Dynamics and energetics of solute permeation through the *Plasmodium falciparum* aquaglyceroporin

Camilo Aponte-Santamaría,[a] Jochen S. Hub[b] and Bert L. de Groot*[a]

The aquaglyceroporin from *Plasmodium falciparum* (PfAQP) is a potential drug target for the treatment of malaria. It efficiently conducts water and other small solutes, and is proposed to intervene in several crucial physiological processes during the parasitic life cycle. Despite the wealth of experimental data available, a dynamical and energetic description at the single-molecule level of the solute permeation through PfAQP has been lacking so far. Here we address this question by using equilibrium and umbrella sampling molecular dynamics simulations. We computed the water osmotic permeability coefficient, the pore geometry and the potential of mean force for the permeation of water, glycerol and urea. Our simulations show that the PfAQP, the human aquaporin 1 (hAQP1) and the *Escherichia coli* glycerol facilitator (GlpF) have nearly identical water permeabilities. The Arg196 residue at the ar/R region was found to play a crucial role regulating the permeation of water, glycerol and urea. The computed free energy barriers at the ar/R selectivity filter corroborate that PfAQP conducts glycerol at higher rates than urea, and suggest that PfAQP is a more efficient glycerol and urea channel than GlpF. Our results are consistent with a solute permeation mechanism for PfAQP which is similar to the one established for other members of the aquaglyceroporin family. In this mechanism, hydrophobic regions near the NPA motifs are the main water rate limiting barriers, and the replacement of water–arg196 interactions and solute-matching in the hydrophobic pocket at the ar/R region are the main determinants underlying selectivity for the permeation of solutes like glycerol and urea.

1. Introduction

The parasite *Plasmodium falciparum* is responsible for the most lethal form of the malaria disease.[1] It expresses one aquaglyceroporin (PfAQP),[2] that has been proposed to play crucial physiological roles during the parasitic life cycle,[2,3] and is therefore a potential antimalarial drug target.[4–6]

PfAQP has been intensively studied experimentally over the last decade. Functional studies[2] established that PfAQP conducts water and glycerol at high permeability rates, comparable to the rates of the human aquaporin 1 (hAQP1) in the case of water. These experiments, therefore, suggested that PfAQP is implicated in tasks such as osmotic stress regulation, and glycerol uptake for lipid synthesis and oxidative stress regulation.[2] In addition, PfAQP was found to be permeable to urea,[2] ammonia,[7] other polyols up to five carbons long,[2] carbonyl compounds[3] and arsenite,[2] suggesting that PfAQP may also be involved in the release of toxic products. The biological role of aquaglyceroporins for the *Plasmodium* parasites was further investigated in experiments deleting the orthologue of PfAQP in the rodent malaria parasite, *Plasmodium berghei* (PbAQP), that showed a slower proliferation of the parasite in infected mice.[8] Furthermore, mutation studies demonstrated that a glutamate residue located at the C loop (Glu125) near the conserved Arg196 is critical for water permeation.[9]

The structure of PfAQP was recently determined by X-ray crystallography at 2.05 Å resolution.[10] It revealed the 3D architecture of the channel and provided molecular insights into the efficient water and glycerol permeation. PfAQP has the same fold as compared to other members of the family of aquaglyceroporins.[11] The pore geometry is remarkably similar to the *Escherichia coli* glycerol facilitator (GlpF),[10] the closest homologous aquaglyceroporin (with a known structure) with a sequence similarity of 50%.[2] PfAQP arranges in a tetrameric structure, where each monomeric unit constitutes a conduction pore. The conduction pore contains two constriction regions: one located near the NPA motifs, that are replaced in PfAQP by unusual NLA and NPS motifs, and a second one at the ar/R region, where Arg196 is located facing the hydrophobic Trp50 and Phe190 residues.[10] In addition, the C loop is anchored to the extracellular vestibule with Glu125 located in proximity to Arg196, highlighting on the importance of this region for the solute permeation,[10] as previously hypothesized from functional studies.[9]

Despite the wealth of data available, a dynamical and energetic description at the single-molecule level of solute permeation through PfAQP is lacking so far. Furthermore, a systematic comparison of PfAQP with hAQP1 and GlpF (the most studied aquaporin and aquaglyceroporin, respectively), in terms of the solute permeation and their dynamical causes, is incomplete. The goal of this paper is to address these two questions employing molecular dynamics simulations, based on the X-ray crystallographic structure of PfAQP. Initially, the water permeation through PfAQP was quantified by

[a] *Computational Biomolecular Dynamics Group, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany*
[b] *Computational and Systems Biology, Dept. of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden*

computing the single molecule osmotic permeability coefficient. Subsequently, the pore geometry and the energetics for water transport through PfAQP were analyzed by computing the radius and potential of mean force profiles from equilibrium simulations. Finally the energetics of permeation of glycerol and urea was studied by computing potentials of mean force by using the technique of umbrella sampling simulations.

The paper is organized as follows: section 2.1 lists the parameters and algorithms used for the equilibrium molecular dynamics simulations; section 2.2 to 2.4 describes the methods to derive the water permeability coefficients, the pore geometry and potentials of mean force from equilibrium MD simulations, and section 2.5 describes the umbrella sampling simulations. The results are presented and discussed in sections 3 and 4. Finally, section 5 summarizes our conclusions.

2. Theory and methods

2.1 Equilibrium molecular dynamics simulations

Equilibrium molecular dynamics simulations were carried out starting with the aquaporin tetramer in a fully solvated Dipalmitoylphosphatidylcholine (DPPC) lipid bilayer (Fig. 1). Three independent simulations of PfAQP aquaporin (labeled from I to III) were carried out: a control simulation of the wild-type form (I); a mutant where Glu125 (located at the C loop) was mutated into serine (II), and a second mutant where Arg196 (located at the pore face) was replaced by alanine (III). The latter two simulations were performed to study the effect of these two residues on the water permeation (Fig. 1(c)). In addition, simulations of hAQP1 (IV) and GlpF (V) aquaporins were also carried out for comparison.

The simulation boxes contain the protein tetramer, 265 DPPC lipids (263 in simulation V) and around 21 500 SPC/E water molecules.[12] The PfAQP and GlpF structures were taken from the Protein Data Bank (PDB ID codes 3C02[10] and 1FX8,[13] respectively). The starting structure of hAQP1 was modeled based on the X-ray structure of bovine AQP1 (PDB ID code 1J4N)[14] by mutating differing residues by using the WHAT IF modeling software.[15] The tetramer was inserted into the lipid bilayer by using the g_membed software.[16] Crystallographic water molecules were kept in the structures and ions were added to neutralize the simulation systems. The amber99SB all-atom force field[17] was used for the protein, and lipid parameters were taken from Berger et al.[18] The simulations were carried out using the GROMACS simulation software.[19,20] Long-range electrostatic interactions were calculated with the particle-mesh Ewald method.[21,22] Short-range repulsive and attractive interactions were described by a Lennard-Jones potential, which was cut off at 1.0 nm. The Settle algorithm[23] was used to constrain bond lengths and angles of water molecules and Lincs[24] was used to constrain all other bond lengths. The fastest angular degrees of freedom involving hydrogen atoms were removed by using the virtual interaction-sites algorithm,[25] allowing a time step of 4 fs. The temperature was kept constant by coupling the system to a velocity rescaling thermostat[26,27] at 300 K with a coupling constant $t = 0.5$ ps. The pressure was kept constant by coupling the system to a semiisotropic Parrinello-Rahman
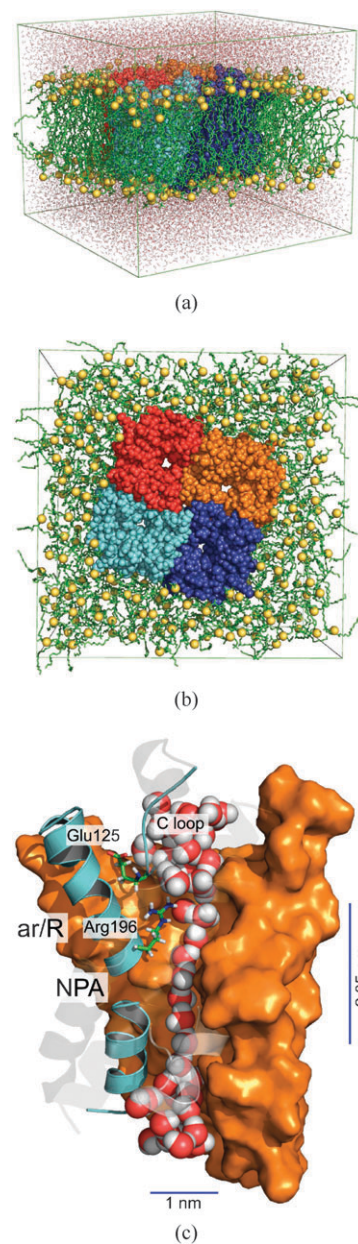

Fig. 1 Molecular dynamics simulations of PfAQP. Side (a) and top views (b) of the simulation box illustrating the tetramer (red, orange, cyan and blue), fully embedded in a DPPC lipid bilayer (yellow phosphorous atoms and green tails) and solvated by water (red, white). (c) Conduction pore showing the helices HB and HE; the C loop (cyan), and Glu125 and Arg196 (green). The blue lines indicate the region considered to compute the water permeability coefficient p_f.

barostat[28] at 1 bar with a coupling constant of $t = 5.0$ ps. All simulations were equilibrated for 4 ns before production. During this equilibration time the coordinates of the protein were harmonically restrained, with a harmonic force constant of 1000 kJ mol^{-1}/nm^2. The simulation length of the production runs was 100 ns discarding the first 10 ns as equilibration.

2.2 Water permeability calculations

The single-channel osmotic permeability, p_f, was chosen to quantify the water permeation through the simulated

This journal is © the Owner Societies 2010

Phys. Chem. Chem. Phys., 2010, **12**, 10246–10254 | 10247

aquaglyceroporins. A solute concentration difference, ΔC_s, between two media connected by a channel causes a net water flux, j, through the channel, and the permeability coefficient, p_f, relates these two quantities:

$$j = -p_f \Delta C_s. \tag{1}$$

The p_f was independently calculated for every monomer, based on the collective diffusion model proposed by Zhu et al.[29] A collective variable n is defined as

$$dn = \sum_i dz_i / L, \tag{2}$$

where dz_i are the displacements along the pore coordinate of the water molecules inside the channel (labeled as i) during a time dt, and L is the length of the channel. The variable n obeys an Einstein relation,

$$\langle n^2(t) \rangle = 2D_n t, \tag{3}$$

and the proportionality constant, $2D_n$, is related with the permeability coefficient:[29]

$$p_f = v_w D_n, \tag{4}$$

where $v_w \simeq 30$ Å^3 is the volume occupied by one water molecule.

Water molecule displacements, dz_i, were computed every 2 ps, within a cylindrical region (of length $L = 2.05$ nm and radius $r = 0.5$ nm) centered at the pore axis and spanning -0.7 nm down and 1.35 nm up from the NPA region (Fig. 1(c)). $\langle n^2(t) \rangle$ was obtained by averaging over 450 time windows of 200 ps length each. Finally the p_f was obtained from the slope of the curve $\langle n^2(t) \rangle$ versus time. An effective p_f value was obtained by averaging the values of the four monomers and the error was estimated as the standard error.

2.3 Pore dimensions

The geometry of the channels was monitored by computing pore radius profiles with the HOLE software,[30] averaging over snapshots taken every 50 ps, yielding a standard error always smaller than 0.4 Å

2.4 Potential of mean force for water

The potential of mean force (PMF) for water was calculated to identify rate limiting regions and relevant water–protein and water–water interaction sites inside the pore. It was independently calculated for every monomer by using the formula[31]

$$G_i(z) = -k_B T \ln \langle n_i(z) \rangle, \tag{5}$$

where k_B is the Boltzmann constant, T is the simulation temperature and $\langle n_i(z) \rangle$ is the average (over the whole trajectory) of the number of water molecules at the z position along the ith pore. The number of water molecules is referred to a cylinder (of radius 0.5 nm) that is aligned with the pore coordinate (Fig. 1(c)). Therefore, the entropy at the bulk regions is underestimated, and consequently, the bulk free energy is also underestimated. To relate $G(z)$ with the area of one aquaporin monomer and account for such understimations, a trapezoidal correction was applied in the entrance and

exit regions. In consequence, the final PMF refers to a density of one channel per membrane area occupied by an aquaporin monomer.[32] The correction reads $\Delta G_{cyl} = k_B T \ln(A_{mono}/A_c)$, where A_{mono} denotes the membrane cross section area of the aquaporin monomer and A_c the cross section area of the cylinder ($\pi \times 0.25$ nm^2). The computed correction was 6.1 kJ/mol for PfAQP, 6.0 kJ/mol for hAQP1 and 6.2 kJ/mol for GlpF. Prior to the analysis the protein monomers were superimposed to a reference structure. The effective PMF, $G_{eff}(z)$, was computed by combining the four monomer $G_i(z)$ as

$$\begin{aligned} \exp(-G_{eff}(z)/k_B T) &= \frac{1}{4} \sum_{i=1}^{4} \exp(-G_i(z)/k_B T) \\ &= \frac{1}{4} \sum_{i=1}^{4} \langle n_i(z) \rangle, \end{aligned} \tag{6}$$

and the error of G_{eff} was estimated by propagating the standard errors of $\langle n_i(z) \rangle$, yielding a maximum uncertainty of approximately 1.5 kJ/mol.

2.5 Umbrella sampling simulations

The starting frames for the umbrella simulation were taken from a 10 ns equilibrium simulations of PfAQP. The simulation system contained the PfAQP tetramer, 264 POPE lipids, 20484 TIP4P water molecules,[33] and four chloride ions. Protein and glycerol interactions were described using the OPLS all-atom force field,[34,35] lipid parameters were taken from from Berger et al.,[18] and urea parameters from ref. 36 and 37. All simulation parameters were chosen as described in the section 2.1, except that hydrogen atoms were not described as virtual sites, requiring a time step of 2 fs. The PfAQP channel was divided into 0.25 Å (1 Å for urea) wide equidistant sections parallel to the membrane with the center of each section representing an umbrella center. Subsequently, the solute molecules (glycerol or urea) were placed into the channel at the umbrella center. To enhance sampling, several solute molecules were placed in each pore keeping a distance of 20 Å between the solute molecules. Water molecules that overlapped with the solute were removed. The central carbon atom of glycerol or urea was restrained in z direction by a harmonic umbrella potential ($k = 1000$ kJ mol^{-1}/nm^2). In addition, the solute molecules were restrained into cylinders that were centered along the respective channel. Accordingly, a flat-bottom quadratic potential in the xy-plane was applied on the restrained atom (radius $r_c = 6$ Å, $k_c = 400$ kJ mol^{-1}/nm^2). After energy minimization, each system was simulated for 2 ns (4 ns for urea PMFs).

After removing the first 500 ps (1 ns for urea) for equilibration from each umbrella simulation, 1440 umbrella histograms (360 for urea) were extracted from the z-coordinate of the restrained atom. Subsequently, the umbrella positions were corrected with respect to the center of mass of the corresponding monomer. This procedure avoids a possible unphysical flattening of the PMF due to fluctuations of the PfAQP monomers within the tetramer. Visual inspection of the umbrella histograms showed that multiple histograms overlapped at each z coordinate. The PMF was calculated using a cyclic

implementation of the weighted histogram analysis method (WHAM).[38] The WHAM procedure incorporated the integrated autocorrelation times (IACT) of the umbrella windows. The IACTs were derived by fitting a double exponential to the autocorrelation function of each window, allowing one to analyically compute the IACTs. Because the IACT is subject to large uncertainty in case of limited sampling (such as inside the channel), we have subsequently smoothed the IACT along the reaction coordinate using a moving average filter with a width of 5 Å.

Due to the cylindric flat-bottom restraint the umbrella samplings yield a PMF which refers to a channel density of one channel per cross section area of the cylinder. We also corrected the PMF by a trapezoidal correction like the water PMF above. In this case, A_c was chosen such that the entropy of the solute in the flat-bottom quadratic potential equals the entropy of a solute in a cylindrical well-potential of area A_c. We found that this condition approximately holds if A_c is computed via $A_c = \pi(r_c + 2\sigma_c)^2$ where $\sigma_c = (k_B T/k_c)^{1/2}$ is the width of the Gaussian-shaped solute distribution at the edge of the flat-bottom quadratic potential. A_{mono} was estimated to equal 10.7 nm^2 and A_c was computed to 1.80 nm^2, yielding a correction of $\Delta G_{cyl} = 4.4$ kJ/mol. The statistical uncertainty of the PMFs was estimated using bootstrap analysis as described before,[39] yielding an uncertainty of 2 kJ/mol (3 kJ/mol for urea) at the main barrier and/or the main energy well.

3. Results

To quantify the water permeation through the different studied aqua(glycero)porins, the permeability coefficient, p_f, was computed from equilibrium molecular dynamics simulations (Fig. 2). The computed p_f value of the wild type PfAQP is $3.3 \pm 0.2 \times 10^{-14}$ cm^3/s. This value is nearly equal to the computed value for hAQP1 and slightly smaller than that of GlpF. In addition, the mutation of Glu125 (at the C loop, Fig. 1) into serine shows a reduction in the p_f of 18% compared to the wild type simulation. In contrast, the mutation of Arg196

(inside the narrow ar/R region of the pore, Fig. 1) into alanine reveals a substantial increase (by a factor of two) in the water permeability compared to the wild type simulation.

To identify the rate limiting regions in the pore that determine the water permeation characteristics observed above, the averaged pore radius and the PMF for water were calculated (Fig. 3). In PfAQP, hAQP1 and GlpF, the pore geometry remains essentially unchanged compared with the hole profiles calculated from the X-ray structures,[10] implying a relatively rigid pore. In PfAQP (orange line) and GlpF (cyan line) the pore profiles were found to be remarkably similar, narrowing in the ar/R region down to 1.4 Å radius, and wider than the profile in hAQP1 (blue line). In addition, the point mutation Glu125Ser (green line) does not modify the pore geometry, whereas the Arg196Ala mutation (red line) induces a widening of the pore to 1.7 Å at the ar/R region.

The PMF for water permeation is also illustrated in Fig. 3. In PfAQP (orange line), the PMF displays two major barriers of 14.6 kJ/mol and 15.4 kJ/mol near the NPA region which is not the most constricted region, and a third smaller barrier of 13.8 kJ/mol at the end of the ar/R region. In hAQP1 (blue line) and GlpF (cyan line), the PMFs also have a main free energy barrier at the NPA region, but they are smaller than the barrier in PfAQP by 2.4 kJ/mol and 1.3 kJ/mol, respectively. Furthermore, in PfAQP, the PMF displays eight local minima inside the channel, and water molecules are predominantly found at the positions of these minima. Consequently, the effective water–water distance along the pore, computed as the average distance between minima, is 2.6 ± 0.3 Å. At the same region, the PMF for hAQP1 (GlpF) shows seven (eight) minima separated by a water–water distance of 2.9 ± 0.6 Å (2.8 ± 0.6 Å). In hAQP1, the water–water distance increment is mainly observed between the NPA and the ar/R region. Additionally, the Glu125Ser mutant (green line) does not reveal significant changes in the PMF profile, whereas the Arg196Ala mutant (red line) appears to have a complete flattened profile at the ar/R region with a corresponding decrease in the free energy barrier to a value less than 11 kJ/mol.

Finally, Fig. 4 depicts the PMFs for glycerol and urea, derived from umbrella sampling simulations, for PfAQP (orange line), and compared with previous molecular dynamics studies for GlpF[32] (cyan line). In PfAQP, the free energy barrier at the ar/R region for glycerol is substantially smaller than the existing for urea (approximately 21 kJ/mol smaller). Moreover, for both solutes, glycerol and urea, the free energy barrier at the ar/R region decreases in PfAQP compared to that observed in GlpF. In addition, the PMF for glycerol reveals new features in PfAQP that are not observed in GlpF: first, there is a potential well at the intracellular vestibule ($z < -1.0$ nm) instead of a free energy barrier. Second, the potential wells at an intermediate region between the intracellular vestibule and the NPA motifs are deeper in PfAQP than in GlpF. Contrary to glycerol, the PMFs for urea in PfAQP and GlpF were found to be similar between the intracellular vestibule and the NPA motifs. Furthermore, for both PfAQP and GlpF, glycerol crystallographic positions (indicated with dots) agree favorably with the potential well positions observed in the simulations.
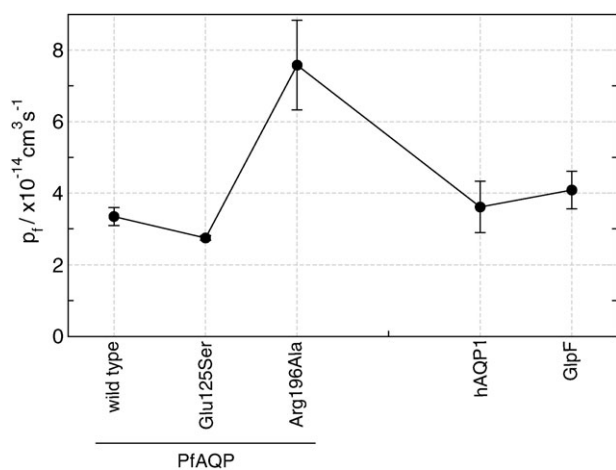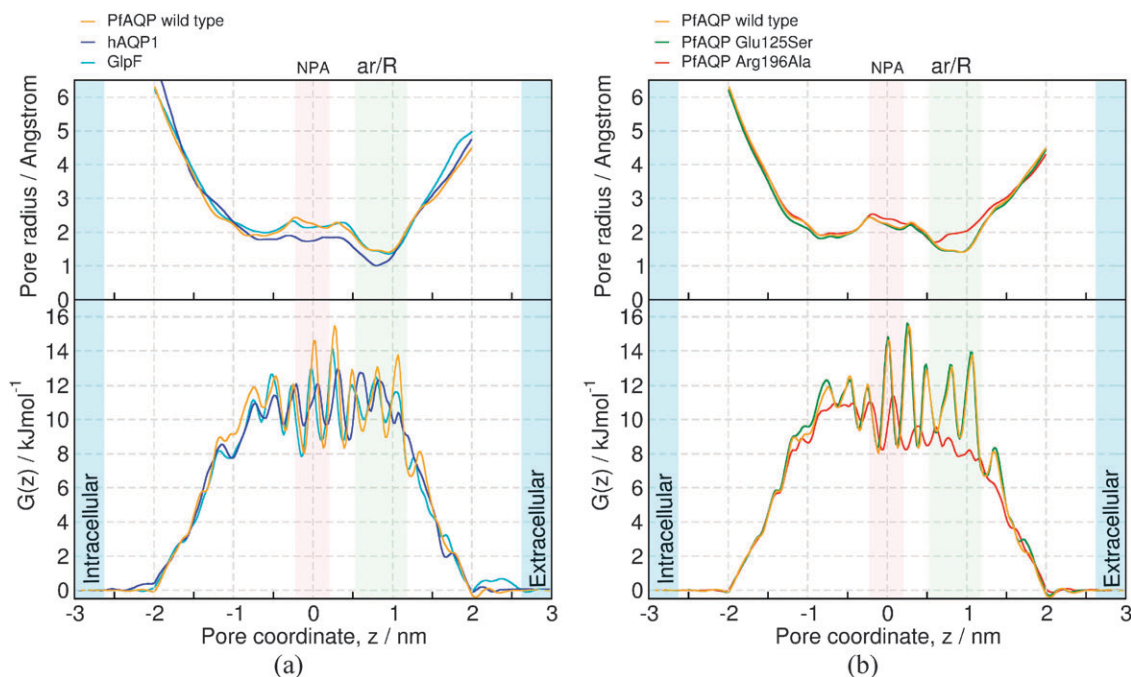

Fig. 2 Single molecule water permeability coefficients derived from equilibrium molecular dynamics simulations for the indicated proteins and mutants.

Fig. 3 Pore radius profiles (upper panels) and potential of mean force for water permeation profiles (lower panels) derived from equilibrium molecular dynamics simulations. (a) Comparison between PfAQP (orange), hAQP1 (blue) and GlpF (cyan). (b) Effect of the point mutations Glu125Ser (green) and Arg196Ala (red).
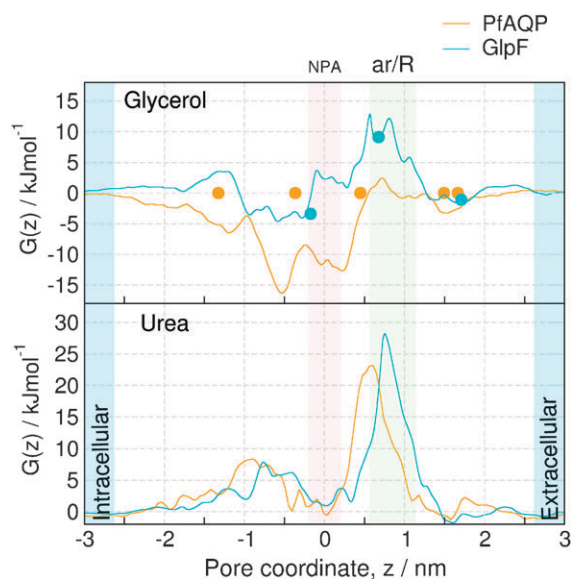

Fig. 4 PMFs for glycerol and urea permeation through PfAQP (orange) compared with previous molecular dynamics studies for GlpF[32] (cyan). The dots correspond to the glycerol positions in the PfAQP[10] (orange dots) and Glpf[13] (cyan dots) crystal structures.

4. Discussion

Here we present a computational study to characterize at the molecular level the permeation of water, glycerol and urea through the *Plasmodium falciparum* aquaglyceroporin (PfAQP). Initially, the water permeation through PfAQP was quantified by means of the single molecule osmotic permeability coefficient, p_f. Subsequently, the pore geometry

and the energetics for water, glycerol and urea transport through PfAQP was analyzed by computing radius and potential of mean force profiles.

4.1 Quantifying the water transport activity

Let us first consider the absolute p_f values predicted in our simulations in light of permeation measurements and previous computational studies. First, our computed values for the reference (aquaporin) hAQP1 and (aquaglyceroporin) GlpF (Fig. 2) are close to the experimental values, spanning between 4.6×10^{-14} cm^3/s to 11.7×10^{-14} cm^3/s for hAQP1[40–43] and 0.7×10^{-14} cm^3/s for GlpF.[44] Moreover, our predicted p_f for PfAQP is also close to these measurements. It should be noted that there are no single-molecule p_f measurements for PfAQP reported so far to compare to directly. Second, our derived p_f values are smaller than the values predicted in previous computational studies on hAQP1 using the GROMOS force field and the SPC water model[31,45] by a factor of two. Similarly, they are smaller than the values derived in studies on hAQP1 and GlpF using the CHARMM force field and the TIP3P water model[46–48] at least by a factor of three. These differences can be mainly attributed to the water diffusion constant. For the SPCE water model (used in this study) the diffusion constant is very close to the experimental value (2.4×10^{-5} cm^2/s). In contrast, for the SPC[49] and the TIP3P[33] water models (used in the mentioned previous computational studies) the diffusion constant is overestimated by a factor of 1.5 and 2, respectively.

Let us now to consider the relative differences in the p_f for PfAQP compared to hAQP1 and GlpF. The computed p_f for PfAQP is close to the value predicted for hAQP1 (Fig. 2), suggesting that PfAQP conducts water at a similar rates as

hAQP1. This result is in excellent agreement with experimental data from oocyte swelling assays,[2] and supports therefore that PfAQP is a highly efficient water channel. Interestingly, the p_f for GlpF was observed to be higher than the values for PfAQP and hAQP1, suggesting that GlpF also conducts water at appreciable rates (comparable to hAQP1). A water permeability for GlpF has been observed in previous computational studies,[31,46–48] experimental assays with proteoliposomes[50] and reconstituded planar bilayers,[44] but with the experiments showing a lower permeation rate than the computational studies. The reason for this discrepancy remains unknown. A possible explanation would be simulation inaccuracies. However, given the excellent agreement between the simulation and the experiment for AQP1, and the consistent results for GlpF in several different and independent simulations, this appears unlikely. Another possible explanation could lie in the problematic estimation of the reconstitution efficiency, required to derive the copy number to estimate the experimantal single-channel p_f.

4.2 The energetics of water permeation

Further calculations of the pore geometry and the PMF for water (Fig. 3) allowed us to identify the rate limiting regions that determine the water permeation properties presented above. The averaged radius profiles demonstrate that PfAQP and GlpF have practically identical pore geometries, wide enough to allow the passage of water and other solutes such as glycerol and urea. Accordingly, the narrower pore observed for hAQP1 constitutes one of the main factors for the exclusion of large solute molecules in this aquaporin. The average pore geometry from the simulations is remarkably similar to that from the initial crystal structures[10,13,14] (all compared in ref. 10), indicating a rigid channel.

The pore geometry itself is not sufficient to explain the water transport rates but the full energetics must be taken into account. The major rate limiting region (highest barriers in the PMF) for PfAQP is observed at the NPA region, located exactly at the same place as for hAQP1 and GlpF (Fig. 3). In all three cases, two rings of hydrophobic residues sit there (Fig. 5), and they constitute the main barriers for water passage. This is in perfect agreement with previous simulations of hAQP1 and GlpF,[31,51] and confirms that PfAQP has similar water regulation mechanisms as other members of the family of aquaglyceroporins. Furthermore, in PfAQP, the Leu192 residue may force the permeating water molecules to interact more strongly with the Asn residue than in GlpF (where this Leu192 is replaced by Met202, a less hydrophobic residue, Fig. 5), leading to the observed higher free energy barrier, for PfAQP as compared to GlpF. Surprisingly, hAQP1 has smaller free energy barriers than PfAQP and GlpF in this critical region, despite the fact that the protein-water interactions are much stronger due to a narrower pore and the presence of the larger Phe24 and Phe56 residues in this region (Fig. 5).[31]

Finally, from the free energy barriers, it could be expected that PfAQP, GlpF and hAQP1 conduct water at rates in the following order: PfAQP < GlpF < hAQP1. Our p_f calculations predict the lowest rate for PfAQP. However, hAQP1 did not show the highest rate but GlpF did. This apparent contradiction between the p_f and the PMF for hAQP1 and GlpF may be due to the fact that a permeation event contributing to the p_f (the translocation of one effective water molecule from one aqueous medium to the other) implies the collective motion of all molecules inside the channel,[29,52] and therefore multiple free energy barriers and not only the highest should be taken into account when relating permeation rates to permeation energetics.
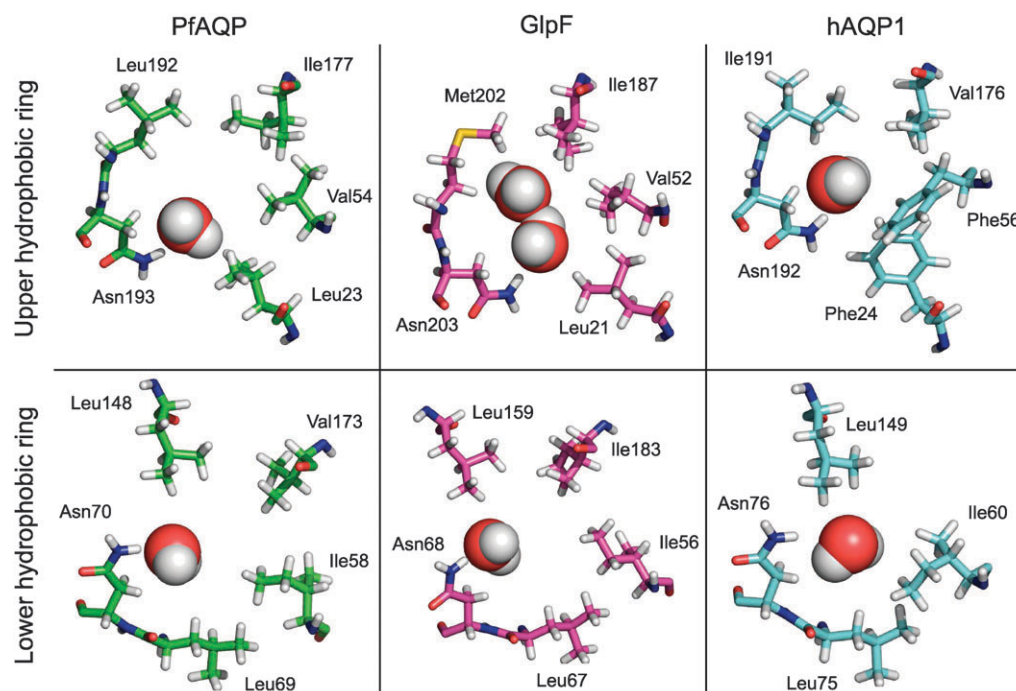

Fig. 5 Top view of the two hydrophobic rings (upper and lower panels) located near the NPA motifs in the indicated proteins.

This journal is © the Owner Societies 2010

Phys. Chem. Chem. Phys., 2010, **12**, 10246–10254 | 10251

4.3 The effect of point mutations

We investigated the role of Glu125 and Arg196 on the water permeation of PfAQP. On one hand, the Glu125Ser mutant shows neither a substantial reduction in the p_f nor an increment in the free energy barriers compared to the wild type simulation (Fig. 2 and 3(b)), contrary to what previous mutation experiments demonstrated.[9] The reduction in the water permeability due to the Glu125Ser mutation was attributed to the destabilization of the C loop and further disruption of the hydrogen bonds between Arg196 and Trp124.[9,10] However, the conformational change of the C loop may occur on a longer time scale than the simulated time (100 ns), impeding to be detected in our equilibrium simulations. In fact, the C loop was observed to be rigidly anchored to the extracellular vestibule (with an average backbone rmsd of 0.97 Å with a standard deviation of 0.26 Å) in the wild type simulation, and just slightly more flexible (with an average backbone rmsd of 1.06 Å and a standard deviation of 0.36 Å) in the Glu125Ser simulation.

On the other hand, the Arg196Ala mutant reveals a widening of the pore at the Ar/R region and a corresponding reduction in the free energy barrier to water permeation, leading to a significant increase in the water permeability (Fig. 2 and 3(b)). This result confirms that Arg196 is a crucial residue for the water conduction through PfAQP.

4.4 Permeation of other solutes

We studied the energetics of glycerol and urea conduction (Fig. 4). Let us first analyze the free energy barriers at the ar/R region for the different simulated solutes. Our simulations predict that PfAQP conducts urea at a lower rate than glycerol which is in excellent agreement with oocyte swelling assays.[2] In PfAQP, like in other aquaglyceroporins, glycerol molecules arrange at the ar/R in such a way that its hydroxyl groups can interact with Arg196, replacing water–Arg196 hydrogen bonds, and its apolar backbone orients towards the hydrophobic face (Trp50 and Phe190 residues).[10,13] In contrast, urea molecules cannot find this compensatory effect (dictated by the amphiphilic nature of the pore), because increasing the carbonyl–oxygen interactions with Arg196 would lead to unfavourable orientations of the amine groups facing the hydrophobic residues, and therefore would increase the free energy barrier as indeed observed in our simulations.

Focussing at the differences between PfAQP and GlpF, it is intriguing to note that PfAQP has a reduced free energy barrier for both glycerol and urea compared to GlpF at the ar/R region. This suggests that PfAQP is a more efficient glycerol and urea channel than GlpF, despite the fact that both have identical residues in this critical region. A possible explanation for this would be that fewer water–arginine hydrogen bonds need to be replaced by the permeating solute molecule in the case of PfAQP compared to GlpF, and therefore the energetic cost to replace water with a solute at that position is reduced. Consequently, in PfAQP, the arginine residue instead would form more additional hydrogen bonds with the surrounding residues (namely, Ser200 and Trp124) than in GlpF, as proposed by Newby et al.[10] Indeed, the histograms of the number of hydrogen bonds between the
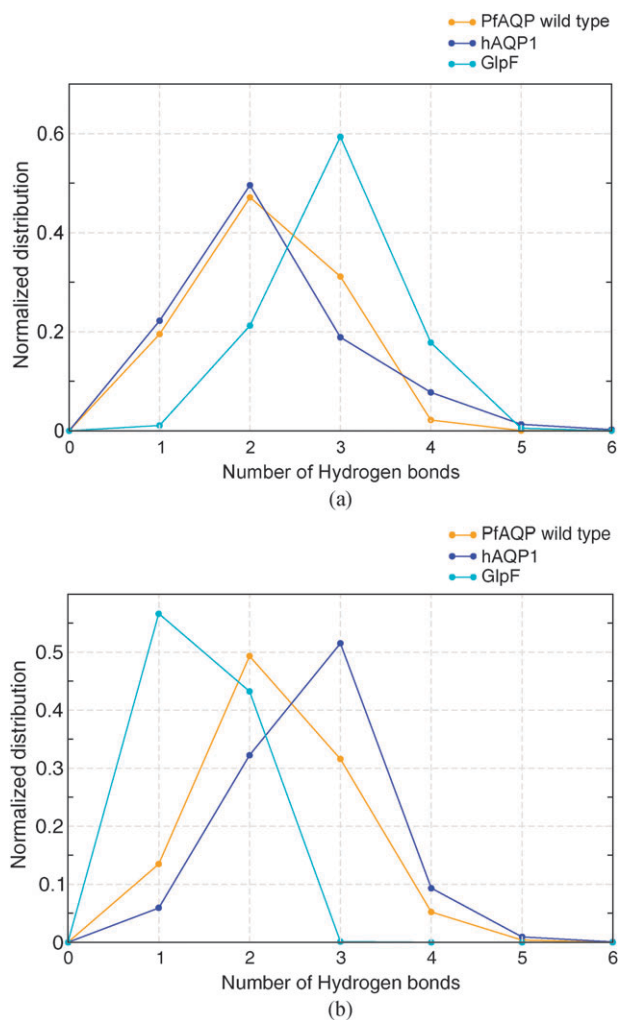

Fig. 6 Water–Arg196 (a) and protein–Arg196 (b) hydrogen bond distribution in PfAQP (orange), hAQP1 (blue) and GlpF (cyan).

solvent and the arginine residue computed in the equilibrium simulations (Fig. 6(a)) show on average fewer hydrogen bonds for PfAQP than for GlpF, and accordingly, more hydrogen bonds between the protein and the arginine residue for PfAQP than for GlpF (Fig. 6(b)).

In addition, the glycerol PMF for PfAQP reveals two features not observed in GlpF: a potential well at the intracellular vestibule and a putative binding site inside the channel near the NPA region This binding site is expected to affect the water permeability, particularly, at high glycerol concentrations. It is therefore highly interesting to see if this glycerol-dependent water permeability can be observed experimentally.

5. Conclusions

Here we present the first molecular dynamics study to characterize at the molecular level the permeation of water and other solutes through the *Plasmodium falciparum* malaria aquaglyceroporin (PfAQP). Our simulations confirm that PfAQP is a highly efficient water channel, that is able to conduct water at single-molecule permeability rates comparable to the hAQP1 and GlpF rates. Furthermore, we identified the

10252 | *Phys. Chem. Chem. Phys.*, 2010, **12**, 10246–10254

This journal is © the Owner Societies 2010

hydrophobic regions near the NPA motif as the main water rate-limiting barriers, confirming therefore that PfAQP has a similar water regulation mechanism as other members of the family of aquaglyceroporins. Our simulations also support that GlpF is a highly efficient water channel, with a water permeability higher than that for hAQP1, as previously observed in several different and independent simulation studies.

We demonstrate that Arg196 located at the ar/R selectivity filter plays a crucial role regulating the permeation of water and other solutes such as glycerol or urea, as was previously hypothesized from functional and crystallographic studies. Our simulation results are consistent with a general permeation mechanism for aquaglyceroporins in which water is rate-limited by hydrophobic interactions in the NPA region. Other solutes, such as glycerol or urea, must accommodate in the ar/R selectivity filter, replacing water–arg196 hydrogen bonds interactions and facing the hydrophobic Trp50 and Phe190 residues, a process that is favorable for glycerol but less favorable for urea. In addition, in light of this mechanism, our simulations suggest that PfAQP is able to conduct both glycerol and urea at higher permeabilities than GlpF does, because fewer water–Arg196 hydrogen bonds need to be replaced by a permeating solute molecule in the case of PfAQP compared to GlpF, a prediction to be further investigated and validated in experimental studies.

The effect on the water permeability due to the mutation of Glu125 (located at the C loop) to serine was found to be less severe in our simulations than observed experimentally. This could be attributed to the high stability of the C loop, rigidly anchored to the extracellular vestibule, that would slow C loop motions to time scales longer than the simulated time due to the mentioned mutation.

This study is expected to guide further computational and experimental studies in the search of putative blockers of PfAQP, understanding of their mechanism of action, that can hopefully be used to interrupt crucial physiological processes of the malaria parasite such as the water regulation and glycerol uptake.

Acknowledgements

References

1 M. J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. A. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser and B. Barrell, *Nature*, 2002, **419**, 498–511.
2 M. Hansen, J. A. F. J. Kun, J. E. Schultz and E. Beitz, *J. Biol. Chem.*, 2002, **277**, 4874–4882.
3 S. Pavlovic-Djuranovic, J. F. J. Kun, J. E. Schultz and E. Beitz, *Biochim. Biophys. Acta, Biomembr.*, 2006, **1758**, 1012–1017.
4 E. Beitz, *ChemMedChem*, 2006, **1**, 587–592.
5 E. Beitz, *Biol. Cell*, 2005, **97**, 373–383.
6 K. Kirk, *Acta Trop.*, 2004, **89**, 285–298.
7 T. Zeuthen, B. Wu, S. Pavlovic-Djuranovic, L. M. Holm, N. L. Uzcategui, M. Duszenko, J. F. J. Kun, J. E. Schultz and E. Beitz, *Mol. Microbiol.*, 2006, **61**, 1598–1608.
8 D. Promeneur, Y. Liu, J. Maciel, P. Agre, L. S. King and N. Kumar, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 2211–2216.
9 E. Beitz, S. Pavlovic-Djuranovic, M. Yasui, P. Agre and J. E. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 1153–1158.
10 Z. E. R. Newby, J. O'Connell III, Y. Robles-Colmenares, S. Khademi, L. J. Miercke and R. M. Stroud, *Nat. Struct. Mol. Biol.*, 2008, **15**, 619–625.
11 T. Walz, Y. Fujiyoshi and A. Engel, *Handbook of Experimental Pharmacology: Aquaporins*, 2009, pp. 31–56.
12 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.
13 D. Fu, A. Libson, L. J. W. Miercke, C. Weitzman, P. Nollert, J. Krucinski and R. M. Stroud, *Science*, 2000, **290**, 481–486.
14 H. Sui, B.-G. Han, J. K. Lee, P. Walian and B. K. Jap, *Nature*, 2001, **414**, 872–878.
15 G. Vriend, *J. Mol. Graphics*, 1990, **8**, 52–56.
16 M. G. Wolf, M. Hoefling, C. Aponte-Santamaría, H. Grubmüller and G. Groenhof, *Journal of Computational Chemistry*, 2010, http://dx.doi.org/10.1002/jcc.21507.
17 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712–725.
18 O. Berger, O. Edholm and F. Jähnig, *Biophys. J.*, 1997, **72**, 2002–2013.
19 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
20 D. V. D. Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.
21 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.
22 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
23 S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 952–962.
24 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
25 K. A. Feenstra, B. Hess and H. J. C. Berendsen, *J. Comput. Chem.*, 1999, **20**, 786–798.
26 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 14101.
27 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
28 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
29 F. Zhu, E. Tajkhorshid and K. Schulten, *Phys. Rev. Lett.*, 2004, **93**, 224501.
30 O. S. Smart, J. M. Goodfellow and B. A. Wallace, *Biophys. J.*, 1993, **65**, 2455–2460.
31 B. L. de Groot and H. Grubmüller, *Science*, 2001, **294**, 2353–2357.
32 J. S. Hub and B. L. de Groot, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 1198–1203.
33 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
34 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.
35 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. B*, 2001, **105**, 6474–6487.
36 E. M. Duffy, D. L. Severance and W. L. Jorgensen, *Israel Journal of Chemistry*, 1993, **33**, 323–330.
37 L. J. Smith, H. J. C. Berendsen and W. F. van Gunsteren, *J. Phys. Chem. B*, 2004, **108**, 1065–1071.
38 S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 1011–1021.
39 J. S. Hub and B. L. de Groot, *Biophys. J.*, 2006, **91**, 842–848.

This journal is © the Owner Societies 2010

Phys. Chem. Chem. Phys., 2010, **12**, 10246–10254 | 10253

40 M. L. Zeidel, S. V. Ambudkar, B. L. Smith and P. Agre, *Biochemistry*, 1992, **31**, 7436–7440.

41 M. L. Zeidel, S. Nielsen, B. L. Smith, S. V. Ambudkar, A. B. Maunsbach and P. Agre, *Biochemistry*, 1994, **33**, 1606–1615.

42 T. Walz, B. L. Smith, M. L. Zeidel, A. Engel and P. Agre, *Journal of Biological Chemistry*, 1994, **269**, 1583–1586.

43 B. Yang and A. S. Verkman, *J. Biol. Chem.*, 1997, **272**, 16140–16146.

44 S. M. Saparov, S. P. Tsunoda and P. Pohl, *Biol. Cell*, 2005, **97**, 545–550.

45 B. L. de Groot and H. Grubmüller, *Curr. Opin. Struct. Biol.*, 2005, **15**, 176–183.

46 F. Zhu, E. Tajkhorshid and K. Schulten, *Biophys. J.*, 2002, **83**, 154–160.

47 M. Jensen and O. G. Mouritsen, *Biophys. J.*, 2006, **90**, 2270–2284.

48 M. Hashido, A. Kidera and M. Ikeguchi, *Biophys. J.*, 2007, **93**, 373–385.

49 W. F. van Gusteren, J. Hermans, H. J. C. Berendsen and J. P. M. Postma, *Intermolecular forces*, Reidel, Dordrecht, Holland, 1981.

50 M. J. Borgnia and P. Agre, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 2888–2893.

51 E. Tajkhorshid, P. Nollert, M. O. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud and K. Schulten, *Science*, 2002, **296**, 525–530.

52 F. Zhu, E. Tajkhorshid and K. Schulten, *Biophys. J.*, 2004, **86**, 50–57.