Supporting Information for:

Modeling Separation of Lanthanides via Heterogeneous Ligand Binding

Kevin Leung* and Anastasia G. Ilgen

Sandia National Laboratories, Albuquerque, NM 87185, USA

Contents

S1. Effect of van der Waal's Corrections (DFT-D3) ..... page S2

S2. Water at T=0 K ..... page S5

S3. Hydration numbers ..... page S8

S4. Pair correlation Functions and Other Details ..... page S11

S5. Final Energies and Configurations ..... page S15

S1. EFFECT OF VAN DER WAAL'S CORRECTIONS (DFT-D3)

To address the possible effect of dispersion forces, not included in the main text, spot checks using DFT-D3[1] with zero-damping, based on the PBE functional, are added in two Ce-to-Lu series: $Ln^{3+}(H_2O)_8$ clusters (part of Fig. 5a of the main text) and Ln^{3+} coordinated to silica surfaces functionalized with two oxalates (Fig. 6c and Fig. 7c of the main text).

Applying D3 to configurations previously optimized using PBE functional yields significant changes in single point total energies in the simulation cells (up to ∼18 eV for the 300+ atom slabs in Fig. 6c of the main text), but minimal changes in the atomic forces. Further optimization of the atomic configurations using PBE-D3 yields only minor differences in the calculated energies, within 0.02-0.12 eV per unit cell, even for the 300+ atom cells.

The trends associated with Ln^{3+} coordinated to 8 water molecules (Fig. 5a of the main text) and Ln^{3+} at an oxalate binding site on silica surfaces (Fig. 6c of the main text) are depicted in Fig. S1a and S1b-c, respectively. The magnitude of $\Delta\Delta E$ generally increases with increasing Ln^{3+} atomic mass; the same behavior is seen with and without D3 corrections.

However, in the middle of the Ln^{3+} series, Gd and Tb now exhibit a noticeable discontinuity in the slope of $\Delta\Delta E$ in water clusters (Fig. S1a); $\Delta\Delta E$ also becomes non-monotonic there. The oxalate ligands exhibit a smaller discontinuity in the slope of $\Delta\Delta E$ at the same position (Fig. S1b). $\Delta\Delta\Delta E$ (Fig. S1d), which involves subtracting Fig. S1a from Fig. S1b, retains this non-monotonic feature at the Gd/Tb location; otherwise it is similar to Fig. 8c of the main text. Qualitatively, the slope going down the Ln series remain steeper on functionalized silica surfaces than with Ln^{3+} coordinated to 8 water molecules, with or without D3. Quantitatively, the $\Delta\Delta E$'s differ by at most 0.25 eV after adding D3 corrections.

The non-monotonicity at the Gd/Tb positions for water clusters (Fig. S1a) may reflect the "the gadolinium break," observed in some properties like ionic radii, formation constants of complexes, and solvent extraction equilibria.[2] However, non-monotonic trends in hydration free energies have not been previously predicted or observed in experiments,[4,5] to the best of our knowledge. Instead, Marcus's analysis yields a non-monotonic step at the Pm/Nd position (Fig. S1a; the rise from Yb to Lu there is a possible typo in reported data).[5] The non-monotonic behavior in Fig. S1a may or may not be an artifact of PBE-D3, and should be the subject of future studies. Finally, a recent work has shown that revPBE-D3 yields more significant disagreement with CCSD(T) predictions than revPBE (no dispersion correction)

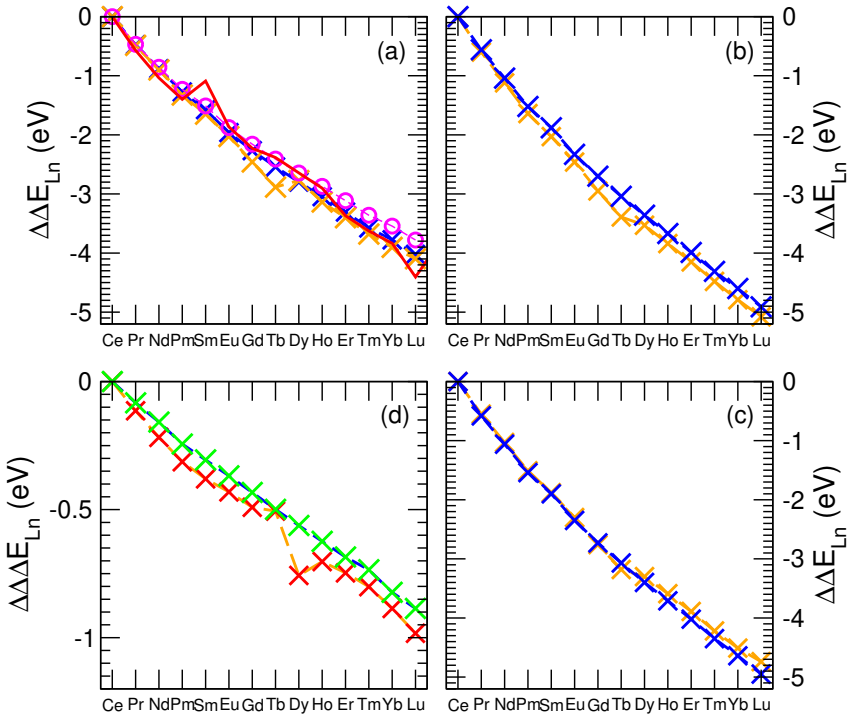

FIG. S1: Clockwise: (a) Orange and blue crosses: $\Delta\Delta E$ with/without D3 corrections for $Ln^{3+}(H_2O)_8$; purple circles: $\Delta\Delta E$ for $Ln^{3+}(H_2O)_9$; red line: experimental data.[5] (b) $\Delta\Delta E$ associated with Ln^{3+} at a binding site where two oxalate ligands are attached to β-crystobalite (110) surfaces, starting from Ce^{3+}-based AIMD configurations. Orange and blue: with/without D3 corrections. (c) $\Delta\Delta E$ associated with Ln^{3+} at a binding site where two oxalate ligands are attached to β-crystobalite (110) surfaces, starting from Lu^{3+}-based AIMD configurations Orange and blue: with/without D3 corrections. (d) $\Delta\Delta\Delta E$ for the Ce^{3+}-generated configurations (subtracting (b) from (a)). Orange line/red crosses and blue line/green crosses: with/without D3 corrections, respectively.

for interactions between water and some cations.[6] The hydration numbers calculated without the D3 correction are also in better agreement with experiments.

In summary, including D3 corrections does not qualitatively change our conclusions in the main text.

S2. WATER AT T=0 K

This section discusses the optimization of water configurations. We speculate in the main text that adding multiple layers of water beyond the first hydration shell is likely to yield amorphous ice in the limit of a large number of H_2O molecules, despite the fact that a perfect optimization module should yield ice I_h. Such trapping in local minima of the energy landscape is often observed in systems with three-dimensional hydrogen bond networks – even though molecular codes which model isolated molecular system (no periodic boundary conditions, PBC) tend to allow rotation of molecular sub-groups and permits more efficient optimization (Fig. S2a). In contrast, PBC DFT codes generally relax atomic configurations by moving all atoms separately according to forces on each of them, without regard for molecular connectivity (Fig. S2b), and are even less efficient in escaping from metastable states.

To partially demonstrate the propensity for water becoming trapped in glassy metastable configurations, we perform molecular dynamics (MD) simulations of liquid water at 1.0 g/cc density, using 256 SPC/E water molecules[7] in a cubic periodic cell of dimensions (19.71 Å)^3 and the LAMMPS code.[8] After equilibration, a $K^+(Br^-)$ complex with a constrained 2.86 Å interionic distance (gas phase optimal distance according to DFT/PBE) replaces two neighboring H_2O. This system, which mimicks the existence of multiple hydration shells around a contact ion pair, is equilibrated for 2 ns at T=300 K. At the 1- and 2-ns points, the snapshots are optimized using the conjugate gradient module of the LAMMPS code with all constraints removed. The final configurations are depicted in Figs. S2c-d. They are indeed consistent with the ion pair embedded in amorphous ice. The total energies of the two simulation cells differ by 1.34 eV, which illustrates the variability of energies in amorphous systems.

If the conjugate gradient module can predict the most favorable molecular configuration (including lattice constants), in the limit of a large number of water molecules ("N hydration shells" at large N), it should end up predicting ice Ice I_h, the most stable structure for the SPC/E water model.[9] First we optimize a simulation cell with 288 SPC/E water molecules in a 15.21×26.34×21.43 Å^3 Ice 1_h 2×4×3 supercell. Then we replace a neighboring water pair with a K^+-Br^- complex and re-optimize. To properly compare with the Fig. S2c-d configurations which have 254 water molecules, we subtract the per-H_2O energy, computed

in the pure ice I_h simulation cell, for the 32 excess 32 H_2O molecules in Fig. S2e. The corrected energy of this simulation cell is indeed more favorable than the more stable of the amorphous ice configuration quenched from liquid state snapshot by a very significant 17.12 eV.

We conclude that (1) optimizing "multi-layer water" configurations generally yield ions embedded in glassy ice; and (2) embedding the same ionic complex in ice I_h is far more favorable and would have been the most energetically favorable result if the optimization module were able to arrive at I_h geometry spontaneously.

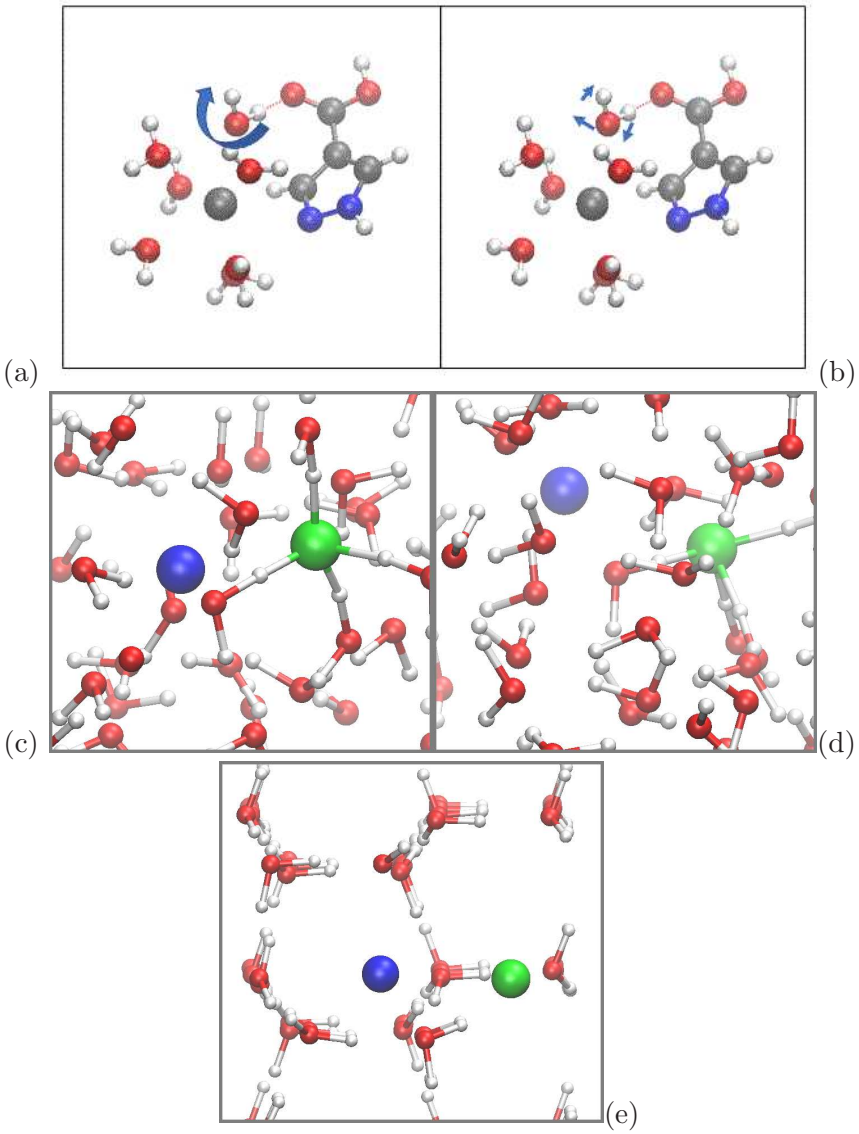

FIG. S2: Schematics of configuration optimization in (a) quantum chemistry codes (left) vs. that in (b) periodic boundary conditions DFT codes (right). Red, white, blue, grey, and green denote O, H, N, C, and Ce atoms. The arrow indicates hindered H_2O rotation motion which can break the hydrogen bond shown as a dashed line. (c)-(d) Two optimized K^+Br^-/254 H_2O configurations that have different, initial MD-generated starting points. Their relative energies are 0.00 eV and 1.34 eV, respectively. (e) Simulation cell with K^+Br^- in ice I_h, with a relative energy of -17.12 eV after correcting for number of H_2O molecules. Red, white, pink, and green, denote O, H, K, and Br atoms.

S3. HYDRATION NUMBERS

AIMD and DFT predictions of hydration (or coordination) numbers (CN) are discussed in this section. Our assertion is that predicting CN that differ from measurements by ±1 is almost inevitable, and is expected to lead to small errors in hydration free energy ($\Delta G_{\rm hyd}$) differences.

The schematic in Fig. S3a illustrates the competing factors that determine CN's. Ion-water attraction favors large CN's while water-water repulsion in the first hydration shell favores small CN's. An intermediate CN value optimizes the free energy. In our experience,[10] changing CN from the optimal value by ± 1 typically increases the free energy by only $<2\ k_{\rm B}T$ at room temperature (∼0.05 eV). As DFT functionals do not have "chemical accuracy" (i.e., accuracy within 1 kcal/mol or ∼0.04 eV) compared with experiments, ±1 errors in CN become difficult to avoid. For the same reason, the PBE functional typically predicts a CN which is smaller than that calculated using BLYP.[10] The implicit solvent approximation used also exhibits systematic error. But the corresponding relative error in $\Delta G_{\rm hyd}$ is expected to be small.

Experimentally, it is known that the early Ln^{3+} are 9-coordinated and the late Ln^{3+} are 8-coordinated.[3] If we assume that $\Delta\Delta E$s for CN=8 and CN=9 both linearly increase in magnitudes with progression down the series, the two restricted $\Delta G_{\rm hyd}$ curves must cross (Fig. S3b). Near the crossing point 8- and 9-coordination must have very similar $\Delta G_{\rm hyd}$.

Next we use the gaussian suite of programs[11] to estimate the CN=8 and CN=9 curves using the PBE functional and two implicit solvent approximations: PCM[12] (Fig. S3c) and SMD (Fig. S3d).[13] The SDD basis set is adopted.[14] Only La^{3+} and Lu^{3+} hydration free energies are explicitly calculated; the empty and full $4f$ shells of these end-point elements make the calculations much easier than for the Ln^{3+} in between. The CN=9 configuration is taken to be the reference energy for each Ln^{3+}, while the hydration free energy of CN=8 is

$$\Delta G_{\rm CN=8} = \Delta G_{\rm CN=9} - \Delta G_{\rm water}. \tag{S1}$$

An entropic correction of 0.186 eV is added to convert the gas phase translational entropy reported in gaussian for gaseous water at 1.0 atm. pressure, to that appropriate to liquid water density.

In the PCM calculations, after interpolating $\Delta G_{\rm hyd}$ for both CN=8 and CN=9 as straight

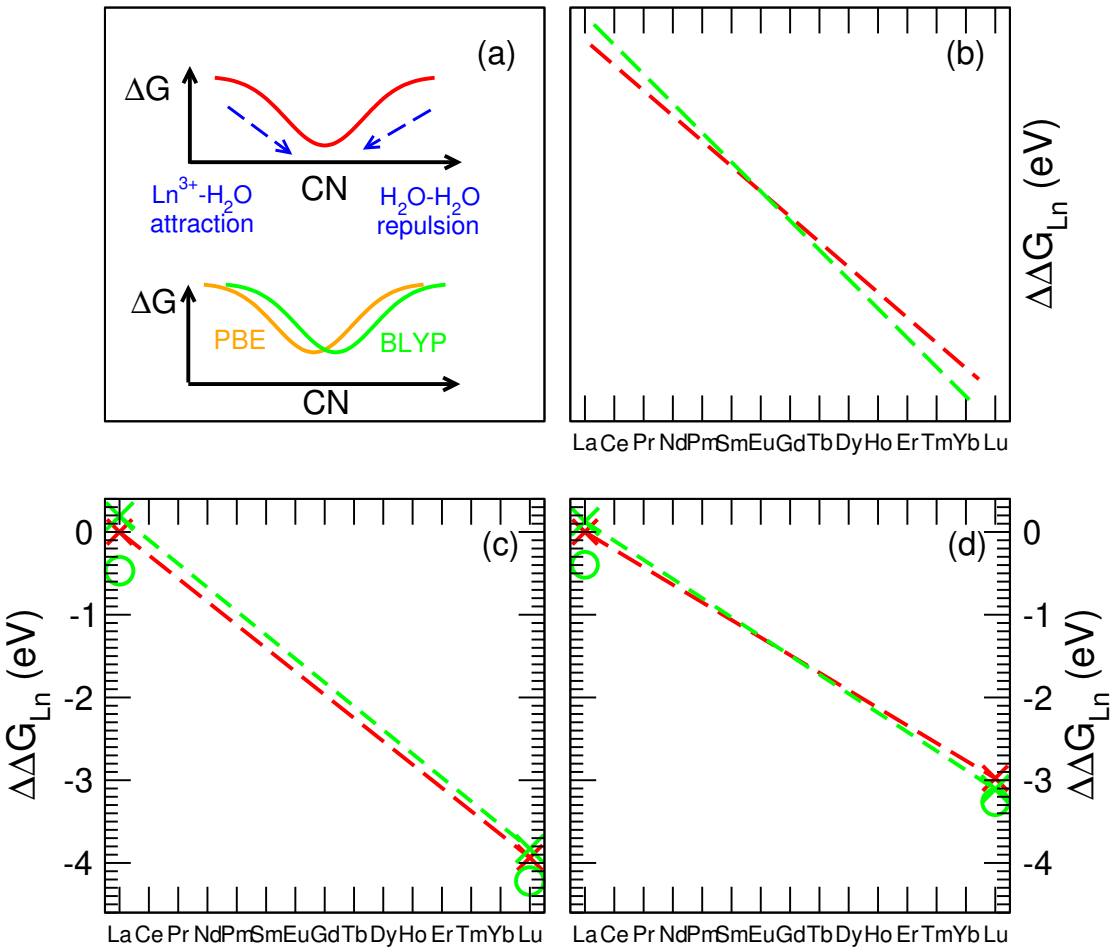

FIG. S3: (a) Schematic of competing factors that determine CN's. (b) Schematic of experimental $\Delta\Delta G_{\text{hyd}}$ at fixed CN=8 and CN=9. (c)-(d) Interpolated $\Delta\Delta G_{\text{hyd}}$ at fixed CN=8 and CN=9 for the PCM and SMD implicit solvation methods. The crosses and dashed lines represent all H_2O in the first hydration shell (both CN=8 and CN=9). The empty circles are for 9 H_2O molecules in the simulation cell, but with one H_2O in the second hydration shell.

lines, it is estimated that the entire Ln^{3+} series should have CN=9 (Fig. S3c). The SMD results appear to exhibit cross-over behavior, with 9-coordination favored for La^{3+} and 8-coordination favored for Lu^{3+}, ensuring that, for some intermediate Ln^{3+}, the ΔG_{hyd} for CN=8 and CN=9 must be similar (Fig. S3d).

But in fact, the situation is more complex. Our CN=9 coordination are initiated using $U(H_2O)_9$ configurations,[15] and optimized via the gaussian quasi-Newton module. If we use the conjugate-gradient method for optimization, in at least one case an H_2O molecule is spontaneously ejected from the first hydration shell, forming hydrogen bonds with other first hydration shell H_2O molecules. Restarting from this latter configuration, the $\Delta G_{\rm hyd}$, now for a 9-water cluster but in reality a CN=8 configuration, proves more favorable (circles in Fig. S3c-d) than both the inner sphere CN=9 and the CN=8-plus-free-water. This is true of both La^{3+} and Lu^{3+}, and of both PCM and SMD implicit solvent methods. Intepolating from the end-point elements, we argue CN=8 is favored for all Ln^{3+} for the PBE functional.

The above describes the results of static DFT calculations with implicit solvent approximations. The CN's in a number of dynamical AIMD simulations of La^{3+} in water has also been reported (Ref. 4, Table 1). Both CN=8 and CN=9 have been predicted. The EXAFS value is 9.1.[16] We have not considered La^{3+} in our AIMD/PBE simulations, but judging from our CN=7.3 and 7.4 for Dy^{3+} and Tb^{3+} in water (Table 1 of the main text)– both less than CN=8 – we expect the AIMD/PBE CN to be slightly underestimated compared with the accepted experimental values of 8-9. In the literature, the combination of B3LYP and PCM have been shown to yield Ln^{3+} CN's which cross over from 9 to 8,[3] in agreement with experiments. However, hybrid DFT functionals like B3LYP would have been prohibitively expensive to use for the DFT and AIMD simulation cells with 300+ atoms herein. Instead we have chosen the arguably less accurate PBE functional for our large scale simulations.

In conclusion, due to the finite accuracy of DFT functioals, the hydration or coordination numbers (CN) predicted for Ln^{3+} can easily deviate from the experimental value by ± 1.

S4. PAIR CORRELATION FUNCTIONS AND OTHER DETAILS

Fig. S4a-h depict the pair correlation functions ($g(r)$) between Dy^{3+} or Tb^{3+} and the oxygen sites of H_2O and ligands described in Table 1 of the main text. Integrating the first peak yields the hydration numbers listed in Table 1 of the main text.

Table S1 describes the trajectory lengths and number of water molecules associated with various simulation cells associated with Fig. 1 of the main text. In AIMD simulations with only water and Ln^{3+}, the cell dimensions are 12.42×12.42×12.42 $Å^3$ with 64 H_2O molecules; such simulation cells have $+3|e|$ net charges, where $|e|$ is the electronic charge. AIMD simulation cells with ligands are charge-neutral, have 18×12×12 $Å^3$ dimensions, and 85 H_2O molecules. AIMD simulation cells used in potential-of-mean-force simulations (Fig. 2 of the main text) are charge-neutral, have 24×12×12 $Å^3$ dimensions, and 111 H_2O molecules.

We have performed additional calculations to compare the AIMD $\Delta\Delta G$ calculations used for Fig. 1 and Fig. 3 of the main text, and the $\Delta\Delta E$ method associated with Fig.6-Fig. 8 of the main text. Fig. S5 depicts an optimized cluster taken from the trajectory associated with Fig. 1b of the main text. Most H_2O molecules except those coordinated to the Dy^{3+} and to the two sulfate anions are removed, and atomic optimization is carried out in "vacuum;" this procedure is similar to the Figs. 6-8 calculations of the main text. The lanthanide cation is 6-coordinated in Fig. S5. $\Delta\Delta E$ for switching between Tb^{3+} and Dy^{3+} for is -0.33 eV, very similar to the $\Delta\Delta G$ calculated using AIMD perturbation with the entire simulation cell filed with water (-0.30 eV). This illustrates the internal consistency of the two methods.

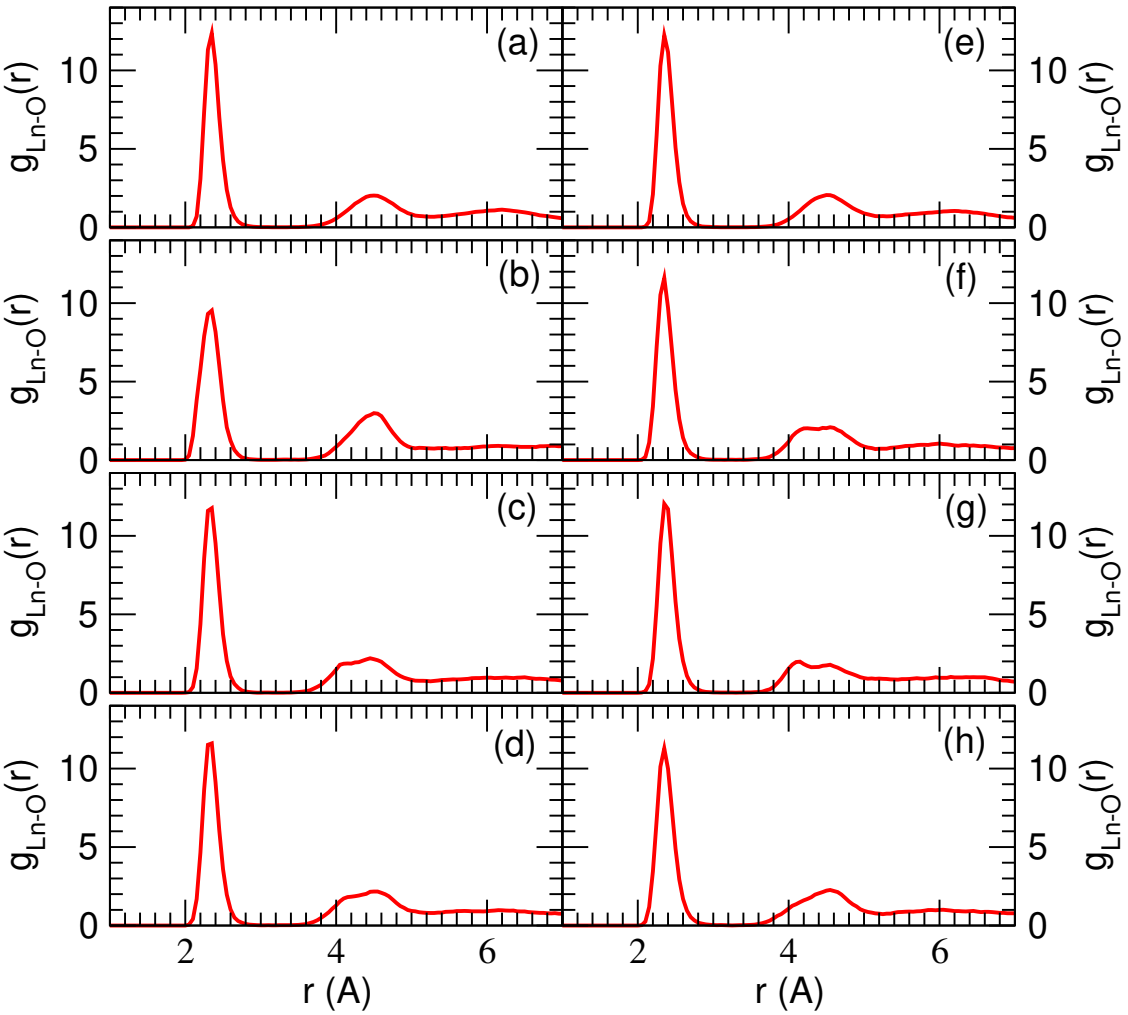

FIG. S4: $g(r)$ between Ln^{3+} and oxygen atoms in both water and ligands. (a)-(d): Dy^{3+} in pure water, coordinated with $HPO_4^{2-}/H_2PO_4^-$, with $SO_4^{2-}/H_2PO_4^-$, and with SO_4^{2-}/HSO_4^-, respectively. (e)-(h): same as (a)-(d) but with Tb^{3+} rather than Dy^{3+}.

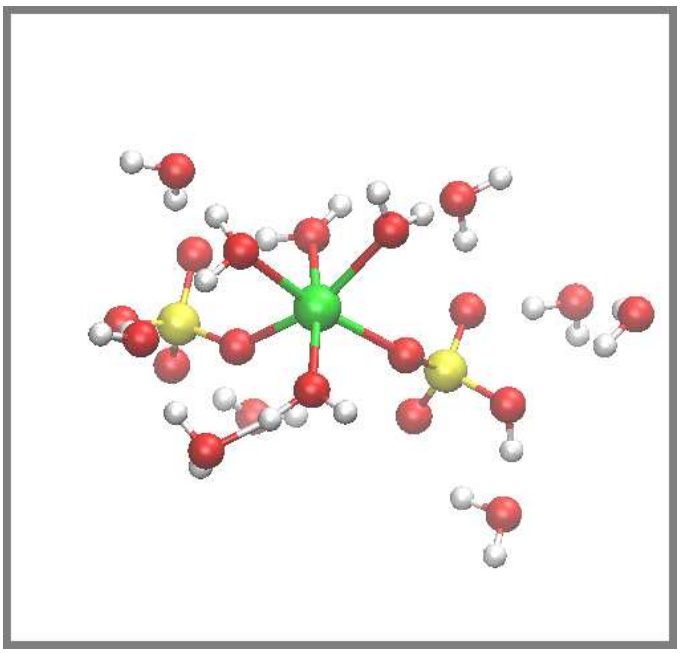

FIG. S5: Cluster configuration taken from an AIMD snapshot associated with the trajectory depicted in Fig. 1b of the main text, with a SO_4^{2-} and a $SO_4^{2-}H$ anion coordinated to Dy^{3+}. Red, white, yellow, and green denote O, H, S, and Dy atoms.

system	time	system	time
Dy^{3+} in water	263.8	Tb^{3+} in water	257.3
$Dy^{3+}(PO_4^{3-})_2$	54.5	$Tb^{3+}(PO_4^{3-})_2$	67.4
$Dy^{3+}(SO_4^{2-})_2$	74.0	$Tb^{3+}(SO_4^{2-})_2$	76.6
$Dy^{3+}(SO_4^{2-})/PO_4^{3-}$	83.4	$Tb^{3+}(SO_4^{2-})/PO_4^{3-}$	62.5
Dy^{3+} in UiO-66	18.8	Tb^{3+} in UiO-66	19.6
PMF (none)	33.2	PMF 2.80	27.2
PMF 2.95	38.7	PMF 3.25	11.7
PMF 3.50	17.2	PMF 3.80	24.1
PMF 4.15	16.4	PMF 4.50	22.5
PMF 4.85	36.6	PMF 5.15	22.8
PMF 5.50	23.3	PMF 5.85	18.3
PMF 6.20	22.5	PMF 6.55	16.4
PMF 6.90	16.7	PMF 7.55	16.7

TABLE S1: AIMD trajectory lengths in picoseconds for different systems. "PMF" refers to potential-of-mean-force simulations of the dissociation of $C_2H_5NH_2$ from Lu^{3+}; the number is the constraint distance R_o in the constraint potential $V(R) = A(R - R_o)^2$, in units of Angstrom; A is 2 eV/Å^2. Extra protons are added to SO_4^{2-} and PO_4^{3-} to ensure charge neutrality. The exceptions are simulation cells with only Ln^{3+} and water, which have net $+3|e|$ charges.

S5. FINAL ENERGIES AND CONFIGURATIONS

The $\langle \Delta E_b \rangle_a$ and $\langle \Delta E_a \rangle_a$ described in Eq. 1 of the main text and utilized in Fig. 1 and Fig. 3 of the main text are listed in Table S2. Note that Eq. 1 yields small $\Delta\Delta G$ from substracting large numbers. This is justified because the $\langle \Delta E_b \rangle_a$ and $\langle \Delta E_a \rangle_a$ are correlated and are generated using the configurations along the same AIMD trajectory with pseudopotential "a". The correlated sampling greatly reduces the statistical uncertainty.

	Fig. 1a	Fig. 1b	Fig. 1c	Fig. 1d	Fig. 3
$\langle \Delta E_{\mathrm{Dy}} \rangle_{\mathrm{Dy}}$	-967.4183	-1344.9576	-1360.7031	-1353.2604	-5316.1160
$\langle \Delta E_{\mathrm{Tb}} \rangle_{\mathrm{Dy}}$	-967.1237	-1344.6452	-1360.3903	-1352.9453	-5315.8053
$\langle \Delta E_{\mathrm{Dy}} \rangle_{\mathrm{Tb}}$	-967.6740	-1344.8968	-1360.7322	-1352.8316	-5316.4056
$\langle \Delta E_{\mathrm{Tb}} \rangle_{\mathrm{Tb}}$	-967.3918	-1344.6054	-1360.4337	-1352.5394	-5316.1042

TABLE S2: $\langle \Delta E_b \rangle_a$ associated with Fig. 1 and Fig. 3 of the main text, in eV. a and b are alternatively Dy and Tb. The energies of the bare Dy^{3+} and Tb^{3+} are subtracted.

Ln^{3+}	8 H_2O	7 H_2O/OH^-	7 H_2O/PYC	9 H_2O
Ce	-109.3056	-112.0641	-174.1694	-124.4979
Pr	-109.2928	-112.1041	-174.1572	-124.4751
Nd	-109.2714	-112.1225	-174.1318	-124.4387
Pm	-109.2574	-112.1483	-174.1096	-124.4041
Sm	-109.1943	-112.1058	-174.0384	-124.3238
Eu	-109.1895	-112.1262	-174.0188	-124.2965
Gd	-109.1822	-112.1550	-173.9953	-124.2648
Tb	-109.1445	-112.1417	-173.9498	-124.2054
Dy	-109.0987	-112.1185	-173.8968	-124.1387
Ho	-109.0508	-112.0910	-173.8401	-124.0682
Er	-109.0135	-112.0645	-173.7955	-124.0132
Tm	-108.9926	-112.0580	-173.7644	-123.9709
Yb	-108.9327	-112.0580	-173.7118	-123.8909
Lu	-108.8780	-111.9939	-173.6612	-123.8249

TABLE S3: E_{tot} associated with Fig. 4 of the main text, for $Ln^{3+}(H_2O)_8$, $Ln^{3+}(H_2O)_7(OH^-)$, $Ln^{3+}(H_2O)_7$(PYC), and $Ln^{3+}(H_2O)_9$.

The optimized energies in simulation cells (E_{tot}, Eq. 2) of the text) associated with Figs. 4a-c, Fig. 6a-f, and Fig. 9b are listed in Tables S3-S7. Individually E_{tot} are meaningless and cannot be reproduced by other DFT codes. Only the difference of E_{tot} for different Ln^{3+} in the same environment, which becomes Eq. 3 of the main text via Eq. 2 of the main text, are physical quantities.

Ln^{3+}	Fig. 6a Ce	Fig. 6a Lu	Fig. 6b Ce	Fig. 6b Lu
Ce	-2324.6132	-2324.5378	-2340.8677	-2369.8888
Pr	-2324.6901	-2324.6335	-2340.9295	-2369.9639
Nd	-2324.7414	-2324.7036	-2340.9672	-2370.0145
Pm	-2324.7967	-2324.7817	-2341.0142	-2370.0708
Sm	-2324.7758	-2324.7790	-2340.9987	-2370.0574
Eu	-2324.8236	-2324.8541	-2341.0395	-2370.1067
Gd	-2324.8678	-2324.9244	-2341.0840	-2370.1549
Tb	-2324.8743	-2324.9508	-2341.0923	-2370.1679
Dy	-2324.8701	-2324.9671	-2341.0915	-2370.1706
Ho	-2324.8641	-2324.9828	-2341.0863	-2370.1702
Er	-2324.8576	-2324.9929	-2341.0799	-2370.1672
Tm	-2324.8697	-2325.0334	-2341.0961	-2370.1898
Yb	-2324.8586	-2325.0516	-2341.0908	-2370.1978
Lu	-2324.8395	-2325.0557	-2341.0716	-2370.1881

TABLE S4: $E_{\rm tot}$ associated with Fig. 6a-b of the main text. "Ce" and "Lu" refer to configurations started with Ce-based and Lu-based AIMD runs. Recall that Ce- and Lu-configurations can have different number of water molecules, and their energies cannot be directly compared.

The optimized atomic configurations associated with Fig. 4, Fig. 6-8, and Fig. 9b of the main text are found in a separate file (finalconfig.zip). In Fig. 4 of the main text, the lattice dimensions are 16×16×16 Å^3. For systems involving silica substrates (Fig. 6-8, Fig. 9b), the lattice dimensions are 20.25×14.32×36.00 Å^3. Configurations for each of the 14 Ln^{3+} for each of the 6 silica binding sites associated with Fig. 6 of the main text are listed twice – once for the Ce-initiated AIMD configurations, and once for the Lu-initiated configurations. Note that these starting points can lead to different number of H_2O molecules present.

The AIMD simulations associated with Figs. 1-3 generate many gigabytes of data. These will be available upon request to the authors.

Ln^{3+}	Fig. 6c Ce	Fig. 6c Lu	Fig. 6d Ce	Fig. 6d Lu
Ce	-2234.6401	-2234.6068	-2259.4795	-2262.4338
Pr	-2234.7081	-2234.6905	-2259.5494	-2262.5264
Nd	-2234.7652	-2234.7510	-2259.5974	-2262.5907
Pm	-2234.8345	-2234.8205	-2259.6450	-2262.6598
Sm	-2234.8328	-2234.8182	-2259.6236	-2262.6547
Eu	-2234.8903	-2234.8788	-2259.6620	-2262.7105
Gd	-2234.9545	-2234.9465	-2259.7173	-2262.7735
Tb	-2234.9812	-2234.9733	-2259.7296	-2262.8034
Dy	-2234.9990	-2234.9928	-2259.7326	-2262.8179
Ho	-2235.0153	-2235.0082	-2259.7287	-2262.8310
Er	-2235.0314	-2235.0232	-2259.7223	-2262.8358
Tm	-2235.0680	-2235.0665	-2259.7363	-2262.8692
Yb	-2235.0878	-2235.0901	-2259.7426	-2262.8886
Lu	-2235.0984	-2235.1017	-2259.7320	-2262.8912

TABLE S5: E_{tot} associated with Fig. 6c-d of the main text. "Ce" and "Lu" refer to configurations started with Ce-based and Lu-based AIMD runs. Recall that Ce- and Lu-configurations can have different number of water molecules, and their energies cannot be directly compared.

Ln^{3+}	Fig. 6e Ce	Fig. 6e Lu	Fig. 6f Ce	Fig. 6f Lu
Ce	-2574.6407	-2533.5817	-2172.8261	-2172.5092
Pr	-2574.6522	-2533.6697	-2172.9017	-2172.6040
Nd	-2574.7033	-2533.7350	-2172.9550	-2172.6756
Pm	-2574.7672	-2533.8028	-2173.0143	-2172.7597
Sm	-2574.7628	-2533.7966	-2173.0016	-2172.7605
Eu	-2574.8244	-2533.8559	-2173.0530	-2172.8291
Gd	-2574.8753	-2533.9137	-2173.1167	-2172.9126
Tb	-2574.8892	-2533.9393	-2173.1349	-2172.9531
Dy	-2574.9361	-2533.9553	-2173.1451	-2172.9800
Ho	-2574.9480	-2533.9742	-2173.1168	-2173.0039
Er	-2574.9458	-2533.9852	-2173.1617	-2173.0212
Tm	-2574.9912	-2534.0328	-2173.2006	-2173.0670
Yb	-2575.0065	-2534.0529	-2173.2163	-2173.1050
Lu	-2575.0222	-2534.0558	-2173.2163	-2173.1254

TABLE S6: E_{tot} associated with Fig. 5c-d of the main text. "Ce" and "Lu" refer to configurations started with Ce-based and Lu-based AIMD runs. Recall that Ce- and Lu-configurations can have different number of water molecules, and their energies cannot be directly compared.

Ce	-2264.7720	Pr	-2264.8630	Nd	-2264.9381
Pm	-2265.0212	Sm	-2265.0251	Eu	-2265.0922
Gd	-2265.1687	Tb	-2265.1981	Dy	-2265.2185
Ho	-2265.2385	Er	-2265.2510	Tm	-2265.2925
Yb	-2265.2990	Lu	-2265.3234		

TABLE S7: E_{tot}, in eV, associated with Fig. 9b of the main text.

[1] S. Grimme, J. Antony, S. Ehrlich and S. Krieg, *J. Chem. Phys.* 2010, **132**, 154104.

[2] R.M. Smith and A.E. Martell, *Sci. Total Environ.* 1987, **64**, 125-147.

[3] J. Kuta and A.E. Clark, *Inorg. Chem.* 2010, **49**, 7808-7817.

[4] P. D'Angelo and R. Spezia, *Eur. J. Chem.* 2012 **18**, 11162-11178.

[5] Y. Marcus, *Biophys. Chem.* 1994, **51**, 111-127.

[6] V. Kostal, P.E. Mason, H. Matinez-Seara and P. Jungwirth, *J. Phys. Chem. Lett.* 2023, **14**, 4403.

[7] H.J.C. Berendsen, J.R. Grigera and T.P. Straatsma, *J. Phys. Chem.* 1987, **91**, 6269-6271.

[8] S.J. Plimpton, *J. Comp. Phys.* 1995, **117**, 1.

[9] C. Vega, E. Sanz and J.L.F. Abascal, *J. Chem. Phys.* 2005, **122**, 114507.

[10] K. Leung and S.B. Rempe, *J. Am. Chem. Soc.* 2004, **126**, 344-351.

[11] Frisch, M. J. *et al.*, Gaussian 09, Gaussian Inc., Wallingford, CT, 2004.

[12] G. Scalmani and M.J. Frisch, *J. Chem. Phys.* 2010, **132**, 114110.

[13] A.V. Marenich, C.J. Cramer and D.G. Truhlar, *J. Phys. Chem. B* 2009, **113**, 6378-6396.

[14] T.H. Dunning, *J. Chem. Phys.* 1977, **66**, 1382-1383.

[15] K. Leung and T.M. Nenoff, *J. Chem. Phys.* 2012, **137**, 074502.

[16] P. D'Angelo, A. Zitolo, V. Migliorati, G. Chillemi, M. Duvail, P. Vitorge, S. Abadie and R. Spezia, *Inorg. Chem.* 2011, **50**, 4572-4579.

[17] A.G. Ilgen, N. Kabengi, K. Leung, P. Ilani-Kashkouli, A.W. Knight and L. Loera, *Environ. Sci. Nano* 2021, **8**, 432-443.