

# Architecture and Selectivity in Aquaporins 2.5: Å X-Ray Structure of Aquaporin Z

David F. Savage<sup>1,2</sup>, Pascal F. Egea<sup>1</sup>, Yaneth Robles-Colmenares<sup>1</sup>, Joseph D. O'Connell III<sup>1</sup>, Robert M. Stroud<sup>1\*</sup>

**1** Department of Biochemistry and Biophysics, University of California School of Medicine, San Francisco, California, United States of America, **2** Graduate Group in Biophysics, University of California, San Francisco, California, United States of America

**Aquaporins are a family of water and small molecule channels found in organisms ranging from bacteria to animals. One of these channels, the *E. coli* protein aquaporin Z (AqpZ), has been shown to selectively conduct only water at high rates. We have expressed, purified, crystallized, and solved the X-ray structure of AqpZ. The 2.5 Å resolution structure of AqpZ suggests aquaporin selectivity results both from a steric mechanism due to pore size and from specific amino acid substitutions that regulate the preference for a hydrophobic or hydrophilic substrate. This structure provides direct evidence on the molecular mechanisms of specificity between water and glycerol in this family of channels from a single species. It is to our knowledge the first atomic resolution structure of a recombinant aquaporin and so provides a platform for combined genetic, mutational, functional, and structural determinations of the mechanisms of aquaporins and, more generally, the assembly of multimeric membrane proteins.**

## Introduction

The aquaporin family, composed of transmembrane water-conducting channels (aquaporins) and glycerol (and water)-conducting channels (aquaglyceroporins), is a group of highly selective passive transporters (Heller et al. 1980; Preston et al. 1992; Park and Saier 1996). The diversity of the aquaporin family is embodied by the human proteome, where at least ten different aquaporins are expressed in tissues and cells such as brain, kidneys, and erythrocytes. Aquaporins play a fundamental role in osmoregulation, and mutations are responsible for human diseases ranging from diabetes insipidus to congenital cataract formation (Borgnia et al. 1999b).

The structural architecture of aquaporins was first determined by electron microscopy (Walz et al. 1997; Murata et al. 2000). Later, high-resolution X-ray structures of the recombinant *Escherichia coli* aquaglyceroporin glycerol facilitator (GlpF) and bovine aquaporin 1 (AQP1), obtained from red blood cells, elucidated the mechanisms by which aquaporins preserve the electrochemical membrane potential and selectively conduct water and linear polyalcohols (Fu et al. 2000; Sui et al. 2001). Until now, there was no structure of a recombinantly expressed water-selective aquaporin, which would allow the systematic analysis of amino acid substitution by mutagenesis, structure, and function.

Besides GlpF, the *E. coli* genome contains a second aquaporin, aquaporin Z (AqpZ). AqpZ is a highly efficient water channel and conducts water at rates six times that of GlpF (Calamita et al. 1995; Borgnia et al. 1999a). AqpZ has been used to probe substrate selectivity and has shown promise as a structural target, providing one of the early electron microscopy studies of an aquaporin (Calamita et al. 1998; Ringler et al. 1999; Borgnia and Agre 2001). The pair of AqpZ and GlpF exists in the same organism, implying similar lipid, chemical, and osmotic environments, and thus presents a unique opportunity to study aquaporin structure and function in a genetically and biochemically tractable system.

In anticipation of the mutagenic probes of function and structure, we report the X-ray structure of wild-type AqpZ to 2.5 Å resolution.

## Results

### Structure of AqpZ

Three-dimensional crystals of AqpZ were grown and diffraction data to better than 2.5 Å were collected under cryoconditions (space group P4, a = 93.6 Å, c = 80.4 Å, with two protomers per asymmetric unit, henceforth called protomers A and B). The structure was solved by molecular replacement and refined to  $R_{\text{cryst}}$  of 22.7% and  $R_{\text{free}}$  of 26.8% using reflections to 2.5 Å (Table 1; Table 2). There are two tetramers in the unit cell composed of four copies of protomer A as a tetramer and four of protomer B, respectively. The positioning of these tetramers in the unit cell shows quasi-I4 symmetry in which the body-centered tetramer is slightly rotated around the 4-fold axis. Protomers A and B are involved in different crystal-packing interactions and display slightly different conformations, particularly near the periplasmic vestibule; their root mean square difference (RMSD) is 0.44 Å. Pore electron density was stronger for protomer A. Noncrystallographic symmetry did not prove useful in model building, presumably because of the differences between protomer A and protomer B.

Received August 29, 2003; Accepted October 10, 2003; Published December 22, 2003

DOI: 10.1371/journal.pbio.0000072

Copyright: © 2003 Savage et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abbreviations: AQP1, aquaporin 1; AqpZ, aquaporin Z; GlpF, glycerol facilitator; OG, n-octyl-D-glucopyranoside; NPA motif, asparagine-proline-alanine motif; RMSD, root mean square difference

Academic Editor: Sanford Simon, The Rockefeller University

\* To whom correspondence should be addressed. E-mail: stroud@msg.ucsf.edu



**Table 1.** Data Collection Statistics

Dataset	Advanced Light Source Beamline 8.3.1 Native
Wavelength (Å)	1.1
Resolution (Å)	50–2.3
Total reflections	58,536
Unique reflections	27,408
Redundancy	2.1
Completeness (last shell) (%)	89.9 (90.7)
$R_{sym}$ (%)	6.38
$I/\sigma$	12.0

DOI: 10.1371/journal/pbio.0000072.t001

The protomer structure of AqpZ displays the canonical aquaporin fold of six transmembrane helices and two half-membrane-spanning helices (M1–M8) in a right-handed helical bundle. As shown previously, the protomer oligomerizes to form a homotetramer (Figure 1A). The amino terminus begins on the cytoplasmic side. M1 crosses the membrane and loops to M2, which recrosses the membrane (Figure 1B). M2 is followed by a loop from residue 54 to residue 62 that contains the four carbonyls that project into the pore near the cytoplasmic side. Following this loop is helix M3, which contains the signature asparagine–proline–alanine (NPA) motif and is oriented in such a way as to point its positive dipole towards the central water position in the channel. The first domain of AqpZ ends with M4 ending on the periplasmic side. This is followed by a loop from residue 103 to residue 131 that descends into the periplasmic vestibule and leads into the carboxy-terminal segment. M5–M8 reiterates the amino-terminal topology, except now beginning on the periplasmic and ending on the cytoplasmic side. This pseudo-2-fold symmetry creates a general architecture in which the main chain carbonyls establish water-binding sites along the channel and in which sidechain variation determines channel size and chemistry (Figure 1B).

## The Channel

The aquaporin channel is a long (approximately 28 Å) and narrow (less than 4 Å diameter) pore that widens out to periplasmic and cytoplasmic vestibules. The channel is formed by the packing of helices M1–M3 and M5–M7 and is amphipathic, establishing a single-file water conduction pathway. The hydrophilic nature of the channel results from the four adjacent carbonyls of G59 (GlpF number 64), G60(65), H61(66), and F62(67) from the amino-terminal domain and the quasi-2-fold related N182(199), T183(200), S184(201), and V185(202) from the carboxy-terminal domain (Figure 1B). The hydrophobic nature of the channel results from an abundance of valines, phenylalanines, and isoleucines within the channel.

The channel contains two highly conserved regions, the selectivity filter and the NPA region. Located approximately 7 Å inside the periplasmic vestibule, the selectivity filter is the narrowest point (diameter of approximately 2 Å) in the entire channel (Figure 2). It is formed by the sidechains of F43(48),

**Table 2.** Refinement Statistics

Refinement Statistics	AqpZ
Resolution (Å)	50.0–2.5
Reflections in working set	20,216
Reflections in test set (6.9%)	1,504
$R_{cryst}$ (%)	22.7 %
$R_{free}$ (%)	26.8 %
RMSD bonds (Å)	0.008
RMSD angles (°)	1.24
<b>Protomer A and Protomer B</b>	
Nonhydrogen protein atoms	1661 and 1641
Nonhydrogen heteroatoms	12 and 88
Solvent molecules	131
Average B-factors (Å <sup>2</sup> )	66.9

RMSD is the root-mean square deviation from ideal geometry.  
 $R_{sym} = \frac{\sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl,i} \rangle|}{\sum_{hkl} \sum_i I_{hkl,i}}$ , where  $\langle I_{hkl,i} \rangle$  is the average intensity of the multiple  $hkl_i$  observations for symmetry-related reflections.

$R_{cryst} = \frac{\sum |F_{obs} - F_{calc}|}{\sum F_{obs}}$ , where  $F_{obs}$  and  $F_{calc}$  are observed and calculated structure factors,  $R_{free}$  is calculated from a randomly chosen 6.9% of reflections, and  $R_{cryst}$  is calculated over the remaining 93.1% of reflections.

DOI: 10.1371/journal/pbio.0000072.t002

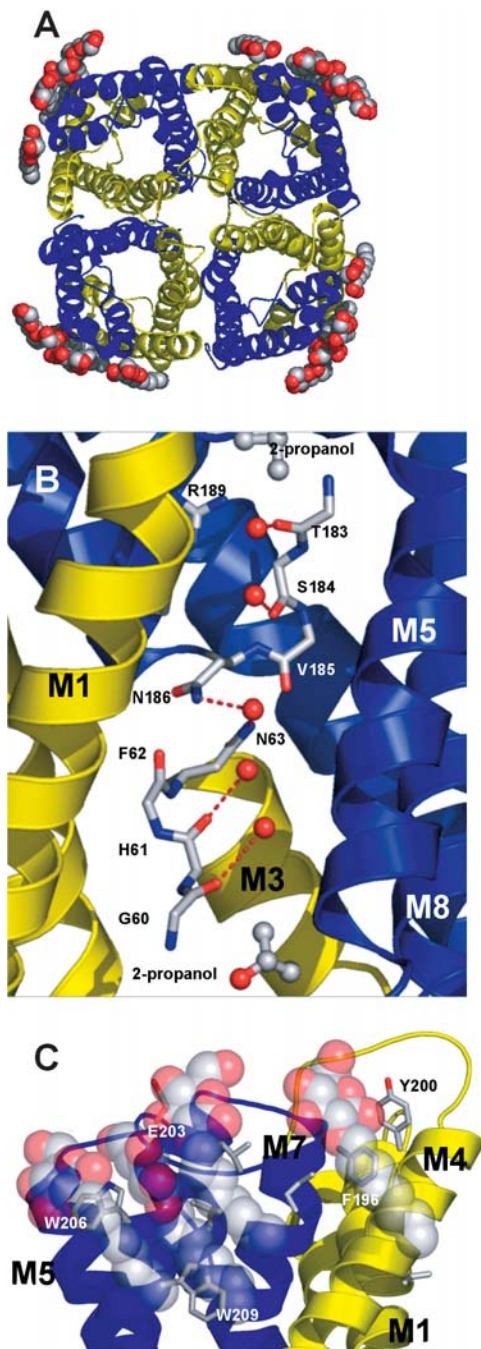
H174(191), and R189(206) and the carbonyl of T183(200). The trio of H174(191), T183(200), and R189(206) create a hydrophilic triangle opposite the hydrophobic F200.

The NPA sequences from each M1–M4 and M5–M8 domain form a constrained and interlocked junction around the quasi-2-fold axis, based on asparagine, proline, and alanine from the amino-terminal ends of M3 and M7 (Figure 3). The alanine sidechain and the proline ring make a head-to-tail, twinned, largely hydrophobically driven contact with the proline and alanine of the other domain. Each asparagine sidechain is oriented by two almost ideal hydrogen bonds. For N63(68), these bonds are one from OD1 to the NH of A65(70) and one from NH2 to the carbonyl of V185(202). Similar interactions occur at N186(203). This highly constrains and orients both asparagine sidechains to project their ND2 groups strictly into the pore, which are hydrogen-bond donors to the central water molecule.

Five waters are unambiguously located in the channel (Figure 4). The water is arranged in single file, with hydrogen bonding as donors to the projecting carbonyls from AqpZ and as donors to neighboring waters. From the periplasmic side to the cytoplasmic side, there are waters located adjacent to the carbonyls of T183(200) (OO distance of 3.0 Å), S184(201) (3.2 Å), H61(66) (3.0 Å), and G60(65) (3.4 Å). The waters are at appropriate (less than 3.2 Å) distances from each other for hydrogen bonding. No electron density was observed adjacent to the carbonyls of G59(64) or V185(202).

In protomer B, four n-octyl-D-glucopyranoside (OG) molecules are positioned at the potential location of the periplasmic membrane leaflet (see Figure 1A and 1C). The detergent head groups pack against the aromatic residues F196(224), W200(228), and W206(234) near helix M8 and the lipid tails run towards the centerline of AqpZ. Their conformation suggests a belt-like micelle surrounding the full tetramer.





**Figure 1.** Structure of AqpZ

Three-dimensional fold of AqpZ with the quasi-2-fold related segments in yellow (residues 1–117) and blue (residues 188–231).

(A) Cartoon representation of the AqpZ tetramer with OG detergent molecules represented as spheres; view is from the periplasmic side. Atoms are colored according to atom type (red, oxygen; gray, carbon; blue, nitrogen; yellow, sulfur).

(B) Cartoon representation of the AqpZ monomer, with M2 and M6 removed for ease of viewing. Single-file water is shown hydrogen-bonding to carbonyls of main chain. Central water is shown accepting a hydrogen bond from the NH<sub>2</sub> group of Asn63 and Asn186. Sidechains of the selectivity filter are also shown. Isopropanol molecules located in density are shown as sticks, just outside the channel.

(C) A view from the membrane plane of the OG micelle interactions with the periplasmic segment of AqpZ. OG molecules pack against the aromatic sidechains, while making hydrogen bonds with the main chain carbonyls, E203, and each other. All figures were made with PyMOL (DeLano 2003).

DOI: 10.1371/journal.pbio.0000072.g001

Three isopropanol molecules are located in the cytoplasmic and periplasmic vestibules, just outside the channel (see Figure 1B). The propyl groups are packed against hydrophobic sidechains, while the hydroxyl groups participate in hydrogen bonding with vestibule waters.

## Discussion

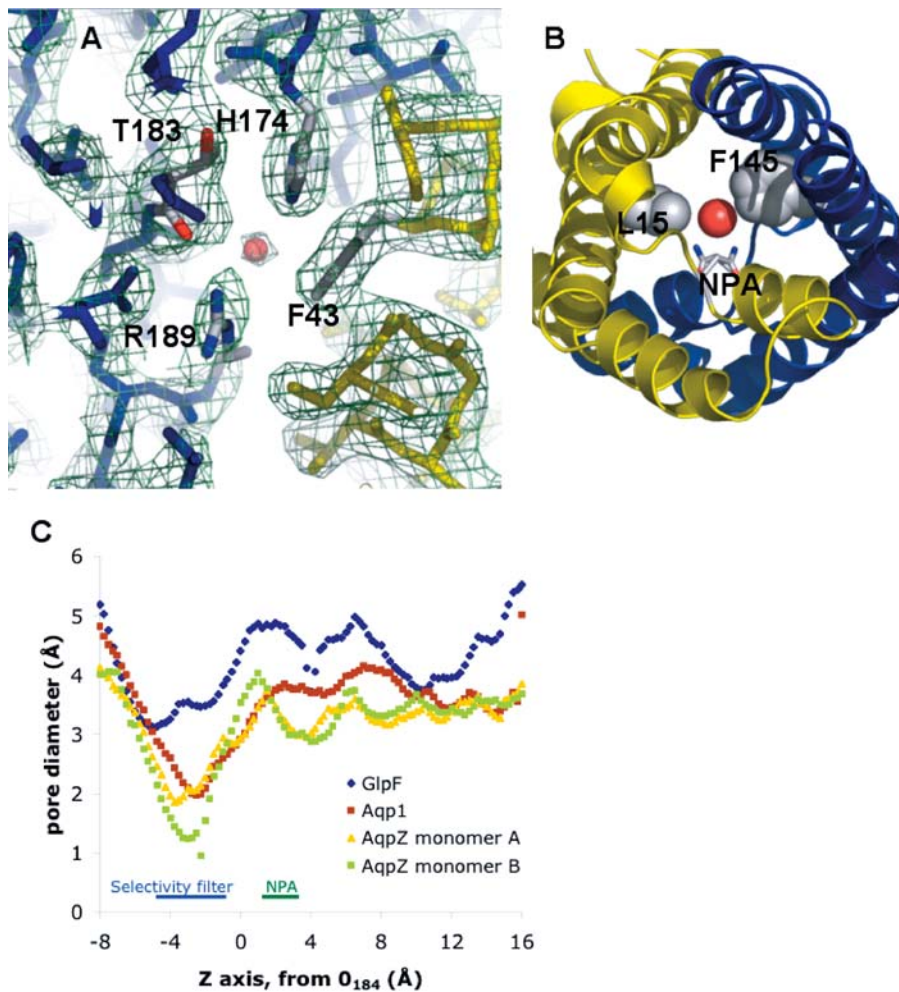
### Selectivity of Aquaporins

The *E. coli* genome encodes two aquaporins, GlpF and AqpZ. These two channels represent the functional diversity of the aquaporin family and, as both are *E. coli* transmembrane proteins, exist in the same lipid, chemical, and osmotic environments. Both channels preserve the electrochemical gradient and display selective transport, yet have a different biological function. A comparison of the two structures delineates the nature of aquaporin selectivity, uncluttered by species differentiation.

In vitro and in vivo functional experiments demonstrate AqpZ's preference for water transport and GlpF's preference for glycerol (Maurel et al. 1994; Borgnia and Agre 2001). To first approximation, this preference is due to channel size, as depicted in Figure 2 for the known structures. This calculation shows that aquaporins have a smaller pore size than aquaglyceroporins and that the selectivity filter is the narrowest point in the channel for all three proteins. In AqpZ, this selectivity filter is formed by the sidechains of F43(48), H174(191), and R189(206) and the carbonyl of T183(200) (see Figure 2A). The presence of a bound water molecule (distances: 2.7 Å to NE2 H174, 3.0 Å to O T183, 2.6 Å NH<sub>2</sub> R189) confirms the selectivity filter's preference for a small hydrophilic substrate (see Figure 4). The AQP1 selectivity filter is nearly identical, with a cysteine substituted for threonine, which is also the basis of inhibition by mercury (Preston et al. 1992). Not surprisingly, AQP1 and AqpZ have similar in vitro function. In sharp contrast, GlpF contains the typical aquaglyceroporin substitutions of F43W, H199G, and T200F; the GlpF wild-type structure contains both a water and glycerol molecule bound at the selectivity filter.

The GlpF selectivity filter, larger and more hydrophobic than in AqpZ, is reminiscent of the maltoporin “greasy slide” sugar-binding sites (Dutzler et al. 1996; Van Gelder et al. 2002). Maltoporins, a family of bacterial outer membrane transporters, facilitate the translocation of maltooligosaccharides using a “greasy slide” hydrophobic path of seven aromatic residues along the central pore. Such a path, with a preference for nonpolar groups, can increase the effective concentration of ligand near the channel and thereby increase the probability of a transport event. The periplasmic vestibule of GlpF also has a hydrophobic patch of residues leading into the selectivity filter, while in AqpZ the polar sidechain of N182 chemically and structurally caps off the already hydrophilic selectivity filter.

Three molecules of isopropanol, present at 4% in the crystallization solution, are located just outside the channel in the vestibule regions (see Figure 1B). They pack against hydrophobic sidechains, forming favorable van der Waals contacts with A27, F36, V39, F43, A62, T153, and I178 and hydrogen bonding with nearby water. Based on its vestibular location, isopropanol is seemingly too big for transport, though this idea has not been tested experimentally. Despite its presence at high concentration, one substrate not seen in



**Figure 2.** Channel Constriction in Aquaporins

(A) A view of the aquaporin selectivity filter from the periplasmic side. Experimental electron density ( $2F_{\text{obs}} - F_{\text{calc}}$ ) is contoured at  $1.1 \sigma$ .

(B) Secondary constriction at the NPA motif due to F145 and L15. The drawn water is HOH1032, hydrogen-bonded to the NPA motif asparagines.

(C) Pore diameters for the aquaporin X-ray structures, calculated with HOLE2. The AqpZ monomers (protomers) A and B refer to the crystallographically distinct monomers in the unit cell.

DOI: 10.1371/journal.pbio.0000072.g002

the AqpZ structure is glycerol. This contrasts with the structure of GlpF, where ordered glycerol was located at three sites in the channel, including the NPA motif. This absence confirms previous functional data and suggests a steric mechanism of selectivity.

There are five well-oriented waters in the AqpZ channel, forming a chain of water nearly the length of the channel. Four waters are hydrogen-bond donors to the carbonyls of G60(65), H61(66), T183(200), and S184(201). The fifth central water molecule is a hydrogen-bond acceptor from the ND2 groups of the NPA motif asparagines (see Figure 1B and Figure 4). Normally, a single-file column of water such as we observe should conduct protons. In 1806, de Grotthuss, based on electrolysis experiments, postulated that polar water molecules could align themselves in long chains from cathode to anode, in essence forming a wire (de Grotthuss 1806). Bernal and Fowler (1933), using quantum mechanics to explain de Grotthuss's qualitative hypothesis, postulated that protons could easily jump between neighboring waters, thereby making protons highly mobile in solution. It is therefore remarkable that aquaporins, which inherently contain a chain of water, preserve the electrochemical gradient (de Groot and Grubmüller 2001). A possible mechanism for this disparity is disruption of the proton jumping mechanism at the NPA region.

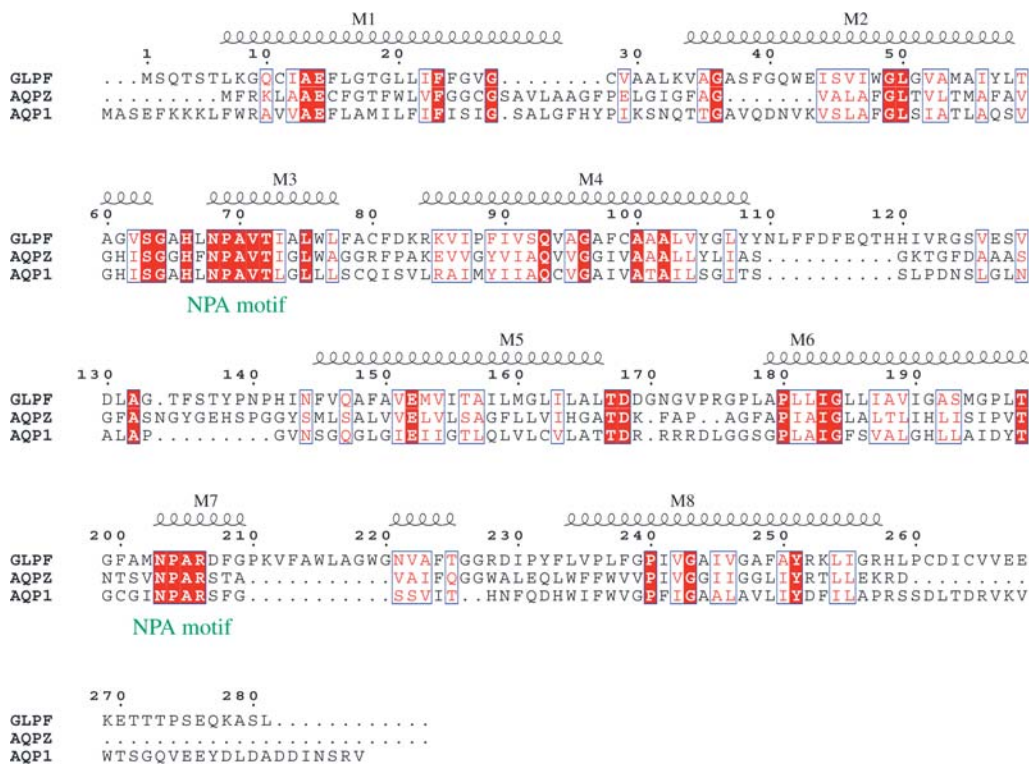
The NPA asparagine ND2 groups act as hydrogen-bond

donors to the central water, locking it in a conformation such that it can only donate hydrogen bonds to nearby single-file waters. Therefore, while the central water can readily donate a proton, it can never accept one. This prevents adjacent water from performing the reorientation necessary to conduct protons, and the proton-conducting "wire" is broken. This effect, termed global orientational tuning (Tajkhorshid et al. 2002), is also aided by the positive dipoles of M3 and M7 (Murata et al. 2000), which are aimed directly at the central water in a manner reminiscent of potassium channels (Doyle et al. 1998).

For water flux to occur, the central water must quickly be replaced by another, and the uniquely oriented carbonyls proximal to the NPA motif, those of F62 and V185, may reorient the NPA region water as it moves away from the center (see Figure 4). While the carbonyls of G59(64), G60(65), H61(66), N182(199), T183(200), and S184(201) are nearly orthogonal to the channel axis, those of F62(67) and V185(202) run parallel to the axis. As these carbonyls are proximal to the asparagines of the NPA motif, their unique conformation may be necessary to allow water to reorient as it passes the NPA region at the quasi-2-fold axis. This hypothesis is supported by molecular dynamics simulations (unpublished data).

Besides breaking the proton wire, the NPA region is likely to play a role in selectivity. Sequence analysis has shown that



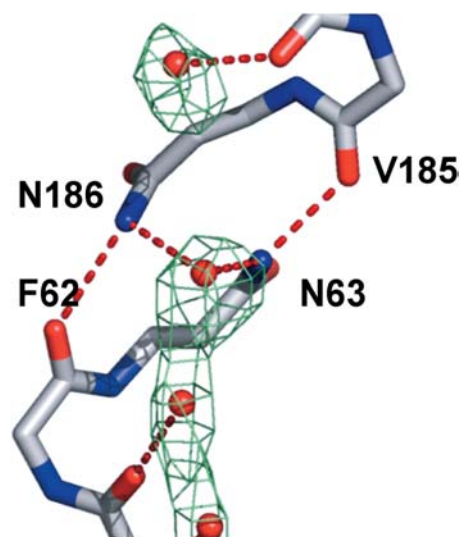


**Figure 3.** Sequence Alignment of Aquaporins with Known Structure

Alignment of GlpF, AqpZ, and Aqp1 is numbered according to GlpF. Helices are shown and labeled M1–M8. Residue positions of similar chemical nature are shown in blue boxes; identical residues are shown in red.

DOI: 10.1371/journal.pbio.0000072.g003

positions 15(21) and 145(159), residues from helix M1 and M5, respectively, are correlated; aquaglyceroporins typically contain leucine at both positions, while aquaporins have an aromatic at one of the two (Heymann and Engel 2000) (see



**Figure 4.** Water at the NPA Region

N63 and N186 donate hydrogen bonds to the central water by projecting their NH2 moieties into the pore. This conformation is aided by a hydrogen bond from the adjacent carbonyls of V185 and F62, respectively. Experimental electron density ( $2F_{\text{obs}} - F_{\text{calc}}$ ) is contoured at 0.7  $\sigma$ .

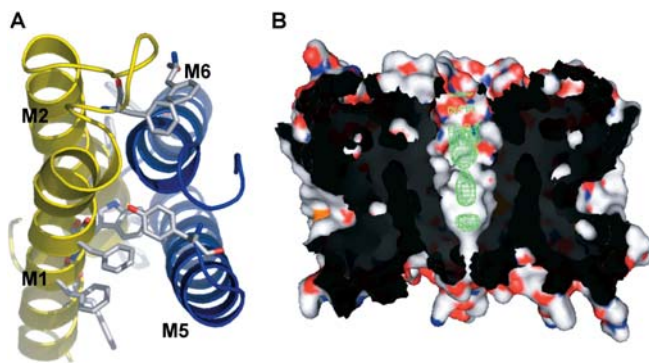
DOI: 10.1371/journal.pbio.0000072.g004

Figure 3). In AqpZ, the sidechains of L15(21) and F145(159) project into the pore and narrow it to a diameter of 3.0 Å; the GlpF diameter is 4.0 Å (see Figure 2B) and AQP1 is 3.5 Å. In AqpZ, the central water located opposite the NH2 of N63(68) and N186(203) is better resolved and has a shorter hydrogen-bond distance (approximately 2.8 Å versus approximately 3.5 Å) than the corresponding water in the GlpF structure without glycerol in the crystallization buffer. Furthermore, with glycerol in the buffer, GlpF readily crystallizes with glycerol bound at this position. Thus, this secondary constriction emphasizes the preference for a small hydrophilic substrate in aquaporins.

Continuing along the pore axis towards the cytoplasm, aquaporins display a narrower channel. This pore difference does not result from helix rearrangement, as the main chain RMSD for AqpZ and GlpF is 1.6 Å and that of AqpZ and AQP1 is 1.2 Å, but comes from sidechain variation. Strict conservation of helical tertiary structure suggests one can effectively apply methods such as homology modeling to predict sidechain conformation and function in the pore region (Marti-Renom et al. 2000).

### C $\alpha$ H $\cdots$ O Bonds in Aquaporins

The strict tertiary conservation of aquaporins underscores the unique features of helix packing in membrane proteins. A survey of  $\alpha$ -helical membrane proteins has revealed that transmembrane helices are often packed at distances close enough for C $\alpha$ H $\cdots$ O hydrogen-bond formation (less than 3.5 Å) (Wahl and Sundaralingam 1997; Senes et al. 2001). This type of interaction is facilitated by glycine, an amino acid that



**Figure 5.** Oligomerization of AqpZ

(A) Protomer-protomer interface. The interface is composed of helices M1 and M2 from one protomer (yellow) and M5 and M6 from a second protomer (blue). The helices participate in knobs into holes packing, and the interface is extensive ( $3,340 \text{ \AA}^2$ ) owing to the large number of aromatic residues.

(B) Abstruse electron density along the 4-fold axis (colored in green). Two protomers are display in surface rendering. The experimental electron density ( $2F_{\text{obs}} - F_{\text{calc}}$ ) is contoured at  $1.1 \sigma$ .

DOI: 10.1371/journal.pbio.0000072.g005

is overrepresented in transmembrane segments, because it allows short interhelix distances (Senes et al. 2000). In an analysis of AqpZ, we identified 15 potential bonds. With an estimated energy of 2.5–3.0 kcal/mol per bond (in vacuo), this is a partial explanation for the stability of AqpZ in denaturing conditions (Borgnia et al. 1999a; Scheiner et al. 2001). These bonds are also likely to play a role in the dynamics of other  $\alpha$ -helical membrane proteins.

The structure of lactose permease was recently solved using a thermostable cysteine-to-glycine mutant (Abramson et al. 2003). The site-directed mutation occurs at the interface between two transmembrane helices and appears to lock the protein in a conformation such that it can tightly bind substrate, but not translocate it. Engineering ultrastable glycine mutants may prove useful in structural studies of other, less robust, membrane proteins.

### Detergents in Aquaporins

Four OG detergent molecules were located bound to the periplasmic surface of each AqpZ molecule in a belt-like fashion (see Figure 1C). The detergents are situated near the carboxyl terminus of M7 and the amino terminus of M8, capping off the helices and forming hydrogen bonds with the carbonyls of the loop between helix M7 and helix M8. The hydrophobic sugar rings and tails pack against the nearby residues, many of which are aromatic.

The GlpF and AQP1 structures each contained three detergent molecules at virtually identical location on the outside surface, presumably owing to the abundance of aromatic sidechains in all three proteins at this location. This abundance is present in all aquaporins and may be important for lipid interaction. In AqpZ, there are also both acidic and basic residues interacting with the OG head group, suggesting the native lipid may be a zwitterion-like phosphatidyl ethanolamine, the most common *E. coli* lipid (Neidhardt et al. 1996). The importance of native lipids has been demonstrated in the folding and function of ion channels and may

be important in designing future aquaporin functional assays (Valiyaveetil et al. 2002).

### Tetramer Axis

While it is clear the monomer is the functional unit, the existence of aquaporin tetramers in nature reinforces the importance of oligomerization. Fundamentally, tetramerization is driven by the energetically favorable assembly of four protomers. The protomer-protomer interface is large, tightly packed, and formed by helices M1 and M2 of one protomer and the quasi-2-fold related M5 and M6 of the neighboring protomer (Figure 5A). This interface is therefore repeated four times. In AqpZ, the interface is  $3,340 \text{ \AA}^2$ , in large part owing to the presence of 11 aromatic residues. The GlpF and AQP1 interfaces are  $3,060 \text{ \AA}^2$  and  $3,180 \text{ \AA}^2$ , respectively, with five and three aromatic residues, respectively. Strikingly, the interface surface area correlates positively with biochemical stability; GlpF tends to aggregate in solution, AQP1 is well-behaved, and AqpZ is a stable tetramer in even mild denaturing conditions. In this protomer-protomer interface, helices M2 and M6 form the tetramer pore.

There is remarkable tertiary and quaternary structure conservation, and the tetramer pore remains nearly constant in shape between the known structures. All structures appear open towards the periplasm but closed towards the cytoplasm, a necessity in preserving the electrochemical gradient, and contain a large (7–10 Å diameter) central cavity. Except for a glutamate near the periplasmic opening in AQP1 and GlpF, the residues lining the pore are hydrophobic, suggesting a very large energetic barrier of translocation for any polar substrate. Notably, all three X-ray structures contained electron density along the pore, signifying the presence of multiple molecules (Figure 5B). It remains to be seen whether this pore is functional or primarily a structural necessity to facilitate monomer function.

### Materials and Methods

**Expression and purification.** AqpZ was cloned by PCR from isolated *E. coli* genomic DNA into the pET28 expression vector with kanamycin selection and an amino-terminal 6xHis affinity tag (Novagen, Madison, Wisconsin, United States). The *E. coli* strain C43 (Miroux and Walker 1996) was transformed, grown to 0.6–1 OD at 600 nm in LB with 20 mg/l kanamycin, and induced with 1 mM isopropyl-D-thiogalactoside. Cells were harvested and lysed by sonication in 20 mM Tris (pH 7.4), 100 mM NaCl, 0.5 mM phenylmethylsulfonyl fluoride, and 5 mM 2-mercaptoethanol. Cellular debris were pelleted at  $10,000 \times g$  for 45 min and discarded. Membranes were recovered from supernatant by  $100,000 \times g$  centrifugation for 90 min. AqpZ was solubilized from membranes by agitation in 20 mM Tris (pH 7.4), 100 mM NaCl, 5 mM 2-mercaptoethanol, 10% glycerol, and 270 mM OG (Anatrace, Maumee, Ohio, United States) for 12–16 h. Solubilized protein was bound in batch to Ni-NTA resin (Qiagen, Valencia, California, United States), washed, and eluted with 20 mM Tris (pH 7.4), 500 mM NaCl, 5 mM 2-mercaptoethanol, 10% glycerol, 40 mM OG, and 250 mM imidazole. Imidazole was removed using a Bio-Rad (Hercules, California, United States) Econo-Pac DG10 desalting column, and the histidine tag was removed following the protocol of Borgnia et al. (1999a). The final purification step was performed on a Pharmacia Superose 12 column (Pfizer, New York, New York, United States).

**Crystallization.** Purified AqpZ was concentrated to approximately 20 mg/ml and crystallized in 28% polyethylene glycol monomethyl ether 2000, 100 mM sodium cacodylate (pH 6.5), 200 mM  $\text{MgCl}_2$ , 4% isopropanol in hanging drop plates (Nextal Biotechnologies, Montreal, Quebec, Canada) by vapor diffusion at room temperature. Crystals grew to  $300 \mu\text{m} \times 300 \mu\text{m} \times 150 \mu\text{m}$  in several days. Crystals were flash frozen in a 90K nitrogen gas stream.

**Data collection and model building.** Diffraction intensities were



collected at the Advanced Light Source Beamline 8.3.1 (Lawrence Berkeley National Laboratory, Berkeley, California, United States) on a Quantum CCD detector (Area Detector Systems, Poway, California, United States). Data were processed with DENZO/SCALEPACK (Otwinowski and Minor 1997). The structure was solved by molecular replacement using the AQP1 structure as a search model. The model was refined with CNS and built using Moloc (Gerber and Müller 1995; Brünger 1996).

## Supporting Information

### Accession Numbers

Coordinates of the structure have been deposited in the Research Collaboratory for Structural Bioinformatics' Protein Data Bank (PDB) (accession code IRC2), found at <http://www.rcsb.org/pdb/>.

### References

- Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, et al. (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 301: 610–615.
- Bernal JD, Fowler RH (1933) A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J Chem Phys* 1: 515–548.
- Borgnia MJ, Agre P (2001) Reconstitution and functional comparison of purified GlpF and AqpZ, the glycerol and water channels from *Escherichia coli*. *Proc Natl Acad Sci U S A* 98: 2888–2893.
- Borgnia MJ, Kozono D, Calamita G, Maloney PC, Agre P (1999a). Functional reconstitution and characterization of AqpZ, the *E. coli* water channel protein. *J Mol Biol* 291: 1169–1179.
- Borgnia M, Nielsen S, Engel A, Agre P (1999b) Cellular and molecular biology of the aquaporin water channels. *Annu Rev Biochem* 68: 425–458.
- Brünger AT (1996) X-PLOR version 3.843. New Haven, Connecticut: Yale University.
- Calamita G, Bishai WR, Preston GM, Guggino WB, Agre P (1995) Molecular cloning and characterization of AqpZ, a water channel from *Escherichia coli*. *J Biol Chem* 270: 29063–29066.
- Calamita G, Kempf B, Bonhivers M, Bishai WR, Bremer E, et al. (1998) Regulation of the *Escherichia coli* water channel gene *aqpZ*. *Proc Natl Acad Sci U S A* 95: 3627–3631.
- de Groot BL, Grubmüller H (2001) Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and GlpF. *Science* 294: 2353–2357.
- de Groot CJT (1806) Sur la décomposition de l'eau et des corps qu'elle tient en dissolution à l'aide de l'électricité galvanique. *Ann Chim* 58: 54–74.
- DeLano WL (2003) The PyMOL molecular graphics system. San Carlos, California: DeLano Scientific. Available at <http://pymol.sourceforge.net/> via the Internet. Accessed October 25, 2003.
- Doyle DA, Morais Cabral J, Pfuetzner RA, Kuo A, Gulbis JM, et al. (1998) The structure of the potassium channel: Molecular basis of K<sup>+</sup> conduction and selectivity. *Science* 280: 69–77.
- Dutzler R, Wang YF, Rizkallah P, Rosenbusch JP, Schirmer T (1996) Crystal structures of various maltooligosaccharides bound to maltoporin reveal a specific sugar translocation pathway. *Structure* 4: 127–134.
- Fu D, Libson A, Miercke LJ, Weitzman C, Nollert P, et al. (2000) Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* 290: 481–486.
- Gerber PR, Müller K (1995) MAB: A generally applicable molecular force field for structural modeling in medicinal chemistry. *J Comput Aided Mol Des* 9: 251–268.
- Heller KB, Lin EC, Wilson TH (1980) Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. *J Bacteriol* 144: 274–278.
- Heymann JB, Engel A (2000) Structural clues in the sequences of the aquaporins. *J Mol Biol* 295: 1039–1053.
- Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, et al. (2000)

## Acknowledgments

We thank the staff of Beamline 8.3.1 at the Advanced Light Source, Lawrence Berkeley National Laboratory, and Chris Waddling for X-ray support. We also thank Janet Finer-Moore, John D. Anderson, and Michael J. LaFemina for intellectual discussion. This work was supported by National Institutes of Health (NIH) grant GM24485 (to RMS). DFS received predoctoral support from NIH grant GM08284.

**Conflicts of interest.** The authors have declared that no conflicts of interest exist.

**Author contributions.** DFS conceived and designed the experiments. DFS, PFE, YRC, and JDO performed the experiments. DFS, PFE, and RMS analyzed the data. DFS contributed reagents/materials/analysis tools. DFS and RMS wrote the paper. ■

- Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* 29: 291–325.
- Maurel C, Reizer J, Schroeder JI, Chrispeels MJ, Saier MH Jr (1994) Functional characterization of the *Escherichia coli* glycerol facilitator, GlpF, in *Xenopus* oocytes. *J Biol Chem* 269: 11869–11872.
- Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260: 289–298.
- Murata K, Mitsuoka K, Hirai T, Walz T, Agre P, et al. (2000) Structural determinants of water permeation through aquaporin-1. *Nature* 407: 599–605.
- Neidhardt FC, Curtiss R, Lin EC, editors (1996) *Escherichia coli* and *Salmonella*. 2<sup>nd</sup> edition. Washington, DC: American Society for Microbiology Press. 3,008 p.
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276: 307–326.
- Park JH, Saier MH Jr (1996) Phylogenetic characterization of the MIP family of transmembrane channel proteins. *J Membr Biol* 153: 171–180.
- Preston GM, Carroll TP, Guggino WB, Agre P (1992) Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* 256: 385–387.
- Ringler P, Borgnia MJ, Stahlberg H, Maloney PC, Agre P, et al. (1999) Structure of the water channel AqpZ from *Escherichia coli* revealed by electron crystallography. *J Mol Biol* 291: 1181–1190.
- Scheiner S, Kar T, Gu Y (2001) Strength of the C $\alpha$ -H $\cdots$ O hydrogen bond of amino acid residues. *J Biol Chem* 276: 9832–9837.
- Senes A, Gerstein M, Engelman DM (2000) Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J Mol Biol* 296: 921–936.
- Senes A, Ubarretxena-Belandia I, Engelman DM (2001) The C $\alpha$ -H $\cdots$ O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci U S A* 98: 9056–9061.
- Sui H, Han BG, Lee JK, Walian P, Jap BK (2001) Structural basis of water-specific transport through the AQP1 water channel. *Nature* 414: 872–878.
- Tajkhorshid E, Nollert P, Jensen MO, Miercke LJ, O'Connell J, et al. (2002) Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science* 296: 525–530.
- Valiyaveetil FI, Zhou Y, MacKinnon R (2002) Lipids in the structure, folding, and function of the KcsA K<sup>+</sup> channel. *Biochemistry* 41: 10771–10777.
- Van Gelder P, Dumas F, Bartoldus I, Saint N, Prilipov A, et al. (2002) Sugar transport through maltoporin of *Escherichia coli*: Role of the greasy slide. *J Bacteriol* 184: 2994–2999.
- Wahl MC, Sundaralingam M (1997) C–HO hydrogen bonding in biology. *Trends Biochem Sci* 22: 97–102.
- Walz T, Hirai T, Murata K, Heymann JB, Mitsuoka K, et al. (1997) The three-dimensional structure of aquaporin-1. *Nature* 387: 624–627.

