

## Structural Basis of Water Specific Transport through AQP1 Water Channel

Haixin Sui<sup>1+</sup>, Bong-Gyoon Han<sup>1+</sup>, John K. Lee<sup>1,2</sup>, Peter Walian<sup>1++</sup> and Bing K. Jap<sup>1,2++</sup>

<sup>1</sup>Life Sciences Division, Lawrence Berkeley National Laboratory  
University of California, Berkeley, CA 94720

<sup>2</sup>Graduate Group in Comparative Biochemistry  
University of California, Berkeley  
Berkeley, CA 94720

The structure of AQP1 water channel was determined to 2.2 Å resolution. The channel consists of three topological elements, an extracellular and a cytoplasmic vestibule connected by an extended narrow pore or selectivity filter. At the extracellular end of the selectivity filter is the constriction region, which establishes the steric upper limit of the channel and has an effective solvent accessible diameter of ~2.8 Å. Within the selectivity filter, four bound waters are localized along three hydrophilic nodes which punctuate an otherwise extremely hydrophobic pore segment. This novel combination of a long hydrophobic pore and a minimal number of solute binding sites facilitates rapid water transport. Residues of the constriction region, in particular histidine 182 which is conserved among all known water specific channels, are critical in establishing water specificity. Analysis of the AQP1 pore also indicates that the transport of protons through this channel is highly energetically unfavorable.

Water channels facilitate the rapid transport of water across cell membranes in response to osmotic gradients. These channels are believed to be involved in many physiological processes that include renal water conservation, neuro-homeostasis, digestion, regulation of body temperature and reproduction<sup>1,2</sup>. Members of the water channel superfamily have been found in a range of cell types from bacteria to human. In mammals, there are currently 10 families of water channels, referred to as the aquaporins (AQP): AQP0 -AQP9, which can be divided into two major groups<sup>3,4</sup>; AQP0, AQP1, AQP2, AQP4-AQP6 and AQP8, permeable to water but not to small organic and inorganic molecules<sup>5</sup>, and AQP3, AQP7 and AQP9, permeable to glycerol or urea as well as water<sup>3,5</sup>.

AQP1 (M.W. 28 kDa) was initially found in red blood cells and renal proximal tubules<sup>2,6</sup>. AQP1 water channels allow water, but not ions including protons, to freely and bidirectionally move across the cell membrane<sup>5,7</sup>. Sequence analysis shows high homology among members of the AQP1 family and that the two halves of the sequence exhibit a high degree of similarity<sup>4</sup>. Each sequence half contains an NPA (asn, pro, ala) motif which is conserved throughout the AQP superfamily including the glycerol facilitators. Moderate resolution projection and low resolution 3-D maps of AQP1 derived from electron crystallographic studies provided the first structure based evidence for a general architecture consisting of six helices surrounding two putative helical structures within the membrane bilayer<sup>8-11</sup>. Models of AQP1 derived from electron crystallographic structural studies at about 4 Å resolution have recently been reported and independently confirmed the presence of two non-membrane spanning helices<sup>12-14</sup>.

We report here the structure of AQP1 from bovine red blood cells at 2.2 Å resolution as determined by x-ray crystallography. At this resolution the positions of the side chains establishing the properties of the transmembrane channel pathway, as well as water molecules captured in transit, are clear. Critical differences exist between this AQP1 structure and those determined by electron crystallography, most notably between the pore profiles<sup>12-14</sup>. For example, the constriction region in models derived from these previous studies is approximately 8 Å away from this region in the model of AQP1 presented here. The constriction region in our model is also much smaller than that inferred from the high-resolution structure of the *E. coli* glycerol facilitator, GlpF<sup>15</sup>. The high-resolution structure of AQP1 now reveals the structural basis for water specificity; together with the structure of GlpF it also provides the molecular details necessary to understand the mechanisms regulating water and other solute selectivity within the aquaporin superfamily.

### **Architectural Overview**

The functional unit of AQP1 is a tetramer with each monomer providing an independent water pore (Figures 1a,c). Each monomer contains six transmembrane helices packed to form part of a trapezoid-like structure when viewed normal to the membrane plane. Two membrane-inserted but non-membrane-spanning helices, which define a major portion of the pore, are partially enclosed by this structure and form an integral part of the outer wall as well. This general folding topology is similar to that reported for the structure of the *E. coli* glycerol facilitator GlpF<sup>15</sup> and to those reported from electron crystallographic studies of AQP1<sup>12-14</sup>.

Residues comprising the six transmembrane and two membrane inserted non-spanning helices (M1 - M8) are depicted in Figure 1b. The N-terminus of the molecule is located at the cytoplasmic side of the cell membrane, as determined from earlier biochemical studies<sup>16,17</sup>, and leads into the first of two transmembrane helices. These two helices (M1, M2) are followed by a membrane inserted loop that contains an NPA motif and leads into the non-membrane-spanning helix (M3). The loop exiting this very short helix enters the cytoplasm and turns back into the membrane to connect to transmembrane helix M4 that ends the N-terminal half of the structure on the extracellular side of the cell membrane. The second half of the structure is essentially an inverted repeat of the first half topology and results in the location of the C-terminus at the cytoplasmic face. Three nonylglucoside detergent molecules have been located on a region of the monomer surface in contact with the extracellular leaflet of the lipid bilayer. Each monomer is ~40 Å across and ~60 Å long.

### **Structure of the Pore**

The pore of AQP1 is dumbbell-like in shape when viewed in profile (Figures 2a,b), consisting of three general elements, an extracellular vestibule, an extended narrow pore or selectivity filter containing the constriction region, and a cytoplasmic vestibule. Helices M4 and M8 do not contribute residues to the pore. About half of the channel wall along the selectivity filter can be considered hydrophobic and the other half hydrophilic. The hydrophilic face provides the chemical groups essential for displacing certain waters of hydration and therefore establishes a pathway for coordinating water transport.

The steep crossing angles of the monomer helices aid in the formation of the extracellular and cytoplasmic vestibules. Shaping of the vestibule mouths is completed by loop regions at each monomer face and additionally by the N- and C-terminal residues at the cytoplasmic face (Figures 1a,c). The extracellular vestibule is roughly conical in shape with a mouth diameter of about 15 Å (pore size estimates are based on the van der Waals radii used in the AMBER program suite<sup>18</sup>). The population of polar moieties along the surface of the extracellular vestibule consists predominantly of polar residues, only a small number of which are charged, and the solvent exposed backbone from extended loop regions.

Over a distance of about 20 Å, the extracellular vestibule tapers down to its narrowest point, ~2.8 Å in diameter, forming the constriction region and the beginning of the ~20 Å long selectivity filter (Figure 3). A series of solvent-accessible carbonyl oxygens forms a path leading from the extracellular vestibule through the constriction regions of both AQP1 and GlpF; in the case of bovine AQP1 these groups are provided by residues G190, C191, G192 and I193 of the connecting loop leading into non-transmembrane helix M7 (Figure 4). This architectural scheme which positions a helix linker loop so as to form a key element of the selectivity filter is reminiscent of the structural motif employed in the KcsA K<sup>+</sup> channel for the removal of solute hydration waters<sup>19</sup>.

At the constriction region, residues H182 and R197 along with the solvent accessible carbonyl oxygen of residue C191 form the hydrophilic face of the pore (Figure 5). The imidazole ring of H182 is fully extended into the pore while the bulk of R197 is pointed upwards almost parallel to the pore axis in a manner similar to that observed for the equivalent arginine of

the GlpF channel<sup>15</sup>. Opposite the hydrophilic face at the constriction region is the hydrophobic face defined here by residue F58.

Three of the four residues defining the constriction region of the AQP1 pore (R197, H182 and F58) are conserved across the water specific aquaporins<sup>20</sup>. Access to the carbonyl group of the fourth residue, C191, appears critical at that location. The conservation of arginine, histidine and phenylalanine side chains at their respective locations within the constriction region in the known water channels is a strong indicator of channel water specificity and may be useful in assigning function to the sequences of candidate aquaporins for which physiological assays have not been performed.

In GlpF the H182 of AQP1 is replaced by glycine. This substitution provides the additional room needed to accommodate a second substitution, phenylalanine for C191 (Figure 5). These substitutions have two critical effects on the characteristics of the GlpF constriction region; they increase its size and hydrophobicity.

Past the constriction region the pore opens up again averaging about 4 Å in diameter over the next ~15 Å. About 8 Å past the constriction point, residues from the two highly conserved NPA motifs are brought into close proximity due to the end-to-end packing of the two short membrane-inserted non-spanning helices M7 and M3 placing the terminal amine groups of the two NPA motif asparagine residues, N194 and N78 into the pore. The pseudo two-fold axis of the molecule runs parallel to the membrane plane and is located approximately halfway between these two short helices.

As in the case of GlpF, short helices M3 and M7 are situated end-to-end and lengthwise along one side of the channel so that the positive ends of their helical dipoles point inward toward the pore. The aquaporins therefore differ from KcsA in their use of such non-membrane-spanning helices; in KcsA, four such helices are positioned radially around the pore axis and the direction of the dipoles is reversed establishing a net negative zone in the middle of the pore<sup>19</sup>. The two-helix aquaporin motif is found in both water and glycerol selective channels, suggesting that this structure is not used in discriminating between water and glycerol. The examples of KcsA and the aquaporins suggest that pore located non-membrane-spanning helices may serve as functionally critical structural motifs in other classes of transport associated membrane proteins.

A second "string" of carbonyl groups lines the pore this time extending away from the NPA motifs towards the cytoplasmic vestibule and are provided by residues (L77, H76, A75 and G74) of the connecting loop from M2 into the other membrane-inserted non-spanning helix, M3 (Figure 4). The complete set of pore exposed backbone oxygens and asparagine amine groups are arranged in a long-pitched helical pattern forming much of the hydrophilic half of the selectivity filter (Figures 2b, 4). A similar distribution of pore-lining groups was reported for the structure of GlpF<sup>15</sup>.

Near the cytoplasmic end of the selectivity filter is another pore accessible histidine, H76. Unlike H182, this histidine is highly conserved across the aquaporin superfamily and glycerol facilitators. Instead of extending maximally into the pore, as does the constriction defining

H182, H76 is fixed against the side of the pore wall. The GlpF counterpart to H76 is oriented within the pore in a similar fashion.

In the last 8-10 Å of the channel the pore flares out to form a cytoplasmic vestibule with an approximately 15 Å wide mouth. As its walls are somewhat more uniform in height this vestibule is more conical in shape than the extracellular vestibule and its mouth is essentially perpendicular to the pore axis. Here the concentration of polar residues increases significantly from that occurring in the selectivity filter and establishes a zone within the inner half of this vestibule more hydrophilic than that of GlpF (Figure 3).

### **Location of Waters in the Channel**

Waters have been identified at four locations within the AQP1 selectivity filter (Figure 4). Density attributable to a single water molecule is located adjacent to the constriction region about 7 Å from the pseudo two-fold axis toward the extracellular surface. At this location water is coordinated by hydrogen bonds established with the E2 nitrogen of H182 and the backbone carbonyl oxygen of G192. The next two waters are centered about the pseudo two-fold axis, one nearest and hydrogen bonded to the D2 nitrogen of N194 and the other to nitrogen D2 of N78. The fourth water visible in the channel is located near the cytoplasmic end of the selectivity filter. The backbone carbonyl oxygens of residues H76 and A75 form the coordinating hydrogen bonds with this water. These four waters do not form a contiguous hydrogen bonded chain as only the middle two are close enough to each other to form a water-water hydrogen bond. The hydrophobicity profile of the residues lining the pore indicates that there are three hydrophilic

nodes distributed along the length of the selectivity filter (Figure 3). As would be expected, the four waters identified within the selectivity filter are all located at these nodes.

### **Mechanisms of Water Selectivity**

Through a combination of channel sterics and solute binding sites, AQP1 facilitates the rapid and highly selective throughput of water. A steric limit of  $\sim 2.8 \text{ \AA}$  is established at the constriction region while the chemical properties of the residues forming this structure provide additional criteria for solute selection. From the steric limit alone it is now clear why the transport of glycerol by AQP1 is highly unfavorable. Even for a molecule of water to diffuse across the narrow AQP1 constriction region its effective diameter must be reduced by shedding waters of hydration. For this to be an energetically favorable process interactions with primary hydration shell waters must be replaced by interactions with residues lining the channel wall over as small a diffusion distance as possible. In AQP1, sufficient hydrogen bond forming groups are available so that water molecules can readily move through the constriction region. These bond forming groups are provided by constriction region residues H182 (conserved across the water specific aquaporins) and R197 (conserved throughout most of the aquaporin superfamily), and the backbone carbonyl oxygens of residues G190, C191 and G192.

In addition to the bound water found at the constriction region, there are three additional waters in the selectivity filter. Averaging only about  $4 \text{ \AA}$  in diameter, the selectivity filter is also rather hydrophobic but is punctuated by water binding regions at several hydrophilic nodes (Figure 3). The availability of water binding sites at these nodes reduces the energy barrier to water transport across this predominantly hydrophobic pathway while the relatively low number

of such sites keeps the degree of solute-pore interaction to a minimum. In balancing these opposing factors the aquaporins are able to transport water selectively while optimizing permeability.

While the aquaporins are optimized for the rapid transport of water, some members of the family facilitate the transport of other solutes. The bacterial aquaglyceroporin homologue, GlpF, is optimized for the rapid transport of glycerol. Although water has favorable steric accessibility through both the AQP1 and GlpF channels there is a critical difference in glycerol steric accessibility between these two pores; the constriction region of GlpF is almost 1 Å wider than in AQP1. The dramatic difference in glycerol permeability between AQP1 and GlpF is effectively bridged in aquaglyceroporin AQP3, which transports both water and glycerol at moderate rates<sup>21</sup>. The hydrophathy profiles for the selectivity filter regions of AQP1, AQP3 and GlpF are quite similar. Most of the amino acid differences here are relatively moderate such as swapping one type of hydrophobic residue for another, yet changes to only one or two residues of the constriction regions in these channels are sufficient to radically alter solute selectivity. It appears that in AQP1 the residue most critical in supporting rapid water throughput, while hindering the passage of glycerol, is H182 (Figure 5). Throughout the aquaporin family the choice of amino acid at this location appears to be essential in defining whether an aquaporin will be specific for water or additionally selective for other solutes such as glycerol. In GlpF the replacement of this histidine by glycine serves to both significantly increase the size of the constriction region and sterically support a second residue change, C191 to phenylalanine. These changes in turn alter the polar nature of this area and result in improved channel interactions with the hydrophobic backbone of glycerol. H182 of AQP1 is likely also replaced

by glycine in AQP3. As with GlpF, such a substitution should provide the additional room needed to accommodate the side chain of a second probable substitution, in this case a tyrosine for C191. In contrast to the substitution for phenylalanine occurring at this position in GlpF, the switch to tyrosine potentially provides an additional location for solute hydrogen bonding. As a consequence of these two substitutions (H182G, C191Y) it is expected that the constriction region of AQP3 is similar in size to that of GlpF, although somewhat more polar, thereby allowing for the moderately rapid transport of glycerol as well as water.

While the amphipathic nature of the selectivity filter plays a key role in the rapid transport of water in AQP1, this property is preserved across the superfamily of aquaporins and has been shown to be particularly beneficial in the case of glycerol transport. The hydrophobic face of the GlpF selectivity filter provides an ideal match for the carbon backbone side of glycerol while the opposing hydrophilic face is equipped with hydrogen bonding groups to replace waters of hydration in the vicinity of glycerol hydroxyl groups<sup>15</sup>. Both glycerol and water molecules were visualized within the GlpF selectivity filter indicating that at high concentrations of glycerol there should be almost one-to-one transport of glycerol and water. Although GlpF permeability to water is significant, it is substantially less than that measured for its bacterial aquaporin counterpart and AQP1 homologue, AqpZ<sup>22</sup>. Since the constriction region of GlpF is wider than that of AQP1 it would appear that the more hydrophobic nature of GlpF is the primary factor responsible for its relatively reduced ability to transport water.

Preselection of solutes by the vestibules may also be an important factor in maximizing rates of transport. The relatively stronger hydrophobic nature of the vestibules in GlpF should

improve the transport rates of solutes that contain a significant hydrophobic component, such as glycerol, although potentially at the expense of rapid water throughput. Conversely the more hydrophilic vestibules of AQP1 favor the preselection of water.

### **Barriers to Ion Transport**

In the aquaporin and KcsA K<sup>+</sup> channel selectivity filters, carbonyl groups extending from inner helix linker loops face the pore staggered along the channel axis. In AQP1, and as reported in the structure of GlpF<sup>15</sup>, these pore accessible carbonyl oxygens (ten) are distributed along a ~25 Å stretch along one side of the pore. For KcsA this zone is significantly shorter, confined to one-half of the transmembrane pore and yet contains 16 carbonyl oxygens arranged in a stacked ring configuration, four carbonyls per ring<sup>19</sup>; hydrated K<sup>+</sup> ions are closely and symmetrically surrounded by carbonyl oxygens so that making coordination site transfers between hydrating water and the carbonyl oxygens is energetically favorable. Hydrated ions encountering the type of selectivity filter utilized by the aquaporins would find that the available carbonyl groups and water coordinating residues, such as R197 and H182 in AQP1, could only effectively substitute for less than half of the hydrating waters at any point along this segment of the channel. Such partially hydrated ions would still be too large to pass through the constriction region or most of the selectivity filter. Positively charged ions would find additional resistance to their transport from repulsive forces generated through interactions with the dipoles of helices M3 and M7 (oriented with their positive ends pointed into the pore), the highly conserved R197 (located near the constriction region), the amines of N78 and N194 (located about the pseudo two-fold axis) and by the histidines H182 and H76 located at opposite ends of the selectivity filter. Negatively

charged ions would experience repulsive forces from the many pore accessible carbonyl oxygens lining the selectivity filter and inner regions of the vestibules.

Water networks or chains have been seen in the high-resolution structures of proteins such as the photosynthetic reaction center<sup>23</sup> and cytochrome  $b_6f$ <sup>24</sup> and it has been suggested that protons could be transported along a continuous linear network of suitably oriented hydrogen-bonded waters (“proton-wire”) by a transfer process termed the Grotthuss mechanism<sup>25,26</sup>. Under this mechanism a proton can be “instantaneously” shuttled along this special network of water molecules. However, no suitable chain of hydrogen-bonded water molecules spanning the selectivity filter has been located in the density map of the AQP1 crystal structure. Factors preventing the establishment of such a hydrogen-bonded water chain along the string of carbonyl oxygens lining the AQP1 channel include the following highly conserved elements: arginine located at the extracellular side of the selectivity filter, the two asparagines and helical dipole end charges located in the pore about the pseudo two-fold axis as suggested recently by others<sup>13,14</sup> and the histidines located at the ends of the selectivity filter. The two positive helical dipole end charges are by themselves sufficient to disrupt any suitably oriented alignment of water dipoles along the channel pore.

## Methods

### Crystallization of AQP1

Native crystals were obtained as described previously<sup>27</sup>. Soaking of crystals in a variety of heavy-atom compounds greatly affected diffraction quality even for very low concentrations of heavy atom compounds. Co-crystallization yielded thallium derivatized crystals suitable for providing initial phasing information.

### Data Collection and Processing

MAD (four wavelength) data sets were collected from Tl-AQP1 crystals at the Advanced Light Source (ALS) of Lawrence Berkeley National Laboratory. Diffraction data were processed with DENZO/SCALEPACK<sup>28</sup>. Crystals belong to space group I422 ( $a = b = 93.3$ ,  $c = 180.5$  Å,  $\alpha = \beta = \gamma = 90^\circ$ ) and contain one molecule per asymmetric unit. The best crystals diffracted to 2.1 Å. A Patterson map based on dispersive and anomalous differences showed strong heavy atom peaks located near the crystal's four-fold symmetry axis. These positions were refined with the use of the SHARP program<sup>29</sup>. The refined heavy atom positions were used to obtain an initial density map. An improved density map was produced with the application of solvent flattening through the program DM<sup>30</sup> and was used to build an initial model at 2.8 Å resolution.

### Model Building

Model building was performed using O<sup>31</sup> while model refinement was conducted using *Crystallographic and NMR system (CNS)*<sup>32</sup>. Extension of phases and model refinement to 2.2 Å was accomplished with a data set obtained from crystals grown in the presence of gold cyanide but for which no multi-wavelength anomalous differences were observed. Others have also obtained improvements in crystal order using gold cyanide as a crystallization additive<sup>33</sup>. While the diffraction data from these crystals extended to 2.1 Å it was significantly anisotropic and resulted in a reduction in the percentage of statistically reliable data in the direction of the h, k plane. CNS based refinement of the model (residues 1-249 of a possible 271) using data to 2.2 Å resolution resulted in an improved model with an  $R_{\text{free}}$  value of 30.8% and a crystallographic R of 26.6%. The 2Fo-Fc electron density map obtained from these data however is of good quality (Figure 1d). The relatively high values of R and  $R_{\text{free}}$  are a consequence of the diffraction set anisotropy; such an effect was expected and larger R factors have been reported for similar resolution anisotropic data sets which also yielded good quality maps as reported by Huber and colleagues<sup>34</sup>.

Table 1.

Data Set	Resolution (Å)	Complete (%)	R <sub>sym</sub> (%)	I / σI
Au-AQP1	2.2 (2.30 - 2.20)	99.7 (100.0)	6.2 (58.2% *)	41.8 (5.3)
TI-AQP1:				
λ1 (0.97871)	2.8 (2.85 - 2.80)	99.2 (99.8)	5.5 (41.8%)	50.0 (9.7)
λ2 (0.97954)	2.8 (2.85 - 2.80)	99.2 (100.0)	5.0 (29.2%)	53.5 (13.8)
λ3 (1.00800)	2.8 (2.85 - 2.80)	99.4 (100.0)	5.1 (30.2%)	54.5 (13.6)
λ4 (0.95593)	2.8 (2.85 - 2.80)	99.2 (99.8)	6.6 (47.9%)	45.8 (8.3)

## Refinement Statistics

Resolution (Å)	15 - 2.2	Nonglycine residues in most favorable	
Reflections in working/test set	16934/1235	region of Ramachandran plot (%)	97.2
R (%)	26.6	Outliers and generously allowed regions (%)	2.8
R <sub>free</sub> (%)	30.8		

## RMS deviation from ideality

Bond lengths (Å)	0.0066
Bond angles (degree)	1.25

R<sub>sym</sub>=

$\frac{\sum |I - \langle I \rangle|}{\sum \langle I \rangle}$ , where I is the measurement of intensity.  $R = \frac{\sum |F_{obs} - F_{calc}|}{\sum F_{obs}}$  for all values and R<sub>free</sub> is the R value for the reflections that were excluded in refinement, where F<sub>obs</sub> is the observed native amplitude and F<sub>calc</sub> is the one calculated from atomic models. The values for the highest resolution zone are shown in parenthesis. \* This large R<sub>sym</sub> value is the consequence of significant anisotropy. When the data was filtered at a 1σ level, the percentage of reflections removed from the quadrants of maximum anisotropy was proportionally the highest and the following values were obtained: R<sub>sym</sub>=34.9%, completeness=78.0%, and I / σI = 6.7.

## References

1. Verkman, A.S. Water channels in cell membranes. *Ann. Rev. Physiol.* **54**, 97-108 (1992).
2. Agre, P. *et al.* Aquaporin CHIP: the archetypal molecular water channel. *Am. J. Physiol.* **265**, F463-F476 (1993).
3. Ishibashi, K. *et al.* Cloning and functional expression of a new aquaporin (AQP9) abundantly expressed in the peripheral leukocytes permeable to water and urea, but not to glycerol. *Biochem. Biophys. Res. Comm.* **244**, 268-274 (1998).
4. Borgnia, M. *et al.* Cellular and Molecular Biology of the aquaporin water channels. *Ann. Rev. Biochem.* **68**, 425-458 (1999).
5. Meinild, A.K. *et al.* Bidirectional water fluxes and specificity for small hydrophilic molecules in aquaporins 0-5. *J. Biol. Chem.* **273**, 32446-32451 (1998).
6. Denker, B.M. *et al.* Identification, purification, and partial characterization of a novel Mr 28,000 integral membrane protein from erythrocytes and renal tubules. *J. Biol. Chem.* **263**, 15634-15642 (1988).
7. Zeidel, M.L. *et al.* Ultrastructure, pharmacologic inhibition, and transport selectivity of aquaporin channel-forming integral protein in proteoliposomes. *Biochemistry* **33**, 1606-1615 (1994).
8. Jap, B.K. and Li, H.-L. Structure of Osmo-regulated H<sub>2</sub>O Channel, AQP-CHIP, in Projection at 3.5Å Resolution. *J. Mol. Biol.* **251**, 413-420 (1995).
9. Li, H.-L., Lee, S. and Jap, B.K. Molecular design of aquaporin-1 water channel as revealed by electron crystallography. *Nature Struct. Biol.* **4**, 263-265 (1997).
10. Cheng, A. *et al.* Three-dimensional organization of a human water channel. *Nature* **387**, 627-630 (1997).
11. Walz, T. *et al.* The three-dimensional structure of aquaporin-1. *Nature* **387**, 624-627 (1997).
12. Ren, R. *et al.* Three-dimensional fold of the human AQP1 water channel determined at 4 Å resolution by electron crystallography of two-dimensional crystals embedded in ice. *J. Mol. Biol.*, **301**, 369-387 (2000).
13. Murata, K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599-605 (2000).
14. Ren, R. *et al.* Visualization of a water-selective pore by electron crystallography in vitreous ice. *Proc. Natl. Acad. Sci.* **98**, 1398-1403 (2001).
15. Fu, D. *et al.* Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**, 481-486 (2000).
16. Smith, B.L. and Agre, P. Erythrocyte Mr 28,000 transmembrane protein exists as a multisubunit oligomer similar to channel proteins. *J. Biol. Chem.* **266**, 6407-6415 (1991).

17. Nielsen, S. *et al.* Distribution of the aquaporin CHIP in secretory and resorptive epithelia and capillary endothelia. *Proc. Natl. Acad. Sci. USA* **90**, 7275-7279 (1993).
18. Weiner, S.J. *et al.* A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765-784 (1984).
19. Doyle, D.A. *et al.* The structure of the potassium channel: molecular basis of K<sup>+</sup> conduction and selectivity. *Science* **280**, 69-77 (1998).
20. Park, J.H. and Saier, M.H. Jr. Phylogenetic characterization of the MIP family of transmembrane channel proteins. *J. Membr. Biol.* **153**, 171-180 (1996).
21. Zeuthen, T. and Klaerke, D.A. Transport of water and glycerol in aquaporin 3 is gated by H<sup>+</sup>. *J. Biol. Chem.* **274**, 21631-21636 (1999).
22. Borgnia, M.J. and Agre, P. Reconstitution and functional comparison of purified GlpF and AqpZ, the glycerol and water channels from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **98**, 2888-2893 (2001).
23. Baciou, L. and Michel, H. Interruption of the water chain in the reaction center from *Rhodobacter sphaeroides* reduces the rates of the proton uptake and of the second electron transfer to QB. *Biochemistry* **34**, 7967-7972 (1995).
24. Ponamarev, M.V. and Cramer, W.A. Perturbation of the internal water chain in cytochrome f of oxygenic photosynthesis: loss of the concerted reduction of cytochromes f and b<sub>6</sub>. *Biochemistry* **37**, 17199-17208 (1998).
25. N. Agmon. The Grotthuss mechanism. *Chem. Phys. Lett.* **224**, 456-462 (1995).
26. Pomés, R. and Roux, B. Free energy profiles for H<sup>+</sup> conduction along hydrogen-bonded chains of water molecules. *Biophys. J.* **75**, 33-40 (1998).
27. Sui, H. *et al.* Crystallization and preliminary X-ray crystallographic analysis of water channel AQP1. *Acta Crystallogr. D* **56**, 1198-1200 (2000).
28. Otwinoski, Z. and Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326 (1997).
29. de La Fortelle, E. *et al.* In *Crystallographic Computing 7*, edited by P. Bourne and K. Watenpaugh. Oxford: IUCr/Oxford University Press (1997).
30. Cowtan, K.D. *Joint CCP4 ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34 (1994).
31. Jones, T.A. *et al.* Improved methods for building protein models in electron density maps. *Acta Crystallogr. A* **47**, 110-119 (1991).
32. Brünger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905-921 (1998).
33. Chang, G. *et al.* Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel. *Science* **282**, 2220-2226 (1998).
34. Groll, M. *et al.*, The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study. *Proc. Natl. Acad. Sci. USA* **96**, 10976-10983 (1999).

35. Thompson, J.D. *et al.* CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680 (1994).
36. Kraulis, P.J. Molscript - a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950 (1991).
37. Merritt, E.A. and Bacon, D.J. Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524 (1997).
38. Smart, O.S. *et al.* The pore dimensions of gramicidin A. *Biophys. J.* **65**, 2455-2460 (1993).
39. Kyte, J. and Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132 (1982).

## Acknowledgments

This work is supported by funding from the National Institutes of Health and by Office of Health and Environmental Research, US Department of Energy. Preliminary data collection and screening for heavy atom derivatives were conducted at beamlines X25 at National Synchrotron Light Source and 1-5 at Stanford Synchrotron Radiation Laboratory; data sets used in determining the model were collected at beamline 5.0.2, Advanced Light Source, Lawrence Berkeley National Laboratory. We would like to thank all staff members of these beamlines for their valuable assistance and Dr. B.-C. Wang for helpful discussions on heavy atom position refinement. The coordinates have been deposited in the Protein Data Bank under accession number 1J4N.

+ Indicates equal contribution.

++ To whom correspondence should be addressed. E-mail: [BKJap@lbl.gov](mailto:BKJap@lbl.gov) or [PJWalian@lbl.gov](mailto:PJWalian@lbl.gov)

## FIGURE CAPTIONS:

Figure 1. Models of AQP1, sequence alignment of selected superfamily members and a view of the density map. (a) Combined ribbon diagram and space filling model of AQP1 monomer viewed parallel to the membrane (b) Amino acid sequence alignment of bovine AQP1, human AQP1, *E. coli* AqpZ, human AQP3 and *E. coli* GlpF produced using ClustalW<sup>35</sup>. AQP1 membrane-embedded helices are colored coded as shown in (a) and labeled M1 - M8. M3 and M7 are membrane inserted non-spanning helices. Residues lining the extended narrow pore (selectivity filter) of bovine AQP1 are indicated in red italicized letters. (c) AQP1 tetramer viewed normal to the membrane;

two of the monomers also contain a space-filling model to show the pore entrance. The black arrow points along the direction of the cut-line used to produce the monomer view in Figure 2b. (d) Stereo view of the 2Fo-Fc electron density map ( $1.5\sigma$  contour) and corresponding region of the AQP1 model. View is from the extracellular side, down the pore, and centered about the constriction region. The spherical densities near the pore center are those of the two waters located in the extracellular half of the selectivity filter. Figures produced using MOLSCRIPT<sup>36</sup> and Raster3D<sup>37</sup>.

Figure 2. Side-views of AQP1. (a) Backbone in ribbon format, residues depicted in ball-and-stick representation; residues lining the pore are shown in CPK color format. The pore profile is highlighted by an array of cyan colored dots generated by the program HOLE<sup>38</sup>. The constriction region is visible as the pinched-in area in the extracellular half of the profile. (b) Side-view cut-away of a space-filling model made along the axis indicated by the black arrow in Figure 1c. Residues not in direct contact with the pore are colored dark gray. The pore lining in the top figure consists primarily of polar groups distributed along the length of the selectivity filter while the complementary half is predominantly hydrophobic in nature. The constriction region and pseudo two-fold axis (centered about the NPA motifs) are indicated by cyan and black arrows respectively. The pore profile (cyan colored dots) is superimposed on each channel half. Figure produced using MOLSCRIPT and Raster3D.

Figure 3: The effective pore diameter and hydrophobicity of the AQP1 and GlpF channels.

Green and blue arrows indicate the locations of the constriction region and the pseudo two-fold axis respectively. Below these arrows a cyan colored bar indicates the location of the extracellular vestibule, a red bar the selectivity filter and a green bar the cytoplasmic vestibule. The three black bars in the hydrophobicity profile identify hydrophilic nodes within the selectivity filter. Pore diameters were determined utilizing AMBER based van der Waals radii<sup>18</sup> and analyzed using the program HOLE<sup>38</sup>. Hydrophobicity was characterized using the Kyte and Doolittle amino acid hydrophobicity scale<sup>39</sup> and a four point averaging window across the pore lining residues.

Figure 4. Selectivity filter waters and residues forming the hydrophilic face of the channel pore. Cut-away side view of the channel with secondary structure shown in ribbon format. Side chains critical to establishing the long-pitched helical, hydrophilic path across the length of the selectivity filter are shown. The four waters located within the selectivity filter are depicted as green translucent spheres. Of these four waters only the middle two are close enough to form a water-water hydrogen bond. The constriction region and pseudo two-fold axis (centered about the NPA motifs) are indicated by cyan and black arrows respectively. Figure produced using MOLSCRIPT and Raster3D.

Figure 5 Residues defining the constriction region. Residues involved in the formation of the AQP1 constriction region (H182, R197, F58 and C191) are depicted in conventional CPK style while the translucent side chains are those of the equivalent residues found in the structure of GlpF (G191, R206, W48 and F200). These side chain differences result in a larger and more hydrophobic constriction region in GlpF. Alignment of the residues was performed using a least square fit of their backbone atoms. Figure produced using MOLSCRIPT and Raster3D.