



# Structural and thermodynamic properties of $\text{La}^{3+}$ in chloride-bearing hydrothermal fluids – insights from a new *ab initio*-based polarizable force field

Rajorshi Chattopadhyay<sup>a, b, \*</sup>, Sandro Jahn<sup>a, b</sup>

<sup>a</sup> Institute of Geology and Mineralogy, University of Cologne, Zùlpicher Str. 49b, Köln, 50674, Germany

<sup>b</sup> Department of Earth and Environmental Sciences, Ludwig-Maximilians University Munich, Theresienstr. 41, München, 80333, Germany

## HIGHLIGHTS

- New polarizable force field is developed for rare earth elements (REE) in hydrothermal fluids.
- Changes in relative stability of complexes are observed from subcritical to supercritical conditions.
- Effects of simulation box size on calculated thermodynamic properties are negligible at high temperatures.

## ARTICLE INFO

### Keywords:

Rare earth elements  
Hydrothermal fluids  
Speciation  
Reaction thermodynamics  
Molecular dynamics  
Enhanced sampling techniques  
Polarizable force field

## ABSTRACT

Rare earth elements (REEs) are an important group of elements both geologically and economically. The ability of hydrothermal fluids to mobilize REEs in natural, ore-forming environments depends on the chemical composition and the presence of suitable ligands such as chloride and fluoride. Here, we use molecular dynamics (MD) simulations to study the molecular structure and thermodynamic stability of  $\text{La}^{3+}$  species in Cl-bearing hydrothermal fluids. We develop a new polarizable force field for this system optimized by reference to density functional theory (DFT) calculations. The structural and thermodynamic data obtained with the new potential using the well-tempered metadynamics (WMetaD) technique reproduce experimental and *ab initio* MD simulation data well. Polarization effects are shown to be essential for predicting realistic association and stability constants, which were not achieved with simpler non-polarizable interaction potentials. Simulations with different box sizes indicate that the effects of simulation box sizes on calculated thermodynamic quantities are almost negligible at high temperatures.

## 1. Introduction

Rare earth elements (REEs) are a group of 17 (including scandium (Sc) and yttrium (Y)) silvery-white heavy metals. Compounds containing REEs have immense applications in electronics and electrical equipment, lasers, glass, semiconductors, and magnetic materials. Due to their increasing importance in economic development and the potential risk to their supply [1,2], they have been classified as critical elements by the U.S. Geological Survey (USGS).

Hydrothermal processes control the mineralogy and grade distribution of several important REE deposits, such as Gallinas Mountain in New Mexico [3] or Thor Lake in Canada [4]. The mobility of elements in hydrothermal fluids is primarily controlled by chemical composition

and speciation, which have been shown to be controlled by temperature and pressure [5]. Commonly formed REE complexes include carbonates, sulfates, fluorides, and chlorides [6,7]. Therefore, accurate knowledge of the speciation of REEs in dilute to concentrated hydrothermal brines is necessary to improve existing geochemical models and to complement existing mineral recovery and extraction processes.

Lanthanum (La), which is classified as a light rare earth element (LREE), is generally enriched by hydrothermal processes that form primary ore deposits. Several minerals such as bastnasite, monazite, zircon, and apatite are important sources of REEs in such deposits [8]. Although LREEs form stronger complexes with  $\text{F}^-$  with increasing temperature,  $\text{Cl}^-$  is considered the main transport ligand of LREEs mainly due to

\* Corresponding author at: Department of Earth and Environmental Sciences, Ludwig-Maximilians University Munich, Theresienstr. 41, München, 80333, Germany.  
Email addresses: [R.Chattopadhyay@lmu.de](mailto:R.Chattopadhyay@lmu.de) (R. Chattopadhyay), [s.jahn@lmu.de](mailto:s.jahn@lmu.de) (S. Jahn).

its much higher availability in hydrothermal fluids [9]. Several experimental and theoretical studies have been conducted to determine the properties of  $\text{La}^{3+}$  aqua and chloride complexes at room temperature (Table 1 in Guan et al. [9]). Experimental and molecular dynamics (MD) simulation studies [10–17] have shown that the hydration structure of  $\text{La}^{3+}$  at room temperature is a tricapped trigonal prism with nine water molecules. Weak chloride complexation was found by Raman spectroscopy in  $\text{LaCl}_3$  solutions with a solute concentration above 0.1 mol/kg [18], which was confirmed by *ab initio* molecular dynamics simulation (AIMD) simulations [19]. Petit et al. [20] used a combination of pair potential-based MD and AIMD simulations and found an average configuration of  $\text{La}(\text{H}_2\text{O})_8\text{Cl}_2^-$  in concentrated 14 mol/kg solutions. However, fully hydrated  $\text{La}^{3+}$  complexes were seen in 1 mol/kg solutions [17].

Although aqueous lanthanum chloride solutions have been well studied at room temperature [17–20], studies targeting hydrothermal temperatures and pressures are quite scarce in the available literature. In the absence of experiments at high temperatures, earlier studies relied on extrapolations of room temperature data using a species-based equation of state (EOS), such as the Helgeson-Kirkham-Flowers (HKF) model [21]. Haas et al. [22] fitted a set of parameters for the HKF model at 25 °C, 1 bar. Mayanovic et al. [23] performed *in situ* X-ray absorption spectroscopy (XAS) of  $\text{La}^{3+}$  in Cl-bearing solutions up to 500 °C. They predicted a gradual increase in chloride complexation from zero at 25 °C to 3.1 at 500 °C. Migdisov et al. [24] performed high-temperature solubility experiments and published a set of HKF parameters that can be used up to 300 °C. As far as computational studies are concerned, Guan et al. [9] is the only study available in the literature. They performed AIMD simulations up to 500 °C and reported speciation and association constants over a wide range of temperatures, pressures, and concentrations.

As the forces between atoms in AIMD simulations are calculated by electronic structure methods, usually within DFT, such simulations require extremely large computational resources. Parameterized force fields are much more efficient and can overcome some of the shortcomings of AIMD while still providing sufficient geochemical insights into fluid properties at hydrothermal conditions [25]. However, an accurate description of the ionic and molecular interactions requires an analytical model with sufficient chemical complexity. The explicit account of polarizability becomes particularly important when dealing with highly charged ions that have high polarization power in concentrated solutions [26]. Many different polarizable force fields exist, each having a different representation of polarizability [27–29], but none of them are suitable for the system and the thermodynamic conditions of interest here.

In this paper, we fit an *ab initio*-based polarizable ion model (PIM) for the  $\text{LaCl}_3\text{-H}_2\text{O}$  system. We use the new potential to study the speciation of  $\text{La}^{3+}$  and derive association constants for different  $\text{La-Cl}$  complexes in dilute to concentrated brines under two different hydrothermal conditions. To understand the importance of ionic polarizability, a comparative study is performed with non-polarizable pair potentials [12,30,31]. The results are also compared to available MD simulation and experimental data.

## 2. Methods

### 2.1. Polarizable ion model

The PIM describes the interactions between the charged particles of the system and has been used in previous studies [26,32,33], so we only summarize its functional form here. The total potential energy  $V_{\text{total}}$  of the system comprises four terms:

$$V_{\text{total}} = V_{\text{charge}} + V_{\text{disp}} + V_{\text{rep}} + V_{\text{pol}} \quad (1)$$

The  $V_{\text{charge}}$  term is calculated between ion pairs  $I$  and  $J$  using Coulomb's law

$$V_{\text{charge}} = \sum_{I,J} \frac{q_I q_J}{r_{IJ}} \quad (2)$$

where  $q_I, q_J$  are the charges of ions  $I$  and  $J$  and  $r_{IJ}$  is the distance between the ions. Ions are treated as rigid particles with fixed formal charges of +3 for La and −1 for Cl. The  $V_{\text{disp}}$  term includes contributions from the dipole-dipole and dipole-quadrupole dispersion interactions:

$$V_{\text{disp}} = - \sum_{I,J>I} \left( f_6^{IJ}(r_{IJ}) \frac{C_6^{IJ}}{r_{IJ}^6} + f_8^{IJ}(r_{IJ}) \frac{C_8^{IJ}}{r_{IJ}^8} \right) \quad (3)$$

represented by the dispersion parameters  $C_6^{IJ}$  and  $C_8^{IJ}$ . Short range corrections to these interactions are described using Tang-Toennies functions  $f_n^{IJ}(r_{IJ})$ , which are of the form [34]

$$f_n^{IJ}(r_{IJ}) = 1 - c_n^{IJ} e^{-b_n^{IJ} r_{IJ}} \sum_{k=0}^n \frac{(b_n^{IJ} r_{IJ})^k}{k!} \quad (4)$$

with  $n = 6$  for  $f_6^{IJ}$  and  $n = 8$  for  $f_8^{IJ}$ . Furthermore, we set  $c_6^{IJ} = c_8^{IJ} = 1$  and  $b_6^{IJ} = b_8^{IJ} = b_{\text{disp}}^{IJ}$ .

The  $V_{\text{rep}}$  term accounts for the short range repulsion and is modeled by an exponential function with two parameters  $A^{IJ}$  and  $B^{IJ}$  as

$$V_{\text{rep}} = \sum_{I,J} A^{IJ} e^{-B^{IJ} r_{IJ}} \quad (5)$$

Finally, many-body electrostatic effects are described by induced dipole ( $\vec{\mu}_I$ ) terms. These are treated as additional dynamic quantities and obtained at each MD step by minimizing the polarization energy,  $V_{\text{pol}}$

$$V_{\text{pol}} = \sum_I \frac{|\vec{\mu}_I|^2}{2\alpha^I} + \sum_{I,J} \left( \sum_{\alpha} \mu_{\alpha}^I T_{IJ}^{\alpha} q^J f_{\text{pol}}^{JJ}(r_{IJ}) - \sum_{\alpha} q^I T_{IJ}^{\alpha} \mu_{\alpha}^J f_{\text{pol}}^{IJ}(r_{IJ}) - \sum_{\alpha,\beta} \mu_{\alpha}^I T_{IJ}^{\alpha\beta} \mu_{\beta}^J \right) \quad (6)$$

where  $\alpha^I$  is the ion polarizability,  $T_{IJ}^{\alpha}$  and  $T_{IJ}^{\alpha\beta}$  are multipole interaction tensors [33,35] and  $\alpha, \beta = x, y, z$  are Cartesian coordinates. For the short-range corrections in  $V_{\text{pol}}$ , the sum in Eq. (4) has  $n = 4$ .

The solvent water molecules are described by the polarizable model developed by Dang and Chang (DC) [29], which is compatible with PIM. DC water is a rigid 4-site model, with an additional virtual site M along the symmetry axis of the molecule, which carries a negative partial charge and the induced dipole. The short range repulsion and dispersion interactions are described by Lennard-Jones (LJ) interactions on the oxygen atom only. The parameters of the DC model are summarized in Table 1. The functional form of the water potential is included in SI-1. The choice of this model is motivated by the fact that although the DC model and PIM have different functional forms of  $V_{\text{disp}} + V_{\text{rep}}$ , they have the same representations of interactions arising due to polarization (the induced dipole model). Although this model underestimates the absolute values of water density and overestimates the dielectric constant at ambient conditions, it correctly reproduces the trends in several thermophysical properties of bulk water from ambient to high temperature conditions [36].

### 2.2. Potential parameter optimization

The PIM potential parameters are determined sequentially. First, we derive the ionic polarizability ( $\alpha^I$  in Eq. (6)) of  $\text{La}^{3+}$  and  $\text{Cl}^-$  ions. For  $\text{La}^{3+}$ , DFT calculations are performed with an isolated gas phase cation. The dipole induced on the cation by an electric field perturbation is analyzed by constructing maximally localized Wannier functions (MLWF) using the method of Molina et al. [32]. Since the effect of the coordination environment on the polarizability of anions is significant [32,33],  $\alpha^I$  of  $\text{Cl}^-$  ion is determined from analogous simulations of the solvated anion with box #3 in Table 2.

To determine the other adjustable potential parameters, the  $\text{LaCl}_3\text{-H}_2\text{O}$  system is split into 3 different subsystems: crystalline  $\text{LaCl}_3$  to

**Table 1**Parameters of the DC water model with partial charges,  $\epsilon_O$  and  $\sigma_O$  are LJ parameters.

| $r_{OH}$ (Å) | $r_{OM}$ (Å) | H-O-H angle (°) | $\epsilon_O$ (kcal/mol) | $\sigma_O$ (Å) | $q_H$ (e) | $q_O$ (e) | $\alpha^M$ (Å <sup>3</sup> ) |
|--------------|--------------|-----------------|-------------------------|----------------|-----------|-----------|------------------------------|
| 0.9752       | 0.215        | 104.52          | 0.1825                  | 3.234          | 0.5190    | -1.038    | 1.444                        |

**Table 2**Details of simulation boxes ( $\rho$  – density in g/cm<sup>3</sup>,  $m_{Cl^-}$  – molality of Cl<sup>-</sup> ions in mol/kg,  $m_{La^{3+}}$  – molality of La<sup>3+</sup> in mol/kg,  $T$  – temperature in K).

| Box | System                                                       | Edge lengths (Å)     | $\rho$ | $m_{Cl^-}$ | $m_{La^{3+}}$ | $T$          |
|-----|--------------------------------------------------------------|----------------------|--------|------------|---------------|--------------|
| #1  | 32 LaCl <sub>3</sub> (crystal)                               | 26.13 × 15.09 × 8.74 | 3.77   | –          | –             | 298 K, 773 K |
| #2  | 84 H <sub>2</sub> O, 1 La <sup>3+</sup>                      | 14.28                | 0.87   | –          | 0.66          | 773 K        |
| #3  | 84 H <sub>2</sub> O, 1 Cl <sup>-</sup>                       | 14.28                | 0.87   | 0.66       | –             | 773 K        |
| #4  | 227 H <sub>2</sub> O, 1 La <sup>3+</sup> , 3 Cl <sup>-</sup> | 20.0                 | 0.90   | 0.73       | 0.24          | 473 K, 773 K |
| #5  | 55 H <sub>2</sub> O, 1 La <sup>3+</sup> , 3 Cl <sup>-</sup>  | 12.57                | 1.03   | 3.0        | 1.0           | 773 K        |
| #6  | 55 H <sub>2</sub> O, 1 La <sup>3+</sup> , 3 Cl <sup>-</sup>  | 12.24                | 1.11   | 3.0        | 1.0           | 473 K        |
| #7  | 339 H <sub>2</sub> O, 1 La <sup>3+</sup> , 3 Cl <sup>-</sup> | 22.0                 | 0.98   | 0.50       | 0.17          | 473 K        |

determine the La-La, La-Cl and Cl-Cl potential parameters, hydrated La<sup>3+</sup> in liquid water to determine the La-O potential parameters and hydrated Cl<sup>-</sup> in liquid water to determine the Cl-O potential parameters, both at 773 K, 5 kbar. This pressure and temperature are representative of supercritical hydrothermal fluids at mid-crustal conditions. In addition, they allow for a direct comparison of our results with the AIMD data by Guan et al. [9], and the DC water model has already been tested at the corresponding water density [36]. The simulation boxes used for the simulations of the individual subsystems are listed in Table 2 (boxes #1 to #3). AIMD simulations are performed for each of the subsystems with the PBE exchange and correlation (XC) functional [37] to generate structural configurations that are used as a training set for the subsequent dipole and force fitting. The PBE XC functional gives reliable results for pure water [38,39], lanthanides [40,41] as well as aqueous solutions at high temperatures [42]. For each configuration, MLWFs and atomic forces are computed. MLWFs are used to calculate the dispersion interaction terms ( $C_6^{IJ}$  and  $C_8^{IJ}$  in Eq. (3)) and the ionic or molecular dipole moments. More technical details on the DFT-related and AIMD simulations and derived properties are given in SI-1.

In a next step, the DFT-derived dipole moments are used to determine the short-range damping parameters of  $V_{pol}$ ,  $b_{pol}^{IJ}$  and  $c_{pol}^{IJ}$ , using a dipole fitting approach [26,43]. They are optimized to minimize the difference between dipole moments calculated from PIM and those obtained from DFT for the different training configurations:

$$\chi_\mu^2 = \frac{1}{N_{conf} N_{atom}} \sum_{conf} \sum_{atom} \frac{|\bar{\mu}^{PIM} - \bar{\mu}^{DFT}|^2}{|\bar{\mu}^{DFT}|^2} \quad (7)$$

where  $N_{conf}$  is the number of configurations in training set,  $N_{atom}$  is the number of atoms in each snapshot,  $\bar{\mu}^{PIM}$  are the dipole moments obtained from PIM and  $\bar{\mu}^{DFT}$  are the reference dipole moments obtained from MLWFs.  $\chi^2$  is the relative mean square error (MSE).

Finally, the parameters of the short-range repulsion term ( $V_{rep}$ ) and  $b_{disp}^{IJ}$  in  $V_{disp}$  are optimized following a similar procedure [26,43] by fitting the forces calculated from PIM to those obtained from DFT:

$$\chi_F^2 = \frac{1}{N_{conf} N_{atom}} \sum_{conf} \sum_{atom} \frac{|\bar{F}^{PIM} - \bar{F}^{DFT}|^2}{|\bar{F}^{DFT}|^2} \quad (8)$$

$\bar{F}^{PIM}$  are the atomic forces obtained from PIM and  $\bar{F}^{DFT}$  are the reference atomic forces obtained from DFT.  $b_{disp}^{IJ}$  are set equal to  $B^{IJ}$  for La-La, La-Cl and Cl-Cl interactions because they are found to be very close to each other. For Cl-O interactions,  $b_{disp}^{IJ}$  is taken as 0, consistent with Tazi et al. [26].

### 2.3. MD simulations

The fitted PIM potential is used to investigate the speciation and association constants of La-Cl complexes under two different conditions representing supercritical (773 K, 5 kbar) and subcritical (473 K, 400 bar) fluids at similar solvent densities in both cases. MD simulations are performed in the  $NVT$  ensemble, that is, with a constant number of particles,  $N$ , volume,  $V$  and temperature,  $T$ , and with a time step of 0.5 fs. As a dedicated EOS for aqueous La-Cl systems at high temperatures and high pressures is not available, we use the correlation formulae for the H<sub>2</sub>O-NaCl system from Driesner et al. [44] for the same molality of Cl<sup>-</sup> ions to estimate the dimensions of the cubic simulation boxes, which are compiled in Table 2 (boxes #4 to #7). The boxes for MD simulations are set up from a random configuration of water molecules followed by replacement of some of the water molecules with a geometry optimized LaCl<sub>3</sub> molecule (boxes #4 to #7) or a single water molecule by single La<sup>3+</sup> or Cl<sup>-</sup> ions (boxes #2 and #3). These simulation boxes are then used for the equilibration and production simulations under the conditions mentioned in Table 2. We also calculate stress tensors from our production simulations at both studied conditions. An evaluation of stress tensors obtained from PIM is shown in SI-2. All simulations are performed with the CP2K code [45].

To compare the fitted PIM potential against available non-polarizable force fields, MD simulations are carried out with the same boxes using 3 different pair potentials, referred to as N-POL-1-3 in the following. Details of the N-POL potentials are described in Tables S2 and S3. These potentials have been shown to accurately predict the structural and thermodynamic properties of aqueous La<sup>3+</sup> at ambient conditions [12,30,31].

Structural analysis of the MD trajectories is performed by computing partial radial distribution functions,  $g_{IJ}(r)$ , for different pairs of species  $I$  and  $J$  using the TRAVIS code [46]

$$g_{IJ}(r) = \left\langle \frac{dn_r^{IJ}}{4\pi r^2 \rho_J dr} \right\rangle \quad (9)$$

where  $dn_r^{IJ}$  is a function that calculates the number of particles of type  $J$  within a shell of thickness  $dr$  at a distance  $r$  from a particle of type  $I$ .  $\rho_J$  is the number density ( $N_J/V$ ) of species  $J$ . The angular brackets denote ensemble average over the entire trajectory. Coordination numbers (CN) are determined by counting the number of particles up to the first minimum of the respective  $g_{IJ}(r)$ .

In order to calculate the association constants of different La-Cl complexes, well-tempered metadynamics (WMetaD) [47,48] simulations are performed in the  $NVT$  ensemble for ~6 ns. Hills of height 1 kJ/mol and width 0.02 are deposited every 10<sup>2</sup> timesteps to achieve a good sampling of the associated and dissociated states within a reasonable simulation time. To allow a direct comparison with the AIMD data from Guan

**Table 3**  
Ionic polarizabilities obtained from DFT calculations.

| Ion              | Polarizability ( $\text{\AA}^3$ ) |
|------------------|-----------------------------------|
| $\text{Cl}^-$    | 3.05                              |
| $\text{La}^{3+}$ | 1.12                              |

et al. [9] under their studied conditions, boxes #5 and #6 in Table 2 are used for these simulations. WMetaD simulations are also performed with boxes #4 and #7 to study the effects of box size on the association constant. These simulations are conducted using the PLUMED package together with the CP2K code [45]. Association constants are obtained from the free energy surface (FES) using the method described by Chialvo et al. [49]. More details of this method are provided in SI-1. The integration cutoff of the radial distribution function,  $r_c$  in Equation S11 is taken as half of the edge length of the respective simulation boxes, i.e.,  $\sim 6.10$ – $6.30$   $\text{\AA}$  for boxes #5 and #6 and  $\sim 10.00$ – $10.80$   $\text{\AA}$  for boxes #4 and #7 (see Table 2).  $r_i$  in the same equation is set to  $\sim 2.7$ – $2.8$   $\text{\AA}$  in all biased simulations corresponding to the distance of the first minimum in FES.

### 3. Results

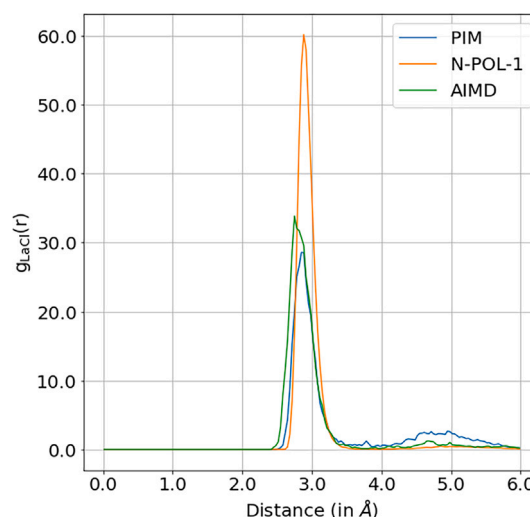
#### 3.1. PIM parameters

The ionic polarizability values of the  $\text{La}^{3+}$  and  $\text{Cl}^-$  ions obtained from the DFT simulations are listed in Table 3. These values agree well with previously calculated polarizability values for these ions [32,50,51]. The other PIM parameters obtained with the methods described above are given in Table 4. The MSEs obtained for forces and dipole moments of the different subsystems are listed in Table S4. Plots comparing fitted and computed dipole moments and forces for different subsystems are shown in Figures S3–S8.

#### 3.2. Speciation and structure

For a brief evaluation of the PIM at room temperature, readers may refer to the section titled “Brief evaluation of polarizable ion model (PIM) at room temperature” in SI-2. In the remaining paper we discuss the results from supercritical and subcritical conditions.

To test the performance of the new PIM potential, we first compare structural data from PIM and AIMD simulations using the boxes used as training sets. The calculated bond lengths and coordination numbers from the PIM agree well with the AIMD data as shown in Table 5. For La-Cl, the AIMD data from Guan et al. [9] are used as a reference. Note that the latter study was performed with a different DFT exchange-correlation functional.



**Fig. 1.** Comparison of the radial distribution functions  $g_{\text{LaCl}}(r)$  obtained from PIM, N-POL-1 and AIMD at 773 K and density of  $1.03 \text{ g/cm}^3$  (box #5).

Next, La-Cl radial distribution functions,  $g_{\text{LaCl}}(r)$ , from PIM, N-POL-1, and AIMD are compared in Fig. 1. While the PIM data reproduce the AIMD results reasonably well, N-POL-1 predicts strong La-Cl attraction that allows only negligible ion exchange throughout the simulation. This point is illustrated and discussed in greater detail later. La-O and Cl-O pair correlation functions from AIMD and PIM are shown in Figure S9. Both models predict similar sizes and overall structures of the hydration shells.

The predictions of ion pair distances and coordination numbers from PIM and all N-POL MD models under both conditions are listed in Table 6 and compared to AIMD data.  $g_{\text{LaCl}}(r)$  and  $g_{\text{LaO}}(r)$  obtained at 773 K and 473 K from box #4 with both models are shown in Figures S10 and S11 respectively.

For both conditions, N-POL-1 predicts longer La-Cl and La-O distances compared to PIM. In contrast, the N-POL-2 and N-POL-3 models predict shorter La-Cl and La-O distances compared to PIM. The number of nearest neighbors of the  $\text{La}^{3+}$  ion is  $\sim 7$  for PIM, while it is  $\sim 8$  for all N-POL MD models at supercritical conditions. At subcritical conditions, the total CN is  $\sim 8$  for PIM, while it is  $\sim 8$ – $9$  for the different N-POL MD models. In contrast to PIM, N-POL MD shows almost no La-Cl dissociation, resulting in high CN at both conditions. This is reflected in sharp high peaks in  $g_{\text{LaCl}}(r)$  at  $\sim 2.7$ – $2.8$   $\text{\AA}$  in Figures S10 and S11 (left). At

**Table 4**  
Fitted parameters of PIM.

| Subsystem                | Ion pair ( $I, J$ ) | $A^{IJ}$ (Ha) | $B^{IJ}$ ( $\text{\AA}^{-1}$ ) | $C_6^{IJ}$ ( $\text{Ha}\text{\AA}^6$ ) | $C_8^{IJ}$ ( $\text{Ha}\text{\AA}^8$ ) | $b_{\text{disp}}^{IJ}$ ( $\text{\AA}^{-1}$ ) | $b_{\text{pol}}^{IJ}$ ( $\text{\AA}^{-1}$ ) | $c_{\text{pol}}^{IJ}$ |
|--------------------------|---------------------|---------------|--------------------------------|----------------------------------------|----------------------------------------|----------------------------------------------|---------------------------------------------|-----------------------|
| $\text{LaCl}_3$ crystal  | La-La               | 1.5643        | 5.3542                         | 2.8889                                 | 5.7830                                 | 5.3542                                       | –                                           | –                     |
| $\text{LaCl}_3$ crystal  | La-Cl               | 194.6102      | 2.9768                         | 4.4840                                 | 10.1427                                | 2.9768                                       | 2.4498                                      | 1.0                   |
| $\text{LaCl}_3$ crystal  | Cl-Cl               | 18.6488       | 2.4573                         | 7.1163                                 | 17.9533                                | 2.4573                                       | –                                           | –                     |
| aqueous $\text{La}^{3+}$ | La-O                | 132.1557      | 3.3156                         | 1.4961                                 | 2.7856                                 | 2.8447                                       | 3.7756                                      | 1.0242                |
| aqueous $\text{Cl}^-$    | Cl-O                | 973.7683      | 3.8222                         | 2.2225                                 | 4.3390                                 | 0                                            | 3.0998                                      | –0.7204               |
| aqueous $\text{Cl}^-$    | Cl-H                | –             | –                              | –                                      | –                                      | –                                            | 5.7893                                      | 3.8516                |

**Table 5**  
Comparison of bond lengths and coordination numbers from PIM and AIMD simulations at 773 K and density of  $1.03 \text{ g/cm}^3$  (box #5).

| Ion pair | Box | Bond length from PIM ( $\text{\AA}$ ) | Bond length from AIMD ( $\text{\AA}$ ) | CN from PIM | CN from AIMD |
|----------|-----|---------------------------------------|----------------------------------------|-------------|--------------|
| La-O     | #2  | 2.52                                  | 2.55                                   | 8.0         | 8.5(5)       |
| Cl-O     | #3  | 3.18                                  | 3.08                                   | 10.5(5)     | 10.5(5)      |
| La-Cl    | #5  | 2.85                                  | 2.78 [9]                               | 1.5(5)      | 2.0 [9]      |



**Table 6**

Nearest neighbour distances and running coordination numbers for different ion pairs obtained from simulations of box #4 (for PIM and N-POL MD models) at 773 K and 473 K.

| Temperature | Model     | Ion pair | 1st maximum (Å) | 1st minimum (Å) | CN      | 2nd maximum (Å) | 2nd minimum (Å) | CN     |
|-------------|-----------|----------|-----------------|-----------------|---------|-----------------|-----------------|--------|
| 773 K       | PIM       | La-Cl    | 2.83            | 4.0             | 1.5(2)  | 4.84            | 6.0             | 2.2(3) |
|             |           | La-O     | 2.52            | 3.5             | 6.0(1)  | 5.00            | 6.2             | 27(3)  |
|             | AIMD [9]  | La-Cl    | 2.77            | –               | 2.0     | –               | –               | –      |
|             |           | La-O     | 2.58            | –               | 5.8     | –               | –               | –      |
|             | EXAFS [9] | La-Cl    | 2.90–3.02       | –               | 1.4     | –               | –               | –      |
|             |           | La-O     | 2.57–2.59       | –               | 5.0     | –               | –               | –      |
|             | N-POL-1   | La-Cl    | 2.88            | 4.0             | 2.3(2)  | –               | –               | –      |
|             |           | La-O     | 2.58            | 3.5             | 6.2(3)  | 5.15            | 6.7             | 35(5)  |
|             | N-POL-2   | La-Cl    | 2.71            | 4.0             | 2.92(2) | –               | –               | –      |
|             |           | La-O     | 2.50            | 3.5             | 5.0(1)  | 5.13            | 6.5             | 31(5)  |
|             | N-POL-3   | La-Cl    | 2.71            | 4.0             | 2.7(2)  | –               | –               | –      |
|             |           | La-O     | 2.50            | 3.5             | 5.0(3)  | 5.15            | 6.5             | 32(5)  |
| 473 K       | PIM       | La-Cl    | 2.87            | 4.0             | 0.1(2)  | 4.88            | 6.0             | 1.4(4) |
|             |           | La-O     | 2.53            | 3.5             | 8.0(3)  | 4.82            | 6.0             | 28(5)  |
|             | AIMD [9]  | La-Cl    | 2.86            | –               | 1.0     | –               | –               | –      |
|             |           | La-O     | 2.55            | –               | 7.2     | –               | –               | –      |
|             | EXAFS [9] | La-Cl    | 2.90–3.02       | –               | 0.5     | –               | –               | –      |
|             |           | La-O     | 2.57–2.59       | –               | 7.5     | –               | –               | –      |
|             | N-POL-1   | La-Cl    | 2.90            | 4.0             | 2.2(2)  | –               | –               | –      |
|             |           | La-O     | 2.58            | 3.5             | 7.0(2)  | 4.78            | 6.5             | 33(3)  |
|             | N-POL-2   | La-Cl    | 2.72            | 4.0             | 2.0(2)  | –               | –               | –      |
|             |           | La-O     | 2.50            | 3.5             | 6.5(5)  | 4.72            | 6.3             | 30(2)  |
|             | N-POL-3   | La-Cl    | 2.73            | 4.0             | 3.0(0)  | –               | –               | –      |
|             |           | La-O     | 2.50            | 3.5             | 5.0(3)  | 5.00            | 6.3             | 29(5)  |

473 K, PIM predicts an almost completely hydrated  $\text{La}^{3+}$  ion with  $\sim 8$  water molecules in the first hydration shell. This is seen in Figure S11 (left) where  $g_{\text{LaCl}}(r)$  from PIM is characterized by a narrow low peak at  $\sim 2.8$  Å and a broad peak at  $\sim 5.0$  Å. These two peaks indicate the formation of two different types of ion pairs, contact ion pairs (CIPs) and solvent-shared ion pairs (SShIPs).

### 3.3. Association constants

Reaction constants and energy barriers from the biased simulations using WMetaD associated with different La-Cl association reactions are presented in Table 7. Potential of mean force (PMF) data from Guan et al. [9] is used to recalculate  $\log K_{\text{eq}}^{\infty}$  for the AIMD reference in Table 7 using the method described in SI-1. Making these corrections leads to a maximum difference of about one log unit in  $\log K_{\text{eq}}^{\infty}$  of the studied reactions. To make a credible comparison, PIM association constants obtained from the small simulation boxes are compared with the AIMD data. The effects of box size are discussed subsequently. La-Cl complexes show step-wise association reactions as



The reactions for successive values of  $n$  starting from  $n = 1$  in Eq. (10) are referred to as reactions 1, 2 and 3 respectively in the remaining paper.

Energy barriers for dissociation are approximated by  $\Delta G$  between the first energy maximum and the first energy minimum in the FES, whereas relative stabilities of different metastable states are determined as free energy differences between them as shown in Fig. 2 (right). Note that since the MD simulations are performed in the  $NVT$  ensemble, the free energies obtained with the biased simulations are the Helmholtz free energies. However, as shown in Schulze et al. [52], the volume change due to the addition of an associated ion pair or a dissociated ion pair is negligibly small at solvent density of  $\sim 0.8$  g/cm<sup>3</sup>. These effects can therefore be neglected and the Helmholtz free energies can be considered to be approximately equal to the Gibbs free energies.

WMetaD is performed with the La-Cl interionic distance of the dissociating  $\text{Cl}^-$  as the collective variable (CV). The other  $\text{Cl}^-$  ligands are prevented from dissociating or associating by restraining them to a fixed distance either close to or far away from the  $\text{La}^{3+}$  ion. For example, in case of reaction 2, one  $\text{Cl}^-$  ligand is restrained at  $\sim 3$  Å, the other is allowed to associate while the third one (to be associated in the next step)

is restrained at  $\sim 8$ – $10$  Å. This is illustrated in Fig. 2 (left, inset). In dilute solutions ions generally exist in equilibrium as contact ion pairs (CIPs – ion pairs where anion and cation are in close contact), solvent-shared ion pairs (SShIPs – ion pairs where anion and cation share hydration water molecules) and solvent-separated ion pairs (SSIPs – ion pairs where anion and cation form their individual hydration shells) as shown in Fig. 3. These can be distinguished from each other using the La-Cl interionic distance, thereby allowing proper sampling of all the metastable states.

Fig. 2 shows the FES of reaction 2 obtained from PIM and N-POL-1 models at supercritical and subcritical conditions. FES for the other reactions are shown in Figures S12 and S13. All FES primarily show two metastable states indicated by energy minima at  $\sim 3$  Å and at  $\sim 5$  Å (marked by 1 and 2 in Fig. 2 right). State 1 is attributed to the CIPs whereas state 2 is populated by SShIPs and these metastable states are separated by a high-energy transition state (TS). The relative stabilities of CIPs and SShIPs show a change from subcritical to supercritical conditions with PIM. However, no such change in relative stability is observed with N-POL MD.  $\Delta G_{2-1}$  obtained from the different models at 773 K and 473 K is shown in Figure S14.

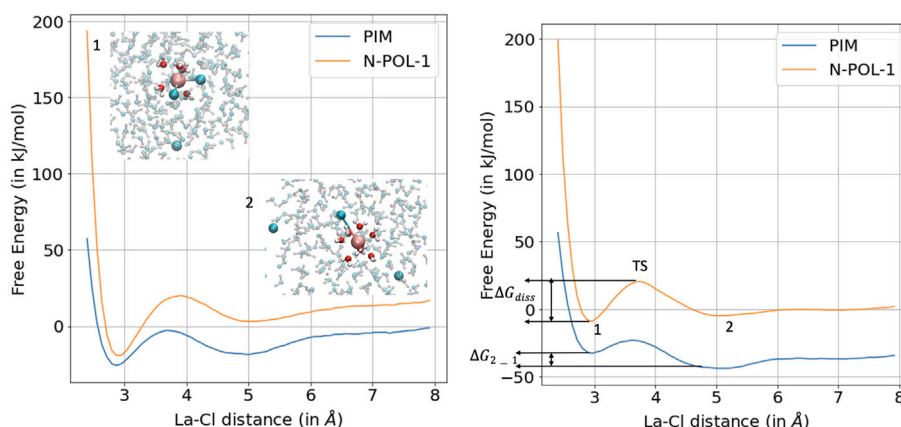
In Tables 7 and S5, high values of the energy barrier for dissociation can be seen compared to the thermal energy,  $k_B T$ . This creates a kinetic bottleneck in unbiased simulations, which underlines the motivation to use WMetaD for this system. Furthermore, it can be inferred that PIM predicts lower energy barriers for dissociation compared to those from all N-POL MD models for all reactions at subcritical as well as supercritical conditions. PIM predicts increasing energy barriers for successive dissociation reactions at both conditions. All N-POL MD models predict almost constant energy barriers at 773 K whereas no systematic trends are observed at 473 K.

All models predict positive values of  $\log K_{\text{eq}}^{\infty}$  for all stepwise association reactions due to the tendency of ions to associate in water at high temperatures. In general, the calculated values of  $\log K_{\text{eq}}^{\infty}$  for PIM are lower than those for N-POL MD for all association reactions at subcritical and supercritical conditions. This is because of the deep energy basins predicted by N-POL MD for CIPs, which make them more exothermic than those from PIM. All models show an increase in  $\log K_{\text{eq}}^{\infty}$  with increasing temperature. Note that FES and  $\log K_{\text{eq}}^{\infty}$  have a many-to-one relation, which means that the same  $\log K_{\text{eq}}^{\infty}$  value can be obtained

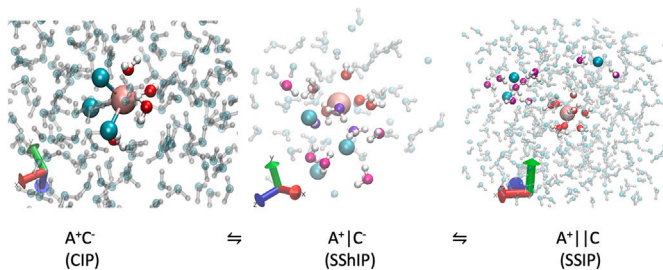
**Table 7**

Energy barriers and reaction constants of La-Cl association reactions from PIM, N-POL-1 and AIMD at supercritical conditions (773 K, density – 1.03 g/cm<sup>3</sup> with box #5) and subcritical conditions (473 K, density – 1.11 g/cm<sup>3</sup> with box #6). Data from N-POL-2 and N-POL-3 are shown in Table S5. The standard deviation of energy barriers ( $\sigma_{\Delta G(r)}$ ) is  $0.5 k_B T$  (conservative value) for all studied reactions under all conditions. The error estimation method is described in SI-1.

| Temperature | Reaction                                    | Method   | Energy barrier (in $k_B T$ ) | $\log K_{eq}^\infty$ |
|-------------|---------------------------------------------|----------|------------------------------|----------------------|
| 773 K       | $La^{3+} + Cl^- \longrightarrow LaCl^{2+}$  | PIM      | 2.97                         | $2.86 \pm 0.01$      |
|             |                                             | N-POL-1  | 6.19                         | $3.51 \pm 0.03$      |
|             |                                             | AIMD [9] | 4.65                         | 2.72                 |
|             | $LaCl^{2+} + Cl^- \longrightarrow LaCl_2^+$ | PIM      | 3.57                         | $2.20 \pm 0.01$      |
|             |                                             | N-POL-1  | 6.17                         | $2.64 \pm 0.03$      |
|             |                                             | AIMD [9] | 4.10                         | 1.66                 |
|             | $LaCl_2^+ + Cl^- \longrightarrow LaCl_3$    | PIM      | 4.07                         | $1.53 \pm 0.02$      |
|             |                                             | N-POL-1  | 6.22                         | $1.77 \pm 0.03$      |
|             |                                             | AIMD [9] | 3.51                         | 0.53                 |
| 473 K       | $La^{3+} + Cl^- \longrightarrow LaCl^{2+}$  | PIM      | 2.25                         | $2.29 \pm 0.02$      |
|             |                                             | N-POL-1  | 7.12                         | $2.45 \pm 0.03$      |
|             |                                             | AIMD [9] | 3.73                         | 1.81                 |
|             | $LaCl^{2+} + Cl^- \longrightarrow LaCl_2^+$ | PIM      | 2.46                         | $1.60 \pm 0.03$      |
|             |                                             | N-POL-1  | 7.47                         | $1.60 \pm 0.01$      |
|             |                                             | AIMD [9] | 4.60                         | 1.19                 |
|             | $LaCl_2^+ + Cl^- \longrightarrow LaCl_3$    | PIM      | 3.10                         | $0.97 \pm 0.02$      |
|             |                                             | N-POL-1  | 6.67                         | $0.76 \pm 0.02$      |
|             |                                             | AIMD [9] | 3.90                         | 0.18                 |



**Fig. 2.** Left: FES of the  $LaCl_2^+ + Cl^- \longrightarrow LaCl_2^+$  reaction at 773 K. Inset 1 shows the associated state where two  $Cl^-$  ions (green) are coordinated to the  $La^{3+}$  ion (brown) while the third  $Cl^-$  is restrained at a long distance. Inset 2 shows the dissociated state where one of the previously coordinating  $Cl^-$  ions is at a long distance. Right: FES of the  $LaCl_2^+ + Cl^- \longrightarrow LaCl_2^+$  reaction at 473 K.  $\Delta G_{diss}$  is the energy barrier for dissociation calculated as the difference in free energy between state 1 (associated state) and the high energy transition state (TS).  $\Delta G_{2-1}$  is the energy difference between states 2 and 1. FES from PIM are rigidly shifted by a small offset for better visual clarity.



**Fig. 3.** Snapshots of different ion pairs obtained from simulations. CIPs (left): brown –  $La^{3+}$ , green –  $Cl^-$ , red – hydration water molecules of  $La^{3+}$ . SSHIPs (middle): brown –  $La^{3+}$ , green –  $Cl^-$ , red – hydration water molecules of  $La^{3+}$ , magenta – hydration water molecules of  $Cl^-$ , violet – hydration water molecules shared by  $La^{3+}$  and  $Cl^-$  ions. SSIPs (right): brown –  $La^{3+}$ , green –  $Cl^-$ , red – hydration water molecules of  $La^{3+}$ , magenta – hydration water molecules of  $Cl^-$ .

from different FES. Therefore, an accurate prediction of the underlying FES is of paramount importance to retrieve reliable association constants.

## 4. Discussion

### 4.1. Speciation and structure: comparison with available AIMD and experimental data

The new PIM potential predicts La-Cl bond distances closer to AIMD data compared to N-POL MD (Table 6) at both conditions studied here. However, distances from both PIM and all N-POL MD models are shorter than the 2.90–3.02 Å range predicted by experiments at similar conditions [9]. Shorter bond lengths from PIM are a consequence of shorter bond lengths predicted by DFT which is used to generate the training configurations. All models predict a decrease in La-Cl bond length with increasing temperature, which is consistent with experiments [9] and AIMD [9] at similar conditions. N-POL MD simulations with box #4 at both conditions show almost no La-Cl dissociation resulting in  $CN > 2$ . The presence of high-energy barriers for dissociation ( $\sim 7$ – $10 k_B T$ , Tables 7, S5) at both conditions is the reason why La-Cl dissociation is kinetically unfavorable in N-POL MD. Similar conclusions were drawn by Zhang and Yan [53] with a different set of non-polarizable pair potentials at similar temperature conditions. As the existing N-POL potentials were optimized to reproduce structures at ambient conditions,

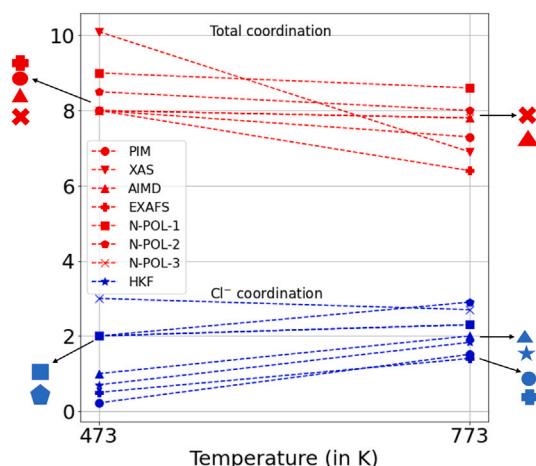


Fig. 4. Comparison of coordination numbers from PIM, N-POL MD with XAS [23], EXAFS [9], AIMD [9] and HKF model [6] data.

we attempted to refit the LJ potential from Migliorati et al. [12] with atomic force data at supercritical conditions. However, this refitted potential also predicted unfavorable La-Cl dissociation. The observed inability of LJ-type potentials to predict the hydration free energies and ion-oxygen distances of divalent cations at room temperature and pressure conditions was already addressed by Li et al. [31], who attributed this behavior to the lack of ion-dipole interactions. These problems are overcome with PIM, which softens the La-Cl interactions resulting in significant dissociation at both conditions, consistent with AIMD simulations and extended X-ray absorption fine structure (EXAFS) data [9] at similar conditions. At 473 K,  $\Delta G_{2-1}$  is negative for PIM reactions (Figure S14), which means that SSHIP is the more stable state compared to CIP. This is also seen in the unbiased simulations (PIM with box #4, Table 6) where almost all  $\text{Cl}^-$  ligands preferentially populate the second coordination shell at a distance of  $\sim 5 \text{ \AA}$ . La-O distances from PIM are shorter compared to those from EXAFS data [9] at similar conditions.

Fig. 4 shows a comparison of the total number of ligands and the number of  $\text{Cl}^-$  ligands in the first coordination shell. With increasing temperature, the static dielectric constant of water decreases, which promotes ion association. This effect is seen in PIM where a change from fully hydrated  $\text{La}^{3+}$  ion to  $\text{LaCl}(\text{H}_2\text{O})_6$  (increase of about one  $\text{Cl}^-$  ion) is observed from subcritical to supercritical conditions, consistent with AIMD data [9], EXAFS studies [9], XAS studies [23] and HKF model data [6]. All N-POL MD models predict higher  $\text{Cl}^-$  coordination with temperature. In terms of the absolute number of  $\text{Cl}^-$  ligands, the PIM data agree well with the HKF model [6] generated using solubility data from Migdisov et al. [24] and EXAFS data [9], whereas an overestimation of  $\sim 1$ – $2$  ligands is observed for the N-POL MD models. With decreasing temperature, PIM shows an increase in the total number of ligands of about one, in agreement with EXAFS data [9]. XAS data [23] predict an increase of about three ligands. However, XAS generally overestimates coordination numbers as indicated by the 10-fold coordinated  $\text{La}^{3+}$  predicted at subcritical conditions [23]. N-POL MD predicts negligible change in the total number of ligands with temperature. In terms of the absolute number of ligands in the first coordination shell, PIM data agree well with EXAFS [9] and AIMD [9] data, while N-POL MD overestimates the number of ligands by about one (except N-POL-3 at 473 K). PIM also shows better agreement with XAS studies [23] in terms of the total number of ligands, compared to N-POL MD at 773 K. In general, La-O distances predicted by PIM are shorter than those in AIMD and experiments [9] at similar conditions.

Table 8

Cumulative association constants ( $\log \beta$ ) obtained from N-POL-1, PIM and AIMD (\*values are from the two data sets shown in Migdisov et al. [24]). Data for N-POL-2 and N-POL-3 models are shown in Table S5.

| Temperature | Reaction                                                         | Method     | $\log \beta$ |
|-------------|------------------------------------------------------------------|------------|--------------|
| 773 K       | $\text{La}^{3+} + \text{Cl}^- \longrightarrow \text{LaCl}^{2+}$  | PIM        | 2.86         |
|             |                                                                  | N-POL-1    | 3.51         |
|             |                                                                  | AIMD [9]   | 2.72         |
|             | $\text{La}^{3+} + 2 \text{Cl}^- \longrightarrow \text{LaCl}_2^+$ | PIM        | 5.06         |
|             |                                                                  | N-POL-1    | 6.15         |
|             | $\text{La}^{3+} + 3 \text{Cl}^- \longrightarrow \text{LaCl}_3$   | AIMD [9]   | 4.38         |
| 473 K       | $\text{La}^{3+} + \text{Cl}^- \longrightarrow \text{LaCl}^{2+}$  | PIM        | 6.59         |
|             |                                                                  | N-POL-1    | 7.92         |
|             |                                                                  | AIMD [9]   | 4.91         |
|             | $\text{La}^{3+} + 2 \text{Cl}^- \longrightarrow \text{LaCl}_2^+$ | PIM        | 2.29         |
|             |                                                                  | N-POL-1    | 2.45         |
|             |                                                                  | AIMD [9]   | 1.81         |
|             |                                                                  | Model [24] | 2.29, 2.38*  |
|             | $\text{La}^{3+} + 3 \text{Cl}^- \longrightarrow \text{LaCl}_3$   | PIM        | 3.89         |
|             |                                                                  | N-POL-1    | 4.05         |
|             |                                                                  | AIMD [9]   | 3.00         |
|             |                                                                  | Model [24] | 3.86         |
|             |                                                                  | PIM        | 4.86         |
|             |                                                                  | N-POL-1    | 4.81         |
|             |                                                                  | AIMD [9]   | 3.18         |

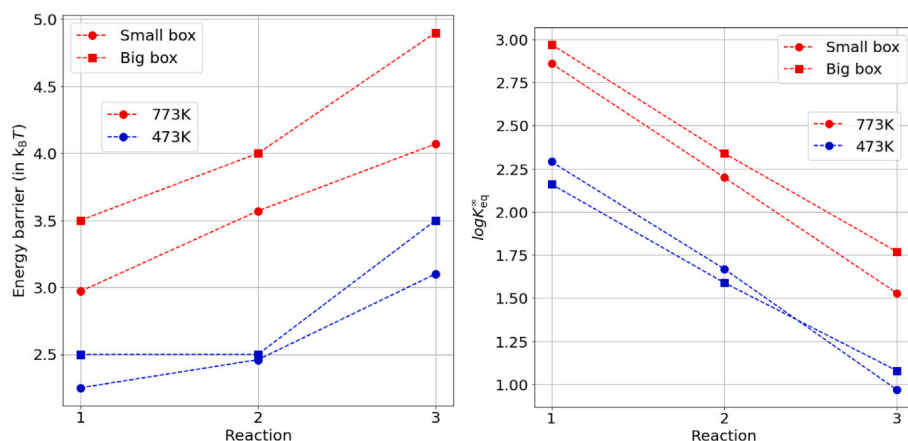
#### 4.2. Ion association thermodynamics: comparison with available AIMD and experimental data

In general, association constants predicted by AIMD, N-POL MD and PIM follow the order,  $\log K_{\text{eq,N-POL}}^{\infty} > \log K_{\text{eq,PIM}}^{\infty} > \log K_{\text{eq,AIMD}}^{\infty}$ . This is seen in Tables 7, S5 and also illustrated in Figure S15. Although PIM gives better  $\log K_{\text{eq}}^{\infty}$  values compared to N-POL, the disagreement with AIMD is still relatively large, i.e. about one log unit in some cases. Tables 8, S5 show the cumulative association constants (also called formation constants),  $\log \beta$  obtained from the N-POL MD models, PIM, AIMD [9] and the HKF model parameterized with solubility data from Migdisov et al. [24] up to  $\sim 350^\circ \text{C}$ . They are obtained by cumulative addition of  $\log K_{\text{eq}}^{\infty}$  values shown in Tables 7, S5 and hence, follow the same descending order as association constants.  $\log \beta$  predicted by PIM for reactions 2 and 3 at 473 K are in better agreement with the thermodynamic model and AIMD than all N-POL MD models. Due to the increased stability of La-Cl complexes at higher temperatures, the step-wise addition of  $\text{Cl}^-$  into the coordination sphere of  $\text{La}^{3+}$  becomes more exothermic at higher temperatures. As a result, all models predict an increase in  $\log \beta$  values of  $\text{LaCl}^{2+}$ ,  $\text{LaCl}_2^+$  and  $\text{LaCl}_3$  with increasing temperature, in agreement with data from the HKF model up to  $\sim 400^\circ \text{C}$  [24] and with AIMD data [9].

The relative stabilities of the metastable CIPs and SSHIPs predicted by PIM are in agreement with AIMD data as shown in Figure S14. Both predict CIPs as the more stable state at 773 K and SSHIPs as the more stable state at 473 K in contrast to the N-POL models that always predict CIPs as the more stable state under both conditions. Zhang and Yan [53] also predict similar trends in energy barriers and relative stabilities with a different set of non-polarizable pair potentials at similar temperature conditions. Absolute values of  $\Delta G_{2-1}$  from PIM are generally lower than those from AIMD.

##### 4.2.1. Effect of box size on association constant

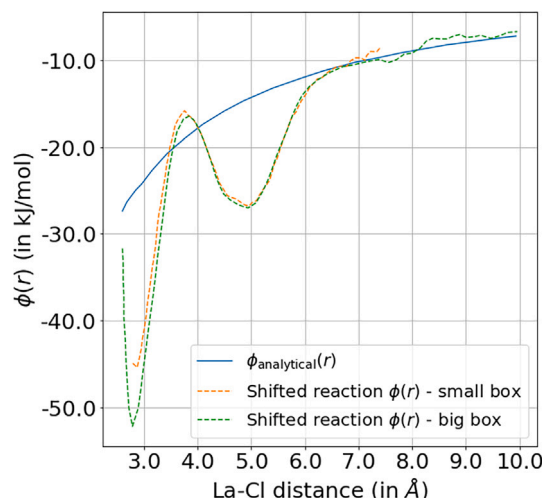
MD simulations use periodic boundary conditions to replicate bulk liquid systems. The success of this assumption depends on the size of the simulation box which determines the distance between the particles and their nearest images. A small simulation box would mean smaller distances between particles and their periodic images leading to higher interactions between them, which is not desirable as this imposes correlations not expected in fluids. To understand the effect of box size on the calculated values of  $\log K_{\text{eq}}^{\infty}$  and energy barriers, WMetaD simulations



**Fig. 5.** Comparison of energy barriers (left) and  $\log K_{\text{eq}}^{\infty}$  of association reactions (right) from small and big simulation boxes. Integration limits of  $\sim 6.30 \text{ \AA}$  and  $\sim 11 \text{ \AA}$  were used for small and big boxes, respectively.

are performed with boxes #4 and #7 in Table 2 with the new PIM potential. The results are presented in Fig. 5. This analysis is particularly important to assess the quality of association constant data obtained from AIMD simulations with small simulation boxes.

Fig. 5 shows minor disagreements (maximum difference of  $\sim 0.9 k_B T$  in energy barriers and  $\sim 0.25 \log$  units in association constants) between results from the two simulation boxes. The technique used to calculate the association constants in this work boils down to calculating  $g^{\infty}(r)$ , which is the radial distribution function in the limit of infinite dilution. For highly dilute solutions, obtaining  $g^{\infty}(r)$  from MD simulations is difficult as sampling deficiencies occur due to rather short simulation times, leading to poor statistics. Therefore, the more practical way is to calculate  $g^{\infty}(r)$  from  $\Delta G$  obtained from biased simulations using Equation S12. If it were possible to simulate a 'theoretically' infinitely dilute solution (the simulation box is big enough with adequate number of water molecules and La, Cl particles to reach bulk behavior),  $g^{\infty}(r)$  would converge to 1 and hence the potential of mean force ( $\phi(r)$ , also called the solvent averaged McMillan-Mayer potential) would converge to 0 at  $r \rightarrow \infty$ , which is the dissociated state in the ideal case. However, due to the extremely long-range nature of Coulomb interactions and limited number of La, Cl particles, this is difficult to achieve even in simulation boxes with edge lengths of  $\sim 30 \text{ \AA}$  [54]. Hence, the dissociated state is taken to be the limiting continuum behavior where charged particles interact via the screened Coulomb interaction given by  $\phi_{\text{analytical}}(r) = q_i q_j / 4\pi\epsilon_0 \epsilon_r r$ . PMF is calculated for simulation boxes of different sizes by adding the configurational entropy term to the corresponding FES (Equation S15). PMFs are then rigidly shifted so that their tails coincide with  $\phi_{\text{analytical}}(r)$  at long distances, as shown in Fig. 6. The difference in the  $\log K_{\text{eq}}^{\infty}$  values arises due to the difference in length of the tail coinciding with the analytical solution. However, once the PMFs properly converge with the analytical solution (indicating efficient sampling of all the metastable states), dependence of  $\log K_{\text{eq}}^{\infty}$  on the simulation box size (and hence the cutoff distance ( $r_c$  in Equation S11)) is weak beyond  $\sim 5 \text{ \AA}$  (corresponding to SShIPs) at high temperatures, as shown in Figure S16. Boxes used in AIMD simulations with edge lengths of  $\sim 13 \text{ \AA}$  are big enough to efficiently sample the two important metastable states (corresponding to CIPs and SShIPs) as well as the limiting continuum at the studied conditions. This is indicated by the convergence of PMFs obtained from the small simulation boxes (Fig. 6) with  $\phi_{\text{analytical}}(r)$  at long distances. As  $\log K_{\text{eq}}^{\infty}$  shows a long-tail behavior once limiting continuum is sampled (Figure S16), association constants calculated from the small simulation boxes used in AIMD simulations show rather weak dependence on box size at the studied conditions. The effect of box size may however become important at lower temperatures where SSIPs are stabilized in addition to CIPs and SShIPs. Since



**Fig. 6.**  $\phi(r)$  obtained from small simulation box (at 773 K and density of  $1.03 \text{ g/cm}^3$  (box #5)) and big simulation box (at 773 K and  $0.9 \text{ g/cm}^3$  (box #4)) aligned to the analytical solution.  $\phi(r)$  from larger box shows a longer tail coinciding with the analytical solution.

SSIPs are formed at  $\sim 7.5\text{--}8 \text{ \AA}$  efficient sampling of these ion pairs is no longer possible in small simulation boxes. Bigger simulation boxes with minimum edge lengths of  $\sim 16\text{--}17 \text{ \AA}$  are required to calculate reliable association constants at lower temperatures.

## 5. Conclusion

We fit a new polarizable potential for  $\text{La}^{3+}$  in  $\text{Cl}^-$ -bearing hydrothermal systems from first-principles AIMD simulations. Production simulations were performed at 773 K and 473 K to derive structural and thermodynamic properties of  $\text{La}^{3+}$  in dilute to concentrated solutions.

The new PIM interaction potential gives good estimates of La-Cl and La-O distances under the studied conditions. PIM provides better estimates of total coordination and  $\text{Cl}^-$  coordination numbers compared to N-POL MD in dilute solutions. It also reproduces the increase in La-Cl coordination with increasing temperature as observed in experiments. However, the La-O distances predicted by PIM are shorter than the experimental data under similar conditions.

Calculation of association constants reveals increasing stability of  $\text{LaCl}_2^+$ ,  $\text{LaCl}_2^+$  and  $\text{LaCl}_3$  complexes with increasing temperature. In general, the formation constants calculated from PIM simulations are closer



to the data from the AIMD and HKF models than those from N-POL MD. This is despite the fact that no energy values are used in fitting the PIM. Relative stabilities of CIPs and SSHIPs predicted by PIM are in agreement with AIMD at both conditions. Altogether, the newly fitted PIM gives more precise structural and thermodynamic insight into dilute La-Cl-bearing hydrothermal fluids at both the studied temperature and pressure conditions than those from simple pair potentials, with reasonable computational cost. We also show that the effects of simulation box size on calculated thermodynamic properties are negligibly small at high temperatures. Our results underpin the importance of an explicit description of ionic polarizability to retrieve reliable structural and thermodynamic information in hydrothermal systems containing highly polarizable anions such as  $\text{Cl}^-$  and cations with high polarization power such as  $\text{La}^{3+}$ .

The explicit polarizability term gives a better description of the charge-dipole and dipole-dipole interactions, thereby stabilizing the SSHIPs at room temperature and subcritical conditions by reducing the total potential energy of the system. Although non-polarizable pair potentials have successfully predicted the hydration structure of  $\text{La}^{3+}$  cation at room temperature, an implicit description of charge-dipole and dipole-dipole interactions is not good enough to correctly predict the relative stability of the different ion pairs, particularly in the presence of highly polarizable anions such as  $\text{Cl}^-$ . We tried to optimize an LJ potential with atomic data at supercritical conditions and found this potential to over-stabilize the CIPs as well. So, this problem seems to arise due to the relatively simple functional form of the LJ potential and not due to the thermodynamic conditions or the target properties on which these potentials are optimized. Adding the explicit polarizability term offers a relatively straightforward solution to this problem at reasonable computation costs. Of course, it remains to be tested how the explicit charge-dipole and dipole-dipole interactions affect other structural and transport properties of dilute and concentrated aqueous  $\text{LaCl}_3$  solutions. Additionally, most of the non-polarizable interaction potentials are optimized to reproduce macroscopic experimental results like hydration free energies or EXAFS signatures, which severely limits the transferability of these potentials beyond the thermodynamic conditions at which they are optimized. Using atomic force and dipole moment data to optimize the potential, as in the case of PIM, makes the model much more transferable, at least for the thermodynamic and structural properties studied in this work.

#### CRedit authorship contribution statement

**Rajorshi Chattopadhyay:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sandro Jahn:** Writing – review & editing, Supervision, Resources.

#### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered potential competing interests:

Rajorshi Chattopadhyay reports that financial support was provided by Deutsche Forschungsgemeinschaft (DFG). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

This study was supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of project JA 1469/13-1. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. ([www.gauss-centre.eu](http://www.gauss-centre.eu)) for funding this-project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). The authors would like to thank the editor and the reviewers for their constructive reviews which improved the overall quality of the paper.

#### Appendix A. Supplementary data

Supplementary data for this article can be found online at doi:[10.1016/j.molliq.2025.128546](https://doi.org/10.1016/j.molliq.2025.128546).

#### Data availability

The pair correlation functions, FES and example CP2K input files are hosted in an open-source repository - DOI [10.17605/OSF.IO/E93GQ](https://doi.org/10.17605/OSF.IO/E93GQ).

#### References

- [1] R.G. Skirrow, D.L. Huston, T.P. Mernagh, J.P. Thorne, H. Duffer, A. Senior, *Critical Commodities for a High-Tech World: Australia's Potential to Supply Global Demand*, Geoscience Australia Canberra, 2013.
- [2] I. González-Álvarez, F. Stoppa, X.Y. Yang, A. Porwal, Introduction to the special issue, insights on carbonatites and their mineral exploration approach: a challenge towards resourcing critical metals, *Ore Geol. Rev.* 133 (2021) 104073, <https://doi.org/10.1016/j.oregeorev.2021.104073>.
- [3] A. Williams-Jones, I. Samson, G. Olivo, The genesis of hydrothermal fluorite-REE deposits in the Gallinas Mountains, New Mexico, *Econ. Geol.* 95 (2000) 327–341, <https://doi.org/10.2113/gsecongeo.95.2.327>.
- [4] E. Sheard, A. Williams-Jones, M. Heiligmann, C. Pederson, D. Trueman, Controls on the concentration of zirconium, niobium, and the rare earth elements in the thor lake rare metal deposit, northwest territories, Canada, *Econ. Geol.* 107 (2012) 81–104, <https://doi.org/10.2113/econgeo.107.1.81>.
- [5] J. Brugger, W. Liu, B. Etschmann, Y. Mei, D.M. Sherman, D. Testemale, A review of the coordination chemistry of hydrothermal systems, or do coordination changes make ore deposits? *Chem. Geol.* 447 (2016) 219–253, <https://doi.org/10.1016/j.chemgeo.2016.10.021>.
- [6] A. Migdisov, A.E. Williams-Jones, J. Brugger, F.A. Caporuscio, Hydrothermal transport, deposition, and fractionation of the REE: experimental data and thermodynamic calculations, *Chem. Geol.* 439 (2016) 13–42, <https://doi.org/10.1016/j.chemgeo.2016.06.005>.
- [7] D.A. Banks, B.W.D. Yardley, A.R. Campbell, K.E. Jarvis, REE composition of an aqueous magmatic fluid: a fluid inclusion study from the Capitan Pluton, New Mexico, U.S.A., *Chem. Geol.* 113 (3) (1994) 259–272, [https://doi.org/10.1016/0009-2541\(94\)90070-1](https://doi.org/10.1016/0009-2541(94)90070-1).
- [8] L. Lintjeras, I. Setiawan, Mobility of rare earth element in hydrothermal process and weathering product: a review, *IOP Conf. Ser. Earth Environ. Sci.* 118 (1) (2018) 012076, <https://doi.org/10.1088/1755-1315/118/1/012076>.
- [9] Q. Guan, Y. Mei, B. Etschmann, M. Louvel, D. Testemale, R. Spezia, J. Brugger, Speciation and thermodynamic properties of La(III)-Cl complexes in hydrothermal fluids: a combined molecular dynamics and in situ x-ray absorption spectroscopy study, *Geochim. Et Cosmochim. Acta* 330 (2022) 27–46, <https://doi.org/10.1016/j.gca.2022.02.032>.
- [10] T. Ikeda, M. Hirata, T. Kimura, Hydration structure of  $\text{Y}^{3+}$  and  $\text{La}^{3+}$  compared: an application of metadynamics, *J. Chem. Phys.* 122 (24) (2005) 244507, <https://doi.org/10.1063/1.1940029>.
- [11] C. Terrier, P. Vitorge, M.-P. Gaigeot, R. Spezia, R. Vuilleumier, Density functional theory based molecular dynamics study of hydration and electronic properties of aqueous  $\text{La}^{3+}$ , *J. Chem. Phys.* 133 (4) (2010) 044509, <https://doi.org/10.1063/1.3460813>.
- [12] V. Migliorati, A. Serva, F. Terenzio, P. D'Angelo, Development of Lennard-Jones and Buckingham potentials for lanthanoid ions in water, *Inorg. Chem.* 56 (May 2017) <https://doi.org/10.1021/acs.inorgchem.7b00207>.
- [13] M. Duval, M. Souaille, R. Spezia, T. Cartailleur, P. Vitorge, Pair interaction potentials with explicit polarization for molecular dynamics simulations of  $\text{La}^{3+}$  in bulk water, *J. Chem. Phys.* 127 (3) (2007) 034503, <https://doi.org/10.1063/1.2751503>.
- [14] M. Duval, P. Vitorge, R. Spezia, Building a polarizable pair interaction potential for lanthanoids(III) in liquid water: a molecular dynamics study of structure and dynamics of the whole series, *J. Chem. Phys.* 130 (10) (2009) 104501, <https://doi.org/10.1063/1.3081143>.
- [15] R. Spezia, M. Duval, P. Vitorge, P. D'Angelo, Molecular dynamics to rationalize EXAFS experiments: a dynamical model explaining hydration behaviour across the lanthanoid(III) series, *J. Phys. Conf. Ser.* 190 (1) (2009) 012056, <https://doi.org/10.1088/1742-6596/190/1/012056>.
- [16] C. Clavaguéra, R. Pollet, J.M. Soudan, V. Brenner, J.P. Dognon, Molecular dynamics study of the hydration of lanthanum(III) and europium(III) including many-body effects, *J. Phys. Chem. B* 109 (16) (2005) 7614–7616, <https://doi.org/10.1021/jp051032h>.
- [17] C. Beuchat, D. Hagberg, R. Spezia, L. Gagliardi, Hydration of lanthanide chloride salts: a quantum chemical and classical molecular dynamics simulation study, *J. Phys. Chem. B* 114 (47) (2010) 15590–15597, <https://doi.org/10.1021/jp105590h>.
- [18] W.W. Rudolph, G. Irmer, Hydration and ion pair formation in common aqueous La(III) salt solutions – a Raman scattering and DFT study, *Dalton Trans* 44 (2015) 295–305, <https://doi.org/10.1039/C4DT03003F>.
- [19] M. Bühl, N. Sieffert, A. Partouche, A. Chaumont, G. Wipff, Speciation of La(III) chloride complexes in water and acetonitrile: a density functional study, *Inorg. Chem.* 51 (24) (2012) 13396–13407, <https://doi.org/10.1021/ic302255a>.
- [20] L. Petit, R. Vuilleumier, P. Maldivi, C. Adamo, Molecular dynamics study of the coordination sphere of trivalent lanthanum in a highly concentrated LiCl aqueous solution: a combined classical and ab initio approach, *J. Phys. Chem. B* 112 (34) (2008) 10603–10607, <https://doi.org/10.1021/jp8017106>.

- [21] J.C. Tanger, H.C. Helgeson, Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures; revised equations of state for the standard partial molal properties of ions and electrolytes, *Am. J. Sci.* 288 (1) (1988) 19–98, <https://doi.org/10.2475/ajs.288.1.19>.
- [22] J.R. Haas, E.L. Shock, D.C. Sassani, Rare earth elements in hydrothermal systems: estimates of standard partial molal thermodynamic properties of aqueous complexes of the rare earth elements at high pressures and temperatures, *Geochim. Et Cosmochim. Acta* 59 (21) (1995) 4329–4350, [https://doi.org/10.1016/0016-7037\(95\)00314-P](https://doi.org/10.1016/0016-7037(95)00314-P).
- [23] R. Mayanovic, A. Anderson, W. Bassett, I.-M. Chou, Steric hindrance and the enhanced stability of light rare-earth elements in hydrothermal fluids, *Am. Mineral.* 94 (2009) 1487–1490, <https://doi.org/10.2138/am.2009.3250>.
- [24] A.A. Migdisov, A.E. Williams-Jones, T. Wagner, An experimental study of the solubility and speciation of the rare earth elements (iii) in fluoride- and chloride-bearing aqueous solutions at temperatures up to 300°C, *Geochim. Et Cosmochim. Acta* 73 (23) (2009) 7087–7109, <https://doi.org/10.1016/j.gca.2009.08.023>.
- [25] D. Sherman, M.D. Collings, Ion association in concentrated NaCl brines from ambient to supercritical conditions: results from classical molecular dynamics simulations, *Geochim. Trans.* 3 (2002) 102–107, <https://doi.org/10.1039/b208671a>.
- [26] S. Tazi, J.J. Molina, B. Rotenberg, P. Turq, R. Vuilleumier, M. Salanne, A transferable ab initio based force field for aqueous ions, *J. Chem. Phys.* 136 (11) (2012) 114507, <https://doi.org/10.1063/1.3692965>.
- [27] W. Zhang, A.C.T. van Duin, Improvement of the ReaxFF description for functionalized hydrocarbon/water weak interactions in the condensed phase, *J. Phys. Chem. B* 122 (14) (2018) 4083–4092, <https://doi.org/10.1021/acs.jpcc.8b01127>.
- [28] G. Lamoureux, E. Harder, I.V. Vorobyov, B. Roux, A.D. MacKerell, A polarizable model of water for molecular dynamics simulations of biomolecules, *Chem. Phys. Lett.* 418 (1) (2006) 245–249, <https://doi.org/10.1016/j.cplett.2005.10.135>.
- [29] L.X. Dang, T.-M. Chang, Molecular dynamics study of water clusters, liquid, and liquid–vapor interface of water with many-body potentials, *J. Chem. Phys.* 106 (19) (1997) 8149–8159, <https://doi.org/10.1063/1.473820>.
- [30] V. Miglioni, A. Serva, F. Sessa, A. Lapi, P. D'Angelo, Influence of counterions on the hydration structure of lanthanide ions in dilute aqueous solutions, *J. Phys. Chem. B* 122 (10) (2018) 2779–2791, pMID: 29481749, <https://doi.org/10.1021/acs.jpcc.7b12571>.
- [31] P. Li, L.F. Song, K.M.J. Merz, Parameterization of highly charged metal ions using the 12-6-4 lj-type nonbonded model in explicit water, *J. Phys. Chem. B* 119 (3) (2015) 883–895, <https://doi.org/10.1021/jp505875v>.
- [32] J.J. Molina, S. Lectez, S. Tazi, M. Salanne, J.-F. Dufrêche, J. Roques, E. Simoni, P.A. Madden, P. Turq, Ions in solutions: determining their polarizabilities from first-principles, *J. Chem. Phys.* 134 (1) (2011) 014511, <https://doi.org/10.1063/1.3518101>.
- [33] R.J. Heaton, P.A. Madden, S.J. Clark, S. Jahn, Condensed phase ionic polarizabilities from plane wave density functional theory calculations, *J. Chem. Phys.* 125 (14) (2006) 144104, <https://doi.org/10.1063/1.2357151>.
- [34] K.T. Tang, J.P. Toennies, An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients, *J. Chem. Phys.* 80 (8) (1984) 3726–3741, <https://doi.org/10.1063/1.447150>.
- [35] A. Aguado, P.A. Madden, Ewald summation of electrostatic multipole interactions up to the quadrupolar level, *J. Chem. Phys.* 119 (14) (2003) 7471–7483, <https://doi.org/10.1063/1.1605941>.
- [36] J. Wang, P. Cieplak, Q. Cai, M.-J. Hsieh, J. Wang, Y. Duan, R. Luo, Development of polarizable models for molecular mechanical calculations. 3. Polarizable water models conforming to thole polarization screening schemes, *J. Phys. Chem. B* 116 (28) (2012) 7999–8008, <https://doi.org/10.1021/jp212117d>.
- [37] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* 77 (1996) 3865–3868, <https://doi.org/10.1103/PhysRevLett.77.3865>.
- [38] V. Rozsa, G. Galli, Solvation of simple ions in water at extreme conditions, *J. Chem. Phys.* 154 (14) (2021) 144501, <https://doi.org/10.1063/5.0046193>.
- [39] D. Pan, L. Spanu, B. Harrison, D.A. Sverjensky, G. Galli, Dielectric properties of water under extreme conditions and transport of carbonates in the deep earth, *Proc. Natl. Acad. Sci.* 110 (17) (2013) 6646–6650, <https://doi.org/10.1073/pnas.1221581110>.
- [40] S. Grimm, G. Schoendorff, A.K. Wilson, Gauging the performance of density functionals for lanthanide-containing molecules, *J. Chem. Theory Comput.* 12 (3) (2016) 1259–1266, pMID: 26848891, <https://doi.org/10.1021/acs.jctc.5b01193>.
- [41] R.C. Shier, J.L. Fulton, M. Balasubramanian, M.-T. Nguyen, J.-B. Lu, J. Li, R. Rousseau, V.-A. Glezakou, D.C. Cantu, Coordination sphere of lanthanide aqua ions resolved with ab initio molecular dynamics and x-ray absorption spectroscopy, *Inorg. Chem.* 60 (5) (2021) 3117–3130, <https://doi.org/10.1021/acs.inorgchem.0c03438>.
- [42] S.J. Fowler, D.M. Sherman, The nature of NaCl–H<sub>2</sub>O deep fluids from ab initio molecular dynamics at 0.5–4.5 gpa, 20–800 °C, and 1–14 m NaCl, *Geochim. Et Cosmochim. Acta* 277 (2020) 243–264, <https://doi.org/10.1016/j.gca.2020.03.031>.
- [43] M.S. Yoshiki Ishii, S. Kasai, N. Ohtori, Transport coefficients and the Stokes–Einstein relation in molten alkali halides with polarisable ion model, *Mol. Phys.* 113 (17–18) (2015) 2442–2450, <https://doi.org/10.1080/00268976.2015.1046527>.
- [44] T. Driesner, C.A. Heinrich, The system H<sub>2</sub>O–NaCl. Part I: correlation formulae for phase relations in temperature–pressure–composition space from 0 to 1000°C, 0 to 5000bar, and 0 to 1 X<sub>NaCl</sub>, *Geochim. Et Cosmochim. Acta* 71 (20) (2007) 4880–4901, <https://doi.org/10.1016/j.gca.2006.01.033>.
- [45] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, cp2k: atomistic simulations of condensed matter systems, *WIREs Comput. Mol. Sci.* 4 (1) (2014) 15–25, <https://doi.org/10.1002/wcms.1159>.
- [46] M. Brehm, M. Thomas, S. Gehrke, B. Kirchner, TRAVIS—a free analyzer for trajectories from molecular simulation, *J. Chem. Phys.* 152 (16) (2020) 164105, <https://doi.org/10.1063/5.0005078>.
- [47] A. Barducci, G. Bussi, M. Parrinello, Well-tempered metadynamics: a smoothly converging and tunable free-energy method, *Phys. Rev. Lett.* 100 (2008) 020603, <https://doi.org/10.1103/PhysRevLett.100.020603>.
- [48] A. Barducci, M. Bonomi, M. Parrinello, Metadynamics, *WIREs Comput. Mol. Sci.* 1 (5) (2011) 826–843, <https://doi.org/10.1002/wcms.31>.
- [49] A.A. Chialvo, P.T. Cummings, H.D. Cochran, J.M. Simonson, R.E. Mesmer, Na<sup>+</sup> Cl<sup>−</sup> ion pair association in supercritical water, *J. Chem. Phys.* 103 (21) (1995) 9379–9387, <https://doi.org/10.1063/1.470707>.
- [50] C. Clavaguéra, J.-P. Dognon, Accurate static electric dipole polarizability calculation of +3 charged lanthanide ions, *Chem. Phys.* 311 (2005) 169–176, <https://doi.org/10.1016/j.chemphys.2004.10.014>.
- [51] E.E. Wait, J. Gourary, C. Liu, E.D. Spörke, S.B. Rempe, P. Ren, Development of amoeba polarizable force field for rare-earth La<sup>3+</sup> interaction with bioinspired ligands, *J. Phys. Chem. B* 127 (6) (2023) 1367–1375, <https://doi.org/10.1021/acs.jpcc.2c07237>.
- [52] M. Schulze, T. Driesner, S. Jahn, Assessing the validity and limits of linear density models for predicting dissociation–association equilibria in supercritical water, *Geochim. Et Cosmochim. Acta* (2024) <https://doi.org/10.1016/j.gca.2024.10.008>.
- [53] W. Zhang, T. Yan, A molecular dynamics investigation of La<sup>3+</sup> and Lu<sup>3+</sup>-ligand speciation in aqueous solution, *J. Mol. Liq.* 347 (2022) 118367, <https://doi.org/10.1016/j.molliq.2021.118367>.
- [54] J. Aurfot, P. Raiteri, J.D. Gale, Computational insights into Mg<sup>2+</sup> dehydration in the presence of carbonate, *ACS. Earth. Space. Chem.* 6 (3) (2022) 733–745, <https://doi.org/10.1021/acsearthspacechem.1c00389>.