Supporting Information

Selective Complexation of K^+ and Na^+ in Simple Polarizable Ion-ligating Systems

David L. Bostick and Charles L. Brooks III*
The University of Michigan
Department of Chemistry and Program in Biophysics
930 N. University Ave.
Ann Arbor, Michigan 48109

* To whom all correspondence should be addressed: brookscl@umich.edu

I. Supporting Discussion

A. Causality and Thermodynamics in Selective Ion-ligating Systems

In the main text, we discuss how the observed ion selectivity is *"caused"* by the external restraints (or field) of the studied models, and is *not "caused"* and *not "controlled"* by the ligands (or molecules) of which the simple model is composed. This word choice is deliberate and has specific meaning. To avoid misconstruation, we provide here an explanation for this word choice, which is grounded in statistical thermodynamic formalism.

The relative favorability of K^+ versus Na^+ in either a real binding site or simple ion-ligating model is quantified by the Helmholtz free energy difference, $\Delta F^{\text{site}}_{\text{K}\to\text{Na}} = F^{\text{site}}_{\text{Na}} - F^{\text{site}}_{\text{K}}$, where $F^{\text{site}}_{\text{X}}$ is the free energy for ion X at the site (or model). Given the free energy difference between K^+ and Na^+ in bulk water, $\Delta F^{\text{bulk}}_{\text{K}\to\text{Na}} = F^{\text{bulk}}_{\text{Na}} - F^{\text{bulk}}_{\text{K}}$, the selective free energy for K^+ over Na^+ in the site (or model) is

$$\Delta\Delta F = \Delta F^{\text{site}}_{\text{K}\to\text{Na}} - \Delta F^{\text{bulk}}_{\text{K}\to\text{Na}} \tag{S1}$$

Clearly, if $\Delta\Delta F > 0$, then the model/host favors binding K^+, if $\Delta\Delta F < 0$, then the model/host favors binding Na^+, and if $\Delta\Delta F = 0$, then the model/host favors binding either ion equivalently.

The free energy of ion X at the site (or model) may be expressed, in a canonical ensemble, as $F^{\text{site}}_{\text{X}} = -RT\ln Q^{\text{site}}_{\text{X}}$, where R is the gas constant and T is the temperature. The partition function is defined as $Q^{\text{site}}_{\text{X}} = \text{Tr}\{\exp[-(K + V_{\text{X}} + W^{\text{ext}}_{\text{X}})/RT]\}$, where $\text{Tr}\{\}$ is the classical trace operator, K is the kinetic energy of the system, $V_{\text{X}} = V_{\text{X}}(\boldsymbol{R})$ is a potential energy function encoding ion-ligand and ligand-ligand interactions, and $W^{\text{ext}}_{\text{X}} = W^{\text{ext}}_{\text{X}}(\boldsymbol{R})$ is an external potential (in the case of a simple ion-ligating model) or potential of mean force (in the case of a real binding site) encoding the ion-ligating system remainder of a

host (i.e., those elements of the system that are neither ion nor ligand). The degrees of freedom of the available ligands, $\boldsymbol{R}$, are written with respect to the position of the ion. Thus the dependence of $Q_{\mathrm{X}}^{\mathrm{site}}$ on the position of the ion is implicit. It is also worth pointing out that both V_{X} and $W_{\mathrm{X}}^{\mathrm{ext}}$ depend on the identity of the particular ion (K^+/Na^+) bound. While this fact may be obvious in the case of V_{X}, one should note that in a real host, $W_{\mathrm{X}}^{\mathrm{ext}}$ encodes not only ion-host interactions, but also accounts for the response of the entire system remainder to all of the ligands interacting with X. As such, one cannot generally assume $W_{\mathrm{K}}^{\mathrm{ext}} = W_{\mathrm{Na}}^{\mathrm{ext}}$. To do so would invoke a form of high-order superposition approximation[1]. When constructing simple models, therefore, one may choose $W_{\mathrm{X}}^{\mathrm{ext}}$ to suit whatever design principle is of interest, which may require[2] $W_{\mathrm{K}}^{\mathrm{ext}} \neq W_{\mathrm{Na}}^{\mathrm{ext}}$.

The so-called "intrinsic" free energy of the model (or host system) with ion X bound is defined as[2-4]

$$F_{\mathrm{X}}^{\mathrm{intrinsic}} = F_{\mathrm{X}}^{\mathrm{site}} - \left\langle W_{\mathrm{X}}^{\mathrm{ext}} \right\rangle = F_{\mathrm{X}}^{\mathrm{site}} - U_{\mathrm{X}}^{\mathrm{rest}} \tag{S2}$$

such that $U_{\mathrm{X}}^{\mathrm{rest}} = \left\langle W_{\mathrm{X}}^{\mathrm{ext}} \right\rangle$ is the average interaction energy of the external potential (in this work, a restraining potential) with the ion and ligands. We may further break $F_{\mathrm{X}}^{\mathrm{site}}$ into entropic and internal energy components

$$F_{\mathrm{X}}^{\mathrm{site}} = U_{\mathrm{X}}^{\mathrm{rest}} + F_{\mathrm{X}}^{\mathrm{intrinsic}} = U_{\mathrm{X}}^{\mathrm{rest}} + U_{\mathrm{X}}^{\mathrm{IL}} + U_{\mathrm{X}}^{\mathrm{LL}} + U_{\mathrm{X}}^{\mathrm{intra}} + \left\langle K \right\rangle - TS \tag{S3}$$

where $U_{\mathrm{X}}^{\mathrm{IL}}$, $U_{\mathrm{X}}^{\mathrm{LL}}$, $U_{\mathrm{X}}^{\mathrm{intra}}$ are the internal energy components from ion-ligand, ligand-ligand, and intramolecular (intra-ligand) interactions, $\left\langle K \right\rangle$ is the average kinetic energy, and S is the entropy. With the intrinsic free energy, $F_{\mathrm{X}}^{\mathrm{intrinsic}}$, defined, it is possible to show two important identities[2-4]

$$\delta F_{\mathrm{X}}^{\mathrm{site}} / \delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}) = \rho_{\mathrm{X}}(\boldsymbol{R}) \quad \text{(S4a)} \qquad \text{and,} \qquad \delta F_{\mathrm{X}}^{\mathrm{intrinsic}} / \delta \rho_{\mathrm{X}}(\boldsymbol{R}) = -W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}) \quad \text{(S4b)}$$

where $\rho_{\mathrm{X}}(\boldsymbol{R})$ is the density function describing the configurational space (i.e., the "structure") available to the ligands (for the benefit of the reader, these identities are derived in the next section, ***I.B.***, of this Supporting Discussion). For a fixed ligand composition, described by V_{X}, these functional derivatives dictate that $W_{\mathrm{X}}^{\mathrm{ext}}$ uniquely determines the following: (a) the configurational space available to the ligands (Eq. S4a) and (b) $F_{\mathrm{X}}^{\mathrm{intrinsic}}$ and its components (Eq. S4b).

With the above, we arrive at the only logical conclusion – for a given molecular composition, $V_{\mathrm{K/Na}}$, the selective free energy, $\Delta\Delta F$, of a simple model (or a site in a host) is uniquely determined by the external field, $W_{\mathrm{K/Na}}^{\mathrm{ext}}$. This is because the model field $W_{\mathrm{K/Na}}^{\mathrm{ext}}$ uniquely determines both $F_{\mathrm{K/Na}}^{\mathrm{intrinsic}}$ and $U_{\mathrm{K/Na}}^{\mathrm{rest}}$, and therefore determines $F_{\mathrm{K/Na}}^{\mathrm{site}}$. This is the definition of "topological" or "architectural" or "structural" control over the selectivity of a binding site[2, 5-10].

We give the discussion above in order to dispel a confusion, presented in a variety of works[11-17], that can lead to the incorrect interpretation of results from simple ion-ligating models. Several studies[11-17] assign a causative relationship (with respect to ion selectivity) to individual components of the free energy difference

$$\Delta F^{\text{site}}_{\text{K}\to\text{Na}} = \Delta U^{\text{rest}}_{\text{K}\to\text{Na}} + \Delta U^{\text{IL}}_{\text{K}\to\text{Na}} + \Delta U^{\text{LL}}_{\text{K}\to\text{Na}} + \Delta U^{\text{intra}}_{\text{K}\to\text{Na}} - T\Delta S_{\text{K}\to\text{Na}} \tag{S5}$$

For example, neglecting the entropic contribution, $-T\Delta S_{\text{K}\to\text{Na}}$, one might observe that the sum, $\Delta U^{\text{IL}}_{\text{K}\to\text{Na}} + \Delta U^{\text{LL}}_{\text{K}\to\text{Na}}$, makes the largest contribution to $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ [11-17]. Noting that this sum originates from ligand interactions with one another and the ion, it could be tempting to conclude that the ligands *cause* or *control* the selectivity, $\Delta\Delta F$ [11-17]. Another common example occurs when one finds the contribution from the external field (or protein host), $\Delta U^{\text{rest}}_{\text{K}\to\text{Na}} = \langle W^{\text{ext}}_{\text{Na}} \rangle - \langle W^{\text{ext}}_{\text{K}} \rangle$, to be negligible compared to other contributions. In this case, it might be tempting to say that the protein host or the simple model's external field (described by $W^{\text{ext}}_{\text{K/Na}}$) *does not cause* or *does not control* the selectivity, $\Delta\Delta F$ [11-17]. Such conclusions miss the point entirely (i.e., they are *ignoratio elenchi*). Their fallacy lies in the fact that the external potential (or protein host), uniquely determines every component in the thermodynamic breakdown of $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ in Eq. S5, thereby *"controlling"* $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ and ultimately *"causing"* the selectivity $\Delta\Delta F$. Attributing selectivity to any subset of components in Eq. S5 does not in any way, by itself, constitute a design principle or statement of *"causality."*

A straightforward illustration of this concept is found in the simplified models of this work where eight (NMA/formamide/water) molecules are included (N_{I} = 8). The selectivities, $\Delta\Delta F$, and contributions to $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ of these models are summarized in Table S1. It is clear that in the absence of any external restraints (i.e. in gas phase), NMA and formamide models yield $\Delta\Delta F < 0$ (–1.5 and –1.9 kcal/mol, respectively) and water models yield $\Delta\Delta F > 0$ (+0.74 kcal/mol). When $W^{\text{ext}}_{\text{K/Na}}$ is designed such that all eight ligands are forced to coordinate the ion (i.e. $N_{\text{I}} = N_{\text{C}} = 8$) all models are K^+-selective ($\Delta\Delta F > 0$), whether composed of NMA, formamide, or water. Also, all models are K^+-selective when $W^{\text{ext}}_{\text{K/Na}}$ is designed to provide generic 3.5 Å confinement. These results suggest two different design principles, i.e. *causes*, for K^+ selectivity: (a) enforcement of 8-fold coordination of K^+/Na^+, in the absence of other constraints, by NMA, formamide, or water and (b) enforcement of generic 3.5 Å confinement for eight NMA, formamide, or water molecules, in the absence of other constraints, around K^+/Na.

Table S1 shows that the component of $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ due to the external field, $W^{\text{ext}}_{\text{K/Na}}$, is negligible ($\Delta U^{\text{rest}}_{\text{K}\to\text{Na}} < 0.12$ kcal/mol, at best) compared to the net selectivity of all models. In fact, the magnitude of the external field contribution, $\Delta U^{\text{rest}}_{\text{K}\to\text{Na}}$, can be considered effectively zero compared to ion-ligand ($\Delta U^{\text{IL}}_{\text{K}\to\text{Na}}$), ligand-ligand ($\Delta U^{\text{LL}}_{\text{K}\to\text{Na}}$), and entropic ($-T\Delta S_{\text{K}\to\text{Na}}$) contributions without significantly changing the results of this work. This observation says nothing whatsoever about the *"cause"* of selectivity in these models. If we impose $W^{\text{ext}}_{\text{K/Na}} \equiv 0$ (i.e. gas phase), the eight C=O-containing compounds, NMA and formamide, are Na^+-selective, and if we impose $W^{\text{ext}}_{\text{K/Na}}$ consistent with 8-fold coordination or 3.5 Å confinement, the eight C=O-containing compounds are K^+-selective. In other

words, upon imposing the external potential, $W^{\text{ext}}_{\text{K/Na}}$, we observe that the larger contributions to $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ – $\Delta U^{\text{IL}}_{\text{K}\to\text{Na}}, \Delta U^{\text{LL}}_{\text{K}\to\text{Na}}$, and $-T\Delta S_{\text{K}\to\text{Na}}$ – change in order to summarily yield a clear qualitative switch in the sign of $\Delta\Delta F$ (from $\Delta\Delta F < 0$ without the potential to $\Delta\Delta F > 0$ with the potential). Moreover, via Eqs. S4a and S4b we know that $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$, and therefore $\Delta\Delta F$ is "caused" by $W^{\text{ext}}_{\text{K/Na}}$. Thus, the considerations of this study demonstrate, analytically and through demonstration, that $W^{\text{ext}}_{\text{K/Na}}$ determines which ion is selected, and to what extent, for a given set of ligands (NMA, formamide, or water). Previous works[11-17] have asserted that, as a rule, the quantity, $-T\Delta S_{\text{K}\to\text{Na}}$, should be negligible, that eight C=O ligands should provide $\Delta U^{\text{IL}}_{\text{K}\to\text{Na}} \approx \Delta F^{\text{bulk}}_{\text{K}\to\text{Na}}$ (~ –17.3 kcal/mol for the AMOEBA model[18]), and that, as a result, $\Delta\Delta F \approx \Delta U^{\text{LL}}_{\text{K}\to\text{Na}}$. Table S1 does not support this proposition as a general rule. In fact, the entropic contribution to $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ was found to range from –4.9 to +6.1 kcal/mol, depending on the model (Table S1). Although it is clear that $W^{\text{ext}}_{\text{K/Na}}$ uniquely determines the breakdown of free energy components in Eq. S5, it is unclear from our data how the individual components of $\Delta F^{\text{site}}_{\text{K}\to\text{Na}}$ may be predicted given a specified external field, $W^{\text{ext}}_{\text{K/Na}}$.

In closing this discussion, we provide one final note on hypothesis testing in the context of design, methodology, and interpretation of simple models such as those in this work and elsewhere[2, 5-13, 19]. The designs of these models test specific hypotheses, and require undistracted scrutiny. For example, although coordinating models such as those in Figure 2a fix the number of coordinating oxygen atoms around K^+/Na^+ (i.e. $N_{\text{I}} \equiv N_{\text{C}}$) they do not explicitly impose *any* other restraints on the included molecular species (e.g. coordination radius, molecular orientation, etc.). Therefore, just because one observes some particular value of $\Delta\Delta F$ for a given N_{I}, it does *not* follow that *all* binding sites with N_{I} will provide that particular selectivity, $\Delta\Delta F$. Models can be designed to impose other constraints (e.g. molecular orientation, coordination radius, etc. – see previous work[6,10]). However, we do not test hypotheses associated with such constraints here.

Let us consider additional specific examples of possible fallacious interpretations of results from simple ion-ligating models. Consider the statement, "six coordinating C=O groups yield Na^+ selectivity according to Figure 2a, but valinomycin binds cations with six ligands and is highly selective for K^+." Statements like this in the literature[16] would appear to reveal an inconsistency in the use of simplified models, but only if one chooses to hyperbolically twist the hypothesis that the models of Figure 2a test. If the 6-fold coordinating model of Figure 2a imposed exactly the same restraints on the ligands that valinomycin does, then *perhaps* it *would* be K^+-selective, but the coordinating models do not test that hypothesis. As another example, consider using 3.5 Å volume-confining models (such as those in Figure 2b) to conclude that coordination by eight C=O ligands causes K^+-selectivity[11-17]. Of course, this *could* be true, but Figure S5 shows that such volume-confining models do not sample 8-fold coordination around both K^+ and Na^+. One might also consider using either unconstrained gas-phase models (Figure 1) or 3.5 Å volume-confining models (Figure 2b) to conclude that coordination by eight water molecules does not cause K^+ selectivity[15, 16]. Again, this *could* be true, but Figures 1b and S5 clearly show that such models do not sample 8-fold coordination around both K^+ and Na^+. Also, although 3.5 Å volume-confining models with eight water molecules provide

smaller $\Delta\Delta F$ than those with eight organic C=O-containing compounds, the opposite is true of unrestrained gas-phase models (i.e. in gas-phase models, water is always more selective than NMA or formamide). To test hypotheses pertaining to the dependence of $\Delta\Delta F$ on coordination, one must design experiments/models that probe what is intended. We point out the above examples because there has been work built around such *non sequitur* to argue for or against specific design principles of ionic selectivity from the standpoint of simple ion-ligating model systems. It is impossible to predict all variations on the illogic one could concoct to explain the results of these systems. The analyses and presentation of this work are designed to eschew such illogic.

B. The Functional Derivative Identities of Eqs. S4a and S4b

The tautologous identities contained in Eqs. S4a and S4b are well known[2-4], however the reader may explicitly verify them by recalling the definition of the free energy, $F_{\mathrm{X}}^{\mathrm{site}} = -RT \ln Q_{\mathrm{X}}^{\mathrm{site}}$, where $Q_{\mathrm{X}}^{\mathrm{site}} = \mathrm{Tr}\{\exp[-(K + V_{\mathrm{X}} + W_{\mathrm{X}}^{\mathrm{ext}})/RT]\}$. To proceed, we begin by noting that, to within a constant, C, we may write $F_{\mathrm{X}}^{\mathrm{site}}$ in terms of a configuration integral alone (taken over the macroscopic volume of the system).

$$F_{\mathrm{X}}^{\mathrm{site}} = -RT \ln \int d\boldsymbol{R} e^{-\left[V_{\mathrm{X}}(\boldsymbol{R}) + W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})\right]/RT} + C \tag{S6}$$

Then we may take the functional derivative

$$\frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})} = \frac{(-RT)}{\int d\boldsymbol{R} e^{-\left[V_{\mathrm{X}}(\boldsymbol{R}) + W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})\right]/RT}} \times \int d\boldsymbol{R}' e^{-\left[V_{\mathrm{X}}(\boldsymbol{R}') + W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}')\right]/RT} \frac{\delta(\boldsymbol{R} - \boldsymbol{R}')}{(-RT)} \tag{S7}$$

where we have used, $\delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}')/\delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}) = \delta(\boldsymbol{R} - \boldsymbol{R}')$. It is then clear that we obtain the relation in Eq. S4a

$$\frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})} = \frac{e^{-\left[V_{\mathrm{X}}(\boldsymbol{R}) + W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})\right]/RT}}{\int d\boldsymbol{R} e^{-\left[V_{\mathrm{X}}(\boldsymbol{R}) + W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})\right]/RT}} = \rho_{\mathrm{X}}(\boldsymbol{R}) \tag{S8}$$

To obtain Eq. S4b, it is useful to realize that one may regard $F_{\mathrm{X}}^{\mathrm{site}}$ as being a functional of either the density (i.e. "structure"), $\rho_{\mathrm{X}}(\boldsymbol{R})$, or the external field, $W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})$.

$$\delta F_{\mathrm{X}}^{\mathrm{site}} = \int d\boldsymbol{R} \frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} \delta \rho_{\mathrm{X}}(\boldsymbol{R}) = \int d\boldsymbol{R} \frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})} \delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}) \tag{S9}$$

This, together with Eq. S8, implies that

$$\frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} = \int d\boldsymbol{R} \rho_{\mathrm{X}}(\boldsymbol{R}) \frac{\delta W_{\mathrm{X}}^{\mathrm{ext}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} \tag{S10}$$

Now, the intrinsic free energy, $F_{\mathrm{X}}^{\mathrm{intrinsic}}$, of the system is that part of $F_{\mathrm{X}}^{\mathrm{site}}$ that *does not* include the internal energy contribution, $U_{\mathrm{X}}^{\mathrm{rest}} = \left\langle W_{\mathrm{X}}^{\mathrm{ext}} \right\rangle$, from the external potential (in this work, a restraining potential – see Eqs. S2 and S3): $F_{\mathrm{X}}^{\mathrm{intrinsic}} = F_{\mathrm{X}}^{\mathrm{site}} - \left\langle W_{\mathrm{X}}^{\mathrm{ext}} \right\rangle = F_{\mathrm{X}}^{\mathrm{site}} - U_{\mathrm{X}}^{\mathrm{rest}}$. It is clear that, by definition,

$$U_{\mathrm{X}}^{\mathrm{rest}} = \left\langle W_{\mathrm{X}}^{\mathrm{ext}} \right\rangle = \int d\boldsymbol{R} \rho_{\mathrm{X}}(\boldsymbol{R}) W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}) \tag{S11}$$

and therefore,

$$\begin{aligned}\frac{\delta F_{\mathrm{X}}^{\mathrm{intrinsic}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} &= \frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} - \frac{\delta \overbrace{\left[\int d\boldsymbol{R}\rho_{\mathrm{X}}(\boldsymbol{R})W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})\right]}^{\left\langle W_{\mathrm{X}}^{\mathrm{ext}}\right\rangle}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} \\ &= \frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} - \left[\int d\boldsymbol{R}'\delta(\boldsymbol{R}-\boldsymbol{R}')W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}') + \int d\boldsymbol{R}\rho_{\mathrm{X}}(\boldsymbol{R})\frac{\delta W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})}\right] \\ &= \frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} - \left[W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}) + \frac{\delta F_{\mathrm{X}}^{\mathrm{site}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})}\right]\end{aligned} \tag{S12}$$

where we have used $\delta\rho_{\mathrm{X}}(\boldsymbol{R}')/\delta\rho_{\mathrm{X}}(\boldsymbol{R}) = \delta(\boldsymbol{R}-\boldsymbol{R}')$. Finally, taking Eq. S10 and S12 together, we obtain the result of Eq. S4b as follows:

$$\frac{\delta F_{\mathrm{X}}^{\mathrm{intrinsic}}}{\delta \rho_{\mathrm{X}}(\boldsymbol{R})} = -W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R}) \tag{S13}$$

This serves as analytical proof that the free energy, $F_{\mathrm{X}}^{\mathrm{site}}$, of the ion-bound site is uniquely determined by the structure (i.e., configurational distribution) of the site, $\rho_{\mathrm{X}}(\boldsymbol{R})$, which is, in turn, uniquely determined by the external field (in this work, a restraining potential), $W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})$. This tautology holds true even if the internal energy contribution from the external field (in this work, a restraining potential) of a model binding site is negligible, or even uniquely zero (i.e., $U_{\mathrm{X}}^{\mathrm{rest}} \equiv \left\langle W_{\mathrm{X}}^{\mathrm{ext}}\right\rangle \equiv 0$). Likewise, the other contributions to the free energy, $F_{\mathrm{X}}^{\mathrm{site}}$, (e.g. ion-ligand, $U_{\mathrm{X}}^{\mathrm{IL}}$, or ligand-ligand, $U_{\mathrm{X}}^{\mathrm{LL}}$, contributions, or the entropic contribution, $-TS$) have no bearing on the *"cause"* of the observed $F_{\mathrm{X}}^{\mathrm{site}}$. For a given ligand composition, $V_{\mathrm{X}}(\boldsymbol{R})$, such contributions are, without exception, determined or *"caused"* by the field $W_{\mathrm{X}}^{\mathrm{ext}}(\boldsymbol{R})$.

C. Structural Analysis and Design of Ion-ligating Systems

We analyzed the structure of all simple models by investigating ion-oxygen radial probability density functions (or radial distribution functions – RDFs). Figure S1 shows RDFs from unrestrained (gas-phase) models. In these models, we found that one singular coordination shell can be maintained around the central K^+/Na^+ only when N_I is ~4 molecules or less. When gas-phase models were comprised of ~5-8 ligands, a second (and in some cases, a third) maximum in the probability density could be observed (Figure S1). At a given time (snapshot) in each unrestrained model trajectory, the ion-oxygen coordination number was defined in the usual way – by locating the first minimum, R_{min}, in the model's RDF, and counting the number of oxygen atoms whose ion-oxygen distance is less than R_{min}. A summary of R_{min} values for each unrestrained model in this work is given in Table S2. Using this definition for all configurations in the trajectory, we calculated the coordination number average and standard deviation for each model (Figure 1b).

To determine the dependence of selective free energy on the coordination number for a given molecule type, we designed models that allowed only one, singular coordination shell around K^+/Na^+. To achieve this, we used an iterative procedure to determine the equilibrium position, R_{eq}, for a half-harmonic ion-oxygen restraint in 8-ligand models (see Figure S2 and Methods). For each molecule type (NMA, formamide, or water) and ion (K^+ or Na^+), **(*a*)** we began by setting R_{eq} equal to the first minimum in the RDF of the corresponding unrestrained gas-phase model. **(*b*)** We then simulated the resulting restrained model, recalculated the RDF, and again determined the first minimum in the RDF. Resetting R_{eq} to the position of the minimum, we continued to repeat step ***b***, above, until no secondary peaks in the resulting RDF were observed, implying only one singular maximum in the RDF and, therefore, only one "coordination" shell. Figure S3 shows RDFs for all of the models of this type. It is clear from these data that the models provide for only one RDF maximum, implying only one ion-oxygen coordination shell.

Figure S4 shows RDFs from all of the 3.5 Å volume-confining models in this work. The first minimum in each RDF is summarized in Table S3. The average coordination number for each model was calculated in the same way as in the unrestrained gas-phase models, and is summarized in Figure S5. Note that only one singular coordination shell around the central ion is seen for all models with 1-5 included molecules. However, secondary maxima (i.e. secondary solvation shells) are seen to emerge, especially around Na^+, when 6-8 molecules are included. This is particularly noticeable in the model with eight water molecules around Na^+. It is also interesting to notice that all of the models with organic C=O-containing compounds around K^+ (top panels) provide one singular coordination shell; however, in the 3.5 Å volume-confining models with 7-8 water molecules around K^+, a secondary solvation shell begins to emerge.

II. Methods

Simulation/Analysis of K^+/Na^+ in Simplified Ion-ligating Model Systems

All simplified models were comprised of a single fixed cation, K^+ or Na^+, surrounded by $N_I \leq 8$ neutral molecular species (NMA, formamide, or water) under the influence of an applied external field, $W_{K/Na}^{ext}$. Molecular interactions were described by the AMOEBA force field[18, 20], and were simulated using the TINKER simulation package[21] at constant temperature (298 K) with Langevin dynamics. Calculations were performed with a time step of 1 fs. Full interactions were included for all molecular species in each model. Configurations were saved every 0.1 ps for later analysis.

To determine $\Delta F_{K\rightarrow Na}^{site}$ for each model, we performed free energy perturbation (FEP) calculations by alchemically transforming the Hamiltonian describing each of our various models over a series of eleven simulations[2]. Each simulation in the perturbation series was 12 ns in length, and carried out at coupling parameter values, $\lambda = \{0.0,0.1,0.2,\cdots,1.0\}$. This allowed a 20-step perturbation for each toy. The first 6 ns half of each trajectory was treated as an equilibration run, and thrown away. The latter 6 ns production segment of each trajectory was used for structural and free energy analyses. For the FEP analyses, we divided the 6 ns production runs into 10 equally sized (600 ps) blocks for averaging and calculation of standard error. The resulting value of $\Delta F_{K\rightarrow Na}^{site}$ was used along with the value for $\Delta F_{K\rightarrow Na}^{bulk}$ from previous work[18] (–17.3 ± 0.1 kcal/mol) to calculate the selective free energy, $\Delta\Delta F$ (Eq. S1).

The effects of three different classes of external field, $W_{K/Na}^{ext}$, were investigated. The first class was entirely unrestrained (i.e. gas phase, with $W_K^{ext} = W_{Na}^{ext} = 0$), the second was a generic 3.5 Å volume confinement as described in previous works[2, 5, 11-13], and the third class was designed to ensure that all included ligands coordinated both K^+ and Na^+. The external potential in the second and third classes was a half-harmonic restraint (harmonic constant of 10^3 kcal/mol/Å^2) employed by the "restrain-distance" directive in TINKER acting between the central ion and the oxygen atoms of the included molecules. For the 3.5 Å volume-confining (second class) models, the half-harmonic restraining potential was set to have an equilibrium distance of 3.5 Å from the central cation. This distance served as a spherically symmetric boundary, preventing the oxygen atoms of the included model compounds from traveling farther than ~3.5 Å (see Figure S4). The coordinating (third class) models employed different harmonic potential equilibrium distances for K^+ and Na^+. Thus, calculation of $\Delta F_{K\rightarrow Na}^{site}$ involved the simultaneous alchemical transformation of $K^+ \rightarrow Na^+$ and $W_K^{ext} \rightarrow W_{Na}^{ext}$ over the series of eleven FEP simulations. The equilibrium distance for the NMA, formamide, and water-based models with K^+/Na^+ were determined individually by placing the half-harmonic equilibrium distance such that formation of outer solvation shells was prevented (see Figure S2 and Supporting Discussion, part B). The positions of the equilibrium distance, $R_{K/Na}$, for the coordination restraint models were as follows: for NMA-based models $R_K = 4.19$ Å and $R_{Na} = 3.13$ Å, for formamide-based models $R_K = 3.77$ Å and $R_{Na} = 3.06$ Å, and for water-based models $R_K = 3.32$ Å and $R_{Na} = 2.74$ Å.

Supplementary Tables

Table S1. Thermodynamic analysis of simple ion-ligating model systems composed of eight included (NMA, formamide, or water) molecules. Quantities listed are K^+ selective free energy ($\Delta\Delta F$), Helmholtz free energy difference between K^+ and Na^+ ($\Delta F^{site}_{K\rightarrow Na}$ or ΔF_{model}), total internal energy difference between K^+ and Na^+ (ΔU), entropic contribution to ΔF_{model} ($-T\Delta S$), ion-ligand contribution to ΔU (ΔU_{IL}), ligand-ligand contribution to ΔU (ΔU_{LL}), external restraint potential contribution to ΔU (ΔU_{rest}), and ligand intramolecular contribution to ΔU (ΔU_{intra}). Values are reported in kcal/mol. Values in parenthesis (…) represent standard errors from block averaging.

Molecule (Ligand) Type	Simulated Model System	$\Delta\Delta F$	ΔF_{model}	ΔU	$-T\Delta S$	ΔU_{IL}	ΔU_{LL}	ΔU_{rest}	ΔU_{intra}
NMA	**No Restraints (Gas Phase)**	**–1.5 (0.4)**	–18.8 (0.4)	–13.8 (1.3)	–4.9 (0.9)	–22 (1)	7.7 (1.3)	0 (0)	0.8 (1.1)
	Coordination Restraint	**4.4 (0.2)**	–12.9 (0.2)	–19.1 (0.8)	6.1 (0.8)	–33 (2)	15 (2)	0.030 (0.007)	–1.2 (1.3)
	Generic 3.5 Å Restraint	**3.2 (0.2)**	–14.14 (0.14)	–17.5 (0.9)	3.4 (0.9)	–27 (2)	10 (2)	0.012 (0.005)	–0.6 (1.3)
Formamide	**No Restraints (Gas Phase)**	**–1.9 (0.2)**	–19.2 (0.1)	–23 (2)	4 (2)	–30 (6)	7 (6)	0 (0)	0 (3)
	Coordination Restraint	**5.0 (0.2)**	–12.31 (0.09)	–16.8 (0.8)	4.5 (0.8)	–33.7 (0.8)	17.4 (1.5)	0.035 (0.004)	–0.5 (1.2)
	Generic 3.5 Å Restraint	**3.3 (0.2)**	–14.0 (0.1)	–15.3 (1.2)	1.2 (1.2)	–26.6 (1.1)	11 (2)	0.014 (0.004)	0 (2)
Water	**No Restraints (Gas Phase)**	**0.74 (0.15)**	–16.56 (0.05)	–18 (2)	2 (2)	–23 (2)	5 (3)	0 (0)	0 (3)
	Coordination Restraint	**6.4 (0.2)**	–10.89 (0.07)	–16.3 (0.2)	5.4 (0.2)	–26.1 (0.6)	9.8 (0.8)	0.118 (0.006)	0.0 (0.4)
	Generic 3.5 Å Restraint	**1.55 (0.14)**	–15.75 (0.03)	–18.3 (0.2)	2.5 (0.2)	–21.9 (0.5)	3.6 (0.7)	0.013 (0.006)	0.1 (0.3)

Table S2. Position of the first minimum of the RDF from unrestrained (gas-phase) model systems composed of N_I = 1-8 (NMA, formamide, or water) molecules around K^+/Na^+. All values are in Å.

Molecule (Ligand) Type		$N_I = 1$	$N_I = 2$	$N_I = 3$	$N_I = 4$	$N_I = 5$	$N_I = 6$	$N_I = 7$	$N_I = 8$
NMA	RDF 1st Minimum (Å) K+	3.1	3.3	3.4	4.8	4.2	4.3	4.8	4.6
	RDF 1st Minimum (Å) Na+	2.6	2.7	2.9	3.3	3.3	5.2	3.9	4.4
Formamide	RDF 1st Minimum (Å) K+	3.2	3.4	3.8	4.6	4.3	4.1	4.8	4.5
	RDF 1st Minimum (Å) Na+	2.6	2.8	3.1	3.2	5.0	4.5	3.9	4.1
Water	RDF 1st Minimum (Å) K+	3.7	3.7	3.7	3.8	3.6	3.6	3.6	3.6
	RDF 1st Minimum (Å) Na+	2.8	3.0	3.3	3.3	3.3	3.2	3.2	3.2

Table S3. Position of the first minimum of the RDF from 3.5 Å volume-confining model systems composed of N_I = 1-8 (NMA, formamide, or water) molecules around K^+/Na^+. All values are in Å.

Molecule (Ligand) Type		$N_I = 1$	$N_I = 2$	$N_I = 3$	$N_I = 4$	$N_I = 5$	$N_I = 6$	$N_I = 7$	$N_I = 8$
NMA	RDF 1st Minimum (Å) K+	3.0	3.2	3.4	3.6	3.6	3.6	3.6	3.6
	RDF 1st Minimum (Å) Na+	2.6	2.7	2.8	3.4	3.3	3.4	3.4	3.2
Formamide	RDF 1st Minimum (Å) K+	3.2	3.4	3.5	3.6	3.6	3.6	3.6	3.6
	RDF 1st Minimum (Å) Na+	2.6	2.8	3.0	3.4	3.4	3.4	3.4	3.2
Water	RDF 1st Minimum (Å) K+	3.5	3.5	3.5	3.5	3.5	3.5	3.4	3.3
	RDF 1st Minimum (Å) Na+	3.5	3.5	3.5	3.5	3.3	3.2	3.1	3.0

Supplementary Figures

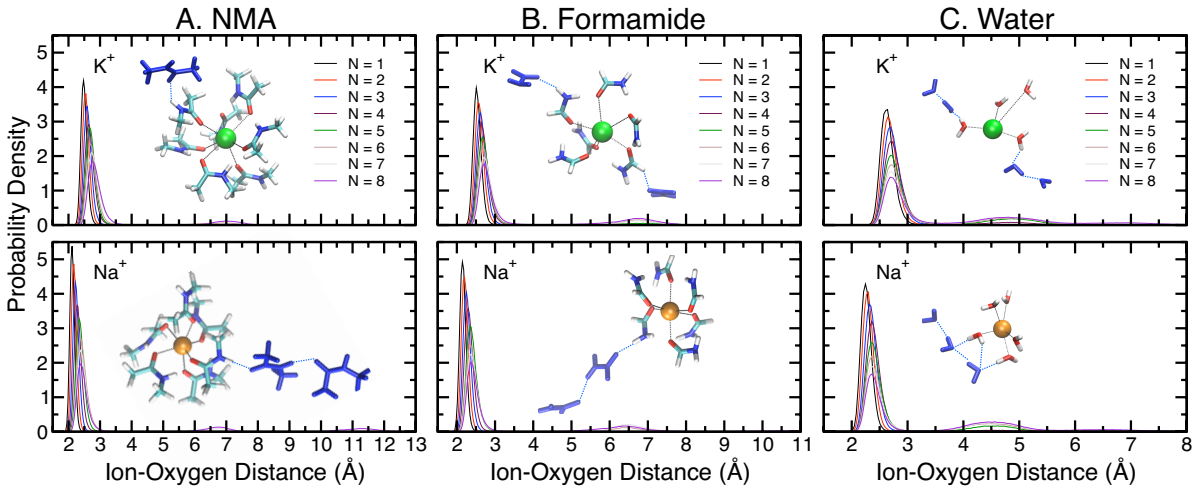

Fig S1: Ion-oxygen RDFs from "unrestrained" (gas-phase) ion-ligating models with 1-8 included molecular species around K^+ (top panels) and Na^+ (bottom panels). Inset configurations show example snapshots from models containing eight included molecules. Ion-ligand coordination interactions are black; outer solvation shell molecules and hydrogen bonds are blue. The first minimum in each RDF is summarized in Table S2, and the average coordination number for each model is summarized in Figure 1b. Note the outer (second and/or third) solvation shells around the central cation in models including more than ~4 molecules.

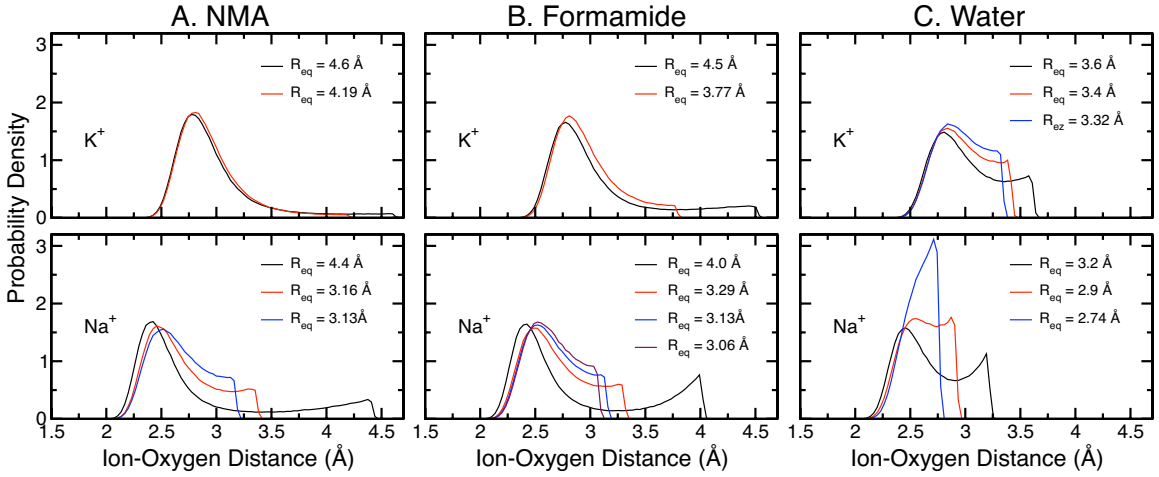

Fig S2: Ion-oxygen RDFs from the iterative process of designing the ion-coordinating models for K^+ (top panels) and Na^+ (bottom panels). All models included eight molecules. Each curve represents the result from a tested half-harmonic ion-oxygen restraint at equilibrium position, R_{eq}, with respect to the central cation. The first restraint tested in each panel (black curve) had the largest R_{eq}, corresponding with the first minimum in the corresponding unrestrained model (see Table S2). The position of R_{eq} was then decreased by placing it at the first minimum of the RDF of the resulting model. This process was repeated until the resulting RDF displayed no second maximum.

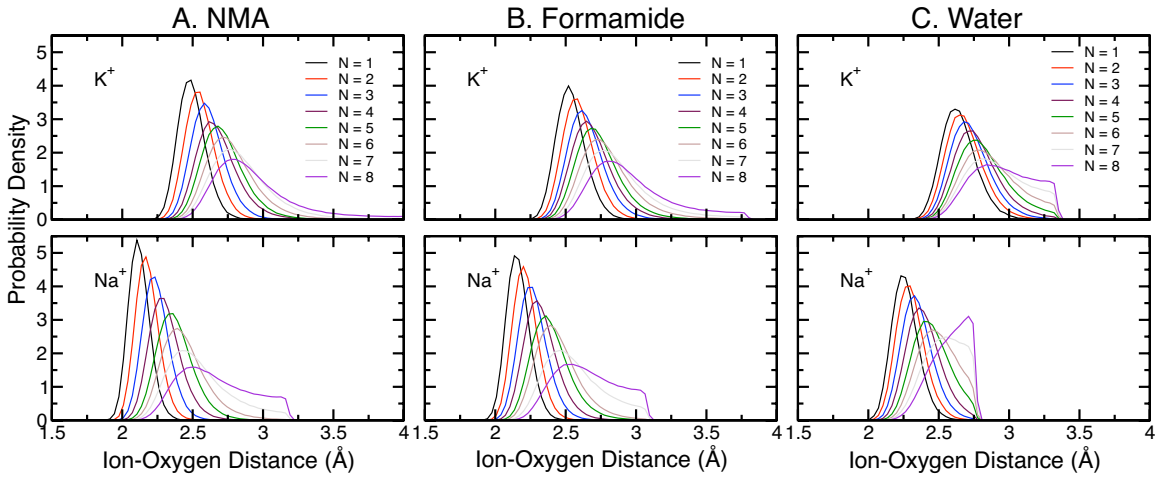

Fig S3: Ion-oxygen RDFs from "coordinating" ion-ligating models with 1-8 included molecular species around K^+ (top panels) and Na^+ (bottom panels). All models were seen to provide one singular coordination shell around the central ion, consistent with the design of the model, which was intended to probe the selectivity in an environment that provides ion-oxygen coordination by $N_I = N_C$ molecules, in the absence of any other explicit restraints on the ligands. No secondary maxima (i.e. secondary solvation shells) are seen to emerge.

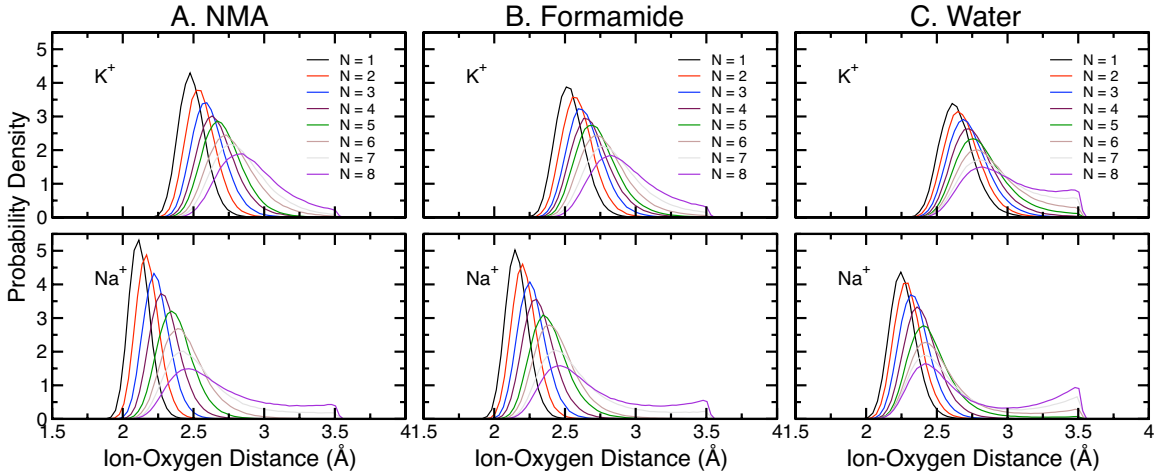

Fig S4: Ion-oxygen RDFs from 3.5 Å "volume-confining" models with 1-8 included molecular species around K^+ (top panels) and Na^+ (bottom panels). The first minimum in each RDF is summarized in Table S3, and the average coordination number for each model is summarized in Figure S5. Note that only one singular coordination shell around the central ion is seen for all models with 1-5 included molecules. However, secondary maxima (i.e. secondary solvation shells) are seen to emerge, especially around Na^+, when 6-8 molecules are included. This is particularly noticeable in the model with eight water molecules around Na^+. It is also interesting to notice that all of the models with organic C=O-containing compounds around K^+ (top panels) provide one singular coordination shell, however in the models with 7-8 water molecules around K^+, a secondary solvation shell begins to emerge.

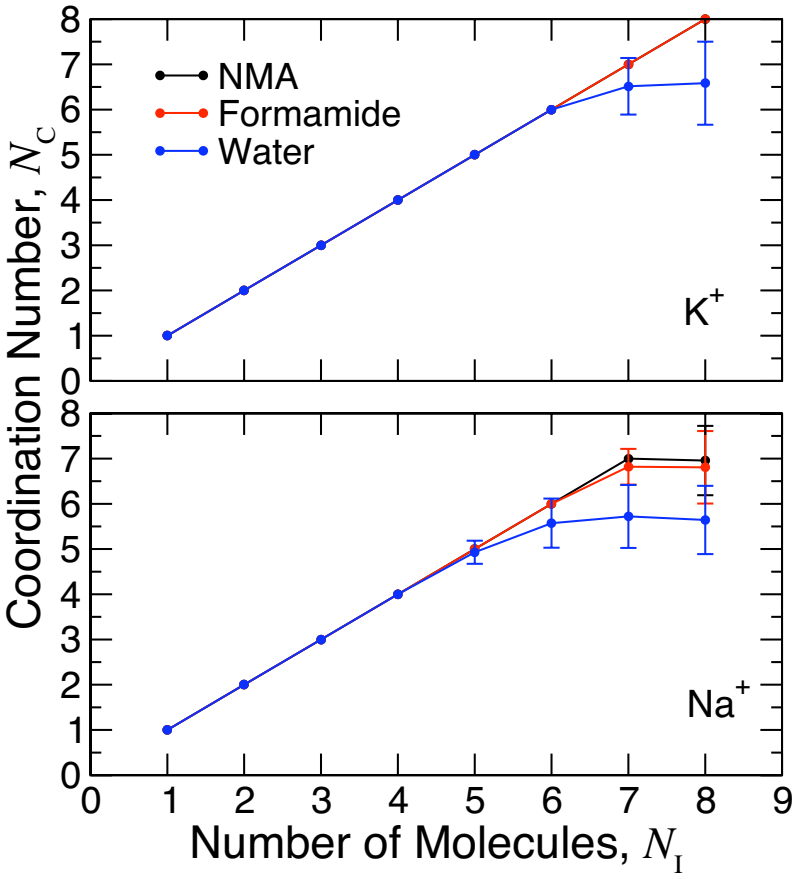

Fig S5: Average coordination number, N_C, as a function of the number of included molecules, N_I, included in 3.5 Å "volume-confining" models. Vertical bars represent standard deviation of the sample.

References

1. Matsuda, H., Physical Nature of Higher-Order Mutual Information: Intrinsic Correlations and Frustration. **2000**.
2. Bostick, D. L.; Arora, K.; Brooks III, C. L., K^+/Na^+ Selectivity in Toy Cation Binding Site Models Is Determined by the 'Host'. *Biophys. J.* **2009,** 96, (10), 3887-3896.
3. Evans, R., The Nature of the Liquid-Vapor Interface and Other Topics in the Statistical Mechanics of Non-Uniform, Classical Fluids. *Adv. Phys.* **1979,** 28, (2), 143-200.
4. Henderson, J. R.; Schofield, P., Statistical Mechanics of a Fluid Drop. *Proc. R. Soc. London, A* **1982,** 380, (1778), 211-227.
5. Bostick, D. L.; Brooks III, C. L., Selectivity in K^+ Channels Is Due to Topological Control of the Permeant Ion's Coordinated State. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** 104, (22), 9260-9265.
6. Bostick, D. L.; Brooks III, C. L., Statistical Determinants of Selective Ionic Complexation: Ions in Solvent, Transport Proteins, and Other 'Hosts'. *Biophys. J.* **2009,** 96, (11), 4470-4492.
7. Thomas, M.; Jayatilaka, D.; Corry, B., The Predominant Role of Coordination Number in Potassium Channel Selectivity. *Biophys. J.* **2007,** 93, (8), 2635-2643.
8. Varma, S.; Rempe, S. B., Tuning Ion Coordination Architectures to Enable Selective Partitioning. *Biophys. J.* **2007,** 93, (4), 1093-1099.
9. Varma, S.; Rempe, S. B., Structural Transitions in Ion Coordination Driven by Changes in Competition for Ligand Binding. *J. Am. Chem. Soc.* **2008,** 130, (46), 15405-15419.
10. Varma, S.; Sabo, D.; Rempe, S. B., K^+/Na^+ Selectivity in K Channels and Valinomycin: Over-Coordination Versus Cavity-Size Constraints. *J. Mol. Biol.* **2008,** 376, 13-22.
11. Noskov, S. Y.; Bernèche, S.; Roux, B., Control of Ion Selectivity in Potassium Channels by Electrostatic and Dynamic Properties of Carbonyl Ligands. *Nature* **2004,** 431, 830-834.
12. Noskov, S. Y.; Roux, B., Ion Selectivity in Potassium Channels. *Biophys. Chem.* **2006,** 124, (124), 279-291.
13. Noskov, S. Y.; Roux, B., Importance of Hydration and Dynamics on the Selectivity of the Kcsa and Nak Channels. *J. Gen. Physiol.* **2007,** 129, (2), 135-143.
14. Noskov, S. Y.; Roux, B., Control of Ion Selectivity in Leut: Two Na^+ Binding Sites with Two Different Mechanisms. *J. Mol. Biol.* **2008,** 377, (3), 804-818.
15. Roux, B., Exploring the Ion Selectivity Properties of a Large Number of Simplified Binding Site Models. *Biophys. J.* **2010,** 98, (12), 2877-2885.
16. Yu, H.; Noskov, S. Y.; Roux, B., Hydration Number, Topological Control, and Ion Selectivity. *J. Phys. Chem. B.* **2009,** 113, (25), 8725-8730.
17. Yu, H.; Roux, B., On the Utilization of Energy Minimization to the Study of Ion Selectivity. *Biophys. J.* **2009,** 97, (8), L15-L17.
18. Grossfield, A.; Ren, P.; Ponder, J. W., Ion Solvation Thermodynamics from Simulation with a Polarizable Force Field. *J. Am. Chem. Soc.* **2003,** 125, (50), 15671-15682.
19. Fowler, P. W.; Tai, K.; Sansom, M. S. P., The Selectivity of K^+ Ion Channels: Testing the Hypotheses. *Biophys. J.* **2008,** 95, (11), 5062-5072.
20. Ren, P.; Ponder, J. W., Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B.* **2003,** 107, (24), 5933-5947.
21. Ponder, J. W. *Tinker: Software Tools for Molecular Design*, 4.2; Saint Louis, MO, 2004.