that the PRE_{amp} values can directly translate to plane restraints for the CT region of the TMD-CT^{LLP2}. **f,** Assignment of plane restraints for residues 711-788 of the TMD-CT^{LLP2} using data from Gd-DOTA (top) and 16-DSA (bottom) titrations. First, the PRE_{amp} vs. *residue number* of the TMD-CT^{LLP2} was converted to PRE_{amp} vs. r_Z using the known TMD trimer structure (PDB ID: 5JYN). The PRE_{amp} vs. r_Z profiles were then fitted to the sigmoidal function (Eq. 2) (left). Second, the best-fit curves (red) were used to assign r_Z for residues 711-788 (middle). Only r_Z in the PRE-sensitive regions of the curves were retained (white area). Finally, the calculated r_Z from Gd-DOTA and 16-DSA curves were averaged, merged into one single data set and used as plane restraints for the TMD-CT^{LLP2} structure calculation. Once the structure was obtained, the correlation between experimental PRE_{amp} and those predicted from the new structure was determined. For both Gd-DOTA and 16-DSA data sets, the agreement is good as high Pearson linear correlation coefficient (R) were obtained (note: R is smaller for 16-DSA than for Gd-DOTA because the magnitude of PRE_{amp} variation is smaller for the former (~0.8 - 1.0) than for the latter (~0.6 - 1.0)). In addition, the membrane partition of the TMD-CT^{LLP2} trimer was derived by fitting the PRE_{amp} vs. r_Z profile (generated this time using the new atomic coordinates) to Eq. 2 and found to be virtually identical to that of the TMD, validating the use of the latter as internal calibration. In all plots, error bars represent the uncertainty derived from fitting error. **g,** Same as (f) but for the MPER-TMD-CT^{LLP2}. In this case, only the Gd-DOTA data were used. Source data are provided as a Source Data file.



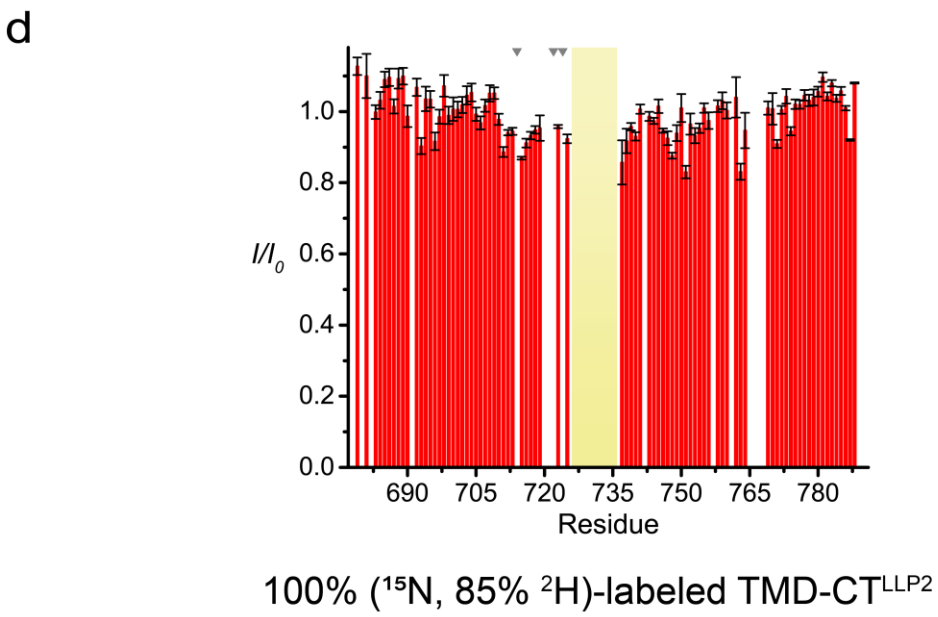
Supplementary Fig. 6 (continue in the next page)



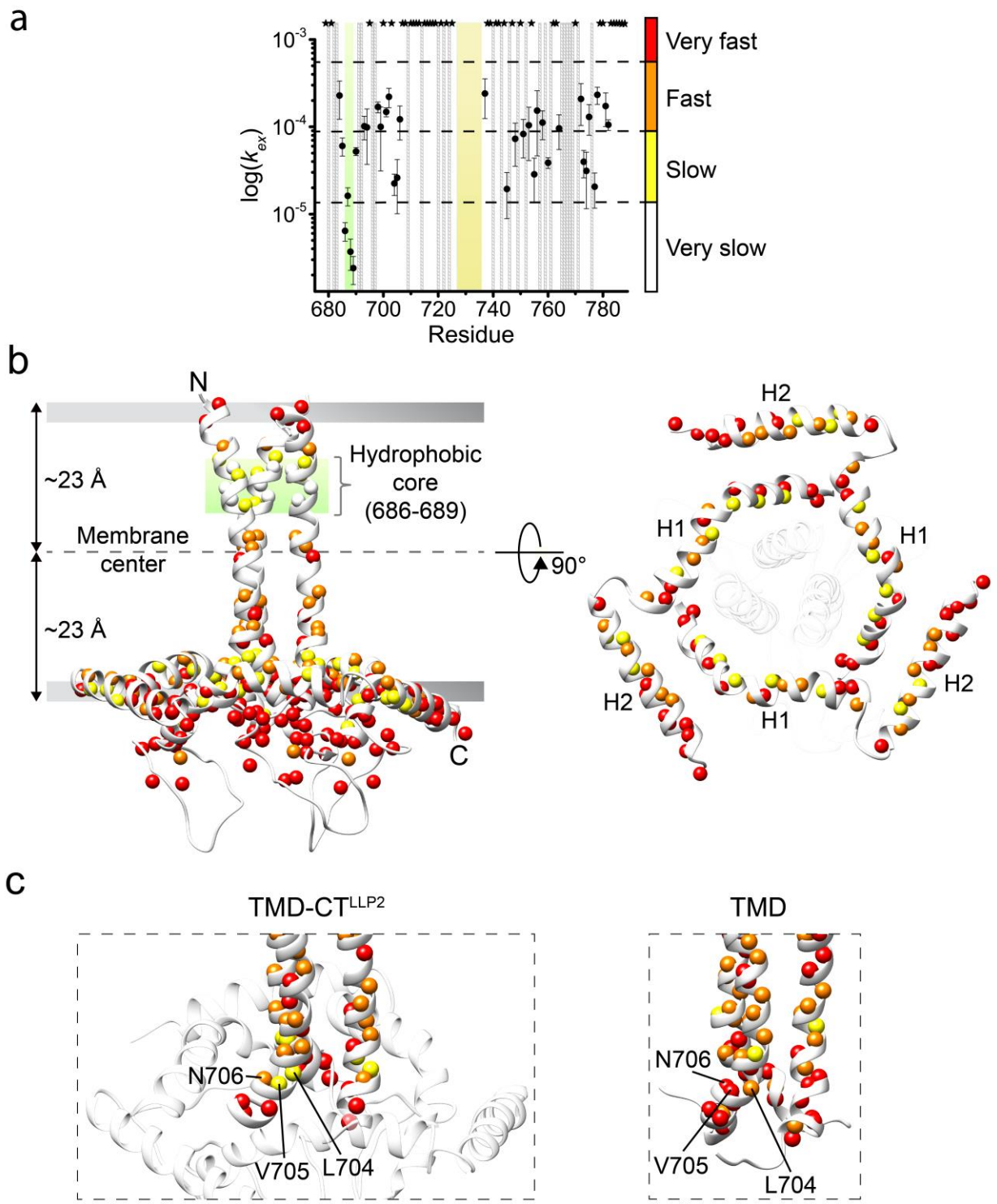
Supplementary Fig. 6 (continue in the next page)



Supplementary Fig. 6 (continue in the next page)

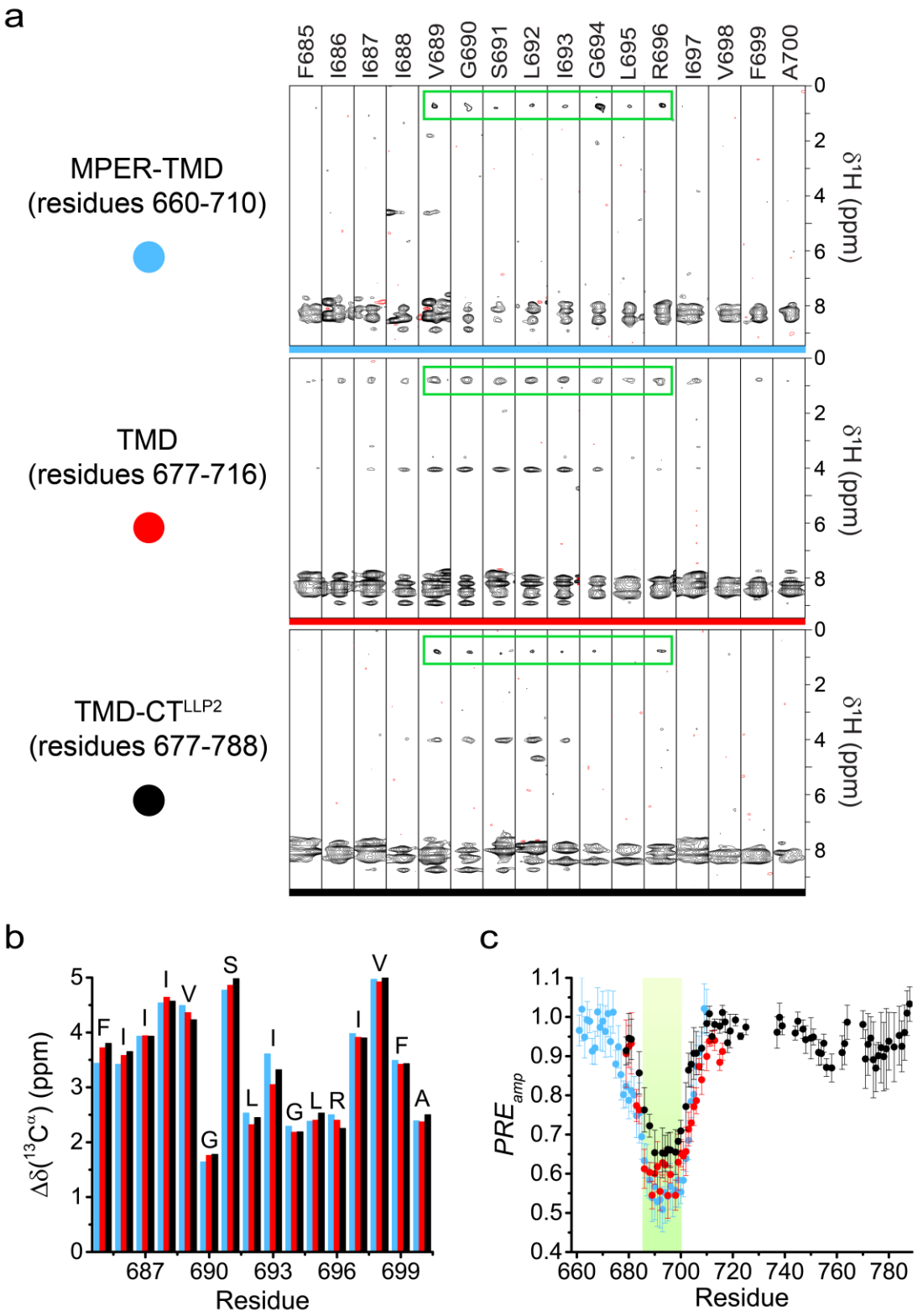


Supplementary Fig. 6 | Validation of the TMD-CT^{LLP2} structure by inter-chain PRE analysis. **a,** *Top*: Residue-specific PRE (I/I_0) plot of (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2} mixed with the TMD-CT^{LLP2} G738C mutant MTSL-labeled at C738 at ~1:1 molar ratio. Error bars represent the uncertainty derived from cross-peaks signal to noise. Missing bars are due to prolines or residues with overlapping peaks. The yellow box corresponds to residues 726-736 (deleted in the constructs), the grey triangles mark the proline residues and the asterisk indicates the position of the MTSL label. The horizontal dash lines define the three PRE regimes used to map onto the protein structure (bottom): strong ($I/I_0 \leq 0.6$; blue), medium ($0.6 < I/I_0 \leq 0.8$; light blue) and weak ($I/I_0 > 0.8$; white). PRE of 0.6 and 0.8 correspond approximately to a distance of 10 and 25 Å, respectively, from the paramagnetic tag. *Bottom*: Mapping of the PRE regimes onto the protein structure using the color scheme above. The (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2} and the MTSL-labeled G738C mutant are represented as white and yellow ribbons, respectively. Amide protons for which accurate PRE measurement was feasible are shown as spheres and colored according to the PRE regimes defined above. The yellow and red spheres cover 25 and 10 Å distance from the paramagnetic tag (highlighted in red), respectively. **b,** Same as (a) but for MTSL labeled at C764. **c,** Same as (a) but for MTSL labeled at C789. **d,** Control experiment in which only (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2} was used. The absence of PRE effect indicates that the sample preparation procedure completely removed excessive free MTSL (note: this is important as MTSL is hydrophobic and partitions in micelles or bicelles). Source data are provided as a Source Data file.



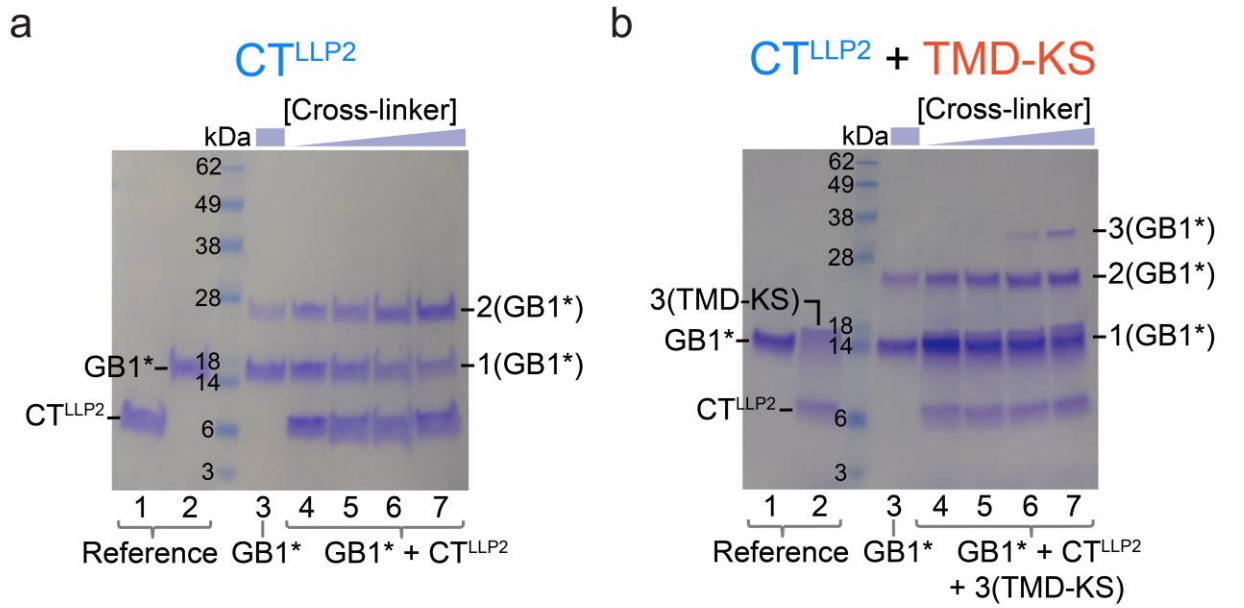
Supplementary Fig. 7 | H-D exchange of the TMD-CT^{LLP2} in q = 0.5 bicelles at 303 K. a, Residue-specific k_{ex} reported in logarithmic scale. Error bars represent the uncertainty derived from fitting error. Amide protons with exchange too fast to be measured are marked as ★. Residues that could not be analyzed (overlapping peaks and prolines) are marked with striped gray bars. The yellow box corresponds to deleted residues 726-736. The green box marks the hydrophobic core

(residues 686-689) that exhibits very slow exchange. The color code on the right represents the four different exchange regimes: very fast (red), fast (orange), slow (yellow), and very slow (white: very slow). **b,** Mapping of the four exchange regimes defined in (a) onto the TMD-CT^{LLP2} trimer structure, placed in lipid bilayer according to the PPT-derived membrane partition. The location of the hydrophobic core (residues 686-689) is highlighted with the green box. Slow-exchanging residues are mostly facing the bilayer interior, whereas very fast-exchanging residues are mostly outside the bilayer or within the TMD hydrophilic core[4]. **c,** Comparison of the water accessibility of the TM hydrophilic core between the TMD-CT^{LLP2} (left) and TMD [4] (right). Notably, residues 704-706, rapidly exchanging in the TMD, show significantly slower exchange in the TMD-CT^{LLP2} due to protection by the H1 helix of the CT^{LLP2}.

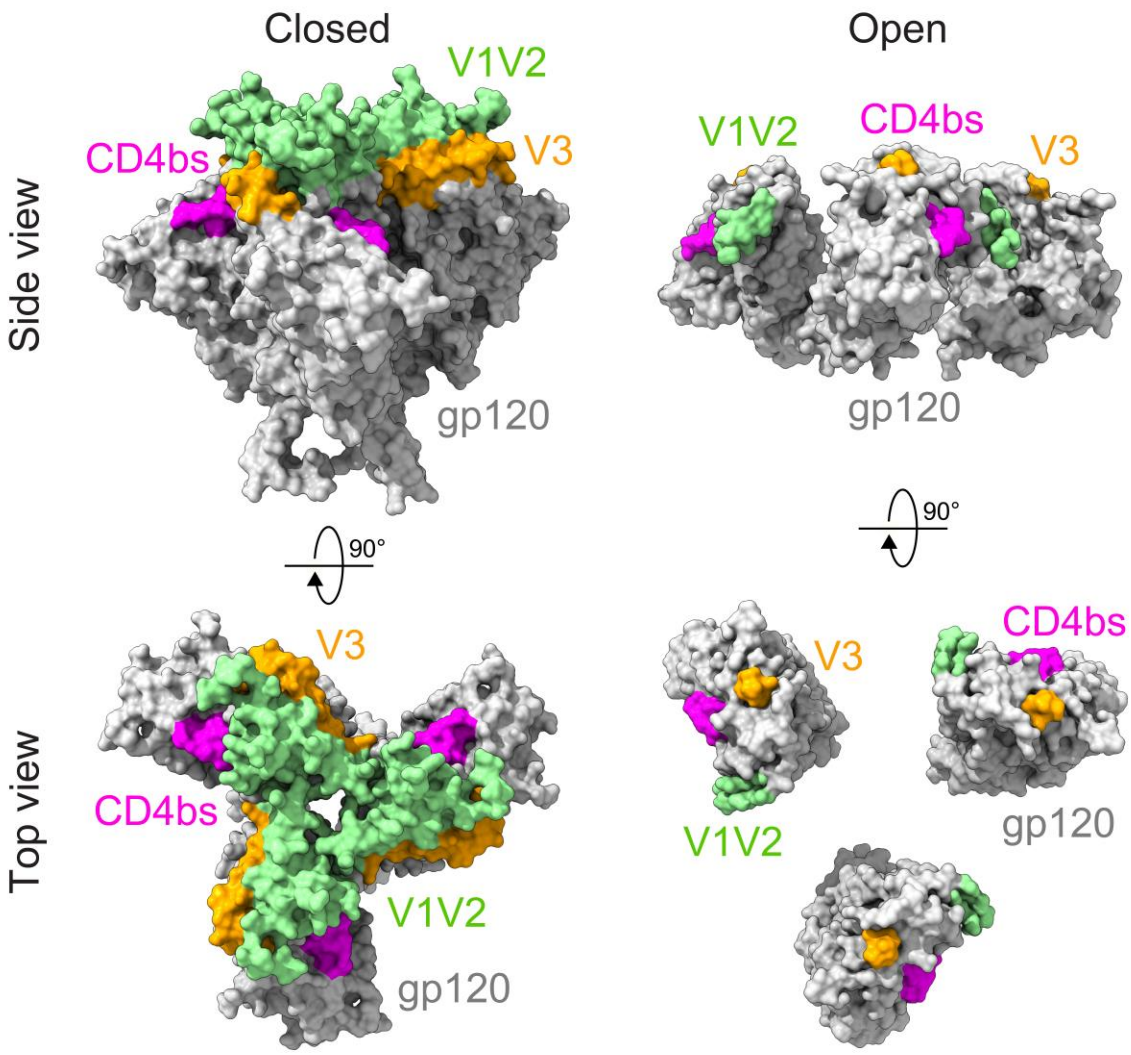


Supplementary Fig. 8 | Structural similarity of the TM core (residues 685-700) among the MPER-TMD, TMD and TMD-CT^{LLP2}. a, The strips corresponding to residues 685-700 taken from the 3D ^{15}N-edited NOESY-TROSY-HSQC spectra of the mixed samples of MPER-TMD (top), TMD (center) and TMD-CT^{LLP2} (bottom), showing similar pattern of inter-chain NOE cross-

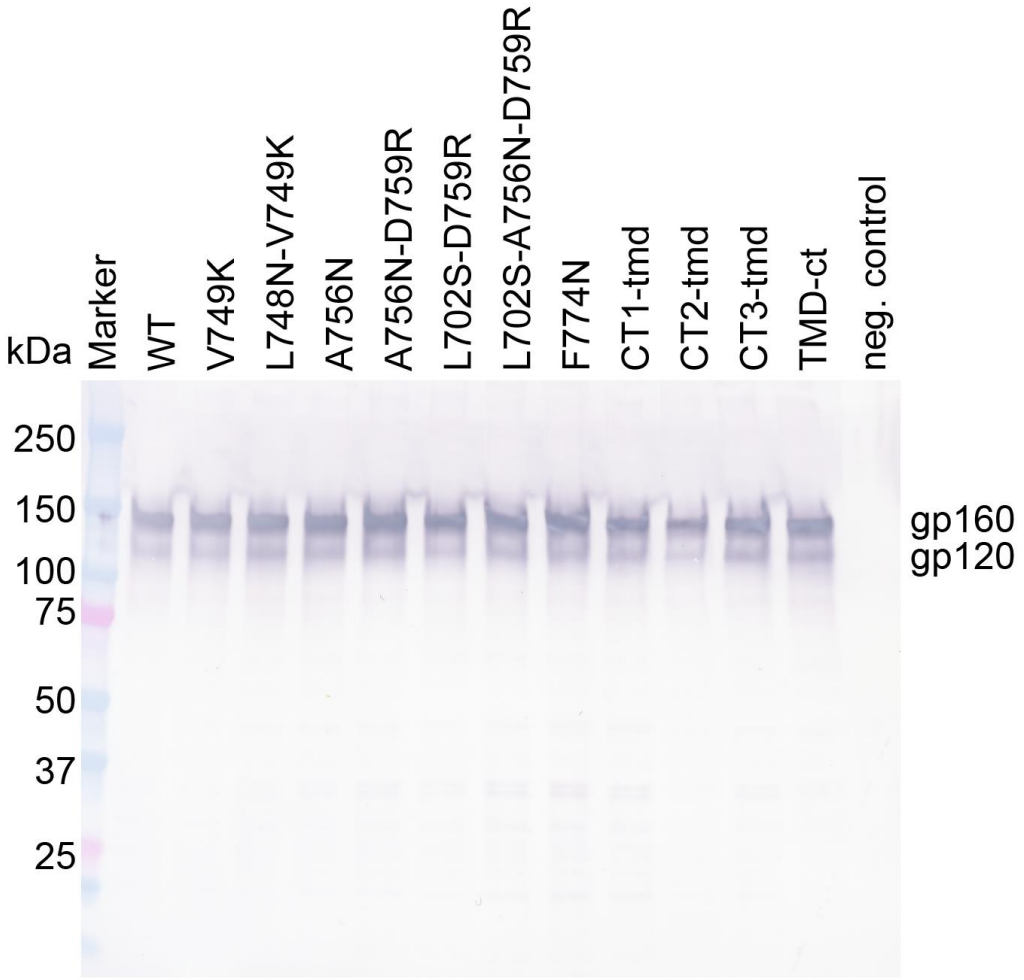
peaks. All spectra were recorded with mixed samples containing ~1:1 ratio of (^{2}H, ^{15}N)-labeled and ^{13}C-labeled chains. The same NOE mixing time of 200 ms was employed for all three experiments. The strongest inter-chain NOEs are highlighted in the green box (residues 689-696). The different peak intensities are due to the different size and relaxation properties of the three constructs. The spectra of the MPER-TMD and TMD were taken from previous studies[3,5]. **b,** Comparison of secondary $^{13}C^{\alpha}$ chemical shifts ($\Delta\delta$, experimental chemical shift minus the corresponding "random coil" shift) for residues 685-700 among the MPER-TMD (light blue), TMD (red) and TMD-CT^{LLP2} (black). The remarkable similarity in $\Delta\delta(^{13}C^{\alpha})$ indicates that the three constructs have the same secondary structure in the TM region. **c,** Overlay of the PRE_{amp} vs. *residue number* plots of the three constructs derived from Gd-DOTA titration, showing essentially the same membrane partition profiles for residues 685-700 (green box). The color definition of the three plots is the same as in (b). The notable agreement of the position of the PRE_{amp} minima indicates that the three constructs have virtually identical membrane partition. Error bars represent the uncertainty derived from fitting error. Source data are provided as a Source Data file.



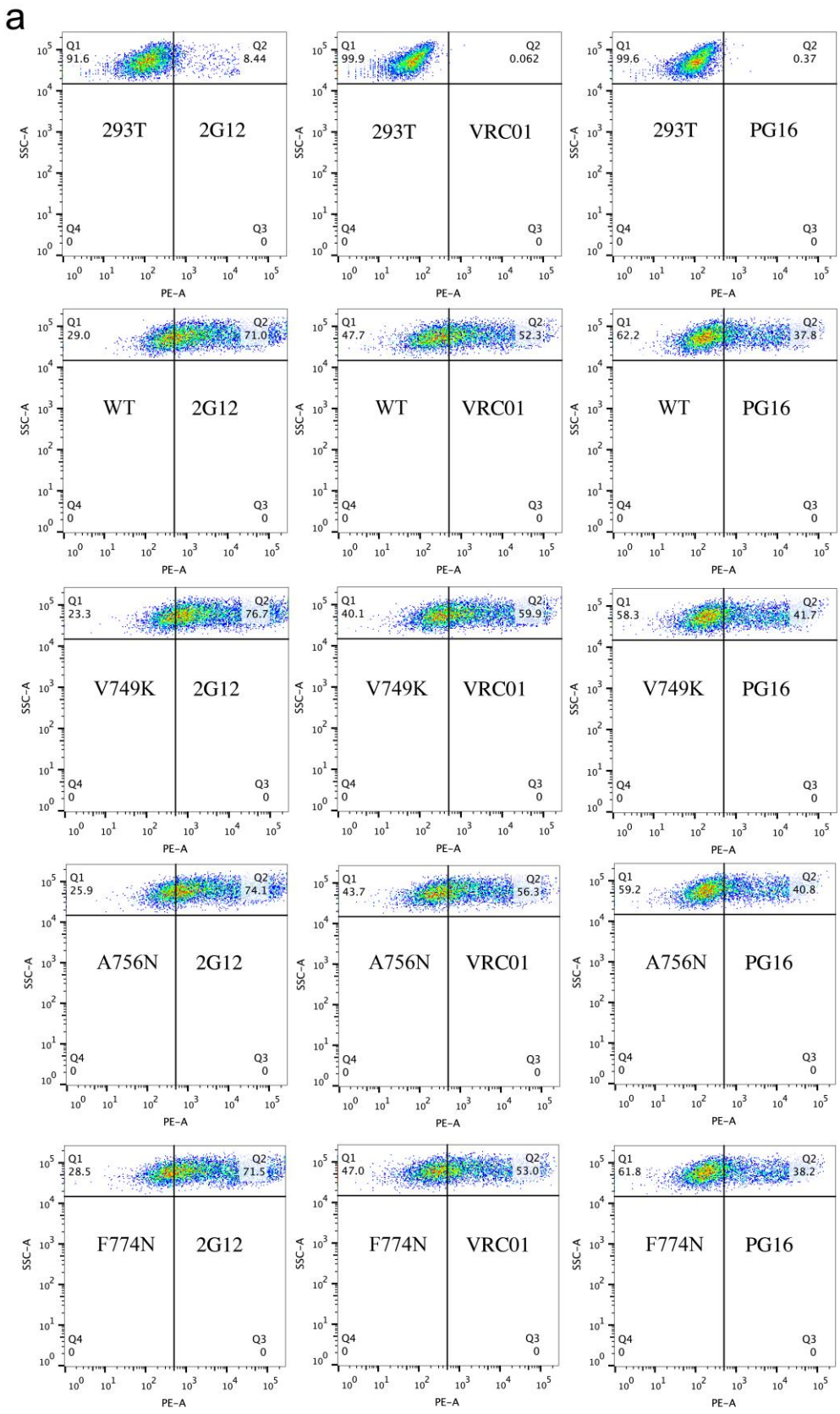
Supplementary Fig. 9 | Evidence of CT^{LLP2}–TMD interaction from OG-label analysis. a, Determination of the oligomerization state of the CT^{LLP2} reconstituted in bicelles by OG-label analysis (see Methods and Supplementary Fig. 1b for OG-label description). The SDS-PAGE gel lanes are: 1 – CT^{LLP2} (monomeric in SDS); 2 – GB1*; 3 – GB1* cross-linked alone; 4-7 – GB1* cross-linked in the presence of CT^{LLP2} at increasing cross-linker concentration. The increasing and decreasing strength of the dimer and monomer bands in lanes 4-7, respectively, compared to those in the control lane 3, suggests that the CT^{LLP2} can form dimers. The experiment was performed once. **b,** Oligomerization state of the CT^{LLP2} reconstituted in bicelles in the presence of the TMD-KS by OG-label analysis. The gel lanes are: 1 – GB1*; 2 – TMD-KS and CT^{LLP2} (trimeric and monomeric is SDS); 3 – GB1* cross-linked alone; 4-7 – GB1* cross-linked in the presence of CT^{LLP2} and TMD-KS at increasing cross-linker concentration. The experiment was performed once. Source data are provided as a Source Data file.



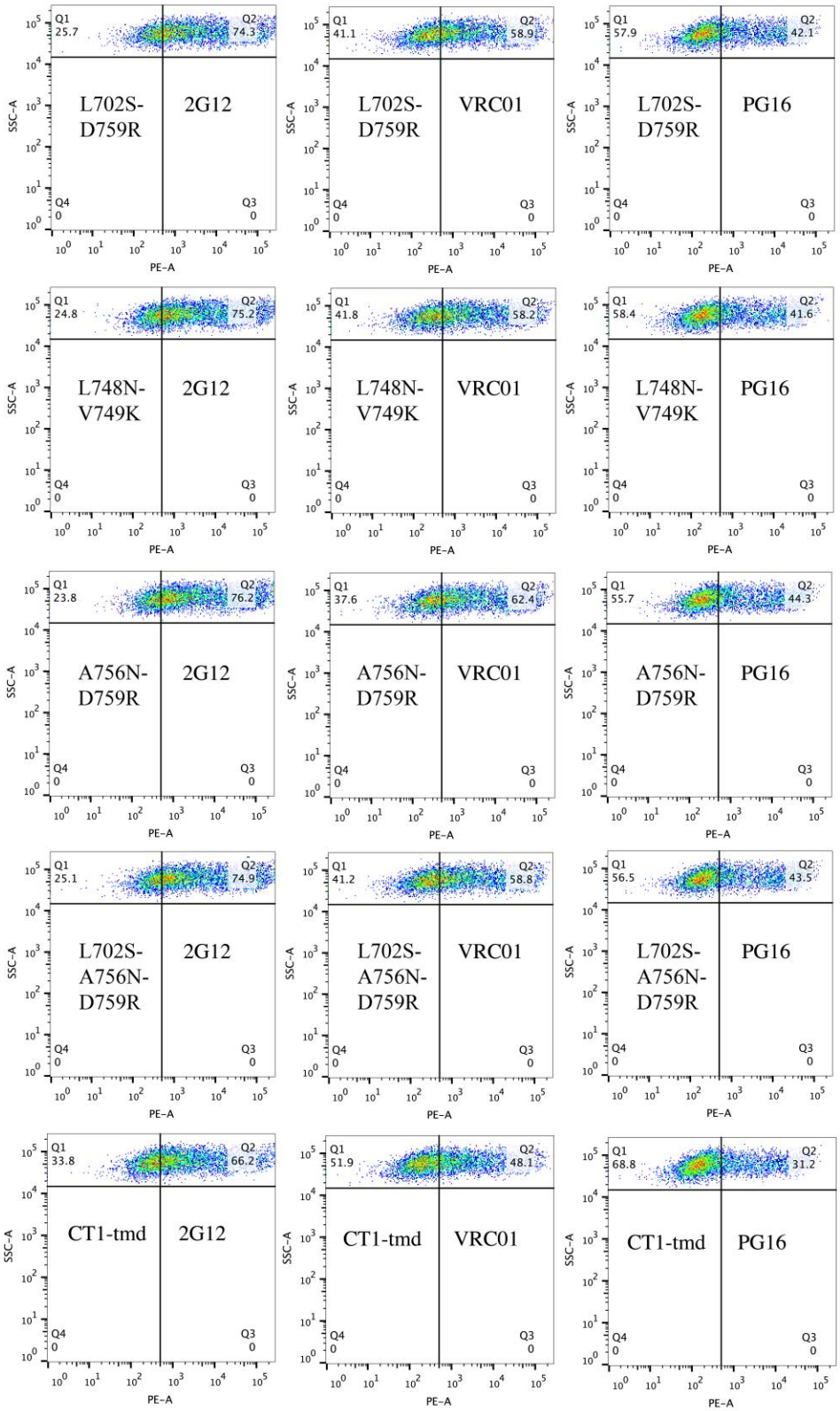
Supplementary Fig. 10 | Epitope mapping of the antibodies used in the pseudovirus-based neutralization assay. Side and top views of gp120 in the closed (prefusion state, left)[6] and open (CD4-bound)[7] conformations are shown. Locations of V1V2 loops, V3 loop and CD4 binding site (CD4bs) are highlighted in green, orange and magenta, respectively. Trimer-specific bnAbs such as PG9, PG16 and PGT145 can bind to the V1V2 loops only when the Env is in the closed conformation, while nonneutralizing or strain-specific neutralizing antibodies such as b6 (CD4bs), 3791 (V3) and 17b (CD4-induced) can bind to the Env only in the open conformation.



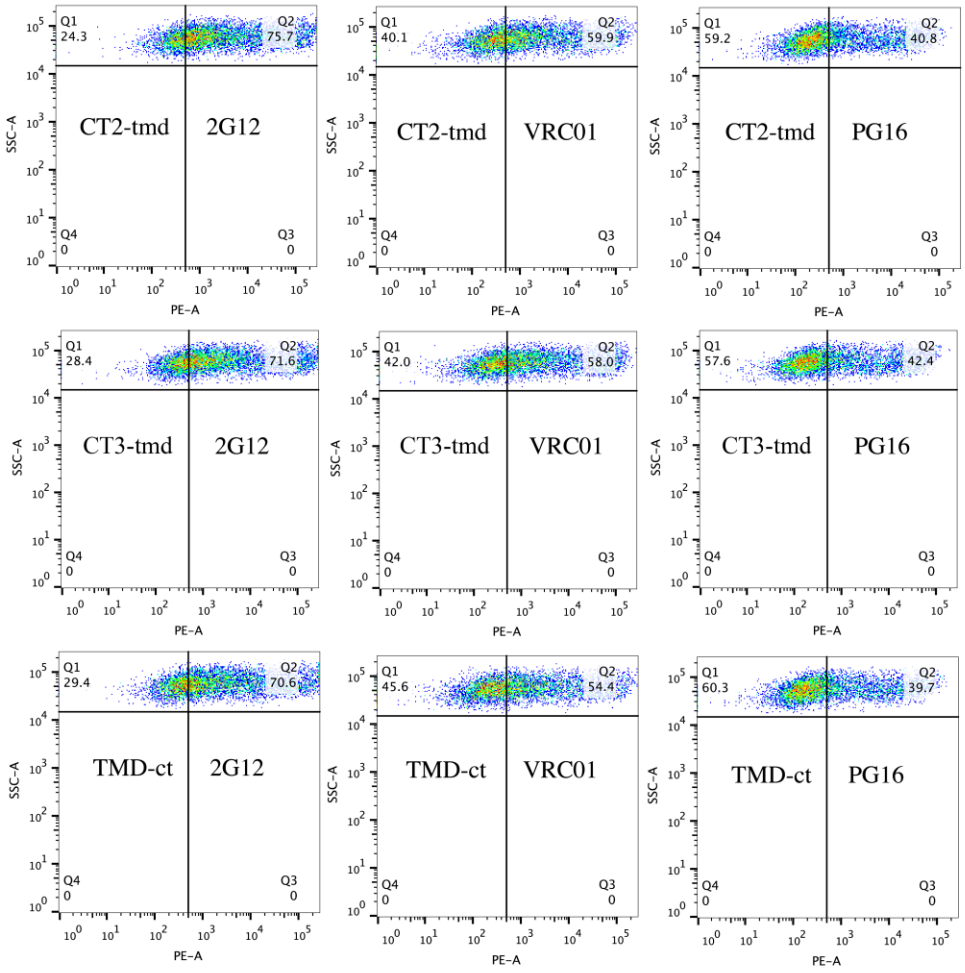
Supplementary Fig. 11 | Expression and processing of Env mutants expressed in 293T cells. Env samples prepared from 293T cells transiently transfected with 1 μg of the HIV-1 gp92UG037.8 gp160 expression plasmid or of each CT mutant were detected by an anti-V3 antibody 3791. The experiment was repeated independently with similar results twice. Source data are provided as a Source Data file.



Supplementary Fig. 12 (continue in the next page)



Supplementary Fig. 12 (continue in the next page)

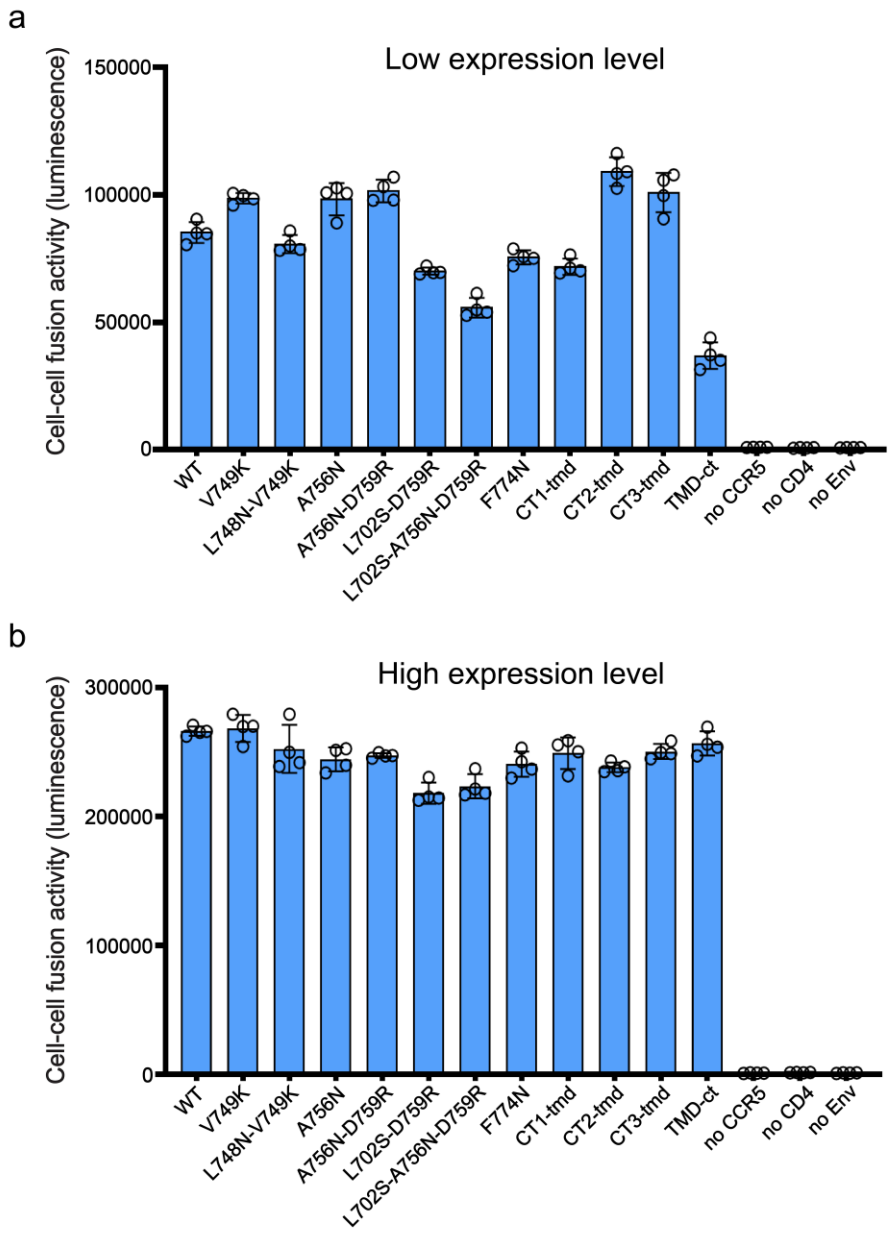


b

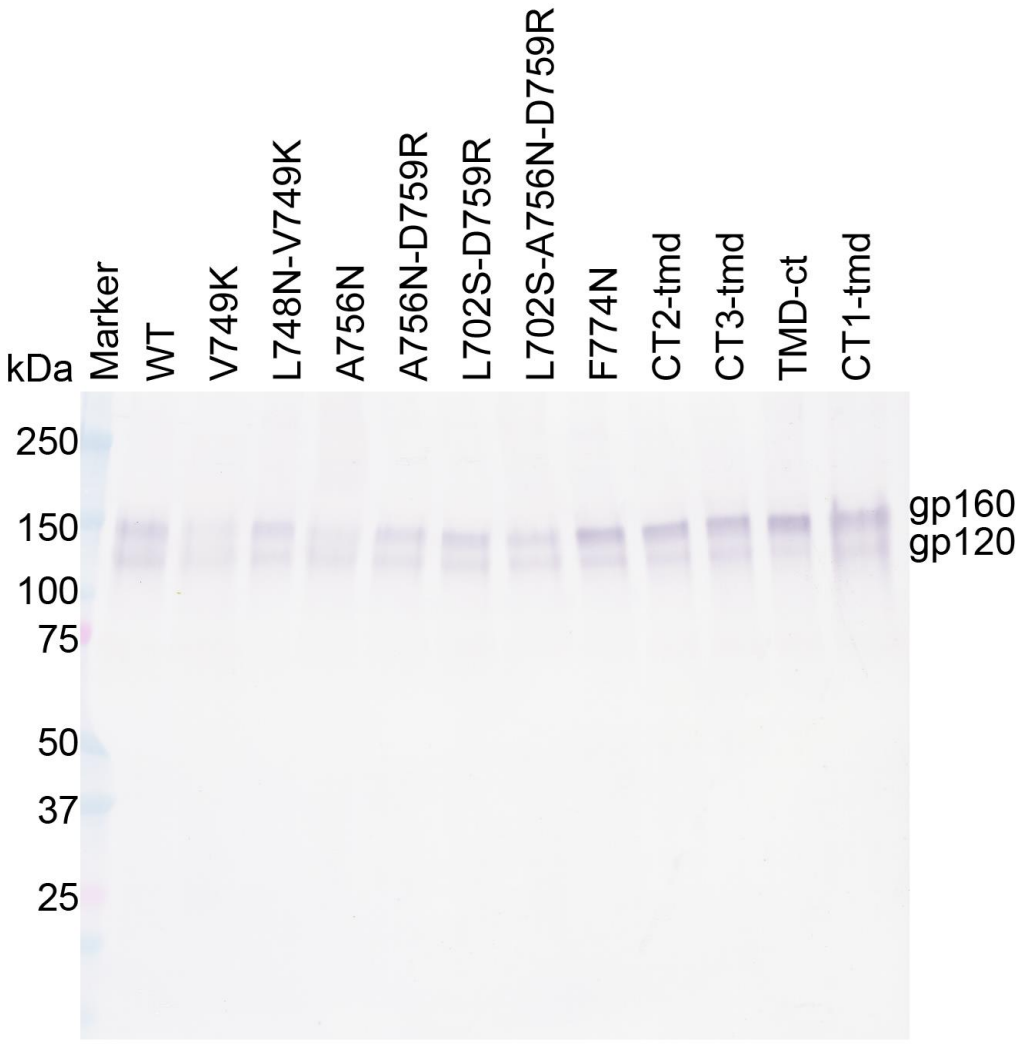
Mutation	Mean Fluorescence Intensity (MFI)		
	2G12	VRC01	PG16
WT	7991	5833	5232
V749K	8736	7238	5697
A756N	9405	6825	6217
F774N	8481	6903	5490
L702S-D759R	9287	7567	5440
L748N-V749K	8947	6856	6161
A756N-D759R	10161	8356	6352
L702S-A756N-D759R	9555	7297	6217
CT1-tmd	8170	4706	3157
CT2-tmd	8299	7376	5401
CT3-tmd	8061	7714	5988
TMD-ct	6907	6817	5436
293T	567	73	125

Supplementary Fig. 12 | Cell-surface expression of Env mutants detected by flow cytometry. **a,** Representative dot plots for negative control 293T cells, cells expressing the wildtype Env

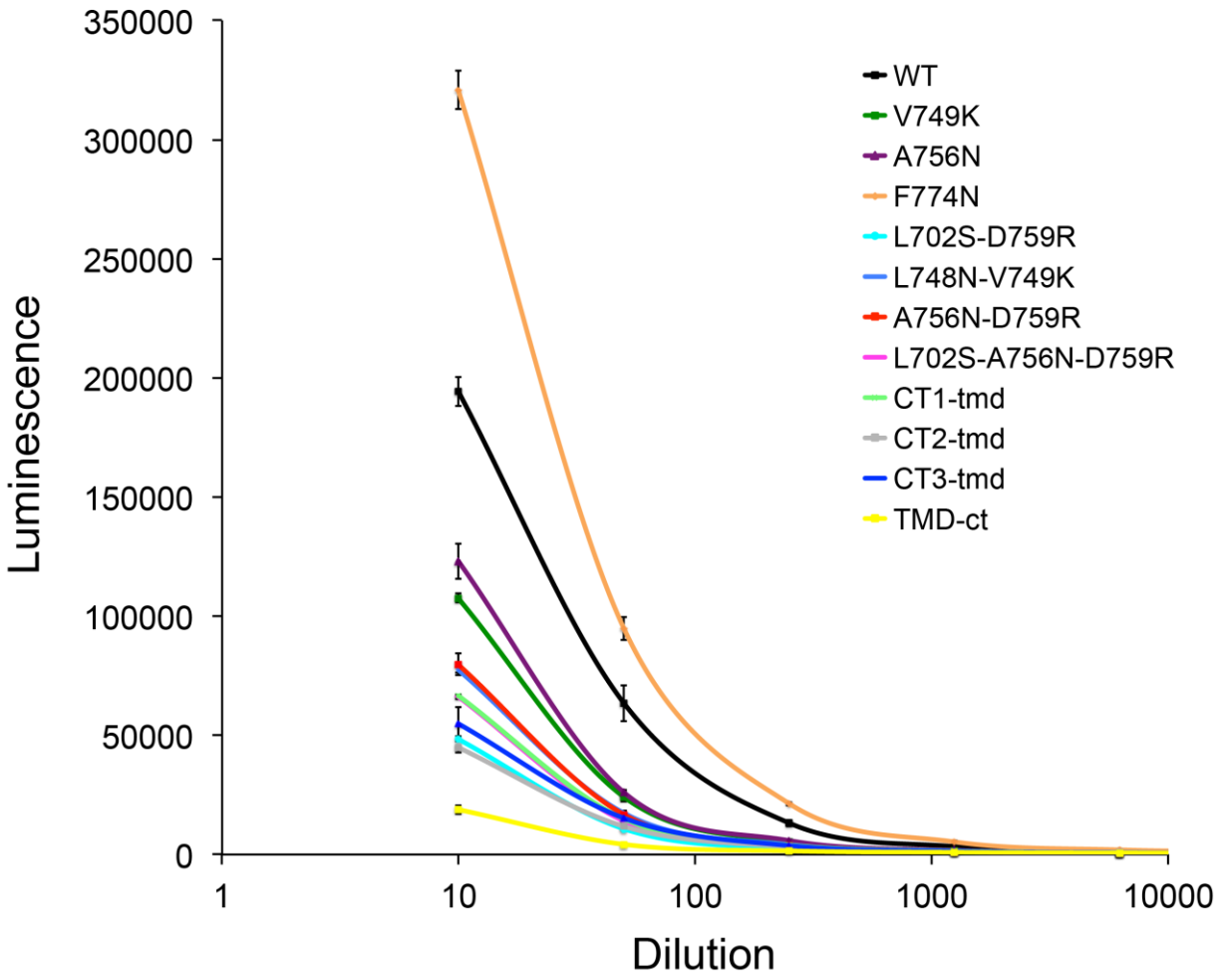
(WT), or CT mutants, measured by flow cytometry using monoclonal antibodies 2G12, VRC01 and PG16. **b,** Summary of mean fluorescence intensity (MFI) from data as shown in (a) for the WT and the 11 CT mutants, as well as the 293T cell control.



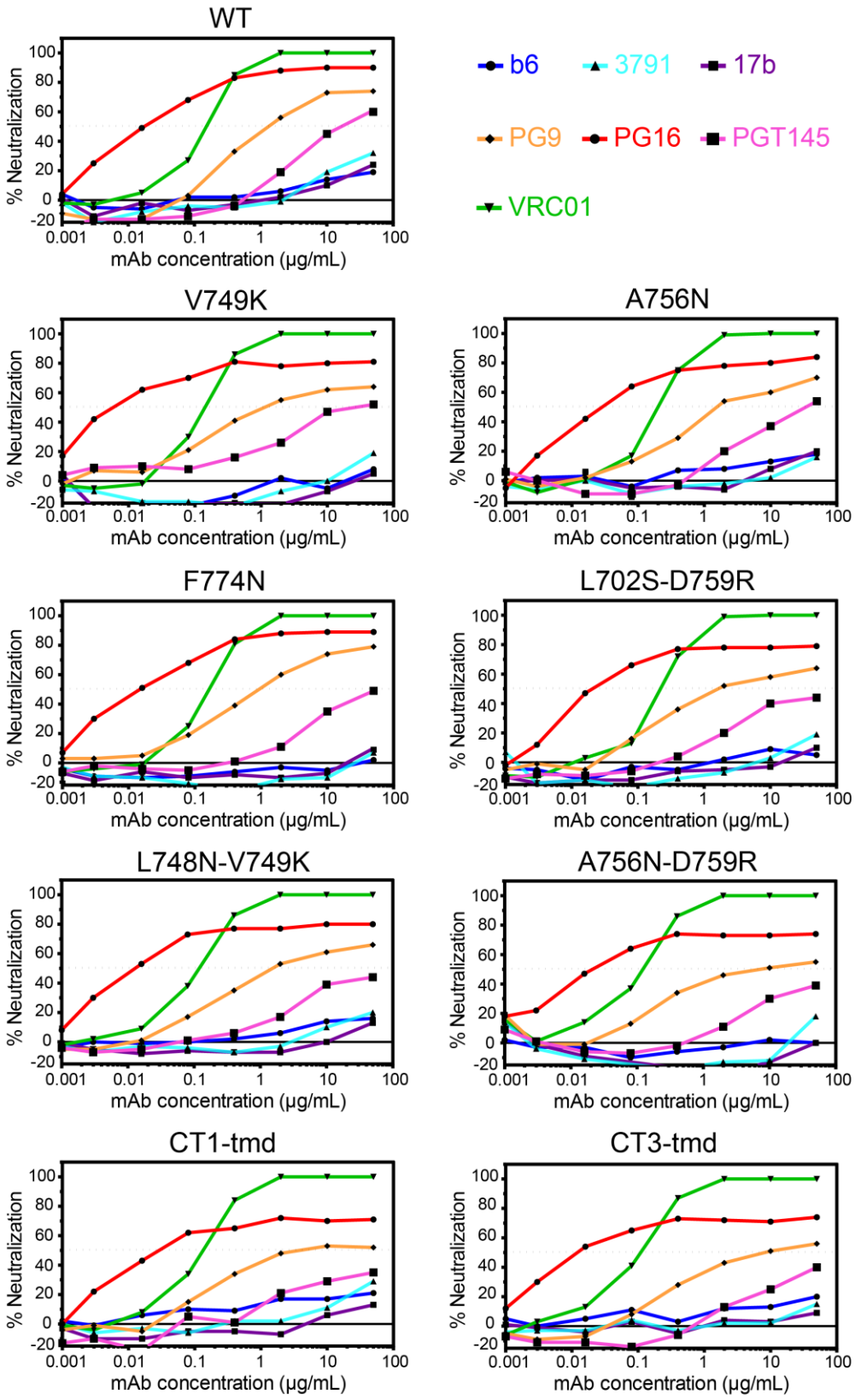
Supplementary Fig. 13 | Cell-cell fusion of HIV-1 Env and its CT mutants at different expression levels. a, Cell-cell fusion of HIV-1 Env and its CT mutants at a low expression level. 293T cells transfected with 75 ng of the 92UG037.8 Env expression plasmid or of each CT mutants were mixed with CD4- and CCR5-expressing cells. Cell-cell fusion led to reconstitution of active β-galactosidase and the fusion activity was quantified by a chemiluminescent assay. **b,** Cell-cell fusion of HIV-1 Env and its CT mutants at a high expression level. 293T cells transfected with 10 μg of the 92UG037.8 Env expression plasmid or of each CT mutants were fused with CD4- and CCR5-expressing cells. In (a) and (b), no Env, no CD4 and no CCR5 were negative controls. The experiments were 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 ± standard deviations from triplicate measurements. Source data are provided as a Source Data file.



Supplementary Fig. 14 | Incorporation of Env mutants into pseudoviruses analyzed by western blot. Env samples prepared from p24-normalized pseudoviruses containing either HIV-gp92UG037.8 Env or each of its CT mutants were detected by an anti-V3 antibody 3791. The experiment was repeated independently with similar results twice. Source data are provided as a Source Data file.

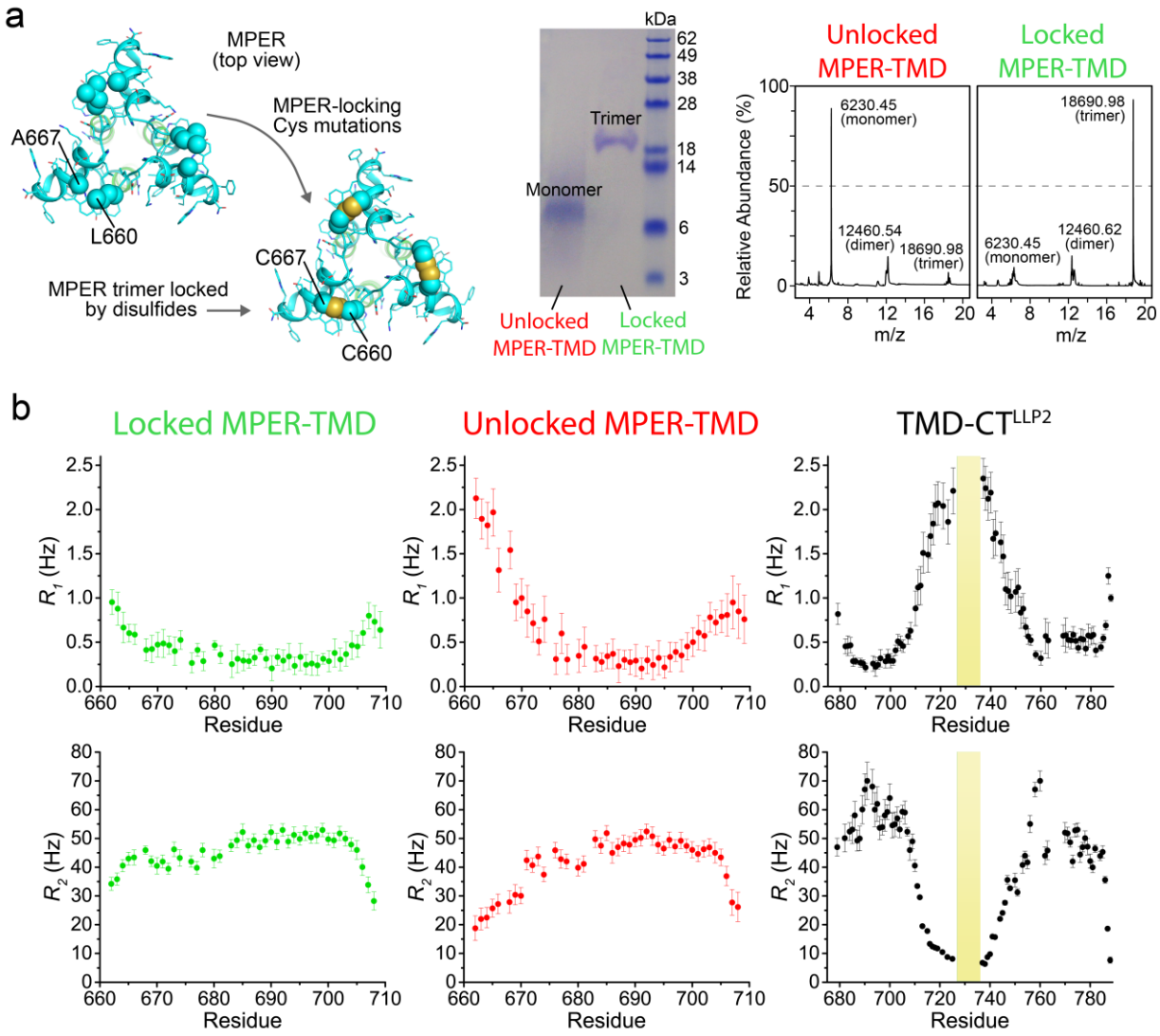


Supplementary Fig. 15 | Viral infectivity of Env mutants. Pseudoviruses containing Env CT mutants were normalized by p24-antigen, titrated 10 times using 5-fold dilution series, and tested for viral infectivity in TZM.bl cells. The experiment was carried out in quadruplicate. Data are presented as mean values and error bars indicate the standard deviation calculated by *Excel*. Source data are provided as a Source Data file.

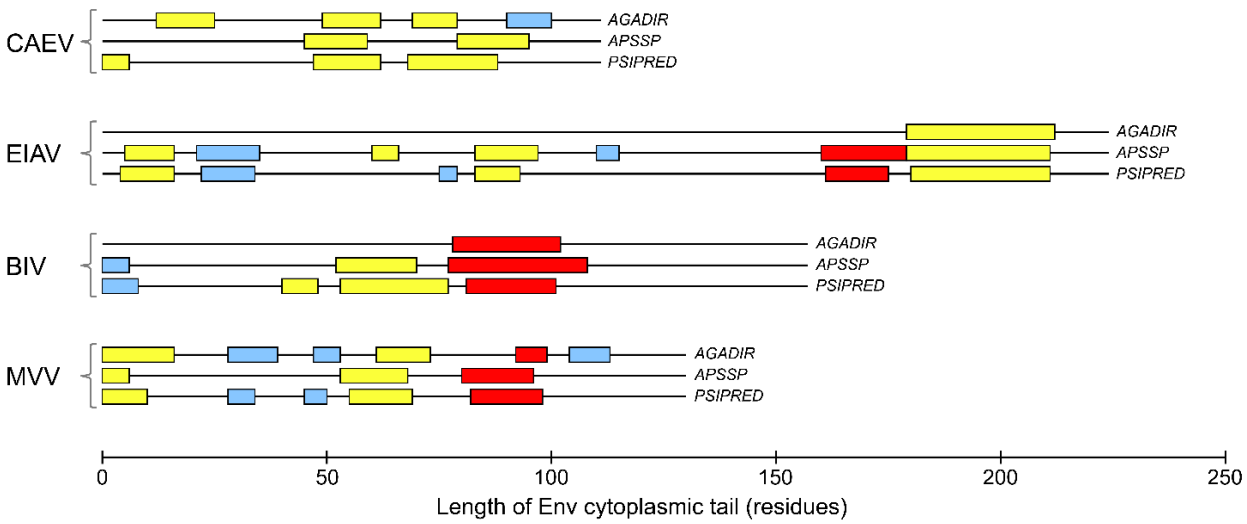


Supplementary Fig. 16 | Effect of mutations in the CT on Env antibody sensitivity. Antibody neutralization of pseudoviruses containing either the 92UG037.8 Env or the CT mutants was

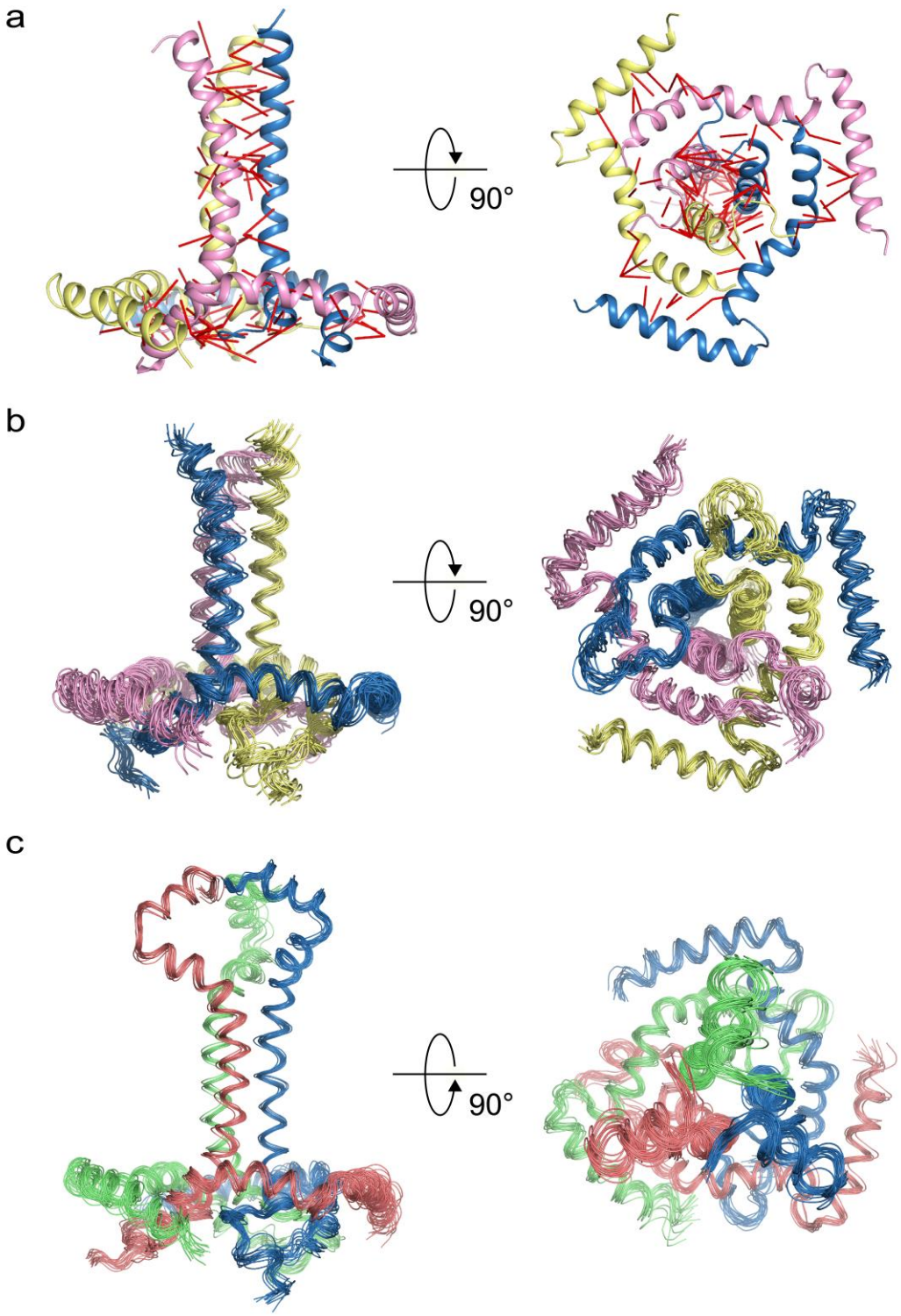
determined for non-neutralizing antibodies, including b6 (CD4 binding site; blue), 3791 (V3; cyan) and 17b (CD4-induced; purple), and trimer-specific bnAbs, including PG9 (orange), PG16 (red) and PGT145 (magenta). The CD4 binding site bnAb VRC01 was a control antibody shown in green. The experiment was performed in duplicate. Source data are provided as a Source Data file.



Supplementary Fig. 17 | Locking of the MPER-TMD in the prefusion state and ^{15}N backbone dynamics of the locked MPER-TMD, unlocked MPER-TMD and TMD-CT^{LLP2}. **a,** Top view of the MPER-TMD trimer with L660 of one chain in close contact with A667 of a neighboring chain (Left). Mutating L660 and A667 to Cys results in disulfide bonds that lock the trimer structure. Urea-PAGE (center) and mass spectrometry (right) analyses of the unlocked (reduced) and locked (oxidized) MPER-TMD mutant showing that it is a covalent trimer when oxidized. The experiment was performed once. **b,** Residue-specific ^{15}N R_1 and R_2 relaxation rates of the locked MPER-TMD (green), unlocked MPER-TMD (red) and TMD-CT^{LLP2} (black). Error bars represent the uncertainty derived from fitting error. Missing data points are due to proline or residues with overlapping peaks. The yellow box marks the deleted residues 726-736 in the TMD-CT^{LLP2} construct. Source data are provided as a Source Data file.

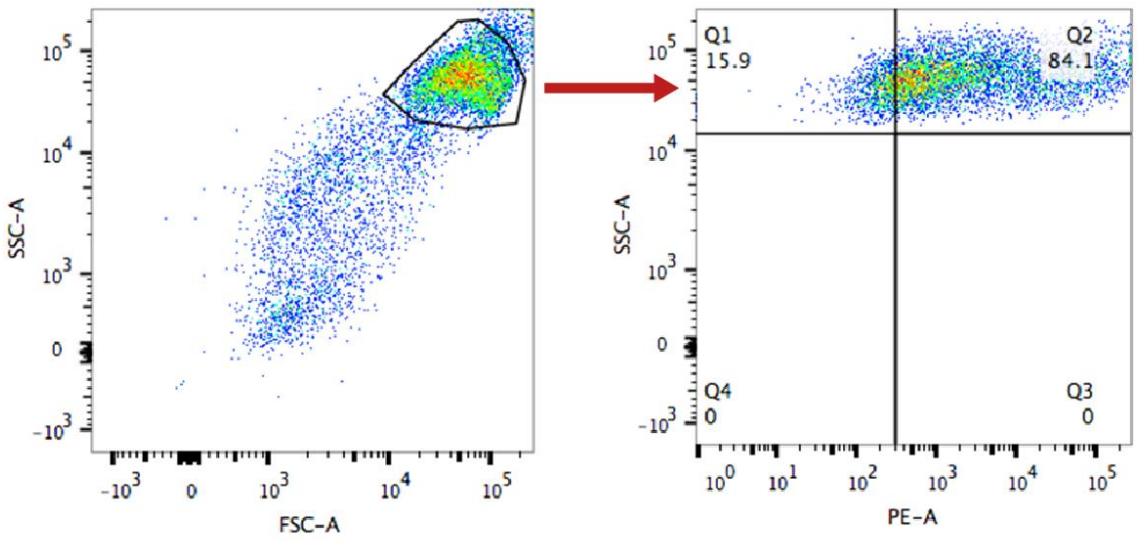


Supplementary Fig. 18 | Relative length of the cytoplasmic tails of fusion proteins from other lentiviruses. For each of the viruses, including the CAEV (caprine arthritis encephalitis virus), EIAV (equine infectious anemia virus), BIV (bovine immunodeficiency virus), and MVV (maedi/visna virus), the cytoplasmic tail (CT) sequence of the viral fusion protein was subject to protein secondary structure prediction using three different programs: *AGADIR*[8], *APSSP*[9], and *PSIPRED*[10]. The sources of the sequences are: UniProtKB D1M7T1 (CAEV), UniProtKB P32541 (EIAV), UniProtKB P19557 (BIV), and UniProtKB P03379) (MVV). The CT sequences are aligned so that the TM helix ends at position 0. The predicted helical segments are represented as boxes colored in red, yellow, and blue coding for hydrophobic, amphipathic or hydrophilic helices, respectively.



Supplementary Fig. 19 | NMR-derived TMD-CT^{LLP2} and MPER-TMD-CT^{LLP2} structural ensembles. a, Side and bottom view of the TMD-CT^{LLP2} average structure, represented as ribbon, showing NOE-derived inter-chain restraints (red lines). The unstructured KS (residues 711-736) has been omitted for clarity. **b,** Side and bottom view of the ensemble of the 15 lowest-energy TMD-CT^{LLP2} structures calculated using NMR-derived structural restraints (Supplementary Table

11). Structures are represented as thin ribbons and Residues 726-736 have been omitted for clarity. **c,** Side and top view of the NMR ensemble of the 15 lowest-energy MPER-TMD-CT^{LLP2} structures calculated using NMR-derived structural restraints (Supplementary Table 12). Structures are represented as thin ribbons and residues 726-736 have been omitted.



Supplementary Fig. 20 | Flow cytometry gating strategy.

Supplementary Table 1 | List of CT mutants. Mutants (#1-21) are named borrowing the residue type and number from the 92UG037.8 and 92UG024.2 sequence, respectively. Mutants (#22-28) are named indicating the site of the mutation and the interface affected in uppercase and lowercase, respectively. The mutants that have been thoroughly characterized are highlighted in yellow.

Mutant	#	Description		Control exps.
		92UG037.8 sequence (functional data)	92UG024.2 sequence (NMR constructs)	
L702Q	1	L697Q	L702Q	
L702S	3	L697S	L702S	
I746K	7	I741K	I746K	
R747D	5	R742D	R742D	
L748N	2	L743N	L748N	
L748Q	4	L743Q	L748Q	
V749K	6	V744K	V749K	✓
G751Y	8	G746Y	G751Y	
L755N	9	L750N	L755N	
A756N	10	A751N	I756N	✓
D758R	11	D753R	D758R	
D759R	12	D754R	D759R	
L771K	13	L766K	L771K	
F774N	14	F669N	L774N	✓
R780D	15	R775D	R780D	
L702S-D759R	16	L697S-D754R	L702S- D759R	✓
L748N-V749K	17	L743N-V744K	L748N- V749K	✓
A756N-D759R	18	A751N-D754R	I756N- D759R	✓
L771K-F774N	19	L766K-F769N	L771K-L774N	
L702S-A756N-D759R	20	L697S-A751N-D754R	L702S-I756N-D759R	✓
A756N-L760S-L763N-C764E	21	A751N-L755S-L758N-C759E	I756N-L760S-L763N-C764E	
polyA(748-762)	22	polyA(743-757)	polyA(748-762)	
Δ(748-762)	23	Δ(743-757)	Δ(748-762)	
CT1-tmd	24	I741N-L743S-V744R-L750A-D753K-D754R	I746N-L748S-V749R-L755A-D758K-D759R	✓
CT2-tmd	25	L743S-L750S-D753A-D754A-N757A	L748S-L755S-D758A-D759A-S762A	✓
CT3-tmd	26	I741S-L743S-V744S-L750S-D753A-D754A-N757A	I746S-L748S-V749S-L755S-D758A-D759A-S762A	✓
CT4-tmd	27	I741N-L743S-V744R-L750A-A751N-D753K-D754R-L755S-L758N-C759E	I746N-L748S-V749R-L755A-I756N-D758K-D759R-L760S-L763N-C764E	
TMD-ct	28	V699S-I700S-N701A-V703S-R704A-Q705A	L704S-V705S-N706A-V708S-R709A-Q710A	✓

Supplementary Table 2 | Antibody neutralization of CT mutants.

Titer in TZM.bl Cells (μg/mL)												
Mutant		TCID50	b6 IgG			17b			3791			
WT		69,877	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
			>50	>50	19	>50	>50	24	>50	>50	32	
VRC01			PG9			PG16			PG145			
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
0.143	0.318	100	1.075	>50	74	0.017	0.210	90	18.278	>50	61	
Mutant		TCID50	b6 IgG			17b			3791			
L702Q		18,275	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
			>50	>50	4	>50	>50	8	>50	>50	0	
VRC01			PG9			PG16			PG145			
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
0.226	0.599	100	2.508	>50	65	0.034	1.070	82	>50	>50	43	
Mutant		TCID50	b6 IgG			17b			3791			
L702S		69,877	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
			>50	>50	4	>50	>50	12	>50	>50	14	
VRC01			PG9			PG16			PG145			
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
0.215	0.722	100	1.521	>50	74	0.032	0.277	88	31.473	>50	53	
Mutant		TCID50	b6 IgG			17b			3791			
I746K		349,386	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
			>50	>50	23	>50	>50	25	>50	>50	34	
VRC01			PG9			PG16			PG145			
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
0.115	0.320	100	1.012	>50	69	0.013	0.462	84	24.266	>50	57	
Mutant		TCID50	b6 IgG			17b			3791			
R747D		18,275	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
			>50	>50	0	>50	>50	0	>50	>50	2	
VRC01			PG9			PG16			PG145			
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
0.140	0.374	100	0.854	>50	72	0.013	0.219	87	46.006	>50	52	
Mutant		TCID50	b6 IgG			17b			3791			
L748N		53,437	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
			>50	>50	3	>50	>50	0	>50	>50	23	
VRC01			PG9			PG16			PG145			
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	
0.113	0.266	100	1.058	>50	72	0.011	0.573	84	31.535	>50	57	

Supplementary Table 2 (continued)

Mutant		TCID50	b6 IgG			17b			3791		
L748Q		781,250	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	17	>50	>50	12	>50	>50	17
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.129	0.356	100	0.767	>50	76	0.010	0.233	87	33.764	>50	53
Mutant		TCID50	b6 IgG			17b			3791		
V749K		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	8	>50	>50	5	>50	>50	19
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.135	0.297	100	0.964	>50	64	0.006	1.700	81	20.588	>50	52
Mutant		TCID50	b6 IgG			17b			3791		
G751Y		69,877	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	3	>50	>50	3	>50	>50	18
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.096	0.258	100	1.318	>50	66	0.008	>50	76	>50	>50	48
Mutant		TCID50	b6 IgG			17b			3791		
L755N		6,250	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	21	>50	>50	11	>50	>50	34
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.106	0.225	100	1.444	>50	66	0.014	0.884	83	12.172	>50	51
Mutant		TCID50	b6 IgG			17b			3791		
A756N		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	18	>50	>50	20	>50	>50	16
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.189	0.503	100	1.835	>50	70	0.028	11.168	84	32.972	>50	54
Mutant		TCID50	b6 IgG			17b			3791		
D758R		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	7	>50	>50	5	>50	>50	23
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.088	0.294	100	0.801	>50	70	0.009	0.491	83	26.853	>50	52

Supplementary Table 2 (continued)

Mutant		TCID50	b6 IgG			17b			3791		
D759R		2,795	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	5	>50	>50	0	>50	>50	27
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.115	0.248	100	0.920	>50	66	0.014	0.424	79	49.309	>50	50
Mutant		TCID50	b6 IgG			17b			3791		
L771K		53,437	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	0	>50	>50	0	>50	>50	14
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.092	0.249	100	0.672	>50	77	0.009	0.080	88	28.761	>50	55
Mutant		TCID50	b6 IgG			17b			3791		
F774N		1,746,928	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	2	>50	>50	10	>50	>50	7
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.152	0.409	100	0.742	>50	79	0.015	0.200	89	>50	>50	49
Mutant		TCID50	b6 IgG			17b			3791		
R780D		18,275	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	1	>50	>50	12	>50	>50	24
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.082	0.304	100	0.573	>50	78	0.007	0.095	90	15.588	>50	61
Mutant		TCID50	b6 IgG			17b			3791		
L702S-D759R		3,655	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	5	>50	>50	10	>50	>50	19
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.205	0.541	100	1.153	>50	64	0.021	>50	79	>50	>50	44
Mutant		TCID50	b6 IgG			17b			3791		
L748N-V749K		267,184	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	16	>50	>50	13	>50	>50	20
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.111	0.303	100	1.647	>50	66	0.013	36.106	80	>50	>50	44

Supplementary Table 2 (continued)

Mutant		TCID50	b6 IgG			17b			3791		
A756N-D759R		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	0	>50	>50	0	>50	>50	18
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.102	0.294	100	3.277	>50	55	0.023	>50	74	>50	>50	39
Mutant		TCID50	b6 IgG			17b			3791		
L771K-F774N		18,275	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	8	>50	>50	7	>50	>50	20
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.104	0.284	100	0.504	39.024	82	0.005	0.108	91	19.283	>50	62
Mutant		TCID50	b6 IgG			17b			3791		
L702S-A756N-D759R		10,687	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	12	>50	>50	6	>50	>50	16
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.140	0.310	100	>50	>50	48	0.018	>50	67	>50	>50	30
Mutant		TCID50	b6 IgG			17b			3791		
A756N-L760S-L763N-C764E		1,250	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	1	>50	>50	0	>50	>50	17
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.085	0.219	100	0.280	7.796	83	0.006	0.044	91	7.238	>50	64
Mutant		TCID50	b6 IgG			17b			3791		
polyA(748-762)		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	12	>50	>50	15	>50	>50	32
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.104	0.295	100	1.266	>50	72	0.011	0.545	83	33.704	>50	56
Mutant		TCID50	b6 IgG			17b			3791		
Δ(748-762)		91,376	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	20	>50	>50	17	>50	>50	24
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.108	0.298	100	1.406	>50	70	0.011	1.220	83	>50	>50	50

Supplementary Table 2 (continued)

Mutant		TCID50	b6 IgG			17b			3791		
CT1-tmd		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	21	>50	>50	13	>50	>50	29
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.126	0.342	100	2.533	>50	52	0.027	>50	71	>50	>50	35
Mutant		TCID50	b6 IgG			17b			3791		
CT2-tmd		53,437	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	20	>50	>50	9	>50	>50	23
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.095	0.266	100	46.920	>50	51	0.063	>50	66	>50	>50	37
Mutant		TCID50	b6 IgG			17b			3791		
CT3-tmd		91,376	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	20	>50	>50	9	>50	>50	15
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.088	0.308	100	8.109	>50	56	0.012	>50	74	>50	>50	40
Mutant		TCID50	b6 IgG			17b			3791		
CT4-tmd		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			>50	>50	11	>50	>50	10	>50	>50	26
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.085	0.247	100	0.322	6.742	84	0.007	0.095	91	8.695	>50	60
Mutant		TCID50	b6 IgG			17b			3791		
TMD-ct		13,975	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
			1.148	>50	60	>50	>50	45	1.208	>50	70
VRC01			PG9			PG16			PG145		
IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
0.056	0.156	100	>50	>50	0	>50	>50	2	>50	>50	0

Supplementary Table 3 | Chronic HIV+ serum samples for tier phenotyping.

Titer in TZM.bl Cells (μg/mL)												
Sample ID	**WT**			**CT2-tmd**			**CT3-tmd**			**TMD-ct**		
	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
HIV-018	<20	<20	22	<20	<20	48	<20	<20	43	335	28	85
HIV-019	31	<20	60	21	<20	51	28	<20	59	155	29	85
HIV-021	<20	<20	6	<20	<20	26	<20	<20	29	89	<20	80
HIV-023	<20	<20	0	<20	<20	0	<20	<20	0	<20	<20	44
HIV-024	486	179	97	821	264	99	986	313	98	1,776	500	100
HIV-025	<20	<20	0	<20	<20	13	<20	<20	20	227	39	85
HIV-026	<20	<20	0	<20	<20	0	<20	<20	0	90	<20	71

Supplementary Table 4 | Additional bnAbs epitope phenotyping.

Titer in TZM.bl Cells (μg/mL)												
bnAbs (epitope)	**WT**			**CT2-tmd**			**CT3-tmd**			**TMD-ct**		
	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI	IC50	IC80	MPI
3BNC117 (CD4bs)	0.015	0.044	100	0.010	0.029	100	0.008	0.030	100	0.005	0.025	100
PGT121 (V3-glycan)	0.097	0.266	100	0.058	0.202	100	0.052	0.185	100	0.052	0.147	100
10-1074 (V3-glycan)	0.101	0.331	100	0.100	0.342	100	0.072	0.241	100	0.096	0.318	100
PGDM1400 (V3-glycan)	0.103	0.347	99	0.066	0.231	99	0.063	0.221	99	0.061	0.318	96
PGT151 (gp120/gp41 interface)	>50	>50	14	>50	>50	14	>50	>50	24	>50	>50	0
8ANC194 (gp120/gp41 interface)	1.044	4.636	98	0.697	3.357	98	0.537	3.749	98	0.545	1.893	99
4E10 (MPER)	0.467	3.012	99	0.046	0.289	100	0.049	0.424	100	0.031	0.186	100

Supplementary Table 5 | Amino acid sequence of the NMR and OG-label constructs used in this study. The gp41 sequences are based on that of clade D HIV-1 isolate 92UG024.2. Residues of the MPER, TMD-KS and CT^{LLP2} are shown in blue, red and green, respectively. The 10 residues of the KS that were deleted are show in purple. Residues highlighted in bold and black indicate the relevant modifications introduced to each construct.

WT gp41 (residues 677-788)	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRGPDRP EGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVE LLGRR
NMR constructs	
TMD-CT^{LLP2} (residues 677-725; 737-788)	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRGGERD RDRSIRLVNGSLALIWDDLRSL**S**LFSYHRLRDLLLIVTRIVELLGRR Note: residues 726-736 were deleted; C764S
TMD-CT^{LLP2} G738C	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRG**C**ERD RDRSIRLVNGSLALIWDDLRSLSLFSYHRLRDLLLIVTRIVELLGRR
TMD-CT^{LLP2} S764C	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRGGERD RDRSIRLVNGSLALIWDDLRSL**C**LFSYHRLRDLLLIVTRIVELLGRR Note: the native Cys was reintroduced
TMD-CT^{LLP2} C789	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRGGERD RDRSIRLVNGSLALIWDDLRSLSLFSYHRLRDLLLIVTRIVELLGRR**C**
TMD-KS (residues 677-725)	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPR**M** Note: the additional C-terminal Met was introduced to allow CNBr cleavage of TMD-KS from CT^{LLP2} during protein purification
CT^{LLP2} (residues 737-788)	GGERDRDRSIRLVNGSLALIWDDLRSLSLFSYHRLRDLLLIVTRIVELLGRR
CT^{LLP2} C789	GGERDRDRSIRLVNGSLALIWDDLRSLSLFSYHRLRDLLLIVTRIVELLGRR**C**
CT2-tmd	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRGGERD RDRSIR**S**VNGSLA**S**IW**AA**LR**A**LSLFSYHRLRDLLLIVTRIVELLGRR Note: L748S-L755S-D758A-D759A-S762A mutations were introduced
CT2-tmd S764C	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRGGERD RDRSIRSVNGSLASIWAALRAL**C**LFSYHRLRDLLLIVTRIVELLGRR Note: the native Cys was reintroduced to the TMD-CT^{LLP2} CT2-tmd
MPER-TMD L660C A667C (residues 660-710)	**C**LELDKW**C**SLWNWFDITNWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQ Note: the MPER was locked or unlocked controlling on the oxidation state of the Cys
OG-label constructs	
His_6-tagged TMD-CT^{LLP2}	NWLWYIRIFIIIVGSLIGLRIVFAVLSLVNRVRQGYSPLSFQTHLPTPRGGERD RDRSIRLVNGSLALIWDDLRSLSLFSYHRLRDLLLIVTRIVELLGRR**HHHHH H**
His_6-tagged CT^{LLP2}	GGERDRDRSIRLVNGSLALIWDDLRSLSLFSYHRLRDLLLIVTRIVELLGRR **HHHHHH**

Supplementary Table 6 | DNA sequence of the synthetic genes and primers used in this study.

	Synthetic genes
#1	AATTGGCTGTGGTATATTCGTATTTTCATTATCATAGTGGGAAGCCTGATTGGCCTGCGTATTG TGTTTGCGGTTCTGAGCCTGGTGAATCGGGTGCGTCAGGGTTATAGCCCGCTGAGCTTTCAGAC CCATCTGCCTACCCCGCGTGGCCCGGATCGTCCGGAAGGGATAGAAGAAGAGGGCGGAGAAC GCGACCGTGATCGTAGCATTCGTCTGGTTAATGGCAGCCTGGCACTGATTTGGGATGATCTGCG TAGCCTGTGCCTTTTTAGCTATCATCGTCTGCGTGATTTACTGCTGATTGTGACCCGTATCGTGG AACTGCTGGGCCGTCGT
#2	AATTGGCTGTGGTATATTCGTATTTTCATTATCATAGTGGGAAGCCTGATTGGCCTGCGTATTG TGTTTGCGGTTCTGAGCCTGGTGAATCGGGTGCGTCAGGGTTATAGCCCGCTGAGCTTTCAGAC CCATCTGCCTACCCCGCGTGGCTGCGAACGCGACCGTGATCGTAGCATTCGTCTGGTTAATGGC AGCCTGGCACTGATTTGGGATGATCTGCGTAGCCTGAGCCTTTTTAGCTATCATCGTCTGCGTG ATTTACTGCTGATTGTGACCCGTATCGTGGAACTGCTGGGCCGTCGT
#3	AATTGGCTGTGGTATATTCGTATTTTCATTATCATAGTGGGAAGCCTGATTGGCCTGCGTATTG TGTTTGCGGTTCTGAGCCTGGTGAATCGGGTGCGTCAGGGTTATAGCCCGCTGAGCTTTCAGAC CCATCTGCCTACCCCGCGTATGGGCGGAGAACGCGACCGTGATCGTAGCATTCGTCTGGTTAA TGGCAGCCTGGCACTGATTTGGGATGATCTGCGTAGCCTGAGCCTTTTTAGCTATCATCGTCTG CGTGATTTACTGCTGATTGTGACCCGTATCGTGGAACTGCTGGGCCGTCGT
#4	AATTGGCTGTGGTATATTCGTATTTTCATTATCATAGTGGGAAGCCTGATTGGCCTGCGTATTG TGTTTGCGGTTCTGAGCCTGGTGAATCGGGTGCGTCAGGGTTATAGCCCGCTGAGCTTTCAGAC CCATCTGCCTACCCCGCGTGGCGGAGAACGCGACCGTGATCGTAGCATTCGTTCTGTTAATGG CAGCCTGGCATCTATTTGGGCGGCGCTGCGTGCGCTGAGCCTTTTTAGCTATCATCGTCTGCGT GATTTACTGCTGATTGTGACCCGTATCGTGGAACTGCTGGGCCGTCGT
#5	AATTGGCTGTGGTATATTCGTATTTTCATTATCATAGTGGGAAGCCTGATTGGCCTGCGTATTG TGTTTGCGGTTCTGAGCCTGGTGAATCGGGTGCGTCAGGGTTATAGCCCGCTGAGCTTTCAGAC CCATCTGCCTACCCCGCGTGGCGGAGAACGCGACCGTGATCGTAGCATTCGTTCTGTTAATGG CAGCCTGGCATCTATTTGGGCGGCGCTGCGTGCGCTGTGCCTTTTTAGCTATCATCGTCTGCGT GATTTACTGCTGATTGTGACCCGTATCGTGGAACTGCTGGGCCGTCGT
#6	TGTCTGGAGCTGGACAAATGGTGTAGCCTGTGGAATTGGTTTGATATTACCAACTGGCTGTGGT ATATTCGTATTTTCATCATTATTGTGGGCAGCCTGATTGGCCTGCGTATTGTGTTTGCGGTGCTG AGCCTGGTGAATCGTGTGCGTCAG
#7	AATTGGCTGTGGTATATTCGTATTTTCATTATCATAGTGGGAAGCCTGATTGGCCTGCGTATTG TGTTTGCGGTTCTGAGCCTGGTGAATCGGGTGCGTCAGGGTTATAGCCCGCTGAGCTTTCAGAC CCATCTGCCTACCCCGCGTGGCGGAGAACGCGACCGTGATCGTAGCATTCGTCTGGTTAATGG CAGCCTGGCACTGATTTGGGATGATCTGCGTAGCCTGAGCCTTTTTAGCTATCATCGTCTGCGT GATTTACTGCTGATTGTGACCCGTATCGTGGAACTGCTGGGCCGTCGTGGCCGTCGTCATCACC ATCACCATCAC
#8	GGCGGAGAACGCGACCGTGATCGTAGCATTCGTCTGGTTAATGGCAGCCTGGCACTGATTTGG GATGATCTGCGTAGCCTGAGCCTTTTTAGCTATCATCGTCTGCGTGATTTACTGCTGATTGTGA CCCGTATCGTGGAACTGCTGGGCCGTCGTCATCACCATCACCATCAC
	Primers
#1	GATCTGCGTAGCCTGAGCCTTTTTAGCTATCATCGTCTG
#2	CAGACGATGATAGCTAAAAAGGCTCAGGCTACGCAGATC
#3	CCGGATCGTCCGGAA GGCGGAGAACGCGA
#4	TCGCGTTCTCCGCCTTCCGGACGATCCGG
#5	CTTTCAGACCCATCTGCCTACCCCGCGTGGCGGAGAACGCGACCGTGATCGTAGCATTC
#6	GAATGCTACGATCACGGTCGCGTTCTCCGCCACGCGGGGTAGGCAGATGGGTCTGAAAG
#7	AACTGCTGGGCCGTCGTTGCTAAGGATCCGAATTCGAGC
#8	GCTCGAATTCGGATCCTTAGCAACGACGGCCCAGCAGTT

Supplementary Table 7 | NMR acquisition parameters.

Type of experiment	Spectral widths and chemical shift evolution times	# of scans	Inter-scan delay (s)	Duration of the experiment	Notes
Backbone assignment					
~ 0.7 mM (^{15}N, ^{13}C, 85% ^{2}H)-labeled TMD-CT^{LLP2}; magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 62.5 ms (^{15}N)	40	1.2	3 h	
3D TROSY HNCO	10800 Hz, 95.0 ms ($^{1}H^{N}$) 1600 Hz, 31.3 ms (^{15}N) 1700 Hz, 13.5 ms ($^{13}C'$)	32	1.2	2 d 10 h	
3D TROSY HNCA	10800 Hz, 95.0 ms ($^{1}H^{N}$) 1600 Hz, 31.3 ms (^{15}N) 5300 Hz, 5.3 ms ($^{13}C^{\alpha}$)	64	1.2	5 d 22 h	
~ 0.7 mM (^{15}N, ^{13}C, 85% ^{2}H)-labeled TMD-CT^{LLP2}; magnetic field: 17.6 T					
3D TROSY HN(CA)CO	10500 Hz, 97.0 ms ($^{1}H^{N}$) 1900 Hz, 30.0 ms (^{15}N) 1950 Hz, 16.4 ms ($^{13}C'$)	24	1.2	3 d	
3D TROSY HN(CO)CA	10500 Hz, 49.0 ms ($^{1}H^{N}$) 1800 Hz, 27.4 ms (^{15}N) 5000 Hz, 6.6 ms ($^{13}C^{\alpha}$)	20	1.2	2 d 3 h	
3D TROSY HNCACB	10500 Hz, 97.0 ms ($^{1}H^{N}$) 1900 Hz, 26.3 ms (^{15}N) 4700 Hz, 4.3 ms ($^{13}C^{\alpha/\beta}$)	32	1.3	2 d 8 h	
~ 0.8 mM (^{15}N, ^{2}H)-labeled TMD-CT^{LLP2}; magnetic field: 17.6 T					
3D ^{15}N-edited NOESY-TROSY-HSQC	10500 Hz, 95.0 ms ($^{1}H^{N}$) 1900 Hz, 28.9 ms (^{15}N) 2250 Hz, 20.0 ms (^{1}H)	20	1.2	3 d 15 h	τ_{mix} = 180 ms
Side-chain and NOE assignment					
~ 1.0 mM (^{15}N, ^{13}C)-labeled TMD-CT^{LLP2}; magnetic field: 18.8 T					
2D ^{1}H-^{15}N TROSY HSQC	11150 Hz, 92.0 ms ($^{1}H^{N}$) 1900 Hz, 134.7 ms (^{15}N)	32	1.2	6 h	
2D ^{1}H-^{13}C HSQC	11150 Hz, 46.0 ms ($^{1}H^{N}$) 4400 Hz, 27.3 ms (^{13}C)	32	1.2	3 h	
3D ^{15}N-edited NOESY-TROSY-HSQC	11150 Hz, 92.0 ms ($^{1}H^{N}$) 1900 Hz, 32.1 ms (^{15}N) 9600 Hz, 10.4 ms (^{1}H)	28	1.2	11 d 16 h	τ_{mix} = 80 ms
3D ^{13}C-edited NOESY	11150 Hz, 92.0 ms ($^{1}H^{N}$) 4400 Hz, 19.3 ms (^{13}C) 9600 Hz, 10.4 ms (^{1}H)	16	1.0	8 d 4 h	τ_{mix} = 150 ms

Supplementary Table 7 (continued)

Type of experiment	Spectral widths and chemical shift evolution times	# of scans	Inter-scan delay (s)	Duration of the experiment	Notes
~ 0.5 mM (^{15}N, ^{2}H)-labeled TMD-CT^{LLP2} (control sample); magnetic field: 21.1 T					
2D ^{1}H-^{15}N TROSY HSQC	12600 Hz, 81.0 ms ($^{1}H^{N}$) 1900 Hz, 73.1 ms (^{15}N)	8	1.3	1 h	
3D ^{15}N-edited NOESY-TROSY-HSQC	12600 Hz, 81.0 ms ($^{1}H^{N}$) 1900 Hz, 26.1 ms (^{15}N) 10800 Hz, 8.3 ms (^{1}H)	24	1.2	7 d 22 h	τ_{mix} = 200 ms
~ 0.5 mM (^{15}N, ^{2}H)-labeled TMD-CT^{LLP2} mixed with ~ 0.5 mM (^{13}C)-labeled TMD-CT^{LLP2} (~ 1:1); magnetic field: 21.1 T					
2D ^{1}H-^{15}N TROSY HSQC	12600 Hz, 81.0 ms ($^{1}H^{N}$) 1900 Hz, 73.1 ms (^{15}N)	8	1.3	1 h	
2D ^{1}H-^{13}C HSQC	12600 Hz, 40.5 ms ($^{1}H^{N}$) 4100 Hz, 27.8 ms (^{13}C)	8	1.3	1h	
3D ^{15}N-edited NOESY-TROSY-HSQC	12600 Hz, 81.0 ms ($^{1}H^{N}$) 1900 Hz, 26.1 ms (^{15}N) 10800 Hz, 8.3 ms (^{1}H)	24	1.2	7 d 22 h	τ_{mix} = 200 ms
~ 0.5 mM (^{15}N, ^{2}H)-labeled TMD-CT^{LLP2} mixed with ~ 0.5 mM (^{13}C)-labeled TMD-CT^{LLP2} (~ 1:1); magnetic field: 18.8 T					
2D ^{1}H-^{15}N TROSY HSQC	11150 Hz, 92.0 ms ($^{1}H^{N}$) 1900 Hz, 79.0 ms (^{15}N)	32	1.2	4 h	
2D ^{1}H-^{13}C HSQC	11150 Hz, 46.0 ms ($^{1}H^{N}$) 4400 Hz, 27.3 ms (^{13}C)	32	1.2	3 h	
3D J_{CH}-modulated NOESY	11150 Hz, 92.0 ms ($^{1}H^{N}$) 1900 Hz, 26.3 ms (^{15}N) 9600 Hz, 7.3 ms (^{1}H)	24	1.2	6 d 9 h (each spectrum)	τ_{mix} = 200 ms; two interleaved spectra were recorded (J_{CH} = 0 and 8 ms)
PPT					
~ 0.5 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2}; magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 56.3 ms (^{15}N)	56	3.5	11 h (each spectrum)	[Gd-DOTA] = 0, 1.0, 2.0, 4.0, 6.0, 8.0, 10.0, 15.0, 20.0 mM
~ 0.9 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2}; magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 56.3 ms (^{15}N)	32	3.5	6 h (each spectrum)	[16-DSA] = 0, 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, 4.2 mM

Supplementary Table 7 (continued)

Type of experiment	Spectral widths and chemical shift evolution times	# of scans	Inter-scan delay (s)	Duration of the experiment	Notes
Inter-protomer PREs					
~ 0.3 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2} (control sample); magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 93.8 ms (^{15}N)	56	3.5	17 h (each spectrum)	[Sodium ascorbate] = 0, 20.0 mM
~ 0.4 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2} mixed with ~ 0.4 mM unlabeled TMD-CT^{LLP2} G738C (~ 1:1); magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 93.8 ms (^{15}N)	136	3.5	1 d 18 h (each spectrum)	[Sodium ascorbate] = 0, 20.0 mM
~ 0.4 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2} mixed with ~ 0.4 mM unlabeled TMD-CT^{LLP2} S764C (~ 1:1); magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 93.8 ms (^{15}N)	196	3.5	1 d 6 h (each spectrum)	[Sodium ascorbate] = 0, 20.0 mM
~ 0.4 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2} mixed with ~ 0.4 mM unlabeled TMD-CT^{LLP2} C789 (~ 1:1); magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 71.9 ms (^{15}N)	136	3.5	1 d 8 h (each spectrum)	[Sodium ascorbate] = 0, 20.0 mM
~ 0.5 mM (^{15}N, 85% ^{2}H)-labeled TMD-KS mixed with ~ 0.5 mM unlabeled CT^{LLP2} C789 (~ 1:1); magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 94.0 ms (^{15}N)	64	3.5	20 h (each spectrum)	[Sodium ascorbate] = 0, 20.0 mM
~ 0.15 mM (^{15}N, 85% ^{2}H)-labeled CT^{LLP2} mixed with ~ 0.35 mM unlabeled CT^{LLP2} C789 and ~ 0.5 mM unlabeled TMD-KS (~ 1:2:3); magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 94.0 ms (^{15}N)	64	3.5	20 h (each spectrum)	[Sodium ascorbate] = 0, 20.0 mM
~ 0.5 mM (15N, 85% 2H)-labeled CT2-tmd mixed with ~ 0.5 mM unlabeled CT2-tmd S764C (~ 1:1); magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 94.0 ms (^{15}N)	104	3.5	1 d 8 h (each spectrum)	[Sodium ascorbate] = 0, 20.0 mM
H-D exchange					
~ 0.8 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2}; magnetic field: 14.1 T					
2D ^{1}H-^{15}N TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 75.0 ms (^{15}N)	32	1.2	3 h (each spectrum)	

Supplementary Table 7 (continued)

Type of experiment	Spectral widths and chemical shift evolution times	# of scans	Inter-scan delay (s)	Duration of the experiment	Notes
^{15}N backbone dynamics					
~ 0.5 mM (^{15}N, 85% ^{2}H)-labeled TMD-CT^{LLP2}; magnetic field: 14.1 T					
2D ^{1}H-^{15}N T_1 TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 75.0 ms (^{15}N)	40	1.5	6 h (average time of each spectrum)	Relaxation delays: 10, 50, 100, 200, 300, 600, 800, 1000 ms
2D ^{1}H-^{15}N T_2 TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1600 Hz, 75.0 ms (^{15}N)	40	1.5	5 h (each spectrum)	Relaxation delays: 6.4, 10, 20, 30, 40, 50, 64, 80 ms
~ 0.6 mM (^{15}N, 85% ^{2}H)-labeled Unlocked MPER-TMD; magnetic field: 14.1 T					
2D ^{1}H-^{15}N T_1 TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1500 Hz, 68.0 ms (^{15}N)	64	1.5	6 h (average time of each spectrum)	Relaxation delays: 10, 50, 100, 200, 300, 600, 800, 1000 ms
2D ^{1}H-^{15}N T_2 TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1500 Hz, 68.0 ms (^{15}N)	80	1.5	9 h (each spectrum)	Relaxation delays: 6.4, 10, 20, 30, 40, 50, 64, 80 ms
~ 0.6 mM (^{15}N, 85% ^{2}H)-labeled Locked MPER-TMD; magnetic field: 14.1 T					
2D ^{1}H-^{15}N T_1 TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1500 Hz, 68.0 ms (^{15}N)	64	1.5	6 h (average time of each spectrum)	Relaxation delays: 10, 50, 100, 200, 300, 600, 800, 1000 ms
2D ^{1}H-^{15}N T_2 TROSY HSQC	9600 Hz, 106.0 ms ($^{1}H^{N}$) 1500 Hz, 68.0 ms (^{15}N)	80	1.5	9 h (each spectrum)	Relaxation delays: 6.4, 10, 20, 30, 40, 50, 64, 80 ms

Supplementary Table 8 | Residue specific PRE_{amp} of TMD-CT^{LLP2} in q = 0.5 bicelles. The values were determined by fitting I/I_0 vs. [paramagnetic agent] to Eq. 1. PRE_{amp} of non-proline residues not shown were discarded because of signal overlap or poor data fitting.

Residue	PRE_{amp} (Gd-DOTA)	PRE_{amp} (16-DSA)	Residue	PRE_{amp} (Gd-DOTA)	PRE_{amp} (16-DSA)
679	0.914±0.036	0.871±0.011	716	1.011±0.019	0.883±0.023
680	0.946±0.063	-	717	0.987±0.026	0.795±0.012
681	0.943±0.052	0.9015±0.0080	718	0.934±0.025	0.772±0.022
682	-	0.912±0.024	719	0.964±0.043	-
684	0.857±0.054	0.930±0.012	721	0.992±0.015	0.733±0.017
685	-	0.966±0.020	723	0.951±0.009	-
686	0.763±0.057	0.9869±0.0093	725	0.974±0.020	-0.208±0.027 (*)
688	0.722±0.029	0.974±0.023	737	0.961±0.040	-
689	-	0.987±0.013	738	0.999±0.037	-2.66±0.44 (*)
690	0.654±0.058	0.9973±0.0082	739	0.977±0.023	-0.636±0.021 (*)
693	0.653±0.063	0.980±0.021	744	0.959±0.018	0.380±0.016 (*)
694	0.652±0.046	0.9930±0.0047	745	0.989±0.024	0.793±0.031
695	0.662±0.041	0.990±0.013	746	-	0.856±0.018
696	0.661±0.045	0.970±0.019	747	0.969±0.023	0.871±0.024
698	0.655±0.057	0.975±0.022	748	0.942±0.035	0.929±0.021
699	0.683±0.025	0.981±0.013	750	0.946±0.053	0.957±0.014
700	0.709±0.027	0.966±0.020	751	0.949±0.021	0.907±0.028
701	-	0.987±0.013	753	0.910±0.033	0.960±0.015
702	0.771±0.061	0.981±0.012	754	0.907±0.028	0.921±0.022
703	0.864±0.043	0.950±0.022	755	0.933±0.023	0.946±0.023
704	0.879±0.057	0.944±0.021	756	0.871±0.023	0.955±0.020
705	0.907±0.066	0.945±0.019	758	0.870±0.036	0.925±0.028
706	0.908±0.055	0.904±0.015	759	-	0.923±0.020
707	-	0.891±0.010	762	0.909±0.043	0.949±0.022
708	0.920±0.055	0.889±0.016	763	0.932±0.054	-
709	-	0.895±0.032	764	0.987±0.044	0.950±0.019
710	0.983±0.024	0.812±0.011	770	0.981±0.035	0.963±0.011
711	1.008±0.019	0.844±0.016	771	0.892±0.075	0.918±0.023
712	0.950±0.014	0.8611±0.0099	772	0.930±0.079	0.938±0.026
713	0.981±0.014	0.858±0.015	773	0.946±0.055	0.865±0.022
715	0.977±0.018	0.909±0.017	774	0.891±0.097	0.922±0.020

Supplementary Table 8 (continued)

775	0.869±0.060	0.967±0.020	782	0.923±0.088	0.931±0.021
776	0.901±0.034	0.9264±0.0092	784	0.953±0.075	0.835±0.015
777	0.922±0.090	0.908±0.017	785	0.925±0.065	0.8591±0.0079
778	0.899±0.088	0.959±0.024	786	0.960±0.071	0.877±0.020
779	0.92±0.10	0.926±0.027	787	1.100±0.057	0.820±0.011
780	0.938±0.074	0.870±0.020	788	1.033±0.044	0.754±0.022

(*) These PRE_{amp} were not used in the PPT analysis since their values were dramatically dominated by the fast timescale of the highly flexible loop.

Supplementary Table 9 | Residue-specific k_{ex} of the TMD-CT^{LLP2} in bicelles. k_{ex} of non-proline residues not shown were discarded because of signal overlap or poor data fitting. '*' indicates residues with k_{ex} too fast to be measured. The colored boxes 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).

Residue (H^N)	k_{ex} (Hz)	τ_{ex}		Residue (H^N)	k_{ex} (Hz)	τ_{ex}	
679	*	*		738	*	*	
681	*	*		739	*	*	
684	$(2.3\pm1.1)\cdot10^{-4}$	1.2±0.6 hours		741	*	*	
685	$(6.1\pm1.4)\cdot10^{-5}$	4.6±1.1 hours		742	*	*	
686	$(6.4\pm1.6)\cdot10^{-6}$	1.8±0.4 days		744	*	*	
687	$(1.6\pm0.4)\cdot10^{-5}$	17.1±4.0 hours		745	$(1.9\pm0.1)\cdot10^{-5}$	14.3±7.8 hours	
688	$(3.7\pm1.5)\cdot10^{-6}$	3.1±1.2 days		747	*	*	
689	$(2.4\pm0.9)\cdot10^{-6}$	4.8±1.7 days		748	$(7.3\pm3.8)\cdot10^{-5}$	3.8±2.0 hours	
690	$(5.2\pm0.5)\cdot10^{-5}$	5.3±0.6 hours		750	*	*	
693	$(1.0\pm0.3)\cdot10^{-4}$	2.7±0.8 hours		751	$(8.3\pm3.9)\cdot10^{-5}$	3.4±1.6 hours	
694	$(9.9\pm6.1)\cdot10^{-5}$	2.8±1.8 hours		753	$(1.0\pm0.6)\cdot10^{-4}$	2.7±1.6 hours	
695	*	*		754	*	*	
698	$(1.7\pm0.2)\cdot10^{-4}$	1.6±0.2 hours		755	$(2.9\pm1.5)\cdot10^{-5}$	9.7±5.1 hours	
699	$(1.0\pm0.7)\cdot10^{-4}$	2.8±1.9 hours		756	$(1.5\pm1.1)\cdot10^{-4}$	1.8±1.3 hours	
700	*	*		758	$(1.1\pm0.4)\cdot10^{-4}$	2.5±0.9 hours	
701	$(1.5\pm0.2)\cdot10^{-4}$	1.9±0.2 hours		760	$(3.9\pm0.5)\cdot10^{-4}$	7.2±1.0 hours	
702	$(2.2\pm0.6)\cdot10^{-4}$	1.3±0.3 hours		762	*	*	
703	*	*		763	*	*	
704	$(2.3\pm0.6)\cdot10^{-5}$	12.3±3.4 hours		764	$(9.6\pm4.1)\cdot10^{-5}$	2.9±1.2 hours	
705	$(2.6\pm1.6)\cdot10^{-5}$	10.6±6.5 hours		770	*	*	
706	$(1.2\pm0.5)\cdot10^{-4}$	2.3±1.0 hours		772	$(2.1\pm1.1)\cdot10^{-4}$	1.3±0.7 hours	
707	*	*		773	$(4.0\pm1.4)\cdot10^{-5}$	7.0±2.4 hours	
708	*	*		774	$(3.1\pm2.0)\cdot10^{-5}$	8.9±5.6 hours	
710	*	*		775	$(1.3\pm0.5)\cdot10^{-4}$	2.1±0.8 hours	
711	*	*		777	$(2.1\pm0.9)\cdot10^{-5}$	13.4±5.8 hours	
712	*	*		778	$(2.3\pm0.5)\cdot10^{-4}$	1.2±0.3 hours	
713	*	*		779	*	*	
715	*	*		780	*	*	
716	*	*		781	$(1.7\pm0.7)\cdot10^{-4}$	1.6±0.7 hours	
717	*	*		782	$(1.1\pm0.1)\cdot10^{-4}$	2.6±0.4 hours	
718	*	*		783	*	*	
719	*	*		784	*	*	
721	*	*		785	*	*	
723	*	*		786	*	*	
725	*	*		787	*	*	
737	$(2.4\pm1.2)\cdot10^{-4}$	1.2±0.6 hours		788	*	*	

Supplementary Table 10 | Plane restraints used for the TMD-CT^{LLP2} and MPER-TMD-CT^{LLP2} structure calculation. r_Z = 0 corresponds to the bicelle center. Plane restraints for the MPER and CT are shown in light blue and green, respectively; those for the TMD structure (internal calibration) are shown in red.

Residue	TMD-CT^{LLP2} r_Z (Å)	MPER-TMD-CT^{LLP2} r_Z (Å)	Residue	TMD-CT^{LLP2} r_Z (Å)	MPER-TMD-CT^{LLP2} r_Z (Å)
678	-	17.8±1.0	709	-22.8±1.0	-22.8±1.0
679	24.1±1.0	18.4±1.0	710	-23.8±1.0	-23.8±1.0
680	23.5±1.0	17.5±1.0	712	-23.3±1.0	-23.3±1.0
681	21.5±1.0	18.1±1.0	713	-23.5±1.0	-23.5±1.0
682	19.7±1.0	17.8±1.0	715	-20.3±1.0	-20.3±1.0
683	19.1±1.0	16.5±1.0	716	-22.0±1.0	-22.0±1.0
684	17.7±1.0	16.7±1.0	746	-23.6±1.0	-23.6±1.0
685	15.5±1.0	15.3±1.0	747	-22.7±1.0	-22.7±1.0
686	14.0±1.0	13.5±1.0	748	-18.9±1.0	-18.9±1.0
687	13.0±1.0	12.6±1.0	751	-20.4±1.0	-20.4±1.0
688	11.2±1.0	-	753	-18.7±1.0	-18.7±1.0
689	9.3±1.0	-	754	-19.0±1.0	-19.0±1.0
690	8.5±1.0	-	755	-17.5±1.0	-17.5±1.0
691	6.7±1.0	-	756	-16.8±1.0	-16.8±1.0
692	4.3±1.0	-	758	-18.0±1.0	-18.0±1.0
693	3.0±1.0	-	759	-19.4±1.0	-19.4±1.0
694	1.8±1.0	-	762	-18.0±1.0	-18.0±1.0
695	-0.4±1.0	-0.4±1.0	771	-18.8±1.0	-18.8±1.0
696	-2.4±1.0	-2.4±1.0	772	-19.2±1.0	-19.2±1.0
697	-4.3±1.0	-4.3±1.0	773	-23.1±1.0	-23.1±1.0
698	-6.6±1.0	-6.6±1.0	774	-18.6±1.0	-18.6±1.0
699	-8.4±1.0	-8.4±1.0	775	-16.8±1.0	-16.8±1.0
700	-10.0±1.0	-10.0±1.0	776	-18.7±1.0	-18.7±1.0
701	-11.7±1.0	-11.7±1.0	777	-20.0±1.0	-20.0±1.0
702	-13.1±1.0	-13.1±1.0	778	-18.1±1.0	-18.1±1.0
703	-15.2±1.0	-15.2±1.0	779	-19.2±1.0	-19.2±1.0
704	-17.6±1.0	-17.6±1.0	780	-22.7±1.0	-22.7±1.0
705	-18.2±1.0	-18.2±1.0	782	-19.2±1.0	-19.2±1.0
706	-18.9±1.0	-18.9±1.0	785	-23.4±1.0	-23.4±1.0
707	-21.3±1.0	-21.3±1.0	786	-22.3±1.0	-22.3±1.0
708	-23.2±1.0	-23.2±1.0			

Supplementary Table 11 | TMD-CT^{LLP2} NMR structure calculation and refinement statistics.

NMR distance, dihedral and plane restraints [a]	**TMD (677-710)**	**TMD-CT^{LLP2} (711-788)**
Distance restraints from NOE [b]	510	585
Short-range intramolecular (|*i-j*|≤4)	459	498
Long-range intramolecular (|*i-j*|≥5)	0	24
Intermolecular	51	63
Total plane restraints [c]	96	84
Total dihedral angle restraints [d]	480	
Φ (TALOS)	240	
ψ (TALOS)	240	
Structure statistics [e]		
Violations (mean ± s.d.)		
Distance restraints (Å)	0.131±0.006	
Plane restraints (Å)	0.063±0.026	
Dihedral angle restraints (°)	0.686±0.065	
Deviations from idealized geometry		
Bond lengths (Å)	0.008±0.000	
Bond angles (°)	0.917±0.015	
Impropers (°)	0.621±0.022	
Average pairwise r.m.s. deviation (Å) [f]		
Heavy	1.763	
Backbone	1.296	

[a] The numbers of restraints are summed over all three subunits.

[b] For residues 711-788, all NOE restraints were obtained in the current study. For residues 679-710, the NOE restraints were taken from our earlier NMR study of the TMD[3].

[c] For residues 711-788, plane restraints were derived using PRE_{amp} values obtained in the current study. For residues 679-710, plane restraints were directly taken from our earlier NMR study of the TMD (internal calibration)[4].

[d] Backbone φ and ψ restraints and their respective uncertainties were obtained from the "GOOD" dihedrals generated by *TALOS+* based on the backbone chemical shift values.

[e] Statistics are calculated and averaged over an ensemble of the 15 lowest energy structures out of 150 calculated structures.

[f] The precision of the atomic coordinates is defined as the average r.m.s. difference between the 15 final structures and their mean coordinates (flexible residues 711-740 have been omitted from the comparison).

Supplementary Table 12 | MPER-TMD-CT^{LLP2} NMR model calculation and refinement statistics.

NMR distance, dihedral and plane restraints [a]	**MPER-TMD (660-693)**	**TMD (694-710)**	**TMD-CT^{LLP2} (711-788)**
Distance restraints from NOE [b]	333	429	585
Short-range intramolecular ($\|i-j\|\leq 4$)	303	387	498
Long-range intramolecular ($\|i-j\|\geq 5$)	6	0	24
Intermolecular	24	42	63
Total plane restraints [c]	30	48	84
Total dihedral angle restraints [d]		540	
Φ (TALOS)		270	
ψ (TALOS)		270	
Structure statistics [e]			
Violations (mean ± s.d.)			
Distance restraints (Å)		0.118±0.003	
Plane restraints (Å)		0.074±0.020	
Dihedral angle restraints (°)		0.728±0.038	
Deviations from idealized geometry			
Bond lengths (Å)		0.008±0.000	
Bond angles (°)		0.886±0.018	
Impropers (°)		0.606±0.024	
Average pairwise r.m.s. deviation (Å) [f]			
Heavy		1.240	
Backbone		0.941	

[a] The numbers of restraints are summed over all three subunits.

[b] For residues 711-788, all NOE restraints were obtained in the current study. For the MPER and first half of the TMD (residues 660-693), the NOE restraints were taken from our earlier NMR study of the MPER-TMD[5]; for the second half of the TMD (residues 694-710), the NOE restraints were taken from our earlier NMR study of the TMD[3].

[c] For the residues 711-788, plane restraints were derived using the PRE_{amp} values obtained in the current study. For the MPER and first half of the TMD (residues 660-693), plane restraints were derived using the PRE_{amp} values taken from our earlier NMR study of the MPER-TMD[5]; for the second half of the TMD (residues 694-710), plane restraints were directly taken from our earlier NMR study of the TMD (internal calibration)[4].

[d] Backbone φ and ψ restraints and their respective uncertainties were obtained from the "GOOD" dihedrals generated by *TALOS+* based on the backbone chemical shift values. For residues 694-788, all dihedral angle restraints were obtained in the current study; for residues 660-693, all dihedral angle restraints were taken from our earlier NMR study of the MPER-TMD[5].

[e] Statistics are calculated and averaged over an ensemble of the 15 lowest energy structures out of 150 calculated structures.

[f] The precision of the atomic coordinates is defined as the average r.m.s. difference between the 15 final structures and their mean coordinates (flexible residues 711-740 have been omitted from the comparison).

Supplementary Table 13 | DNA sequence of the synthetic genes used to generate the expression constructs of antibodies PG9, PG16, and PGT145.

Synthetic genes	
PG9	See reference [11].
PG16	See reference [11].
PGT145	*Heavy chain*: GGGGACAAGTTTGTACAAAAAAGCAGGCTGAGCTCTCTAGAGAATTCGCCACCATGG GCTGGTCCTGCATCATCCTGTTTCTGGTGGCCACAGCCACCGGCGTGCACAGCCAGGC CTCTACCATGGACTGGATCTGGCGGATTCTGTTCCTGGTGGCTGCCGCCACAAGCGCC CATTCTCAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAACCCGGCAGCAGC GTGAAGGTGTCCTGCAAGGCCAGCGGCAACAGCTTCAGCAACCACGACGTGCACTGG GTGCGCCAGGCTACAGGACAGGGCCTGGAATGGATGGGCTGGATGAGCCACGAGGGC GACAAGACAGGACTGGCCCAGAAATTCCAGGGCAGAGTGACCATCACCCGGGACAGC GGAGCCAGCACCGTGTACATGGAACTGAGAGGCCTGACCGCCGACGACACCGCCATC TACTACTGTCTGACCGGCAGCAAGCACCGGCTGCGGGACTACTTCCTGTACAACGAGT ACGGCCCCAACTACGAGGAATGGGGCGACTACCTGGCCACCCTGGATGTGTGGGGAC ACGGAACAGCCGTGACCGTGTCTAGCGCCAGCACAAAGGGCCCCAGCGTGTTCCCTCT GGCCCCTAGCAGCAAGAGCACATCTGGCGGCACAGCCGCCCTGGGCTGCCTCGTGAA GGATTACTTCCCCGAACCCGTGACAGTGTCCTGGAACTCTGGCGCCCTGACCAGCGGA GTGCACACCTTTCCAGCCGTGCTGCAGAGCAGCGGCCTGTACTCTCTGAGCAGCGTCG TGACTGTGCCCAGCAGCAGCCTGGGAACCCAGACCTACATCTGCAACGTGAACCACA AGCCCAGCAACACCAAGGTGGACAAGAAGGTGGAACCCAAGAGCTGCGACTGATGAG GATCCGCTAGCGGTACCACCCAGCTTTCTTGTACAAAGTGGTCCCC
	Light chain: GGGGACAAGTTTGTACAAAAAAGCAGGCTGAGCTCTCTAGAGAATTCGCCACCATGG GCTGGTCCTGCATCATCCTGTTTCTGGTGGCCACAGCCACCGGCGTGCACAGCATGAG ACTGCCTGCTCAGCTGCTGGGCCTGCTGATGCTGTGGGTGTCAGGATCTGGCGCCGAG GTCGTGATTACCCAGAGCCCCCTGTTCCTGCCTGTGACACCTGGCGAAGCCGCCAGCC TGAGCTGCAAGTGTAGCCACAGCCTGCAGCACTCCACCGGCGCCAATTACCTGGCCTG GTATCTGCAGCGGCCTGGCCAGACACCCAGACTGCTGATTCACCTGGCCACCCACAGA GCCTCTGGCGTGCCCGATAGATTTTCCGGCTCTGGCAGCGGCACCGACTTCACCCTGA AGATCAGCCGGGTGGAAAGCGACGACGTGGGCACCTACTACTGTATGCAGGGCCTGC ACAGCCCCTGGACCTTTGGCCAGGGCACCAAGGTGGAAATCAAGCGGACAGTGGCCG CTCCCAGCGTGTTCATCTTCCCACCTAGCGACGAGCAGCTGAAGTCCGGCACAGCCTC TGTCGTGTGCCTGCTGAACAACTTCTACCCCCGCGAGGCCAAGGTGCAGTGGAAGGTG GACAATGCCCTGCAGAGCGGCAACAGCCAGGAAAGCGTGACCGAGCAGGACAGCAA GGACTCCACCTACAGCCTGTCCAGCACCCTGACCCTGAGCAAGGCCGACTACGAGAA GCACAAGGTGTACGCCTGCGAAGTGACCCACCAGGGCCTGTCTAGCCCCGTGACCAA GAGCTTCAACCGGGGCGAGTGCTGATGAGGATCCGCTAGCGGTACCACCCAGCTTTCT TGTACAAAGTGGTCCCC

Supplementary References

1 Chen, W. *et al.* The Unusual Transmembrane Partition of the Hexameric Channel of the Hepatitis C Virus. *Structure* **26**, 627-634 e624 (2018).

2 Shen, Y. & Bax, A. Identification of helix capping and b-turn motifs from NMR chemical shifts. *Journal of biomolecular NMR* **52**, 211-232 (2012).

3 Dev, J. *et al.* Structural basis for membrane anchoring of HIV-1 envelope spike. *Science* **353**, 172-175 (2016).

4 Piai, A., Dev, J., Fu, Q. & Chou, J. J. Stability and Water Accessibility of the Trimeric Membrane Anchors of the HIV-1 Envelope Spikes. *Journal of the American Chemical Society* **139**, 18432-18435 (2017).

5 Fu, Q. *et al.* Structure of the membrane proximal external region of HIV-1 envelope glycoprotein. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E8892-E8899 (2018).

6 Gristick, H. B. *et al.* Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. *Nature structural & molecular biology* **23**, 906-915 (2016).

7 Liu, J., Bartesaghi, A., Borgnia, M. J., Sapiro, G. & Subramaniam, S. Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455**, 109-113 (2008).

8 Munoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. *Nat Struct Biol* **1**, 399-409 (1994).

9 Raghava, G. P. S. Protein secondary structure prediction using nearest neighbor and neural network approach. (2000).

10 Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202 (1999).

11 Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**, 285-289 (2009).