Multiple modes of cholesterol translocation in the human Smoothened receptor

Prateek Bansal[1], **Maia Kinnebrew**[2], **Rajat Rohatgi**[2,3], **Diwakar Shukla**[1,4,5,6*]

***For correspondence:**
diwakar@illinois.edu

[1]Department of Chemical and Biomolecular Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States; [2]Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305, United States; [3]Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, United States; [4]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL, 61801; [5]Department of Bioengineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States; [6]Cancer Center at Illinois, University of Illinois at Urbana-Champaign, Urbana, IL, 61801

Abstract Smoothened (SMO), a member of the G Protein-Coupled Receptor superfamily, mediates Hedgehog signaling and is linked to cancer and birth defects. SMO responds to accessible cholesterol in the ciliary membrane, translocating it via a longitudinal tunnel to its extracellular domain. Reaching a complete mechanistic understanding of the cholesterol translocation process would help in the development of cancer therapies. Experimental data suggest two modes of translocation to support entry of cholesterol from outer and inner membrane leaflets, but the exact mechanism of translocation remains unclear. Using atomistic molecular dynamics simulations (~2 millisecond simulations) and biochemical assays of SMO mutants, we assess the energetic feasibilities of the two modes. We show that the highest energetic barrier for cholesterol translocation from the outer leaflet is lower than that from the inner leaflet. Mutagenesis experiments and complementary simulations of SMO mutants validate the role of critical amino acid residues along the translocation pathways. Our data suggests that cholesterol can take either pathway to enter SMO, thus explaining experimental observations in the literature. Thus, our results illuminate the energetics and provide a first molecular description of cholesterol translocation in SMO.

Introduction

Smoothened (SMO) is a member of the G Protein-Coupled Receptor (GPCR) superfamily with a typical heptahelical transmembrane fold (***Wang et al., 2013***, ***2014***). SMO has been identified as an oncoprotein, and mutations that cause overactivity of SMO drive tumorigenesis in basal cell carcinoma and medulloblastoma (***Raleigh and Reiter, 2019***). SMO antagonists are validated anti-cancer drugs (***Axelson et al., 2013***; ***Jain et al., 2017***), but are limited by drug resistance and side effects (***Meani et al., 2014***). SMO functions as a transmembrane signal transducer in the Hedgehog (HH) signaling pathway, which is known to mediate cell differentiation during embryonic development (***Nüsslein-Volhard and Wieschaus, 1980***; ***Ingham, 2022***). It transduces signals across the cell membrane, particularly in the primary cilia (***Rohatgi et al., 2007***; ***Corbit et al., 2005***). However, the mechanism of endogenous activation of SMO is still a debated question in the field. SMO activa-

tion has been linked to membrane sterols in numerous studies (***Cooper et al., 2003***; ***Stottmann et al., 2011***; ***Nachtergaele et al., 2012***; ***Myers et al., 2013***; ***Blassberg et al., 2015***; ***Luchetti et al., 2016***; ***Huang et al., 2016***; ***Byrne et al., 2016***; ***Myers et al., 2017***; ***Kinnebrew et al., 2019***). Cholesterol, a steroidal molecule found abundantly in the plasma membranes of vertebrates, has been identified as the endogenous agonist for SMO (***Luchetti et al., 2016***; ***Huang et al., 2016***; ***Byrne et al., 2016***). Membrane cholesterol has been shown to interact with and modulate GPCR activity (***Kumar and Chattopadhyay, 2019***; ***Kumar et al., 2021***; ***Serdiuk et al., 2022***), but in the case of SMO, cholesterol has been uniquely shown to be necessary and sufficient for SMO activation (***Luchetti et al., 2016***; ***Huang et al., 2016***). Recent work has shown that SMO is activated by a minor pool of membrane cholesterol, termed accessible cholesterol, at the primary cilium (***Kinnebrew et al., 2019***; ***Radhakrishnan et al., 2020***; ***Kinnebrew et al., 2021***; ***Siebold and Rohatgi, 2023***). Patched-1 (PTCH1), a twelve-pass transmembrane protein, sits directly upstream of SMO in the HH pathway, and inhibits SMO (***Ingham et al., 1991***; ***Denef et al., 2000***; ***Ingham et al., 2000***; ***Taipale et al., 2002***). PTCH1 modulates the availability of accessible cholesterol at the primary cilium, thereby regulating SMO, with models suggesting effects on both the CRD and 7TM pockets (***Kinnebrew et al., 2019***; ***Radhakrishnan et al., 2020***; ***Siebold and Rohatgi, 2023***).

The mechanism by which SMO is activated has been the subject of much discussion over recent years, with multiple studies theorizing the mechanism of PTCH1's inhibition on SMO. Previous studies have suggested that PTCH1 could function as a sterol transporter (***Davies et al., 2000***; ***Taipale et al., 2002***; ***Bidet et al., 2011***; ***Zhang et al., 2018***; ***Kinnebrew et al., 2019***, ***2021***). A recent study using coarse-grained Molecular Dynamics simulations investigated the possibility of this process, concluding that the overall process might occur at an energetic cost of $\sim$ 3-5 kcal/mol (***Ansell et al., 2023***). Recently resolved structures of PTCH1 (***Gong et al., 2018***; ***Qi et al., 2018***; ***Rudolf et al., 2019***; ***Qi et al., 2019a***; ***Zhang et al., 2020***) have reported the presence of a Sterol Binding Domain (SBD) and a hydrophobic conduit that extends from the outer leaflet to the extracellular space. Heterologous expression of PTCH1 in membranes reduces outer leaflet cholesterol accessibility, suggesting that PTCH1 reduces the access of SMO to membrane cholesterol (***Kinnebrew et al., 2021***).

The CRD of SMO contains a binding site for steroidal molecules (***Nachtergaele et al., 2012***, ***2013***; ***Nedelcu et al., 2013***; ***Myers et al., 2013***). The endogenous agonist cholesterol, (***Byrne et al., 2016***) as well as the naturally occurring alkaloidal agonist cyclopamine, (***Nachtergaele et al., 2013***; ***Huang et al., 2016***) bind in the CRD. In addition, SMO has a pocket in the Transmembrane Domain (TMD), which is known to bind to multiple antagonists such as LY2940680 (***Wang et al., 2013***), SANT1 and AntaXV (***Wang et al., 2014***), cyclopamine (***Weierstall et al., 2014***), TC114 (***Zhang et al., 2017***), Vismodegib (***Byrne et al., 2016***), the synthetic agonist SAG (and variants SAG1.3, SAG1.5, SAG21k) (***Wang et al., 2014***; ***Qi et al., 2020***; ***Deshpande et al., 2019***; ***Kinnebrew et al., 2022***), the steroidal agonists 24S,25-epoxy cholesterol (***Qi et al., 2019b***), and cholesterol (***Deshpande et al., 2019***; ***Qi et al., 2020***).

While cholesterol binding to both the TMD and CRD sites is required for full SMO activation, our work focuses on how cholesterol gains access to the CRD site, perched above the outer leaflet of the membrane (***Luchetti et al., 2016***; ***Kinnebrew et al., 2022***). Multiple lines of evidence suggest that PTCH1-regulated cholesterol binding to the CRD plays an instructive role in SMO regulation both in cells and animals. Mutations in residues predicted to make hydrogen bonds with the hydroxyl group of cholesterol bound to the CRD reduced both the potency and efficacy of SHH in cellular signaling assays (***Kinnebrew et al., 2022***; ***Byrne et al., 2016***) and, more importantly, eliminated HH signaling in mouse embryos (***Xiao et al., 2017***). Experiments using both covalent and photocrosslinkable sterol probes in live cells directly show that PTCH1 activity reduces sterol access to the CRD (***Kinnebrew et al., 2022***; ***Xiao et al., 2017***). Notably, our simulations evaluate a path of cholesterol translocation that includes both the TMD and CRD sites: cholesterol first enters the 7-transmembrane domain bundle from the membrane; it then engages the TMD site before continuing along a conduit to the CRD site. Thus, we analyze translocation energetics and residue-level contacts along a path that includes both the TMD and the CRD.

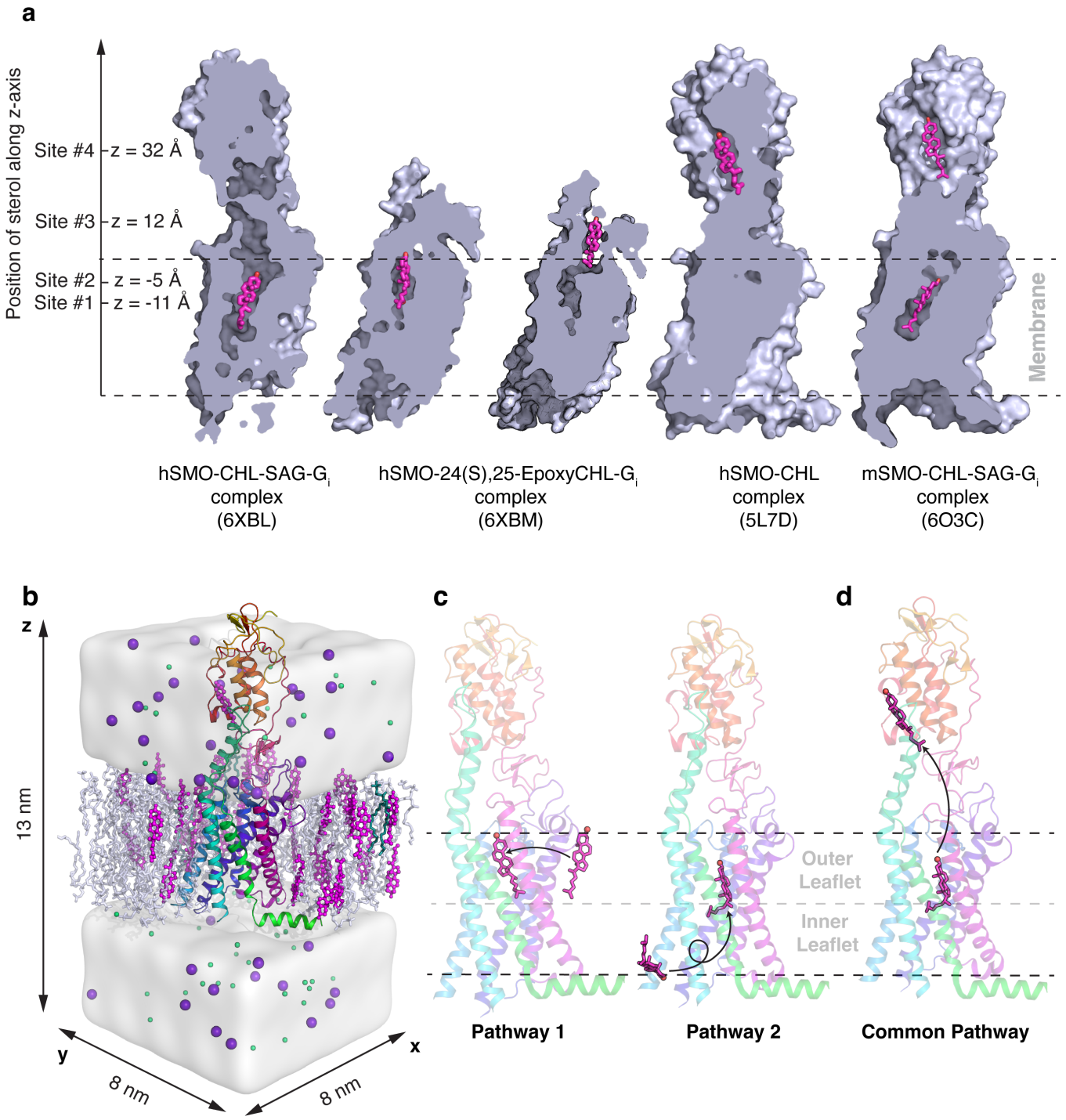


Figure 1. (a) The binding sites of the sterols along the hypothesized tunnel in SMO. Sterol binding sites have been identified deep in the TMD (6XBL) (***Qi et al., 2020***), at the interface of CRD and TMD (6XBM) (***Qi et al., 2020***), at the CRD sterol-binding site (5L7D) (***Byrne et al., 2016***), and in a dual-bound mode where cholesterol is bound to both the TMD and CRD (6O3C) (***Deshpande et al., 2019***). (b) Example simulation system showing SMO (5L7D, cyan) embedded in a membrane (white/magenta). Water is shown as a white surface, while sodium (purple) and chloride (green) ions are shown as spheres. (c) Pathway 1 and Pathway 2 investigate the translocation of cholesterol from the membrane to SMO's TMD. (d) The Common Pathway follows the translocation of cholesterol from the TMD to the CRD. Snapshots in (a) are made from structures in the PDB, while (b-d) are frames taken from MD simulations.

Figure 1—figure supplement 1. Comparison of TM6 in Class F versus Class A

Huang et al. (***2018***) theorized that SMO's activation involved the translocation of cholesterol from the membrane via a hydrophobic conduit to the binding site in the CRD. This study resolved the structure of active *Xenopus laevis* SMO, which showed outward movements of transmembrane helices 5 and 6 (TM5-TM6) on the intracellular end, which opened a cavity in SMO extending to the inner leaflet laterally, between TM5 and TM6. This led to the hypothesis that the entry of cholesterol into SMO could happen from the inner leaflet of the membrane, between TM5 and TM6. There is also additional evidence showing that the activity of SMO is regulated by cholesterol in the outer leaflet, which enters SMO between TM2 and TM3. ***Hedger et al.*** (***2019***) reported a cholesterol binding site present at the outer leaflet, between the TM2 and TM3 helices. Using coarse-grained simulations, the authors tested a variety of membrane compositions around SMO and found that a cholesterol binding site existed in human SMO (hSMO). ***Kinnebrew et al.*** (***2021***) used total internal reflection fluorescence microscopy (TIRFM) to assess the effect of PTCH1's activity on membrane cholesterol. They concluded that PTCH1 activity caused a decrease in the accessibility of cholesterol in the outer leaflet, suggesting that outer leaflet cholesterol is sensed by SMO. Furthermore, in our previous work (***Bansal et al., 2023***), we observed that cholesterol accumulated outside TM2-TM3 in the outer leaflet of inactive SMO. This was supported by the observation that when SMO was bound to its agonist SAG, we observed a tunnel opening between TM2-TM3 in the outer leaflet, which may facilitate the translocation of cholesterol into the core of SMO's TMD. We therefore hypothesize that cholesterol shows two modes of translocation to enter the TMD from the membrane - (1) starting at the outer leaflet, between TM2-TM3, or (2) starting at the inner leaflet, between TM5-TM6. Alternatively, (3) cholesterol could use both pathways if they show similar energetic behaviors.

How cholesterol moves from the membrane into the core of the TMD of SMO is still poorly understood. Cholesterol traverses SMO to ultimately reach the TMD and CRD binding sites, but the mechanism of cholesterol perception has not been elucidated yet, which gives us an opportunity to explore the mechanistic aspects of this process from both computational and experimental viewpoints. In this study, we simulate SMO by embedding it in a membrane (***Figure 1***b). We report the entire translocation path of cholesterol from the membrane to SMO's CRD for both modes of translocations (***Figure 1***c) - between TM2-TM3 in the outer leaflet of the membrane (hereafter referred to as "Pathway 1"), and between TM5-TM6 in the inner leaflet of the membrane (hereafter referred to as "Pathway 2"). We observe that cholesterol can translocate via both pathways, and the free energy barriers associated with Pathway 1 are lower than those of Pathway 2. We test mutations in SMO that can disrupt the movement along either pathway, and show that the experimental results are further supported by simulations of cholesterol translocation in SMO mutants.

After cholesterol has entered the TMD of SMO in our simulations, we observe cholesterol moving along TM6 to the TMD-CRD interface (Common Pathway, ***Figure 1***d) to access the binding site in the CRD (***Huang et al., 2018***; ***Kinnebrew et al., 2022***). One of the unique features of SMO is the presence of a long helix 6 (TM6) (***Wang et al., 2013***), (***Figure 1—figure Supplement 1***) which acts as a connector between the CRD and the TMD. We test mutations in SMO that can disrupt cholesterol movement along this Common Pathway, and show that these mutants can halt the translocation process by loss of hydrophobic contacts. These results for SMO mutants are further validated by additional MD simulations. Therefore, in this study, the entire process of cholesterol translocation was observed using aggregate 2 milliseconds of unbiased all-atom MD simulations. Exploring the mechanism by which SMO translocates cholesterol would provide insights into the endogenous regulation of HH activity and suggest strategies for the next generation of drugs targeting SMO.

Results and Discussion

The entry of cholesterol from the outer leaflet into the TMD exhibits the highest energetic barrier along Pathway 1

We first simulated the translocation mode in which cholesterol enters the TMD of SMO from the outer leaflet (Pathway 1). To model this process, cholesterol was placed outside of the TMD in the membrane outer leaflet, between the TM2-TM3 interface. The entry of cholesterol from the membrane into the protein was steered towards the TMD to sample the entry pathway. The frames generated were then used to seed unbiased simulations for adaptive sampling (see Methods section for details). This was done to estimate the energetic barriers involved in the cholesterol entry from the outer leaflet.

To enter the TMD from the membrane, cholesterol must overcome the entropic barrier of being restrained inside the protein. This can make the entry process energetically expensive. To identify the barrier associated with entry, we projected the entire dataset on different translocation-related metrics and identified the mechanism of cholesterol translocation. The z-coordinate of cholesterol's center of mass (z-axis is perpendicular to the plane of the membrane) was used as a proxy for the progress of the overall translocation from the membrane to the CRD. The angle of cholesterol with the x-y plane (the plane of the membrane) was calculated to identify cholesterol pose. We observe that this angle (***Figure 2***a and ***Figure 2—figure Supplement 1***a) shows multiple minima ($\boldsymbol{\alpha}/\boldsymbol{\beta}$ along the Pathway 1). This provides evidence for multiple stable poses along the pathway as observed in the multiple stable poses of cholesterol in Cryo-EM structures of SMO bound to sterols (***Deshpande et al., 2019***; ***Qi et al., 2019b***, ***2020***). A reliable estimate of the barriers comes from using the time-lagged Independent Components (tICs), which project the entire dataset along the slowest kinetic degrees of freedom. Overall, the highest barrier along Pathway 1 is $\sim 5.8 \pm 0.7$ kcal/mol, and it is associated with the entry of cholesterol into the TMD (***Figure 2—figure Supplement 2***). Several factors contribute to this high barrier; first, cholesterol needs to be at the correct position - just outside the protein. In addition, cholesterol needs to be in the correct orientation - angled, with the isooctyl tail pointing towards the protein core. Additionally, the steric hindrances from the hydrophobic residues at the interface provide further obstacles for entry. Therefore, the TM2/3 residues at the entry point need to undergo conformational change to facilitate cholesterol entry – making it a rare event.

In the first stage of translocation, cholesterol is in the membrane ($\boldsymbol{\alpha}$, ***Figure 2***), just outside the TM2/3 helices of SMO as identified in literature (***Hedger et al., 2019***). At this cholesterol recognition site, SMO primarily contains hydrophobic residues (***Figure 2***c-f), which preferentially interact with the hydrophobic isooctyl tail of cholesterol. Therefore, the cholesterol entry involves the insertion of the tail into the TMD first, which is then followed by the hydrophilic androsterolic moiety. For the sake of clarity, we divide this entry pocket at the TM2-TM3-membrane interface into two parts - 'lower' and 'upper' pocket. The 'upper' pocket corresponds to the residues that coordinate with the androsterolic moiety of cholesterol (***Figure 2***c, d), and the 'lower' pocket corresponds to residues that lie closer to the isooctyl tail of cholesterol (***Figure 2***e, f). The upper (M286^{ECL1}, A289^{ECL1}, I^{ECL2} and A^{ECL2}) and the lower (W281$^{2.58f}$, F391^{ECL2} and M525$^{7.45f}$) pocket residues undergo conformational changes to open the space for the entry of cholesterol (***Figure 2***c-f). Here, the superscript refers to the Wang numbering scheme (***Wang et al., 2014***). This flexible motion is facilitated by G280$^{2.57f}$ in the lower pocket and G288^{ECL1} in the upper pocket. Upon the entry of the cholesterol tail, an intermediate state, $\boldsymbol{\beta}$ (***Figure 2***d, f) is observed where cholesterol lies flat with respect to the membrane plane and forms extensive hydrophobic contacts with A283$^{2.60f}$, I317$^{3.28f}$, F318$^{3.29f}$ and the disulfide bridge forming residues C314$^{3.25f}$ - C390^{ECL2}.

To reach the TMD binding site, cholesterol must first "rock" back to its upright pose ($\boldsymbol{\alpha}^*$) from its flat conformation in the state $\boldsymbol{\beta}$. This rocking motion is facilitated by the polar interactions between S313$^{3.24f}$, Q284$^{2.63f}$, and the alcoholic oxygen in cholesterol. The entry of cholesterol is thus captured by a $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta} \rightarrow \boldsymbol{\alpha}^*$ transition. For further clarity, we have plotted the minimum energy

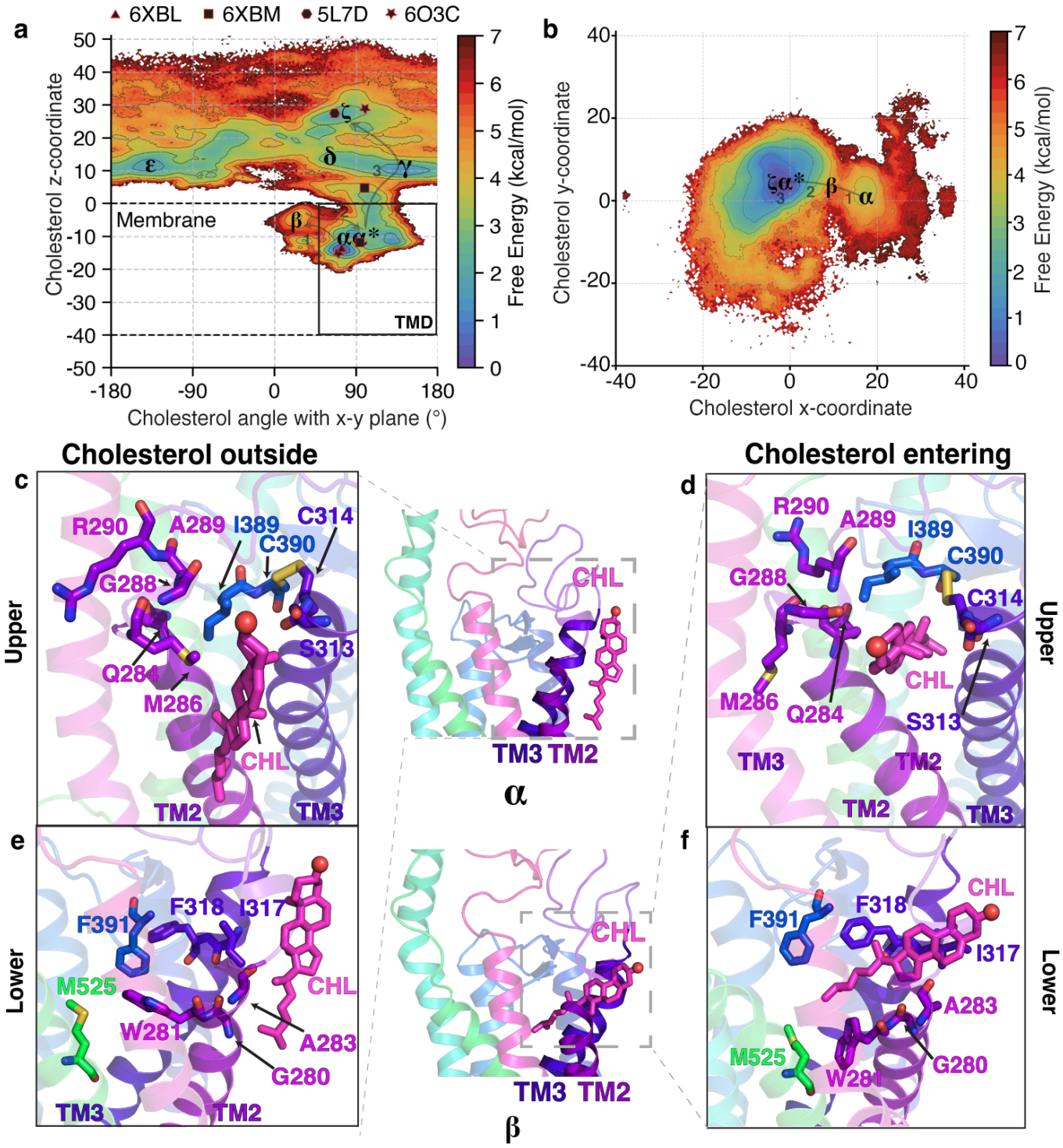


Figure 2. The molecular events as cholesterol enters the core of SMO's TMD from the outer leaflet of the membrane. (a) Free energy plot showing the angle of cholesterol with the x-y plane, the plane of the membrane, versus the z-coordinate of cholesterol for Pathway 1. The pathway followed by cholesterol is $\alpha \rightarrow \beta \rightarrow \alpha^*$. The experimental structures of SMO are shown as black polygons. (b) Free energy landscape of cholesterol's y-coordinate plotted versus cholesterol's x-coordinate. Cholesterol interacts with residues in TM2-TM3 while entering the core TMD of SMO. (c-f) Insets show cholesterol's interactions with residues at the membrane-protein interface for Pathway 1. (c,e) show cholesterol outside the protein (α), while (d,f) show cholesterol entering the protein (β). All snapshots presented are frames taken from MD simulations.

Figure 2—figure supplement 1. Error in Free Energies Fig. 2 and Fig. 4

Figure 2—figure supplement 2. TICAplot for SMO Cholesterol transport

Figure 2—figure supplement 3. Minimum energy path taken by cholesterol for both pathways

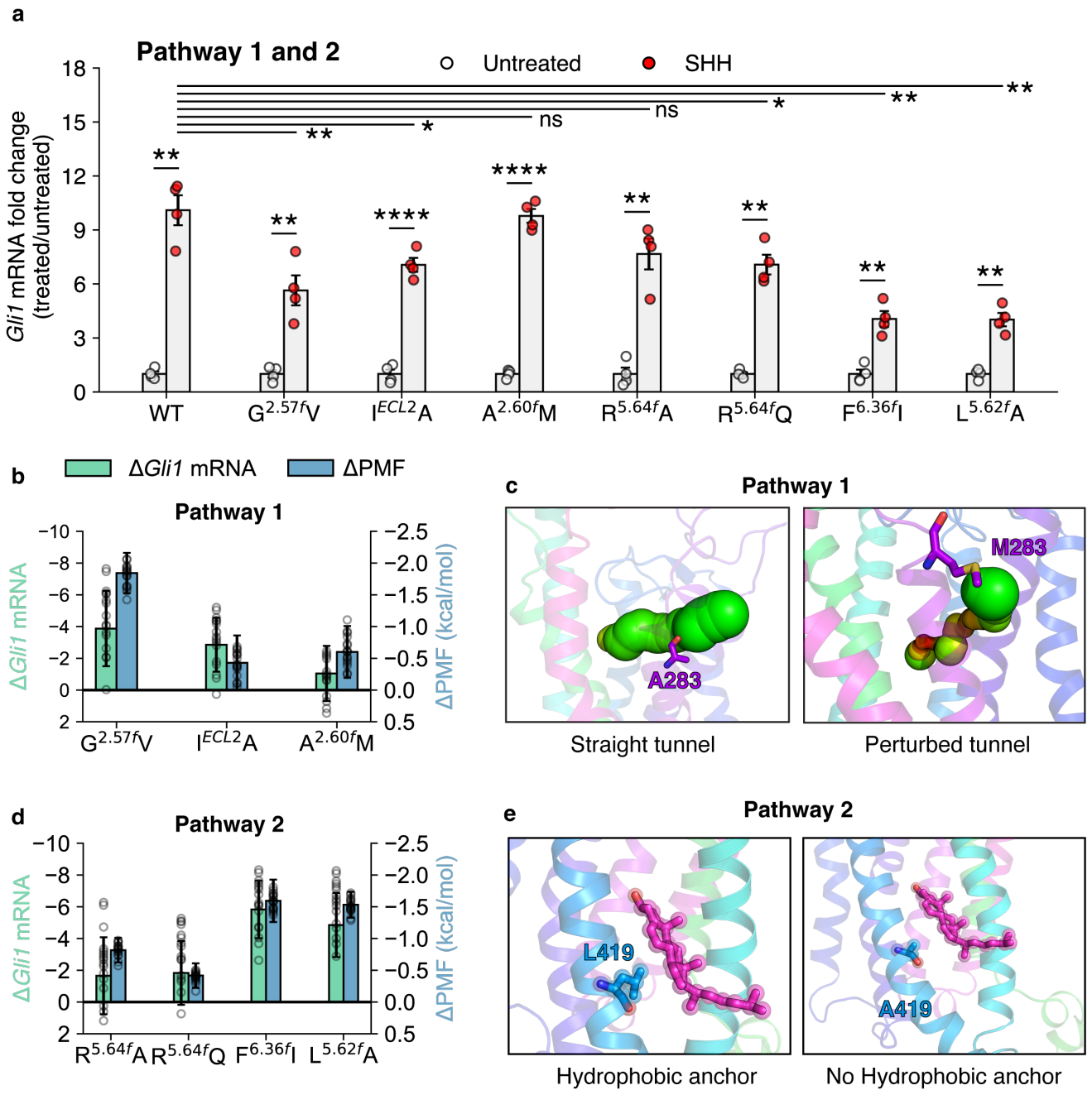


Figure 3. Effects of mutations along Pathways 1 and 2 on the activation of SMO. (a) *Gli1* mRNA fold changes show the responsiveness of SMO mutants to SHH. Untreated Gli1 levels indicate low SMO activity, while SHH-treated values correspond to the level of SMO activation induced by SHH ligand. A t-test with Welch's correction was used to compute statistical significance. (P values: untreated vs treated: WT: 1.327×10^{-3}, $G^{2.57f}V$: 9.212×10^{-3}, $I^{ECL2}A$: 4.2×10^{-5}, $A^{2.60f}M$: 7.1×10^{-5}, $R^{5.64f}A$: 2.062×10^{-3}, $R^{5.64f}Q$: 1.192×10^{-3}, $F^{6.36f}I$: 2.163×10^{-3}, $L^{5.62f}A$: 1.948×10^{-3}, treated WT vs treated mutant: $G^{2.57f}V$: 9.1×10^{-3}, $I^{ECL2}A$: 0.02734, $A^{2.60f}M$: 0.7477, $R^{5.64f}A$: 0.08858, $R^{5.64f}Q$: 0.02766, $F^{6.36f}I$: 1.923×10^{-3}, $L^{5.62f}A$: 2.306×10^{-3} key: Not significant (ns) $P > 0.05$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$, N=4 for all experiments.) (b) Δ*Gli1* mRNA fold change (SHH vs untreated) and Δ PMF (difference of peak PMF, calculated as PMF_{WT} - PMF_{mutant}) plotted for the mutants in Pathway 1. (c) Example mutant $A^{2.60f}M$ shows that cholesterol is able to enter SMO through Pathway 1 even on a bulky mutation. (d) Same as (b) but for Pathway 2 (e) Example mutant $L^{5.62f}A$ shows that cholesterol can enter SMO through Pathway 2 due to lesser steric hindrance. All snapshots presented are frames taken from MD simulations.

Figure 3—figure supplement 1. Mutation data for Pathway 1

Figure 3—figure supplement 2. PMF data for Pathway 1

Figure 3—figure supplement 3. Tunnel diameter profile for Mutant A283M

Figure 3—figure supplement 4. Mutation data for Pathway 2

Figure 3—figure supplement 5. PMF data for Pathway 2

path taken by cholesterol as it translocates along this pathway (***Figure 2—figure Supplement 3***a,b). To further elucidate the position of cholesterol as it enters the protein, we projected the x and y coordinates of cholesterol on a free energy landscape (***Figure 2***b, ***Figure 4***b and ***Figure 2—figure Supplement 1***a,b). In this figure, the state β clearly marks the transition state between cholesterol outside and inside the TMD. Overall, the angle between cholesterol and the x-y plane transitions from 90° → 0° → 90° for cholesterol to enter the protein.

To further dive into the details of cholesterol translocation, we designed mutations along Pathway 1, and measured the change in activity with respect to Wild-Type (WT) mouse Smoothened (mSMO). mSMO was chosen since there are no human SHH-responsive cell lines that can be passaged, edited, or transduced with genes. In addition, mSMO shows 92.8% sequence identity and 94.6% sequence similarity with hSMO for full-length sequences. Previous studies that have resolved hSMO structures have used mSMO for their structure-guided mutagenesis studies to comment on SMO activity (***Byrne et al., 2016***). Activity was measured using *Gli1* mRNA fold change in the presence of SHH (***Figure 3***a). To further validate the mutations, simulations were performed on hSMO to compute the Potential of Mean Force (PMF) of cholesterol entry in the presence of mutations (***Darve et al., 2008***; ***Hénin et al., 2009***) (additional details presented in Methods). Here, the PMF can characterize the barriers associated with the translocation of cholesterol, and a difference in the peak value of PMF is presented for each mutant (***Figure 3***b). We mutated G280$^{2.57f}$ to valine - G$^{2.57f}$V to test whether reducing the flexibility of TM2 prevents cholesterol entry into the TMD. Consequently, the activity of mSMO showed a decrease (***Figure 3—figure Supplement 1***). However, this decrease could also be attributed to the steric hindrance added by the presence of a bulky propyl group in valine. We designed I^{ECL2}A to check the importance of hydrophobic contacts during translocation along Pathway 1. However, this mutant did not affect the activity nor the barrier for translocation significantly (***Figure 3***b, ***Figure 3—figure Supplement 2***). Finally, we mutated A283$^{2.60f}$ to methionine - A$^{2.60f}$M to test whether the presence of a bulkier residue would block translocation. Surprisingly, the effect on activity was not significant. When we calculated the PMF for cholesterol entry, A$^{2.60f}$M mutant showed a restricted tunnel, but it did not fully block the tunnel (***Figure 3—figure Supplement 3***). Therefore, the change in the PMF and experimentally measured activity was not significant (***Figure 3***b, c, ***Figure 3—figure Supplement 2***).

Cholesterol 'flipping' corresponds to the highest barrier in its translocation from the inner leaflet along Pathway 2

To quantitatively assess the validity of cholesterol translocation from the inner leaflet, we performed adaptive sampling simulations to obtain the associated free energy barriers. For Pathway 2, cholesterol first binds at the interface between TM5 and TM6 in the inner leaflet (***Deshpande et al., 2019***; ***Huang et al., 2018***). In a structure resolved in 2022, cholesterol was observed at the interface between the protein and the membrane, in the inner leaflet, between TM5 and TM6 (***Zhang et al., 2022***). A striking observation is that this cholesterol binding site pose was never used as a starting point for simulations and was discovered independently from the pose described in ***Zhang et al.*** (***2022***) (***Figure 4—figure Supplement 1***). The cholesterol in the inner leaflet has a downward orientation, with the polar hydroxyl group pointing intracellularly (η, ***Figure 4***, ***Figure 4—figure Supplement 1***). Thus, if cholesterol has to translocate from the inner leaflet, it has to undergo "flipping" motion, as the resolved structure with the cholesterol completely inside the TMD (***Qi et al., 2020***) shows the alcoholic moiety pointing towards the CRD (***Deshpande et al., 2019***) (***Figure 1***a). The energetic barrier associated with flipping the motion of cholesterol can be observed by estimating the angle cholesterol makes with the x-y plane, and the barrier associated with translocation towards the TMD binding site can be estimated by projecting the data on the z-coordinate of cholesterol, similar to ***Figure 2***a for Pathway 1.

In ***Figure 4***a, multiple free energy minima are observed. The state η corresponds to cholesterol outside of the TMD and pointing downwards, forming a -90° angle with the x-y plane. This observed pose is similar to the binding pose for cholesterol outlined in a previously resolved structure, fur-

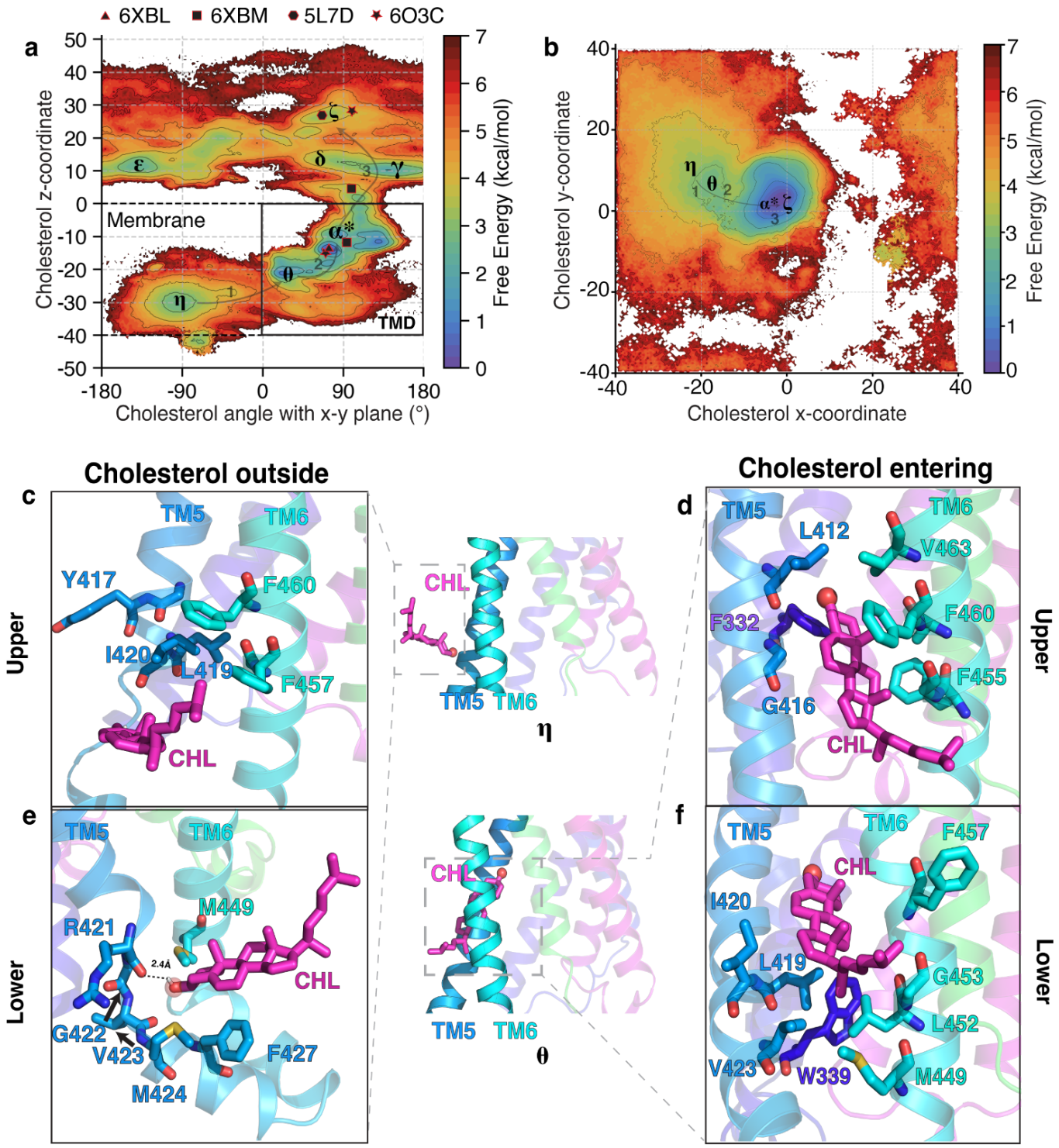


Figure 4. The molecular events as cholesterol enters the TMD from the inner leaflet in Pathway 2. (a) Free energy plot showing the angle of cholesterol with the x-y plane, the plane of the membrane, versus the z-coordinate of cholesterol for Pathway 2. The pathway followed by cholesterol is $\eta \rightarrow \theta \rightarrow \alpha^*$. The experimental structures of SMO are shown as black polygons. (b) Free energy landscape of cholesterol's y-coordinate plotted versus cholesterol's x-coordinate for Pathway 2. Cholesterol interacts with residues in TM5-TM6 for Pathway 2 while entering the SMO core TMD. (e-h) Insets show cholesterol's interactions with residues at the membrane-protein interface for Pathway 2. (c,e) show cholesterol outside the protein (η), while (d,f) show cholesterol entering the protein (θ) for Pathway 2. All snapshots presented are frames taken from MD simulations.

Figure 4—figure supplement 1. Comparison of cholesterol entry with existing resolved structure (***Zhang et al., 2022***)

ther corroborating our observations (***Zhang et al., 2022***). The state ϵ corresponds to cholesterol at the membrane-TMD interface and forms a 45° angle with the x-y plane. Finally, the state α^* corresponds to cholesterol inside the TMD and forming a +90° angle with the x-y plane. The state α^*, which represents the TMD binding site, is the most stable pose of cholesterol inside SMO. According to the free-energy landscapes (***Figure 2***a, ***Figure 4***a and ***Figure 2***c,d and ***Figure 4***c,d), the entry of cholesterol from the inner leaflet is associated with the highest barrier for the entire translocation process, ~6.5 ± 0.8 kcal/mol and it corresponds to the transition between the states η and θ. These values are comparable to ATP-Binding Cassette (ABC) transporters of membrane lipids, which use ATP hydrolysis (-7.54 ± 0.3 kcal/mol) (***Meurer et al., 2017***) to drive lipid transport from the membrane to an extracellular acceptor. Some of these transporters share the same mechanism as SMO, where the lipid from the inner leaflet is flipped and transported to the extracellular acceptor protein (***Tarling et al., 2013***). Additionally, for secondary active transporters that do not use ATP for the transport of substrates, a thermodynamic barrier of 5-6 kcal/mol has been reported in literature (***Chan et al., 2022***; ***Selvam et al., 2019***; ***McComas et al., 2023***; ***Thangapandian et al., 2025***).

The computed values of free energy barriers are dependent on the projections of the data. Upon projecting the data using the z-coordinate of cholesterol versus the angle between the cholesterol and the x-y plane (***Figure 4***a) and the y-coordinate of cholesterol versus the x-coordinate (***Figure 4***b), the barrier between η and θ is different. On plotting the first two components of tICs, (***Figure 2—figure Supplement 2***a,c), we observe that the energetic barrier between η and θ is ~6.5 ± 0.8 kcal/mol. We further validate that the slowest degrees of freedom in the model correspond to the entry of cholesterol from the inner leaflet of the membrane into SMO TMD (***Figure 2—figure Supplement 2***b,d).

Contrary to Pathway 1, the entry of cholesterol into the TMD happens via the alcoholic moiety first, which forms a hydrogen bond with the backbone oxygen of R421$^{5.64f}$ (***Figure 4***e). Furthermore, the TM5 loses some of its helicity, between residues R421$^{5.64f}$-M424$^{5.67f}$ due to the flexibility provided by G422$^{5.65f}$, which is a part of the conserved WGM motif implicated in SMO activation (***Bansal et al., 2023***). Here, we also designate the "upper" and "lower" pocket, which coordinate with the isooctyl and androsterolic moieties of cholesterol, respectively (***Figure 4***c-f). In the upper pocket, the isooctyl tail entry is blocked by the strong hydrophobic patch formed by Y417$^{5.60f}$, L419$^{5.62f}$, I420$^{5.63f}$, F457$^{6.38f}$, and F460$^{6.41f}$ (***Figure 4***c,e). Flexible residues such as glycine are also not present, which would allow for the fluctuations leading to the opening of the hydrophobic patch and enable entry of the cholesterol tail. Therefore, the androsterolic moiety of cholesterol enters first, followed by the isooctyl tail of cholesterol to reach the state ϵ. To stabilize and facilitate the androsterolic moiety's entry, F460$^{6.41f}$ and F455$^{6.36f}$ form $\pi-\pi$ contacts, and L419$^{5.62f}$, L452$^{6.33f}$ form hydrophobic contacts (***Figure 4***d,f) in state θ. Once cholesterol has flipped in θ, it allows for further translocation towards the TMD binding site (α^*) and finally arrival at the CRD site (ζ) with the hydroxyl group pointing extracellularly. The translocation of cholesterol from the membrane to the CRD binding site via Pathway 2 is captured by a $\eta \rightarrow \theta \rightarrow \alpha^* \rightarrow \zeta$ transition. For further clarity, we have plotted the minimum energy path taken by cholesterol as it translocates along this pathway (***Figure 2—figure Supplement 3***c,d). Overall, the angle between cholesterol and the x-y plane involves an entire cycle of -90°→ 0°→ 90° and cholesterol must flip, for cholesterol to enter the protein.

Interestingly, mutants along Pathway 2 showed a significant decrease in activity compared to Pathway 1 (***Figure 3***a, ***Figure 3***d, ***Figure 3—figure Supplement 4***), along with an increased thermodynamic barrier for translocation (***Figure 3***d, ***Figure 3—figure Supplement 5***). Mutating R421$^{5.64f}$ to alanine or glutamine did not decrease SMO activity significantly (***Figure 3***d), because the interaction with cholesterol is mediated by the protein backbone, and not the side-chain (***Figure 4***e). However, mutations like F$^{6.36f}$I and L$^{5.62f}$A reduce SMO activity. Their expression levels of these mutants are comparable to the wild-type mSMO (***Figure 3—figure Supplement 4***). The mutants compared to WT SMO showed a significant increase in PMF, due to the lack of the hydrophobic π-stacking pro-

vided by F455$^{6.36f}$ and hydrophobic contacts provided by L419$^{5.62f}$ during cholesterol translocation. Overall, we report that the mutants for Pathway 2 show a decrease in the activity of SMO and show a strong correlation between the reduction in activity and the barrier for cholesterol translocation (***Figure 3***b,d). These results validate the role of critical residues involved in cholesterol translocation from the inner leaflet as observed in the simulations.

Based on our experimental and computational data, we conclude that cholesterol translocation can happen via either pathway. This is supported on the basis of the following observations: mutations along Pathway 2 affect SMO activity more significantly, and the presence of a direct conduit that connects the inner leaflet to the TMD binding site. In addition, a resolved structure of SMO in the presence of cholesterol (***Zhang et al., 2022***) shows a cholesterol situated at the entry point from the membrane into the protein between TM5 and TM6, in the inner leaflet. However, we also observe that Pathway 1 shows a lower thermodynamic barrier (5.8 $\pm$ 0.7 kcal/mol vs. 6.5 $\pm$ 0.8 kcal/mol, $p = 0.0013$) . Additionally, PTCH1 controls cholesterol accessibility in the outer leaflet (***Kinnebrew et al., 2021***). This shows that there is a possibility for transport from both leaflets. One possibility that might alter the thermodynamic barriers is native membrane asymmetry, particularly the anionic lipid-rich inner leaflet. This presents as a limitation of our current model.

The pathway connecting the TMD to the CRD binding sites shows off-pathway intermediates

According to the dual-site model, to reach the binding site in the CRD (ζ), cholesterol translocate along the TMD-CRD interface from the TM binding site (α^*) is required. This Common Pathway is shared by cholesterol molecules translocating from the inner and outer leaflets. The translocation of cholesterol from the TMD to the CRD binding site involves a linear movement of the androsterolic moeity through the extracellular end of the TMD, where cholesterol maintains a primarily upright position, with the polar androsterolic moeity pointing towards the CRD site. The energetic barrier associated with the transition (α^*(TMD site) $\rightarrow$ ζ(CRD site)) can be visualized by plotting the z-coordinate of cholesterol versus the angle it forms with the x-y plane (***Figure 2***a, ***Figure 4***a). We observe that the highest barrier along the Common Pathway is ~5.1 $\pm$ 0.3 kcal/mol, which is lower than the highest energetic barrier for cholesterol entry from the inner and outer leaflet. Another interesting observation is that the stability of the cholesterol in the TMD binding site (α^*) is higher than the CRD binding site (ζ). This observation can be explained on the basis of the thermodynamic driving forces in the two pockets. The TMD binding site is composed mainly of hydrophobic residues, which form strong interactions with cholesterol (***Figure 5—figure Supplement 1***) in contrast to the CRD binding site, where the cholesterol is exposed to the aqueous environment. The CRD binding site being exposed to the solvent increases the conformational entropy associated with the cholesterol compared to the restrained TMD binding site. In addition, the CRD binding site is more flexible than the TMD binding site in the absence of cholesterol, as reported previously (***Kinnebrew et al., 2022***). This increased flexibility of the CRD binding site further leads to the formation of multiple conformational states between the CRD and cholesterol.

Once cholesterol reaches the TMD-CRD interface, it can adopt multiple poses before reaching the CRD binding site (***Figure 2***a, ***Figure 4***a). Cholesterol at this position can be upright (δ, ***Figure 2***a, ***Figure 4***a), where it interacts with F484$^{6.65f}$, W480$^{6.61f}$, V488$^{6.69f}$ and L221^{LD}, forming hydrophobic contacts (***Figure 5***a). However, there exists a thermodynamic barrier to take a completely upright path to the CRD binding site (angle ~+90°). This is due to the presence of the long beta sheet of the linker domain (residues L197^{LD}-I215^{LD}) of SMO at the TMD-CRD interface, blocking the direct upright translocation of cholesterol (***Figure 5—figure Supplement 2***). Hence, the major conformation of cholesterol at this position is slightly tilted, away from the plane of the membrane (γ, ***Figure 2***a, ***Figure 4***a). Additionally, in γ, Y207^{LD} creates a hydrophobic interaction with cholesterol, stabilizing the bent pose (***Figure 5***b). γ has been identified as the binding site of the synthetic agonist SAG in SMO (***Qi et al., 2020***). This provides further validation for γ being the major intermediate state in the Common Pathway.

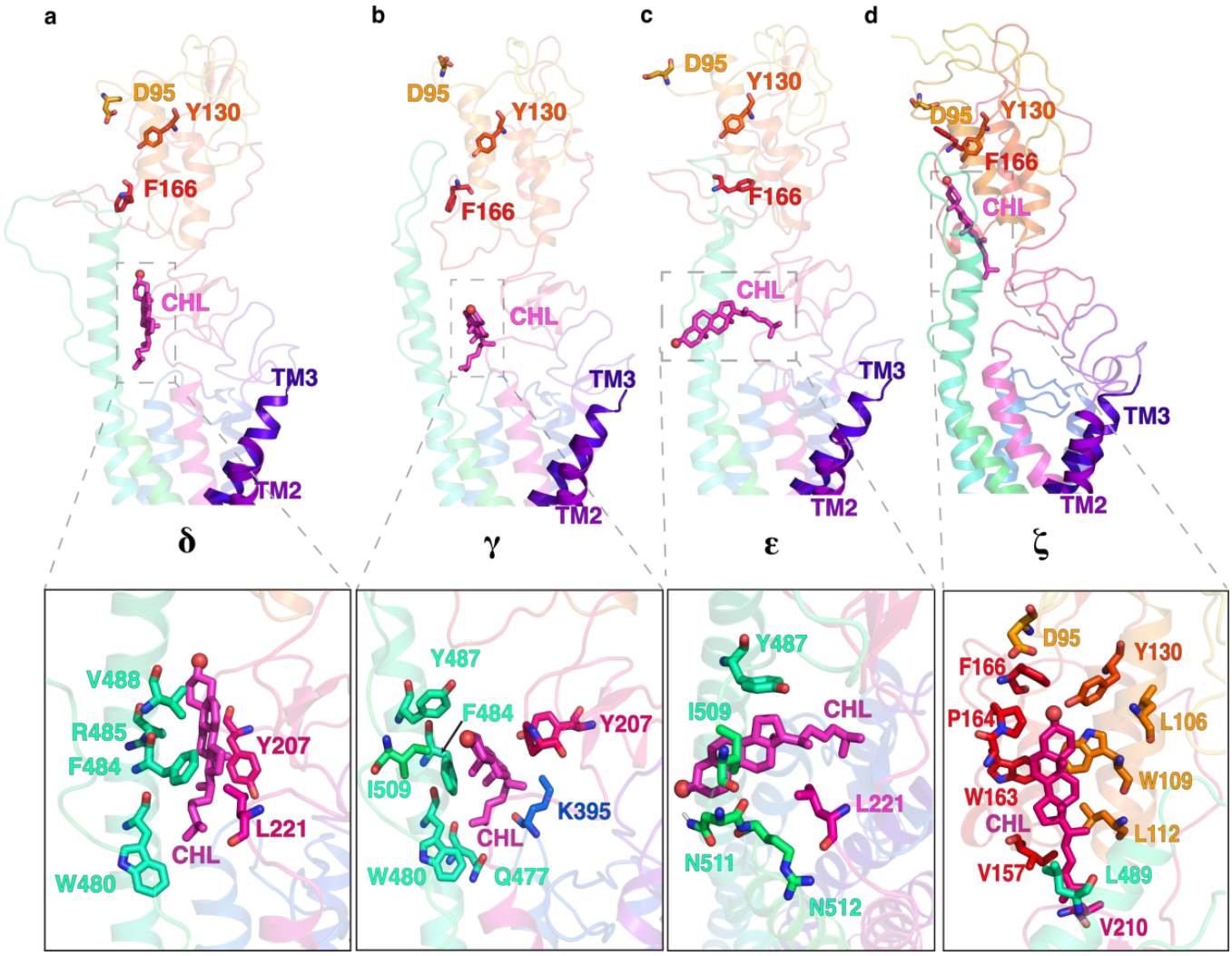


Figure 5. Multiple positions of cholesterol as it translocates through the Common Pathway, including the off-pathway intermediate. (a) upright (δ), (b) tilted (γ), (c) the overtilted off-pathway intermediate(ϵ), and (d) cholesterol at the CRD binding site (ζ). All snapshots presented are frames taken from MD simulations.

Figure 5—figure supplement 1. Hydrophobic Tunnel inside SMO

Figure 5—figure supplement 2. Titled cholesterol due to Linker Domain

Since the degrees of freedom accessible to cholesterol at this point in the pathway are higher than at the TMD, cholesterol can undergo "overtilting " as it approaches the CRD (ϵ, ***Figure 5***c). This state (ϵ) is attributed to an off-pathway intermediate state in the cholesterol translocation process, raising the timescales required for cholesterol to translocate to the CRD binding site. This overtilted pose is stabilized by hydrophobic interactions between the sterol and Y487$^{6.78f}$, L221^{LD}, and I509^{ECL3}, and a hydrogen bond between the side-chain of N511^{ECL3} and the alcoholic oxygen (***Figure 5***c). Once the cholesterol has crossed the TMD-CRD interface, it can reach the CRD sterol-binding site (ζ, ***Figure 5***d) in the CRD. In summary, we identify numerous conformational states of cholesterol-bound SMO, which are distinct from the available structures of sterol-bound SMO.

We performed experimental mutagenesis to validate the critical residues along this pathway, from the TMD to the CRD binding site. These mutants - Y^{LD}A, F$^{6.65f}$A, and I^{ECL3}A, all showed a significant decrease in activity compared to WT SMO (***Figure 6***a,b,***Figure 6—figure Supplement 1***) as well as an increase in peak PMF, suggesting that the force required to translocate cholesterol is higher than the WT residue in this position (***Figure 6***b, ***Figure 6—figure Supplement 2***). This fur-

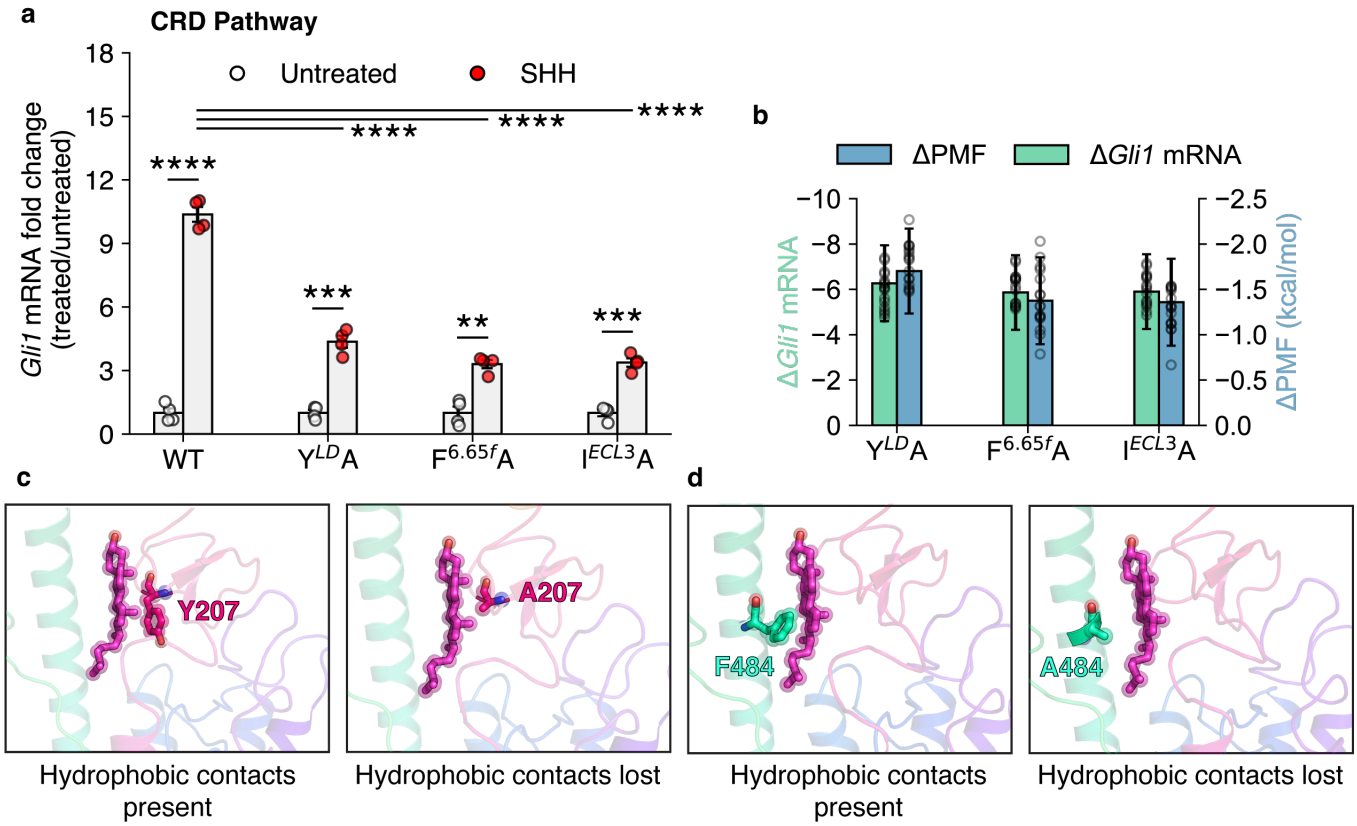


Figure 6. The effects of mutations along the translocation pathway connecting the TMD and CRD binding sites on the activation of SMO. (a) *Gli1* mRNA fold changes show the responsiveness of SMO mutants to SHH. Untreated Gli1 levels indicate low SMO activity, while SHH-treated values correspond to the level of SMO activation induced by SHH ligand. A t-test with Welch's correction was used to compute statistical significance. (P values: untreated vs treated: WT: 3×10^{-6}, $Y^{LD}A$: 2.46×10^{-4}, $F^{6.65f}A$: 1.08×10^{-3}, $I^{ECL3}A$: 1.12×10^{-4}, treated WT vs treated mutant: $F^{6.65f}A$: 1.6×10^{-5}, $I^{ECL3}A$: 1.6×10^{-5}, $Y^{LD}A$: 1.4×10^{-5}, key: Not significant (ns) $P > 0.05$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$, N=4 for all experiments.) (b) Δ*Gli1* mRNA fold change (SHH vs untreated) and Δ PMF (difference of peak PMF, calculated as PMF_{WT} - PMF_{mutant}) are plotted for mutants along the TMD-CRD pathway. (c, d) Example mutants $Y^{LD}A$ and $F^{6.65f}A$ show that cholesterol is unable to translocate through this pathway because of the loss of crucial hydrophobic contacts provided by Y207 and F484 and along the solvent-exposed pathway.

Figure 6—figure supplement 1. Mutation data for Common Pathway between TMD and CRD

Figure 6—figure supplement 2. PMF data for Common Pathway between TMD and CRD

Figure 6—figure supplement 3. Cholesterol positions along the entire transport pathway from membrane to CRD

ther implies that the mutants reduce the activity of SMO by increasing the energetic barriers for cholesterol translocation. In particular, the effect is pronounced for $Y^{LD}A$, which forms hydrophobic contacts along the pathway (***Figure 6***c). $F^{6.65f}A$ showed a significant decrease in SMO activity, which can be attributed to a reduction in hydrophobic stabilizing contacts that enable cholesterol's entry to the CRD (***Figure 6***d).

Overall, the entire cholesterol translocation process can be divided into two parts - entry from the membrane to the TMD binding site via Pathway 1 or 2, and translocation from the SMO TMD to the CRD binding site via the Common Pathway. The entire process is mapped out in ***Figure 6—figure Supplement 3***. The computational investigation shown here covers the dual-site model, where cholesterol reaches the CRD site via binding to the TM binding site first. In comparison to the CRD site, the TM site is more stable by ~ 2 kcal/mol (***Figure 2—figure Supplement 3***b, d). The experimental and computational analysis shows that both Pathway 1 and Pathway 2 are thermodynamically feasible pathways and exhibit similar energetic barriers for cholesterol to take from the membrane to the TMD binding site.

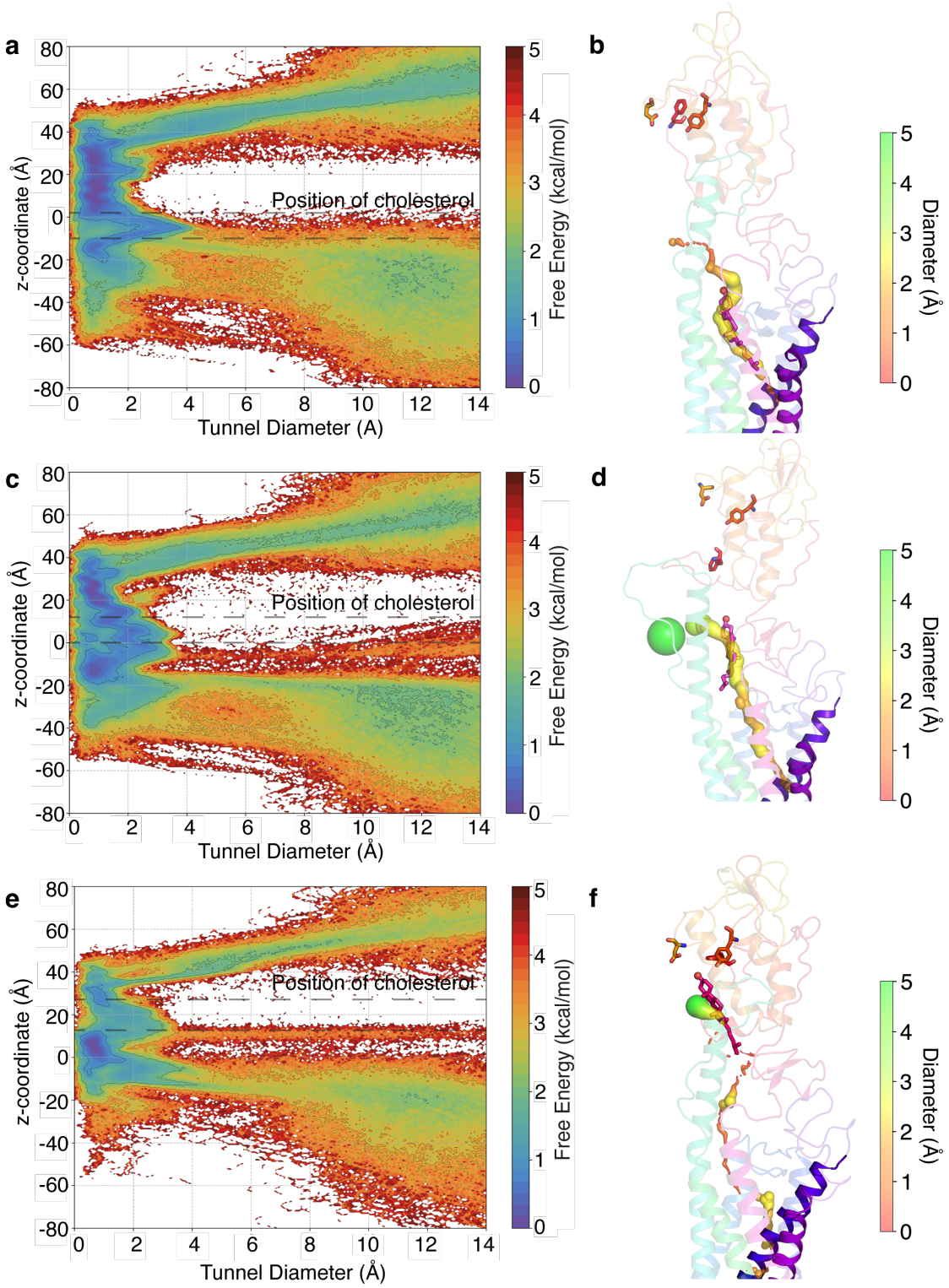


Figure 7. The tunnel profile during cholesterol translocation in SMO. (a) Free energy plot of the z-coordinate versus the tunnel diameter when cholesterol is present in the core TMD. The tunnel shows a spike in the radius in the TMD domain, indicating the presence of a cholesterol-accommodating cavity. (b) Representative figure for the tunnel when a cholesterol molecule is in the TMD. (c) Same as (a), when cholesterol is at the TMD-CRD interface. (d) same as (b), when cholesterol is at the TMD-CRD interface. (e) same as (a), when cholesterol is at the CRD binding site. (f) same as (b), when cholesterol is at the CRD binding site. Tunnel diameters are shown as spheres. Cholesterol positions are marked on plots using dotted lines. All snapshots presented are frames taken from MD simulations.

Figure 7—figure supplement 1. Error in Free Energies Fig 7

Figure 7—figure supplement 2. Average tunnel radius in SMO

A squeezing mechanism for translocation of cholesterol in SMO

To further elucidate the structural changes that happen in SMO during cholesterol translocation, we sought to characterize the conformation of the hydrophobic tunnel inside SMO. We calculated the tunnel diameter along the channel as cholesterol traverses through the protein (***Figure 7*** and ***Figure 7—figure Supplement 1***). The tunnel calculations were done with cholesterol at different points along the pathway as indicated by the metastable states in the Markov state models - with cholesterol in the TMD binding site, cholesterol at the CRD-TMD interface, and cholesterol present at the binding site in the CRD. When we plot cholesterol's position at different points along its transit through the TMD, we observed that the tunnel diameter varies along the z-coordinate as cholesterol moves through the channel. This corroborates that the tunnel radius is a function of the position of cholesterol in SMO. Furthermore, we observe that the tunnel radius shows a peak at z $\sim$ -7 when cholesterol is present in the core of the TMD (***Figure 7*** a,b). In the rest of the tunnel, the average radius remains relatively small (***Figure 7—figure Supplement 2***). This peak in diameter is seen to move along with the cholesterol position (***Figure 7*** c,d). This provides evidence that SMO uses a squeezing mechanism to translocate cholesterol. The term 'squeezing' here implies that the tunnel remains open only around cholesterol, and it closes as cholesterol moves away from its current position in the tunnel. Additionally, recent experimental work (***Zhang et al., 2022***) also suggests a squeeze-type mechanism for SMO. Once cholesterol has reached the binding site in the CRD, the tunnel in the TMD region is closed (***Figure 7*** e,f, ***Figure 7—figure Supplement 2***).

The cholesterol translocation mechanism of SMO shows similarities to the alternating access model proposed for substrate transport in membrane transporter proteins, where the transport tunnel closes behind the substrate to facilitate substrate transport across the membrane. Membrane transporters, including lipid transporters, use either an ion-based gradient or ATP hydrolysis to facilitate substrate transport. An ion binding site for SMO is currently unknown in contrast to Class A receptors, which have a sodium binding site (***Katritch et al., 2014***; ***Selvam et al., 2018***; ***Dutta et al., 2022b***). There is no experimental evidence of ion-coupling with the cholesterol export via SMO. Therefore, we posit that cholesterol translocation through SMO involves passive or concentration-dependent diffusion driven by a shift in the pool of accessible cholesterol, which rises once PTCH1 is inhibited. Thus, we provide a mechanistic overview of the dynamics of the cavity during cholesterol translocation in SMO.

The translocation of cholesterol occurs on a millisecond timescale

To give a perspective on the overall translocation process from a kinetic standpoint, we sought to calculate the timescales associated with the translocation of cholesterol from the membrane to the binding site in the CRD. Using a combination of Transition Path Theory and Markov state models, the reactive fluxes associated with each stage of the translocation cycle can be calculated. This enables us to calculate the mean first passage time (MFPT), which gives us an estimate of the timescales associated with the process (***Figure 8***). Pathway 1 was divided into two stages - the first being the translocation of cholesterol from the outer leaflet (***Figure 8***d) to the TMD - yielding a mean first passage time of 700 $\pm$ 122 μs. Here, the cholesterol tail has entered the TMD (***Figure 8***a). This was followed by cholesterol reaching the TMD binding site, with a timescale of 205 $\pm$ 41 μs (***Figure 8***b). For Pathway 2, cholesterol starts in the inner leaflet (***Figure 8***f). In the first stage, cholesterol undergoes flipping, which leads to a higher timescale than stage 1 of Pathway 1 - 823 $\pm$ 132 μs. This is followed by 122 $\pm$ 22 μs for cholesterol to reach the TMD binding site (***Figure 8***b).

Once cholesterol has reached the TMD binding site, it is followed by translocation of cholesterol from the TMD to the CRD binding site with a timescale of 255 $\pm$ 36 μs. Overall, the calculated timescales for cholesterol to reach the CRD site from the inner/outer leaflet of the membrane are 1023 $\pm$ 223 μs (Pathway 1) and 1134 $\pm$ 188 μs (Pathway 2). These timescales are comparable to the substrate transport timescales of Major Facilitator Superfamily (MFS) transporters (***Chan et al., 2022***). Furthermore, several experimental studies have also resolved the millisecond-scale kinetics of MFS transporters (***Blodgett and Carruthers, 2005***; ***Körner et al., 2024***; ***Bazzone et al., 2022***;

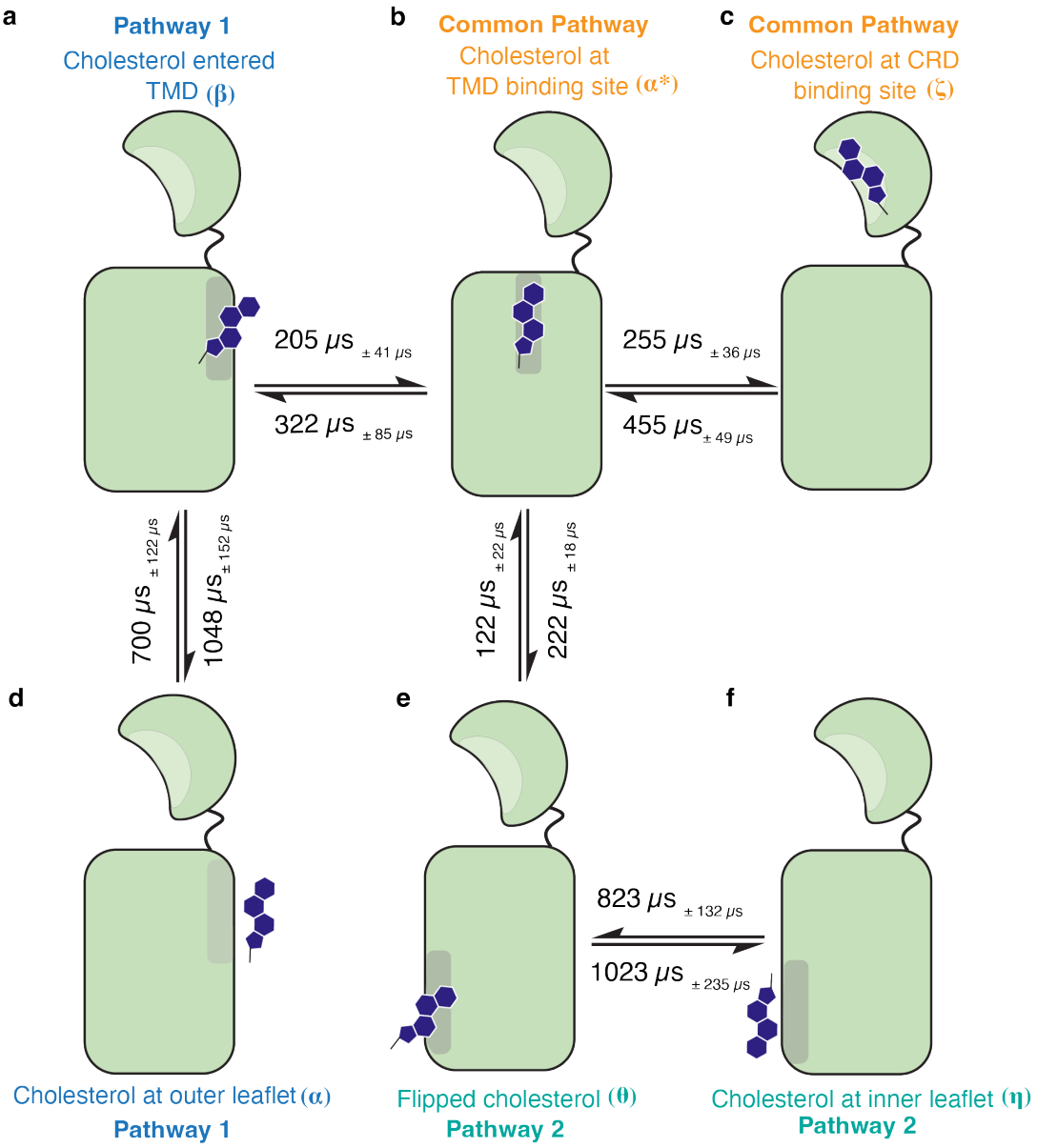


Figure 8. The timescales associated with the translocation of cholesterol through SMO. Each major intermediate state has been marked (a-f). Timescales were obtained by calculating the mean first passage time (MFPT) using the Markov state model. Errors in timescales are shown as subscripts. The arrows represent the relative flux for the translocation between subsequent steps. The overall process occurs at a timescale of ∼ 1ms.

Smirnova et al., 2014; ***Zhu et al., 2019***), further corroborating the results from our study. Interestingly, the timescales for the reverse process of translocating cholesterol from the CRD binding site to the membrane are higher for each step (***Figure 8***), indicating that the reverse process has a higher energetic barrier associated with it. This provides kinetic evidence that the overall translocation process from the membrane to the CRD binding site is thermodynamically favorable. Thus, it can be concluded that SMO facilitates the translocation of cholesterol.

Conclusions

In this study, we have used a combination of millisecond-scale atomistic molecular dynamics simulations, Markov state modeling, and experimental mutagenesis to describe the step-by-step pro-

cess of cholesterol translocation through SMO. Previous structural studies have delineated multiple translocation pathways for cholesterol transport via SMO. In this study, we have examined the mechanism of cholesterol translocation from the membrane to the CRD sterol-binding site of SMO via two modes - the outer leaflet pathway, between TM2 and TM3, and the inner leaflet pathway, between TM5 and TM6. We quantitatively assess the thermodynamic barriers of the translocation pathways and estimate the timescales associated with the process. The key intermediate cholesterol-bound conformations of SMO and the role of specific residues in the translocation process were identified computationally and validated using both experimental and *in silico* mutagenesis.

We observe that cholesterol can move through two distinct conduits, starting at either the outer leaflet of the membrane (Pathway 1) or the inner leaflet (Pathway 2), followed by translocation to the cholesterol binding site in the CRD. In the first mode (Pathway 1), cholesterol enters the protein between TM2 and TM3 in the outer leaflet, and then it is translocated along the extended TM6 to reach the CRD binding site. The highest barrier along this first mode is associated with the cholesterol entry from the outer leaflet of the membrane to the TMD binding site. Similarly, the highest barrier for the second mode (Pathway 2) also involves cholesterol translocation from the inner leaflet membrane to the TMD of SMO. We show that the barriers associated with the pathway starting from the outer leaflet are lower by ~0.7 kcal (p=0.0013). We also provide evidence that cholesterol can enter SMO via both leaflets, considering that multiple computational and experimental studies have found cholesterol entry sites and activation modulation via the outer leaflet, between TM2-TM3 (***Hedger et al., 2019***; ***Kinnebrew et al., 2021***; ***Bansal et al., 2023***). Other experimental and computational studies have proposed that cholesterol can enter SMO via the inner leaflet (***Huang et al., 2018***; ***Zhang et al., 2022***). Overall, our work shows that cholesterol translocation from either pathway is feasible.

The second-highest barrier in cholesterol translocation is at the TMD-CRD interface, but it is ~1 kcal/mol lower than the barrier for cholesterol entry. We also show that the cholesterol translocation process occurs via a squeezing mechanism that maintains the forward flux of cholesterol from the membrane to the CRD binding site. Overall, the translocation process takes place on the millisecond timescale with multiple intermediate states identified using simulations and supported by the sterol-bound structures of SMO.

Despite the extensive MD simulations reported in this study, there is still a need to further probe the endogenous "activation" of SMO by cholesterol in a position-dependent manner. Here, we have explored the role the CRD-site plays in SMO activation. In addition, through simulating the CRD site-dependent SMO activation hypothesis, we have also simulated the TMD site-dependent activation. We show that the overall stability of cholesterol in the TMD site is higher than the stability of cholesterol in the CRD site by ~ 2 kcal/mol. An alternative possibility states that the flexibility associated with the CRD would allow it to directly access the membrane and, consequently, cholesterol. In the extensive simulations reported in this study, the binding site of cholesterol in the CRD remains at least 20 Å away from the nearest lipid head group in the membrane, suggesting that such direct extraction and the bending of the CRD do not occur within the timescales sampled (Appendix 2 – ***Figure 6***). The mechanistic details of this process are still unexplored and form the basis of future work. Additionally, GPCRs exist in a conformational equilibrium of the active and inactive states, and the job of the agonist (cholesterol) is to lower the barrier to activation and shift the equilibrium towards the active state. Our recent work (***Kim et al., 2024***) discusses how the binding position of cyclopamine modulates SMO activity. Cyclopamine acts as an antagonist when bound to SMO TMD and acts as an agonist when bound to SMO CRD. Cholesterol binding at different positions along the translocation pathway leads to a position-dependent modulation of SMO activity (***Kinnebrew et al., 2022***; ***Kumari et al., 2023***). ***Kinnebrew et al.*** propose that binding of cholesterol to the TMD determines basal SMO activity, whereas binding to both the TMD and CRD is required for full, SHH-induced activity. In this study, we only focused on the cholesterol movement from the membrane to the CRD binding site. Therefore, a future investigation is needed to fully sample the

activation process of SMO with cholesterol bound at different positions.

Our study provides computational and experimental evidence for the translocation of cholesterol along the conduit within SMO and outlines the entire translocation mechanism from a kinetic and thermodynamic perspective. Our findings provide a plausible model for how an increase in membrane cholesterol accessibility can activate SMO: by promoting the entry and movement of cholesterol along a conduit that traverses the center of the protein and ends at the CRD. Our study provides a framework for the development of drugs against oncogenic SMO that could function by blocking intermediate states of transport.

Methods

Molecular Dynamics (MD) Simulations

Simulation setup

Structures of SMO bound to sterols - 6XBL (***Qi et al., 2020***) (SMO-CHL-1), 6XBM (***Qi et al., 2020***) (SMO-CHL-2, SMO-CHL-3) and 5L7D (***Byrne et al., 2016***) (SMO-CHL-4) were used as the starting structures for simulations. For SMO-CHL-1, the bound agonist SAG was removed. The two sterols occupying different positions in the tunnel in 6XBM were used to build 2 separate systems with cholesterol at different sites in the pathway (SMO-CHL-2, SMO-CHL-3). For SMO-CHL-2 and SMO-CHL-3, to account for the lack of CRD in the structure 6XBM, the sterol positions were aligned to the full-length SMO (6XBL, 0.8 Å RMSD from 6XBM) and the 24(S),25-epoxycholesterol for each system was replaced by cholesterol. For SMO-CHL-4, the inactivating mutation $V^{3.40}F$ was mutated back to Wild Type. For all systems, any stabilizing antibodies and bound G proteins were removed. The missing residues in the intra/extracellular loop for every protein were modeled using MODELLER (***Webb and Sali, 2016***) (Appendix 1 – ***Table 1***). Termini for all proteins were capped using acetyl (ACE) and N-methylamino (NME) at the N- and C-termini to ensure neutrality. All four protein systems with cholesterols at different points along the translocation path were embedded in a lipid bilayer. The composition of the bilayer was set similar to mice cerebellum (***Scandroglio et al., 2008***) (Appendix 1 – ***Table 2***), to mimic physiological conditions, using CHARMM-GUI (***Jo et al., 2008***; ***Lee et al., 2018***). Interactions between the atoms - bonded and non-bonded- were modeled using the CHARMM36 Force Field (***Klauda et al., 2010***; ***Best et al., 2012***). TIP3P water (***Jorgensen et al., 1983***) and 0.15 M NaCl were used to solvate the system, to mimic physiological conditions. Non-protein hydrogen masses were repartitioned to 3.024 Da, to enable use of a longer timestep (4fs) (***Hopkins et al., 2015***). Starting points for Pathways 1 and 2 were chosen from already simulated data, according to the closest cholesterol distance from the respective helices (outer leaflet, TM2-TM3 for Pathway 1, lower leaflet, TM5-TM6 for Pathway 2). This was done once the rest of the pathway was completely explored.

Pre-Production MD

All systems were subject to 50,000 steps of initial minimization. Further, minimization was performed for another 10000 steps by constraining the hydrogens using SHAKE (***Andersen, 1983***). Systems were then heated to 310K to mimic physiological conditions, at NVT for 10 ns. This was followed by equilibration at NPT and 1 bar for 5 ns. Backbone constraints of 10 kcal/mol/$Å^2$ were applied during NVT and NPT. Next, systems were equilibrated at NPT for 40 ns without constraints to ensure system stability. All pre-production steps were performed using AMBER18 (***Case et al., 2020***; ***Salomon-Ferrer et al., 2012***; ***Case et al., 2005***; ***Götz et al., 2012***; ***Salomon-Ferrer et al., 2013***).

Production MD

Systems were then subjected to some initial sampling of 100ns each. This was followed by clustering and performing adaptive sampling to enable a divide-and-conquer approach to sample the conformational landscape. 3 rounds of sampling were performed on each system. This was followed by similar rounds of adaptive sampling on the distributed computing project Folding@Home

(http://foldingathome.org). OpenMM 7.5.1 (***Eastman et al., 2017***) was used for running simulations on Folding@Home.

In all simulations, 4fs was the chosen integration timestep. Particle Mesh Ewald (PME) (***Darden et al., 1993***) method was used to account for long-range electrostatics. The cutoff for considering non-bonded interactions was set to 10 Å. Temperature was maintained using the Langevin Thermostat (***Davidchack et al., 2009***). Pressure was maintained using the Monte-Carlo Barostat (***Åqvist et al., 2004***). All hydrogen bonds were constrained using SHAKE (***Andersen, 1983***).

Steered MD

Steered MD was performed to steer cholesterol from the membrane into the protein, for generating the starting frames for Pathway 1 and Pathway 2. This was done by first finding frames from the existing data, where cholesterols were closest to the respective starting points. Then, cholesterols were steered towards the center of TMD with the end point being the cholesterol binding site deep in the TMD, as resolved in the structure 6XBL. This was done using a steering force of 20 kcal/(mol $Å^2$), over a course of 500ns. The entire protein, except the helices involved at the entry (TM2-TM3: Pathway 1; TM5-TM6: Pathway 2), was constrained using two RMSD restraints during simulations to prevent any unphysical effects. The force constants used for RMSD restraints are as follows: 10 kcal/(mol $Å^2$) for restraining residues from CRD to ICL1, and 35 kcal/(mol $Å^2$) for restraining residues from ICL2 to helix8 for Pathway 1. For Pathway 2, an RMSD restraint of 35 kcal/(mol $Å^2$) was used to restrain residues from CRD to ECL2, and 10 kcal/(mol $Å^2$) to restrain residues from ECL3 to helix8. Three replicates of steered MD were performed to ensure that the pathway explored converged. The frames generated from these runs were then used as seed frames to start simulations for exploring the entry of cholesterol into SMO. Steered MD was performed using NAMD (***Phillips et al., 2020***, ***2005***). Each frame generated by Steered MD was minimized for 50000 steps, and then equilibrated for 40 ns, using AMBER using the same methodology for these seed frames as described in section Pre-Production MD.

Adaptive Biasing Force-Based Sampling for PMF generation

To elucidate the effect of mutations on the cholesterol translocation barriers, we used an adaptive-biasing force (ABF) based sampling for generating the potential of mean force (PMF) profiles for each case, and compared them to the wild-type (WT) translocation barriers. For generating the starting files for every mutant system, psfgen (***Humphrey et al., 1996***; ***Stone, 1998***) was used. SMO WT PMF profiles were generated for each pathway, and mutant profiles were generated for their respective pathways. A biasing potential of 45 kcal/(mol $Å^2$) was used for both lower wall and upper wall of the ABF potential. Each mutant was run for 3 replicates, and 10^7 samples were generated for each pathway to compute the PMF. NAMD was used for this purpose (***Phillips et al., 2020***, ***2005***).

Adaptive sampling, feature selection and clustering

The cholesterol translocation process in SMO was simulated in a stage-wise process, with the four starting points having cholesterol present at different points along the translocation pathway (***Figure 1***). To accelerate the sampling of the entire translocation process, simulations were performed using an approach that parallelizes the exploration of conformational space. Adaptive sampling (***Hinrichs and Pande, 2007***; ***Bowman et al., 2010***; ***Kleiman et al., 2023***) was utilized to achieve this acceleration, which uses an iterative sampling approach involving picking the next round of simulation starting points based on the current data. Several different types of machine learning and heuristic-based adaptive sampling approaches have been proposed in the literature (***Kleiman et al., 2023***; ***Kleiman and Shukla, 2023***, ***2022***; ***Shamsi et al., 2018***, ***2017***; ***Zimmerman and Bowman, 2015***; ***Weber and Pande, 2011***). However, least count-based sampling, where the least visited states are chosen as the starting points for the next set of simulations, is among the best sampling strategies for exploration of the conformational free energy landscapes (***Weber and Pande, 2011***; ***Nadeem and Shukla, 2025***). The following approach was used to collect the data:

1. Initially, sampling was started from the 4 starting points with cholesterol at different points along the tunnel (***Figure 1***). Each starting point was simulated as a separate system.
2. All systems were subject to Pre-Production MD (Refer Pre-Production MD).
3. Following Pre-Production, all systems were subject to 200 ns of Production MD (Refer Production MD).
4. The Production data until this point was combined and clustered on the basis of Adaptive Sampling metrics (Appendix 1 – ***Table 3***) using k-means clustering (for the first 4 rounds) and then mini-batch k-means later on. pyEMMA (***Scherer et al., 2015***) was used for this purpose.
5. Frames were chosen from the clusters with the least populations, and were used as seeds for the next round of simulations. This enabled a parallel iterative approach for sampling the conformational landscape (Appendix 1 – ***Table 4***, Appendix 2 – ***Figure 1***)
6. Steps 4-5 were repeated until the entire landscape was explored.

The entire process of cholesterol translocation - from cholesterol present in the membrane to cholesterol bound to the CRD sterol-binding site - was simulated. A total of 2 milliseconds of unbiased simulation data was collected using this approach.

Dimensionality reduction using tICA

Dimensionality reduction was performed on the high-dimensional data set using Time-Lagged Independent Component Analysis (tICA) (***Schwantes and Pande, 2013***). tICA uses a linear transformation to project the input data into a lower-dimensional basis set, the components of which approximate the slowest processes observed in the simulations. In our case, the slowest process simulated was identified by tICA as the translocation of cholesterol (***Figure 2—figure Supplement 2***). The dataset was divided into 2 different groups, separately for Pathways 1 and 2. To further gain insights from the simulations, we constructed a Markov state model from the tICA-projected simulation data.

Markov state model construction

Markov State Model (MSM) is a kinetic modeling technique that uses short trajectory data that sample local transitions to provide a global estimate of the thermodynamics and kinetics of the physical process (***Pande et al., 2010***; ***Noé and Rosta, 2019***; ***Prinz et al., 2011***; ***Husic and Pande, 2018***; ***Wang et al., 2017***; ***Shukla et al., 2015***). A MSM discretizes the dataset into kinetically distinct microstates and calculates the rates of transitions among such microstates. This methodology has been used extensively to investigate the conformational dynamics of membrane proteins, including G-protein coupled receptors (***Deganutti et al., 2024***; ***Fleetwood et al., 2021***; ***Dutta et al., 2022a***,b; ***Chan et al., 2022***; ***Kohlhoff et al., 2013***; ***Kim et al., 2024***; ***Bansal et al., 2023***). To construct the Markov state model, the data collected from simulations were first featurized (Appendix 1, – ***Table 3***). Pathways 1 and 2 were treated independently of each other. Time-lagged independent component analysis (tICA) was performed on the data to reduce the dimensionality of the input data and identify the slowest processes observed in the simulations. A Markov state model was then constructed on the 42 and 28 components from the tIC space for Pathways 1 and 2, respectively. To construct the models for both pathways, the following approach was used:

1. The data was clustered into different numbers of clusters, ranging from 100-1000, and the implied timescales were calculated as a function of the MSM lag time. The lag time after which the implied timescales converged (30ns) in both cases was chosen as the MSM lag time (Appendix 2, – ***Figure 2***a, Appendix 2, – ***Figure 3***a).
2. Once the MSM lag times were chosen, a grid-search-based approach was chosen to compute the optimal number of clusters. Each constructed MSM was evaluated using a VAMP2 score (***Wu and Noé, 2019***), where the sum of squares of the top five eigenvalues was computed. Each pathway's dataset was clustered into several clusters (200 - Pathway 1, 400 - Pathway 2) that gave the highest VAMP2 score (Appendix 2, – ***Figure 2***b, Appendix 2, – ***Figure 3***b).

3. Once the MSMs were constructed, bootstrapping was performed, with 200 rounds and 80% of the data in each round, to compute the error associated with the probabilities.
4. To validate the final MSM, a Chapman-Kolmogorov test was performed (Appendix 2, – ***Figure 4***, Appendix 2, – ***Figure 5***) to show the long-timescale validity of the Markovian property followed by the constructed models.

This information can be used to calculate the probability of each state, which can be used to recover the thermodynamic and kinetic properties of the entire ensemble. One of the use cases of the probabilities is that they can be used to reweigh the projected free energy of each datapoint along a said reaction coordinate, which has been used in ***Figure 2***, ***Figure 4***, and ***Figure 7***. The errors in free energies were computed using the errors in the projected probabilities from the bootstrapped MSMs (***Figure 2—figure Supplement 1***, and ***Figure 7—figure Supplement 1***).

Trajectory Analysis and Visualization

Trajectories were stripped of water and imaged before analysis to allow faster computation. cpptraj (***Roe and Cheatham, 2013***) was used for this purpose. For constructing the figures, VMD (***Humphrey et al., 1996***; ***Stone, 1998***) and open-source PyMOL (***Schroedinger, LLC, 2022***) (rendering) MDTraj (***McGibbon et al., 2015***) (computing observables from trajectories), matplotlib (***Hunter, 2007***) and seaborn (***Waskom, 2021***) (Plot rendering), Numpy (***Harris et al., 2020***) (numerical computations), HOLE (***Smart et al., 1993***) (tunnel diameter calculations) were used.

Cell culture and cell line generation

Mouse embryonic fibroblasts (MEFs) lacking Smoothened (*Smo*$^{-/-}$) were tested to ensure lack of endogenous SMO protein using immunoblotting, as described previously (***Rohatgi et al., 2007***). These *Smo*$^{-/-}$ MEF cells were used to generate stable cell lines expressing SMO mutants, which were then authenticated by immunoblotting to ensure stable expression of the transgene (***Nachtergaele et al., 2013***). Cell lines were confirmed to be negative for mycoplasma infection.

MEF cells were grown in high-glucose DMEM (Thermo Fisher Scientific, catalog no. SH30081FS) containing 10% FBS (Sigma-Aldrich, catalog no. S11150) and the following supplements: 1 mM sodium pyruvate (Gibco, catalog no. 11-360-070), 2 mM L-glutamine (GeminiBio, catalog no. 400106), 1× minimum essential medium NEAA solution (Gibco, catalog no. 11140076), penicillin (40 U/ml), and streptomycin (40 μg/ml) (GeminiBio, catalog no. 400109). This media, hereafter referred to as supplemented DMEM, was sterilized through a 0.2-μm filter and stored at 4°C.

To measure Hedgehog responsiveness by quantitative polymerase chain reaction (PCR) or Western blotting, cells were seeded in 10% FBS-supplemented DMEM and grown to confluence. To induce ciliation, a requirement for Hedgehog signaling, cells were serum-starved in 0.5% FBS-supplemented DMEM and simultaneously treated with Sonic Hedgehog (SHH) for 24 hours before analysis.

Western blotting was carried out to assess SMO protein expression for all mutants. Briefly, whole-cell extracts were prepared in lysis buffer containing 150 mM NaCl, 50 mM Tris-HCl (pH 8), 1% NP-40, 1× protease inhibitor (SigmaFast Protease inhibitor cocktail, EDTA-free; Sigma-Aldrich, catalog no. S8830), 1 mM $MgCl_2$, and 10% glycerol. After lysate clarification by centrifugation at 20,000g, samples were resuspended in 50 mM tris(2-carboxyethyl)phosphine and 1× Laemmli buffer for 30 min at 37°C. Samples were then subjected to SDS-polyacrylamide gel electrophoresis, followed by immunoblotting with antibodies against GLI1 [anti-GLI1 mouse monoclonal (clone L42B10); Cell Signaling Technology, catalog no. 2643, RRID: AB_2294746], SMO (rabbit polyclonal) (***Rohatgi et al., 2007***), or GAPDH [anti-GAPDH mouse monoclonal (clone 1E6D9); Protein tech, catalog no. 60004-1-Ig, RRID: AB_2107436].

Measuring Hedgehog signaling with quantitative PCR

Gli1 mRNA transcript levels were measured using the Power SYBR Green Cells-to-CT kit (Thermo Fisher Scientific). *Gli1* levels relative to *Gapdh* were calculated using the Delta-Ct method (CT(*Gli1*)

- CT(*Gapdh*). The RT-PCR was carried out using custom primers for *Gli1* (forward primer: 5′-ccaagccaactttatgtcaggg-3′ and reverse primer: 5′-agcccgcttctttgttaatttga-3′), and *Gapdh* (forward primer: 5′-agtggcaaagtggagatt-3′ and reverse primer: 5′-gtggagtcatactggaaca-3′).

Acknowledgements

The authors thank The Blue Waters Petascale Computing Facility and National Center for Supercomputing Applications, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. D.S. acknowledges support from NIH grant R35GM142745 and the Cancer Center at Illinois. R.R. acknowledges support from NIH grant GM118082. The authors also acknowledge support from citizen scientists providing computing hours for simulations performed on Folding@Home, which have enabled us to collect data on a large scale for this project.

Author Contributions

D.S. and R.R. acquired funding for the project. D.S. supervised the project. D.S. and P.B. designed the research. P.B. performed simulations and analyzed the computational data. M.K. designed and performed experiments. P.B. wrote the manuscript with inputs from M.K., R.R., and D.S.

Competing interests

The authors declare no competing interests.

Data Availability

Relevant trajectory files have been shared via Box (https://uofi.box.com/s/zldrqparb39ax72c637amcxi43hv2yaq) and are available publicly. Python scripts used for analysis are available on GitHub (https://github.com/ShuklaGroup/Bansal_et_al_Cholesterol_Smoothened_2024).

References

Andersen HC. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. Journal of Computational Physics. 1983 Oct; 52(1):24–34. doi: 10.1016/0021-9991(83)90014-1.

Ansell TB, Corey RA, Viti LV, Kinnebrew M, Rohatgi R, Siebold C, Sansom MSP. The energetics and ion coupling of cholesterol transport through Patched1. Science Advances. 2023 Aug; 9(34):34. doi: 10.1126/sciadv.adh1609.

Åqvist J, Wennerström P, Nervall M, Bjelic S, Brandsdal BO. Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. Chemical Physics Letters. 2004 Jan; 384(4-6):288–294. doi: 10.1016/j.cplett.2003.12.039.

Axelson M, Liu K, Jiang X, He K, Wang J, Zhao H, Kufrin D, Palmby T, Dong Z, Russell AM, Miksinski S, Keegan P, Pazdur R. U.S. Food and Drug Administration Approval: Vismodegib for Recurrent, Locally Advanced, or Metastatic Basal Cell Carcinoma. Clinical Cancer Research. 2013 Apr; 19(9):2289–2293. doi: 10.1158/1078-0432.ccr-12-1956.

Bansal PD, Dutta S, Shukla D. Activation mechanism of the human Smoothened receptor. Biophysical Journal. 2023 Apr; 122(7):1400–1413. doi: 10.1016/j.bpj.2023.03.007.

Bazzone A, Tesmer L, Kurt D, Kaback HR, Fendler K, Madej MG. Investigation of sugar binding kinetics of the E. coli sugar/H+ symporter XylE using solid-supported membrane-based electrophysiology. Journal of Biological Chemistry. 2022 Feb; 298(2):101505. doi: 10.1016/j.jbc.2021.101505.

Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. Journal of Chemical Theory and Computation. 2012 Aug; 8(9):3257–3273. doi: 10.1021/ct300400x.

Bidet M, Joubert O, Lacombe B, Ciantar M, Nehmé R, Mollat P, Brétillon L, Faure H, Bittman R, Ruat M, Mus-Veteau I. The Hedgehog Receptor Patched Is Involved in Cholesterol Transport. PLoS ONE. 2011 Sep; 6(9):e23834. doi: 10.1371/journal.pone.0023834.

Blassberg R, Macrae JI, Briscoe J, Jacob J. Reduced cholesterol levels impair Smoothened activation in Smith–Lemli–Opitz syndrome. Human Molecular Genetics. 2015 Dec; 25(4):693–705. doi: 10.1093/hmg/ddv507.

Blodgett DM, Carruthers A. Quench-Flow Analysis Reveals Multiple Phases of GluT1-Mediated Sugar Transport. Biochemistry. 2005 Jan; 44(7):2650–2660. doi: 10.1021/bi048247m.

Bowman GR, Ensign DL, Pande VS. Enhanced Modeling via Network Theory: Adaptive Sampling of Markov State Models. Journal of Chemical Theory and Computation. 2010 Feb; 6(3):787–794. doi: 10.1021/ct900620b.

Byrne EFX, Sircar R, Miller PS, Hedger G, Luchetti G, Nachtergaele S, Tully MD, Mydock-McGrane L, Covey DF, Rambo RP, Sansom MSP, Newstead S, Rohatgi R, Siebold C. Structural basis of Smoothened regulation by its extracellular domains. Nature. 2016 Jul; 535(7613):517–522. doi: 10.1038/nature18934.

Case DA, Belfon K, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TEI, Cruzeiro VWD, Darden TA, Duke RE, Giambasu G, Gilson MK, Gohlke H, Goetz AW, Harris R, Izadi S, Izmailov SA, Kasavajhala K, Kovalenko A, Krasny R, Kurtzman T, et al. AMBER 2020. University of California, San Francisco; 2020.

Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. Journal of Computational Chemistry. 2005; 26(16):1668–1688. doi: 10.1002/jcc.20290.

Chan MC, Selvam B, Young HJ, Procko E, Shukla D. The substrate import mechanism of the human serotonin transporter. Biophysical Journal. 2022 Mar; 121(5):715–730. doi: 10.1016/j.bpj.2022.01.024.

Cooper MK, Wassif CA, Krakowiak PA, Taipale J, Gong R, Kelley RI, Porter FD, Beachy PA. A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. Nature Genetics. 2003 Mar; 33(4):508–513. doi: 10.1038/ng1134.

Corbit KC, Aanstad P, Singla V, Norman AR, Stainier DYR, Reiter JF. Vertebrate Smoothened functions at the primary cilium. Nature. 2005 Aug; 437(7061):1018–1021. doi: 10.1038/nature04117.

Darden T, York D, Pedersen L. Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. Journal of Chemical Physics. 1993 Jun; 98(12):10089–10092. doi: 10.1063/1.464397.

Darve E, Rodríguez-Gómez D, Pohorille A. Adaptive biasing force method for scalar and vector free energy calculations. The Journal of Chemical Physics. 2008 Apr; 128(14). doi: 10.1063/1.2829861.

Davidchack RL, Handel R, Tretyakov MV. Langevin thermostat for rigid body dynamics. Journal of Chemical Physics. 2009 Jun; 130(23):234101. doi: 10.1063/1.3149788.

Davies JP, Chen FW, Ioannou YA. Transmembrane Molecular Pump Activity of Niemann-Pick C1 Protein. Science. 2000 Dec; 290(5500):2295–2298. doi: 10.1126/science.290.5500.2295.

Deganutti G, Pipitò L, Rujan RM, Weizmann T, Griffin P, Ciancetta A, Moro S, Reynolds CA. Hidden GPCR structural transitions addressed by multiple walker supervised molecular dynamics (mwSuMD). elife. 2024 May; p. doi: 10.7554/elife.96513.1. doi: 10.7554/elife.96513.1.

Denef N, Neubüser D, Perez L, Cohen SM. Hedgehog Induces Opposite Changes in Turnover and Subcellular Localization of Patched and Smoothened. Cell. 2000 Aug; 102(4):521–531. doi: 10.1016/s0092-8674(00)00056-8.

Deshpande I, Liang J, Hedeen D, Roberts KJ, Zhang Y, Ha B, Latorraca NR, Faust B, Dror RO, Beachy PA, Myers BR, Manglik A. Smoothened stimulation by membrane sterols drives Hedgehog pathway activity. Nature. 2019 Jul; 571(7764):284–288. doi: 10.1038/s41586-019-1355-4.

Dutta S, Selvam B, Das A, Shukla D. Mechanistic origin of partial agonism of tetrahydrocannabinol for cannabinoid receptors. Journal of Biological Chemistry. 2022 Apr; 298(4):101764. doi: 10.1016/j.jbc.2022.101764.

Dutta S, Selvam B, Shukla D. Distinct Binding Mechanisms for Allosteric Sodium Ion in Cannabinoid Receptors. ACS Chemical Neuroscience. 2022 Jan; 13(3):379–389. doi: 10.1021/acschemneuro.1c00760.

Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang LP, Simmonett AC, Harrigan MP, Stern CD, Wiewiora RP, Brooks BR, Pande VS. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLOS Computational Biology. 2017 Jul; 13(7):e1005659. doi: 10.1371/journal.pcbi.1005659.

Fleetwood O, Carlsson J, Delemotte L. Identification of ligand-specific G protein-coupled receptor states and prediction of downstream efficacy via data-driven modeling. eLife. 2021 Jan; 10:e60715. doi: 10.7554/elife.60715.

Gong X, Qian H, Cao P, Zhao X, Zhou Q, Lei J, Yan N. Structural basis for the recognition of Sonic Hedgehog by human Patched1. Science. 2018 Aug; 361(6402). doi: 10.1126/science.aas8935.

Götz AW, Williamson MJ, Xu D, Poole D, Grand SL, Walker RC. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. Journal of Chemical Theory and Computation. 2012 Apr; 8(5):1542–1555. doi: 10.1021/ct200909j.

Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, et al. Array programming with NumPy. Nature. 2020 Sep; 585(7825):357–362. doi: 10.1038/s41586-020-2649-2.

Hedger G, Koldsø H, Chavent M, Siebold C, Rohatgi R, Sansom MSP. Cholesterol Interaction Sites on the Transmembrane Domain of the Hedgehog Signal Transducer and Class F G Protein-Coupled Receptor Smoothened. Structure. 2019 Mar; 27(3):549–559.e2. doi: 10.1016/j.str.2018.11.003.

Hénin J, Fiorin G, Chipot C, Klein ML. Exploring Multidimensional Free Energy Landscapes Using Time-Dependent Biases on Collective Variables. Journal of Chemical Theory and Computation. 2009 Dec; 6(1):35–47. doi: 10.1021/ct9004432.

Hinrichs NS, Pande VS. Calculation of the distribution of eigenvalues and eigenvectors in Markovian state models for molecular dynamics. The Journal of Chemical Physics. 2007 Jun; 126(24):244101. doi: 10.1063/1.2740261.

Hoffmann M, Scherer MK, Hempel T, Mardt A, de Silva B, Husic BE, Klus S, Wu H, Kutz JN, Brunton S, Noé F. Deeptime: a Python library for machine learning dynamical models from time series data. Machine Learning: Science and Technology. 2021; doi: 10.1088/2632-2153/ac3de0.

Hopkins CW, Grand SL, Walker RC, Roitberg AE. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. Journal of Chemical Theory and Computation. 2015 Mar; 11(4):1864–1874. doi: 10.1021/ct5010406.

Huang P, Nedelcu D, Watanabe M, Jao C, Kim Y, Liu J, Salic A. Cellular Cholesterol Directly Activates Smoothened in Hedgehog Signaling. Cell. 2016 Aug; 166(5):1176–1187.e14. doi: 10.1016/j.cell.2016.08.003.

Huang P, Zheng S, Wierbowski BM, Kim Y, Nedelcu D, Aravena L, Liu J, Kruse AC, Salic A. Structural Basis of Smoothened Activation in Hedgehog Signaling. Cell. 2018 Jul; 174(2):312–324.e16. doi: 10.1016/j.cell.2018.04.029.

Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. Journal of Molecular Graphics. 1996 Feb; 14(1):33–38. doi: 10.1016/0263-7855(96)00018-5.

Hunter JD. Matplotlib: A 2D graphics environment. Computing in Science & Engineering. 2007; 9(3):90–95. doi: 10.1109/MCSE.2007.55.

Husic BE, Pande VS. Markov State Models: From an Art to a Science. Journal of the American Chemical Society. 2018 Feb; 140(7):2386–2396. doi: 10.1021/jacs.7b12191.

Ingham PW. 1. In: Hedgehog signaling Elsevier; 2022. p. 1–58. doi: 10.1016/bs.ctdb.2022.04.003.

Ingham PW, Nystedt S, Nakano Y, Brown W, Stark D, van den Heuvel M, Taylor AM. Patched represses the Hedgehog signalling pathway by promoting modification of the Smoothened protein. Current Biology. 2000 Oct; 10(20):1315–1318. doi: 10.1016/s0960-9822(00)00755-7.

Ingham P, Taylor A, Nakano Y. Role of the Drosophila patched gene in positional signalling. Nature. 1991; 353(6340):184–187. doi: 10.1038/353184a0.

Jain S, Song R, Xie J. Sonidegib: mechanism of action, pharmacology, and clinical utility for advanced basal cell carcinomas. OncoTargets and Therapy. 2017 Mar; Volume 10:1645–1653. doi: 10.2147/ott.s130910.

Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A web-based graphical user interface for CHARMM. Journal of Computational Chemistry. 2008 Mar; 29(11):1859–1865. doi: 10.1002/jcc.20945.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics. 1983 Jul; 79(2):926–935. doi: 10.1063/1.445869.

Katritch V, Fenalti G, Abola EE, Roth BL, Cherezov V, Stevens RC. Allosteric sodium in class A GPCR signaling. Trends in Biochemical Sciences. 2014; 39(5):233–244. doi: 10.1016/j.tibs.2014.03.002.

Kim K, Bansal PD, Shukla D. Cyclopamine modulates smoothened receptor activity in a binding position dependent manner. Communications Biology. 2024 Sep; 7:1207. doi: 10.1038/s42003-024-06906-y.

Kinnebrew M, Iverson EJ, Patel BB, Pusapati GV, Kong JH, Johnson KA, Luchetti G, Eckert KM, McDonald JG, Covey DF, Siebold C, Radhakrishnan A, Rohatgi R. Cholesterol accessibility at the ciliary membrane controls hedgehog signaling. eLife. 2019 Oct; 8:e50051. doi: 10.7554/elife.50051.

Kinnebrew M, Luchetti G, Sircar R, Frigui S, Viti LV, Naito T, Beckert F, Saheki Y, Siebold C, Radhakrishnan A, Rohatgi R. Patched 1 reduces the accessibility of cholesterol in the outer leaflet of membranes. eLife. 2021 Oct; 10:e70504. doi: 10.7554/elife.70504.

Kinnebrew M, Woolley RE, Ansell TB, Byrne EFX, Frigui S, Luchetti G, Sircar R, Nachtergaele S, Mydock-McGrane L, Krishnan K, Newstead S, Sansom MSP, Covey DF, Siebold C, Rohatgi R. Patched 1 regulates Smoothened by controlling sterol binding to its extracellular cysteine-rich domain. Science Advances. 2022 Jun; 8(22):22. doi: 10.1126/sciadv.abm5563.

Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD, Pastor RW. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. The Journal of Physical Chemistry B. 2010 May; 114(23):7830–7843. doi: 10.1021/jp101759q.

Kleiman DE, Nadeem H, Shukla D. Adaptive Sampling Methods for Molecular Dynamics in the Era of Machine Learning. The Journal of Physical Chemistry B. 2023 Dec; 127(50):10669–10681. doi: 10.1021/acs.jpcb.3c04843.

Kleiman DE, Shukla D. Multiagent Reinforcement Learning-Based Adaptive Sampling for Conformational Dynamics of Proteins. Journal of Chemical Theory and Computation. 2022 Aug; 18(9):5422–5434. doi: 10.1021/acs.jctc.2c00683.

Kleiman DE, Shukla D. Active Learning of the Conformational Ensemble of Proteins Using Maximum Entropy VAMPNets. Journal of Chemical Theory and Computation. 2023 Apr; 19(14):4377–4388. doi: 10.1021/acs.jctc.3c00040.

Kohlhoff KJ, Shukla D, Lawrenz M, Bowman GR, Konerding DE, Belov D, Altman RB, Pande VS. Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways. Nature Chemistry. 2013 Dec; 6(1):15–21. doi: 10.1038/nchem.1821.

Körner A, Bazzone A, Wichert M, Barthmes M, Dondapati SK, Fertig N, Kubick S. Unraveling the kinetics and pharmacology of human PepT1 using solid supported membrane-based electrophysiology. Bioelectrochemistry. 2024 Feb; 155:108573. doi: 10.1016/j.bioelechem.2023.108573.

Kumar GA, Chattopadhyay A. Statin-Induced Chronic Cholesterol Depletion Switches GPCR Endocytosis and Trafficking: Insights from the Serotonin Receptor. ACS Chemical Neuroscience. 2019 Dec; 11(3):453–465. doi: 10.1021/acschemneuro.9b00659.

Kumar GA, Sarkar P, Stepniewski TM, Jafurulla M, Singh SP, Selent J, Chattopadhyay A. A molecular sensor for cholesterol in the human serotonin receptor. Science Advances. 2021 Jul; 7(30):30. doi: 10.1126/sciadv.abh2922.

Kumari S, Mitra A, Bulusu G. Putative Role of Cholesterol in Shaping the Structural and Functional Dynamics of Smoothened (SMO). The Journal of Physical Chemistry B. 2023 Oct; 127(44):9476–9495. doi: 10.1021/acs.jpcb.3c02255.

Lee J, Patel DS, Ståhle 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. CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. Journal of Chemical Theory and Computation. 2018 Dec; 15(1):775–786. doi: 10.1021/acs.jctc.8b01066.

Luchetti G, Sircar R, Kong JH, Nachtergaele S, Sagner A, Byrne EF, Covey DF, Siebold C, Rohatgi R. Cholesterol activates the G-protein coupled receptor Smoothened to promote Hedgehog signaling. eLife. 2016 Oct; 5:e20304. doi: 10.7554/elife.20304.

McComas SE, Reichenbach T, Mitrovic D, Alleva C, Bonaccorsi M, Delemotte L, Drew D. Determinants of sugar-induced influx in the mammalian fructose transporter GLUT5. eLife. 2023 jul; 12:e84808. http://dx.doi.org/10.7554/eLife.84808, doi: 10.7554/elife.84808.

McGibbon RT, Beauchamp KA, Harrigan MP, Klein C, Swails JM, Hernández CX, Schwantes CR, Wang LP, Lane TJ, Pande VS. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. Biophysical Journal. 2015; 109(8):1528 – 1532. doi: 10.1016/j.bpj.2015.08.015.

Meani RE, Lim SW, Chang ALS, Kelly JW. Emergence of chemoresistance in a metastatic basal cell carcinoma patient after complete response to hedgehog pathway inhibitor vismodegib (GDC-0449). Australasian Journal of Dermatology. 2014 Aug; 55(3):218–221. doi: 10.1111/ajd.12196.

Meurer F, Do HT, Sadowski G, Held C. Standard Gibbs energy of metabolic reactions: II. Glucose-6-phosphatase reaction and ATP hydrolysis. Biophysical Chemistry. 2017 Apr; 223:30–38. doi: 10.1016/j.bpc.2017.02.005.

Myers BR, Neahring L, Zhang Y, Roberts KJ, Beachy PA. Rapid, direct activity assays for Smoothened reveal Hedgehog pathway regulation by membrane cholesterol and extracellular sodium. Proceedings of the National Academy of Sciences USA. 2017 Dec; 114(52). doi: 10.1073/pnas.1717891115.

Myers BR, Sever N, Chong YC, Kim J, Belani JD, Rychnovsky S, Bazan JF, Beachy PA. Hedgehog Pathway Modulation by Multiple Lipid Binding Sites on the Smoothened Effector of Signal Response. Developmental Cell. 2013 Aug; 26(4):346–357. doi: 10.1016/j.devcel.2013.07.015.

Nachtergaele S, Mydock LK, Krishnan K, Rammohan J, Schlesinger PH, Covey DF, Rohatgi R. Oxysterols are allosteric activators of the oncoprotein Smoothened. Nature Chemical Biology. 2012 Jan; 8(2):211–220. doi: 10.1038/nchembio.765.

Nachtergaele S, Whalen DM, Mydock LK, Zhao Z, Malinauskas T, Krishnan K, Ingham PW, Covey DF, Siebold C, Rohatgi R. Structure and function of the Smoothened extracellular domain in vertebrate Hedgehog signaling. eLife. 2013 Oct; 2. doi: 10.7554/elife.01340.

Nadeem H, Shukla D. Ensemble Adaptive Sampling Scheme: Identifying an Optimal Sampling Strategy via Policy Ranking. Journal of Chemical Theory and Computation. 2025 Apr; 21(9):4626–4639. doi: 10.1021/acs.jctc.4c01488.

Nedelcu D, Liu J, Xu Y, Jao C, Salic A. Oxysterol binding to the extracellular domain of Smoothened in Hedgehog signaling. Nature Chemical Biology. 2013 Jul; 9(9):557–564. doi: 10.1038/nchembio.1290.

Noé F, Rosta E. Markov Models of Molecular Kinetics. The Journal of Chemical Physics. 2019 Nov; 151(19):190401. doi: 10.1063/1.5134029.

Nüsslein-Volhard C, Wieschaus E. Mutations affecting segment number and polarity in Drosophila. Nature. 1980 Oct; 287(5785):795–801. doi: 10.1038/287795a0.

Pande VS, Beauchamp K, Bowman GR. Everything you wanted to know about Markov State Models but were afraid to ask. Methods. 2010 Sep; 52(1):99–105. doi: 10.1016/j.ymeth.2010.06.002.

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry. 2005; 26(16):1781–1802. doi: 10.1002/jcc.20289.

Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, Buch R, Fiorin G, Hénin J, Jiang W, McGreevy R, Melo MCR, Radak BK, Skeel RD, Singharoy A, Wang Y, Roux B, Aksimentiev A, Luthey-Schulten Z, Kalé LV, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. Journal of Chemical Physics. 2020 Jul; 153(4):044130. doi: 10.1063/5.0014475.

Prinz JH, Wu H, Sarich M, Keller B, Senne M, Held M, Chodera JD, Schütte C, Noé F. Markov models of molecular kinetics: Generation and validation. The Journal of Chemical Physics. 2011 May; 134(17):174105. doi: 10.1063/1.3565032.

Qi C, Di Minin G, Vercellino I, Wutz A, Korkhov VM. Structural basis of sterol recognition by human hedgehog receptor PTCH1. Science Advances. 2019 Sep; 5(9). doi: 10.1126/sciadv.aaw6490.

Qi X, Friedberg L, Bose-Boyd RD, Long T, Li X. Sterols in an intramolecular channel of Smoothened mediate Hedgehog signaling. Nature Chemical Biology. 2020 Sep; 16(12):1368–1375. doi: 10.1038/s41589-020-0646-2.

Qi X, Liu H, Thompson B, McDonald J, Zhang C, Li X. Cryo-EM structure of oxysterol-bound human Smoothened coupled to a heterotrimeric Gi. Nature. 2019 Jun; 571(7764):279–283. doi: 10.1038/s41586-019-1286-0.

Qi X, Schmiege P, Coutavas E, Li X. Two Patched molecules engage distinct sites on Hedgehog yielding a signaling-competent complex. Science. 2018 Oct; 362(6410):6410. doi: 10.1126/science.aas8843.

Radhakrishnan A, Rohatgi R, Siebold C. Cholesterol access in cellular membranes controls Hedgehog signaling. Nature Chemical Biology. 2020 Nov; 16(12):1303–1313. doi: 10.1038/s41589-020-00678-2.

Raleigh DR, Reiter JF. Misactivation of Hedgehog signaling causes inherited and sporadic cancers. Journal of Clinical Investigation. 2019 Feb; 129(2):465–475. doi: 10.1172/jci120850.

Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. Journal of Chemical Theory and Computation. 2013 Jun; 9(7):3084–3095. doi: 10.1021/ct400341p.

Rohatgi R, Milenkovic L, Scott MP. Patched1 Regulates Hedgehog Signaling at the Primary Cilium. Science. 2007 Jul; 317(5836):372–376. doi: 10.1126/science.1139740.

Rudolf AF, Kinnebrew M, Kowatsch C, Ansell TB, El Omari K, Bishop B, Pardon E, Schwab RA, Malinauskas T, Qian M, Duman R, Covey DF, Steyaert J, Wagner A, Sansom MSP, Rohatgi R, Siebold C. The morphogen Sonic hedgehog inhibits its receptor Patched by a pincer grasp mechanism. Nature Chemical Biology. 2019 Sep; 15(10):975–982. doi: 10.1038/s41589-019-0370-y.

Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. Wiley Interdisciplinary Reviews: Computational Molecular Science. 2012 Sep; 3(2):198–210. doi: 10.1002/wcms.1121.

Salomon-Ferrer R, Götz AW, Poole D, Grand SL, Walker RC. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. Journal of Chemical Theory and Computation. 2013 Aug; 9(9):3878–3888. doi: 10.1021/ct400314y.

Scandroglio F, Venkata JK, Loberto N, Prioni S, Schuchman EH, Chigorno V, Prinetti A, Sonnino S. Lipid content of brain, brain membrane lipid domains, and neurons from acid sphingomyelinase deficient mice. Journal of Neurochemistry. 2008 Jul; 107(2):329–338. doi: 10.1111/j.1471-4159.2008.05591.x.

Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, Wehmeyer C, Prinz JH, Noé F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. Journal of Chemical Theory and Computation. 2015 Oct; 11(11):5525–5542. doi: 10.1021/acs.jctc.5b00743.

Schroedinger, LLC. The PyMOL Molecular Graphics System. Schrödinger, LLC; 2022, version 1.8. Available from: https://pymol.org/2/.

Schwantes CR, Pande VS. Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. Journal of Chemical Theory and Computation. 2013 Mar; 9(4):2000–2009. doi: 10.1021/ct300878a.

Selvam B, Shamsi Z, Shukla D. Universality of the Sodium Ion Binding Mechanism in Class A G-Protein-Coupled Receptors. Angewandte Chemie. 2018 Feb; 57(12):3048–3053. doi: 10.1002/anie.201708889.

Selvam B, Yu YC, Chen LQ, Shukla D. Molecular Basis of the Glucose Transport Mechanism in Plants. ACS Cent Sci. 2019 May; 5(6):1085–1096. doi: 10.1021/acscentsci.9b00252.

Serdiuk T, Manna M, Zhang C, Mari SA, Kulig W, Pluhackova K, Kobilka BK, Vattulainen I, Müller DJ. A cholesterol analog stabilizes the human β_2-adrenergic receptor nonlinearly with temperature. Science Signaling. 2022 Jun; 15(737):737. doi: 10.1126/scisignal.abi7031.

Shamsi Z, Cheng KJ, Shukla D. Reinforcement Learning Based Adaptive Sampling: REAPing Rewards by Exploring Protein Conformational Landscapes. The Journal of Physical Chemistry B. 2018 Aug; 122(35):8386–8395. doi: 10.1021/acs.jpcb.8b06521.

Shamsi Z, Moffett AS, Shukla D. Enhanced unbiased sampling of protein dynamics using evolutionary coupling information. Scientific Reports. 2017 Oct; 7(1):12700. doi: 10.1038/s41598-017-12874-7.

Shukla D, Hernández CX, Weber JK, Pande VS. Markov State Models Provide Insights into Dynamic Modulation of Protein Function. Accounts of Chemical Research. 2015 Jan; 48(2):414–422. doi: 10.1021/ar5002999.

Siebold C, Rohatgi R. The Inseparable Relationship Between Cholesterol and Hedgehog Signaling. Annual Review of Biochemistry. 2023 Jun; 92(1):273–298. http://dx.doi.org/10.1146/annurev-biochem-052521-040313, doi: 10.1146/annurev-biochem-052521-040313.

Smart OS, Goodfellow JM, Wallace BA. The pore dimensions of gramicidin A. Biophysical Journal. 1993 Dec; 65(6):2455–2460. doi: 10.1016/s0006-3495(93)81293-1.

Smirnova I, Kasho V, Kaback HR. Real-time conformational changes in LacY. Proceedings of the National Academy of Sciences. 2014 May; 111(23):8440–8445. doi: 10.1073/pnas.1408374111.

Stone J. An Efficient Library for Parallel Ray Tracing and Animation. Master's thesis, Computer Science Department, University of Missouri-Rolla; 1998.

Stottmann RW, Turbe-Doan A, Tran P, Kratz LE, Moran JL, Kelley RI, Beier DR. Cholesterol Metabolism Is Required for Intracellular Hedgehog Signal Transduction In Vivo. PLoS Genetics. 2011 Sep; 7(9):e1002224. doi: 10.1371/journal.pgen.1002224.

Taipale J, Cooper MK, Maiti T, Beachy PA. Patched acts catalytically to suppress the activity of Smoothened. Nature. 2002 Aug; 418(6900):892–896. doi: 10.1038/nature00989.

Tarling EJ, Vallim TQdA, Edwards PA. Role of ABC transporters in lipid transport and human disease. Trends in Endocrinology & Metabolism. 2013 Jul; 24(7):342–350. doi: 10.1016/j.tem.2013.01.006.

Thangapandian S, Fakharzadeh A, Moradi M, Tajkhorshid E. Conformational free energy landscape of a glutamate transporter and microscopic details of its transport mechanism. Proceedings of the National Academy of Sciences. 2025 Mar; 122(10):e2416381122. http://dx.doi.org/10.1073/pnas.2416381122, doi: 10.1073/pnas.2416381122.

Wang C, Wu H, Evron T, Vardy E, Han GW, Huang XP, Hufeisen SJ, Mangano TJ, Urban DJ, Katritch V, Cherezov V, Caron MG, Roth BL, Stevens RC. Structural basis for Smoothened receptor modulation and chemoresistance to anticancer drugs. Nature Communications. 2014 Jul; 5(1):4355. doi: 10.1038/ncomms5355.

Wang C, Wu H, Katritch V, Han GW, Huang XP, Liu W, Siu FY, Roth BL, Cherezov V, Stevens RC. Structure of the human smoothened receptor bound to an antitumour agent. Nature. 2013 May; 497(7449):338–343. doi: 10.1038/nature12167.

Wang W, Cao S, Zhu L, Huang X. Constructing Markov State Models to elucidate the functional conformational changes of complex biomolecules. WIREs Computational Molecular Science. 2017 Oct; 8(1):e1343. doi: 10.1002/wcms.1343.

Waskom ML. seaborn: statistical data visualization. Journal of Open Source Software. 2021; 6(60):3021. doi: 10.21105/joss.03021.

Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. Current Protocols in Bioinformatics. 2016 Jun; 54(1):5–6. doi: 10.1002/cpbi.3.

Weber JK, Pande VS. Characterization and Rapid Sampling of Protein Folding Markov State Model Topologies. Journal of Chemical Theory and Computation. 2011 Sep; 7(10):3405–3411. doi: 10.1021/ct2004484.

Weierstall U, James D, Wang C, White TA, Wang D, Liu W, Spence JCH, Doak RB, Nelson G, Fromme P, Fromme R, Grotjohann I, Kupitz C, Zatsepin NA, Liu H, Basu S, Wacker D, Han GW, Katritch V, Boutet S, et al. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. Nature Communications. 2014 Feb; 5(1):3309. doi: 10.1038/ncomms4309.

Wu H, Noé F. Variational Approach for Learning Markov Processes from Time Series Data. Journal of Nonlinear Science. 2019 Aug; 30(1):23–66. doi: 10.1007/s00332-019-09567-y.

Xiao X, Tang JJ, Peng C, Wang Y, Fu L, Qiu ZP, Xiong Y, Yang LF, Cui HW, He XL, Yin L, Qi W, Wong CCL, Zhao Y, Li BL, Qiu WW, Song BL. Cholesterol Modification of Smoothened Is Required for Hedgehog Signaling. Molecular Cell. 2017 Apr; 66(1):154–162.e10. http://dx.doi.org/10.1016/j.molcel.2017.02.015, doi: 10.1016/j.molcel.2017.02.015.

Zhang K, Wu H, Hoppe N, Manglik A, Cheng Y. Fusion protein strategies for cryo-EM study of G protein-coupled receptors. Nature Communications. 2022 Jul; 13(1):4366. http://dx.doi.org/10.1038/s41467-022-32125-2, doi: 10.1038/s41467-022-32125-2.

Zhang X, Zhao F, Wu Y, Yang J, Han GW, Zhao S, Ishchenko A, Ye L, Lin X, Ding K, Dharmarajan V, Griffin PR, Gati C, Nelson G, Hunter MS, Hanson MA, Cherezov V, Stevens RC, Tan W, Tao H, et al. Crystal structure of a multidomain human smoothened receptor in complex with a super stabilizing ligand. Nature Communications. 2017 May; 8(1):15383. doi: 10.1038/ncomms15383.

Zhang Y, Bulkley DP, Xin Y, Roberts KJ, Asarnow DE, Sharma A, Myers BR, Cho W, Cheng Y, Beachy PA. Structural Basis for Cholesterol Transport-like Activity of the Hedgehog Receptor Patched. Cell. 2018 Nov; 175(5):1352–1364.e14. doi: 10.1016/j.cell.2018.10.026.

Zhang Y, Lu WJ, Bulkley DP, Liang J, Ralko A, Han S, Roberts KJ, Li A, Cho W, Cheng Y, Manglik A, Beachy PA. Hedgehog pathway activation through nanobody-mediated conformational blockade of the Patched sterol conduit. Proceedings of the National Academy of Sciences USA. 2020 Nov; 117(46):28838–28846. doi: 10.1073/pnas.2011560117.

Zhu Y, He L, Liu Y, Zhao Y, Zhang XC. smFRET Probing Reveals Substrate-Dependent Conformational Dynamics of E. coli Multidrug MdfA. Biophysical Journal. 2019 Jun; 116(12):2296–2303. doi: 10.1016/j.bpj.2019.04.034.

Zimmerman MI, Bowman GR. FAST Conformational Searches by Balancing Exploration/Exploitation Trade-Offs. Journal of Chemical Theory and Computation. 2015 Nov; 11(12):5747–5757. doi: 10.1021/acs.jctc.5b00737.

Appendix 1

Modelled residues in SMO

Appendix 1—table 1. Modelled residues in 5L7D-inactive-Apo-SMO starting structure. The helical content of K440–I445 was modelled based on the structure of SANT1-bound SMO (PDB: 4N4W) (***Wang et al., 2014***).

Modelled Residues in 5L7D-inac-Apo-SMO	Constraints	Location
I429	None	ICL3
K430	None	ICL3
S431	None	ICL3
N432	None	ICL3
H433	None	ICL3
P434	None	ICL3
G435	None	ICL3
L436	None	ICL3
L437	None	ICL3
S438	None	ICL3
E439	None	ICL3
K440	α-helical	TM6
A441	α-helical	TM6
A442	α-helical	TM6
S443	α-helical	TM6
K444	α-helical	TM6
I445	α-helical	TM6

Lipid Composition

Appendix 1—table 2. Composition of the membrane used for embedding the protein in simulations.

Lipid Name	Upper Leaflet	Lower Leaflet
Cholesterol (CHL1)	21	21
1-palmitoyl-2-oleoylphosphatidylcholine (POPC)	76	76
Palmitoylsphingomyelin (PSM)	4	4
Total	101	101

Distances used for trajectory featurization in Adaptive Sampling

Appendix 1—table 3. Adaptive Sampling metrics used for clustering during the iterative landscape exploration process.

W281	M525	V321	M525	W281	V321
S278	F526	L325	M525	T241	F274
D473	E518	D384	S387	D384	Y394
Y322	F391	G277	V321	W281	V321
V319	Y323	V276	I320	I316	V319
N219	P513	L221	I509	Y487	I509
S483	N511	Y207	K395	I215	M301
N219	W480	W109	Y130	W109	R161
E160	R485	I156	E160	L112	I156
N114	V210	D209	A492	E208	V488
CHL	H470	CHL	N521	CHL	F391
CHL	Y394	CHL	S387	CHL	W281
CHL	L522	CHL	N521	CHL	L325
CHL	V404	CHL	T466	CHL	I408
CHL	T528	CHL	V329	CHL	F462
CHL	V463	CHL	F274	CHL	F332
CHL	E518	CHL	D473	CHL	N521
CHL	F391	CHL	Y394	CHL	R400
CHL	N521	CHL	L522	CHL	W281
CHL	M525	CHL	P220	CHL	F484
CHL	S387	CHL	H470	CHL	D384
CHL	L515	CHL	V386	CHL	K395
CHL	N219	CHL	F222	CHL	W480
CHL	P513	CHL	L221	CHL	W109
CHL	R161	CHL	I156	CHL	D95
CHL	Y130	CHL	I496	CHL	L112
CHL	N114	CHL	K105	CHL	F166
CHL	V210	CHL	L489	CHL	V157
CHL	E160	CHL	L108	CHL	P164

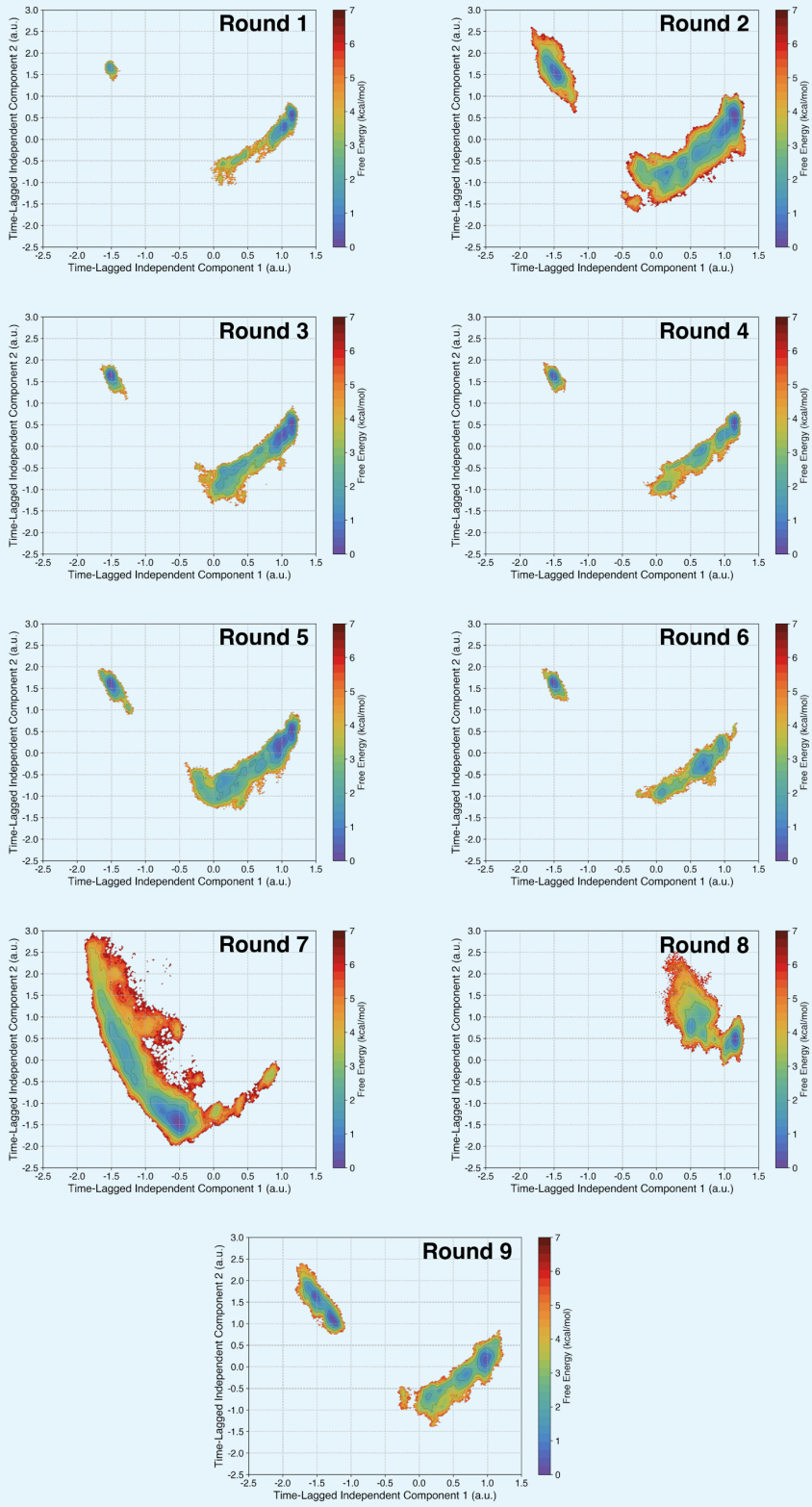
Round-wise data collection cholesterol translocation

Appendix 1—table 4. Round wise data collection for Cholesterol transport of SMO.

Simulation Round	Amount of Data
Round 1	20.38μs
Round 2	460.15μs
Round 3	71.12μs
Round 4	82.17μs
Round 5	43.26μs
Round 6	26.70μs
Round 7	508.16μs
Round 8	184.63μs
Round 9	118.69μs
Round 10	550.59 μs
Total	2066.227 μs

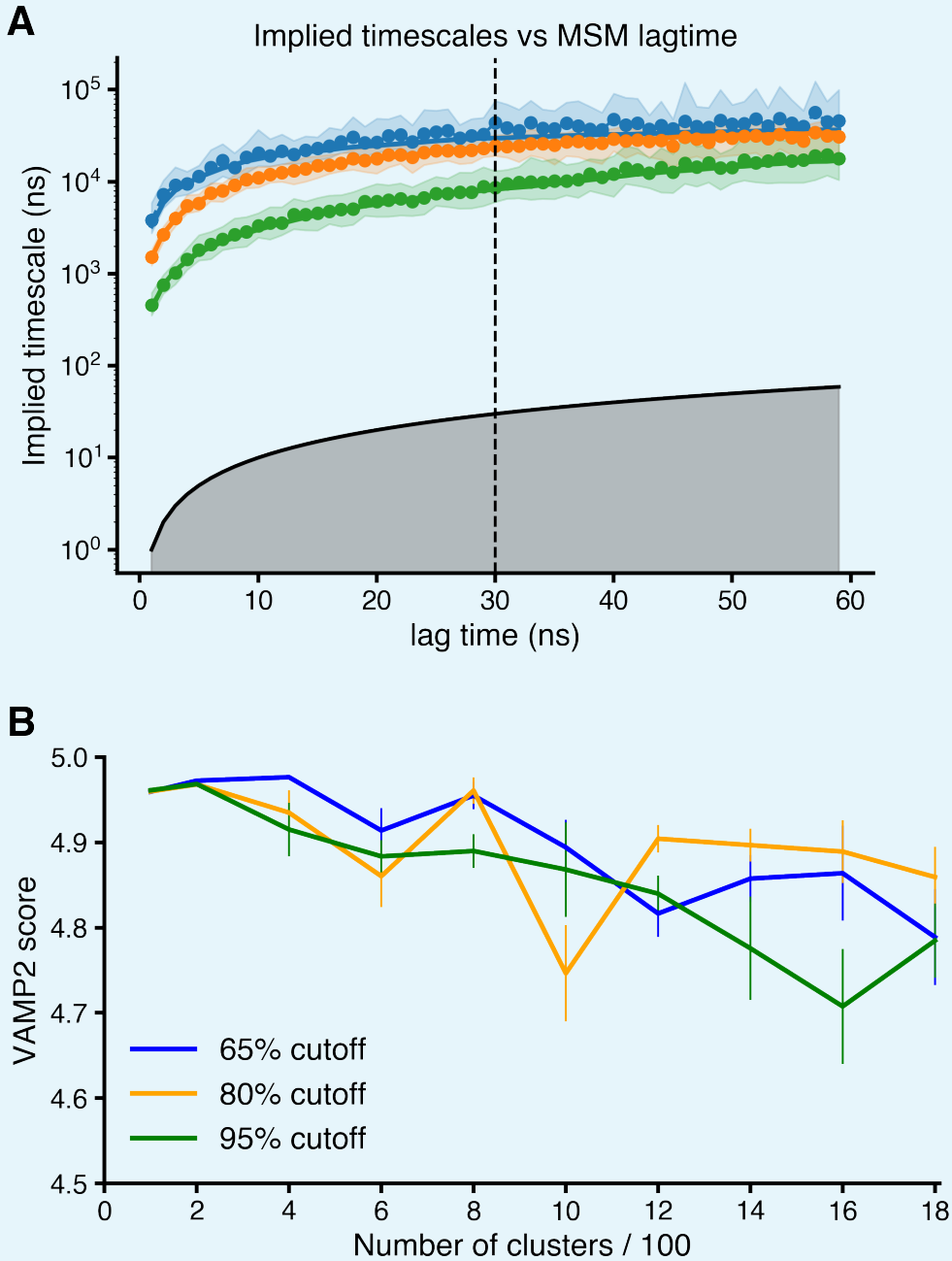
Appendix 2

Roundwise data collection for SMO

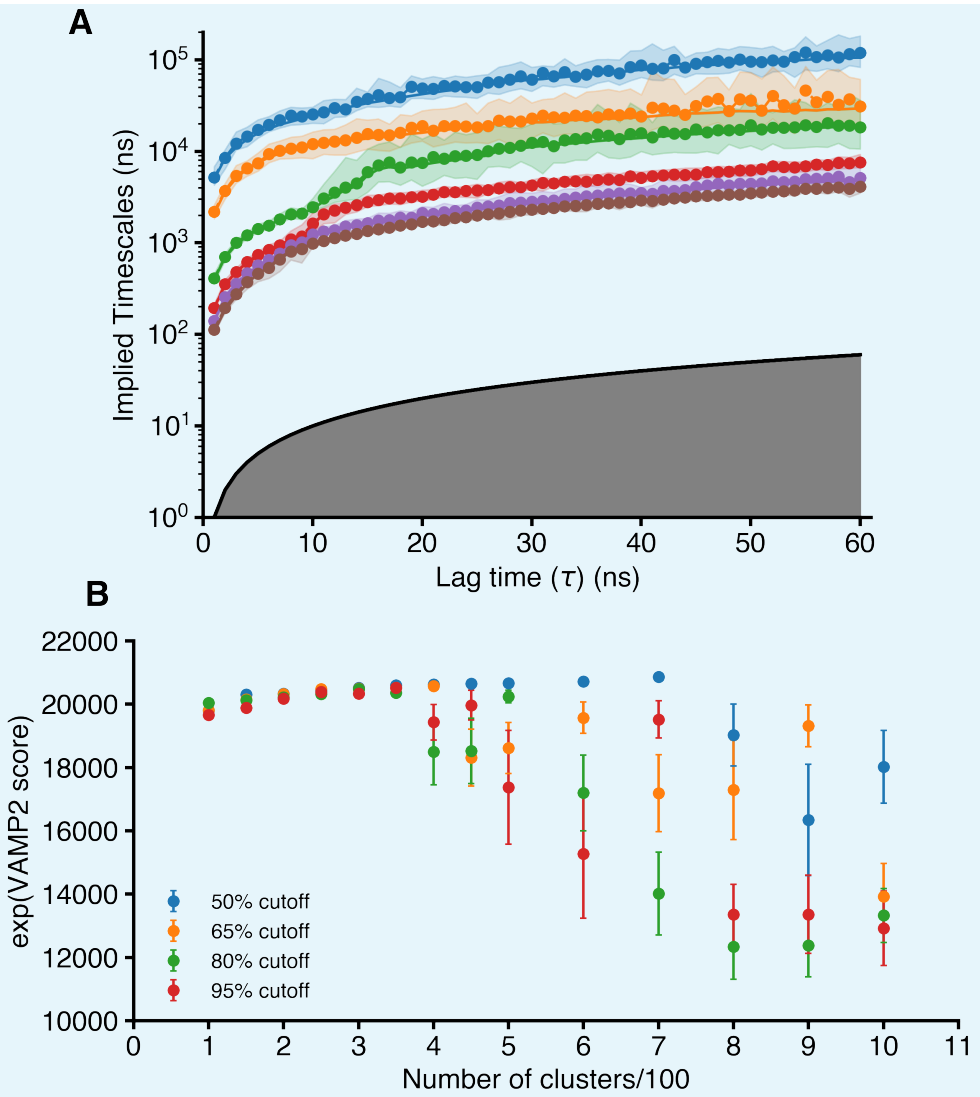


Appendix 2—figure 1. Roundwise data collection for SMO Cholesterol transport as projected along the first two time-lagged independent components (tICs). Round numbers are specified in the respective plots.

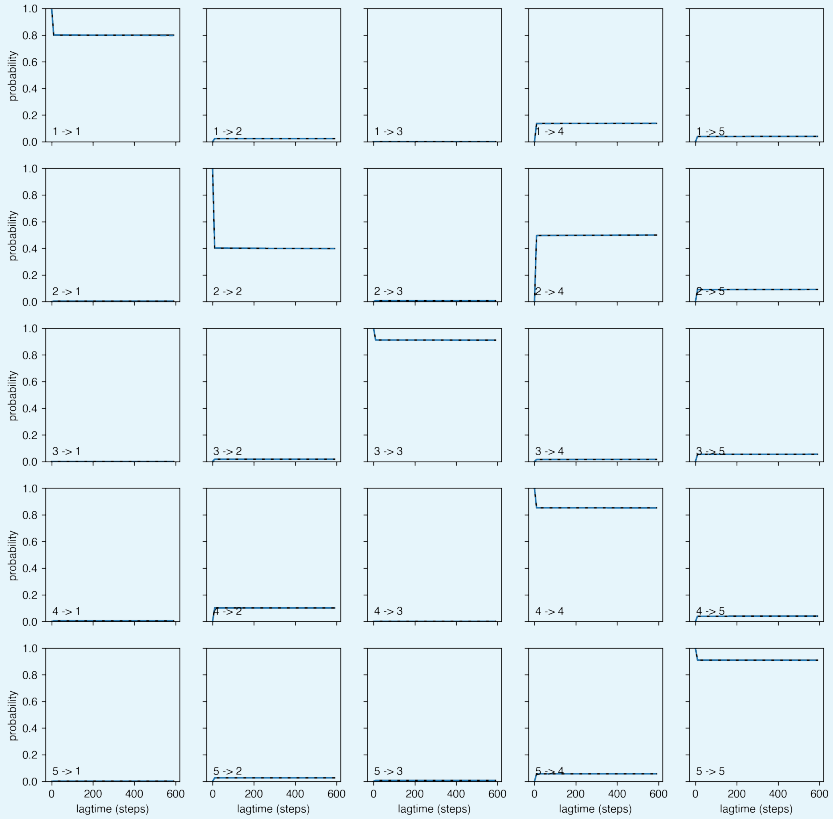
MSM Construction



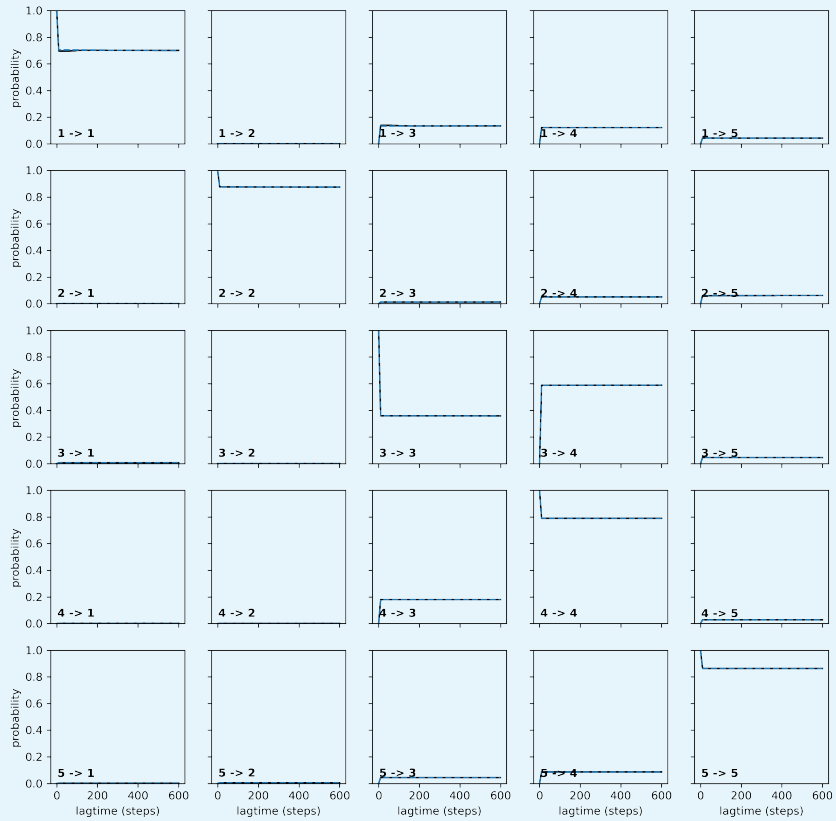
Appendix 2—figure 2. MSM Construction for SMO Cholesterol transport - Pathway 1. (A) Implied Timescales versus MSM Lagtime plot shows the convergence of timescales. A lagtime of 30ns was chosen for the MSM construction. (B) VAMP2 score versus Number of Clusters used for clustering the TICA-reduced data at three different variational cutoffs. The final MSM was made using 200 clusters and a 95 % cutoff (corresponding to 42 tIC dimensions).



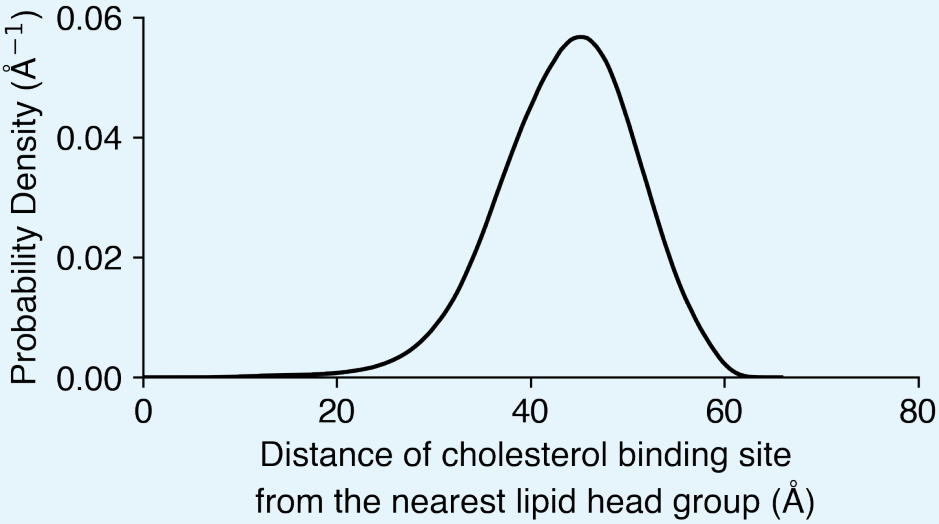
Appendix 2—figure 3. MSM Construction for SMO Cholesterol transport - Pathway 2. (A) The Implied Timescales versus MSM Lagtime plot shows the convergence of timescales. A lagtime of 30ns was chosen for the MSM construction. (B) VAMP2 score versus Number of Clusters used for clustering the TICA-reduced data at three different variational cutoffs. The final MSM was made using 400 clusters and a 65 % cutoff (corresponding to 28 tIC dimensions).



Appendix 2—figure 4. Chapman Kolmogorov Test for MSM validation - Pathway 1. Chapman Kolmogorov test was performed using the deeptime library (***Hoffmann et al., 2021***).



Appendix 2—figure 5. Chapman Kolmogorov Test for MSM validation - Pathway 2. Chapman Kolmogorov test was performed using the deeptime library (***Hoffmann et al., 2021***).



Appendix 2—figure 6. Distance of the binding site of cholesterol (marked by D95) from the nearest lipid headgroups in the membrane. Kernel density plotted for 2.1ms of simulation data. The binding site is at least 20 Å away from the nearest lipid headgroup in simulations.

1069

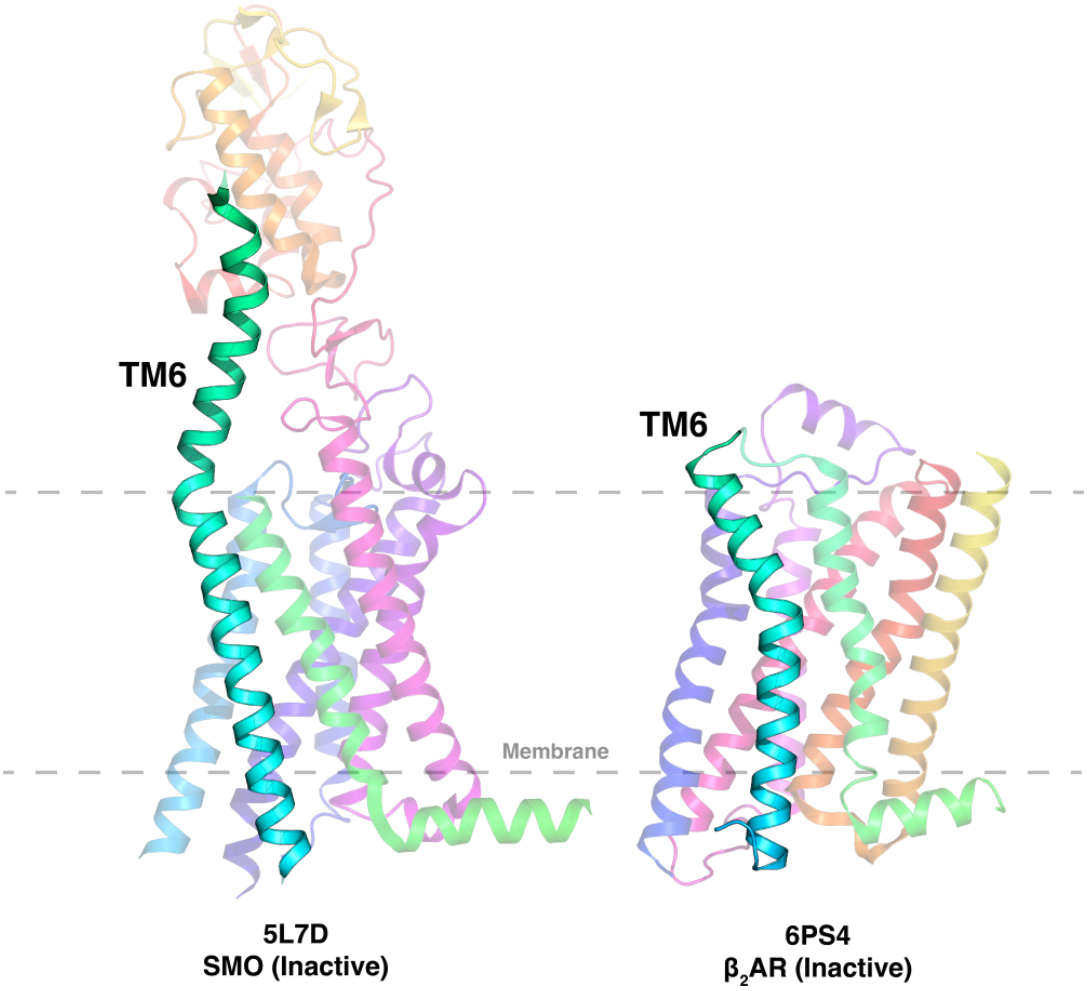


Figure 1—figure supplement 1. Comparison of TM6 length for SMO (A) and β2AR (B). SMO shows an elongated TM6, which emerges into the extracellular space.

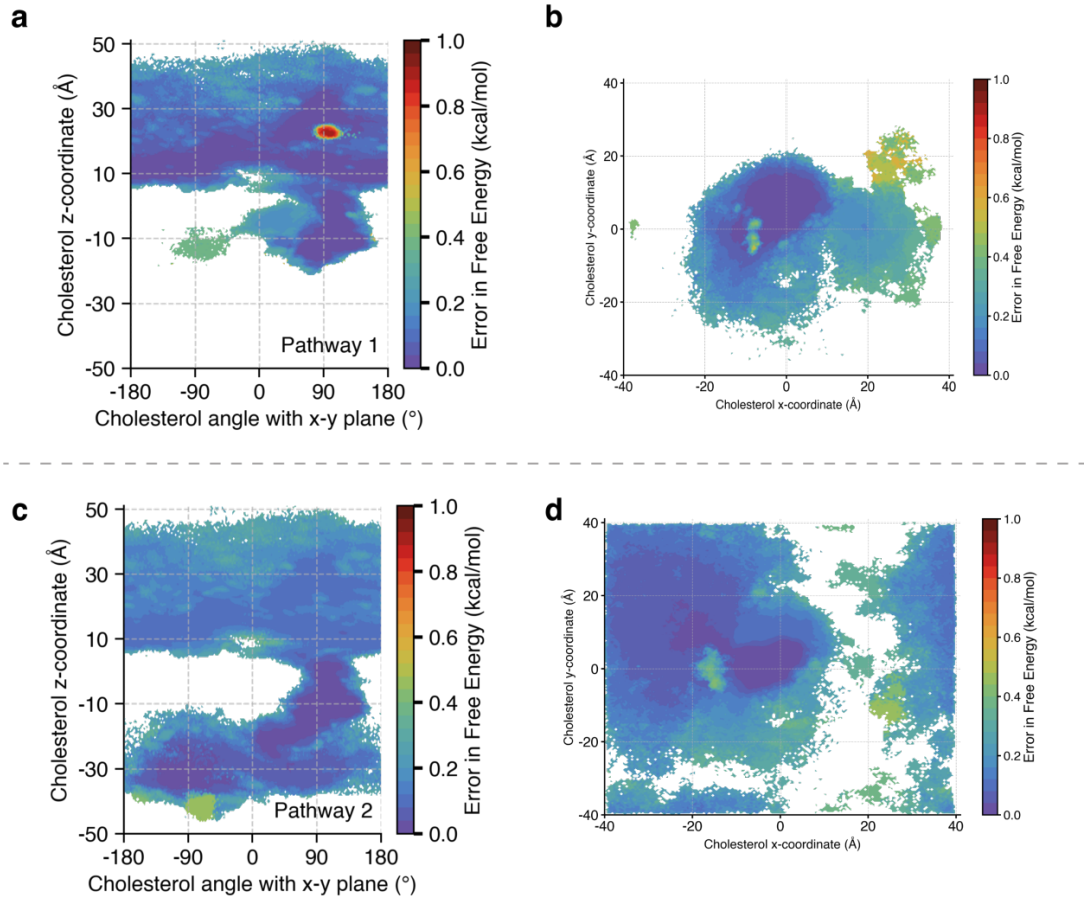


Figure 2—figure supplement 1. Error in Free Energies for Fig. 2 and Fig. 4. Errors were computed using a bootstrapping strategy, where the MSM probabilities were computed from 80% of the data 200 times. (a) Error in Fig. 2(a). (b) Error in Fig. 2(b). (c) Error in Fig. 4(a). (d) Error in Fig. 4(b)

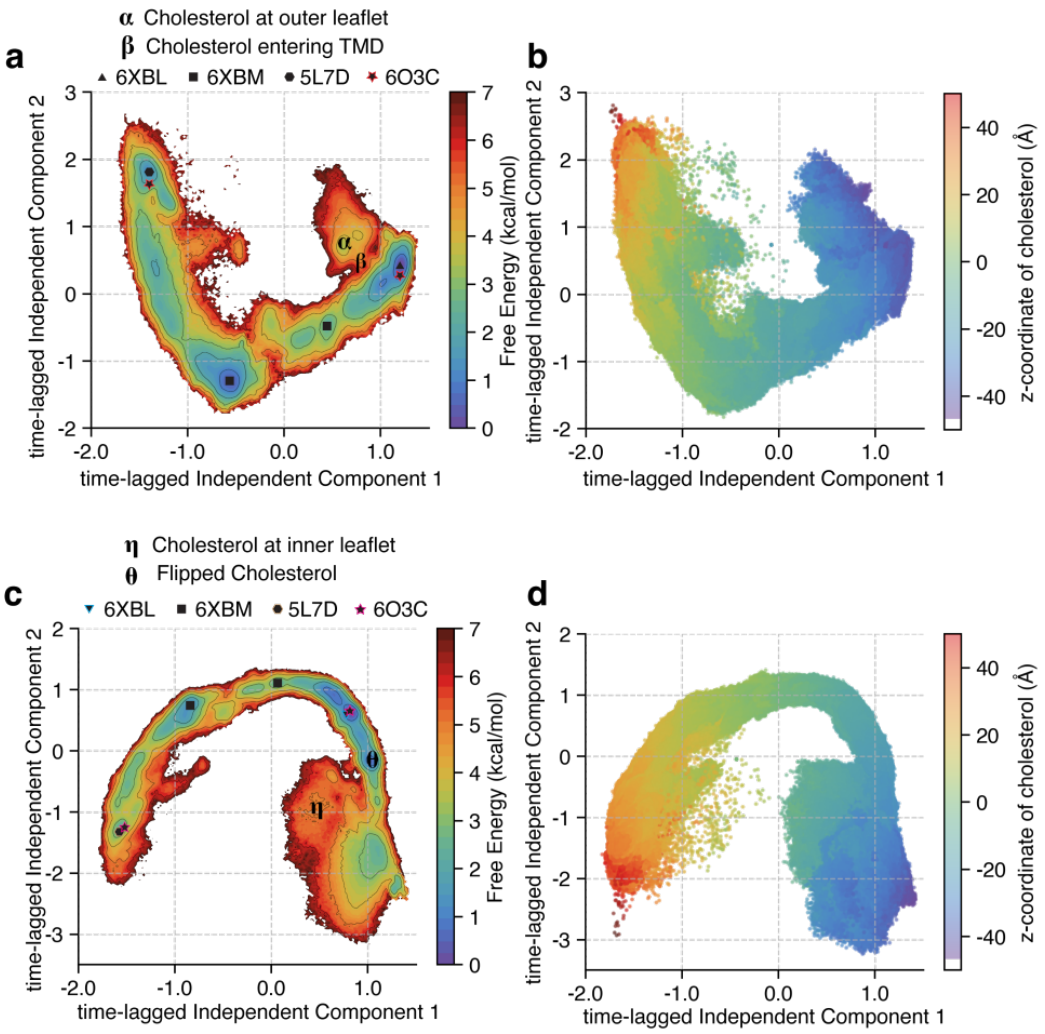


Figure 2—figure supplement 2. TICA (Time-lagged Independent Component Analysis) plot for SMO Cholesterol transport. - Pathway 1. (A) The entire data projected along the first two time-lagged independent components (tICs). Free energies were calculated for each datapoint and reweighed using the Markov State Model probabilities. (B) The same data as (A), except the z-coordinate of the cholesterol plotted against the first two tICs. The translocation of cholesterol is the slowest process observed, as shown by the gradient in the plot. (C) Same as (A), but for Pathway 2. (D) Same as (B), but for Pathway 2.

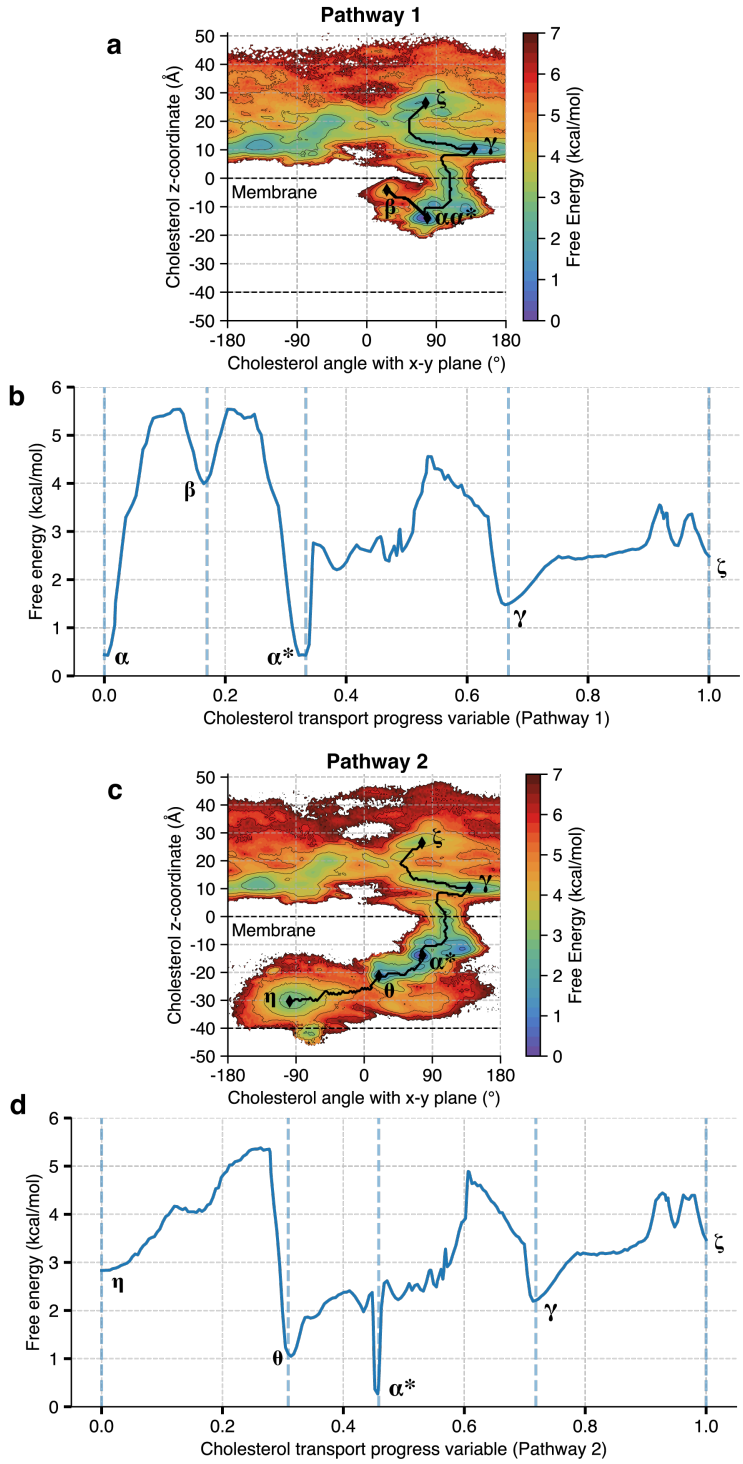


Figure 2—figure supplement 3. Minimum energy path taken by cholesterol for both pathways. (a) Minimum energy pathway for Pathway 1 (black), marked on Figure 2(a). Note that the order of transitions is $\alpha \rightarrow \beta \rightarrow \alpha^* \rightarrow \gamma \rightarrow \zeta$. (b) Free energy profile as cholesterol moves along Pathway 1. The progress is mapped by the cholesterol translocation progress variable. (c) Minimum energy pathway for Pathway 2 (black), marked on Figure 4(a). Note that the order of transitions is $\eta \rightarrow \theta \rightarrow \alpha^* \rightarrow \gamma \rightarrow \zeta$. (d) Free energy profile as cholesterol moves along Pathway 2. The progress is mapped by the cholesterol translocation progress variable.

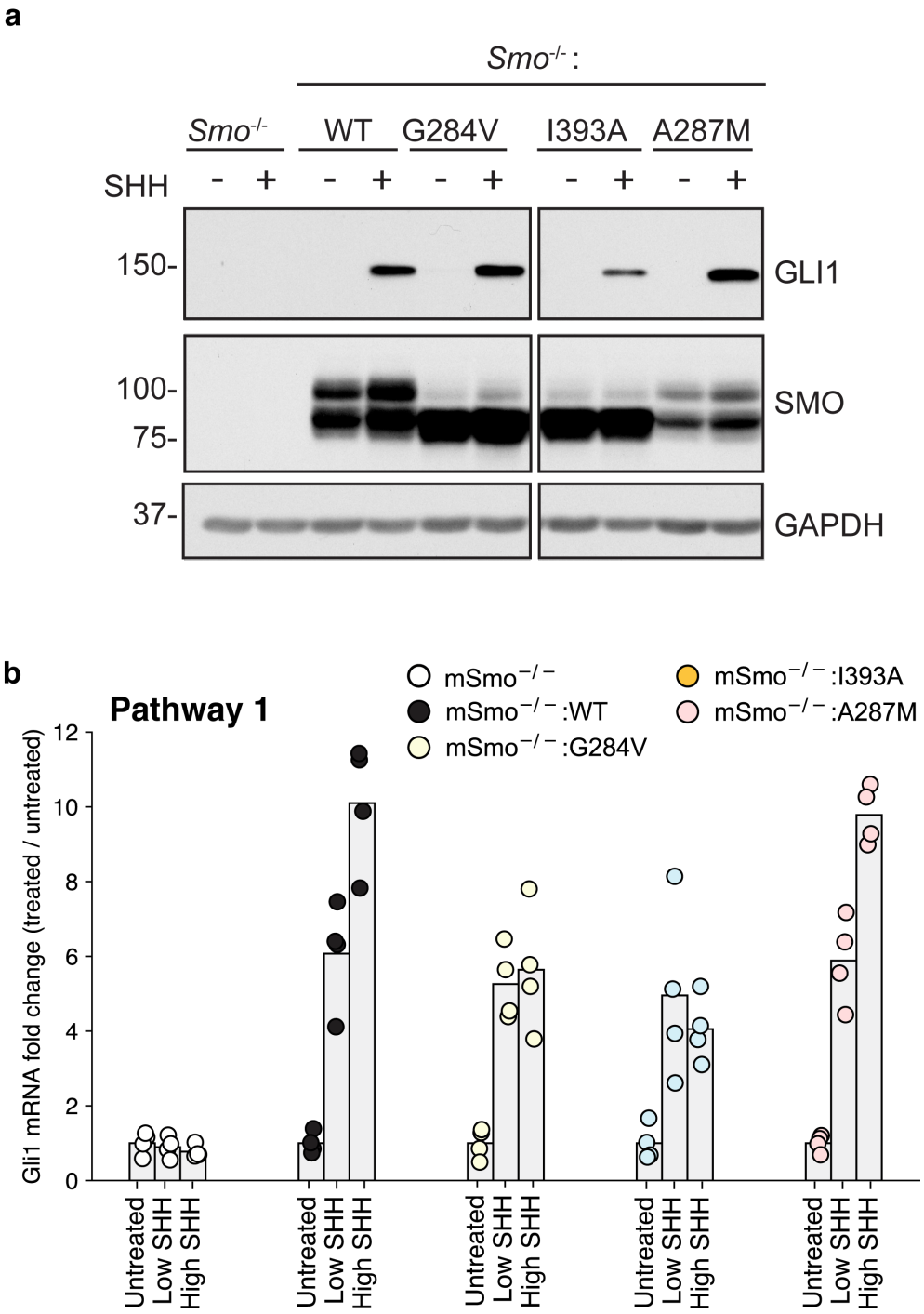


Figure 3—figure supplement 1. Experimental data for mutants along Pathway 1. (a) Immunoblotting was used to measure abundance of mSMO and GLI1 proteins in SMO$^{-/-}$ cells stably expressing either mSMO-WT or Pathway 1 mutants after treatment with SHH. (b) *Gli1* mRNA fold change plotted for SMO mutants, showing fold change when the mutants are untreated, treated with a low concentration of SHH, and treated with a saturating concentration of SHH.

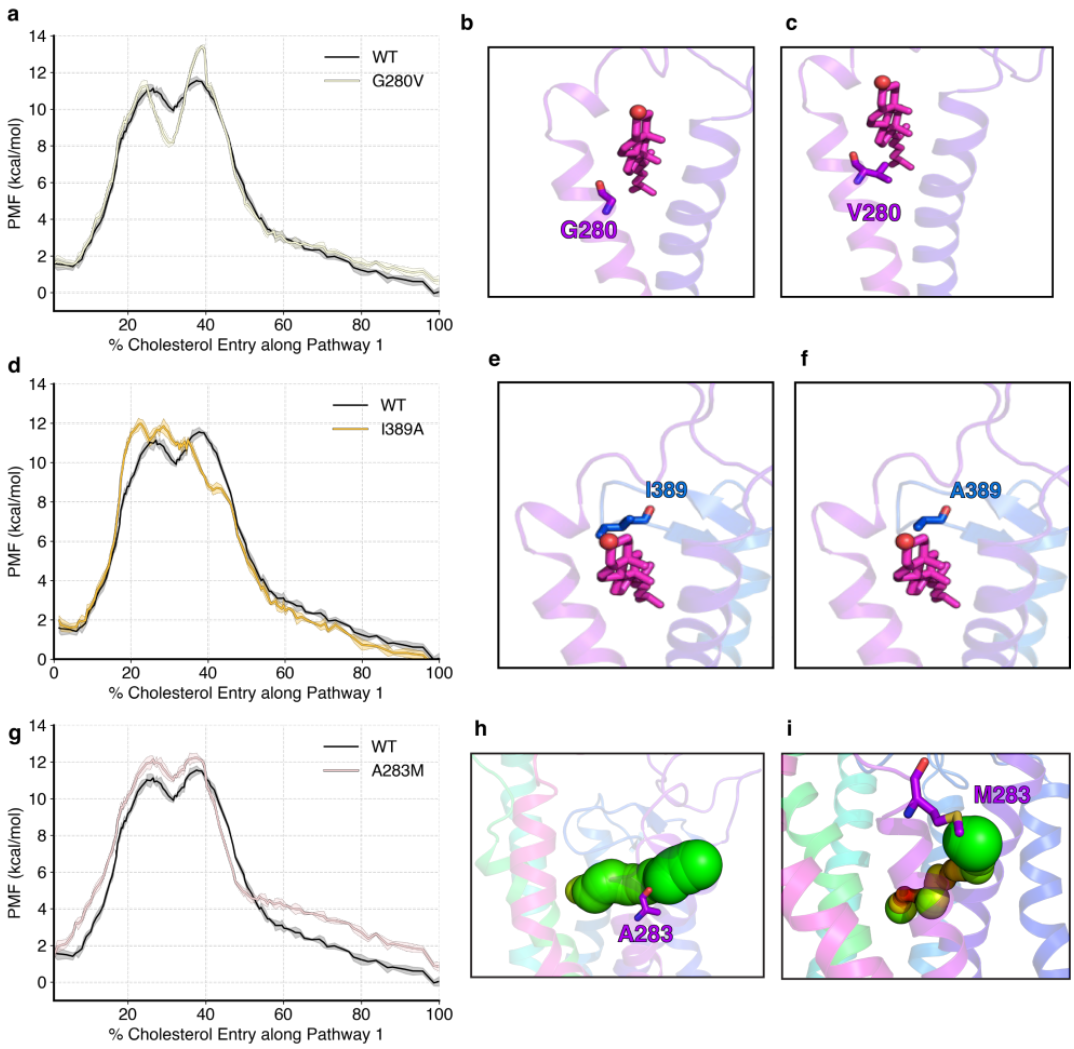


Figure 3—figure supplement 2. PMF data for mutants along Pathway 1. (a) PMF for cholesterol translocation was computed for *hSMO* G280V and compared with WT *hSMO* using adaptive biasing force-based simulations. Representative figures showing the mutation for WT G280 (b) versus V280 (c). (d, g) - Same as (a) but for I389A and A283M, respectively. (e, h) same as (b) but for I389 and A283, respectively. (f, i) - same as (c) but for A389 and M283, respectively.

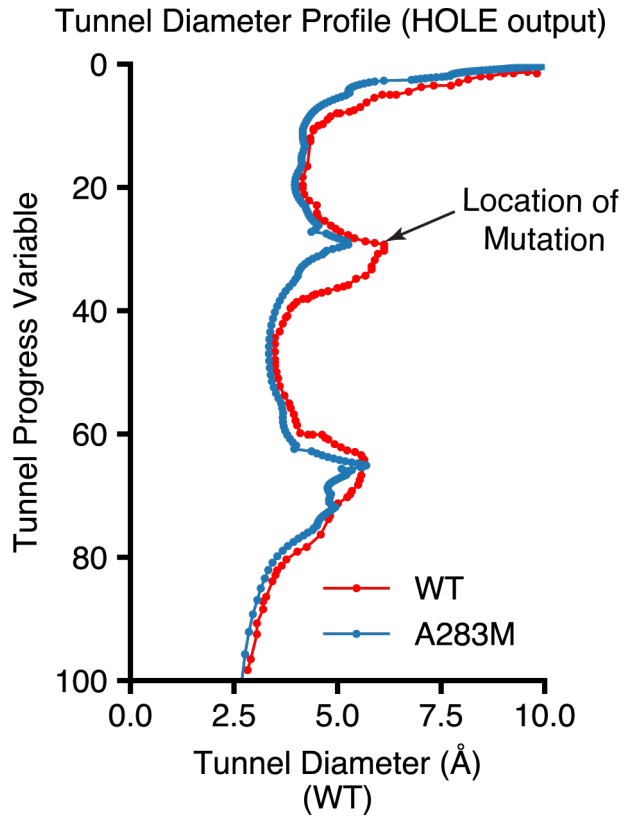


Figure 3—figure supplement 3. Tunnel diameter profile for WT versus mutant A283M.

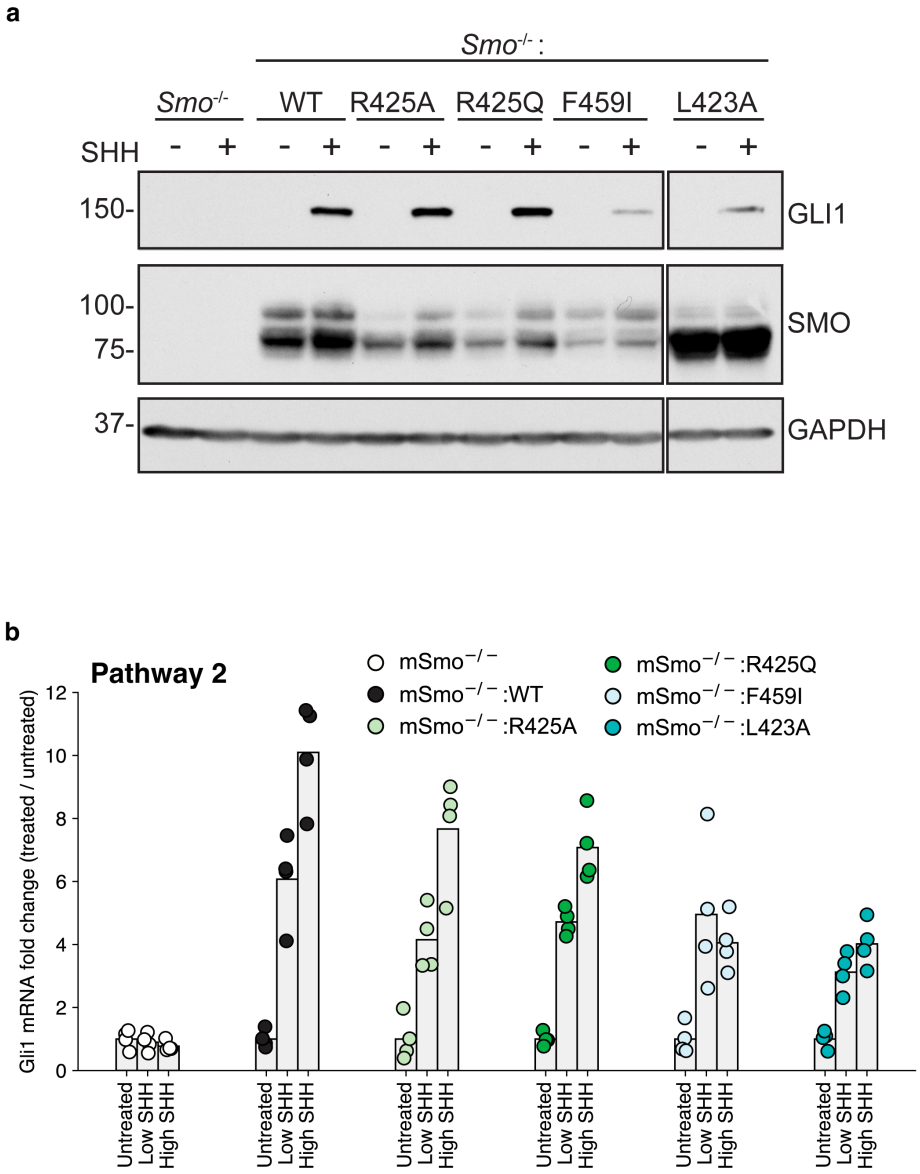


Figure 3—figure supplement 4. Experimental data for mutants along Pathway 2. (a) Immunoblotting was used to measure abundance of mSMO and GLI1 proteins in SMO$^{-/-}$ cells stably expressing either mSMO-WT or Pathway 2 mutants after treatment with SHH. (b) *Gli1* mRNA fold change plotted for SMO mutants, showing fold change when the mutants are untreated, treated with a low concentration of SHH, and treated with a saturating concentration of SHH.

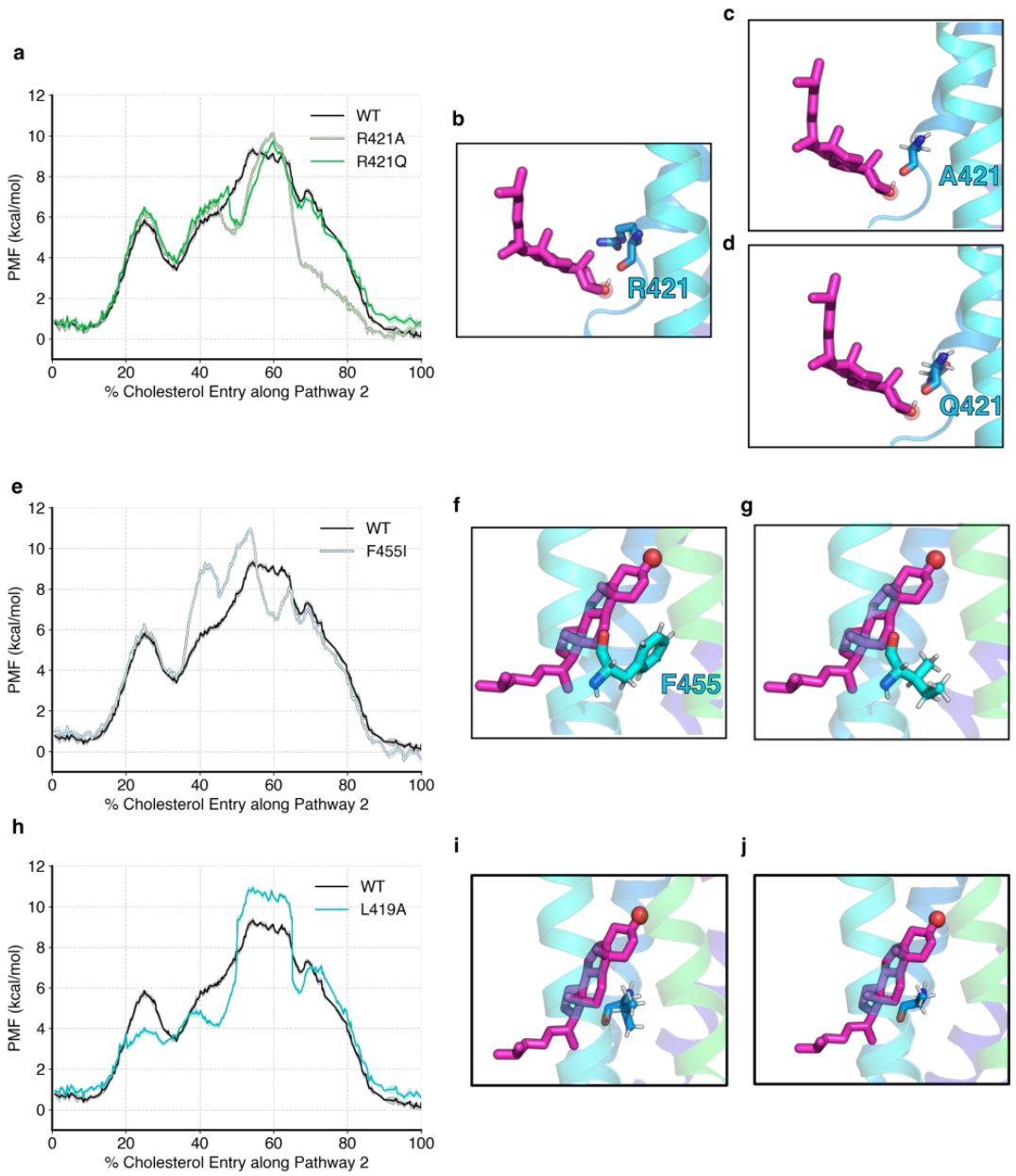


Figure 3—figure supplement 5. PMF data for mutants along Pathway 2. (a) PMF for cholesterol translocation was computed for *hSMO* R421A/R421Q and compared with WT *hSMO* using adaptive biasing force-based simulations. Representative figures showing the mutation for WT R421 (b) versus A421 (c) and Q421 (d). (e, h) - Same as (a) but for F455I and L419A, respectively. (f, i) same as (b) but for F455 and L419, respectively. (g, j) - same as (c, d) but for I455 and A419, respectively.

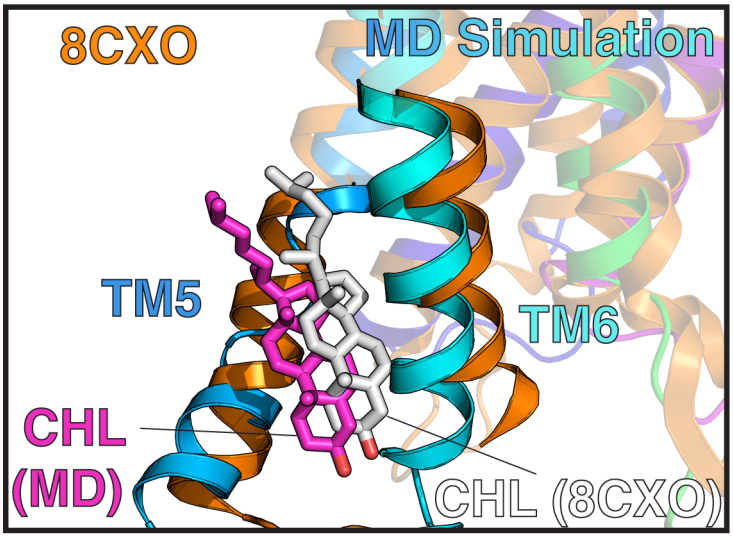


Figure 4—figure supplement 1. Comparison of cholesterol entry with existing resolved structure. Frame from MD simulation showing cholesterol entry (cyan-blue), and resolved structure (PDB 8CXO, ***Zhang et al.*** (***2022***), orange). Cholesterol from the MD simulation frame is magenta, cholesterol from the resolved structure is white.

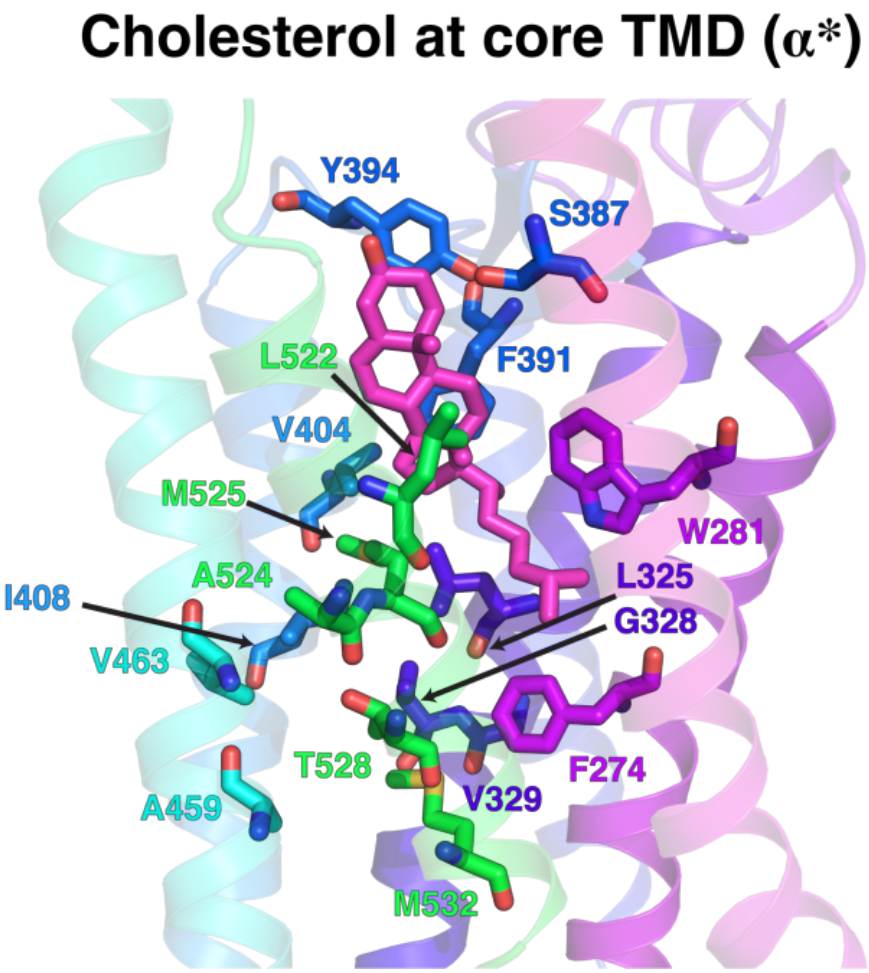


Figure 5—figure supplement 1. Hydrophobic Tunnel inside SMO core TMD. Cholesterol interacts with the hydrophobic core of the TMD along the transport pathway. Figure corresponds to pose δ in Figure 4.

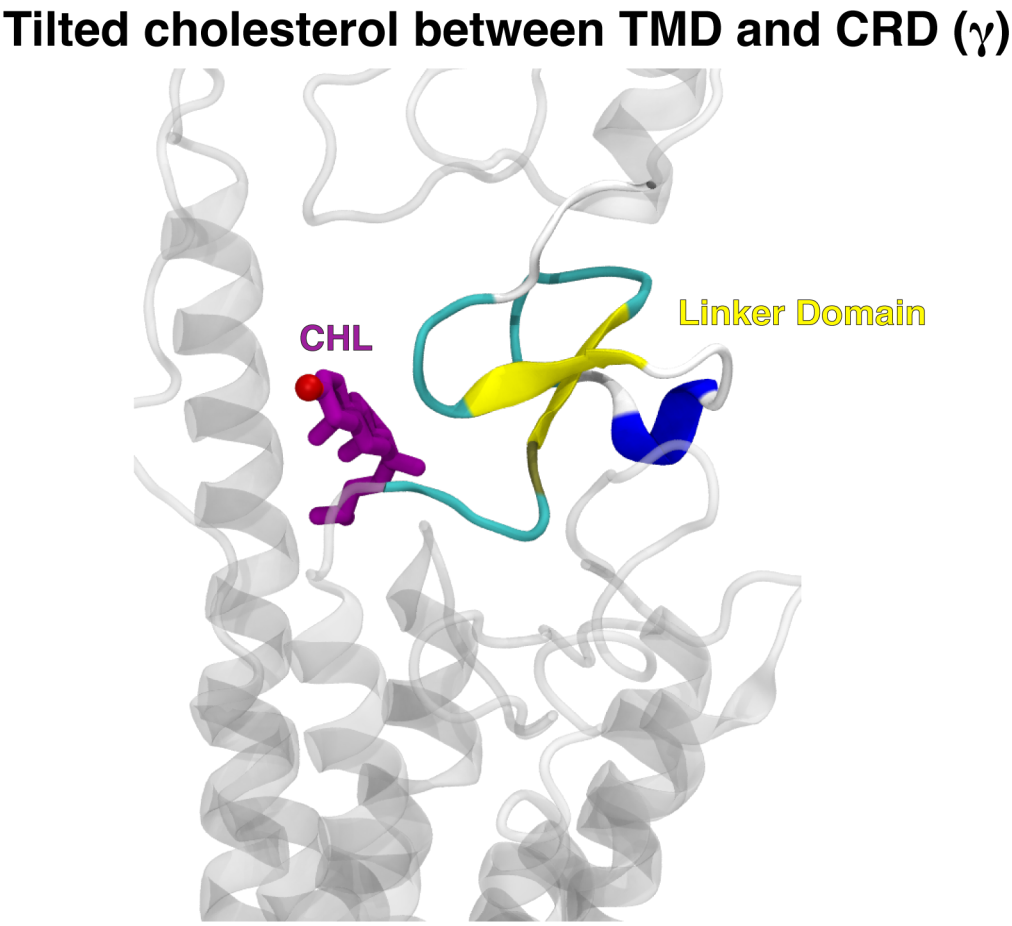
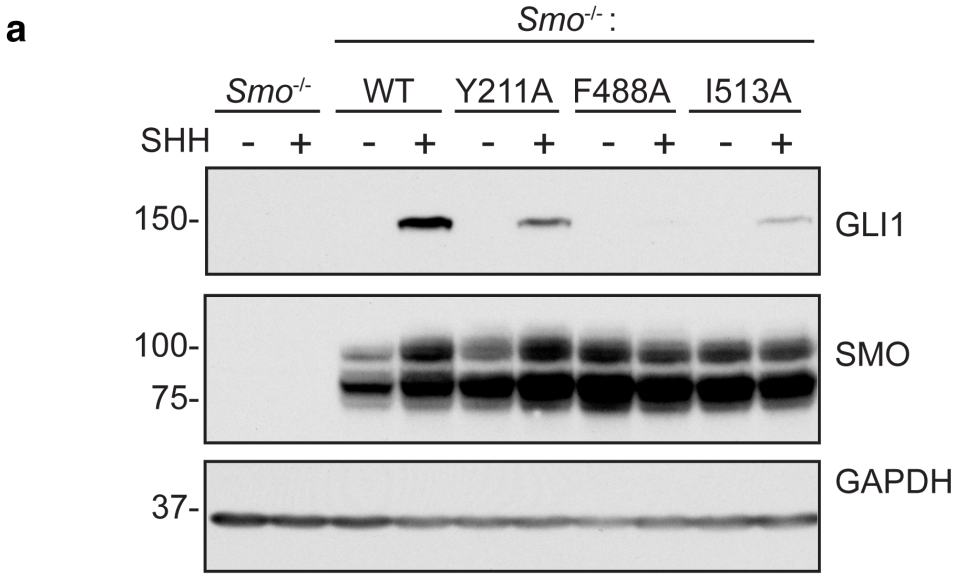


Figure 5—figure supplement 2. Cholesterol (purple) forms a tilted pose to enter the CRD binding site. This is due to the presence of the Linker Domain (yellow).



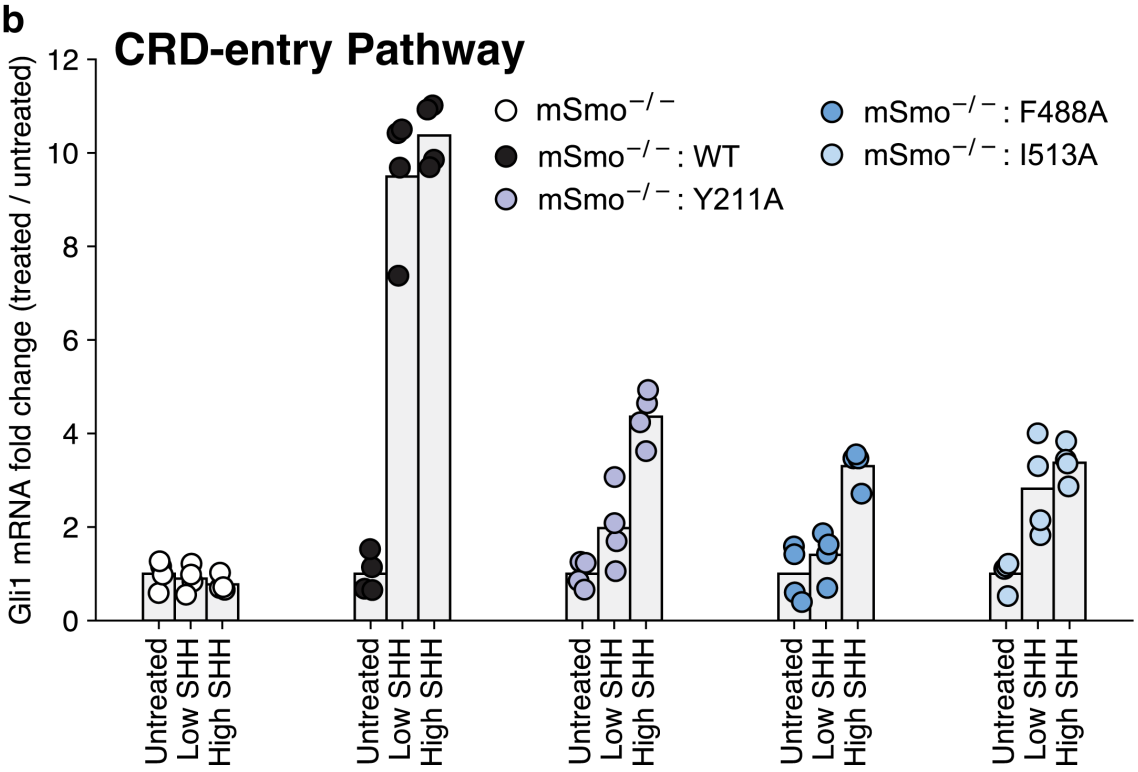


Figure 6—figure supplement 1. Experimental data for mutants along the Common Pathway between the TMD and the CRD. (a) Immunoblotting was used to measure the abundance of mSMO and GLI1 proteins in $SMO^{-/-}$ cells stably expressing either mSMO-WT or Common Pathway mutants after treatment with SHH. (b) *Gli1* mRNA fold change plotted for SMO mutants, showing fold change when the mutants are untreated, treated with a low concentration of SHH, and treated with a saturating concentration of SHH.

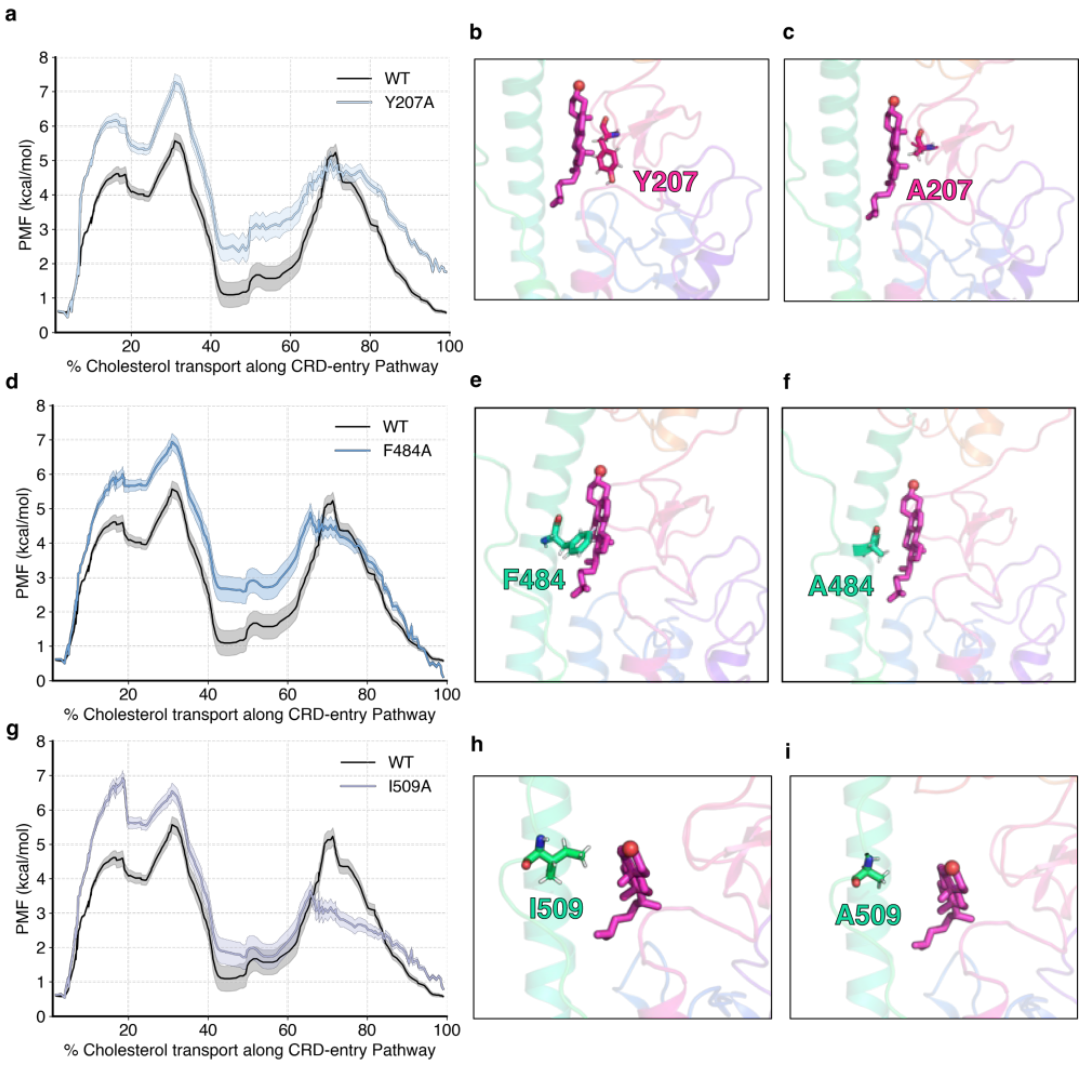


Figure 6—figure supplement 2. PMF data for mutants along the Common Pathway between the TMD and the CRD. (a) PMF for cholesterol translocation was computed for *hSMO* Y207A and compared with WT *hSMO* using adaptive biasing force-based simulations. Representative figures showing the mutation for WT Y207 (b) versus A207 (c). (d, g) - Same as (a) but for F484A and I509A, respectively. (e, h) same as (b) but for F484 and I509, respectively. (f, i) - same as (c) but for A484 and A509, respectively.

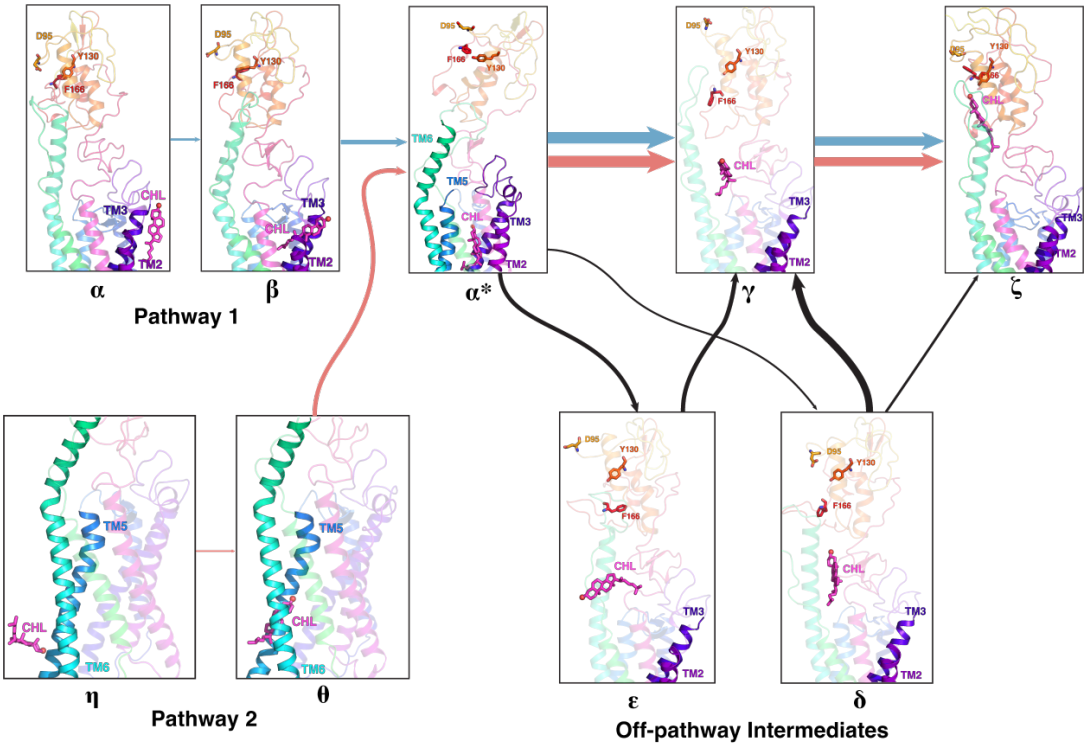


Figure 6—figure supplement 3. Cholesterol positions along the entire transport pathway from membrane to CRD. Pathway 1 is shown as ($\alpha \rightarrow \beta \rightarrow \alpha^*$), while Pathway 2 is shown as ($\eta \rightarrow \theta \rightarrow \alpha^*$) from the membrane to the TMD are shown separately in the left. The Common Pathway from the TMD to the CRD is $\alpha^* \rightarrow \gamma \rightarrow \zeta$, with off-pathway intermediates ϵ and δ.

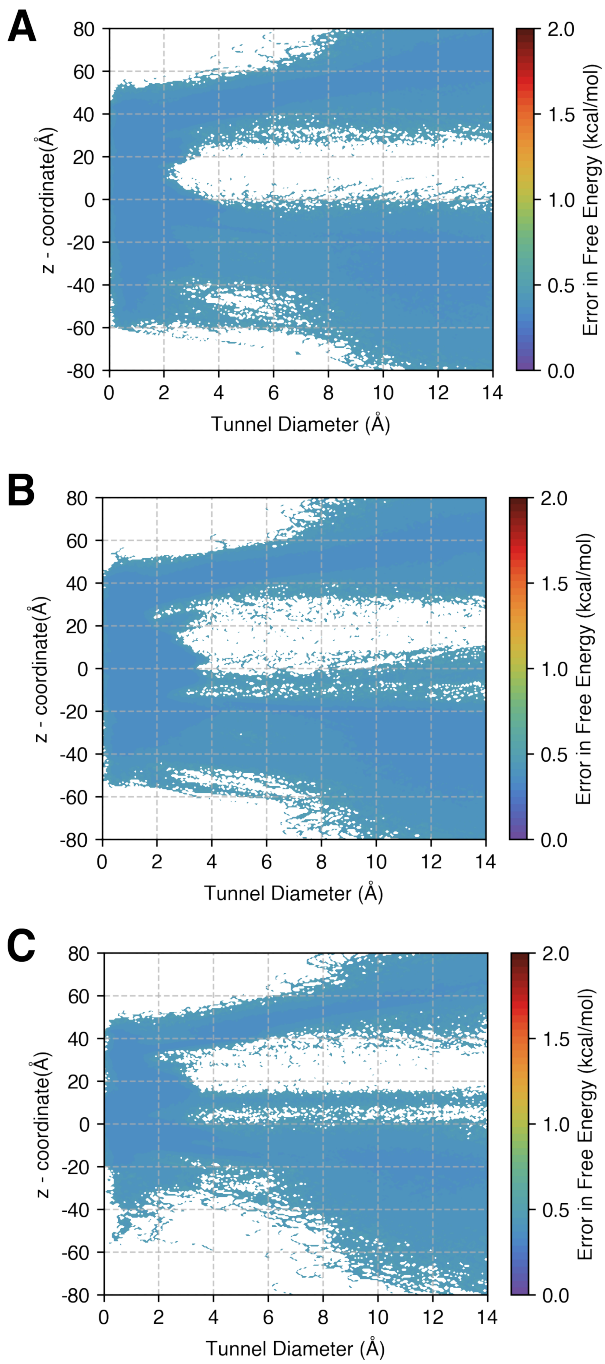


Figure 7—figure supplement 1. Error in Free Energies for Figure 7. Errors were computed using a bootstrapping strategy, where the MSM probabilities were computed from 80% of the data 200 times. (a) Plot showing error for Figure 7(a). (b) Plot showing error for Figure 7(b). (c) Plot showing error for Figure 7(c).

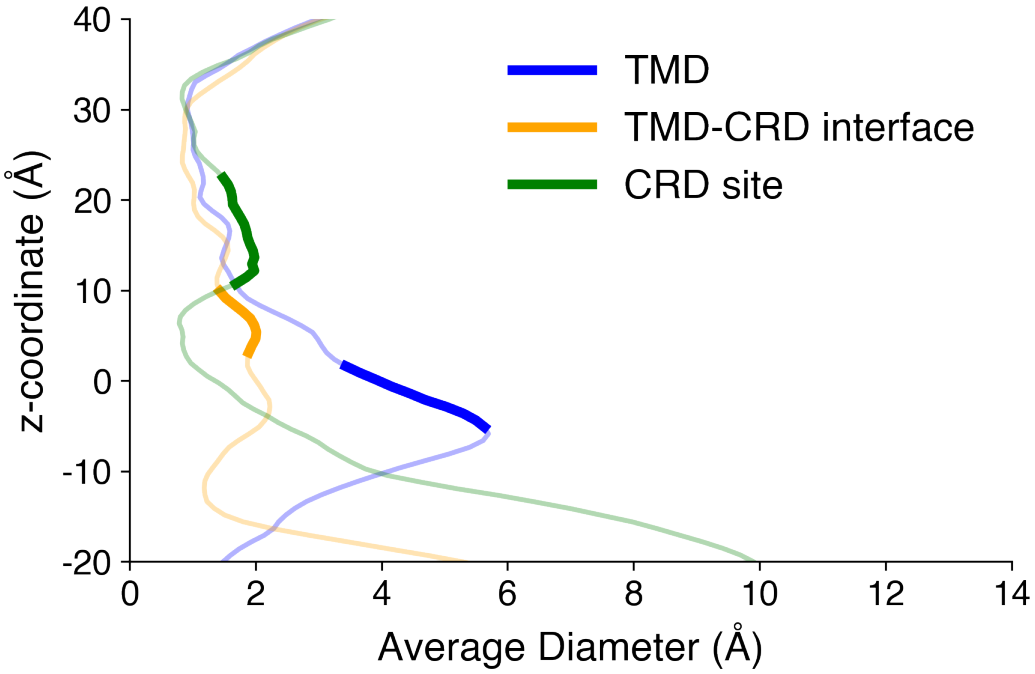


Figure 7—figure supplement 2. Average tunnel radius for Figure 7. The different positions of the cholesterol are marked in bold, and the overall tunnel profile has been marked as translucent.