Conformational Variability of HIV-1 Env Trimer and Viral Vulnerability

Yiwei Cao[1] and Wonpil Im[1]*

[1]Department of Biological Sciences, Lehigh University, 111 Research Dr, Bethlehem, PA 18015, USA

*Corresponding Author: Wonpil Im (wonpil@lehigh.edu)

Abstract

HIV-1 envelope glycoprotein (Env) is critical for viral fusion and entry into host cells and remains a primary target for vaccine and antiviral drug development. Advances in soluble gp140 trimer design have provided insight into the ectodomain structure and dynamics. However, the membrane-proximal external region (MPER) and transmembrane domain (TMD) are comparatively understudied, and knowledge of the cytoplasmic tail (CT) is virtually absent. Additionally, the ectodomain and TMD have been investigated separately in previous studies. To investigate the trimeric gp120–gp41 as a complete entity and its structural flexibility, we built a full-length model of the gp120–gp41 trimer that is fully glycosylated with N-linked glycans and embedded in a lipid bilayer, and performed all-atom molecular dynamics simulations. Our results show that the ectodomain maintains a rigid internal structure stable in the prefusion state, whereas the intrinsic flexibility of the MPER enables the ectodomain to adopt a range of tilted orientations, potentially enhancing spatial alignment for receptor engagement. The centrally positioned R696 residue in the TMD interacts with lipid headgroups, ions, and the CT residues, resulting in conformational variability in the TMD and perturbations in the surrounding membrane that may facilitate the fusion process. Finally, we demonstrate how simulation trajectories can be leveraged to evaluate the accessibility of antibody epitopes across different regions of the protein.

Introduction

Human immunodeficiency virus type 1 (HIV-1) is the most prevalent strain of HIV responsible for the development of acquired immunodeficiency syndrome (AIDS). The HIV-1 envelop (Env) consists of a host cell-derived lipid membrane and viral glycoproteins that play a crucial role in mediating viral entry into host cells. The Env glycoprotein is initially synthesized in the endoplasmic reticulum (ER) as a precursor gp160 and cleaved by furin into two subunits, gp120 and gp41. The non-covalently associated gp120–gp41 complex is transported to the cell surface in the form of a trimer, where it is subsequently incorporated into the envelope of nascent virions during viral assembly. The exposure of Env protein is essential for binding to the primary receptor CD4 and the co-receptors CCR5 or CXCR4, triggering membrane fusion and viral entry. However, this exposure also renders the virus susceptible to immune attack. In response to host immune pressure, Env is densely coated with N-linked glycans added during post-translational modification in the ER and Golgi apparatus, which effectively shield vulnerable epitopes from immune recognition.

Since HIV-1 was identified as the cause of AIDS in the early 1980s, extensive research has been conducted to elucidate the mechanisms of viral infection and immune response. A major breakthrough was the design of soluble gp140 trimers, composing gp120 and part of the gp41 subunit. Native-like Env trimers, such as SOSIP (Sanders et al., 2013), native flexibly linked (NFL) (Sharma et al., 2015), and uncleaved prefusion-optimized (UFO) constructs (Kong et al., 2016), mimic the structure of viral spike and serve as valuable antigen targets for developing small molecule inhibitors and broadly neutralizing antibodies (bNAbs). With advances in structure determination techniques, such as X-ray crystallography and cryo-electron microscopy (cryo-EM), numerous high-resolution structures of gp140, both unliganded and antibody-bound, are now available in the Protein Data Bank (PDB). Even with such progresses, however, the remaining portions of gp41, including the membrane-proximal external region (MPER), transmembrane domain (TMD), and cytoplasmic tail (CT), are relatively understudied. The MPER is a highly conserved region targeted by several bNAbs, including 10E8, 2F5, 4E10, and Z13e1 (Ofek et al., 2004; Cardoso et al., 2005; Pejchal et al., 2009; Huang et al., 2012). However, in most studies, MPER is examined as a monomeric peptide bound to antibodies or as part of the Env trimer in membrane mimetics such as bicelles and nanodiscs (Rantalainen et al., 2020; Yang et al., 2022; Qi et al., 2025), which do not capture the structure and dynamics of unliganded MPER in the context of the gp120–gp41 trimer embedded in a membrane bilayer. Despite the critical roles of the MPER and TMD in mediating the fusion of viral and host cell membranes (Salzwedel et al., 1999; Miyauchi et al., 2005), they are often excluded from structural studies due to difficulties in crystallization of hydrophobic TM segments while maintaining their native conformations. Therefore, nuclear magnetic resonance (NMR) spectroscopy remains one of the few viable methods for obtaining structural information of MPER and TMD. However, the NMR studies from different research groups have yielded conflicting conclusions regarding the oligomeric state of the TMD. Reported TMD structures include trimeric coiled coils (Dev et al., 2016; Kwon et al., 2018; Piai et al., 2020; Piai et al., 2021), monomeric helices (Apellaniz et al., 2015), and trimers that are not tightly bundled (Reardon et al., 2014). Similarly, there are different conclusions about the conformation and orientation of MPER. Some studies suggest that gp41 bends at the MPER-TMD boundary (around residue 673), with the entire MPER adopting a helical conformation that lies nearly parallel to the membrane (Sun et al., 2008; Kwon et al., 2018), while others show that the C-terminal residues of MPER forms a continuous helix with the TMD (Dev et al., 2016; Chiliveri et al., 2018; Piai et al., 2020; Piai et al., 2021). In addition, the exceptionally long CT plays an important role in facilitating the incorporation of the Env glycoprotein into virions (Checkley et al., 2011). While it is known to contain three conserved amphipathic α-helical segments, referred to as lentiviral lytic peptides (LLPs), proposed models differ in the arrangement of LLPs (Murphy et al., 2017; Piai et al., 2021), and the complete structure of CT and its location in the membrane remains inconclusive.

In this work, we built a model of full-length gp120–gp41 trimer embedded in a lipid bilayer mimicking the lipid composition of the mammalian plasma membrane (Pogozheva et al., 2022) (Figure 1). We prepared simulation systems varying in the presence of the cleavage site and

CT, as well as the initial position of protein in the membrane. Multiple microsecond-long all-atom molecular dynamics (MD) simulations were performed for each system to explore the motions of individual protein domains and the membrane, and to examine how their conformational variability is affected by the difference in the initial configurations. Our results show that the ectodomain undergoes substantial tilting relative to the membrane plane while maintaining a rigid internal structure. In contrast, the MPER and TMD display highly diverse conformations, and their structural variations are influenced by the presence of the CT and the initial TMD position in the membrane. Moreover, we selected several bNAbs targeting the epitopes across different regions of the Env protein and demonstrate that the simulation trajectories can be used to assess the epitope accessibility.

Results

The ectodomain maintains a rigid internal structure and tilts independently of the TMD

As described in the Methods section, multiple simulation systems were constructed, varying in cleavage, TMD positioning, and CT truncation. Simulations trajectories are denoted as $CH^{CT}1$, $UL^{CT}2$, $CL^{\Delta CT}3$, $UH^{\Delta CT}3$, etc., where the first letter (C/U) indicates cleaved or uncleaved, the second letter (H/L) indicates the high or low initial TMD position, CT/ΔCT indicates the presence or absence of the CT, and the numeric suffix specifies the trajectory index among three replicas (Supplementary file 1–Supplementary Table 1). In all simulations, both the ectodomain and TMD adopted variable orientations relative to the bilayer plane. To quantify the tilt of two domains, we defined the tilt angles for the ectodomain (θ_{EC}) and TMD (θ_{TM}) (Figure 2A). θ_{EC} is measured between the bilayer normal and the vector from the center of mass (COM) of G594 (on all three protomers) to the COM of D167, while θ_{TM} is measured between the bilayer normal and the vector from the COM of I684 to the COM of V705. Across trajectories, θ_{EC} typically ranges from 0° to 40°, with occasional tilting up to 50°. Representative structures for different θ_{EC} values are shown in Figure 2A. In contrast to the wide range of θ_{EC}, θ_{TM} generally remained within 0° to 20° with rare excursions to 30° observed in a small fraction of trajectories. We split each 1-μs trajectory into four 0.25-μs intervals, and the variations of θ_{EC} and θ_{TM} over time show that large conformational changes occurred during the first 0.5 μs, followed by convergence of θ_{EC} and θ_{TM} distributions during the second 0.5 μs of most trajectories (Figure 2—figure supplements 1–4). For $CH^{\Delta CT}$, $CL^{\Delta CT}$, $UH^{\Delta CT}$ and $UL^{\Delta CT}$ systems, the θ_{EC} and θ_{TM} distributions calculated from the second 0.5 μs of three independent trajectories are shown in Figure 2B, and the data for CH^{CT}, CL^{CT}, UH^{CT} and UL^{CT} systems are shown in Figure 2—figure supplement 5. Although the combined dataset spans a wide range of θ_{EC} (0°–50°) and θ_{TM} (0°–30°), each individual trajectory explores only a portion of this space. For instance, $CH^{\Delta CT}1$ samples large θ_{EC} with small θ_{TM}, whereas $CH^{\Delta CT}3$ samples small θ_{EC} with medium-to-large θ_{TM}. Thus, multiple independent simulations are essential to capture a wide region of the conformational space. Most importantly, no consistent correlation was observed between θ_{EC} and θ_{TM}, either within single trajectories or across all trajectories combined. Pearson correlation coefficients of θ_{EC} and θ_{TM} in single trajectories varies between -0.5 and 0.5, with examples such as $CH^{\Delta CT}1$ and $CH^{\Delta CT}2$ showing similar θ_{TM} but distinct θ_{EC} values, and $CL^{\Delta CT}1$ and $CL^{\Delta CT}3$ showing similar θ_{EC} but distinct θ_{TM} values.

Despite the considerable tilting of the ectodomain relative to the membrane, its internal structure remains rigid and well-preserved throughout the simulations. The root-mean-square fluctuation (RMSF) and root-mean-square deviation (RMSD) of the ectodomain were calculated after the snapshots from each trajectory were aligned to the initial structure by maximizing the overlap in the ectodomain. The majority of the ectodomain displays low RMSF (< 2 Å), and the RMSD stabilized around 4 Å after an initial rise during the first 250 ns (Figure 3A, Figure 3—figure supplement 1). Higher RMSF values were observed in the residues missing from the cryo-EM structure highlighted in red in Figure 1A, B, and in part of HR1 (Q551-H564) that forms a flexible loop at the interface between two neighboring protomers. In addition, the RMSF of the MPER was also calculated with the trajectories aligned by the ectodomain. The entire MPER (K665–R683) and the adjacent HR2 segment (L660–D664) at the C-terminus of the ectodomain exhibited elevated fluctuations. The gp120–gp41 model was built based on

the NMR structure in which the MPER adopts a bent conformation consisting of two helices joined by a sharp turn. The resulting trimer widens from the HR2 helix to the midpoint of the MPER (F673) and narrows from F673 to the TMD. However, this specific conformation was not maintained throughout the simulations. We measured the inter-chain distances between the Cα atoms of corresponding residues (G644, E654, D664, and F673) on neighboring protomers to characterize the structural variation along the HR2 helix and the MPER (Figure 3B). The inter-chain distances of G644 and E654 maintained narrow distributions centered on their initial values, while those of D664 and F673 exhibited broader distributions, reflecting inward shifts of three protomers. This effect was more pronounced in uncleaved systems. In cleaved systems, the HR2 helix interacted with residues M530-N543 and L619-N625 of neighboring protomers, which helps to stabilize the bent MPER conformation (Figure 3C, D). However, these interactions were not consistently observed across all protomers and all simulations, and therefore we still observed considerable conformational variability in this region. In uncleaved systems, the closed loop at the cleavage site occupied the space between the HR2 helix and the neighboring protomer, disrupting their interactions and facilitating inward shifts of the HR2 helix (Figure 3E, F).

The energetically unfavorable R696 in the hydrophobic core results in asymmetric, kinked TMD conformations and disrupts membrane integrity

Unlike the predominantly hydrophobic TMDs commonly found in many viral envelop proteins, the gp41 TMD contains multiple charged residues: R683 at the N-terminal boundary, R707 and R709 at the C-terminal boundary, and a central arginine, R696 (Figure 4A). In simulations, R683 consistently interacted with lipid headgroups in the exoplasmic leaflet, while R707 and R709 interacted with those in the cytoplasmic leaflet, together acting as anchors that secure the TMD within the bilayer. In the cryo-EM structure (PDB ID: 7LOI), the side chain of R696 forms cation-π interaction with the side chain of F699 and a hydrogen bond with the backbone carbonyl group of L692, but it is oriented outward from the helical bundle. When embedded in the bilayer, however, this configuration is energetically unfavorable, as the positively charged side chain directly contacts hydrophobic lipid tails (Cheng et al., 2012). Early in the simulations, the TMD rapidly rearranged to allow R696 to interact with more favorable partners, including negatively charged lipid headgroups from either leaflet, ions and water molecules diffusing into the bilayer center, as well as the polar and positively charged groups in the CT when it is present. Because the limited space at the TMD core can only accommodate at most two inward-facing arginine residues, at least one R696 is forced outward to interact with lipid headgroups or CT residues. The differences in arginine orientation and interacting partners give rise to asymmetric protomer conformations and distinct TMD tilts (Figure 4A–F). When R696 points outward, its interactions with lipid head groups or CT residues can be strong enough to destabilize the local helix, introducing a kink into the TMD. Representative snapshots from different trajectories illustrate these asymmetric kinked conformations (Figure 4—figure supplements 1–8). Beyond local deformation of the TMD, R696–lipid interactions perturb bilayer organization, inducing the translocation of lipid headgroups and water molecules toward the bilayer center (Figure 4G). We calculated the interaction frequencies of each TMD residue with lipid headgroups, lipid tails, and water molecules. The results show that the membrane disruption is persistent since many non-terminal TMD residues maintain frequent contacts with water and lipid headgroups throughout the simulation (Figure 4H).

To explore whether R696 exhibits a preference for the exoplasmic versus cytoplasmic leaflet, or interacts with either randomly, we generated two initial structures (high and low) with the TMD positioned at two distinct positions, separated by 4 Å along the membrane normal (see Methods). In simulations initiated from the “high” TMD configuration, R696 residues in three protomers interacted with the lipid headgroups in either leaflet (Figure 4B, C, Figure 4—figure supplements 1, 3) In contrast, in simulations initiated from the “low” TMD configuration, R696 residues interacted exclusively with the headgroups in the cytoplasmic leaflet (Figure 4E and F, Figure 4—figure supplements 2, 4). In the full-length systems, the plate-shaped CT occupies substantial space in the cytoplasmic leaflet, displacing lipids during model construction.

Because the CT is not thick enough to fully span the cytoplasmic leaflet, an empty gap remained between the CT and the exoplasmic leaflet (Figure 1—figure supplement 1D). Over time, lipids in the exoplasmic leaflet shifted downward while the CT residues moved upward to fill this space, causing local bilayer thinning. Under these conditions, upward-oriented R696 can still interact with the headgroups in the exoplasmic leaflet, but downward-oriented R696 primarily contacted CT residues, with rare cases of lipids migrating upward from the cytoplasmic leaflet and approaching the protein. A comparison of all full-length systems (Figure 4—figure supplements 5–8) shows that R696 preferentially adopted downward orientations in the simulations initiated from the "low" TMD configuration (Figure 4—figure supplements 6, 8). In these cases, upward-shifted CT residues interact with the C-terminal half of the TMD, rather than solely with R696, resulting in deeper burial of the TMD in the membrane.

MPER adopts diverse conformations, and its exposure depends on both MPER and TMD conformations

Starting from the initial helix-turn-helix conformation consisting of two separate helical segments, the N-terminal half (MPER-N) and the C-terminal half (MPER-C), MPER underwent rapid rearrangements, and a wide variety of conformations were sampled across all trajectories. In the initial structure, the trimeric MPER was positioned perpendicular to the membrane, with the helical MPER-N tilted inward and MPER-C tilted outward (Figure 4A). Such conformation and orientation were maintained in some trajectories such as $CL^{\Delta CT}3$ (Figure 4—figure supplement 2C), but in others the helix-turn-helix MPER shifted into a horizontal orientation parallel to the membrane surface (Figure 4—figure supplement 6A) or a more vertical arrangement with both MPER-N and MPER-C tilted outward (Figure 4E, Figure 4—figure supplement 4A). It was also observed that the HR2 helix in the ectodomain, MPER, and TMD merged into a continuous long helix (Figure 4C, F, Figure 4—figure supplement 3C). In addition, it was common for part of the MPER, particularly the MPER-C, to lose its helical structure and become a random coil. Therefore, the distinct MPER conformations reported in various experimental studies can be all valid as each captures a possible state within the highly flexible conformational landscape of the MPER. It is noteworthy that the MPER in three protomers can adopt different conformations and orientations, resulting in asymmetric local structures.

Because the MPER is a target of multiple bNAbs, we next examined how its exposure is affected by TMD conformation. As described above, interactions between R696 and lipid headgroups affect the burial depth of the TMD, thereby influencing the positioning of the adjacent MPER. To quantify MPER exposure, we measured the vertical distance from the Cα of F673, approximately the midpoint of MPER, to the highest point of the neighboring lipid headgroups, denoted by d_{F673}. Positive values indicate that F673 lies above the bilayer surface, and negative values indicate its membrane burial. In the initial 'low' and 'high' TMD configurations, d_{F673} was 6.1 Å and 9.1 Å, respectively, but across simulations it spanned a wide range from -15 Å to 20 Å (Figure 5A, B). Two examples illustrate this variability. In the first example, all three R696 residues interacted with the cytoplasmic leaflet, drawing the TMD deeper into the membrane. Consequently, the entire MPER-C and most of MPER-N were buried in the membrane, with one F673 positioned 11.3 Å below the membrane surface, thus bringing the ectodomain in close proximity to the membrane (Figure 5C, E). In the second example, two R696 residues interacted with the exoplasmic leaflet, while one interacted with the cytoplasmic leaflet. In the protomer with upward-oriented R696, the MPER-N, MPER-C, and TMD formed a continuous helix. As a result, most of the MPER extended outside the bilayer, with d_{F673} reaching 18.0 Å and the ectodomain displaced farther from the membrane (Figure 5D, F). The d_{F673} distributions across all cleaved CT-truncated systems suggest that d_{F673} tends to be smaller when the simulations started from the "low" TMD configuration (Figure 5B). In both "high" and "low" configurations, the mean d_{F673} calculated from the simulation trajectories is smaller than its value calculated from the initial structure, due to both protein conformational change and lipid diffusion. In the "low" configuration, d_{F673} decreased by 5.0 Å (from 6.1 Å to 1.1 Å), a slightly larger reduction than in the "high" configuration where d_{F673}

dropped by 3.5 Å (from 9.1 Å to 5.6 Å), indicating that deeper MPER burial arises from a larger-scale protein motion rather than merely from a lower initial placement.

Ectodomain epitopes are conditionally accessible, whereas MPER epitopes are virtually inaccessible in the closed prefusion state

To access the accessibility of epitopes on different parts of the trimeric Env protein in the prefusion state, we quantified how frequent their epitopes were exposed without steric clashes from neighboring protein residues, glycans or membrane lipids to six selected antibodies: PGT128 targeting the V3 loop and the N332 glycan (PDB ID: 5JSA); PG9 targeting the V1/V2 loop (PDB ID: 3U2S); VRC01 targeting the CD4 binding site (PDB ID: 4LST); 35O22 targeting the gp120–gp41 interface (PDB ID: 4TVP); and 10E8 (PDB ID: 6VPX) and 4E10 (PDB ID: 1TZG) targeting the MPER (Figure 6A, Supplementary file 1–Supplementary Tables 2–5).

The epitope of PGT128 consists of the V3 loop and N332 glycan, and it is widely recognized that PGT128 binding is mediated by the protein-protein interaction with V3 loop and the protein-glycan interaction with N332 glycan, facilitated by its extra-long HCDR3 loop penetrating the glycan shield (Figure 6—figure supplement 1). Due to the absence of defined secondary structures, glycans exhibit greater conformational flexibility than the protein, although the crowded surface of gp120–gp41 may partially restrict the motion of glycans. N332 glycan must adopt specific conformations to enable specific interactions with PGT128, whereas most conformations block antibody approach. Additional glycans on N137, N156, and N301 can further occlude the site (Figure 6B). Epitope accessibility to PGT128 varied widely across protomers and trajectories, with many cases exceeding 35%. In about half trajectories, at least one protomer exhibited >35% accessibility (Supplementary file 1–Supplementary Table 2). For instance, in system $CH^{\Delta CT}1$, two protomers showed accessibility >40%, whereas in system $CL^{CT}2$, all three protomers are <5%. The second antibody, PG9, targets the V1/V2 apex, where binding can be hindered by six glycans, N156, N160, and N185E on the same protomer, and N160, N185E, and N185H on the neighboring protomer (Figure 6—figure supplement 2). In most trajectories, the epitopes on all three protomers were occluded in >95% of the snapshots, with the exception of a few systems ($CH^{\Delta CT}1$, $CH^{\Delta CT}2$, and $CL^{\Delta CT}3$) where accessibility was non-negligible. (Supplementary file 1–Supplementary Table 3). The third antibody, VRC01, targets the CD4 binding site where six glycans, N185H, N197, N276, N363, and N462 on the same protomer and N301 on the neighboring protomer, can interfere with antibody binding (Figure 6—figure supplement 3). Similar to PGT128, the VRC01 epitope is moderately to highly accessible to at least one protomer in many trajectories, whereas in some trajectories, the epitopes on all three protomers are nearly completely shielded (Supplementary file 1–Supplementary Table 4).

35O22 targets the gp120–gp41 interface that is not intrinsically membrane-proximal. However, its binding orientation combined with ectodomain tilting can lead to clashes of 35O22 with the membrane (Figure 6—figure supplement 4). Comparing accessibility with and without including the steric effects of membrane lipids shows that the membrane can hinder 35O22 binding, particularly when it approaches from the direction of ectodomain tilt (Figure 6C). For example, in system $UH^{\Delta CT}2$, the accessibility frequencies of the three protomers were 49%, 5%, and 25%, when only the shielding of three glycans (N88, N234, and N618) was considered. When the steric effects of both glycans and membrane lipids were included, the first protomer's frequency dropped from 49% to 1%, while the other two remain essentially unchanged (Supplementary file 1–Supplementary Table 5).

The last two antibodies, 10E8 and 4E10, target the MPER. In the PDB structures of 10E8, the epitope is a 17-residue peptide corresponding to residues 671-687 in this work. In the PDB structure of 4E10, the epitope is a 12-residue peptide corresponding to residues 669-680. Superposition of these antibodies onto the initial simulation structure revealed extensive clashes with the protein and glycans linked to N611, N618, and N637 on the neighboring protomer, as well as slight steric clashes with the membrane (Figure 6—figure supplements 5, 6). Using the strict criteria (≤10 heavy-atom clashes with protein/glycan and ≤20 with lipids),

no snapshot was identified in which the epitope of either antibody was accessible. Even with relaxed criteria (≤20 heavy-atom clashes with protein/glycan or ≤40 with lipids), accessible cases remain rare, which require either large ectodomain tilts to create space on the opposite side or the MPER transitioning into continuous helices to displace the ectodomain from the membrane (Figure 6D, Figure 6—figure supplements 7–10). Given the extremely low frequency of such specific conformations, MPER epitopes are effectively inaccessible in the prefusion trimer, suggesting that MPER-targeting antibodies such as 10E8 and 4E10 act at later stages of viral entry.

Discussion

The hydrophobic and flexible nature of the MPER has made it difficult to crystallize, and most structural insights have come from NMR studies. However, structural information on the MPER and TMD has been inconsistent in the literature. For example, one study of the MPER in DPC micelles suggested that the MPER adopted a distorted helical structure lying parallel to the membrane surface (Sun et al., 2008), while NMR studies using gp41 fragments containing part of the MPER and TMD embedded in bicelles or bilayers reached conflicting conclusions. Some proposed a kink (Apellaniz et al., 2015) or turn (Kwon et al., 2018) between the MPER and TMD, whereas others suggested that the MPER and TMD formed a continuous helix aligned perpendicular to the membrane (Dev et al., 2016; Chiliveri et al., 2018). In the NMR structure (PDB ID: 7LOI) used to build our model structure, a kink appears between the MPER and TMD, along with a turn in the middle of the MPER (Fu et al., 2018; Piai et al., 2021). Crystal structures of MPER peptides bound to various antibodies show that in addition to the helix-turn-helix conformation, MPER can also adopt alternative structures in which the MPER-C remains helical while part of the MPER-N becomes unstructured (Williams et al., 2017). Such variability underscores the intrinsic structural flexibility of the MPER, with each experimental method capturing only a subset of possible states under specific conditions. In this study, we employed MD simulations to sample a broader conformational landscape and gain a more comprehensive view of the structural heterogeneity of the MPER. The results reveal that the MPER can adopt all of the conformations reported experimentally. While MPER plasticity has been linked to its role in virus-host membrane fusion because it enables the ectodomain and TMD to adopt distinct orientations during large-scale structural rearrangements, our results show that this flexibility is already present in the prefusion state. At this stage, the MPER functions as a flexible hinge that facilitates ectodomain tilting, which can be critical for the spatial alignment of the CD4-binding site with the host receptor for efficient receptor engagement.

To probe antibody recognition, we assessed epitope accessibility using snapshots extracted from simulation trajectories. Unlike static experimental structures, this approach incorporates protein dynamics as well as the influence of glycans and membrane lipids, yielding a frequency-based measure indicating how often each epitope is accessible to its corresponding antibody, rather than a binary accessible/occluded classification. It is worth noting that viral glycans are highly variable in both site occupancy and the specific glycoform present at each glycosylation site, as indicated by mass spectrometry data (Behrens et al., 2016; Cao et al., 2017). For structure modeling, however, we assumed full occupancy at all glycosylation sites and selected a single representative glycoform per site. As a result, the estimated glycan shield may not fully reflect the in vivo situation, and the calculated frequency of antibody accessibility should be regarded as an approximation. Our results indicate that epitopes on the ectodomain, though heavily shielded by glycans, can still become transiently accessible in the closed prefusion state. The frequency provides a quantitative measurement of how vulnerable each epitope is to antibody binding. When estimating steric clashes between the membrane and antibodies, we adopted relaxed cutoffs to account for lipid fluidity, but the membrane due to fast lipid dynamics may in reality accommodate even greater steric overlap. Although the MPER can transiently extend out of the membrane when gp120–gp41 remains in the prefusion state, MPER-targeting antibodies still encounter substantial steric hindrance from the bulky gp120 subunit, surrounding glycans, and the membrane throughout nearly all simulations.

Therefore, these antibodies are unlikely to adopt the proper orientations required to bind their epitopes. This is consistent with experiment studies indicating that MPER-targeting antibodies bind effectively only after the gp120–gp41 trimer undergoes major conformational rearrangements toward a fusion-intermediate or post-fusion state (Alam et al., 2009; Chen et al., 2014; Lee et al., 2016).

Overall, the data presented here demonstrate that structural modeling integrated with molecular dynamic simulations can be applied to complex biomolecular systems, such as the gp120–gp41 trimer embedded in a membrane bilayer, which enables detailed characterization of protein dynamics at the molecular level and facilitates the investigation of antigen-antibody interactions, thereby offering potential guidance for rational vaccine design.

Materials and methods

System building

Modeling of full-length gp120–gp41 trimer

The full-length gp120–gp41 trimeric model was generated by combining the crystal structure of the ectodomain with the NMR structure containing the MPER, TMD, and CT (Figure 1A, Figure 1—figure supplement 1A). Considering structure resolution and the number of missing residues, we selected the crystal structure of a soluble Env trimer exhibiting a closed-from, native-like prefusion conformation (PDB ID: 6B0N) (Sarkar et al., 2018), which contains the entire ectodomain including the V1-V5 loops, fusion peptide (FP), heptad repeat 1 (HR1), and heptad repeat 2 (HR2) (Figure 1B). In contrast to the ectodomain well represented in the PDB with numerous crystal and cryo-EM structures, structural data for the TMD are scarce. For this region, we chose the NMR structure (PDB ID: 7LOI) (Piai et al., 2021) containing the MPER, TMD, and CT. Both 6B0N and 7LOI are trimeric structures and the C-terminus of 6B0N includes five residues (L660–D664) overlapping with the N-terminus of 7LOI. Measurement of the Cα–Cα distances for D664 in the three protomers revealed that the inter-protomer spacing in 7LOI (16 Å) was much shorter than in 6B0N (39 Å) (Supplementary file 1–Supplementary Figure 1A). To reconcile this difference, we performed short MD simulations on 6B0N with residues E32–Q640 fixed and a distance restraint applied to D664 in order to reduce the Cα–Cα distance to 33 Å. Similarly, for 7LOI, residues I675–L856 were fixed, and a distance restraint was applied to D664 to increase the distance to 33 Å (Supplementary file 1–Supplementary Figure 1B). The two adjusted structures were then combined by aligning the overlapping residues. The missing loops in the ectodomain (T63, D149, E185A–N185I, S401–G409) and the unresolved region between the TM and CT (F717–G738) were grafted from a modeled structure generated by I-TASSER (Yang et al., 2015) (Supplementary file 1–Supplementary Figure 1C). 6B0N is a native flexibly linked (NFL) trimer in which the furin cleavage site ^{508}REKR^{511} is substituted with a 10-residue linker ^{508}GG<u>GGSGGGG</u>S^{511}, but the underlined residues are unresolved in the structure. To construct the cleaved gp120–gp41 model, we removed the initial “GG” and final “S” from 6B0N. For the uncleaved gp160 model, we closed the cleaved loop by adding ^{508}REKR^{511} back to the structure. If we simply connect G507 and A512 without adjusting the conformation of flanking loops on both sides of the cleavage site, two neighboring protomers become entangled, yielding a knot-like fold (Supplementary file 1–Supplementary Figure 2A, B). In 6B0N, the helical HR2 region lies at the C-terminus of each protomer, allowing it to thread through a loop formed by the neighboring protomer (Supplementary file 1–Supplementary Figure 2C). However, in the context of the full-length gp120–gp41, it is implausible that the three protomers are entangled as it does not conform to a realistic folding pathway. Therefore, we performed short MD simulations with restraints to adjust the conformations of the flanking loops and the HR2 helix before closing the loop with REKR fragment (Supplementary file 1–Supplementary Figure 2D, E). Two mutations (S764C and S837C) were applied to CT to reintroduce the palmitoylation sites, and lipid tails oriented towards the hydrophobic core of the bilayer were added to the palmitoylation sites (Figure 1—figure supplement 1D). In addition to the full-length model, we built the CT-

truncated model by removing the modeled residues F717–G738 that are originally unresolved in the PDB structure 7LOI, and the residues E739–L856 that form the large CT plate.

Glycosylation

N-linked glycans were modeled using *Glycan Reader & Modeler* (Jo et al., 2011; Park et al., 2017; Park et al., 2019) in CHARMM-GUI (Jo et al., 2008). The reported mass spectrometry (MS) experiments have revealed the probability of different types of N-linked glycans at each glycosylation site (Behrens et al., 2016; Cao et al., 2017). For each site, we selected the one with the highest probability. As the MS data provide only the glycan type (high-mannose, hybrid, and complex) and composition (number of each monosaccharide unit), a representative isomer was chosen when multiple isomers corresponded to the same composition. For example, HexNAc(2)Hex(9), i.e., the Man_9 glycan, has a single isomer, whereas HexNAc(2)Hex(8), i.e., the Man_8 glycan, has multiple isomers depending on which non-reducing terminal mannose is trimmed from Man_9. For complex N-linked glycans, one composition can correspond to multiple isomers differing in the number of branches on the α1-3 and α1-6 arms and in the placement of non-reducing terminal neuraminic acid (Neu5Ac). In the absence of linkage-specific information, one isomer was arbitrarily selected in cases of multiple possibilities. The selected glycan sequences for the 27 glycosylation sites in each promoter are summarized in Supplementary file 1–Supplementary Table 6.

Membrane

The full-length and CT-truncated gp120–gp41 models were embedded into an asymmetric lipid bilayer with the lipid composition corresponding to a mammalian plasma membrane (Pogozheva et al., 2022), which consists of phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylserine (PS), phosphatidic acid (PA), sphingomyelin (SM), cholesterol (CHOL), and glucosylceramide (GlcCer). In the CT-truncated system, the exoplasmic and cytoplasmic leaflets contain similar numbers of lipids, although their compositions differ. In the full-length system, the cytoplasmic leaflet contains approximately 100 fewer lipids than the exoplasmic leaflet due to the space occupied by the CT. The lipid composition is summarized in Supplementary file 1–Supplementary Table 7. Although it is recognized that the charged residues R683, R707, and R709, flanking the N- and C-termini of the TMD, are anchored within the lipid headgroups, there remains some flexibility to shift the TMD slightly up or down within the membrane. Assuming the lipid bilayer is aligned parallel to the xy-plane, the protein was positioned at two distinct depths along the z-axis, differing by approximately 4 Å, referred to as the “high” and “low” TMD configurations. To examine whether this subtle difference in the initial configurations leads to distinct conformational changes during simulations, we initiated independent simulations from each configuration. The glycoprotein-membrane system was solvated in a box of approximately 210 × 210 × 260 $Å^3$ using the TIP3P water model (Jorgensen et al., 1983), and KCl was added at a concentration of 0.15 M to neutralize the system. Final simulation input files were generated using *Membrane Builder* (Jo et al., 2007; Jo et al., 2009; Wu et al., 2014; Lee et al., 2016; Lee et al., 2019) in CHARMM-GUI.

Simulation details

The combination of cleaved vs. uncleaved, full-length vs. CT-truncated, and high vs. low TMD positions in the membrane results in eight distinct configurations, and we performed three independent 1-μs all-atom MD simulations for each configuration. The CHARMM36(m) force field (Guvench et al., 2009; Klauda et al., 2010; Raman et al., 2010; Guvench et al., 2011; Huang et al., 2017) was used for proteins, carbohydrates, and lipids. The total number of atoms is approximately 1,100,000 (including ~300,000 water molecules and ~1,700 ions), with slight variations between systems. van der Waals interactions were smoothly switched off over 10–12 Å using a force-based switching function (Steinbach et al., 2004), and long-range

electrostatic interactions were calculated using the particle-mesh Ewald method (Essmann et al., 1995) with a mesh size of ~1 Å. All simulations were performed using GROMACS (Van Der Spoel et al., 2005). Bond lengths and angles involving hydrogens were constrained using the LINCS algorithm (Hess et al., 1997). The system was equilibrated in the canonical (NVT) ensemble at 310.15 K for 2 × 1.25 ns with a 1-fs time step using the Berendsen thermostat (Berendsen et al., 1984) (coupling constant τ_t = 1 ps), and then switched to the isothermal–isobaric (NPT) ensemble for 1.25 ns with a 1-fs time step, followed by 3 × 5 ns with a 2-fs time step using the Berendsen thermostat (τ_t = 1 ps) and barostat (reference pressure = 1 bar, coupling constant τ_p = 5 ps, compressibility = 4.5×10^{-5} bar^{-1}). Positional and dihedral restraints were applied to proteins, glycans, and lipids, with force constants progressively reduced over successive intervals. In the production run, temperature was maintained using the Nosé–Hoover thermostat (Hoover, 1985; Nosé, 2006) and the pressure coupling was applied using the semi-isotropic Parrinello–Rahman barostat (Parrinello et al., 1981; Nosé et al., 2006). A 4-fs time step was used with the hydrogen mass repartitioning technique(Hopkins et al., 2015) and all restraint potentials were removed. The Python library MDTraj (McGibbon et al., 2015) was used to analyze the simulation trajectories.

Assessment of Antibody Epitope Accessibility

We selected six antibodies based on epitope location and the availability of PDB structures containing both antibody and epitope: PGT128 targeting the V3 loop and the N332 glycan (PDB ID: 5JSA) (Kong et al., 2016); PG9 targeting the V1/V2 loop (PDB ID: 3U2S) (McLellan et al., 2011); VRC01 targeting the CD4 binding site (PDB ID: 4LST) (Zhou et al., 2013); 35O22 targeting the gp120–gp41 interface (PDB ID: 4TVP) (Pancera et al., 2014); 10E8 (PDB ID: 6VPX) (Rantalainen et al., 2020) and 4E10 (PDB ID: 1TZG) (Cardoso et al., 2005) targeting the MPER. For each antibody–epitope complex, TM-align (Zhang et al., 2005) was used to identify the optimal alignment between the epitope from the PDB structure and each protomer of the trimeric protein, and the resulting rotation matrix was applied to place the antibody relative to each protomer. Steric clashes were then assessed by counting any heavy atom of protein, glycan, or lipid within 2 Å of a heavy atom of the antibody. An epitope was considered occluded by protein and glycans if more than 10 heavy-atom clashes occurred with these components. To account for membrane flexibility, we adopted a more permissive cutoff for lipids: the epitope was classified as occluded by the membrane if more than 20 lipid heavy-atom clashes were detected. To calculate the frequency of epitope accessibility, snapshots were extracted from simulation trajectories, and structural alignment and assessment of steric clashes were performed on each snapshot.

Data availability

The modeled structures, force field parameters, and GROMACS input files are available on Zenodo [https://doi.org/10.5281/zenodo.17259909]. The code used in this article can be found at https://doi.org/10.5281/zenodo.17259951.

References

Alam SM, Morelli M, Dennison SM, Liao HX, Zhang R, Xia SM, Rits-Volloch S, Sun L, Harrison SC, Haynes BF, Chen B (2009) **Role of HIV membrane in neutralization by two broadly neutralizing antibodies** *PNAS* **106**: 20234-20239.

Apellaniz B, Rujas E, Serrano S, Morante K, Tsumoto K, Caaveiro JM, Jimenez MA, Nieva JL (2015) **The Atomic Structure of the HIV-1 gp41 Transmembrane Domain and Its Connection to the Immunogenic Membrane-proximal External Region** *Journal of Biological Chemistry* **290**: 12999-13015.

Behrens AJ, Vasiljevic S, Pritchard LK, Harvey DJ, Andev RS, Krumm SA, Struwe WB, Cupo A, Kumar A, Zitzmann N, Seabright GE, Kramer HB, Spencer DI, Royle L, Lee JH, Klasse PJ, Burton DR, Wilson IA, Ward AB, Sanders RW, Moore JP, Doores KJ, Crispin M (2016) **Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein** *Cell Reports* **14**: 2695-2706.

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) **Molecular dynamics with coupling to an external bath** *The Journal of Chemical Physics* **81**: 3684-3690.

Cao L, Diedrich JK, Kulp DW, Pauthner M, He L, Park SR, Sok D, Su CY, Delahunty CM, Menis S, Andrabi R, Guenaga J, Georgeson E, Kubitz M, Adachi Y, Burton DR, Schief WR, Yates JR, III, Paulson JC (2017) **Global site-specific N-glycosylation analysis of HIV envelope glycoprotein** *Nature Communications* **8**: 14954.

Cardoso RM, Zwick MB, Stanfield RL, Kunert R, Binley JM, Katinger H, Burton DR, Wilson IA (2005) **Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41** *Immunity* **22**: 163-173.

Checkley MA, Luttge BG, Freed EO (2011) **HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation** *Journal of Molecular Biology* **410**: 582-608.

Chen J, Frey G, Peng H, Rits-Volloch S, Garrity J, Seaman MS, Chen B (2014) **Mechanism of HIV-1 neutralization by antibodies targeting a membrane-proximal region of gp41** *Journal of Virology* **88**: 1249-1258.

Cheng X, Im W (2012) **NMR observable-based structure refinement of DAP12-NKG2C activating immunoreceptor complex in explicit membranes** *Biophysical Journal* **102**: L27-29.

Chiliveri SC, Louis JM, Ghirlando R, Baber JL, Bax A (2018) **Tilted, Uninterrupted, Monomeric HIV-1 gp41 Transmembrane Helix from Residual Dipolar Couplings** *Journal of the American Chemical Society* **140**: 34-37.

Dev J, Park D, Fu Q, Chen J, Ha HJ, Ghantous F, Herrmann T, Chang W, Liu Z, Frey G, Seaman MS, Chen B, Chou JJ (2016) **Structural basis for membrane anchoring of HIV-1 envelope spike** *Science* **353**: 172-175.

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) **A smooth particle mesh Ewald method** *The Journal of Chemical Physics* **103**: 8577-8593.

Fu Q, Shaik MM, Cai Y, Ghantous F, Piai A, Peng H, Rits-Volloch S, Liu Z, Harrison SC, Seaman MS, Chen B, Chou JJ (2018) **Structure of the membrane proximal external region of HIV-1 envelope glycoprotein** *PNAS* **115**: E8892-E8899.

Guvench O, Hatcher ER, Venable RM, Pastor RW, Mackerell AD (2009) **CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses** *Journal of Chemical Theory and Computation* **5**: 2353-2370.

Guvench O, Mallajosyula SS, Raman EP, Hatcher E, Vanommeslaeghe K, Foster TJ, Jamison FW, 2nd, Mackerell AD, Jr. (2011) **CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling** *Journal of Chemical Theory and Computation* **7**: 3162-3180.

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) **LINCS: A linear constraint solver for molecular simulations** *Journal of Computational Chemistry* **18**: 1463-1472.

Hoover WG (1985) **Canonical dynamics: Equilibrium phase-space distributions** *Physical Review A, General Physics* **31**: 1695-1697.

Hopkins CW, Le Grand S, Walker RC, Roitberg AE (2015) **Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning** *Journal of Chemical Theory and Computation* **11**: 1864-1874.

Huang J, Ofek G, Laub L, Louder MK, Doria-Rose NA, Longo NS, Imamichi H, Bailer RT, Chakrabarti B, Sharma SK, Alam SM, Wang T, Yang Y, Zhang B, Migueles SA, Wyatt R, Haynes BF, Kwong PD, Mascola JR, Connors M (2012) **Broad and potent neutralization of HIV-1 by a gp41-specific human antibody** *Nature* **491**: 406-412.
Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmuller H, MacKerell AD, Jr. (2017) **CHARMM36m: an improved force field for folded and intrinsically disordered proteins** *Nature Methods* **14**: 71-73.
Humphrey W, Dalke A, Schulten K (1996) **VMD: visual molecular dynamics** *Journal of Molecular Graphics and Modelling* **14**: 33-38, 27-38.
Jo S, Kim T, Im W (2007) **Automated builder and database of protein/membrane complexes for molecular dynamics simulations** *PLoS One* **2**: e880.
Jo S, Kim T, Iyer VG, Im W (2008) **CHARMM-GUI: a web-based graphical user interface for CHARMM** *Journal of Computational Chemistry* **29**: 1859-1865.
Jo S, Lim JB, Klauda JB, Im W (2009) **CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes** *Biophysical Journal* **97**: 50-58.
Jo S, Song KC, Desaire H, MacKerell AD, Jr., Im W (2011) **Glycan Reader: automated sugar identification and simulation preparation for carbohydrates and glycoproteins** *Journal of Computational Chemistry* **32**: 3135-3141.
Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) **Comparison of simple potential functions for simulating liquid water** *The Journal of Chemical Physics* **79**: 926-935.
Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD, Jr., Pastor RW (2010) **Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types** *The Journal of Physical Chemistry B* **114**: 7830-7843.
Kong L, He L, de Val N, Vora N, Morris CD, Azadnia P, Sok D, Zhou B, Burton DR, Ward AB, Wilson IA, Zhu J (2016) **Uncleaved prefusion-optimized gp140 trimers derived from analysis of HIV-1 envelope metastability** *Nature Communications* **7**: 12040.
Kwon B, Lee M, Waring AJ, Hong M (2018) **Oligomeric Structure and Three-Dimensional Fold of the HIV gp41 Membrane-Proximal External Region and Transmembrane Domain in Phospholipid Bilayers** *Journal of the American Chemical Society* **140**: 8246-8259.
Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, Wei S, Buckner J, Jeong JC, Qi Y, Jo S, Pande VS, Case DA, Brooks CL, 3rd, MacKerell AD, Jr., Klauda JB, Im W (2016) **CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field** *Journal of Chemical Theory and Computation* **12**: 405-413.
Lee J, Patel DS, Stahle J, Park SJ, Kern NR, Kim S, Lee J, Cheng X, Valvano MA, Holst O, Knirel YA, Qi Y, Jo S, Klauda JB, Widmalm G, Im W (2019) **CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans** *Journal of Chemical Theory and Computation* **15**: 775-786.
Lee JH, Ozorowski G, Ward AB (2016) **Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer** *Science* **351**: 1043-1048.
McGibbon RT, Beauchamp KA, Harrigan MP, Klein C, Swails JM, Hernandez CX, Schwantes CR, Wang LP, Lane TJ, Pande VS (2015) **MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories** *Biophysical Journal* **109**: 1528-1532.
McLellan JS, Pancera M, Carrico C, Gorman J, Julien JP, Khayat R, Louder R, Pejchal R, Sastry M, Dai K, O'Dell S, Patel N, Shahzad-ul-Hussan S, Yang Y, Zhang B, Zhou T, Zhu J, Boyington JC, Chuang GY, Diwanji D, Georgiev I, Kwon YD, Lee D, Louder MK, Moquin S, Schmidt SD, Yang ZY, Bonsignori M, Crump JA, Kapiga SH, Sam NE, Haynes BF, Burton DR, Koff WC, Walker LM, Phogat S, Wyatt R, Orwenyo J, Wang LX, Arthos J, Bewley CA, Mascola JR, Nabel GJ, Schief WR, Ward AB, Wilson IA, Kwong PD (2011) **Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9** *Nature* **480**: 336-343.
Miyauchi K, Komano J, Yokomaku Y, Sugiura W, Yamamoto N, Matsuda Z (2005) **Role of the specific amino acid sequence of the membrane-spanning domain of human immunodeficiency virus type 1 in membrane fusion** *Journal of Virology* **79**: 4720-4729.

Murphy RE, Samal AB, Vlach J, Saad JS (2017) **Solution Structure and Membrane Interaction of the Cytoplasmic Tail of HIV-1 gp41 Protein** *Structure* **25**: 1708-1718.e1705.
Nosé S (2006) **A molecular dynamics method for simulations in the canonical ensemble** *Molecular Physics* **52**: 255-268.
Nosé S, Klein ML (2006) **Constant pressure molecular dynamics for molecular systems** *Molecular Physics* **50**: 1055-1076.
Ofek G, Tang M, Sambor A, Katinger H, Mascola JR, Wyatt R, Kwong PD (2004) **Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope** *Journal of Virology* **78**: 10724-10737.
Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, Huang J, Acharya P, Chuang GY, Ofek G, Stewart-Jones GB, Stuckey J, Bailer RT, Joyce MG, Louder MK, Tumba N, Yang Y, Zhang B, Cohen MS, Haynes BF, Mascola JR, Morris L, Munro JB, Blanchard SC, Mothes W, Connors M, Kwong PD (2014) **Structure and immune recognition of trimeric pre-fusion HIV-1 Env** *Nature* **514**: 455-461.
Park SJ, Lee J, Patel DS, Ma H, Lee HS, Jo S, Im W (2017) **Glycan Reader is improved to recognize most sugar types and chemical modifications in the Protein Data Bank** *Bioinformatics* **33**: 3051-3057.
Park SJ, Lee J, Qi Y, Kern NR, Lee HS, Jo S, Joung I, Joo K, Lee J, Im W (2019) **CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates** *Glycobiology* **29**: 320-331.
Parrinello M, Rahman A (1981) **Polymorphic transitions in single crystals: A new molecular dynamics method** *Journal of Applied Physics* **52**: 7182-7190.
Pejchal R, Gach JS, Brunel FM, Cardoso RM, Stanfield RL, Dawson PE, Burton DR, Zwick MB, Wilson IA (2009) **A conformational switch in human immunodeficiency virus gp41 revealed by the structures of overlapping epitopes recognized by neutralizing antibodies** *Journal of Virology* **83**: 8451-8462.
Piai A, Fu Q, Cai Y, Ghantous F, Xiao T, Shaik MM, Peng H, Rits-Volloch S, Chen W, Seaman MS, Chen B, Chou JJ (2020) **Structural basis of transmembrane coupling of the HIV-1 envelope glycoprotein** *Nature Communications* **11**: 2317.
Piai A, Fu Q, Sharp AK, Bighi B, Brown AM, Chou JJ (2021) **NMR Model of the Entire Membrane-Interacting Region of the HIV-1 Fusion Protein and Its Perturbation of Membrane Morphology** *Journal of the American Chemical Society* **143**: 6609-6615.
Pogozheva ID, Armstrong GA, Kong L, Hartnagel TJ, Carpino CA, Gee SE, Picarello DM, Rubin AS, Lee J, Park S, Lomize AL, Im W (2022) **Comparative Molecular Dynamics Simulation Studies of Realistic Eukaryotic, Prokaryotic, and Archaeal Membranes** *Journal of Chemical Information and Modeling* **62**: 1036-1051.
Qi Y, Zhang S, Wang K, Ding H, Zhang Z, Anang S, Nguyen HT, Kappes JC, Sodroski J, Mao Y (2025) **The membrane-proximal external region of human immunodeficiency virus (HIV-1) envelope glycoprotein trimers in A18-lipid nanodiscs** *Communications Biology* **8**: 442.
Raman EP, Guvench O, MacKerell AD, Jr. (2010) **CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses** *The Journal of Physical Chemistry B* **114**: 12981-12994.
Rantalainen K, Berndsen ZT, Antanasijevic A, Schiffner T, Zhang X, Lee WH, Torres JL, Zhang L, Irimia A, Copps J, Zhou KH, Kwon YD, Law WH, Schramm CA, Verardi R, Krebs SJ, Kwong PD, Doria-Rose NA, Wilson IA, Zwick MB, Yates JR, 3rd, Schief WR, Ward AB (2020) **HIV-1 Envelope and MPER Antibody Structures in Lipid Assemblies** *Cell Reports* **31**: 107583.
Reardon PN, Sage H, Dennison SM, Martin JW, Donald BR, Alam SM, Haynes BF, Spicer LD (2014) **Structure of an HIV-1-neutralizing antibody target, the lipid-bound gp41 envelope membrane proximal region trimer** *PNAS* **111**: 1391-1396.
Salzwedel K, West JT, Hunter E (1999) **A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity** *Journal of Virology* **73**: 2469-2480.
Sanders RW, Derking R, Cupo A, Julien JP, Yasmeen A, de Val N, Kim HJ, Blattner C, de la Pena AT, Korzun J, Golabek M, de Los Reyes K, Ketas TJ, van Gils MJ, King CR, Wilson IA, Ward AB, Klasse PJ, Moore JP (2013) **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 Pathogens* **9**: e1003618.

Sarkar A, Bale S, Behrens AJ, Kumar S, Sharma SK, de Val N, Pallesen J, Irimia A, Diwanji DC, Stanfield RL, Ward AB, Crispin M, Wyatt RT, Wilson IA (2018) **Structure of a cleavage-independent HIV Env recapitulates the glycoprotein architecture of the native cleaved trimer** *Nature Communications* **9**: 1956.

Sharma SK, de Val N, Bale S, Guenaga J, Tran K, Feng Y, Dubrovskaya V, Ward AB, Wyatt RT (2015) **Cleavage-independent HIV-1 Env trimers engineered as soluble native spike mimetics for vaccine design** *Cell Reports* **11**: 539-550.

Steinbach PJ, Brooks BR (2004) **New spherical-cutoff methods for long-range forces in macromolecular simulation** *Journal of Computational Chemistry* **15**: 667-683.

Sun ZY, Oh KJ, Kim M, Yu J, Brusic V, Song L, Qiao Z, Wang JH, Wagner G, Reinherz EL (2008) **HIV-1 broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane** *Immunity* **28**: 52-63.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) **GROMACS: fast, flexible, and free** *Journal of Computational Chemistry* **26**: 1701-1718.

Williams LD, Ofek G, Schatzle S, McDaniel JR, Lu X, Nicely NI, Wu L, Lougheed CS, Bradley T, Louder MK, McKee K, Bailer RT, O'Dell S, Georgiev IS, Seaman MS, Parks RJ, Marshall DJ, Anasti K, Yang G, Nie X, Tumba NL, Wiehe K, Wagh K, Korber B, Kepler TB, Munir Alam S, Morris L, Kamanga G, Cohen MS, Bonsignori M, Xia SM, Montefiori DC, Kelsoe G, Gao F, Mascola JR, Moody MA, Saunders KO, Liao HX, Tomaras GD, Georgiou G, Haynes BF (2017) **Potent and broad HIV-neutralizing antibodies in memory B cells and plasma** *Science Immunology* **2**: eaal2200.

Wu EL, Cheng X, Jo S, Rui H, Song KC, Davila-Contreras EM, Qi Y, Lee J, Monje-Galvan V, Venable RM, Klauda JB, Im W (2014) **CHARMM-GUI Membrane Builder toward realistic biological membrane simulations** *Journal of Computational Chemistry* **35**: 1997-2004.

Yang J, Zhang Y (2015) **I-TASSER server: new development for protein structure and function predictions** *Nucleic Acids Research* **43**: W174-181.

Yang S, Hiotis G, Wang Y, Chen J, Wang JH, Kim M, Reinherz EL, Walz T (2022) **Dynamic HIV-1 spike motion creates vulnerability for its membrane-bound tripod to antibody attack** *Nature Communications* **13**: 6393.

Zhang Y, Skolnick J (2005) **TM-align: a protein structure alignment algorithm based on the TM-score** *Nucleic Acids Research* **33**: 2302-2309.

Zhou T, Zhu J, Wu X, Moquin S, Zhang B, Acharya P, Georgiev IS, Altae-Tran HR, Chuang GY, Joyce MG, Kwon YD, Longo NS, Louder MK, Luongo T, McKee K, Schramm CA, Skinner J, Yang Y, Yang Z, Zhang Z, Zheng A, Bonsignori M, Haynes BF, Scheid JF, Nussenzweig MC, Simek M, Burton DR, Koff WC, Program NCS, Mullikin JC, Connors M, Shapiro L, Nabel GJ, Mascola JR, Kwong PD (2013) **Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies** *Immunity* **39**: 245-258.

Article and author information

Author details

Yiwei Cao
Department of Biological Sciences, Lehigh University

Contribution: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing

Competing interests: No competing interests declared

Wonpil Im
Department of Biological Sciences, Lehigh University

Contribution: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing

For correspondence: wonpil@lehigh.edu

Competing interests: No competing interests declared

Funding
National Institutes of Health (R35-GM153458)

Figures

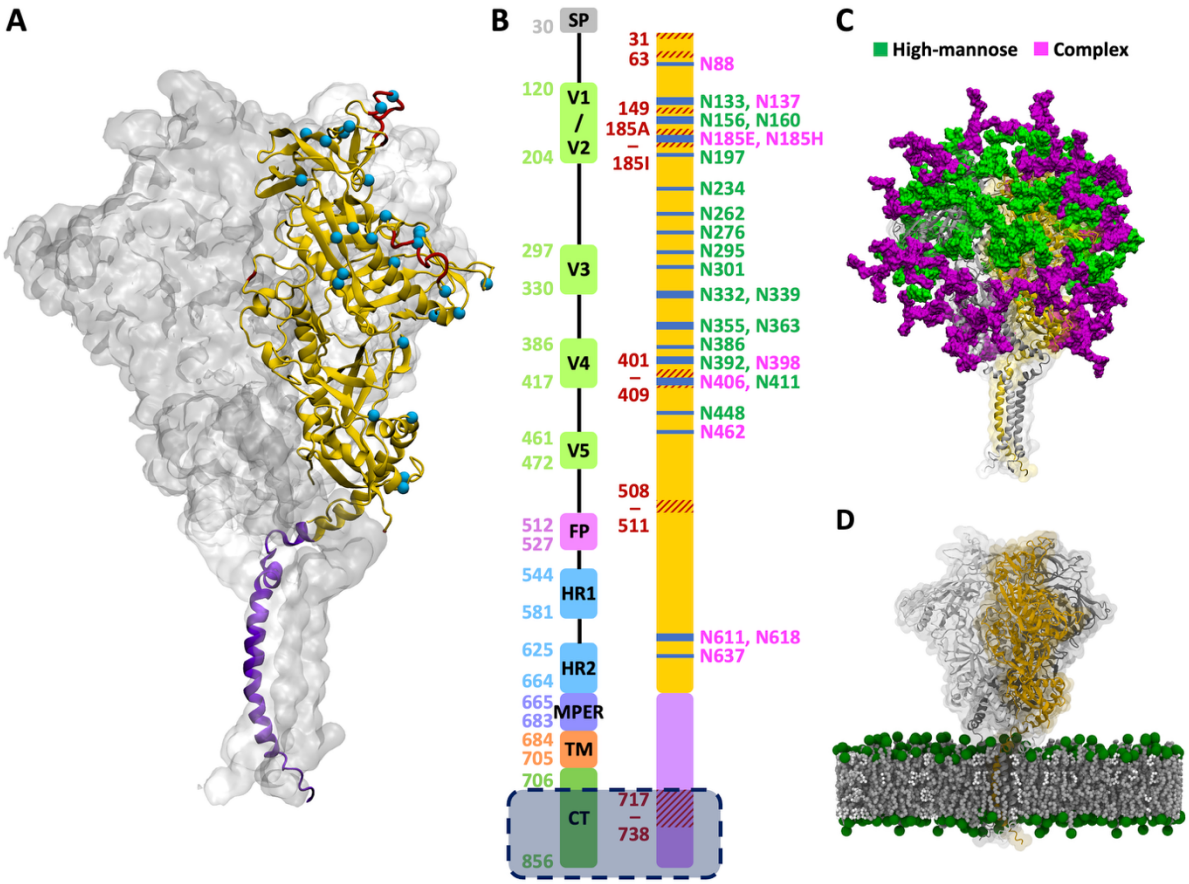

Figure 1. Model structure of a fully glycosylated full-length HIV Env trimer embedded in a membrane. (**A**) The model structure built by combining the cryo-EM structure of the ectodomain (yellow, PDB ID: 6B0N) with the NMR structure of the MPER and TMD (purple, PDB ID: 7LOI). The missing loops in the PDB structures are highlighted in red, and the glycosylation sites are marked by cyan spheres. (**B**) Left: assignment of functional domains with boundary residue numbers, including signal peptide (SP), variable regions (V1-V5), fusion peptide (FP), heptad repeats (HR1 and HR2), membrane-proximal external region (MPER), transmembrane domain (TMD), and cytoplasmic tail (CT). Right: missing loops (red) and glycosylation sites (blue). The shaded region at the bottom marks CT residues excluded in the CT-truncated model (see Figure 1—figure supplement 1 for a model with the full CT included). (**C**) N-linked glycans with high-mannose (green) and complex (magenta) types. (**D**) Env trimer embedded in a membrane. Lipid headgroups are highlighted by green spheres and glycans are omitted for visual clarity. Molecular illustrations were prepared using Visual Molecular Dynamics (VMD) (Humphrey et al., 1996).

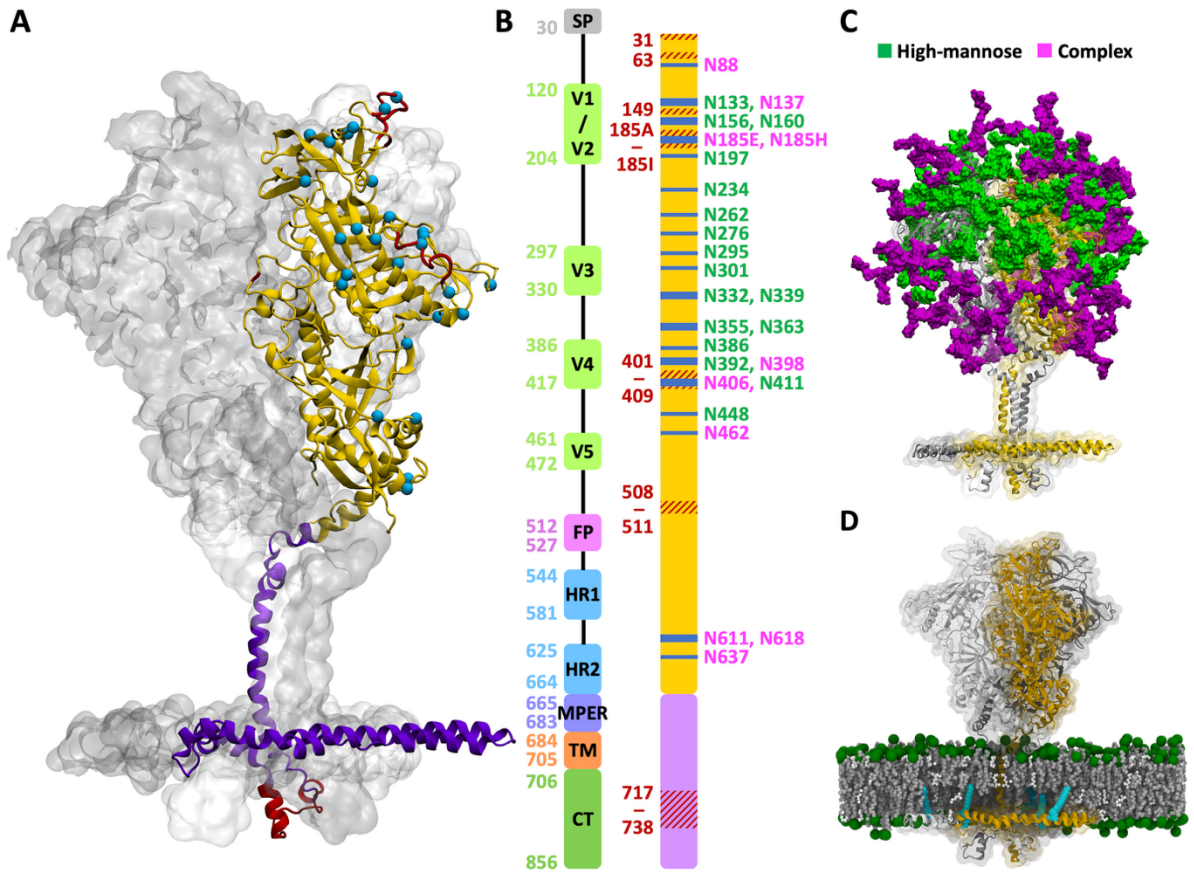

Figure 1—figure supplement 1. Model structure of the fully glycosylated full-length HIV Env trimer embedded in a membrane, including the cytoplasmic tail. Labeling and color coding are the same as in Figure 1. **(D)** The palmitoyl groups covalently attached to C764 and C837 are shown in cyan.

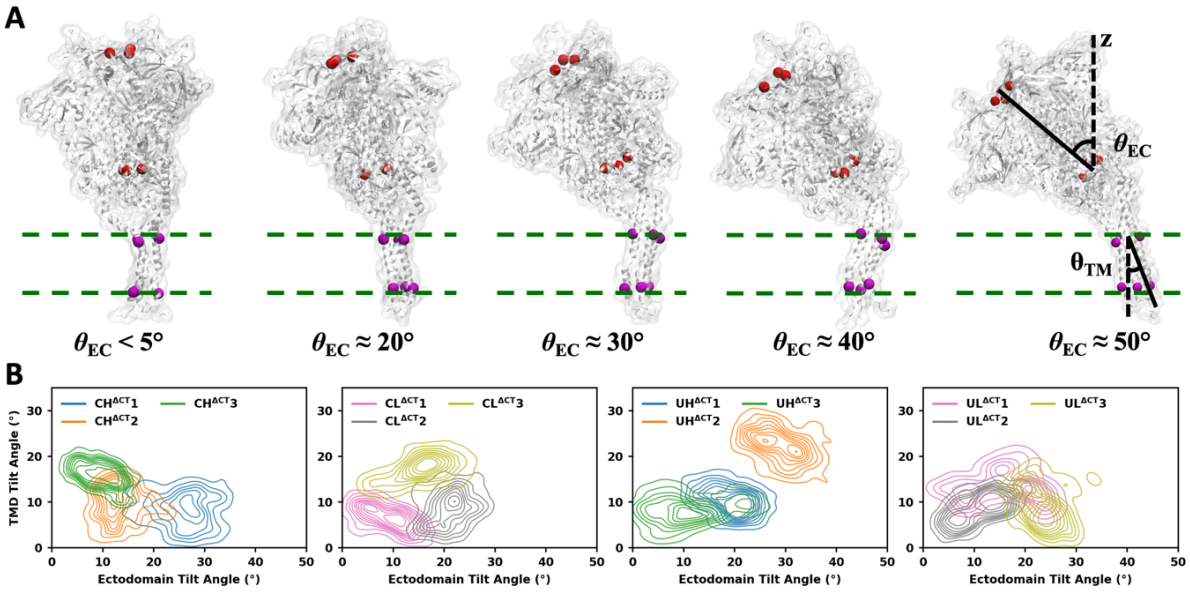

Figure 2. Tilting motions of the ectodomain and TMD are independent. (**A**) Representative structures illustrating different ectodomain tilt angles and the schematic showing how tilt angles are calculated. (**B**) Ectodomain tilt versus TMD tilt, calculated from CT-truncated systems with various initial configurations.

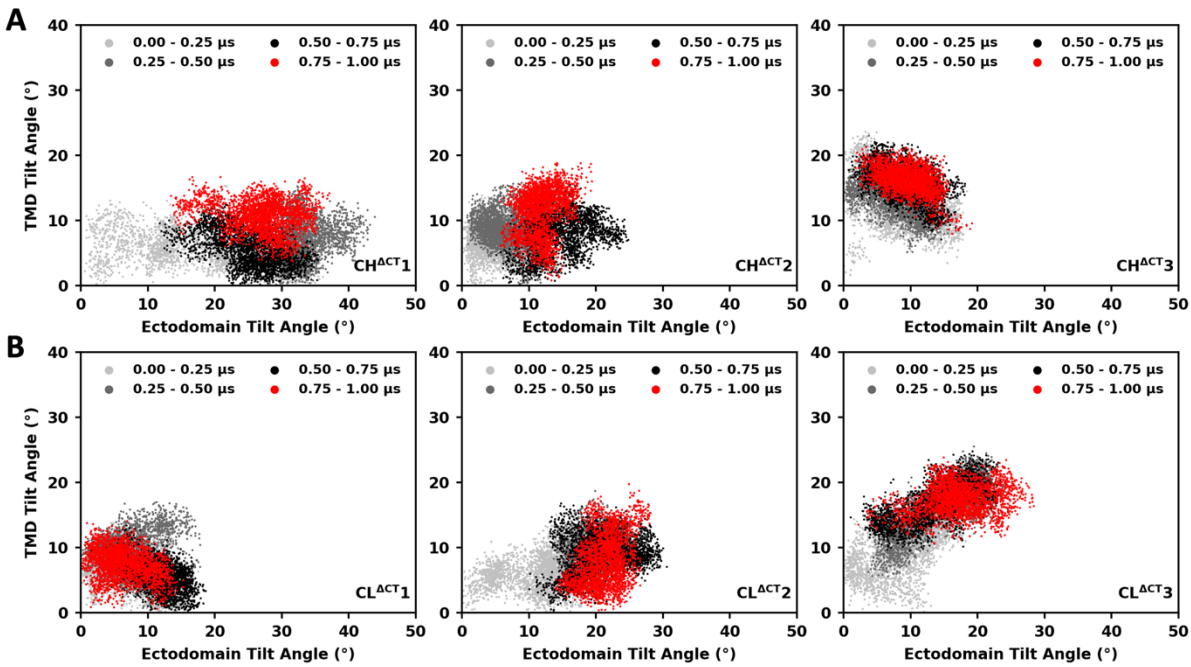

Figure 2—figure supplement 1. Ectodomain tilt versus TMD tilt, grouped by time intervals (cleaved CT-truncated systems). (**A**) Three trajectories starting from the "high" TMD configuration. (**B**) Three trajectories starting from the "low" TMD configuration. The 1-µs trajectory was divided into four intervals, with values from each interval shown in light gray, dark gray, black, and red, respectively.

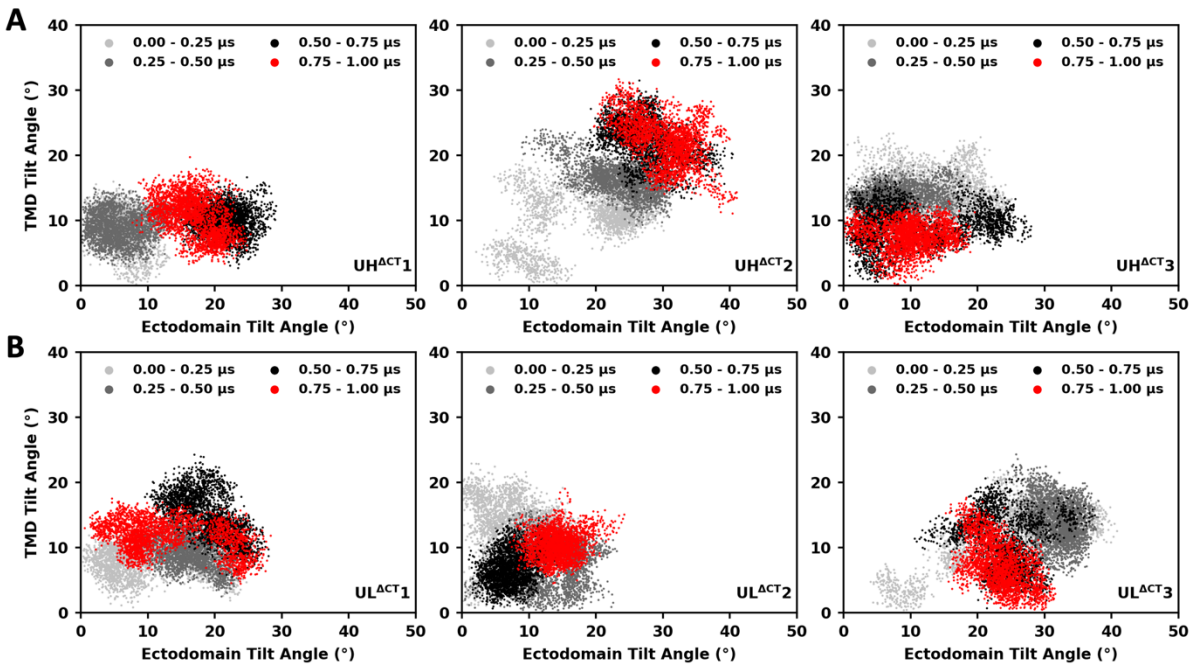

Figure 2—figure supplement 2. Ectodomain tilt versus TMD tilt, grouped by time intervals (uncleaved CT-truncated systems). Labeling and color coding are the same as in Figure 2—figure supplement 1.

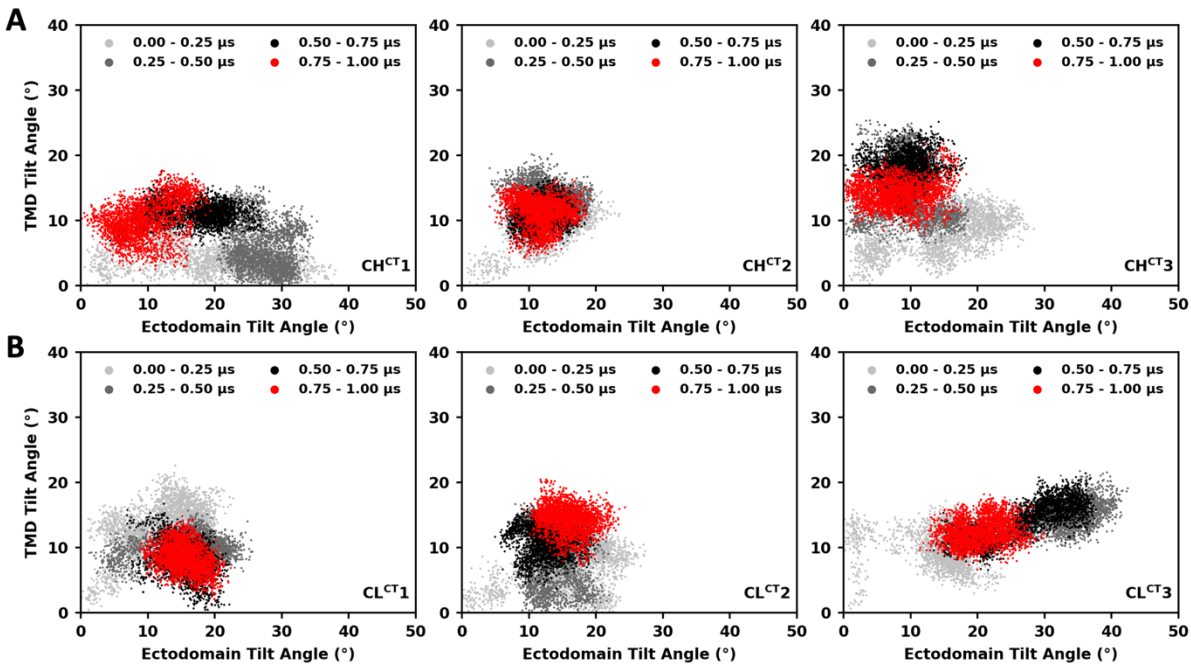

Figure 2—figure supplement 3. Ectodomain tilt versus TMD tilt, grouped by time intervals (cleaved full-length systems). Labeling and color coding are the same as in Figure 2—figure supplement 1.

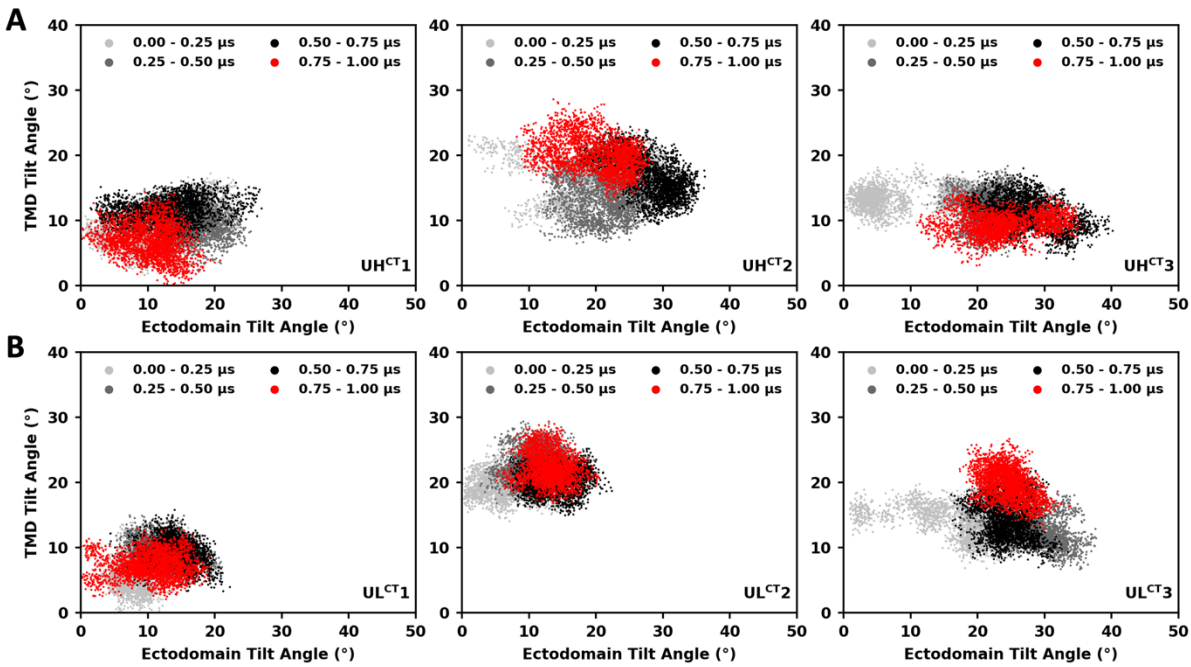

Figure 2—figure supplement 4. Ectodomain tilt versus TMD tilt, grouped by time intervals (uncleaved full-length systems). Labeling and color coding are the same as in Figure 2—figure supplement 1.

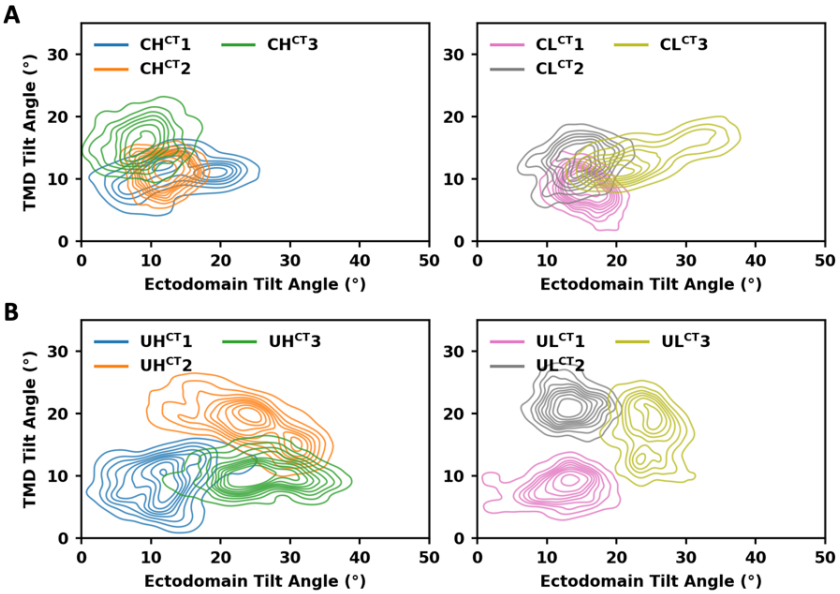

Figure 2—figure supplement 5. Ectodomain tilt versus TMD tilt (full-length systems).

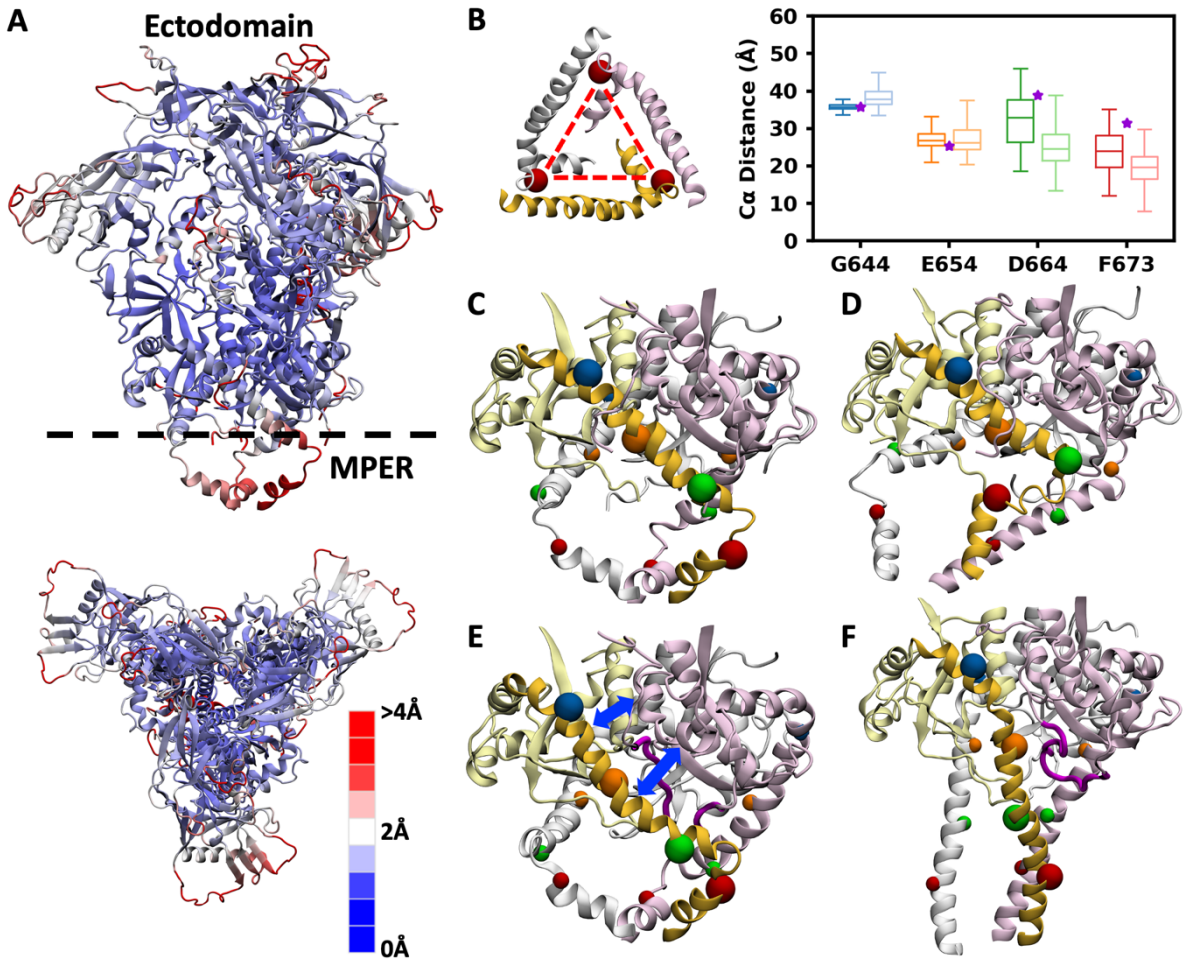

Figure 3. Ectodomain is rigid, whereas the MPER is highly flexible and adopts diverse conformations. (**A**) Top and side views of the ectodomain and MPER in the cleaved system, with RMSF indicated by color. (**B**) Schematic illustrating the calculation of interchain distance and its distributions measured at the Cα atoms of G644, E654, D664, and F673. Cleaved and uncleaved systems are represented by solid and transparent colors, respectively. The initial values of interchain distances are marked by purple stars. (**C–F**) Local structures of the ectodomain C-terminus and MPER. The HR2 helix and MPER in one protomer are highlighted in dark yellow, with the Cα atoms of four selected residues marked by blue, orange, green and red spheres. (**C**) The initial conformation and (**D**) representative snapshot from simulations of the cleaved system. (**E**) The initial conformation and (**F**) representative snapshot from simulations of the uncleaved system.

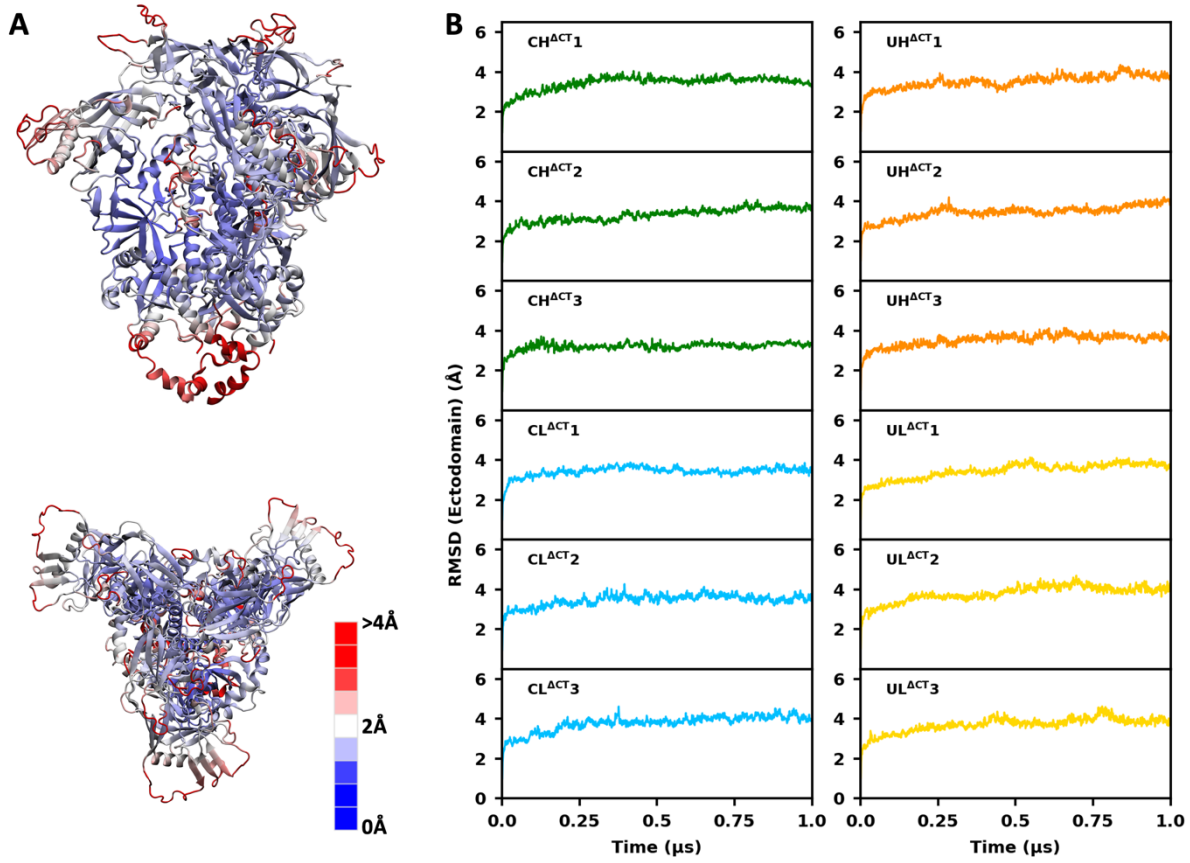

Figure 3—figure supplement 1. RMSF and RMSD of the ectodomain. (**A**) Top and side views of the ectodomain and MPER in the uncleaved system, with RMSF indicated by color. (**B**) RMSD relative to the initial model as a function of time, calculated from the trajectories of CT-truncated systems.

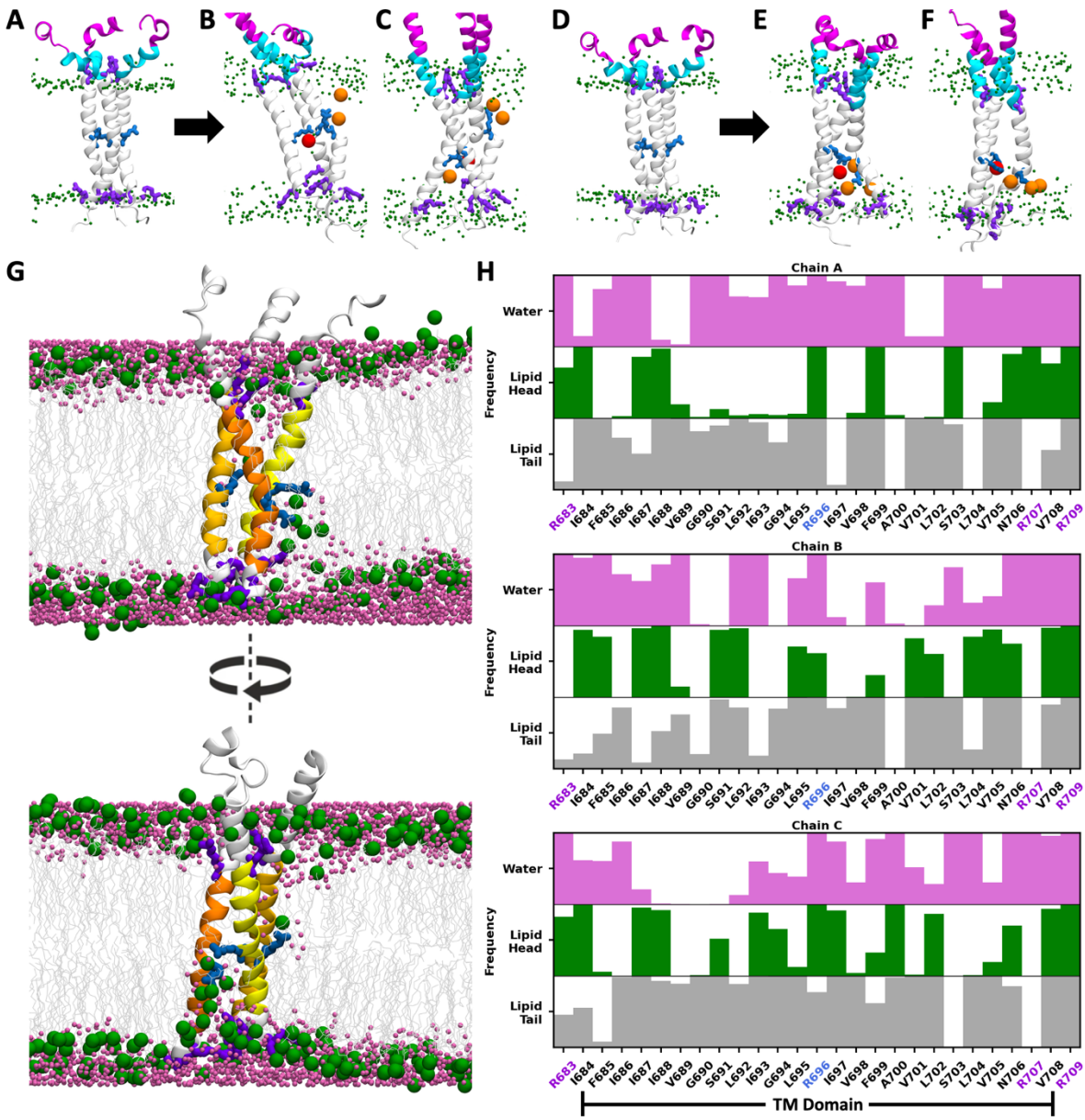

Figure 4. R696 interacts with lipid headgroups and disrupts membrane integrity. (**A–C**) MPER and TMD in the CT-truncated system with the "high" TMD configuration. MPER-N, MPER-C, and TMD are shown in magenta, cyan, and white, respectively. Lipid headgroups, R696, and the residues anchored in the lipid headgroups (R683, R707 and R709) are shown in green, blue, and purple, respectively. Lipid headgroups and ions interacting with R696 are highlighted in orange and red, respectively. (**A**) Initial conformation. (**B, C**) Representative snapshots from different trajectories. (**D–F**) MPER and TMD in the CT-truncated system with the "low" TMD configuration. (**G**) Two side views of the same snapshot where R696 of one protomer interacts with lipid headgroups in the exoplasmic leaflet and R696 of two protomers interact with lipid headgroups in the cytoplasmic leaflet. Lipid headgroups and tails are shown in green and gray, and water molecules in magenta. TMD of three protomers (i.e., chains A, B and C) are shown in light yellow, dark yellow and orange, respectively. (**H**) Frequency of TMD residues interacting with lipid headgroups, lipid tails, and water. Bar shading reflects interaction frequency, from fully filled (100%) to unfilled (0%).

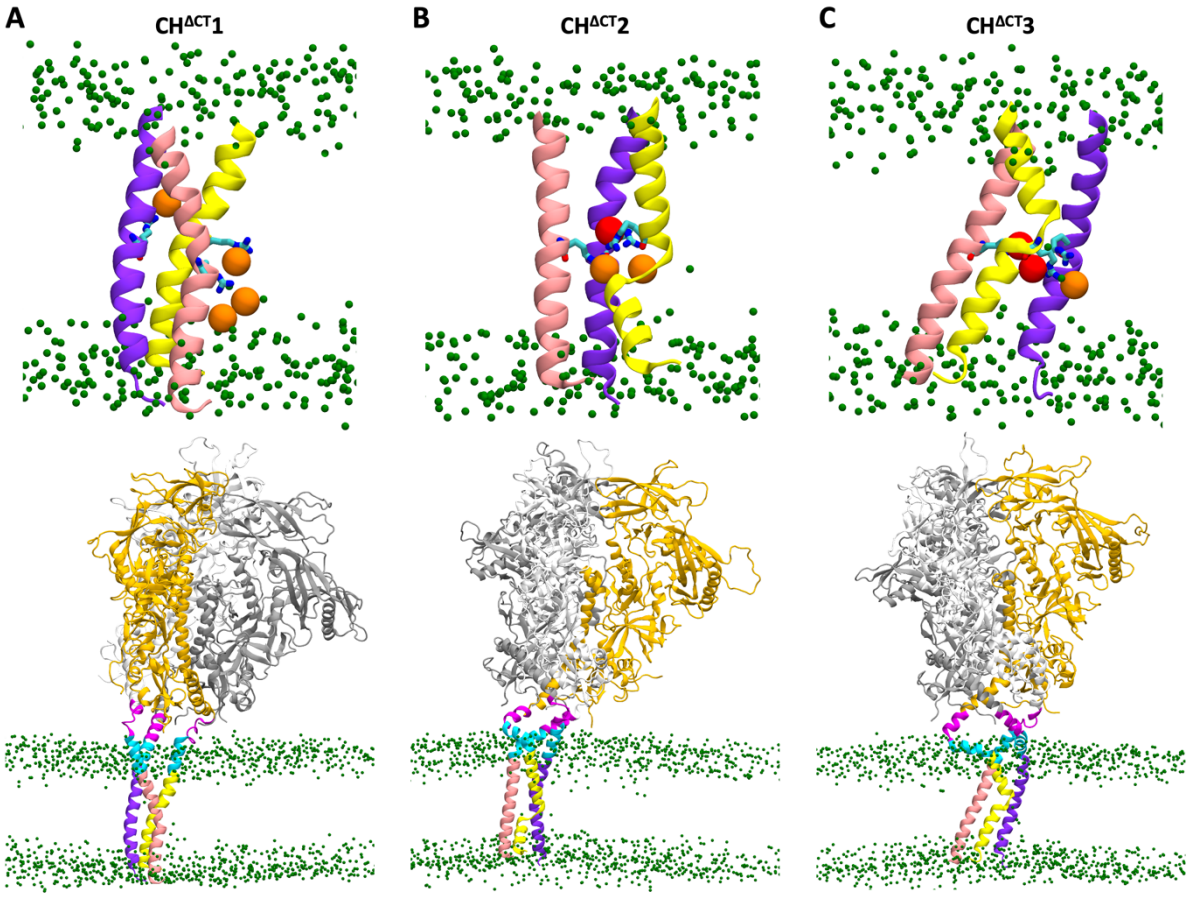

Figure 4—figure supplement 1. Local conformations of the TMD and global conformations of protein and membrane (CH$^{\Delta CT}$ systems). Three protomers of the TMD are shown in light yellow, purple, and pink; three protomers of the ectodomain in dark yellow, gray, and white; MPER-N and MPER-C in magenta and cyan, respectively; and lipid headgroups in green. Lipid headgroups interacting with R696 are highlighted in orange, and the ions interacting with R696 in red.

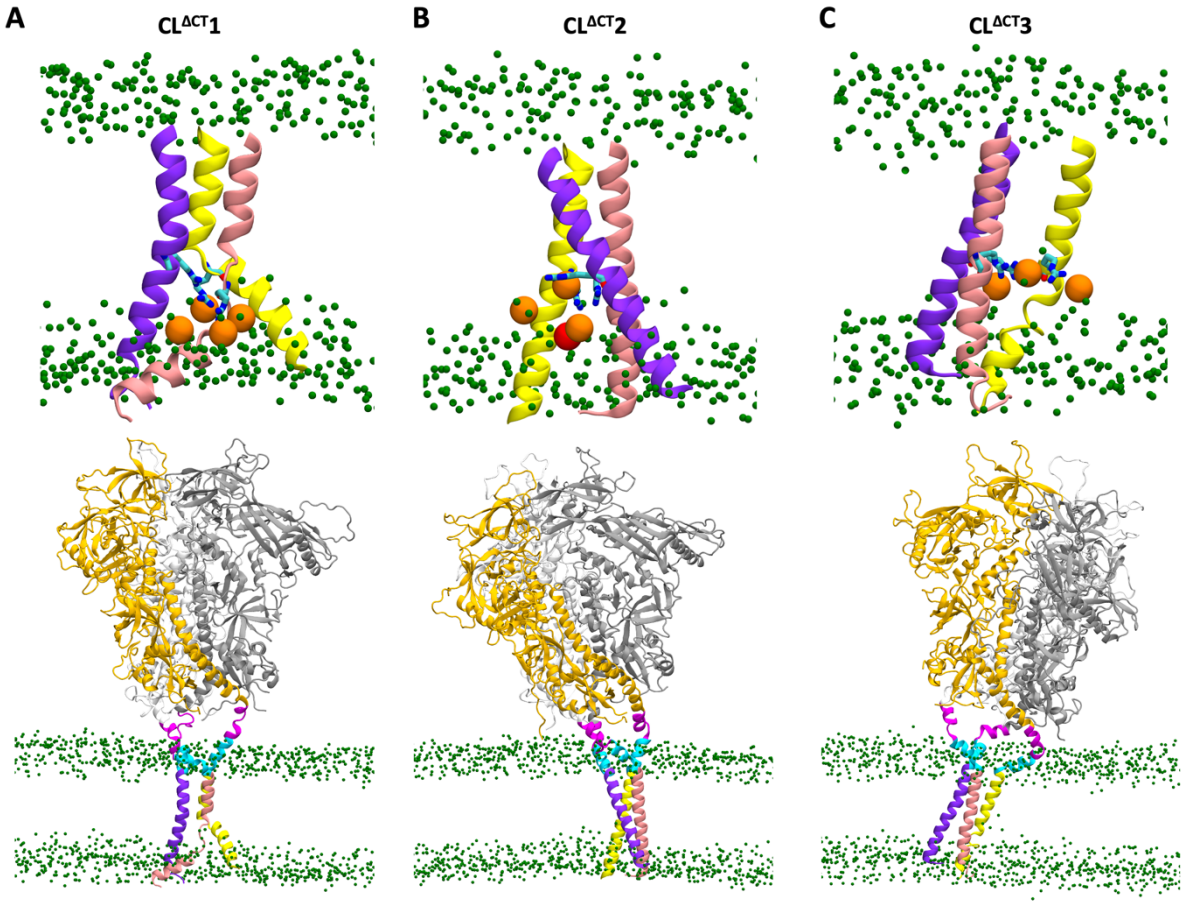

Figure 4—figure supplement 2. Local conformations of the TMD and global conformations of protein and membrane (CL$^{\Delta CT}$ systems). Labeling and color coding are the same as in Figure 4—figure supplement 1.

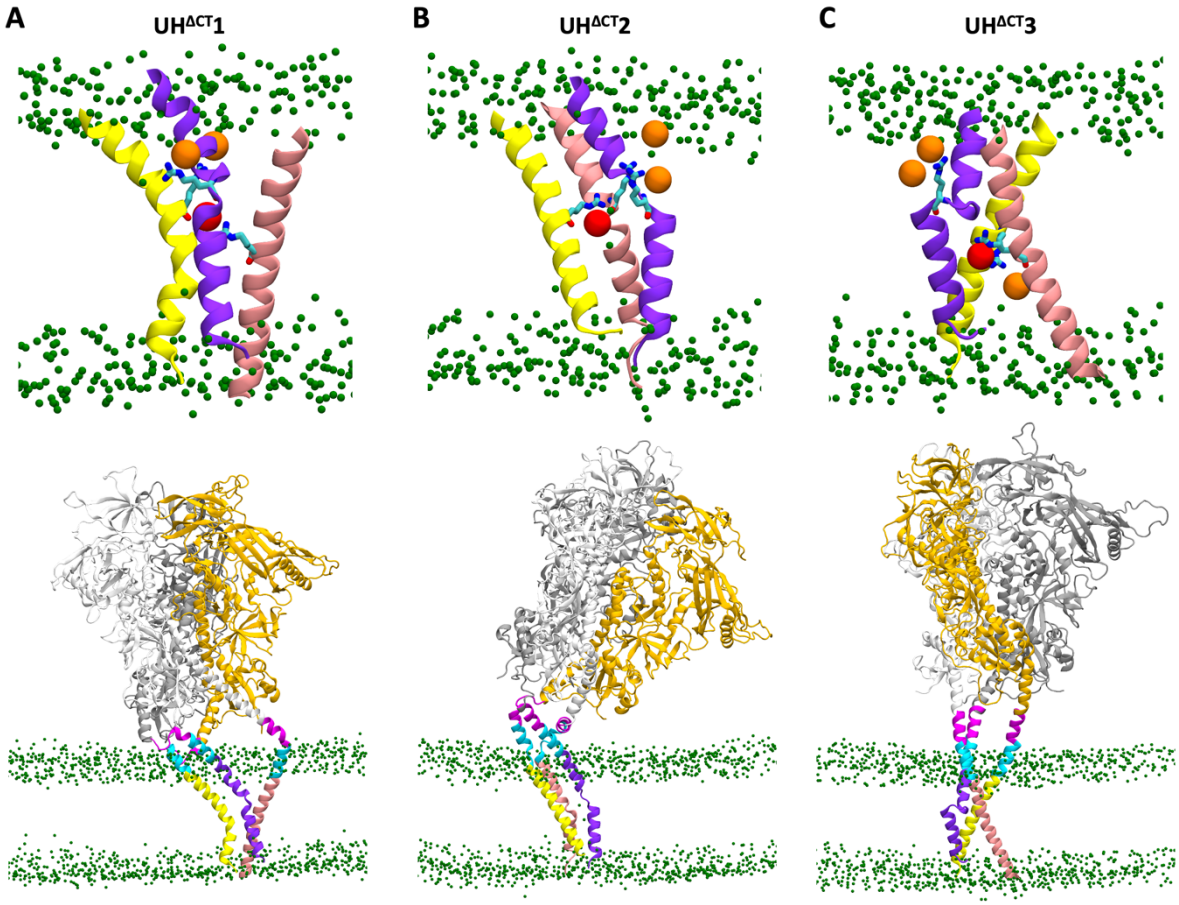

Figure 4—figure supplement 3. Local conformations of the TMD and global conformations of protein and membrane (UH$^{\Delta CT}$ systems). Labeling and color coding are the same as in Figure 4—figure supplement 1.

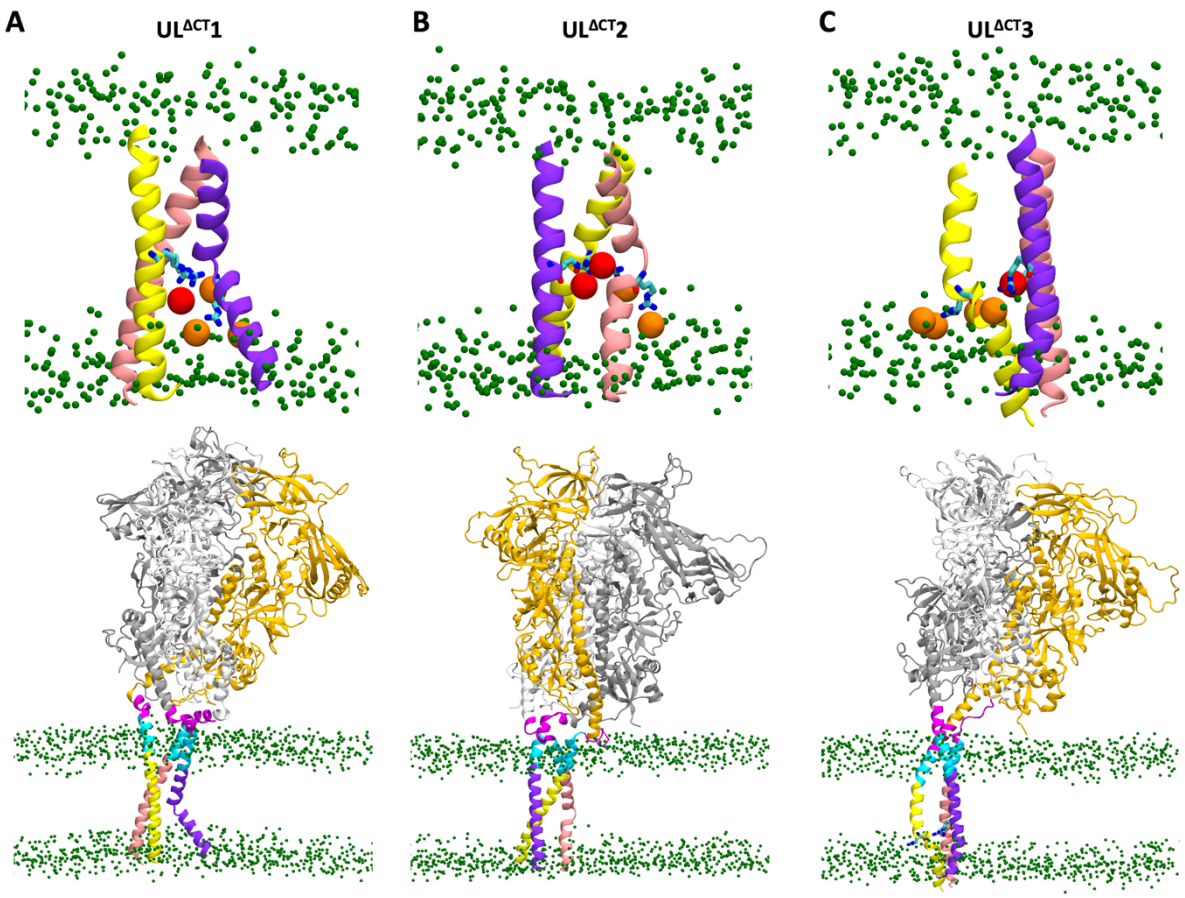

Figure 4—figure supplement 4. Local conformations of the TMD and global conformations of protein and membrane (UL$^{\Delta CT}$ systems). Labeling and color coding are the same as in Figure 4—figure supplement 1.

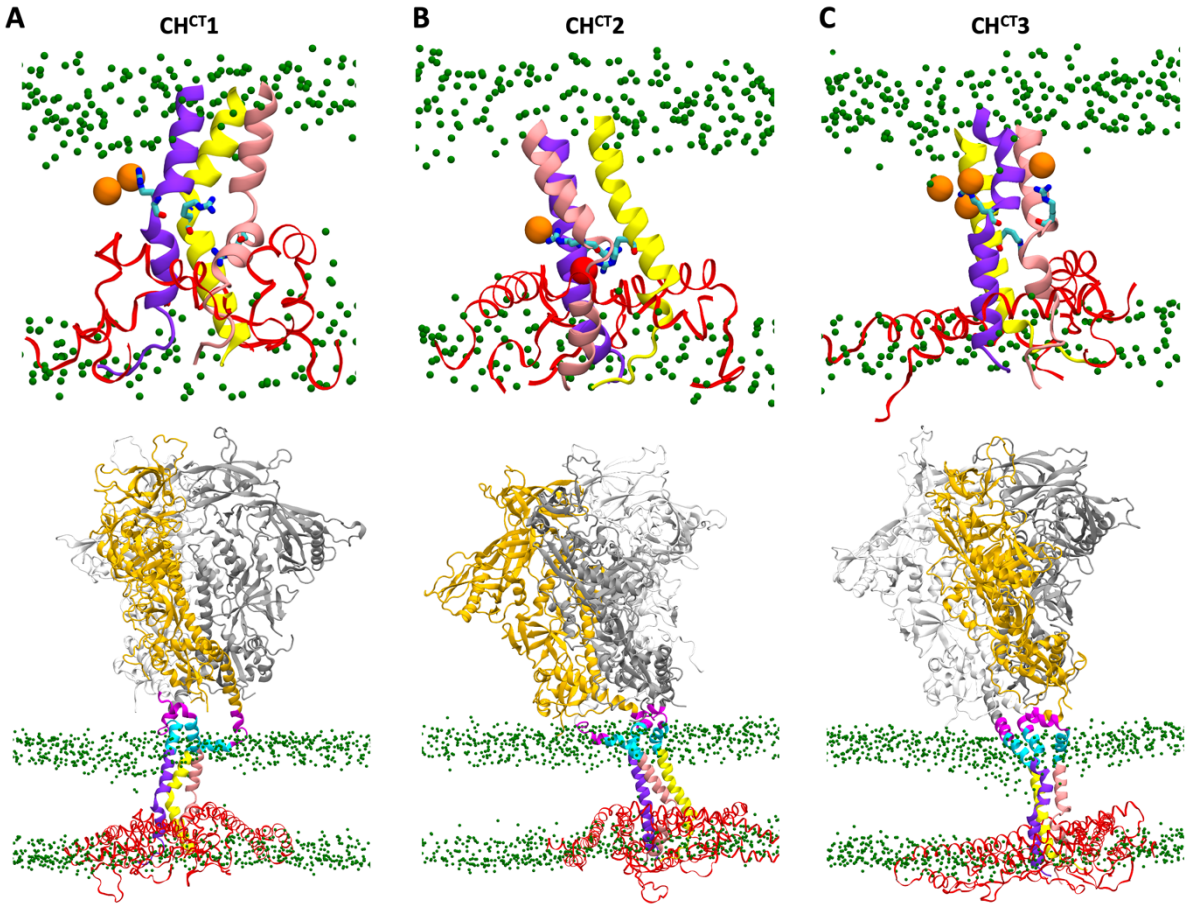

Figure 4—figure supplement 5. Local conformations of the TMD and global conformations of protein and membrane (CH^{CT} systems). Labeling and color coding are the same as in Figure 4—figure supplement 1, with the CT additionally shown in red.

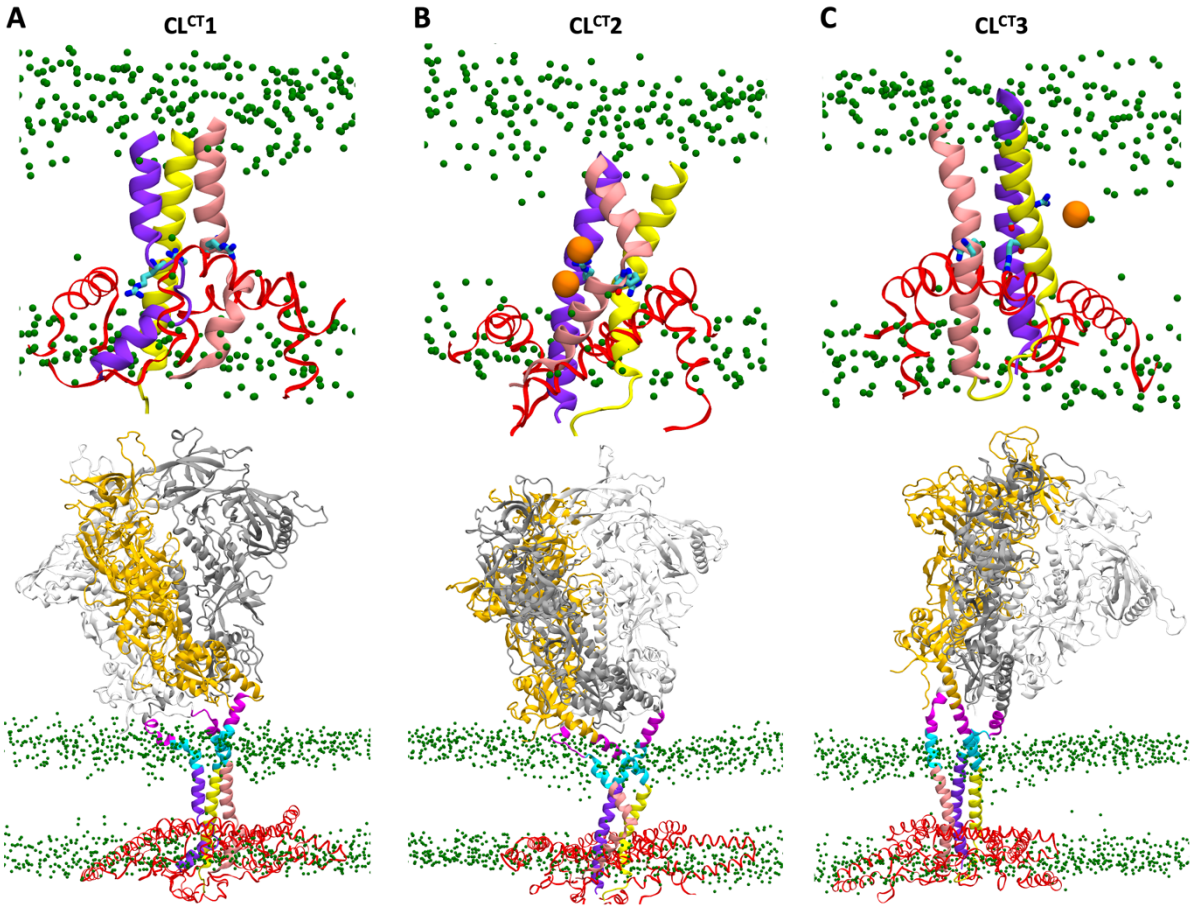

Figure 4—figure supplement 6. Local conformation of the TMD and global conformation of protein and membrane (CL^{CT} systems). Labeling and color coding are the same as in Figure 4—figure supplement 5.

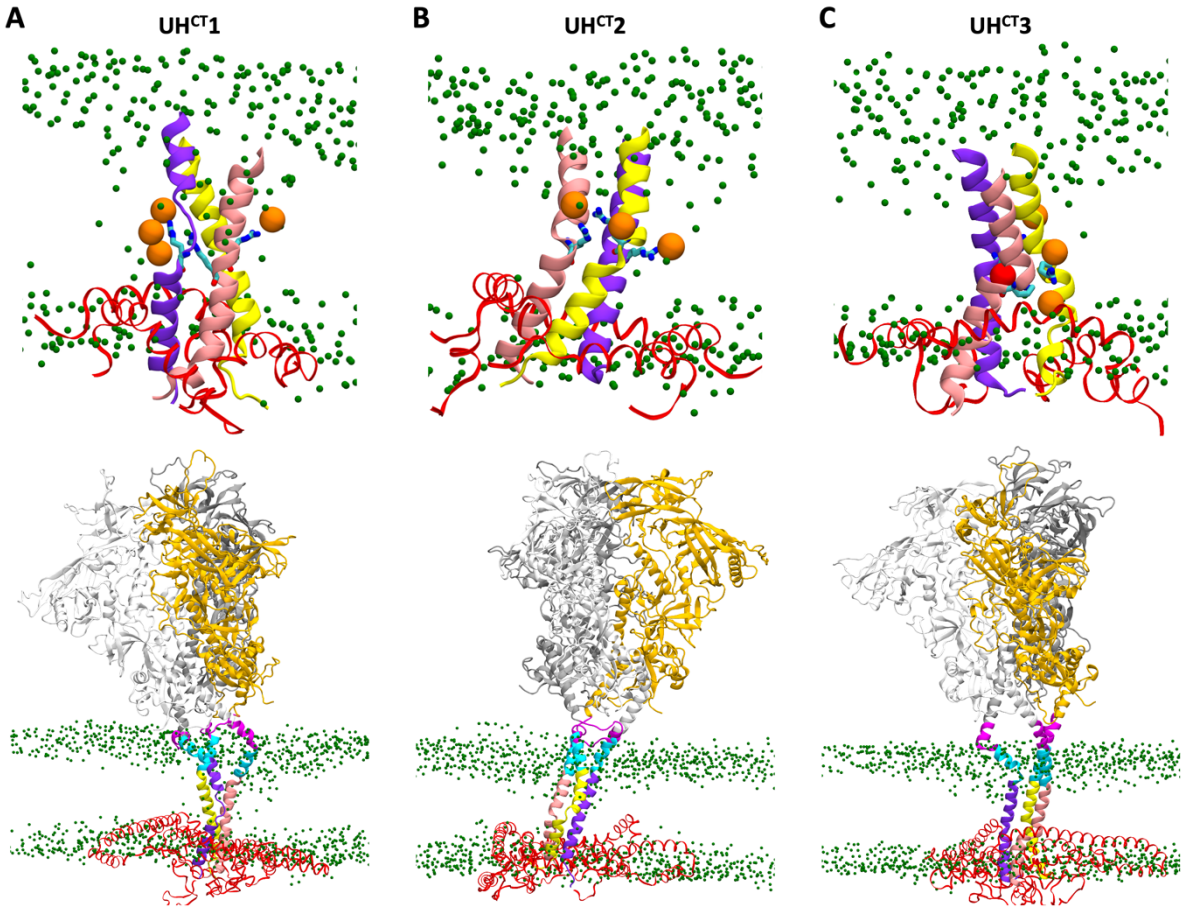

Figure 4—figure supplement 7. Local conformations of the TMD and global conformations of protein and membrane (UH^{CT} systems). Labeling and color coding are the same as in Figure 4—figure supplement 5.

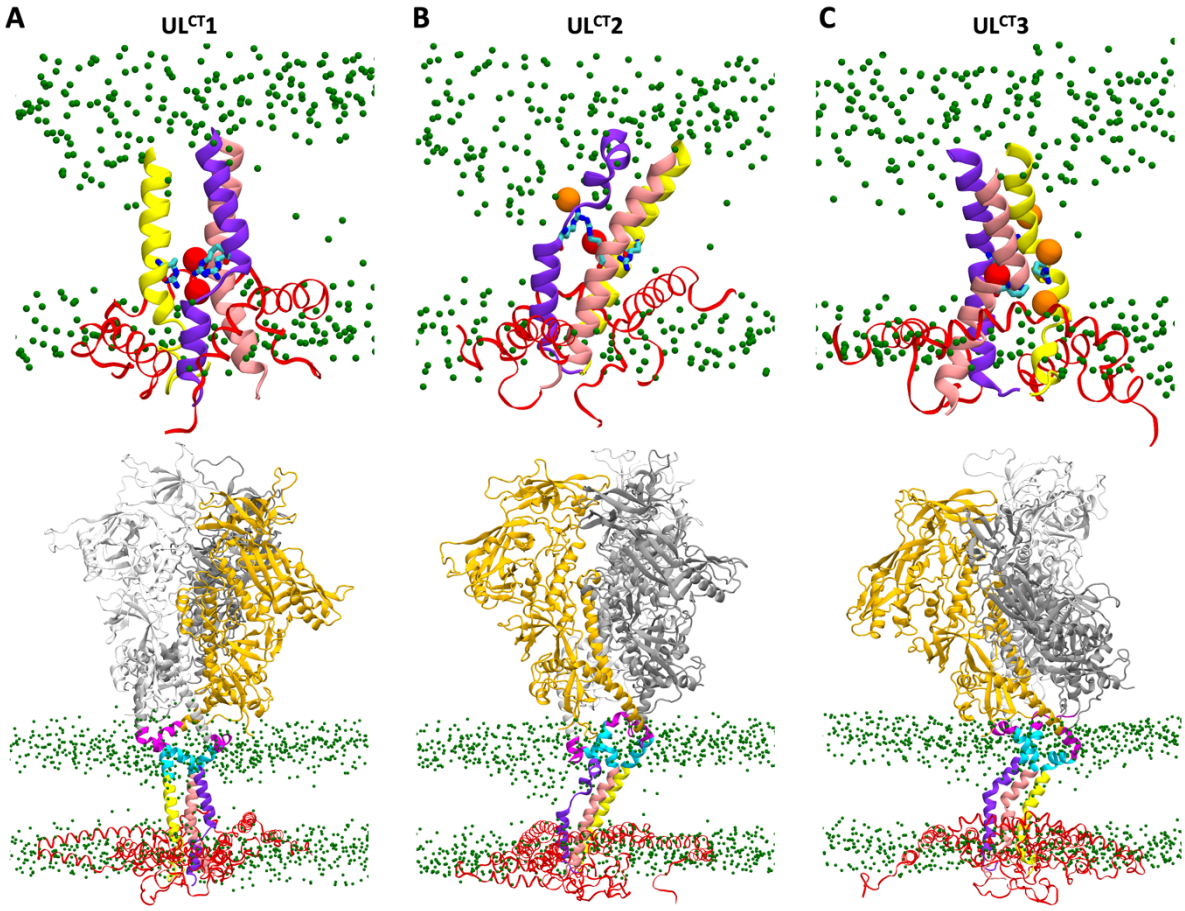

Figure 4—figure supplement 8. Local conformations of the TMD and global conformations of protein and membrane (UL^{CT} systems). Labeling and color coding are the same as in Figure 4—figure supplement 5.

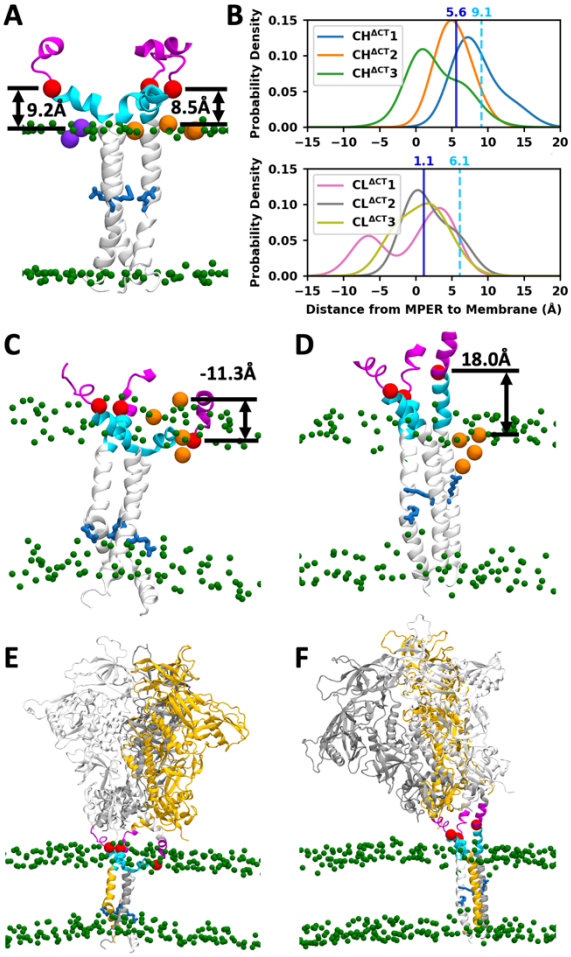

Figure 5. MPER exhibits diverse conformations, and its exposure depends on both MPER and TMD. (**A**) The initial structure of the $CH^{\Delta CT}$ system, where d_{F673} of two promoters equals 8.5 Å and 9.2 Å. Lipid headgroups are shown in green and R696 in blue. d_{F673} is defined as the distance from the Cα of F673 (red) to the highest among the adjacent lipid headgroups (orange and purple). (**B**) Distribution of d_{F673} in the $CL^{\Delta CT}$ and $CH^{\Delta CT}$ systems. The cyan dashed line indicates the mean d_{F673} of three protomers in the initial structure, and the blue solid line indicates the mean across all data sampled from simulations. (**C, D**) Representative snapshots illustrating the buried (**C**) and exposed (**D**) MPER. (**E, F**) The entire trimer structures corresponding to (**C**) and (**D**), respectively.

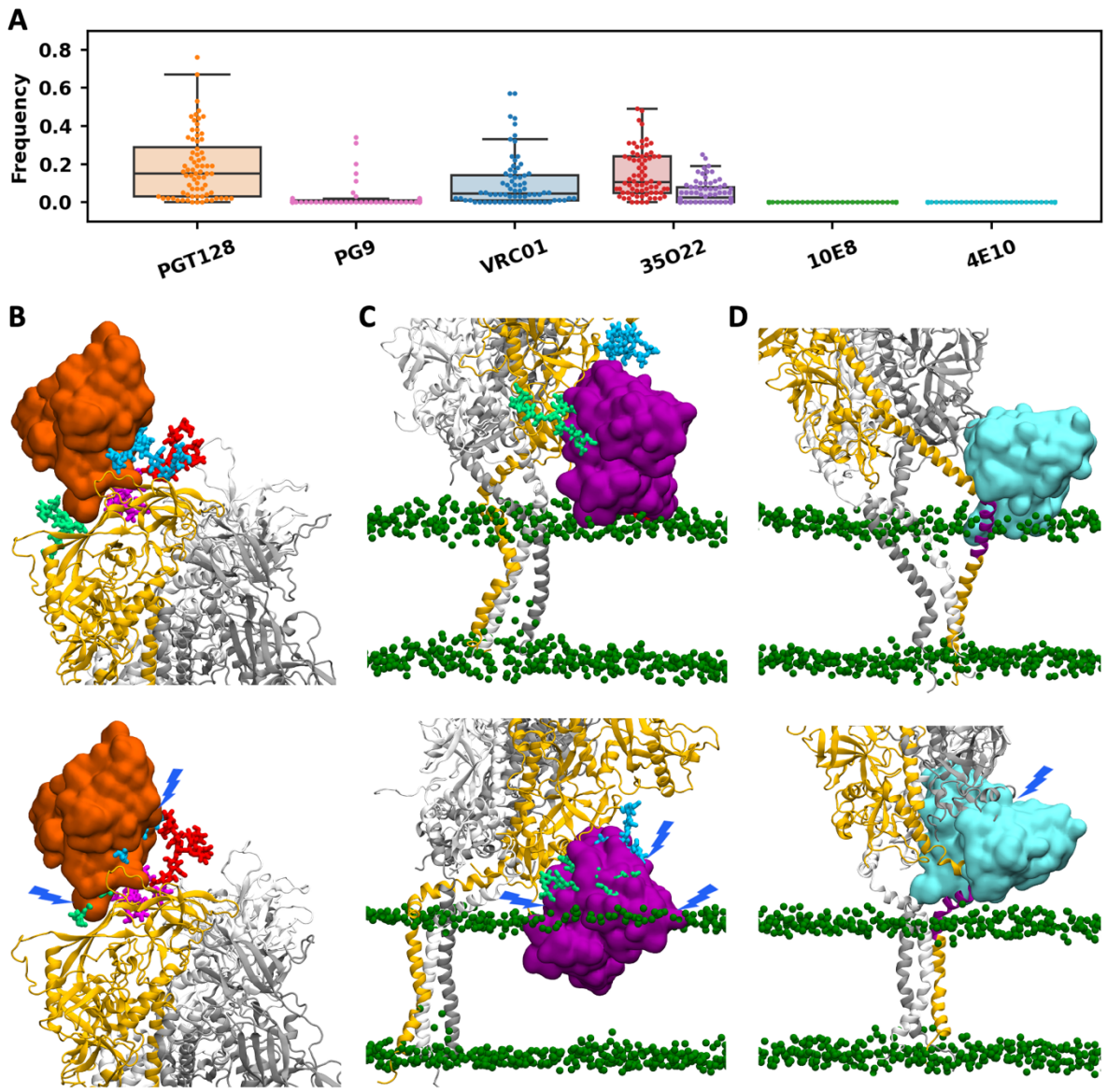

Figure 6. Antibody epitope accessibility. (**A**) The frequency of accessibility. Each marker represents the epitope on one of the three protomers across all trajectories. For 35O22, red indicates the accessibility frequency without considering steric clashes with the membrane, while purple indicates the frequency accounting for clashes with the membrane. (**B–D**) Representative snapshots showing conformations with the epitope exposed (upper) and shielded (lower) for antibodies PGT128, 35O22, and 4E10, respectively. The antibody V_H and V_L domains are shown in surface representation, with lipid head groups in green spheres and glycans that may interfere with the antibody in distinct colors.

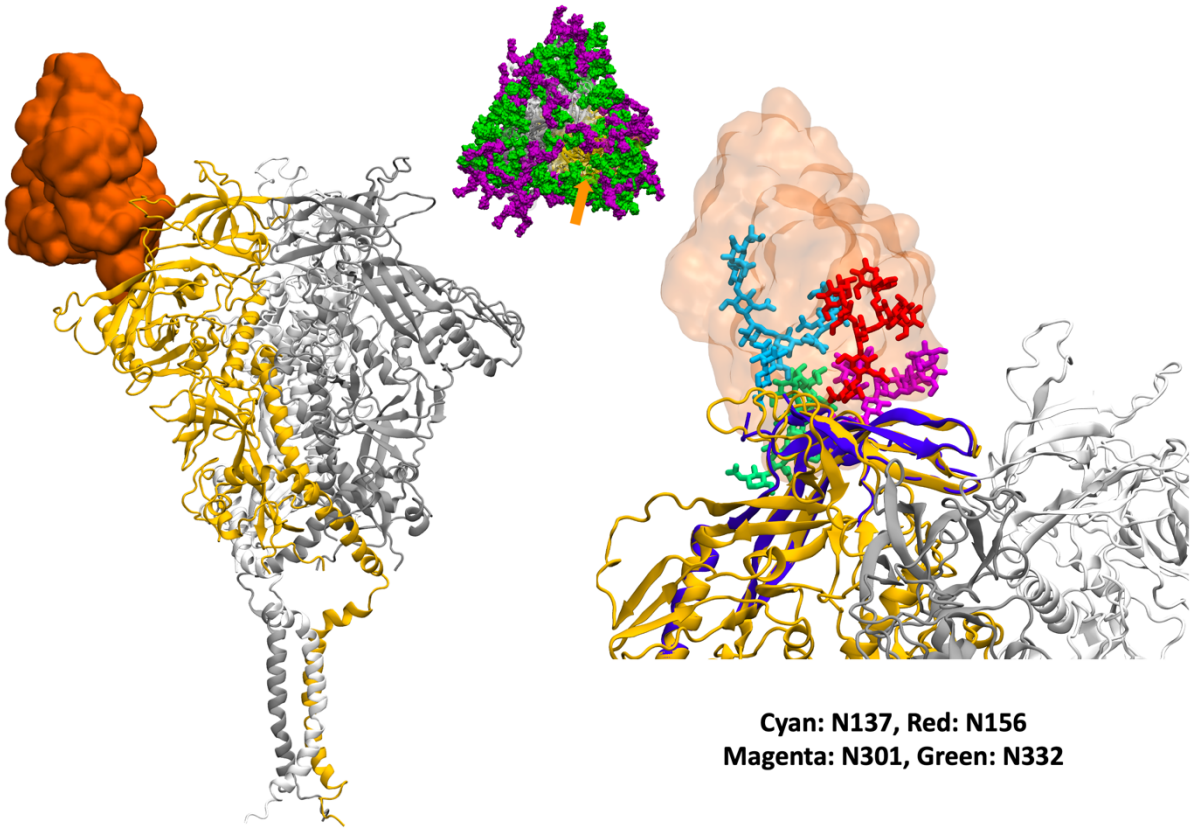

Figure 6—figure supplement 1. Shielding of antibody PGT128 epitope. (**Left**) Variable domains of the heavy and light chains aligned onto our modeled structure. (**Middle**) Top view of the glycosylated trimeric protein, with the orange arrow indicating the epitope location. (**Right**) Part of the epitope in the antibody-epitope complex PDB structure was used for structural alignment and is highlighted in purple. Glycans capable of hindering antibody binding are shown in various colors.

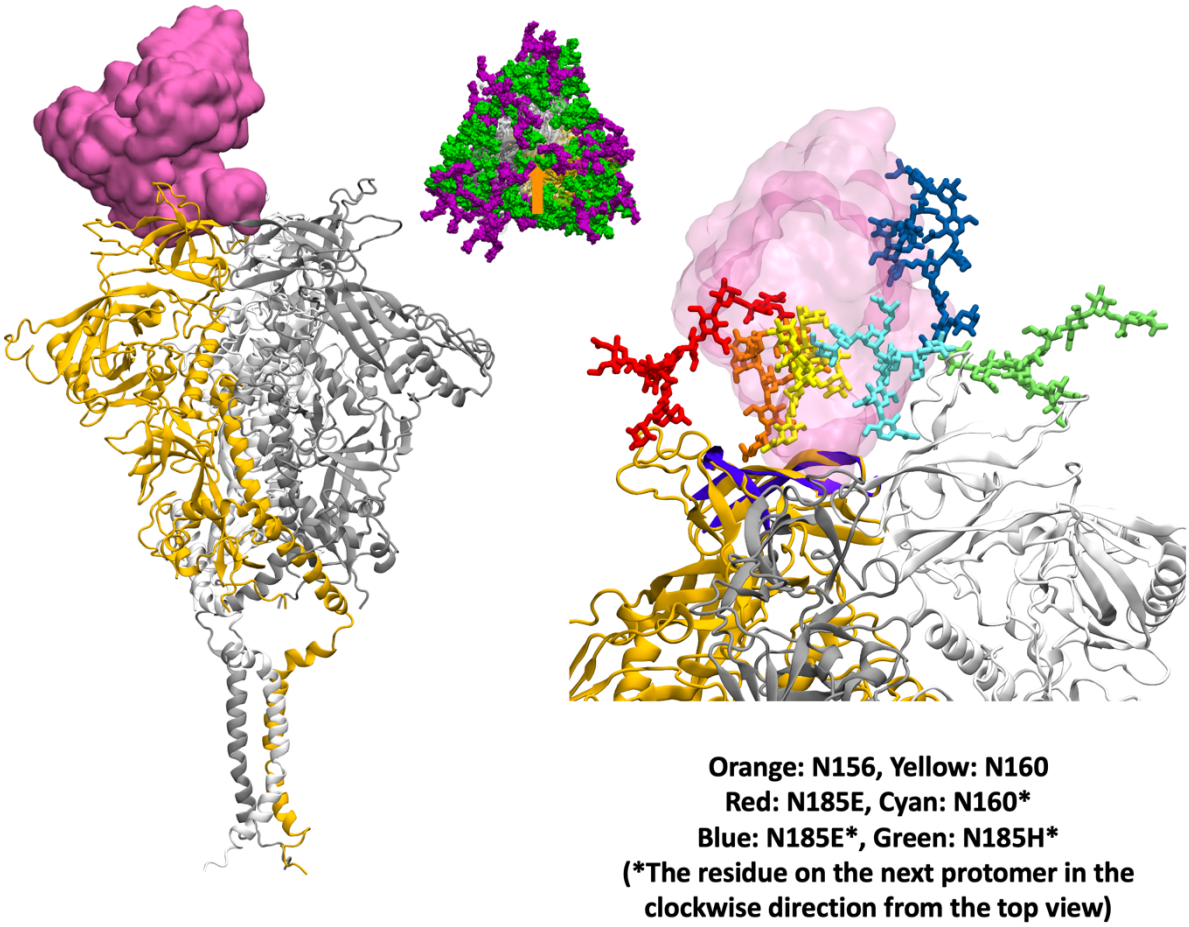

Figure 6—figure supplement 2. Shielding of antibody PG9 epitope. Labeling and color coding are the same as in Figure 6—figure supplement 1.

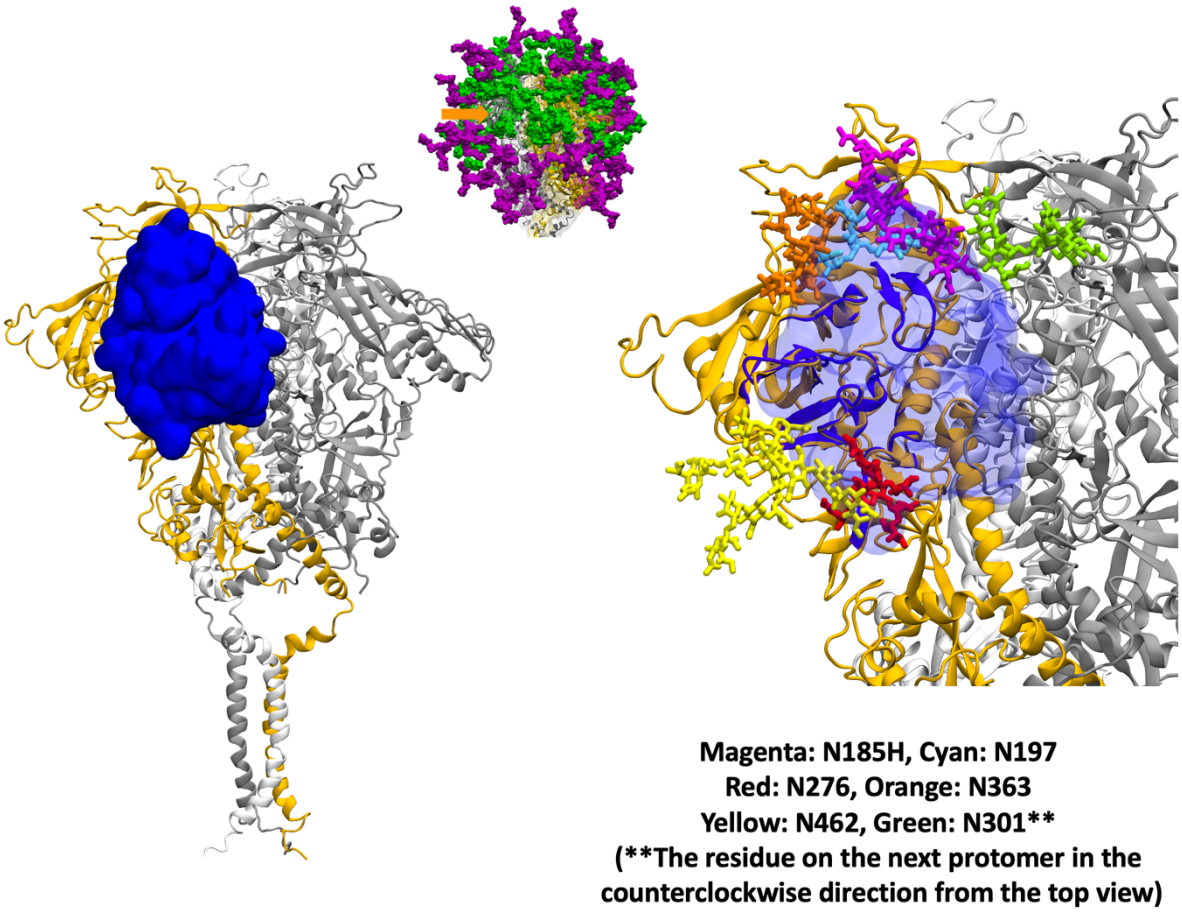

Figure 6—figure supplement 3. Shielding of antibody VRC01 epitope. Labeling and color coding are the same as in Figure 6—figure supplement 1. (**Middle**) The glycosylated trimeric protein is shown in a side view instead of the top view.

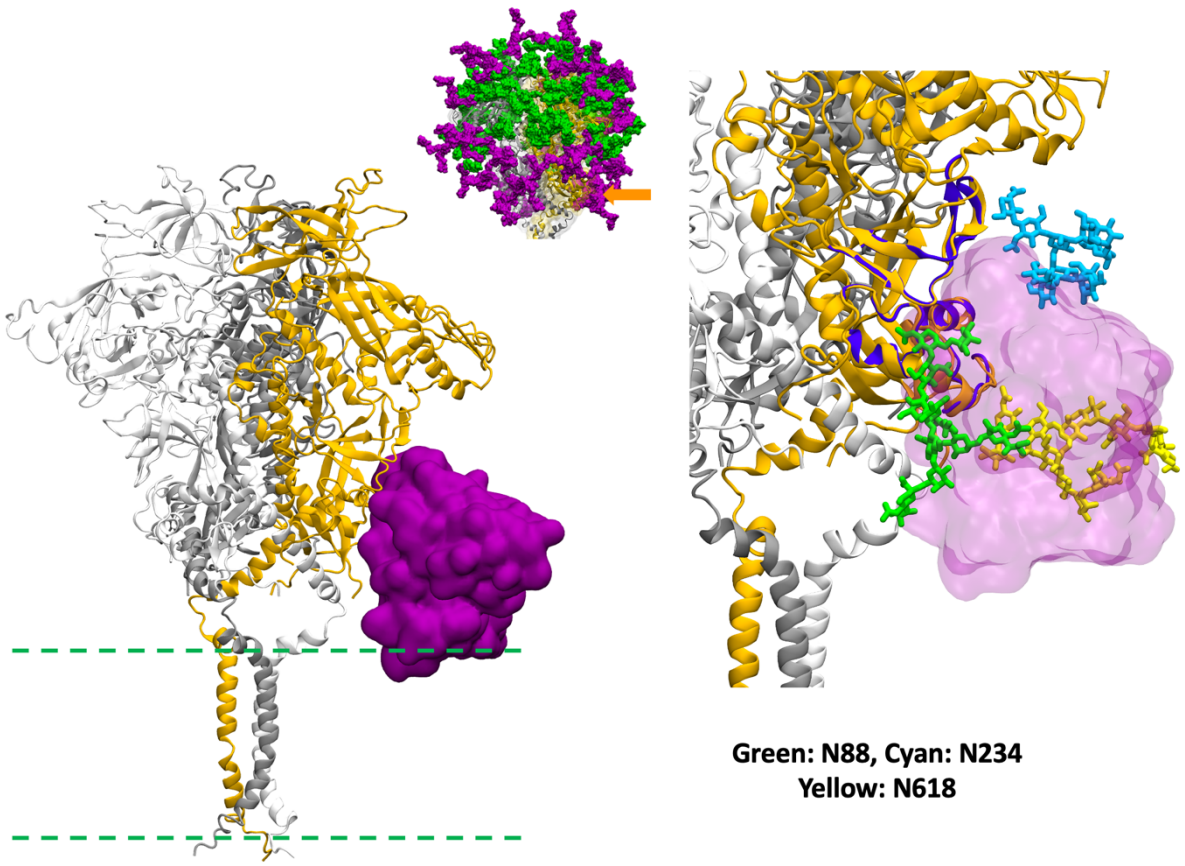

Figure 6—figure supplement 4. Shielding of antibody 35O22 epitope. Labeling and color coding are the same as in Figure 6—figure supplement 3. (**Left**) The dashed greens indicate the approximate location of the lipid headgroups.

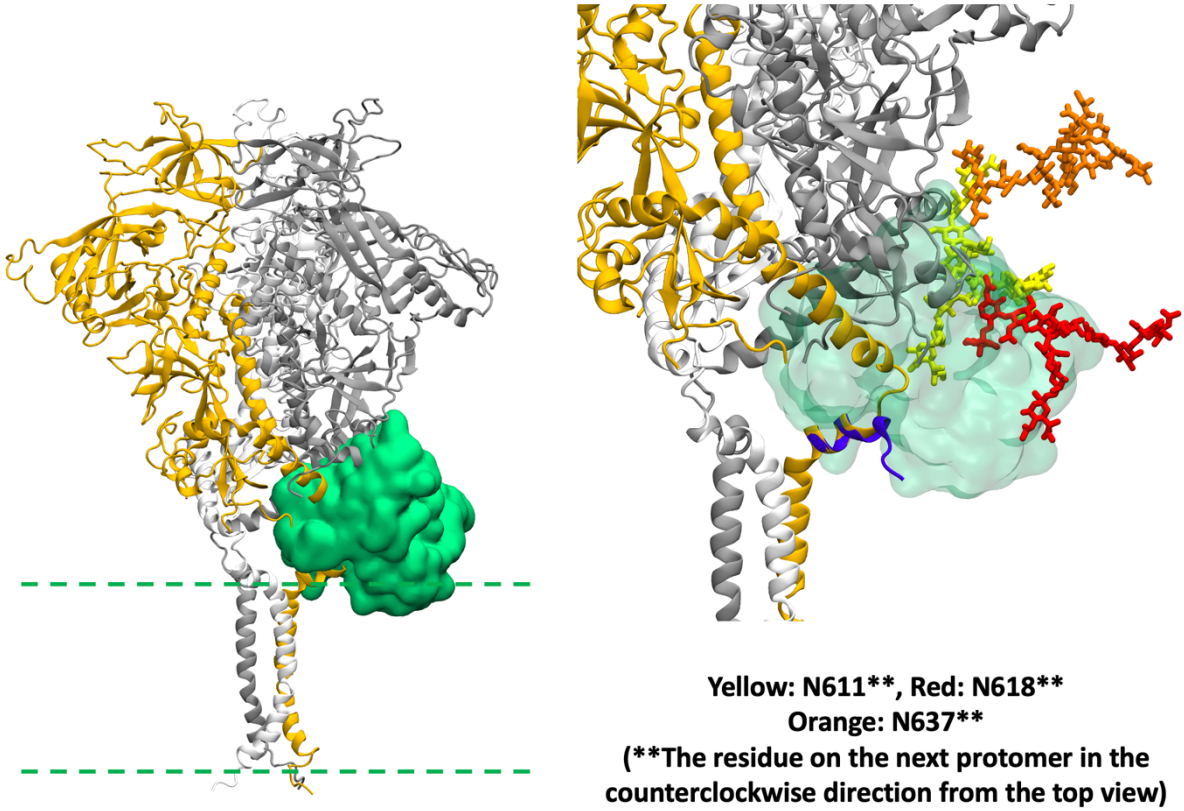

Figure 6—figure supplement 5. Shielding of antibody 4E10 epitope. Labeling and color coding are the same as in Figure 6—figure supplement 4.

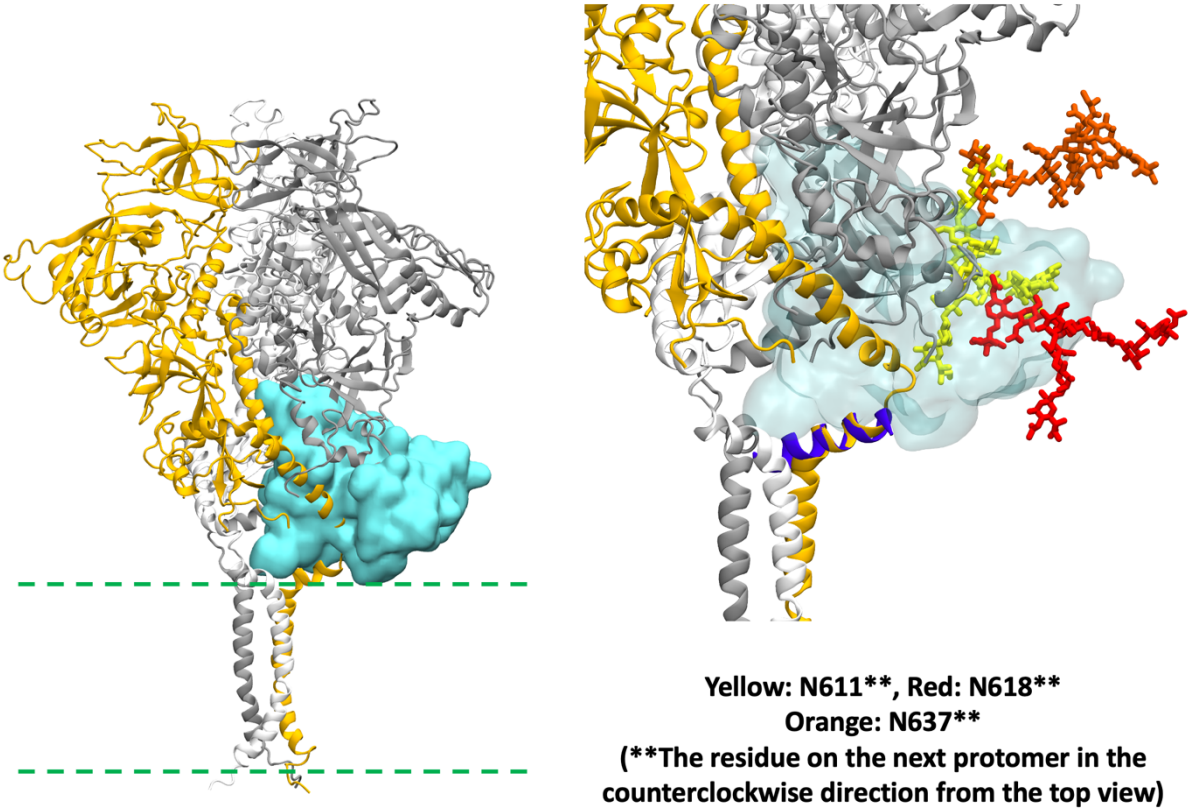

Figure 6—figure supplement 6. Shielding of antibody 10E8 epitope. Labeling and color coding are the same as in Figure 6—figure supplement 4.

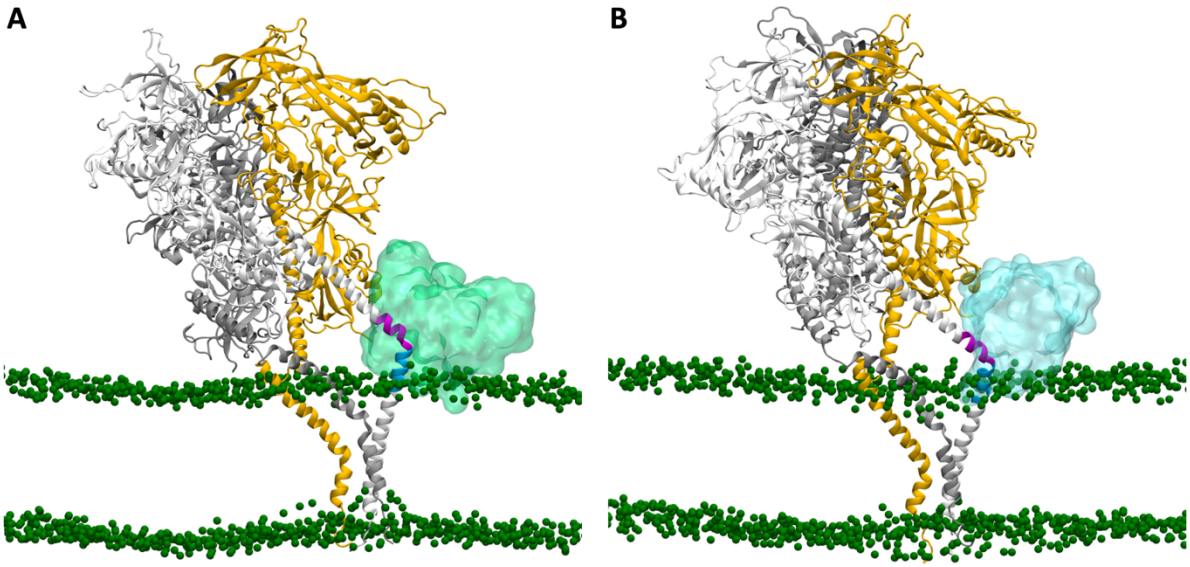

Figure 6—figure supplement 7. Snapshots showing the MPER of one protomer accessible to either 4E10 or 10E8. (**A**) Snapshot in which the MPER of the white protomer is accessible to 4E10 (cyan transparent surface) but not to 10E8. (**B**) Snapshot in which the MPER of the white protomer is accessible to 10E8 (green transparent surface) but not to 4E10.

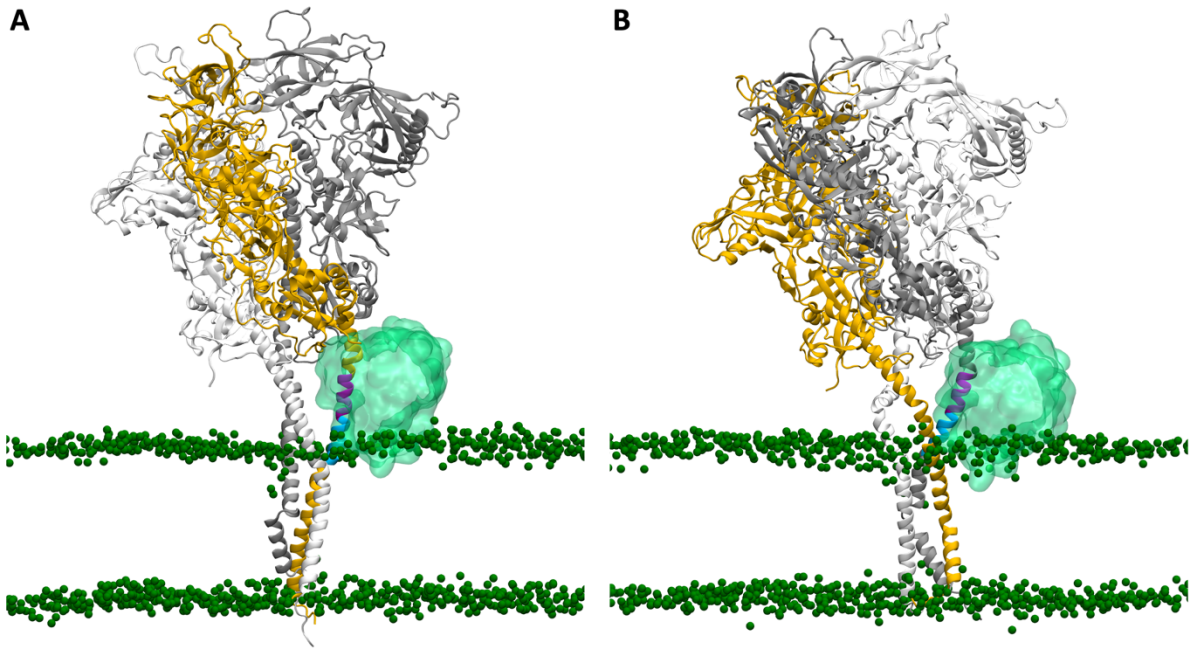

Figure 6—figure supplement 8. Snapshots showing the MPER of two protomers are accessible to 4E10. The MPER is accessible to 4E10, (**A**) exclusively in the yellow protomer, and (**B**) exclusively in the gray protomer.

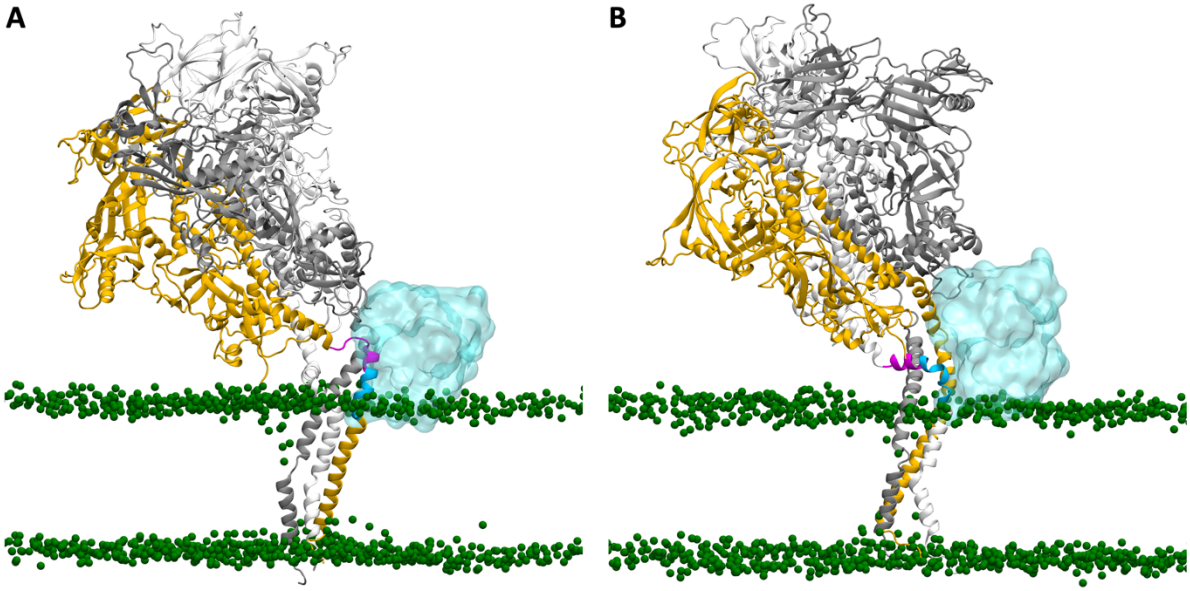

Figure 6—figure supplement 9. Snapshots showing the MPER of two protomers are accessible to 10E8. The MPER is accessible to 4E10, (**A**) exclusively in the yellow protomer, and (**B**) exclusively in the white protomer.

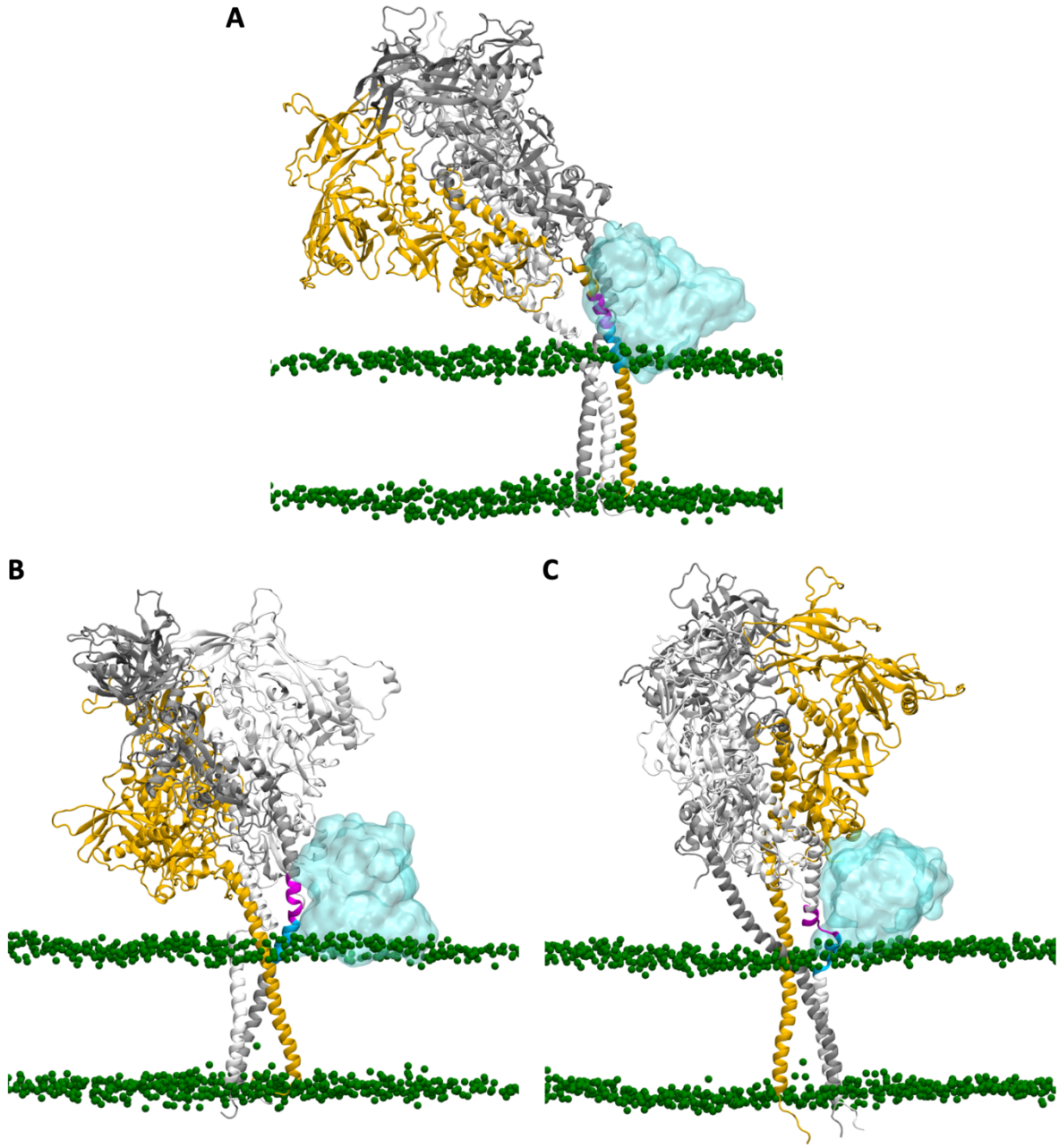

Figure 6—figure supplement 10. Snapshots showing the MPER of three protomers are accessible to 10E8. The MPER is accessible to 4E10, (**A**) exclusively in the yellow protomer, (**B**) exclusively in the gray protomer, and (**C**) exclusively in the white protomer.