

Mechanisms of Selectivity in Channels and Enzymes Studied with Interactive Molecular Dynamics

Paul Grayson, Emad Tajkhorshid, and Klaus Schulten
Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois

ABSTRACT Interactive molecular dynamics, a new modeling tool for rapid investigation of the physical mechanisms of biological processes at the atomic level, is applied to study selectivity and regulation of the membrane channel protein GlpF and the enzyme glycerol kinase. These proteins facilitate the first two steps of *Escherichia coli* glycerol metabolism. Despite their different function and architecture the proteins are found to employ common mechanisms for substrate selectivity: an induced geometrical fit by structurally homologous binding sites and an induced rapid dipole moment reversal. Competition for hydrogen bonding sites with water in both proteins is critical for substrate motion. In glycerol kinase, it is shown that the proposed domain motion prevents competition with water, in turn regulating the binding of glycerol.

INTRODUCTION

Computer simulation provides a powerful means for investigation of biochemical reactions and biological systems at an atomic level. In silico experiments can now be done on a wide variety of systems without the practical limitations that one may face in experimental approaches. A researcher is not only able to examine a native system of interest, he can also set up simulations with exaggerated or artificial conditions to further test proposed mechanisms. The main obstacles in atomistic simulations, however, have been the restriction to very short timescales and the inability of equilibrium simulations to provide statistically satisfactory samplings. These limitations have prevented us from providing a complete description, including energetics, of many biological events. The problem is greatly magnified when one wants to explore a variety of hypotheses and mechanisms.

Equilibrium molecular dynamics (MD) simulations of average-size biological systems are nowadays limited to a few tens of nanoseconds. The timescale is obviously much shorter than what one would need to model major events, such as major conformational changes of macromolecules, permeation of solutes through channels, and (un)binding of ligands to their binding sites. Due to the described limitations, so-called biased, or guided, MD simulations were developed in which a particular event is induced through the application of biasing potentials or external forces that effectively guide the system toward a predefined configuration. When performed slowly and carefully enough, such simulations allow one to study the response of the system and calculate changes in its free energy. Two such methods, umbrella sampling (Torrie and Valleau, 1977; McCammon and Harvey, 1987; Mezei, 1989; Kumar et al., 1992) and steered molecular dynamics (Izrailev et al., 1998; Lu et al., 2000; Isralewitz et al., 2001a,b) simulations, have been suc-

cessfully applied to a variety of molecular systems. Although the precision of the calculated free energy changes is dependent on sufficient sampling, the resulting picture of the dynamics of the system has always proven useful in understanding the details of the event. In practice, steered molecular dynamics (SMD) and umbrella sampling simulations require a predefined reaction coordinate that connects the initial and the final configurations, e.g., the position of an atom along a vector or its angle of rotation about a defined axis.

Interactive molecular dynamics (IMD) is a novel simulation method for interactively manipulating biomolecular systems to probe dynamic processes (Nelson et al., 1995; Rapaport, 1997; Ai and Frölich, 1998; Prins et al., 1999; Vormoor, 2001; Stone et al., 2001). The method uses the same force fields and classical dynamics as conventional MD and is most closely related to SMD: in IMD and SMD, external forces are added to the forces between atoms (Isralewitz et al., 2001a). External forces are used in both IMD and SMD to probe the mechanical functions of biomolecules or to accelerate processes that are otherwise too slow to model within the timescale accessible to MD simulations. In contrast to SMD, the applied forces in IMD simulations are continuously adjusted by the researcher, who monitors the dynamics through a visual display of the simulation. This interactive adjustment allows the researcher to apply forces of varying direction and magnitude to different atoms at every step of the simulation, to induce the simulated event. The simulation speed required for real-time interaction is provided readily by modern computer platforms. In fact, IMD simulations described in this article have been made possible mainly by the development of software that integrates MD with molecular graphics, not through raw computer speed. IMD simulations can be effectively carried out on small clusters of commodity computers or even on single personal computers (Stone et al., 2001).

Though IMD and SMD are closely related techniques, their differences suggest how each is best applied. SMD simulations have been employed most successfully in studies of

Submitted November 15, 2002, and accepted for publication February 25, 2003.

Address reprint requests to Klaus Schulten, kschulte@ks.uiuc.edu.

© 2003 by the Biophysical Society

0006-3495/03/07/36/13 \$2.00

biomechanical processes where simulations directly mimicked cellular function and experiment, for example, rupture of the biotin-avidin bond (Grubmüller et al., 1996; Izrailev et al., 1997), unbinding of antigen and antibody (Heymann and Grubmüller, 1999), or stretching of an immunoglobulin domain of titin (Lu et al., 1998; Marszalek et al., 1999). SMD was also employed successfully in exploration of putative reaction pathways, for example, a lipophilic route of retinal binding in bacteriorhodopsin (Izrailev et al., 1999) and back door mechanisms in actin (Wriggers and Schulten, 1999) and nuclear hormone receptors (Kosztin et al., 1999). SMD simulations proceed in three stages. First, a hypothesis is developed regarding the mechanism of the biological process of interest. The hypothesis is based on the investigator's a priori understanding and leads to a specification of external forces to apply to the system. Second, the simulation is conducted with the weakest force possible to complete the process within an affordable amount of computer time, usually days or weeks. Third, analysis of the simulation relates the applied force to the progress of the system along the chosen reaction path. IMD simulations differ in all three respects. First, simulations are unplanned, the goal being to develop a hypothesis. Second, simulations are fast, because the applied forces are strong enough to complete a simulation in minutes. Third, simulations are rough and do not usually produce data suitable for an accurate analysis of molecular properties. Quantitative analysis is readily provided by further MD or SMD simulations that build on the hypotheses developed during an IMD session. The main advantage of IMD simulations is that the intuition of the researcher is part of the simulation, generating and quickly testing hypotheses for reaction mechanisms as well as speeding up substrate transport, docking, or other manipulations that are difficult to cast into numerical programs.

Fig. 1 shows an IMD session. Our system for IMD (Stone et al., 2001) connects a molecular visualization program, VMD (Humphrey et al., 1996), to an MD simulation program, NAMD2 (Kalé et al., 1999). In this study, we used a personal computer for the visualization, and a cluster of 32 1333-MHz Athlon CPUs running Scyld Linux for the simulations. These computational resources allow the simulation of systems containing several thousand atoms at a speed sufficient for interaction. As shown in Fig. 1, steering forces are issued through a haptic device, allowing the researcher to apply any force in three dimensions and, through forces felt in his hand, feel how the atoms move in response.

This article describes the first application of IMD to a research problem. We use IMD to investigate the function and selectivity mechanisms of two biological systems, the *E. coli* glycerol uptake facilitator (GlpF) and glycerol kinase (GK), two very different proteins which facilitate the first two steps of *E. coli* glycerol metabolism. In both GlpF and GK, the part of the protein that interacts directly with its substrate contains only ~ 1000 atoms. The small scale of this



FIGURE 1 Our system for IMD. A customizable view of a running simulation is displayed on a desktop PC as it is computed. The PC displays a pointer that corresponds to the position of the haptic device used to interact with the system. A researcher uses this haptic device to apply a downward force to a simulated molecule of glycerol, feeling resistance in his hand as it moves through the channel.

region allows 0.1 ps of dynamics to be simulated on the computing cluster every second, making meaningful real-time interaction with the simulation feasible.

GlpF is a member of the aquaporin family of channels used for water transport across the cell membrane. Aquaporins are present in all life forms, and more than 100 members of the family have been found to date (Borgnia et al., 1999). Eleven human aquaporins have been identified so far, and their impaired function is 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 transporting water, GlpF facilitates the conduction of small linear sugar molecules such as glycerol. GlpF is very selective in its conduction. Large molecules, ions, and even protons are completely excluded, and the channel favors certain small linear sugars over others by two orders of magnitude, even distinguishing between stereoisomers (Fu et al., 2000). GlpF is thought to be used by *E. coli* to accelerate the absorption of glycerol from the environment for use in metabolism, especially when there is little glycerol available (Richey and Lin, 1972).

The recent three-dimensional structure of GlpF (Fu et al., 2000) has allowed the selectivity of this channel to be investigated. The inner lining of the channel consists of a hydrophobic face complemented by hydrogen binding groups on the opposite side, corresponding to the amphiphilic structure of glycerol. The narrowest region of the channel is located close to the periplasmic mouth of the channel and is proposed to function as its *selectivity filter* (Fu et al., 2000; Nollert et al., 2001; Law and Sansom, 2002; Fujiyoshi et al., 2002). More recently, aquaporin-1 (AQP1), a water channel that does not conduct glycerol, has also yielded to high resolution structure determination (Sui et al., 2001). Its structure shows that AQP1 is narrower and more hydrophilic than GlpF, sacrificing glycerol conductivity to

achieve a maximal rate of water transport (Murata et al., 2000; de Groot et al., 2001; Sui et al., 2001). The three-dimensional structures have already led to several MD simulations that explain the properties of conduction and selectivity in aquaporins (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). In particular, MD simulations have resolved the long-standing question of proton exclusion, showing how aquaporins differ from water channels such as gramicidin A (Roux, 2002) that conduct protons: a strict bipolar orientation of water molecules dictated by the electric field of the channel was proposed to prevent proton conduction via the Grothüss mechanism (Tajkhorshid et al., 2002). Other researchers attributed the proton exclusion to the disruption of the water file, mainly at the selectivity filter (de Groot and Grubmüller, 2001).

GK is an enzyme that catalyzes the phosphorylation of glycerol after it is conducted into the cytoplasm (Lin, 1996). Phosphorylation is necessary for glycerol to enter the glycolytic pathway and prevents the molecule from being conducted back into the extracellular space, because phosphorylated glycerol is not conducted through GlpF. However, glycerol can only be used at a certain maximum rate by the cell, and excessive production of glycerol phosphate leads to poisoning and cell death. In addition, glucose is preferred to glycerol, so glycerol kinase must be shut down when glucose is readily available (Zwaig et al., 1970; Holtman et al., 2001). The catalytic activity of GK is controlled through binding to the regulatory protein IIA^{Glc} (Roseman and Meadow, 1990) and fructose 1,6-bisphosphate (Zwaig et al., 1970), which are present whenever further phosphorylation of glycerol is not desirable. GK is thought to adopt two distinct conformational states, the *open* and *closed* forms. A dimer with at least one *open* monomer is thought to represent the active form of the enzyme, and the inhibitors are thought to act by stabilizing a *closed* tetrameric form (Byström et al., 1999). There are also indications that the activity of GK is increased by the presence of GlpF (Voegle et al., 1993), but no specific mechanism has been suggested.

Over the past decade, thirteen different crystallographic structures for GK have been solved. Most of the structures (Hurley et al., 1993; Feese et al., 1994; Ormö et al., 1998; Mao et al., 1999) are thought to represent the inactive, *closed* conformation. A single GK structure (PDB accession code 1GLJ) shows what may be a partially open form of GK that is similar to the *open* active form. The *closed* and *open* forms are related by a simple rigid-body motion similar to those observed in hexokinase and actin (Byström et al., 1999; Holmes et al., 1993; Pettigrew et al., 1998). The structures have been used to explain the function of GK in *E. coli* and its role in human GK deficiency (Dipple et al., 2001), but there have been no attempts in these efforts to use simulation.

In this article we demonstrate how IMD can be applied to study the interaction of sugar molecules with GlpF and GK. Through IMD, we reveal several interesting properties that allow GlpF and GK to act selectively and we clarify the mechanism of regulation of GK. A previous SMD study (Jensen et al., 2002) characterized the energetics of conduction of glycerol through GlpF. We extend this study by examining the conduction properties of two pentatols, ribitol and arabitol, that are known to be affected by strong stereoselectivity in GlpF (Fu et al., 2000). Since the pentatols are significantly longer than glycerol, unfavorable steric interactions or hydrogen bonding mismatches are amplified. As a result, we expect the mechanisms of stereoselectivity to appear more clearly for the pentatols than for glycerol. GK has not previously been studied through simulation, so we examine the binding properties of its natural substrate, glycerol.

METHODS

Simulations of GlpF

Simulations of GlpF used the structure reported by Fu et al. (2000). As described earlier, a GlpF tetramer was inserted into a membrane, fully solvated, and equilibrated for 1 ns (Jensen et al., 2001, 2002). To simulate the system at the speed necessary for interaction, it was necessary to make several simplifications. All but a single GlpF monomer and 119 water molecules were removed. The number of water molecules included in the system ensured the availability of enough water to replace substrate along its movement through the channel. Since the outer residues of the channel do not move significantly during the conduction of sugar (Jensen et al., 2001), it was not necessary to allow them to move in the simulation. A single molecule of ribitol or arabitol was positioned in the periplasmic space at the beginning of each interactive simulation. The simplified system, containing 4218 atoms, 923 of which were free to move, is shown in Fig. 2. The interactive simulations used the Charmm27 force field (MacKerell Jr. et al., 1992, 1998). The system was maintained at a temperature of 310 K through Langevin damping with a coefficient of 10 ps^{-1} , and the nonbonded interactions were simulated with a 10-Å cutoff.

Four interactive simulations of GlpF were performed, as listed in Table 1. Ribitol and arabitol were each pulled slowly through the channel, once in each direction, over $\sim 1 \text{ h}$ of real time (100–160 ps of simulated time). Constraints were applied to distances between neighboring water molecules in the channel to ensure that their single file remained intact. The sugar molecules were forced into a variety of conformations along the channel, and potentially important steps in their conduction pathways were selected from these conformations. Four steps in the conduction pathway of ribitol are shown in Fig. 2; a total of 11–16 steps were selected from each of the four interactive simulations as the starting points of the noninteractive simulations described below.

Additional 100-ps noninteractive simulations were performed to sample the dynamics of ribitol and arabitol at each step in the conductions. Only the last 88 ps of these simulations were used for analysis. For the noninteractive simulations, the constraints on the water molecules were removed, and one additional constraint was added: the center of mass of the sugar molecule was restricted to move no farther than 0.5 Å from its starting position along the channel axis. Table 1 lists the times required for the noninteractive simulations. The noninteractive simulations were performed under the same conditions as the interactive simulations, except that the Langevin damping coefficient was reduced to 5 ps^{-1} . The simulations were run on single 500-MHz Alpha processors running Tru64 UNIX.

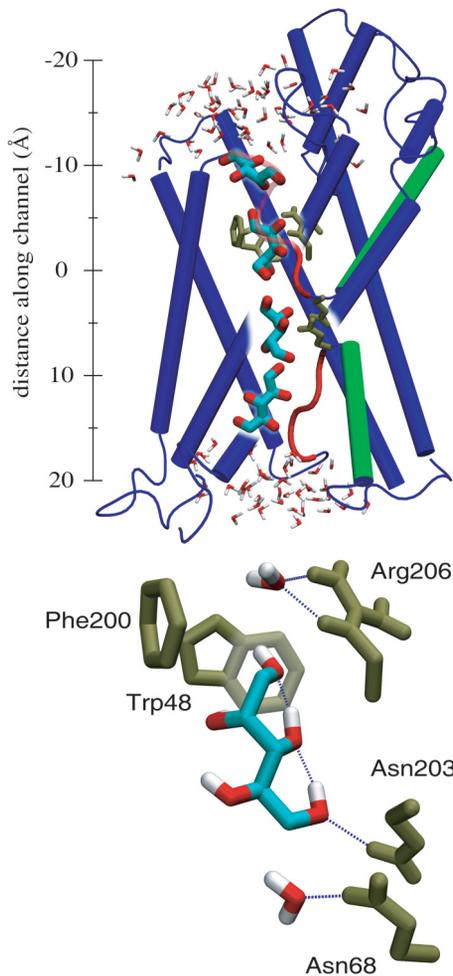


FIGURE 2 Pathway for pentitol conduction through GlpF, revealed through IMD. (Top) Pathway for ribitol (C1 first). The half-helices M3 and M7 are shown as green cylinders, connected through the NPA motifs to the half-membrane-spanning nonhelical sections (red tubes). Water in the periplasm (top) and cytoplasm (bottom) is shown in licorice representation. Snapshots of ribitol are shown in a larger licorice representation at four key locations: entrance from the periplasmic vestibule (top), selectivity filter (Trp48, Phe200, and Arg206 shown in gold), NPA motifs (Asn68 and Asn203, also in gold); and exit into the cytoplasm. (Bottom) An expanded view of the selectivity filter and NPA motifs, showing one step in the C5-first conduction of arabititol, at which arabititol adopts an alternating conformation. The two neighboring water molecules are visible, and hydrogen bonds are drawn as dotted lines. Bonds are formed between arabititol's hydroxyl groups and from hydroxyl groups to protein. Motion of the sugar up or down would cause arabititol to exchange hydrogen bonds with the water molecules. All figures of steps in conduction in this article use the equilibrated structures from the end of the noninteractive simulations.

Glycerol kinase system construction

Simulations of GK were based on the high-resolution crystal structure reported by Feese et al. (1998) (PDB code 1GLF). The structure represents a tetramer of GK; the most complete monomer was used to construct the system. The 59 bound water molecules and a single bound glycerol in the crystal structure were preserved. The monomer was solvated and neutralized in a periodic rhombic dodecahedral box containing 6607 water molecules and 12 sodium ions, bringing the system to a total of 27,822 atoms. In

TABLE 1 Simulated pentitol conduction and glycerol unbinding events

| Event name | IMD | | MD | |
|----------------|------------|-------|------------|-------|
| | Simulation | Real | Simulation | Real |
| Ribitol-1 | 143 ps | 1.1 h | 1.2 ns | 28 h |
| Ribitol-5 | 134 ps | 1.0 h | 1.6 ns | 35 h |
| Arabitol-1 | 158 ps | 1.0 h | 1.1 ns | 24 h |
| Arabitol-5 | 96 ps | 0.8 h | 1.1 ns | 20 h |
| GK preparation | – | – | 3.0 ns | 166 h |
| GK-closed | 237 ps | 0.6 h | 1.5 ns | 41 h |
| GK-open | 124 ps | 0.3 h | 1.2 ns | 34 h |
| GK-reversal | 49 ps | 0.1 h | 0.1 ns | 2 h |
| Totals | 0.9 ns | 5.0 h | 10.8 ns | 350 h |

Simulated events are, from top to bottom: ribitol conduction, with carbon 1 (C1) entering the channel first; ribitol, C5 first; arabitol, C1 first; arabitol, C5 first; equilibration, opening, and closing of GK; glycerol unbinding from closed GK; glycerol unbinding from open GK; and the end-for-end reversal of glycerol in closed GK. The GK interactive simulations ran significantly faster, mainly due to the use of a 2-fs timestep and a smaller system size.

simulations of the entire monomer, the system was maintained at a temperature of 310 K (through Langevin damping with a coefficient of 5 ps^{-1}) and a pressure of 1 atm (using the Langevin piston method). Van der Waals interactions were calculated using a 10-Å cutoff, and the Particle Mesh Ewald method was employed to compute electrostatic interactions. Fig. 3 shows the system after 1 ns of equilibration in preparation for IMD.

Comparison of the closed (PDB code 1GLF) and partially open GK (PDB code 1GLJ) revealed that a hinge rotation by ~ 10 degrees links the two structures. The rotating domains were defined by inspection, and the hinge axis was determined with independent root-mean-square deviation (RMSD) alignments of the two domains between 1GLF and 1GLJ. The domains are shown in Fig. 3.

To establish a model for a completely open GK the equilibrated system, here denoted as $1GLF_c$, was subjected for 1 ns to angular constraints applied to the $C\alpha$ atoms that rotated the domains around the hinge axis by 20° . The open structure produced after 1 ns is denoted by $1GLF_c(1 \text{ ns})$. To check that the constrained rotation did not irreversibly distort GK, $1GLF_c(1 \text{ ns})$ was subjected for another 1 ns to angular constraints that reversed the former 20° rotation. Fig. 4 shows that GK opened smoothly and reversibly. Also shown is an RMSD comparison of the structure of $1GLF_c(t)$ ($1GLF_c$ during its opening and closing motion) with the structures 1GLF (closed) as well as with 1GLJ (partially open). The RMSD values were determined here only for the $C\alpha$ atoms. As expected, $1GLF_c(t)$ differs equally from 1GLF and 1GLJ at time $t \approx 250 \text{ ps}$, i.e., when the domains of $1GLF_c(t)$ have been rotated apart by 5° . The structure of $1GLF_c(t)$ corresponds most closely to 1GLJ at 1 ns, and is again closer to 1GLF than to 1GLJ at 2 ns. Opening and closing added only $\sim 0.5 \text{ \AA}$ to the RMSD value. However, there were several local changes visible during opening and closing. In particular, after $\sim 0.1 \text{ ns}$ of closing, several water molecules moved between the bound glycerol and Asp245, and remained in the binding pocket through the closing simulation.

The simulations required for the MD and SMD simulations of GK were performed on a cluster of 32 1333-MHz Athlon CPUs running Scyld Linux and the Terascale Computing System at the Pittsburgh Supercomputing Center.

Simulations of GK

The interactive simulations of GK were conducted with simulation parameters identical to those of GlpF, except that covalent bonds to hydrogen atoms were constrained to fixed lengths with SHAKE (Ryckaert

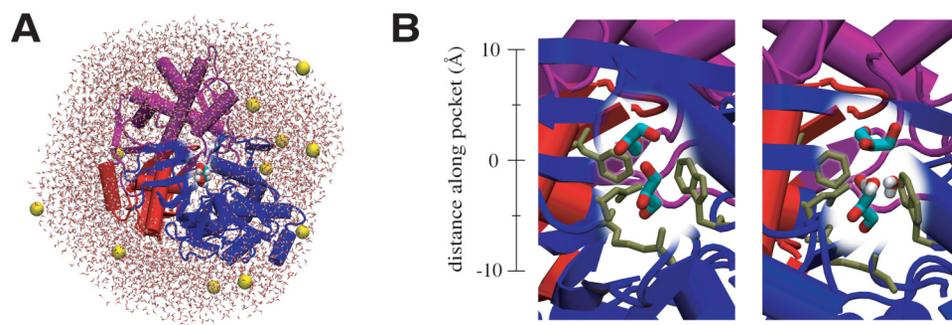


FIGURE 3 Pathway of glycerol unbinding from GK, as revealed through IMD. (A) The equilibrated monomer of glycerol kinase used in this study, in its *closed* (inactive) form. GK is drawn in a cartoon representation, colored according to its division into three domains: the central hinge (*red*) and the two regions (*purple* and *blue*) that were tilted apart into the *open* (active) form. Ions are drawn as yellow VDW spheres, and glycerol is visible in VDW representation near the center of the

protein. All of the water molecules solvating the protein are drawn, using thin sticks. (B) An expanded view of the GK binding pocket in the *closed* (left) and *open* (right) conformations. Glycerol is drawn in licorice representation using two snapshots from the interactive simulation, the initial bound state (below) and the final unbound state (above), to illustrate its unbinding pathway. The residues forming the selectivity filter of the GK binding pocket clockwise from the right, in *gold*: Trp103, Arg83, Asp245, and Phe270. All water molecules within 3 Å of the second hydroxyl group of the initial glycerol are drawn in licorice representation. No water molecules are this near the group in the *closed* conformation, although two are present in the *open* conformation.

et al., 1977) so that the system could be simulated with a 2-fs time step. For these simulations, the *molten-zone method* of Prins et al. (1999) was used: only atoms near to glycerol were “molten,” i.e., allowed to move, and atoms farther than 8 Å were removed entirely from the simulation. This method and the small size of the GK binding pocket relative to the GlpF pore allowed a molten zone containing as few as 200–300 atoms.

Two interactive simulations of GK were performed, as listed in Table 1: one on the *closed* structure 1GLF_c and one on the modeled *open* structure 1GLF_o (1 ns). In each case, the glycerol molecule was slowly pulled out of the binding region over ~200 ps of simulated time (about one half hour of real time) and forced into various conformations along the unbinding pathway. Key steps during the unbinding were selected from these conformations and used as the starting points for noninteractive simulations described below. The first and last steps that were selected are shown in Fig. 3. The binding of glycerol in its reversed orientation was investigated with an additional IMD simulation in which glycerol was pulled out of the *closed* ar/R region, reversed, and returned to a bound state.

Additional 100-ps noninteractive simulations were performed on each step of the unbinding of glycerol from GK, in both the *open* and *closed* conformations. Water and protein residues near the unbinding pathway, including ~1000 atoms, were allowed to move in the noninteractive simulations. All water and protein atoms within 8 Å of the molten-zone atoms, ~3000 atoms, were included in the simulation but held fixed. The simulation parameters used here were identical to those used for the noninteractive simulations of GlpF. Table 1 lists the times required for the noninteractive simulations, which were each run on single processors of SunBlade 2000 workstations.

Software tools for IMD

As mentioned earlier, our IMD simulations on systems of biological scale were made possible through the integration of available software. Visualization and the control of applied forces was mediated entirely by VMD, hiding NAMD almost entirely from the user. The VRPN software package (Taylor II, 1998) was used to connect VMD to a commercially available haptic device, the SensAble Technologies PHANToM Desktop, which must be controlled by a separate computer. A variety of other input devices are supported by VRPN. VMD allows the user to perform IMD within a full-featured molecular visualization environment while an advanced MD simulation program is running in the background. A recent addition to VMD, AutoIMD, automates the construction of interactive simulations, especially those using the molten-zone method. Particularly useful is the ability to select the “molten” section of a system, initiate an IMD simulation, stop,

save the results, and continue, entirely within a graphical interface. All of the software required to perform IMD, including AutoIMD and compatibility with VRPN, is integrated into the VMD and NAMD packages and available on our website: <http://www.ks.uiuc.edu/>.

RESULTS AND DISCUSSION

IMD was used to characterize the conduction pathways of ribitol and arabitol through GlpF (Fig. 2) and the unbinding pathways of glycerol from *open* and *closed* GK (Fig. 3). This required, altogether, 0.9 ns of simulation performed over 5 h of interaction. Essential steps in these pathways were identified, and the dynamics of the substrate was sampled at each position through 100-ps noninteractive MD simulations. The noninteractive simulations, 4.9 ns in total, were used to analyze the selectivity mechanisms in GlpF and GK. To prepare a structure representing *open* GK, a 2-ns SMD simulation of its opening and closing motion was performed. Table 1 lists the interactive and noninteractive simulations carried out in this study.

Conduction of ribitol and arabitol in GlpF

The IMD simulations of sugar conduction through GlpF begin with a molecule of ribitol or arabitol positioned in vacuum, near the water on the periplasmic side of the channel, as shown at the top of Fig. 2. Initially, a minute is typically spent arranging the molecular representations shown in the graphical display. The scale of the displayed coordinates (γ) and the scale of the force on the haptic device (β) are adjusted so that several hundred piconewtons can be comfortably applied to the molecule by hand. It is desirable for a user of IMD to be able to feel the resistance of the atoms pulled, but too much resistance will make it physically difficult to manipulate the system. The apparent mass m^* of a simulated atom of mass m is strongly dependent on α , the scale of time in the simulation (Stone et al., 2001):

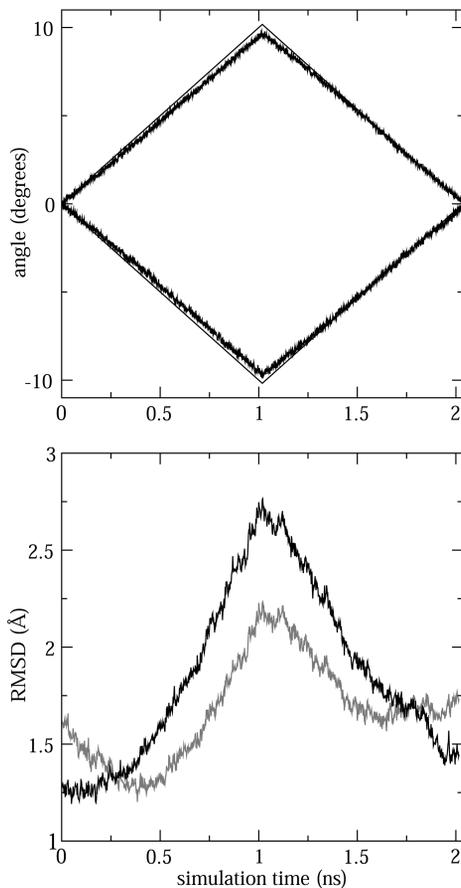
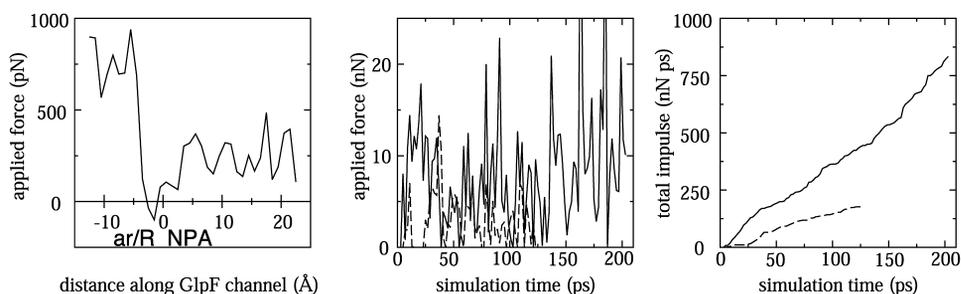


FIGURE 4 Simulation of the opening and closing of GK. (Top) The constraint angles (straight lines) and average relative angles of the constrained C α carbons (rough lines) in GK domains I (above) and II (below) during the simulation. (Bottom) The aligned C α -only RMSD between the simulated structure and the closed (black) and open (gray) crystal structures (from 1GLF and 1GLJ). The minima, maxima, and crossings of the two curves can be used to assess the progress of the simulation (see text).

$$m^* = \frac{\alpha^2}{\beta\gamma} m. \quad (1)$$

Faster simulations therefore allow much lower forces to be applied to the atoms, improving the accuracy of the results. By the time these initial adjustments are made, several picoseconds have already elapsed in the simulation—enough



to extract glycerol from its binding site, averaged in bins of 2-ps width. (Right) The integral of the total force shown in the middle graph over time, with the total impulse applied to the system. Significantly more force needed to be applied for a longer time to extract glycerol from closed GK than from open GK.

time for the sugar molecule to diffuse toward the water and begin forming hydrogen bonds. After several more minutes of simulation, glycerol is completely dissolved by the water. However, thermal motion is unlikely to cause the sugar to enter the channel and diffuse across in a reasonable amount of time. To accelerate this process, the pointer of the haptic device is moved to the position in which its three-dimensional cursor points to the sugar molecule on the screen, as shown in Fig. 1, and a button on the pointer is pressed. Immediately, a simulated spring links the pointer to the simulation. Pulling the pointer downward causes a force to be applied to the sugar, so that it begins to move toward the channel. The force applied to the sugar is represented on the screen with a three-dimensional arrow and transmitted to the user's hand by the motors of the haptic device. This allows the sugar molecule to be steered to the entrance of the GlpF channel in seconds.

As the molecule of ribitol or arabitol is pulled into the channel, a strong resistance is felt in the user's hand, and he must push forcefully to overcome it. If the button on the haptic device is released at this point, detaching the simulated spring, the sugar molecule is observed to drift back out of the entrance and into the periplasmic water, pushed by repulsive interactions with the channel walls. However, after the sugar molecule has been pulled ~ 10 Å through the channel, the resistance lessens noticeably. This region of the channel is wider, as can be seen by pulling the sugar molecule from side to side to feel the channel walls. Even less downward resistance is felt another 20 Å later as the sugar is pulled out of the channel into the cytoplasmic water. Data collected during the IMD session (see Fig. 5) confirm that the maximum downward force must be applied to steer the sugar molecule past a region immediately below the periplasmic entrance. This is the narrowest part of the channel, a region that is proposed to act as its selectivity filter (Fu et al., 2000). The finding that the largest force is required to pass the selectivity filter is consistent with the results of a recent study in which the potential of mean force for glycerol conduction was constructed using a multinano-second SMD simulation of a membrane-embedded GlpF tetramer (Jensen et al., 2002). The potential of mean force exhibits the largest free energy barriers in the selectivity

FIGURE 5 Forces applied during IMD. (Left) The component of the force on ribitol and arabitol along the channel as a function of the distance along the channel, averaged in bins of width 1 Å. The scale of the graph corresponds to that in Fig. 2. Significantly more force needed to be applied before passing the selectivity filter (at ~ 5 Å) than after. (Middle) The magnitude of the total force applied to the closed (solid) and open (dashed) GK

filter, so it is natural that large forces are required to pass this region. IMD allows the largest barriers to be identified, however, during just a few minutes of interactive simulation.

The longer IMD simulations of GlpF described here, typically lasting ~ 1 h, allowed an initial identification of the important interactions involved in the conduction of ribitol and arabitol. All four simulated conduction events followed a similar path in which motion down the channel was accompanied by rotations of C-C bonds necessary to maintain hydrogen bonds with as many locations on the hydrophilic side of the channel as possible, similarly to the path described earlier (Jensen et al., 2001). As a result, two types of steering forces were applied with the haptic device. The primary force was oriented downward (toward the cytoplasm) and applied intermittently to encourage the substrate to move along the channel. Secondary forces pulled specific hydroxyl groups toward hydrogen bonding sites on the channel interior. These forces caused the substrate to change its conformation at each step along the channel and were often strong enough to push the sugar molecule into an unexpected position. Constant monitoring of the graphical display and adjustment of the forces applied with the haptic device was necessary to keep the substrate moving along a pathway deemed suitable.

Passage through the selectivity filter is the key step in the conduction of ribitol and arabitol. Four hydrogen bonding sites are available in this region: two hydrogen donor groups on the side chain of *Arg206* and two acceptor sites provided by the carbonyl groups of *Gly199* and *Phe200*. The numerous binding locations and the small size of the selectivity filter meant that much attention had to be paid at this point to the conformation of the substrates. After the first HCOH group entered the selectivity filter, the sugar molecules were moved downward step by step, so that only one or two hydrogen bonds needed to be broken at a time. Several conformations of the sugar molecules were attempted, to find those that maximized hydrogen bonding at each step.

The crystal structure of GlpF includes a glycerol bound to the selectivity filter with carbon 3 nearest to the cytoplasm (a C3-first orientation). In this orientation, glycerol is capable of making hydrogen bonds to all four available sites. Fu et al. (2000) suggested that the observed differences in the conduction rates of longer sugars are due to their differing abilities to form this optimal number of hydrogen bonds in the selectivity filter. The IMD simulations indeed did not indicate a stable conformation of arabitol in the selectivity filter that matched the binding of glycerol, explaining arabitol's low rate of conduction. Similarly, when ribitol passes the channel in its C5-first orientation, its glycerol-matching conformations are not stable. However, the C1-first conduction of ribitol included conformations with HCOH groups bound to all available sites in the selectivity filter, accounting for the high rate of conduction of ribitol through GlpF.

After passing the selectivity filter, the substrates move into contact with the NPA motifs at the channel center. In all four IMD simulations, the substrates assumed a conformation that

takes advantage of the increased space available in this region. The substrates straightened out by rotating around their C-C bonds into a fully extended configuration like that exhibited by arabitol in Fig. 2, with the plane of the carbon backbone parallel to the side of the channel containing the NPA motifs. The ability of the substrates to make hydrogen bonds with the NPA motifs (*Asn68* and *Asn203*) was then determined by their molecular geometry. Various conformations were tried as the substrates were pulled, one step at a time, through this region. Arabitol and ribitol could be steered into either of two possible parallel orientations by twisting their middle bond. Fig. 2 shows a sample conformation near the NPA motifs that was investigated. Interestingly, the molecules' extended conformation allowed hydroxyl groups that did not bind to the NPA motifs to bind to other hydroxyl groups of the substrate: the first hydroxyl group could bind to the third, the second group to the fourth, or the third group to the fifth. This resulted in several of the hydroxyl groups always pointing vertically, either up or down the channel. As the substrate was pulled past the NPA motifs, the favored direction for the hydroxyl groups reversed. This reversal is discussed further below.

Past the NPA motifs, the channel grows wider and there are two further significant hydrogen bonding sites for ribitol and arabitol: the exposed carbonyl oxygens of *Ala65* and *Gly64*. The sugar molecules were pulled one step at a time past these two binding sites, keeping the substrates in their extended conformations and making adjustments to maintain two hydrogen bonds with *Ala65* and *Gly64* whenever possible. As the substrates neared the cytoplasmic exit, water surrounded them on all sides, and they could be easily removed from the channel.

Unbinding of glycerol from *closed* and *open* GK

Before the IMD simulations of GK are described, the importance of water needs to be considered. In previous MD simulations, water was already found to play a key role in the transport of glycerol through GlpF: the competition between glycerol and water molecules for hydrogen bonding to the channel lining was identified as the stochastic driving force for glycerol conduction (Jensen et al., 2001). However, unlike the channel of GlpF, the GK binding pocket has only one opening, so glycerol blocks water from competing for hydrogen bonds at one end. The opening of GK could allow water to play the same role as in GlpF: in *open* GK, water molecules could facilitate glycerol unbinding after having gained access to the binding site, while unbinding from *closed* GK would not be aided by water. That opening the protein could release and admit glycerol was suggested as one of two possible regulatory mechanisms in Feese et al. (1998). Our SMD simulations support this hypothesis, indicating that opening motion of GK can help water reach the binding pocket. After GK was opened by 16° , the C3 end of glycerol rotated into a new position that allowed water to

come into contact with the second hydroxyl group of glycerol. Fluctuations between this conformation and the original conformation continued as GK was opened further, breaking and restoring the original hydrogen bond to Asp245. Eventually, several water molecules moved farther into the binding pocket. Once inside the pocket, they can facilitate glycerol unbinding by competing for hydrogen bonds with the first (most deeply buried) HCOH group of glycerol. To reproduce the unbinding of glycerol in IMD simulations, it was therefore necessary to apply forces not only to glycerol, but also to the surrounding water molecules.

At the beginning of the IMD simulation of glycerol unbinding from *open* GK, glycerol was visible in the binding pocket, held tightly by surrounding protein residues. In the base of the binding pocket, the side chains of Tyr135, Glu84, and Arg83 were in close contact with the hydroxyl group of C1, clearly preventing it from moving out of its bound location. The hydroxyl group belonging to C2 formed hydrogen bonds both to Asp245 and the backbone of Arg83, fixing the orientation of glycerol. The aromatic residues Phe270 and Trp103 made contact with the hydrophobic side of glycerol, so that the substrate was held tightly from all sides. At the exit of the pocket, near the ATP binding site, C3 was found in either of two positions: with or without a hydrogen bond to Asp245. Forces applied with the haptic device easily broke or established this hydrogen bond.

To accelerate the unbinding of glycerol from *open* GK, a water molecule was pulled between glycerol and Asp245 and then further into the pocket, causing it to break the hydrogen bonds between the C1 hydroxyl group and the protein side chains. Strong forces between the water molecule, glycerol, and the residues lining the binding pocket could be felt as the water molecule was moved into this region, guiding it into the original location of the C1 hydroxyl group. Further water molecules were forced to follow the same path, moving along the glycerol-GK hydrogen bonds and into the pocket, as glycerol moved in the opposite direction. As each water molecule was added to the binding pocket, forces were applied that pulled glycerol out to leave space for it. Each outward step corresponded to a shift in the hydrogen bonding pattern of glycerol with Asp245. The first water molecule caused C3 to unbind from Asp245, being replaced from below by C1. The second water molecule similarly removed C2 from Asp245, and after a third was added to the binding pocket, glycerol was no longer hydrogen-bonded to *open* GK.

The extraction of glycerol thus required three different types of forces. As in GlpF, force was applied intermittently to the sugar molecule to encourage it to move in the direction of the unbinding path; force was also applied to the substrate's hydroxyl groups to cause the conformational changes necessary at each step to maintain best hydrogen bonding. The most important type of applied force may have been the third—that which drove water molecules into the binding pocket.

In *closed* GK, a tighter association between glycerol and Asp245 was observed. Again, at the beginning of the simulation the C3-Asp245 hydrogen bond was broken, and a molecule of water was moved past the hydrophilic side chains into the bottom of the pocket. Glycerol was moved outward at each step, but did not break its hydrogen bonds to Asp245 until the fourth water molecule was added. The tighter binding in the *closed* form meant that the extraction of glycerol was more difficult during IMD, suggesting that the closure of GK locks glycerol in place. A more detailed analysis of the IMD simulations supports the hypothesis that *closed* GK prevents glycerol from reaching or leaving its binding pocket in two ways: first, as shown in Fig. 5, the applied forces required to extract glycerol were much higher in the *closed* form than the *open* form; second, analysis of MD simulation trajectories (see below) showed that glycerol is more flexible in the *open* form, increasing the likelihood for nearby water molecules to pass between it and the protein to reach the base of the binding pocket.

Selection for dipole flexibility in GlpF and GK

MD simulations (Tajkhorshid et al., 2002) of aquaporins revealed that water molecules are strictly aligned in the channels, but with a direction that reverses midway. This dipole alignment is the result of several structural elements contributing to a strong electrostatic field that swiftly reverses around the NPA motifs (Tajkhorshid et al., 2002). The electrostatic field attracts water molecules with a matching dipole moment, aligning them and counteracting the repulsive forces of the channel walls. It has been suggested that the dipole reversal of GlpF selectively attracts small molecules, such as water, because their dipole moments can easily reverse during conduction (de Groot and Grubmüller, 2001). This suggestion raises the question of whether dipole reversal is also an important feature for glycerol transport in GK and for conduction of linear sugars in GlpF.

The IMD simulations described here revealed that both GlpF and GK align the hydroxyl group dipoles of their sugar substrates, as shown in Fig. 6. One can discern a clear orientational preference for both systems that is reversed along the two pathways. Apparently, both proteins attract substrates with strong dipole moments but require the alignment to reverse over a short distance.

In GlpF, a strong dipole reversal is made possible by the extended conformation of the sugars that is adopted near the NPA motifs, orienting the hydroxyl groups along the length of the channel. As each hydroxyl group of ribitol and arabitol passes the NPA motifs, its alignment reverses to match the electrostatic field of the channel. The dipole moment orientation and reversal at the NPA motifs is clearly similar to that found for channel water (Tajkhorshid et al., 2002). We see here that even a large molecule such as a pentitol can have groups of sufficiently small, flexible dipoles to pass through the channel.

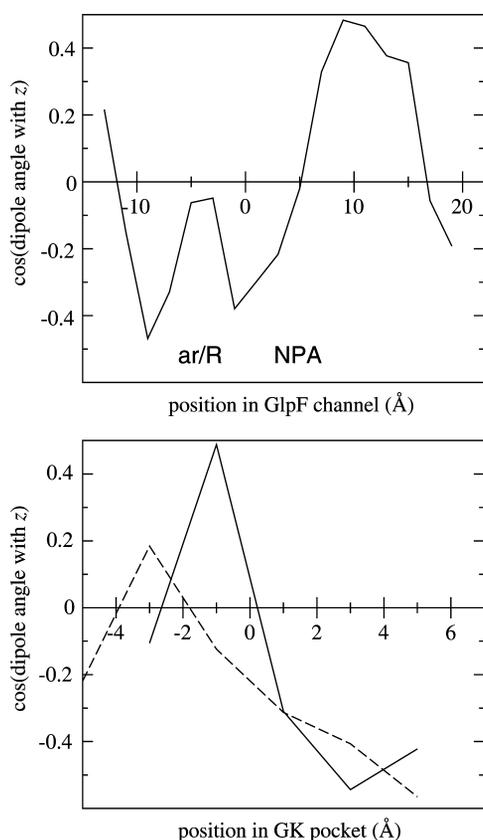


FIGURE 6 Orientation of the HCOH group dipole moments. The dipole moments were sampled in bins of 2-Å width from 88-ps simulations that started after the equilibration of steps in the pathways. In GlpF (*top*) one can discern a reversal of the dipole moments at the NPA motifs and a disruption of the alignment in the selectivity filter. In *closed* GK (*bottom, solid line*) there is a large dipole moment reversal at 0 Å; this corresponds to the location of the negatively charged residue *Asp245*. In *open* GK (*bottom, dashed line*) a dipole moment reversal is also present, but its magnitude is reduced. The distance axes of the graphs correspond to those in Figs. 2 and 3.

In GK, a dipole reversal, similar in magnitude to that in GlpF, arises at the location of the negatively charged *Asp245*. Fig. 6 shows that the strength of the alignment is greatly reduced in the *open* form of GK, in which glycerol was observed to have a weaker interaction with *Asp245*. Presumably, the electric field points toward *Asp245*, making it energetically favorable for the dipole moment of glycerol to also point toward this residue from either side. As was the case in GlpF, the hydroxyl groups are individually reversed as they pass this region, allowing glycerol to reach its binding location. A pathway optimized for an ion or a fixed-dipole substrate would have to be constructed quite differently than the pathway in GK.

Mechanisms of substrate selectivity in the ar/R region

To carry out their biological functions, GlpF and GK need to select their substrates with high specificity. In GlpF,

selectivity prevents undesirable molecules and ions from being conducted into the cell. In GK, selectivity is used to assure that only glycerol is bound and that glycerol is positioned properly for the formation of glycerol-3-phosphate. Selectivity is achieved incrementally during several stages of conduction (GlpF) and binding (GK), but the key stage is when the proteins establish the tightest fit around the substrate: for GlpF this occurs at the narrowest segment of the channel, aptly termed the selectivity filter (Fu et al., 2000); for GK this fit occurs at the bottom of the binding pocket. In both cases the protein presents hydrophobic and hydrophilic linings that match the amphiphilic structure of glycerol. Fig. 7 shows an amazing similarity between the GK binding pocket and the GlpF selectivity filter: both contain a hydrophobic corner on one side, formed by a pair of tryptophan and phenylalanine residues, and an arginine residue on the opposite side, that together form a ring around the bound sugar. Because of the combination of aromatic

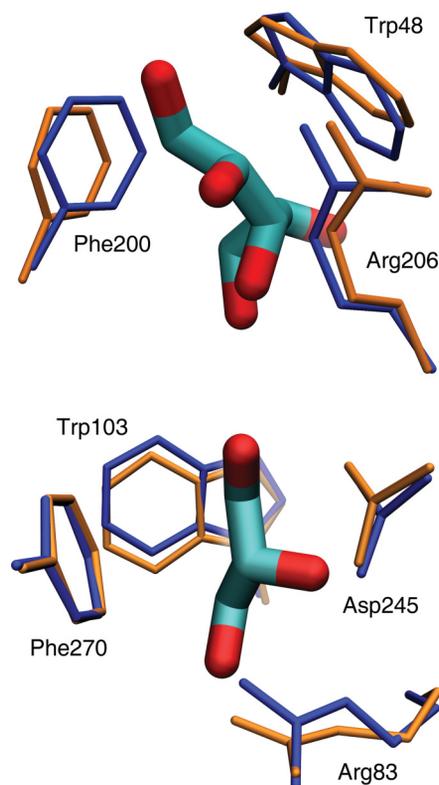


FIGURE 7 Orientation of the side groups that form the selectivity filter in GlpF (*top*) and the binding pocket in the *closed* form of GK (*bottom*), with (*orange*) and without (*blue*) ribitol and glycerol. Ribitol is shown with C1 entering first; glycerol is shown in its normal bound state, with C3 (*above*). The figure shows significant motion of the side groups in both proteins: the groups shown generally moved more in response to the motion of glycerol than any other protein side chains. The RMSD of the protein atoms shown was 0.6 Å for GlpF and 0.9 Å for GK. In GlpF, the motion is consistent with a widening necessary to hold ribitol. In GK, the extraction of glycerol causes *Trp103* to move toward the exit of the binding pocket, *Asp245* to tilt outward, and *Arg83* to move inward.

(*ar*) rings and an arginine (*R*) residue, we will refer to both as *ar/R* regions as suggested in de Groot and Grubmüller (2001). In GK an important aspartate interacts with glycerol in the *ar/R* region, but this aspartate may also be viewed as similar to two exposed carbonyl oxygens in the GlpF selectivity filter. Interestingly, a mutation study revealed that replacing the aspartate with arginine increases GK's binding affinity for glycerol (Pettigrew et al., 1998).

The IMD simulations described above indicate that the *ar/R* region facilitates precise substrate selectivity in both proteins. Selectivity in the *ar/R* region is provided by two cooperating features in this region, an induced fit and optimal hydrogen bonding of the proper substrate. Understanding the selectivity mechanism permits one to answer, for example, why GlpF conducts glycerol and ribitol very well but conducts arabitol at a 100-fold reduced rate (Fu et al., 2000).

Induced fit

Our IMD simulations revealed that the substrates fit more tightly into the *ar/R* region than any other part of the channel, showing that this region is unreachable by larger molecules and that it is the region where the most substrate-specific interactions are possible. The freedom of motion was quantified by analyzing the individual HCOH groups through 100-ps MD simulations at each step in the substrate pathways. Low-frequency, i.e., poorly sampled, motions were removed with a high-pass filter, and the root-mean-square deviation (RMSD) of the filtered motion was calculated. Fig. 8 shows the RMSD along the pathways in GlpF and GK. In both GlpF and GK, the substrate was held most tightly in the *ar/R* region. For GK, Fig. 8 shows that the protein binds glycerol more tightly in its *closed* conformation than in its *open* conformation, which suggests that the *closed* conformation fixes glycerol in a strictly optimal position for catalysis.

At the *ar/R* regions of GlpF and GK, the surrounding side chains move more than side chains in any other parts of the pathway to accommodate the substrates and establish a tight fit. This *induced fit*, as seen in our IMD simulations, is shown in Fig. 7. The motions observed in GlpF closely match those observed in the crystal structures of GlpF with (1FX8) and without (1LDA, 1LDI) bound glycerol (Tajkhorshid et al., 2002). This flexibility corresponds to a widening motion caused by steric repulsion between the substrate and the channel walls, consistent with our observation during IMD that ribitol and arabitol are typically pushed out of the selectivity filter in several picoseconds, if not constrained in the channel by an external force.

Attempts to prepare glycerol-free structures of GK have been unsuccessful (Huang et al., 2001), so our present IMD and MD simulations are necessary to quantify the response of GK to glycerol binding. As shown in Fig. 7, glycerol does not cause a significant widening of the *ar/R* region as in GlpF. The *open-closed* transition of GK explains why the

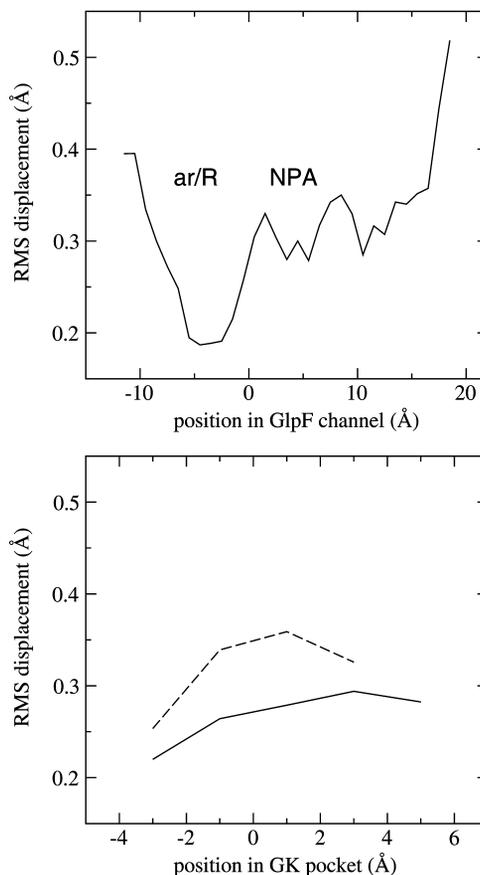


FIGURE 8 Fluctuation of individual HCOH groups perpendicular to the conduction/unbinding pathways. The root-mean-square (RMS) motion was calculated from 88-ps simulations that started after the equilibration of steps in the pathways. A high-pass filter was applied to eliminate motions with a period longer than 8.8 ps. (Top) The combined values for ribitol and arabitol conduction through GlpF, sampled in bins of width 1 Å. (Bottom) The values for the *open* (dashed lines) and *closed* (solid lines) forms of GK, sampled in bins of 2-Å width. In GlpF, fluctuations are reduced in the selectivity filter and to a lesser degree near the NPA motif and a third binding site at 11 Å. In GK, fluctuations are smallest in the binding pocket (<0 Å). The *open* form of GK shows consistently larger fluctuations than the *closed* form. The overall range of fluctuations in GK is close to the range in GlpF. The distance axes of the graphs correspond to those in Figs. 2 and 3.

ar/R region does not become narrower when glycerol unbinds. Experiments indicate that the presence of glycerol stabilizes GK in its *closed* conformation (Thorner and Paulus, 1973). When glycerol unbinds, GK reverts to its *open* conformation; i.e., there is no reason to expect a clear narrowing. The small movements seen in Fig. 7 may just be the beginning of this transition, leading ultimately to an induced fit much more dramatic than that in GlpF.

Optimal hydrogen bonding

The second cause of selectivity in the *ar/R* region results from the presence of multiple hydrogen bonding sites. For substrates capable of bonding to all available sites, the hydrogen bonds balance unfavorable steric interactions

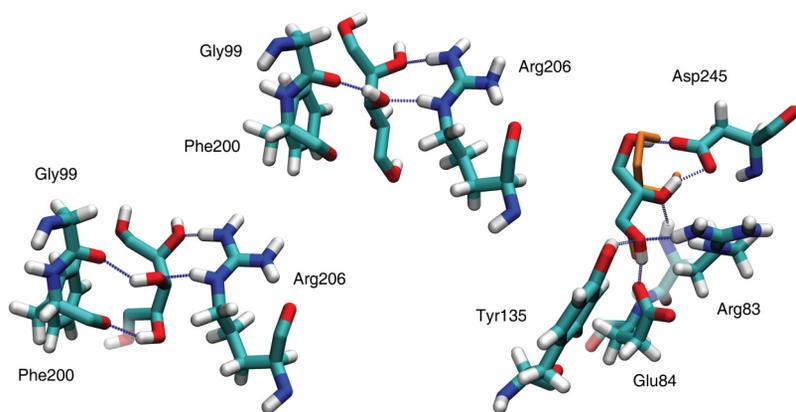


FIGURE 9 The hydrogen bonding patterns of sugar molecules in the ar/R regions of GlpF (*left, middle*) and GK (*right*). GlpF is shown with ribitol (C1 first, *left*) and arabinol (C5 first, *middle*). Ribitol and arabinol both make hydrogen bonds to Arg206 and the carbonyl oxygen of Gly199. Ribitol, however, also makes a hydrogen bond to the carbonyl oxygen of Phe200. GK is shown with glycerol in its natural orientation (C3, *above*) and reversed (C3, *below*; glycerol, in orange). The same hydrogen bonds are formed in both orientations. The two HCOH groups at the bottom of the GK binding pocket assume almost the same conformation in both orientations, whereas the HCOH group nearest to the ATP binding site (*above the visible region*) is rotated slightly in the reversed orientation.

resulting from the tightness of this region. Fu et al. (2000) suggested that the multiple hydrogen binding sites in the ar/R region of GlpF hold multiple parts of a passing sugar in place, and that this locking mechanism is responsible for the fact that GlpF has high conductivity for glycerol and ribitol, but not for arabinol. In addition, the stereocenter recognition model (Sundaresan and Abrol, 2002) suggests that five binding sites are necessary for specific stereoselectivity between molecules (such as the pentatols) with three stereocenters. Fig. 9 shows indeed that three hydroxyl groups in ribitol can form hydrogen bonds to the ar/R region of GlpF at once. Including bonds at both ends to water molecules in the channel, ribitol achieves hydrogen bonding with all five of its hydroxyl groups, showing that the requirements for the stereocenter recognition model can be met. Arabinol, however, was seen to be forced to break at least one hydrogen bond while passing the ar/R region in any orientation, explaining why arabinol is conducted more slowly than ribitol.

The IMD simulations of GK suggest a possible reason for its exclusive production of glycerol-3-phosphate instead of glycerol-1-phosphate. To be converted into glycerol-3-phosphate, glycerol must enter the binding pocket, C1-first, so that C3 will be nearest to the ATP binding site. Glycerol is shown in its C1-first orientation in all crystal structures of GK, so it was necessary to use IMD to investigate the binding of the reversed, C3-first glycerol. The simulations show that glycerol binds in nearly the same position when reversed. All of the same glycerol-GK hydrogen bonds are also made in the C3-first orientation, and the two most deeply buried HCOH groups overlap almost exactly, as shown in Fig. 9. The only difference between the two orientations is a rotation of the HCOH group at the end nearest to the ATP binding site. This rotation is accommodated by a corresponding rotation of Asp245. Since the simulations did not include ATP, they do not reveal how its binding may be affected by the reversal of glycerol. If, however, the binding of ATP is unaffected, this slight rotation will affect phosphate transfer in some way, possibly slowing phosphorylation.

CONCLUSION

Selectivity and regulation are essential principles for a wide range of biomolecular processes, but the underlying mechanisms are often still poorly understood at the atomic level. These principles are exhibited by the channel protein GlpF and the enzyme glycerol kinase (GK), the first proteins involved in *E. coli* glycerol metabolism. Here, we use interactive molecular dynamics (IMD), a new computational methodology widely applicable to biomolecular processes, to study these proteins. It is shown that, despite their different function and architecture, a need to physically transport glycerol leads to common mechanisms of substrate selectivity and transport.

Substrate transport in both proteins depends critically on competition with water for protein hydrogen bonding sites. In the *closed* conformation of GK, water molecules are unable to enter the glycerol binding pocket, impeding the binding and release of glycerol. Stabilizing the *closed* conformation, regulatory factors slow glycerol-binding, in turn slowing phosphorylation.

The IMD simulations reveal that tight binding to numerous hydrogen bonding sites causes selectivity in both proteins. GK binds glycerol in a narrow binding pocket, offering hydrogen bonding sites that hold glycerol tightly in a precise position. The narrowest part of the GlpF channel provides many hydrogen bonding sites, holding passing molecules tightly in place. Glycerol and ribitol pass this region in a conformation that allows the molecules to make a maximum number of hydrogen bonds at each step. The hydrogen bonding ability of glycerol and ribitol explains their experimentally observed high rate of passage relative to arabinol, which shows reduced binding at all steps in conduction.

Both proteins have an additional source of selectivity: an electrostatic field that points in varying directions along the substrate pathway. The desired substrates, the linear sugar molecules, align their flexible dipoles with the fields. Undesirable molecules without flexible dipoles are unable to follow the pathway.

The authors thank Cory Bystrom for helpful advice about glycerol kinase. Molecular images in the article were created with the program VMD (Humphrey et al., 1996).

This work was supported by the National Institutes of Health Research grants PHS-5-P41RR05969 and R01-GM60946, as well as the National Science Foundation supercomputer time grant NRAC MCA93S028.

REFERENCES

- Ai, Z., and T. Fröhlich. 1998. Molecular dynamics simulation in virtual environments. *Comp. Graph. Forum*. 17:C267–C273.
- 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.
- Bystrom, C. E., D. W. Pettigrew, B. P. Branchaud, P. O'Brien, and S. J. Remington. 1999. Crystal structures of *Escherichia coli* glycerol kinase variant S58→W in complex with nonhydrolyzable ATP analogues reveal a putative active conformation of the enzyme as a result of domain motion. *Biochemistry*. 38:3508–3518.
- 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.
- Dipple, K. M., Y.-H. Zhang, B.-L. Huang, L. L. McCabe, J. Dallongeville, T. Inokuchi, M. Kimura, H. J. Marx, G. O. Roederer, V. Shih, S. Yamaguchi, I. Yoshida, and E. R. B. McCabe. 2001. Glycerol kinase deficiency: evidence for complexity in a single gene disorder. *Hum. Genet.* 109:55–62.
- Feese, M., D. W. Pettigrew, N. D. Meadow, S. Roseman, and S. J. Remington. 1994. Cation-promoted association of a regulatory and target protein is controlled by protein phosphorylation. *Proc. Natl. Acad. Sci. USA*. 91:3544–3548.
- Feese, M. D., H. R. Faber, C. E. Bystrom, D. W. Pettigrew, and S. J. Remington. 1998. Glycerol kinase from *Escherichia coli* and an Ala65→Thr mutant: the crystal structures reveal conformational changes with implications for allosteric regulation. *Structure*. 6:1407–1418.
- 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.
- Grubmüller, H., B. Heymann, and P. Tavan. 1996. Ligand binding and molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*. 271:997–999.
- Heymann, B., and H. Grubmüller. 1999. An02/dnp-hapten unbinding forces studied by molecular dynamics atomic force microscopy simulations. *Chem. Phys. Lett.* 303:1–9.
- Holmes, K. C., C. Sander, and A. Valencia. 1993. A new ATP-binding fold in actin, hexokinase, and Hsc70. *Trends Cell Biol.* 3:53–59.
- Holtman, C. K., A. C. Pawlyk, N. D. Meadow, and D. W. Pettigrew. 2001. Reverse genetics of *Escherichia coli* glycerol kinase allosteric regulation and glucose control of glycerol utilization in vivo. *J. Bacteriol.* 183:3336–3344.
- Huang, H.-S., T. Inoue, K. Ito, and T. Yoshimoto. 2001. Preliminary crystallographic study of *Thermus aquaticus* glycerol kinase. *Acta Crystallogr. D*. 57:1030–1031.
- Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
- Hurley, J. H., H. R. Faber, D. Worthylake, N. D. Meadow, S. Roseman, D. W. Pettigrew, and S. J. Remington. 1993. Structure of the regulatory complex of *Escherichia coli* III^{Glc} with glycerol kinase. *Science*. 259:673–677.
- Isralewitz, B., J. Baudry, J. Gullingsrud, D. Kosztin, and K. Schulten. 2001a. Steered molecular dynamics investigations of protein function. *J. Mol. Graph. Modeling*. 19:13–25.
- Also *In Protein Flexibility and Folding*. Biological Modeling Series L. A. Kuhn and M. F. Thorpe, editors. Elsevier.
- Isralewitz, B., M. Gao, and K. Schulten. 2001b. Steered molecular dynamics and mechanical functions of proteins. *Curr. Op. Struct. Biol.* 11:224–230.
- Izrailev, S., A. R. Crofts, E. A. Berry, and K. Schulten. 1999. Steered molecular dynamics simulation of the Rieske subunit motion in the cytochrome *bc₁* complex. *Biophys. J.* 77:1753–1768.
- Izrailev, S., S. Stepanians, M. Balsera, Y. Oono, and K. Schulten. 1997. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* 72:1568–1581.
- Izrailev, S., S. Stepanians, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, and K. Schulten. 1998. Steered molecular dynamics. *In Computational Molecular Dynamics: Challenges, Methods, Ideas*. Vol. 4, Lecture Notes in Computational Science and Engineering. P. Deuffhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel, editors. Springer-Verlag, Berlin, Germany. pp.39–65.
- 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.
- Kosztin, D., S. Izrailev, and K. Schulten. 1999. Unbinding of retinoic acid from its receptor studied by steered molecular dynamics. *Biophys. J.* 76:188–197.
- 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.
- Kumar, S., D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comp. Chem.* 13:1011–1021.
- Law, R. J., and M. S. P. Sansom. 2002. Water transporters: how so fast, yet so selective? *Curr. Biol.* 12:R250–R252.
- Lin, E. C. C. 1996. Dissimilatory pathways for sugars, polyols, and carboxylates. *In Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. F. C. Neidhart, editor. ASM Press, Washington, DC. pp.307–326.
- Lu, H., B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* 75:662–671.
- Lu, H., A. Krammer, B. Isralewitz, V. Vogel, and K. Schulten. 2000. Computer modeling of force-induced titin domain unfolding. *In Elastic Filaments of the Cell*. J. Pollack and H. Granzier, editors. Kluwer Academic/Plenum Publishers, New York, NY. pp.143–162.
- MacKerell, A. D., Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-hydrogen empirical potential for molecular modeling and dynamics studies of proteins using the CHARMM22 force field. *J. Phys. Chem. B*. 102:3586–3616.
- MacKerell, A. D., Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus.

1992. Self-consistent parameterization of biomolecules for molecular modeling and condensed phase simulations. *FASEB J.* 6:A143.
- Mao, C., Z. Ozer, M. Zhou, and F. Uckun. 1999. X-ray structure of glycerol kinase complexed with an ATP analog implies a novel mechanism for the ATP-dependent glycerol phosphorylation by glycerol kinase. *Biochem. Biophys. Res. Commun.* 259:640–644.
- Marszalek, P. E., H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, and J. M. Fernandez. 1999. Mechanical unfolding intermediates in titin modules. *Nature.* 402:100–103.
- McCammon, J. A., and S. C. Harvey. 1987. Dynamics of Proteins and Nucleic Acids. Cambridge University Press, Cambridge, UK.
- Mezei, M. 1989. Evaluation of the adaptive umbrella sampling method. *Mol. Sim.* 3:301–313.
- Murata, K., K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. B. Heymann, A. Engel, and Y. Fujiyoshi. 2000. Structural determinants of water permeation through aquaporin-1. *Nature.* 407:599–605.
- Nelson, M., W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R. Skeel, K. Schulten, and R. Kufryn. 1995. MDScope—a visual computing environment for structural biology. *Comp. Phys. Comm.* 91:111–134.
- 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.
- Ormö, M., C. E. Bystrom, and S. J. Remington. 1998. Crystal structure of a complex of *Escherichia coli* glycerol kinase and an allosteric effector fructose 1,6-biphosphate. *Biochemistry.* 37:16565–16572.
- Pettigrew, D. W., G. B. Smith, K. P. Thomas, and D. C. Dodds. 1998. Conserved active site aspartates and domain-domain interactions in regulatory properties of the sugar kinase superfamily. *Arch. Biochem. Biophys.* 349:236–245.
- Prins, J. F., J. Hermans, G. Mann, L. S. Nyland, and M. Simons. 1999. A virtual environment for steered molecular dynamics. *Fut. Gen. Comp. Sys.* 15:485–495.
- Rapaport, D. C. 1997. Interactive molecular dynamics. *Physica A.* 240:246–254.
- Richey, D. P., and E. C. C. Lin. 1972. Importance of facilitated diffusion for effective utilization of glycerol by *Escherichia coli*. *J. Bacteriol.* 112:784–790.
- Roseman, S., and N. D. Meadow. 1990. Signal transduction by the bacterial phosphotransferase system. *J. Biol. Chem.* 265:2993–2996.
- Roux, B. 2002. Computational studies of the gramicidin channel. *Acc. Chem. Res.* 35:366–375.
- Ryckaert, J.-P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comp. Phys.* 23:327–341.
- Stone, J., J. Gullingsrud, P. Grayson, and K. Schulten. 2001. A system for interactive molecular dynamics simulation. In *The 2001 ACM Symposium on Interactive 3D Graphics*. J. F. Hughes and C. H. Séquin, editors. New York. ACM SIGGRAPH, New York. pp.191–194.
- Sui, H., B.-G. Han, J. K. Lee, P. Walian, and B. K. Jap. 2001. Structural basis of water-specific transport through the AQP1 water channel. *Nature.* 414:872–878.
- Sundaresan, V., and R. Abrol. 2002. Towards a general model for protein-substrate stereoselectivity. *Prot. Sci.* 11:1330–1339.
- 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.
- Taylor II, R. M. 1998. VRPN: virtual reality peripheral network. <http://www.cs.unc.edu/Research/vrpn>.
- Thomer, J. W., and H. Paulus. 1973. Catalytic and allosteric properties of glycerol kinase from *Escherichia coli*. *J. Biol. Chem.* 248:3922–3932.
- Torrie and Valleau. 1977. Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J. Comp. Phys.* 23:187–199.
- 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.
- Vormoor, O. 2001. Quick and easy interactive molecular dynamics using Java3D. *Comp. Sci. Eng.* 3:98–104.
- Wriggers, W., and K. Schulten. 1999. Investigating a back door mechanism of actin phosphate release by steered molecular dynamics. *Prot. Struct. Func. Gen.* 35:262–273.
- 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.
- Zwaig, N., W. S. Kistler, and E. C. C. Lin. 1970. Glycerol kinase, the pacemaker for the dissimilation of glycerol in *Escherichia coli*. *J. Bacteriol.* 102:753–759.