Biophysical Journal, Volume 99

Supporting Material

Title: Voltage-regulated water flux through aquaporin channels <I>in silico</I>

Authors: Jochen S Hub, Camilo Aponte-Santamaría, Helmut Grubmüller, and Bert L de Groot

Supplementary material for

Voltage-regulated water flux through aquaporin channels *in silico*

Jochen S. Hub*, Camilo Aponte-Santamaría‡, Helmut Grubmüller,‡ and Bert L. de Groot‡

*Department for Cell and Molecular Biology, Uppsala University, Uppsala, Sweden
‡Department of Theoretical and Computational Biophysics, Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany

Two-state model for the conserved arginine

To estimate if the arginine displacement is sufficient to explain the tuning of P_{open} in response to $\Delta\Phi$, let us assume a simple open/closed two-state model. When switching between the *up* and the *down* state, the arginine (Arg195 and Arg216 for AQP1 and AQP4, respectively) moved by $\delta z \sim$1.5Å in the z-direction with respect to the center of mass of the respective monomer (data not shown). Let us in the following in addition assume a homogeneous electric field across the membrane and a membrane thickness of $d = 4\,\text{nm}$. Then, $\Delta\Phi$ affects the potential energy difference between *up* and *down* state by $\Delta V(\Delta\Phi) = e\,\delta z\,\Delta\Phi/d$, where e denotes the unit charge. Note that Figures 2E and 3B suggest that the open and closed state are equally likely at $\Delta\Phi = 0$. The relative populations of the open and closed states thus follow $P_{\text{open}}/P_{\text{closed}} = \exp(\Delta V/k_BT)$, where k_B and T denote the Boltzmann constant and the temperature, respectively. Using $P_{\text{open}} + P_{\text{closed}} = 1$, the probability of the open state is given by

$$P_{\text{open}}(\Delta\Phi) = \left[1 + \exp\left(-\frac{e\,\delta z}{d\,k_BT}\,\Delta\Phi\right)\right]^{-1} \tag{1}$$

$P_{\text{open}}(\Delta\Phi)$ is plotted as dashed black curves in Figures 2E and 3B. The curves are in reasonable agreement with the data points and the fitted lines, suggesting that electrostatic interactions of the arginine with the external field are indeed sufficient to explain the tuning of P_{open}.

p_f of AQP4 versus membrane potential $\Delta\Phi$

The single-channel permeabilities p_f converge slowly with simulation time. Besides the influence of the Arginine-Histidine distance in the ar/R region, p_f is affected by fluctuations of other residues

along the pore. In addition, we observed that the pore is at times only partly filled by water, or the water file does hardly moves for several nanoseconds without any "obvious" reason. Such effects add substantial noise the p_f-versus-$\Delta\Phi$ functions, as derived from finite simulations. As a consequence, p_f-versus-$\Delta\Phi$ for AQP4 (Supporting Figure S1) turned out to be more noisy than the respective signal for AQP1 (Fig. 2A of the main text). The curve is compatible with a voltage-sensitive p_f, but cannot fully prove the significance of the signal. Recall that the openness of the ar/R region of AQP4, contrast to p_f, robustly correlates with $\Delta\Phi$ (Fig. 3B).

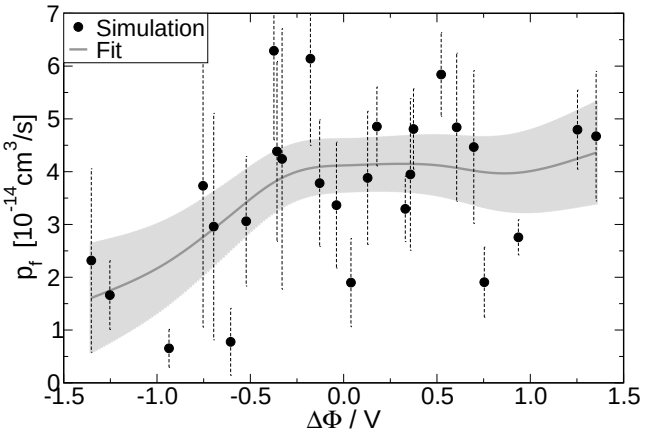

Figure S1: Single-channel permeability p_f for AQP4, as a function of membrane potential $\Delta\Phi$.

Materials and Methods

Simulation setup

The simulations systems of aquaporin-1 (AQP1) and aquaporin-4 AQP4 were set up as follows. The AQP4 crystal structure was taken from the protein data bank (pdb code 3GD8 (1)). The AQP4 tetramer including crystal water was placed into a simulation box. The protein was solvated in a lipid membrane of 265 palmitoyloleoylphosphatidylethanolamine (POPE) lipids using the g_membed software (2), and subsequently solvated in explicit water. A pre-equilibrated simulation system of an AQP1 tetramer, embedded in a lipid membrane and solvated in explicit water was taken taken from a previous study (3). For each of the AQP systems, two copies of the simulation box were stacked on top of each other, with the extracellular side of each tetramer pointing in positive z direction (membrane normal). Because all simulations were carried out with periodic boundary conditions, the stacked system thus formed two water compartments separated by two membranes. The water compartment in the center of the simulation box (compare Fig. 1A) is in the following referred to as 'central' water compartment, and the water compartment formed by the water at the top and the bottom of the simulation box as 'outer' water compartment. Approx-

imately 150 mM sodium chloride was added to each water compartment separately. To neutralize the AQP1 simulation cell, 16 more chloride as compared to sodium ions we added to each compartment. The AQP1 simulation system contained two tetramers, 542 POPE lipids, 38444 water molecules, 108 chloride ions, and 76 sodium ions, summing up to a total of 214968 atoms. The AQP4 system contained two tetramers, 530 POPE lipids, 32420 water molecules, 88 chloride and 88 sodium ions, summing up to 187160 atoms.

An electrostatic potential $\Delta\Phi$ across the two membranes was generated by placing additional cations and anions into the central and outer water compartments, respectively. Due to the limited size of the simulation system, the addition of one cation/anion pairs induces a substantial $\Delta\Phi$ of $\sim$200 mV. To allow for the simulation of smaller increments in $\Delta\Phi$, we placed, in addition to ions with integer charge, ions with fractional charges of $\pm e/2$ where e denotes the unit charge. Lennard-Jones parameters for the cations and anions with fractional charge were taken from the sodium and chloride ions, respectively. The electrostatic potential $\Phi(z)$, averaged over the simulation trajectory, was derived from the Poisson equation $\Delta\Phi(\mathbf{x}) = -\rho(\mathbf{x})/\epsilon_0$, where Δ denotes the Laplace operator, $\Phi(\mathbf{x})$ the electrostatic potential, $\rho(\mathbf{x})$ the time-averaged charge density, and ϵ_0 the vacuum permittivity. Accordingly, the charge density in the simulation was averaged over the $x-y$ plane and over the simulation frames, yielding a one-dimensional time-average density $\rho(z)$. $\Phi(z)$ was straight-forwardly computed by twice integrating $\rho(z)$. Because of the double integration, slightly non-converged charge densities may result in a non-physical offset in $\Phi(z)$ between the top and the bottom of the simulation box. In order to account for the periodic boundary conditions such offsets were removed by a linear correction in $\Phi(z)$.

All simulation were carried out using the Gromacs simulation software (4, 5). The OPLS all-atom force field (6, 7) and the TIP4P model (8) were employed for the protein and water, respectively, and lipid parameters were taken from Berger *et al.* (9). Electrostatic interactions were calculated at every step with the particle-mesh Ewald method (10, 11). Short-range repulsive and attractive dispersion interactions were described together by a Lennard-Jones potential, which was cut off at 1.0 nm. The Settle (12) algorithm was used to constrain bond lengths and angles of water molecules, and LINCS (13) was used to constrain all other bond lengths. In addition, the fastest angular degrees of freedom involving hydrogen atoms were removed by using the virtual interaction-sites algorithm (14), allowing a time step of 4 fs. The temperature was kept at 300 K through velocity rescaling (15) (τ = 2.5 ps, 0.5 ps in AQP4 simulations), and the pressure was controlled at 1 bar using the Parrinello-Rahman coupling scheme (τ = 5 ps) (16).

Simulation analysis

Before analyzing the trajectories, the first 10 ns for AQP1 and 5 ns for AQP4 were removed from all simulations for equilibration. The single-channel permeabilities p_f were computed independently for each monomer based on the collective diffusion model proposed by Zhu *et al.* (17). Accordingly,

p_f was computed from the slope of the mean-square displacement (MSD) of the collective water coordinate. The slope was computed by fitting a line to the MSD up to a time displacement of 200 ps and 100 ps for AQP1 and AQP4, respectively. For each monomer, the analysis was carried out on water molecules within a cylinder of length 1.15 Å and 0.84 Å for AQP1 and AQP4, respectively, between the NPA and the aromatic/arginine region, that is within the narrowest part of the channel.

The smooth splines in Fig. 2A and S1 and the linear line in Figs. 2E and 3B were fitted to the data points from all individual monomers (not shown) using gnuplot. Statistical errors for the fitted smooth splines were computed using the Bayesian bootstrap (18). Accordingly, the fit was carried out 200 times, with randomly selected weights for the individual data points for each fitting procedure. According to the Bayesian bootstrap, the weights are generated as follows (18). Given n data points, draw $n-1$ uniform random variables between 0 and n, and let $u_{(1)}, u_{(2)}, \ldots, u_{(n-1)}$ denote their values in increasing order. In addition, let $u_{(0)} = 0$ and $u_{(n)} = n$. The random weights w_i are then defined by the gaps between two consecutive random numbers, i.e. $w_i = u_{(i)} - u_{(i-1)}$, where $i = 1, \ldots, n$. The statistical error for the fit is given by the standard deviation computed from the 200 fitted splines (gray shaded areas in Figs. 2A and S1 show two standard deviations). In statistical error in the linear least-square fit to P_{open} (gray curve and shaded area in Fig. 2E and 3B) was computed using standard error propagation, based on the standard errors of the slope and the y-intercept, which are derived by the fitting procedure. All statistical errors were plotted as is as two standard deviations, that is as 95% confidence intervals.

Statistical errors for the data points from individual simulations (black dots with dashed error bars in in Figs. 2A, 2E, and 3B) were computed from the the four values (p_f, or P_{open}) taken from the four monomers via $\sigma/\sqrt{N_{\text{m}} - 1}$, where σ denotes the standard deviation and $N_{\text{m}} = 4$ the number of monomers. Because these errors are based on only four values, we stress they are not meant to represent a robust errors but rather as a semi-quantitative estimates, allowing to visualize the scattering of the monomeric values. Note that the statistical error in the fitted curves are substantially smaller than the error for the individual data points, because the fit effectively incorporates many data points at different $\Delta\Phi$, whereas the error for each data point is based on only $N_{\text{m}} = 4$ values.

The distance between the arginine and the histidine $d_{\text{R-H}}$ in the aromatic/arginine (ar/R) constriction site was defined as the smallest distance between the C_ζ atom of the arginine to the nearest heavy atom of the histidine. $d_{\text{R-H}}$ equals 6.30Å and 5.88Å in the crystal structures of AQP1 and AQP4, respectively (1, 19).

Simulations with the conserved arginine restrained in open/closed state

To confirm that the arginine state indeed regulates p_f we have carried out additional simulations of the stacked AQP1 and AQP4 system at zero membrane potential, but with the arginine of the ar/R region restrained in the open or in the closed state. The restraint was implemented by

applying an additional harmonic potential between the C_ζ atom of the arginine, and the N_ϵ of the histidine (Arg195/His180 for AQP1 and Arg216/His201 for AQP4). The equilibrium distance of the harmonic potential was set to 4.3Å and 6.1Å for the closed and open state, respetively. These values were motivated from the distribution of $d_{\text{R-H}}$ in simulations which clearly switched between open and closed states (compare also Figure 3A). The force constant was set to $2800\,\text{kJ}\,\text{mol}^{-1}\text{nm}^{-2}$, because that force constant allows for thermal fluctuations on the same scale as found in the open and closed state in unrestrained simulations. Each of the four setups, that is AQP1 and AQP4, either restrained in the open or closed state, was simulated 3 times for 30 ns. Because 8 monomers are present in each simulation setup, p_f was averaged from 24 independent values. The p_f values derived from these simulations are (in $10^{-14}\,\text{cm}^3/\text{s}$)

- AQP1, open: 3.0 ± 0.7;
- AQP1, closed: 0.13 ± 0.02;
- AQP4, open: 6.8 ± 0.6;
- AQP4, closed: 0.16 ± 0.02.

Hence, the arginine position strongly regulates the water flux, suggesting that the voltage tuning of P_{open} (Fig. 2E and 3B) consequently also tunes p_f.

Supplementary material references

[1] Ho, J. D., R. Yeh, A. Sandstrom, I. Chorny, W. E. C. Harries, R. A. Robbins, L. J. W. Miercke, and R. M. Stroud, 2009. Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance. *Proc Natl Acad Sci U S A* 106:7437–7442.

[2] Wolf, M. G., M. Hoefling, C. Aponte-Santamaría, H. Grubmüller, and G. Groenhof, 2010. g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J Comput Chem* 31:2169–2174.

[3] Hub, J. S., and B. L. de Groot, 2008. Mechanism of selectivity in aquaporins and aquaglyceroporins. *Proc Natl Acad Sci U S A* 105:1198–1203.

[4] Van der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, 2005. GROMACS: Fast, Flexible and Free. *J. Comp. Chem.* 26:701–1719.

[5] Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl, 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4:435–447.

[6] Jorgensen, W. L., D. S. Maxwell, and J. Tirado-Rives, 1996. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* 118:11225–11236.

[7] Kaminski, G. A., R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, 2001. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* 105:6474–6487.

[8] Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

[9] Berger, O., O. Edholm, and F. Jähnig, 1997. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* 72:2002–2013.

[10] Darden, T., D. York, and L. Pedersen, 1993. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

[11] Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, 1995. A smooth particle mesh ewald potential. *J. Chem. Phys.* 103:8577–8592.

[12] Miyamoto, S., and P. A. Kollman, 1992. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithms for Rigid Water Models. *J. Comp. Chem.* 13:952–962.

[13] Hess, B., 2008. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* 4:116122.

[14] Feenstra, K. A., B. Hess, and H. J. C. Berendsen, 1999. Improving Efficiency of large time-scale of Molecular Dynamics Simulations of Hydrogen-rich Systems. *J. Comput. Chem.* 20:786–798.

[15] Bussi, G., D. Donadio, and M. Parrinello, 2007. Canonical sampling through velocity rescaling. *J Chem Phys* 126:014101.

[16] Parrinello, M., and A. Rahman, 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* 52:7182–7190.

[17] Zhu, F., E. Tajkhorshid, and K. Schulten, 2004. Collective diffusion model for water permeation through microscopic channels. *Phys Rev Lett* 93:224501.

[18] Rubin, D. B., 1981. The Bayesian Bootstrap. *Ann. Statist.* 9:130–134.

[19] Sui, H., B.-G. Han, J. K. Lee, P. Walian, and B. K. Jap, 2001. Structural basis of water-specific transport through the AQP1 water channel. *Nature* 414:872–878.