

The Mechanism of Glycerol Conduction in Aquaglyceroporins

Morten Ø. Jensen,^{1,2} Emad Tajkhorshid,¹ and Klaus Schulten^{1,3}

¹Beckman Institute

University of Illinois at Urbana-Champaign

405 N. Mathews

Urbana, Illinois 61801

²Department of Chemistry

Technical University of Denmark

DTU 207

DK-2800 Lyngby

Denmark

Summary

Background: The *E. coli* glycerol facilitator, GlpF, selectively conducts glycerol and water, excluding ions and charged solutes. The detailed mechanism of the glycerol conduction and its relationship to the characteristic secondary structure of aquaporins and to the NPA motifs in the center of the channel are unknown.

Results: Molecular dynamics simulations of GlpF reveal spontaneous glycerol and water conduction driven, on a nanosecond timescale, by thermal fluctuations. The bidirectional conduction, guided and facilitated by the secondary structure, is characterized by breakage and formation of hydrogen bonds for which water and glycerol compete. The conduction involves only very minor changes in the protein structure, and cooperativity between the GlpF monomers is not evident. The two conserved NPA motifs are strictly linked together by several stable hydrogen bonds and their asparagine side chains form hydrogen bonds with the substrates passing the channel in single file.

Conclusions: A complete conduction of glycerol through the GlpF was deduced from molecular dynamics simulations, and key residues facilitating the conduction were identified. The nonhelical parts of the two half-membrane-spanning segments expose carbonyl groups towards the channel interior, establishing a curve-linear pathway. The conformational stability of the NPA motifs is important in the conduction and critical for selectivity. Water and glycerol compete in a random manner for hydrogen bonding sites in the protein, and their translocations in single file are correlated. The suggested conduction mechanism should apply to the whole family.

Introduction

Aquaporins (AQPs) dissipate the osmotic gradient across the cell membrane by selectively conducting water. AQPs constitute a class of membrane proteins with more than 150 members, including a glycerol-conducting subclass (aquaglyceroporins) [1, 2]. Almost all AQPs

completely exclude passage of ions and charged solutes and thereby preserve the electrochemical potential across the cell membrane; only AQP6 has been reported to be ion conducting at low pH [3]. Two structures of human aquaporin 1 (AQP1) determined by electron crystallography were previously reported [4, 5]. The structures are atomic models derived from 2D crystallography data at 3.8 Å and 3.7 Å in plane and 4.4 Å and 6.0 Å normal resolution, respectively, and both identify AQP1 as a homotetrameric membrane channel [4, 5]. The X-ray structure of the *Escherichia coli* glycerol facilitator, GlpF, a homotetramer highly similar to AQP1, was reported at a resolution of 2.2 Å [6]. The GlpF structure included the positions of three glycerol molecules and two water molecules per monomer, confirming the GlpF selectivity for both substrates. GlpF also stereo- and enantio-selectively conducts other polyalcohols (alditols) [6, 7]. Water and glycerol influxes in reconstituted liposomes are increased by GlpF 10-fold and 100- to 1000-fold, respectively, relative to the pure liposome [2, 6–8]. In terms of water permeability, GlpF is found to be about one order of magnitude less efficient than AqpZ, the *E. coli* variant of human AQP1 [2, 8].

The GlpF monomer has two characteristic half-membrane-spanning repeats that are related by a quasi-two-fold symmetry [6]. Approximately half of the membrane part of the repeat has α -helical structure, whereas the other half has a characteristic nonhelical conformation. The repeats are unique in the sense that the N-terminals of the helical part of these repeats (M3 and M7) meet almost at the center of the channel, an architecture also found in AQP1 [4, 5]. An Asn-Pro-Ala (NPA) motif of each segment is found at this place, the motif being conserved in all AQPs [2, 4–6]. The diameter of the channel opening measures less than 3.5 Å at its narrowest point, and the constriction region of the channel is approximately 28 Å long [6]. Due to the geometric constraints, the substrates can only pass in single file [6]. The polarity of the amino acids in the channel interior is believed to determine the selectivity in GlpF and other AQPs. The hydrophobic backbone of glycerol is oriented toward apolar residues; particularly significant in this respect is the wedging of glycerol between Trp48 and Phe200 in the selectivity filter [6]. The hydroxyl groups of glycerol are, on the other hand, hydrogen bonded to polar residues and to water molecules [6].

While static structural characteristics of AQPs have been revealed [4–6], little is yet known in atomic detail regarding the mechanism and dynamics of the transport in AQPs. Here, we report a molecular dynamics study of GlpF that very surprisingly revealed spontaneous glycerol conduction, the underlying mechanism as well as the role of the conserved secondary structure in the AQP family. The mechanistic details of the glycerol transport indicate a single-file translocation where water competitively replaces glycerol along an array of ex-

³Correspondence: kschulte@ks.uiuc.edu

Key words: aquaporin; membrane protein; water channel; molecular dynamics; hydrogen bond

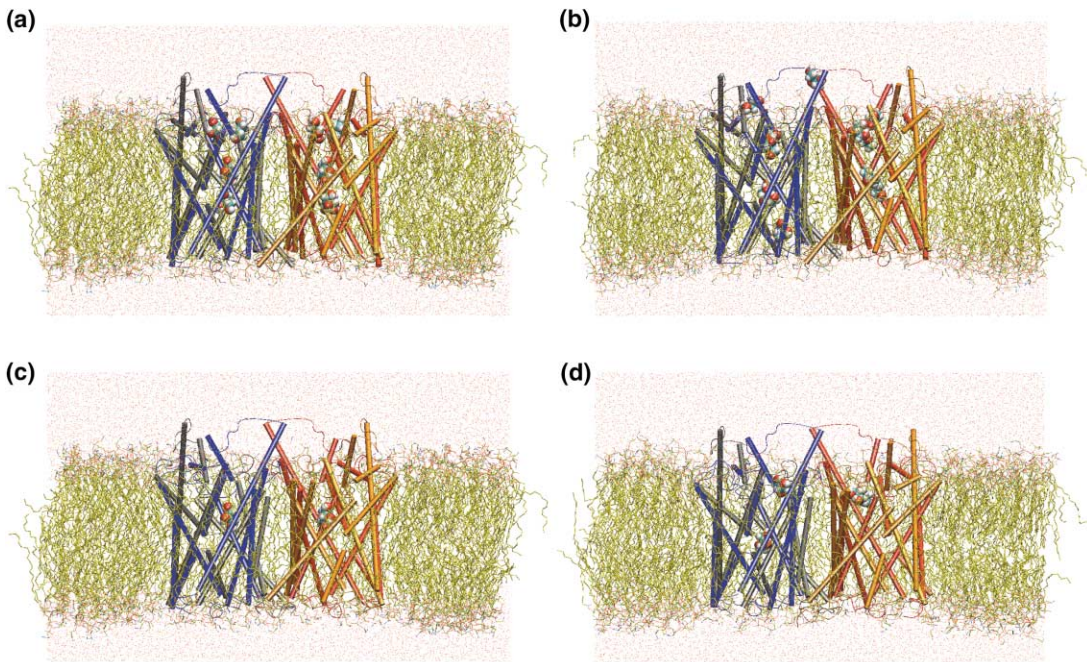


Figure 1. Simulated Systems

System 1 with three glycerol molecules per monomer: (a) glycerols at initial symmetrical positions and (b) glycerols after 1 ns of molecular dynamics. System 2 with one glycerol per monomer: (c) glycerol at initial symmetrical position, and (d) glycerol after 1 ns of molecular dynamics. The direction of periplasm and cytoplasm is upwards and downwards, respectively. The glycerol molecules in system 1 are enumerated G1, G2, and G3 toward the cytoplasm. Glycerol molecules are rendered in van der Waals representation; glycerols from different monomers partly overlap due to choice of perspective. GlpF monomers 1–4 are rendered in blue, red, gray, and orange, respectively. The hydrophobic parts of the lipid molecules are rendered in yellow; N and O atoms are shown in blue and red, respectively. Water molecules are presented as red points. One can recognize the hydrophobic matching of the lipid bilayer and the protein through the decrease of the membrane thickness toward the protein-lipid interface, particularly in (b) and (d).

posed carbonyl groups. Well-defined binding sites inside the channel are formed by the selectivity filter and the NPA motifs.

Results and Discussion

Two systems were simulated, one with three glycerol molecules per monomer and one with a single glycerol molecule per monomer. We denote these systems, shown in Figure 1, as system 1 and system 2, respectively. System 1 was constructed using only information from the observed crystallographic structure (see Experimental Procedures). Due to the high concentration of glycerol (15% v/v) used in the crystallization medium [6], each monomer in the crystal structure contains three glycerol molecules (Figure 1a). However, under physiological glycerol concentrations, the number of glycerol molecules per monomer is smaller. Therefore, we have also carried out simulations on system 2, which has only one glycerol per monomer.

The Protein-Lipid Interface and Hydrophobic Matching

To account for the membrane environment, GlpF was embedded in a (16:0/18:1c9) palmitoyloleoylphosphatidylethanolamine (POPE) bilayer, which is being considered as a good model for the *E. coli* membrane [9]. Although GlpF was crystallized from a surfactant medium [6], anal-

ysis of the resulting structure suggested the position of the protein relative to the lipid membrane [6]. The POPE membrane employed here had an initial vertical separation of 39.6 Å between the P atoms ($d_{p,p}$) of the two lattice monolayers, corresponding to an initial hydrophobic thickness (d_L) of 30.6 Å. After 1 ns of equilibration, the average $d_{p,p}$ value of the lipids located within a 5 Å lateral distance from the protein measured about 38.0 Å, corresponding to $d_L = 28.0$ Å. The laterally distant lipid molecules, which were more than 15 Å away from the protein, became vertically separated with a $d_{p,p}$ value of about 42.0 Å ($d_L = 32.0$ Å). Therefore, compared to the initial configuration ($d_L(t = 0 \text{ ns}) = 30.6$ Å), the membrane became compressed at the protein-lipid interface and expanded significantly distant from the protein with a membrane thickness that is in agreement with the reported value of the bulk POPE membrane of 32.0 Å [9–11].

The described membrane relaxation and hydrophobic matching at the protein-lipid interface are manifested in the membrane curvature clearly discernable in Figures 1b and 1d. Despite the short length scale of the simulated system, the position of GlpF in the membrane is correctly described along with the bilayer hydrophobic thickness, d_L , measured at the box boundaries. Due to hydrophobic dispersion interactions at the protein-lipid interface, d_L decreases about 2.5 Å during the simulation implying $d_L(t = 1 \text{ ns})$ matches the hydrophobic thickness

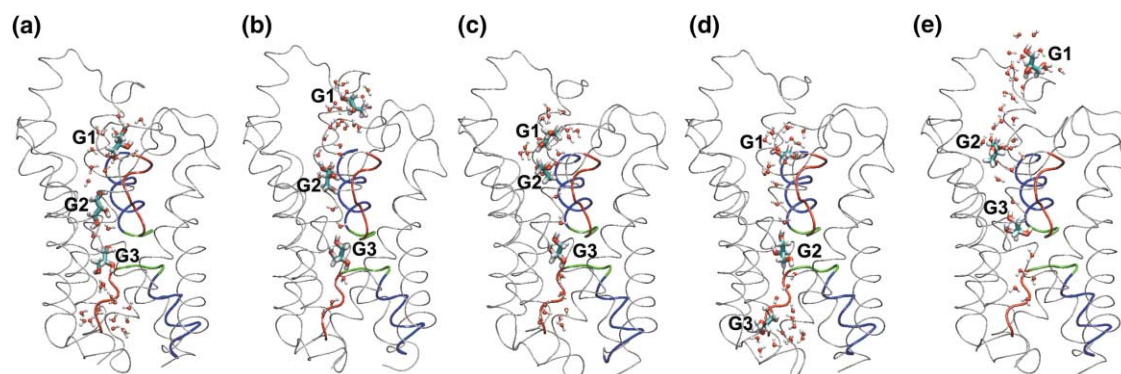


Figure 2. Glycerol Diffusion in Different Monomers after 1 ns

Simulation snapshots of the four monomers in system 1: (a) glycerol molecules located as in the crystal structure at $t = 0$; (b–e) monomers 1–4 at $t = 1$ ns. Water molecules located close to glycerol and/or inside the channel are shown. The two half-membrane-spanning repeats are presented; the nonhelical parts are in red, the NPA motifs are in green, and the α -helical parts are in blue.

of the protein. None of the polar residues at the periplasmic and cytoplasmic faces of the protein are in contact with the apolar part of the lipid membrane. A similar lipid relaxation has been previously observed [10] and is also in accord with theoretical models [11–15].

Spontaneous, Bidirectional Glycerol Conduction

The simulations revealed conduction of glycerol both in system 1 and in system 2. The rather unexpected finding constitutes the main result of our study and permitted us to investigate the conduction mechanism in atomic detail. The conduction can be ascribed to thermal fluctuations arising spontaneously in our simulations. In the constriction region of the channel, centered around the NPA motifs and starting from the position of the selectivity filter (residues Trp48, Phe200, and Arg206), glycerol undergoes purely one-dimensional diffusion along the channel axis (coordinate z). Figure 2 compares the glycerol positions after 1 ns of simulation of system 1 to the initial positions in the GlpF crystal structure. The complete glycerol pathway deduced from the spontaneous diffusion in monomers 1–4 of this system is illustrated in Figure 3. The conduction of glycerol occurred independently in each GlpF monomer (Figure 2). Cooperativity of conformational changes implying coordinated glycerol conduction between the monomers was not discernible. In fact, the glycerol conduction occurred without noticeable changes in secondary, tertiary, and quaternary protein structures. In the vestibules of the channel, rotation of the glycerol molecules was observed. In the constriction part of the channel, rotation along the glycerol long axis was restricted, and the aliphatic backbone of all glycerol molecules was predominantly facing the apolar part of the channel interior in accord with the stereo- and enantio-selectivity of GlpF [6, 7]. This is evident from Figure 3 where the positions of the displaced glycerol molecules are shown relative to the glycerol molecules in the crystal structure.

In simulations of system 1, most glycerol molecules were displaced relative to their initial positions in the crystal structure. Figures 2b–2e display the final glycerol positions in monomers 1–4 of system 1, which can be compared to the initial glycerol positions in the crystal

structure in Figure 2a. The displaced glycerol molecules were found to move randomly and independently in different monomers. In monomers 1 and 4 of system 1, glycerol 1 and 2 (Figures 2b and 2e) diffused toward the periplasm; glycerol 1 moved completely into the aqueous phase. Glycerol 1 was initially (in the crystal structure; see Figure 2a) hydrogen bonded to only a

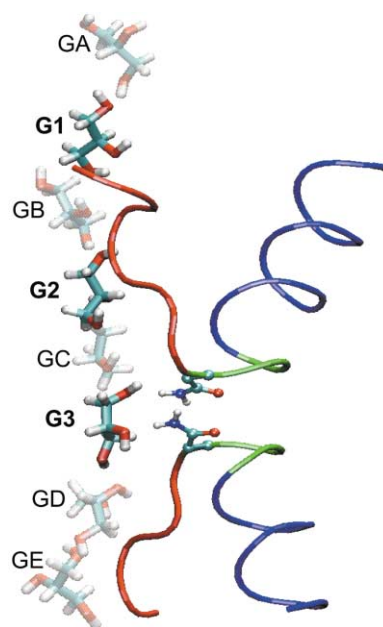


Figure 3. Construction of the Complete Conduction Pathway

Glycerol positions taken from monomer 4 (GA, GB, and GC) and monomer 3 (GD and GE) of system 1 superimposed to illustrate the complete glycerol conduction. The glycerol positions in the crystal structure (G1, G2, and G3) are rendered as solid molecules. Displaced glycerols are shown as transparent molecules. The two half-membrane-spanning repeats are also shown; nonhelical parts (residues 64–67 and 197–202) are in red; NPA motifs (residues 68–70 and 203–205) are in green; and the α -helical parts (residues 71–77 and 206–215) are in blue. Side chains of Asn68 and Asn203 are rendered as CPK models. Residues 64–77 and 197–215 are located in the cytoplasmic (lower) and periplasmic (upper) repeats, respectively.

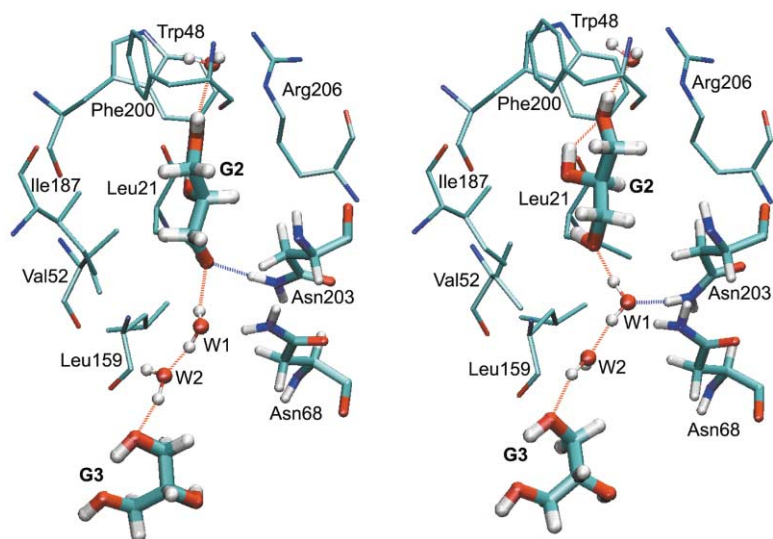


Figure 4. Competition of Water and Glycerol for Hydrogen Bonding Sites

Two snapshots during the upward conduction of glycerol toward the periplasm (system 1), revealing a water molecule substituting glycerol at the NPA motifs from left to right panel. One can recognize the single-file arrangement of the water and glycerol substrates, as well as the orientation of Asn68:H_δ and Asn203:H_δ pointing toward the passing substrates. During the simulation, water molecules W1 and W2 are located between G2 and G3 in the narrow part of the channel. Shown without hydrogen atoms are the residues in the selectivity filter (Trp48, Phe200, and Arg206) as well as the apolar residues opposing the NPA motifs (Leu21, Ile187, Val52, and Leu159).

single residue, Tyr138:O, which is located in the periplasmic vestibule of the channel. The rapid diffusion of this glycerol indicates relatively weak (nonspecific) binding of the substrate to GlpF in the periplasmic vestibule. In contrast, glycerol 3 remained closer to its initial site located in the region of the two NPA motifs (Figure 2b). In the crystal structure, this binding is characterized by two hydrogen bonds between glycerol and Asn203:H_δ as well as Asn68:H_δ of the two NPA motifs and a third hydrogen bond to His66:O [6]. In monomer 2 (Figure 2c), glycerol 1 shifted only slightly, whereas glycerol 2 moved a significant distance toward the periplasm. In monomers 3 and 4 (Figures 2d and 2e), all glycerol molecules diffused toward the cytoplasm and periplasm, respectively. In monomer 4, the NPA region was completely devoid of glycerols at the end of the simulation.

In system 2, with only one glycerol per monomer, three out of four glycerol molecules moved from their initial positions in the selectivity filter toward the periplasmic vestibule. No glycerol molecules moved completely out of the channel. In the fourth monomer, the glycerol molecule moved to the region between the two NPA motifs and remained at this position. It is interesting to note that the narrowest part of the channel is the selectivity filter (Figures 2a and 4). Fu et al. [6] suggested a hydrogen bond between Arg206:H_η and a glycerol oxygen lowers the energy barrier in this region sufficiently to enable glycerol to pass this restrictive site, which is furthermore defined by the wedging of glycerol between Trp48 and Phe200 (Figure 4). In system 2, we observed long-time population at this site, i.e., the glycerol molecules required hundreds of picoseconds (ps) to pass completely. Nevertheless, at the end of the simulation, no glycerols were bound anymore to the selectivity filter.

Binding of glycerol was also found to be strong in the NPA region. In fact, this site remained the most populated one in system 1. The glycerol molecules above the NPA site showed a tendency to move toward the periplasm. Interestingly, in the crystal structure, no glycerol molecules are located closer to the cytoplasm than at the NPA site. However, we observed in monomer 3 of system 1 that glycerol 3 approached the cytosolic

side and remained bound to a site below the NPA region for a significant amount of time (Figure 1d), indicating another site with weak binding. During the simulations, the glycerol molecules were nearly continuously involved in one to three hydrogen bonds with protein atoms and water molecules, though the bonding pattern changed over time. According to our simulations, the energy barriers connected with changes of hydrogen bond patterns can amount only to a few kT , since they were overcome in periods of less than one nanosecond. An exception is the binding to the selectivity filter. At this position, glycerol engages in three or more strong hydrogen bonds with Trp48:H_ε, Phe200:O, Arg206:H_ε, and Arg206:H_η. The energetic cost of passage through the selectivity filter, therefore, must be larger than for the rest of the channel. The energy barriers to be surpassed in order to access the NPA site seem to be also significant. We conclude that at least two stable binding sites are present in the channel.

The observed bidirectional conduction of both water and glycerol is in accord with the physiological function of GlpF, since GlpF dissipates an osmotic gradient and also provides nutrition to the cell [8], though the net uptake of glycerol by the cell is generally positive. Since the simulations were carried out by using periodic boundary conditions, any deviation from bidirectionality must be ascribed to statistical fluctuations.

The Role of Water in Glycerol Conduction

Water molecules were found to be intimate participants in the conduction of glycerol. Both water molecules and glycerol molecules are observed in the constriction part of the channel in single file only (Figures 2, 4, and 5). The simulations revealed that glycerol molecules are conducted through an inherent competition with water for hydrogen bonds to the channel interior. The atomic radial distribution functions $r(O_{\text{GlpF}} - H_{\text{Glycerol}})$, $r(O_{\text{H}_2\text{O}} - H_{\text{Glycerol}})$, and $r(H_{\text{H}_2\text{O}} - O_{\text{Glycerol}})$ all feature their first peaks at 1.9 Å, which is in accord with this observation. By being close to glycerol, one to three water molecules are continuously in a position to substitute glycerol and form hydrogen bonds to the protein (Figures 4 and 5).

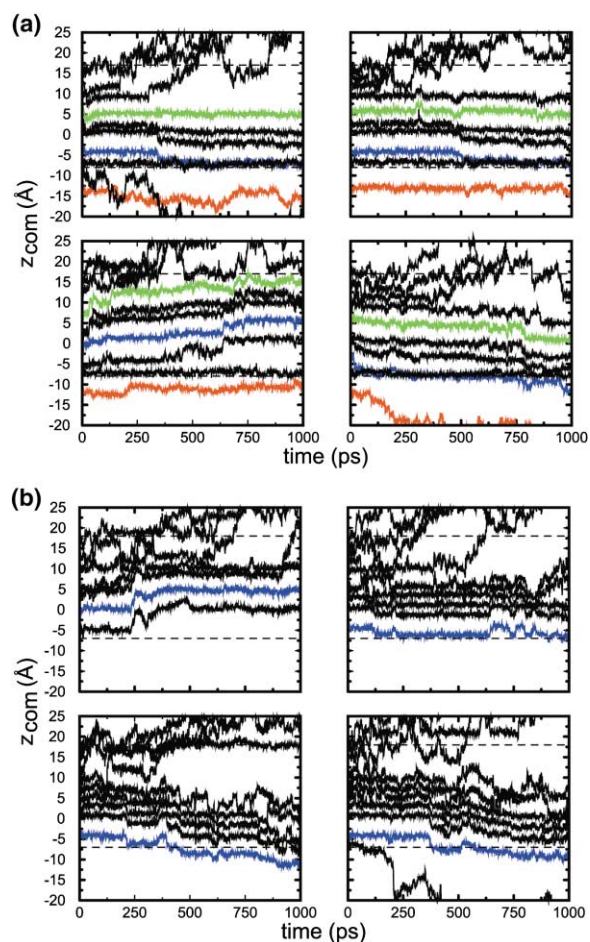


Figure 5. Correlation of Displacement of Glycerol and Water inside the Channel

Time evolution of the z component of the center of mass, z_{com} , of glycerol (green, blue, and red curves) and of water (black curves) along the channel axis in the four monomers of system 1 (a) and system 2 (b). Monomers 1–4 are shown in the top left, top right, bottom left, and bottom right panels, respectively. The constriction region is within $-7 \leq z \leq 18$ Å, (indicated by dashed lines), and the periplasmic outlet is toward negative z values. The origin of the channel axis ($z = 0$) was taken at the position of the oxygen atom of Ala201 at each t ; the selectivity filter is located about $-5.0 < z < -2.5$ Å.

Since one to three hydrogen bonds between the same glycerol and GlpF typically are present simultaneously in the constriction part of the channel, the water molecules need to act in a coordinated fashion when substituting glycerol (Figure 5). A coordinated event is of greater energetic cost and less likely to occur than breakage of a single hydrogen bond in accord with glycerol residing at the selectivity filter and at the NPA motifs for longer times than in the vestibules.

Details in the dynamics of glycerol-water diffusion in single file are revealed by the time evolution of the z component of their center of mass coordinates, z_{com} , which reflects the molecular displacement along the channel axis (Figure 5). Clearly, the time evolution of z_{com} indicates that glycerol diffusion is uncorrelated between the individual monomers, whereas strongly correlated

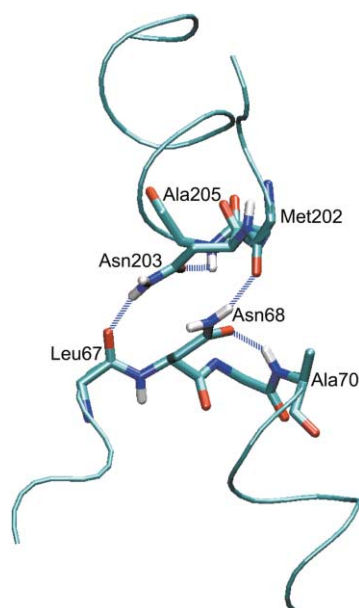


Figure 6. NPA Motifs

Central part of the channel involved in four strong hydrogen bonds between the NPA motifs: Asn68:O $_{\delta}$ /Ala70:HN (lower right), Asn68:H $_{\delta}$ /Met202:O (upper right), Asn203:H $_{\delta}$ /Leu67:O (lower left), and Asn203:HN/Ala205:O (upper left).

displacements occur within the same monomer in the constriction region of the channel, $-7 \leq z \leq 18$ Å. For example, in Figure 5a, glycerols are collectively displaced in monomer 1 at about 350 ps, in monomer 2 at about 500 and 850 ps, in monomer 3 at about 125 and 700 ps, and in monomer 4 at about 200 and 800 ps. The translocation occurs with discrete steps of about 2.5 Å, reflecting the exchange of one set of hydrogen bonds for another set. The persistence time of these hydrogen bonds is seen to vary significantly from tens to hundreds of picoseconds. The glycerol diffusion in the vestibules

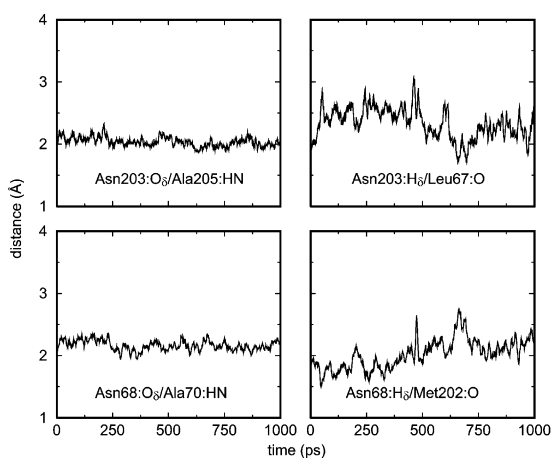


Figure 7. Stability of the Hydrogen Bonds between the NPA Motifs
The distances of the four hydrogen bonds in Figure 6 are shown along the trajectory (monomer 1, system 2). Similar results were found for the other three monomers in system 2 as well as for all monomers in system 1.

Table 1. Hydrogen Bonding Sites Participating in Conduction

Acceptors			Donors		
Asp130	O _δ	-18.6 ± 0.8	Lys33	H _ζ	-15.7 ± 0.6
His142	N _δ	-18.0 ± 0.4	Asn140	H _δ	-14.3 ± 0.6
Thr167	O _γ	-16.9 ± 0.3	Asn140	HN	-11.9 ± 0.3
His142	N	-16.6 ± 0.3	Tyr138	HN	-9.8 ± 0.3
Asn140	O _δ	-15.6 ± 0.4	Gln41	H _ε	-8.1 ± 0.4
Asn140	N _δ	-14.9 ± 0.5	Gly199	HN	-6.8 ± 0.5
Pro139	O	-13.6 ± 0.3	Arg206	H _η	-4.9 ± 0.4
Pro196	O	-11.8 ± 0.3	Arg206	H _ε	-2.9 ± 0.4
Pro139	N	-11.7 ± 0.3	Trp48	H _ε	-0.1 ± 0.4
Gly133	O	-10.9 ± 0.5	Asn203	HN	1.4 ± 0.2
Leu197	O	-10.4 ± 0.3	Asn203	H _δ	4.5 ± 0.4
Gln41	O _ε	-10.1 ± 0.4	Asn68	H _δ	6.4 ± 0.2
Thr198	O	-9.0 ± 0.5	His66	H _δ	15.8 ± 0.3
Ser136	O	-8.8 ± 0.3	His142	H _δ	17.3 ± 0.4
Gly195	O	-8.5 ± 0.4			
Gln41	N _ε	-8.0 ± 0.4			
Gly199	O	-5.6 ± 0.3			
Trp48	N _ε	-4.4 ± 0.3			
Phe200	O	-2.6 ± 0.4			
Ala201	O	0.0 ± 0.0			
Trp48	O	1.6 ± 0.4			
Asn203	N _δ	4.5 ± 0.3			
Asn68	N _δ	6.1 ± 0.2			
Leu67	O	7.2 ± 0.3			
His66	O	10.7 ± 0.3			
Tyr138	O	10.9 ± 0.5			
Ala65	O	13.6 ± 0.4			
His66	N _δ	15.1 ± 0.4			
Gly64	O	16.5 ± 0.4			

Hydrogen donors and acceptors in GlpF participating in hydrogen bonding to glycerol and water, arranged according to their position along the channel axis, z. Shown in bold are the Asn residues of the NPA motifs, as well as the Trp48, Phe200, and Arg206 residues of the selectivity filter. Average atomic positions and fluctuations (in Å) along z are listed for monomer 1 of system 1, which is representative for both systems. The average atomic positions are relative to the position of the oxygen atom of Ala201, since this residue is subject to minor fluctuations only (see also Figure 7).

(and in bulk) is clearly not correlated with glycerol translocation in the constriction region, where water and glycerol displacements indeed are correlated, in particular, for those water molecules that are located near to (or between) the glycerol molecules. Again, this reflects the competitive nature of the mechanism translocating the glycerol and water in single file along the channel. In addition, one can recognize water-water correlated displacements in Figure 5b, illustrating single-file translocation of the water molecules in the constriction region of the channel.

The glycerol-water competition for hydrogen bonds with GlpF, accordingly, is a strict necessity for glycerol conductance. Without this competition, glycerol would not move along the channel, and conduction would not occur. Water and glycerol are not cotransported to fulfill stoichiometric balances; rather, the conduction mechanism is defined by the intrinsic, random competition for the same hydrogen bonds.

The GlpF Selectivity Filter

Presence of a high-affinity binding site near one outlet of the channel has been reported for the sugar-conducting maltoporin, LamB [16–18]. The physiological rationale for such a site is to extract the metabolite from the environment, where it is present in low concentration. The presence of a so-called greasy slide [16, 17] in GlpF, realized through the selectivity filter and the hydropho-

bic lining of one side of the channel interior, can explain why the protein conducts glycerol, in contrast to, e.g., human AQP1 [4, 5]. In fact, during the simulations, the respective parts of the glycerol and the protein predominantly faced each other (Figure 3). This would not be the case in AQP1 where the hydrophobic lining of the channel interior is “contaminated” by hydrophilic residues, which interact unfavorably with the aliphatic backbone of glycerol. Furthermore, Trp48 and Phe200 of the GlpF selectivity filter are replaced by Phe and His residues, respectively, in AQP1. These side chains in AQP1 [4, 5] do not attract glycerol from a low-concentration environment by means of strong hydrophobic dispersion interactions, since the glycerol backbone cannot be wedged between these rings as it can be in GlpF. Initial binding of glycerol, therefore, can not be accomplished in AQP1, thereby disfavoring glycerol transport.

The NPA Motifs

The two opposing NPA motifs are conserved in the AQP family [2]. Simulations showed that these motifs consistently participate in the conduction process via formation of hydrogen bonds to the substrate. In the crystal structure, Asn68 and Asn203 are confined by four hydrogen bonds: Asn68:O_δ/Ala70:HN, Asn68:H_δ/Met202:O, Asn203:H_γ/Leu67:O, and Asn203:O_δ/Ala205:HN [6]. These hydrogen bonds keep the two half-membrane-spanning

repeats close together in a conformation where the amide groups of Asn68 and Asn203 side chains are oriented almost parallel to each other and parallel to the membrane plane (Figure 6; see also Figure 3); these hydrogen bonds also ensure that in each of the two Asn residues, Asn68 and Asn203, one of the two H_δs is facing the channel and, hence, is available as a hydrogen bond donor to glycerol and water. As shown in Figure 7 and in contrast to the results from molecular dynamics simulations of AQP1 [19], based on the published structure [4], the highly restricted conformation of the NPA motifs is preserved during the simulation, implying that the H_δs are permanently available with their hydrogen bond donor capacity. A structure of AQP1 in a revised form [20] also exhibited stable NPA motifs.

Being lined up in single file after passing a constriction point 8 Å above the quasi-two-fold axis, all substrates entering from the periplasmic side consequently become hydrogen bonded to the two exposed H_δs of Asn68 and Asn203. This is the case also when the substrates pass the NPA motifs from the cytoplasmic side, since a single file is formed here as well (Figures 4 and 5). No hydrogen bond can be formed to the nonpolar residues opposing the NPA motifs. Mutations to polar residues at this position could possibly form hydrogen bonds with the substrate, which might then alter the selectivity of the channel.

A Curve-Linear Conduction Pathway

The lifetime of the hydrogen bonds between specific residues and glycerol vary from tens to hundreds of picoseconds (Figure 5). For system 1, we compiled all GlpF hydrogen donor/acceptor atoms that were found during the simulation within hydrogen bonding distance (≤ 3.0 Å) to the donor/acceptor atoms of glycerol (Table 1). Strikingly, Table 1 reveals several backbone oxygen atoms involved in the conduction. These atoms are mainly located in the nonhelical parts of the two conserved half-membrane-spanning repeats. Together, these atoms trace out a curve-linear (spiral) pathway inside the channel along which glycerol and water molecules are conducted (Figure 8, see also Figure 3). Adoption of this characteristic, nonhelical secondary structure over approximately half of each half-membrane-spanning repeat effectively provides a set of permanent hydrogen acceptors subject to only minor fluctuations along the channel axis. In fact, atomic fluctuations along the channel axis in general are small. One may recognize that the atomic fluctuations are of equal magnitude in the periplasmic ($z < 5.3$ Å) and cytoplasmic ($z > 5.3$ Å) parts of the channel, in contrast to the fluctuations observed for, e.g., the KcsA channel [21].

This critical part of the secondary structure is clearly left unaltered during the simulation, in accord with our observation that the glycerol conduction does not involve significant protein conformational changes (see below). The nonhelical parts of the half-membrane-spanning repeats are fixed via two Glu residues (14 and 152) that each form two hydrogen bonds to two backbone atoms in the repeats; Glu152:O ϵ forms hydrogen bonds to Ala201:HN and Met202:HN in the periplasmic repeat, and Glu14 forms hydrogen bonds to His66:HN

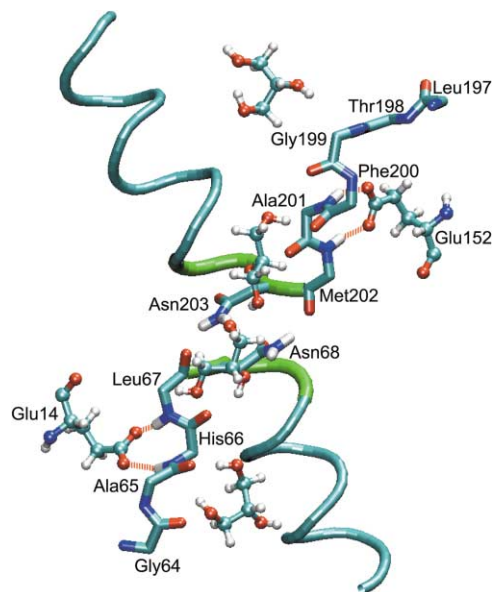


Figure 8. Secondary Structure and Glycerol Conduction

The curve-linear (spiral) pathway of glycerol conduction as established by backbone oxygen atoms of residues 64–66 and 199–201 in lower and upper half-membrane-spanning repeats, respectively. Carbonyl oxygens are shown in red. Leu67:O and Met202:O are seen to stabilize the conformation of the NPA motifs (see Figure 6 and Table 1). This peculiar secondary structure of the repeats is stabilized by hydrogen bonds from Glu14:O ϵ atoms to His 66:HN and Leu67:HN, and from Glu152:O ϵ atoms to Ala201:HN and Met202:HN. Also shown are the NPA motifs (green), the α -helical parts of the half-membrane-spanning repeats (blue; residues 71–78 and 206–214), and the side-chains of Asn68 and Asn203. Snapshots of four glycerol positions from the simulation of system 1 are included to indicate the glycerol pathway. Compared to the glycerol-protein interactions in the constriction region of the channel (between Gly64 and Gly199), one can recognize that outside this region the interactions are less specific (as illustrated by the top glycerol).

and Leu67:HN in the cytoplasmic repeat. Hydrogen bonds between glycerol oxygen atoms and protein are predominant in the region of the NPA motifs (Asn68:H_δ and Asn203:H_δ) as well as at the selectivity filter (Arg206:H ϵ and Arg206:H η). Table 1 reveals that, apart from those hydrogens, relatively few side chain atoms come within hydrogen bonding distance to glycerol.

The KcsA potassium channel has a secondary structure in its cytoplasmic part that is similar to the channel in GlpF; four half-membrane-spanning repeats together form a hydrogen bonding funnel where the hydrated ion enters [21–23]. The water molecules around the ion can form hydrogen bonds with the exposed donor/acceptor atoms along the funnel and gradually become stripped off before the ion is delivered to the constriction point of the channel at the center of the membrane [22].

Conformational Responses to Glycerol Passage

The conformational response of GlpF to glycerol conduction can be directly examined by comparing the root-mean-square deviation (rmsd) between the crystal and simulated structures (Figure 9). The rmsds of the C α atoms of the tetramer (Figure 9a, I and II) is just above 1.0 Å in both systems, which is 0.5–2.0 Å lower than the

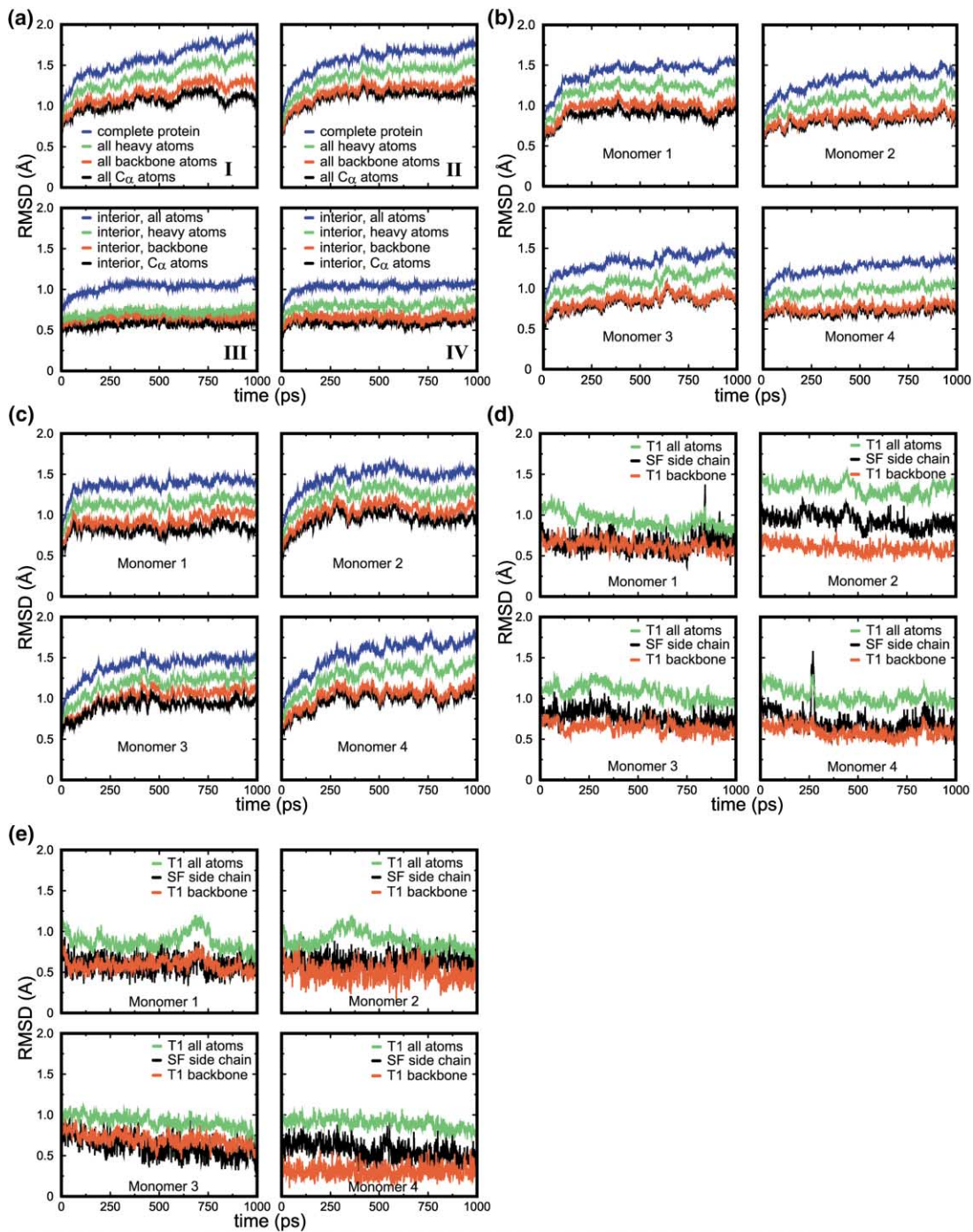


Figure 9. Conformational Stability of GlpF

(a) Root-mean-square deviation (rmsd) of the GlpF tetramer in system 1 (I), system 2 (II), channel interior of system 1 (III), and channel interior of system 2 (IV). Rmsd of the GlpF monomers in system 1 (b) and system 2 (c). Color-coding of the individual entries is as in (a). Rmsd of the GlpF residues in Table 1 (T1) and rmsd of the selectivity filter (SF) only are shown for system 1 (d) and system 2 (e). In determining the rmsd values, the structures along the trajectories were least squares fitted to the initial (X-ray) structure.

rmsd typically observed in molecular dynamics simulations of proteins. Even after the inclusion of side chains, the rmsds are remarkably low, indicating minor side chain movements. For the residues lining the channel interior along the constriction region, the rmsds are even lower (Figure 9a, III and IV), again reflecting that only

very slight conformational changes are accompanying glycerol conduction. In the monomers, the rmsds of C_{α} atoms and those of backbone atoms hardly exceed 1.0 Å (Figures 9b and 9c). These rmsds are clearly uncorrelated, in accord with the absence of cooperativity between monomers with regard to glycerol conduction

(Figure 5). The differences between rmsd of the tetramer and the monomers are small, e.g., less than 0.2 Å for C_α atoms. Focusing on the residues listed in Table 1, one finds slight conformational changes; for example, the conformational changes of the selectivity filter region is comparable to the average conformational change of the other atoms involved in the glycerol conduction (Figures 9d and 9e).

In system 2, glycerol is initially located in the selectivity filter (Figure 1); in monomer 1, glycerol diffuses during the simulation out of the filter toward the cytoplasm, whereas in monomers 2, 3, and 4, it diffuses toward the periplasm (Figure 5). By combining the individual displacements of glycerol in monomer 1 and in one of the other monomers, a complete glycerol passage through the selectivity filter can be reconstructed. At the beginning and at the end of the passage events, glycerol has neither hydrogen bonding nor dispersion interactions with the filter. We monitored the average and rmsd values of the dihedral angle $\angle C_{\alpha}-C_{\beta}-C_{\gamma}-C_{\delta}$ capturing the conformational changes of the side chains of Trp48 and Phe200 during this glycerol passage; the values are $98.6^{\circ} \pm 13.0^{\circ}$ and $-74.8^{\circ} \pm 9.0^{\circ}$ for Trp48 and Phe200, respectively. Recalling that Phe200 exposes its carbonyl oxygen as part of the glycerol-conducting pathway (Table 1), whereas Trp48 is donated from the adjacent transmembrane helix (M2), it is not surprising that the side chain of Phe200 is subject to less fluctuation. Again, these conformational changes are modest relative to the crystal structure (103.5° and -59.6°).

Altogether, these findings consistently indicate minor conformational changes in the secondary, tertiary, and quaternary structures in clear accord with the transmembrane helices being packed in an almost optimal manner [6]. A slight tilt of the outer transmembrane helix M4 [6] was visually recognized in both systems, which most likely is due to a slight relaxation from the crystal structure due to altered packing constraints and the altered environment.

Tetrameric Form and Ion Conductance through Central Pore

It has been speculated that the central pore is an ion or water channel [2, 6]. However, no widening of the central hydrophobic pore formed by the four monomers in the tetrameric structure was observed in the simulations. Furthermore, no water molecules were found to diffuse through the pore. It is unlikely that any charged substrate or ion could pass this pore without a significant conformational change that has not been observed. It appears more likely that metal ion(s), which may be present in the central pore, facilitate assembly of the tetramer [8]; in fact, the presence of Mg²⁺ was found to increase the fraction of tetrameric GlpF, but the ion seems not to be required for conformational stability.

It can be recognized that the tetrameric form of GlpF provides a funnel-like shape exposed toward the periplasm that may help capture and funnel glycerol into the channels. Regarding the lower concentration of glycerol in the periplasm as compared to water, such a funnel mechanism might be of fundamental significance. AQP1 does not exhibit pronounced funnel characteris-

tics that would be in keeping with the fact that water is present in high concentration, and a funnel is therefore not needed. We also note that in the tetrameric form several carbonyl oxygen atoms are exposed toward the periplasm, increasing the GlpF-glycerol contact area.

Biological Implications

We have reported the first molecular dynamics investigation on the high-resolution X-ray structure of the *E. coli* glycerol facilitator GlpF. The results provide fundamental insight into the mechanism of glycerol and water conduction in aquaporins. The conduction mechanism identified explains the peculiar secondary structure of GlpF that exposes backbone carbonyl oxygen atoms to the interior of the channel, defining a curve-linear conduction pathway. Considering the conservation of this secondary structure in the aquaporin family, we suggest that the conduction mechanism reported here applies to the whole aquaporin family.

The mechanism of glycerol conduction in GlpF was deduced from spontaneous displacement of glycerol molecules, which was found to occur due to thermal fluctuations on the nanosecond timescale. Glycerol is conducted by a mechanism driven by an inherent competition between water and glycerol for hydrogen bonds with specific residues in GlpF. Accordingly, water is strictly required for glycerol transport at physiological glycerol concentrations, but there is no reason that water should be stoichiometrically cotransported; rather, the presence of water is part of the conduction mechanism in a stochastic manner.

The conduction takes place amid only very minor conformational changes of the protein. The two conserved NPA motifs of GlpF are consistently hydrogen bonded, and no substrate passes the motifs without hydrogen bonding to the asparagine side chains of these motifs. Any mutation that would alter the conformation of the NPA motifs or compete for hydrogen bonds may affect the profound selectivity of the channel. This observation might be relevant in regard to clinical disorders known to originate from defective AQPs [2], as well as in regard to further investigations of the AQP selectivity through point mutation studies.

Experimental Procedures

System 1: The Full System

The crystal structure of monomeric GlpF [6] was obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) [24], PDB entry code 1FX8. Atoms of Arg257 absent from the X-ray structure [6] were added with Insight II [25]. Hydrogen atoms were added to the monomer with XPLOR [26]. The tetrameric structure of GlpF was generated with VMD [27] by using the transformation matrices provided in the PDB file. The symmetry-related atoms of occupancy 0.25, i.e., atom numbers 300, 301, 302, and 303, all being oxygen atoms of water molecules, were not included in the transformation but were added after the tetramer was constructed. These atoms are located at the position where two Mg²⁺ might be present [6]; however, since no ions are deposited with the entry 1FX8, we did not include Mg²⁺ in the simulations.

Additional water molecules buried inside the tetramer but not present in the resolved structure were introduced by iteratively invoking the program DOWSER [28] until no more water molecules could be inserted. In this procedure, the interaction energy between the water molecules introduced by DOWSER and the rest of the

system is evaluated. According to an energy threshold of -12 kcal/mol, it was determined whether the introduced water molecule was retained or not. Following this procedure, 87 water molecules were added to the tetramer, whereas only a few were added inside the channel.

For the lipid membrane, POPE lipids were used, since this lipid models well the *E. coli* membrane [9]. Initially, 480 POPE lipids, all in *cis* conformation (16:0/18:1c9), were placed on a regular hexagonal lattice with a lattice vector of 8.1 Å. The membrane plane was taken parallel to the *xy* plane, the membrane normal being along the *z* direction. To position GlpF into the membrane, the tetramer, including resolved and added water molecules, and the membrane were centered at the Cartesian origin. The location of tyrosine residues was used to adjust the normal position of the membrane relative to GlpF. The thickness of the membrane, measured by the vertical P-P separation ($d_{p,p}$) between the two monolayers, was separately adjusted according to the proposed midmembrane residues (Thr18, Thr156, Gly49, Asn68, Pro69, Ala70, Gly96, Gly184, and Gly243) [6]. From this adjustment, equal distances from the lipid head groups to the above residues resulted, ensuring that the proposed hydrophobic part of the protein [6] was embedded by the hydrophobic part of the membrane only, with $d_{p,p}$ initially measuring 39.6 Å. Upon visual inspection in VMD [27], lipids overlapping with the protein were removed; 317 lipids were retained.

The program SOLVATE [29] was subsequently used to add a 20 Å ellipsoidal solvation shell around the membrane and the outer, solvent-exposed parts of GlpF. From this system, a rectangular periodic box, based on the positions of the nitrogen atoms of the ethanolamine head groups of the lipids, was cut out. In the resulting periodic box, all water molecules that were located inside the hydrophobic part of the membrane were removed.

The initial dimensions of the full system were 122.0 Å, 112.7 Å, and 77.0 Å in the *x*, *y*, and *z* directions, respectively. The corresponding dimensions of the pure GlpF tetramer are 73.4 Å, 73.4 Å, and 57.0 Å. The normal dimension of the full system ensures approximately 10 Å hydration of each POPE monolayer. Finally, four water molecules were replaced by chloride ions to bring the system to neutrality. The resulting total system size was 106,261 atoms: 15,356 GlpF atoms, 12 glycerol molecules (168 atoms), 317 POPE lipids (39,625 atoms), 17,036 water molecules (51,108 atoms), and 4 chloride ions.

For the molecular dynamics simulations, the NAMD2 program [30] and the CHARMM 27 parameter set [31, 32] were used. Parameters for glycerol were taken from the same parameter set. Initially, the system was minimized for 1700 conjugate gradient steps, while keeping the protein, glycerol, and crystal water molecules fixed. While still keeping this part fixed, the minimized system was equilibrated for 200 ps at 310 K (physiological temperature and above the gel-liquid crystalline phase transition temperature of POPE) and at a constant pressure of 1 atm (NPT), with a time step of 1 fs and the Langevin piston method [33] with a damping coefficient of 5 ps $^{-1}$ and a piston period of 100 fs. After the 200 ps equilibration period, the system was subjected to 200 conjugate gradient minimization steps now releasing all fixed atoms. The full system was then equilibrated under the same conditions as above for 1 ns, saving coordinates at every 0.25 ps. The particle mesh Ewald method (PME) was used for the computation of the electrostatics without cutoff [34]. The grid spacing was kept below 1.0 Å, and a fourth order spline was used for the interpolation, the long-range part of the electrostatics being evaluated every fourth fs. The van der Waals interactions were cut off at 12 Å by using a switching function starting at 10 Å. Full periodic boundary conditions were imposed.

System 2: One Glycerol per Monomer

A system retaining only the middle glycerol molecule of each monomer was prepared by removing the two outer glycerol molecules of the crystal structure. The program DOWSER [28] was then used to replace the removed glycerol molecules by water. The crystal structure of a single monomer and the crystal water were taken into account in these calculations. Two water molecules replaced the glycerol molecule nearest the periplasm, and one water molecule replaced the glycerol nearest the cytoplasm. All three water molecules were located so close to the two removed glycerol molecules that they substituted the hydrogen bonds formerly formed between protein and glycerol.

System 2 was minimized for 650 steps and equilibrated for 100 ps while fixing the protein and glycerol. The configuration of lipid and bulk water, equilibrated earlier over 200 ps (in system 1 for fixed protein, crystal water, and glycerol, see above), was used to constitute the total system. After 100 ps of equilibration, the protein and glycerol molecules were released, and 700 steps of minimization were done prior to 1 ns of NPT simulation using the protocol outlined above.

Molecular dynamics simulations were carried out on 128 processors (Compaq Alpha ES40, four processors per node) of the Tera-scale Computing System (TCS1) at the Pittsburgh Supercomputing Center, and on a local 32 processor Linux (Athlon) cluster. The time required for 1 ns simulation was 3 and 10 days on these platforms, respectively.

Analysis

The protein-substrate atomic radial distribution functions were calculated by using the exposed carbonyl atoms of the residues 64–67 and 198–202 (Table 1) along with glycerol and water. Because the spatial arrangement of the substrate molecules is linear (single file, Figure 5), no radial normalization was invoked. The functions were averaged over the last 500 ps of the simulation, and one representative monomer was chosen.

To monitor the translocation of glycerol and water through the channel, the time evolution of the *z* component of the center of mass, z_{com} , of glycerol and water, defined relative to the position of the oxygen atom of Ala201, was monitored. Only those water molecules initially located within $-7 < z < 18$ Å were included.

GlpF residues that could hydrogen bond to glycerol and water during conduction were identified by calculating the distances between the oxygen/hydrogen atoms in the glycerol hydroxyl groups and all potential donors/acceptors in each monomer of GlpF, at every 0.25 ps along the 1 ns trajectory of system 1. A distance of 3.0 Å was considered as the cut-off length for hydrogen bonds. The complementary information from monomers 1–4 was combined to compile, in Table 1, all hydrogen donors and acceptors in the conduction pathway in GlpF (Figure 3).

Atomic fluctuations along the channel axis (coordinate *z*, Table 1) were calculated using the *z* position of the carbonyl oxygen atom of Ala201 as origin of the *z* axis. Being close to the NPA motif, we found this atom subject to very small fluctuations along *z*. Atomic positions and fluctuations were averaged over the last 500 ps of the simulation of system 1, which is representative for both systems.

The root-mean-square deviation (rmsd) was calculated along the trajectories of all systems using the crystal structure of the GlpF tetramer or individual monomers for the alignment. To calculate the rmsd of the atoms in the channel interior, atoms within a 30 Å long region along the *z* axis centered around the NPA motifs, and those within 13 Å intervals along the *x* and *y* axes of each monomer centered around the channel axis were taken into account.

Acknowledgments

This work was supported by the National Institutes of Health grants PHS 5 P41 RR05969 and R01 GM60946. The authors acknowledge the computer time provided by the grant NRAC:MCA93S028. M.O.J. and E.T. acknowledge financial support from the Danish Natural Science Council and the Human Frontier Science Program Organization. The molecular images in this paper were created with the molecular graphics program VMD [27].

Received: June 7, 2001

Revised: September 11, 2001

Accepted: September 25, 2001

References

1. Preston, G.M., Piazzza-Carroll, P., Guggino, W.B., and Agre, P. (1992). Appearance of water channels in *xenopus oocytes* expressing red cell chip28 water channel. *Science* 256, 385–387.
2. Borgnia, M., Nielsen, S., Engel, A., and Agre, P. (1999). Cellular and molecular biology of the aquaporin water channels. *Annu. Rev. Biochem.* 68, 425–458.

3. Yasui, M., Kwon, T.H., Knepper, M.A., Nielsen, S., and Agre, P. (1999). Aquaporin-6: an intracellular vesicle water channel protein in renal epithelia. *Proc. Natl. Acad. Sci. USA* 96, 5808–5913.
4. Murata, K., et al., and Fujiyoshi, Y. (2000). Structural determinants of water permeation through aquaporin-1. *Nature* 407, 599–605.
5. Ren, G., Reddy, V.S., Cheng, A., and Mitra, A.K. (2001). Visualization of a water-selective pore by electron crystallography in vitreous ice. *Proc. Natl. Acad. Sci. USA* 98, 1398–1403.
6. Fu, D., et al., and Stroud, R.M. (2000). Structure of a glycerol conducting channel and the basis for its selectivity. *Science* 290, 481–486.
7. Heller, K.B., Lin, E.C., and Wilson, T.H. (1980). Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. *J. Bacteriol.* 144, 274–278.
8. Borgnia, M.J., and Agre, P. (2001). Reconstitution and functional comparison of purified GlpF and AqpZ, the glycerol and water channels from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 98, 2888–2893.
9. Tieleman, D.P., and Berendsen, H.J.C. (1998). A molecular dynamics study of the pores formed by *Escherichia coli* OmpF Porin in a fully hydrated palmitoyl-oleoyl-phosphatidylcholine bilayer. *Biophys. J.* 74, 2786–2801.
10. Tieleman, D.P., Forrest, L.R., Sansom, M.S.P., and Berendsen, H.J.C. (1999). Lipid properties and the orientation of aromatic residues in OmpF, influenza and alamethicin systems: molecular dynamics simulations. *Biochem.* 37, 17554–17561.
11. Dumas, F., Lebrun, M.C., and Tocanne, J.-F. (1999). Is the protein/lipid hydrophobic matching principle relevant to membrane organization and functions? *FEBS Lett.* 458, 271–277.
12. Mouritsen, O.G., and Bloom, M. (1984). Mattress model of lipid protein interactions in membranes. *Biophys. J.* 46, 141–153.
13. Mouritsen, O.G., and Bloom, M. (1993). Models of lipid protein interactions in membranes. *Annu. Rev. Biophys. Biomol. Struct.* 22, 145–171.
14. Mouritsen, O.G., and Sperotto, M. (1993). Thermodynamics of lipid-protein interactions in lipid membranes. In *The Hydrophobic Matching Condition*, chapter 4, M. Jackson, ed. (Boca Raton, FL: CRC Press, Inc.).
15. Mouritsen, O.G., et al., and Zukermann, M.J. (1995). The computer as a laboratory for the physical chemistry of membranes. *Biophys. Chem.* 55, 55–66.
16. Schirmer, T., Keller, T.A., Wang, Y.-F., and Rosenbusch, J.P. (1995). Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution. *Science* 267, 512–514.
17. Dutzler, R., Wang, Y.-F., Rizkallah, P.J., Rosenbusch, J.P., and Schirmer, T. (1996). Crystal structures of various maltooligosaccharides bound to maltoporin reveal a specific sugar translocation pathway. *Structure* 6, 127–134.
18. Hilty, C., and Winterhalter, M. (2001). Facilitated substrate transport through membrane protein. *Phys. Rev. Lett.* 24, 5624–5627.
19. Zhu, F., Tajkhorshid, E., and Schulten, K. (2001). Molecular dynamics study of aquaporin-1 water channel in a lipid bilayer. *FEBS Lett.* 504, 212–218.
20. de Groot, B.L., Engel, A., and Grubmüller, H. (2001). A refined structure of human aquaporin-1. *FEBS Lett.* 504, 206–211.
21. Bernèche, S., and Roux, B. (2000). Molecular dynamics of the KcsA K⁺ channel in a bilayer membrane. *Biophys. J.* 78, 2900–2917.
22. Doyle, D.A., et al., and MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity. *Science* 280, 69–77.
23. Nollert, P., Harries, W.E.C., Fu, D., Miercke, L.J.W., and Stroud, R.M. (2001). Atomic structure of a glycerol channel and implications for substrate permeation in aqua(glycerol)porins. *FEBS Lett.* 504, 112–117.
24. Bernstein, F.C., et al., and Tasumi, M. (1977). The protein data bank: a computer based archival file for macromolecular structures. *J. Mol. Biol.* 112, 535–542.
25. Molecular Simulations (1998). *Insight II, Version 98.0—Molecular Modeling System*. (San Diego, CA: Molecular Simulations, Inc.).
26. Brünger, A.T. (1992). *XPLOR, Version 3.1: a system for x-ray crystallography and NMR*. (New Haven, CT: The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University).
27. Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD—visual molecular dynamics. *J. Mol. Graph.* 14, 33–38.
28. Hermans, J., Xia, X., and Cavanaugh, D. (1998). DOWSER. Department of Biochemistry and Biophysics School of Medicine University of North Carolina Chapel Hill, NC 27599–7260. (<http://femto.med.unc.edu/DOWSER/>).
29. Grubmüller, H. (1996). SOLVATE v. 1.0. Theoretical Biophysics Group, Institut für Medizinische Optik, Ludwig-Maximilians-Universität München, München, Germany. (<http://www.mpibpc.gwdg.de/abteilungen/071/solvate/docu.html>).
30. Kale, L., et al., and Schulten, K. (1999). NAMD2: greater scalability for parallel molecular dynamics. *J. Comp. Phys.* 151, 283–312.
31. Schlenkerich, M., Brickmann, J., MacKerell, A.D., Jr., and Karplus, M. (1996). Empirical potential energy function for phospholipids: criteria for parameter optimization and applications. In *Biological Membranes: A Molecular Perspective from Computation and Experiment*, K.M. Merz, and B. Roux, eds. (Boston: Birkhäuser), pp. 31–81.
32. MacKerell, A.D., Jr., et al., and Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102, 3586–3616.
33. Feller, S.E., Zhang, Y.H., Pastor, R.W., and Brooks, B.R. (1995). Constant pressure molecular dynamics simulation—the Langevin piston method. *J. Chem. Phys.* 103, 4613–4621.
34. Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an *Nlog(N)* method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092.