Supporting Information for "Polarization Effects in Water-Mediated Selective Cation Transport Across a Narrow Channel"

By Van Ngo, Hui Li, Alexander D. MacKerell, Jr., Toby W. Allen,

Benoît Roux, and Sergei Noskov

Impact of other simulation parameters on the computed PMFs

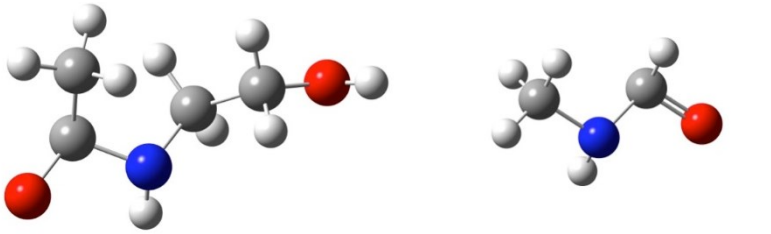
The use of CMAP corrections have been shown to impact PMFs for ion permeation across gA.[1] **Figure S2** shows the two-dimensional (2D) PMFs of alanine dipeptide without and with the L-CMAP and D-CAMP corrections. The dihedrals (φ, ψ) of the backbone of the beta-helical structure should be on the top-left corner and near bottom-right corners of the PMFs. **Figure S4** shows the overall stable regions of the torsional angles of gA. Clearly, the torsional parameters of the protein backbones without any CMAP corrections capture well the stable states of both L- and D-amino acids of gA simultaneously, even though they don't have correct free-energy barriers between minima. **Figure S5** shows that the ions may be sensitive to the free-energy landscapes of the gA backbone. The PMF of K^+ without CMAP has a free-energy barrier (with respect to the bulk) ~3 kcal/mol higher than the PMF shown in Fig. 3 (main text), but about 2 kcal/mol lower than the barrier of C27 PMF. A similar conclusion can be drawn for the PMF of Na^+ without CMAP. We observed that the PMFs without CMAP appear to quickly converge within 5 ns/window as compared to the other PMFs with CMAP, which require more than 6 ns/window to converge. This can be explained by the fact that the areas of the stable states with (φ, ψ) = (60,-150:-60), corresponding to the conformations of the D-amino acids in the absence of the D-CMAP in the force-field are less accessible compared to simulations where D-CMAP correction turned off. To keep the parameters consistent for the ions with carbonyl oxygen atoms of both the backbones and termini, we used the NBFIX and THOLE parameters in **Table 1** for the PMFs in the main text, even though they are slightly different from the parameters obtained separately from the parametrization steps (see **Figure S5**).

It is known that system finite-size effects would contribute to the energy barriers calculation for ion permeation.[2] Allen et al. showed that different sizes of a lipid membrane for embedding gA would have different effects on the crossing free-energy barrier of K^+ using C27. Specifically, using the Poisson-Boltzmann equation, the size-correction[3] due to the finite-size can be estimated by

$$\Delta G_{\text{size}}(\text{z}) = G(z; \varepsilon_{\text{m}}; L_0) - G(z; \varepsilon_{\text{m}}; L{\sim}\infty), \quad \text{(S1)}$$

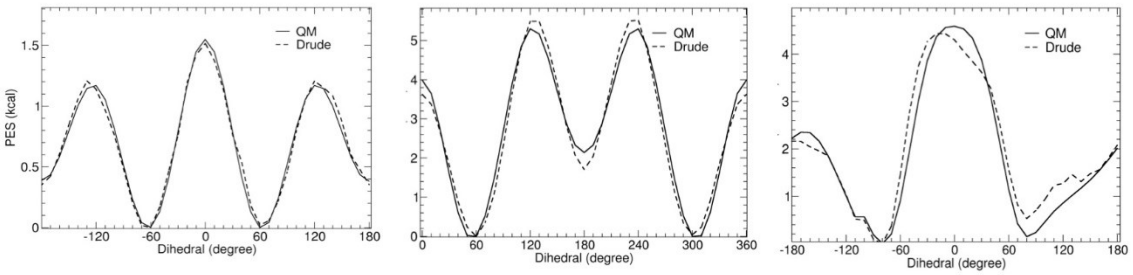
where ε_{m} is the dielectric of the lipid membrane, and $L = L_0$ is the periodic length of the finite system. The estimate is carried out by stripping off water molecules, lipids, and other ions. As a result, the system for the Poisson-Boltzmann equation contains only one Drude ion, the two peptide chains of Drude gA, and a slab of a continuum lipid membrane with a dielectric constant ε_{m}=2-7. We used PB parameters from Refs.[3-4] (water dielectric constant = 80, dielectric constant of the protein = 1 within the cylinder of radius 3 Å, and along $-12.5 \leq z \leq 12.5$ Å), L_0 = 60.5 Å representing an orthorhombic gA system with a similar size to our system, and L = 181.5 Å, representing a very large system. The relative values of ΔG_{size} (**Figure S7**) with respect to $\Delta G_{\text{size}}(\text{z} = -20)$ is at a maximum of 0.03 kcal/mol, which is much smaller than the uncertainty of 2 kcal/mol. This means that the finite-size effect in our simulations is negligible.

Supporting Figures



(a)2-(methylamino)ethanol (b) N-methylformamide

Figure S1. Model compounds (a) 2-(methylamino)ethanol and (b) N-methylformamide (formyl) used for FF development of the ethanolamine and formyl termini, respectively.



(a)HDP1A OD31B CD32A CD32A (b)OD31B CD32A CD32A ND2A2 (c)CD2O1A ND2A2 CD32A CD32A

Figure S2. Potential energy surfaces obtained from scanning dihedral angles in the model compound 2-(methylamino)ethanol, calculated with the optimized Drude parameters and quantum mechanical (QM) calculations.

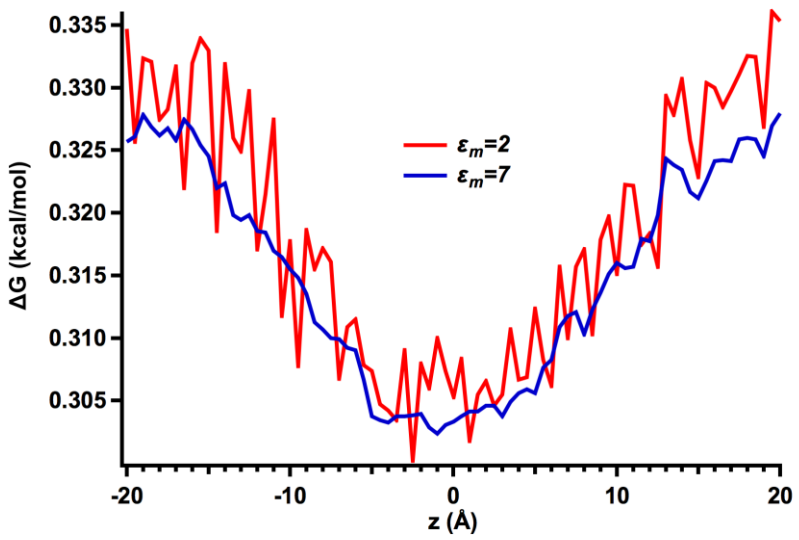


Figure S3. Poisson-Boltzmann finite size corrections.

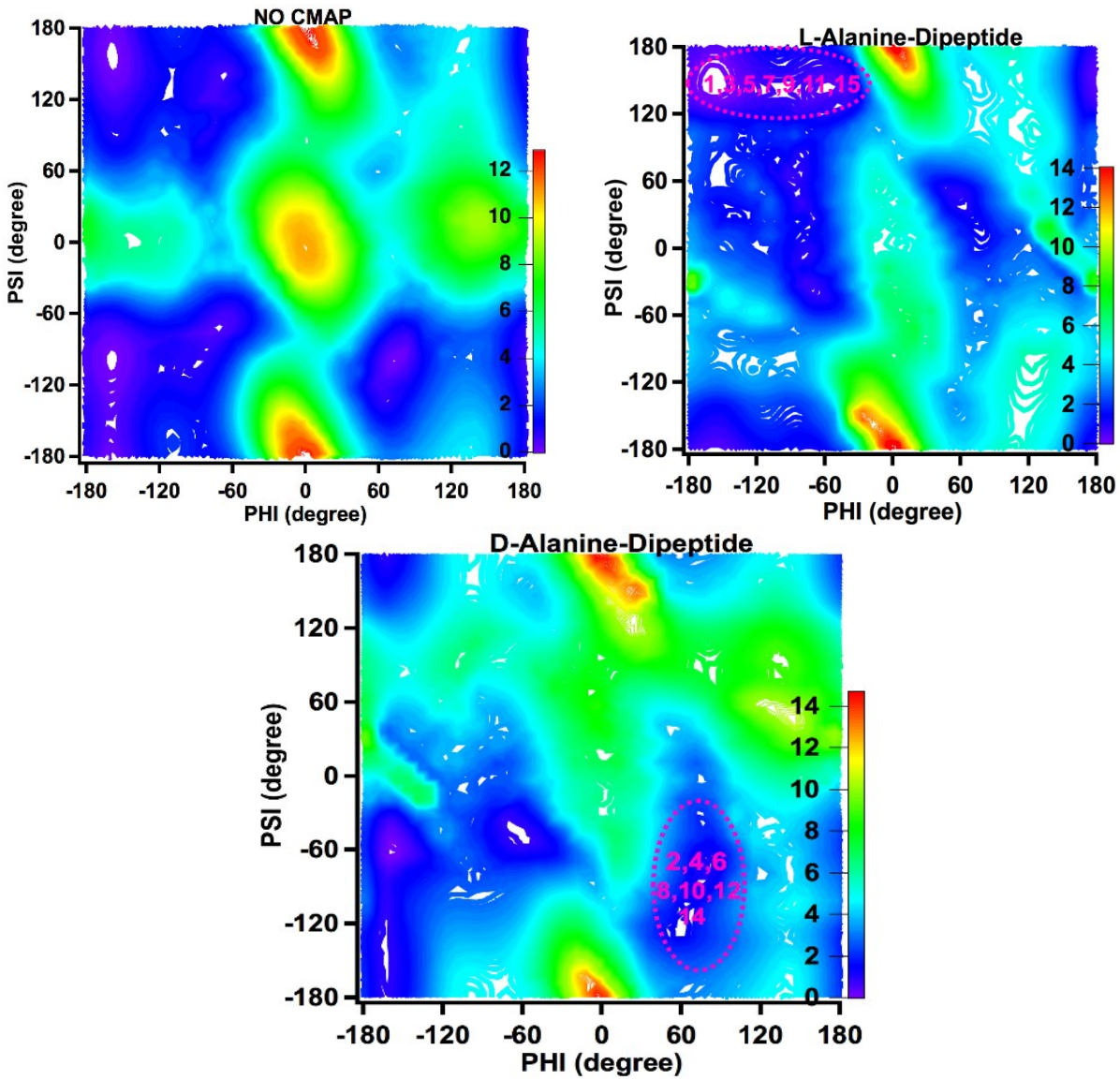


Figure S4. 2D PMF of ALA-Dipeptide without CMAP (left), with the L-CMAP (center) and the D-CMAP (right). The odd and even numbers enclosed by the ovals indicate the regions of the Phi, Psi dihedrals sampled by the L- and D- amino acids in gA, respectively.

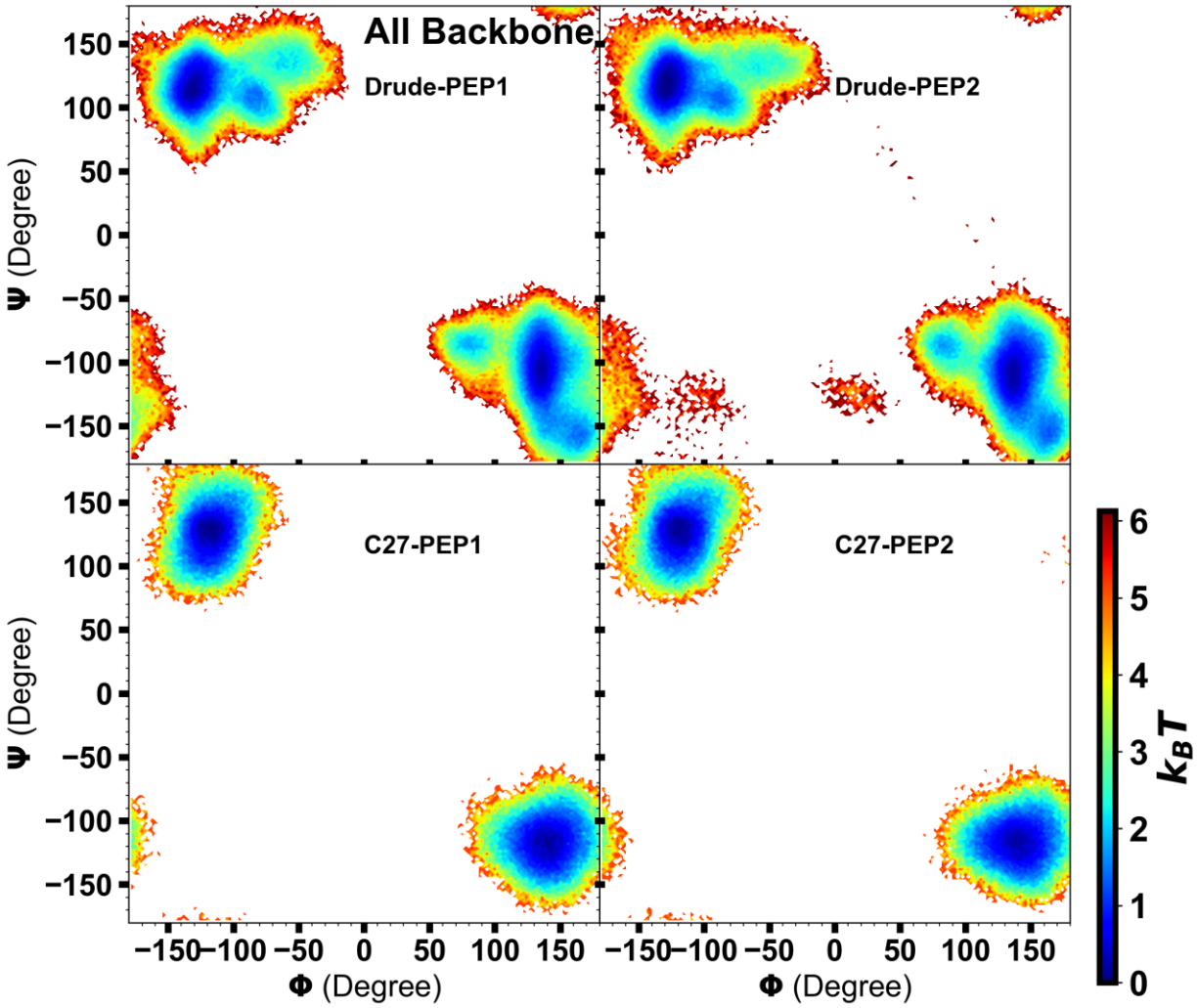


Figure S5. Ramachandran plot of all residues in gA. C27 does not include CMAP, which was shown to increase the free-energy barriers for ions. Drude simulations include the L-CMAP and D-CMAP for the L- and D-residues. Ramachandran plots of all individual residues are shown in Figure 3. Sequence of the gA peptide: formyl-Val_1-Gly_2-Ala_3-$DLeu_4$-Ala_5-$DVal_6$-Val_7-$DVal_8$-Trp_9-$DLeu_{10}$-Trp_{11}-$DLeu_{12}$-Trp_{13}-$DLeu_{14}$-Trp_{15}-ethanolamine.

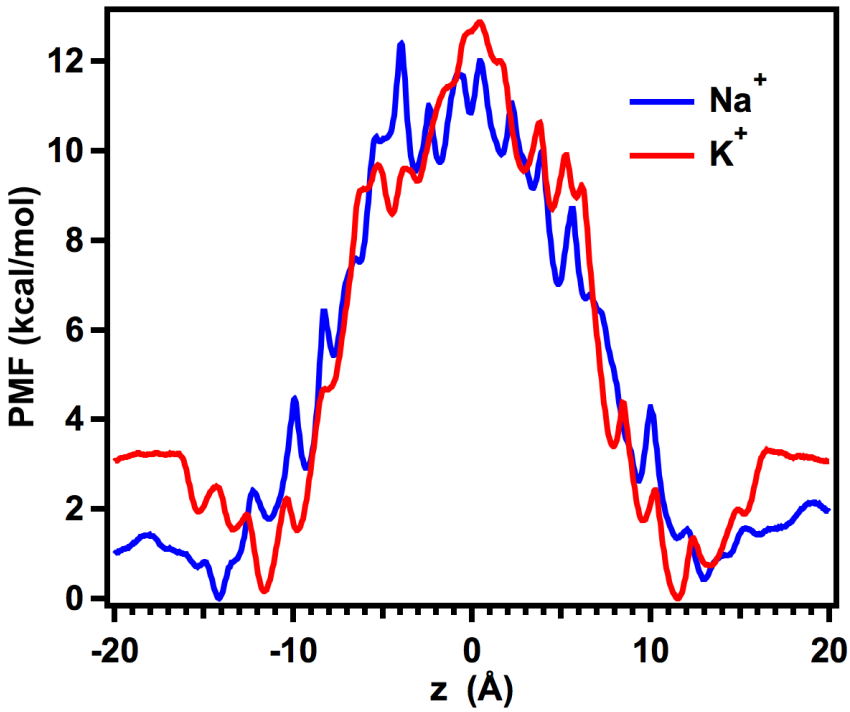


Figure S6. PMFs calculated without CMAP correction (converged at ~2.2-4 ns/window).

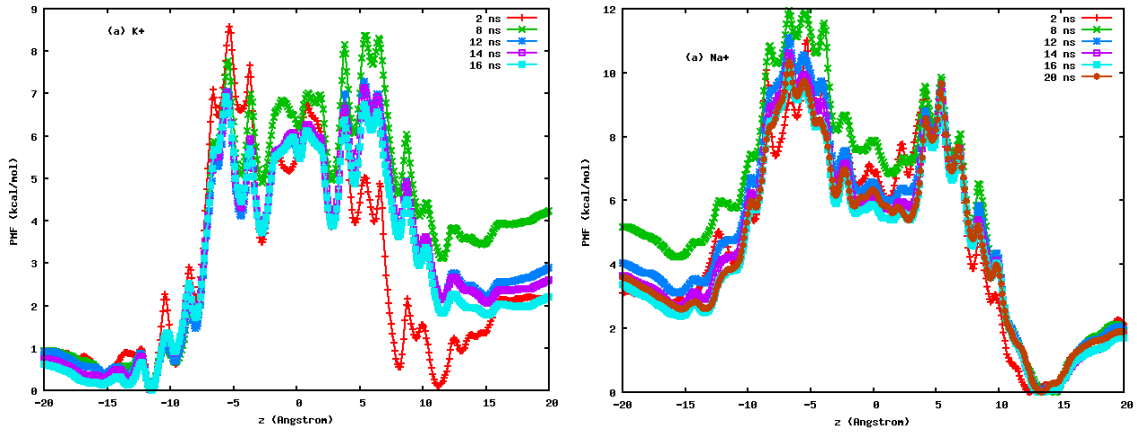


Figure S7. PMF convergence for unsymmetrized PMFs for K^+ and Na^+ (T_{Drude} = 0.1 K). The times listed in the panels indicate simulated timescales per window in H-REMD. The uncertainty is estimated from the final PMFs as as $\varepsilon(|z|) = W(z > 0) - W_1 - (W(z < 0) - W_2)$, where $W(z > 0)$ and $W(z < 0)$ are the parts of the PMF with z > 0 and z < 0, respectively; W_1 and W_2 are the minimum values of $W(z > 0)$ and $W(z < 0)$, respectively. From this way of estimating the uncertainty, we can see that the largest uncertainty actually arises near the middle of channel, whose terminal groups are formyl termini.

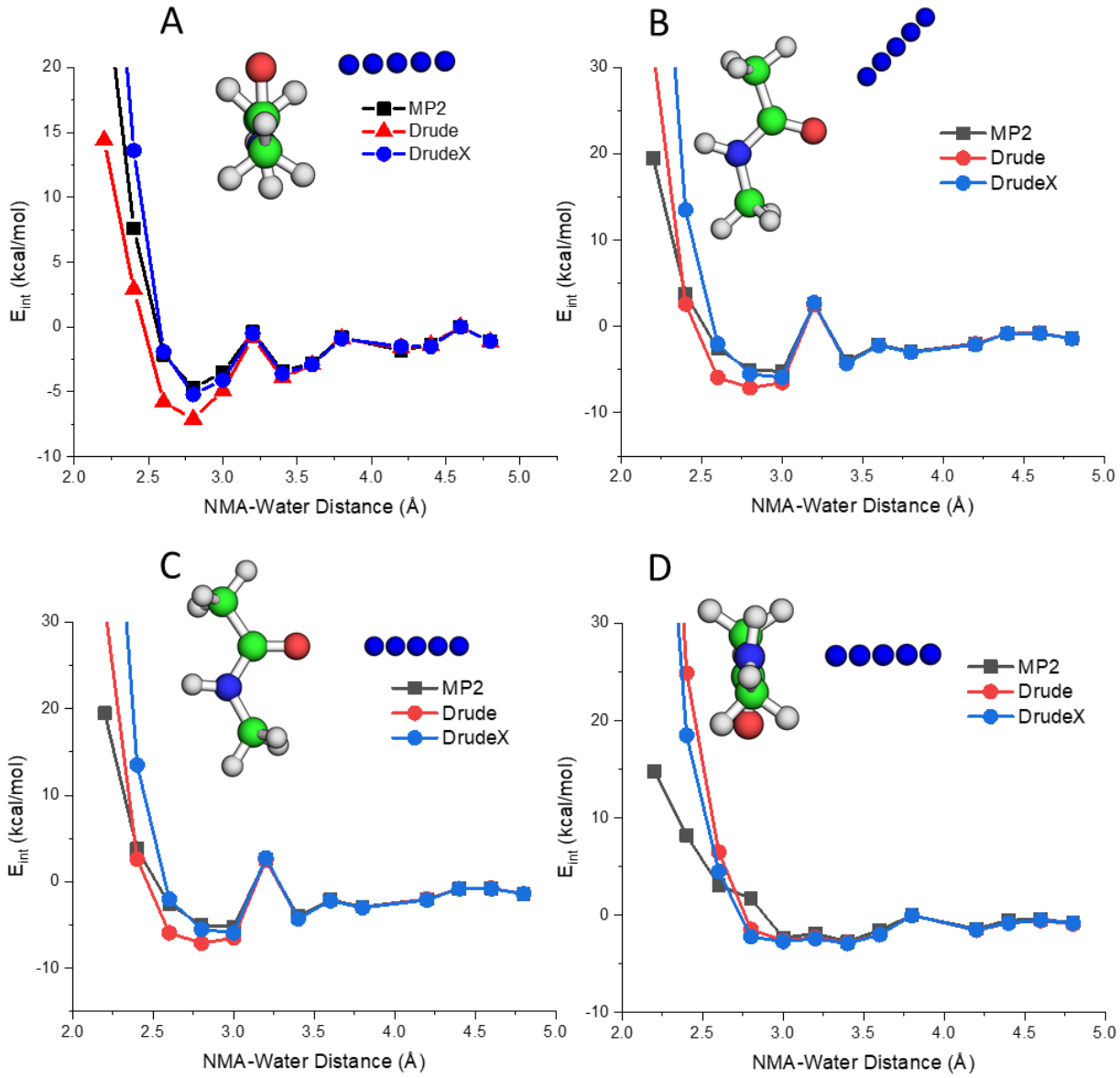


Figure S8. Gas-Phase interaction energies between NMA and H_2O molecules from ab-initio, Drude and DrudeX calculations. The water molecule was translated along the scan axis with 0.2 Å interval from 2.0 to 5.0 Å distance between heavy atom of NMA (O or N) and an oxygen of water. The single-point energies were calculated for the optimized geometries of NMA-water complex. The position of the water oxygen in each energy scans are illustrated by the blue spheres. Four scans were considered with A) Water probe was translated along the axis perpendicular to the C=O bond; B) Water probe was translated along the axis representing position of a lone pair (120° to C=O bond); C) Water probe was translated along the vector defined by C=O bond (180°) and D) hydrogen-bond to an amide nitrogen scan.

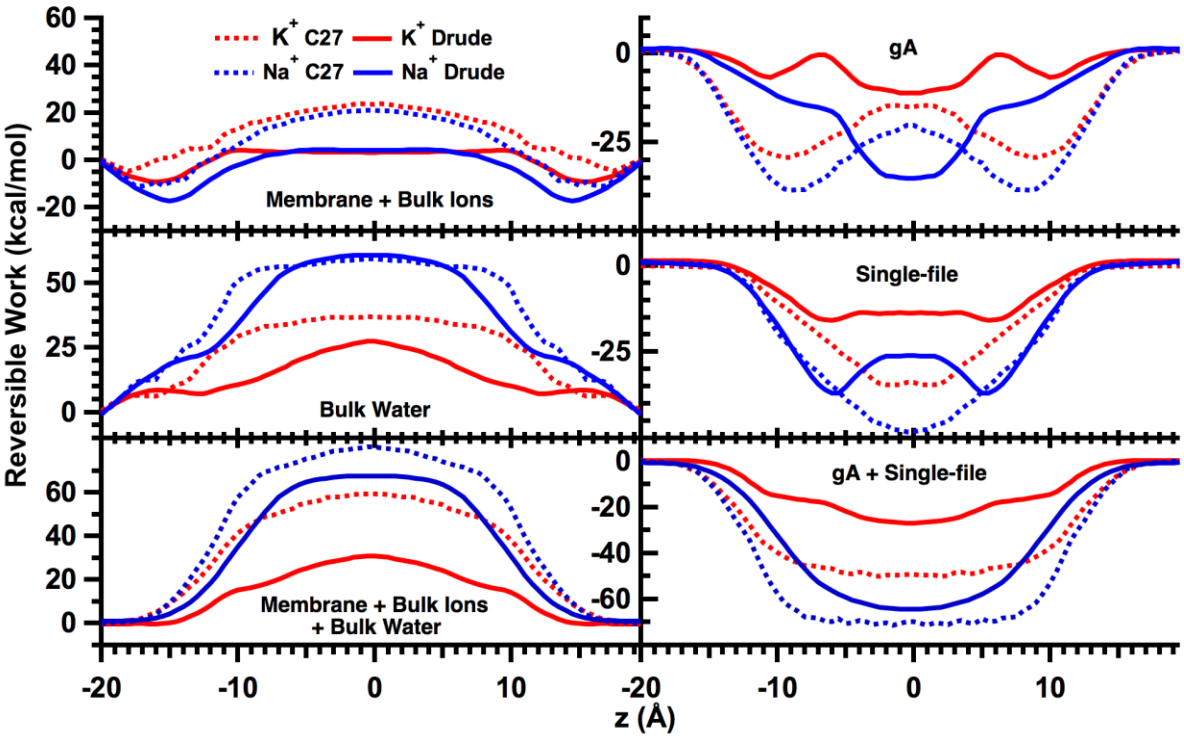


Figure S9. Mean force decompositions for C27 versus Drude.

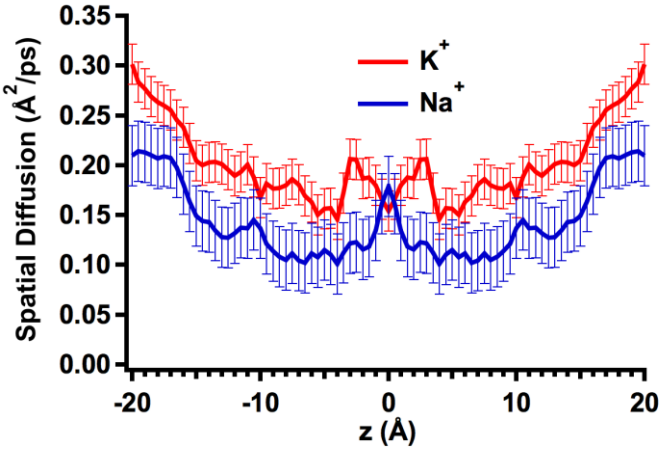


Figure S10. Spatial diffusion coefficient of the ions (see Eq. (3) in the main text).

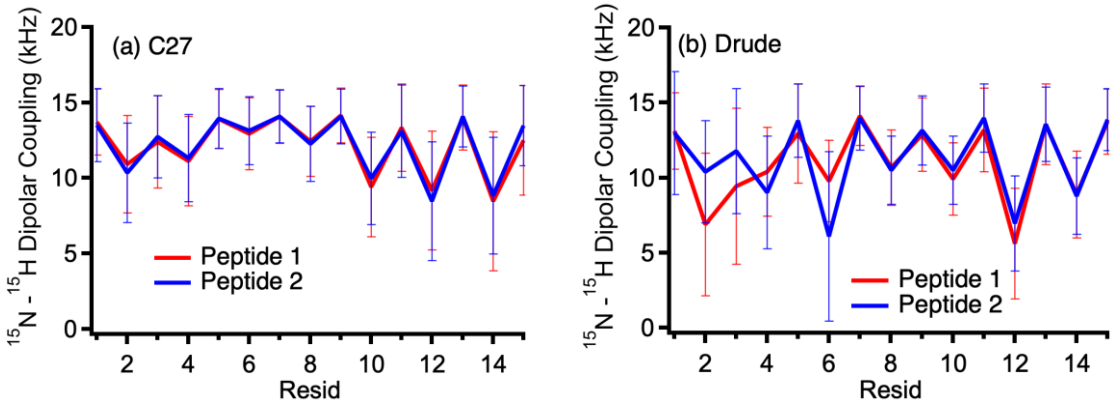


Figure S11. Dipolar coupling of the backbone using the 300-ns data. Note that these data are few kHz smaller than the previous estimates using 0.5-ns data.[13]

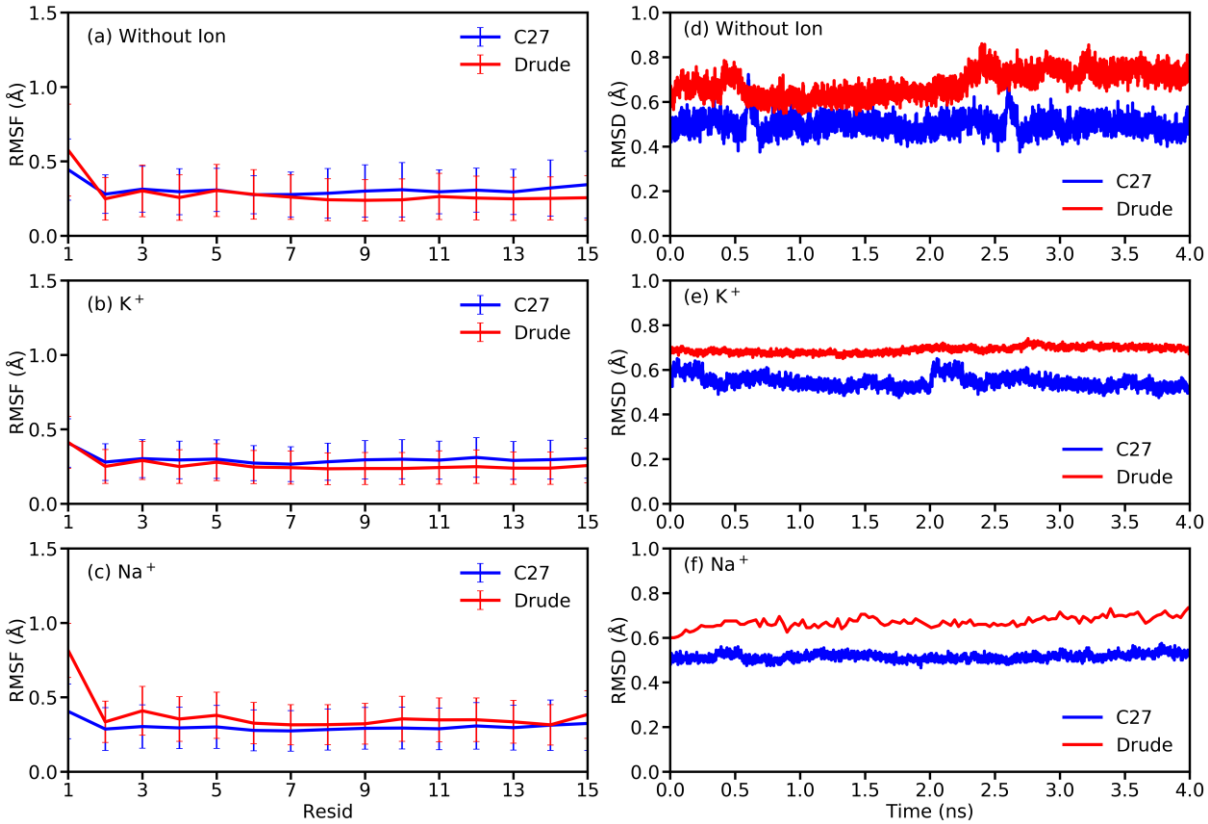


Figure S12 (a-c) Root mean square fluctuation (RMSF) per residue. For each system, we divide a trajectory into five blocks of data. For each block of data, we computed the average structure of each peptide, which was used to align each frame of the corresponding peptide. So, the error bars can be computed from the values obtained from each block of the trajectory. d-f) Root mean square deviations (RMSD) of the peptide backbones with respect to the same initial structure during the last 4 ns. Each of the peptide backbones were computed separately after alignment, and then two backbones of the two identical peptides were combined for computing the average RMSD as plotted.

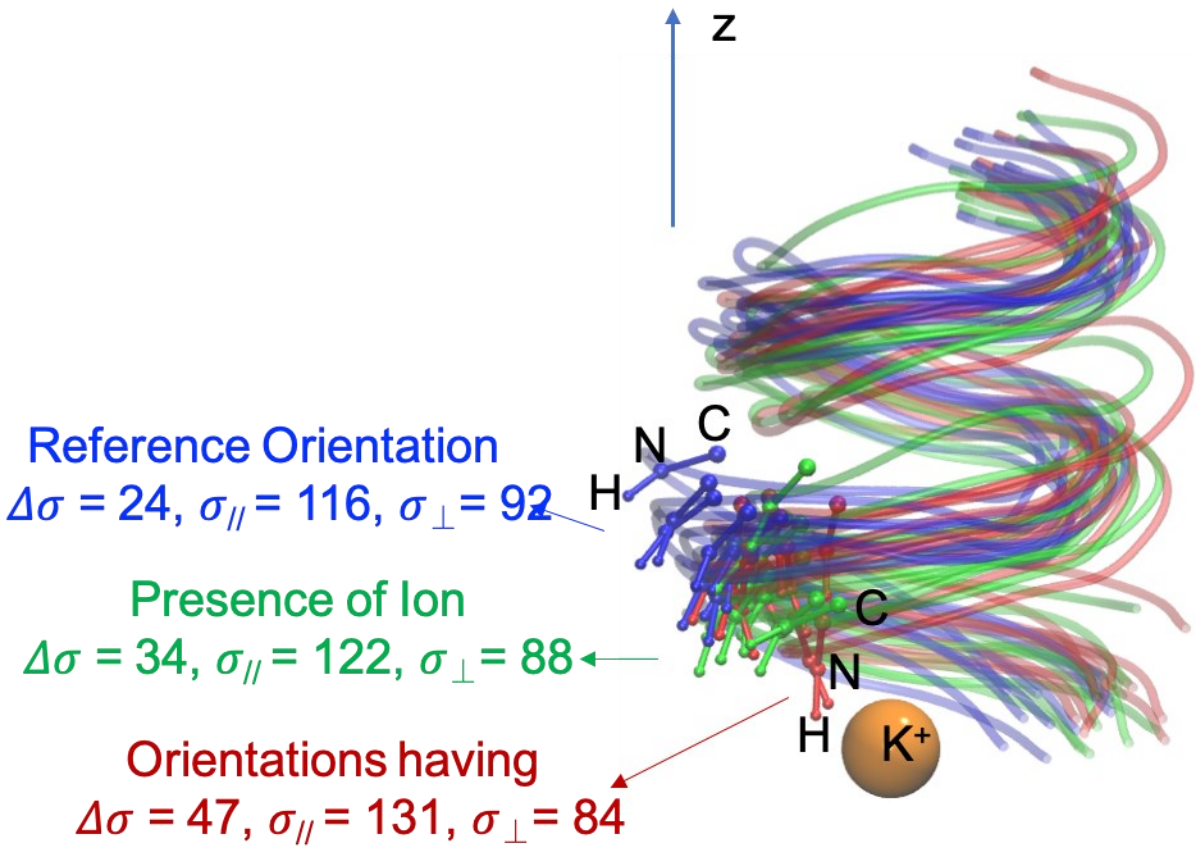


Figure S13. Snapshots with corresponding chemical shifts computed for ^{15}N-H of D-Leu12. The colored-blue structures are reference orientations. In the presence of an ion, the structures were colored in green. Another set of orientations colored in red were also plotted for comparison.

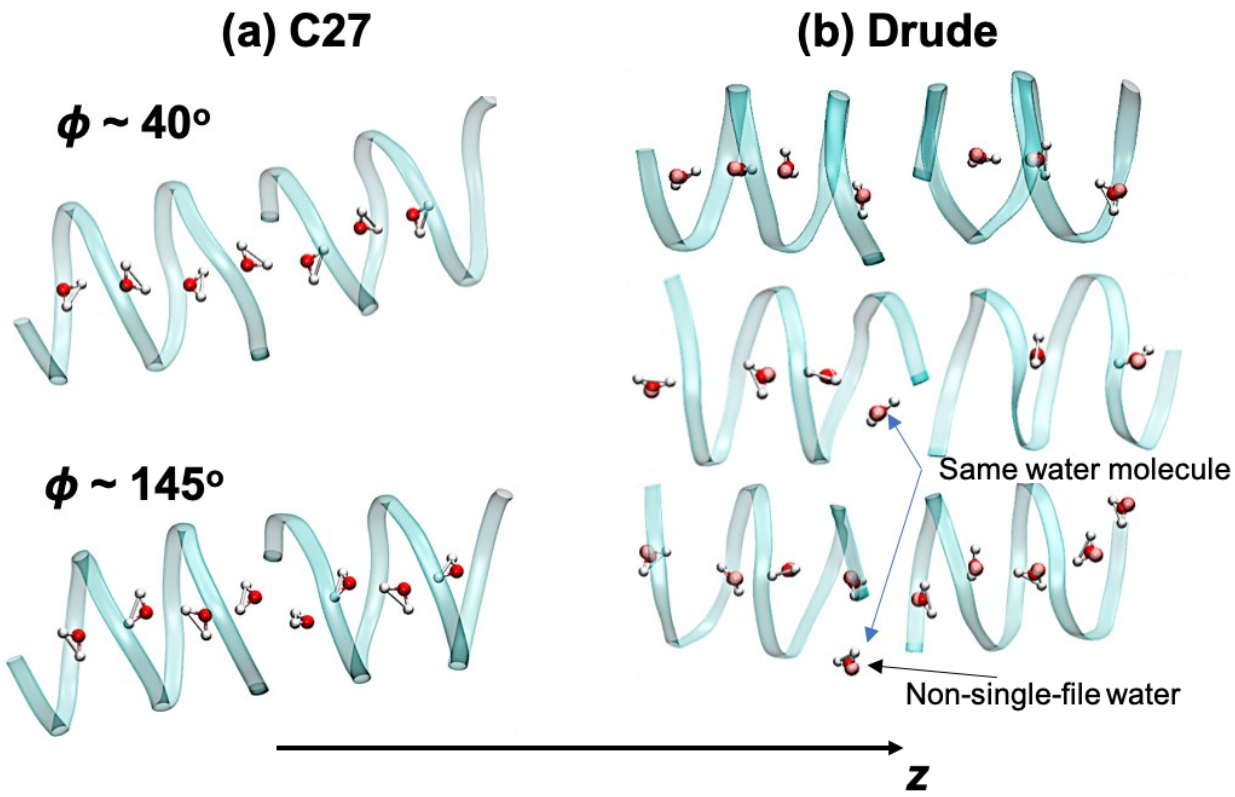


Figure S14. Snapshots of confined water molecules in the gA channel.

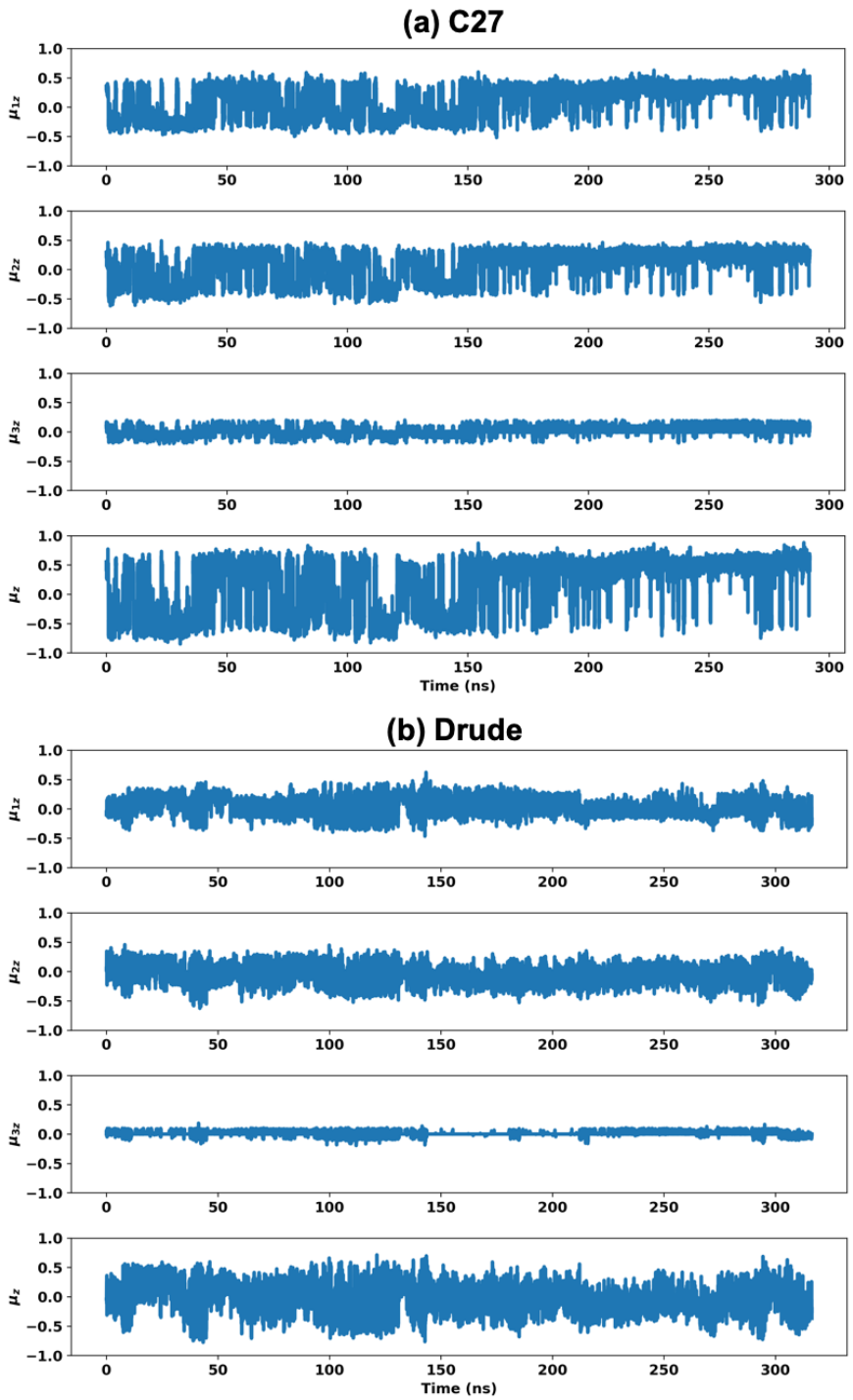


Figure 15. Dipole moment decomposition of the total dipole moment along the z-axis, μ_z. The components μ_{1z}, μ_{2z}, and μ_{3z} are the dipole moment components of water molecules lining up the peptides 1, 2, and at the middle between the two peptides, respectively. $\mu_z = \mu_{1z} + \mu_{2z} - \mu_{3z}$ since both μ_{1z} and μ_{2z} share the same water molecules at the middle of the channel.

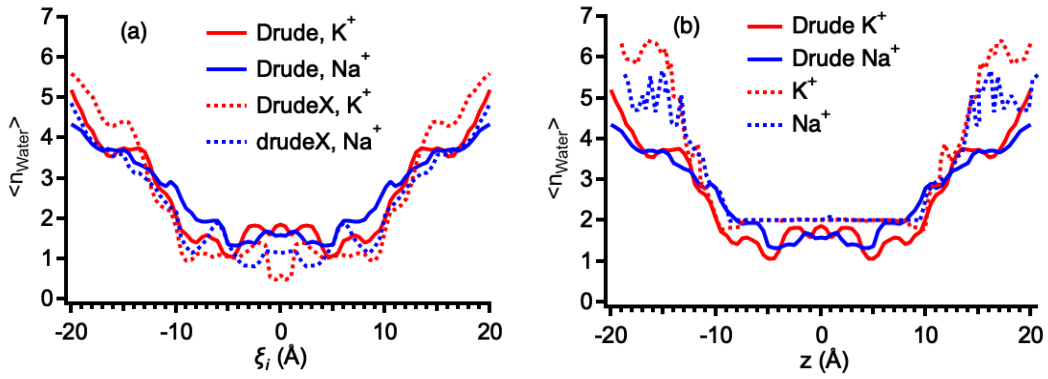


Figure S16. Average coordination number of water around the ions in (a) Drude versus DrudeX and (b) Drude versus C27.

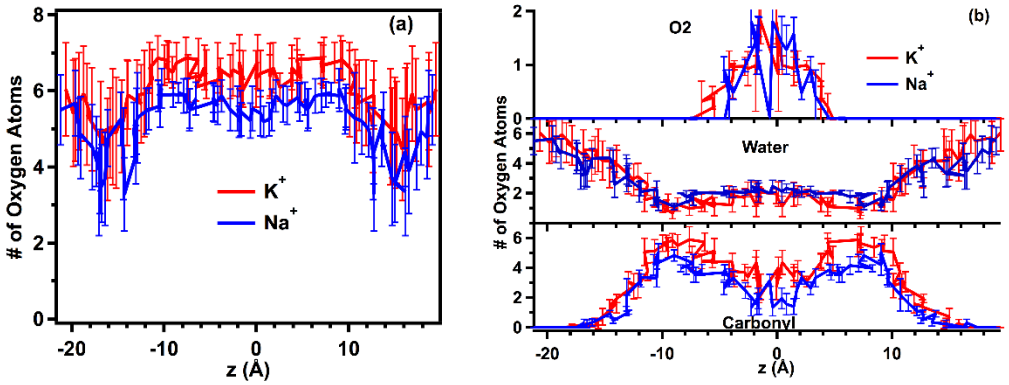


Figure S17. Coordination numbers of oxygen atoms around Na within 3.1 Å and K within 3.5 Å for (a) total of all oxygen atoms, and (b) three oxygen-atom types, in which O2 is the name of the oxygen atom in the formyl terminus. The Drude parameters were used.

Table S1. Parameters for Ethanol and Formyl termini

```
! Ethanol terminus
RESI EAM       0.000
! Corresponding patch is PRES TEAM
! having CL-HL(1-3) replaced with CA
!     HL1     O  HNT
!        \     ||     |
! HL2--CL--C---NT           H1A
!        /          |              |
!     HL3         C2-----------C1------O1--HO1
!                  /    \           |
!              H2A  H2B      H1B
!
GROUP
ATOM CL   CD33C   -0.102  ALPHA -1.894  THOLE 2.122
ATOM HL1  HDA3A    0.048
ATOM HL2  HDA3A    0.048
ATOM HL3  HDA3A    0.048
ATOM C    CD2O1A   0.497  ALPHA -0.675  THOLE 0.295
ATOM O    OD2C1A   0.000  ALPHA -0.651  THOLE 0.310
ATOM NT   ND2A2   -0.382  ALPHA -1.942  THOLE 0.250
ATOM HNT  HDP1A    0.272
GROUP
ATOM C1   CD32A   -0.060  ALPHA -1.000  THOLE 1.3
ATOM O1   OD31G    0.000  ALPHA -1.028  THOLE 1.3
ATOM HO1  HDP1C    0.360
ATOM H1A  HDA2A    0.080
ATOM H1B  HDA2A    0.080
GROUP
ATOM C2   CD32A   -0.060  ALPHA -1.200  THOLE 1.3
ATOM H2A  HDA2A    0.085
ATOM H2B  HDA2A    0.085
GROUP
ATOM LP1A LPD     -0.230
ATOM LP1B LPD     -0.230
GROUP
ATOM LPA  LPDO1   -0.312
ATOM LPB  LPDO1   -0.227

BOND  C1  C2   C1  O1   C1  H1A  C1  H1B  O1  HO1
BOND  C2  H2A  C2  H2B
BOND  HL1 CL   HL2 CL   HL3 CL
BOND  CL  C    C   NT   NT  C2
BOND  C   O    NT  HNT
BOND  O   LPA  O   LPB
BOND  O1  LP1A O1  LP1B

IMPR  C  CL N  O   N  C  C2  H

LONEPAIR relative LPA O C CL distance 0.30 angle 91.0 dihe 0.0
LONEPAIR relative LPB O C CL distance 0.30 angle 91.0 dihe 180.0
ANISOTROPY O C LPA LPB  A11  0.82322 A22  1.14332

LONEPAIR relative LP1A O1 C1 HO1 distance 0.35 angle 110.0 dihe
91.0
LONEPAIR relative LP1B O1 C1 HO1 distance 0.35 angle 110.0 dihe
269.0
ANISOTROPY O1 C1 LP1A LP1B A11 0.8108 A22 1.2162
```

```
! Parameters
! Lennard-Jones Parameters
NONBONDED
CDEH    0.00    -0.050000    1.775000
HDEH    0.00    -0.046000    0.850000

NBFIX
POTD    OD2C1A     -0.180       3.21
POTD    LPDO1      -0.07000     3.02000
POTD    LPDO2      -0.07000     3.02000
SODD    OD2C1A     -0.09000     2.88999
SODD    LPDO1      -0.06000     2.77700
SODD    LPDO2      -0.06000     2.77700
THOLE TCUT 5.0 MAXNBTHOLE 5000
POTD OD2C1A      2.19
SODD OD2C1A      1.04
ODW     OD2C1A -0.20540  3.60690
ODW     ND2A2  -0.20540  3.63690
!NBFIX parameters of DrudeX
!ODW   OD2C1A -0.20540 3.7162899
!ODW   ND2A2 -0.20540 3.5662899
```

```
! Formyl terminus
RESI NMF     0.0000   !
! Corresponding patch is PRES TCHO
! having CAT-HTC(1-3) replaced with CA.
!      HTC1      HNT
!          \        |
! HTC2--CAT-NT --C2 = O2
!           /              |
!       HTC3           H2

ATOM CAT  CD33C   -0.055  ALPHA -1.639  THOLE 2.122
ATOM HTC1 HDA3A    0.055
ATOM HTC2 HDA3A    0.055
ATOM HTC3 HDA3A    0.055
ATOM NT   ND2A2   -0.382  ALPHA -1.942  THOLE 0.250
ATOM HNT  HDP1A    0.272

ATOM C2   CDNMF    0.559  ALPHA -1.739  THOLE 1.300
ATOM O2   OD2C1A   0.000  ALPHA -0.579  THOLE 1.300
ATOM H2   HDNMF    0.049
ATOM LPAN LPDO2   -0.312
ATOM LPBN LPDO2   -0.296

BOND CAT HTC1   CAT  HTC2  CAT HTC3  CAT  NT   NT HNT
BOND NT  C2     C2   O2    C2    H2
BOND O2  LPAN   O2 LPBN

IMPR    NT  CAT  C2  HNT     C2   H2   NT   O2

LONEPAIR relative LPAN O2 C2 H2 distance 0.30 angle 91.0 dihe   0.0
LONEPAIR relative LPBN O2 C2 H2 distance 0.30 angle 91.0 dihe 180.0
ANISOTROPY O2 C2 LPAN LPBN  A11  0.82322 A22  1.14332
```

Table S2. Heights of the free-energy barriers in calculated PMFs relative to the global minima (kcal/mol)

2013 Drude FF		
	K+	**Na+**
without CMAP	12.5	12.5
With LDCMAP (Figure 3, maintext)	5.5	8

References

1. Ingolfsson, H. I.; Li, Y.; Vostrikov, V. V.; Gu, H.; Hinton, J. F.; Koeppe, R. E., 2nd; Roux, B.; Andersen, O. S., Gramicidin a Backbone and Side Chain Dynamics Evaluated by Molecular Dynamics Simulations and Nuclear Magnetic Resonance Experiments. I: Molecular Dynamics Simulations. *J Phys Chem B* **2011**, *115*, 7417-26.
2. Levitt, D. G., Electrostatic Calculations for an Ion Channel. I. Energy and Potential Profiles and Interactions between Ions. *Biophys J* **1978**, *22*, 209-19.
3. Allen, T. W.; Andersen, O. S.; Roux, B., Ion Permeation through a Narrow Channel: Using Gramicidin to Ascertain All-Atom Molecular Dynamics Potential of Mean Force Methodology and Biomolecular Force Fields. *Biophys J* **2006**, *90*, 3447-3468.
4. Allen, T. W.; Andersen, O. S.; Roux, B., Energetics of Ion Conduction through the Gramicidin Channel. *P Natl Acad Sci USA* **2004**, *101*, 117-122.
5. Huang, J.; Lemkul, J. A.; Eastman, P. K.; MacKerell, A. D., Molecular Dynamics Simulations Using the Drude Polarizable Force Field on Gpus with Openmm: Implementation, Validation, and Benchmarks. *Journal of Computational Chemistry* **2018**, *39*, 1682-1689.
6. Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D., An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem Rev* **2016**, *116*, 4983-5013.
7. Lamoureux, G.; Roux, B., Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J Chem Phys* **2003**, *119*, 3025-3039.
8. Huang, J.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *Journal of Physical Chemistry Letters* **2014**, *5*, 3144-3150.
9. Li, H.; Ngo, V.; Da Silva, M. C.; Salahub, D. R.; Callahan, K.; Roux, B.; Noskov, S. Y., Representation of Ion–Protein Interactions Using the Drude Polarizable Force-Field. *The Journal of Physical Chemistry B* **2015**.
10. Ngo, V.; da Silva, M. C.; Kubillus, M.; Li, H.; Roux, B.; Elstner, M.; Cui, Q.; Salahub, D. R.; Noskov, S. Y., Quantum Effects in Cation Interactions with First and Second Coordination Shell Ligands in Metalloproteins. *J Chem Theory Comput* **2015**, *11*, 4992-5001.
11. Mai, W.; Hu, W.; Wang, C.; Cross, T. A., Orientational Constraints as 3-Dimensional Structural Constraints from Chemical-Shift Anisotropy - the Polypeptide Backbone of Gramicidin-a in a Lipid Bilayer. *Protein Science* **1993**, *2*, 532-542.
12. Tian, F.; Cross, T. A., Cation Transport: An Example of Structural Based Selectivity. *J Mol Biol* **1999**, *285*, 1993-2003.
13. Woolf, T. B.; Roux, B., Molecular Dynamics Simulation of the Gramicidin Channel in a Phospholipid Bilayer. *Proc Natl Acad Sci U S A* **1994**, *91*, 11631-5.