Glycerol facilitator GlpF and the associated aquaporin family of channels

Robert M Stroud*, Larry JW Miercke, Joseph O’Connell, Shahram Khademi, John K Lee, Jonathan Remis, William Harries, Yaneth Robles and David Akhavan

The aqua (glycero) porins conduct water (and glycerol) across cell membranes. The structure of these channels reveals a tripathic channel that supports a hydrophobic surface and, opposite to this, a line of eight hydrogen-bond acceptors and four hydrogen-bond donors. The eight carbonyls act as acceptors for water (or glycerol OH) molecules. The central water molecule in the channel is oriented to polarize hydrogen atoms outward from the center. This arrangement suggests how the structure prevents the potentially lethal conduction of protons across the membrane. The structure also suggests the mechanism behind the selectivity of aquaglyceroporins for glycerol, the basis for enantioselectivity among alditols, and the basis for the prevention of any leakage of the electrochemical gradient.

Addresses
Department of Biochemistry & Biophysics, School of Medicine, 600 16th Street, University of California San Francisco, CA 94143-2240, USA
*e-mail: stroud@msg.ucsf.edu


This review comes from a themed issue on Membranes
Edited by Eric Gouaux and Stephen H White

0959-440X/$ – see front matter
© 2003 Elsevier Ltd. All rights reserved.
DOI 10.1016/S0959-440X(03)00114-3

Abbreviation
AQP aquaporin

Introduction
Every cell is surrounded by an insulating membrane. It is essential that cells insulate themselves from the high electrochemical gradient of the 100,000 volt/cm electric field across the membrane. Every cell also has to capture metabolic nutrients. How cells bring in the essential nutrients and how the cell membrane absolutely excludes leakage of even protons out of the cell is a fundamental question in membrane biology. The structures of channels in the aquaporin (AQP) family open the way to address the larger issue of how strict membrane selectivity is determined.

One of the earliest recognized transmembrane channels
One hundred years ago, Alfred Fischer deduced that glycerol channels existed. When certain species of pathogenic bacteria were placed in hyperosmotic solutions, about half of them failed to undergo lysis in glycerol solutions [1]. Fischer correctly concluded that this subgroup of pathogens had to have membranes that were highly permeable to glycerol to relieve the hyperosmotic stress.

Beginning in the mid 1960s, a brilliant series of genetic analyses pioneered by ECC Lin defined the nature of and the conducting specificity of GlpF [2]. GlpF is a highly selective transmembrane channel that conducts glycerol, water and small uncharged organic molecules. Once inside the bacterial cell, glycerol is rapidly phosphorylated by glycerol kinase to produce glycerol 3-phosphate. GlpF conducts urea, glycine and DL-glyceraldehyde in addition to glycerol. It is also stereo and enantio selective in the conductance of linear carbohydrates (called alditols) [2]. Aldoses — cyclized alditols — are not conducted through the GlpF channel. The structure of the GlpF channel shows that it is indeed too small to conduct cyclic molecules, beautifully demonstrating its stereo and enantio selectivity [3,4].

GlpF is now known to be a member of the AQP channel family. The first member of this family to be identified was the human water channel AQP1 from red blood cells and renal tubules [5]. Peter Agre and colleagues first identified the cDNA of this water-conducting channel [6], and named the family ‘aquaporins’ or AQPs [7]. AQPs comprise functionally distinct subgroups that include transmembrane water-conducting channels, which conduct water but do not conduct linear alcohols or glycerol (aquaporins), and channels that conduct water and also conduct glycerol (aquaglyceroporins). These latter channels also conduct urea, DL-glyceraldehyde, alditols and other small organic molecules [8–10].

The AQP family arose by tandem intragenic duplication [11] to give proteins that have an internal repeat; the N-terminal segment displays ~20% conservation with the C-terminal segment [12] (Figures 1 and 2). This duplication event appeared early in evolution, as bacteria such as Escherichia coli contain both a water- and glycerol-conducting channel (GlpF), and a separate water channel (AQPZ). Each segment contains a conserved -Asn-Pro-Ala- (NPA) signature sequence near its center and several other conserved residues, including Glu14 and Glu152 near the beginning of each segment.
Eleven human AQPs, numbered AQP0 to AQP10, play key roles in human physiology. In humans, AQP3, AQP7, AQP9 and AQP10 are aquaglyceroporins [13,14]. The atomic structures of GlpF and AQP1 offer clues to the mechanisms of AQP selectivity [3,15,16]. Both intuition and molecular mechanics can clearly be flawed, making it crucial to understand the limitations of each. But, in principle, molecular mechanics can be used to identify and subsequently evaluate the contributions of each factor to selectivity by ‘turning them off’ in a simulation [17]. The factors refer to such aspects as channel amino acid identity and charge nature, amino acid proximity to the channel, the hydrophilic/hydrophobic nature of the channel residues, and the spatial orientation of specific residues that provide hydrogen acceptors/donors for the water and glycerol molecules that traverse the channel. For these membrane channels, such simulations have been particularly instructive and serve to direct the next round of experiment.

AQPs are all based on arrangements of four channels as tetramers that seem to stick together well, even in detergent-based purifications, as they do in their host membrane [18] (Figure 3). Within each monomer, a right-hand twisted arrangement of six α helices and two half-spanning α helices (M1–M8) surrounds the channel [14]. This was first visualized by electron crystallography [19–23] at increasing resolutions down to 4.5 Å in-plane (4.5 × ~9 Å in three dimensions) [14,24] and later by X-ray crystallography at atomic resolution [3,4,16].

Glycerol conductance and the three-dimensional structure of GlpF

In GlpF, an amphipathic transmembrane pathway shows three specific intermediate glycerol-binding sites (G1–G3). Glycerol molecules therefore move down the channel from the extracellular side in single file, with their alkyl backbone sequestered along a two-sided hydrophobic corner (Figure 4). This orientation is demonstrated by glycerol bound in the ‘selectivity filter’ close to the narrowest portion of the channel and in contact with conserved Arg206 (Figure 5). Diametrically opposed to the hydrophobic corner are two NHs from the amidine of Arg206 that serve as hydrogen-bond donors to two successive glycerol OH groups (OH1 and OH2) and two atoms from Gly199 and Phe200 that serve as hydrogen-bond acceptors to OH1 and OH2. This arrangement forms a ‘tripathic’ channel, as it presents the channel lumen with three different characteristics: a hydrophobic corner; hydrogen-bond donors; and hydrogen-bond acceptors. Together, these three elements orient three counterparts on the permeant molecule.

Channel selectivity

The channel is an ~28 Å long selective tripathic channel (r<3.5 Å). The pathway is defined by several water
Figure 2

The amino acid sequence of GlpF shows how the twofold relationship derived from the early gene duplication of sequence is reflected in the organization of structure. Black indicates residues that interact with the glycerol molecules observed within the channel. Red indicates the residues that contribute carbonyls to the central channel. Purple indicates sidechains that contribute hydrocarbon to the channel. The extended chains that precede M3 and M7 contribute the essential line of eight carbonyls within the 28 Å long channel that serve to accept hydrogen bonds from the permeant molecule.

Figure 3

The tetramer of channels is common to all AQP4s. This arrangement optimizes the packing between molecules, rather than within one molecule. Two aromatic rings (Trp48 and Phe200) form a two-sided corner in the channel. With two NH donors from the guanidinium sidechain of Arg206 on another side and two mainchain carbonyl oxygens on the fourth side, the selectivity filter at G2 is highly tripathic. G2 binds in the Trp-Phe-Arg triad, leaving no free space around it, such that van der Waals, hydrogen bond and electrostatic forces each play a role. In G2, OH1 and OH2 act
as both hydrogen-bond acceptors from NHs of the guanidinium group of Arg206 (2.9 Å and 2.7 Å) and hydrogen-bond donors to the carbonyl oxygens of Gly199 (2.6 Å) and Phe200 (2.8 Å), respectively. There is no space for substitution at the CH hydrogen positions because the G2 alkyl backbone is packed flush against the aromatic corner.

As is likely in all AQP family members, the two mainchain amide nitrogens of residues 200–201 are held in place by hydrogen bonds to the entirely buried carboxyl of Glu152 (Figure 6). Likewise, amide nitrogens of Leu67 and His66 are pinned by a second highly conserved and equally buried carboxyl of Glu14, which is quasi-twofold related to Glu152. These serve as powerful restraining forces. Four surrounding carbonyls in the entry vestibule, of Gly199, Phe200, Ala201 and Met202, and the gene-duplication related carbonyls of Gly64, His65, Ala66 and Leu67 are oriented by these carboxyls and provide the line of hydrogen-bond acceptors, each separated by ∼3.0 Å apart conducting hydrogen-bond donor OH groups (Figure 7).

G3 lies opposite invariant Asn203 and Asn68, between helices M3 and M7. With ideal geometry, NH$_2$ of Asn203 is donor to OH1 of G3 (3.0 Å) and NH$_2$ of Asn68 is donor to OH2 of G3 (3.1 Å). Thus, glycerol molecules can pass through the G2–G3 region only in single file, with successive CHOH groups within one molecule following each other in single file.

**Stereoselectivity among linear alditols**

The influx of carbohydrates into proteoliposomes reconstituted using purified GlpF has been measured. Ribitol, a five-carbon alditol, generated exponential reswelling with a time constant of ∼0.35 s$^{-1}$. By contrast, D-arabitol, a chiral stereoisomer of ribitol with a mixed arrangement of hydroxyls, shows an approximately tenfold reduction in transport relative to ribitol. This demonstrates the stereoselectivity of the channel. The channel is stereoselective for different chain lengths and different chiralities.

**Simulated rates through the channel**

Twenty hydrogen-bond acceptors and twenty-three hydrogen-bond donors participate in the conductance of glycerol [4**.25**]. Molecular mechanics assesses agreement with the experimental values of measured properties as a first validation. In the context of GlpF and now AQP1, such calculations seem to be quite robust. Very little change is induced in the structure of the channel as seen by crystallography when glycerol versus water is presented to the channel, and little change occurs during the passage of either during the simulation [17**].

**Simulation and rates of water passage**

The mechanism of water conduction through AQP channels is suggested by the crystal structure of GlpF with water alone in the channel. Visualized at 2.7 Å resolution, the observed water positions can be compared with those in a molecular simulation. Seven to nine water molecules form a single file in the 20 Å constriction region of the channel. The calculated diffusion constant corresponds to a flux of 2.4 × 10$^9$ s$^{-1}$, which is just five times larger.
than the experimental flux for GlpF \([26,27,28]\) of \(0.5 \times 10^8 \text{s}^{-1}\). The comparison between simulation and crystal structure is remarkably good, indicating that seven to nine water molecules form a stable line through the entire channel.

**Insulation from proton conductance?**

How do channels that support a line of water molecules exclude the conductance of protons, which would be lethal to the cell? The Grotthuss mechanism involves the protonation of one water from a donor at one side of the membrane and the concerted transfer of another proton from the water molecule to its neighbor through a line of waters. This mechanism could in theory leak protons through the membrane without generating any charge within the 28 Å narrow region of the channel.

The crystal structure of GlpF in the absence of glycerol \([17]\) revealed water molecules lined up through the channel. Each water molecule is a hydrogen-bond donor to one of the eight carbonyls that line the channel and to a neighboring water. However, the central water molecule is oriented by the two central asparagine sidechain NH donors to polarize the entire line of waters facing outward from the central position, as postulated by Murata *et al.* from their 3.5 Å electron microscopic structure of AQP1 \([29]\). The helix dipoles of M3 and M7 place their positively charged ends together near the NPA regions and thus ‘bipolarize’ a line of water \([29]\). The full charge present on Arg206 also plays a role in polarizing the central water molecule, as proposed \([30]\) and subsequently validated by molecular simulations. This arrangement bipolarizes the entire line of waters, with each water
acting as a hydrogen-bond acceptor from its inner neighbor on its inside relative to the center of the bilayer at water 522.

The effectiveness of each of these three factors in polarizing the central water could be measured by selectively turning individual contributors off in the calculation [17**]. Eliminating partial charges on the NH$_2$ groups of the sidechains of Asn68 and Asn203 shows that their ability to determine the water configuration is only partial. When the partial charges on the amide backbone atoms of half-helices M3 and M7 are turned off, eliminating the helix dipoles, again only a partial effect on water orientation is seen. When this phenomenon is combined with the elimination of the partial charges on the NH$_2$ groups of the Asn68 and Asn203 sidechains, the bipolar orientation of the single file arrangement breaks down, potentially leading to a proton wire. Channel-lining carbonyl group dipoles maintain the bipolar orientation of water molecules. Thus, each of these factors contributes to the abrogation of any proton conductance through the channel.

Protons can speed through a line of water molecules, hopping cooperatively down a line of hydrogen-bonded water molecules in what is known as the Grotthuss mechanism [31]. This effect is observed in proton tunneling through ice. If protons were to move cooperatively, no charge even needs to exist in the center of the channel. We propose that the polarized center of the line of water is key to insulation [17**,30**], as implied by crystallography [3*,4*,16**,32]. Presumably fully charged Arg206 might also become a major repulsive element to any positive charge within the transbilayer luminal region of the channel.

Against passing ions
GlpF and AQPs absolutely reject ions and charged solutes. This is because the GlpF channel cannot accommodate hydrated ions due to the large size of the hydration shell. The cost of dehydrating an ion is enormous and must be compensated by the carbonyls or NHs of amides within the walls of the channel. Anions and cations are excluded because they require groups that are polarized with both δ$^+$ donor and δ$^-$ acceptor moieties, respectively. The two charged sidechains of Arg206 and Glu152 might impose electrostatic energy barriers to the passage of a charged ion or solute molecule, including OH$^-$ or H$_2$O$^+$. Such electrostatic repulsion would play the same role in all AQPs, as Arg206 and Glu152 are conserved across almost all AQPs.

Selectivity: glycerol versus water
The conductivity of water-preferring AQPs such as AQPZ and AQP1 for glycerol, urea or other uncharged small molecules is essentially undetectable. GlpF conducts water at a relative rate that is one-sixth that for the water channel AQPZ [28**].

Mutagenesis of two residues, Tyr236Pro and Trp237Leu, in AQP1 allowed it to pass glycerol [33]. In GlpF, Pro236 lies in the protein interior in contact with the Ala201 sidechain behind G2. Leu237 contacts Pro204 of the second NPA sequence at the G3 site. Mutational changes at these sites can be relayed to the filter to favor passage of glycerol.

Substitution of the conserved Arg206 in the water channel AQPZ (Arg189Val/Ser) inactivated water transport [27]. A mutation of the conserved Arg206 in the related AQP2 (Arg187) is known to cause nephrogenic diabetes insipidus in humans. It is believed that the disease is caused by impaired intracellular membrane traffic [34]. However, the location of Arg206 in the selectivity filter must alter the water transport of AQP2 in collecting duct principal cells in the kidney.
Conclusions
The structures of GlpF and AQP5 show how they are specific to the conduction of linear alditols and water, respectively. Molecular mechanics calculations show how they also prevent protons and any ions from leakage through the AQP5 channels. Future directions will seek to design and express mutations in GlpF that alter the conductance properties. The conductances of other AQP5s and their selectivities will be measured and compared to measured properties of mutants that seek to address the basis of conductance and regulation. Another major direction will be the determination of the crystal structures and properties of human AQP5s, many of which are associated with human health and disease.

References and recommended reading
Papers of particular interest, published within the annual period of review, have been highlighted as:
• of special interest
**of outstanding interest
29. Determination of the rates of conductance of glycerol and water through GlpF and AQPZ is reported. Although both AQPZ and GlpF have similar primary amino acid sequences, including the NPA motifs, differences in the quaternary structure explain differences in their selectivity for water.

The selectivity filter of GlpF is similar to that of AQP1 in that the constriction of both pores prevents the passage of ions, protons and charged solutes, but permits water passage. NPA regions in both are important in forming the surface of the aqueous pore in this regard. The mechanisms for passing water through the channel in both GlpF and AQP1 are also similar in that breakage of hydrogen bonds is involved (NPA regions are critical in this process). In GlpF, water molecules form a single-file line, becoming hydrogen-bond donors to carbonyls, with the exception of the central water molecule, which becomes a hydrogen-bond acceptor from NPA region asparagines. Another important effect of this is that interruption of these hydrogen bonds prevents proton conductance.


Water permeation in GlpF and AQP1 occurs through a two-stage filter: the NPA region conserved selectivity-determining region and the aromatic/arginine filter, which is proposed to act as a proton filter. Also, hydrophobic regions near the NPA region are rate-limiting water barriers. The fine-tuned water dipole rotation in AQP1 during water passage is similar to how helix dipoles orient water molecules at the center of GlpF. Although GlpF contains a wider pore than AQP1, its water permeation is lower than AQP1.


