Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facilitator GlpF

Deyu Lu, Paul Grayson, and Klaus Schulten
Department of Physics and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois

ABSTRACT The aquaglyceroporin GlpF is a transmembrane channel of *Escherichia coli* that facilitates the uptake of glycerol by the cell. Its high glycerol uptake rate is crucial for the cell to survive in very low glycerol concentrations. Although GlpF allows both influx and outflux of glycerol, its structure, similar to the structure of maltoporin, exhibits a significant degree of asymmetry. The potential of mean force characterizing glycerol in the channel shows a corresponding asymmetry with an attractive vestibule only at the periplasmic side. In this study, we analyze the potential of mean force, showing that a simplified six-step model captures the kinetics and yields a glycerol conduction rate that agrees well with observation. The vestibule improves the conduction rate by 40% and 75% at 10-μM and 10-mM periplasmic glycerol concentrations, respectively. In addition, neither the conduction rate nor the conduction probability for a single glycerol (efficiency) depends on the orientation of GlpF. GlpF appears to conduct equally well in both directions under physiological conditions.

INTRODUCTION

Escherichia coli, like all bacteria, has evolved mechanisms that allow it to survive in the most nutrient-poor conditions. One such mechanism is a very efficient glycerol uptake system that allows rapid growth in glycerol solution at a concentration of 5 μM (Mitscherlich and Marth, 1984) or perhaps even 1 μM (Neidhart, 1996). Glycerol molecules entering the cell pass through two types of passive channel proteins: first the porins, located in the outer membrane; then the inner-membrane aquaglyceroporins, including the glycerol facilitator GlpF. Upon entering the cytoplasm, glycerol molecules are phosphorylated by glycerol kinase and are unable to diffuse out of the cell (Neidhart, 1996). This article presents a model of the second step: conduction through GlpF.

GlpF belongs to the aquaporin family, a family of channels for water and small hydrophilic solutes which is found in most organisms (Hohmann et al., 2001). More than 100 different aquaporins have been discovered (Borgnia et al., 1999), including 11 in humans, some of which are implicated in diseases such as nephrogenic *diabetes insipidus* and congenital cataract of the eye (Borgnia et al., 1999; Kozono et al., 2002). In addition to water, GlpF conducts linear polyalcohols with a high degree of stereoselectivity (Heller et al., 1980; Borgnia and Agre, 2001). Glycerol, in particular, is conducted very quickly; 100–1000× faster than would be expected for a transporter (Heller et al., 1980).

The recent discovery of the structure of GlpF at a resolution of 2.2 Å (Fu et al., 2000; Nollert et al., 2001) has revealed important features of the transport mechanism. The channel walls match the hydrophilic and hydrophobic sides of glycerol, so that glycerol can be dehydrated without an energetic penalty. The entrance to the channel contains a narrow (4 Å wide) region that is thought to function as a selectivity filter. Interestingly, GlpF is asymmetric—a large (10 Å wide) hydrophilic vestibule extends into the water on the external side of the membrane (see Fig. 1; see also Fu et al., 2000; Nollert et al., 2001; Law and Sansom, 2002; Fujiyoshi et al., 2002). This feature is also seen in the three-dimensional structures of two general porins (Weiss et al., 1990; Kreusch et al., 1994), the porin OmpF (Cowan et al., 1992), maltoporin (Wang et al., 1997), and sucrose porin (Forst et al., 1998). GlpF has been studied in several molecular dynamics (MD) simulations (Zhu et al., 2001; de Groot et al., 2001; Jensen et al., 2001; de Groot and Grubmüller, 2001; Tajkhorshid et al., 2002; Jensen et al., 2002; Zhu et al., 2002; Grayson et al., 2003), including one in which the potential of mean force (PMF) guiding glycerol conduction was computed (Jensen et al., 2002). The PMF shows that the vestibule binds glycerol tightly, apparently acting as a trap for glycerol molecules. One naturally thinks that the vestibule benefits *E. coli* by pulling glycerol from the periplasm and that such a trap should face the source of glycerol. This intuitive reasoning can be examined using the language of statistical mechanics.

In this article, we address the questions of why the attractive vestibule exists and why it is located on the external side of the membrane. The answers to these questions may reveal the mechanistic strategies of aquaporin function. A simple statistical-mechanical model can answer the first question by quantifying the effect of an attractive vestibule on glycerol conduction. However, calculations of glycerol conduction cannot easily answer the second question. As we will demonstrate later, when the channel orientation is reversed, i.e., the attractive vestibule is on the cytoplasmic side, the glycerol uptake rate is no less than that with the original channel orientation. It therefore appears that simple calculations of the steady-state flux cannot explain why the attractive vestibule is on the periplasmic side. Much of the rest of the article is a detailed exploration of this argument.

In addition, both simple diffusion arguments and stochastic modeling demonstrate that GlpF conducts with equal

Submitted April 17, 2003, and accepted for publication June 27, 2003.

Address reprint requests to Klaus Schulten, Tel.: 217-244-1604; Fax: 217-244-6078; E-mail: kschulte@ks.uiuc.edu.

0006-3495/03/11/2977/11 $2.00

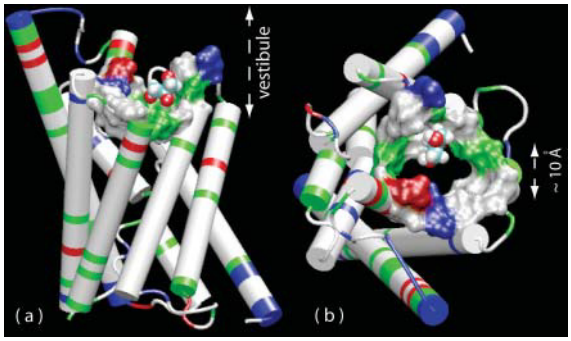

FIGURE 1 Periplasmic vestibule of GlpF holding a glycerol molecule. (*a*) Side view of the vestibule. (*b*) Top view of the vestibule. The protein close to the glycerol molecule is shown in surface rendering and the rest of the protein is shown in cartoon representation. The color of the protein corresponds to residue type (*white*, nonpolar; *green*, polar; *blue*, positively charged; *red*, negatively charged). This figure was produced with VMD (Humphrey et al., 1996).

efficiency in both directions, where efficiency is defined as the probability that a glycerol molecule entering the channel will be conducted. As a result, an appeal to the stochastic nature of glycerol conduction events does not explain the orientation of GlpF. One possible explanation is that if the attractive vestibule faces the cytoplasm, it may get clogged by a variety of solutes inside the cell, thus becoming less permeable to glycerol. However, we show that cytoplasmic clogging also appears unlikely to be an important factor.

It is believed that conduction of glycerol through GlpF is necessary for *E. coli* growth in concentrations of glycerol $<\sim 1$ mM (Rickey and Lin, 1972). However, differences in the kinetics of glycerol phosphorylation in a facilitator-minus-mutant suggest that the presence of GlpF directly activates glycerol kinase (Voegle et al., 1993). Without ignoring this important effect, we will focus on the conduction properties of GlpF. We assume in this article that GlpF has been optimized for glycerol conduction under physiological conditions.

We will first develop a six-step kinetic rate model, which is similar to the proposed four-step mechanism of the potassium channel KcsA (Morais-Cabral et al., 2001), from the PMF of the GlpF channel and define the flux in the steady state. To simplify our analysis, we introduce a two-step model, where quantities can be expressed analytically. Next we present the results of our kinetic rate model and compare the flux to the experimentally observed value. We also outline a stochastic model of channel efficiency, in which the efficiencies of opposite channel orientations are compared. We conclude with a discussion of the role of the physical asymmetry of GlpF.

METHODS

Biological channels have been studied extensively. Theoretical methods at different levels include continuum electrostatic Poisson-Boltzmann theory, Brownian dynamics, Possion-Nernst-Planck electrodiffusion theory, the kinetic rate model, and molecular dynamics (MD) simulations (Roux, 2002). In the kinetic rate model, the diffusing molecules are localized in the free energy minima. They can hop into vacant neighboring minima when enough thermal energy is acquired (Cooper et al., 1985; Hille, 1992). Recently, progress has been made to incorporate information obtained from MD simulations, such as the PMF and diffusion coefficient, into traditional kinetic rate models (Schumaker et al., 2000, 2001).

In this section, we develop a one-dimensional six-step kinetic rate model to study the dynamics of glycerol conductance through the GlpF channel. The model is inspired by the results of SMD simulations (Jensen et al., 2002) and crystallographic data (Fu et al., 2000; Nollert et al., 2001), which indicate that glycerol makes a series of discrete steps as it passes through the channel. For GlpF we use six steps that are similar to the four steps used in Morais-Cabral et al. (2001) to model the potassium channel KcsA. Since positions of the local minima of the PMF are found in accord with the locations of the hydrogen-bonding sites at the interior of the channel, revealed both from the crystal structure (Fu et al., 2000; Nollert et al., 2001) and the equilibrium MD simulation (Jensen et al., 2001), we believe that the PMF that resulted from the SMD simulation captures the intrinsic conduction mechanism of GlpF.

Despite its simple nature, the six-step model can provide a great deal of insight into the glycerol transport mechanism. The time-dependent conduction rate can be easily simulated with the Monte Carlo method. The steady-state flux, which is experimentally observable, can be computed efficiently using numerical methods. Furthermore, the physical asymmetry of the channel can also be conveniently explored by tuning a single parameter. Another significance of the six-step model is that the interaction of the diffusive particles—for example, the exclusion of more than one glycerol molecule occupying the same state simultaneously—can be suitably accounted for in the calculation.

Besides the six-step model, we also investigate a two-step model that has the benefit of being analytically solvable. The efficiency of the channel is also examined and we conclude that it is independent of the channel orientation.

Construction of a six-step model

As shown in Fig. 2 *b*, nine local free energy minima, from A to I, and eight local maxima, from A′ to H′, can be clearly identified in the interior of the channel. The 17 minima and maxima are denoted by the subscript *i*, which takes values from 1 to 17. For each of the minima and maxima, we denote the free energy by E_i and the glycerol population by ρ_i. The boundary regions of the channel are defined by $z < z_A$ and $z > z_I$, where z_A and z_I are the z-coordinates (measured parallel to the channel axis) of states A and I.

To calculate the flux, one needs to know the transition rates $Q_{i,j}$ between two neighboring locations i and j as a function of their free energy difference $\Delta E_{i,j} = E_i - E_j$. In thermodynamic equilibrium, when the average concentrations ρ_i and ρ_j are very low, the Boltzmann relation,

$$\frac{\rho_i}{\rho_j} = \exp(\beta \Delta E_{j,i}), \tag{1}$$

holds, where $\beta = 1/k_B T$ and k_B is the Boltzmann constant. All calculations in this study were carried out at physiological temperature (37°C). Under these conditions, the Boltzmann relation is established by the following choice of rates:

$$Q_{i,j} = \frac{\tilde{D}}{\Delta z_{i,j}^2} \exp\left(\frac{1}{2}\beta \Delta E_{i,j}\right), \tag{2}$$

where $\tilde{D}$ is the diffusion coefficient of glycerol inside the channel, and $\Delta z_{i,j}$ is the distance between the two locations. Since diffusion properties inside narrow pores are quite different from those in bulk water (Breed et al., 1996; Sansom et al., 1996; Tieleman and Berendsen, 1998), we distinguish $\tilde{D}$ from D, the diffusion coefficient of glycerol in bulk water. Our method of determining $\tilde{D}$ in this study is explained later. Though other definitions of $Q_{i,j}$ are possible, this one has the desirable property of reproducing the

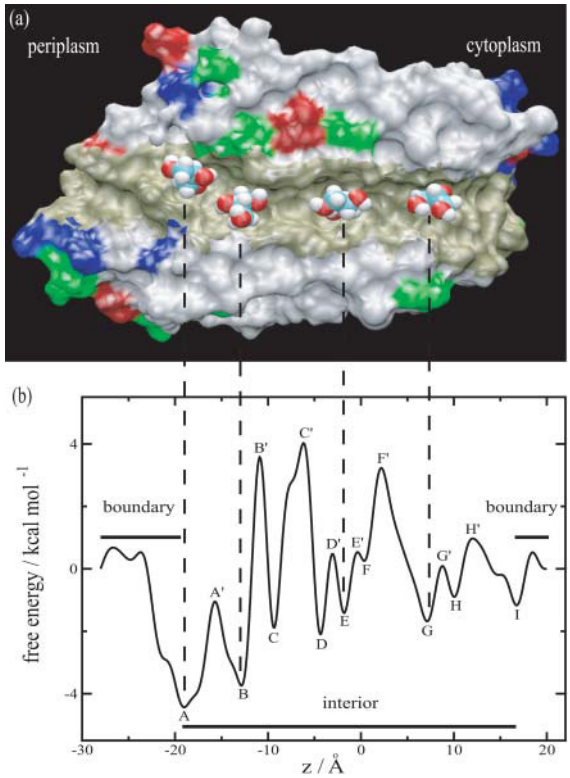

FIGURE 2 (*a*) Slice of GlpF along the glycerol conduction pathway. Four simulated glycerols are displayed in VDW representation. The color of the protein corresponds to residue type (see Fig. 1), with the interior of the channel indicated in tan. This figure was produced with VMD (Humphrey et al., 1996). (*b*) PMF of glycerol in GlpF as reported in Jensen et al. (2002). The local free energy minima A, B, E, and G correspond to the locations of glycerols in *a*. The channel is classified into interior region and boundary regions according to both the PMF and the width of the channel.

correct continuous diffusion dynamics when the limit $\Delta z_{i,j} \rightarrow 0$ is taken (Agmon and Hopfield, 1983; Schumaker et al., 2001, 2002). The rate expression in Eq. 2 will be used throughout this article.

The periplasm and the cytoplasm, labeled by the subscripts *out* and *in*, respectively, are regarded as two ends of the channel. Starting from the periplasm, there are altogether 18 transitions that need to be made for a sugar molecule to reach the cytoplasm. Though such a detailed description is possible in principle, it needs to be simplified to allow efficient computation.

Due to optimal hydrogen bonding with the channel protein, glycerols are most stable at the free energy minima (Jensen et al., 2001). Twelve glycerol trajectories of MD simulation, as seen in Fig. 3 *a* (Jensen et al., 2001), reveal the kinetics of glycerol conduction: long occupation of the free energy minima separated by rapid transitions between neighboring minima. This behavior allows us to faithfully describe the conduction by only modeling the transitions of glycerols between the neighboring minima through a kinetic model. If steps $i-1$ and $i+1$ are local minima, and step i is a local maximum, the steady-state rate $R_{i-1,i+1}$ can be expressed

$$R_{i-1,i+1} = \frac{Q_{i-1,i}\, Q_{i,i+1}}{Q_{i,i-1} + Q_{i,i+1}}, \tag{3}$$

as long as ρ_{i-1}, ρ_i, and ρ_{i+1} are small. The kinetic model is then reduced to nine steps. The resulting rate constants $R_{i,j}$ are shown in Fig. 3 *b*.

The rate constants $R_{i,j}$ can be divided into two categories: fast transitions within clusters of close neighbors (indicated by *ellipses*); and slow transitions between the clusters. The conduction is governed by the slowest processes and, hence, it is a good approximation to treat the states within

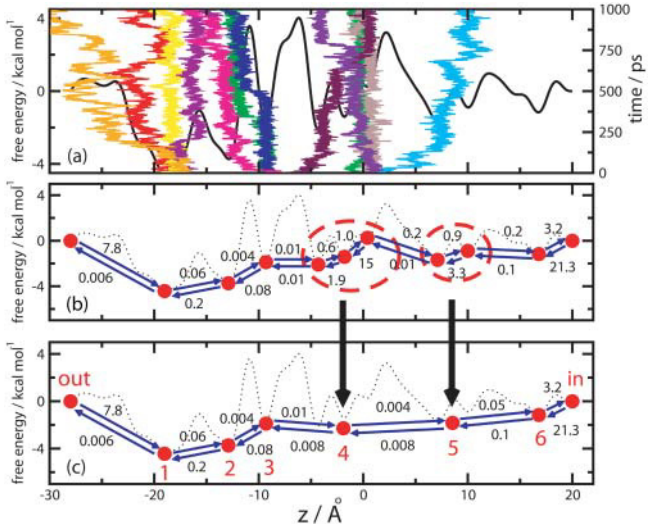

FIGURE 3 Construction of the six-step model. (*a*) Trajectories $z(t)$ of glycerol center of mass (*right*) in GlpF. The places where glycerol molecules spend most of their time correspond to the local free energy minima of the PMF (*left*). (*b*) Transition rates of glycerol, in 1/ns, between adjacent valleys. There are two clusters (indicated by *ellipses*), in which the rates are much faster than outside. (*c*) Transition rates, in 1/ns, of the six-step model after each cluster is converted into a single state.

a cluster as always in equilibrium. With this assumption, each cluster is effectively replaced by a single state. The kinetic model is finally simplified to six channel states as well as states "in" and "out" as shown in Fig. 3 *c*. Assuming that a cluster α contains n states, $\alpha_1, \ldots, \alpha_n$, the effective free energy of α is

$$E_\alpha = -\frac{1}{\beta}\log\left[\sum_{j=1}^{n}\exp(-\beta E_{\alpha_j})\right]. \tag{4}$$

The transition rates between α and its left neighbor i and right neighbor j are

$$R'_{i,\alpha} = R_{i,\alpha_1}\exp\left(\frac{1}{2}\beta\Delta E_{\alpha,\alpha_1}\right), \tag{5}$$

$$R'_{j,\alpha} = R_{j,\alpha_n}\exp\left(\frac{1}{2}\beta\Delta E_{\alpha,\alpha_n}\right), \tag{6}$$

where $\Delta E_{\alpha,\alpha_1}$ and $\Delta E_{\alpha,\alpha_n}$ denote the free energy difference between the cluster α and the states α_1 and α_n.

The flux of glycerol through the channel is computed in the steady state. The boundary conditions are $\rho_{out} = \rho_0$ and $\rho_{in} = 0$, where ρ_0 is related to the external glycerol concentration C as described in Appendix A, and the zero boundary condition applied at the cytoplasm accounts for glycerol depletion by phosphorylation. The flux J is expressed as

$$J = \rho_6 R_{6,in}, \tag{7}$$

where ρ_6 labels the glycerol population at step 6, and $R_{6,in}$ denotes the rate constant from step 6 to the cytoplasm ("in" state) in the six-step model. The parameters used for the six-step model are summarized in Table 1.

In the six-step model, it is assumed that each step is either empty or occupied by a single glycerol molecule. The total number of configurations possible for the entire channel is $2^6 = 64$. The dynamics of the model can therefore be represented as a finite-state Markov process. Our method of simulating this process is described in the next section.

One potential error in the six-step model described above is that it is derived from the PMF of a single glycerol molecule. When more than one

TABLE 1 Coordinates *z* and free energies *E* of 17 states (A–I and A′–H′ in Fig. 2 *b*) used to derive the six-step model

	A	B	C	D	E	F	G	H	I
z (Å)	−19.0	−12.9	−9.3	−4.3	−1.8	0.4	7.1	10.0	16.8
E (k_BT)	−7.2	−6.1	−3.1	−3.4	−2.3	0.4	−2.7	−1.5	−1.9
	A′	B′	C′	D′	E′	F′	G′	H′	
z (Å)	−15.7	−10.9	−6.2	−3.1	−0.4	2.2	8.8	12.0	
E (k_BT)	−1.7	5.8	6.5	0.8	0.9	5.2	0.1	1.6	

$\tilde{D} = 2.2 \times 10^{-6}$ cm^2/s, $T = 310$ K

glycerol is present in the channel, the potential felt by each glycerol differs from the PMF used in the model. To estimate this error, we divide the real PMF into two components (Jensen et al., 2002): one from the channel protein, and the other from water. The conformational response of GlpF to glycerol conduction was exhaustively studied in Jensen et al. (2001) by comparing the root mean-square deviation (RMSD) between the crystal and simulated structures. The RMSD of the C_α atoms of the tetramer is ~1.0 Å, which is 0.5–2.0 Å lower than the typical RMSD observed in the MD simulations of proteins. Considering all atoms at the channel interior, the RMSD value is still ~1.0 Å, which suggests a minor conformational change of protein at the channel interior. As the protein is very rigid and the glycerol molecules interact with different parts of the channel, the error in the channel-protein–glycerol component is negligible.

As for the water–glycerol component, the correction depends on the distance between two glycerols. At most locations in the six-step model, the glycerol molecules are separated by >6 Å, so there is enough room for water molecules to form hydrogen bonds with both glycerols. However, steps 2 and 3 are separated by only 3.5 Å, which does not leave enough room for water, possibly changing the PMF. To test the effect of this error, we purposely prevent glycerol molecules from occupying steps 2 and 3 simultaneously and compare the results of this refined model to the six-step model.

Steady-state and Monte Carlo simulations

The steady-state glycerol flux through the six-step model channel under physiological conditions was calculated directly from the transition rates given by Eqs. 2–6. We denote the 64 configurations of the entire channel by the subscript i, $i = 1, 2, \ldots, 64$. In the steady state, each of the 64 possible states of the channel has a probability, P_i. The rate of transition from state i to state j is proportional to P_i and also to the associated transition rate, which we call $\tilde{R}_{j,i}$. If the configurations j and i are not related by a transition of a single glycerol molecule, $\tilde{R}_{j,i}$ is zero. Otherwise, the particular transition that relates i to j determines $\tilde{R}_{j,i}$. For example, if moving one glycerol molecule from step six into the cytoplasm brings the system from configuration i to configuration j, we have $\tilde{R}_{j,i} = R_{6,in}$. We further define $\tilde{R}_{i,i} = -\sum_{j\neq i}\tilde{R}_{j,i}$, so that the total rates of change of the probabilities are given by a matrix product:

$$\frac{d}{dt}P_i = \sum_{j=1}^{64}\tilde{R}_{i,j}P_j. \tag{8}$$

The probability distribution in the steady state is constant with time, so this product must be zero for every i. The set of vectors P_i for which this product is zero is called the *kernel*: numerical methods exist to efficiently compute the kernel (Wolfram, 1996). Since there is only one steady-state solution, the kernel has exactly one (properly normalized) member, specifying a particular set of P_i. In our steady-state simulations, the kernel was computed numerically, and the glycerol distribution ρ_i, $i = 1, 2, \ldots, 6$, was derived from the resulting P_i.

In the refined steady-state calculation, configurations with two glycerols at states 2 and 3 are excluded, leaving $2^6 - 2^4 = 48$ valid conformations. The way of computing flux in this refined model is the same as described above.

Monte Carlo (random walk) simulations similar to those in Schumaker et al. (2000, 2001) were used to verify the steady-state rates. In these simulations, the system is assumed to be in a definite state at all times, instead of a steady superposition of states. Time is divided up into small time-steps Δt, and the probability $\Gamma_{i,j}$ of a transition from the current state i to each new state j is computed once at every time-step:

$$\Gamma_{i,j} = \tilde{R}_{j,i}\Delta t. \tag{9}$$

The system can make at most one transition during a time-step, so Δt must be small enough that

$$\sum_j \Gamma_{i,j} \leq 1 \tag{10}$$

is satisfied (Binder, 1986). Simulations of a random walk can be greatly accelerated by scaling up Δt so that

$$\sum_j \Gamma_{i,j} = 1. \tag{11}$$

In this limit, Δt becomes the expected time to make a single transition, and the $\Gamma_{i,j}$ becomes the probability that the next transition will bring the system to the state j. The system makes one simulated transition at every time-step, allowing the fastest possible computation of conduction.

Construction of a two-step model

To investigate the role of the channel asymmetry due to the periplasmic attractive well, we consider the two-step model shown in Fig. 4. In this model, all details of the PMF are neglected except for the well and a barrier, the latter being indispensable for a highly selective channel (de Groot and Grubmüller, 2001). The simplicity of the model permits us to describe the channel conduction properties analytically.

As seen in Figs. 2 *b* and 4, the well region includes valleys A and B, while the barrier region represents peaks B′, C′, D′, and F′. The free energies of the well and the barrier, given by the exponential average of their PMF, are −7.48 and 7.11 k_BT, respectively. Then there exists an overall barrier of ~15 k_BT on the glycerol uptake pathway, which is supported by the 15.6 k_BT (9.6 kcal/mol) Arrhenius activation energy measured on affinity-purified GlpF (Borgnia and Agre, 2001). Following the same notions used in the six-step model, the forward and the backward flux between states i and j are

$$J_{ij} = \rho_i(1-\rho_j)T_{ij}, \tag{12}$$

$$J_{ji} = \rho_j(1-\rho_i)T_{ji}, \tag{13}$$

and the net flux is

$$J = J_{ij} - J_{ji}, \tag{14}$$

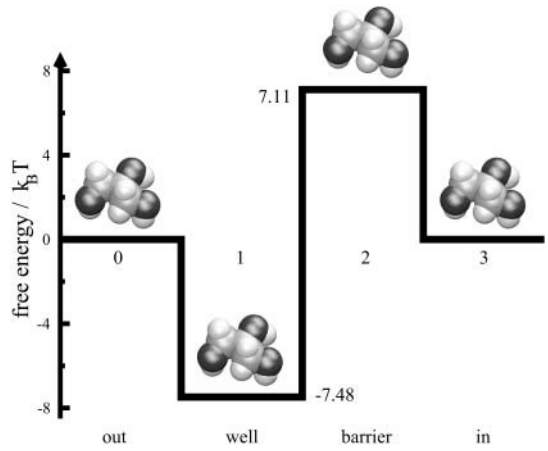

FIGURE 4 Schematic representation of the two-step model with a well of 7.48 k_BT and a barrier of 7.11 k_BT.

which is constant along the channel in the steady state. Again the transition rate T_{ij} is related to the free energy difference by

$$T_{ij} = T_0 \exp\left(\frac{1}{2}\beta\Delta E_{i,j}\right), \tag{15}$$

where T_0 can be fixed by fitting the calculated flux to the six-step model.

Under the same boundary conditions as adopted for the six-step model, the master equation for the two-step model reads

$$\frac{d}{dt}\rho_w = -\rho_w[(1-\rho_b)T_{12} + (1-\rho_0)T_{10}] + (1-\rho_w)(\rho_0 T_{01} + \rho_b T_{21}), \tag{16}$$

$$\frac{d}{dt}\rho_b = -\rho_b[(1-\rho_w)T_{21} + T_{23}] + (1-\rho_b)\rho_w T_{12}, \tag{17}$$

where ρ_0, ρ_w, and ρ_b are the probabilities at the periplasm, the well, and the barrier, respectively. The steady-state solution to these equations can be expressed analytically:

$$\rho_b = \frac{1}{2}(-t_1 + \sqrt{t_1^2 + 4t_2}), \tag{18}$$

$$\rho_w = \frac{\rho_0 T_{01} - \rho_b T_{23}}{T_{10}(1-\rho_0) + \rho_0 T_{01}}, \tag{19}$$

$$J = T_0 \rho_b \exp\left(\frac{1}{2}\beta U_b\right), \tag{20}$$

where the parameters t_1 and t_2 are defined as

$$t_1 = (1-\rho_0)\frac{T_{10}(T_{23}+T_{21})}{T_{23}T_{21}} + \frac{T_{12}}{T_{21}} + \rho_0\frac{T_{01}}{T_{21}}, \tag{21}$$

$$t_2 = \rho_0\frac{T_{01}T_{12}}{T_{21}T_{23}}, \tag{22}$$

and U_w and U_b are the absolute values of well depth and barrier height. The derivation of Eqs. 18–22 is given in Appendix B.

By matching the flux of this model to that of the six-step model at various glycerol concentrations, we determine $T_0 = 1.8 \pm 0.1\ \text{ns}^{-1}$. Since T_0 only varies by ~5%, it can indeed be interpreted as a constant, allowing the two-step model to account fairly precisely for the glycerol conduction rate. To simplify our later discussion, we define the effective flux $\tilde{J}$ as

$$\tilde{J} \equiv J/T_0. \tag{23}$$

However, in general, the parameter T_0 depends on U_w and U_b. Under certain circumstances, T_0 cannot be obtained by a single fit. For example, we want to discover the role of the attractive well by investigating the relationship between the flux and the depth of the well. When U_w is varied, T_0 is no longer a constant. But T_0 is fairly constant when the well is either shallower than 10 k_BT or deeper than 20 k_BT. In these ranges, we will not use the two-step model to predict the absolute flux J, but to study the glycerol distribution and the effective flux $\tilde{J}$.

Conduction probability for individual glycerol molecules

In the previous sections we have assumed that the most important feature of GlpF is the steady-state flux of glycerol at a given concentration. One possible objection to this reasoning is that glycerol may be found in very small quantities that are distributed nonuniformly around the cell. The concentration and flux of glycerol are ill-defined when glycerol molecules arrive at the channel one at a time and irregularly. Under such conditions a more reasonable quantity to characterize the channel is its *efficiency*: the probability that a glycerol molecule arriving at the channel entrance will be conducted to the cytoplasm. We will denote the efficiency by η and express it as a conditional probability,

$$\eta = P\,(conduction|arrival). \tag{24}$$

Defined this way, η is independent of the concentration or distribution of glycerol.

We will first relate η to the steady-state flux, and show that it is independent of the channel orientation. Next, we evaluate η directly from a stochastic model. In this model, η of opposite channel orientations are compared, and found to be identical.

Efficiency from the steady state

Under steady state, low concentration conditions, the arrival and conduction events both happen at well-defined rates, which we denote by $R(arrival)$ and $R(conduction)$. Since arrival at the channel is required for conduction, η can be re-expressed as

$$\eta = \frac{R(conduction)}{R(arrival)}. \tag{25}$$

The rate of conduction is simply the steady-state flux J. The steady-state flux is in turn proportional to the glycerol concentration C, so that

$$\eta = \frac{J}{R(arrival)} = \frac{\alpha \times C}{R(arrival)}, \tag{26}$$

where α, the proportionality constant, is independent of the channel orientation. A more detailed explanation of this linear relationship is given in the Discussion section. Since diffusion to the entrance of the channel does not depend on forces within the channel, $R(arrival)$ is also independent of channel orientation. Therefore, in addition to the flux, the efficiency is also independent of the channel orientation.

Efficiency from time-dependent solution

Let us begin with a channel described by discrete states i, ($i = 1, \ldots, n$), which allows only single file transport, where states 1 and n correspond to the solute positioned at the two ends of the channel. From state n, the solute molecule is absorbed with a rate K_a by the cell interior. On the other side, the solute molecule can drift out of the channel from state 1 with rate K_d. K_a and K_d linearly depend on the opening areas of the channel and should be actually identical (see Appendix A). Once the solute escapes from the channel, it may diffuse back, but we neglect the chance of recapture.

In this model, we are looking at low solute concentrations, such that no more than one solute molecule is in the channel at a given time. For the normal channel orientation, a solute molecule is initially placed at state 1. The efficiency is defined as the overall probability of the molecule being absorbed at state n. To obtain the efficiency of the reversed orientation, one simply starts with the solute molecule at state n, and looks for the probability of being absorbed at state 1. The efficiencies of both orientations are related to a time integral (Ritz et al., 2001; Şener et al., 2002)

$$\eta_{nor} = \int_0^\infty dt\, K_a\, \rho_n(t), \tag{27}$$

$$\eta_{rev} = \int_0^\infty dt\, K_d\, \rho_1(t), \tag{28}$$

where we denote by $\rho_n(t)$ and $\rho_1(t)$ the probability to find the solute at state n and state 1, respectively.

Let us now define the probability vector, $|\rho(t)\rangle = (\rho_1(t), \rho_2(t), \ldots, \rho_n(t))^T$, as being the time-dependent probability distribution at each state. The time evolution of $\rho(t)$ is governed by

$$\partial_t|\rho(t)\rangle = K|\rho(t)\rangle, \tag{29}$$

where K is an n-by-n rate matrix. The matrix elements of K can be constructed similarly as the kinetic rate model in Eq. 2. The off-diagonal matrix elements,

$$K_{i,j} = \frac{\tilde{D}}{\Delta z_{i,j}^2}\exp\left(\frac{1}{2}\beta\Delta E_{j,i}\right), \tag{30}$$

give the rates from state j to state i, and the diagonal matrix elements,

$$K_{i,i} = -(\delta_{i,1}K_d + \delta_{i,n}K_a) - \sum_{j\neq i} K_{j,i}, \tag{31}$$

describe the total rate leaving state i. The absorption rate K_a and the drift rate K_d are included at the two ends.

The solution of Eq. 29 reads

$$|\rho(t)\rangle = e^{Kt}|\rho(0)\rangle, \tag{32}$$

where $|\rho(0)\rangle$ is the initial distribution. If we define the localized states $|1\rangle$ and $|n\rangle$ as

$$|1\rangle \equiv (1, 0, \ldots, 0)^T, \tag{33}$$

$$|n\rangle \equiv (0, 0, \ldots, 1)^T, \tag{34}$$

where only the state 1 or n is occupied and all the other states are empty, for the normal orientation, the initial condition is $|\rho(0)\rangle = |1\rangle$; for the reversed orientation, $|\rho(0)\rangle = |n\rangle$. Using Eqs. 27–28 and 32–34, the efficiencies can be written

$$\eta_{nor} = -K_a\langle n|K^{-1}|1\rangle = -K_a(K^{-1})_{n,1}, \tag{35}$$

$$\eta_{rev} = -K_d\langle 1|K^{-1}|n\rangle = -K_d(K^{-1})_{1,n}, \tag{36}$$

where the matrix elements $(K^{-1})_{n,1}$ and $(K^{-1})_{1,n}$ can be calculated from the cofactors of $K_{1,n}$ and $K_{n,1}$, and the following relation holds:

$$\frac{\eta_{nor}}{\eta_{rev}} = \frac{K_a\prod_{i=2}^{n}K_{i,i-1}}{K_d\prod_{i=1}^{n-1}K_{i,i+1}}. \tag{37}$$

From the definition of rate elements in Eq. 30, we have

$$\frac{\prod_{i=2}^{n}K_{i,I-1}}{\prod_{i=1}^{n-1}K_{i,i+1}} = \frac{\exp\left(\frac{1}{2}\beta(E_1 - E_n)\right)}{\exp\left(\frac{1}{2}\beta(E_n - E_1)\right)}. \tag{38}$$

If the free energy is the same at both sides,

$$\frac{\eta_{nor}}{\eta_{rev}} = \frac{K_a}{K_d}. \tag{39}$$

Since $K_d = K_a$, one finds $\eta_{nor}/\eta_{rev} = 1$, which agrees with what we concluded from stationary rates.

The diffusion coefficient inside the channel

The diffusion of water in nano-scale channels is largely reduced compared to bulk (Breed et al., 1996; Sansom et al., 1996; Tieleman and Berendsen, 1998). As glycerol molecules conduct through GlpF in single file separated by water molecules (Fu et al., 2000; Jensen et al., 2001), the diffusion of glycerol is also slowed down inside the pore. Since there is no direct experimental measurement of the corresponding diffusion coefficient $\tilde{D}$, we need to derive it for the steady-state model.

Our first attempt was to evaluate it from a MD simulation by calculating the velocity autocorrelation function. This approach has been successfully applied in Nadler et al. (1987) to model the dynamics of lysine side chains in the active site of RNase A. Details of this method are given in Appendix C. However, $\tilde{D}$ calculated from a 50-ps MD simulation is about twice as large as the bulk glycerol diffusion coefficient, which means the randomness controlling the long time (ns, see Fig. 3 *c*) glycerol conduction is different from the randomness that controls the short time (ps) transport.

Alternatively, a natural upper bound on $\tilde{D}$ is set by the diffusion coefficient of water inside GlpF, 4.6×10^{-6} cm^2/s (Tajkhorshid et al., 2002). This choice is based on the following assumption about the glycerol conduction mechanism: glycerol forms a single file with water molecules in GlpF such that the motion of glycerol is correlated with that of its two neighboring water molecules. Only when these two water molecules move is the glycerol able to move. In such a concerted motion, the diffusion of the glycerol is driven by the diffusion of the water next to it, which in turn sets the upper limit for the diffusion coefficient of glycerol. Rather than take the water diffusion coefficient itself, we consider the glycerol molecule, being heavier than water, to move more slowly than water molecules, so that its diffusion coefficient should also be smaller. As the bulk diffusion coefficient of glycerol is about one-half of that of water, we roughly estimate $\tilde{D} = 2.2 \times 10^{-6}$ cm^2/s, which is also about one-half of the water diffusion coefficient in GlpF.

In the channel interior, from states 1 to 6 in the six-step model, we use the reduced diffusion coefficient $\tilde{D}$. At the boundaries, however, the glycerol is diffusing in bulk water. We therefore use here the bulk value, $D = 1.1 \times 10^{-5}$ cm^2/s (Lide, 1994).

RESULTS AND DISCUSSION

In this section, the glycerol conduction rates of GlpF are calculated at various glycerol concentrations based on our kinetic models. The theoretical results are compared with observation, yielding remarkable agreement. The role of the attractive well is revealed by its effect on glycerol conduction rates. Further effort is made to explore the puzzle of why the attractive well is located on the periplasmic side, by testing the clogging effect of solutes in the cytoplasm. However, our calculations indicate that this effect is unlikely to be important.

Flux at various glycerol concentrations

The flux curves are computed using both the six-step model and the two-step model. To test our modeling, they are compared with a Monte Carlo simulation in Fig. 5. One can see that the Monte Carlo simulation results agree well with the steady-state results of the kinetic models, indicating that the steady-state description indeed captures the long term conduction behavior. The flux of the two-step model also reproduces well the results of more accurate models, which confirms our assumption that the channel is characterized mainly by the periplasmic well and an overall barrier. From Fig. 5, it can be seen immediately that the channel conducts nonlinearly at high concentration. For example, the flux at 60 mM is about one-half of that at 0.5 M. This nonlinear effect results from the clogging of the well at high glycerol concentration, as illustrated in the inset: the well is one-half occupied at $\sim$2 mM, and is almost fully occupied at a glycerol concentration $>$0.1 M.

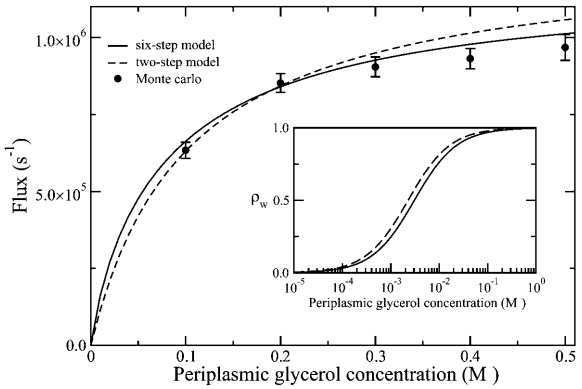

FIGURE 5 Glycerol flux through GlpF versus periplasmic glycerol concentration calculated from six-step model (*solid*), two-step model (*dashed*), and Monte Carlo simulation (*disks*). The saturation of the channel flux is due to clogging of the attractive vestibule at the periplasmic side. Glycerol population of the vestibule is shown in the inset.

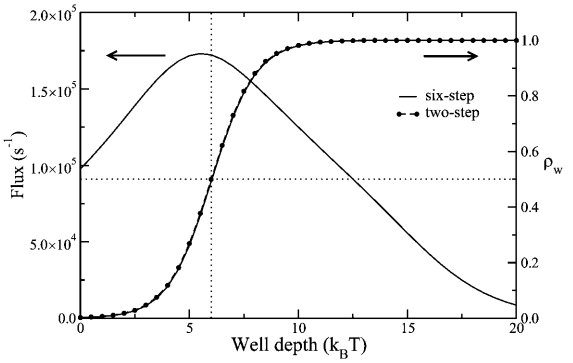

FIGURE 6 The flux of glycerol in the six-step model of GlpF (*left*) and the glycerol population in the attractive vestibule (*right*) in the presence of 10 mM glycerol. The flux initially increases with the well depth, reaches its maximum at a well depth of 5.5 k_BT, and drops rapidly thereafter. The well population of the two-step model (*dotted*) agrees well with that of the six-step model (*solid*). The flux is nearly optimized at 50% well occupation.

Next we examine the absolute magnitude of the glycerol flux. The calculated glycerol flux of GlpF is $1.0 \times 10^6\ s^{-1}$ at 0.5 M, which agrees rather well with the observed value of $2 \times 10^5\ s^{-1}$ (Heller et al., 1980). When states 2 and 3 are prohibited to be occupied simultaneously, the flux at 0.5 M in this refined model is $4.3 \times 10^5\ s^{-1}$, which is even closer to observation. Due to the difficulty of experimentally measuring the density of GlpF, the flux reported in Heller et al. (1980) is actually estimated based on the assumption that the total amount of GlpF per cell is the same as the amount of glycerol kinase, which is encoded by the same operon. An error due to this estimate together with the roughness of our model could account for the difference between our model and the experimental data.

Conduction for varying well depth

To answer the question of how the attractive vestibule affects the glycerol flux, we examine the flux against the depth of the well at high (10 mM) and low (10 μM) concentrations. According to the connection rule described in Appendix A, $\rho_0 \ll 1$ holds for both scenarios. In this regime, the probability that both states 2 and 3 are occupied is extremely low, and the refined steady-state calculation yields the same results as the six-step model. To be concise, we will only present the results from the simple six-step model.

Let us first look at two extreme cases in the two-step model: the well being nearly empty, i.e., $\rho_w \ll 1$, and the well being nearly clogged, i.e., $1 - \rho_w \ll 1$. As one can see from Figs. 6 and 7, these restrictions correspond to very shallow or deep wells, where the fitting parameter T_0 remains constant, and we can analyze the two extremes using $\tilde{J}$.

In the former case, Eq. 20 reduces to

$$\tilde{J} = \frac{\rho_0 \exp\left(-\frac{1}{2}\beta U_b\right)}{1 + \exp\left(-\frac{1}{2}\beta U_w\right) + \exp\left[-\frac{1}{2}\beta(U_w + U_b)\right]}. \quad (40)$$

According to Eq. 40, the effective flux increases as the well gets deeper and reaches its maximum of $\rho_0 \exp(-(1/2)\beta U_b)$ asymptotically as the depth of the well goes to infinity. Since we assume $\rho_w \ll 1$, the flux saturates for large values of U_w. Given that the uptake of glycerol molecules is diffusion-limited, one can also estimate, from Eq. 40, how much the flux is reduced by the presence of the selectivity filter. In other words, we can estimate how the glycerol conduction rate of GlpF is compared with an artificial channel with flat potential as $U_w = U_b = 0$. The two-step model suggests that to overcome the free energy barrier, the flux is slowed down by an order of magnitude.

In the latter case, when the well is clogged, the effective flux is

$$\tilde{J} = \exp\left[-\frac{1}{2}\beta(U_w + U_b)\right]. \quad (41)$$

Based on the above discussion, the effective flux is described by an initial increase, saturation, and an exponential decay as

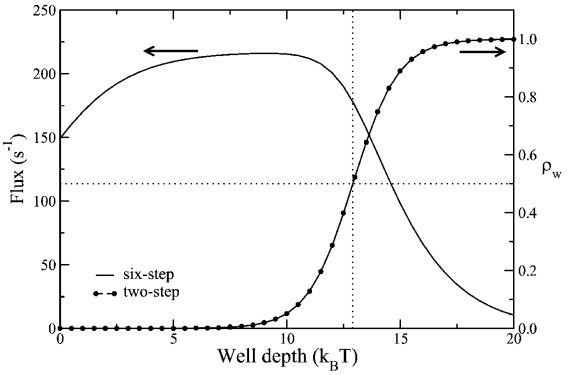

FIGURE 7 The flux of glycerol described through the six-step model of GlpF (*left*) and the glycerol population in the attractive vestibule (*right*) in the presence of 10 μM glycerol. Instead of the peak in the previous figure, the flux curve exhibits a broad plateau and is optimized at <50% well occupation.

the depth of the well increases. From the flux curves in Figs. 6 and 7, one can immediately recognize these features. The absolute flux of the two-step model is not shown here since the scaling parameter T_0 depends on U_w. The glycerol populations ρ_w in the first well in both models, independent of T_0, are plotted together, and the results exhibit remarkable agreement.

When the glycerol concentration is 10 mM, the flux hits its maximum at a well depth of $\sim$5.5 k_BT, achieving a 75% improvement compared with a channel without a well. Under 10 μM glycerol concentration, the attractive vestibule helps to increase the flux by >40% and the flux is optimized at the well depth between $\sim$5 and 10 k_BT. While the curve is very flat over this range, the actual well depth of 7.2 k_BT indicates that GlpF has been optimized by evolution for low glycerol concentration. We see that the attractive well plays a key role in catching glycerol molecules, which is the primary function of GlpF. These results are independent of the particular choice of the diffusion coefficient, since it does not affect the relative fluxes.

Channel with reversed orientation

Our study above stresses the importance of the attractive vestibule for the glycerol conduction. But why is the vestibule on the periplasmic side, and not on the cytoplasmic side? To address this question, we consider a channel with an inverted PMF so that the attractive well is facing the cytoplasmic side. In other words, regarding the PMF in Fig. 2 *b*, the fluxes of the normal and reversed orientations correspond to glycerol influx and outflux. We will stick to the six-step model in comparing the conduction rates of both channel orientations.

Results in Fig. 8 show that at <1 mM glycerol concentration, the fluxes in both directions are about the same, and that at >1 mM concentration, the flux of the reversed orientation is larger.

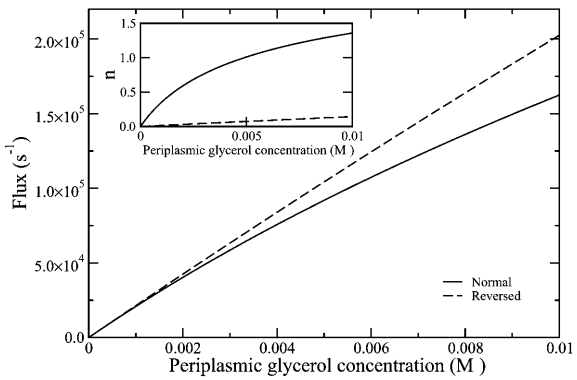

FIGURE 8 Comparison of the glycerol flux through GlpF with its attractive vestibule placed in the periplasm, i.e., the normal orientation (*solid*), or in the cytoplasm, i.e., the reversed orientation (*dashed*). Below 1 mM, the fluxes coincide. But the channel with reversed orientation conducts better at higher glycerol concentration. The total number of glycerols inside the channel in each orientation is shown in the inset.

These results can be understood in the context of a simple diffusion argument. For low enough concentrations, interactions between glycerol molecules can be neglected. If we define the direction of the net flux to be from periplasm to cytoplasm, the net flux is given by subtracting an outflux proportional to the cytoplasmic concentration from an influx proportional to the periplasmic concentration. In static equilibrium, equal concentrations of glycerol are found on both sides and there is no net flux, so the two coefficients of proportionality must be equal. Accordingly, the net flux is proportional to the concentration difference ($C_{out} - C_{in}$), with a positive coefficient of proportionality. In a living cell, the channel conducts low concentrations of periplasmic glycerol into the cytoplasm, which is nearly devoid of glycerol. Because the flux depends only on the concentration difference across the channel, maintaining glycerol concentrations on both sides and reversing the channel will not change the flux. Therefore, there seems to be no reason to prefer one direction over the other. At higher concentrations, however, interactions become important, and conduction is no longer linear. Since glycerol molecules held in the vestibule prevent others from being conducted, the optimal direction can be estimated by putting the vestibule where it will be most rarely occupied: as close to the cytoplasm as possible. As a result, the reversed GlpF has an increased conduction rate due to reduced clogging in the vestibule.

One explanation for the channel orientation of GlpF is that if the attractive vestibule would face the cytoplasm, it would be clogged by a variety of solutes inside the cell. To test this hypothesis, we calculate the total number of glycerols, n, inside the channel, against the glycerol concentration for both orientations. As one can see from the inset of Fig. 8, n is <0.5 for glycerol concentrations less than a few mM, in both the normal and reversed orientations. For a cytoplasmic solute to significantly clog the reversed channel, three conditions must be satisfied:

1. The solute must be found in concentrations of at least a few mM. Glycerol phosphate, in particular, is maintained at a concentration of 200 μM (Neidhart, 1996).
2. The solute must bind to the vestibule as strongly as glycerol, despite the specialization of GlpF for glycerol.
3. The solute must be found in relatively low concentrations in the periplasm.

The solutes found in the periplasm are unfortunately not well-characterized (Neidhart, 1996). This hypothesis is additionally complicated by the possibility of a direct interaction between glycerol kinase and GlpF, which would take place on the cytoplasmic side. For all of these reasons, it seems unlikely that cytoplasmic solutes can provide a simple explanation for the orientation of GlpF. However, one cannot exclude the possibilities that cytoplasmic solutes other than glycerol could clog a cytoplasmic vestibule.

According to our model, the glycerol flux through GlpF does not depend on its orientation in the membrane of *E. coli*

at <1 mM. Therefore, we conclude that no argument aiming only at optimization of glycerol conduction can explain why the attractive vestibule is on the periplasmic side of the channel. It is worth noting that we have consistently ignored the dynamics of glycerol molecules in the periplasm. We expect that one-dimensional motion along the channel is by far the slowest step, but we cannot exclude the possibility that the periplasm affects conduction in an important way.

A possible answer of the channel asymmetry might be related to cell protection. Though the glycerol conduction rate needs to be enhanced at low concentrations, such as $\sim 1\ \mu$M, it is also important for the cell to keep the activity of glycerol kinase below certain limits, since the rapid accumulation of dihydroxyacetone-phosphate, an intermediate product of the glycerol metabolic pathway, results in lethal formation of the highly chemically reactive methylglyoxal (Neidhart, 1996). Though the protection mechanism is usually maintained by catabolite repression of glycerol kinase, a slowdown of glycerol uptake relative to the reversed channel orientation, when experiencing a glycerol shock, may become a supplementary self-protection mechanism to prevent the formation of methylglyoxal.

CONCLUSIONS

The crystal structure of GlpF clearly exhibits a hydrophilic vestibule on the periplasmic side of the glycerol conduction pathway. This structural asymmetry is reflected in an attractive well in the PMF, which also shows a high degree of asymmetry. Kinetic rate models are built on the basis of the PMF determined in Jensen et al. (2002) to investigate the glycerol conduction rate and the role of asymmetry. The glycerol conduction rate calculated compares favorably with the observed value. The attractive well is found to increase the flux at 10 μM and 10 mM glycerol concentrations by >40% and 75%, respectively, which is beneficial for nutrient uptake. The question of why the attractive well faces the periplasm and not the cytoplasm cannot be directly explained by either flux or efficiency. The clogging hypothesis also appears not to answer this question. Some factor other than the optimization of glycerol conduction, such as cell protection, may be responsible for selecting the preferred orientation. As the structural asymmetry is also present in other porins, our answers given should also be relevant for other membrane channels.

APPENDIX A: CONNECTION RULE AT THE BOUNDARIES

In our model study, the population of glycerol in the interior of the channel is represented by the occupation, ρ_i, but at the periplasm and cytoplasm, it corresponds to the glycerol concentration, C. The two quantities must be connected at the boundaries.

According to the continuum model (Cooper et al., 1985), the flux can be written as

$$J = CS\left(\int_{z_1}^{z_2} \frac{e^{\beta E(z)}}{D}\, dz\right)^{-1}, \tag{42}$$

where S is the opening area of the channel. In our rate model,

$$J = \frac{D}{(z_2 - z_1)^2}\rho_0 e^{(1/2)\beta[E(z_1)-E(z_2)]}. \tag{43}$$

By matching Eq. 42 with Eq. 43, we conclude

$$\rho_0 = C\,S(z_2 - z_1)^2 e^{(1/2)\beta[E(z_2)-E(z_1)]}\left(\int_{z_1}^{z_2} e^{\beta E(z)}\, dz\right)^{-1}. \tag{44}$$

The opening area of GlpF is roughly 100 Å^2, and the ratio of ρ_0/C is 407 Å^3 at the periplasmic side and 415 Å^3 at the cytoplasmic side. We use the average value 411 Å^3 in our calculation.

APPENDIX B: STEADY-STATE SOLUTION OF THE TWO-STEP MODEL

Starting from Eqs. 16–17, we assume $\rho_b \ll 1$. In this case the master equation under stationary conditions reduces to

$$\rho_w = \rho_b[(1-\rho_w)T_{21} + T_{23}]/T_{12}, \tag{45}$$

$$\rho_w = (1-\rho_w)(\rho_0 T_{01} + \rho_b T_{21})/[T_{12} + (1-\rho_0)T_{10}]. \tag{46}$$

The assumption $\rho_b \ll 1$ is in general valid considering the barrier is as high as 7 k_BT, and it can be justified by our six-step model calculation. After eliminating ρ_w, we obtain the second-order equation of ρ_b,

$$\rho_b^2 + t_1\rho_b - t_2 = 0. \tag{47}$$

The parameters t_1 and t_2 are defined in Eqs. 21–22. The solution of Eq. 47 is well-known.

APPENDIX C: DERIVATION OF DIFFUSION COEFFICIENT FROM MD SIMULATION

As the relevant timescale of glycerol's motion in our study is many times longer than the microscopic collision times, the motion of single glycerol inside GlpF is described by the one-dimensional Langevin equation (Risken, 1989) with a potential of mean force $U(x)$,

$$\ddot{x} + \gamma\dot{x} + \frac{1}{m}\nabla U(x) = A(t), \tag{48}$$

where γ is the friction coefficient, and $A(t)$ is the Langevin force. The properties of $A(t)$ are characterized through the ensemble averages

$$\langle A(t)\rangle = 0, \tag{49}$$

$$\langle A(t)A(t')\rangle = q\delta(t-t'). \tag{50}$$

The diffusion coefficient is related to γ and q by Einstein relation,

$$\tilde{D} = \frac{q}{2\gamma^2}. \tag{51}$$

When a glycerol molecule is trapped at the free energy minima, as we can see from Fig. 3 *a* is true for most of the time, the potential $U(x)$ is approximated well by a harmonic potential. One can then replace $U(x)$ by $(1/2)kx^2$ in Eq. 48, and by introducing $\beta = k/m$, the Langevin equation assumes the simple form

$$\ddot{x} + \gamma\dot{x} + \beta x = A(t). \tag{52}$$

We denote the velocity by $v = \dot{x}$ and define the velocity autocorrelation function as

$$C(t) = \frac{\langle [v(t) - \langle v\rangle][v(0) - \langle v\rangle]\rangle}{\langle v^2\rangle}. \tag{53}$$

$\tilde{C}(\omega)$, the spectral density of $C(t)$, can be evaluated from Eq. 52 as

$$\tilde{C}(\omega) = \int_{-\infty}^{\infty} dt\, C(t) e^{i\omega t} \tag{54}$$

$$= \frac{q}{\langle v^2\rangle \left[\gamma^2 + \left(\omega - \frac{\beta}{\omega}\right)^2\right]}. \tag{55}$$

Substituting Eq. 51 into Eq. 55, one immediately gets

$$\tilde{D} = \frac{1}{2}\tilde{C}(\sqrt{\beta})\langle v^2\rangle. \tag{56}$$

To record the velocity trajectories of glycerol molecules, a 50-ps MD simulation was performed for an NPT ensemble using NAMD2 (Kalé et al., 1999). In this simulation, we took advantage of a previous simulation of a GlpF tetramer (see system 1 in Jensen et al., 2001) by using exactly the same system and simulation parameters. Accordingly there were three glycerols per GlpF monomer and 12 glycerols in total. We also employed the restart files from Jensen et al. (2001), so that the starting point of our simulation corresponds to the 250-ps instance in Fig. 3 *a*. Velocity trajectories are discarded where the potentials are far from a harmonic potential. This results in nine trajectories used in our analysis. The spectra of the resulting velocity autocorrelation functions were fitted to Eq. 55 using the NonlinearFit package of Mathematica (Wolfram, 1996). The diffusion coefficient, the average of nine fits, was found to be $\tilde{D} = (2.2 \pm 0.3) \times 10^{-5}\ \mathrm{cm}^2/\mathrm{s}$.

The authors are grateful to Prof. Mark F. Schumaker for his help to refine the steady-state model and bring the work of Agmon and Hopfield, 1983, to their attention. They also thank Dr. Mario Borgnia, Dr. Emad Tajkhorshid, Mr. Sanghyun Park, Mr. Fangqiang Zhu, Ms. Yan Li, Mr. Paul Wiggins, and Dr. Ioan Kosztin for valuable discussions.

This work was supported by the National Institutes of Health (PHS5 P41RR05969) and the National Science Foundation (MCB-9982629). The authors also acknowledge computer time provided by the National Resource Allocations Committee grant MCA93S028.

REFERENCES

Agmon, N., and J. J. Hopfield. 1983. Transient kinetics of chemical reactions with bounded diffusion perpendicular to the reaction coordinate: intramolecular processes with slow conformational changes. *J. Chem. Phys.* 78:6947–6959.

Binder, K. (editor). 1986. Monte Carlo Methods in Statistical Physics, 2nd Ed. Springer-Verlag, Berlin, Germany.

Borgnia, M., S. Nielsen, A. Engel, and P. Agre. 1999. Cellular and molecular biology of the aquaporin water channels. *Annu. Rev. Biochem.* 68:425–458.

Borgnia, M. J., and P. Agre. 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.

Breed, J., R. Sankararamakrishnan, I. D. Kerr, and M. S. P. Sansom. 1996. Molecular dynamics simulations of water within models of ion channels. *Biophys. J.* 70:1643–1661.

Cooper, K., E. Jakobsson, and P. Wolynes. 1985. The theory of ion transport through membrane channels. *Prog. Biophys. Mol. Biol.* 46:51–96.

Cowan, S. W., T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R. A. Pauptit, J. N. Jansonius, and J. P. Rosenbusch. 1992. Crystal structures explain functional properties of two *E. coli* porins. *Nature.* 358:727–733.

de Groot, B. L., A. Engel, and H. Grubmüller. 2001. A refined structure of human aquaporin-1. *FEBS Lett.* 504:206–211.

de Groot, B. L., and H. Grubmüller. 2001. Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. *Science.* 294:2353–2357.

Forst, D., W. Welte, T. Wacker, and K. Diederichs. 1998. Structure of the sucrose-specific porin ScrY from *Salmonella typhimurium* and its complex with sucrose. *Nat. Struct. Biol.* 5:37–46.

Fu, D., A. Libson, L. J. W. Miercke, C. Weitzman, P. Nollert, J. Krucinski, and R. M. Stroud. 2000. Structure of a glycerol conducting channel and the basis for its selectivity. *Science.* 290:481–486.

Fujiyoshi, Y., K. Mitsuoka, B. L. de Groot, A. Philippsen, H. Grubmüller, P. Agre, and A. Engel. 2002. Structure and function of water channels. *Curr. Opin. Struct. Biol.* 12:509–515.

Grayson, P., E. Tajkhorshid, and K. Schulten. 2003. Mechanisms of selectivity in channels and enzymes studied with interactive molecular dynamics. *Biophys. J.* 85:36–48.

Heller, K. B., E. C. Lin, and T. H. Wilson. 1980. Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. *J. Bacteriol.* 144:274–278.

Hille, B. 1992. Ionic Channels of Excitable Membranes, 2nd Ed. Sinauer Associates, Sunderland, MA.

Hohmann, S., S. Nielsen, and P. Agre. 2001. Aquaporins. Academic Press, San Diego, CA.

Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.

Jensen, M. Ø., S. Park, E. Tajkhorshid, and K. Schulten. 2002. Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. USA.* 99:6731–6736.

Jensen, M. Ø., E. Tajkhorshid, and K. Schulten. 2001. The mechanism of glycerol conduction in aquaglyceroporins. *Structure.* 9:1083–1093.

Kalé, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: greater scalability for parallel molecular dynamics. *J. Comp. Phys.* 151:283–312.

Kozono, D., M. Yasui, L. S. King, and P. Agre. 2002. Aquaporin water channels: atomic structure and molecular dynamics meet clinical medicine. *J. Clin. Invest.* 109:1395–1399.

Kreusch, A., A. Neubuser, E. Schiltz, J. Wechesser, and G. Schulz. 1994. Structure of the membrane channel porin from *rhodopseudomonas blastica* at 2.0 Å resolution. *Prot. Sci.* 3:58–63.

Law, R. J., and M. S. P. Sansom. 2002. Water transporters: how so fast yet so selective? *Curr. Biol.* 12:R250–R252.

Lide, D. R. (editor). 1994. CRC Handbook of Chemistry and Physics, 75th Ed. CRC Press, Boca Raton, FL.

Mitscherlich, E., and E. H. Marth. 1984. Microbial Survival in the Environment. Springer-Verlag, Berlin, Germany.

Morais-Cabral, J. H., Y. Zhou, and R. MacKinnon. 2001. Energetic optimization of ion conduction rate by the K^+ selectivity filter. *Nature.* 414:37–42.

Nadler, W., A. Brünger, K. Schulten, and M. Karplus. 1987. Molecular and stochastic dynamics of proteins. *Proc. Natl. Acad. Sci. USA.* 84:7933–7937.

Neidhart, F. C. (editor). 1996. *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. ASM Press, Washington, DC.

Nollert, P., W. E. C. Harries, D. Fu, L. J. W. Miercke, and R. M. Stroud. 2001. Atomic structure of a glycerol channel and implications for substrate permeation in aqua(glycero)porins. *FEBS Lett.* 504:112–117.

Rickey, D. P., and E. E. C. Lin. 1972. Importance of facilitated diffusion for effective utilization of glycerol by *Escherichia coli*. *J. Bacteriol.* 112:784–790.

Risken, H. 1989. The Fokker-Planck Equation: Methods of Solution and Applications, 2nd Ed. Springer, New York.

Ritz, T., S. Park, and K. Schulten. 2001. Kinetics of excitation migration and trapping in the photosynthetic unit of purple bacteria. *J. Phys. Chem. B*. 105:8259–8267.

Roux, B. 2002. Theoretical and computational models of ion channels. *Curr. Op. Struct. Biol.* 12:182–189.

Sansom, M. S. P., I. D. Kerr, J. Breed, and R. Sankararamakrishnan. 1996. Water in channel-like cavities: structure and dynamics. *Biophys. J.* 70: 693–702.

Schumaker, M. F. 2002. Boundary conditions and trajectories of diffusion processes. *J. Chem. Phys.* 117:2469–2473.

Schumaker, M. F., R. Pomès, and B. Roux. 2000. A combined molecular dynamics and diffusion model of single proton conduction through gramicidin. *Biophys. J.* 79:2840–2857.

Schumaker, M. F., R. Pomès, and B. Roux. 2001. Framework model for single proton conduction through gramicidin. *Biophys. J.* 80:12–30.

Şener, M. K., D. Lu, T. Ritz, S. Park, P. Fromme, and K. Schulten. 2002. Robustness and optimality of light harvesting in cyanobacterial Photosystem I. *J. Phys. Chem. B*. 106:7948–7960.

Tajkhorshid, E., P. Nollert, M. Ø. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud, and K. Schulten. 2002. Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science*. 296:525–530.

Tieleman, D. P., and H. J. C. Berendsen. 1998. A molecular dynamics study of the pores formed by *Escherichia coli* OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. *Biophys. J.* 74:2786–2801.

Voegle, R. T., G. D. Sweet, and W. Boos. 1993. Glycerol kinase of *Escherichia coli* is activated by interaction with the glycerol facilitator. *J. Bacteriol.* 175:1087–1094.

Wang, Y., R. Dutzler, P. Rizkallah, J. Rosenbusch, and T. Schirmer. 1997. Channel specificity: structural basis for sugar discrimination and differential flux rates in maltoporin. *J. Mol. Biol.* 272:56–63.

Weiss, M. S., T. Wacker, J. Weckesser, W. Welte, and G. E. Schulz. 1990. The three-dimensional structure of porin from *Rhodobacter capsulatus* at 3 Å resolution. *FEBS Lett.* 267:268–272.

Wolfram, S. 1996. The Mathematica Book, 3rd Ed. Wolfram Media/ Cambridge University Press, NY.

Zhu, F., E. Tajkhorshid, and K. Schulten. 2001. Molecular dynamics study of aquaporin-1 water channel in a lipid bilayer. *FEBS Lett.* 504:212–218.

Zhu, F., E. Tajkhorshid, and K. Schulten. 2002. Pressure-induced water transport in membrane channels studied by molecular dynamics. *Biophys. J.* 83:154–160.