Supporting Information

Cholesterol Interaction with the Trimeric HIV Fusion Protein gp41 in Lipid Bilayers Investigated by Solid-State NMR Spectroscopy and Molecular Dynamics Simulations

Byungsu Kwon [1], Taraknath Mandal [2], Matthew R. Elkins [1], Younghoon Oh [2], Qiang Cui [2,3,4] and Mei Hong [1]*

[1] Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139

[2] Department of Chemistry, [3] Department of Physics and [4] Department of Biomedical Engineering, Boston University, 590 Commonwealth Avenue, Boston, MA 02215

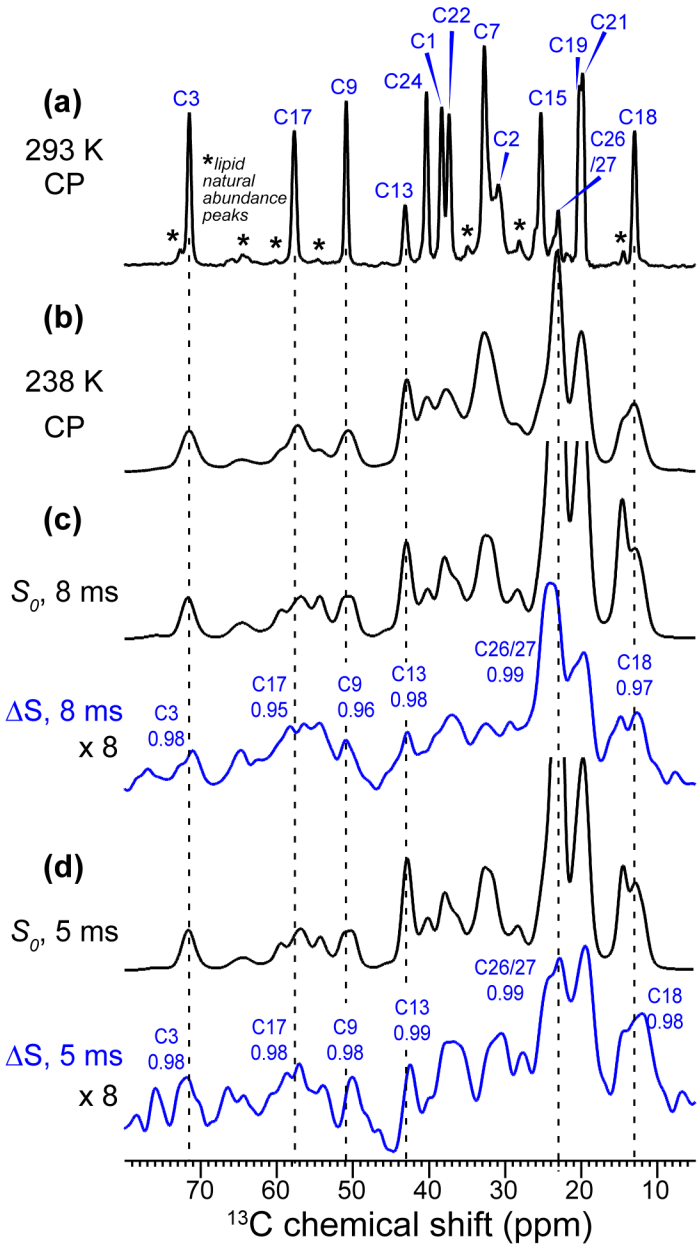

Figure S1. 1D ^{13}C spectra of 1-^{13}C CHOL bound to gp41-containing VM+ membranes. (**a**) ^{13}C CP spectrum measured at 293 K, showing well resolved ^{13}C signals for 1-^{13}C labeled sites at this temperature. (**b-d**) ^{13}C spectra measured at 238 K in the frozen membrane. (**b**) ^{13}C CP spectrum shows the expected line broadening at this low temperature. However, several carbons, including C17 and C9, remain resolved. (**c**) ^{13}C-^{19}F REDOR S_0 and ΔS spectra of 1-^{13}C CHOL at 8 ms mixing. ^{19}F dephasing was caused by 4-^{19}F-F673 and 5-^{19}F-W680 in gp41. The spectra were measured under 10 kHz MAS. The S/S_0 values for selected sites are indicated in the ΔS spectrum. At the P : L : C molar ratio of 1 : 16 : 7, because each gp41 trimer has 3 cholesterol molecules in close proximity, 18 out of 21 cholesterol are too far from the peptide to be dephased. Thus the lowest S/S_0 value is expected to be 0.86. (**d**) ^{13}C-^{19}F REDOR S_0 and ΔS spectra at 5 ms mixing.

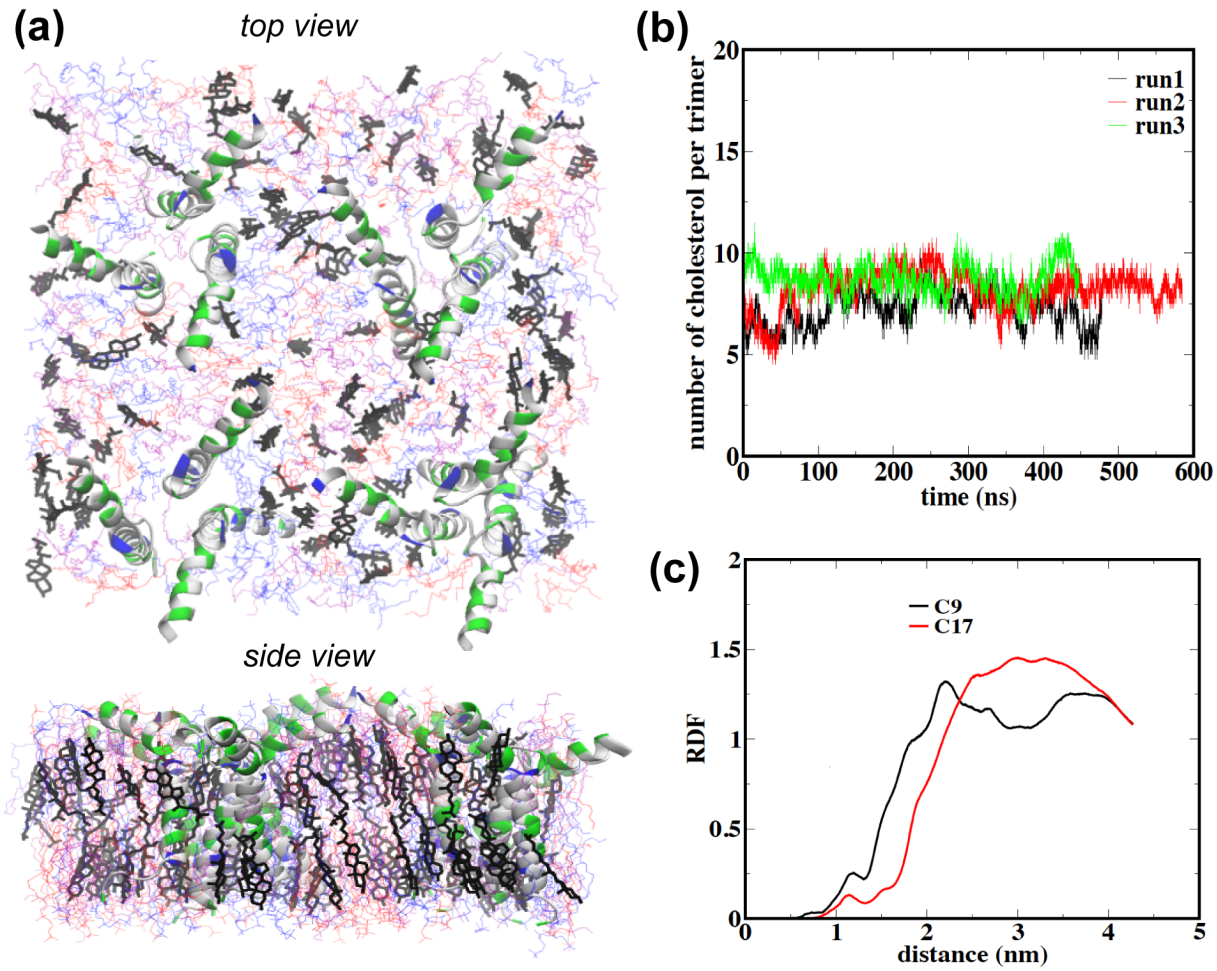

Figure S2. (**a**) Top and side view of an equilibrated protein-lipid-cholesterol system: Light blue, purple, and red color represent POPC, POPE and PSM lipids, respectively. Highlighted black color represents cholesterols. Protein is represented by ribbon structure. Water and ions are not shown for clarity. (**b**) Number of cholesterol molecules per gp41 trimer (defined as cholesterols within 3 Å of any protein atom) as a function of time from three independent simulations. (**c**) Radial distribution functions of the C9 and C17 cholesterol atoms from W680 Cζ3.

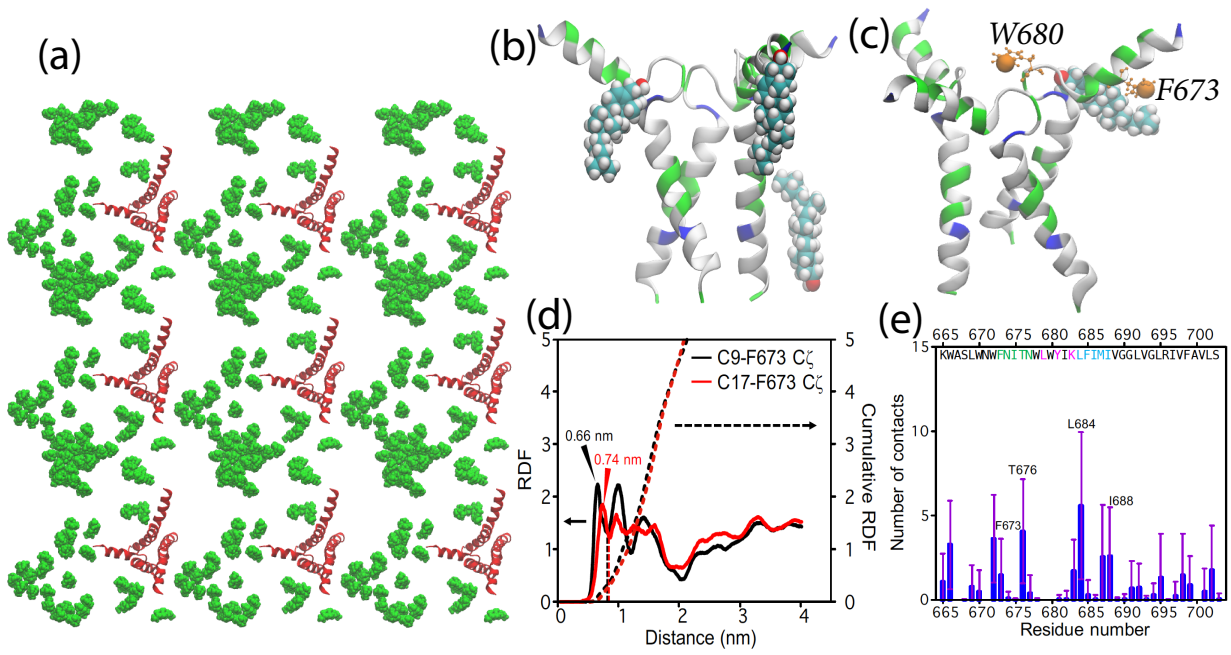

Figure S3. For the low peptide-concentration case, we conducted an independent set of simulations that prepared the system using a similar protocol as the high peptide-concentration simulations; i.e., 200 ns of HMMM simulation was used to equilibrate the lipids and cholesterols, after which the lipid tails were restored to full atomistic models, and then the system was simulated for 500 ns; three independent runs were conducted. The results are presented in the same format as **Fig. 6** in the main text. While HMMM simulations have been shown to be effective at sampling lipid distributions, the lack of full lipid tails might promote artificial cholesterol clustering when the protein concentration is low, and these clusters do not break up at the time scale of 500 ns in the subsequent full-atomistic-scale simulations (compare **Fig. S3a** with **Fig. 6a** in the main text). The clustering of cholesterols makes them less accessible to the gp41 protein; as a result, the computed radial distribution functions (**Fig. S3d**) give a slightly lower stoichiometry of 1.1 cholesterol near MPER for each gp41 trimer, as compared to the simulations discussed in the main text (**Fig. 6d**), which give a stoichiometry of 1.5 cholesterol near MPER for each gp41 trimer. Nevertheless, a similar set of protein residues are engaged in cholesterol interactions (compare **Fig. S3e** and **Fig. 6e** in the main text). The comparison of the two sets of low peptide concentration simulations highlights that care needs to be exercised to properly sample the distribution of cholesterols for systems that contain a low fraction of protein and a high concentration of cholesterol.

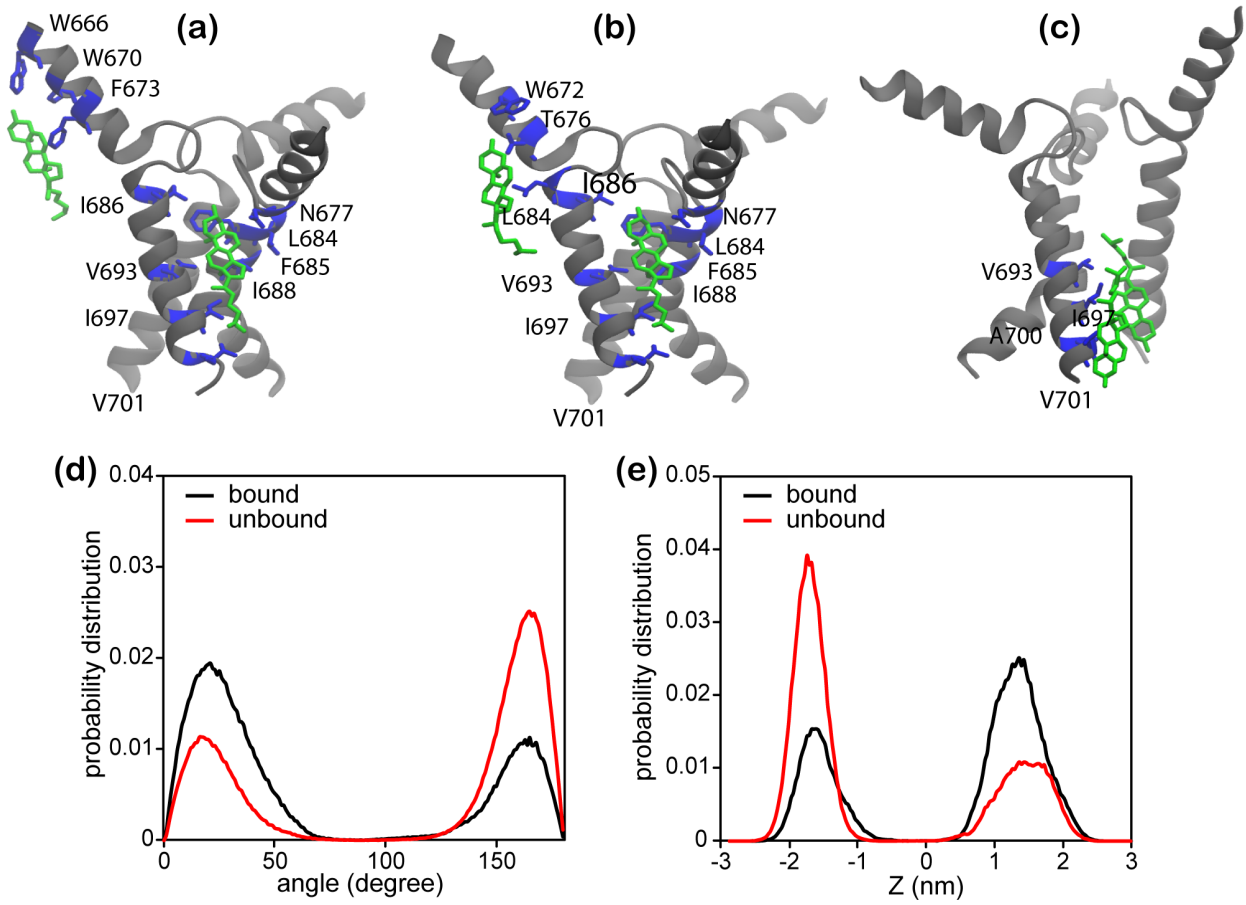

Figure S4. Representative snapshots of different modes of protein-cholesterol interactions. (**a**) A cholesterol bound to the MPER only and another cholesterol deeply inserted inside the membrane. (**b**) Two cholesterols bound to both the MPER and TMD regions. (**c**) Two cholesterols bound to the TMD only. Cholesterol-contacting protein residues are highlighted in blue. (**d**) Probability distribution of the orientations of bound and unbound cholesterols with respect to the membrane normal. (**e**) Positions of the O3 oxygen atoms of bound and unbound cholesterols along the direction of the membrane normal (z). Panels (**a-c**) suggest that the orientation of the bound cholesterols can vary significantly depending on their binding site in the gp41 trimer. For example, the MPER bound cholesterol shown in (a) is tilted away from the membrane normal while the cholesterols bound to TMD are almost parallel to the membrane normal. The latter are also inserted into the membrane more deeply than the former. Accordingly, we calculated the probability distribution of the orientations of bound and unbound cholesterols (**Fig. S2d**). The angle is defined between a vector joining C13 to C3 of the cholesterol and the membrane normal, so the peak values close to 0° and 180° correspond to the upper and lower leaflet cholesterols, respectively. **Fig. S2d** shows that the angle distribution of the bound cholesterols in the upper leaflet is broader than the unbound cholesterols, because many of the MPER-bound cholesterols are tilted. However, the angle distribution of the bound cholesterols in the lower leaflet is narrower than the unbound cholesterols, suggesting that these cholesterols pack tightly with the TMD. We also show the distribution of the cholesterol 'O3' atom position in **Fig. S2e**. The membrane center is set to $z = 0$, and a positive z corresponds to the upper leaflet. The positional distribution of the bound cholesterols in the upper leaflet is slightly broader than the unbound ones due to the varying insertion depth of the cholesterols, depending on the binding region. The bound cholesterols in the lower leaflet also insert slightly more deeply compared to the unbound ones, due to the long TMD region that binds the cholesterol tails.