Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance

Joseph D. Ho,a,b Ronald Yehb, Andrew Sandstromc, Ilya Chornyb, William E. C. Harriesb, Rebecca A. Robbinc, Larry J. W. Miercede, and Robert M. Strouda,b,1

aGraduate Program in Chemistry and Chemical Biology and bDepartment of Biochemistry and Biophysics, Genentech Hall, University of California, 600 16th Street, San Francisco, CA 94158-2517

Communicated by James A. Wells, University of California, San Francisco, CA, March 15, 2009 (received for review January 9, 2009)

Aquaporin (AQP) 4 is the predominant water channel in the mammalian brain, abundantly expressed in the blood–brain and brain–cerebrospinal fluid interfaces of glial cells. Its function in cerebral water balance has implications in neuropsychopathological disorders, including brain edema, stroke, and head injuries. The 1.8-Å crystal structure reveals the molecular basis for the water selectivity of the channel. Unlike the case in the structures of water-selective AQPs AqpZ and AQP1, the asparagines of the 2 Asn-Pro-Ala motifs do not hydrogen bond to the same water molecule; instead, they bond to 2 different water molecules in the center of the channel. Molecular dynamics simulations were performed to ask how this observation bears on the proposed mechanisms for how AQPs remain totally insulating to any proton conductance while maintaining a single file of hydrogen bonded water molecules throughout the channel.

The aquaporin (AQP) family includes both AQPs that conduct water, but not glycerol, and aquaglyceroporins that mediate diffusion of water, glycerol, and certain other small molecules in their neutral form across biological membranes. In humans, 13 different AQPs (AQP0–12) provide for transport in different tissues, each of which has broad clinical importance (1, 2). Besides AQP4, AQP1 and AQP9 are also expressed in the brain (3); AQP1 is expressed in the epithelial cells of the choroid plexus, and has a role in cerebrospinal fluid production, whereas AQP4 is localized to the endfeet of astrocytes in contact with the blood vessels of the blood–brain barrier and in astrocytic processes in contact with synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragonal arrays, and the improved response to water intoxication or stroke in AQP4 closely packed tetragonal arrays, and the improved response to synapses. From its tissue-specific concentrated localization in closely packed tetragon
contacts in the crystal, the interaction between the short 310 helices plane contacts of the latticed tetramers. Based on the molecular work of water molecules found at the intracellular side of the central pore. The extracellular side down the crystallographic 4-fold symmetry axis. (c) The network of water molecules found at the intracellular side of the central pore. The central pore is at the crystallographic 4-fold symmetry axis, and is formed by the tetramer. The 2Fo-FC density of the water molecules is shown in black, contoured at 1.2σ. The backbone amides of Ser-188 and Gly-189 are colored yellow in diagram representation. Phe-195 is shown as brown stick and cyan surface. (D–F) Diagram representation of the C loop of all of the AQP X-ray structures solved to date. (D) E. coli GlpF (brown), archael AqpM (magenta), spinach AQP SoPp2;1 (blue), and PIAQP (green). (E) Rat AQP4 (yellow), human AQP4 (black), human AQP5 (red), E. coli AqpZ (cyan), bovine AQP0 (green), and bovine AQP1 (purple). (F) Comparison of the 3α helix of rat AQP4 (yellow) with the 2-turn helix of AqpM (magenta) and spinach AQP (blue). All structural renderings were made with PyMOL (http://www.pymol.org).

molecules observed. The X-ray structure of hAQP4 at 1.8-Å resolution shows water molecules throughout the channel, 5 glycerol, and 1 octyl glucoside molecule. Each monomer, surrounded by 6 and 2 half-length alpha-helices (M1 to M8), tetramerizes along the crystallographic 4-fold c axis (Fig. 1 A and B) (26).

Central 4-Fold Axis. The physiological 4-fold axis insulates against all solutes and water. On the cytoplasmic side, a 4-fold arrangement of water molecules is stabilized by the backbone amides of Ser-188 and Gly-189 (Fig. 1C). Throughout ~21 Å of the midmembrane section, Phe-195, Leu-191, and Leu-75, repeated 4 times, create a hydrophobic block. This observation contrasts with the 4-fold axis in Plasmodium falciparum aquaglyceroporin, PIAQP, where the region is blocked by 4 aliphatic chains of phospholipids or fatty acids (27), and the human AQP5 where a single lipid molecule is found (25).

Rat AQP4 and Human AQP4. Although hAQP4 and the electron diffraction structure of rAQP4 in lipid bilayers crystallize in the same space group (P422), their crystal lattice contacts lie on different surfaces of the protein. The hAQP4 3D crystal contains head-to-head contacts only, because tetramers within the horizontal plane are too far apart (a = 82.1 Å) to make contact with each other (Fig. S2A). The rAQP4 2D crystal lattice has tetramers closer together (a = 69.0 Å), and contains both in-plane and between-plane contacts of the lattice tetramers. Based on the molecular contacts in the crystal, the interaction between the short 310 helices in the C loop was proposed to be a possible mechanism for AQP4-mediated cell–cell adhesion (Fig. S2B) (13). Although the sequences of the C loop are the same in hAQP4 and rAQP4 (Fig. S3), hAQP4 does not adopt the short 310 helix in this region (Fig. S4).

Extracellular Vestibule, Selectivity Filter, and Conducting Pore. AQP4 is a water-selective channel. Signature to the water-selective channels, His-201 lies directly in the selectivity filter, reducing the channel diameter to ~1.5 Å, sterically excluding the passage of glycerol (Fig. 2). AQP4 was purified and crystallized in the presence of 5% (vol/vol) glycerol (0.7 M), and 3 glycerol molecules are found in the extracellular vestibule, although not in the selectivity filter where the 2 glycerol-conducting AQPs, GlpF, and PIAQP, bind glycerol identically to one another (26, 27) (Fig. S5). In the water-selective rAQP1, the double mutant Phe56Ala and His180Ala (Phe-77 and His-201 in hAQP4) (Fig. S5) allows for the passage of glycerol, showing that steric occlusion is one mechanism for exclusion of larger solutes (28).

The ~25-Å long conducting pore contains a line of water molecules and no solute molecule. However, the electron density of the water molecules are distributed along the pore with residual positive Fo – Fe density observed in between water positions indicating increased anisotropic distribution along the channel axis, implying low-energy barriers between the water molecules along the direction of the channel (Fig. 3). As in other AQPs, the pathway through the channel is amphipathic. The hydrophobic sides are formed by the side chains of Phe-77, Ile-81, Val-85, Leu-170, Ile-174, and Val-197. The 8 backbone carbonyls of Gly-93, Gly-94, His-95, and Ile-96, from the cytoplasmic side and Gly-209, Ala-210, Ser-211, and Met-212 form the hydrophilic hydrogen bond acceptors for 8 positions of water molecules in transit. This arrangement allows bidirectional conductance of water from either side of membrane.

The asparagines 213 and 97 of the 2 almost totally conserved Asn-Pro-Ala (NPA) motifs form the canonical “fireman’s grip-like” structure in the center of the pore (26), and provide the defining force that orients water as it passes through the midpoint of the channel. However, in hAQP4, each asparagine donates its single, highly oriented hydrogen bond to a separate water molecule (Fig. 3). This arrangement is a key variant, because in 3 other water-selective AQP structures (23, 24, 29), the 2 asparagines donate hydrogen bonds to a single water molecule.

Conductance of hAQP4 in Proteoliposomes. Both full-length hAQP4 and trypsinized hAQP4 were reconstituted into proteoliposomes, and water and glycerol conductance were measured. The proteoli-
demonstrate that the protein expressed in AQP4 conduct water, but do not conduct glycerol. The data also show that trypsinolysis of the protein does not have an apparent effect on water conductivity.

**Fig. 3.** Electron density. Residues that form the wall of the pore are shown in sticks. Water molecules are shown as red spheres. The glycerol molecule is shown in green stick. The 2Fo-Fc density is shown in black, contoured at 1.2σ. Positive Fo–Fc density is shown in green, contoured at 3σ. There is no negative Fo–Fc density.

Protoposomes containing the full-length protein conducts water at the rate of 120.3 ± 9.5 s⁻¹, 12 times the rate for protein-free liposomes 9.80 ± 0.70 s⁻¹, and do not conduct any glycerol, with measured rate of 0.071 ± 0.001, the same as for protein-free liposomes 0.071 ± 0.0001. Proteoliposomes containing the trypsinized protein conducts water at the rate (relative rates) of 107.4 ± 0.0001. Proteoliposomes containing the trypsinized protein conduct water at the rate (relative rates) of 107.4 ± 0.0001. Proteoliposomes containing the trypsinized protein conducts water at the rate (relative rates) of 107.4 ± 0.0001. Proteoliposomes containing the trypsinized protein conducts water at the rate (relative rates) of 107.4 ± 0.0001.

Fig. 3. Electron density. Residues that form the wall of the pore are shown in sticks. Water molecules are shown as red spheres. The glycerol molecule is shown in green stick. The 2Fo-Fc density is shown in black, contoured at 1.2σ. Positive Fo–Fc density is shown in green, contoured at 3σ. There is no negative Fo–Fc density.

Discussion

**Implication of the C Loop in Cell–Cell Adhesion and NMO.** X-ray structures of AQPs from all 3 kingdoms, bacteria, archaea, and eukaryote, have been determined. The C loop that connects M4 and M5 and forms part of the extracellular vestibule shows remarkable features across all kingdoms. GlpF is unique in that part of the C loop folds as 2 α-helices. Both AqpM (30) and the spinach AQP SoPIP2;1 have a 2-turn helix entering the C loop. PIAQP has a short helix that is very similar to its bacterial aquaglyceroporin homolog, GlpF (Fig. 1D). Among others, rAQP4 stands out in that it has a 1-turn 3₁₀ helix (Fig. 1E) and is similar to the 2-turn helix seen in AqpM and the spinach AQP SoPIP2;1 (Fig. 1F). Hiroaki et al. (13) proposed from the rAQP4 electron diffraction structure that the 3₁₀ helix in the C loop (residues 139 to 142; Fig. S4) is the main region that provides for AQP4-mediated cell–cell adhesion (13). The hAQP4 and rAQP4 share the same protein sequences in this loop (Fig. S3), but in the X-ray structure of hAQP4, even though this region has the hydrogen bonds of an α-helix, it is a well-defined loop and does not adopt the dihedral angles (ϕ, ψ) of a 3₁₀ helix (Fig. S4). Therefore, the 3₁₀ helix in the C loop is a conformation that may be induced when 2 cells expressing AQP4 on the surface are in close proximity bringing 2 C loops into direct contact. It is also possible that the 3₁₀ helix is nonphysiological, induced by the packing of the rAQP4 double-layered structure in the electron microscopy samples (Fig. S2b) (13). In 2 different cell types, each with or without expression of orthogonally arrayed AQP4, separate cell–cell adhesion and dynamic light scattering assays demonstrated that AQP4 does not have any cell–cell adhesion property (31). Because in our 1.8-Å hAQP4 structure the C loop is not involved in the crystal packing, the lack of any helix in the C loop, therefore, supports the conclusion that AQP4 does not strongly drive adhesion, and that the 3₁₀ helix in the C loop observed in the rAQP4 structure may be induced by crystal contact.

Nevertheless, given the diversity and the flexibility of the C loop, it is possible that the C loop could have an important role in the pathogenesis of NMO. It has been proposed that the 3 extracellular loops (A, C, and E) could independently or jointly form the antigenic epitopes for the NMO-IgG autoantibody (32). Therefore, it is possible that the variable conformations of the C loop could aid in the affinity of the autoantibody to AQP4, so strengthening the downstream cascades of complement-mediated inflammatory response.

**Recruitment of Solutes in the Extracellular Vestibule.** In the crystal structures of GlpF and PIAQP, glycerol molecules are located in the selectivity filter, as well as in the entry vestibule (Fig. S6). In aquaglyceroporins, the vestibule may serve as a place for glycerol recruitment and desolvation of solutes for transport through the channel. For the water-specific bAQP1, the spinach AQP SoPIP2;1, bAQP0, and hAQP5, glycerol was not included in the crystallization conditions, but for the water-specific AqpZ where 5% glycerol (0.7 M) was present during crystallization, glycerol molecules were also not found in the vestibule (Fig. S6) (29). The 1.8-Å structure of the water-selective hAQP4 (crystallized in the presence of 5% glycerol), has no glycerol in the selectivity filter, but has 3 glycerol molecules nonspecifically bound in the extracellular vestibule. The extracellular vestibule, a canyon defined primarily by the A and C loop, interacts with the glycerol molecules through hydrogen bonds of the amide backbone with the hydroxyl groups of glycerol. It is tantalizing to speculate that the locations of the glycerol molecules...
in the vestibule could represent "fragment binding sites" for defining inhibitors that would bind from the extracellular side.

Because there are tremendous prospects for drugs that inhibit AQP4, and indeed many inhibitors of AQP4 have been described in the literature to date, we co-crystallized hAQP4 with 5 mM of 3 such compounds, tetraethylammonium (TEA), acetazolamide, and rizatriptan, and determined their ocystal structures. In oocytes swelling assays, these compounds were reported to be AQP4 inhibitors with IC50 in the micromolar range (6–9); however, we have not been able to detect any compound bound in the structures. We then reconstituted purified hAQP4 into liposomes, and measured water conductance in the presence of these compounds. TEA up to 10 mM had no effect on hAQP4 water conductance. Acetazolamide and rizatriptan do inhibit water conductance with an approximate IC50 in the low millimolar range (∼3 and ∼1 mM, respectively; Fig. S1b). We argue that water conductance measurement using purified proteoliposomes is more reliable than in oocytes. In a similar experiment, acetazolamide was also found to inhibit rAQP4 in the mM range, but not hAQP1 (33). These results call into question the previous interpretation of TEA, acetazolamide and rizatriptan as micromolar inhibitors of AQP4.

**Arginine 216 Environment Determines Conductance of Water in AQP4.**

Because the glycerol conducting GlpF, with low water conductance, the selectivity filter arginine (Arg-206) has only one of the NH satisfied (Fig. 4), so it is possible that the extra cost of desolvating the fully charged guanidinium leaves GlpF "holding on" to water molecules in transit. This higher degree of the guanidinium cation buffering by the protein may be the basis for higher efficiency water transport among the AOPs.

**Variant Role of the Dual NPA Motifs in Proton Exclusion.**

The interlocking of 2 almost totally conserved NPA motifs in AQP4 provides 2 highly oriented donors from NH2 of the 2 asparagines to the center of the channel. We proposed that they have a key role in insulation against any conduction of protons or ions while supporting a single file of hydrogen bonded water molecules throughout the entire length of the channel (34). In the crystal structures of the water-selective *Escherichia coli* AqpZ and bAQP1, the central water is positioned in between the 2 asparagines of the NPA motifs, and receives 2 hydrogen bonds from the 2 asparagines NH2s (represented by bAQP1 in Fig. 5A and B). Such orientation of the central water may be a factor in preventing proton conduction through the channel by the highly cooperative hop-and-turn Grothuss relay mechanism (35, 36), because the central water cannot rotate to accept or exchange its proton (26, 37). However, the hAQP4 crystal structure does not have a water molecule centrally located between the 2 NPA motifs. Instead, each asparagine of the NPA motifs (Arg-197 and Arg-191) receives 1 additional hydrogen bond from the protein, and 2 nearly totally conserved NPA motifs in AQPs provide insulation against any conduction of protons or ions while supporting a single file of hydrogen bonded water molecules throughout the entire length of the channel (34). In the crystal structures of the water-selective *Escherichia coli* AqpZ and bAQP1, the central water is positioned in between the 2 asparagines of the NPA motifs, and receives 2 hydrogen bonds from the 2 asparagines NH2s (represented by bAQP1 in Fig. 5A and B). Such orientation of the central water may be a factor in preventing proton conduction through the channel by the highly cooperative hop-and-turn Grothuss relay mechanism (35, 36), because the central water cannot rotate to accept or exchange its proton (26, 37). However, the hAQP4 crystal structure does not have a water molecule centrally located between the 2 NPA motifs. Instead, each asparagine of the NPA motifs (Arg-197 and Arg-191) receives 1 additional hydrogen bond from the protein, and 2 nearly totally conserved NPA motifs in AQPs provide insulation against any conduction of protons or ions while supporting a single file of hydrogen bonded water molecules throughout the entire length of the channel (34).

![Diagram of AQP4 and AQP1](image-url)

**Fig. 5.** The NPA motifs. (A) Schematic representation of the hydrogen bonding network through the channels of hAQP4 and bAQP1. The distances are between heavy-atom to heavy-atom. (B) Stick representation of the NPA motifs. Distances that are too long to be a hydrogen bond are colored in red. (C) Plot of the MD simulations of hAQP4 from 4 different experiments. Details are described in Discussion.
mechanistic terms, molecular dynamic (MD) simulations were performed to assess the probability of finding a single central water molecule bonding to both NPA asparagines simultaneously. The results of the MD simulations were used to calculate the number density of water molecules along the channel near the 2 NPA motifs (Fig. S7). Four simulations were carried out. In the first simulation (thick solid line), the system commenced from the protein and crystallographic waters; the protein positions were kept frozen throughout the simulation. The resulting number density is consistent with the observed water positions in the hAQP4 crystal structure (solid red spheres), and suggests that this particular configuration of the protein supports 2 energy minima for water in the NPA region. To test whether this result was biased by the introduction of the crystallographic water molecules into the simulation, a second simulation (dotted line) was performed, in which the protein positions were kept frozen, but the crystallographic waters were removed from the simulation, and the pore evacuated. After equilibration with the remaining simulation water molecules, the resulting number density is consistent with the experimentally observed water positions in the hAQP4 crystal structure, and gives further support that the calculation is stable and reiterates the particular observed configuration of hAQP4. In the third simulation (dashed line), the system contained both the protein and the crystallographic waters, but instead of freezing the atomic positions of the protein, the heavy atoms (N, C, and O) were restrained to their crystallographic positions using a harmonic restraint. This slight release of the protein atoms from their crystallographic positions allows the number density to shift such that a water molecule positions itself in the middle of the NPA motifs, as is observed in the crystal structures of bAQP1 (Fig. 5 A and B; Fig. S7). Last, in the forth simulation (dashed line), all of the positional restraints were removed from the protein, allowing the protein to move freely. Again, the resulting number density suggests the presence of a water molecule in the middle of the NPA region. Thus, we conclude that the observed position of the 2 waters near the NPA region, in the hAQP4 crystal structure, do not conform to a functionally distinct variant, but rather, signal the lowered barriers of the protein, the heavy atoms (N, C, and O) were restrained to their crystallographic positions using a harmonic restraint. This slight release of the protein atoms from their crystallographic positions allows the number density to shift such that a water molecule positions itself in the middle of the NPA motifs, as is observed in the crystal structures of bAQP1 (Fig. 5 A and B; Fig. S7).

Analogous sets of simulations were performed on bAQP1. In all 4 simulations, the number density supports the mechanism that as a water molecule travels through an AQP or aquaglyceroporin channel, it does transition into and through a central position where it simultaneously accepts hydrogen bonds from both NPA asparagine donors, and that this in transition state, the line of waters throughout the channel are polarized such that all waters have their arginine donors, and that in this transition state, the line of waters was transitioned into and through a central position where a water molecule moves freely. Again, the resulting number density suggests the presence of a water molecule in the middle of the NPA region. Thus, we conclude that the observed position of the 2 waters near the NPA region, in the hAQP4 crystal structure, do not conform to a functionally distinct variant, but rather, signal the lowered barriers to conductance between these sites.

Our first AQP structure, for GlpF, led to our suggestion that the bipolar line of water, the charge on the conserved valence bond (MS-EVB) models for proton transfer, Ilan et al. (37) concludes that other factors including the NPA region within the channel are important as well (42). Using multiscale empirical valence bond (MS-EVB) models for proton transfer, Ilan et al. (37) suggests that the bipolar line of water, the charge on the conserved arginine of the selectivity filter, and ion desolvation penalties are all relevant as we proposed initially (34).

### Perspective

Because of the flexibility of the termini, the N-terminal 19 aa and the C-terminal 64 aa were removed from the M1 isoform by trypsinolysis, as determined by MS and N-terminal sequencing. Thus, the crystal structure we report at 1.8-Å resolution contained Glu-20 to Lys-259 with visible electron density from Gln-32 to Pro-254. In contrast, the electron diffraction structure of rAQP4 was expressed as the M23 isoform, and revealed density for residues 31–254, consistent with a natively unstructured region at each of the termini. Thus, whereas the structure fully elucidates all features of the water transport pathway, it does not yet address the question of how the M1 isoform works together with the M23 isoform to limit the extent of orthogonal array formation. It also remains to be determined how the C terminus interacts with α-synthrophin in the context of polarized expression in astrocytes.

Questions also remain on the mechanism of how Ser-111 and Ser-180 phosphorylation affect AQP4 conductance. Studies have shown that the phosphorylation state of Ser-111 on the B loop and Ser-180 on the D loop affect water conductance of AQP4 (Fig. S3) (18–20). Phosphorylated Ser-180 has been speculated to interact with Lys-259, Arg-260, and Arg-261 as a means of tethering the C-terminal domain to block the channel (13), and phosphorylated Ser-111 has been proposed to mimic the gating mechanism of the spinach AQP SoPIP2;1 by disrupting the network of hydrogen bonds that anchor the D loop to the N terminus; thus, opening the channel (20). In our 1.8-Å crystal structure, the channel is open and, although the density for the B and D loops is clear, no phosphate is observed on either serine. Because the protein was trypsinized, neither the N- and C-terminal domains are present. Although we

### Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th>P42 2</th>
<th>Space group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell dimensions a, c, Å</td>
<td>82.1, 76.4</td>
<td>1.8 (1.86–1.8)</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>9.0 (79.6)</td>
<td>20 (3.3)</td>
</tr>
<tr>
<td>Rwork/Rfree</td>
<td>99.9 (100)</td>
<td>11.9 (12.0)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>23,583</td>
<td>16.0/16.5</td>
</tr>
<tr>
<td>No. of atoms</td>
<td>27</td>
<td>42</td>
</tr>
<tr>
<td>Protein</td>
<td>20</td>
<td>55</td>
</tr>
<tr>
<td>Glycerol</td>
<td>61</td>
<td>45</td>
</tr>
<tr>
<td>Water</td>
<td>1,667</td>
<td>0.01</td>
</tr>
<tr>
<td>Avg B-factors, Å²</td>
<td>30</td>
<td>Bond lengths, Å</td>
</tr>
<tr>
<td>Protein</td>
<td>20</td>
<td>1.047</td>
</tr>
<tr>
<td>Glycerol</td>
<td>42</td>
<td>Bond angles, °</td>
</tr>
<tr>
<td>Water</td>
<td>55</td>
<td></td>
</tr>
</tbody>
</table>

Values in parentheses are for the highest resolution shell. Rfree is calculated from 5% of reflections chosen randomly.
cannot comment on the interaction of Ser-180 with the C-terminal domain, we think it is unlikely that the gating of AQP4 at Ser-111 is similar to the Ser-OH interaction. The loop of hAQP4 is 4 residues shorter than in the M1-21, and residues in the N-terminal domain that are involved in binding to the loop (Asp-28 and Glu-31) are different amino acids in hAQP4 (Fig. S3).

The high resolution structure of the phosphorylated protein, the M1 isoform, and complexes with binding elements of syntrophin may address these issues of larger scale assemblies beyond that of the tetrameric AQP4.

Materials and Methods

Materials and methods for protein expression and purification, crystallization, proteoliposome assay, and MD simulations are described in SI Materials and Methods.

For data collection and model building, diffraction data were collected using a wavelength of 1.1 Å at Beamline 8.3.1 at the Advanced Light Source (Lawrence Berkeley National Laboratory). Data were processed by using HKL2000 (43). Molecular replacement was performed with Phaser (44) using the rAQP4 electron diffraction structure (SI Materials and Methods).