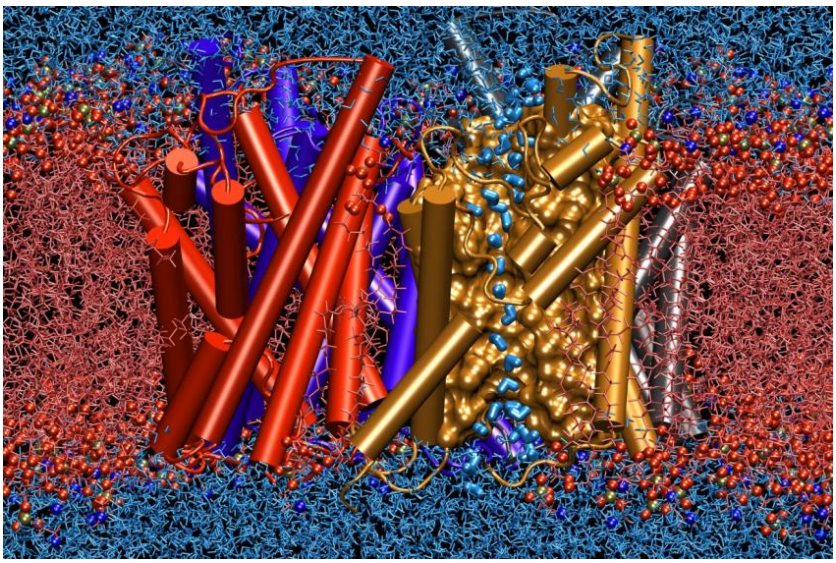
Case Study: Aquaporins

Yi Wang, Jin Yu and Emad Tajkhorshid
October 23, 2012



1 Universal Water Channels: Aquaporins

The organization of water is critical to most biological processes, whether on a cellular or multicellular level. Although cell membranes are to some extent water-permeable, they cannot facilitate the rapid exchange of large volumes of water, as required by the kidneys in the human body. Aquaporins (AQPs) are a family of specialized integral membrane proteins which function mainly as water channels. They usually conduct water at a rate of $\sim 10^9$ molecules per second, which is almost comparable to the free diffusion of water. These highly efficient water channels can explain how we secrete tears, saliva and sweat, how our kidneys concentrate urine, and how our brains maintain spinal fluids [1].

AQPs are present in all domains of life, including archaea, bacteria, plants, and all phyla of animals. Members of the AQP family can be di-

vided into two subfamilies based on their permeability characteristics: "classic AQPs" which conduct water exclusively, and "aquaglyceroporins" which possess the extended ability to conduct small linear carbohydrates, in particular, glycerol, a metabolic intermediate [1]. So far, more than 150 sequences of the AQP family have been identified [1, 2].

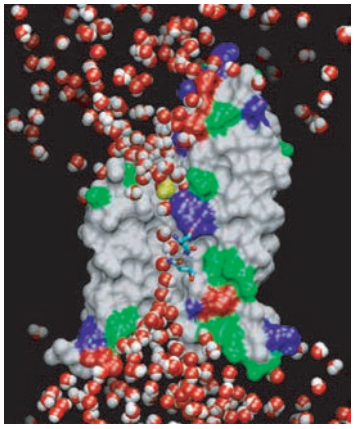


Figure 1: Snapshot of water permeation through aquaporin GlpF. The protein is displayed in surface representation and colored by residue type (red: acidic, blue: basic, white: nonpolar, green: polar). Water molecules are represented in red and white, with the exception of one highlighted in yellow.

Fig. 1 is a snapshot taken from a molecular dynamics (MD) simulation (using the program NAMD2 [3]) showing in atomic detail how water molecules permeate through an AQP channel. Watch the movie *waterpermeation.mpg* generated from the simulation. You will notice that water molecules form a single file and move back and forth along the channel. These spontaneous movements, caused by thermal fluctuations, may result in complete transport of water through the channel, as demonstrated by the movement of the water molecule colored in yellow.

In the human body, AQPs are found in various organs including the brain, lens, lungs, kidneys, etc. Fig. 2 shows schematically the distribution of AQPs throughout the human body. Mutations of these proteins are responsible for human diseases ranging from nephrogenic diabetes insipidus (NDI) to congenital cateract formation [1, 2]. In NDI, the misfolding of AQP2 creates water reabsorption problems in the collecting ducts of the kidneys. In the case of congenital cataract, mutations in the structurally important residues of AQP0 may result in opacity in the lens.

AQPs in plants participate in various physiological processes, such as photosynthesis. One of the major ingredients of the photosynthesis, carbon dioxide, passes through leaf aquaporins [4]. Interestingly, yeast aquaporins confer resistence to freezing, and a yeast aquaglyceroporin protects the organisms against hypo-osmotic shock [5].

Since the discovery of the first aquaporin AQP1 in rabbit red blood cells [6], researchers have been extensively studying the AQP family. Nowa-

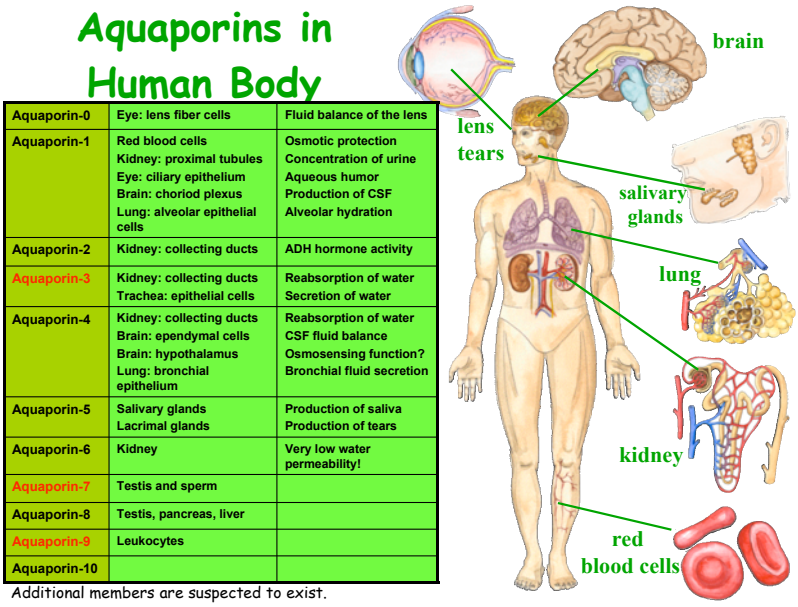
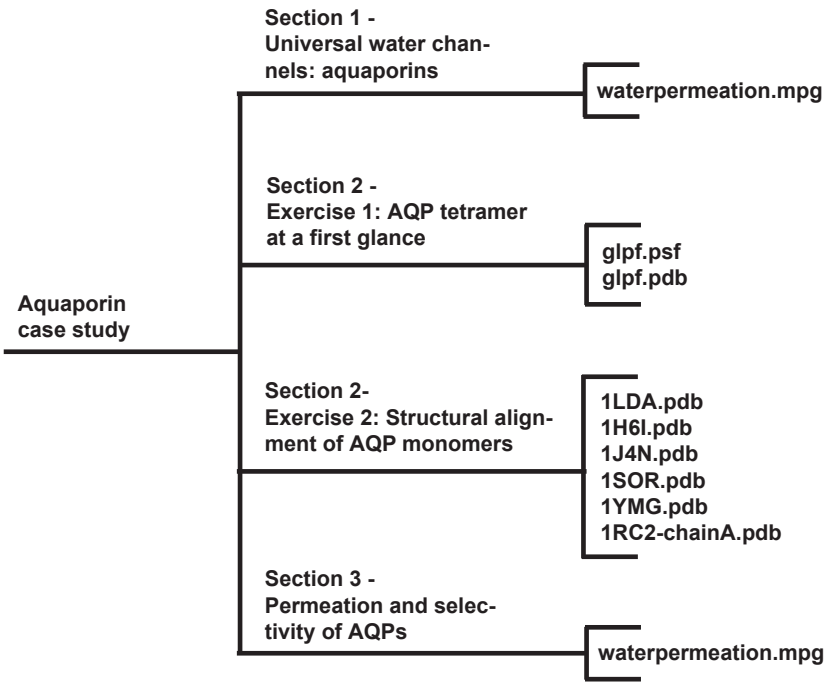


Figure 2: The distribution of 11 different AQPs in the human body. The table shows the specific locations and functions of each aquaporin in the human body.

days, more than 11 AQP structures are available in the Protein Data Bank (PDB). The physical structure and dynamical characteristics of AQPs can therefore be explored in atomic detail with advanced visualization and computation technologies.

In the following sections, we will first explore the structural features of AQPs, and then examine the permeating mechanism and selectivity of AQPs as revealed by MD simulations.

In this case study, we'll use the following files as shown below.



2 Structure and Sequence Features of AQPs

AQPs organize as tetramers in cell membranes, with each monomer forming a functionally independent water pore. This tetrameric structure is universal to the entire AQP family, although why these water channles decided to choose such a configuration is still unknown. In this section, we will first look at the AQP tetramer embedded in a lipid membrane, and focus then on the AQP monomer.

Exercise 1: AQP tetramer at a first glance. In this exercise you will explore the tetrameric structure of the aquaporin GlpF.

- Start a new VMD session and load the files glpf.psf and glpf.pdb. Create seperate representations for "water", "lipid" and "protein" using the style "Lines", "Licorice" and "Cartoon" respectively. Then change the "Coloring Method" of the protein to "SegName".
- Turn off the "water" representations by double clicking it and your system should now look like the one shown in Fig. 3 a. Rotate the system in VMD and get a side view of the system, as Fig. 3 (b).
- Turn off the "lipid" and "protein" representations and turn on the "water" representation. Rotate the system and identify those water molecules forming single files across the membrane. Fig. 3 c gives you a better view of the four single water files by highlighting them in vdW representations.
- Turn off the "water" representations, turn on the "protein" representation and color it with "ResType" using the "Surf" style. Turn on the "lipid" representation as well and you will get a similar picture as we have in Fig. 3 d. Pay attention to the part of the protein colored in white and the part of lipids colored in cyan.

You may notice that the two parts roughly match in length. This results from the protein-lipid interaction mechanism known as "hydrophobic matching", which determines that the hydrophobic thickness of the lipids be matched with the hydrophobic length of the transmembrane domain of a membrane protein.

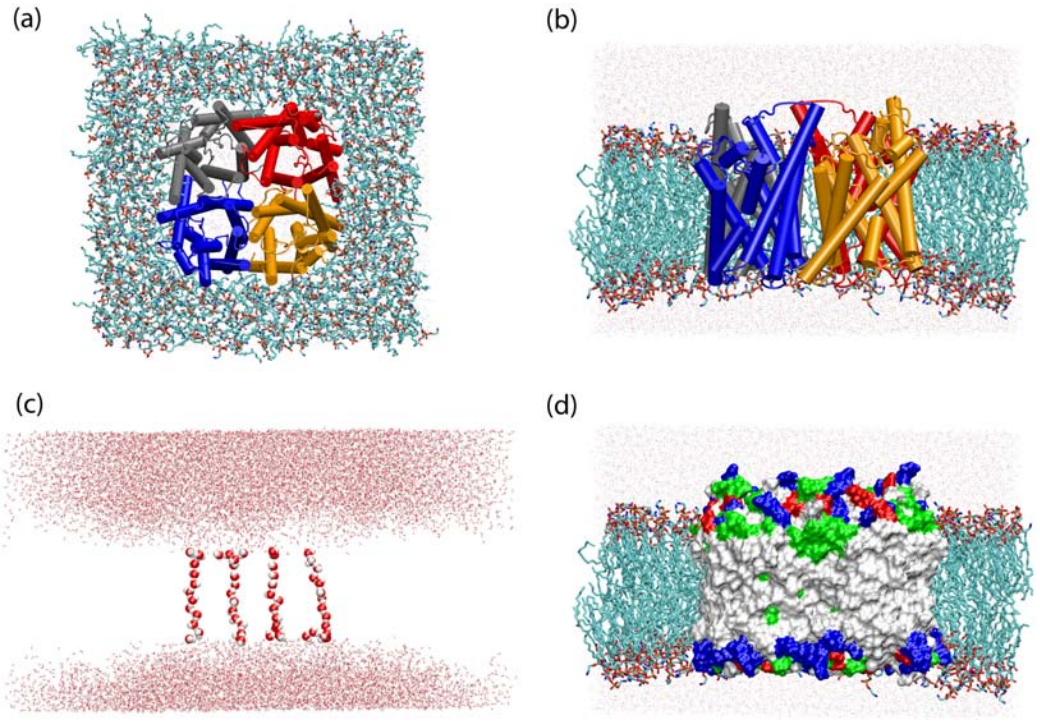


Figure 3: GlpF tetramer embedded in a lipid membrane and surrounded by water. (a) Top view and (b) side view. The four monomers are distinguished by color. Water molecules appear as red points. (c) Single files of water connecting the cellular and extracellular bulk water. Water inside the GlpF monomers are shown in vdW representation while water on the two sides of the membrane (i.e., bulk water) are represented in lines. (d) GlpF tetramer shown in surface representation. The reader is encouraged to reproduce the figure with VMD using the supplied files glpf.psf and glpf.pdb (See Exercise 1).

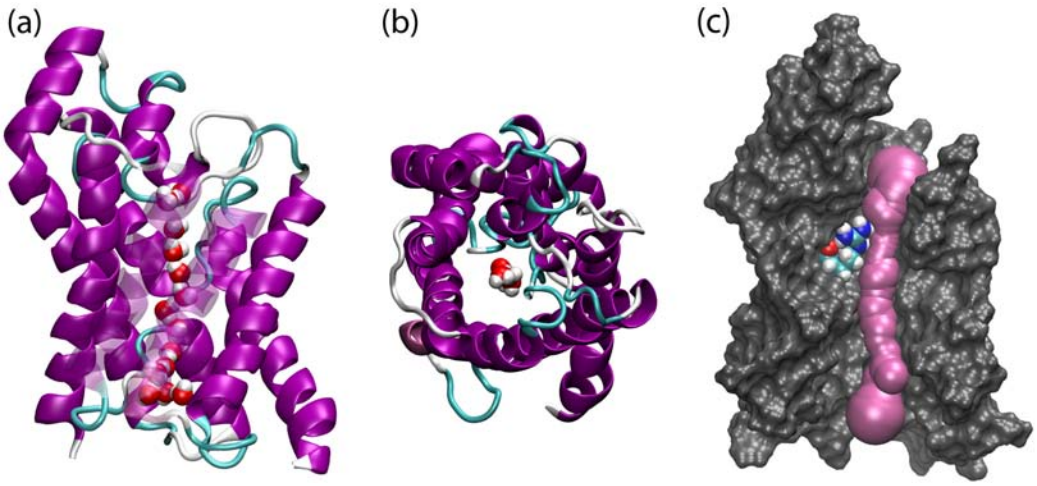


Figure 4: Side (a) and top view (b) of the GlpF monomer. The monomer is colored by structure type (purple: α-helix, mauve: 3_{10}-helix, cyan: turn, white: coil). The single file of water molecules inside the GlpF monomer is shown in vdW representation. (c) Pore of GlpF. Part of the protein is shown in gray surface representation with the Arg206 at the selectivity filter highlighted in vdW representation. (a) and (b) are created with VMD from mono.psf and mono.pdb, the figures can be obtained by representing the protein in "New Cartoon" and coloring it by "Structure".

2.1 AQP monomer

Each monomer in an AQP tetramer is a functionally independent water channel. The secondary structure of an AQP monomer consists of six transmembrane helices and two reentrant loops. The helices are arranged such that a pore forms in between. Fig. 4 a shows one of the four monomers of the GlpF system, with a single file of water molecules inside the channel.

Fig. 4b provides a top view of the monomer with a cross-section of the water file. Although it appears that there's plenty of space left in the channel even with the presence of water, the pore, in fact, is rather narrow. To better deduce the shape and size of an AQP channel, we analyzed the pore dimension of a GlpF monomer with the program HOLE [7]. Along the channel axis of GlpF, HOLE looks for the largest sphere that can be accommodated by the channel without overlap with the van der Waals surface of protein atoms. [7]. The result was then visualized by VMD, as shown in Fig. 4 c.

The narrowest region in an AQP channel is the selectivity filter (SF). The "classic" aquaporin, AqpZ has a 1.2 Å radius at the SF, just wide enough for

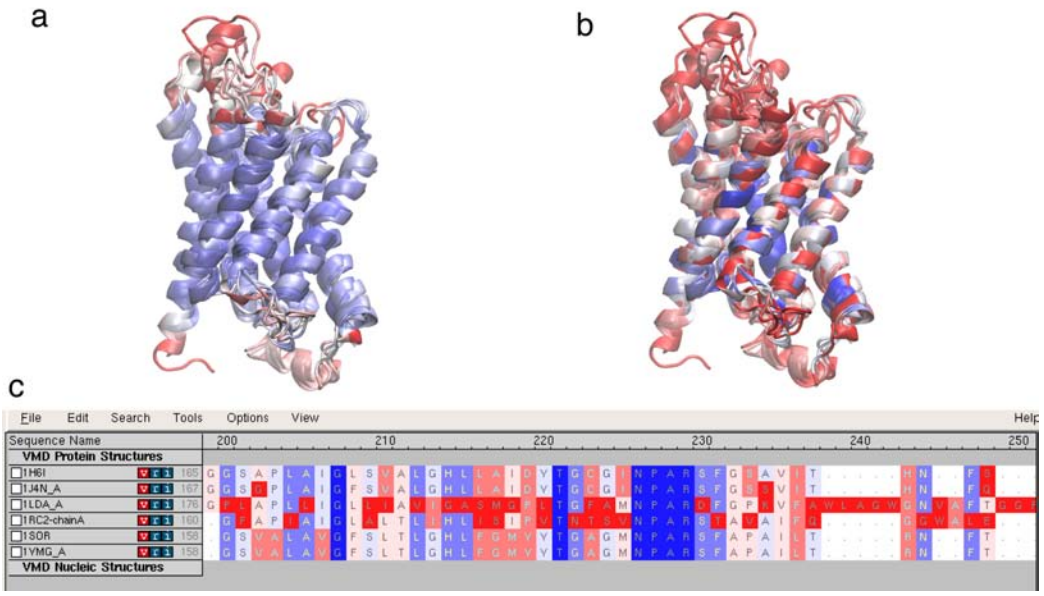


Figure 5: Aligned structures of GlpF (1LDA), human AQP1 (1H6I), cow AQP1 (1J4N), AqpZ (1RC2), sheep AQP0 (1SOR) and cow AQP0 (1YMG). The structures are colored by Q values from structural alignment in (a) and by sequence identity per residue in (b). Red represents the least similar segments and blue represents the most similar ones. (c) Sequence display window of VMD Multiple Alignment. Some highly conserved residues including the arganine at SF and the NPA motif are highlighted in yellow.

one water molecule to pass through. For GlpF, the channel radius is $\sim$1.7 Å at the SF. Interestingly, both AqpZ and GlpF are from the bacterium *E. coli*. While AqpZ is a pure water channel, GlpF allows larger substrates such as glycerol to pass through. The difference in their pore dimensions is therefore closely related with their distinct functions.

Aside from examining the structure of a GlpF monomer, one can align and compare the structures of different AQP monomers with VMD. The structural alignment provides insight into how these proteins evolved into membrane channels for various organisms. Here pdb files for six AQPs from human, cows, sheep, and *E. coli* are obtained from the PDB. You can perform a structural alignment on them following Exercise 2.

Fig. 5 a and b present the structural alignment in two different coloring schemes. In both figures blue signifies high levels of similarity, red signifies low levels. Fig. 5 a is colored by *Qres*, a parameter indicating the structural

similarity — higher Q values represent closer structures. Fig. 5 a shows that aligned AQP monomers are mostly blue in the center, suggesting they share a highly conserved core structure. As a matter of fact, the universal core structure of AQPs consists of 208 residues with some highly conserved sequence motifs [8].

In Fig. 5 b AQPs are colored by *sequence identity*. As revealed by the mostly red color of this figure, there is less sequence similarity than structural similarity among various AQPs. However, upon closer examination, there are segments of blue on the helices in the central part of the channel, revealing some highly conserved sequence patterns. They can be better identified in the sequence alignment window, as shown in Fig. 5 c.

Exercise 2: Structural Alignment of AQP Monomers. In this VMD exercise you will align and compare the structures of six AQP monomers.

- Start a new VMD session and load files 1H6I.pdb, 1J4N.pdb, 1LDA.pdb, 1RC2-chainA.pdb, 1SOR.pdb, and 1YMG.pdb. Change the "Style" of each molecule to "Tube" in the "Graphical Representations" window.
- In the VMD main menu, click "Extensions→Analysis→MultiSeq", highlight all non-protein sequences and press delete to remove them, and then choose "Tools→Stamp Structural Alignment" in the "MultiSeq" window.
- Now the six AQPs should be aligned in your VMD graphic window. Go to "View→Coloring" in the "MultiSeq" window and choose "Qres". Your system should now look like the one shown in Fig. 5 a. To get Fig. 5 b, simply choose "Sequence Identity" instead.

Q is a parameter that indicates structural identity. Q=1 implies that structures are identical. When Q has a low score (<0.3), structures are not aligned well [9].

2.2 Taking a closer look at AQPs

Some of the most conserved structural features appearing in Fig. 5 are of great importance to the proper function of AQPs. The "reentrant loops" are such a structural feature conserved among all AQPs. As shown in Fig. 6 a, they are the two loops entering the center of a GlpF monomer from both ends of the channel. Each reentrant loop is formed by a short helix and an extended polypeptide.

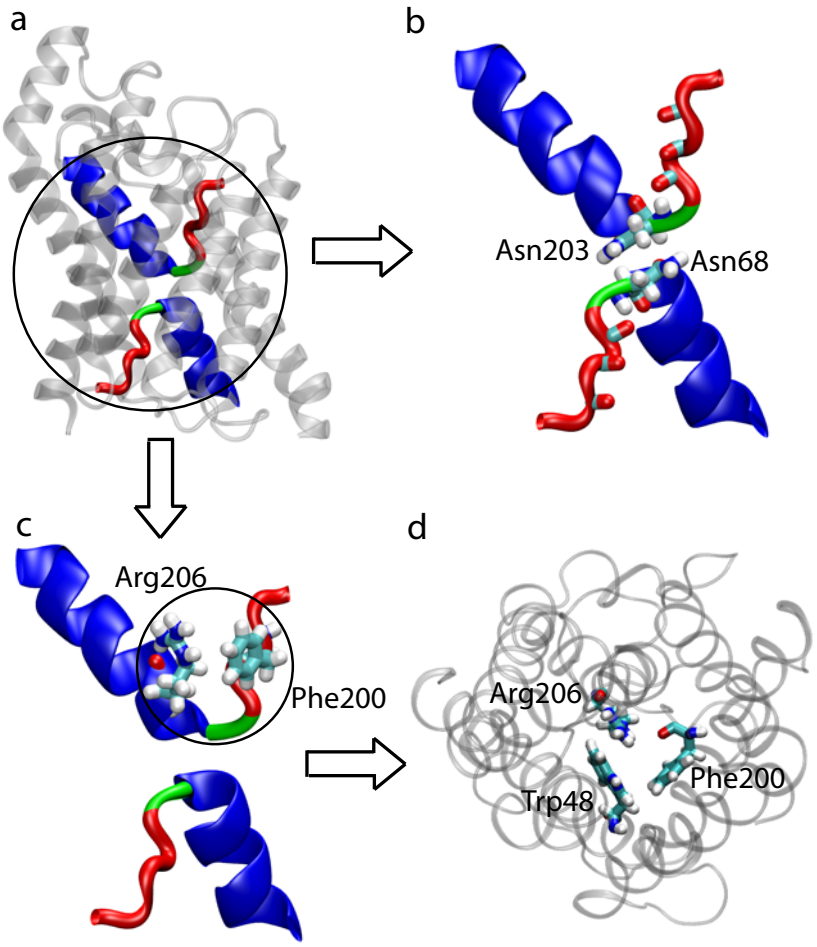


Figure 6: (a) GlpF's reentrant loops, composed of a short helix (blue) and an extended polypeptide (red, except for the two asparagines of NPA motifs, which are colored in green). The rest of the monomer is shown in transparent gray. (b) Asn203 and Asn68 of the NPA motifs and the carbonyl groups of residues on reentrant loops are shown in Licorice representation. (c) Arg206 and Phe200 at the selectivity filter are shown in Licorice representation. (d) Top view of the selectivity filter, which is composed of Arg206, Phe200 and Trp48.

On each of the two reentrant loops, near the center of the channel, there is a signature motif of AQPs, named the "NPA motif". It is composed of three consecutive residues, asparagine-proline-alanine. The two asparagine residues of the two NPA motifs are colored in green in Fig. 6 a to give you an idea about their positions. NPA motifs are indispensable to the primary function of AQPs: water conduction. Futhermore, as we will show later, they are also crucial in the exclusion of protons by AQPs.

Opposite the NPA motifs across the channel there are only hydrophobic residues. The carbonyl groups on the backbones of these residues form the sole hydrophilic side of the channel in an otherwise purely hydrophobic environment, as shown in Fig. 6 b. This hydrophilic side helps water molecules maintain certain orientations inside the channel.

The SF(shown in Fig. 6 c and d) is the narrowest section of the channel. In GlpF, the SF is composed of three residues: arginine, phenylalanine and tryptophan. The arginine at SF is highly conserved among all AQP family members.

3 Permeation and Selectivity of AQPs

Most biomolecular processes, i.e., the folding of proteins, occur on a time scale that cannot be readily approached by current computational powers. However, water permeation events through AQP channels happen on a nanosecond time scale and can thus be studied in real time using advanced computational technologies. In the following sections we will focus on a few of these computational studies that have significantly contributed to our understanding of the AQP family.

3.1 Water permeation in AQPs

As shown in the movie *waterpermeation.mpg*, water diffuses through AQP channels even under equilibrium conditions. Further studies indicate that water molecules translocate through the AQP channel in a highly correlated way: the channel is almost always occupied by about seven water molecules and the entire single file of water moves in hops simultaneously. The diffusion coefficient is found to be $\sim 0.4 \times 10^{-5}\mathrm{cm}^2\mathrm{s}^{-1}$, millions of times higher than the diffusion coefficient of water through pure lipid bilayers [10].

However, under the equilibrium conditions mentioned above, there is no

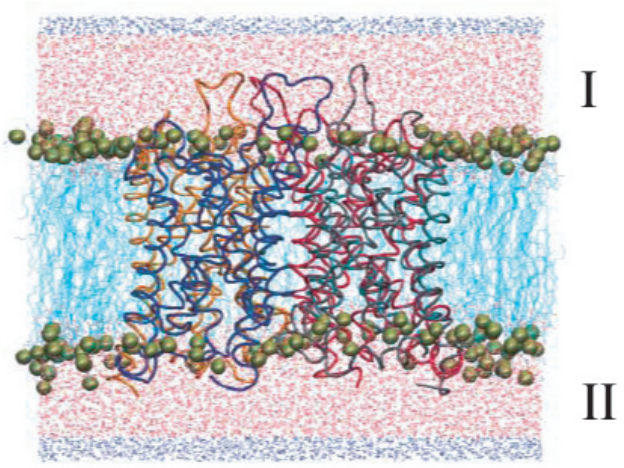


Figure 7: Schematic illustration of pressure induced water conduction through aquaporin. The figure shows aquaporin in a lipid bilayer membrane, the membrane separating water at pressure $p_{\rm I}$ (top, region I) and water at pressure $p_{\rm II}$ (bottom, region II). The pressure difference $p_{\rm I} - p_{\rm II}$ drives the water from top to bottom.

net water flux through AQP channels. Then how do AQPs carry out water transport across the membrane? The answer is that physiological conditions are usually far from equilibrium, i.e., different concentrations of impermeable solutes on the two sides of a membrane induce an osmotic pressure difference across the membrane, driving water to flow from the side with high osmotic pressure to the side with low osmotic pressure. Fig. 7 illustrates schematically the water, membrane, aquaporin system. Aquaporins can be observed and simulated under these conditions and the water flow can be measured or calculated. The ability of aquaporin to conduct water can be stated best through the so-called osmotic permeability p_f that obeys the relationship

$$\text{water flux (water conducted per channel every second)} \;=\; p_f \;\cdot\; \frac{(p_{\rm I} - p_{\rm II})}{k_B T}$$

where k_B is the Boltzmann constant and T denotes the temperature. The osmotic permeability measured and calculated is about ($\sim 7{\times}10^{-14}\mathrm{cm}^3\mathrm{s}^{-1}$) [11].

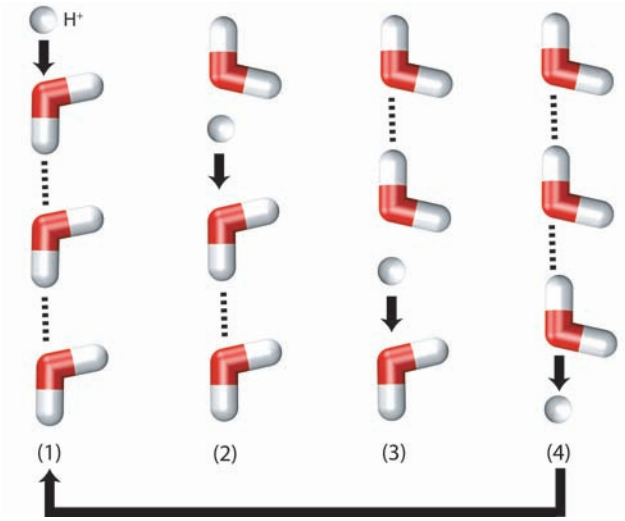


Figure 8: Schematic proton conduction mechanism. The proton is conducted in the direction shown by black arrows. Hydrogen bonds are indicated in black dashed lines. A proton is transferred by the rearrangement of hydrogen bonds in a hydrogen-bonded water file.

3.2 Proton exclusion mechanism of AQPs

While allowing water to pass through rapidly, AQPs exclude ions, protons and other charged solutes. Long before scientists understood the mechanism of this remarkable feature of AQPs, they were aware of its great importance. The proton concentration gradient across the membrane plays a key role in metabolism. By excluding protons, AQPs help a cell to maintain the electrochemical balance across cell membranes, which would be easily impaired if these highly efficient water channels allowed protons to pass through.

Before we move on to look at the proton exclusion mechanism of AQPs, it is necessary for us to first understand how protons are conducted. As shown in Fig. 8, by using water molecules as stepping stones, a proton can readily "jump" from one water molecule to another in a hydrogen-bonded water file. This mechanism, namely the cleavage and rearrangement of hydrogen bonds between neighboring water molecules is called the "Grotthuss Mechanism" [12]. It allows protons to transfer across several water molecules on the picosecond time scale, thousand times faster than the nanosecond time scale needed for water itself to fully permeate an AQP channel.

The conduction of protons through protein channels shares a similar mechanism. In other words, given that a properly hydrogen-bonded chain of water molecules (named a "proton wire") exists in a protein channel, protons can be conducted through the channel. Now an interesting question arises why protons are blocked by AQPs, since such a proton wire was well expected to exist in these water channels. The answer was found by MD simulations.

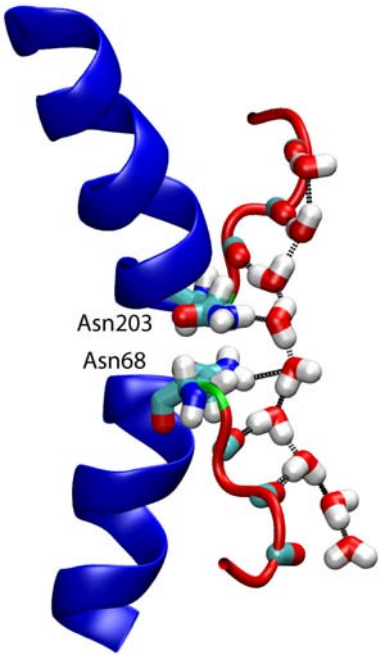


Figure 9: Water orientations in the GlpF channel. The asparagines at the NPA motifs and carbonyl groups of residues on reentrant loops are highlighted. Hydrogen bonds are shown in black dashed line.

Referring back to Fig. 1, a four nanosecond simulation was performed on the GlpF tetramer. The movie shows a highlighted water molecule passing through the channel of a monomer. Watch the movie *waterpermeation.mpg* again and pay attention to the water orientations along the whole channel. Can you see any difference in molecule orientations between water inside the GlpF channel and water shown in Fig. 8?

As shown in Fig. 9, starting from NPA motifs, water molecules are oriented in opposite directions in the two halves of GlpF channel, with their oxygen atoms all facing toward the center from both entrances of the channel. A water molecule near the two NPA motifs assumes an almost perpendicular orientation relative to the channel axis. Thereby a "bipolar" water file is formed. As we mentioned above, proton conduction requires a uniform orientation of hydrogen-bonded water molecules that permits reorientation during proton transfer. This requirement is not met by the bipolar water file in GlpF. Water molecules are highly restricted in their orientations inside an AQP channel, making the rearrangement of hydrogen-bonds extremely difficult.

This amazing bipolarity of the water file is the reason for proton exclusion by AQPs. The bipolarity breaks a potential proton wire and effectively excludes protons. This mechanism for proton exclusion seems to apply to the entire AQP family [13].

As shown in Fig. 9, in the center of the water file, a water molecule is hydrogen bonded to the two asparagines of NPA motifs. Other hydrogen bonding patterns are prohibited because there are only hydrophobic residues across the channel around NPA motifs. Since the oxygen of the central water molecule can only form H-bonds with the two asparagines, the molecule's orientation is highly constrained and only its two hydrogen atoms are free to form hydrogen bonds with neighboring water molecules. Therefore, successive O-H bonds are formed and oriented away from this central water molecule as shown in Fig. 9. The bipolar orientation is propagated outward along the water file.

In addition to the effect of NPA motifs, reentrant loops also help to maintain the configuration of the bipolar water file: the two α-helices, as shown in Fig. 9, generate two electrostatic dipoles which point toward the center of the channel, forcing water dipoles to orient in the opposite direction. Under their influence, water molecules tend to point their oxygens toward the center of the channel.

4 Glycerol Conduction in AQPs: a Stereoselective Process

Aquaporins, as considered throughout the case study, have the ability to conduct water while excluding charged solutes. However, a subfamily of AQPs, aquaglyceroporins, have the extended capability of conducting small linear carbohydrates, which can be largely attributed to an increased pore radius as discussed before.

The aquaglyceroporin GlpF conducts glycerol along with water in the bacterium *E. coli*. Interestingly, it was found that glycerol molecules in certain orientations are preferred to pass through the channel. In other words, the GlpF channel has a stereoselectivity for glycerol [16].

As shown in Fig. 4 b left, when glycerol enters the GlpF channel in the favorable orientation, a series of hydrogen bonds are made between hydroxyl groups of the glycerol and the protein. Besides, the aliphatic backbone of glycerol is optimally pointed to the hydrophobic side of GlpF.

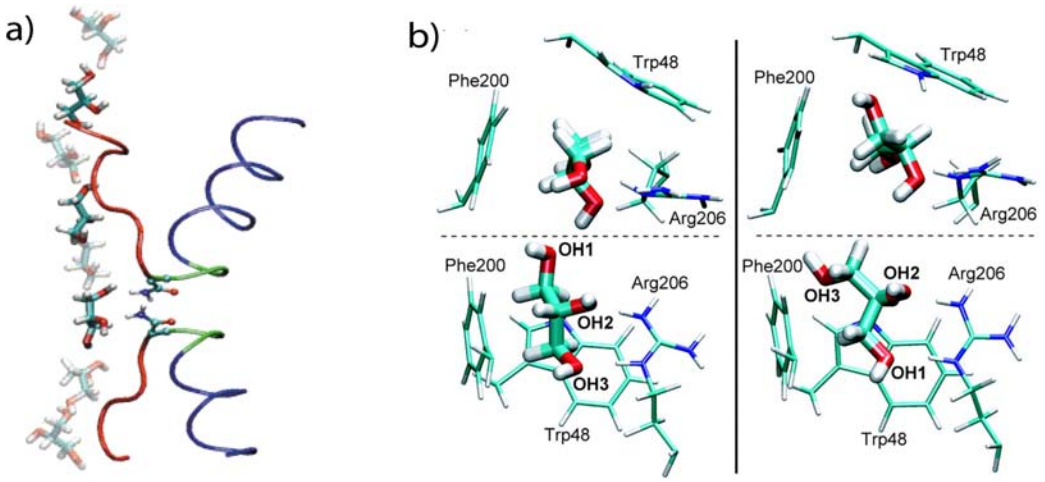


Figure 10: a) Complete glycerol conduction pathway in the GlpF channel [14]. The glycerol positions in the crystal structure are shown as solid molecules. b) Representative snapshots from simulations [15], demonstrating favorable (left) and unfavorable (right) orientations of glycerol at the SF, seen from the top (upper) and the side (lower).

5 Remaining Questions on AQPs

With the help of modern computation and visualization technologies, our understanding of the AQP family has been greatly enhanced after more than a decade's extensive study. However, there still remains some interesting questions.

Although AQPs generally appear as passive water channels, channel gating might actually play an important role in regulating the permeability of small substrates for several members of the AQP family. AQP4 phosphorylation by a protein kinase may lead to a reduction in water permeability [17]. Phosphorylation mediated gating of water pores has also been demonstrated in a plant AQP [18]. The closed water pore of the bovine lens AQP0 can be activated at specific pH value due to the pH sensitivity conferred by the histidine residues within each pore [19].

Experiments also suggested that the central pore, which is the fifth pore in an AQP tetramer, may conduct ions after being gated [20]. It is known that ion channels use the central pore of a tetramer as a pathway for gated ion flux. Interestingly enough, the tetrameric organization of AQPs is reminiscent of that in ion channels.

More and more discoveries have been made regarding AQPs in recent years, suggesting that the original concept of AQPs as an exclusively unregulated water channel may only be a starting point in understanding this protein family. AQPs are emerging as more sophisticated and regulated proteins with multiple functional capabilities.

References

[1] Mario Borgnia, Søren Nielsen, Andreas Engel, and Peter Agre. Cellular and molecular biology of the aquaporin water channels. *Annu. Rev. Biochem.*, 68:425–458, 1999.

[2] L.S. King and P. Agre. From structure to disease: The evolving tale of aquaporin biology. *Nat. Rev. Mol. Cell Biol.*, 5:687–698, 2004.

[3] Laxmikant Kalé, Robert Skeel, Milind Bhandarkar, Robert Brunner, Attila Gursoy, Neal Krawetz, James Phillips, Aritomo Shinozaki, Krishnan Varadarajan, and Klaus Schulten. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comp. Phys.*, 151:283–312, 1999.

[4] N. Uehlein, C. Lovisolo, F. Siefritz, and R. Kaldenhoff. The tobacco aquaporin NtAQP1 is a membrane CO_2 pore with physiological functions. *Nature*, 425:734–737, 2003.

[5] S. Hohmann. Osmotic stress signaling and osmoadaptation in yeasts. *Micro. Mol. Biol. Rev.*, 66:300–372, 2002.

[6] P. Agre, S. Sasaki, and J. Chrispeels. Aquaporins: a family of water channels proteins. *Am. J. Physiol. Renal Physio.*, 265:F461, 1993.

[7] O.S. Smart, J.M. Goodfellow, and B.A. Wallace. The pore dimensions of Gramicidin A. *Biophys. J.*, 65:2455–2460, 1993.

[8] J.B. Heymann and A. Engel. Structural clues in the sequences of the aquaporins. *JMB*, 295:1039–1053, 2000.

[9] M. Eastwood, C. Hardin, Z. Luthey-Schulten, and P. Wolynes. Evaluating protein structure-prediction scheme using energy landscape theory. *IBM J. Res. Dev.*, 45:475–497, 2001.

[10] Morten Ø. Jensen, Emad Tajkhorshid, and Klaus Schulten. Electrostatic tuning of permeation and selectivity in aquaporin water channels. *Biophys. J.*, 85:2884–2899, 2003.

[11] Fangqiang Zhu, Emad Tajkhorshid, and Klaus Schulten. Theory and simulation of water permeation in aquaporin-1. *Biophys. J.*, 86:50–57, 2004.

[12] N. Agmon. The grotthuss mechanism. *Chem. Phys. Lett.*, 244:456–462, 1995.

[13] Emad Tajkhorshid, Peter Nollert, Morten Ø. Jensen, Larry J. W. Miercke, Joseph O'Connell, Robert M. Stroud, and Klaus Schulten. Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science*, 296:525–530, 2002.

[14] Morten Ø. Jensen, Sanghyun Park, Emad Tajkhorshid, and Klaus Schulten. Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. USA*, 99:6731–6736, 2002.

[15] Sanghyun Park, Fatemeh Khalili-Araghi, Emad Tajkhorshid, and Klaus Schulten. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.*, 119:3559–3566, 2003.

[16] Morten Ø. Jensen, Emad Tajkhorshid, and Klaus Schulten. The mechanism of glycerol conduction in aquaglyceroporins. *Structure*, 9:1083–1093, 2001.

[17] M. Zelenina, S. Zelenin, A. A. Bondar, H. Brismar, and A. Aperia. Water permeability of aquaporin-4 is decreased by protein kinase C and dopamine. *Am. J. Physiol. – Ren. Physiol.*, 283(2):F309–F318, 2002.

[18] S. Törnroth-Horsefield, Y. Wang, K. Hedfalk, U. Johanson, M Karlsson, E. Tajkhorshid, R. Neutze, and P. Kjellbom. Structural mechanism of plant aquaporin gating. *Nature*, 2005. in press.

[19] KL Nemeth-Cahalan and JE Hall. pH and calcium regulate the water permeability of aquaporin 0. *J. Biol. Chem.*, 275:6777–82, 2000.

[20] Andrea J. Yool and Alan M. Weinstein. New roles for old holes: Ion channel function in aquaporin-1. *News Physio. Sci.*, 17:68–72, 2002.