Structure of a Glycerol-Conducting Channel and the Basis for Its Selectivity

Daxiong Fu, Andrew Libson, Larry J. W. Miercke, Cindy Weitzman, Peter Nollert, Jolanta Krucinski, Robert M. Stroud*

Membrane channel proteins of the aquaporin family are highly selective for permeation of specific small molecules, with absolute exclusion of ions and charged solutes and without dissipation of the electrochemical potential across the cell membrane. We report the crystal structure of the *Escherichia coli* glycerol facilitator (GlpF) with its primary permeant substrate glycerol at 2.2 Ångstrom resolution. Glycerol molecules line up in an amphipathic channel in single file. In the narrow selectivity filter of the channel the glycerol alkyl backbone is wedged against a hydrophobic corner, and successive hydroxyl groups form hydrogen bonds with a pair of acceptor, and donor atoms. Two conserved aspartic acid–proline–alanine motifs form a key interface between two gene-duplicated segments that each encode three-and-one-half membrane-spanning helices around the channel. This structure elucidates the mechanism of selective permeability for linear carbohydrates and suggests how ions and water are excluded.

Members of the aquaporin family (AQPs) are composed of functionally distinct subgroups that include transmembrane water-conducting channels (aquaporins) and glycerol-conducting channels (aquaglyceroporins) that include the *E. coli* glycerol facilitator (GlpF) (1–3). The electrochemical potential across the cell membrane is preserved by sharp selectivity that excludes all ions including hydrogen, and hydroxide ions (4). GlpF also conducts linear polyalcohols (called alditols), for which it is stereo- and enantio-selective (5).

Found in species from bacteria to yeast, plants, and humans, the sequences of more than 150 AQPs are known. This ancient family arose by tandem intragenic duplication (6), resulting in proteins that have an internal repeat where the NH2-terminal segment displays ~20% conservation with the COOH-terminal segment (7). Each segment contains a conserved Asn-Pro-Ala signature sequence (NPA) near its center. Human family members include at least 10 aquaporins (AQP0 to AQP9), all of which are permeable to water. AQP3, AQP7, and AQP9 are also permeable to glycerol. In eukaryotes many of the family members are regulated by phosphorylation, pH, osmolarity, or binding of other proteins or ligands (8, 9).

The first AQP to be characterized was a water-conducting channel, AQP1 (1). Monomers of AQP1 assemble in membranes as tetramers of four independent channels (10). In reconstituted membranes, the tetramers come together alternately cytoplasmic and extracellular side up. Electron crystallography of AQPs (11–15) at increasing resolutions down to 4.5 Å in-plane, 4.5 by ~9 Å in three dimensions (16) reveals a right-hand twisted arrangement of six transmembrane and two half-membrane–spanning helices around a central channel (17).

The rate of GlpF-mediated glycerol influx is 100- to 1000-fold greater than expected for a transporter and is nonsaturable to a glycerol concentration of >200 mM (5). GlpF acts as a highly selective channel, as indicated by lack of counterflow of radioactive substrate against a concentration gradient (18).

**GlpF structure.** The structure of GlpF was determined by isomorphous replacement and anomalous dispersion (Table 1). GlpF crystallizes as a symmetric arrangement of four channels with three glycerol molecules (G1 to G3) in each (Fig. 1, A and B). The plane of the bilayer would be perpendicular to the fourfold axis. Each tetramer is circled by a hydrophobic surface matching the dimensions of the lipid bilayer. The periplasmic leaflet is indicated by the interactions of 12 octylglucoside molecules. Layers that include tryptophan side chains can productively interact with the lipid head-group regions and are capped by charged residues that result in net positive charge on the cytoplasmic side consistent with the trend in other membrane proteins.

Six transmembrane and two half-membrane–spanning α helices (M1 to M8) form a right-handed helical bundle around each channel. The helices M1 to M8 diverge outward from the central plane to generate a vestibule on each surface. The interhelix angles of ~+35° to ~+40° are not in the categories most favored by "knobs into holes" packing (19). However, helices pack with near optimal packing angles of ~—20°, both between M1 and M2 of one monomer and M5 and M6 of its neighbor, and between (four copies of) M2 and M6 of all monomers around the fourfold axis.

The NH2-terminus of GlpF was found to be cytoplasmic by thrombin accessibility, as in AQP1 (20). The two genetic repeats form segments related by a quasi twofold axis that would pass through the center of the bilayer and intersect the fourfold axis. The NH2-terminal segment begins with two membrane-spanning helical segments, M1 and M2, followed by residues 61 to 67 that project into the center of the channel as an extended polypeptide chain and provide three successive carboxyls to the cytoplasmic vestibule. The half-spanning helix M3 begins with the first NPA motif and returns to the cytoplasmic side. M4 is packed between M3 and M1 (Fig. 1, B and D).

The COOH-terminal segment reiterates a similar transmembrane topology beginning from the periplasmic side that likewise provides three carboxyls to the periplasmic vestibule from the extended chain 195–202. Thus, the quasi twofold relation applies to elements of the amphipathic walls of the channel. The acceptor carboxyls in the vestibules, and the hydrogen bond donors of Arg206, Asn63, and Asn203 at the interface of the two half-helices M3 and M7 form a left-handed helical polar stripe that connects the periplasmic and cytoplasmic vestibules. The region that links the two segments forms two short α helices (109–120 and 126–134) that are protrusions on the periplasmic side at the COOH-terminal end of M4. An extended chain (137–143) returns to the walls of the channel and contributes three more carboxyls to the periplasmic vestibule.

A unique feature of this membrane protein structure is that two helices meet at their NH2-terminal ends in the center of the membrane. Their conserved NPA motifs create an intimate interface across the quasi twofold axis in which the proline rings are in Van der Waals contact cupped between the proline and alanine side chains of the opposite helix (Fig. 1C). Each asparagine side chain is constrained by two hydrogen bonds that precisely orient side-chain NHs toward acceptors on the permeant substrate. One side of each half-spanning helix M3, M7 contacts the lipid-accessible exterior perimeter of the tetramer. The helical axes of M3, M7 intersect at an angle of ~—146°, thereby narrowing the cross section of the channel in the center of the bilayer region.

Two transmembrane helices, M2 and M6,
positioned nearest the fourfold axis of the tetramer are not long enough to span the bilayer, consistent with the tetramer being the stable physiological quaternary structure. The interfaces between subunits are almost as hydrophobic as the exterior, suggesting that a monomer could be transiently stable in the membrane upon synthesis, before forming tetramers.

The high conservation throughout the aquaporin family indicates functional and structural roles in the context of a common fold (Table 2). Of eight residues that are conserved in the two repeats (group III), six lie close to the central plane of the bilayer related by the quasi twofold axis of symmetry to corresponding residues in the second repeat. They contact each other in pairs and maintain packing around the narrow region of the channel and around the quasi twofold axis in GlpF. Four of these are glycines within helices that allow the closest approaches of helices in the center of the bundle (21).

The channel. The channel has a ~15 Å wide vestibule on the periplasmic surface and

Fig. 1. Three-dimensional fold of GlpF with the quasi twofold related segments colored in yellow (residues 6 to 108) and blue (residues 144 to 254). Glycerols G1, G2, and G3 lined up within the channel pathway are colored magenta. (A) Ribbon representation of the GlpF tetramer viewed from the periplasmic side. The dotted line represents the clipping plane used in Fig. 2B. (B) A stereo view (crossed eyed) of the monomeric glycerol channel viewed down the quasi twofold axis. The vertical bar (35 Å) represents the position of the membrane. Helices are labeled M1 to M8. (C) The two NPA motifs at the segment interface where the two half-helices M3 (yellow) and M7 (blue) meet form a clasplike junction. Each proline ring lies in-between the proline and alanine side chains of the opposite helix. The asparagine amide groups of one helix form two symmetry-related hydrogen bonds—one with the alanine nitrogen of that same half-helix and one with a carbonyl oxygen of the opposite half-helix. This structure presents a highly constrained pair of donor hydrogen bonds to successive OHs at G3. In this and succeeding figures, atoms are colored according to atom type (oxygen, red; carbon, gray; nitrogen, blue; and sulfur, yellow). (D) The amino acid sequence of GlpF is arranged topologically as in the structure, with helices viewed as if from inside the channel, viewed looking away from the fourfold axis. The NPA motifs are in the center of the figure. Three-and-one-half helices compose each segment, labeled M1 to M4 and M5 to M8. Related helices M1 and M5, M2 and M6, and so forth are boxed in similar colors to represent their homologous roles in the structure. Residues in black circles interact with glycerol. Residues in red circles contribute carbonyl oxygen or amide NHs to the channel. Residues in purple circles contribute hydrocarbon to the channel. Gray circles represent residues that are not seen in the structure. The deduced location of the cell membrane is illustrated in gray. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
reaches its constriction of \( \sim 3.8 \text{ Å} \) by 3.4 Å, 8 Å above the quasi twofold axis. This constriction lies at the start of a \( \sim 28 \text{ Å} \) long selective channel (radius \( <3.5 \text{ Å} \)) that extends to the cytoplasmic surface. Fig. 2). G1 is hydrogen bonded to the O of Tyr \( 138 \) in the entry vestibule. G2 and G3 are both tightly organized in a site we term the “selectivity filter,” a region only large enough in cross section to accommodate a single CH-OH group (Fig. 3). Therefore, CH-OH groups of alditol molecules can only pass through the region of G2 and G3 in single file. The water molecule between G2 and G3 suggests that, at least at high concentrations of glycerol, glycerol and water may be stoichiometrically cotransported.

The selectivity filter is strongly amphipathic, with the planes of two perpendicular aromatic rings (Trp \( 48 \) and Phe \( 200 \)) forming a hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A).

The selectivity filter is strongly amphipathic, with the planes of two perpendicular aromatic rings (Trp \( 48 \) and Phe \( 200 \)) forming a hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A). The alkyl backbone of G2 is tightly packed against this corner, leaving no space for any substitution at the C-H hydrophobic corner (Fig. 3A).

Table 1. Crystallographic statistics to 2.2 Å resolution. GlpF, which includes an NH\(_2\)-terminal six-histidine tag followed by a thrombin cleavage site, was cloned from \( E. coli \) strain K12 and purified by nickel affinity and size-exclusion chromatography. GlpF (15 to 20 mg/ml) was crystallized from a solution that contained 28% (w/v) polyethylene glycol 2000, 100 mM Bicine, 15% (v/v) glycerol, 35 mM n-octyl-\( \beta \)-D-glucoside, 300 mM MgCl\(_2\), and 5 mM dithiothreitol (DTT) (pH 8.9). Crystals were in space group I\( 422 \) (\( a = 96.4 \text{ Å}, c = 184.5 \text{ Å} \)). Data were collected at ALS beamline 5.0.2 with a charge-coupled device detector (Quantum) and integrated and scaled with MOSFLM and SCALA (37). Phases were calculated by multiple isomorphous replacement (MIR) and anomalous scattering with SHARP (38). After solvent flattening and phase extension to 3.0 Å, the initial map showed a single, continuous, main-chain density with strong side-chain features that allowed unambiguous sequence assignment for 80% of the residues in GlpF. The model was completed and refined with CNS (39). PCMBs, P-chloromercuribenzenesulfonate.

Table 2. Sequence-structure relationship in GlpF. The roles of the most conserved residues, based on sequence alignment of 52 representative aquaporin family members (3). Group I includes residues found only in the first genetic repeat; group II, residues in the second repeat only; and group III, residues conserved in both repeats I and II. Group I and II therefore may be related to structure and function specific to each respective repeat, whereas group III reflects the genetic relation between the two genetic repeats.
The pathway can be visualized as a series of activation barriers that separate intermediate binding sites through the channel, including the ones we observe at G2 and G3 (22). The narrowest constriction in the channel lies between G1 and G2. Thus, it constitutes a putative “transition state” in the pathway. At this narrowest position the hydrogen bond geometry from the donor NH₃⁺ of Arg²⁰⁶ to an OH group would become ideal. Thus, it acts as in an enzyme to lower the activation energy, in this case for transport through the most restricted region.

**Stereoselective preference for glycerol and linear carbohydrates.** To assess the selectivity of GlpF we adapted an osmotic method (23) to measure influx of carbohydrates into reconstituted proteoliposomes (Fig. 4). Upon addition of carbohydrates, proteoliposomes shrink owing to osmosis. The rate of carbohydrate influx is indicated by the rate of recovery of the initial volume of proteoliposomes. Ribitol, with all OH groups having the same stereospecific relation to the carbon backbone, generated an exponential reswelling with a time constant of 3.0 s (Fig. 4, inset). In contrast D-arabitol, a chiral stereoisomer of ribitol with a mixed arrangement of hydroxyls, shows a 10-fold reduction in transport relative to ribitol, demonstrating stereoselectivity of the channel. The structure shows how this can be visualized because two successive –CHOH groups are oriented at the G2 site. The carbon backbone must be lined up along the channel axis. Depending on the enantiomer, the C-H hydrogen(s) is in contact with the aromatic ring of either Phe200 or Trp48 in similar environments. However, any CHOH groups adjacent to these two will place the carbon in one of two tetrahedrally disposed sites that have quite different environments. The permeation of D-arabitol is slowed considerably, most likely owing to the difference in accommodation at these adjacent positions. There must therefore be some selectivity, even between the two prochiral (chirality with respect to direction in which the alditol enters the channel) carbon orientations.

Membranes have a high intrinsic permeability for glycerol. When GlpF is present, the reswelling rate increases such that the vesicles recover by a factor similar to that seen with ribitol. Measured as a function of the amount of GlpF used in the reconstitution, the reswelling rate increases linearly with the amount of GlpF.}

---

**Fig. 2.** The glycerol-conducting channel in GlpF. (A) The surface of the glycerol channel is colored according to atom type, showing the hydrophobic surface (left) and the polar surface (right). The length of the narrow region through the channel is indicated by the bar. (B) The GlpF tetramer sectioned diagonally as indicated by the dotted line in Fig. 1A in front of the plane containing the two glycerol channels indicated by orange arrows, and the fourfold axial space in the center of the figure (green arrow). The surfaces are color-coded by atom type. The left and right images of the channel represent views that are 180° apart. The left-handed helical stripe corresponding to the amphipathic pathway through the channel is indicated by orange lines. A 35 Å vertical bar indicates the deduced position of the bilayer. Cross-section widths of the glycerol channel and space around the fourfold axis are indicated.

**Fig. 3.** (A) Stereo view of experimental electron density (2Fo – Fo) in the selectivity region of the channel viewed from the periplasmic side. The orientation of G2 is clearly defined in the amphipathic channel. The density is contoured at 1.5σ and rendered with SPDBV (40) and POV-ray (41). (B) The hydrogen bonding network (dotted lines) showing interactions of G2, G3, and associated water molecules with the selectivity filter. Hydrophobic contacts are indicated as radial lines around atoms or residues. The view is rotated 90° about a vertical axis from the view in Fig. 1B. Distances (in angstroms) between heavy atoms are indicated. (C) Cross section through the channel (space-filling rendering) showing the interactions with G2OH2, O2 (orange) and its associated CH group (C2; purple), and hydrogen (green). The cross section shows the donor NEH of Arg²⁰⁶, acceptor carbonyl oxygen of Phe²⁰⁰, and hydrophobic corner formed by Trp48 and Phe²⁰⁰ viewed edge-on to the aromatic rings.
tion, the rate increases proportionately, indicating that transport is a GlpF-dependent process. The rates for a series of alditols demonstrate stereoselectivity of the channel. Of those tested, chiral molecules had lower rates than nonchiral molecules. The rates we observed precisely follow the trend observed for GlpF in whole cells (25), showing that in vivo characteristics are represented in reconstituted proteoliposomes.

Selectivity for glycerol versus water. GlpF has a much lower conductivity for water than do water channels (2). In GlpF conductivity for water may be impaired by the hydrophobic wall of the channel. The geometric requirements for hydration of water are stringent with typically four or five coordinate waters (24, 25). To pass the filter in single file, waters could not retain more than two water molecules as neighbors, and thus the energetic barrier for such dehydration is high.

Water channels conduct water and not glycerol; however, the side chains that form the hydrophobic corner are smaller and more polar. Trp28 is typically replaced by Phe or His; Phe230 is typically replaced by Ala, Thr, or Cys. The larger size of the channel should allow a greater degree of water coordination than in GlpF. Replacement of the indole of Trp28, and the phenyl ring of Phe230, also removes the hydrophobic corner so that there is no longer the potential for Van der Waals interactions with the alkyl chain of glycerol, consequently disfavoring alditol transport by water AQPs.

In AQPs, mercury blocks water transport at the thiol of Cys189 (26). The homolog of Cys189 in GlpF is Phe230, which lies at the narrowest position of the channel. Water transport has also been inactivated by substitution of the counterpart of Arg230 in AQPZ (Arg189) by Val or Ser (27). A mutation of this residue in human AQP2 causes nephrogenic diabetes insipidus. Although it has been suggested that the disease may be caused by impaired intracellular membrane traffic on the basis of expression in Xenopus oocytes (28), any mutation here must alter the water transport in the principal cells of the collecting duct in the kidney.

Regulated ion channels formed by AQPs. Certain AQPs act as ion channels under regulated conditions. Cation conductance has been induced in AQP1 by activation of cyclic GMP–dependent pathways (8, 29). In the latter case, cation conductance was blocked by Hg2+, suggesting that a conformational change induced in AQP1 can lead to cation conductance through the normal water channel. AQP6 conducts anions at pH’s below 5.5. However, Hg2+ leads to an increase in water and anion conductance (30), suggesting a different mechanism.

The fourfold symmetry axis: A potential ion channel in AQPs. Electron density on the fourfold axis, near the periplasmic surface, suggests coordination of two cations. The most external is coordinated octahedrally by four glutamates (Glu43, 2.9 Å) and water (2.1 Å) (Fig. 5). The most internal lies in between four indole nitrogens of Trp28 (3.6 Å) and 2.3 Å from the water. The electrostatics, bond lengths, and the presence of 300 mM Mg2+ suggest that the extra electron densities correspond to Mg2+ ions. The remainder of the pathway is surrounded by aliphatic side chains (Fig. 2B). This pathway is notable because many multimeric channels form ion channels on their symmetry axis. The dimensions along the axis are similar to those of the tetrameric, KcsA potassium channel (31). Like the KcsA channel, it also has a hydrophobic ~10 Å diameter cavity in the center of the bilayer and, overall, the axial pathway in GlpF is ~10 Å shorter than in the KcsA channel. However, the narrow cytoplasmic vestibule in KcsA is polar, whereas in GlpF it is nonpolar, closing it to all hydrated ions and polar molecules larger than water. Nevertheless, this pathway must be considered as a possible candidate for an ion channel in regulated AQPs.

Maintaining an electrochemical gradient. The channel in GlpF is too narrow to accommodate a hydrated ion, and the energetic cost of removing even a single water of hydration from an ion is prohibitive. Cation channels achieve
selective conductance—for example, in the case of the KcsA or gramicidin channels, by provision of multiple oriented carbonyl groups that each carry a $\delta^+$ charge, within narrow selective regions. These can replace or compensate for waters that are displaced from the first hydration shell (32–36).

Removal of water from, and fixation of a specific conformation of glycerol in the GlpF channel are also energetically costly processes. The channel achieves selectivity and conductance by provision of an amphipathic pathway that closely matches successive CH-OH moieties. In GlpF, polar interactions are only required at a permeant group that is polarizable and the hydration shell around an ion cannot be compensated for at all on the hydrophobic side. Consequently, all ions including OH$^-$ or H$_2$O$^+$ are excluded from the channel. By also requiring a permeant group that is polarizable with both $\delta^+$ donor and $\delta^-$ acceptor characteristics at each position in the selectivity filter, carbohydrates composed of CH-OH moieties are efficiently transported. Thus, the structure shows the precise mechanism for sharp selectivity and high conductance of alditols while excluding all charged species and water.

References and Notes

Structure of a Glycerol-Conducting Channel and the Basis for Its Selectivity

Daxiong Fu, Andrew Libson, Larry J. W. Miercke, Cindy Weitzman, Peter Nollert, Jolanta Krucinski and Robert M. Stroud

Science 290 (5491), 481-486.
DOI: 10.1126/science.290.5491.481