Localizing the charged side chains of ion channels within the crowded charge models

Justin J. Finnerty,[†] Robert Eisenberg,[‡] and Paolo Carloni[*,¶]

Computational Biophysics, German Research School for Simulation Sciences, 52425 Jülich, Germany, Department of Molecular Biophysics and Physiology, Rush University, Chicago IL 60612, United States of America, and Computational Biophysics, German Research School for Simulation Sciences , 52425 Jülich, Germany and Institute for Advanced Simulation IAS-5, Computational Biomedicine, Forschungszentrum Jülich, 52425 Jülich, Germany

E-mail: <p.carloni@grs-sim.de>

Abstract

The simplified coarse grained models of selectivity of Nonner, Eisenberg and co-workers predict ion selectivity for a variety of different ion channels. The model includes the charged atoms of the channel's charged residues and permeant ions. However its MC implementation does not take advantage of the increasingly large body of structural information available. Here, we introduce the location of the channel's charged residues in to the model's Hamiltonian. In the DEKA Na^+ channel this allow us to correlate the lysine's topological location directly with the predicted selectivity. In the NanC channel, from *Escherichia coli*, the dramatic variation in the resulting ion

[*]To whom correspondence should be addressed

[†]Computational Biophysics, German Research School for Simulation Sciences, 52425 Jülich, Germany

[‡]Department of Molecular Biophysics and Physiology, Rush University, Chicago IL 60612, United States of America

[¶]Computational Biophysics, German Research School for Simulation Sciences , 52425 Jülich, Germany and Institute for Advanced Simulation IAS-5, Computational Biomedicine, Forschungszentrum Jülich, 52425 Jülich, Germany

population predicts novel selectivity regions and binding sites that can be directly correlated with structural information. These results have well defined thermodynamic properties that are significantly modified by structural detail allowing new insights with molecular detail.

Introduction

Ion selective channels are transmembrane proteins that allow the flow of ions such as Na^+ [1–10], K^+ [11–20], Ca^{2+} [21–28] and Cl^- [29–32] into and out of a cell. Measured selectivities for ion channels are typically small, ranging from a modest $5-10:1$ for Na^+ over Ca^{2+} in sodium channels to $1000:1$ for Ca^{2+} over Na^+ in calcium channels [33]. This means an ion channel achieves selectivity by very small (a few k_BT) energy differences compared to the energies of ordinary covalent bonds. Experimental work carried out in ionic concentrations ranging from 10^{-7} M to 1.0 M has greatly facilitated the investigation of molecular mechanisms governing the selectivity.

Computations have also helped us understand the mechanism and energetics of selectivity. On one hand, a variety of molecular simulation methods have been used to estimate the free energy of ion permeation through ion channels, for which structural information is available (such as K^+ channels [17–20] and recently a bacterial Na^+ channel [34,35]). However, these simulations have not studied the wide range of ion concentrations used in experiment. On the other hand, coarse-grain methods such as the charge space competition (CSC) models of Nonner and Eisenberg [36–44], some using the induced charge computation (IC) method [45], employ a Hamiltonian based only on electrostatics and volume exclusion in a Monte Carlo (MC) simulation, to provide quantitative predictions of selectivity in Na^+ [46], L-type Ca^{2+} [47–50] and RyR Ca^{2+} [51–56] channels. These predictions come from the output ion density profiles simulated over a wide range of well-defined concentrations. The fact that the CSC model reproduces the selectivity of such a large class of ion channels shows the critical involvement of electrostatic interactions in determining selectivity. However, because the relationship of the simplified structure used so far in the charge space competition models to the real structure is not known, the molecular origin of the energetics seen from this model [36,44,55,57–66]

is still unclear.[67]

The charge space competition model uses drastic approximations to structure; representing the ion channel protein as a thick walled tube with rounded edges. These simplifications to the geometry allow the CSC model Hamiltonian to use the Poisson-Boltzmann equation directly, instead of a generalized Born model, when computing the electrostatic interactions (U_{IC}) between the induced charge on the protein-water dielectric boundary and all charged ions (in a computationally efficient manner.) The Hamiltonian also includes contributions for the exact pair-wise screened Coulomb potential (U_{C}) and hard object overlap (U_{overlap}) (see Methods section below for more details).

Charged amino acid residues from the ion channel (*structural ions*, ions modeled as charged spheres) are restricted inside a hard-walled cylinder representing the selectivity filter (shown as the light blue cylinder in Fig. 1a and Fig. 2a and b). We refer to this model as the flexible charge space competition (FCC) model (named from Giri et al.[60]) to indicate the structural ions are confined only inside this cylinder. Taking into account *where* the charge bearing atoms of the side-chains are located could allow the molecular mechanism of selectivity determined from the model to be more directly compared with atomistic models. Here we present a localization procedure similar to Yu et al[19], which we call the localized charge space competition (LCC) model, that extends the FCC implementation to introduce a parabolic potential well on the positions of the structural ions ($U_{R,i}$). We use the LCC in two ways, to study the impact of particle location in an abstract model on the prediction of selectivity (focussing on the DEKA Na^+ channel) and to take an atomistic structure from experiment to screen for areas of potential interest for further experimental and computational studies (focussing on the NanC channel[68]).

FCC calculations have been previously carried out[46] on the DEKA Na^+ channel from eukaryotes, with a geometry[36,69,70] derived from sieving experiments using Pauling radii ions[71]. The importance of the lysine residue on selectivity is known, although the mechanism of action is still debated[72]. Here we use LCC to investigate how the location of the lysine charge in our model is correlated with selectivity. The measure of selectivity[73] used specifically for this is to define

selectivity as the equilibrium constant of the reaction[74]

$$\mathrm{M_1} + (\mathrm{X}\cdot\mathrm{M_2})_{\mathrm{filter}} \quad \overset{\mathrm{K_{M_1,M_2}}}{\rightleftharpoons} \quad \mathrm{M_2} + (\mathrm{X}\cdot\mathrm{M_1})_{\mathrm{filter}} \tag{1}$$

where $\mathrm{M_1}$ and $\mathrm{M_2}$ are two ions outside channel X, competing to occupy, $(\mathrm{X}\cdot\mathrm{M_1})_{\mathrm{filter}}$ and $(\mathrm{X}\cdot\mathrm{M_2})_{\mathrm{filter}}$ respectively, the selectivity filter region of channel X (Eqn. 1). This definition is appropriate here as the model evaluates the system under equilibrium conditions; we make no assumption of a relationship between this measure and the selectivities measured from experiment. The equilibrium constant $\mathrm{K_{M_1,M_2}}$ reads

$$\mathrm{K_{M_1,M_2}} = \frac{[\mathrm{X}\cdot\mathrm{M_1}]\,[\mathrm{M_2}]}{[\mathrm{X}\cdot\mathrm{M_2}]\,[\mathrm{M_1}]} \tag{2}$$

where $[\mathrm{X}\cdot\mathrm{M}_i]$ is evaluated as the occupancy of ion M_i in sub-intervals of the selectivity filter model (shown as a brown line in Fig. 1c,d). The sub-interval is scanned along the channel pore to determine which part of the ion density profile most closely predicts the experimental selectivities. In this way the region of our model that contains the features responsible for selectivity can be found.

The reported FCC model of the DEKA Na^+ channel has structural ion density profiles[46] that clearly show a $-+-$ charge pattern ($+$ refers to the lysine residue, $-$ to E or D; patterns described in Fig. 1e) dominates the structural ion density profile of the model, although small peaks of positive ion density at either end indicate that populations of $--+$ and $+--$ charge patterns (see Fig. 1e) are also present. The self-organization into a $-+-$ pattern is hardly surprising as it maximizes both the separation of same charged side-chains and interaction between the oppositely charged side-chains. In the FCC model the key sub-interval corresponds to the center of the channel which is also the main location of the positive lysine ion. However, the center of the model is also the key sub-interval found by FCC for the L-type Ca^{2+} channel[47] which has no positive structural ions. Therefore the correlation between the lysine position and selectivity can only be inferred with

FCC. Using LCC we can now investigate these different topological charge patterns separately and study the impact they have on the selectivity, noting that the symmetry of our simulation makes $--+$ and $+--$ equivalent here. In this way LCC can demonstrate whether the selectivity is associated with the central part of the pore model or the position of the lysine. Without localization the importance of the lysine in any correlation with selectivity can only be inferred by our model.

We also applied LCC to a quite different channel that has many more charged residues within the pore. The X-ray structure[68] of N-acetylneuraminic acid-inducible outer-membrane channel (NanC) from *Escherichia coli*, a member of the oligogalacturonate-specific monomeric porin (KdgM) family, was recently published. We therefore construct our simulation with structural ion localisation based directly on the experimental structural information (see Methods for more details).

In contrast to L-type Ca^{2+} channels[49,50], the DEKA Na^{+} channel and the RyR Ca^{2+} receptor[51–56], so far no *signature selectivity motif* related to NanC's function has been identified in its primary structure. This channel features a wide, long pore lined with an approximately equal number of positively and negatively charged amino acid side chains, thirty in total (Fig. 2a and Table S3 in supplementary information). We can use LCC as a tool to rapidly screen for areas and residues that lie within the channel pore that are of potential interest for further experimental and computational studies. As we show in the next section, these predictions could not have been made without the high resolution structural information.

Results/Discussion

We now discuss applying the FCC and LCC methods on the DEKA Na^{+} channel to investigate the role of the lysine in the DEKA sequence (for simulation details refer to the Methods section). In sodium channels, the selectivity of Na^{+} over Li^{+} and K^{+} is largely in inverse proportion to the volume of the ions[33]. Our results for selectivity in the FCC and LCC with $-+-$ pattern (Table 1) reproduce the experimental results using the sub-interval $[-2.5 : 2.5\text{Å}]$ (shown as a brown

line in Fig. 1b, the zero point is the center of the channel and the selectivity filter occupies the region $[-5:5\text{Å}]$). We find that localizing the side chains to give the $--+$ pattern (see Fig. 1e) dramatically changes the Na^+, K^+, Ca^{2+} and Li^+ ion density along the channel axis compared to the original model. For the LCC with $--+$ pattern the key sub-interval is located where the lysine is localized at -2Å (interval $[-3:-1\text{Å}]$, brown line below Fig. 1d). The LCC simulations on the $-+-$ patterns (interval $[-2.5:2.5\text{Å}]$) and $--+$ (interval $[-3:-1\text{Å}]$) patterns had similar selectivity for Na^+ over Ca^{2+} (Table 1). This selectivity was different from our value from the FCC model. With LCC we see Ca^{2+} ion density is smaller than Na^+ density at the location of the lysine side-chain (Fig. 1c and d; 0 and -2Å respectively), while in the FCC profiles there is only an increase in Ca^{2+} ion density at the location of D/E residues (Fig. 1b; +2 and -2Å). On the other hand, localizing the sides chains to give the $-+-$ pattern changes the predicted selectivity with minimal change in the ion density along the channel axis. We conclude that the improvement in the predicted calcium selectivity is due, at least in part, to the localization scheme making it more difficult for the D and E side-chains to mitigate the repulsion between the doubly charged cation and the now-localized lysine side-chain as the cation passes along the channel, regardless of the charge pattern. Using LCC we can therefore demonstrate that selectivity is associated with the the position of the lysine. As is particularly noticeable for the $--+$ pattern (Fig. 1d), the sub-interval is associated with a minima in the solute ion profile (at -2Å) and not the maxima (at 1.2Å), demonstrating that the model predicts that selectivity is based on passing an energy barrier associated with repulsion between the lysine and the solute cation and not on a binding of the solute cation to the anionic side chains.

We also found the density profiles from the FCC model can be approximated as the weighted sum of the density profiles from the LCC model (as seen in Fig. 4),

$$\rho(\text{FCC}) \approxeq 2\times\rho(\text{LCC},-+-)+\frac{1}{2}\rho(\text{LCC},--+)+\frac{1}{2}\rho(\text{LCC},+--) \tag{3}$$

where ρ is the predicted density ($\rho(\text{LCC},+--)$ is the $\rho(\text{LCC},--+)$ density reversed along the

pore axis). Because it can be approximated as the sum of $(-+-)$ and $(--+)$ profiles the FCC model can therefore be used as a common point for comparing the two LCC results. The similar magnitude of the weighting factors demonstrates the $(--+)$ profile is not a direct response to the presence of a solute ion which are only present in the channel in less than one hundredth of the MC configurations. While the approximations of the model are too crude to obtain a quantitative difference in energy, the low weighting factor implies that, from the electrostatic perspective, a number of significantly different side-chain conformations may contribute to the mechanism of selectivity in the Na^+ channel.

In the NanC channel (See Fig. 2), we find that the population of the permeating ions show dramatic differences between the FCC and LCC calculations (Fig. 2c). The anion population from the LCC method (2c right) has a minima between -2 and 6Å and maxima that are higher than the chloride concentration in the bulk solution at 14 and -10Å. Our results from the Na^+ channel suggest the minima in the anion concentration could be a putative anion selectivity region. This means that the residues ARG6, GLU17, ASP54, ARG107, ASP109, ARG129, GLU168, ASP188, and ARG208 may be involved in selectivity for different anions. The maxima with higher than bulk concentration suggest the presence of anion binding sites at 13Å based on LYS111 and/or LYS153 and at -10Å based on LYS91. There may also be a peak in the region of -15Å suggesting another binding site is associated with ARG72.

The ion density profiles in the LCC calculations differ from those of FCC. We first saw this in the density of the cations in the Na^+ channel model with the $--+$ localization pattern. This change is even more striking in the NanC channel where LCC gives a very definite pattern (Fig. 2c left). In this case we see a variation in the chloride profile that is a direct response to the more realistic distribution of the pore's charged side chains.

The results demonstrate that the inclusion of structural information from experiment directly leads to changes in the predicted population density within the model pore. These changes allow us to postulate residues of potential interest to other researchers investigating the selectivity of NanC. Even without structural information from experiment, localization allows the position of

the measuring interval required to reproduce the experimental selectivities in the Na^+ channel model to be directly related to a particular charged side-chain. This demonstrates the utility of the LCC method to incorporate structural information which can lead to mechanistic insights in the study of ion channels.

Methods

Theoretical model

Each channel is modeled as a thick walled tube with rounded edges that provides a pore (Fig. 1a) separating two baths, essentially as done in a series of papers by Nonner, Eisenberg and colleagues (see Boda et al.[46]). The solute ions can travel to any region of the system. For FCC, the structural ions (from the channel's charged side-chains) are found anywhere in a hard-walled cylinder representing the *selectivity filter* of the pore (shown in light blue in Fig. 1a). Hard spheres of radius R_i and bearing a charge of q_i represent both the structural ions and the permeating ions (see Table 2). The CSC models compute the distribution of ions as outputs of the model. The only energies in the FCC model are the electrostatics in the Hamiltonian and the steric repulsion implicit in the rejection of overlapped spheres. Thus the coarse-grain potential energy of the Hamiltonian, $\mathscr{U}$, (Eqn. 4) is made of the ion–ion pair-wise Coulomb electrostatic interaction (U_{C}, Eqn. 5), the electrostatic interaction between ions and the induced charge at the dielectric boundary between the protein and aqueous media in the model (U_{IC}, Eqn. 6) and the overlap repulsion energy (See Sec. S2 in supplementary information for more details). The effect of uncharged polar and non-polar residues are considered implicitly by the dielectric constant of the protein that, along with the solvent, defines the permittivity change at the dielectric boundary used in U_{IC}. A continuum model is used for the solvent based on the solvent's permittivity screening of the charge–charge interactions (using the dielectric constant ε of 80 for water). Thus, $\mathscr{U}$ reads

$$\mathscr{U} = \begin{cases} U_{\mathrm{C}} + U_{\mathrm{IC}} & : \text{no overlap} \\ \infty & : \text{otherwise.} \end{cases} \tag{4}$$

where U_{C} is a standard screened Coulomb interaction defined for particle i as

$$U_{\mathrm{C},i} = \sum_j \frac{q_i q_j \left(\frac{1}{\varepsilon_i} + \frac{1}{\varepsilon_j}\right)}{2\left|\mathbf{r}_i - \mathbf{r}_j\right|} : \text{for all particles } j \neq i \tag{5}$$

where q_x is the charge, p_x is the position, and ε_x is the dielectric constant of the environment around particle x. The dielectric boundary potential (U_{IC}) is represented by discretizing the boundary surface into a series of tiles. This gives the Coulomb interaction contribution for particle i and each tile k as

$$U_{\mathrm{IC},i} = \sum_k \frac{q_i q_k}{2\left|\mathbf{r}_i - \mathbf{r}_k\right|} : \text{for all protein surface tiles } k \tag{6}$$

where $\mathbf{r}_k$ is the centre-point of the surface tile and q_k is the induced charge computed using the Poisson-Boltzmann equation using the IC discretized integral method[45].

The LCC method introduced here puts additional constraints on the motion of the structural ions. It differs from the hard-walled cylinder restriction of the FCC by the addition of a new term in the potential energy function to localize each structural ion i. The localization closely follows that of the confined micro-droplet model of Yu et al[19] and consists of a harmonic potential $U_{R,i}$ defined by a global constant k_f with per-ion localization centre-points $r_{0,i}$ and maximum displacements $R_{f,i}$

$$U_{R,i} = \begin{cases} -\frac{k_f}{R_{f,i}^2}\left|\mathbf{r}_i - \mathbf{r}_{0,i}\right|^2 & : \left|\mathbf{r}_i - \mathbf{r}_{0,i}\right| \leqslant R_{f,i} \\ \infty & : \left|\mathbf{r}_i - \mathbf{r}_{0,i}\right| > R_{f,i} \end{cases} \tag{7}$$

Here we use a spring-like constant of $\frac{k_f}{R_{f,i}^2}$ compared to the single λ_g used by Yu et al.. This was considered more reasonable than a single constant as a particle with a smaller maximum radius $R_{f,i}$ would be expected to have a steeper potential well. The potential energy function is now extended

to read

$$\mathscr{U} = \begin{cases} U_{\mathrm{C}} + U_{\mathrm{IC}} + \sum_i U_{R,i} & : \text{no overlap} \\ \infty & : \text{overlap} \end{cases} \tag{8}$$

where the sum is taken over all the localized structural ions.

The global k_f modulates the influence that the localization potential has on the total potential energy, and the cut-off, $R_{f,i}$, ensures that the localization is never completely lost if k_f approaches zero. The total localization potential is therefore bounded between zero and the number of localized ions times k_f. Yu et al. estimated a value for λ_g of 0.5-1.0 kcal/mol/Å^2 from MD of the supposedly more rigid selectivity filter of KcsA K^+ channel. Without access to an X-ray structure for the DEKA Na^+ channel, the value k_f to use was assessed empirically, with preliminary calculations on the Na^+ channel model carried out to determine a reasonable value for k_f (See Sec. S1 in the supplementary information). From these preliminary simulations the k_f value of 1.24 (in units of k_BT; ≈ 3kJmol^{-1}) was selected for the DEKA Na^+ channel and also used in the NanC simulations reported here. As $R_{f,i}$ values used were 2Å and above, this is equivalent to a λ_g value of less than 0.2 kcal/mol/Å^2 which Yu identified as allowing solute ion-structural ion and structural ion-structural ion interactions to both significantly contribute to the potential energy. Correspondingly, we found changes in k_f have little effect on the qualitative results for the Na^+ channel provided that $U_{R,i}$ was about two orders of magnitude less than the electrostatic potential terms ($k_f \leq 3$, Sec. S1 in the Supporting Information).

FCC and LCC calculations of a model DEKA Na^+ channel

To calculate ion selectivity, Metropolis Monte Carlo simulations based on the potential $\mathscr{U}$ were carried out in the grand canonical ensemble (TVμ_{M_1},μ_{M_2})[42,46,49,75,76]. CSC evaluates the concentration of the ion occupying the channel ($[\mathrm{X}\cdot\mathrm{M}_x]\,(x=1,2)$ in Eqn. 2) as a function of varying bulk concentrations $[\mathrm{M}_x]\,(x=1,2)$. In these calculations the chemical potentials (μ_{M_1}, μ_{M_2}) are varied

so that the $[M_1]$ and $[M_2]$ vary systematically while maintaining the total ionic strength constant. $[X \cdot M_1]$ and $[X \cdot M_2]$ are then represented by mean occupancy values b_{M_1} and b_{M_2} (b and concentration ratio R notation from[4]) of M_1 and M_2, respectively, in a measurement interval within the selectivity filter (See Eqn. S2 in the supplementary information). The measurement intervals used are selected to best reproduce the experimental selectivities.

For the DEKA Na^+ channel structural information is not available. Therefore we chose the $\mathbf{r}_{0,i}$ to match the respective density maxima from the reported FCC model[46]; which gave us the $-+-$ pattern (see Fig. 1e) with a carboxylic acid localized at $+2$ and -2Å and the lysine ammonium at 0Å. We exchanged the axial positions of the lysine and one carboxylic acid to produce the $--+$ model. $R_{f,i}$ was set to the same value as the pore radius; this was considered reasonable here because the pore radius (3Å) is small.

While K_{M_1,M_2} can be calculated from Eqn. 2 at a single concentration ratio R, fitting b_{M_i} to the standard *dose-response curve*:

$$b_{M_i,i} = \frac{b_{M_i,\max}}{1 + 10^{\left(\log(R_{50,M_i}) - \log(R_i)\right)}} \tag{9}$$

and then taking the point where $b_{M_1} = b_{M_2}$ (Eqn. S5 in supplementary information) gives the following quadratic expression for K_{M_1,M_2}:

$$(K_{M_1,M_2})^2 \cdot \frac{b_{M_1,\max}}{R_{50,M_2}} + K_{M_1,M_2} \cdot (b_{M_1,\max} - b_{M_2,\max}) - b_{M_2,\max} \cdot R_{50,M_1} = 0 \tag{10}$$

where $b_{x,\max}$ are the fitted maximum occupancy and $R_{50,x}$ the fitted ratio at fifty percent occupancy (where R_{50,M_2} is expressed in terms of the $[M_1]/[M_2]$, as per R_{50,M_1}.)

The FCC set up as used in previous studies[46] was used in this study and is only summarized here. LCC calculations were performed with the exact same parameters as for the FCC set up except for the addition of localization parameters for the the structural ions. The ion channel was defined by a tube with internal radius of 3Å and external radius of 20Å. The selectivity region was represented by a 10Å length of the tube. The vestibule zone, where the environment of the filter

region changes to the bulk environment, was modeled by rounding the ends of the channel using an arc of radius 5Å. This gave an overall length of the ion channel of 20Å. The dielectric constant of the solvent was 80 and the protein was 10. The simulation temperature was 298K. The simulation cell is a cylinder coaxial with the ion-channel pore. There was no periodicity used in the cell so any interactions are based on the simple Cartesian displacement between objects. Sampling of the ion concentrations in the *bulk* region was performed in a sub-region chosen to minimize effects from the simulation cell boundaries (See supplementary information of Boda et al.[49].)

Five particles in the pore represent the structural ions of the side-chains from the three charged side chains in the pore. Two half-negative oxygen particles represent each carboxylic acid side chain and one ammonium particle represents the side chain of the lysine. Simulations with concentration ratios up to 1:10 started with 150 solute ions; for example a [Na]/[Ca] ratio of 1:10 gives 98 Cl^-, 5 Na^+ and 47 Ca^{2+} ions. The number of solute particles, and correspondingly the cell volume, was increased for larger concentration ratios so that none of the ion species started the simulation with less than 2 particles. Because the number of particles changes due to the grand-canonical trials, the number of each ion species was monitored and a warning given if the number diverged by more than twice the square-root of the initial number. The ion densities used in the mean for calculating $[\mathrm{X}\cdot\mathrm{M}_x]$ for the FCC and LCC $-+-$ pattern were measured in the interval $[-2.5:2.5\text{Å}]$ and for LCC $--+$ pattern were measured in $[-3:-1\text{Å}]$, both relative to the channel pore center point (See Sec. S3 in supplementary information.) 1×10^9 MC trials or more were carried out.

A series of simulations of the DEKA Na^+ channel model were performed with various concentrations of LiCl, $CaCl_2$ and KCl in the presence of NaCl. The concentration ratio R (as $[\mathrm{M}]/[\mathrm{Na}^+]$) for each cation M was varied at least from 0.01 to 100.0 while the ionic strength was maintained at 220mM.

FCC and LCC calculations of the NanC Channel from *Escherichia coli*

As structural information is available for the NanC Channel, $\mathbf{r}_{0,i}$ was set to the mean position of the charged atom from the X-ray structure (PDB 2WJQ and 2WJR)[68], and $R_{f,i}$ was set to the root mean square displacements (RMSD), derived from X-ray B-factors (using the classical $B_i \approxeq 8\pi^2 \mathrm{RMSD}^2$). The translation of the NanC channel structure to our model involved some approximations. The determination of the cylindrical core length, of 36Å, and the vestibule regions, of 10Å, was determined by simple observation of the structure. The internal radius used, of 7Å, was chosen from the radius of C_α atoms of the cylinder backbone (12.5Å) from which we subtracted an estimate of the volume of the non-charged parts of the side chains. We included all the charged residues within the channel pore ($\approx \pm 18$Å of the protein centre point, shown in Fig. 2a and listed in table S3 in the supplementary information). Those center-points that were outside the 7Å radius were individually rescaled to bring them to about 1Å inside the channel model. The same simulation procedure as for the Na^+ channel above was used, differing only in the specification of the channel geometry and locations of the structural ions.

Acknowledgements

The authors acknowledge the following help in the production of this work. The implementation of localization into the LCC model was performed with the assistance of D. Boda and was based on his FCC implementation. We are most grateful for his generous continual help. The parameters used for the charged particles in the NanC channel were derived from MD simulations performed by Paul Strodel in our lab. Financial funding from the Deutsche Forschungsgemeinschaft (DFG) CA 973/1-1 is gratefully acknowledged.

Supporting Information Available.

Further information is available on the following. Discussion of the results of the simulations with varying k_f that were used to determine the value used here. A more detailed description of the potential energy terms. A more detailed description of the calculation of ion selectivity used here, including the derivation of Eqn. 10. A table of the individual ion selectivities used to generate the averages quoted in Tab. 1. A table of the coordinates of the charged atoms on the side chains of NanC derived from the X-ray structure as used as input into our simulation. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) Chandler, W. K.; Meves, H. *J Physiol-London* **1965**, *180*, 788.

(2) Binstock, L.; Lecar, H. *J Gen Physiol* **1969**, *53*, 342–361.

(3) Hille, B. *J Gen Physiol* **1971**, *58*, 599–619.

(4) Feldman, H. A. *Anal Biochem* **1972**, *48*, 317–338.

(5) Hille, B. *J Gen Physiol* **1972**, *59*, 637–658.

(6) Meves, H.; Vogel, W. *J Physiol-London* **1973**, *235*, 225–265.

(7) Binstock, L. *J Gen Physiol* **1976**, *68*, 551–562.

(8) Campbell, D. T. *J Gen Physiol* **1976**, *67*, 295–307.

(9) Ebert, G. A.; Goldman, L. *J Gen Physiol* **1976**, *68*, 327–340.

(10) Heinemann, S. H.; Teriau, H.; Stuhmer, W.; Imoto, K.; Numa, S. *Nature* **1992**, *356*, 441–443.

(11) Hille, B. *J Gen Physiol* **1973**, *61*, 669–686.

(12) Hagiwara, S.; Takahash.K, *J Membr Biol* **1974**, *18*, 61–80.

(13) Gay, L. A.; Stanfield, P. R. *Pflugers Archiv-Eur J Physiol* **1978**, *378*, 177–179.

(14) Reuter, H.; Stevens, C. F. *J Membr Biol* **1980**, *57*, 103–118.

(15) Blatz, A. L.; Magleby, K. L. *Biophys J* **1984**, *45*, A306.

(16) Taylor, P. S. *J Physiol-London* **1987**, *388*, 437–447.

(17) Aqvist, J.; Luzhkov, V. *Nature* **2000**, *404*, 881–884.

(18) Dudev, T.; Lim, C. *J Am Chem Soc* **2009**, *131*, 8092–8101.

(19) Yu, H.; Noskov, S. Y.; Roux, B. *Proc Natl Acad Sci U S A* **2010**, *107*, 20329–20334.

(20) Kim, I.; Allen, T. W. *Proc Natl Acad Sci U S A* **2011**, *108*, 17963–17968.

(21) Hess, P.; Lansman, J. B.; Tsien, R. W. *J Gen Physiol* **1986**, *88*, 293–319.

(22) Tsien, R. W.; Hess, P.; McCleskey, E. W.; Rosenberg, R. L. *Ann Rev Biophys Biophys Chem* **1987**, *16*, 265–290.

(23) Yang, J.; Ellinor, P. T.; Sather, W. A.; Zhang, J. F.; Tsien, R. W. *Nature* **1993**, *366*, 158–161.

(24) Ellinor, P. T.; Yang, J.; Sather, W. A.; Zhang, J. F.; Tsien, R. W. *Neuron* **1995**, *15*, 1121–1132.

(25) Koch, S. E.; Bodi, I.; Schwartz, A.; Varadi, G. *J Biol Chem* **2000**, *275*, 34493–34500.

(26) Wu, X. S.; Edwards, H. D.; Sather, W. A. *J Biol Chem* **2000**, *275*, 31778–31785.

(27) Sather, W. A.; McCleskey, E. W. *Annu Rev Physiol* **2003**, *65*, 133–159.

(28) Ramachandran, S.; Serohijos, A. W. R.; Xu, L.; Meissner, G.; Dokholyan, N. V. *PLoS Comput Biol* **2009**, *5*, e1000367.

(29) Bormann, J.; Hamill, O.; Sakmann, B. *J Physiol-London* **1987**, *385*, 243–286.

(30) Franciolini, F.; Nonner, W. *J Gen Physiol* **1987**, *90*, 453–478.

(31) Li, M.; McCann, J. D.; Welsh, M. J. *Am J Physiol* **1990**, *259*, C295–C301.

(32) Maduke, M.; Miller, C.; Mindell, J. A. *Annu Rev Biophys Biomol Struct* **2000**, *29*, 411–438.

(33) Hille, B. *Ionic channels of excitable membranes*, 2nd ed.; Sinauer Associates, Inc: Sunderland, Massachusetts, 1992.

(34) Furini, S.; Domene, C. *PLoS Comput Biol* **2012**, *8*, e1002476.

(35) Payandeh, J.; Scheuer, T.; Zheng, N.; Catterall, W. A. *Nature* **2011**, *475*, 353–358.

(36) Nonner, W.; Catacuzzeno, L.; Eisenberg, B. *Biophys J* **2000**, *79*, 1976–1992.

(37) Nonner, W.; Chen, D. P.; Eisenberg, B. *Biophys J* **1998**, *74*, 2327 – 2334.

(38) Nonner, W.; Eisenberg, B. *Biophys J* **1998**, *75*, 1287–1305.

(39) Boda, D.; Busath, D. D.; Henderson, D.; Sokolowski, S. *J Phys Chem B* **2000**, *104*, 8903–8910.

(40) Boda, D.; Henderson, D. D.; Busath, D. D. *J Phys Chem B* **2001**, *105*, 11574–11577.

(41) Boda, D.; Busath, D.; Eisenberg, B.; Henderson, D.; Nonner, W. *Phys Chem Chem Physics* **2002**, *4*, 5154–5160.

(42) Boda, D.; Gillespie, D.; Nonner, W.; Henderson, D.; Eisenberg, B. *Phys Rev E Stat Nonlin Soft Matter Phys* **2004**, *69*, 046702.

(43) Gillespie, D.; Nonner, W.; Henderson, D.; Eisenberg, R. S. *Phys Chem Chem Physics* **2002**, *4*, 4763–4769.

(44) Nonner, W.; Gillespie, D.; Henderson, D.; Eisenberg, B. *J Phys Chem B* **2001**, *105*, 6427–6436.

(45) Allen, R.; Hansen, J.-P.; Melchionna, S. *Phys Chem Chem Physics* **2001**, *3*, 4177–4186.

(46) Boda, D.; Nonner, W.; Valiskó, M.; Henderson, D.; Eisenberg, B.; Gillespie, D. *Biophys J* **2007**, *93*, 1960–1980.

(47) Boda, D.; Valisko, M.; Henderson, D.; Eisenberg, B.; Gillespie, D.; Nonner, W. *J Gen Physiol* **2009**, *133*, 497–509.

(48) Boda, D.; Nonner, W.; Henderson, D.; Eisenberg, B.; Gillespie, D. *Biophys J* **2008**, *94*, 3486–3496.

(49) Boda, D.; Valiskó, M.; Eisenberg, B.; Nonner, W.; Henderson, D.; Gillespie, D. *J Chem Phys* **2006**, *125*, 34901.

(50) Boda, D.; Valiskó, M.; Eisenberg, B.; Nonner, W.; Henderson, D.; Gillespie, D. *Phys Rev Lett* **2007**, *98*, 168102.

(51) Gillespie, D.; Chen, H.; Fill, M. *Cell Calcium* **2012**, *51*, 427–433.

(52) Gillespie, D.; Giri, J.; Fill, M. *Biophys J* **2009**, *97*, 2212–2221.

(53) Gillespie, D.; Fill, M. *Biophys J* **2008**, *95*, 3706–3714.

(54) Gillespie, D.; Boda, D. *Biophys J* **2008**, *95*, 2658–2672.

(55) Gillespie, D. *Biophys J* **2008**, *94*, 1169–1184.

(56) Wang, Y.; Xu, L.; Pasek, D. A.; Gillespie, D.; Meissner, G. *Biophys J* **2005**, *89*, 256–265.

(57) Boda, D.; Giri, J.; Henderson, D.; Eisenberg, B.; Gillespie, D. *J Chem Phys* **2011**, *134*, 055102.

(58) Eisenberg, B. *Fluct Noise Lett* **2012**, *11*, 76–96.

(59) Eisenberg, B. *Trans Faraday Soc* **2012**, *in*, press.

(60) Giri, J.; Fonseca, J. E.; Boda, D.; Henderson, D.; Eisenberg, B. *Phys Biol* **2011**, *8*, 026004.

(61) Eisenberg, B. *Advances in Chemical Physics*; John Wiley & Sons, Inc., 2011; Chapter Crowded Charges in Ion Channels, pp 77–223.

(62) Krauss, D.; Eisenberg, B.; Gillespie, D. *Eur Biophys J* **2011**, *40*, 775–782.

(63) Krauss, D.; Gillespie, D. *Eur Biophys J* **2010**, *39*, 1513–1521.

(64) Eisenberg, R. S. B. Self-organized model of selectivity. 2008; `http://www.ima.umn.edu/2008-2009/W12.8-12.08/abstracts.html`.

(65) Burger, M.; Eisenberg, R. S.; Engl, H. *SIAM J Applied Math* **2007**, *67*, 960–989.

(66) Eisenberg, B. *Biophys Chem* **2003**, *100*, 507–517, Special Issue in honor of John T. Edsall (1902–2002).

(67) This contrasts to other coarse-graining approachs which have a well-defined relationship between structural information and model parameters: Noid, W. G.; Chu, J.-W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. *J Chem Phys* **2008**, *128*, 244114.

(68) Wirth, C.; Condemine, G.; Boiteux, C.; Bernèche, S.; Schirmer, T.; Peneff, C. M. *J Mol Biol* **2009**, *394*, 718–731.

(69) Boda, D.; Henderson, D.; Eisenberg, B.; Gillespie, D. *J Chem Phys* **2011**, *135*, 064105.

(70) Gillespie, D.; Eisenberg, R. S. *Eur Biophys J* **2002**, *31*, 454–466.

(71) Dwyer, T. M.; Adams, D. J.; Hille, B. *J Gen Physiol* **1980**, *75*, 469–492.

(72) Lipkind, G. M.; Fozzard, H. A. *J Gen Physiol* **2008**, *131*, 523–529.

(73) This definition of selectivity is commonly used in Pharmacology and Chemistry for describing competitive processes, although it is less widely used in Biophysics.

(74) We reduce the confusion of having the one letter symbol 'K' for potassium, lysine and equilibrium constants by adding a superscript plus 'K^+' for potassium ions, writing 'Lys' for lysine unless used in a residue sequence such as DEKA and using a subscript 'K_{M_1,M_2}' for equilibrium constants.

(75) Valleau, J. P.; Cohen, L. K. *J Chem Phys* **1980**, *72*, 5935–5941.

(76) Malasics, A.; Boda, D. *J Chem Phys* **2010**, *132*, 124102.

(77) Canessa, C. M.; Schild, L.; Buell, G.; Thorens, B.; Gautschi, I.; Horisberger, J. D.; Rossier, B. C. *Nature* **1994**, *367*, 463–467.

(78) Tomaselli, G. F.; Chiamvimonvat, N.; Nuss, H. B.; Balser, J. R.; Pérez-García, M. T.; Xu, R. H.; Orias, D. W.; Backx, P. H.; Marban, E. *Biophys J* **1995**, *68*, 1814–1827.

(79) Sheng, S.; Li, J.; McNulty, K. A.; Avery, D.; Kleyman, T. R. *J Biol Chem* **2000**, *275*, 8572–8581.

(80) Kellenberger, S.; Auberson, M.; Gautschi, I.; Schneeberger, E.; Schild, L. *J Gen Physiol* **2001**, *118*, 679–692.

(81) Li, J.; Sheng, S.; Perry, C. J.; Kleyman, T. R. *J Biol Chem* **2003**, *278*, 13867–13874.

(82) Anantharam, A.; Palmer, L. G. *J Gen Physiol* **2007**, *130*, 55–70.

(83) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Ithaca, NY: Cornell University Press., 1960.

Figure Legends

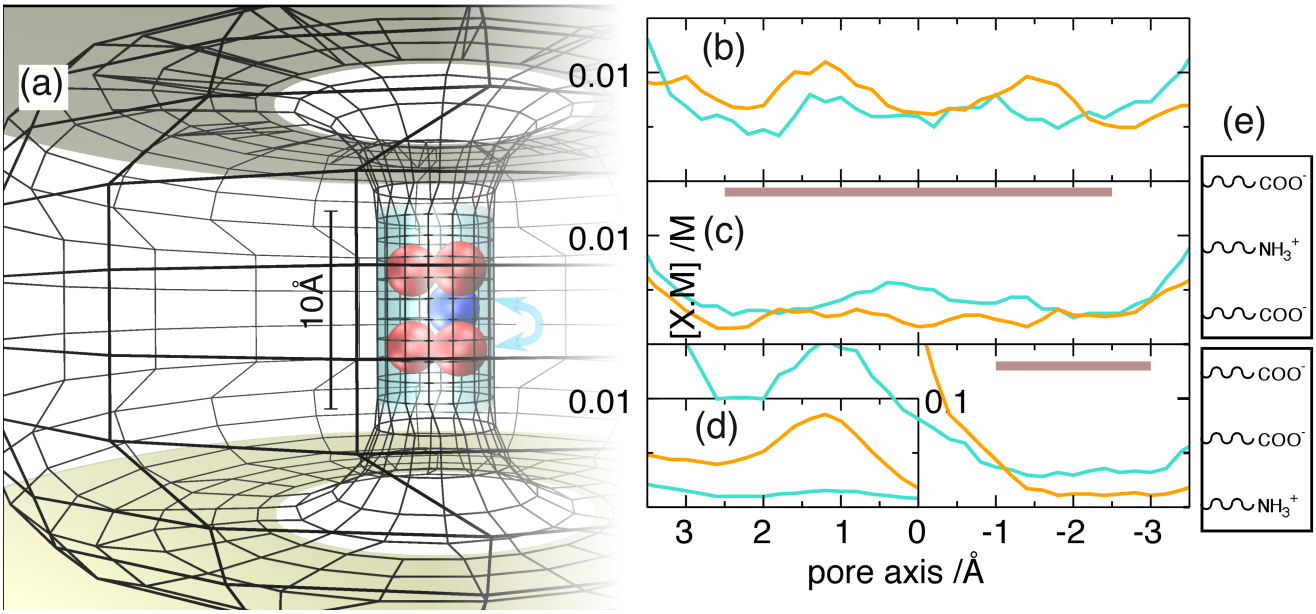

Figure 1: Simulation of Na^+, Ca^{2+}, K^+ and Li^+ occupancy in the eukaryotic DEKA Na^+ channel. (a) Schematic of the CSC model: The surface mesh used to calculate the Possion-Boltzmann induced charge is represented as black lines and lies on the interface between the aqueous media and the channel protein. The D, E and lysine charged side-chains lie within the aqueous zone highlighted in light blue. Lysine is represented as a single positively charged ammonium ion (blue) and the carboxylic acid groups of D and E are represented by two half-negatively charged oxygen ions each (red). The channel axis zero point is the center of the channel with the positive direction pointing up. The graphs show the Na^+ (light-blue) and Ca^{2+} (orange) ion concentrations along the filter region as found by FCC (b) and LCC (c-d) calculations. The Cl^- ion concentrations are zero throughout the channel and are not shown. The brown lines in (c) and (d) show the two sub-intervals used to derive the selectivity. (c) and (d) are the results for the $-+-$ and $--+$ patterns, respectively with the schematic (e) showing how the charged side-chains are arranged in the light blue zone in (a). Shown in (a) is the $-+-$ pattern (e upper), the $--+$ pattern (e lower) corresponds to an exchange, as indicated by the arrow, of the central blue sphere with the two red spheres at the bottom. The concentration profiles are from a simulation of equal concentration of $CaCl_2$ and NaCl with a total ionic strength of 0.22M (220mM). The Inset in (d) shows the very large Ca^{2+} density in the region of the two negative charged side-chains reduced tenfold.

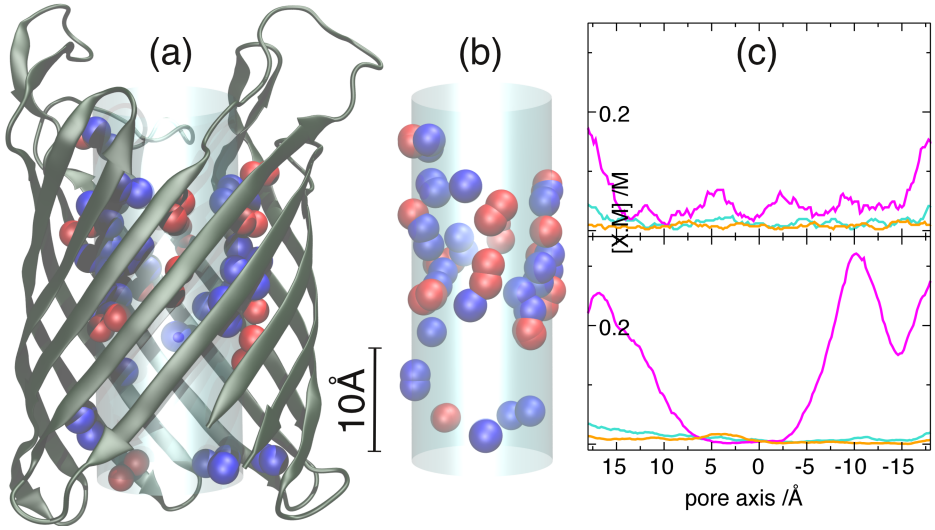

Figure 2: NanC channel from *E. Coli*: from the structural determinants to model calculations. (a) The backbone and all of the charged residues inside the channel; represented as ammonium (blue) and oxygen (red) structural ions analogous to Fig. 1. The channel axis zero point is the center of the channel with the positive direction pointing up. The residues used are listed in Table S3 in the supplementary information and are shown based on available structural information[68]. (b) The pore region cylinder used in the model showing spheres at the LCC *localization centre-points*. Note that in the simulation the structural ions corresponding to any of the spheres that overlap the blue cylinder in the figure will never be able to reach their localization centre-point because they must always be entirely within the cylinder. (c) Na^+ , Ca^{2+} and Cl^- ion concentrations (Color coding as in Fig. 1 plus Cl^- (red)) along the pore region as found by FCC (upper) and LCC (lower) calculations. The bath concentration of the ions are the same as in Fig. 1.

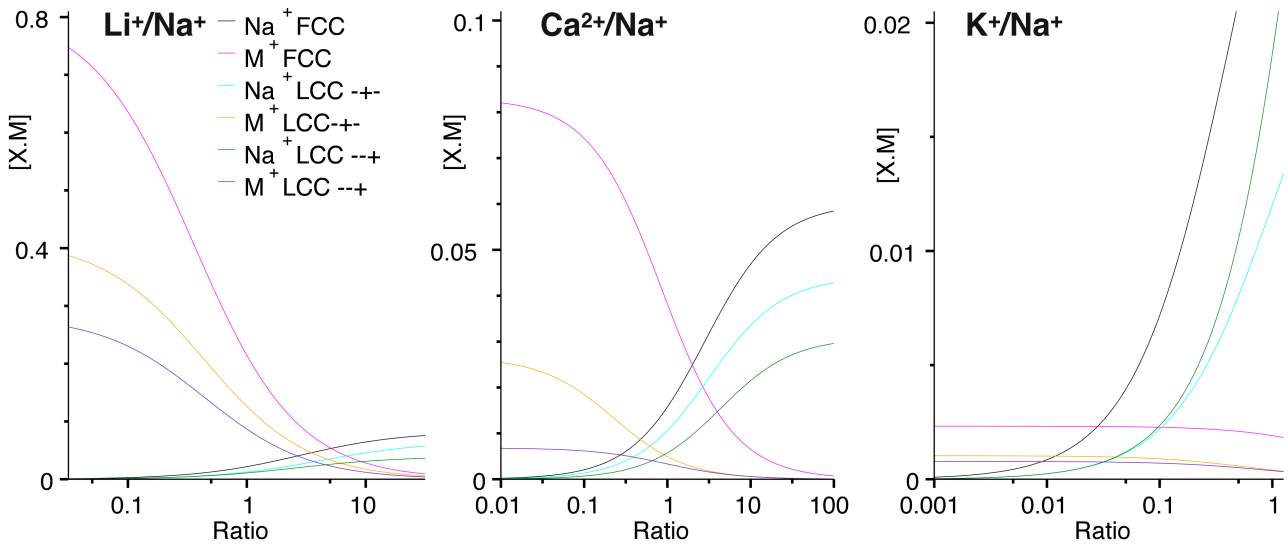

Figure 3: Na^+ channel: Titration of sodium and lithium, calcium or potassium occupancies (normalized to 220mM ionic strength) against [Na]/[Li], [Na]/[Ca] or [Na]/[K] ratio respectively. Note the difference in the occupancy scales between different panels showing different ion pairs, with Na^+ reaching similar occupancy maxima in all three graphs (eg. black lines). Lines represent a fit to Eqn. 9 (see Fig. S4 in the supporting information for calculated values).

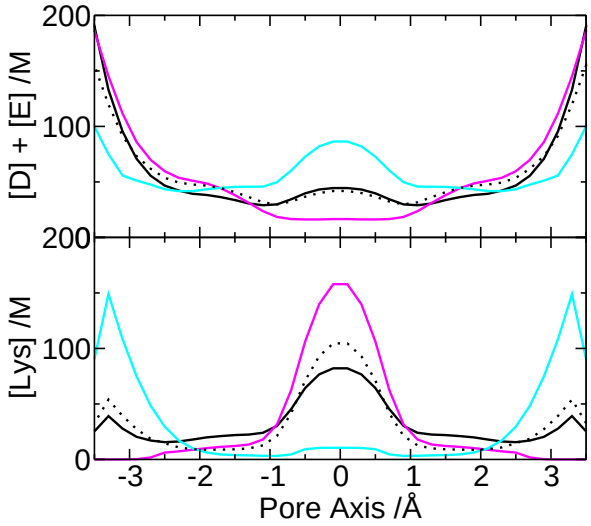

Figure 4: Na^+ channel: Comparison of the concentration of the structural ions along the pore axis with FCC and LCC methods. The black solid line is from FCC model, the red line from LCC with $-+-$ and light blue from LCC with combined ($--+$ and $+--$). The ($--+$ and $+--$) result was the mean of the $--+$ density with a copy of itself reversed along the pore axis. The black dashed line is the scaled sum of the $-+-$ and $--+$ in the ratio two to one, respectively.

Tables

Table 1: Calculated model DEKA Na^+ channel selectivity compared by localization method. The $--+$ and $-+-$ labels show the charge and order of the structural ions along the channel (See Fig. 1e). Interval refers to the sub-region of the channel (shown as a brown line in Fig. 1c–d) used for computing selectivity, with zero being the channel center-point and the selectivity filter region being $[-5:5\text{Å}]$. (See Tab. S2 for full list of the experimental results.)

Method	Na^+ Selectivity (K_{M,Na^+})		
	Li^+	Ca^{2+}	K^+
experimental results			
$\langle P_{M^+}/P_{Na^+}\rangle^{1–10}$	1.0	0.13	0.06
$\langle I_{M^+}/I_{Na^+}\rangle^{77–82}$	1.7	–	<0.01
simulation (interval $[-2.5:2.5]$)			
FCC	4.9	2.0	0.03
LCC $(-+-)$	4.9	0.6	0.04
LCC $(--+)$	6.1	2.7	0.02
simulation (interval $[-3:-1]$)			
LCC $(--+)$	4.6	0.7	0.03

Table 2: Model ion parameters at physiological pH. Model particle charge and radius[83] parameters of the structural and solute ions at physiological pH. Solute ions and lysine residues all have a single charged atom and are modeled as a single hard sphere with a full charge. The carboxylic acid groups of D and E residues and the guanidinium of R residues have two charge bearing atoms and are modeled using two hard spheres having $-\frac{1}{2}$ and $+\frac{1}{2}$ charges, respectively. Because histidines are 10% positively charged at physiological pH their imidazole group should be modeled with $+\frac{1}{10}$ charge. However, NanC has only three histidines in the pore which would give a non-integral total charge, so in this case the charge on each histidine was increased to $+\frac{1}{3}$.

Ion or residue	Modeled using	Charge (q_i) /e	Radius (R_i) /Å
Ca^{2+}		$+2$	0.99
Cl^-		-1	1.81
K^+		$+1$	1.33
Li^+		$+1$	0.60
Na^+		$+1$	1.0
NH_4^+			1.5
O^-			1.4
R(Arg)	$2 \times NH_4$	$2 \times +\frac{1}{2}$	
D(Asp)	$2 \times O$	$2 \times -\frac{1}{2}$	
E(Glu)	$2 \times O$	$2 \times -\frac{1}{2}$	
H(His)	NH_4	$+\frac{1}{3}$	
K(Lys)	NH_4	$+1$	

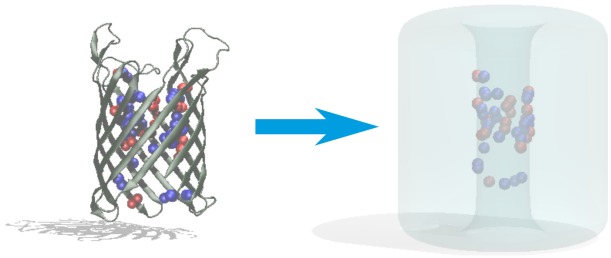

Figure 5: Table of contents graphic for “Localizing the charged side chains of ion channels within the crowded charge models.”