Electrostatics of aquaporin and aquaglyceroporin channels correlates with their transport selectivity

Romina Oliva[a,b,1], Giuseppe Calamita[c], Janet M. Thornton[b], and Marialuisa Pellegrini-Calace[b,2]

[a]Department of Applied Sciences, University "Parthenope" of Naples, Centro Direzionale Isola C4, I-80143, Naples, Italy [b]European Bioinformatics Institute, The Wellcome Trust Genome Campus, CB10 1SD, Hinxton, Cambridge, United Kingdom [c]Department of General and Environmental Physiology, University of Bari, Via Amendola 165/A, I-70126 Bari, Italy

Edited by Robert M. Stroud, University of California, San Francisco, San Francisco, Ca, and approved January 14, 2010 (received for review September 22, 2009)

Aquaporins are homotetrameric channel proteins, which allow the diffusion of water and small solutes across biological membranes. According to their transport function, aquaporins can be divided into "orthodox aquaporins", which allow the flux of water molecules only, and "aquaglyceroporins", which facilitate the diffusion of glycerol and other small solutes in addition to water. The contribution of individual residues in the pore to the selectivity of orthodox aquaporins and aquaglyceroporins is not yet fully understood. To gain insights into aquaporin selectivity, we focused on the sequence variation and electrostatics of their channels. The continuum Poisson-Boltzmann electrostatic potential along the channel was calculated and compared for ten three-dimensional-structures which are representatives of different aquaporin subfamilies, and a panel of functionally characterized mutants, for which high-accuracy three-dimensional-models could be derived. Interestingly, specific electrostatic profiles associated with the main selectivity to water or glycerol could be identified. In particular: (*i*) orthodox aquaporins showed a distinctive electrostatic potential maximum at the periplasmic side of the channel around the aromatic/Arg (ar/R) constriction site; (*ii*) aquaporin-0 (AQP0), a mammalian aquaporin with considerably low water permeability, had an additional deep minimum at the cytoplasmic side; (*iii*) aquaglyceroporins showed a rather flat potential all along the channel; and (*iv*) the bifunctional protozoan PfAQP had an unusual all negative profile. Evaluation of electrostatics of the mutants, along with a thorough sequence analysis of the aquaporin pore-lining residues, illuminated the contribution of specific residues to the electrostatics of the channels and possibly to their selectivity.

channel proteins | electrostatic potentials | pore analysis | comparative analysis

Aquaporins are homotetrameric channel proteins with each monomer defining a single pore that appears to have evolved for the selective passage of water and glycerol, although recently a number of other small molecules have been shown to diffuse through specific aquaporins (reviewed in (1)). Discovered in all the domains of life, from bacteria to mammals, with few exceptions they are all characterized by two conserved Asn-Pro-Ala (NPA) sequence motifs. The conserved NPA motifs are located, with the two Asn side-chains pointing into the pore, at the end of two half helices which lie at the middle of the permeation channel, where they form a constriction. Another narrower constriction, known as aromatic/Arg (ar/R) selectivity filter, is found at the periplasmic side of the channel and is formed by four residues including aromatic residues and a widely conserved Arg.

Aquaporins have been grouped by phylogenetic analyses into about 30 major subfamilies, characterized by their specific transport selectivity and efficiency, source, localization, and physiological function (2). According to their main transport selectivity, aquaporin subfamilies can be further classified into "orthodox" aquaporins, which are mainly water channels, and aquaglyceroporins, which additionally allow the transit of small organic compounds like glycerol and urea. The number of aquaporin subfamilies per species is extremely variable. In mammals and higher plants several orthodox aquaporins and glycerol channels are found (3, 4). Yeast genomes may encode either one/two orthodox aquaporins and glycerol channels or one single orthodox aquaporin, like the Aqy1 from *Pichia pastoris* (5). *Escherichia coli* has two aquaporins, the orthodox AqpZ (6) and the glycerol channel GlpF (7), whereas in archaea genomes only one aquaporin gene is found at most, such as AqpM in the *Methanothermobacter marburgensis*, which appears to be much more permeable to glycerol than to water (8). Interestingly, in the protozoan *Plasmodium falciparum* a single bifunctional aquaporin, PfAQP, is able to conduct both water and glycerol at high rates (9).

Since their discovery in the early 1990s (10), many efforts have been made to characterize the selectivity mechanism of the different aquaporin subfamilies and to explain the exclusion of protons, which are expected to shuttle rapidly across the hydrogen bond network of water molecules according to the Grotthuss mechanism. Prompted by the elucidation of the first experimental three-dimensional-structures of aquaporin-1 (AQP1) (11, 12) and the glycerol facilitator GlpF (13), valuable computational studies have contributed important insights into the dynamics and energetics of water and glycerol conduction in aquaporins (14, 15). The NPA and ar/R constriction regions have been shown to act as filters, and water molecules have been shown to cross the channel in single file oriented in opposite directions in the two halves of the channel, with their hydrogen atoms pointing toward the nearest exit. The conclusion of extensive molecular dynamics/quantum mechanical simulations is that electrostatic effects, both from the opposite dipoles of the two half helices at the middle of the channel and from the channel-lining residues (especially the Arg from the Ar/R site), dominate the water orientation observed in aquaporin channels (reviewed in (16)). The interruption of the H-bond network due to the bipolar water orientation, the electrostatic effects, and the ion dehydration free energy penalty have then been hypothesized in various papers to restrict the proton shuttling (15–19). Explicit calculations of the channel proton permeation free energy profile for wild-type AQP1 (20) and the H180A/R195V proton conducting AQP1 double mutant (21) have shown conclusively that all of these effects are important to varying degrees, with the electrostatics of the selectivity filter region (especially the R195 residue) and the opposing helix dipoles being the most important.

As for the selectivity to either water or glycerol, a comparative steered molecular dynamics study of the *E. coli* AqpZ and GlpF, an orthodox aquaporin and a glycerol channel respectively,

Author contributions: R.O., G.C., J.M.T., and M.P.-C. designed research; R.O. performed research; R.O., G.C., J.M.T., and M.P.-C. analyzed data; R.O., J.M.T., and M.P.-C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: romina.oliva@uniparthenope.it.

[2]To whom correspondence may be addressed. E-mail: marial@ebi.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0910632107/DCSupplemental.

directly addressed the issue, the main result being that energy barriers against the permeation of glycerol in the orthodox AqpZ are steric in nature, due to its narrower pore size as compared to GlpF, especially at the ar/R selectivity filter but also overall along the channel (22). Therefore large solutes are sterically excluded in orthodox aquaporins. More recently, Hub & de Groot applied molecular dynamics (MD) to determine permeation barriers of water, glycerol, ammonia, and other small neutral solutes through human AQP1 and *E. coli* GlpF, as representatives of all aquaporins and glycerol channels, aiming to derive a unifying picture of the aquaporins selectivity. Very interestingly, they found that the hydrophobicity of small solutes is remarkably anticorrelated with their permeability in the orthodox AQP1 but not in the GlpF glycerol channel (23). This suggests that the solute transport is driven in AQP1 by electrostatic features and calls for a thorough exploration of the amino acid composition and electrostatic features of the different aquaporin channels.

Currently, experimental three-dimensional-structure representatives of ten different aquaporin subfamilies are available, that include water and glycerol channels, the low water conductance AQP0, and the bifunctional PfAQP. This variety of aquaporin structures represents a valuable resource for understanding the structure-function relationships within this protein family. In fact, notwithstanding the wealth of information coming both from computational studies and from the effective approach based on the in vitro functional characterization of targeted aquaporin mutants (reviewed in (24)), to date the contribution of individual residues in the channel to the pore selectivity is not fully understood. To contribute and explain the different physiological selectivity of the large duplicated family of aquaporins, here we focused on the sequence variation of their channels and relative electrostatics. In particular, the electrostatic potential along the channels was calculated and compared for the ten representative aquaporin structures and a number of functionally characterized mutants (25, 26) for which high-accuracy three-dimensional-models could be derived. Although electrostatic effects are only one of the factors contributing to the energy barrier for water/solute diffusion across the aquaporin channels, here we show that specific electrostatic profiles are associated with the two main transport functions of aquaporins (i.e. water and/or glycerol) and, on the basis of a comprehensive sequence and structure analysis, can illuminate the role of individual pore-lining and "second-shell" residues in affecting the channel electrostatics and possibly selectivity.

Therefore, electrostatic profiles of aquaporin channels are potentially predictive for the selectivity of new uncharacterized aquaporins, and may be used to assist the design of unique aquaporin mutants.

Table 1. The ten aquaporin representative three-dimensional-structures.

PDB-code	Subfamily	Source	Resolution (Å)	Main selectivity
1fx8 (13)	GlpF	*Escherichia coli*	2.20	Glycerol, Water (low) (7)
1j4n (40)	AQP1	*Bos taurus*	2.20	Water (29)
1rc2 (41)	AqpZ	*Escherichia coli*	2.50	Water (6)
1ymg (32)	AQP0	*Bos taurus*	2.24	Water (low) (29)
1z98 (28)	PIP2	*Spinacia oleracea*	2.10	Water (42)
2f2b (43)	AqpM	*Methanothermobacter marburgensis*	1.68	Glycerol, Water (low) (8)
2w2e (5)	Aqy1	*Pichia pastoris*	1.15	Water (5)
3c02 (44)	PfAQP	*Plasmodium falciparum*	2.05	Water, Glycerol* (9)
3d9s (45)	AQP5	*Homo sapiens*	2.00	Water (29)
3gd8 (46)	AQP4	*Homo sapiens*	1.80	Water (29)

*Sugar alcohols up to five carbons and urea have also been shown to permeate the pore (9).

Results

Selection of Aquaporin Three-Dimensional-Structures. Ten three-dimensional-structures with resolution between 1.15 and 2.50 Å were selected from the Protein Brookhaven Data Bank (wwPDB, (27)) as representatives of ten aquaporin and aquaglyceroporin subfamilies: the mammalian AQP0, AQP1, AQP4, and AQP5 (PDB-code: 1ymg, 1j4n, 3gd8, 3d9s, respectively), the plant PIP2 (PDB-code 1z98), the yeast Aqy1 (PDB-code 2w2e), the protozoan (*P. falciparum*) PfAQP (PDB-code 3c02), the eubacterial AqpZ and GlpF (PDB-codes 1rc2 and 1fx8), and the archaeal AqpM (*M. marburgensis*; PDB-code 2f2b) (see Table 1). Both plant PIP2 (1z98) and yeast Aqpy1 (2w2e) channels are closed at the cytoplasmic side, the former because of a rearrangement of loop D, the latter closed by the N terminus, with Tyr31 occluding the channel entrance. For PIP2, however, a lower resolution structure is also available in its open form (PDB-code: 2b5f, resolution 3.9 Å (28)).

Among the ten three-dimensional-structures in the set, six are of orthodox aquaporins, AQP1, AQP4, AQP5, PIP2, AqpZ, and Aqy1, one of a low water conductance aquaporin, AQP0 (29, 30), and three of aquaglyceroporins, GlpF, AqpM, and PfAQP. They all have water molecules observed throughout the channel, while the three aquaglyceroporins also have glycerol molecules. It is worth highlighting here that, whereas GlpF and AqpM have a low efficiency in transporting water (7, 8), PfAQP is a particularly interesting aquaporin since it can conduct both water and glycerol at high efficiency rates (9). Moreover, PfAQP is also to date the only aquaporin of known three-dimensional-structure with mutations in both the NPA *consensus* motifs.

Protein Electrostatics Inside the Pore at Water and Glycerol Crystallographic Positions. Continuum Poisson-Boltzmann electrostatic potentials for the ten representative structures were calculated with the APBS program (31). The electrostatic profile along the solute pathway was analyzed by evaluating the electrostatic potential at the experimental positions occupied by water and glycerol molecules inside the channel of each PDB structure, as shown in Fig. 1.

Low efficiency AQP. The channel electrostatic potential of the low water conductance AQP0 has a strong positive trend in the periplasmic half, very similar to that of orthodox aquaporins, and of AQP1 in particular (see below). In the cytoplasmic half an unusual trend was found, with negative values reaching a minimum of −15 kT/e (Fig. 1*G*). This could be assigned to a tyrosine specific to the AQP0 subfamily (Y149, see below) that protrudes into the channel on the cytoplasmic side, forming a narrow constriction in addition to the ar/R site (32).

Orthodox aquaporins. The analyzed orthodox aquaporins (Fig. 1*A–F*) show a similar electrostatics profile all along the channel, with a fairly flat area usually with slightly negative values (< -6 kT/e) in the cytoplasmic half and a strong positive trend in the periplasmic half, with a peak at the ar/R constriction site, shifted toward the NPA site in the case of AqpZ (maximum values between 5 and 20 kT/e). Yeast Aqy1 whose channel in the analyzed structure is closed to the water passage with Y31 occluding its cytoplasmic entrance, has instead deep negative electrostatic values at the cytoplasmic side, resulting in a bipolar electrostatic profile similar to that of the low conductance AQP0. This is not surprising as the blocking of the Aqy1 channel by a Tyr residue has been remarked to be reminiscent of AQP0 (5).

Aquaglyceroporins. The two aquaglyceroporins GlpF and AqpM show a rather "neutral" channel, with electrostatic potential, calculated at water and glycerol experimental positions in the core

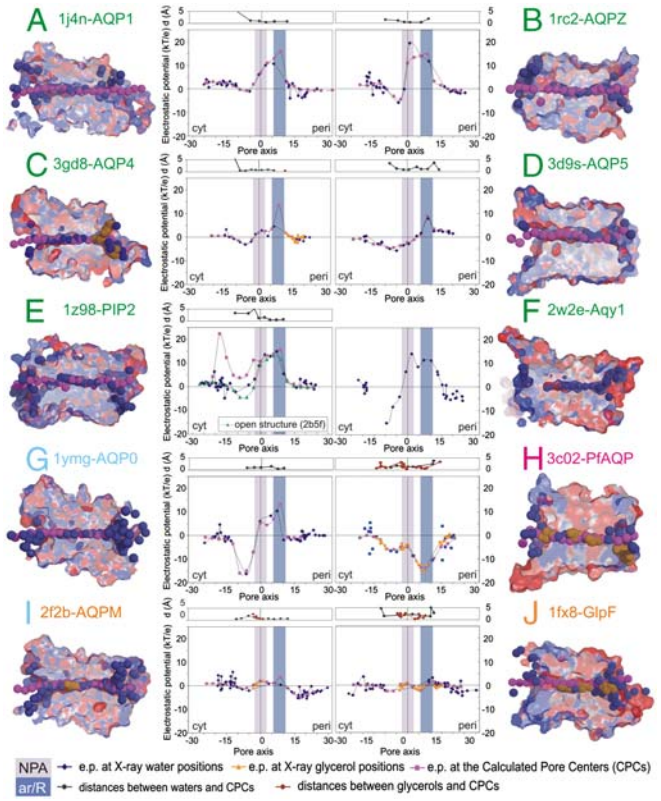

Fig. 1. Electrostatic potentials into the channels, calculated at the crystallographic positions of water/glycerol molecules and at the pore centers predicted by the PoreWalker program (33). Orthodox aquaporins are labeled in *green*, the low water conductance AQP0 in *cyan*, aquaglyceroporins in *orange*, and the bifunctional PfAQP in *magenta*. Sections of the corresponding PDB structures are shown colored according to the electrostatic potentials. X-ray water and glycerol atoms inside the channels are shown as *deep blue* and *orange surfaces*, respectively. Dummy atoms at the geometrical centers of the pores are shown as *magenta surfaces*.

region, never higher than 3.5 kT/e (Fig. 1*I–J*). In contrast, the PfAQP channel shows very distinctive electrostatic features. In fact, its amino acid composition results in negative electrostatic potential all along the channel, with a moderate relative minimum of −7 kT/e on the cytoplasmic side and an absolute minimum in the ar/R site, where the potential falls to ∼ − 15 kT/e (Fig. 1*H*). Interestingly, unlike orthodox aquaporins and AQP0, the electrostatic peaks observed for PfAQP do not correspond to particularly narrow regions of the pore. Note also that the potentials calculated at both water and glycerol positions follow the same trend, except for the position of one single water molecule at the level of the ar/R site for which a surprisingly negative potential (−29 kT/e) is found. Visual inspection showed that the water molecule is H-bonded to both E28 and the ar/R R197, with an average distance of 3.8 and 4.4 Å from the carboxylate and from the guanidinium group, respectively.

Protein Electrostatics at the Geometric Centers of the Pore. The channel axis and geometrical centers of the pore in 3 Å steps were predicted for the representative structures using PoreWalker (33), and relative electrostatic profiles were derived and compared with those obtained at the observed "solute" crystallographic positions (Fig. 1). This comparison was made to assess whether the method can be confidently extended to experimental three-dimensional-structures where water/solute molecules are not visible. As Aqy1 is occluded at the cytoplasmic side, with a gap of over 9 Å between two consecutive water molecules, it was not possible to define the pore centers all along its channel.

Distances between experimental and predicted positions are in the range from 0.1 to 2.6 Å both for orthodox aquaporins and aquaglyceroporins, at least in the channel region where solute molecules are forced to be in single file (deviations may rise to 5 Å when approaching the membrane surfaces) (Fig. 1). Our results show that the calculated electrostatic features inside the channel are not significantly affected by 1–2 Å shifts in the positions used for the measurements of the potential. In fact all the electrostatic profiles obtained at predicted positions are overall very similar to those calculated at crystallographic positions, with PIP2 being the only exception. Therefore correct qualitative trends of the electrostatic potentials can be also obtained by using the geometrical centers of the pore. However, position shifts may well affect the electrostatics experienced in the case of very narrow regions of the channel, especially if ionisable groups are involved, like for the ar/R selectivity filter.

As for PIP2, the analyzed structure (1z98) represents PIP2 in its "closed" state, with the cytoplasmic D loop reoriented to provide a gating mechanism (28, 34). Since there is no clear way out of the pore at this side of the channel, the path identified by PoreWalker corresponds to a side gateway. This led to a discrepancy between the location of the geometric centers predicted by PoreWalker and the experimental positions of water molecules and consequently between the electrostatic values observed in the cytoplasmic loops region. In fact, electrostatic potentials calculated at the PoreWalker geometric centers for the PIP2 structure in its open form (PDB-code 2b5f, resolution 3.9 Å) match correctly values calculated at the water crystallographic positions as in the "closed" channel (Fig. 1*E*).

Electrostatic Potentials of Aquaporin Mutants. The expression of targeted mutants and their experimental characterization is a very effective approach to the study of aquaporin selectivity (24). We therefore calculated the channel electrostatic profile for a number of in vitro characterized AQP1 and PfAQP mutants, in the attempt to correlate changes in the electrostatics with changes in the observed permeabilities. Moreover, the effect on the channel electrostatics of pore-lining residues specific to AQP0 was investigated by targeted *in silico* residue substitutions. They are all reported in Fig. 2. In Fig. 3 a sequence and structure representation of the pore-lining residues in the ten analyzed structures is given, where residues mutated and/or discussed in the following are labeled (see also Table S1 for details on the analyzed mutants and Fig. S1).

rAQP1. Recently, individual or joint mutations of residues F56, H180, and R195 in the selectivity filter of rat AQP1 have been shown to lower its selectivity experimentally (26). In particular the substitution of R195 was unexpectedly shown to allow the passage of protons. Rat AQP1 is 90% identical in sequence to the corresponding bovine homolog, of which the three-dimensional-structure is known (1j4n). Comparative models were thus built for the wild-type rat AQP1 (wt-rAQP1) and all the mutants functionally characterized by Beitz and colleagues (26): H180A, R195V, F56A/H180A, and H180A/R195V. Models were characterized by PoreWalker and electrostatic potentials were calculated at the geometric centers of the pore (Fig. 2*A*).

As expected, the electrostatic profile obtained for wt-rAQP1 is almost identical to that obtained for the *Bos taurus* 1j4n-AQP1 (Figs. 1*A* and 2*A*). The H180A and F56A/H180A mutants also show a similar electrostatic profile but with positive values in the periplasmic region almost halved as compared to the wt protein. On the other hand, a dramatic change is observed in the electrostatic features of the R195V and H180A/R195V mutants. In particular, the R195V electrostatic potential is mostly negative in the periplasmic region of the channel, reaching a minimum of −8.8 kT/e at the ar/R selectivity filter, whereas the double mutant H180A/R195V shows an electrostatic profile completely

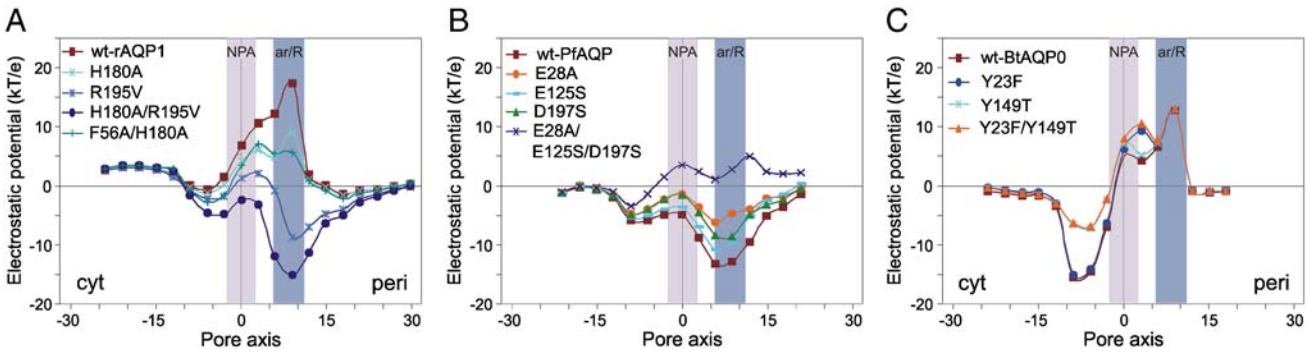

Fig. 2. Electrostatic potentials calculated at pore centers predicted by PoreWalker (33) for: *A*) AQP1 mutants, *B*) PfAQP mutants, and *C*) AQP0 mutants.

opposite to that of the wt protein, with negative values all along the periplasmic side of the pore and a peak value of −15.2 kT/e at the selectivity filter. The double mutation affects the whole transmembrane region of the channel with a particularly strong effect around the selectivity filter, where the change in the electrostatic potential values ranges from 13 to 32 kT/e. Electrostatic potentials were also calculated for the N76D/H180A/R195V and N192D/H180A/R195V triple mutants (see Fig. S3 and *SI Text*).

PfAQP. In the attempt to dissect the permeability of the protein to water and glycerol, Beitz & colleagues mutated some PfAQP pore-lining amino acids and functionally characterized relative mutants in oocyte swelling assays (25). Among the analyzed mutants, the E125S was particularly interesting since it largely abolished permeability to water. The E28A mutation was also shown to reverse the PfAQP swelling rates in favor of glycerol.

A comprehensive analysis of the sequences and structures of aquaporin subfamilies shows that E125 is unique to PfAQP among all aquaporin sequences available in the UniProt Knowledge database (35), whereas E28 is also found in the structure-unknown plant TIP3 subfamily. Comparative analysis of the pore amino acid composition also shows that one more PfAQP acidic residue, D197, which is adjacent in sequence to the ar/R conserved Arg and whose mutation has not yet been functionally tested, is characteristic of aquaglyceroporins. An Asp at this position is indeed found in the eubacterial GlpF and in all the available mammal aquaglyceroporin sequences (AQP3, AQP7, AQP9, and AQP10), whereas it is missing in all the available orthodox aquaporin sequences, where it is mostly substituted by a Ser.

Targeted in *silico* mutations of one or more of these residues were performed to investigate how their negative charges can affect the unusual all-negative PfAQP channel electrostatics. The modeled mutants are: (*i*) E28A, E125S, and E28A-E125S, where the negatively charged amino acids were replaced by the corresponding AqpM and GlpF counterparts, respectively, and (*ii*) D197S and E28A-E125S-D197S, where D197 was replaced by the AQP1 Ser counterpart (Table S1). Electrostatic potentials calculated at predicted pore centers for E28A, E125S, and D197S PfAQP show that each mutation makes the negative electrostatic potential closer to neutrality along the whole channel, with maximum changes of 8 kT/e (Fig. 2*B*). Mutation of both the negatively charged residues (E28A-E125S) smoothed even more strongly the negative potential leading to a trend that is much closer to that of GlpF (Fig. S2). Very interestingly, the additional replacement of D197 (E28A-E125S-D197S mutant) transformed the negative profile into a positive orthodox-like one (Fig. 2*B*), suggesting that this aspartate plays a pivotal role in determining the selectivity of aquaglyceroporins.

Finally, since PfAQP is the only protein in the set where the canonical NPA motifs in the middle of the channel are replaced by one NLA and one NPS motif, the double mutant L71P-S195A was also modeled to restore the canonical NPA motifs. The mutant has an electrostatic profile almost identical to that of the wt protein (Fig. S2), showing that these substitutions, at least in the case of PfAQP, do not affect the electrostatics of the channel.

AQP0. AQP0 is characterized by two highly specific tyrosine residues lining the pore walls (Y23 and Y149) that are very well conserved within the AQP0 family and are not found in any other known aquaporin (Fig. 3*G*). To test the influence of these residues on the electrostatics of the channel, we modeled three AQP1-like AQP0 mutants where they were replaced by the corresponding AQP1 residues: Y23F, Y149T, and Y23F/Y149T (Table S1). Resulting electrostatic potentials within the pore are shown in Fig. 2*C*. The effect of the Y23F and Y149T substitutions seems to be selectively experienced in the periplasmic and cytoplasmic half of the channel, respectively. Y23F substitution causes a moderate increase (5 kT/e) of the positive potential at the periplasmic side, between the NPA and the ar/R constriction sites, whereas Y149T substitution significantly decreases (by 8–9 kT/e) the negative potential in the periplasmic region of the channel, thus restoring an AQP1-like electrostatic profile.

It is interesting that in the closed form of Aqy1, which has an electrostatic profile similar to AQP0, Y31 occupies a channel area that roughly corresponds to that occupied by Y149 in AQP0 (see Fig. 3*F*, *G*). A tyrosine corresponding to Y31 is only additionally found in some NIP2, and the close-in-space Y104 is also unique to Aqy1 among the sequence-known aquaporins.

Discussion

Channel Electrostatics Emerges as a Distinctive Feature Accounting for the Selectivity of Aquaporins and Aquaglyceroporins. Electrostatic effects are obviously only one of the factors contributing to the energy barrier for the water and solute diffusion through the aquaporin channels (14–21, 36, 37). However, several studies have established the important role of the electrostatic field generated by the protein matrix in determining aquaporin selectivity, especially as concerns proton exclusion, although dehydration penalty also plays an important role (15–21). Herein, on the basis of the channel electrostatic potentials calculated for ten aquaporin structural representatives of different subfamilies and different transport functions, we show that each aquaporin exhibits an electrostatic profile closely associated to its selectivity to water and/or glycerol.

The electrostatic profile calculated here for bovine AQP1 matches that obtained by other authors from MD simulations (37) and is consistent with the electrostatic barriers found for proton diffusion in AQP1 by several methods (reviewed in (16) and more recently discussed in (20, 21)). The observed lack of a static electrostatic barrier at the level of the ar/R filter of GlpF agrees with previous MD studies (17, 18), although a barrier is instead found in the overall free energy for proton permeation (20). Finally, the electrostatic features of AQP1 and GlpF converge with the results of a recent MD analysis showing that, for small molecules, the more polar the solute, the higher its permeability in the orthodox AQP1, but not in the GlpF glycerol channel (23).

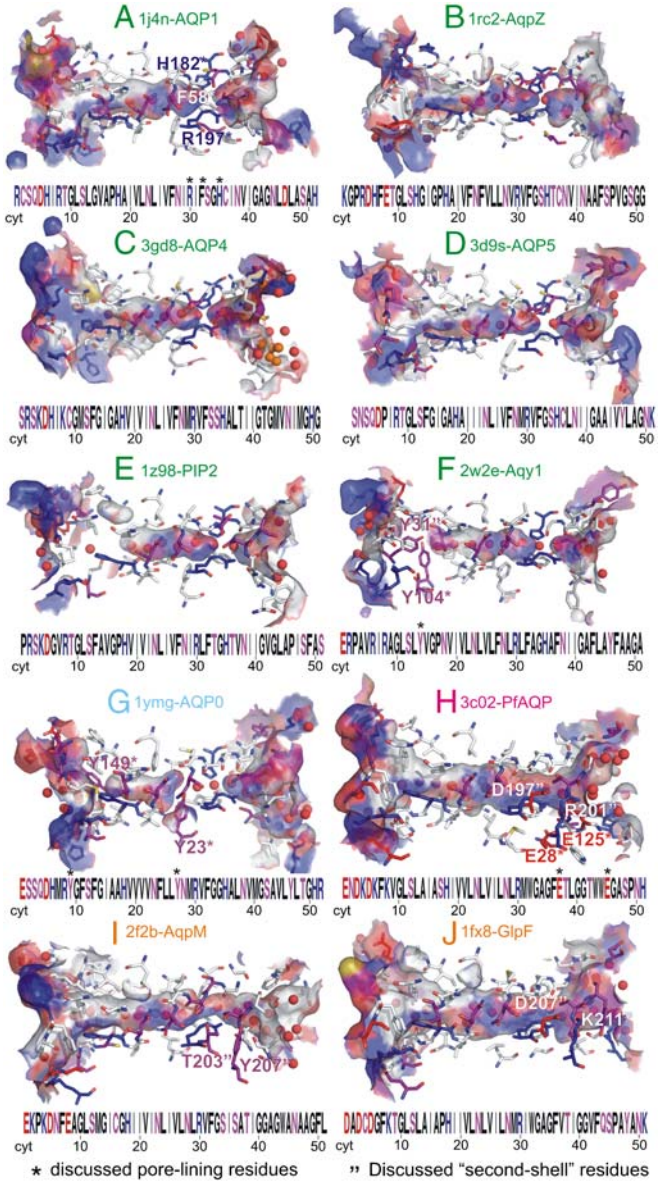

Fig. 3. Representation of sequence and structural features of the pores. In each quadrant, the pore is orientated from the intracellular to the extracellular side and is shown in its sequence (*bottom*) and three-dimensional-structure (*top*). Acidic residues are colored in *red*, basic residues in *blue* and polar residues in *purple*. Three-dimensional-structure representation (*top*): for each structure, all the residues lining the pore are shown and residues mutated and/or discussed in the text are labeled. Sequence representation (*bottom*): corresponding pore-lining residues are shown as sequence logos. Note that logo numbers correspond to the structural position of the residues along the channel axis and not to their sequence position. For details on this representation of the pore sequence see (39). The logo positions 9–44 line the inner region of the channel. The two asparagines of the NPA motifs are found at the logo positions 23 and 28. Arg of the ar/R selectivity filter occupies the logo position 30, while the other residues in the ar/R filter are located at the logo positions 32, 35, and 36.

To our knowledge, this is a unique comparative study of all the available subfamily representative aquaporin structures. In fact, previous computational studies focused on one or two members of the aquaporin family. It is worth outlining that our study was based on a relatively simple and fast protocol, which provides only part of the energetic picture, i.e. the electrostatic field generated by the pore-lining residues and the protein matrix but not structural relaxation phenomena or the energy penalty for the solute desolvation. Nevertheless, it proved to be very effective in outlining electrostatic profiles which correlate well with experimental selectivities and are in agreement with comparable data from MD simulations, in the few cases where they are available. Because of its speed and simplicity, this approach allowed a full comparative analysis of aquaporin structures and the subsequent derivation of important insights on their selectivity.

Our approach can be straightforwardly applied to a large number of three-dimensional structures and therefore facilitates the optimal use of the information arising from the increasing number of three-dimensional structures available for representatives of transmembrane channel proteins.

Mutation of the ar/R Arg Turns the AQP1 Electrostatics into Negative. All known aquaporins have a conserved Arg residue in the ar/R selectivity filter, with plant TIP1 and SIP1 subfamilies being the only exception. The substitution of this conserved Arg with a Val in the selectivity filter of rAQP1, experimentally shown to cause proton leakage (26), is shown here to reverse its positive electrostatic profile especially if associated with the mutation of the ar/R His. This is not a completely new finding. Indeed Chen et al. calculated the free energy profile of proton transport through these AQP1 mutants, showing that replacing the ar/R Arg with a Val results in a dramatic drop of the free energy barrier at the ar/R site and thus indirectly suggesting a negative electrostatics environment around it, which outweighs the ion dehydration penalty (21). Chakrabarti et al. (18) had also found that channel variants of GlpF, in which the charge of the ar/R Arg206 was removed, changed the electrostatic potential from neutral to deeply negative at the selectivity filter region.

Interestingly, we found here that PfAQP shows a negative electrostatic profile similar to that of the Arg-missing AQP1 mutants. Therefore protons or other cations might also be conducted across this aquaglyceroporin. However, ion conductance of PfAQP was investigated and excluded by Beitz & colleagues by electrophysiological recordings in isosmotic medium (25).

The Positive Charge on the ar/R Arg is Tuned by a Network of First- and Second-Shell Pore-Lining Residues in Aquaglyceroporins. Overall amino acids lining aquaporin channels have a conserved hydrophobic character. In fact, other than the conserved Asn of the NPA motifs and the basic residues of the ar/R selectivity filter, only few positions may host charged/polar residues having their side-chains lining the core region of the pores (Fig. 3 and Fig. S1). *P. falciparum* PfAQP has two unique acidic residues lining the channel: E28 and E195 (Fig. 3*H*). The role of such residues on the PfAQP selectivity has been investigated by site-specific mutations and shown to dramatically affect the channel selectivity (25). Both these residues are shown here to contribute to shield the ar/R R196 positive charge; in fact their mutation causes a shift of the electrostatic potential toward neutrality. However, to reverse the electrostatic profile and make it "orthodox-like", the simultaneous substitution of both these residues and of another acidic residue, D197, whose effect has not yet been experimentally tested, is required. D197, adjacent in sequence to the ar/R R196, can be considered a "second-shell" channel residue as it is not exposed on the channel surface but makes a double salt bridge with R196 and with the proximal R201. The electrostatic profile of the triple mutant, although mostly positive, is still different from that of orthodox aquaporins and this is not surprising as, compared to them, the histidine residue in the ar/R site is substituted by a glycine, and therefore its contribution to the positive electrostatic profile, as observed for AQP1 (Fig. 2*A*), is missing.

Analogously to PfAQP, eubacterial GlpF and archaeal AqpM also lack the His residue of the ar/R selectivity filter (substituted by a Gly and an Ile, respectively). Moreover, GlpF also presents one of the three PfAQP key acidic residues, i.e. the second-shell D207, next to the ar/R R206, giving a double salt bridge with R206 and K211. These features help to explain its rather neutral

electrostatic profile. Note that it had already been suggested that acidic residues could possibly compensate the positive charge of the ar/R guanidinium group in GlpF (17). As for AqpM, it hosts two polar residues, T203 and Y207, at positions corresponding to GlpF D207 and K211 (and to PfAQP D197 and K201), which are mostly substituted by a Ser and an Ala in orthodox aquaporins. *In silico* mutations suggest that the proximity of these two large polar residues to the ar/R R202 has an analogous effect in compensating the R202 positive charge and maintaining the neutrality of the channel electrostatics of AqpM. Of course contributions from residues beyond the pore region as well as pore dimension and geometrical shape are also expected to contribute to the fine tuning of all the trends observed for orthodox aquaporins and aquaglyceroporins. Note that water molecules are known to diffuse through the aquaporin channels with their dipoles mostly aligned to the pore axis (14, 15). In contrast, glycerols have recently been shown to transit with their dipoles roughly perpendicular to the pore axis (36). Therefore, the fairly neutral electrostatics of GlpF and AqpM might be seen as an evolutionary strategy to favor the diffusion of glycerol or apolar molecules over that of water.

Y149 is Responsible for the Anomalous AQP0 Electrostatic Profile. The low water conductance AQP0 subfamily is characterized by two unique Tyr residues, Y149 and Y23, protruding into the pore respectively at the cytoplasmic and periplasmic side. Substitution of Y149 with the corresponding AQP1 Thr residue is shown here to dramatically affect the electrostatic potential of the cytoplasmic half of the AQP0 channel, toward more neutral "orthodox-like" values. Residues Y23 and Y149 have been proposed to be responsible for the low AQP0 permeability, on the basis of MD simulations (38). Of course the strong bipolar electrostatic field we observe in the AQP0 channel, together with steric effects, may be of functional significance as it may cause water molecules to reside for longer in a stable configuration at the middle of the channel.

Clearly, the sequences of the large duplicated family of aquaporins have evolved distinct functional properties, according to the specific needs of the organism. Here we show that the electrostatic potentials effectively integrate the effects of these evolutionary developments to show a good correlation with the transport specificity of aquaporins and aquaglyceroporins.

Methods

Details on the sequences and structures collection, prediction of geometric pore centers, identification of pore-lining residues, calculation of electrostatic potentials, and model building are given in *SI Methods*.

ACKNOWLEDGMENTS. We thank Dr. Abdullah Kahraman and Tim Maiwald for helpful discussions. MPC was funded by the UK BBSRC grant number BBE0226421. This work was supported by the Italian MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca; Grant PRIN2006) (to G.C.). This work was also supported by the Federation of Biochemical Societies Short-Term Fellowships programme.

1. Wu B, Beitz E (2007) Aquaporins with selectivity for unconventional permeants. *Cell Mol Life Sci* 64:2413–2421.
2. Zardoya R (2005) Phylogeny and evolution of the major intrinsic protein family. *Biol Cell* 97:397–414.
3. Quigley F, Rosenberg JM, Shachar-Hill Y, Bohnert HJ (2002) From genome to function: The Arabidopsis aquaporins. *Genome Biol* 3 RESEARCH0001.
4. Magni F, et al. (2006) Proteomic knowledge of human aquaporins. *Proteomics* 6:5637–5649.
5. Fischer G, et al. (2009) Crystal structure of a yeast aquaporin at 1.15 angstrom reveals a novel gating mechanism. *PLoS Biol* 7:e1000130.
6. Calamita G, Bishai WR, Preston GM, Guggino WB, Agre P (1995) Molecular cloning and characterization of AqpZ, a water channel from *Escherichia coli*. *J Biol Chem* 270:29063–29066.
7. Maurel C, Reizer J, Schroeder JI, Chrispeels MJ, Saier MH Jr (1994) Functional characterization of the *Escherichia coli* glycerol facilitator, GlpF, in Xenopus oocytes. *J Biol Chem* 269:11869–11872.
8. Kozono D, et al. (2003) Functional expression and characterization of an archaeal aquaporin. AqpM from *Methanothermobacter marburgensis*. *J Biol Chem* 278:10649–10656.
9. Hansen M, Kun JF, Schultz JE, Beitz E (2002) A single, bi-functional aquaglyceroporin in blood-stage *Plasmodium falciparum* malaria parasites. *J Biol Chem* 277:4874–4882.
10. Preston GM, Carroll TP, Guggino WB, Agre P (1992) Appearance of water channels in Xenopus oocytes expressing red cell CHIP28 protein. *Science* 256:385–387.
11. Murata K, et al. (2000) Structural determinants of water permeation through aquaporin-1. *Nature* 407:599–605.
12. de Groot BL, Engel A, Grubmuller H (2001) A refined structure of human aquaporin-1. *FEBS Lett* 504:206–211.
13. Fu D, et al. (2000) Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* 290:481–486.
14. de Groot BL, Grubmuller H (2001) Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and GlpF. *Science* 294:2353–2357.
15. Tajkhorshid E, et al. (2002) Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science* 296:525–530.
16. de Groot BL, Grubmuller H (2005) The dynamics and energetics of water permeation and proton exclusion in aquaporins. *Curr Opin Struct Biol* 15:176–183.
17. Chakrabarti N, Tajkhorshid E, Roux B, Pomes R (2004) Molecular basis of proton blockage in aquaporins. *Structure* 12:65–74.
18. Chakrabarti N, Roux B, Pomes R (2004) Structural determinants of proton blockage in aquaporins. *J Mol Biol* 343:493–510.
19. Kato M, Pisliakov AV, Warshel A (2006) The barrier for proton transport in aquaporins as a challenge for electrostatic models: The role of protein relaxation in mutational calculations. *Proteins* 64:829–844.
20. Chen H, et al. (2007) Charge delocalization in proton channels, I: The aquaporin channels and proton blockage. *Biophys J* 92:46–60.
21. Chen H, Wu Y, Voth GA (2006) Origins of proton transport behavior from selectivity domain mutations of the aquaporin-1 channel. *Biophys J* 90:L73–75.
22. Wang Y, Schulten K, Tajkhorshid E (2005) What makes an aquaporin a glycerol channel? A comparative study of AqpZ and GlpF. *Structure* 13:1107–1118.
23. Hub JS, de Groot BL (2008) Mechanism of selectivity in aquaporins and aquaglyceroporins. *Proc Natl Acad Sci USA* 105:1198–1203.
24. Beitz E, et al. (2009) In vitro analysis and modification of aquaporin pore selectivity. *Handb Exp Pharmacol* 190:77–92.
25. Beitz E, Pavlovic-Djuranovic S, Yasui M, Agre P, Schultz JE (2004) Molecular dissection of water and glycerol permeability of the aquaglyceroporin from *Plasmodium falciparum* by mutational analysis. *Proc Natl Acad Sci USA* 101:1153–1158.
26. Beitz E, Wu B, Holm LM, Schultz JE, Zeuthen T (2006) Point mutations in the aromatic/arginine region in aquaporin 1 allow passage of urea, glycerol, ammonia, and protons. *Proc Natl Acad Sci USA* 103:269–274.
27. Berman HM, et al. (2002) The Protein Data Bank. *Acta Crystallogr D* 58:899–907.
28. Tornroth-Horsefield S, et al. (2006) Structural mechanism of plant aquaporin gating. *Nature* 439:688–694.
29. Yang B, Verkman AS (1997) Water and glycerol permeabilities of aquaporins 1-5 and MIP determined quantitatively by expression of epitope-tagged constructs in Xenopus oocytes. *J Biol Chem* 272:16140–16146.
30. Gonen T, Walz T (2006) The structure of aquaporins. *Q Rev Biophys* 39:361–396.
31. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041.
32. Harries WE, Akhavan D, Miercke LJ, Khademi S, Stroud RM (2004) The channel architecture of aquaporin 0 at a 2.2-A resolution. *Proc Natl Acad Sci USA* 101:14045–14050.
33. Pellegrini-Calace M, Maiwald T, Thornton JM (2009) PoreWalker: A novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional structure. *Plos Comput Biol* 5:e1000440.
34. Hedfalk K, et al. (2006) Aquaporin gating. *Curr Opin Struct Biol* 16:447–456.
35. Apweiler R, et al. (2004) UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res* 32:D115–119.
36. Henin J, Tajkhorshid E, Schulten K, Chipot C (2008) Diffusion of glycerol through *Escherichia coli* aquaglyceroporin GlpF. *Biophys J* 94:832–839.
37. Vidossich P, Cascella M, Carloni P (2004) Dynamics and energetics of water permeation through the aquaporin channel. *Proteins* 55:924–931.
38. Jensen MO, et al. (2008) Dynamic control of slow water transport by aquaporin 0: Implications for hydration and junction stability in the eye lens. *Proc Natl Acad Sci USA* 105:14430–14435.
39. Oliva R, Thornton JM, Pellegrini-Calace M (2009) PoreLogo: A new tool to analyze, visualize, and compare channels in transmembrane proteins. *Bioinformatics* 25:3183–3184.
40. Sui H, Han BG, Lee JK, Walian P, Jap BK (2001) Structural basis of water-specific transport through the AQP1 water channel. *Nature* 414:872–878.
41. Savage DF, Egea PF, Robles-Colmenares Y, O'Connell JD, 3rd, Stroud RM (2003) Architecture and selectivity in aquaporins: 2.5 a x-ray structure of aquaporin Z. *Plos Biol* 1:E72.
42. Fetter K, Van Wilder V, Moshelion M, Chaumont F (2004) Interactions between plasma membrane aquaporins modulate their water channel activity. *Plant Cell* 16:215–228.
43. Lee JK, et al. (2005) Structural basis for conductance by the archaeal aquaporin AqpM at 1.68 A. *Proc Natl Acad Sci USA* 102:18932–18937.
44. Newby ZE, et al. (2008) Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite *Plasmodium falciparum*. *Nat Struct Mol Biol* 15:619–625.
45. Horsefield R, et al. (2008) High-resolution x-ray structure of human aquaporin 5. *Proc Natl Acad Sci USA* 105:13327–13332.
46. Ho JD, et al. (2009) Crystal structure of human aquaporin 4 at 1.8 A and its mechanism of conductance. *Proc Natl Acad Sci USA* 106:7437–7442.