PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Macroscopic electric field inside water-filled biological nanopores

Silvia Acosta Gutierrez, [†] *Igor Bodrenko,*[†] *Mariano Andrea Scorciapino,*[‡] *Matteo Ceccarelli*[†*]

†Department of Physics, University of Cagliari, Cittadella universitaria di Monserrato, S.P.8 – km 0.700, 09042 Monserrato (CA), Italy

‡Department of Biomedical Sciences, Biochemistry Unit, University of Cagliari, Cittadella universitaria di Monserrato, S.P.8 – km 0.700, 09042 Monserrato (CA), Italy

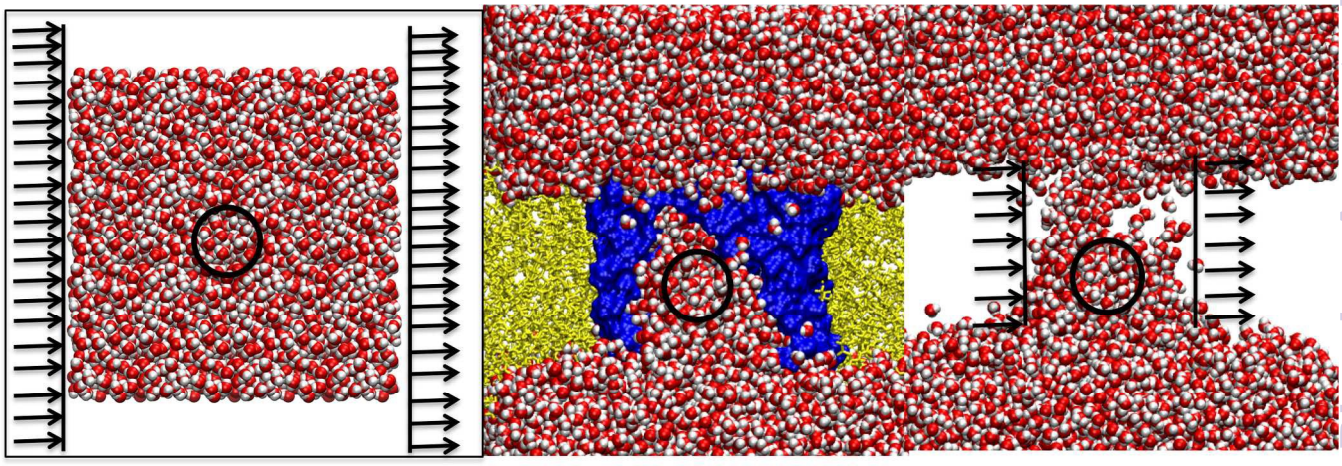
KEYWORDS Water-filled channels, electric field, filtering, polar molecules, bacterial resistance, molecular dynamics simulations

ABSTRACT

Multi-drug resistance bacteria are a challenging problem of contemporary medicine. This is particularly critical for Gram-negative bacteria, where antibiotics are hindered by the outer membrane to reach internal targets. Here more polar antibiotics make use of nanometric water-filled channels to permeate inside. We present in this work a computational all-atom approach, using water as a probe, for the calculation of the macroscopic electric field inside water-filled channels. The method allows one to compare not only different systems but also the same system under different conditions, such as pH and ion concentrations. This provides a detailed picture of

electrostatics in biological nanopores shedding more light on how the charged residues of proteins determine electric field inside, and also how medium can tune it. These details are central to unveil the filtering mechanism behind the permeation of small polar molecules through nanometric water-filled channels.

GRAPHICAL ABSTRACT



We used water as a natural probe to evaluate the macroscopic electric field inside nanometric water-filled channels

INTRODUCTION

Multi-drug resistance bacteria are a challenging problem of contemporary medicine[1]. New synergies and a molecular-based framework are necessary for identifying and developing the next generation of antinfectives[2,3]. This is particularly critical for Gram-negative bacteria where the presence of the additional outer membrane (OM) hinders any antibiotics to access internal targets[4]. In the OM, general diffusion porins are expressed to facilitate the entry of small polar molecules (less than 1,000 Da), and today porins are considered to be the main pathway for polar antibiotics for overcoming the OM barrier[5,6]. Bacteria can develop resistance by reducing the

OM permeability, either by modulating the expression of porins, or by selecting key residues mutations that alter the permeability of the porins themselves[7]. Further, no direct and robust methods to measure permeation/accumulation of drugs are available at present[8], neither *in vitro* nor *in vivo*. This technological gap reduced enormously the research and investment in the field[1,2], leaving the pipeline empty. Therefore, the discovery of new effective antibiotics for Gram-negative bacteria passes through the determination of the mechanism controlling their permeation through porins[9]. The internal electrostatics of porins seem to play a major role[10] and simulations seem to have the potentiality to unravel such molecular mechanism[11,12].

Since the determination of their high-resolution X-ray structures, porin channels showed an hourglass shape and interesting electrostatic properties. Both OmpF and OmpC, the major porins expressed in *Escherichia coli*, show a well-separated distribution of charged residues in the nanometric eyelet region (7x11 Å)[13]. Cell-free electrophysiology experiments at single molecule level pointed out the key role of electrostatics in the interaction of polar molecules with these protein channels[14,15]. Despite similar in size[7], their expression level is usually not the same but profoundly affected by osmolarity, pH and growth conditions[16]. In particular, while OmpF is prevalent in low osmolarity media, OmpC is the most expressed outer membrane porin in high osmolarity.

Several theoretical attempts at characterizing the electrostatics inside porins have been reported[17-20]. Starting from the pioneering work of Karshikoff[17], which employed the Poisson-Boltzmann equation, till the valuable work by Im and Roux[18], which, by using molecular dynamics (MD) simulations, described a screw-like electrostatics for OmpF, although at a rather high ion concentration (1M). In the present work we propose a strategy to quantify the macroscopic electric field inside these nanometric water-filled channels from all-atom MD

simulation. In order to show the detail level achievable, we selected four porins of the same family, namely, OmpF, OmpC and two mutants from *E. coli* clinical strains, OmpC20 and OmpC33, with assessable differences in susceptibility to polar molecules, in spite of the comparable pore size[7]. The same series was recently employed by our group to show the diverse filtering of carbapenem antibiotics[21] and the modulation of water ordering[10] upon charged residue mutations.

The ordering of water molecules is a raising question in many membrane proteins controlling key processes in the cell, such as aquaporins[22,23] and GPCR systems[24]. Here we have taken a step forward by quantifying the internal electric field, not just probing its effects on the water/ions as already done. We have made use of water molecules, since, naturally filling the selected porins, they can be considered an optimal and intrinsic probe. In addition, the small size of water is reflected by a rather low rotational correlation time[25], whereas its relatively large electric dipole ensures sensitivity to the channel's electrostatics. The dielectric constant of water inside the channel is decisive in the calculation of the internal electric field. The non-linear response of water to an applied electric field is widely recognized[26-28] and such nonlinearity of water polarization is reflected by the decrease of the dielectric constant with increasing the applied electric field, the so-called water dielectric saturation effect. In the case of the lumen of protein channels, one faces two problems, the non-linearity and non-uniformity of the media, remarkably complicating the selection of an appropriate dielectric constant[29].

In the present study, the local macroscopic electric field is directly calculated from all-atom MD simulations by using the transfer function that contains the response of the media to the applied electric field. Importantly, the approach is applied without any initial guess of the dielectric constant inside the pore. This allowed us to characterize the electrostatics of the

selected porins in details by calculating the local intrinsic electric field along the channel axis. Subtle differences due to very few amino acid mutations have been identified and highlighted and, with the same level of details, it was possible to investigate the influence of physiologically important modifications of the media, such as the ion concentration and pH. The method is not only relevant to protein or water-filled channels in general, but can be easily extended to solvated proteins for the calculation of the electric field along the external surface or at a water-filled cavity. The only prerequisite is the presence of enough water molecules to be used as probe.

METHODS

Theoretical Background

Complex bio-molecular systems like protein channels are strongly inhomogeneous, as they are embedded into a lipid membrane and solvated. Therefore, the macroscopic (Maxwell) electrostatic field, $\boldsymbol{E}(\boldsymbol{r})$, must be considered locally. According to the standard paradigm of the continuum classical electrostatic theory, the microscopic electric field, $\boldsymbol{e}$, must be statistically averaged over a small space (e.g., a sphere) around the point $\boldsymbol{r}$, where one wants to calculate it, $\boldsymbol{E}(\boldsymbol{r}) = \langle \boldsymbol{e} \rangle_{\boldsymbol{r}}$. One should also determine the local polarization density, $\boldsymbol{P}(\boldsymbol{r})$, i.e., the average dipole moment of the sphere per unit volume, the local electric displacement vector, $\boldsymbol{D}(\boldsymbol{r}) = \boldsymbol{E}(\boldsymbol{r}) + 4\pi\boldsymbol{P}(\boldsymbol{r})$, and the density, $\rho_f(\boldsymbol{r})$, of 'free' charges (unbound or uncompensated, which do not contribute to $\boldsymbol{P}$). All of these quantities are related each other through the Maxwell equations: $\boldsymbol{\nabla} \cdot \boldsymbol{D} = 4\pi\rho_f(\boldsymbol{r})$; $\boldsymbol{\nabla} \times \boldsymbol{E} = 0$. If one knows $\rho_f(\boldsymbol{r})$ as well as the dielectric response function, $\boldsymbol{P}(\boldsymbol{E}, \boldsymbol{r})$, the electric field may be obtained by solving the Maxwell equations with specified boundary conditions.

It is not trivial, however, to determine (both experimentally and theoretically) the local dielectric response for a strongly inhomogeneous system: in general, $\boldsymbol{P}(\boldsymbol{E},\boldsymbol{r})$ is anisotropic (vectors $\boldsymbol{P}$ and $\boldsymbol{E}$ are not parallel) and non-linear; moreover, the response may be non-local, i.e., $\boldsymbol{P}$, is a non-local functional of $\boldsymbol{E}(\boldsymbol{r})$.

In the widely used local, linear and isotropic approximation, one defines the local dielectric constant, $\varepsilon(\boldsymbol{r})$, so that $\boldsymbol{D} = \varepsilon(\boldsymbol{r})\boldsymbol{E}$. Then, the Maxwell equations are reduced to the Poisson equation, $\boldsymbol{\nabla} \cdot \varepsilon(\boldsymbol{r})\boldsymbol{\nabla}\varphi = -4\pi\rho_f(\boldsymbol{r})$, for the electrostatic potential, $\varphi(\boldsymbol{r})$, and the electric field reads, $\boldsymbol{E} = -\boldsymbol{\nabla}\varphi$. One usually uses a piecewise constant function model for $\varepsilon(\boldsymbol{r})$. Thus, in the case of solvated membrane proteins, e.g. in [17], $\varepsilon_p = 4$ is used for the protein part, $\varepsilon_s = 80 -$ for the solvent (water), etc. The non-linear electric field effects on the solvent due to the saturation of the dielectric response in a strong field was considered by taking an effective (smaller) dielectric constant for the regions of strong field[30]. In another study[31], the non-linear dielectric function, $\varepsilon_s(E)$, was approximated by a Langevin-type function with the parameters adjusted to the bulk solvent response and then the non-linear Poisson equation was solved. There, the spatial dependence of the dielectric constant of the solvent comes from the non-linear response and the spatially varying electric field, $\varepsilon_s(\boldsymbol{r}) = \varepsilon_s(E(\boldsymbol{r}))$. The free charge density, $\rho_f(\boldsymbol{r})$, is composed of the fixed charged groups in the proteins and the mobile ions. The ion density at equilibrium is treated by assuming the Boltzmann statistics (the Poisson-Boltzmann method [17]or its linearized version, i.e. the Debye-Hückel method for low ion concentration). If the ion density is not at equilibrium, the diffusion-drift (Nernst-Planck) theory or a more general Brownian dynamics method should be employed [18], coupled with the Poisson equation to calculate the ion density and the electrostatic potential simultaneously.

All the continuum electrostatics methods briefly presented hereinbefore require strong assumptions about the dielectric response of the inhomogeneous medium as well as about the free ions distribution. Differently, it is important to stress here that within the approach presented in this work, the macroscopic electrostatic field is calculated by using all-atom MD simulations with minimal assumptions (besides the standard ones that are fundamentals to the classical MD). In the following, we will consider the electric field in the regions of the system occupied by the solvent (water, in our case).

For the bulk (homogeneous) solvent, the polarization density is completely determined by the thermodynamic parameters, - the density (ρ), the temperature (T), and the electric (macroscopic) field (E), $\boldsymbol{P} = \boldsymbol{P}(\rho, T, \boldsymbol{E})$. Then, we suggest that the local dielectric response over the averaging space in the inhomogeneous regions of the system is the same as for the bulk. This assumption is also accepted and argued in the Poisson equation based applications mentioned above[17,18,30,31]. This is the only approximation we do as the ions, the solvent, the protein, the membrane, etc., are treated at the all-atom level. Within this approximation, the ordering (alignment) of water molecules in the nanopore is assumed to come from the local average electric field in the probe. The surface effects, like the different response of the molecules on the boundary of the probe or the hydrogen bonds of some water molecules with the protein residues are formally not taken into account. However, the local macroscopic electric field is defined within a coarse-grained model that neglects any inhomogeneity within the probe, so that the inhomogeneity in the response should also be neglected. Of course, the number N of water molecules on the probe should be large enough as the surface effects scale as N^2 while the main volume effects scale as N^3. One may estimate the simultaneous effect of all kinds of inhomogeneity, including the surface effects, by calculating the local macroscopic field with various radii of the probe.

The approach we propose is the following. First, the polarization density of the solvent has to be calculated in a localized spherical probe along the MD trajectory,

$$\boldsymbol{P} = \frac{\langle \boldsymbol{\mu} \rangle}{N_w v_w}, \qquad (3)$$

where $\langle \boldsymbol{\mu} \rangle$ is the average dipole moment of the water molecules within the spherical probe, N_w is the average number of water molecules in the probe, and $v_w(\rho)$ is the average volume per solvent molecule at the given water density. Then, the electric field may be estimated by using the bulk response function that relates water polarization to the local Maxwell field $\boldsymbol{E}$ that has generated it, $\boldsymbol{P}(\rho, T, \boldsymbol{E})$, which is calculated within the same MD model. The shape of the probe is important for an inhomogeneous medium as the surface properties may contribute to the electrostatics through the standard boundary conditions for the normal and the tangential components of the electric field. This results in the dependence of the direction of the polarization vector of the probe on the shape of the latter, e.g., the polarization vector in the cylindrical or ellipsoidal probe may be non-parallel with the electric field. For that reason, the sphere seems to be the best choice, as the induced dipole moment is collinear with the electric field, assuming a local isotropic response.

In order to calculate the response function of the (bulk) solvent by using MD simulations of an NPT ensemble in an external electric field, one considers a water box under the same pressure and temperature as in the simulation of the channels. All solvent molecules (we used the TIP3P model for water) feel the same additional external and constant electric field $\boldsymbol{E}_{ext}$. Thus, the total Maxwell field should be the sum of $\boldsymbol{E}_{ext}$ and the averaged electric field E_{ind} generated by the aligned molecules, $\boldsymbol{E}=\boldsymbol{E}_{ind}+\boldsymbol{E}_{ext}$. The corresponding electrostatic potentials, φ, φ_{ind}, φ_{ext}, obey the Laplace equation as there are no 'free' charges in the system. The treatment of the electrostatic interactions within the Particle Mesh Ewald method[32,33] assumes the periodic (constant)

boundary conditions for electrostatic potential φ_{ind}, but not for the external one, $\varphi_{ext} = -(\boldsymbol{r} \cdot \boldsymbol{E}_{ext})$. The Laplace equation with the periodic (constant) boundary condition has only constant solution. Thus, the electric field, $\boldsymbol{E}_{ind}$, generated by the molecules in the box equals zero even if the net polarization density $\boldsymbol{P}$ of the box is not null. The total macroscopic field E then equals the external applied electric field, $\boldsymbol{E}=\boldsymbol{E}_{ext}$, which generates the polarization density $\boldsymbol{P}$.

This result may also be interpreted from the viewpoint of the macroscopic electrostatic theory. Indeed, the constant potential on the boundary corresponds to the conducting (or *tin-foil*) boundary conditions[34]. The conducting boundary generates the reaction field which exactly cancels out the field generated by the polarized solvent molecules, E_{ind}. The remaining macroscopic field is equal to the external applied one.

Finally, the dielectric response function can be calculated from a MD trajectory for a given macroscopic field $\boldsymbol{E}=\boldsymbol{E}_{ext}$ by using eq. 1, where $\langle\boldsymbol{\mu}\rangle(\boldsymbol{E}_{ext})$ is the average dipole moment of the molecules in the spherical probe.

For an inhomogeneous medium, the profile of the probe (the coarse-graining function) affects the statistical averaging of the microscopic charge density and of the microscopic field. In other words, the macroscopic charge density and the macroscopic polarization density depend on the coarse-graining function. Moreover, there are also additional terms in the statistically averaged Maxwell equations depending on the quadrupole and higher-order electrical moments densities [35]. The additional non-dipole terms vanish for the homogeneous fluids [36]. Also, for an inhomogeneous fluid, the quadrupole contribution to the macroscopic Maxwell equation disappears for the specific choice of the coarse-graining function – the one constant within the probe and equal zero out of its boundary [36]. Thus, the use of a constant coarse-graining function in our analysis is justified as the obtained macroscopic electric field obeys the standard Maxwell

equations of continuum electrostatics with common boundary conditions, additivity of the polarization density, etc.

We should also mention a different approach to average the microscopic electrostatic field presented in [37] and implemented in VMD[38]. The authors use gaussians to smooth atomic charges, solve the Poisson equation to obtain the corresponding electrostatic potential at each frame of the MD trajectory and then average the potential along the trajectory. This approach corresponds to a gaussian coarse-graining function. In the homogeneous regions (e.g., in the bulk solvent far outside the membrane) the averaged electrostatic field obtained with this method is the same as that calculated with the constant coarse-graining function in our approach. In the inhomogeneous regions (like those inside the channel), in general, the averaged electrostatic field does not obey the standard continuum Maxwell equations but those containing the quadrupole and the higher-order electric moments.

Computational Details

Production run in the NVT ensemble was performed through the starting from the systems equilibrated in [10], after the NVT run, a suitable number of water molecules in the OmpF and OmpC systems was replaced with ions to have the desired KCl concentration. In the case of OmpF in order to improve statistics four simulations of 100ns each were launched with different seeds and 1 microsecond was run after adding ions. For OmpC only 800ns were run with ions. In the case of pH5 simulations, residues 117 and 121 in the loop L3 were protonated according to the pK_a values evaluated using the PROPKA code[39,40]. All simulations were performed in the NVT ensemble using the ACEMD code[41] compiled for GPUs, by rescaling hydrogen mass to 4 au and increasing the time-step up to 4.0 fs[42]. The Langevin thermostat was used with 1 ps

damping time. SPME was used to treat the electrostatics as for the equilibration stage. The Amber99SB-ILDN force field [43]was used for the protein and lipids, and the TIP3P [44]for waters.

Water analysis was performed separately for each protein monomer. Water molecules within a cylinder of 80 Å height and 17 Å radius, centered in the center of mass of the monomer, were extracted from MD trajectories. Each cylinder was divided in 0.5 Å slices along the channel axis (z-axis) and the distribution of oxygen atoms along the MD trajectory was calculated for each slice, 160 in total. A sphere was then centered at the average position of oxygen over time for each slice, and the water molecules within such spherical probe were extracted for each trajectory frame. Spherical probes of three different size were used (radius, R = 4Å, 5Å and 7Å). The polarization density of each sphere was calculated as $P_{i\alpha} = \frac{1}{N_i v_w}\sum_{j=1}^{N_i}\mu_\alpha \quad \alpha = x, y, z$, where i goes from 1 to 160 and $v_w = 29.9\ \text{Å}^3$ is the volume of a single water molecule. Then, the polarization density of each sphere was averaged over the whole trajectory. Finally, because the protein channels under investigation are trimeric, trajectory averages were also averaged over the three equivalent monomers, which can be considered statistically independent systems in this context.

RESULTS AND DISCUSSION

Dielectric Response of Bulk Water

In order to calculate the dielectric response function of water, we performed different MD simulations in the NVT ensemble as described in the Methods, with 17 different values of the applied electric field, from 0.1 to 500 mV/Å (20 ns each production run). Polarization density was calculated in three different ways in order to test possible artifacts: i) polarization density of

the entire box, ii) spheres of different size centered in the simulation box, and iii) by using a set of spheres along the z axis. No significant differences were found. Results are shown in Fig. 1.

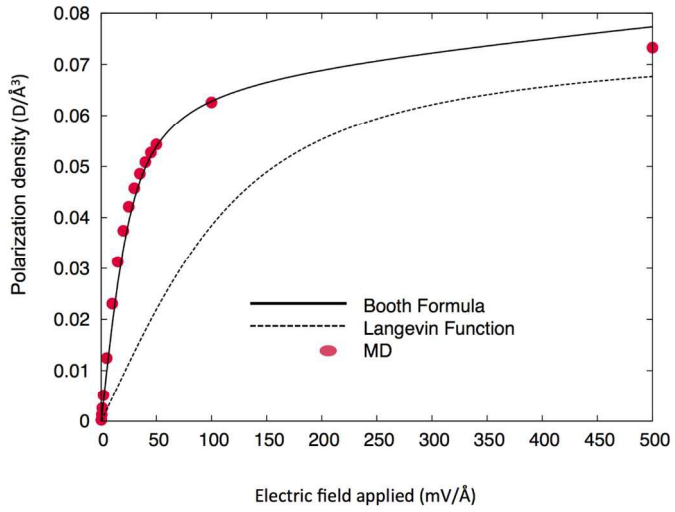


Figure 1. The dielectric response function of (bulk) water. The solid line corresponds to the Booth function[45], the dashed line to the Langevin function with parameters from[46]

At electric field values above 30-40 mV/Å, the calculated dielectric response of water becomes non-linear and starts to saturate at field strengths above 100 mV/Å. This behavior is in excellent agreement with Booth's model[45]

$$P_B(E) = \frac{(n^2-1)}{4\pi} E + \frac{\alpha(n^2+2)\mu}{4v_w} L\left(\frac{\beta(n^2+2)\mu E}{kT}\right) \qquad (1)$$

where $\alpha = \frac{28}{3\sqrt{73}}$; $\beta = \frac{\sqrt{73}}{6}$; n=1.33 is the experimental optical refraction index of water; T is the absolute temperature; k is Boltzmann's constant; v_w – volume per water molecule; is $L(x) = \coth x - \frac{1}{x}$ is the Langevin function.

The dipole moment of water molecule, μ, is the only adjustable parameter of the model used by the author [45]to reproduce the experimental dielectric constant of water in the low-field limit. Although the obtained value, $\mu = 2.1\,D$, is larger than in vacuum (1.86 D), it is still in

reasonable agreement (e.g., in SPC water model the dipole model is 2.27 D, in TIP3P is 2.35 D). The dielectric response within the Booth model, by assuming that T=300 K and $v_w = 29.9$ Å), is shown on Fig. 1 with solid curve. At field strengths below 100 mV/Å, the agreement between MD results and the Booth model is excellent, in accordance with previous studies[29]. The slight deviation at higher electric fields may be due to several factors. First, the field induced structural reordering of water takes place what leads to the phase transition at $E > 1\ V/\text{Å}$ [27], while Booth assumed the field independent model for near order of liquid water. Second, the water model used in current MD simulations is rigid, so that reorientation of molecules is the only mechanism of polarization, while effects of electronic polarization and molecular bond/angle modification, which might become important at higher fields, are neglected. The Booth model actually takes molecular flexibility and the electronic polarization implicitly into account, by introducing in the high-field limit constant from the experimental optical refraction index of water, $\varepsilon_{E\to\infty} = n^2$.

We will not go further into the details of the water dielectric response at high electric fields, as the field strength pertaining to the protein channels investigated here is well below $100\ \text{m}V/\text{Å}$. The important conclusion is that, at the electric field strength we will consider, the dielectric response of water is non-linear but absolutely consistent with the analytic microscopic model that assumes the independence of local structure of liquid water (including the water density) of the applied electric field.

The dielectric response of water at $E < 100\ mV/\text{Å}$ can also be fitted by using the simple Langevin function,

$$P_L(E) = \frac{\mu}{v_w} L\left(\frac{\mu E}{kT}\right) \qquad (2)$$

In this case, however, the fitting parameters, μ and v_w, will not have physically meaningful values[27,30]. The dielectric response function given by Eq. 2 calculated with physical values of the

parameters[46] ($\mu = 2.27$ D, $v_w = 29.9\text{Å}^3, T = 300$ K) is plotted with the dashed line in Fig.1. The polarization density is strongly underestimated and the saturation is shifted to higher electric field strength. This may lead, in turn, to a significant overestimation of the electric field based on the calculated values of the average dipole moment. We will return to this point with an example in the following paragraph.

Electric Field in OmpF: Effects of the Probe Radius and of the Non-Linear Response.

The local polarization density of water within the spherical probes of radius of 5Å centered along the axes of the monomers was calculated for the OmpF porin by following the procedure discussed in the Computational Details section. The Cartesian components of the polarization density vector vs. the position along channel's axis (the z coordinate) are shown in Fig.2a (y-axis points versus the projection of the loop L3 in the CR of channel into the transversal plain). Corresponding electric field's components are shown on Fig. 2b. The parametric plot of the transversal electric field along the axis of the channel is shown in Fig.3.

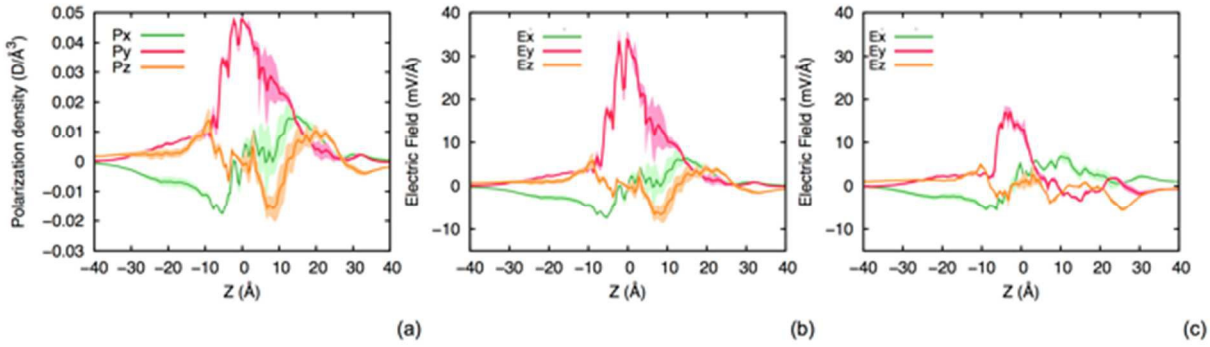


Figure 2. (a) the components of the water polarization density vector averaged over spherical probes with radius of 5 Å, located along OmpF's axis; (b) the corresponding components of the local macroscopic electric field. (c) Macroscopic electric field for OmpC. The standard deviation is shown with the shaded area around the curves.

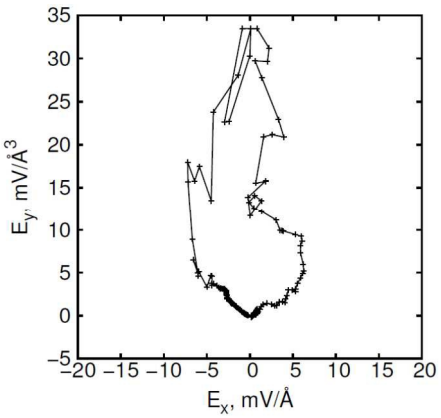


Figure 3. The parametric plot, $\{E_x(z), E_y(z)\}$, of the transversal electric field shown on Fig.2b

First, we note that there is a strong transversal component of the electric field which has the greatest strength in the constriction region (Z around 0 Å), where it points out of the positive basic ladder towards the negative loop L3 (Y direction), as it should be expected. The presented approach to the calculation of the electric field by the way, allows identifying finest details of the channel's electrostatics, hard to be unveiled with other methods. First, it has to be emphasized how the only 'active' component of the electric field in the CR is the Y. Both the X and, most importantly, the Z are absolutely negligible. Before the CR, on the extracellular side of the channel (Z in the range +5 - +10 Å), the Y component is already significant and the Z is actually not negligible and negative. This means that the electric field lines already points down to the CR and towards the loop L3 well before the actual CR. This specific region of the channel has been previously given the name 'preorientation region' in one recent authors' report [10], and the more accurate analysis presented here bolster that concept. Also, upon emerging from the CR on the periplasmic side of the channel (negative Z), the transversal component of the electric field changes direction quite sharply (the Y decreases while the X increases) but the Z is still negligible. The transversal electric field is characterized by a screw-like form and flips by 180^o while passing through the pore (Fig.3). These qualitative details of the OmpF electrostatics were

known from previous studies[17,18], the present analysis has confirmed them and provided a quantitative description.

The profile of the electric field in the CR (Fig.2b) is sharper than that of the polarization density. This effect is due to the nonlinear response of water to the strong electric field in the constriction region[30,31,46].

The values of the *z*-component of the polarization density along the channel axis (Fig. 2a) are in good agreement with previous MD simulations[46] (the average dipole per water molecule presented in Fig. 12 of ref [46] should be divided by $v_w = 29.9\text{Å}^3$ to obtain the polarization density). However, as discussed above, the simple Langevin response function, with the physically chosen parameters used in[46] significantly underestimate the polarization density at any given strength of the electric field (see Fig. 1 and discussion around). That is why the values of electric field strength in the CR reported in [46] are about 2-3 times larger than the present results. The continuum Poisson-Boltzmann models[17,31] with fixed bulk dielectric constant for water in the pore ($\epsilon_w = 78 - 80$) and fixed effective dielectric constant for the protein ($\epsilon_p = 4 - 20$), result in a strength of the electric field in the CR equal to 30-40 mV/Å, which is in good agreement with our calculations (Fig. 2b). On the other hand, the Poisson-Boltzmann model[31] with the nonlinear response (Booth) function and a fixed effective dielectric constant for protein ($\epsilon_p = 10$), results in the electric field approaching 80 mV/Å, i.e. 2-2.5 times larger than the present estimation.

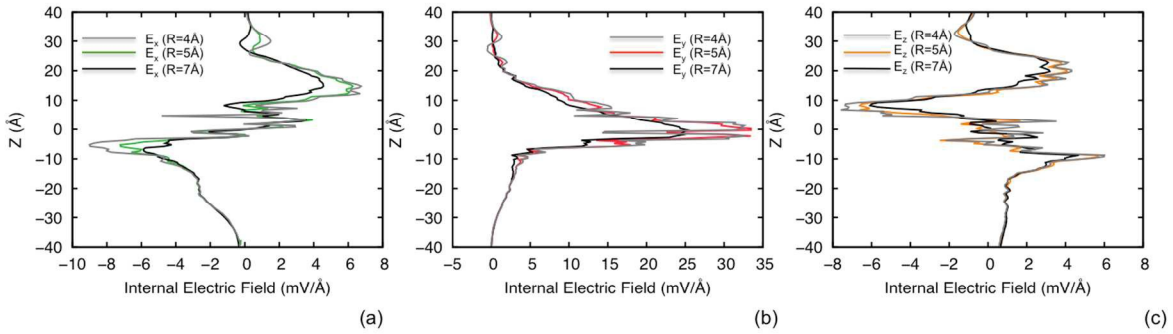


Figure 4. Electric field inside OmpF for different radii of the probe, R=4 Å, R=5 Å, R=7 Å.

Fig. 4 shows the results of the electrostatic field inside OmpF using different probe size. The latter clearly influences the fine structure of the curve because of the non-uniformity of the medium. We selected R=5Å for further analysis, as the field profile is substantially too smoothed at R=7Å. At R=4Å on the other hand, the average number of water molecules in the probe is smaller than 10, leading to the large fluctuations of the averaged values of polarization and, consequently, of the electric field.

There are two kinds of statistical uncertainties in the calculated electric field, the inherent equilibrium thermodynamic fluctuations, and the statistical uncertainty of average values due to the finite time-length of the MD trajectories. Thermodynamic fluctuations (at equilibrium) depend on the size of the probe, the ion density and the other equilibrium thermodynamic variables of the system. In figure 2, for instance, we plot the standard deviation of the thermodynamic fluctuations.

The average value itself has a statistical uncertainty due to the possible insufficient sampling provided by the MD trajectories. This error of the average is reduced with the trajectory duration *T*. The simulation time, *T*, should be larger than the characteristic time of all relaxation processes in the system. We estimated the error of the average by dividing the whole trajectory into three

pieces and by calculating the variance of the average electric field obtained for each piece. This calculation was also performed separately on each of the three protein monomers. The statistical error on the average values of the electric field was always smaller than its inherent thermodynamic (equilibrium) fluctuations.

Comparison of the two main general channels of *E. Coli*, OmpF and OmpC

In Figure 2, the electric field calculated along the channel axis is displayed for OmpF (2.b) and OmpC (2.c). This two proteins have been widely studied[16], and it is still not clear why OmpF is much more permeable than OmpC despite of the similar size[47]. In a recent work[10], we showed how the polarization density of water molecules inside the channel is able to qualitatively capture the main electrostatics differences among the two channels, suggesting very important implications for rational drug design.

The electric field has the maximum value in the CR for both channels. They are also similar in having the most intense component of the electric field onto the y-axis, i.e. pointing from the positive basic ladder towards the negative loop L3. However, while OmpF shows a field strength of ~36 mV/Å in the CR, a significant reduction to ~20 mV/Å is observed in OmpC. In addition, dramatic differences are observed in the so-called 'preorientation region'[10], which is between +5 and +10 Å along the z-axis in the present cases.

Although OmpC has a higher net negative charge, it is not localized in the CR but in the extracellular vestibule[10]. The effect of this additional negative charge (due to specific amino acid mutations[10]) in OmpC, is the almost complete disruption of the opposite charge segregation that characterizes, instead, the OmpF preorientation zone. In turn, all of the three electric field components are negligible (Fig. 2c), meaning that in OmpC there is no effective electric field at the mouth of the CR, driving dipolar molecules to adopt any specific orientation before they get

into the CR. Conversely, OmpF and OmpC are rather similar, as far as the electric field is concerned upon emerging from the CR on the periplasmic side of the channel (negative Z).

OmpF and OmpC, effect of the ionic strength

It is known how relative expression of OmpF and OmpC is strictly dependent upon several environmental parameters, such as the ionic strength of the medium, the pH, the temperature and the level of nutrients [16]. The earlier observations that OmpF is predominant in low osmolarity media, while OmpC is more expressed at high osmolarity, led to the conclusion that the latter had a reduced size to act as a more selective filter in environments rich in nutrients but also in potentially noxious species. On the other hand, OmpF was concluded to be less stringent, which is required when bacteria are grown under nutritional deficiency.

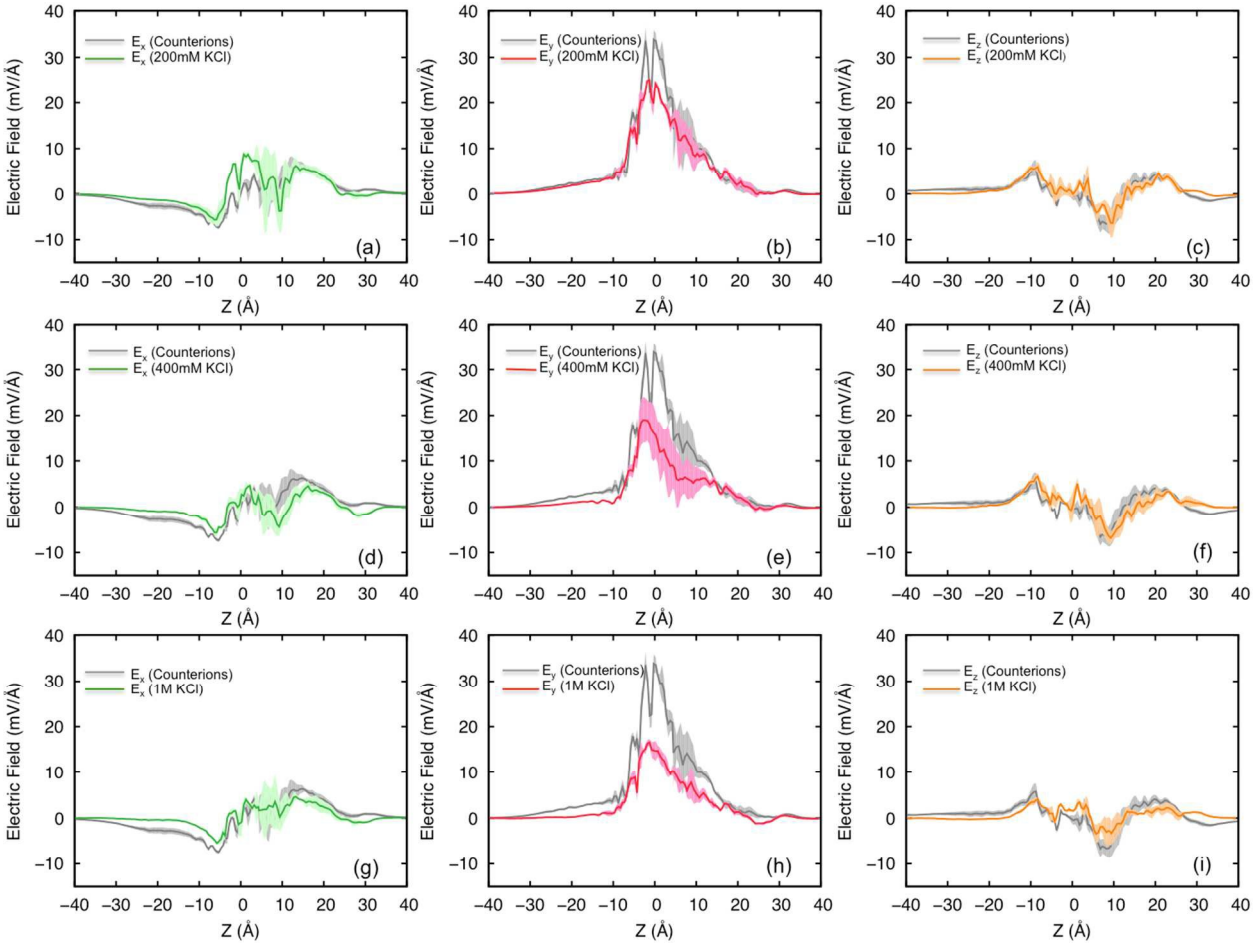


Figure 5. Internal electric field in OmpF in the presence of different electrolyte concentrations, (a-c) 200mM KCl, (d-f) 400mM KCl, and (g-i) 1M KCl.

It has been more recently experimentally shown that although pore size is comparable among OmpF and OmpC, the internal electrostatics might be the major determinant for permeability differences[48]. In order to investigate the electrostatics of these two protein channels in a more physiologically relevant scenario, we performed additional (~5 □s) MD simulations in the presence of different KCl concentrations, namely, 200 and 400 mM. In the case of OmpF, MD was performed also at 1M KCl. The results of electric field calculations are reported in Fig. 5 and Fig. 6 for OmpF and OmpC, respectively.

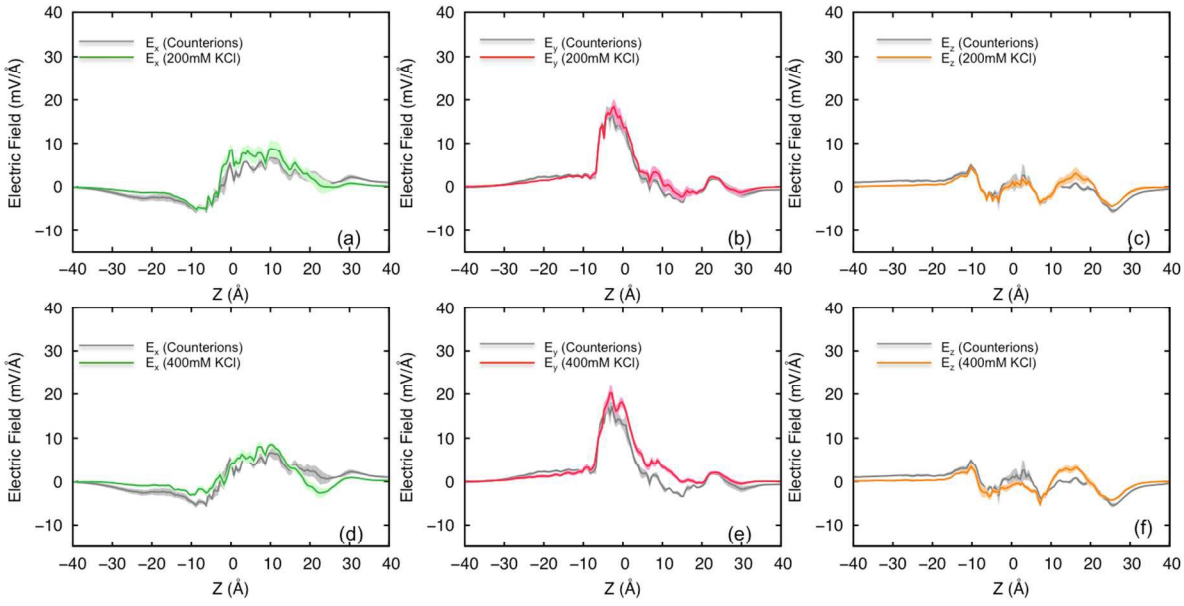


Figure 6. Electric field in OmpC in the presence of different electrolyte concentrations, (a-c) 200mM KCl, and (d-f) 400mM KCl.

OmpF is very sensitive to the electrolyte concentration (Fig. 5). The y-component of the electrostatic field is dramatically reduced already at 200 mM KCl, down to values comparable to the OmpC in the absence of electrolytes (Fig. 2c). On the other hand, both the x and z

components change only slightly, meaning that the overall 'choreography' is maintained, even at 1M KCl. However, not only the CR electrostatics appear to be affected by increased ion concentration, also the preorientation region is perturbed, so that the y-component is significantly reduced there also, whereas the z one is not (Fig. 5).

Conversely, when the same conditions were applied to OmpC, no remarkable differences in the intrinsic electric field were found, up to 400 mM KCl (Fig. 6). However, in the preorientation zone, a closer inspection of the results reveals a significant increase of the y-component in the positive direction. This means that perturbation of the electrostatic field of OmpC due to the presence of physiological concentration of monovalent ions might be helpful to the passage of polar solutes, by creating a favorable preorientation field at the mouth of the CR, which was negligible in the absence of electrolytes.

It is quite astonishing how the present results fairly match the above mentioned general considerations about OmpF and OmpC relative expression. Actually, these two porins look to be designed by evolution to work properly in different environmental conditions. The electrostatics of OmpF, devoted to low osmolarity, is profoundly affected by the presence of electrolytes in the medium. The OmpC, on the other hand, is almost insensitive to the ionic strength, preserving its inherent electrostatic profile in the presence of electrolytes at rather high concentrations.

Now, from the present results it can be speculated that an OmpF-like profile of the electric field (Fig. 2b), with rather high field strength in the CR and a favorable preorientation zone may help the uptake of dipolar molecules. In an environment characterized by nutrients deficiency this should be a must for a general channel, in order to guarantee a sufficient influx of essential species[16,48]. However, most of the commonly employed antibiotics are characterized by a significant electric dipole[49] and should be, thus, similarly able to penetrate through OmpF

possibly too easily, making bacteria expressing OmpF rather susceptible. In the same conditions, OmpC expressing bacteria are known to be generally less susceptible, probably because of the lower field-strength in the CR and, most importantly, to the negligible preorientation (Fig. 2c).

When a high osmolarity environment is considered, a more stringent molecular filter is required, guaranteeing sufficient uptake of nutrients but avoiding penetration by noxious species to the maximum extent[48]. The OmpC-like profile of the electric field in the absence of electrolytes (Fig. 2c) is possibly too selective, but it is possible to suggest that its filtering behavior is finely tuned by the presence of electrolytes, leaving the CR almost unaffected and 'switching on' the preorientation zone (Fig. 6).

At the moment, these are mere speculations and explaining all the more recent experimental results is certainly beyond the aim of this work, but the major role of the electrostatics is clear and widely recognized. It is finally interesting to note how the present results are in agreement with recent experimental findings[48]. Overall permeability was showed to become comparable for OmpF and OmpC with increasing ionic strength of the medium. In our simulations, the electric field profiles do become comparable.

Despite protein channel's electric field is certainly fundamental in defining permeability, other factors cannot be ignored, such as, substrate flexibility and size, net charge and its distribution, H-bond donors/acceptors, desolvation of both protein residues and the substrate upon entering the CR, among others[50], all of which potentially modulate the permeability in a substrate specific fashion, over a channel specific background.

Effects of the pH

As the protonation state of certain amino acid residues might change with pH, their net charge can be modified in turn. This should result, of course, in the modification of channel's electric field.

In Figure 7, the results for OmpF at pH 5 are compared to the ones at pH 7. A dramatic change of the electric field can clearly be seen, pertaining both the CR and preorientation region. Similarly to the effects shown above for increasing KCl concentration, also decreasing the pH, thus 'loosing' some negative charges, reduces the y component of the electric field, leaving the x and z almost unaffected. In the case of pH reduction, however, the negative z component in the preorientation region is eliminated, while the electrolytes were not able to produce this alteration even at the highest concentration tested (Fig. 5).

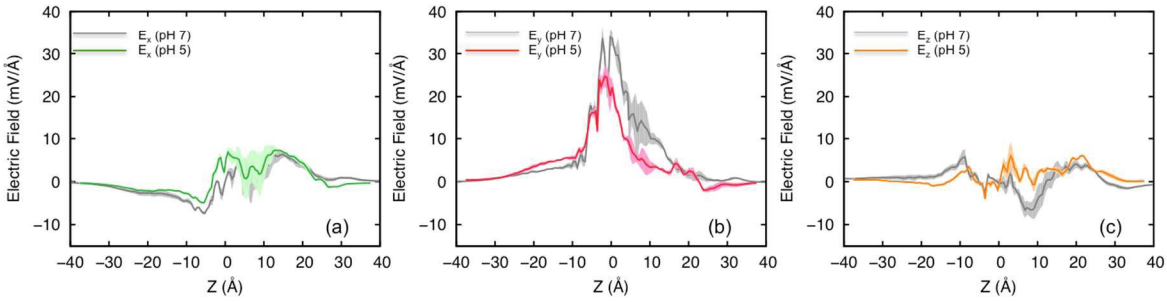


Figure 7. Electric field of OmpF for pH=7 in grey, and pH=5, (a) in green for the x, (b) in red for y, and (c) in orange for the z component.

Similarly to the ionic strength, also the pH of the environment is known to impact relative expression of OmpF and OmpC. In particular, acidic pH favors the expression of OmpC, while that of OmpF is repressed[51]. Coherently, the present work has shown how OmpF appears to be highly sensitive to the pH, as far as the internal electric field is concerned, bolstering the idea that evolution has finely designed these water filled channels to work in rather specific

conditions. Going into further details is beyond the scope of this presentation, which is focused on the method, but the intriguing implications of these results certainly call for specifically targeted investigations.

Mutants from antibiotic resistant clinical isolates: OmpC20 and OmpC33

Quite recently, the crystal structure of a series of OmpC mutants has become available[7]. Seven *E. coli* strains were isolated during two years antibiotics treatment of a patient suffering from Caroli syndrome[52]. The antibiotics resistance of these clinical strains was progressively greater to the antibiotics used for treatment, including imipenem, meropenem, cefotaxime, ceftazidime and ciprofloxacin[52].

As we pointed out in a recent work[10], the electrostatics properties of the corresponding OmpC mutants showed a remarkable progressive variation directly related to the amino acid mutations inside the channels, while the size pore remained unchanged. In particular, we investigated the first and the last mutants of the series, OmpC20 and OmpC33, respectively, finding the more interesting differences in the preorientation, not in the CR region.

Starting from those results, we calculated their internal electric field. Results are shown in figure 8. In agreement with our previous study[10], in what we called the preorientation region, just above the CR on the extracellular side of the channel, the y-component of the electric field is inverted with respect to OmpF (Fig. 5) and OmpC (Fig. 6), i.e. the y is negative in the reference system adopted here, while the other two components are comparable. The OmpC33, in addition, is characterized by a rather low field strength in the CR.

This electric field profile along the channel axis is expected to have a large influence on the permeability of polar species, which are preoriented in the extracellular vestibule with a direction opposite to the one imposed by the electrostatics of the CR. This means that the crossing

molecule has to undertake sharp reorientations while traversing the channel, right at the mouth of the more restricted zone, where the steric hindrance is severe.

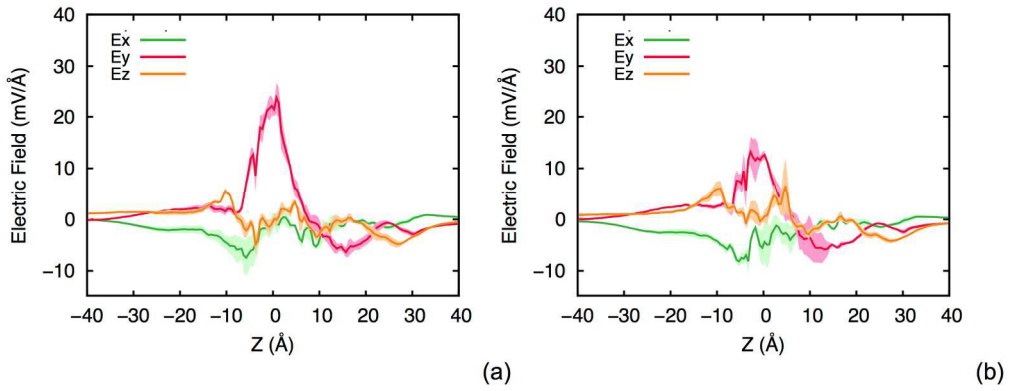


Figure 8. Electric field in (a) OmpC20 and (b) OmpC33. The three components of the electric field are depicted, in green Ex, in red Ey and in orange Ez.

CONCLUSIONS

We have presented an all-atom simulations approach for the calculation of the macroscopic electric field inside nanometric water-filled channels under different environmental scenarios. Intrinsic water is taken as a natural probe, and its local polarization density is connected to the macroscopic field through the electrostatic response function. The method allows one to compare not only different systems but also the same system under different pH and ion concentration conditions.

Our method is enough accurate to be able to capture small differences upon subtle mutations of a few residues only, and for this reason can be used as a way to characterize channels. This is a basic step to design new drugs with enhanced permability but also such knowledge might be crucial for designing engineered biomimetic nanofilters. Indeed, the accurate quantification of

the macroscopic electric field inside channels opens up the way to the definition of an electrostatic model to calculate the interaction energy of crossing polar molecules, which sense the internal macroscopic electric field by virtue of their charges/dipole.

ACKNOWLEDGMENTS

The research leading to these results was conducted as part of the "Translocation" consortium (www.translocation.com) and has received support from the Innovative Medicines Initiatives Joint Undertaking under Grant Agreement no. 115525, resources that are composed of financial contribution from the European Union's seventh framework programme (FP7/2007-2013) and EFPIA companies in kind contribution. S.Acosta-Gutierrez is a Marie Skłodowska-Curie fellow within the "Translocation" Network, project no. 607694 . We also want to thank the financial support of MIUR (PRIN 2010CSJX4F).

REFERENCES

1 R. Tommasi, D. G. Brown, G. K. Walkup, J. I. Manchester and A. A. Miller, *Nat Rev Drug Discov*, 2015, **14**, 529–542.
2 D. J. Payne, L. F. Miller, D. Findlay, J. Anderson and L. Marks, *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 2015, **370**, 20140086–20140086.
3 R. A. Stavenger and M. Winterhalter, *Sci. Transl. Med.*, 2014, **6**, 228ed7–228ed7.
4 J.-M. Pagès, C. E. James and M. Winterhalter, *Nat Rev Micro*, 2008, **6**, 893–903.
5 T. Mach, P. Neves, E. Spiga, H. Weingart, M. Winterhalter, P. Ruggerone, M. Ceccarelli and P. Gameiro, *J Am Chem Soc*, 2008, **130**, 13301–13309.
6 J. Cama, H. Bajaj, S. Pagliara, T. Maier, Y. Braun, M. Winterhalter and U. F. Keyser, *J Am Chem Soc*, 2015, **137**, 13836–13843.
7 H. Lou, M. Chen, S. S. Black, S. R. Bushell, M. Ceccarelli, T. Mach, K. Beis, A. S. Low, V. A. Bamford, I. R. Booth, H. Bayley and J. H. Naismith, *PLoS ONE*, 2011, **6**, e25825.
8 M. Winterhalter and M. Ceccarelli, *Eur J Pharm Biopharm*, 2015.
9 V. M. Isabella, A. J. Campbell, J. Manchester, M. Sylvester, A. S. Nayar, K. E. Ferguson, R. Tommasi and A. A. Miller, *Chem Biol*, 2015, **22**, 535–547.
10 S. Acosta-Gutierrez, M. A. Scorciapino, I. Bodrenko and M. Ceccarelli, *J Phys Chem Lett*, 2015, **6**, 1807–1812.
11 A. Kumar, E. Hajjar, P. Ruggerone and M. Ceccarelli, *J Phys Chem B*, 2010, **114**, 9608–9616.
12 S. Samanta, M. A. Scorciapino and M. Ceccarelli, *Phys Chem Chem Phys*, 2015, 1–10.
13 S. W. Cowan, T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R. A. Pauptit, J. N. Jansonius and J. P. Rosenbusch, *Nature*, 1992, **358**, 727–733.
14 E. M. Nestorovich, C. Danelon, M. Winterhalter and S. M. Bezrukov, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 9789–9794.

15 P. R. Singh, M. Ceccarelli, M. Lovelle, M. Winterhalter and K. R. Mahendran, *J Phys Chem B*, 2012, **116**, 4433–4438.
16 M. Masi and J.-M. Pagès, *Open Microbiol J*, 2013, **7**, 22–33.
17 A. KARSHIKOFF, V. SPASSOV, S. W. Cowan, R. LADENSTEIN and T. Schirmer, *J Mol Biol*, 1994, **240**, 372–384.
18 W. Im and B. Roux, *J Mol Biol*, 2002, **322**, 851–869.
19 M. Aguilella-Arzo, J. J. García-Celma, J. Cervera, A. Alcaraz and V. M. Aguilella, *Bioelectrochemistry*, 2007, **70**, 320–327.
20 A. Philippsen, W. Im, A. Engel, T. Schirmer, B. Roux and D. J. Müller, *Biophys J*, 2002, **82**, 1667–1676.
21 H. Bajaj, M. A. Scorciapino, L. Moynié, M. G. P. Page, J. H. Naismith, M. Ceccarelli and M. Winterhalter, *J Biol Chem*, 2015, jbc.M115.690156.
22 B. L. de Groot and H. Grubmüller, *Science*, 2001, **294**, 2353–2357.
23 E. Tajkhorshid, P. Nollert, M. Ø. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud and K. Schulten, *Science*, 2002, **296**, 525–530.
24 S. Yuan, H. Vogel and S. Filipek, *Angew. Chem. Int. Ed.*, 2013, n/a–n/a.
25 M. Marchi, F. Sterpone and M. Ceccarelli, *J Am Chem Soc*, 2002, **124**, 6787–6791.
26 H. E. Alper and R. M. Levy, *J. Phys. Chem.*, 1990, **94**, 8401–8403.
27 G. Sutmann, *Journal of Electroanalytical Chemistry*, 1998, **450**, 289–302.
28 J. L. Aragones, L. G. MacDowell, J. I. Siepmann and C. Vega, *Phys Rev Lett*, 2011, **107**, 155702.
29 I.-C. Yeh and M. L. Berkowitz, *J Chem Phys*, 1999, **110**, 7935–7942.
30 M. S. Sansom, G. R. Smith, C. Adcock and P. C. Biggin, *Biophys J*, 1997, **73**, 2404–2415.
31 M. Aguilella-Arzo, A. Andrio, V. M. Aguilella and A. Alcaraz, *Physical Chemistry Chemical Physics*, 2009, **11**, 358–365.
32 T. Darden, D. York and L. Pedersen, *J Chem Phys*, 1993, **98**, 10089–10092.
33 L. Perera, U. Essmann and M. L. Berkowitz, *J Chem Phys*, 1995, **102**, 450–456.
34 M. Neumann, *Molecular Physics*, 2006, **50**, 841–858.
35 J. D. Jackson, *Classical Electrodynamics, 3rd Edn Wiley*, New York, 1999.
36 K. K. Mandadapu, J. A. Templeton and J. W. Lee, *J Chem Phys*, 2013, **139**, 054115–11.
37 A. Aksimentiev and K. Schulten, *Biophys J*, 2005, **88**, 3745–3761.
38 W. Humphrey, A. Dalke and K. Schulten, *J Mol Graph*, 1996, **14**, 33–38.
39 C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J Chem Theory Comput*, 2011, **7**, 2284–2295.
40 M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J Chem Theory Comput*, 2011, **7**, 525–537.
41 M. J. Harvey, G. Giupponi and G. De Fabritiis, *J Chem Theory Comput*, 2009, **5**, 1632–1639.
42 I. Buch, T. Giorgino and G. De Fabritiis, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 10184–10189.
43 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, **78**, 1950–1958.
44 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J Chem Phys*, 1983, **79**, 926.
45 F. Booth, *J Chem Phys*, 1951, **19**, 391–394.
46 D. P. Tieleman and H. J. Berendsen, *Biophys J*, 1998, **74**, 2786–2801.
47 A. Kumar, E. Hajjar, P. Ruggerone and M. Ceccarelli, *Journal of Physics: Condensed Matter*, 2010, **22**, 454125.
48 S. Kojima and H. Nikaido, *J Biol Chem*, 2014.
49 G. Malloci, A. Vargiu, G. Serra, A. Bosin, P. Ruggerone and M. Ceccarelli, *Molecules*, 2015, **20**, 13997–14021.
50 Q.-T. Tran, R. A. Pearlstein, S. Williams, J. Reilly, T. Krucker and G. Erdemli, *PROTEINS-STRUCTURE FUNCTION AND BIOINFORMATICS*, 2014, **82**, 2998–3012.
51 M. Heyde and R. Portalier, *Mol Gen Genet*, 1987, **208**, 511–517.
52 A. S. Low, F. M. MacKenzie, I. M. Gould and I. R. Booth, *Mol Microbiol*, 2001, **42**, 619–630.