Minireview

Selectivity and conductance among the glycerol and water conducting aquaporin family of channels

Robert M. Stroud*, David Savage, Larry J.W. Miercke, John K. Lee, Shahram Khademi, William Harries

Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143-2240, USA

Received 25 August 2003; accepted 1 September 2003

First published online 23 October 2003

Edited by Gunnar von Heijne, Jan Rydström and Peter Brzezinski

Abstract The atomic structures of a transmembrane water plus glycerol conducting channel (GlpF), and now of aquaporin Z (AqpZ) from the same species, *Escherichia coli*, bring the total to three atomic resolution structures in the aquaporin (AQP) family. Members of the AQP family each assemble as tetramers of four channels. Common helical axes support a wider channel in the glycerol plus water channel paradigm, GlpF. Water molecules form a single hydrogen bonded file throughout the 28 Å long channel in AqpZ. The basis for absolute exclusion of proton or hydronium ion conductance through the line of water is explored using simulations.

Key words: Channel; Transmembrane; Glycerol; Water; Aquaporin

1. Introduction

The aquaporin/aquaglyceroporin family (AQPs) regulate movement of water and aliphatic alcohols across cell membranes [1–3]. Remarkably the AQPs support a line of water that can move through the channel at close to diffusion limited rates, with an Arrhenius activation energy for AqpZ of only $E_a = 3.7$ kcal/mol, conducting water at rates close to 10^9 s^{-1}, yet they remain totally insulating to leakage of protons or ions that would be lethal to the cell if it happened.

AQPs are found throughout all kingdoms and in all species, emphasizing the essential nature of response to osmolarity, and the need for conductance of glycerol. Glycerol, a three carbon backbone, tri-alcohol is a key component of ~2/3 of all phospholipids, and an important metabolite. Inside the cell glycerol is immediately phophorylated by glycerol kinase, maintaining the inward gradient that drives inward glycerol flux. The water channels, and the glycerol+water channel subfamilies diverged in bacterial times, on the order of ~2 billion years ago [4] and the structures of representatives of a glycerol channel GlpF [5], and a water channel AqpZ (aquaporin Z) [6] from the same species, *Escherichia coli*, here allows an assessment of channel properties, uncluttered by species differentiation.

The roles of the AQPs are rapid adjustment of water and specific nutrients. However the roles are highly specialized such that there are 11 human AQPs, variously expressed in tissue-specific profiles. Of these AQP3, AQP7, AQP9 and AQP10 also conduct glycerol. These eukaryotic AQPs are often regulated by phosphorylation, pH, osmolarity or binding of other proteins or ligands [7,8]. Yet AQPs, while they support a hydrogen bonded line of water molecules, also exclude all ions including hydroxide, hydronium ions, and protons from conductance through the channel. These are essential to preservation of the electrochemical potential across the membrane [9]. This property, especially insulation against proton conductance, offers insight into the essential insulating property of all selective channels in biological membranes.

Simulation of conductance through AQPs is powerful in adding a quantitative element to ideas that emanate from the structure, such as the polarization of the central water molecule and its effect in polarizing all adjacent water molecules. It is also powerful in pre-evaluating mutagenesis as a biochemical test of the role of specific amino acids in function. Caution must always be exercised in interpreting molecular simulations. However, used as a complement to structural determination, functional determination and mutagenesis they can augment structural intuition. This seems to be reasonably applicable to AQPs since the atomic resolution structures of GlpF and AQP1 seem to be very stable around their crystal structures after extensive simulation times [10,11].

2. Materials and methods

The structure of AqpZ was determined to 2.5 Å by X-ray crystallography [6]. In this structure there are two differently surrounded tetramers, each lying on four-fold axes in the unit cell of space group P4, $a = 93.6$ Å and $c = 80.4$ Å, but rotated differently around the four-fold axis implying two independent observations of the structure of the AqpZ tetramer. This structure is compared with that of the glycerol channel GlpF to 2.2 Å resolution [5,12], and with bovine AQP1 to 2.2 Å resolution [13]. Site directed mutagenesis of GlpF was

*Corresponding author. Fax: (1)-415-476 1902.
E-mail address: stroud@msg.ucsf.edu (R.M. Stroud).

Abbreviations: GlpF, glycerol channel from *E. coli*; AqpZ, aquaporin Z from *E. coli*; AQPs, family of water conducting aquaporins and glycerol plus water conducting aquaglyceroporins

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)01195-5

used to probe the roles of amino acids that determine selectivity and conductance. Crystal structures of these mutant channels are compared with conductance measurements made using light scattering of reconstituted vesicles. Molecular mechanics and simulations on GlpF were carried out using Schulten's NAMD program [10,11].

3. Results and discussion

For ease of recognition, in referring to amino acid sequence numbers in other AQPs we always refer to the sequence number in GlpF, after alignment of sequence, as in Fig. 1A. Where residues differ between AqpZ and GlpF, they are de-

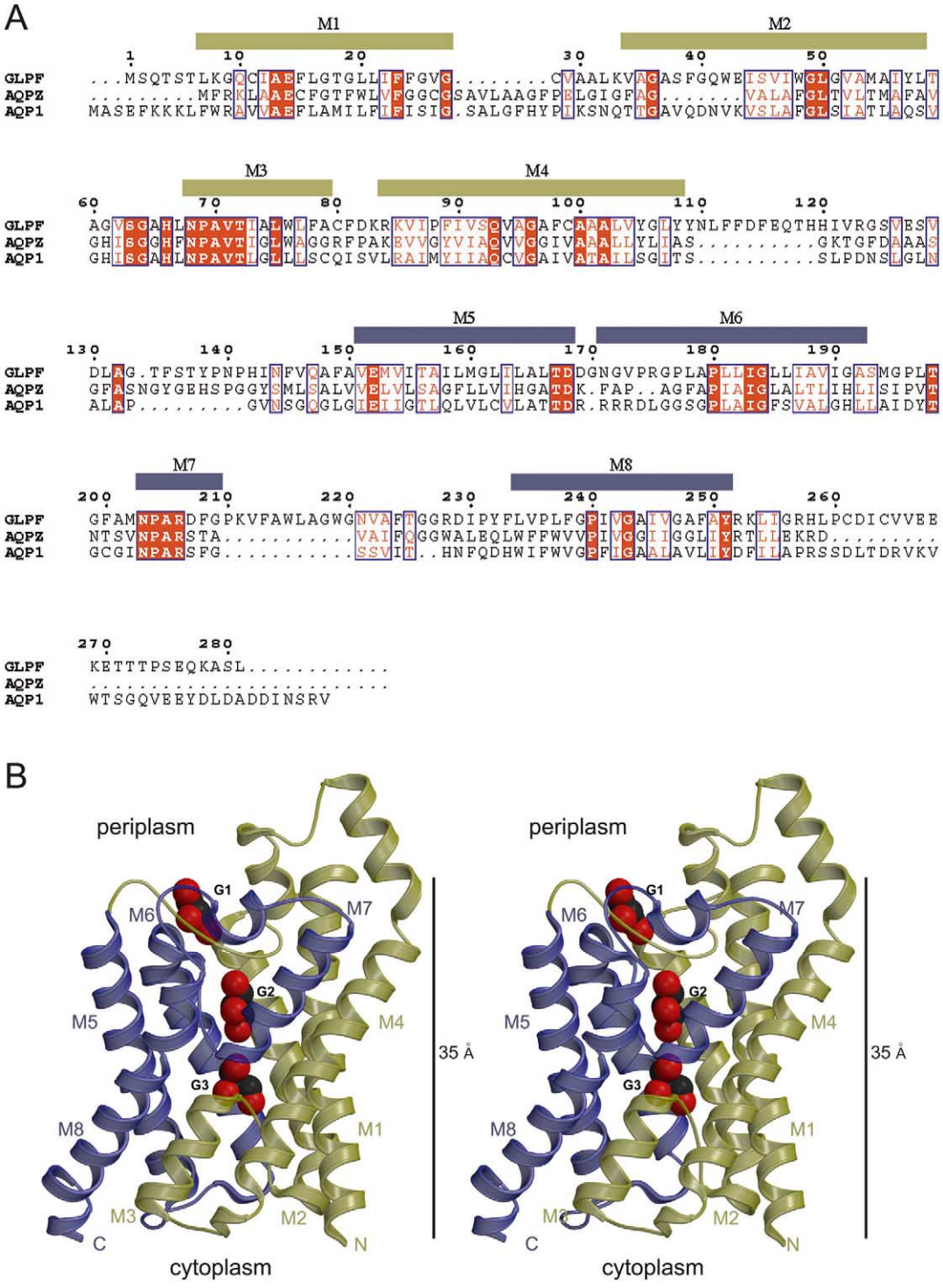


Fig. 1. A: Sequence alignment of GlpF, AqpZ, and AQP1. Color bars indicate the transbilayer helical regions M1–M8. In the N-terminal portion the bars are yellow for M1 to M4. The gene duplicated C-terminal equivalents are blue for M5 to M8. The numbering is for GlpF, and these numbers are used to describe residues in all sequences presented. B: Stereo figure showing GlpF viewed down the quasi-two-fold axis that relates M1 to M4 (yellow-green) to M5 to M8 (blue). Three glycerol molecules seen in the channel are indicated in space filling representation. The 35 Å bar indicates the transmembrane portion.

scribed by the correct amino acid name in AqpZ, and the number in GlpF as in Fig. 1.

3.1. Genetic duplication reflected in quasi-two-fold relationships in AQPs, and in other membrane proteins

Early gene duplication was easily recognized in the AQPs [14], and first seen as a quasi-two-fold axis in the atomic structure of GlpF [5]. The N-terminal residues 6–108 and C-terminal residues 144–254 [15] share ~20% conservation (Fig. 1). This is expressed as a two-fold symmetric arrangement of segments, related by a two-fold axis that lies in the mid-membrane plane (Fig. 1B). In Fig. 1A the transbilayer helices observed in the structure are indicated as yellow bars for the N-terminal segment, and as blue bars for the corresponding C-terminal segment. Once rotated around the quasi-two-fold axis the transbilayer helices related by the two-fold superimpose almost identically [12]. In the AQPs of known structure the quasi-two-fold axis lines up with the diagonal of the tetramer, suggesting that it was important to the assembly of the tetramer from the earliest times in evolution.

An in-plane quasi-two-fold symmetry axis may be quite common among membrane proteins. Four membrane protein families represented in this Nobel symposium series show structural two-fold relatedness where the two-fold axis lies in the plane of the bilayer. These include the AQPs, the chloride channel [16], the lac permease LacY [17], and the major protein of the translocon pore, secY. In all except the aquaporins, the two-fold relatedness in the structures of each monomer was not recognized in their sequences. This principle may be more common in membrane protein structures in general than previously anticipated.

3.2. Structure of the GlpF, AQP1 and AqpZ channels

Each molecule contains one channel at its center. The tetramers of four channels seen in GlpF, AQP1, and now AqpZ crystal structure are closely similar to the tetramers that associate, albeit alternate sides up, in reconstituted membranes with GlpF, AQP1 and AqpZ [18–24]. However there is no evidence to support any cooperativity between the monomeric channels in the tetramer. Therefore the tetramer may simply be to form a stable structural assembly. The monomers each have a hydrophobic interface with neighboring monomers not unlike the surface of a membrane protein in contact with lipids, suggesting that they might be stable initially upon synthesis as monomers, then come together to form the stable tetramer.

Surrounding each AQP channel are six and two half membrane spanning α-helices arranged in a right handed supertwisted fashion. In all cases the N- and C-termini are cytoplasmic [25]. With exception of the N- and C-termini most of the variations in sequence lengths and insertions or deletions are seen on the periplasmic, or outside side. In the glycerol conducting channels there is considerably more sequence on the outside between M4 and M5, and between M7 and M8 that encode a hydrophobic 'tongue' that may contribute to the attraction of glycerol or linear alditols to the lumen of the channel, from the outside.

The first segment (residues 6–108) forms one side of the channel. The second segment (144–254) reiterates the same transmembrane topology beginning from the periplasmic side rather than from the cytoplasmic side. In GlpF, the linking region between genetic repeats forms two helices (109–120 and 126–134) as protrusions on the periplasmic side (Fig. 1B). The two half-spanning helices M3 and M7 are both in contact with the lipid accessible exterior on the outside 'corner' of the tetramer, and form a unique junctional contact between N-terminal ends of the short helices. The NPA-contact is formed as the proline rings are in Van der Waals contact and cup each other between proline and alanine side chains of the opposite helix [5]. Their N-terminal ends related by the quasi-two-fold axis form a bend where their axes intersect. This unique contact maintains a key, double hydrogen bond donor site in the center of the bilayer, which is key to the function of all AQPs.

Of the 25 most conserved residues in all AQPs, six are conserved only in the first segment. These residues are all close packed together in a structural core of the first segment that is very close to and ultimately orients the key Asn 203 (of -NPA- of the second segment) in the center of the membrane bilayer. In the second genetic repeat, Arg 206 is an essential functional determinant whose side chain lines the hydrogen bond donor side of the tripathic channel, near its center and provides two donor hydrogen bonds to acceptor oxygens on the transported water or glycerol molecule. Several conserved glycines are seen near the center of the transbilayer region and facilitate the close crossing contact of helices, with correct geometry for forming C–H...O hydrogen bonds.

3.3. The channel in AQPs

The channel pathway in AQPs has a wide vestibule ~15 Å across on the outer surface, and reaches its constriction of ~3.8 Å×3.4 Å, 8 Å above the central plane of the transmembrane region (Fig. 2). This constriction lies at the start of a ~28 Å long selective channel ($r < 3.5$ Å) that subsequently widens out again to the cytoplasmic surface.

The entry and exit to the AQP channels are based on a line of four remarkable carbonyl oxygens on the periplasmic side, and the quasi-two-fold related four on the cytoplasmic of the channel (Fig. 3A). Together these provide hydrogen bond acceptors down one entire side of the channel, and separated by ~ 3.1 Å in distance down the channel axis. Entering from the cytoplasmic side, three carbonyls of adjacent amides in the sequence, Gly 64, Ala 65, His 66, are oriented into the cyto-

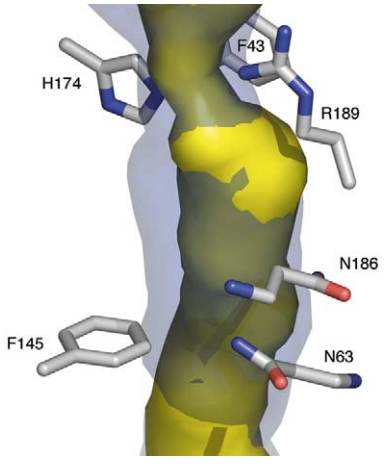
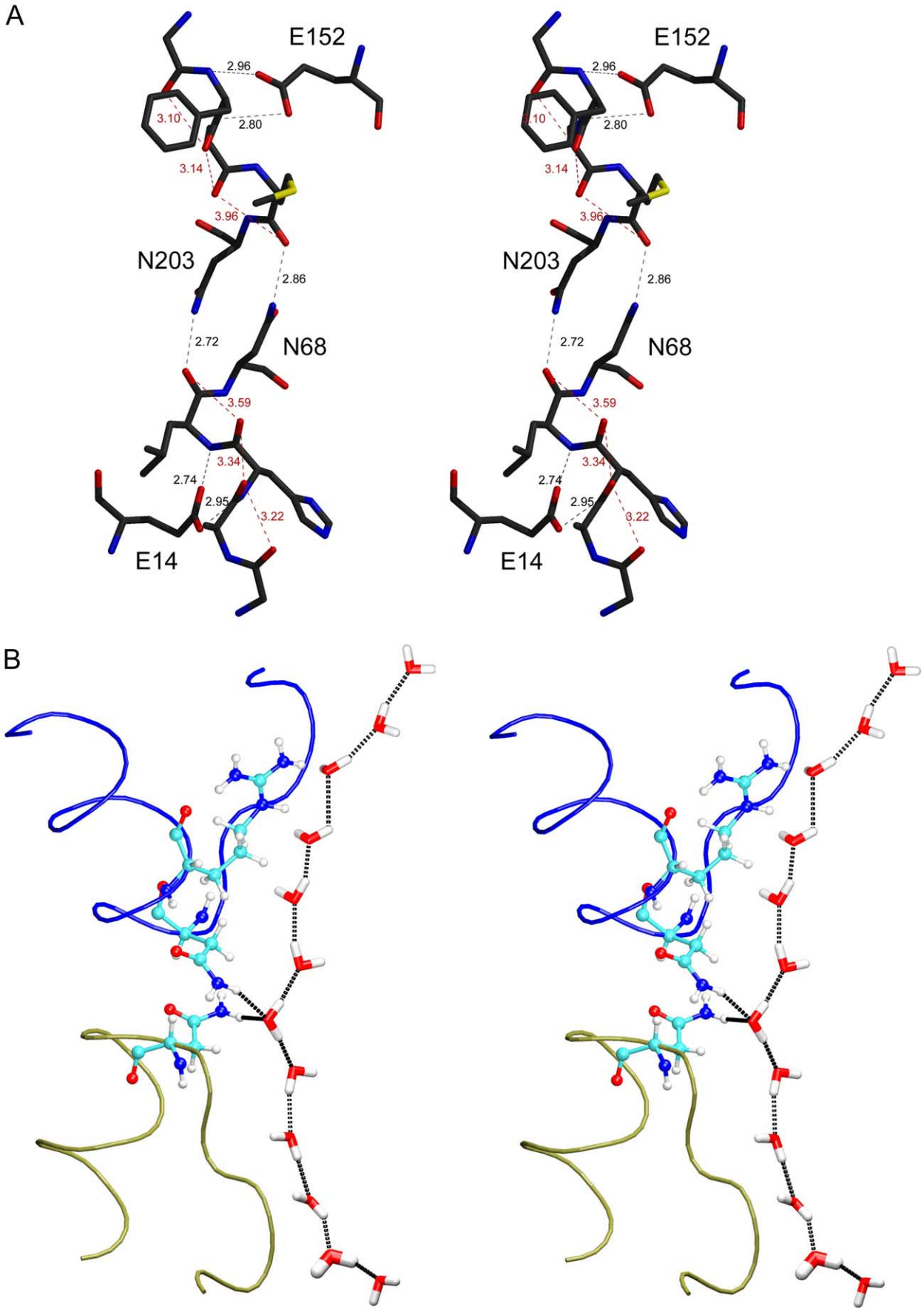


Fig. 2. A side view of the internal surface of the channel in GlpF (blue) AqpZ (yellow) and AQP1 (green) after superimposing the coordinates. This view emphasizes the larger diameter of the glycerol channel with respect to the water channels.


A
E152
2.96
3.10
2.80
3.14
3.96
N203
2.86
2.72
N68
3.59
3.34
2.74
2.95
3.22
E14
B

plasmic vestibule of the channel, while that of Leu 67 follows the same theme, but forms a hydrogen bond to the side chain of conserved Asn 203 (of the other NPA box). The conformation of four residues are alternately in the (usual) right handed α-helical (Gly 64), then the (unusual) left handed (Ala 65), then right handed (His 66) and then left handed (Leu 67) conformation. The four adjacent carbonyls point into, and act as hydrogen bond acceptors for the central line of water molecules. Conservation of sequence across the AQPs suggest that all the AQPs are assembled in this same way.

A similar structure is repeated from the periplasmic surface where Gly 199, (the conserved, duplicate analog of Gly 64) provides for the sharp turn of the chain into the lumen, followed by the carbonyls of Phe 200, Ala 201, Met 202, Asn 203 that line the periplasmic portal. Phe 200 and Met 202 provide key side chains to the hydrophobic wall of the narrow region of the channel.

Maintenance of these two 'ladders' of carbonyls is supported by the two conserved buried glutamates. Glu 14 forms hydrogen bonds to the amide NHs of His 66, Leu 67, stabilizing a pair of amides that are in the right handed configuration. Glu 152 plays the same (quasi-two-fold related) role with the NHs of Ala 201, Met 202. The carbonyl ladder provides adjacent hydrogen bond acceptor sites that are ~3.1 Å apart and line the pathway into and out of the channel.

3.4. Water conductance without any conductance of ions or protons

In AQPs, the conserved -Asn-Pro-Ala-signature sequences (-NPA-) are repeated in each of the two symmetry related segments. They form a key conserved interface between helices M3 and M7 that orients the side chains of both the Asn residues 68, 203. Each is doubly hydrogen bonded to highly conserved interface residues and together they provide the only two hydrogen bond donating groups from the protein, the donor NHs of the Asn side chains, to the central water molecule in the channel [5]. In AqpZ a central water molecule has shorter hydrogen bonds from these two donor hydrogen bonds than in GlpF, consistent with the generally smaller diameter channel in AqpZ. From the first structure of an AQP, for GlpF, we suggested that this ordered water might polarize all neighboring water molecules outward from the center.

This contention is supported by simulations that show that by polarizing the central water, the polarization permeates through the entire line of water molecules (Fig. 3B). We also argue that this polarization could help abrogate any proton conductance through the channel [11]. Five waters are seen in single file in the AqpZ channel, eight in the GlpF channel. The waters are hydrogen bonded to each other, and are hydrogen donors to the carbonyls. Thus there would appear to be no breaks in the water conducting line. The channels are not gated and are open. In the presence of high concentrations of glycerol, three molecules of glycerol are seen in transit 'intermediate' sites in GlpF, and removal of glycerol evokes only small side chain adjustments in the lining of the channel showing that they too do not close when a natural permeant molecule is replaced by water.

3.5. Selectivity properties of the GlpF versus AQP channels that do not conduct glycerol

GlpF is five-fold less conducting for water than the water channel AqpZ. In GlpF conductivity for water in the absence of glycerol may be impaired by the amphipathic nature of the channel, since the hydration shell of water itself is quite stringent, with many more neighbors than found in carbohydrates. Several waters pass the filter in single file and could not retain more than two neighbors in the center of any AQP channel. The geometric requirements for hydration of water are stringent, with as many as four to five coordinate waters [26,27].

At high concentration glycerol binds at the selectivity filter in GlpF, with the CH backbone lying in a corner on one side of the channel formed by two aromatic rings (Trp 48, Phe 200), and with the glycerol OH groups acting as acceptors for the guanidinium side chain of Arg 206, and as donors to two main chain carbonyl oxygens. Glycerol binds in the Trp-Phe-Arg triad leaving no free space around it, such that Van der Waals, hydrogen bond, and electrostatic forces each play a role. Water molecules and linear alcohols must pass through this channel in single file, and with at least one water molecule between each glycerol.

In AqpZ, the selectivity filter lies ~7 Å inside the periplasmic vestibule, and is again the narrowest point (diameter of ~1.5 Å) in the entire channel. It is formed by the side chains of Phe 43(48), His 174(191), Arg 189(206) and the carbonyl of Thr 183(200). In this case His 174(191), Thr 183(200), and Arg 189(206) create a hydrophilic triangle.

GlpF and all AQPs conduct water. The basis for specificity in conducting the linear tri-alcohol glycerol (CH_2OH–$CHOH$–CH_2OH) by the aquaglyceroporins, versus water (HOH) but not glycerol in the aquaporins is atomic. Most notably the channel diameter is larger in GlpF than in either AQP1 or AqpZ water channels (Fig. 2). This is especially true at the selectivity filter which lies opposite the guanidinium of Arg 206, and the planes of the aromatic rings of Phe 200, Trp 48 and Gly 191 in GlpF [5]. In water channels the last three side chains are generally substituted by Cys or Thr at 200, by Phe at 48, and by His at 191 as in AQP1 and AqpZ. The NPA region also plays a role in selectivity. AqpZ positions Leu 21 and Phe 159, residues from helices M1 and M5 respectively, near the NPA region are covariant. The aquaglyceroporins usually have leucine at these positions, while aquaporins have an aromatic at one of the two. In AqpZ, the side chains of Leu 21 and Phe 159 project into the pore and narrow it to a diameter of 3.0 Å. In GlpF the diameter is 4.0 Å and in AQP1 it is 3.5 Å.

3.6. AQP selectivity against passing ions or leaking an electrochemical gradient

Common to GlpF and AQPs is their absolute selectivity against passage of any ions or charged solutes. The channel is too small to accommodate a hydrated ion, and by its am-

←

Fig. 3. A: The structure of the polypeptide that presents the carbonyls of Gly 199 to Asn 203, and Gly 64 to Asn 68 into the lumen of the channel. This is for GlpF, though common to all three structures of AQPs discussed. B: A snapshot from a molecular dynamic simulation shows the polarization of water molecules as they are hydrogen bond donors to the carbonyls seen in A. The atomic centers of the water molecules are seen by X-ray crystallography, the hydrogen positions inferred by hydrogen bond distances seen in the X-ray structure, and from simulations.

phipathic nature provides no replacement for water of hydration on all but one side throughout the channel. Removal of water and fixation of a specific conformation of glycerol in the GlpF channel are also energetically costly. However the structure specifically matches the donor plus acceptor, and amphipathic nature of each of a series of C–OH groups, replacing the hydration shell in water by a precise match within the channel.

3.7. GlpF selectivity for glycerol and linear carbohydrates

GlpF is stereo- and enantioselective for linear alditols [5,28]. Stereoselectivity is seen between different chain lengths, and different chiralities. As an example we find a factor of 10 difference in rate between ribitol and its stereoisomer D-arabitol [5]. The structure shows how this occurs since two successive –CHOH groups will always be triangulated at the main 'G2 site' that lies opposite Arg 206. This in turn forces the next CHOH group to one side or the other in the channel, and these are inherently different: CHOH groups will alternately place the carbon in one of two tetrahedrally disposed sites that have quite different environments [5,28]. We have generated mutations in GlpF that seek to experimentally address the determinants of stereoselectivity, and water versus glycerol selectivity.

3.8. Comparison of the adjacent line of carbonyls with those in K^+ channels

The unusual arrangement of successive carbonyl oxygens that point into the lumen of the channel deserves comparison with the structure found in the KcsA potassium channel [29]. Unlike the AQPs, the KcsA channel is formed in the center of the tetramer. In each KcsA monomer, beginning from the extracellular side, the carbonyls of Tyr 78, Gly 77, Val 76, Thr 75, Thr 74 (together with the side chain OH of Thr 75) and their four-fold related equivalents form the selectivity filter in the center of the tetramer. This contributes 20 carbonyls and four OHs to the lumen of the channel. This backbone structure is maintained by residues that are alternately in the right, left, right, left and right handed α-helical configuration, and these residues are the highly conserved residues of the selectivity filter in potassium channels. The direction of the chain however is opposite to that in AQPs, in that advancing sequence numbers in AQPs encode the carbonyls as they enter the membrane from either surface, whereas in KcsA advancing sequence numbers come out toward the outside surface of the membrane. Thus there is no suggestion of any obvious evolutionary relatedness, in spite of a remarkably similar arrangement of alternately left and right handed helical geometry.

How is the structure in KcsA maintained? In this case the residues in left handed conformation include a Gly (77) and a Thr (75). Thus only Thr 75 lies in an unusual conformation. Interestingly, a buried carboxyl containing residue behind the amide NHs Gly 77 and Tyr 78 is Glu 71 in the KcsA channel. This glutamate may play the same role as do Glu 14, Glu 152 in the GlpF lumenae, in orienting the amide bridge. In the case of KcsA it is around the left handed helical glycine 77, rather than the right handed conformation in GlpF, however the side chain of Glu 71 is not seen in the initial structure, and its position is only suggested by us here because of its striking similarity in relative location to those of the Glu 14 and Glu 152 in AQPs. Glu 77, though conserved in several K^+ channels, including KcsA, romK, hgirK, is not generally conserved in K^+ channels. Gly 79, at the extracellular end of the carbonyl ladder has possibly an analogous role to that of the conserved Gly 64 and Gly 199 in GlpF, in allowing a sharp turn away from the central pore at the aqueous surface, though the glycines in GlpF are at the beginning in linear sequence rather than the end of the carbonyl ladder.

Acknowledgements: This work was supported by the National Institutes of Health (GM24485) to R.M.S.

References

[1] Preston, G.M., Carroll, T.P., Guggino, W.B. and Agre, P. (1992) Science 256, 385–387.
[2] Maurel, C., Reizer, J., Schroeder, J.I., Chrispeels, M.J. and Saier Jr., M.H. (1994) J. Biol. Chem. 269, 11869–11872.
[3] Park, J.H. and Saier Jr., M.H. (1996) J. Membr. Biol. 153, 171–180.
[4] Heymann, J.B. and Engel, A. (2000) J. Mol. Biol. 295, 1039–1053.
[5] Fu, D., Libson, A., Miercke, L.J., Weitzman, C., Nollert, P., Krucinski, J. and Stroud, R.M. (2000) Science 290, 481–486.
[6] Savage, D.F., Egea, P.F., Robles, Y.C., O'Connell, J.D. and Stroud, R.M. (2003) PLoS Biol., in press.
[7] Anthony, T.L., Brooks, H.L., Boassa, D., Leonov, S., Yanochko, G.M., Regan, J.W. and Yool, A.J. (2000) Mol. Pharmacol. 57, 576–588.
[8] Engel, A., Fujiyoshi, Y. and Agre, P. (2000) EMBO J. 19, 800–806.
[9] Finkelstein, A. (1987) Water Movement through Lipid Bilayers, Pores, and Plasma Membranes: Theory and Reality, Wiley, New York.
[10] Jensen, M.O., Park, S., Tajkhorshid, E. and Schulten, K. (2002) Proc. Natl. Acad. Sci. USA 99, 6731–6736.
[11] Tajkhorshid, E., Nollert, P., Jensen, M.O., Miercke, L.J., O'Connell, J., Stroud, R.M. and Schulten, K. (2002) Science 296, 525–530.
[12] Nollert, P., Harries, W.E., Fu, D., Miercke, L.J. and Stroud, R.M. (2001) FEBS Lett. 504, 112–117.
[13] Sui, H., Han, B.G., Lee, J.K., Walian, P. and Jap, B.K. (2001) Nature 414, 872–878.
[14] Pao, G.M., Wu, L.F., Johnson, K.D., Hofte, H., Chrispeels, M.J., Sweet, G., Sandal, N.N. and Saier Jr., M.H. (1991) Mol. Microbiol. 5, 33–37.
[15] Wistow, G.J., Pisano, M.M. and Chepelinsky, A.B. (1991) Trends Biochem. Sci. 16, 170–171.
[16] Dutzler, R., Campbell, E.B., Cadene, M., Chait, B.T. and MacKinnon, R. (2002) Nature 415, 287–294.
[17] Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R. and Iwata, S. (2003) Science 301, 610–615.
[18] Cheng, A., van Hoek, A.N., Yeager, M., Verkman, A.S. and Mitra, A.K. (1997) Nature 387, 627–6230.
[19] Hasler, L., Heymann, J.B., Engel, A., Kistler, J. and Walz, T. (1998) J. Struct. Biol. 121, 162–171.
[20] Li, H., Lee, S. and Jap, B.K. (1997) Nat. Struct. Biol. 4, 263–265.
[21] Hasler, L., Walz, T., Tittmann, P., Gross, H., Kistler, J. and Engel, A. (1998) J. Mol. Biol. 279, 855–864.
[22] Ringler, P., Borgnia, M.J., Stahlberg, H., Maloney, P.C., Agre, P. and Engel, A. (1999) J. Mol. Biol. 291, 1181–1190.
[23] Walz, T. et al. (1997) Nature 387, 624–627.
[24] Mitsuoka, K., Murata, K., Walz, T., Hirai, T., Agre, P., Heymann, J.B., Engel, A. and Fujiyoshi, Y. (1999) J. Struct. Biol. 128, 34–43.
[25] Preston, G.M., Smith, B.L., Zeidel, M.L., Moulds, J.J. and Agre, P. (1994) Science 265, 1585–1587.
[26] Bernal, J.D. and Fowler, R.H. (1933) J. Chem. Phys. 1, 515–520.
[27] Eisenberg, D. and Kauzman, W. (1969) The Structure and Properties of Water, Oxford University Press.
[28] Heller, K.B., Lin, E.C. and Wilson, T.H. (1980) J. Bacteriol. 144, 274–278.
[29] Zhou, Y., Morais-Cabral, J.H., Kaufman, A. and MacKinnon, R. (2001) Nature 414, 43–48.