Mapping the Importance of Four Factors in Creating Monovalent Ion Selectivity in Biological Molecules

Michael Thomas, Dylan Jayatilaka and Ben Corry

Abstract: *The ability of macrocycles, enzymes, ion channels, transporters and DNA to differentiate between ion types is often crucial to their function. Using molecular dynamics simulations on both detailed systems and simple models we quantify the importance of several factors which affect the ion selectivity of such molecules, including the number of coordinating ligands, their dipole moment and their vibrational motion. The information resulting from our model systems is distilled into a series of 'selectivity maps' that can be used to 'read off' the relative free energy associated with binding of different ions, and to provide an estimate of the importance of the various factors. While our maps cannot capture all elements of real systems, it is remarkable that they produce differential site binding energies in line with experiment and more detailed simulations for a variety of systems, making them useful for assisting in understanding the origins of selective binding and transport. The chemical nature of the coordinating ligands is essential for creating thermodynamic ion selectivity in flexible molecules (such as 18c6), but as the binding site becomes more rigid the number of ligands (as in ion channels) and the reduction of thermal fluctuations (as in amino acid transporters) can become important. In the future, our maps could aid in the determination of the local structure from binding energies and assist in the design of novel ion selective molecules.*

Introduction

Many biological molecules, including ion channels, transporters, enzymes, macrocycles and DNA, selectively bind or transport ions. In most cases the differentiation between ion types is critical to the function of the molecule. For example, potassium channels must be able to rapidly move K^+ out of a cell while preventing the passage of Na^+, otherwise the electrochemical gradient across the cell membrane would be lost and the cell would die (1). Ions also play important and specific roles in the structure and function of many enzymes; binding of the wrong ion can inactivate such molecules thus perturbing important regulatory systems (2). Natural macrocyclic ionophores such as valinomycin and nonactin, as well as synthetic counterparts such as crown ethers and cryptands, can also selectively complex ions and have found uses in electrophysiology, catalysis and in building ion selective electrodes.

The discrimination between K^+ and Na^+ is particularly interesting given their prevalence in biology and their identical charge, spherical nature and similar size. There has been a long history of describing selectivity in small macrocyclic ligands which particularly highlighted the role of rigidity in creating a cavity that preferentially bound ions of a certain size (3). Biological macrocycles such as valinomycin, however, are conformationally flexible and pioneering free energy simulations highlighted the importance of the solvation energies of the ions and the strength of the electrostatic interaction with the closest ligands as playing an important role in creating selectivity in addition to structural (steric) considerations (4, 5). Recently, there has been much discussion of the origins of selectivity in potassium channels that are able to discern between K^+ and Na^+ ions with up to 1000 fold preference for K^+ (6–9). Selectivity is achieved in a narrow region of the channel known as the selectivity filter which is lined with carbonyl oxygens that coordinate permeating ions (10, 11) creating a thermodynamic preference for binding K^+ relative to Na^+ in the range of 5-6 kcal/mol (6–9, 12, 13). The preference has been suggested to at least partly arise due to the channel better compensating the energy cost of dehydrating K^+ than that of Na^+.

Many of the explanations for ion selectivity in K^+ channels are thermodynamic in nature and this paper works from this basic premise. For a long time the most widely held view attributed selectivity to the better structural fit of K^+ into the selectivity filter binding sites than could occur for the smaller Na^+ (14, 15), seemingly consistent with the crystal structures (10, 16). However, in its simplest form such an explanation requires that the protein maintains a large degree of rigidity as the two ions differ in ionic

radius by 0.38 Å. Noskov and Roux challenged this structural explanation of selectivity arguing that not only is the channel likely to be too flexible for this mechanism to work, but also that such flexible sites can still achieve K^+ selectivity (12, 17).

They showed that the balance of local ion-ligand and ligand-ligand interactions can create preferences for different ions in flexible/dynamic ion binding sites, in line with ideas presented by Eisenman (18) and previously suggested to operate in valinomycin (4). Thus, the precise selectivity is determined by the chemical nature and the number of ligands coordinating the ion which determine the magnitude of the ion-ligand and ligand-ligand interactions. Considerable subsequent analysis has attempted to assess the relative importance the ligand type and ligand number in potassium channels. While all acknowledge that both can be important, some emphasise the role of restriction of coordination numbers (13, 19–21), others reemphasise the chemical nature of the ligands (22, 23) while others maintain a more agnostic view (24–26). Kinetic factors have also been suggested to be important in preventing intracellular Na^+ from permeating the pore (27).

All the factors discussed for K^+ ion channels are likely to be important to a greater or lesser extent for ion selectivity in a large range of biological ion binding sites. It is our aim to investigate the conditions in which each factor becomes important in creating ion selectivity in a number of specific biological molecules. Although there are an infinite range of factors over which to explore selectivity, here we systematically explore only four factors; (i) the magnitude of the dipole moment of the coordinating ligands, (ii) restrictions on the number of coordinating ligands and the somewhat related concepts of (iii) cavity size of the coordination site and (iv) the thermal fluctuations in the positions of the ligands. Factors (ii), (iii) and (iv) are imposed by the protein scaffold and share some characteristics with a weak interpretation of the early structural fit hypothesis, whereas (i) is an intrinsic property of the coordinating ligands.

Admittedly, the factors chosen for investigation in this paper are arbitrary. For example one could choose instead to examine the ion-ligand and ligand-ligand energy contributions. We chose the above set of factors because we believe they afford the most explanatory power since they can used to directly analyse binding site structures found in the PDB. Furthermore, as defined below, the energetic contributions arising from these factors can be defined in such a way that they additively combine to the total selectivity of the site.

Model systems have been used to isolate each of the factors above in order to investigate the conditions under which each dominates the con-

tribution to ion selectivity. The models range from an ion in a liquid of ligands of variable dipole moment, to an ion surrounded by a fixed number of ligands constrained within a sphere and, finally, to an ion surrounded by ligands tethered with variable force constants to the vertices of classic coordination polyhedra at a fixed distance from the central ion. The resulting information yields insight into the interplay of the various factors in achieving selectivity in a range of biological molecules. The understanding gleaned from these models is tested against additional detailed atomistic molecular dynamics simulations on the full biological molecules in question and with experimental data.

The use of model systems to study selectivity in ion channels was pioneered by Noskov et al (12). In this paper we hope to extend this work and show how the binding energy explicitly depends simultaneously on all the factors described. Bostick and Brooks have also used simplified models and molecular dynamics simulations to investigate ion selectivity (21). Binning techniques with an ion solvated in a liquid were used to derive how the free energy of binding depends on the average ion-ligand separation and the average coordination number. In this paper we explore more parameters, in particular the dipole moment of the ligand and the dynamic flexibility of the coordinating atoms. Like those authors, we produce selectivity maps, but in our case these are obtained by using constrained (biased) sampling of the configuration space. We have previously used simplified molecular dynamics (MD) models and quantum mechanical (QM) models to examine selectivity in KcsA (13). Varma (19) and Dudev (25) examine a larger range of parameters within a QM framework highlighting similar trends to that seen in MD, but such QM approaches cannot easily cover the same range of parameters as the MD simulations carried out here.

The question arises whether the simplified models used in this paper are capable of producing quantitative or even semi-quantitative ion binding energies. This depends firstly on whether the free energy of binding is dominated by local contributions. Secondly, it depends on whether we can calculate accurate energies for the locally interacting atoms; in general this is a question of the accuracy of the force field or other methods used. The veracity of the former assumption can be tested by comparison with large scale atomistic calculations. The latter can be tested against more accurate calculations and/or experimental data.

Non-local effects such as strain energy with atoms not directly coordinated to the ion and cases where different ion types have a different number of coordinating ligands when in the same site could be important and therefore would cause simplified models like ours to break down. A number of

these effects have been examined in detail and are briefly presented here and at length in the supplementary material. Provided one keeps in mind these limitations, the selectivity maps that we produce from our models could be of great value for interpreting and understanding ion selective binding.

Methods

Model Systems. Four local factors affecting ion selectivity (dipole moment of the ligand, number of ligands, cavity size and thermal fluctuations) were investigated using free energy perturbation molecular dynamics (FEP MD) simulations. Ion exchange reactions were used to compare selectivity between ions. Fig. 1 is an overview of the three families of model systems, successively more elaborate and nuanced, used to model a generalised ion binding site. The most complex model considers a FEP between two ions where each of the dipole moment (characterised by q), the number of ligands, n, the distance of the ligands from each ion, r_{K^+} & r_{Na^+} and the thermal fluctuation of the ligands (characterised by k) are systematically varied. By building up from a simple system, each of these factors can be investigated in turn. Once we have chosen to examine particular factors, the only obvious way in which to explore them is to build them up in the sequence shown in the Fig. 1 and described below.

In the first family of model systems, Fig. 1, family 1, we looked at the influence of ligand dipole moment in isolation by considering the exchange of an ion M^+ (where M=Li, Na, Rb, Cs) with K^+ between water and a second solvent, $Sol(q)$, with controllable dipole moment (controlled by the partial charge q):

$$M^+/H_2O + K^+/Sol(q) \quad \xrightarrow{\Delta G_q} \quad K^+/H_2O + M^+/Sol(q) \qquad (1)$$

The change in free energy ΔG indicates the partitioning of the two ions (M^+ and K^+) between the solvents. The ΔG value for the exchange reaction is calculated as the combination of two individual FEP simulations (28) (one for each solvent) in which an ion of one type is slowly morphed into another over a number of steps. Bulk water was represented by a 30×30×30 Å TIP3P periodic water box with a counter ion while the second solvent was a 30×30×30 Å box of abstract ligands based on the structure of formaldehyde molecules. The model ligands are not intended to represent formaldehyde itself, but rather an abstract, simple dipole that we can manipulate as a crude representation of range of ligands of different chemical composition. The dipole moment of the abstract ligands was set by altering the partial

charge on the carbon and oxygen atoms (C=q, O=-q, H1=0, H2=0). The model ligand box simulation was equilibrated with partial charge 0.51 under an NPT ensemble before subsequent FEP simulations were conducted under an NVT ensemble such that the ligand density was the same for all values of the partial charge.

In the second family of binding site models, Fig. 1, family 2, restrictions on the coordination number of the ion were imposed to add to the effect of the ligand dipole moment. This was done by forcing n ligand oxygen atoms into the first solvation shell of the ions, yielding another ion exchange reaction:

$$M^+/H_2O + K^+/model(n,q) \quad \xrightarrow{\Delta G_{nq}} \quad K^+/H_2O + M^+/model(n,q) \quad (2)$$

For this purpose the oxygen atoms were held with a flat bottom, steep harmonic potential within a 3.5 Å sphere (4.0 Å for Cs^+) representing position of the first minimum in the radial distribution function of K^+ in water. Three varieties of this second family of model sites were considered as shown in Fig. 1; all comprised of the n ligands forming the binding site which were surrounded by (a) periodic box of model ligand molecules, (b) finite spherical droplets of model ligand molecules, and (c) a vacuum. The significance of these different models is discussed below.

Whether a K^+ ion is thermodynamically likely to leave the bulk water to enter a binding site was determined from the FEP-MD of another ion exchange reaction:

$$K^+/H_2O + model(n,q) \quad \xrightarrow{\Delta G_{entry}} \quad H_2O + K^+/model(n,q) \quad (3)$$

The free energy was calculated by conducting two FEP calculations: the first where K^+ disappears in a 30×30×30 Å TIP3P periodic water box and the second where K^+ appears in a model ligand system. The electrostatics in these simulations were conducted using a cutoff of 12 Å, so that problems associated with using a lattice sum method could be avoided (29). Two additional sources of error could potentially arise due to the use of the double-annihilation method (30) and the sensitivity of the system to the boundary conditions and the treatment of electrostatics (29). Correction terms were calculated for each of these factors (see supp. mat.). ΔG_{entry} for the other ion types is determined by combining equations (2) and (3).

In a final family of binding sites, Fig. 1, family 3, the influence of cavity size and thermal fluctuations was investigated on top of the ligand dipole and coordination number. In these the n ligands coordinating the ion were placed at the vertices of optimal coordination polyhedra (trigonal

bipyramidal for $n = 5$, octahedral for $n = 6$, cubic for $n = 8$). The size of the binding cavity was controlled by placing a harmonic restraint with force constant k on the oxygen atoms at a distance r_{K^+} and r_{Na^+} from K^+ and Na^+ respectively; and the thermal fluctuations of the coordinating ligands (measured by the RMSD of the oxygen atoms relative to a fixed ion) controlled by the magnitude of k. The oxygen atoms were additionally constrained to the same 3.5 Å sphere about the ion to maintain the desired coordination number, avoiding entropic divergences that could arise in an unconstrained system. To determine the effects of restricting ligand thermal fluctuations on selectivity, two separate sets of FEP were required. This is to account for the changing ion-ligand distance when different ions occupy the binding site.

$$K^+/model(r_{K^+}, k \neq 0) \xrightarrow{\Delta G_{kr_{K^+}}} K^+/model(r_{K^+}, k = 0) \quad (4)$$

$$Na^+/model(r_{Na^+}, k \neq 0) \xrightarrow{\Delta G_{kr_{Na^+}}} Na^+/model(r_{Na^+}, k = 0) \quad (5)$$

There is now a complete description of the ion exchange reaction between K^+ and Na^+ in a hypothetical binding site incorporating each of the local factors:

$$Na^+/H_2O + K^+/model(n, q, r_{K^+}, k) \xrightarrow{\Delta G_{nqkr_{K^+}r_{Na^+}}} K^+/H_2O + Na^+/model(n, q, r_{Na^+}, k) \quad (6)$$

The change in free energy from the exchange reaction of K^+ and Na^+ between water and the binding site that includes all our factors of interest is then (as dictated by (2) + (4) - (5)):

$$\Delta G_{nqkr_{K^+}r_{Na^+}} = \Delta G_{nq} + \Delta G_{kr_{K^+}} - \Delta G_{kr_{Na^+}} \quad (7)$$

Unless otherwise stated, simulations are run using NAMD2 (31) with the CHARMM27 force field (32) at 310 K with 1 fs timesteps. More detailed simulation setups are provided in the supp. mat.

Energy contributions. To determine the contribution to selectivity from the various factors, we combined the results of the ion exchange reactions described above. The free energy contribution from ligand dipole alone is derived from the first family of model systems as ΔG_q (equation 1). In the idealised case, depicted on the left side of Fig. 1, the only difference between

the first and second families of model system is the constraint on the ligand number, and thus the free energy contribution from coordination number restriction, ΔG_n, can be determined from equations 1 and 2:

$$\Delta G_n = \Delta G_{nq} - \Delta G_q \tag{8}$$

The only difference between the second and third families of model system is the inclusion of a harmonic restraint placed on the oxygen atoms of the coordinating ligands and by placing these atoms at r to control cavity size and thermal fluctuations. Thus, the free energy contribution is calculated from equations 4 and 5 as

$$\Delta G_{kr_{K^+}r_{Na^+}} = \Delta G_{kr_{K^+}} - \Delta G_{kr_{Na^+}} \tag{9}$$

Defined in this way, the components are strictly additive, ie, the total free energy difference in model system 3 is the sum of these three free energy contributions.

Unfortunately, using the idealised model systems described above is computationally expensive. As each contribution is added, the number of parameters used to describe the model system increases, and thus the number of simulations needed to survey the parameter space increases significantly. To this end, some approximations were used in our systems in order to decrease the time needed to conduct the FEP, as depicted on the right side of Fig. 1. The ideal model system 2 (using a periodic box of model ligands) was used for one set of parameters only, namely $(n, q) = (8, 0.5)$. To reduce the size of the medium surrounding the binding site, subsequent simulations were conducted by keeping only the model ligand molecules within in a 10.5 Å radius sphere around the ion and simulating the system in a spherical constraint rather than with periodic boundaries (model 2b). To reduce the size of the system even further, a final family of models were used in which all model ligand molecules not directly coordinating the ion were removed. One could argue which model system best represents the environment of a protein binding site and for this reason we include results of all model systems in the results table 2 or in the supp. mat. For our energy components to remain strictly additive, a consistent family of model systems must be used. Since we have introduced approximations our energy components should be seen as indicative rather than as strictly additive, however, as shown in the supp. mat. the difference in the relative contributions of each of the factors found using the ideal and approximate model systems is small.

The fractional contribution to ion selectivity is defined as follows:

$$\chi_q = \frac{\Delta G_q}{|\Delta G_q| + |\Delta G_n| + |\Delta G_{kr_{K^+}r_{Na^+}}|} \tag{10}$$

$$\chi_n = \frac{\Delta G_n}{|\Delta G_q| + |\Delta G_n| + |\Delta G_{kr_{K^+}r_{Na^+}}|} \tag{11}$$

$$\chi_r = \frac{\Delta G_{kr_{K^+}r_{Na^+}}}{|\Delta G_q| + |\Delta G_n| + |\Delta G_{kr_{K^+}r_{Na^+}}|} \tag{12}$$

where χ_q, χ_n and χ_r are portions from dipole moment, coordination number restriction and cavity size/thermal fluctuations respectively. To obtain these values, ΔG_q and ΔG_n were calculated using the 2B family of model systems from Fig. 1. This was done so that the fractional contributions corresponding to these could be broken down in an additive fashion. $\Delta G_{kr_{K^+}r_{Na^+}}$ was calculated using model 3.

Biological Simulations. For comparison with the model systems, FEP simulations are also conducted on full biological systems (eg full protein, solvation box and lipid, periodic boundaries under an NPT ensemble) for each molecule studied. Details of each simulation system are given in the supp. mat. The RMSD information was calculated by Celik et al during simulations investigating substrate binding in in the leucine transporter (33).

Results and Discussion

Influence of ligand dipole moment

The role played by the dipole moment of the coordinating ligands is elucidated by investigating the free energy changes involved in an ion-exchange reaction for the group I ions between water and a second solvent with a dipole moment controlled by adjusting the partial charge on each end of the dipole. For simplicity in our results we refer to only the positive value of the partial charge on the ligands q. As the parameters for Li^+, Rb^+ and Cs^+ are less well developed, results for these ions are included only to show qualitative trends.

As can be seen in Fig. 2, the smaller ions (Li^+ and Na^+) selectively partition into the solvents of greater dipole moment and the larger ions (Rb^+ and Cs^+) into that with smaller dipole moment. A binding site with highly charged ligands is therefore likely to be selective to smaller ions and vice versa, confirming that selectivity can in principle be generated considering only the dipole moment of the ligands. This supports previous studies that suggest the presence of highly charged ligands could be responsible for the Na^+ selectivity of sodium channels (12, 13) and the binding sites within the GluR5 kainate receptor (34). The ligands used in our model systems represent an abstract generalised dipole and this must be taken into account

when equating them to real chemical groups. For example, the dipole moment of the protein backbone carbonyl group will be somewhere between a 'bare' carbonyl group ($\sim$ 3.0 debye, partial charge 0.51) and that with an amide plane, as with n-methyl acetamide ($\sim$4.0 debye, charge $\sim$ 0.65). In the case of the potassium channel KcsA selectivity filter the dipole moment of the carbonyl group could contribute at most 2 kcal/mol (out of a total of 5-6 kcal/mol) toward K^+ selectivity.

Influence of coordination number and dipole moment

The influence that restricting the number of ligands around the ion has on ion selectivity can be examined through ion exchange reactions in the second family of model systems in which we control the ions coordination number n as well as the partial charge on the ligands q as above. This allows us to map the expected ion selectivity of a binding site both qualitatively and quantitatively as shown in Fig. 3. In addition, we show the conditions in which an ion is thermodynamically unlikely to leave the bulk water environment and enter the site by the black 'exclusion zone' ($\Delta G_{entry} > 0$) as such a binding site is not likely to be useful in a biological context.

The 'selectivity maps' shown in Fig. 3A and B enable us to predict the thermodynamic selectivity of a (flexible/liquid like) binding site based only upon the number and charge of the coordinating ligands. It can be seen in Fig. 3B that smaller ions can be selected either by increasing the charge on the ligands, or reducing the number of coordinating ligands; while selecting a large ion is most easily done by having a large number of ligands with small dipole moment (the thin, yellow Cs^+ selective region at low coordination number and large dipole moment is effectively non-selective because of the small free energy difference - see contour lines in Fig. 3A). The applicability of this map can be determined by examining how well it predicts the selectivity of known ion selective structures. Our previous work on the potassium channel, for example (13) indicates that an ion in the selectivity filter has a coordination number of 8. Assuming a carbonyl ligand partial charge of 0.5 - 0.6, the related point on the selectivity map (① in Fig. 3A) shows a free energy difference between K^+ and Na^+ of about 5.5 - 6.5 kcal/mol, similar to both experimental estimates (6–9) and those from detailed molecular dynamics studies of the system (12). Similar predictive success can also be achieved for a number of other ion selective molecules such as a model Na^+ channel, DNA quadruplex and aminoimidazole riboside kinase (ARK), as indicated on Fig. 3A. Some discrepancies arise if the partial charge on all the ligands are not equal. For instance, the presence of

some fully charged ligands favours smaller ions (12, 35), while the presence of water dissipates selectivity (24, 35) (as seen in Fig. 7 in the supp. mat.). However, using the average charge of the ligands gives a good indication of the likely selectivity using Fig. 3A. By extending our model, it is also possible to quantify selectivity when coordination number of Na^+ can be different to that of K^+ as is shown in Fig. 3F and in greater detail in the supp. mat. Remarkably our simple system replicates the trends in ion selectivity seen in a variety of complex molecules.

The origins of selectivity discussed so far consists of the chemical nature of the ligands and the restriction of coordination number. How much each of these contribute toward the total ion selectivity for each value of n and q is shown in Fig. 3C and 3D. It can be seen from the white band in Fig. 3D that when $q \sim 0.6$ the dipole moment does not differentiate between the ions at any coordination number. Any selectivity at this partial charge is thus dominated by coordination number restriction. The vertical blue band at $n = 4 - 6$ in Fig. 3C representing selectivity for Na^+ is expected as these coordination numbers are known to be most favourable for Na^+(13, 20). Similarly coordination numbers $n > 6$ favour K^+. Looking at the parameters for a potassium ion channel, the restriction of coordination number contributes about 60% (3.9 kcal/mol) of total ion selectivity while the dipole moment contributes the remaining 40% (2.6 kcal/mol) (assuming a carbonyl partial charge of 0.51).

Influence of cavity size and ligand thermal fluctuations

There are some structures such as the leucine transporter (LeuT) and aspartate transporter (Glt_{Ph}) for which the map shown in Fig. 3A does not accurately predict selectivity. This is due to the fact that we have yet to consider the third possible origin of ion selectivity: the effects of restricted cavity size or ligand fluctuations which have previously been shown to be important in LeuT (36). To examine in a general context how structural restraints that influence cavity size and ligand fluctuations can affect ion selectivity we constructed the third family of model binding sites, giving us additional control of the equilibrium ion-oxygen distance for each ion type r_{K^+} and r_{Na^+} as well as the size of the thermal fluctuations of the ligands, controlled by the harmonic force constant k.

As we are adding three new parameters to our investigation of selectivity (r_{K^+}, r_{Na^+}, k) only particular combinations can be easily visualised. In the first instance we show selectivity for various values of r_{Na^+} and r_{K^+} for three specific force constants, as shown in Fig. 4A and supp. mat. Fig. 10A

and B. When the force constant is small, cavity size and ligand fluctuations are relatively unconstrained and so there is little contribution from these toward selectivity (supp. mat. Fig. 10A and B). With a larger force constant, Fig. 4A, the situation is a bit more complicated. If $r_{K^+} = r_{Na^+}$ and the force constant is large, the cavity size is being rigidly constrained and this can create selectivity using a classic snug fit mechanism (a complete map of this case is shown in Fig. 9 of the supp. mat.). If the cavity is large ($>\sim$ 2.5 Å), K^+ will be more favoured in the site. If it is small ($<\sim$ 2.5 Å), Na^+ will be more favoured. In our detailed simulations of complete biological molecules, we find that this rigid cavity situation never arises. That is, the size of the binding cavity is always different with Na^+ bound than with K^+, even in the transporters in which cavity size was expected to be restricted. For example in the second ion binding site, S2, in the LeuT, the average ion-ligand distance is 2.69 Å for K^+ and 2.31 Å for Na^+, very close to the average ion-ligand distances found for those ions in a solvent. Thus, we believe that the most useful case to examine in our model systems is where r_{Na^+} and r_{K^+} are taken to be the first peak in the radial distribution function of the equivalent unconstrained system (approximately 2.4 and 2.7 Å respectively). In this case, the cavity size is not being rigidly constrained, that is, the ion-ligand distance can be optimised for the ion type. However the thermal fluctuations of the ligands are still being influenced by the harmonic restraint. At large RMSD (small k) the contribution of limited ligand fluctuations toward ΔG is 0. Decreasing fluctuations (increasing k) contributes to Na^+ selectivity in the cases with 5 or 6 ligands and towards K^+ selectivity in the 8 ligand case, as shown in Fig. 4B. This shows that even when the radius of the cavity is not constrained, reducing the magnitude of the thermal fluctuations of the coordinating ligands can create ion selectivity. The possibility of ligand fluctuations influencing selectivity has been suggested previously (12), but here we show that it can be important even when the equilibrium cavity size is allowed to optimise for each ion type, as well as separating these effects from those of ligand type, number and cavity size.

In the two transporters considered in this study an ion-ligand distance close to that seen in solution is maintained for each ion, however, the ligands display smaller than usual RMSD fluctuations of about 0.3-0.5 Å. In comparison, the ligand RMSD in the enzyme ARK is between 0.55-0.6 Å, while in KcsA it is greater than 0.7 Å (12). Examining Fig. 4B, it is evident that decreased fluctuations of the ligands will have a substantial effect on ion selectivity in the transporters, a small effect in ARK and little effect with KcsA. Thus, it is the last situation considered with respect to cavity

size and ligand flexibility, reduced ligand fluctuations but not constrained cavity size, that we believe is the most informative for the molecules studied here. The influence of reduced ligand thermal vibrations adds to possible influences of the other effects.

From models to reality

Having considered four mechanisms of ion selectivity, we now have the necessary tools to determine in which biological molecules each these mechanisms is likely to be important. To demonstrate the practicality of this model, we compare predictions from the maps with experimental data for a number of different ion selective biological molecules with ion binding sites of differing chemical composition, as shown in table 1. The results of this study are shown in table 2. In addition, we ascertain how well the model systems capture reality by conducting additional detailed MD simulations of each molecule in its natural environment as detailed fully in the supp. mat. For each molecule, n, q, r_{K^+}, r_{Na^+} and RMSD of the coordinating ligands (calculated from the detailed simulations) was used to predict selectivity of the ion binding site (ΔG) via the selectivity maps. MD FEP simulations of each molecule were conducted within the detailed molecular dynamics simulations for comparison, and experimental data was gathered from the literature.

As can be seen in table 2, the trends seen in our selectivity maps are comparable with the values derived through detailed simulation and experiment (except for valinomycin which is discussed below). As well as yielding selectivity for the correct ion, the simple maps give an indication of the magnitude of this selectivity. A strength of this map approach is that an estimate of the relative importance of each mechanism investigated can be determined as shown in the table. It should be noted that such relative contributions should be taken as indicative values only, rather than being strictly quantitative. They are nevertheless very useful for gaining an understanding and intuition of the systems.

¿From these results it is possible to see that different molecules make use of different means to differentiate between ions, but that all the factors described tend to have some influence in each case. As we and others have suggested previously, the ability of potassium channels to enforce a large coordination number around the permeating ions appears to be important in determining the the thermodynamic favourability of the binding sites for K^+. The specific chemical nature of the carbonyl ligands is also important accounting for up to 40% of the total if we assuming a carbonyl partial

charge of 0.51, but less if the effective dipole moment on the backbone carbonyls is larger. The large size of the thermal fluctuations of the carbonyl oxygens seen in the crystallographic and simulation data (12), as well as the ability of the site to shrink around small ions in our simulations suggest that the cavity size and restrictions of the fluctuations of the ligands have little influence on the selectivity of the site.

The enzyme ARK, on the other hand, has a smaller average dipole moment on the ligands, something that favours the binding of K^+ over Na^+. Combined with a moderate reduction in the RMSD of the ligands that also favours K^+, this overcomes the effect of the small number of ligands surrounding the ion would otherwise favour Na^+. As noted above, the ligands in the LeuT and Glt_{Ph} binding sites have smaller than usual fluctuations which appears to to be important in determining the Na^+ selectivity of both binding sites, even though previous work has suggested only one of these is rigid in LeuT (36). The presence of the zwiterionic leucine in site Na1 of LeuT also creates a significant preference for Na^+ over K^+ in accord with previous results (36) and our simple models with some fully charged ligands (supp. mat. Fig. 7). In Glt_{Ph}, site Na2 contains no fully charged ligands and thus has a relatively lower average dipole moment than the other transporter binding sites. For these reasons it seems that reduced ligand fluctuation plays a large role in determining the selectivity in each of the transporter sites, which is further enhanced in the Na1 sites by the presence of acidic residues.

The selectivity filters of sodium channels contain a conserved DEKA motif. The presence of highly charged ligands around the permeating ion would create a preference for Na^+ in accord with our maps and previous suggestions (12, 13). In addition, there is likely to be a smaller number of ligands surrounding the ions in this case than in K^+ channels which also leads to some degree of Na^+ selectivity in our model. Our maps also suggest that the flexible crown ether, 18-crown-6, relies on the dipole moment of coordinating oxygens to achieve K^+ selectivity over Na^+.

Our results also support the understanding of the differing selectivity of K^+ and NaK channels. Crystallographic and computational studies of the NaK channel indicate that the slight differences in the structure of NaK channel enable water molecules to more easily contact the permeating ion (24, 35, 37). As shown in Fig. 3E, and in greater detail in the supp. mat., the substitution of two carbonyl groups with two water molecules reduces selectivity by more than 2 kcal/mol. Previous studies also show that coordination numbers are less constrained in NaK than in KcsA (24, 35, 37). As shown in Fig. 3F this can reduce selectivity by a further 4 kcal/mol.

Together these effects lead to little differentiation between Na^+ and K^+. The number of coordinating ligands also changes with ion type in nonactin, resulting in a small predicted selectivity similar to the case of NaK.

While our model systems enable us to visualise trends in factors that create ion selectivity, there are some elements of real biological systems that they cannot capture. In our model the ligands are not bound to one another and the strain in the molecular scaffold that may arise when the ligands adopt particular configurations, a non-local effect, is only approximated through a harmonic potential. Understanding when molecular strain is likely to influence selectivity requires knowledge of the particular molecule in question. It can be expected to be smaller in cases where the individual ligands have a large degree of flexibility, such as when they are not directly bound to one another as in large proteins and many of the cases studied here. One exception is valinomycin: although Na^+ can optimise its coordination number, this comes at the energy cost of breaking a hydrogen bond in the backbone of the molecule, a non-local effect that is poorly captured by our harmonic strain model. We note that in some situations there are differences between the simple model systems and real molecules. In LeuT one of the ion binding sites abuts the fully charged zwitterionic leucine being transported by the protein. As noted above, our maps become less accurate (in this case the role of dipole moment is underestimated) when the charge on any of the ligands deviates far from the average. The accuracy of the simple model can be improved by including inhomogeneous charges as done in the supp. mat. We also do not include differential polarisations of the ligands in the presence different ions.

To further analyse the validity of our approach we consider how altering the force field model or simulation protocol influences our results. As shown in the supp. mat., use of the OPLS force field (38, 39) in place of CHARMM27 (32) within our model systems changes the predicted selectivity by less than 1 kcal/mol for all the conditions corresponding to the molecules studied. Changing the setup of the model system has a small effect on the exact degree of selectivity seen in our model systems, but does not change the overall trends. For example, the environment of a real ion binding site could influence its selectivity, and it has previously been suggested that this can change the coordination numbers of ions in the site (19). To examine this, selectivity maps were created with the model binding sites surrounded by a bulk medium (shown in supp. mat.) in addition to those in which the site is surrounded by vacuum (as in Fig. 3). The effect of this is to increase the dielectric constant around the site, and as can be seen by the corresponding column in table 2 ($\Delta G\ \epsilon > 1$) this leads to a slight

shift in the predicted selectivity toward Na^+ over K^+ compared to the predictions made in vacuum. This result is in accord with the suggestions of Varma and Rempe (19): increasing the dielectric constant around the site makes it possible for ligands to reduce their interaction with the ion and replace them with interactions with surrounding molecules. This situation is more favourable to Na^+ than K^+ as it allows for an effective reduction in coordination number. The differences in the predictions made with and without a surrounding dielectric medium shown in table 2 gives an indication of the uncertainty inherent to our scheme due to the fact it does not attempt to mimic the specific environment of each of the biological binding sites studied.

We have presented a scheme for understanding the mechanisms behind ion selectivity that involves four factors that can influence the thermodynamics stability of a binding site. By conducting detailed simulations in parallel with our simplified models we hope to have achieved a useful compromise that allows general principles to be drawn while checking this is still a plausible representation of reality. While our scheme cannot capture all the effects that can contribute to ion selectivity, we have shown that it is applicable to a large range of biological and synthetic molecules, each of which employs these factors in different degrees. We believe that our results provide valuable insight into fundamental biological processes and can be used to estimate the nature of uncharacterised sites or in the development of novel ion selective molecules.

Acknowledgements

The authors gratefully acknowledge support of this work from the National Health and Medical Research Council, an award under the Merit Allocation Scheme on the NCI National Facility and additional computer time from iVEC. We would like to thank Prof H.B. Bürgi for his careful reading of the manuscript, Leyla Celik for the LeuT RMSD data and Rasha Ruhayel for help preparing Fig. 1.

References

1. Hille, B., 2001. Ionic Channels of Excitable Membranes. Sinauer Associates Inc., MA., 3^{rd} edition.

2. Page, M., and E. di Cera, 2006. Role of Na^+ and K^+ in enzyme function. *Phyiol. Rev.* 86:1049–1092.

3. Dietrich, B., 1985. Coordination chemistry of alkali and alkali-earth cations with macrocyclic ligands. *J. Chem Ed.* 62:954–964.

4. Åqvist, J., O. Alvarez, and G. Eisenman, 1992. Ion-selective properties of a small ionophore in methanol studied by free energy perturbation simulations. *J. Phys. Chem.* 96:10019–10025.

5. Marrone, T., and K. Merz Jr., 1995. Molecular recognition of K^+ and Na^+ by valinomycin in Methanol. *J. Am. Chem. Soc.* 117:779–791.

6. Yellen, G., 1984. Ionic permeation and blockade in Ca^{2+} activated K^+ channels of bovine chromaffin cells. *J. Gen. Physiol* 84:157–186.

7. Heginbotham, L., and R. MacKinnon, 1993. Conduction properties of the cloned Shaker channel. *Biophys. J.* 65:2089–2096.

8. LeMasurier, M., L. Heginbotham, and C. Miller, 2001. KscA: it's an ion channel. *J. Gen. Physiol* 118:303–314.

9. Nimigean, C. M., and C. Miller, 2002. Na^+ block and permeation in K^+ channel of known structure. *J. Gen. Physiol* 120:323–325.

10. Doyle, D., J. M. Cabral, R. Pfuetzner, A. Kuo, J. Gulbis, S. Cohen, B. Chait, and R. MacKinnon, 1998. The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science* 280:69–77.

11. Zhou, Y., J. H. Morais-Cabral, A. Kaufman, and R. MacKinnon, 2001. Chemistry of ion coordination and hydration revealed by a K^+ channel-Fab complex at 2.0Å resolution. *Nature* 414:43–48.

12. Noskov, S., S. Bernèche, and B. Roux, 2004. Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* 431:830–834.

13. Thomas, M., D. Jayatilaka, and B. Corry, 2007. The predominant role of coordination number in potassium channel selectivity. *Biophys. J.* 93:2635–2643.

14. Mullins, L. J., 1959. The penetration of some cations into muscle. *J. Gen. Physiol.* 42:817–829.

15. Bezanilla, F., and C. Armstrong, 1972. Negative conductance caused by entry of sodium and cesium ions into the potassium channels of squid axons. *J. Gen Physiol* 60:588–608.

16. Lockless, S., M. Zhou, and R. MacKinnon, 2007. Structural and thermodynamic properties of selective ion binding in a K^+ channel. *Plos Biology* 2:1079–1088.

17. Noskov, S., and B. Roux, 2006. Ion selectivity in potassium channels. *Biophys. Chem.* 124:279–291.

18. Eisenman, G., 1962. Cation selective glass electrodes and their mode of operation. *Biophys. J.* 2:259–323.

19. Varma, S., and S. Rempe, 2007. Tuning ion coordination architectures to enable selective partitioning. *Biophys. J.* 93:1093–1099.

20. Bostick, D., and C. Brooks, 2007. Selectivity in K^+ channels is due to topological control of the permeant ion's coordinated state. *PNAS* 104:9260–9265.

21. Bostick, D., K. Arora, and C. Brooks, 2009. K^+/Na^+ selectivity in toy cation binding site models is determined by the 'host'. *Biophys. J.* 96:3887–3896.

22. Dixit, P. D., S. Merchant, and D. Asthagiri, 2009. Ion selectivity in the KcsA potassium channel from the perspective of the ion binding site. *Biophys. J.* 96:2138–2145.

23. Yu, H., S. Noskov, and B. Roux, 2009. Hydration number, topological control, and ion selectivity. *J. Phys. Chem.* 113:8725–8730.

24. Fowler, P., K. Tai, and M. Sansom, 2008. The selectivity of K^+ ion channels: Testing the hypotheses. *Biophys. J.* 95:5062–5072.

25. Dudev, T., and C. Lim, 2009. Determinants of K^+ vs Na^+ selectivity in potassium channels. *JACS* 131:8092–8101.

26. Roux, B., 2010. Exploring the Ion Selectivity Properties of a Large Number of Simplified Binding Site Models. *Biophys. J.* 98:2877–2885.

27. Thompson, A. N., I. Kim, T. D. Panosian, T. M. Iverson, T. W. Allen, and C. M. Nimigean, 2009. Mechanism of potassium-channel selectivity revealed by Na^+ and Li^+ binding sites within the KcsA pore. *Nat. Struct. Mol. Biol.* 16:1317–1324.

28. Zwanzig, R. W., 1954. High temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* 22:1420–1426.

29. Kasterholz, M. A., and P. H. Hünenbuger, 2006. Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation. *J. Chem. Phys.* 124:224501.

30. Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon, 1997. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* 72:1047–1069.

31. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, 2005. Scalable molecular dynamics with NAMD. *J. Comp. Chem.* 26:1781–1802.

32. MacKerell, Jr. A. D. et al, 1998. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* 102:3586–3616.

33. Celik, L., B. Schiøtt, and E. Tajkhorshid, 2008. Substrate binding and formation of an occluded state in the leucine transporter. *Biophys. J.* 94:1600–1612.

34. Vijayan, R., A. J. R. Plested, M. L. Mayer, and P. C. Biggin, 2009. Selectivity and cooperativity of modulatory ions in a neurotransmitter recepter. *Biophys. J.* 96:1751–1760.

35. Noskov, S., and B. Roux, 2007. Importance of hydration and dynamics on the selectivity on the KcsA and NaK channels. *J. Gen. Physiol.* 129:135–143.

36. Noskov, S., and B. Roux, 2008. Control of ion selectivity in LeuT: two Na^+ binding sites with two different mechanisms. *J. Mol. Biol* 377:804–818.

37. Alam, A., and Y. Jiang, 2008. Structural analysis of ion selectivity in the NaK channel. *Nat. Struct. Mol. Biol.* 16:35–41.

38. Jorgensen, W., and J. Tirado-Rives, 1988. The OPLS potential function for protein, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.

39. Jensen, K., and W. Jorgensen, 2006. Halide, ammonium, and alkali metal ion parameters for modeling aqueous solutions. *J. Chem. Theory Comput.* 2:1499–1509.

40. Cahalan, M., and T. Begenisich, 1976. Sodium channel selectivity. Dependence on internal permeant ion concentration. *J. Gen. Physiol.* 68:111–125.

41. Yamashita, A., S. Singh, T. Kawate, Y. Jin, and E. Gouaux, 2005. Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* 437:215–223.

42. Boudker, O., R. Ryan, D. Yernool, K. Shimamoto, and E. Gouaux, 2007. Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature* 445:387–393.

43. Hud, N., F. Smith, F. Anet, and J. Feigon, 1996. The selectivity for K^+ versus Na^+ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by 1H NMR. *Biochemistry* 35:15383–15390.

44. Andersson, C., and S. Mowbray, 2002. Activation of ribokinase by monovalent cations. *J. Mol. Biol.* 315:409–419.

45. Grell, E., T. Funck, and F. Eggers, 1975. Structure and dynamic properties of ion-specific antibiotics. *Membranes* 3:1–126.

46. Rose, M., and R. Henkens, 1974. Stability of sodium and potassium complexes of valinomycin. *BBA* 372:426–435.

47. Izatt, R. M., J. S. Bradshaw, S. A. Nielson, J. D. Lamb, and J. J. Christensen, 1985. Thermodynamic and Kinetic Data for Cation-Macrocycle Interaction. *Chem. Rev.* 85:271–339.

48. Shi, N., S. Ye, A. Alam, L. Chen, and Y. Jiang, 2006. Atomic structure of a Na^+- and K^+-conducting channel. *Nature* 440:570–574.

Structure	Carbonyl	Hydroxyl	Carboxyl	Water	Ether
K^+ Channel Model	8	0	0	0	0
Na^+ Channel Model	1	0	4	0	0
LeuT Na1	4	1	1	0	0
LeuT Na2	3	2	0	0	0
Glt_{Ph} Na1	4	0	2	0	0
Glt_{Ph} Na2	5	0	0	0	0
DNA Quadruplex	8	0	0	0	0
ARK	4	0	1	0	0
Valinomycin	6 (5)	0 (1)	0	0	0
18-crown-6	0	0	0	2 (2)	6 (4)
NaK	6 (4)	0	0	2	0
Nonactin	3	3 (2)	0	0	1

Table 1: The chemical nature of the ligands surrounding the ion within each biding site. Entries with two numbers indicate number of ligands with K^+ (unbracketed) and Na^+ (bracketed).

Structure	ΔG predicted	ΔG simulated	ΔG exp.	ΔG predicted $\epsilon > 1$	% χ_n	% χ_q	% χ_r
1. K^+ Channel Model	6.6	5.6 (13)	5-6 (6–9)	5.5	60	40	0
2. Na^+ Channel Model	-2.6	-5.3	$\sim$ -3 (40)	-3.5	-30	-70	0
3. LeuT Na1	-1.8	-6.1 (36)	$<$ -5 (41)	-3.2	-20	-10	-70
4. LeuT Na2	-0.7	-3.2 (36)		-2.7	-10	-30	-60
5. Glt_{Ph} Na1	-2.2	-1.3	$<$ -3 (42)	-3.6	-20	-5	-75
6. Glt_{Ph} Na2	1.0	-0.6		-0.3	-10	45	-45
7. DNA Quadruplex	6.7	4.0	1.7 (43)	5.5	60	40	0
8. ARK	0.8	1.3	$>$ 0 (44)	-0.4	-45	-25	30
9. Valinomycin	3.0	7.2	5-7 (45, 46)	1.1	-35	65	0*
10. 18-crown-6	3.1	2.6	0-2 (47)	3.1	0	100	0
11. NaK Channel	$\sim$ 0	-1.0 - 1.0 (35)	$\sim$ 0 (48)	-	-	-	-
12. Nonactin	0.2	0.7	0.7 - 1.0 (47)	-	-	-	-

Table 2: Predicted, simulated and experimental K^+/Na^+ ion selectivities for various structures. Predicted selectivities determined with the binding site surrounded by vacuum and by bulk solvent ($\epsilon > 1$) are shown. A percentage breakdown of contributions toward selectivity from coordination number restriction, χ_n; ligand dipole moment, χ_q; and cavity size/ligand thermal motion, χ_r is included. A positive value indicates a contribution toward K^+ preference while a negative value indicates a contribution toward Na^+ preference. Dashes indicate that the coordination number is different for Na^+ and K^+, making an estimate of contribution to selectivity difficult. *Non-local effects that are important in valinomycin are not included in this number in contrast to other studies, as discussed in the text.

Figure Legends

Figure 1.

Families of model systems used in this study. The first family is used to examine the influence of the ligand dipole moment (characterised by the charge on the ligands q) on ion selectivity while families 2 and 3 introduce the effect of the ion coordination number (n) and harmonic restrictions on the ligand position and thermal motion (characterized by r_{K^+}, r_{N^+} and k) respectively. Ideally, a consistent non-overlapping set of systems would be employed (left). However, due to limitations in computational power the simplified systems on the right were used. Squares represent periodic boxes of model ligands, large circles spherical finite droplets of model ligands and small circles constraints on the number of ligands that can coordinate the ion.

Figure 2.

Influence of ligand dipole moment on ion selectivity. Negative values indicate that an ion M^+ is more likely to leave water and enter the alternate solvent than K^+.

Figure 3.

Influence of coordination number and ligand charge on ion selectivity. (A) Selectivity between K^+ and Na^+ is shown with 2 kcal/mol contour differences. Regions that yield K^+ selectivity are indicated in red, Na^+ selectivity in blue and regions where ions will not leave water to enter the site in black. The numbers in circles correspond to the molecules in table 2: ① K^+ channel model ② Na^+ channel model ⑦ DNA quadruplex ⑨ Valinomycin. (B) Selectivity map for multiple group I ions with regions shown that are selective for Li^+ (green), Na^+ (blue), K^+ (red), Rb^+ (cyan) and Cs^+ (yellow). The percentage contribution to selectivity between K^+ and Na^+ by (C) the restriction of coordination number, χ_n and (D) the partial charge on the coordinating ligands, χ_q. Colour indicates whether this contribution is toward K^+ (reds) or Na^+ (blues) selectivity. (E) The effect on Na^+/K^+ selectivity when two of the n ligands are water molcules. (F) The effect on Na^+/K^+ selectivity when the coordination number of Na^+ can be different to that of K^+ with $q = 0.5$.

Figure 4.

Influence of cavity size and ligand thermal fluctuations on ion selectivity. (A) Selectivity between K^+ and Na^+ is shown with 1 kcal/mol contour levels with $k = 3.0$ kcal/mol/Å^2. (B) The effect of decreasing the ligand RMSD on selectivity for 5 different systems with different values of partial charge and ligand number. Cases with 5 or 6 ligands correspond to binding sites in the transporters, while that with 8 to KcsA. Arrows indicate the RMSD at specific k values for the $(n, q) = (6, 0.5)$ system.

Model family	Ideal System	Utilized System
1	 **Periodic box** Controlling: • ligand dipole moment	 **Periodic box** Controlling: • ligand dipole moment
2	 **Periodic box** Controlling: • ligand dipole moment • ion coordination number	 **a. Periodic box** Controlling: • ligand dipole moment • ion coordination number
		 **b. Large solvent sphere** Controlling: • ligand dipole moment • ion coordination number
		 **c. Vacuum** Controlling: • ligand dipole moment • ion coordination number
3	 **Periodic box** Controlling: • ligand dipole moment • ion coordination number • ligand positions and fluctuations	 **Vacuum** Controlling: • ligand dipole moment • ion coordination number • ligand positions and fluctuations

Figure 1:

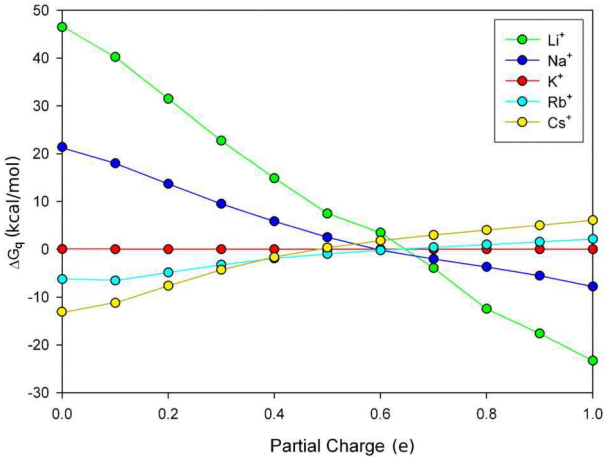
50
40
30
20
10
0
-10
-20
-30
0.0
0.2
0.4
0.6
0.8
1.0
ΔG_q (kcal/mol)
Partial Charge (e)
Li+
Na+
K+
Rb+
Cs+

Figure 2:

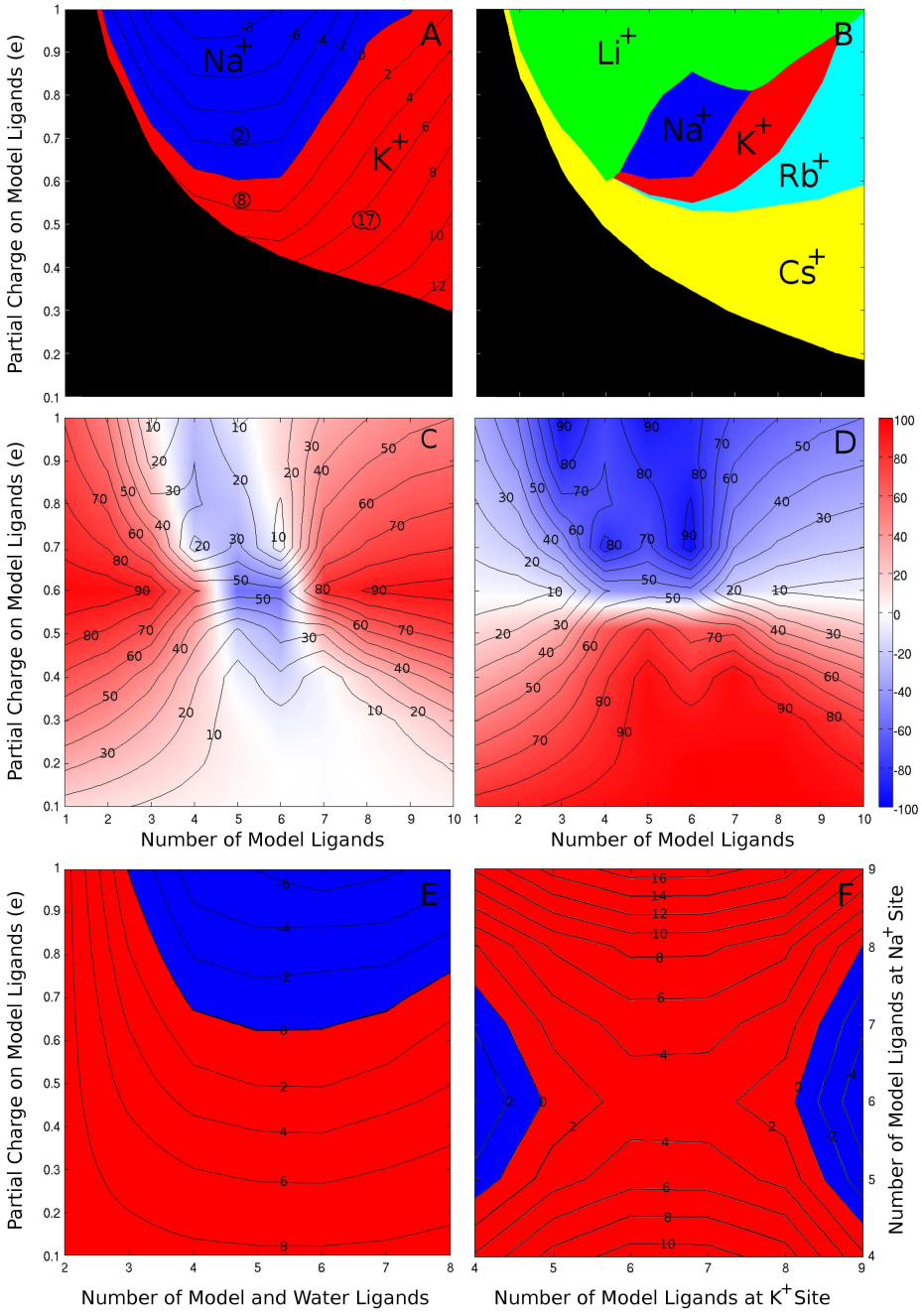
A
Na+
K+
B
Li+
Na+
K+
Rb+
Cs+
C
D
E
F
Partial Charge on Model Ligands (e)
Number of Model Ligands
Number of Model and Water Ligands
Number of Model Ligands at K+ Site
Number of Model Ligands at Na+ Site

Figure 3:

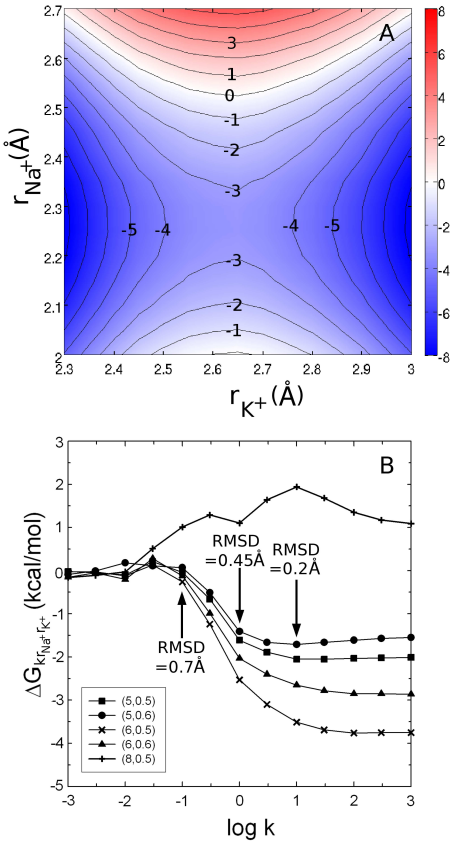
A
B
r_{Na^+}(Å)
r_{K^+}(Å)
$\Delta G_{k_{Na^+}r_{K^+}}$ (kcal/mol)
log k
RMSD =0.45Å
RMSD =0.2Å
RMSD =0.7Å
(5,0.5)
(5,0.6)
(6,0.5)
(6,0.6)
(8,0.5)

Figure 4:

Mapping the Importance of Four Factors in Creating Monovalent Ion Selectivity in Biological Molecules: Supplementary Material

Michael Thomas, Dylan Jayatilaka and Ben Corry

Detailed Simulation Parameters used for Model Systems

Periodic boundary conditions were employed for all bulk systems under an NPT ensemble (except for the bulk abstract ligand systems described earlier) at 1 atmosphere and 310K, (using langevin dynamics) and 1 fs timestep. A cut off distance of 12 Å is used for electrostatic and van der Waals interactions in non-periodic systems. In periodic systems, particle mesh Ewald summation was used to calculate electrostatic interactions unless stated otherwise.

λ windows for thermal fluctuation models FEPs

Although in hindsight it turned out to be unnecessary as the simulations converged well, these FEP simulations were conducted using 34 λ windows, with λ values of 0, then 10^n for $n = -7, -6 \ldots, -2$ then 0.05 then incrementing by 0.05 to $\lambda = 0.95$, then $1 - 10^n$ for $n =, -2, -3, \ldots, -7$ then 1.

Error estimation

Uncertainties in the free energies of the model systems were estimated by obtaining the standard error of 8 replicate FEP simulations (4 forward and 4 reverse morphs) for the situations in which $q = 0.5$, $n = 1, 5, 8$ and $k = 0$. The standard errors in the mean were all below 0.005 kcal/mol.

Verifying Simulation Results

Force Field. The results in this study may be dependent on the choice of force field as well as the environment surrounding the model systems. To study the influence of the choice of force field on the results, the OPLS force field (1) was used to reproduce the selectivity map in Fig. 3 A of the main text. The group I ion parameters (2) and the carbon-oxygen bonding and non-bonded parameters for the model ligands used OPLS, while angle and dihedral parameters for the model ligands were taken from the CHARMM parameter set, as they were not available in OPLS. The OPLS selectivity map, seen in Fig. 5 A, maintains the characteristic shape of the original figure. The difference in ΔG values between the two ranges is plotted in figure 5 C and is between 3 kcal/mol at the extremes, with the majority of useful energies (those that would not fall in the exclusion zone) being less than 2 kcal/mol. The difference in ΔG between CHARMM and OPLS found for the specific molecules studied in the main text is less than 1 kcal/mol.
Simulation Setup. The nature of the medium surrounding the model system may influence the free energy values. Three different model systems were simulated to quantify this effect. Each consists of n model ligands constrained so that the oxygen atoms are in a spherical shell of 3.5 Å (4.0 Å for Cs^+). Enveloping this was (i) vacuum, (ii) a spherical shell of model ligands, 10.5 Å in radius and (iii) a 30×30×30 Å periodic box of model ligands, which for simplicity will be referred to as the 'sphere', 'sphere in a sphere' and 'sphere in a box' models. The additional model ligands have the same partial charges on the carbon and oxygen atoms as those within the inner sphere. Systems (i) and (ii) were simulated using $n = 1 - 10, q = 0.1 - 1$ while (iii) only using (n,q)=(8,0.5) due to the large computational time involved. A comparison of the three systems using (n,q)=(8,0.5) is summarised in table 3 (a comparison for all (n,q) between systems (i) and (ii) is made graphically in Fig. 5 B and D). While there are significant differences in the total predicted selectivity of the model binding site shown in table 3, each is still a reasonable reflection of the experimental selectivity value of 5-6 kcal/mol seen in KcsA (3–6) that has a binding site of similar composition to the model. The percentage contributions to selectivity by coordination number restriction, χ_n, and dipole moment, χ_q differ at most by 7%, which is quite remarkable as it implies that the relative makeup of the total selectivity varies little regardless of the total selectivity.

The point of these additional simulations was to see how the effect of the

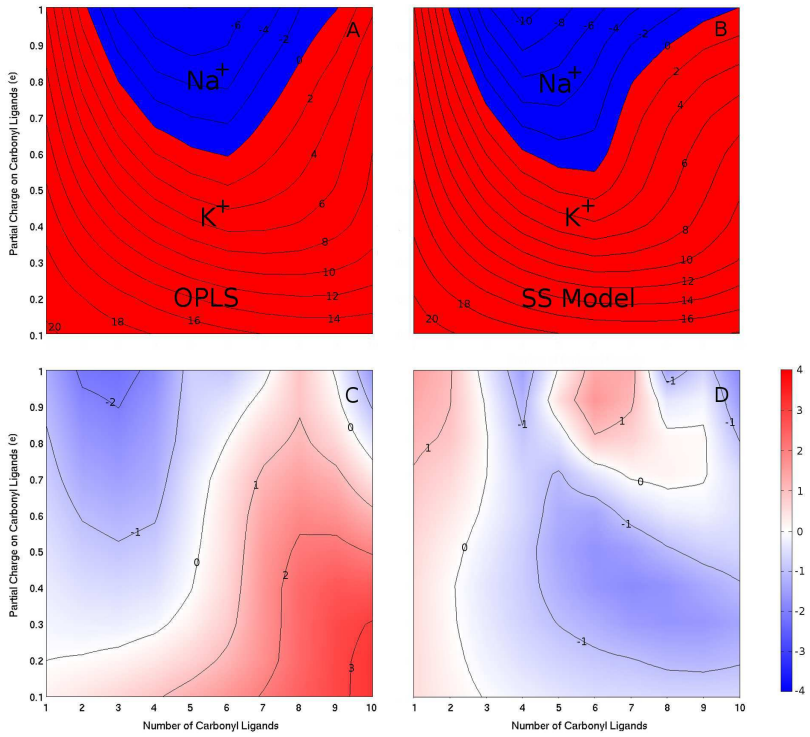

Figure 5: (A) Selectivity map using OPLS force field. (B) Selectivity map using 'sphere in a sphere' (SS) model system. (C) Difference between CHARMM and OPLS selectivity maps. (D) Difference between selectivity maps produced using sphere and SS model systems. All energies in kcal/mol. Red depicts K^+ selectivity, blue depicts Na^+ selectivity. Colours in (C) and (D) are used for clarity.

surrounding medium influences the selectivity of the model binding sites. As each binding site will have a different environment it is difficult to be sure which of our test calculations best mimics the specific site. These studies do, however, give an indication of the range of uncertainty in our predictions that is created by the details of the medium surrounding each site. The 10.5 Å sphere of bulk model ligands in our sphere in a sphere system allows for rapid energy calculations as required for constructing selectivity maps across a wide range of parameters. To test how similar the results of this case are

System	Force Field	ΔG_W	ΔG_{nq}	ΔG_n	ΔG_q	χ_n	χ_q
S	CHARMM	21.06	6.58	N/A	N/A	N/A	N/A
S	OPLS	21.68	5.07	N/A	N/A	N/A	N/A
SS	CHARMM	21.06	5.53	2.16	3.37	60.9%	39.1%
SS	OPLS	21.68	2.82	0.96	1.86	66.0%	34.0%
SB	CHARMM	21.06	8.56	3.02	5.54	64.7%	35.3%
SB	OPLS	21.68	4.50	1.86	2.64	58.7%	41.4%

Table 3: Summary of model systems and associated ΔG and χ_n and χ_q values for (n,q)=(8,0.5). S, SS and SB signify sphere, sphere in a sphere and sphere in a box model systems respectively. ΔG_W is the free energy of morphing Na^+ to K^+ in a 30×30×30 Å period box of TIP3P water. ΔG_{nq} is the total selectivity of the model system. Other ΔG values are described in the methods section of the main text. All energies are in kcal/mol

to those in a larger bulk medium we examined how ΔG_{nq} depends on the size of the bulk sphere for the (n,q)=(8.0.5) sphere in a sphere model system. As shown in table 4 the small sphere gives fairly similar results to those in a larger, more bulk like medium.

Sphere size	ΔG_{nq}
Vacuum	6.58
10.5 Å	5.53
15 Å	5.87
20 Å	5.85
25 Å	5.92
30 Å	5.93

Table 4: ΔG_{nq} (kcal/mol) for Na^+ and K^+ in a sphere in a sphere model system with (n,q)=(8.0.5) for models with differing radii of the outer model ligand containing sphere.

Correction Terms for ΔG_{entry}

ΔG_{entry} is the free energy of moving an ion from bulk water into a model site.

$$K^+/H_2O + model(n,q) \quad \xrightarrow{\Delta G_{entry}} \quad H_2O + K^+/model(n,q) \tag{1}$$

This is constructed from two separate binding free energies:

$$K^+/H_2O \quad \xrightarrow{\Delta G_{entry(A)}} \quad H_2O + K^+ \tag{2}$$

$$K^+ + model(n,q) \quad \xrightarrow{\Delta G_{entry(B)}} \quad K^+/model(n,q) \tag{3}$$

Normally the double-annihilation method is used to calculate the binding free energy (7). However, since K^+ cancels out when equations 2 and 3 are combined, we can reduce the calculation to two FEP calculations:

$$K^+/H_2O \quad \xrightarrow{\Delta G_{entry(A)}} \quad H_2O \tag{4}$$

$$model(n,q) \quad \xrightarrow{\Delta G_{entry(B)}} \quad K^+/model(n,q) \tag{5}$$

The corrections terms that are needed to yield the binding free energies by using the double annihilation method, as described by Gilson et al (8), cancel out.

Additional corrections must be applied when calculating free energies of introducing a charged species into an explicit solvent systems, which in this case is equation 4. These corrections terms are described in detail by Kastenholz and Hünenburger (9). Two of the corrections discussed are applicable to this situation; approximations in the electrostatic interactions introduced from using cut off truncation (type A correction), and artificial polarisation of the solvent molecules from using a periodic boundary condition (type B correction). The corresponding correction terms are $\Delta G_{\mathrm{A}}^{\mathrm{CT}}$ and $\Delta G_{\mathrm{B}}^{\mathrm{CT}}$ (where CT means cut off truncation method). $\Delta G_{\mathrm{A}}^{\mathrm{CT}}$ is further divided into two parts; a correction for electrostatic interactions outside the cut off sphere, $\Delta G_{\mathrm{A1}}^{\mathrm{CT}}$, and a inside the cutoff sphere, $\Delta G_{\mathrm{A2}}^{\mathrm{CT}}$, so that

$$\Delta G_{\mathrm{A}}^{\mathrm{CT}} = \Delta G_{\mathrm{A1}}^{\mathrm{CT}} + \Delta G_{\mathrm{A2}}^{\mathrm{CT}} \tag{6}$$

$\Delta G_{\mathrm{A1}}^{\mathrm{CT}}$ and $\Delta G_{\mathrm{B}}^{\mathrm{CT}}$ are trivial to calculate if $\Delta G_{\mathrm{A}}^{\mathrm{CT}}$ is known.

$$\Delta G_{\mathrm{A1}}^{\mathrm{CT}} = -(8\pi\epsilon_0)^{-1} q_I^2 (1 - \epsilon_S^{-1}) R_C^{-1} \tag{7}$$

$$\Delta G_{\mathrm{B}}^{\mathrm{CT}} = -\frac{R_I}{R_C} 10^{\mu[L/(2R_C)]+\nu} [(8\pi\epsilon_0)^{-1} q_I^2 (1-\epsilon_S^{-1}) R_I^{-1} + \Delta G_{\mathrm{A}}^{\mathrm{CT}}] \quad (8)$$

where ϵ_0 is the permittivity of vacuum, q_I the ionic charge, $\epsilon_S = 78$ the permittivity of the solvent, $R_C = 12$ Å the cut off radius, $R_I = 1.33$ Å the ionic radius, $L = 30$ Å is the edge length of the solvent box and $\mu = -2.20$ and $\nu = 1.29$ are constants. $\Delta G_{\mathrm{A2}}^{\mathrm{CT}}$ is then given by:

$$\Delta G_{\mathrm{A2}}^{\mathrm{CT}} = -(8\pi\epsilon_0)^{-1} q_I^2 (1-\epsilon_S^{-1}) \times [R_I^{-1} - R_C^{-1} + \sum_{k=-1}^{N_a} k a_k R_c^{-k-1} \int_{R_I}^{R_C} dr\ r^{k+1} p(r)] \quad (9)$$

We can define

$$C = R_C^{-1} + \sum_{k=-1}^{N_a} k a_k R_c^{-k-1} \int_{R_I}^{R_C} dr\ r^{k+1} p(r) \quad (10)$$

where C is a constant. We can now determine the value of C is for the case explored by Kastenholz and Hünenburger (9), Na^+ in a periodic water box. We take $\Delta G_{\mathrm{A2}}^{\mathrm{CT}} = 17.61$ kJ/mol from table III. Assuming that this system and our K^+ periodic water box system is similar enough for C to be equal in both cases, we can now solve $\Delta G_{\mathrm{A2}}^{\mathrm{CT}}$ for the K^+ system:

$$\Delta G_{\mathrm{A2}}^{\mathrm{CT}} = -(8\pi\epsilon_0)^{-1} q_I^2 (1-\epsilon_S^{-1}) \times [R_I^{-1} - C] = 17.61\,\mathrm{kcal/mol} \quad (11)$$

Now the binding free energy correction, ΔG_{corr}, is:

$$\Delta G_{\mathrm{corr}} = \Delta G_{\mathrm{A1}}^{\mathrm{CT}} + \Delta G_{\mathrm{A2}}^{\mathrm{CT}} + \Delta G_{\mathrm{B}}^{\mathrm{CT}} = -9.9\,\mathrm{kcal/mol} \quad (12)$$

Ion Parameters

The vdW parameters we have used for $Na^+$and $K^+$are those developed and used by Roux et al in a number of recent works (13–15) (originating from the work of Beglov and Roux (16)). The values used for Rb^+ and Cs^+ also come from Roux et al while only those for Li^+ are developed here. The results we have presented for Rb^+, Cs^+ and Li^+ (eg in Fig 3B) are intentionally only qualitative given that the parameters for these ions have been less thoroughly tested. The vdW parameters used in this study are given in table 5 along with ion-water distances and coordination numbers (table 6) and compared to experimental values. We note that experimentally determining coordination numbers is particularly difficult for Cs^+ given its structural lability in water (17).

Ion	ϵ	$r_{min}/2$	First ion-oxygen RDF peak (Å)	exp. ion-oxygen distance (Å)
Li^+	-0.0030	1.137	1.87 ± 0.02	1.96 (10)
Na^+	-0.0469	1.36375	2.30 ± 0.05	2.39 - 2.42 (11)
K^+	-0.0870	1.76375	2.70 ± 0.05	2.7 - 2.9 (12)
Rb^+	-0.150	1.90	2.90 ± 0.05	2.80 - 3.05 (12)
Cs^+	-0.1900	2.100	3.15 ± 0.05	2.95 - 3.21 (12)

Table 5: Lennard-Jones parameters of the ions used in this study. Included are the distances in the first peak of the RDF in an 30x30x30 Å TIP3 water box and the experimental value of the ion-ligand distance. Error in the simulated value signifies the uncertainty in identifying this peak. Where multiple experiments have been conducted, a range of values is given.

Ion	This Study	Experimental	Molecular Dynamics	Quantum Methods
Li^+	4.0	3-6 (18), 3.0-6.5 (19)	4 (18), 4.1 (20)	4 (18)
Na^+	5.7	4-8 (18), 4.9 (19)	5-7 (18), 5.9 (20)	4-6 (18), 5.5(21)
K^+	7.0	4-8 (18), 6-8 (12), 5.3 (19)	6-8 (18), 7.2 (20)	4-8 (18), 6.2 & 6.8 (21)
Rb^+	7.8	5.6 (22), 6.4-7.4 (23), 6-8 (12), 6.5 (25), 6.9 (19)	7.8 (20), 8.5 (24)	7.1 (24)
Cs^+	9.6	8.2 (26), 3-9 (12), 8.1 (27), 7.9 (28)	9.6 (20)	7.8-9.1 (17)

Table 6: Coordination numbers of the ions used in this study determined in a periodic water box with Cl^- counterion compared with published values derived from experiment, other molecular dynamic simulations and various quantum methods.

Heterogeneous Ligand Partial Charges

As noted in the article, the effects of having ligands with partial charges that deviate substantially from the average partial charge on all the ligands surrounding the ion is a source of disparity between predicted and simulated ΔG values. To investigate this issue we have constructed 8 additional maps involving combinations of two types of ligands. In these we have up to two TIP3 waters as ligands and/or up to two fully charged acidic residues (a

model ligand containing a +1.0 e charge on the carbon and a -1.0 e charge on the oxygen), while the remainder of the ligands making up the total coordination number n are ligands with controllable dipole moment as in the main text.

It is evident from Figs. 6 and 7 that as the number of ligands with unitary charge in a system increases (ie inclusion of acidic residues) the more selective the system becomes for the the smaller ions. Indeed the presence of 2 fully charged ligands makes it essentially impossible to select for K^+ via restricting coordination numbers, indicating the dominance of dipole moment effects on selectivity under such conditions. In contrast, an increase in the number of water molecules in the system dissipates selectivity; it reduces the degree of selectivity but does not appear to change which ion is selected given the nature of the remaining ligands.

These results are particularly interesting given recent discussion of the differences in selectivity between the KcsA channel and the non-selective NaK channel that have similar, but not identical selectivity filters. It has been suggested that the structure of the NaK channel allows for more water to contact ions in the binding sites which act to reduce the selectivity of these sites due to the differing dipole moment of water and carbonyl ligands (14). Our data is certainly consistent with this hypothesis. Even if the number of ligands coordinating an ion in an NaK binding site remains fixed at 8 or more as seen in KcsA, then Fig. 7 indicates that the substitution of two carbonyl groups with two water molecules reduces selectivity by more than 2 kcal/mol. Fowler et al have noted, however, that the coordination number of ions in NaK may be less constrained than in KcsA (29). This observation would provide an alternative route to reduce selectivity in NaK compared to KcsA. It is likely that both the presence of additional water around the ion and the greater freedom in determining coordination numbers in NaK compared to KcsA combine to reduce K^+ selectivity.

Sites where coordination numbers can change

All the model systems discussed so far have assumed that the coordination number remains constant as one ion type is exchanged with another. While this is reasonable in many cases, there are some in which the coordination number changes with ion type. If the coordination number changes between 'optimal' values for each ion type, i.e. there is no restriction on coordination

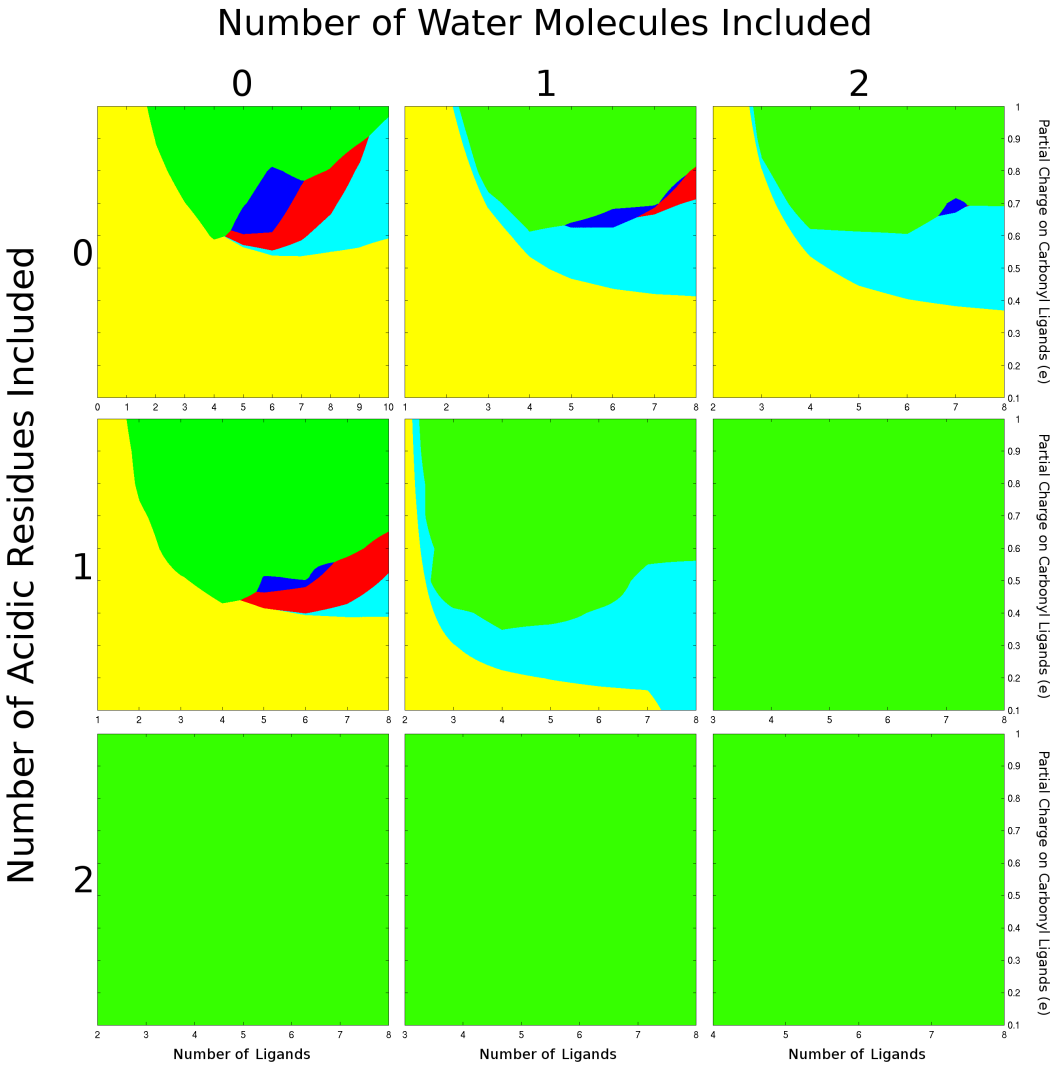

Figure 6: Selectivity maps of (n,q) model systems with water and/or acidic residues participating in ion coordination for group I ions with regions shown that are selective for Li^+ (green), Na^+ (blue), K^+ (red), Rb^+ (cyan) and Cs^+ (yellow). Exclusion zones have not been calculated for these maps.

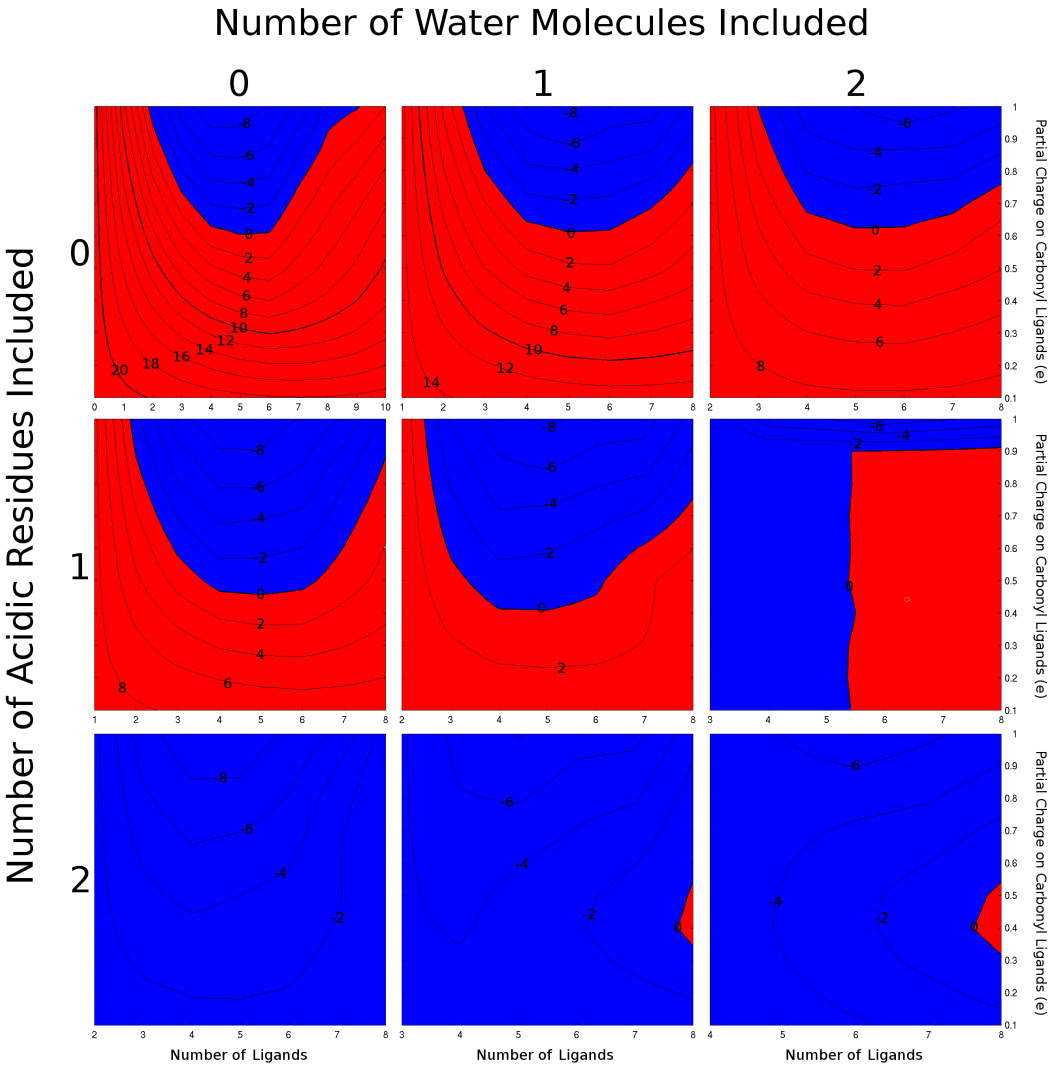

Figure 7: Selectivity maps of (n,q) model systems with water and/or acidic residues participating in ion coordination for K^+ and Na^+ shown with 2 kcal/mol contour differences in ΔG_{nq}. Regions that yield K^+ selectivity are indicated in red, Na^+ selectivity in blue. Exclusion zones have not been calculated for these maps.

number, then coordination number restriction cannot lead to selectivity. In this case selectivity can be accurately predicted using just a study of dipole moment contributions such as in Fig. 2 of the main text. However, it is also possible that the coordination number can alter with ion type, but not be able to obtain the ideal values, in which case coordination number restriction can still create ion selectivity without demanding that the number of coordinating ligands remain constant. These cases can be studied by examining the selectivity of a site that has n fold coordination for one ion type and m fold coordination for another. To create this, an ion exchange equation can be constructed using our model system in equation 2 in the main text, along with additional MD FEP simulations where a model ligand is removed gradually from the model system and added to bulk model ligand;

$$\begin{aligned} Na^+/H_2O + K^+/model(n,q) + CH_3O_{Bulk} \quad &\xrightarrow{\Delta G_{nmq}} \\ K^+/H_2O + Na^+/model(m,q) + CH_3O_{Bulk+(n-m)} \end{aligned} \tag{13}$$

The additional MD FEP simulations are:

$$K^+/model(n,q) \quad \xrightarrow{\Delta G} \quad K^+/model(m,q) \tag{14}$$

and

$$CH_3O_{Bulk} \quad \xrightarrow{\Delta G} \quad CH_3O_{Bulk-1} \tag{15}$$

Fig. 8 shows the results for four different partial charges, allowing selectivity to be predicted in cases where the coordination number changes with ion type. Bostick et al (30) produced graphs using the same reaction but using water instead of model ligands that look very similar to our figure with $q = 0.5$. Our graphs extend their analysis to cover a range of different ligand types.

Restrained cavity size selectivity maps

Six systems were investigated with $r_{Na^+} = r_{K^+}$ and variable RMSD, with different ligand number and partial charge (n, q)= $(5, 0.5)$, $(5, 0.6)$, $(6, 0.5)$, $(6, 0.6)$, $(6, 0.7)$ and $(8, 0.5)$ (Fig. 9). In all six situations, it is observed that cavity size effects play an important role in ion selectivity when the ligands are held firmly in place. At low RMSD (> 0.5 Å) there is an increasingly

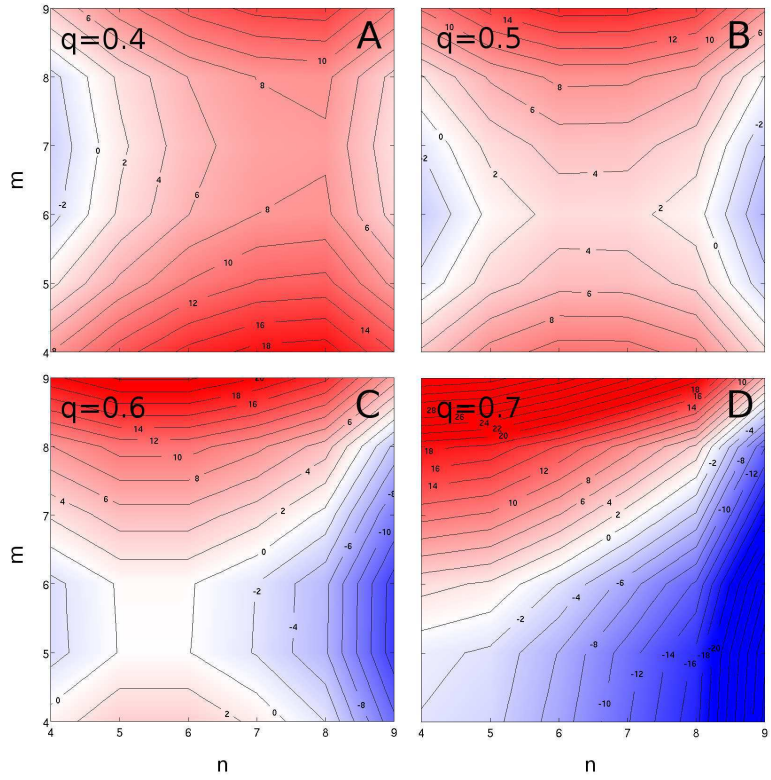

Figure 8: Free energies of selectivity between n ligands coordinating to K^+ and m ligands coordinating to Na^+ for partial charges of (A) 0.4, (B) 0.5, (C) 0.6 and (D) 0.7, according to equation 13. Regions of K^+ selectivity are shown in red, Na^+ in blue. Contours are spaced at 2 kcal/mol intervals.

significant contribution to Na^+ selectivity when $r < 2.5$ Å and to K^+ when $r > 2.5$ Å. However, when the flexibility of the ligands is larger such that they can move to accommodate differently sized ions (ie RMSD > 0.5 Å) the influence of cavity size diminishes. The total selectivity plateaus with increasing RMSD toward the corresponding ΔG value on the selectivity maps shown in Fig. 2B of the main text. As noted in the text we found that restrictions on the thermal fluctuations of coordinating ligands rather than restrictions on the cavity size itself appeared more important in the molecules

studied.

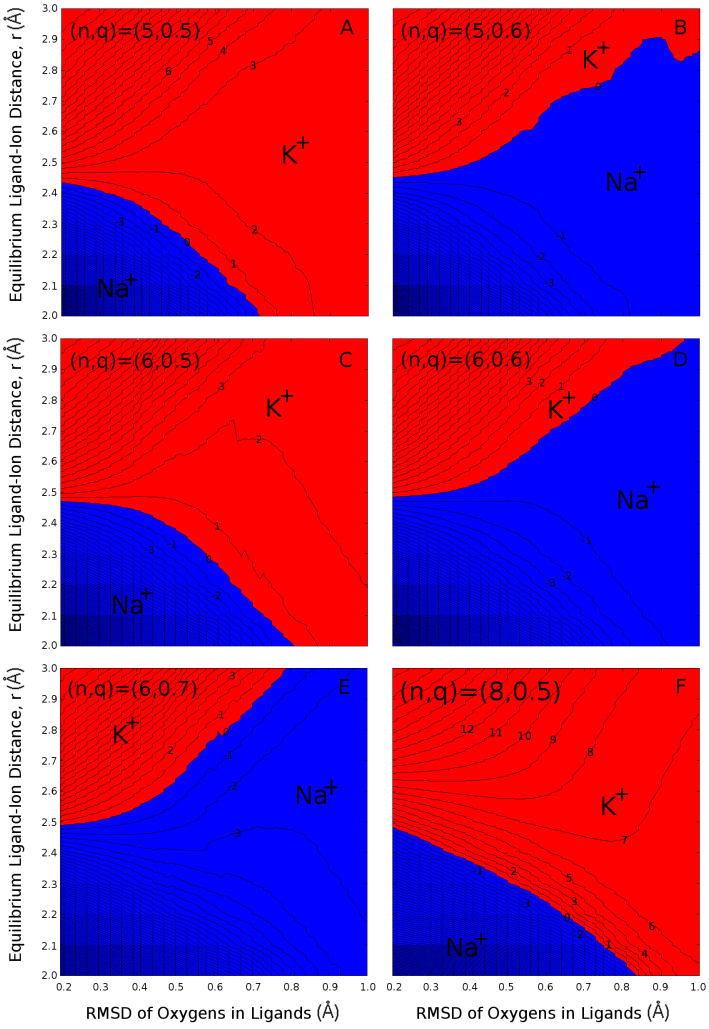

Figure 9: Influence of cavity size (as measured by the equilibrium ion-ligand distance r and RMSD) on ion selectivity for (n, q) systems (A)$(5, 0.5)$, (B)$(5, 0.6)$, (C)$(6, 0.5)$, (D)$(6, 0.6)$, (E)$(6, 0.7)$ and (F) $(8, 0.5)$. Red indicates preference for K^+ while blue indicates preference for Na^+. The interface between the two colours is 0 kcal/mol with each contour line representing 1 kcal/mol.

Additional thermal fluctuation maps

See Fig. 10.

Discussion of specific ion binding sites

As noted in the main text, we validated our predictive scheme by comparison with detailed MD simulations of a number of ion selective molecules. For simplicity, we concerned ourselves only with molecules showing selectivity between Na^+ and K^+. These included three ion channels KcsA (simulations described previously (31)), NaK and a simple model of a sodium channel selectivity filter; two amino acid transporters LeuT and Glt_{Ph}; a DNA quadruplex; the enzyme aminoimidazole riboside kinase; valinomycin, nonactin and 18-crown-6. Details of the parameters used in each simulation are given in table 7 and in the subsequent text. A more detailed discussion of selectivity within some of these binding sites then follows.

Structure	λ step	Total sim. time (ns)
Bulk water	0.05	40.0(Na^+), 8.0(Li^+, Rb^+, Cs^+)
Model systems	0.05	40.0 (S,SS), 8.0(SB)
$k \neq 0$ model systems	variable	136
Na^+ channel Model	0.05	40.0
Glt_{Ph}	0.025 & 0.05	9.6
DNA quadruplex	0.05	4.0
ARK	0.1	2.0
Valinomycin	0.05	4.0
18-crown-6	0.05	4.0

Table 7: Parameters used in alchemical free energy perturbation simulations of model systems and biological molecules. Simulations include solvation boxes, lipid bilayers etc. S, SS and SB indicate sphere, sphere in a sphere and sphere in a box model ligand systems respectively.

The Na^+ channel selectivity filter model utilised the DEKA (aspartic acid, glutamic acid, lysine, alanine) motif in order to produce a near planar

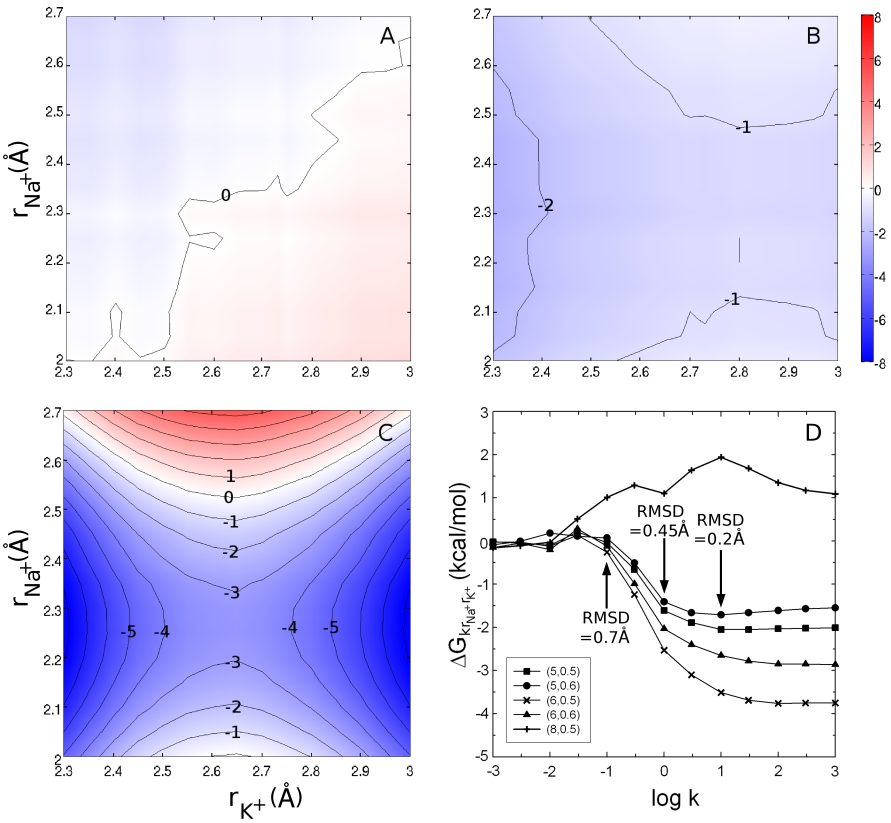

Figure 10: Influence of cavity size and ligand thermal fluctuations on ion selectivity with (A) $k = 0.03$ kcal/mol/Å^2, (B) $k = 0.3$ kcal/mol/Å^2 and (C) $k = 3.0$ kcal/mol/Å^2. Contour lines represent 1 kcal/mol differences in ΔG. (D) The effect of decreasing the ligand RMSD on selectivity for 5 different systems with different values of partial charge and ligand number. Cases with 5 or 6 ligands correspond to binding sites in the transporters, while that with 8 to KcsA. Arrows indicate the RMSD at specific k values for the $(n, q) = (6, 0.5)$ system.

Structure	$\Delta G_{K^+ \rightarrow Na^+}$	$\Delta G_{Na^+ \rightarrow K^+}$
Bulk water	-21.05	21.07
Na^+ channel Model	-25.8	27.0
Glt_{Ph} Na1	-22.2	22.4
Glt_{Ph} Na2	-20.3	23.1
DNA quadruplex	-15.7	18.4
ARK	-19.9	19.6
Valinomycin	-13.8	13.8
Nonactin	-20.9	21.0
18-crown-6	-18.3	18.5

Table 8: Forward and reverse ΔG values for each of the simulated systems. All values are in kcal/mol.

structure of these four amino acids. An approximately 4.4×4.4 Å rectangular gap was left to simulate the pore (32), the boundaries of which were determined by the inward facing side chain groups. Harmonic constraints of 2.2 kcal/mol/Å^2 was placed on the backbone carbon and nitrogen atoms to yield realistic RMSD fluctuations of the atoms. Each amino acid was N and C terminated.

The aspartate transporter (Glt_{Ph}), PDB code 2NWX (33), was placed in a POPC lipid bilayer which was solvated on both sides by an approximately 25 Å deep 120 mM NaCl solution. The transported aspartate was in zwitterionic form (34). The morphing Na^+ and K^+ were held fixed during the simulations. The system was equilibrated for 2 ns prior to FEP calculations. 24 λ windows of 400 ps each were simulated with two window sizes; 0.025 when $0 < \lambda < 0.1, 0.9 < \lambda < 1.0$ and 0.05 elsewhere. The system utilised periodic boundary conditions. There were 144255 atoms in total. The forward and reverse morph differ by 2.8 kcal/mol. As a consequence, the $Na^+ \rightarrow K^+$ morph was simulated an additional two times, and the average of these three simulations was taken.

Valinomycin, Cambridge structural database (CSD) code VALINK (35), was placed in a 30×30×30 Å periodic ethanol box and equilibrated for 4 nanoseconds. The Na^+ and K^+ ion was held in a fixed position at the centre of the box around which the valinomycin was placed. There were 9170 atoms in total.

Nonactin, CSD code NONACS (36), was placed in a 45×45×45 Å periodic

ethanol box, and the Na^+ and K^+ ions were held fixed in the centre as above. There were 9119 atoms in total.

18-crown-6, CSD code KTHOXD (37), was placed in a 30×30×24 Å periodic water box with a Na^+ and K^+ atom fixed at the centre of the crown ether.

The G-quadruplex DNA structure, PDB code 1L1H (38), was solvated in a 55 × 55 × 55 Å box of 300mM KCl solution and equilibrated for 15 ns prior to commencement of FEP calculations. Periodic boundary conditions were used. There were 16987 atoms in total.

The enzyme aminoimidazole riboside kinase (ARK), PDB code 1TYY (39), was solvated in an 80×70×117 Å box of 120mM NaCl solution. Periodic boundary conditions were used. There were 59629 atoms in total.

All simulations were conducted using NAMD (40) with the CHARMM27 (41) force field, unless otherwise noted, at 310 K and 1 atm with 1 fs timesteps.

Valinomycin

Valinomycin is an interesting case. In a previous study, it was demonstrated that the predominant contributor toward K^+ selectivity comes from cavity size effects (42). In our initial studies conducted in vacuum, the Na^+ ion bounces around in the valinomycin cavity, unable to find a favourable Na^+-ligand distance simultaneously for all ligands, whilst K^+ is able to do so giving it a 12.1 kcal/mol selectivity over Na^+ , which is in good agreement with density functional calculations (12.3 kcal/mol) (42). The rigidity of the valinomycin is enforced by intermolecular hydrogen bonding, which maintains its ring like structure (Fig. 11 A & B). In vacuum, therefore, the cavity size of valinomycin appears to give the largest contribution to its K^+ selectivity.

Once valinomycin is placed in a solvent, however, additional factors come into play. While the K^+/valinomycin complex in ethanol behaves similarly to its vacuum counterpart, when Na^+ is present inside the cavity, the valinomycin distorts significantly from a ring into a oval shape (Fig. 11 C). Although the coordination number of 6 is maintained by Na^+ between the two environments, it is not bound to the same ligands. As a result of attempting to optimise the coordination distance, Na^+ coordinates to only 4 of the original 6 oxygens where one of the additional ligands is sourced from a previously hydrogen bonded oxygen and another from an ethanol oxygen

(Fig. 11 D). The breaking of this hydrogen bond allows the valinomycin to distort. The remaining two oxygens that participated in K^+ coordination hydrogen bond to ethanol when Na^+ is bound. The ability of Na^+ to gain a more optimal coordination number reduces the selectivity of the molecule, leading to a K^+ selectivity of 7.2 kcal/mol. This also indicates that coordination number restriction cannot underlie this selectivity. As the cavity size has altered to favourably coordinate both K^+ and Na^+, our analysis attributes the selectivity to the intrinsic dipole moment of the coordinating ligands. The change in selectivity between gas and ethanol, however, is not as large as may be suspected from our selectivity maps due to the added energy cost of breaking a hydrogen bond when coordinating Na^+. This energy cost represents a deformation energy intrinsic to this particular molecule that cannot be captured in our model systems. While this energy would be small in large proteins where the coordinating ligands are not directly attached to one another, it may be more significant in small cyclic molecules. This deformation energy could be considered part of the cavity size effect as it represents the cost incurred in adapting the cavity to the size of each ion, but is not included in our definition. Comparing our predicted and simulated selectivities in valinomycin suggests that the deformation energy may be as great as 4.2 kcal/mol. This equates to 60% of the total selectivity coming from cavity size molecular deformation, leaving 25% and 15% contribution from coordination number restriction, χ_n, and dipole moments, χ_q, respectively. Using both methods, we find the size and importance of this deformation energy is consistent with the findings of Varma and Rempe (42).

18-Crown-6

In 18-crown-6, shown in Fig. 12, the K^+ ion coordinates to the six oxygen molecules in the crown as well as coordinating to two water molecules axially. The Na^+ ion coordinates to only four of the six crown oxygens, plus to two axial water molecules. Thus, both can optimise their coordination numbers. The flexibility of the cavity size means that the overall selectivity of the molecule is attributed to the intrinsic dipole moment of the ligands. But, as with valinomycin, in the presence of Na^+ the crown ether deforms so that the two uncoordinated crown oxygens may coordinate with the bulk water. This again creates a small energy cost not captured in our maps.

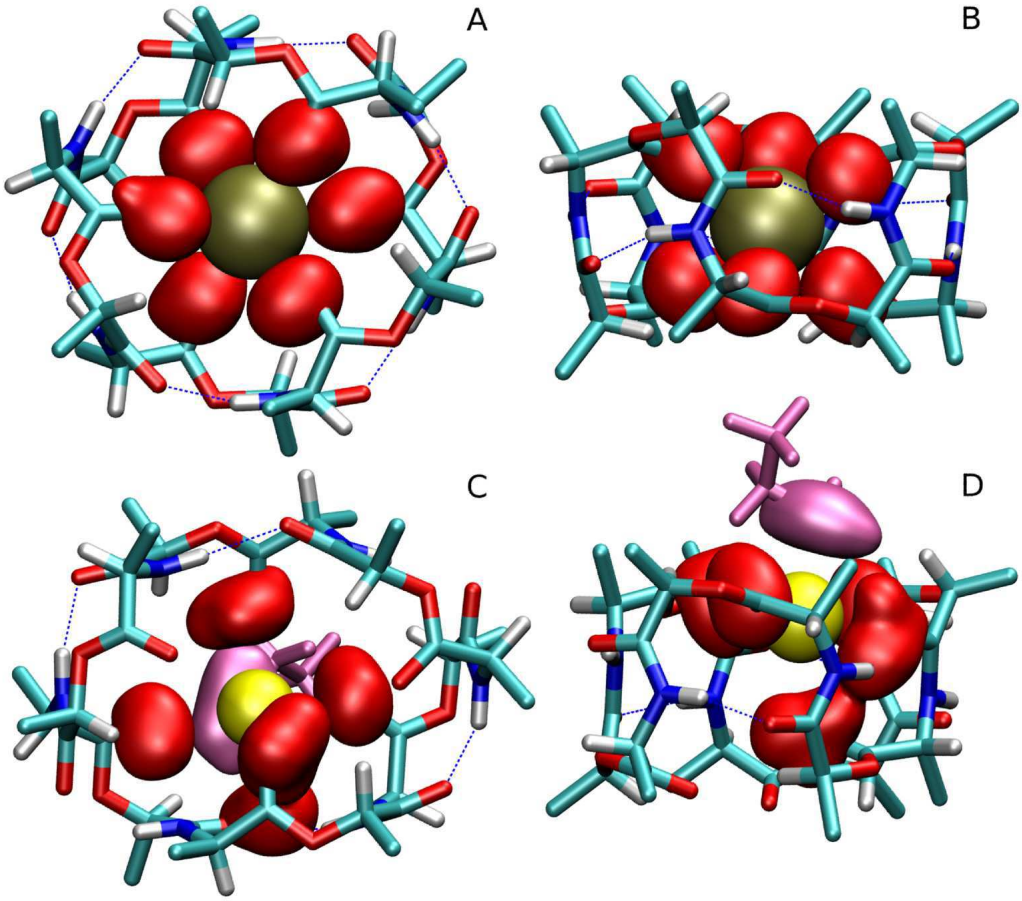

Figure 11: Structures of K^+ and Na^+ complexed with valinomycin in an ethanol box. 'Side chain' atoms in valinomycin have been removed for clarity. K^+ is brown, Na^+ yellow, oxygen red, carbon cyan, nitrogen blue, hydrogen white and ethanol is magenta. The red and magenta surfaces show the regions of space visited by the mobile coordinating oxygens. Hydrogen bonding is denoted by dashed blue lines. (A) Top view of K^+/valinomycin complex. (B) Side view of K^+/valinomycin complex. (C) Top view of Na^+/valinomycin complex showing the distorted shape of the molecule and the participation of ethanol in ion coordination. (D) Side view of Na^+/valinomycin demonstrating the disruption of the hydrogen bonding within valinomycin and the shift of Na^+ to one edge of one of the valinomycin ring.

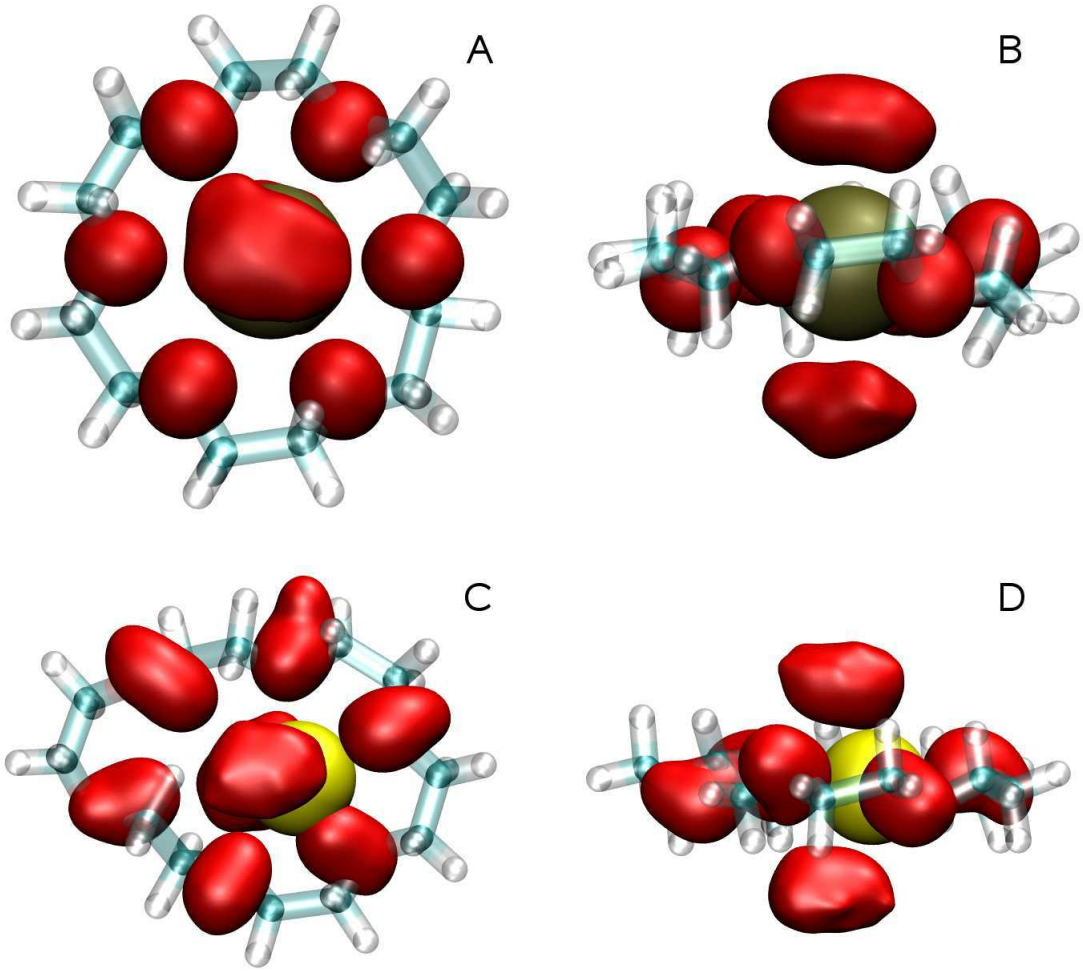

Figure 12: Structures of 18-crown-6 complexed with K^+ or Na^+ in water. Solid red surfaces indicate the regions of space visited by the coordinating oxygen atoms of the crown or water molecules. K^+ is brown and Na^+ yellow. Transparent atoms represent carbon (cyan), hydrogen (white) and oxygen (red). (A) Top view and (B) side view of K^+/18c6 complex. (C) Top view and (D) side view of Na^+/18c6 complex. Note the deformation of the crown structure in the presence of Na^+ compared with that obtained with K^+ . Not all all oxygens coordinate Na^+ simultaneously, as evidenced by the offset position of the Na^+ and the elongated region of space occupied by the oxygen atoms as they approach and recede from the ion.

ARK, LeuT and Glt_{Ph}

Contributions to ion selectivity from restricted ligand mobility were most notable in the two amino acid transporters in which the ligands appear constrained in their movement. The RMSD of the coordinating ligands in Glt_{Ph} with Na^+ bound was 0.30 and 0.30 Å for site 1 and 2 and with K^+ was 0.52 and 0.35 Å. The same RMSD in LeuT was 0.39 and 0.43 Å when Na^+ was bound to site 1 and site 2 (43), while the K^+ bound values are not known. These low RMSD values may be a consequence of the coordinating ligands being part of an α helix and thus held in place through networks of hydrogen bonding.

In contrast to these situations, the ligands forming the binding site of the enzyme ARK are primarily associated with protein loops as shown in Fig. 13. These tend to be much more flexible than α helices, so we would expect that this could produce a more flexible environment in which the binding cavity size and restricted ligand mobility would have a smaller impact on the overall selectivity. This is exactly what we see. The average ion-ligand distance in ARK changes from 2.3 Å for Na^+ to 2.7 Å for K^+ indicating that restrictions on cavity size do not contribute to ion selectivity in this case, and the RMSD of the coordinating atoms is 0.55-0.6 Å.

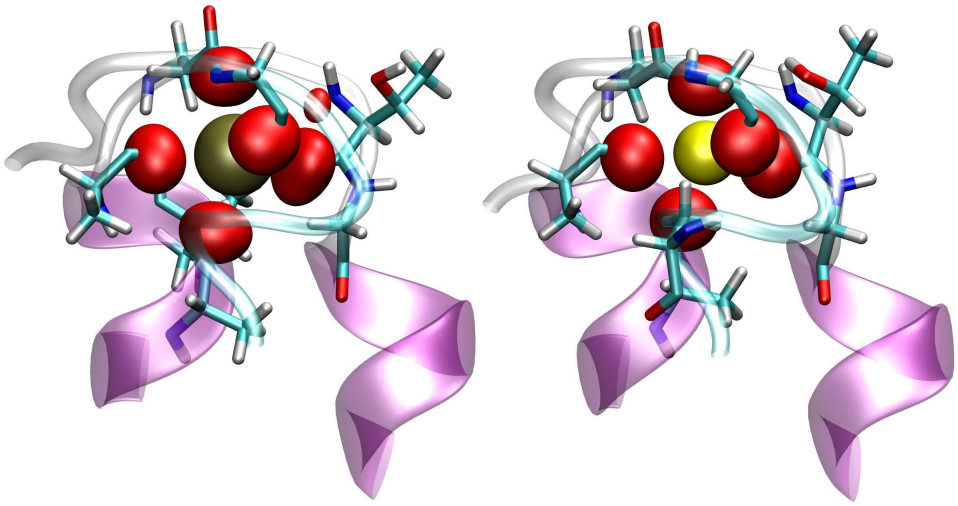

Figure 13: Structures of the ion binding site in ARK in complex with (A) K^+ and (B) Na^+ . Residues directly participating in ion coordination are shown in an all atom representation while the surrounding protein is shown by the cartoon. The regions of space visited by the coordinating oxygen atoms are shown by red surfaces. Note that one oxygen atom is hidden behind the ion in each picture. The small red surface near in the top right in (A) represents the occasional presence of a water molecule coordinating K^+ .

References

1. Jorgensen, W., and J. Tirado-Rives, 1988. The OPLS potential function for protein, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.

2. Jensen, K., and W. Jorgensen, 2006. Halide, ammonium, and alkali metal ion parameters for modeling aqueous solutions. *J. Chem. Theory Comput.* 2:1499–1509.

3. Yellen, G., 1984. Ionic permeation and blockade in Ca^{2+} activated K^{+} channels of bovine chromaffin cells. *J. Gen. Physiol* 84:157–186.

4. Heginbotham, L., and R. MacKinnon, 1993. Conduction properties of the cloned Shaker channel. *Biophys. J.* 65:2089–2096.

5. LeMasurier, M., L. Heginbotham, and C. Miller, 2001. KscA: it's an ion channel. *J. Gen. Physiol* 118:303–314.

6. Nimigean, C. M., and C. Miller, 2002. Na^{+} block and permeation in K^{+} channel of known structure. *J. Gen. Physiol* 120:323–325.

7. Pranata, J., and W. L. Jorgensen, 1991. Monte Carlo Simulations Yield Absolute Free Energies of Binding for Guanine-Cytosine and Adenine-Uracil Base Pairs in Chloroform. *Tetrahedron* 47:2491–2501.

8. Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon, 1997. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* 72:1047–1069.

9. Kasterholz, M. A., and P. H. Hünenbuger, 2006. Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation. *J. Chem. Phys.* 124:224501.

10. Kameda, Y., S. Suzuki, H. Ebata, T. Usuki, and O. Uemura, 1997. The short-range structure around Li^{+} in highly concentrated aqueous LiBr. *Bull. Chem. Soc. Jpn.* 70:47–53.

11. Kameda, Y., K. Sugawara, T. Usuki, and O. Uemura, 1998. Hydration structure of Na^{+} in concentrated aqueous solutions. *Bull. Chem. Soc. Jpn.* 71:2769–2776.

12. Smirnov, P., and V. Trostin, 2007. Structures of the nearest surroundings of the K^+, Rb^+ and Cs^+ ion in aqueous solutions of their salts. *Russ. J. Gen. Chem.* 77:2101–2107.

13. Noskov, S., S. Bernèche, and B. Roux, 2004. Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* 431:830–834.

14. Noskov, S., and B. Roux, 2007. Importance of hydration and dynamics on the selectivity on the KcsA and NaK channels. *J. Gen. Physiol.* 129:135–143.

15. Noskov, S., and B. Roux, 2008. Control of ion selectivity in LeuT: two Na^+ binding sites with two different mechanisms. *J. Mol. Biol* 377:804–818.

16. Beglov, D., and B. Roux, 1994. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J. Chem. Phys.* 100:9050–9063.

17. Schwenk, C., T. Hofer, and B. Rode, 2004. "Structure breaking" effect of hydrated Cs^+. *J. Chem. Phys. A* 108:1509–1514.

18. Varma, S., and S. Rempe, 2006. Coordination numbers of alkali metal ions in aqueous solutions. *Biophys. Chem.* 124:192–199.

19. Ansell, S., A. Barnes, P. Mason, G. Neilson, and S. Ramos, 2006. X-ray and neutron scattering studies of the hydration structure of alkali ion in concentrated aqueous solution. *Biophys. Chem.* 124:171–179.

20. Lee, S., and J. Rasaiah, 1996. Molecular dynamics simulations of ion mobility. 2. Alkali metal and halide ions using the SPC/E model for water at 25 °C. *J. Phys. Chem.* 100:1420–1425.

21. Azam, S., T. Hofer, B. Randolf, and B. Rode, 2009. Hydration of sodium(I) and potassium (I) revisited: A comparitive QM/MM and QMCF MD simulation study of weakly hydrated ions. *J. Phys. Chem.* 113:1827–1834.

22. Fulton, J., D. Pfund, S. Wallen, M. Newville, E. Stern, and Y. Ma, 1996. Rubdium ion hydration in ambient and supercritical water. *J. Chem. Phys.* 105:2161–2166.

23. Ramos, S., A. Barnes, G. Neilson, and M. Capitan, 2000. Anomalous X-ray diffraction studies of hydration effects in concentrated aqueous electrolyte solutions. *Chem. Phys.* 258:171–180.

24. Hofer, T., B. Randolf, and B. Rode, 2005. Stucture-breaking effects of solvated RbI in dilute aqueous solution - an ab initio QM/MM MD approach. *J. Comput. Chem.* 26:949–956.

25. Bertagnolli, H., T. Ertel, M. Hoffmann, and R. Frahm, 1991. EXAFS studies of aqueous solutions of rubidium bromide. *Berich. Bunsen. Gesell.* 95:704–709.

26. Ohtomo, N., and K. Arakawa, 1979. Neutron diffraction study of aqueous ionic solutions I. Aqueous solutions of Lithium Chloride and Caesium Chloride. *B. Chem. Soc. Jpn.* 52:2755–2759.

27. Novikov, A., M. Rodnikova, V. Savostin, and O. Sobolev, 1999. The study of hydration effects in aqueous solutions of LiCl and CsCl by inelastic neutron scattering. *J. Mol. Liq.* 82:83–104.

28. Ramos, S., G. Neilson, A. Barnes, and P. Buchanan, 2005. An anomalous x-ray diffraction study of the hydration structures of Cs^+ and I^- in concentrated solutions. *J. Chem. Phys.* 123:214501 – 214501–10.

29. Fowler, P., K. Tai, and M. Sansom, 2008. The selectivity of K^+ ion channels: Testing the hypotheses. *Biophys. J.* 95:5062–5072.

30. Bostick, D., and C. Brooks, 2007. Selectivity in K^+ channels is due to topological control of the permeant ion's coordinated state. *PNAS* 104:9260–9265.

31. Thomas, M., D. Jayatilaka, and B. Corry, 2007. The predominant role of coordination number in potassium channel selectivity. *Biophys. J.* 93:2635–2643.

32. Vora, T., B. Corry, and S. Chung, 2004. A model of sodium channels. *BBA* 1668:106–116.

33. Boudker, O., R. Ryan, D. Yernool, K. Shimamoto, and E. Gouaux, 2007. Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature* 445:387–393.

34. Wingrove, T., and G. Kimmich, 1988. Low-affinity intestinal L-aspartate transport with 2:1 coupling stoichiometry for Na^+/Asp. *Am. J. Physiol Cell Physiology* 255:C737–C744.

35. Neupert-Laves, K., and M. Dobler, 1975. The crystal structure of a K^+ complex of Valinomycin. *Helv. Chim. Acta* 58:432–442.

36. Dobler, M., and R. Phizackerley, 1975. The crystal structure of the NaNCS complex of nonactin. *Helv. Chim. Acta* 57:664–674.

37. Seiler, P., M. Dobler, and J. Dunitz, 1974. Potassium thiocyanate complex of 1,4,7,10,13,16-Hexaoxacyclooctadecane. *Acta Cryst. Sect. B.* 30:2744–2745.

38. Haider, S., G. Parkinson, and S. Neidle, 2003. Structure of a G-quadruplex-ligand complex. *J. Mol. Biol.* 326:117–125.

39. Zhang, Y., M. Dougherty, D. Downs, and S. Ealick, 2004. Crystal structure of an aminoimidazole riboside kinase from *Salmonella enterica* Implications for the evolution of the ribokinase superfamily. *Structure* 12:1809–1821.

40. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, 2005. Scalable molecular dynamics with NAMD. *J. Comp. Chem.* 26:1781–1802.

41. MacKerell Jr., A. D., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. R. III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, 1998. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* 102:3586–3616.

42. Varma, S., D. Sabo, and S. Rempe, 2008. K^+ and Na^+ selectivity in K channels and valinomycin: over-coordination versus cavity-size constraints. *J. Mol. Biol.* 376:13–22.

43. Celik, L., B. Schiøtt, and E. Tajkhorshid, 2008. Substrate binding and formation of an occluded state in the leucine transporter. *Biophys. J.* 94:1600–1612.