Supporting Online Material for "Simulation of Electrokinetics at the nanoscale: inversion of selectivity in a bio-nanochannel".

Marcel Aguilella-Arzo
Biophysics Group, Department of Physics, Universitat Jaume I, 12080 Castelló, Spain.

Carles Calero and Jordi Faraudo*
Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus de la UAB, E-08193 Bellaterra, Spain
(Dated: October 5, 2010)

I. SIMULATION METHODS

A. Simulation algorithms and force field

We employed the software package NAMD version 2.6 for Linux-ia64[1] to carry out all the Molecular Dynamics simulations. The force-field for protein-lipid simulations supplied by NAMD2 was used, which is a combination of the CHARMM22 force field parameterized to describe protein systems and the CHARMM27 force field, parameterized to describe lipid systems in aqueous media. Water was modeled using the TIP3P model which is the standard choice in the CHARMM force field.

In all simulations we employed periodic boundary conditions in all directions. Lennard-Jones interactions were computed using a smooth $(10-12\mathring{A})$ cutoff, as is standard in NAMD2 simulations. The electrostatic interactions were calculated using the particle-mesh Ewald (PME) method with a precision of 10^{-6} using a $128 \times 128 \times 128$ grid and a $12\mathring{A}$ cutoff for the real space calculation. This unusually large grid was chosen in order to avoid significant spurious drift in the center of mass of the system, observed when using less dense grids. In all our choices we try to ensure very high standards in our calculation of electrostatic interactions which are particularly important in the problem under study.

As described in the next subsections, we have performed simulations under different conditions (NVT, NpT and NVT with external electric field ensembles) In all cases, the temperature T was controlled by using the Langevin thermostat implemented in NAMD[1], which is highly efficient for simulations of large systems. The parameters for the thermostat were standard for protein simulations. We have employed a damping coefficient of 1 ps^{-1} and the Langevin forces were applied to all atoms except for hydrogens. In the case of NpT simulations, the NAMD2 software controls the pressure by using a Langevin piston Nose-Hoover method (see[1] for details). The value of the pressure applied by the piston was set to $p = 1$atm, with a period of the oscillations of 0.1ps and a relaxation constant of 0.05 ps.

In all cases, the equations of motion were solved using a multiple time step in order to speed our very slow simulations. the values employed for the time step were those recommended for the kind of simulations performed here. These are a basic time step of 2fs and a time step of 4fs for the evaluation of k-space contribution to the long range electrostatic forces in the PME technique.

We have carefully checked that our choice of parameters correctly reproduce the bulk properties of the electrolytes employed in the simulations. The results are discussed in detail in Ref.[2]. In particular, we have found an excellent agreement between simulated and measured values of electrical conductivity for solutions of 1M of KCl, 1M of $MgCl_2$ and a mixture containing 1M of KCl and 1M of $MgCl_2$. In the case of divalent cations, to achieve a good agreement between simulated and experimental values the use of a large simulation box is required, as considered in this work.

B. System setup

The building of the initial configurations for the different simulations was made with the Visual Molecular Dynamics (VMD) Software[3]. The atomic coordinates of the OmpF trimer ionic channel were downloaded from the Protein Data Bank database http://www.pdb.org. From the different OmpF structures (with different resolutions and different experimental techniques) available at this database, we have selected the structure with code 2OMF[4], since we found that it is the available structure with a smaller number of missing atomic coordinates (atomic coordinates not resolved in the X-ray procedure and therefore not included in the datafile). As usual with X-ray structures, the downloaded PDB data file contains also spurious atomic coordinates corresponding to atoms from chemical compounds used in the generation of the crystal structures (see details in Ref.[4]). Since these coordinates were clearly identified in the PDB file, they were removed before further processing of the PDB file. Also, a few atomic coordinates of the protein structure were missing in the downloaded structure (oxygen atoms for the residue PHE340). These missing coordinates from the protein structure were reconstructed using the psfgen structure building module supplied with VMD and NAMD2 packages[1,3]. The protonation states of the residues in the channel employed in the simulations were determined according to previous detailed studies on the protonation states of OmpF at pH=7[5,6]. The resulting

TABLE I: **Size and content of the Simulation box for each simulation**. In this Table we show the length of the simulation box (in nm) in each direction after the NpT equilibration procedure and the number of water molecules, number of ions and total number of atoms employed in each simulation. We also show the duration of the production runs and the computer time required per ns of production run using 64 processors "Itanium Montvale" at the CESGA Supercomputing facility.

	Size of Simulation Box			Num. of Water	Number of ions			Atoms in Box	Production	Computer time
Electrolyte	L_x (nm)	L_y (nm)	L_z (nm)	molecules	K^+	Mg^{2+}	Cl^-	(Total)	run	(64 procs)
1M KCl	17.225	17.206	14.530	108270	2033	-	2000	446301	24.9 ns	1.12 days/ns
1M $MgCl_2$	17.251	17.053	14.308	106288	-	2016	3999	442337	36.8 ns	1.22 days/ns
1M $MgCl_2$ + 1M KCl	17.209	17.247	14.113	102256	2016	2016	6015	434273	31.6 ns	1.23 days/ns

trimeric structure is shown in Figure 2 in the main paper, which clearly shows the complexity of the structure of the channel pores.

The OmpF trimer, with its symmetry axis aligned along the z axis, was embedded into a large POPC lipid membrane (17.71 nm long both in the x and y axis) generated with the Membrane Builder supplied with the VMD package. All lipid molecules from the membrane which overlapped with the channel were removed from the system. The resulting system had 763 lipid molecules (with a total of 102242 atoms) and a trimeric protein channel (with 15216 atoms). After that, the system channel-membrane was solvated in a box of preequilibrated TIP3P water molecules and a number of cations (Mg^{2+} and/or K^+) and anions (Cl^-) were added to achieve the desired electrolyte concentration. We have constructed three different initial configurations of the channel-membrane system, each one corresponding to a different ionic solution: 1M of KCl, 1M of $MgCl_2$ and a mixture containing 1M of KCl and 1M of $MgCl_2$. The number of ions and water molecules employed in each case are shown in Table I. The dimensions of the resulting system were $177.1 \times 177.1 \times 145 \text{Å}^3$ for all cases before the equilibration procedure.

C. Equilibration Procedure

We have employed a carefully designed equilibration procedure, based on previous simulations of ion channels[7,8]. In all simulations, the equilibration procedure has the following steps:

- *Energy minimization:* In order to avoid undesired overlaps between atoms, we perform an energy minimization with NAMD of the initial configuration until a constant value for the energy is obtained (typically 2000 steps).
- *Thermalization of the initial configuration*. The configuration resulting from energy minimization is thermalized by first running a NVT simulation at $T = 100$K during 100ps followed by another NVT run at $T = 296$K during 100ps.
- *Pressurization of the system* In order to ensure that the system is at a pressure of $p = 1$ atm, we perform two different simulation runs in the NpT ensemble ($T = 296$K, $p = 1$ atm). In the first NpT run, the system was simulated during 1 ns in the isotropic NpT ensemble ($p = 1$atm, $T = 296$K). In order to avoid problems with the protein, we have the protein restrained during the simulation. After this run, a second NpT run is performed during 3 ns. In this case, the protein is unrestrained and the simulation box is allowed to change in size only in the z direction (perpendicular to the membrane plane). The objective of this run is to adjust the pressure of the electrolyte solution in contact with the membrane, as happens in the experimental situation.
- *Development of ionic current*. After equilibration of the initial configuration, we have performed a run of several ns until an ionic current is developed across the protein channel. These simulations were performed under NVT conditions with an external field with a value equal to that employed in the production runs (see discussion below).

D. Production runs

The production runs were performed in the NVT ensemble under conditions of $T = 296$K and an external field of 14.22 mV/nm magnitude pointing in the direction of the negative z axis (see Figure 3 in the main paper). This value of the electrostatic field corresponds to a potential difference from the top to the bottom of the simulation box of ~ 200 mV. Since the resistivity of the electrolyte solutions is extremely low and the resistance of the protein channel is very high, we expect that almost all the potential drop occurs in the region containing the membrane and the protein. Let us consider for example the case of KCl. The electrical conductivity of 1M KCl electrolyte as obtained from MD simulations[2] is about 11.9 S/m. The water electrolyte slab in the simulation for the protein bathed with 1 M KCl is ≈ 10 nm long (the size of the system excluding the ≈ 4nm length of the membrane+protein system) and has a cross section of 17×17 nm^2, which gives an electrical resistance of about $R \approx 3 \times 10^6\ \Omega$.

TABLE II: Ionic fluxes (number of ions observed to cross the OmpF channel during production runs of duration $t_{\rm run}$) and occupancy numbers defined as the average number of ions of each species inside the protein channel (statistical errors are estimated from 2σ). The results correspond to three different conditions: 1 M KCl, 1 M $MgCl_2$ and a mixture of 1 M $MgCl_2$ and 1 M KCl.

	$t_{\rm run}$	Ionic Flux			Channel Occupancy		
	(ns)	Cl^-	K^+	Mg^{2+}	Cl^-	K^+	Mg^{2+}
KCl	24.9	38	47	-	21.0 ± 0.2	35.5 ± 0.2	-
$MgCl_2$	36.8	32	-	1	47.1 ± 0.9	-	33.9 ± 0.3
Mixture	31.6	35	14	7	24.4 ± 0.2	7.5 ± 0.2	7.2 ± 0.2

Previous MD simulations[8] give a conductance for the OmpF channel between 2-3 nS in 1 M KCl, which corresponds to 3-5$\times 10^8$ Ω. Therefore, the electrical resistance of the electrolyte slab is about 2 orders of magnitude smaller than the electrical resistance of the channel, so it can be safely assumed that the drop of electrostatic potential in the electrolyte solution is negligible as compared with the drop across the protein+membrane system.

We note here that previous MD simulations[7,8] of protein channels were performed by using much larger values for the electrostatic potential drop (around 1 V) in order to obtain good statistics for the flow of ions. Since these high potentials cannot be obtained experimentally, we have decided to employ a smaller electric field in order to obtain a potential drop which is realistic yet gives a ion flux which can be observed in long simulations. The duration of the production runs was selected by ensuring that good statistics in the ion flow were obtained (see next section). In all cases, very long simulations were required, as shown in Table 1. It has to be emphasized that, due to the large number of atoms in the simulation box and the high precision of the calculations, our production runs are extremely slow, so the simulations runs lasted during several months.

II. FURTHER DETAILS IN THE ANALYSIS OF SIMULATION RESULTS

A. Flux of ions

In order to characterize the ionic current, we have monitored the number of ions of each species crossing the protein. These numbers of ions were counted by following the individual trajectories of each ion. An ion was considered to cross the channel if it is initially at one side of the membrane and ends at the opposite side of the membrane bulk electrolyte following a path inside the protein channel. Some ions were found to enter inside the channel and remain inside the channel (this is particularly frequent in the case of Mg^{2+}) or to return again to their initial side of the membrane. These ions were not counted in the number of ions crossing the channel. The accumulated results for each simulation (only during the production run) are shown in Figures 3 and 5 in the main paper. The accumulated numbers at the end of the simulation run are also given in Table II. Let us remark here that all crossing events observed in our simulations follow the direction expected from the applied electric field (cations cross the channel in the direction of the applied electric field and anions in the opposite direction) so all observed crossing events contribute to an electric current in the direction of the applied external field.

B. Number and distribution of ions inside the channel

In the previous subsection, we have computed ionic currents (i.e. the flow of ions across the channel), which characterize the dynamical aspect of the interaction between ions and the protein channel. Now in this subsection we analyze the static aspects of the interaction of the ions with the protein channel. The first question to analyze is the number of ions inside the protein channel. As a definition for the interior of the pore we consider a region 4 nm long given by the β barrel structure (3.2 nm long) plus a wider "entrance" region (0.4 nm long) situated at each ending of the monomer (note that this definition is close to that employed in previous works[9]).

The occupation numbers of the channel, reported in the main paper (Table 1) and repeated here in Table II, were computed by time averaging (during the production runs) the number of ions of each species inside the interior of the channel. The distribution of these ions inside the channel has also interesting features. The most interesting feature is that monovalent cations and anions show well separated pathways (as discovered by Roux[9] using simulations and recently confirmed by experimental methods[10]). This effect is due to the strong transversal field originated in the constriction region (see Figure 1). These different pathways can be nicely demonstrated from isodensity plots of cations and anions inside the channel[8]. This feature is of course also observed in our simulations with 1M KCl, as shown in Figure 1. The same feature is also observed in simulations with 1M of $MgCl_2$: cations and anions follow opposite pathways and in fact Cl^- ions follow the same pathway described in KCl simulations. Also, as discussed in the main text (see Figure 3a in the main paper), the distribution of Mg^{2+} ions shows a new feature not observed for monovalent ions: the presence of very strong inhomogeneities in the cation concentration near acidic groups located at the

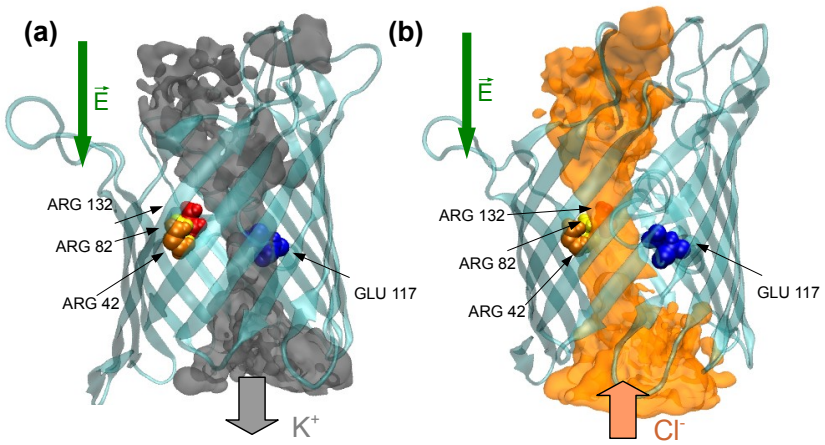

FIG. 1: **Pathways of cations and anions in 1M KCl simulations** Isodensity surfaces of ions obtained from our simulations obtained as described in Ref[8]. In the case of Cl^-, the isodensity surface shown here corresponds to 0.57M. In the case of K^+, the isodensity surface corresponds to 0.83M. It is clear that cations and anions occupy different regions inside the channel, as first discussed by Roux[9]. This effect is largely due to the strong transversal field originated in the constriction region containing two negatively charged groups (ASP113 and GLU117 acids shown in blue) opposite to three positively charged groups (ARG42, ARG82, ARG132). The picture has been produced using VMD[3].

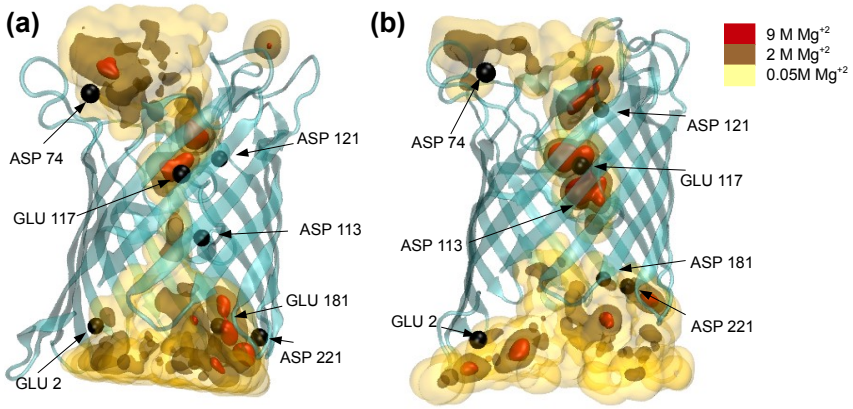

FIG. 2: **Interaction between Mg^{2+} cations and the protein channel in different electrolyte conditions** In this figure we show the concentration of Mg^{2+} cations inside one of the monomers of the protein channel for (a) simulations with 1M $MgCl_2$ and (b) simulations with a mixture of 1M KCl and 1M $MgCl_2$. We show three different isosurfaces averaged over a 10 ns trajectory fragment of the production run. Strong inhomogeneities (with concentrations 1 order of magnitude larger than bulk concentration) are found near positively charged groups located at the protein surface (schematically indicated as black balls). The picture has been produced using VMD[3]

protein walls. We pointed out in the main text that these strong correlations are also observed in the simulations containing a mixture of 1M $MgCl_2$ and 1M KCl. This is illustrated in Figure 2.

C. Examples of Binding

As we said in the main text, Mg^{2+} cations spend long times near negatively charged residues. Although we do not have enough statistics to perform a rigorous analysis of residence and binding times, we can say that the observed trajectories show binding times of the order of 10 ns. Several illustrative examples are shown in Figure 3.

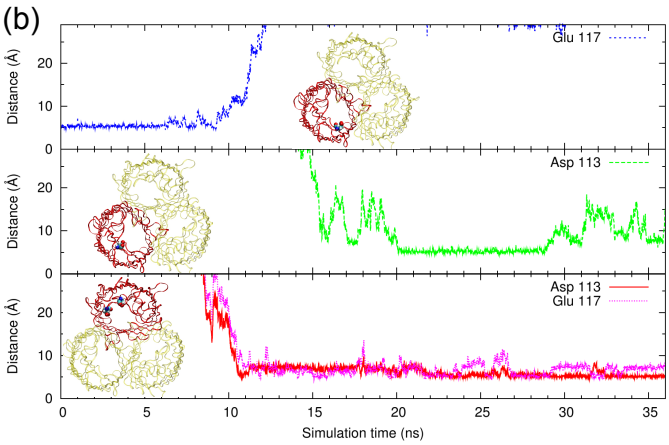

FIG. 3: Examples of long residence times of Mg^{2+} ions near negatively charged residues in simulations with 1 M $MgCl_2$. We show here the time evolution of the distance between three selected Mg^{2+} ions and certain protein residues indicated in the onsets. Top panel: ion close to the residue Glu117 of monomer P2 at the start of the production run which remains there roughly 10 ns. Medium panel: ion initially at bulk solution enters inside the P2 monomer of the channel and remains in contact with the Asp113 residue during 9 ns. Bottom panel: ion initially at bulk solution enters inside the P3 monomer of the channel and remains in contact with both the Asp113 and Glu 117 residues (no detachment was observed).

* Electronic address: jfaraudo@icmab.es

[1] Phillips, J. E., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L. & Schulten, K., Scalable Molecular Dynamics with NAMD *J. Comp. Chem* **26**, 1781-1802 (2005)

[2] Calero, C., Aguilella-Arzo M. & Faraudo, J. *Molecular Simulation* (submitted, preprint available at arXiv:1005.2857v1).

[3] Humphrey, W., Dalke, A. & Schulten, D. J. VMD - Visual Molecular Dynamics. *Molec. Graphics* **14.1**, 33-38 (1996)

[4] Cowan, S.W. The Refined Structure of Ompf Porin from E.Coli at 2.4 Angstroms Resolution. *Protein Data Bank* freely available at http://www.pdb.org code 2OMF (DOI:10.2210/pdb2omf/pdb).

[5] Varma, S., Chiu, S. & Jakobsson, E. The Influence of Amino Acid Protonation States on Molecular Dynamics Simulations of the Bacterial Porin OmpF. *Biophys. J.* **60**, 112-123 (2006).

[6] Alcaraz, A., E. M. Nestorovich, M. Aguilella-Arzo, V. M. Aguilella and S. M. Bezrukov. 2004. Salting out the ionic selectivity of a wide channel: The asymmetry of OmpF. *Biophys. J.* 87:943-957.

[7] Aksimentiev, A. & Schulten, K. Imaging α-Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability, and the Electrostatic Potential Map. *Biophys. J.* **88**, 3745-3761 (2005).

[8] Pezeshki, S., Chimerel, C., Bessonov, A. N., Winterhalter, M. & Kleinekathofer, U. Understanding Ion Conductance on a Molecular Level: An All-Atom Modeling of the Bacterial Porin OmpF. *Biophys. J.* **97**, 1898-1906 (2009).

[9] Im, W. & Roux, B. Ion permeation and Selectivity of OmpF porin: A theoretical Study based on Molecular Dyanmics. *J. Th. Biol.* **322**, 851-869 (2002).

[10] Dhakshnamoorthy B., Raychaudhury S, Blachowicz L. and Roux B. Cation-selective Pathway of OmpF Porin Revealed by Anomalous X-ray Diffraction.*J. Mol. Biol.* **396**, 293-300 (2010).

Description of Movies available for the paper "Simulation of Electrokinetics at the nanoscale: inversion of selectivity in a bio-nanochannel".

Marcel Aguilella-Arzo
Biophysics Group, Department of Physics, Universitat Jaume I, 12080 Castelló, Spain.

Carles Calero and Jordi Faraudo*
Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus de la UAB, E-08193 Bellaterra, Spain
(Dated: September 1, 2010)

I. SIMULATIONS OF THE OMPF CHANNEL BATHED BY 1M KCL

OmpF ionic channel in contact with a 1M KCl ionic solution in the presence of an electric field applied along the axis of the channel.

A. KClcrossK.mov

Potassium ion crossing one of the monomers of the OmpF ionic channel in the direction of the applied electric field.
The potassium ion is represented in blue and the chloride ions in its vicinity are represented in pink. The structure of the monomer is schematically represented by a blue flattened ribbon. In the region where the channel is narrower, the so-called constriction, relevant charged residues of the protein are represented with atomic detail, the positively charged residues (arginine 42, arginine 82 and arginine 132) in green and the negatively charged ones (aspartic acid 113, glutamic acid 117) in orange.

II. SIMULATIONS OF THE OMPF CHANNEL BATHED BY 1M $MGCL_2$

OmpF ionic channel in contact with a 1M $MgCl_2$ ionic solution in the presence of an electric field applied along the axis of the channel.

A. MgCl2crossMg.mov

Magnesium ion crossing one of the monomers of the OmpF channel in the direction of the applied electric field.
The magnesium ion is represented in blue, and the chloride ions in its vicinity in red. The structure of the monomer is schematically represented by a blue flattened ribbon. In the region where the channel is narrower, the so-called constriction, relevant charged residues of the protein are represented with atomic detail, the positively charged residues (arginine 42, arginine 82 and arginine 132) in green and the negatively charged ones (aspartic acid 113, glutamic acid 117) in orange.

B. MgCl2crossCl.mov

Chloride ion crossing one of the monomers of the OmpF ionic channel in the direction opposite to the applied electric field.
The red particle represents the chloride ion which crosses the channel and the blue particles are magnesium ions in the vicinity of the crossing chloride. The structure of the monomer is schematically represented by a blue flattened ribbon. In the region where the channel is narrower, the so-called constriction, relevant charged residues of the protein are represented with atomic detail, the positively charged residues (arginine 42, arginine 82 and arginine 132) in green and the negatively charged ones (aspartic acid 113, glutamic acid 117) in orange.

* Electronic address: jfaraudo@icmab.es