**Mechanism of selectivity in aquaporins and aquaglyceroporins**

**Jochen S. Hub and Bert L. de Groot**

**Computational Biomolecular Dynamics Group, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany**

**Edited by Robert M. Stroud, University of California, San Francisco, CA, and approved October 30, 2007 (received for review August 15, 2007)**

Aquaporins and aquaglyceroporins form a family of pore proteins that facilitate the efficient and selective flux of small solutes across biological membranes. We studied the selectivity of aquaporin-1 (AQP1) and the bacterial glycerol facilitator, GlpF, for $\text{O}_2$, $\text{CO}_2$, $\text{NH}_3$, glycerol, urea, and water. Using molecular dynamics simulations, we calculated potentials of mean force for solute permeation along the aquaporin channels and compared them with the alternative pathway across the lipid bilayer. For small solutes permeating through AQP1, a remarkable anticorrelation between permeability and solute hydrophobicity was observed, whereas the opposite trend was observed for permeation through the membrane. This finding renders AQP1 a selective filter for small polar solutes, whereas GlpF was found to be highly permeable for small solutes and permeable for larger solutes. Surprisingly, not solute-channel but water-channel interactions were found to be the key determinant underlying the selectivity mechanism of aquaporins. Hence, a hydrophobic effect, together with steric restraints, determines the selectivity of aquaporins.

**Results**

**Solute Permeability.** Fig. 2 presents the PMFs for permeation of $\text{O}_2$, $\text{CO}_2$, $\text{NH}_3$, glycerol, and urea permeating through hAQP1 (black curves) and GlpF (red curves). The PMFs $G(z)$ are shown as a function of the pore coordinate $z$, where $z = 0$ corresponds to the center of mass of the conserved NPA motifs.

Fig. 2 shows that the main barrier for each solute permeating through hAQP1 or GlpF is located in the $\alpha R$ region, demonstrating its role as the selectivity filter for uncharged solutes. For all solutes, the barrier in GlpF at the $\alpha R$ site is lower than the barrier in hAQP1. Hence, GlpF is more permeable to the solutes and is less selective than hAQP1. In contrast to the $\alpha R$ region, the highly conserved NPA motifs are not involved in the selectivity for uncharged solutes because they form only a minor barrier against permeation of the rather apolar $\text{O}_2$ and $\text{CO}_2$ molecules. The heights of the main barriers are summarized in Table 1.

---

**Author contributions:** J.S.H. and B.L.d.G. designed research; J.S.H. performed research; J.S.H. analyzed data; and J.S.H. and B.L.d.G. wrote the paper.

The authors declare no conflict of interest.

This article contains supporting information online at www.pnas.org/cgi/content/full/0707662104/DC1.

© 2008 by The National Academy of Sciences of the USA
To address the physiological relevance of permeation through aquaporins, the barriers were compared with barriers against permeation across the lipid bilayer. Fig. 3 shows the PMFs for permeation of urea, glycerol, H₂O, NH₃, CO₂ and O₂ through two phospholipid membranes, one composed of pure 1-palmityl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) and the other of pure 1-palmityl-2-oleoyl-sn-glycero-3-phosphocholine (POPC). The highest barrier for the apolar O₂ and CO₂ molecules permeating through either POPE or POPC is $\leq 6$ kJ/mol, indicating that these membranes are highly permeable to small apolar molecules such as O₂ and CO₂ (Fig. 3, green and yellow curves). The barriers against O₂ and CO₂ permeation through hAQP1 and GlpF are significantly higher than the corresponding membrane barriers. Hence, these aquaporin channels embedded in membranes similar to POPE or POPC are not expected to increase O₂ or CO₂ flux across the membrane.

The barriers against ammonia permeation through POPE (19 kJ/mol) or POPC (15 kJ/mol) are considerably higher than the barriers for apolar gas molecules, reflecting its more hydrophilic nature. Compared with O₂ or CO₂, these barriers refer to a permeability reduced by a factor of 80–400. Here we assumed the permeability to be proportional to $\exp(-\Delta G_{\text{max}}/k_B T)$, where $\Delta G_{\text{max}}$ denotes the maximum barrier height, $k_B$ denotes the Boltzmann constant, and $T$ denotes the temperature. When permeating through hAQP1 and GlpF, ammonia encounters barriers of 18 and 12.5 kJ/mol, respectively, indicating that hAQP1 is unlikely to increase NH₃ permeation, whereas GlpF can significantly enhance the NH₃ flux across the membrane. This observation is in line with an experimental study on ammonia permeation across the oocyte membrane (13). In that study, no influence of AQP1 on NH₃ flux was measured, whereas aquaglyceroporins such as AQP3 and AQP9 were found to increase NH₃ permeation.

Both glycerol and urea encounter substantial barriers between 27 and 34 kJ/mol against permeation through POPE or POPC, underlining the need for membrane channels if a substantial glycerol or urea flux is required by the metabolism. As expected, permeation of urea, glycerol, H₂O, NH₃, CO₂ and O₂ through two phospholipid membranes, one composed of pure 1-palmityl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) and the other of pure 1-palmityl-2-oleoyl-sn-glycero-3-phosphocholine (POPC).
GlpF is permeated efficiently by glycerol with a maximum barrier of 13.5 kJ/mol. This value is significantly lower than the activation energy of 30.5 kJ/mol reported in a previous molecular dynamics (MD) study (9). Compared with membranes of pure POPE or POPC, the barrier of 13.5 kJ/mol results in an increase in the permeability by a factor of 750 or 230, respectively, when assuming dense expression of GlpF. These values compare favorably with the reported increase by a factor of 400 for proteoliposomes reconstituted with GlpF (24). In contrast, urea encounters a large barrier of 29 kJ/mol against permeation through GlpF, confirming that GlpF is not an efficient urea channel. Solvating GlpF in a membrane of POPE or POPC is expected to increase the permeability for urea only 3- or 7-fold, respectively, in agreement with an only 3-fold increase measured for liposomes (24). As expected, hAQP1 does not enhance urea flux across the membrane because its ar/R region forms a barrier against urea permeation of more than 30 kJ/mol.

After inserting glycerol into the ar/R region of hAQP1, we observed a widening of the pore during the energy minimization (EM) of the structure. The widening was visible from the RMSD drift of the protein atoms during the EM and did not occur in any other simulation (data not shown). These findings confirm that the ar/R site of hAQP1 is too narrow to allow passage of glycerol. The unphysical widening of the pore led to an underestimation of the barrier by 24 kJ/mol for glycerol permeation through hAQP1 (Fig. 2, dashed line). The true barrier is likely to be substantially higher, which is in line with experiments that did not observe glycerol permeation through AQP1 (25).

The role of the ar/R region as the selectivity filter for uncharged solutes can be further investigated by point mutations in the ar/R region. Here we studied the H180A/R195V double mutant of hAQP1 (hAQP1-HA/RV) because H180 and R195 in the ar/R region. Here we studied the H180A/R195V double mutant of hAQP1 because H180 and R195 are generally considered key residues for the selectivity of AQP1. A recent experimental study investigated permeation properties of several hAQP1 mutants, including hAQP1-HA/RV (23). The authors found that some of the mutations enhanced NH3 permeability or allowed urea or glycerol permeation through hAQP1. To further rationalize these findings, Fig. 2 shows the PMFs for O2, CO2, NH3, glycerol, and urea permeating through hAQP1-HA/RV as green curves. For all solutes, the barrier in the ar/R region is drastically reduced by the mutation, emphasizing the importance of the ar/R residues for channel selectivity. In contrast to the wild type, the main barriers are not located in the ar/R region anymore. O2 and CO2 can almost freely diffuse through hAQP1-HA/RV. In agreement with ref. 23, the mutation renders hAQP1 an efficient NH3 channel and allows passage of urea, whereas water permeability is hardly affected [Table 1 and supporting information (SI) Fig. 6]. In addition, our results indicate that glycerol is likely to permeate through hAQP1-HA/RV, although ref. 23 reported glycerol permeation through the hAQP1-F56A/H180A mutant only.

**Solute Hydrophobicity.** Fig. 4A displays the height of the main barrier for urea, glycerol, H2O, NH3, CO2, and O2 against permeation through hAQP1 as a function of solute hydrophobicity. The solute hydrophobicity is measured as the logarithm of the hexadecane/water partition coefficient log Khex of the solute, which is a common measure for hydrophobicity. The larger and hydrophilic glycerol and urea molecules are not excluded by hydrophobicity, but rather stericly. Hence, the ar/R region of hAQP1 can be considered as both a hydrophobicity filter and a size filter.

Fig. 4B shows the corresponding plot for GlpF. The larger and more hydrophobic ar/R region of GlpF is less selective than the ar/R region of hAQP1, forming a significant barrier only against permeation of urea, but efficiently allowing the passage of glycerol. The difference between Fig. 4A and B demonstrates that the differing residues in the ar/R regions of hAQP1 and GlpF account for highly different permeation characteristics.

Fig. 4C shows the energetic cost ΔGtails for moving the solute from bulk water to the hydrophobic lipid tail region of a POPC membrane. As expected, the bilayer forms a filter permeated by hydrophobic molecules. The linear dependence between ΔGtails and log Khex favorably compares to Overton’s rule, stating that...
the permeability of a membrane with respect to different solutes is approximately proportional to their oil/water partition coefficients. For comparison, the dashed line indicates the energetic cost \( \Delta G_{\text{hex}} = -k_B T \ln K_{\text{hex}} \) for moving the solute from bulk water into hexadecane.

**Water-Protein Interactions as Selectivity Mechanism.** As demonstrated in the previous section, the ar/R region of hAQP1 is a filter permeated by small polar molecules. A previous MD study showed that the barrier that CO\(_2\) encounters in the ar/R region is not because of unfavorable CO\(_2\)-protein interactions (22). Instead, water–Arg-195 hydrogen bonds were observed to open upon CO\(_2\) passage. Here we investigate the selectivity mechanism of hAQP1 in detail. As an example, we consider a permeation event of the apolar O\(_2\) molecule (Fig. 5). To guide the eye, Fig. 5A shows the PMFs for O\(_2\) permeating through hAQP1 and GlpF, respectively. Fig. 5B presents the interaction between water and key residues of the ar/R region versus the position of an O\(_2\) molecule inside the channel. In Fig. 5B, water–protein interactions in hAQP1 are presented, whereas in Fig. 5B Right, the interactions in GlpF are presented. Likewise, Fig. 5C displays the interaction between the O\(_2\) molecule and the ar/R residues as a function of O\(_2\) position, and Fig. 5D shows the average number of water–protein hydrogen bonds. In Fig. 5E, MD snapshots of the ar/R regions of hAQP1 and GlpF are shown in ball-and-stick representation. The residues are colored according to the curves in the graph, and possible water–protein hydrogen bonds are indicated by dashed lines.

The key finding is that, upon O\(_2\) passage through hAQP1, favorable interactions between water and the ar/R residues (magenta curve) are reduced by \( \approx 60 \) kJ/mol. The loss of water–protein interaction cannot be compensated by O\(_2\)–protein interaction because the apolar O\(_2\) interacts with the ar/R residues on the order of 10 kJ/mol. The reduction in water–protein interaction occurs at the barrier in the PMF. Decomposing the ar/R in the single residues (blue, orange, green, and red curves for Arg-195, Phe-56, Cys-189, and His-180, respectively) shows that the water–protein interaction is mainly affected by reduced interaction to Arg-195 (\( \approx 35 \) kJ/mol). In addition, interaction to Cys-189 and His-180 are reduced by \( \approx 15 \) kJ/mol each. In the wider and more hydrophobic ar/R region of GlpF, water–ar/R interactions are hardly affected by the position of a passing O\(_2\) molecule (Fig. 5B Right). Hence, small solutes pass the ar/R region of GlpF without reducing the water–protein interaction. The same trend can be observed in Fig. 5D, which presents the average number of water–protein hydrogen bonds as a function of O\(_2\) position. Upon O\(_2\) passage through hAQP1, more than one hydrogen bond between water and both His-180 and Cys-189 is significantly reduced (red and green lines). In contrast, water–Arg-206 hydrogen bonds in GlpF are hardly affected by a permeating O\(_2\) molecule.

**Discussion**

From the extensive set of MD simulations presented here, a detailed understanding of the selectivity mechanism of aquaporins can be derived. We showed that, in both aquaporins and aquaglyceroporins, the ar/R region is the selectivity site for uncharged solutes. In hAQP1, the narrow and hydrophilic ar/R site forms a filter permeated by small polar solutes. Upon passage of a solute, favorable interactions between water and protein (mainly to Arg-195 but also to Cys-189 and His-180) are reduced and need to be replaced by solute–protein interactions. The more polar the solute, the stronger it can interact with hydrophilic ar/R residues, rendering a lower energetic cost to replace a water molecule. Larger solutes such as urea or glycerol are sterically excluded in hAQP1.

The hAQP1-HA/RV double mutant was found to be highly permeable to water, with an even lower barrier against water permeation compared with the hAQP1 wild type (SI Fig. 6). Hence, Arg-195 and His-180 are not necessary to isolate single water molecules from the bulk, which is in agreement with experiments that measured high water flux across hAQP1-HA/RV (23). However, we showed that these residues are important for the selectivity of uncharged solutes because mutations in the ar/R region have drastic effects on the channel characteristics. The mutations result in a larger and more hydrophobic pore, thus eliminating the selectivity mechanism for small polar solutes, which is present in the hAQP1 wild type. Without the selectivity mechanism and because of the larger pore size, the hAQP1-HA/RV mutant is highly permeable to apolar solutes, as well as to NH\(_3\) and urea.

In aquaglyceroporins, the ar/R region is wider than in AQP1 and has a more hydrophobic pocket opposite to the arginine. In GlpF, this environment allows passage of polyols, whose hydro-
doxyl groups can interact with Arg-206 replacing water–Arg-206 hydrogen bonds, whereas its apolar backbone matches into the hydrophobic pocket (6). Our results demonstrate that, because of the wider and more hydrophobic ar/R site, the selectivity mechanism present in AQP1 does not apply in aquaglyceroporins similar to GlpF. Therefore, GlpF-like aquaglyceroporins are efficiently permeated by small solutes such as NH3, CO2, or O2.

We stress that, from the permeabilities of the wide range of solutes that were probed in the present study, the permeability of AQP1 and GlpF for other small solutes may be estimated. The results suggest that small hydrophobic solutes such as NO, CO, or N2O are unlikely to permeate through AQP1 with barriers >20 kJ/mol. In contrast, aquaglyceroporins similar to GlpF are expected to be permeable to such solutes.

To address the possible flux of molecules such as CO2 or NO through AQP1, we suggest permeation experiments on GlpF or the hAQP1-HA/RV mutant. If permeation occurs along the AQP1 water channels and not through the lipid bilayer, the different ar/R region should drastically increase the permeability. For GlpF, for example, we would expect a 30-fold increase of the CO2 flux, compared with hAQP1.

Conclusions
A simple filter mechanism governs the permeability of AQP1. For small solutes permeating through AQP1, a hydrophobic effect leads to a remarkable anticorrelation between permeability and hydrophobicity. Large solutes are sterically excluded. Hence, the ar/R region of AQP1 is a filter against both hydrophobicity and size. This filter mechanism does not apply in GlpF, which efficiently permeates all investigated solutes except for urea. Surprisingly, not solute–pore, but water–pore interactions complemented by sterical effects emerged as the key determinants underlying the selectivity mechanism and the barrier height in aquaporins and aquaglyceroporins.

Methods
Simulation Setup. The simulation boxes of hAQP1 or GlpF contained the protein tetramer, 271 POPE lipids, and 19,769 (18,443 for GlpF) TIMP (30) water molecules. The GlpF structure was taken from the Protein Data Bank (PDB ID code 1FX8) (6). The starting structure of hAQP1 was modeled based on the x-ray structure of bovine AQP1 (PDB ID code 1J4N) (5) by mutating differing residues by using the WHAT IF modeling software (31). Crystal water molecules were kept in the structures, and chloride ions were added to neutralize the simulation systems. The OPLS all-atom force field (32, 33) was used for the protein, and lipid parameters were taken from Berger and Rossky (34). The simulations were equilibrated for 2 ns before production. A typical simulation box is shown in Fig. 1B. The simulations were carried out by using the GROMACS simulation software (35, 36). Electrostatic interactions were calculated at every step with the particle-mesh Ewald method (37, 38). Short-range repulsive and attractive dispersion interactions were described together by a Lennard–Jones potential, which was cut off at 1.0 nm. The Settle (39) algorithm was used to constrain bond lengths and angles of water molecules, and Linus (40) was used to constrain all other bond lengths, allowing a time step of 2 fs. The simulation temperature was kept constant by weakly (τ = 0.1 ps) coupling the protein, lipids, and solvent separately to a temperature bath (41) of 300 K. Likewise, the pressure was kept constant by weakly coupling the system to a pressure bath of 1 bar with a coupling constant τ of 1 ps. During umbrella sampling simulations, only the box length in the z direction (perpendicular to the membrane) was kept fixed.

The starting frames for the umbrella simulations were taken from 20-ns equilibrium simulations of hAQP1 and GlpF. The aquaporin channels were divided into 0.25-Å-wide equidistant sections parallel to the membrane, with the center of each section representing an umbrella center. Subsequently, the solute was placed into the channel at the umbrella center. Water molecules that overlapped with the solute were removed. To enhance sampling, two to four solutes were placed at different positions. A distance between the solutes along the pore of at least 25 Å for water, ammonia, CO2, and O2 and at least 30 Å for glycerol and urea was imposed to ensure that interactions between the solute molecules (if any) were negligibly small. During the subsequent EM of the structure, the protein RMSD was observed to ensure that the insertion of a larger solute into the pore did not lead to an unphysical widening of the pore.

Umbrella sampling calculations were carried out by applying a harmonic restraint force along the pore coordinate with force constants between 400 and 4,000 kJ mol–1 nm–2. In ammonia, the nitrogen atom was restrained; in water, the oxygen was restrained; in CO2, urea, and glycerol, the central carbon was restrained; and in O2, a dummy atom centered between the two oxygen atoms was restrained. Additionally, the solutes were restrained to a cylinder of radius r0 = 5 Å whose axis was centered along the pore by applying an additional harmonic force F(r) = −kz(r − r0)2 pointing toward the cylinder axis. Here r denotes the distance from the cylinder axis, kz = 400 kJ mol–1 nm–2 the force constant, and H is the Heaviside step function. The umbrella simulations of O2, CO2, NH3, and water were carried out for 400 ps, and the simulations of urea and glycerol were carried out for 600 ps and 1 ns, respectively.

Umbrella simulations for the lipid bilayers were performed by inserting the solute molecules at various positions into random snapshots taken from a 20-ns equilibrium run of a bilayer patch. The POPE and POPC patches contained 128 lipid molecules each and 4,777 or 5,788 TIP4P water molecules, respectively. Up to 12 solute molecules were inserted into one simulation, keeping a minimum distance of 25 Å to each other parallel and perpendicular to the bilayer. After EM, the simulations ran for ≈500 ps.

Ammonia parameters were taken from ref. 33, and urea parameters were from refs. 42 and 43. Lennard–Jones parameters for O2 were taken from the CHARMM22 force field (44). The O2 quadrupole was calculated from wave functions obtained at the MP2/6–311G* level to Qzz = −0.82 D and modeled by a positively charged dummy atom centered between two negatively charged oxygen atoms. CO2 parameters were taken from ref. 22, and glycerol was modeled from OPLS all-atom force fields. The parameters were validated by comparing the experimental Kmax to Kmax calculated by umbrella simulations (see SI Fig. 7).

Construction of PMFs. After removing the first 100 ps for equilibration, umbrella histograms were extracted from the z coordinate of the restrained atom. Subsequently, the umbrella positions were corrected with respect to the center of mass of the two NPA motifs of the corresponding channel. This procedure avoids a possible unphysical flattening of the PMF because of fluctuations of the monomers within the tetramer. It was ensured that all positions along the channel were well sampled. In case of poor sampling at maxima in the PMF, additional umbrella simulations with higher force constants were performed. In total, 34,432 histograms were collected from 1.6 μs of simulation of the aquaporin systems. The PMFs for permeation through the lipid bilayer were constructed from 15,232 histograms that were extracted from 600 ns of simulation.

The construction of the PMFs is described in more detail in SI Methods and is illustrated in SI Fig. 8. PMFs were calculated separately for each monomeric channel by using the weighted histogram analysis method (45). The single channels were considered as independent pathways. Therefore, the effective PMF Gm(z) was calculated by

\[ G_m(z) = z 
\]

Because of the cylinic restraint, the umbrella simulations yield a PMF that oxygen density of the channel density of the cylinder because of fluctuations of the monomers within the tetramer. It was ensured that all positions along the channel were well sampled. In case of poor sampling at maxima in the PMF, additional umbrella simulations with higher force constants were performed. In total, 34,432 histograms were collected from 1.6 μs of simulation of the aquaporin systems. The PMFs for permeation through the lipid bilayer were constructed from 15,232 histograms that were extracted from 600 ns of simulation.

The construction of the PMFs is described in more detail in SI Methods and is illustrated in SI Fig. 8. PMFs were calculated separately for each monomeric channel by using the weighted histogram analysis method (45). The single channels were considered as independent pathways. Therefore, the effective PMF Gm(z) was calculated by

\[ G_m(z) = z \]

Because of the cylinic restraint, the umbrella simulations yield a PMF that oxygen density of the channel density of the cylinder because of fluctuations of the monomers within the tetramer. It was ensured that all positions along the channel were well sampled. In case of poor sampling at maxima in the PMF, additional umbrella simulations with higher force constants were performed. In total, 34,432 histograms were collected from 1.6 μs of simulation of the aquaporin systems. The PMFs for permeation through the lipid bilayer were constructed from 15,232 histograms that were extracted from 600 ns of simulation.

The construction of the PMFs is described in more detail in SI Methods and is illustrated in SI Fig. 8. PMFs were calculated separately for each monomeric channel by using the weighted histogram analysis method (45). The single channels were considered as independent pathways. Therefore, the effective PMF Gm(z) was calculated by

\[ G_m(z) = z \]

Because of the cylinic restraint, the umbrella simulations yield a PMF that oxygen density of the channel density of the cylinder because of fluctuations of the monomers within the tetramer. It was ensured that all positions along the channel were well sampled. In case of poor sampling at maxima in the PMF, additional umbrella simulations with higher force constants were performed. In total, 34,432 histograms were collected from 1.6 μs of simulation of the aquaporin systems. The PMFs for permeation through the lipid bilayer were constructed from 15,232 histograms that were extracted from 600 ns of simulation.

The construction of the PMFs is described in more detail in SI Methods and is illustrated in SI Fig. 8. PMFs were calculated separately for each monomeric channel by using the weighted histogram analysis method (45). The single channels were considered as independent pathways. Therefore, the effective PMF Gm(z) was calculated by

\[ G_m(z) = z \]

Because of the cylinic restraint, the umbrella simulations yield a PMF that oxygen density of the channel density of the cylinder because of fluctuations of the monomers within the tetramer. It was ensured that all positions along the channel were well sampled. In case of poor sampling at maxima in the PMF, additional umbrella simulations with higher force constants were performed. In total, 34,432 histograms were collected from 1.6 μs of simulation of the aquaporin systems. The PMFs for permeation through the lipid bilayer were constructed from 15,232 histograms that were extracted from 600 ns of simulation.
Combining the histograms of the four channels into an effective PMF turned out to be nontrivial. Side chains along the pore can adopt different conformations corresponding to a partial or even complete closure of the channel. In particular, the conserved Arg-195 in hAQP1 can adopt several conformations, which has also been observed in the AQP-Z crystal structure (1). Such conformations are present in the starting structures of the umbrella runs. Opening and closure transitions are, however, not well sampled within an umbrella window of several hundred picoseconds. In such cases the weighted histogram analysis method (WHAM) (2) tends to overestimate the barrier when histograms from different channels are combined into a single WHAM procedure. To alleviate this problem, the PMF was calculated for each monomer separately using the WHAM procedure. The four channels were considered as independent pathways. Therefore, the effective PMF $G_{\text{eff}}(z)$ was calculated via $\exp(-G_{\text{eff}}(z)/k_BT) = 4^{-1}\sum_{j=1,...,4}\exp(-G_j(z)/k_BT)$. Here, $T$ denotes the temperature and $k_B$ is the Boltzmann constant.

Before combining the four profiles into $G_{\text{eff}}(z)$, the single-channel PMFs were corrected for a possible free-energy offset between the two bulk water regions. A possibility to avoid any offset is to impose the periodicity of the system into the WHAM procedure. However, we observed that a periodic WHAM procedure imposes a rather linear shift into the PMF derived by a nonperiodic WHAM procedure (data not shown). Such a linear shift is unphysical in the bulk water where the PMF should be flat. Therefore, a different procedure was applied. The offset $\Delta G_{\text{off}}$ was removed by introducing corrections $\delta G_{i,i+1}$ between adjacent bins $i$ and $i + 1$, such that $\Delta G_{\text{off}} = -\sum_i \delta G_{i,i+1}$. The $\delta G_{i,i+1}$ were chosen to be proportional to $1/(n_in_{i+1})^\alpha$, where $n_i$ and $n_{i+1}$ are the number of data points in the adjacent bins $i$ and $i + 1$ as collected from all histograms of the channel. Using this procedure with a small value for $\alpha$ (such as $\alpha = 1/2$) yields a rather linear correction between the two bulk water regions, similar to the periodic WHAM. A large value for $\alpha$ induces steps into the PMF at the sharp peak(s) of the PMF, where a completely converged sampling is difficult to achieve. Hence, a sufficiently large value for $\alpha$ yields a PMF that corresponds to an integration of the mean force from the two bulk water regions to the point(s) of lower sampling inside the
pore. A reasonable value was found to be $\alpha = 2$, although the exact choice of $\alpha$ was found to have only minor influence on the main barrier height (<2 kJ/mol).

A one-dimensional PMF for solute permeation across a membrane channel must refer to a given density of channels inside the membrane, or, in turn, to a given membrane cross-section per channel. For example, a higher density of channels (smaller membrane cross section per channel) increases the probability for the solute to “find” the channel and therefore leads to a lower free energy barrier.

Because the solutes were constricted to a cylinder along the pore, the umbrella simulations yield the free-energy profile referring to a channel density of one channel per cross section of the cylinder ($A_c$). However, to render the profile comparable with the profile for permeation through the lipid bilayer (which is independent of the cross-section), we chose to relate the profile to a channel density of one channel per membrane cross-section occupied by the aquaporin monomer ($A_{\text{mono}}$). To this end, the barrier height must to be corrected by $\Delta G = k_B T \ln(A_{\text{mono}}/A_c)$ (3). We estimated $A_{\text{mono}}$ to 10.3 nm$^2$ and 10.7 nm$^2$ for hAQP1 and GlpF, respectively. From the cylinder radius $r_c$ and the force constant $k_c$, $A_c$ can be estimated to $A_c = 1.05$ nm$^2$, yielding a corrections 5.7 and 5.8 kJ/mol for hAQP1 and GlpF, respectively. These corrections were incorporated in the PMFs as a linear trapezoidal correction in the pore entrance and exit region. The construction of a PMF is illustrated in SI Fig. 8.