The heterotrimeric SecY translocon complex is required for the cotranslational assembly of membrane proteins in bacteria and archaea. The insertion of transmembrane (TM) segments during nascent-chain passage through the translocon is generally viewed as a simple partitioning process between the water-filled translocon and membrane lipid bilayer, suggesting that partitioning is driven by the hydrophobic effect. Indeed, the apparent free energy of partitioning of unnatural aliphatic amino acids on TM segments is proportional to accessible surface area, which is a hallmark of the hydrophobic effect [Öjemalm K, et al. (2011) Proc Natl Acad Sci USA 108(31):E359–E364]. However, the apparent partitioning solvation parameter is less than one-half the value expected for simple bulk partitioning, suggesting that the water in the translocon departs from bulk behavior. To examine the state of water in a SecY translocon complex embedded in a lipid bilayer, we carried out all-atom molecular-dynamics simulations of the Pyrococcus furiosus SecYE, which was determined to be in a "primed" open state [Egea PF, Stroud RM (2010) Proc Natl Acad Sci USA 107(40):17182–17187]. Remarkably, SecYE remained in this state throughout our 450-ns simulation. Water molecules within SecY exhibited anomalous diffusion, had highly retarded rotational dynamics, and aligned their dipoles along the SecY transmembrane axis. The translocon is therefore not a simple water-filled pore, which raises the question of how anomalous water behavior affects the mechanism of translocon function and, more generally, the partitioning of hydrophobic molecules. Because large water-filled cavities are found in many membrane proteins, our findings may have broader implications.
Fig. 1. Structure of *P. furiosus* SecYE and cartoon representation of the translocon-to-lipid bilayer partitioning of a TM helix. (A) Structure of SecYE in a lipid bilayer. The lipid headgroups are shown in ice-blue van der Waals representation. The 10 SecY helices are represented in gray cartoon format, except for TM22 (magenta) and TM7 (cyan). The SecE helices are colored green. Hydrophobic ring (HR) residues are drawn as yellow bonds. The plug domain TM22 is colored orange. (B) The apparent free energy of partitioning $\Delta G(z)$ of an alkyl sidechain (orange) with accessible surface area $A_{acc}$ depends upon position $z$ within the TM segment (solid red curve), because the atomic solvation parameter $\sigma$ apparently depends upon position. Our results suggest an opposite behavior (dashed orange curve).

explain the magnitudes and position dependence of the translocon/bilayer partitioning solution parameters? To address this question, we turned to atomistic molecular-dynamics (MD) simulations.

From a number of crystallographic SecY structures (4–8), we chose the SecY structure from *Pyroccotus furiosus* determined by Egea and Stroud (8), because it appears to be in a nearly open state as judged by the separation of the gate helices. We thought at the outset of our simulations that SecYE might be closed. However, SecYE remained stably open, which allowed close examination of the waters inside SecY. We found that waters deep within the translocon diffuse anomalously, have slow rotational dynamics, and have their dipole aligned along the SecY axis. These properties indicate that translocons are not simple water-filled pores, which raises fundamental questions about the nature of translocon/bilayer partitioning of TM segments, and more generally about solute partitioning into compartments of restrained water molecules.

**Results**

**SecYE Remained Stably Open During the Simulation.** We began our study with a 0.45-μs simulation (Sim1) of SecYE embedded in a palmitoyloleoylphosphatidylcholine (POPC) bilayer in excess water using NPT (constant number of particles $N$, pressure $P$, and temperature $T$) conditions (Methods). Egea and Stroud (8) reported the *P. furiosus* SecYE translocon structure to be in a primed open state based upon the dimensions of the HR. We monitored the HR radius throughout the simulation (SI Appendix, Fig. S1) and found it to be quite stable with a mean radius of 6.8 ±0.3 Å (SD) (see Fig. 3D). The same measurement in the closed *Methanococcus janessichii* SecYE10 structure yielded 4.6 ± 0.1 Å, as observed by Egea and Stroud (8) (SI Appendix, Fig. S2). Consistent with the stability of the ring, the SecY TM region was rather rigid (C₆₀ rmss < 2 Å; SI Appendix, Fig. S4A), confirming that SecY did not undergo any major conformational changes during the simulation. We confirmed the stability of the SecY structure by carrying out two additional NPT simulations of 130 ns, described in Supporting Information (SI Appendix, Figs. S4 and S5).

**Water Distribution Within SecYE Conforms to the Hourglass Shape of the Interior.** To examine the water distribution within SecY, we defined a 40 × 40 × 100-Å³ square prism ("the prism") enclosing SecY and centered at the simulation cell origin (Fig. 2B). The placement resulted in the HR ring being located between $z = ±44$ Å (SI Appendix, Fig. S1B). We then determined the time-averaged number of water molecules within the prism in consecutive slices of 1-Å thickness along the $z$ axis (Fig. 2C). The average total number of water molecules within the hourglass between $z = ±20$ Å was about 430.

Dewetting phenomena have been observed in MD simulations of hydrophobic protein channels and gates (23–26). We looked for dewetting in the HR region by monitoring the water occupancy within the volume enclosed by the HR sidechains (Fig. 3A and SI Appendix, Fig. S5). We observed dramatic fluctuations in the number of waters due to lipid acyl chains exploring SecY (Fig. 3B and Movie S2), and they transiently explored the SecY interior, perturbing water passage (Fig. 3C, SI Appendix, Fig. S3, and Movies S1 and S3). These perturbations took place only after the first 0.15 μs of Sim1 (Fig. 3D). To understand how the acyl chains perturbed water within SecY, we examined the time-averaged number of acyl chain carbon atoms inside the prism (SI Appendix, Fig. S3A, Figs. S6 and S7). The major acyl chain incursions occurred between $z = ±10$ Å, but the acyl chain carbon atoms were not able to explore the HR region (SI Appendix, Fig. S3B) during the first 0.15 μs of Sim1, and only to a limited extent after that time. Because the TM helix insertion process involves direct interactions among lipids, water, translocon, and incoming peptides (11, 16), we suppose that these acyl chains—noted in earlier simulations (27, 28)—are likely to play a role in TM helix insertion. However, we do not know whether the lipids we observed are physiologically significant. Similar incursions were seen in Sim2 and Sim3 (SI Appendix, Fig. S5).

**Water Dynamics Within the Hourglass Deviates Dramatically from Bulk Behavior.** For the analysis of the NPT simulations, we stored the system coordinates every 10 ps, but this time interval is too long to reveal the details of water dynamics that occur on the picosecond timescale (29, 30). Therefore, we carried out 40 NVE (constant number of particles $N$, volume $V$, and energy $E$) simulations of 50-ps length, saving configurations every 5 fs (Methods). Each of the simulations was started from independent initial structures selected from five different time intervals (red horizontal lines, Fig. 3A) of Sim1. We determined the translational diffusion and the rotational dynamics of water in the prism enclosing SecY (Fig. 3B). We began by examining the first 0.15 μs of Sim1 (dashed vertical red line,
dipole autocorrelation (SDAC) functions with respect to bulk behavior. As characterized by SecY center, the relaxation time \( \tau_{\mu} \) decreased from 1 to 0.40 as the SecY center was approached (Fig. 4C). In contrast, lipid incursions influenced water rotational dynamics only weakly (Fig. 4D). Approaching the SecY center, the relaxation times are about the same as when water is unperturbed by the lipid acyl chains.

Electrostatic and Confinement Effects Determine Dynamics of Water. What features of SecY cause anomalous water behavior? One possibility is restriction of water movement due to the limited volume of the hourglass (confinement). Another possibility is electrostatic interactions between water and the translocon.

We designed two additional 40 NVE setups (Methods) using the same configurations selected from Sim1. In the first setup, the coordinates of SecY were fixed (SecY-fix; SI Appendix, Figs. S6 C and D and S7 C and D). This approach allowed us to determine the effect of SecY thermal motion on water dynamics. In the second setup (SecY-vdw), in addition to “freezing” SecY, we switched off all SecY-E charges, including partial charges (SI Appendix, Figs. S6 E and F and S7 E and F) to separate confinement from electrostatic effects. For \( t < 0.15 \mu s \), water exhibited essentially the same translational dynamics regardless of whether SecY was mobile or frozen (Fig. 4A, solid black circles and solid green diamonds, respectively), meaning that the thermal motion of SecYE had little effect on diffusion. However, after switching off electrostatic interactions, water still exhibited anomalous diffusion, but the value of \( \alpha \) at the SecY center was 0.72 rather than 0.53 (SecY-vdw in Fig. 4A). For \( t > 0.15 \mu s \), when the lipids make incursions inside SecY, \( \alpha \) reached 0.63 at the SecY center (SecY-vdw in Fig. 4C, blue diamonds) instead of 0.40. These findings indicate that both confinement and the SecY electrostatics make water translational dynamics depart from bulk diffusion.

Water rotational dynamics slowed upon freezing SecY (Fig. 4 B and D; compare black symbols and green diamonds) irrespective of the incursions of the lipid acyl chains inside the SecY pore. We speculate that restraining SecY thermal motions perturbs the

Fig. 3. Fluctuations of water occupancy in the HR region, due in part to incursions of lipid acyl chains. (A) Water occupancy calculated over the 0.45-\( \mu s \) simulation in the region enclosing the HR residues. Black dashed lines indicate the times of the snapshots in B and C; red dashed vertical line defines the time \( t = 0.15 \mu s \) at which the lipid acyl chains start exploring the HR region (SI Appendix, Fig. S3B); red horizontal lines indicate time ranges of NVE simulations. (B and C) Representative snapshots taken at \( t = 0.06 \mu s \) (B) and \( t = 0.28 \mu s \) (C). The color scheme is the same as in Fig. 1, except the plug domain TM2a (orange, new cartoon format). Water is represented using both van der Waals and surface representations. Lipids in front of the lateral gate are shown in blue van der Waals format. The gray dashed lines indicate the region explored by the HR residues along the membrane normal (z axis). (D) Time evolution of the HR radius of SecY. The average value is 6.8 ± 0.3 Å.

Fig. 4. NVE simulations reveal anomalous water motions inside SecY. A and B show data for \( t < 0.15 \mu s \); C and D show data for \( t > 0.15 \mu s \). Black circles (SecY) show values from unrestrained simulations; green diamonds, those from SecY-fix; blue diamonds, those from SecY-vdw. (A and C) Variation of \( \alpha \) exponent variation along the membrane normal estimated from the NVE simulations. (B and D) The variation of the rotational relaxation time \( \tau_{\mu} \) along the membrane normal estimated from the NVE simulations.

We characterized the translational motion of waters by calculating the dependence of mean squared displacements (MSDs) on time (SI Appendix, Fig. S6A). In general, after a few picoseconds, the MSD became proportional to \( \mu^2 \) (29). Linear dependence of the MSD on time (\( \alpha = 1 \)) is a signature of Brownian diffusive motion as observed for bulk water; \( \alpha < 1 \) is a signature of anomalous “subdiffusion,” which is characteristic of confined water and protein hydration water. The MSD of water molecules in the long-time regime (\( t > 2 \, \mu s \)) behaved differently in each of the 5-Å-thick slabs, as characterized by \( \alpha \) (Fig. 4A, solid black circles). Approaching the SecY center, the \( \alpha \) decreased from 1 to 0.53, which is similar to bulk diffusion. However, after switching off electrostatic interactions, water still exhibited anomalous diffusion, but the value of \( \alpha \) at the SecY center was 0.72 rather than 0.53 (SecY-vdw in Fig. 4A). For \( t > 0.15 \mu s \), when the lipids make incursions inside SecY, \( \alpha \) reached 0.63 at the SecY center (SecY-vdw in Fig. 4C, blue diamonds) instead of 0.40. These findings indicate that both confinement and the SecY electrostatics make water translational dynamics depart from bulk diffusion.

Water rotational dynamics slowed upon freezing SecY (Fig. 4 B and D; compare black symbols and green diamonds) irrespective of the incursions of the lipid acyl chains inside the SecY pore. We speculate that restraining SecY thermal motions perturbs the
The simplest way to examine the effect of the SecY electrostatics on water structure and the interpretation of the relaxation time $\tau_\alpha$ in the absence of charges is to determine the average orientation of water dipoles with respect to the SecY $z$ axis using the order parameter $P_2(z) = \langle \cos \theta \rangle (\text{Methods})$. A random distribution of dipole moments corresponds to $P_2(z) = 0$. We first analyzed Sim1 and calculated the distribution of the water dipole orientation over three time windows: 0–0.15 $\mu$s, 0.15–0.3 $\mu$s, and 0.3–0.45 $\mu$s. Our analysis of $P_2(z)$ is displayed in Fig. 5 along with a representative snapshot in which water oxygen atoms are colored according to the dipole orientation calculated between 0 and 0.15 $\mu$s as indicated by the colored bar. SecY is represented using the color scheme of Fig. 1.

Discussion

Our simulations bear strongly on the interplay between the translocon and the lipids. During our 0.45-μs simulation, the hydrophobic pore ring remained stably open independent of lipid intrusions (Fig. 3D). These results suggest that the primed open configuration of SecYE may represent a stable conformational state on the timescale of our simulation. The involvement of the lipids in the nascent-chain integration process has been suggested in previous experimental (11, 40) and computational (27, 28) studies. Our simulations provide evidence for the accessibility of the SecY pore to lipid chains of the surrounding bilayer, supporting the possibility of direct peptide-lipid interactions.

The main finding of our simulations is that water inside SecY does not behave as in bulk: translational diffusion deviates markedly from Brownian motion (Fig. 4 A and C) and the rotational dynamics show strong retardation (Fig. 4 B and D). These features are typical for water in confined environments and at hydrated protein surfaces (32, 33, 41, 42). Moreover, and perhaps more important for the thermodynamics of partitioning TM segments into the membrane, we observe that water dipoles in the HR region ($z = 4 \AA$) are oriented preferentially parallel to the pore axis pointing toward the exterior (Fig. 5). The inhomogeneous nonbulk properties of water within SecY suggest that translocon/membrane partitioning cannot be compared directly to partitioning between bulk aqueous and lipid phases. Contrary to expectations formed in the days when the structures of only a few membrane proteins were known, we now know that the interiors of membrane proteins (43–47), and even soluble proteins (48, 49), can contain considerable amounts of water, which raises the question of how restricted waters might participate in or affect protein function.

Our results reveal the importance of the SecY electrostatics, which might explain, for example, why flanking charges in model TM segments affect insertion by the Sec61 translocon (50). Moreover, the dipole alignment of water molecules within SecY might be a signature of another potential role of water in the membrane insertion process. For instance, it has been shown that water can facilitate or screen the interactions between lipids and a peptide located inside SecY (15). Furthermore, the dipole alignment could be crucial for the interaction of the positively charged N terminus of a signal sequence with the translocon and consequently for membrane protein topology. The presence of a highly polarizing field would likely affect helix insertion. If the idea that incoming peptides pass directly through the SecY hourglass is correct, they would likely replace a substantial number of water molecules in the translocon interior. Hence, they would be exposed to an electrostatic field comparable to that experienced by the water molecules in our simulations. A dipole orientation that mimics these aligned water dipoles could therefore have a substantial stabilizing effect.

Favorable water-to-hydrocarbon partitioning free energies of hydrophobic moieties observed in bulk partitioning arise, in part, from the release of oriented water molecules at hydrophobic residue surfaces (51). Given a hydrophobic $\alpha$-helix inside the SecY translocon that simply partitions parallel to the membrane plane into the bilayer (Fig. 1B), one would expect the solvation parameter $\sigma$ for hydrophilic amino acids at the ends of the helix to be similar to bulk values, because water at the extreme ends of the translocon is bulk-like. Öjemalm et al. (22) observed, however, $\sigma$ at the helix ends was $\approx 6$ kcal·mol$^{-1}$·Å$^{-2}$, which is about one-fourth of the value expected for partitioning from bulk water. Given that water is restrained at the center of the translocon, the value of $\sigma$ would one expect for amino acids at the center of a hydrophobic $\alpha$-helix? To
answer that question, other questions must be answered. Is water present inside SecY when a helix is present? If water is present and restrained due to SecY electrostatics and confinement, how is σ affected? If water is absent, what is the meaning of σ? One scenario might be that amino acids located in the center of a helix in the translocon have neighboring oriented water molecules that remain ordered even after the helix moves into the bilayer. In that case, the apparent solvation parameter at the helix center should be less favorable than at the helix ends. Thus, in experiments such as those of Øjemalm et al. (22), one might expect solvation parameters to be least favorable near the helix center and most favorable near the helix ends (dashed orange curve, Fig. 1B), which is exactly opposite to the results of Øjemalm et al. (solid red line, Fig. 1B). This contradiction raises fundamental questions about the translocon/membrane partitioning process of TM helices. Cymer et al. (52) have suggested that nascent TM helices may preferentially interact with the gate helices and the membrane interface (53, 54) without first passing through the SecY hourglass.

Finally, the crucial question our results raise is whether the concept of hydrophobic effect-driven partitioning even applies when a nonpolar solute moves from a restrained water region into the water-free bilayer interior. Or more simply and fundamentally, what is the partitioning free energy of solutes between water in bulk and water in restraining confined spaces? Although much theoretical and experimental work has been devoted to describing what is the partitioning free energy of solutes between water in bulk (55) and in confined spaces (56) such as inverted micelles (57), this fundamental question remains theoretical and experimental work has been devoted to describing the properties of water in bulk (55) and in confined spaces (56) such as inverted micelles (57), this fundamental question remains unanswered as far as we can establish.

Methods

MD Simulations in the NPT Ensemble. The atomic coordinates of the SecYE translocon were extracted from the Pyrococcus furiosus crystal structure (8) (Protein Data Bank (PDB) ID code 3MP7). From this structure, we set up and carried out three independent NPT simulations: Sim1, Sim2, and Sim3 (SI Appendix, Table S1). In all cases, we embedded SecYE in a POPC bilayer formed by 600 lipids (300 molecules each leaflet). The protocols we used to model residues missing in the original structure, to embed the completed SecYE structure in the POPC lipid bilayer, and to minimize and equilibrate the systems are reported in Supporting Information. We also carried out MD simulations using the M. jannaschii SecYEj translocon crystal structure (4) (PDB ID code 1H2R). The same protocols were followed as for P. furiosus SecYE.

We used NAMD (58, 59), version 2.9, with the CHARMM36 (60) force field for the lipids and the CHARMM22 force field with the CMAP correction for the protein and ions (61, 62). The TIP3P model (63) was used for water molecules. The temperature was kept constant at 300 K using a Langevin dynamics coupling in conjunction with Nosé–Hoover–Langevin piston algorithm (64, 65). Periodic boundary conditions were applied in three dimensions. The electrostatic interactions were computed by means of the smooth particle-mesh Ewald summation method (66, 67) and the short-range real-space interactions were cut off at 12 Å using a switching function between 10 and 12 Å. The equations of motion were integrated with a time step of 4 fs for the long-range electrostatic forces, 2 fs for the short-range nonbonded forces, and 1 fs for the bonded forces by means of a multiple-time step algorithm (68). The SHAKE (69) algorithm was used to constrain the length of the bonds involving hydrogen atoms. Coordinates were saved every 10 ps.

MD Simulations in the NVE Ensemble from Sim1. To characterize water properties within SecY and to investigate how the SecY pore affects water dynamics, we performed five sets of eight NVE simulations, i.e., 40 NVE simulations in total. We selected the eight independent configurations every 2 ns from the following five time ranges of Sim1: 2–16, 60–74, 104–118, 302–316, and 402–416 ns. We used these 40 structures as starting configurations for running NVE simulations of 50-ps length. To have high resolution in time, we collected coordinates every 5 fs.

In addition to these five sets of eight NVE simulations in which the protein complex was unrestrained (labeled SecY in Fig. 4), we used the same starting structures to carry out two more groups of five sets of eight NVE simulations, applying restraints as follows. In the first group, we applied harmonic restraints to freeze SecY (labeled SecY-fx in Fig. 4), and in the second, in addition to freezing SecY, we switched the SecY–water electrostatic interactions off by setting the protein atom charges to zero (labeled SecY-vdw in Fig. 4). These three groups of 40 NVE simulations (SecY, SecY-fx, and SecY-vdw) enabled us to examine how the cavity shape and the electrostatics of SecY affect individually the dynamics of water. In total, we generated and analyzed 120 NVE simulations, summarized in SI Appendix, Table S2.

Water Anomalous Dynamics. We investigated water dynamics in a selected region of the simulation cell characterized by a volume of 40 × 40 × 100 Å³ encompassing SecY (Fig. 2B). We subsequently divided this square prism into 20 5-Å-thick slabs parallel to the membrane (Fig. 2C). For each slab, we first computed and analyzed the MSDs, which characterize the translational diffusion of water molecules. MSDs are defined as follows:

$$\text{MSD}(t) = \langle [r_i(t) - r_i(0)]^2 \rangle,$$

where $$r_i(t)$$ denotes the position of the particle i at time t, and the brackets denote an average over molecules and time origins. In SI Appendix, Figs. S6 A, C, and E, and S7 A, C, and E, we present the MSD of the center of mass of water molecules computed for the SecY, SecY-fx, and SecY-vdw for each slab and for times $$t < 0.15$$ μs and $$t > 0.15$$ μs. In the long-time regime, the time dependence of the MSD can be described with a power law MSD(t) = $$kt^n$$, and the n exponent gives information on the diffusive character of water motion. We fitted the MSD to the power-law function for $$t > 2$$ ps and represented the n values in Fig. 4. Further details of the protocol used for the data analysis are reported in Supporting Information.

Subsequently, we calculated the SDAC functions $$C_2(t)$$, which signify the rotational motion of water molecules and are defined as follows:

$$C_2(t) = \frac{\langle \mu(0) \cdot \mu(t) \rangle}{\langle \mu(0) \cdot \mu(0) \rangle},$$

where $$\mu(t)$$ is the unit vector of the water dipole at time t. In SI Appendix, Figs. S6 B, D, F, and S7 B, D, and F, we have plotted the SDAC functions computed for SecY, SecY-fx, and SecY-vdw for each slab for the time windows $$t < 0.15$$ μs and $$t > 0.15$$ μs. We analyzed consistently all of the 120 NVE simulations and calculated the characteristic relaxation time $$\tau_2$$, as the time at which $$C_2(t)$$ decays to 1/e, that is, $$C_2(t) = 1/e$$ (SI Appendix, Figs. 56 and S7, dashed lines). The $$\tau_2$$ values are reported in Fig. 3. Further details of the data analysis procedure are reported in Supporting Information.

Finally, we examined the water dipole orientation within the SecY channel in terms of the water orientational order parameter defined as follows:

$$P_{\mu}(z) = \cos(\theta) / 2,$$

where θ is the angle between the water dipole and the membrane normal. The θ value varies between 0° and 180°, hence cosθ varies between −1 and 1. We calculated the values of cosθ between 0 and 0.15 μs, 0.15 and 0.3 μs, and 0.3 and 0.45 μs of Sim1 (Fig. 5), and over all of the NVE simulations (SecY, SecY-fx, and SecY-vdw) (SI Appendix, Fig. 58 A and C). The procedure we followed is described in detail in Supporting Information.

ACKNOWLEDGMENTS. We thank J. Alfredo Freites for extremely valuable discussions and Joseph Farran for excellent technical support. We thank Prof. Kim Sharp for useful discussions of restrained waters in proteins. A.-N.B. and S.C. thank the International Office of the Freie Universität, Berlin, for travel support. This research was supported by National Institutes of Health Grants RO1 GM74637 (to S.H.W.) and P01 GM86685 (to D.J.T. and S.H.W.). M.H. was supported by the German National Academy of Science, Leopoldina. A.-N.B. was supported in part by Marie Curie International Reintegration Award IAR 276920. The simulations were performed in part on the High-Performance Computing Cluster at the University of California, Irvine, and in part on Stampede on Extreme Science and Engineering Discovery Environment, supported by National Science Foundation Grant ACI-1053575.


Supporting Information

Capponi et al. 10.1073/pnas.1424483112

Movie S1. Passage of water molecules through the translocon of *Pyrococcus furiosus*. The movie is derived from the water movement captured in Sim1. The movie shows 1.5 ns of the simulation starting at the 276.8-ns point of Sim1. Shown here is the first frame of the movie.

**Movie S1**

Movie S2. Images of lipids at the front of the lateral gate of SecYE and not interfering with water movement in the vicinity of the hydrophobic ring. The movie shows the structure of SecYE and the lipids as SecYE is rotated about the z axis. The movie is essentially a rotating view of Fig. 3B. Shown here is the first frame of the movie.

**Movie S2**
Movie S3. Images of lipids with acyl chains dipping into SecYE and interfering with water movement in the vicinity of the hydrophobic ring. The movie shows the structure of SecYE and the lipids as SecYE is rotated about the z axis. The movie is essentially a rotating view of Fig. 3C. Shown here is the first frame of the movie.

Movie S3

Other Supporting Information Files

SI Appendix (PDF)
Anomalous behavior of water inside the SecY translocon

Sara Capponi\textsuperscript{a}, Matthias Heyden\textsuperscript{b,1}, Ana-Nicoleta Bondar\textsuperscript{c}, Douglas J. Tobias\textsuperscript{b} and Stephen H. White\textsuperscript{a,2}

\textsuperscript{a}Department of Physiology and Biophysics and Center for Biomembrane Systems, University of California at Irvine, Irvine, California 92697-4560

\textsuperscript{b}Department of Chemistry and Center for Biomembrane Systems, University of California at Irvine, Irvine, California, USA 92697-2025

\textsuperscript{c}Department of Physics, Freie Universität Berlin, 14195 Berlin-Dahlem, Germany

\textsuperscript{1}Current address: Theoretical Chemistry, Max-Planck-Institut fur Kohlenforschung, Mulheim an der Ruhr, Germany
Supplementary Methods and Results

**MD simulations in the NPT ensemble.** A number of residues from the *Pyrococcus furiosus* crystal structure (1) (Protein Data Bank code 3MP7) primarily in the helix connecting links, were unresolved. We modeled these into the structure using Phyre2 software (2). The completed SecYE structure was subsequently energy minimized in vacuum, and this minimized conformation was used for setting up three independent simulations: Sim1, Sim2, and Sim3 (See Table S1). In all cases, we embedded SecYE in a POPC bilayer formed by 600 lipids (300 molecules each leaflet). Using the VMD software (3), we aligned the principal axes of SecY along the $x$, $y$ and $z$ directions ($z$-axis normal to the bilayer). The final position of the protein complex within the lipid bilayer was achieved by placing the SecE amphiphilic helix in the membrane interface, avoiding hydrophobic mismatch. The system was then solvated by water molecules, and chloride ions were added to maintain charge neutrality. Several lipid and water molecules were removed to avoid steric clashes. The whole system was relaxed by energy minimization using the conjugate gradient algorithm for 10,000 steps and gradually heated from 25 K to 300 K. Sim1, Sim2, and Sim3 were initiated from the same minimized configuration but using different random initial velocities. Compared with Sim1, in Sim3 (Fig. S4C, Fig. S5B), we also repeated the whole set up protocol described above ‘from scratch’, i.e. from embedding the SecYE complex in a POPC bilayer. The procedure for equilibrating the system is described in the following.

**System Equilibration Protocol.** In order to equilibrate the positions of lipids, ions, water molecules, and protein complex, harmonic restraints were used in a series of five consecutive simulation runs of 1 ns, as follows. In the first equilibration run, restraint force constants of 5 kcal mol$^{-1}$ Å$^{-2}$ were applied to the protein backbone atoms and 0.5 kcal mol$^{-1}$ Å$^{-2}$ to ions, water, and lipid headgroups located farther than $\sim$ 10 Å from the protein. In the subsequent two runs, the restraints on the protein backbone were decreased to 2 and then to 1 kcal mol$^{-1}$ Å$^{-2}$, while restraints on ions, water, and lipid headgroups were 0.2 and to 0 kcal mol$^{-1}$ Å$^{-2}$, respectively. After the last 1 ns equilibration run, in which the restraints on the protein backbone were set to 0.5 kcal mol$^{-1}$ Å$^{-2}$ while the rest of the system was unrestrained, we carried out a further 1 ns run with all the restraints removed. We then began the production runs. The simulations were carried out under NPT conditions (constant number of particles $N$, pressure $P$, and temperature $T$). We also carried out MD simulations using the *M. jannaschii* SecYEβ translocon crystal structure (4) (PDB code 1RHZ). The same protocols were followed as for *P. furiosus* SecYE.

**Dynamics of the ‘primed’ translocon in a POPC bilayer from Sim2 and Sim3.** As we observed in Sim1, SecY remained in the primed open state throughout the whole length of Sim2 (Fig. S4B
and S5A) and Sim3 (Fig. S4C and S5B). Although portions of the lipid acyl chains made incursions into the water-filled interior of SecY, the HR radius was unaffected by the incursions. In all of the simulations, the SecY TM region was rigid with Cα RMSD fluctuating around 2 Å (Fig. S4), and the acyl chains of several lipids—initially located in front of the lateral gate—explored the SecY interior, sometimes perturbing the water occupation (Fig. 3, Fig. S3 and Fig. S5). In Sim2 and Sim3, we detected interruptions in the water permeation through SecY after tens of ns, and in Sim1 after 0.16 µs. Sim1 was thus prolonged for 0.45 µs and 40 independent configurations were used to initiate the NVE simulations described in the next section.

**MD simulations in the NVE ensemble from Sim1.** In order to characterize water properties within SecY, we performed 5 sets of 8 NVE (constant number of particles $N$, volume $V$ and energy $E$) simulations, i.e. 40 NVE simulations in total. We selected the 8 independent configurations every 2 ns from the following 5 time ranges of Sim1: 2 to 16 ns, 60 to 74 ns, 104 to 118 ns, 302 to 316 ns, and 402 to 416 ns. In addition to these 40 NVE simulations, we carried out a group of 40 NVE simulations in which SecY was maintained fixed (labeled SecY-fix in Fig. 4) and another group of 40 NVE simulations, in which we fixed SecY and switched the SecY-water electrostatic interactions off by setting the protein atom charges to zero (labeled SecY-vdw in Fig. 4). These 3 groups of 40 NVE simulations summarized in Table S2 enabled us to examine how the cavity shape and the electrostatics of SecY affect individually the dynamics of water. For the results presented in Figure 4, we grouped together the 3 sets of 8 NVE simulations carried out in the time windows $t < 0.15$ µs and the 2 carried out in the time windows $t > 0.15$ µs (see dashed red line Fig. 3A), we analyzed the data and averaged the results.

**Water anomalous dynamics from the NVE simulations: translational diffusion.** We studied water translational diffusion by computing the mean squared displacements (MSD) of the center-of-mass of the water molecules. We calculated the MSD for each slab and for the SecY, SecY-fix and SecY-vdw simulations and we represented the average of the NVE runs from times $t < 0.15$ µs and $t > 0.15$ µs in the left-hand panels of Fig. S6 and S7, respectively. Colored lines describe water inside SecY and black lines those outside SecY (bulk water). In the long-time regime, the time dependence of the MSD can be described with a power law function $MSD(t) = kt^\alpha$ (5, 6). The $\alpha$ exponent gives information on the diffusive character of water motion; the constant $k$ is a fitting parameter whose value is irrelevant in the present study. A linear dependence on time ($\alpha = 1$) of the MSD is a signature of Brownian motion exhibited by bulk water molecules. A MSD sublinear time-dependence ($\alpha < 1$) indicates anomalous diffusion, which is characteristic of protein hydration water and of confined water. We used the same procedure to analyze all the 120 NVE simulations. We fitted the MSD to the power law function.
for $t > 2$ ps and we estimated the value of the $\alpha$ exponent for each slab and each NVE run. We averaged the results obtained from NVE simulations starting from system configurations sampled from Sim1 at times $t < 0.15$ $\mu$s and times $t > 0.15$ $\mu$s. The results are presented in Fig. 4A-C with the corresponding standard errors.

**Water anomalous dynamics from the NVE simulations: rotational dynamics.** The $C_\mu(t)$ functions reflect the combination of several dynamical processes taking place at different time windows and their decay is usually described by a sum of exponential or stretched exponential functions, whose amplitudes, parameters, and characteristic relaxation times are related to the type of motion involved. We defined the characteristic relaxation time $\tau_\mu$ as the time at which $C_\mu(t_\mu)$ decays to $1/e$, that is $C_\mu(t_\mu) = 1/e$. This definition enabled us to analyze all of the slabs and all of the 120 NVE simulations consistently. In the right-hand panels of Figure S6 and S7, we represented the SDAC functions computed for SecY, SecY-fix, and SecY-vdw for each slab and averaged for NVE simulations starting at times $t < 0.15$ $\mu$s and $t > 0.15$ $\mu$s. The dashed lines represent $1/e$. Colored lines describe water inside SecY and black lines water outside SecY, which correspond to bulk water. We first analyzed individually each slab and each set of NVE simulations and estimated the value of $\tau_\mu$. Subsequently, we grouped the results obtained from the NVE runs starting at time $t < 0.15$ $\mu$s and $t > 0.15$ $\mu$s. The results are shown in Fig. 4B-D with the corresponding standard errors. In the SecY and SecY-fix simulations, the SDAC functions of the water molecules located between $-20$ Å and $+20$ Å did not decay. Therefore, we fitted them with a 3-exponential function and extrapolated them to $1/e$ to determine estimates of $\tau_\mu$. Because of the lack of decay of the SecY and the SecY-fix SDAC functions between $-5$ and $10$ Å, we could not obtain a reliable fit.

**Water dipole orientation in the NPT and NVE simulations.** We examined the orientation of water molecules within the SecY channel in terms of the water orientational order parameter defined as $P_\vartheta(z) = \langle \cos \vartheta \rangle(z)$, where $\vartheta$ is the angle between the water dipole and the membrane normal. The $\vartheta$ value varies between $0^\circ$ and $180^\circ$, hence $\cos \vartheta$ varies between $-1$ and $1$. A random distribution of the water dipoles corresponds to the $<\cos \vartheta> = 0$; non-zero values of $<\cos \vartheta>$ indicate a preferred dipole orientation. For the NPT simulation Sim1, we read 1400 frames in each of the time range analyzed: $0$ to $0.15$ $\mu$s, $0.15$ to $0.3$ $\mu$s, and $0.3$ to $0.45$ $\mu$s. We computed $\cos \vartheta$ by averaging the results obtained every 100 frames. We then calculated the final average and represented it as a function of the membrane normal (Fig. 5A). The error bars represent the standard errors of the means. For the 120 NVE simulations, we read 5000 frames and computed $\cos \vartheta$ by averaging the results obtained every 250 frames. We then calculated the
final average and represented it as a function of the position along the membrane normal in the left-hand panels of Fig. S8. The error bars represent the standard errors of means.

*Lipid incursions inside SecY in the NPT and NVE simulations.* In order to calculate the time-averaged number of acyl chain carbon atoms located inside the square prism enclosing SecY during Sim1 (Figure S3A), we read 1400 frames in each of the time range analyzed: 0 to 0.15 µs, 0.15 to 0.3 µs, and 0.3 to 0.45 µs. We computed the number of the acyl chain carbon atom by averaging the results every 250 frames. We then calculated the final average and represented the results as a function of the distance from the SecY center. The error bars represent the standard errors of the mean. In order to calculate the number of acyl chain carbon atoms located inside the SecY hydrophobic region, we considered the volume defined by the hydrophobic residue positions (Figure S3B) and we followed the procedure previously described. We performed the same analysis for Sim2 (Figure S5A-v) and Sim3 (S5B-v). We represented the time-averaged number of acyl chain carbon atoms located inside the square prism enclosing SecY calculated from the 120 NVE simulations in the right-hand panels of Fig. S8. We read 5000 frames and obtained the number of the acyl chain carbon atom by averaging the results every 250 frames. We then computed the final averages and we presented them as a function of the position along the membrane normal. We grouped together the results obtained from NVE simulations run at times $t < 0.15$ µs (Fig. S8, upper panels) and those obtained at times $t > 0.15$ µs (Fig. S8, lower panels).
Supplementary Tables

Table S1. **MD simulations in the NPT ensemble.** The first and the second columns respectively report the three independent NPT simulations Sim1, Sim2 and Sim3, and their lengths.

<table>
<thead>
<tr>
<th>NPT simulations</th>
<th>Length (µsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sim1</td>
<td>0.45</td>
</tr>
<tr>
<td>Sim2</td>
<td>0.13</td>
</tr>
<tr>
<td>Sim3</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Table S2. **MD simulations in the NVE ensemble.** Summary of the 120 NVE simulations from Sim1 carried out to examine water dynamics. We selected 8 independent configurations every 2 ns from the following 5 time ranges of Sim1: 2 to 16 ns, 60 to 74 ns, 104 to 118 ns, 302 to 316 ns, and 402 to 416 ns. The specific simulated systems are listed in the first column, the number of NVE runs for each time window for each system in the second, time interval at which configurations were saved in the third, and the length of each simulation in the fourth.

<table>
<thead>
<tr>
<th>NVE simulations (from Sim1)</th>
<th>NVE simulations per time window</th>
<th>interval (fs)</th>
<th>length (ps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SecY (SecY)</td>
<td>8 x 5</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Frozen SecY (SecY-fix)</td>
<td>8 x 5</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Frozen SecY and charges off (SecY-vdw)</td>
<td>8 x 5</td>
<td>5</td>
<td>50</td>
</tr>
</tbody>
</table>
**Figure S1.** Dimensions of the 'primed' open hydrophobic ring (HR) of *P. furiosus* and the fluctuations of its six hydrophobic residues observed in Sim1. **A.** Time evolution of the centers-of-mass of the six HR sidechains projected onto the *xy* plane parallel to the lipid bilayer. The dashed circle through the time-evolution data represents the mean opening of the ring, calculated by fitting a circle through the points. The circle has a radius of 6.8±0.3 (s.d.) Å. Shown as a gray disk is the mean opening of the HR of the closed *M. jannaschii* SecYEβ (see Fig. S2). **B.** Time evolution of the centers-of-mass of the HR residues projected onto the *z*-axis.
Figure S2. Dimensions of the closed hydrophobic ring (HR) of the SecYEβ translocon from *M. jannaschii* and the fluctuations of its six hydrophobic residues. **A.** Time evolution of the centers-of-mass of the hydrophobic residues projected onto the *xy* plane parallel to the lipid bilayer. The dashed circle through the time-evolution points represents the time-averaged opening of the HR obtained by fitting a circle through the points. The mean radius of the circle is 4.6±0.1 (s.d.) Å. **B.** Time evolution of the centers-of-mass of the HR residues projected onto the *z*-axis. **C.** Time evolution of the HR radius calculated over the entire length of the simulation.
Figure S3. Average numbers of acyl chain carbon atoms within the aqueous interior of the SecYE translocon during Sim1 calculated by averaging over 1400 frames. The calculation was performed between 0 and 150 ns (black symbols), between 150 and 300 ns (orange symbols), and between 300 and 450 ns (red symbols). The error bars represent the standard errors of the mean. A. Number of acyl chain carbon atoms located inside the 40 Å × 40 Å × 100 Å square prism shown in Figure 2B. Overall, the acyl chain carbon atoms occupy less than the 2 % of the total volume of the prism. B. Average numbers of acyl chain carbons within HR region only, as defined by the hydrophobic residue positions (Fig. S1). Acyl chain carbons are absent or rarely found in the hydrophobic ring region. Note difference in vertical scale relative to panel A.
Figure S4. Time evolution of the root mean-squared deviations (RMSD) of SecYE from *P. furiosus* for the three *NPT* simulations presented in the paper. The panels show the time evolution of the RMSD of all the SecY Cα atoms (*red*), the SecY TM helices (*black*), TM2b (*magenta*), TM7 (*cyan*), and TM2a plug domain (*orange*) calculated from Sim1 (**A**), Sim2 (**B**), and Sim3 (**C**).
Figure S5. Summary of data obtained from Sim2 (A) and from Sim3 (B). (i) The dashed circle through the time-evolution points represents the time-averaged opening of the HR obtained by fitting a circle through the centers of mass of the HR residues. The mean radius is 6.6±0.2 (s.d.) Å and 7.0±0.3 (s.d.) Å for Sim2 and Sim3, respectively. (ii) Time evolution of centers of mass of the hydrophobic residues projected onto the z-axis. (iii) Time evolution of the HR radius calculated over the entire length of the simulation. (iv) Time-averaged number of acyl chain carbon atoms located within the 40 Å × 40 Å × 100 Å square prism (solid circle symbols) and in the hydrophobic region as defined by hydrophobic ring residues (empty circle symbols). The error bars represent standard errors of the mean. (v) Time evolution of the water occupancy in the region enclosed by the six hydrophobic ring residues.
Figure S6. Time evolution of the mean squared displacements (MSD) and of the single-molecule dipole autocorrelation functions (SDAC). The results, averaged over the three types of NVE simulations carried out at times \( t < 0.15 \) µs, are shown under three conditions (see Methods): no restraints, SecY. Coordinates fixed, SecY-fix. Coordinates fixed and the SecY charges (including partial charges) turned off, SecY-vdw. Time evolution of the MSDs and SDACs are shown in the left-hand and right-hand panels, respectively. A and B, SecY. C and D, SecY-fix. E and F, SecY-vdw. The curves are colored according to the legends within the MSD plots describing the position along the membrane normal of the analyzed slab. The black line corresponds to water outside SecY, and the colored lines to water inside SecY. The time dependence of the MSD in the long-time regime was determined from the power function \( t^\alpha \) (see Methods). The dashed lines in the SDAC plots correspond to the value 1/e.
Figure S7. Time evolution of the mean squared displacements (MSD) and of the single molecule dipole autocorrelation functions (SDAC). The results, averaged over the two NVE simulations carried out at times $t > 0.15 \mu s$, are shown under three conditions (see Methods): no restraints, SecY. Coordinates fixed, SecY-fix. Coordinates fixed and the SecY charges (including partial charges) turned off, SecY-vdw. Time evolution of the MSDs and SDACs are shown in the left-hand and right-hand panels, respectively. A and B, SecY. C and D, SecY-fix. E and F, SecY-vdw. The curves are colored according to the legends within the MSD plots describing the position along the membrane normal of the analyzed slab. The black line corresponds to water outside SecY, and the colored lines to water inside SecY. The time dependence of the MSD in the long-time regime was determined from the power function $t^\alpha$ (see Methods). The dashed lines in the SDAC plots correspond to the value $1/e$. 
Figure S8. The effect of acyl chains on the alignment of water dipoles (left-hand panels) and number of acyl chain carbon atoms (right-hand panels) calculated from the NVE simulations. The results are shown for simulations under three conditions (see Methods): no restraints (black symbols), coördinates fixed (green symbols), and coördinates fixed and the SecY charges (including partial charges) turned off (blue symbols). The upper panels (A and B) show the results obtained from the NVE simulations carried out at time t < 0.15 µs, while the lower panels (C and D) show those obtained from the NVE carried out at time t > 0.15 µs.
**Movie 1.** Passage of water molecules through the *Pyrococcus furiosus* SecYE translocon. The movie is derived from the water movement captured in Sim1. The movie shows 1.5 ns of the simulation starting at the 276.8 ns point of Sim1. Shown above is the first frame of the movie.

**Movie 2.** Images of the lipids at the front of the lateral gate of SecYE not interfering with water movement in the vicinity of the hydrophobic ring. The movie shows the structure of SecYE and
the lipids as SecYE is rotated about the z-axis. The movie is essentially a rotating view of Fig. 3B. Shown above is the first frame of the movie.

**Movie 3.** Images of the lipids at the front of the lateral gate of SecYE interfering with water movement in the vicinity of the hydrophobic ring. The movie shows the structure of SecYE and the lipids as SecYE is rotated about the z-axis. The movie is essentially a rotating view of Fig. 3C. Shown above is the first frame of the movie.

**References**


