Aquaporins comprise a family of water-transporting membrane proteins. All aquaporins are efficient water transporters, while sustaining strict selectivity, even against protons, thereby maintaining the proton gradient across the cell membrane. Recently solved structures of these membrane channels have helped us to understand this remarkable property. The structure of the Escherichia coli glycerol facilitator GlpF at 2.2 Å resolution has enabled the refinement of a low-resolution human aquaporin-1 structure. This latter structure has recently been confirmed by the 2.2 Å structure of bovine aquaporin-1. Further insights, particularly with respect to the dynamics of water permeation and the filter mechanism, have come from recent molecular dynamics simulations.

Table 1

<table>
<thead>
<tr>
<th>Pair</th>
<th>Backbone helix</th>
<th>Backbone all</th>
<th>All atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1FQY - 1FX8</td>
<td>1.75</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1FQY - 1H6I</td>
<td>1.41</td>
<td>3.22</td>
<td>4.29</td>
</tr>
<tr>
<td>1FQY - 1J4N</td>
<td>1.46</td>
<td>3.17</td>
<td>-</td>
</tr>
<tr>
<td>1H6I - 1FX8</td>
<td>1.36</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1H6I - 1J4N</td>
<td>0.90</td>
<td>2.44</td>
<td>-</td>
</tr>
<tr>
<td>1FX8 - 1J4N</td>
<td>1.50</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1FQY - 1H5</td>
<td>2.45</td>
<td>3.31</td>
<td>4.47</td>
</tr>
<tr>
<td>1H6I - 1H5</td>
<td>2.26</td>
<td>3.29</td>
<td>4.56</td>
</tr>
<tr>
<td>1FX8 - 1H5</td>
<td>2.67</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1J4N - 1H5</td>
<td>2.34</td>
<td>3.40</td>
<td>-</td>
</tr>
</tbody>
</table>

*PDB codes: 1FQY, EM structure of human AQP1 by Murata et al. [3••]; 1FX8, X-ray structure of E. coli GlpF by Fu et al. [4••]; 1H6I, refined EM structure of human AQP1 by de Groot et al. [5•]; 1H5, EM structure of human AQP1 by Ren et al. [7]; 1J4N, X-ray structure of bovine AQP1 by Su et al. [6••].

The atomic model of AQP1 derived from a 3.8 Å resolution potential map obtained by electron crystallography gave the first insight into water specificity and proton blockage [3••]. This was the first atomic structure of a human membrane protein to be solved. It revealed remarkable structural determinants related to the highly conserved sequence motif Asn–Pro–Ala (NPA) and to the unique aquaporin fold. An important member of the aquaglyceroporin family, the *Escherichia coli* glycerol facilitator GlpF, was analyzed by X-ray crystallography at 2.2 Å resolution, providing a deep insight into the amazing specificity of this channel for linear carbohydrates [4••]. These two independent structures reveal essentially the same fold (Table 1). This enabled the refinement of the lower resolution human AQP1 structure, starting from a homology model based on the sequence-related GlpF structure [5•]. Later, the structure of bovine AQP1 was analyzed by X-ray crystallography at 2.2 Å [6••]. It confirmed the refined structure of human AQP1 (Table 1) and allowed several water molecules to be identified. Another structural model of human AQP1 derived by electron crystallography (PDB code 1H5; [7]) differed significantly from all other AQP1 and GlpF structures for the backbone helices (Table 1). Further insights, particularly into the dynamics of water permeation and into the filter mechanism, came from recent MD simulations [8,9••,10,11••].
Figure 1

AQP1 structure. (a) The AQP1 tetramer, with each monomer rendered in a different representation (clockwise from bottom left: molecular surface, space-filling representation, cartoon and ball-and-stick model). (b) EM potential map at 3.8 Å resolution, showing helices H2 and H5, and highlighting the conserved glycines G57 and G173. (c) The monomeric structure of AQP1, showing the aquaporin fold, helix assignment, and the location of the membrane and the long, extracellular C loop (LC). (d,e) The central pore region (top view), where the NPA motifs contact each other, shown in (d) ball-and-stick and (e) space-filling representations (PDB code 1H6I).
Figure 2

Comparison of (a) the refined cryo-EM structure of human AQP1 (PDB code 1H6I) and (b) the X-ray structure of bovine AQP1 (PDB code 1J4N) in the pore region. The left and right hand figures show the same structure, one rotated with respect to the other. The NPA motifs are highlighted as yellow sticks and the Ar/R filter region (see text) is highlighted in purple. The amino acids in contact with the pore are color coded according to their properties: yellow for aliphatic, green for aromatic, orange for backbone carbonyl and cyan for polar. (c) View onto the NPA motif, showing the stacked prolines and the stabilizing hydrogen bonds. According to the refined EM model 1H6I, the stacked prolines (in green) form the platform from which the short helices HB and HE emanate.
Structure of aquaporin-1

Medium-resolution potential maps of AQP1 obtained by electron crystallography revealed a right-handed bundle of six highly tilted α helices that embrace a central density [12,13]. The large tilt of about 30° with respect to the membrane normal allows the formation of an
hourglass-shaped [14] pore with only six transmembrane helices. Because AQP1 was the first membrane protein to exhibit such a fold, the right-handed arrangement of $\alpha$-helices was initially controversial [12,15]. However, other examples of right-handed membrane channels composed of strongly tilted $\alpha$-helices soon emerged (e.g. the KcsA $K^+$ channel [16], the mechanosensitive MscL channel [17] and a CIC-type Cl$^-$ channel [18]). Subsequently, the right handedness of AQP1 was directly confirmed at a resolution of 3.8 Å, which allowed an atomic model to be constructed [3**] (Figure 1).

AQP1 forms tetramers (Figure 1a) that survive solubilization by detergents, as documented by EM of negatively stained preparations, ultracentrifugation and mass measurements by scanning transmission electron microscopy [19,20]. The aquaporin fold, with its twofold pseudo-symmetry, reflects the tandem repeat in the sequences of all aquaglyceroporins [21] (Figure 1c). The long loop C (LC), which spans the entire molecule at the extracellular side, connects the two tightly intercalated halves (Figure 1c). Many helix–helix interactions stabilize the monomer; one example is the precise fit at the highly conserved glycine residues (G57 and G173 within helices H2 and H5, respectively; Figure 1b). Loops LB and LE, each containing an NPA motif and a short helix (HB and HE), bend into the six-helix bundle to form the channel (Figure 1c). The dipoles of these helices create a functionally important electrostatic field at the membrane center [3**,9**,11**], where the conserved prolines of the NPA motifs are tightly stacked. Unexpectedly, most residues within the central pore are hydrophobic, whereas only a few hydrophilic residues and backbone carbonyls are exposed at strategic sites.

Structure validation statistics [5*] and analysis of structural integrity during MD simulations [5*,10] indicated, however, that a few structural features might require further refinement. AQP1 and GlpF share an overall sequence identity of 30.6% and the 3.8 Å structure of human AQP1 (PDB code 1FQY) fits the X-ray structure of E. coli GlpF (PDB code 1FX8; 2.2 Å) with a root mean square deviation (rmsd) of 1.75 Å for the helical backbone atoms (Table 1). This provided a solid basis for refining the initial human AQP1 model against the 3.8 Å potential map starting from a homology model based on the GlpF structure, yielding the model 1H6I [5*]. The improvement is reflected by the R-factor of 36.7% (39.9% for 1FQY) and the free R-factor of 37.8% (41.7% for 1FQY), but, more importantly, by the stability of the structure during MD simulations. This refined model has recently been confirmed by the 2.2 Å X-ray structure of bovine AQP1 [6**] (compare Figure 2a,b), which fits the structure of the model 1H6I with a rmsd of less than 1 Å for the helical backbone (Table 1).

In the refined EM model of AQP1 (PDB code 1H6I), two highly conserved residues have a different conformation compared to the initial EM model (PDB code 1FQY). First, the new orientation of Trp210 fits more favorably into the potential map in the refined model. This finding is consistent with preferential location at the tryptophan residues at the border between the hydrophobic and hydrophilic zones, a feature observed frequently in membrane proteins [22]. Second, Arg195 deserves particular attention; in GlpF, this residue faces the pore and forms two hydrogen bonds with a bound glycerol molecule, whereas in the 1FQY model, Arg195 was assumed to form a salt bridge with Glu142 because of the strong conservation of these residues [21]. However, the fit of the model to the EM potential map improved when the sidechain of Arg195 adopted an orientation similar to that in GlpF (Figure 2a). This residue is critical, as it provides a strong positive charge at the narrowest region of the pore. Together with the sidechains of His180 and Phe56, it forms a 2 Å wide Ar/R (aromatic/arginine) constriction. A second narrowing of the pore, to less than 3 Å in diameter, occurs at the middle of the membrane, where the NPA motifs are located. Interestingly, the majority of aquaporins exhibit a Phe–Leu pair at this narrowing (Phe24–Leu149 in AQP1, Figure 1d,e), whereas almost all aquaglyceroporins have a Leu–Leu pair [21], which apparently ensures a sufficiently large pore diameter to allow the passage of small solutes.

The NPA motifs with the stacked prolines constitute the platform from which the two short helices HB and HE

---

**Figure 3 legend**

Multinanoscond MD simulation of water permeation through AQP1 [3**]; (a) The simulation system of the protein tetramer (orange/magenta/blue/cyan) embedded in a POPC (palmitoyloleyl-phosphatidylethanolamine) bilayer (yellow/green) surrounded by water (red/white) consisted of approximately 100 000 atoms. (b) The pathway of one of the spontaneous, full permeation events that were observed through the AQP1 pores is depicted. (c) Water molecules were found to be strongly oriented in the channel interior, with their dipoles rotating by about 180° during permeation. The water dipoles (oriented from the oxygen to the midpoint between the two hydrogens) are depicted as yellow (strong orientation) and green (weak orientation) arrows. (d) Hydrogen bond statistics for water molecules show that there are two major interaction sites inside the channel (green line); the NPA and Ar/R regions. Water–water hydrogen bonds, a prerequisite to proton conduction, are weakened most in the Ar/R region (blue and red lines). The black line depicts the total average hydrogen bond energy of a water molecule as a function of the pore axis. (e) The two-stage (NPA and Ar/R) filter, together with the main functional building blocks, as derived from the simulations, is summarized schematically. The channel acts like a two-stage filter: the Ar/R region and the NPA region. The Ar/R constriction region is the narrowest region of the pore. Here, water–water interactions are distorted most with respect to bulk (light blue compared to dark blue). Water dipoles (red arrows) are strongly oriented in the pore with a bipolar distribution through interactions with the NPA residues and the HB and HE helices, causing water molecules to rotate during passage. The midpoint of the rotation of the water dipoles is located in the NPA region. Both stages of the filter add to the remarkable efficiency and selectivity of the aquaporins.
emanate. Importantly, the carbonyl groups of the two asparagine residues in the NPA motifs hydrogen bond to the mainchain NH groups, thereby capping the N-terminal end of the short HB and HE helices (Figure 2c). As a result, the amido groups of Asn76 and Asn192 are fixed to extend into the pore (Figure 3b). This arrangement is further stabilized by several hydrogen bonds contributed mainly from the highly conserved Glu17, Gln101 and Glu142 (Figure 2c).

**Computer simulations of solute permeation through aquaglyceroporins**

Despite its enormous capacity for water conductance, the AQP1 pore also exhibits marked selectivity. Complementing the wealth of information on the mechanism obtained from the atomic structures of AQP1 and GlpF, major aspects of the dynamics involved in actual permeation events have been revealed by MD simulations. Spontaneous partial glycerol permeation events, involving all three glycerol molecules per channel, were observed in simulations of GlpF [8]. This is remarkable because, based on the experimentally determined glycerol permeation rate of ~1 molecule per microsecond [23], spontaneous permeation would not be expected on the simulation timescale of 1 ns. However, the high computational permeation rate might have been induced by the high glycerol concentration that caused three glycerol molecules to be bound to the channel. Glycerol molecules closest to the extracellular face of the protein were found to bind weakest, followed by those bound to the NPA region. The glycerol molecules bound in between, in the narrowest part of the channel, were found to bind strongest. The simulations further showed that glycerol and water compete for hydrogen bonds in the channel interior in a random manner, explaining the observed co-translation of glycerol and water. No evidence for cooperativity between channels was found. Subsequent steered MD simulations allowed the computation of a free energy profile for glycerol motion along the channel [24•].

MD simulations with a timescale of 10 ns [9••], from both the refined AQP1 structure [5*] and the GlpF X-ray structure [4••], showed spontaneous, bi-directional, full water permeation events (Figure 3a,b), with rates in close agreement to the experimentally determined rates. In both simulations, water molecules were found to be strongly oriented in the channel interior, through the alignment of their dipoles with the electric field exerted by the protein, causing water molecules to rotate by 180° upon passage (Figure 3c). Hydrogen bond competition between water molecules and the few polar groups in the pore was found to dominate the permeation process. Two major interaction sites for water molecules were identified inside the channel (Figure 3d, green line): the NPA and Ar/R constriction regions. The two highest enthalpic barriers to the passage of water molecules are located directly adjacent to the NPA region (Figure 3d, black line). This, together with the water rotation that also occurs here, suggested that the NPA region is a major selectivity filter.

Aquaglyceroporins must strictly prevent proton conduction along their pores to maintain the proton gradient across the cell membrane that serves as a major energy storage mechanism. Recently, MD simulations of the glycerol-free GlpF structure confirmed the functional role of the NPA motifs and the dipoles of helices HB and HE [11••]. These simulations report also a bipolar configuration of the water molecules inside the channel. Interestingly, this bipolar configuration of water molecules vanished when the electrostatic interactions between water and the NPA motifs on one hand and the helical dipoles of HB and HE on the other hand were artificially switched off, thus demonstrating their relevance. Furthermore, in the glycerol-free structure, significant water occupancy is seen within the NPA region. In the simulations, frequent simultaneous hydrogen bonding of water molecules to the two NPA asparagines has been observed, thereby weakening interactions among adjacent water molecules in the pore. As contiguous hydrogen bonded water chains are known to be efficient proton conductors, it was suggested that this region is the main proton filter [11••]. In a quantitative analysis of the disruption frequency of the hydrogen bonded water chain along the pore [9••], the proton-conducting chain was also found to be interrupted within the NPA region, but was more severely interrupted within the Ar/R constriction region, where water–water hydrogen bonds were found to be weakest (Figure 3d, red and blue lines). This, together with the positive charge on Arg195, suggested this region to be a major filter for protons. The question of which of the two regions dominates remains open. Figure 3e summarizes these findings.

MD simulations started from the original models of the AQP1 structure [3••,7] were found to behave in a less stable manner, resulting in either the degradation of the channel structure near the NPA region [10] or the formation of multiple water pathways through the protein matrix [25]. These simulations reconfirm the importance of the NPA region not only for the structural integrity of the channel, but also for water conductance: if the asparagines of the NPA motifs are mutated to leucine, water files were found to be interrupted and water flow to be stopped [25].

**Conclusions**

Water regulation is crucially important for every cell and, therefore, for all life forms on earth. Water channels have been identified in almost every living organism — from plants to animals, from prokaryotes to eukaryotes, including humans. Evolutionary processes have developed numerous very efficient water channels [26] based on a unique ancient aquaporin fold. Intramolecular as well as intermolecular helix–helix interactions, together with critical hydrogen bridges, stabilize this unusual fold. Unexpected structural features, such as the right-handed helical bundle and the mostly hydrophobic pore, were revealed by electron crystallography.

Homology modeling of AQP1 using the 2.2 Å structure of GlpF allowed the model based on electron crystallography
to be refined, yielding a structure that has recently been confirmed by the 2.2 Å structure of bovine AQP1. This demonstrated the value of electron crystallography, which has recently been debated. Complementing experimental data, MD simulations by several groups have provided reproducible insights into the dynamics of solute permeation, as well as into the underlying molecular mechanisms, which in turn can be validated experimentally. Structural and dynamic information on the atomic scale is a prerequisite to understanding the function of a channel and this information could become the basis for designing novel therapeutics for the many diseases related to disturbed water balance.

Acknowledgements

This work was supported by the Swiss National Foundation, the ME Müller Foundation of Switzerland, the European Union Quality of Life and Management of Living Resources Project (grants QLRT-2000/00778 and QLRT-2000/00504 to A Engel) and the Human Frontier Science Program (grant RG0021/2000-M103 to A Engel). BL de Groot and H Grubmüller were supported by the BIOTECH Programs of the European Union (grants QLRT-2000/00778 and QLRT-2000/00504). P Agre was supported by grants from the National Institutes of Health. Y Fujiyoshi was supported by a Grant-in Aid for Specially Promoted Research (grant 13001003) and grants from the Japan Biological Informatics Consortium.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

* of special interest
** of outstanding interest


5. de Groot BL, Engel A, Grubmüller H: A refined structure of human aquaporin-1. FEBS Lett 2001, 504:206-211. This work demonstrates how a structure obtained by electron crystallography can be refined by exploiting information from a homologous protein at higher resolution.

6. Sui H, Han BG, Walian P, Jap BK: Structural basis of water-specific transport through the AQP1 water channel. Nature 2001, 414:872-878. The 2.2 Å structure of bovine AQP1 has been obtained by X-ray crystallography. It reveals the site of water molecules in a frozen state. The much improved resolution compared with the 3.8 Å achieved by electron crystallography [3**] confirms the refined model [5*].


9. de Groot BL, Grubmüller H: Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. Science 2001, 294:2353-2357. Multinanoscond MD simulations of an AQP1 tetramer and a GlpF tetramer embedded within a solvated lipid bilayer showed water permeation rates in agreement with the measured rates. A detailed mechanism of efficient yet selective water permeation involving the strong orientation and coordinated rotation of water dipoles is proposed. For GlpF, glycerol-induced gating is suggested.


11. Tajkhorshid E, Nollett P, Jensen MØ, Miercke LJW, O’Connell J, Stroud RM, Schulten K: Control of the selectivity of the aquaporin water channel family by global orientation tuning. Science 2002, 296:525-530. An X-ray structure of GlpF with bound water shows a narrower pore than the structure with bound glycerol. MD simulations of the GlpF tetramer embedded in a solvated lipid bilayer confirm the strong orientation of water molecules permeating through the glycerol channel. The NPA motifs, together with the HB and HE helix dipoles, are identified as the main cause of this strong orientation and the conserved NPA motif is proposed as the main proton filter.


24. Jensen MØ, Park S, Tajkhorshid E, Schulten K: Energectics of glycerol conduction through aquaglyceroporin GlpF. Proc Natl Acad Sci USA 2002, 99:6731-6736. The potential of mean force for glycerol permeation through GlpF has been computed by steered MD simulation, showing the major interaction sites of glycerol within the GlpF pore.


Structure and function of water channels Fujiyoshi et al. 515