

Automated Development of Molten Salt Machine Learning Potentials: Application to LiCl

Ganesh Sivaraman,[†] Jicheng Guo,[‡] Logan Ward,[†] Nathaniel Hoyt,[‡] Mark
Williamson,[‡] Ian Foster,[†] Chris Benmore,[¶] and Nicholas Jackson^{*,§}

[†]*Data Science and Learning Division, Argonne National Laboratory, Lemont, IL 60439,
USA*

[‡]*Chemical and Fuel Cycle Technologies Division, Argonne National Laboratory, Lemont, IL
60439, USA*

[¶]*X-ray Science Division, Argonne National Laboratory, Lemont, IL 60439, USA*

[§]*Department of Chemistry, University of Illinois, Urbana-Champaign, Urbana, IL 61801,
USA*

E-mail: jacksonn@illinois.edu, gsivaraman@anl.gov

Abstract

The *in silico* modeling of molten salts is of crucial importance to emerging "carbon-free" energy applications, but is inhibited by the computational cost of quantum mechanically treating the high polarizabilities characteristic of molten salts. Here, we integrate configurational sampling using classical force-fields with active learning to automate the generation of near-DFT accurate machine learning Gaussian Approximation Potentials (GAP) for molten LiCl using fewer than 600 atomic configurations. Relative to conventional *ab initio* molecular dynamics, the molten LiCl GAP model

exhibits a 19,000x speedup and improved experimental agreement as gauged by calculated R-factors. The accuracy of the GAP parametrization workflow is validated by its ability to reproduce experimental structure factors, densities, self-diffusion coefficients, and ionic conductivities for molten LiCl. This hybrid simulation strategy significantly accelerates the generation of machine learning potentials for molten salts by reducing the expensive *ab initio* calculations required for parameterization to $O(100)$ evaluations, enabling the facile generation of first-principles quality predictions of structural and dynamical properties of molten salts.

Molecular dynamics (MD) simulations constitute a crucial component of modern atomistic modeling efforts.¹ While in principle many-body energies and forces for MD can be derived from solutions of the Schrödinger equation,² the computational costs of *ab initio* MD (AIMD) limits the accessible spatiotemporal scales of simulations. In many atomistic systems the dynamical processes of interest (e.g., structural relaxation, diffusion, transport) extend beyond the picosecond timescale, and alternatives to direct AIMD are required. The most common alternative involves the use of efficient classical pairwise interatomic potentials parametrized via *ab initio* calculations and experiments.³ While these classical potentials exhibit improved computational efficiency, they are approximate and neglect the full complexity of the potential energy surface, resulting in variable and often poor performance. Recently, advances in machine learning (ML) have radically transformed the landscape of dynamical simulations in atomic systems^{4–10} by providing flexible and computationally efficient functional forms for parametrizing force-fields via reference *ab initio* datasets of atomic configurations. These ML potentials (MLP) provide a mechanism for obtaining near AIMD accuracy at a computational cost dramatically reduced from full AIMD simulation, enabling the accurate and scalable dynamical modeling of highly polarizable atomic systems that have been traditionally challenging for either AIMD or classical force-fields.

Molten salts are high temperature ionic liquids with significant potential in energy applications such as liquid metal batteries,^{11,12} concentrated solar power^{13,14} and molten salt reactors.¹⁵ The high chemical reactivity of molten salts makes direct characterization of their structure and properties experimentally challenging. As a result, computational modeling has been a crucial component of efforts to understand and design the structure and thermodynamics of molten salts for over 50 years.^{16–19} Despite the early successes of classical interatomic potentials such as the Rigid Ion Model (RIM), the systematic modeling of molten salts has been inhibited by the dominant role that many-body electronic polarization plays in governing the potential energy surface.^{18,20} Recently, we showed that many-body polarization in NaCl melts can be captured accurately by using MLP simulations scaled to $\sim 10,000$

atoms with near Density Functional Theory (DFT) accuracy.²¹ At these spatiotemporal scales, structural heterogeneity and ensemble averaging can be incorporated meaningfully into calculations, combining the accuracy of DFT with the statistical sampling of classical interatomic potentials.

Despite these successes, the application of MLPs to the broader chemical space of molten salts is challenged by the difficulty of MLP parametrization. In MLP parametrization efforts, one must balance the computational cost of training set generation, the sampling of diverse atomic configurations, and the challenge of ML hyperparameter optimization to develop MLP models of sufficient predictive accuracy and generalizability. A variety of approaches have been put forward to augment MLP training in other disciplines,^{22–24} but comprehensive strategies facilitating the application of MLPs to general molten salt chemistries are absent from the literature. A systematic approach for modeling molten salts via automated bottom-up MLP parametrization would represent a crucial stepping stone towards the understanding of molten salts, as well as the broader field of ionic liquids of importance to numerous energy applications.²⁵

In this letter, we present a systematic approach for automating molten salt MLP development via the combination of configurational sampling using classical force-fields, active learning,²⁶ and Bayesian hyperparameter optimization²⁷ of Gaussian Approximation Potentials (GAP). First, we describe the components of the automated MLP parametrization workflow for molten salts as well as the criteria for convergence. We then apply this automation scheme to generate MLP for molten LiCl, and validate MLP performance at multiple temperatures via comparison to DFT-derived atomic forces and experimentally derived structure factors, densities, diffusivities, and ionic conductivities. The ability of the parametrized MLP to accurately reproduce experimental structural, thermodynamic, and dynamical properties of interest reinforces the efficacy of the proposed modeling paradigm, and we describe further applications and improvements for the method moving forward.

Motivated by the success of previous applications of active learning to MLP genera-

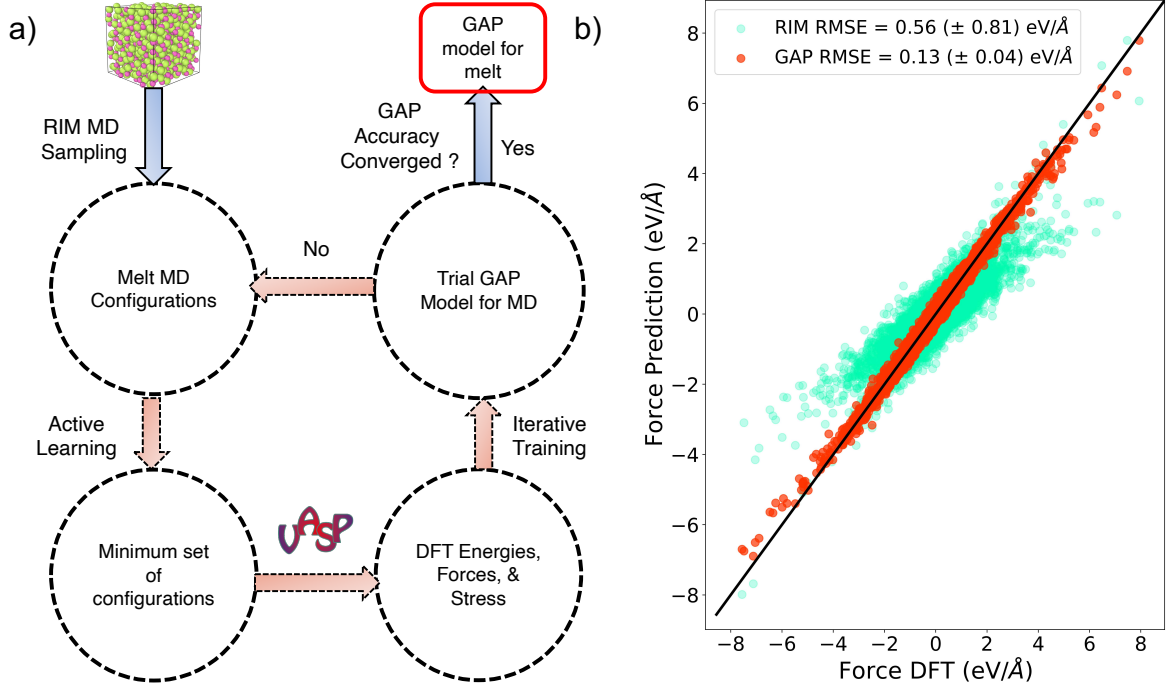


Figure 1: Workflow and performance for LiCl GAP model. **a)** Automated workflow for GAP parametrization incorporating configurational sampling via RIM, active learning, DFT, and GAP model fitting. GAP model fitting proceeds iteratively until prediction accuracy of atomic energies has converged to below a threshold. **b)** Comparison of parametrized GAP model and RIM against exact DFT-derived forces for 50 random configurations drawn from RIM Simulations. Standard deviation of the fit is reported in plot legend parentheses.

tion,^{28,29} we combine configurational sampling with active learning and DFT single point calculations to create an automated workflow that iteratively converges GAP model prediction accuracy to within a targeted threshold. The MLP parametrization workflow developed in this work and applied to molten LiCl is illustrated in Figure 1a). Specifically, a NVT MD simulation using the Born-Mayer-Huggins-Tosi-Fumi potential^{16,17,30,31} is used to melt a 64 atom box of LiCl, with sampling then performed at 1800 K over a density range spanning $\pm 40\%$ of the experimental density. Configurations over this density interval are uniformly sampled and clustered using the HDBSCAN algorithm.³² The clustered configurations are then resampled to identify configurations for labeling (i.e., calculation of energies and forces) via DFT calculations at the PBE+D3 level of theory.^{33–35} This dataset of configurations and DFT-derived labels is then used to fit a GAP model using the Smooth Overlap of Atomic

Positions (SOAP) descriptor,^{36,37} with Bayesian optimization applied to the GAP hyperparameters. This procedure is then iteratively repeated, with the active learning algorithm providing additional configurations for labeling and inclusion in the GAP parametrization until a target accuracy of 5 meV/atom is achieved in the energy predictions. For LiCl this target accuracy is achieved using only 584 training configurations, with Figure 1b) demonstrating the superior performance of the GAP potential as compared to the RIM potential in predicting DFT-derived atomic forces of reference configurations.³⁸ Additional details regarding MD simulations, DFT calculations, active learning,^{28,39} and GAP hyperparameter optimization (Table 3) are provided in the SI.

The LiCl GAP model resulting from the automated parametrization workflow in Figure 1a) is next applied to simulate the structure and dynamics of a larger 1024 atom box of LiCl (512 Li⁺, 512 Cl⁻). GAP MD simulations are performed using the LAMMPS software package compiled with the QUIP pair style.⁴⁰ The simulation box is initialized at the experimental density⁴¹ and equilibrated with NVT MD simulations to a target temperature of 1200 K over 200 ps with a Nose-Hoover thermostat.^{42,43} The volume of the system is then relaxed using a NPT MD simulation⁴⁴⁻⁴⁶ (barostat coupling equal to 1 bar) for 750 ps. The final configurations generated via NPT at 1200 K are then cooled at a rate of 1 K/ps using NPT MD simulations in steps of 100 K until the lowest temperature of interest (900 K) is achieved. Finally, for each target temperature (1200 K, 1100 K, 1000 K, 900 K) configurational sampling is performed for over 8 ns in the NVT ensemble, with the final 7.5 ns of the trajectories being used for calculating the structural and dynamical properties of interest. The diffusion constants derived from our GAP MD simulations (~ 10 nm²/ns) suggest that the Li and Cl nuclei diffuse ~ 10 times the size of an average LiCl bond during our 0.5 ns equilibration time, suggesting that proper system equilibration is achieved.

To validate the performance of the LiCl GAP MLP we compare the structure of GAP MD-derived configurations to those obtained from high energy X-ray diffraction (HEXRD) experiments performed at ~ 900 K. The partial pair distribution functions (PDF) for Li⁺-Li⁺,

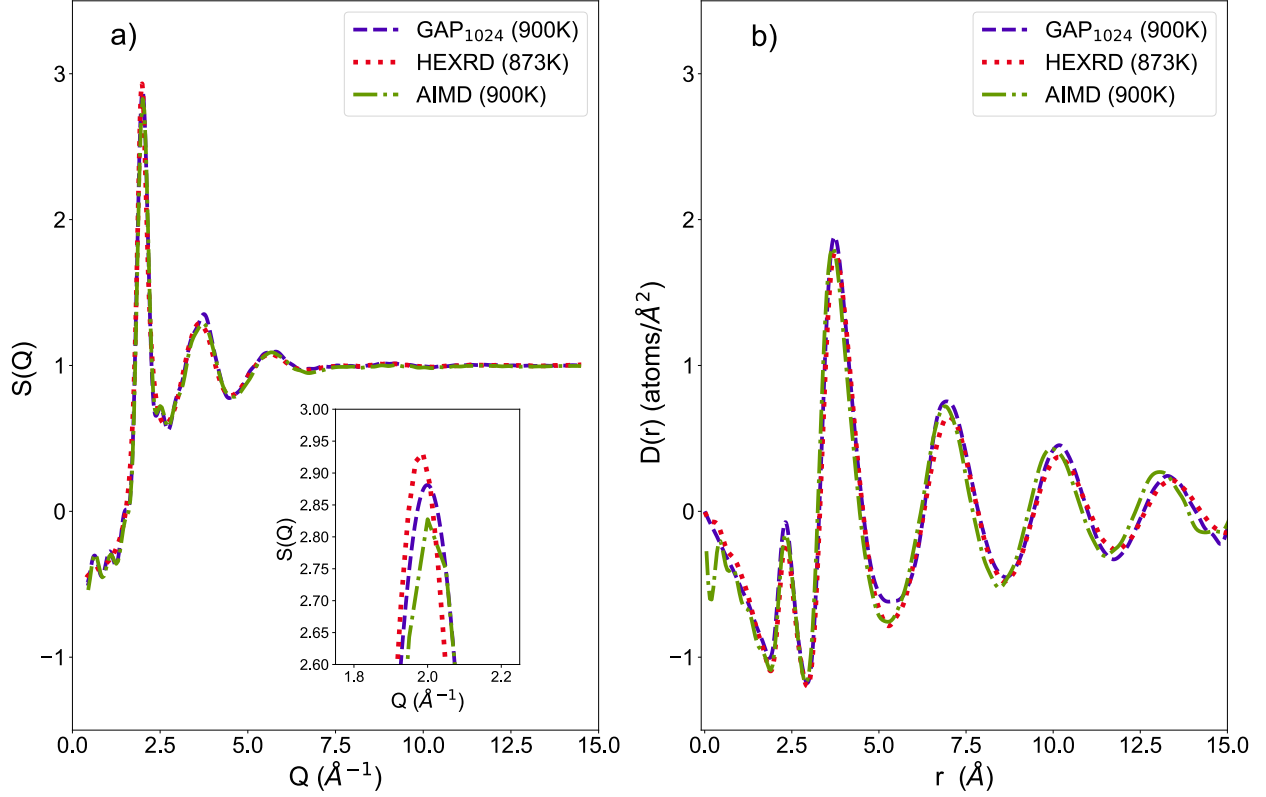


Figure 2: Comparison of structural properties obtained from simulations and experiments.⁴⁷ **a)** Total X-ray structure factor of molten LiCl derived via (blue) GAP MD at 900 K, (red) HEXRD at 873K and (green) conventional AIMD at 900 K. **b)** Total pair distribution function (PDF) of molten LiCl derived via (blue) GAP MD, (red) HEXRD at 873K, and (green) conventional AIMD at 900 K. The subscript in the GAP label refers to the number of atoms used in the simulation.

Li⁺-Cl⁻ and Cl⁻-Cl⁻ atomic pairs are obtained from 900 K GAP MD trajectories of the 1024 atom simulation cell and transformed to the total structure factor, S(Q), via Fourier transformation and use of atomic form factors.^{48,49} The S(Q) obtained via GAP MD simulation is compared in Figure 2a) with that recorded using HEXRD with a transmission geometry at beamline 6-ID-D at the Advance Photon Source (APS),⁴⁷ as well as the results generated with *ab initio* MD simulations.⁴⁷ The GAP MD simulation exhibits excellent agreement with both full AIMD (216 atoms, 15 ps total simulation time) simulation and HEXRD over all Q values. GAP also provides improved accuracy over AIMD in resolving the position of the First Sharp Diffraction Peak (FSDP) at $Q \approx 2 \text{ \AA}^{-1}$, as seen in the insert of Figure 2a). The improved performance at small and intermediate Q values is likely attributable to the significantly larger simulation cell and improved configurational sampling of GAP MD relative to the benchmark AIMD simulation. To quantify the disagreement between experiment and the modeled structure, we calculated the R-factor⁵⁰ (see SI Equation 2 for details) of both AIMD and GAP MD generated results. The GAP MD generated structure showed lower disagreement with respect to experiments at both short ($\sim 4.0\%$) and medium range ($\sim 1.2\%$) compared to that generated from AIMD (Table 1).

In addition to improved predictive accuracy relative to AIMD, GAP MD also exhibits dramatically improved computational scalability, facilitating the simulation of extended spatiotemporal scales. Benchmarking simulations were performed to quantify the speedup associated with GAP MD relative to conventional AIMD. Using two compute nodes a 216 atom LiCl box produced timings of ~ 0.029 seconds/MD timestep and ~ 549 seconds/MD timestep for GAP MD and AIMD, respectively. This constitutes a $\sim 19,000\times$ speedup of GAP MD relative to AIMD, with the additional feature that GAP MD will scale linearly with system size whereas AIMD will scale at best cubically.⁵¹

Provided the significantly improved S(Q)/PDF prediction capability, as well as enhanced computational efficiency, GAP MD was also applied to model the structure of molten LiCl at 1000 K, 1100 K, and 1200 K. The temperature-dependent partial PDF of each atomic

Table 1: R-factor⁵⁰ showing disagreement between experiment and simulations computed using SI Equation 2. Larger R-factor corresponds to more disagreement.

| Range (\AA) | Method | R_X (%) |
|------------------------|--------|-----------|
| $2 < r \leq 5$ | AIMD | 8.31 |
| | GAP | 3.97 |
| $5 < r \leq 14$ | AIMD | 1.98 |
| | GAP | 1.22 |

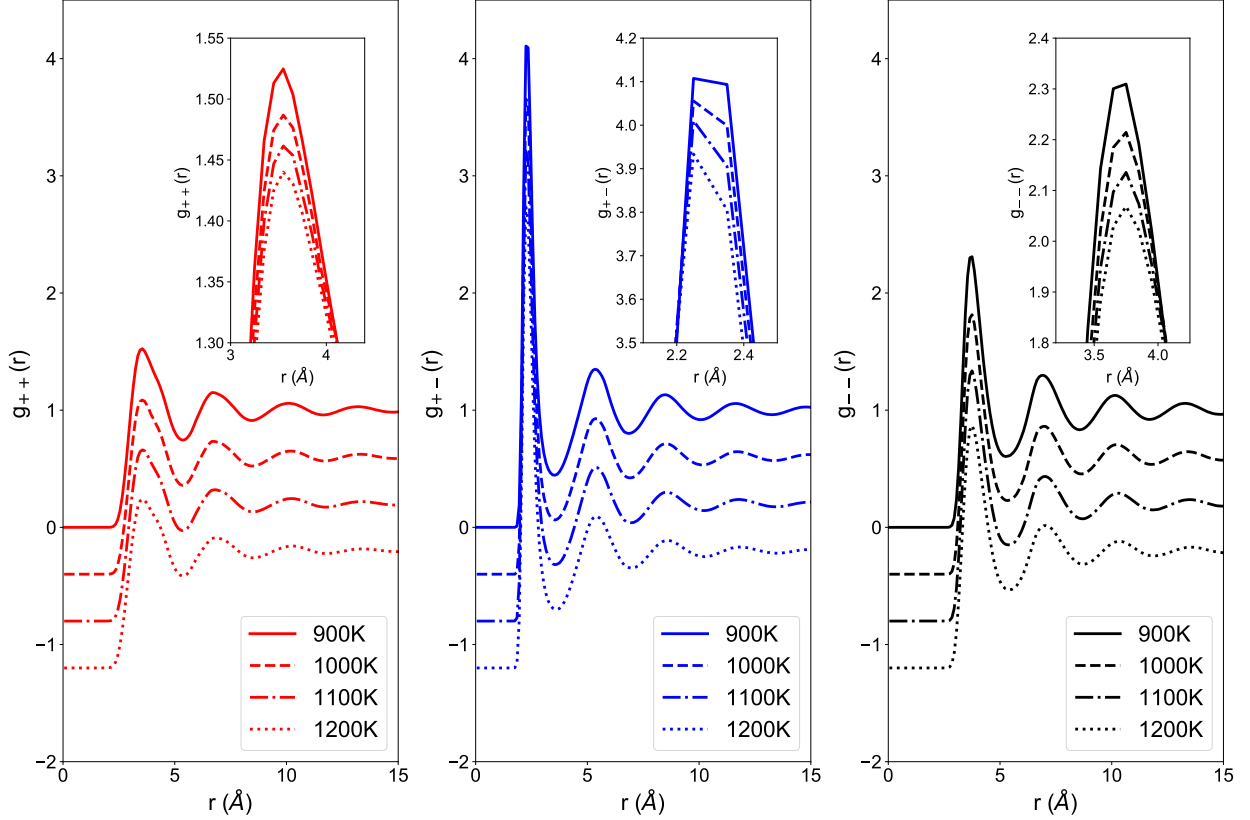


Figure 3: Temperature dependent variation of the total PDF of molten LiCl from 900 K to 1200 K obtained from GAP MD simulations. Curves are shifted vertically for clarity. Inset: zoomed unshifted first peak of PDF.

pair is shown in Figure 3. As expected, the peak intensity for all atomic pairs decreases with increasing temperature, indicating increasing structural disorder. Concomitant with the change in peak intensity are changes in the PDF peak positions, with the distance between like-charged atomic pairs (i.e., Li^+-Li^+ , 3.52 to 3.55 \AA and Cl^--Cl^- , 3.68 to 3.71 \AA) increasing with increasing temperature, and the distance between oppositely charged atomic pairs (i.e., Li^+-Cl^- , 2.30 to 2.29 \AA) decreasing. The shorter average bond length

between Li-Cl pairs is a result of increased binding affinity with increasing temperature (see SI Table 4). This behavior has also been observed in the alkaline chlorides NaCl²¹ and KCl.⁵² In addition to bond length variation, the coordination number between the Li and Cl pairs also decreases from 4.52 to 4.27 as temperature is increased from 900 K to 1200 K (see SI Table 5). These trends agree with previous reports using the reverse Monte Carlo method.^{53,54}

The coordination number of Li-Cl pairs obtained in this study is smaller than that obtained using the RMC method,⁵³ and more importantly is smaller than that of NaCl pairs at similar temperatures,²¹ in contrast to previous work.⁵⁴ The facts that Li is a smaller cation and that the first coordination shell of LiCl is smaller than that of NaCl could lead to the smaller coordination number of Li-Cl pairs compared to that of Na-Cl. Note that the coordination number is a statistical average of neighbouring atoms that fall into the first coordination shell (i.e., from origin to the first minimum on the $g_{\alpha\beta}$ plot): atomic configurations with coordination numbers equal to 4, 5, or 6 co-exist within the melt. The bond angle distribution function of Cl-Li-Cl shows a broad peak near 90° (see SI Figure 5) for all temperatures simulated, indicating that the majority of the structures in molten LiCl retain the octahedral symmetry of the face centered cubic (FCC) crystal structure of its solid phase, which agrees with that reported by McGreevy et al.⁵³

GAP MD simulations across the 900 K–1200 K range were also analyzed to assess the ability of GAP MD to reproduce the experimentally determined temperature-dependent densities of LiCl reported by Janz.⁴¹ The structural prediction accuracy demonstrated in the previous sections manifests in excellent agreement between GAP MD and the experimentally determined densities, with the maximum difference observed over the temperature range being 1.7%. The consistent overprediction of the GAP MD density relative to the experimental density is consistent with previous works observing that PBE+D3 generally overestimates lattice properties.⁵⁵

Next, we use the improved computational efficiency of GAP MD relative to AIMD to

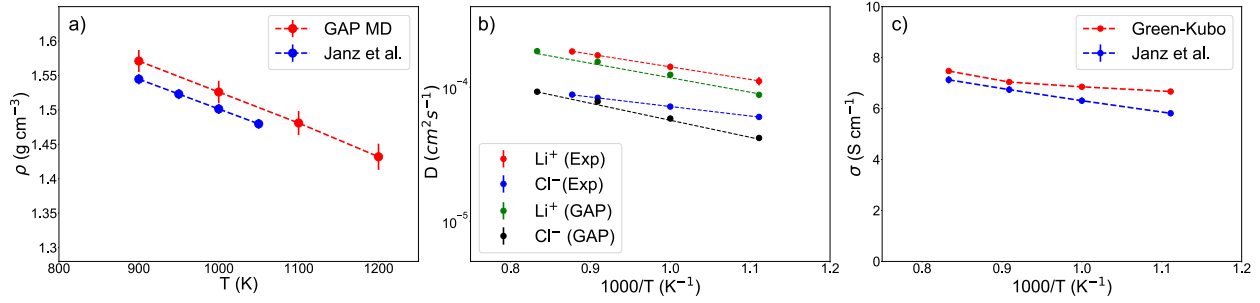


Figure 4: Comparison of temperature-dependent physical properties obtained from GAP MD simulations and experiments.^{41,56–58} **a)** Density, **b)** self-diffusion coefficients for Li⁺, Cl⁻. (Dotted lines are the linear fit between the log scale diffusion coefficient and the reciprocal of temperature) and **c)** ionic conductivities.

examine the dynamical properties of molten LiCl. First we examine the diffusion constants of Li⁺ and Cl⁻, which are important in nuclear fuel reprocessing,⁵⁹ in which Li metal chemically reduces UO₂ to its metal form via the electrochemical reduction of Li⁺, the reaction of which is usually diffusion-limited. The self-diffusion coefficient of Li⁺ and Cl⁻ are computed by fitting the slope of the mean squared displacement vs. time and using the Einstein relation (SI Equation 5). The simulated self-diffusion coefficients are compared with the available experimental data^{56,57} in Figure 4b). The simulated GAP MD self-diffusion coefficient agrees well with experimental values and shows an Arrhenius relationship with the temperature. The activation energy for the self-diffusion of Li⁺ and Cl⁻ is calculated as $E_A = 4.95$ kcal/mol and $E_A = 5.82$ kcal/mol respectively, in line with the experimental results.⁵⁷

Lastly, we compute the ionic conductivity of molten LiCl using the GAP MD trajectories at multiple temperature points. Applications such as liquid metal batteries¹² and pyroprocessing of spent nuclear fuel^{59,60} can benefit from the accurate prediction of the ionic conductivity of LiCl. The ionic conductivity of LiCl is calculated via the Green-Kubo relation (SI Equation 7), which incorporates all ionic correlations in the melt to provide a more realistic estimation of ionic conductivity.⁶¹ The predicted ionic conductivity of LiCl resulting from GAP MD simulations ranges from 6.67 S/cm to 7.47 S/cm over the temperature range 900 K to 1200 K, and is in good agreement with the values reported by Janz.⁵⁸ The accurate prediction of conductivities, in particular, make the approach promising for modeling devices

for thermo or electrochemical energy storage with molten salts.

Many avenues exist for the future improvement of the GAP MD parametrization workflow described in this work. The automated parametrization process does not require generation of expensive AIMD datasets, and the most important configurations are directly sampled from straightforward RIM melt simulations. The active learning identifies that only < 600 single point DFT calculations are required for parametrizing the melt configuration space for fitting the LiCl potential. While PBE+D3-derived forces represent a considerable improvement on those predicted by the RIM model, higher accuracy electronic structure methods are available for parametrization.⁶² While historically the computational cost of advanced hybrid or meta-GGA DFT inhibits their high-throughput use for force-field parametrization, the use of active learning here presents an avenue for their incorporation into the parametrization of MLPs, as the number of required configurations is dramatically reduced relative to random configurational sampling. Further, the sampling strategy proposed in this study can be expanded to account for more complex scenarios traversing thermodynamic configuration and composition space such as modeling mixtures of molten salt melts .

We have developed an automated workflow for the parametrization of molten salt MLP and applied it to examine the structural and dynamical properties of molten LiCl. Parametrized GAP MD MLPs provide significant increases in the spatiotemporal scales of molten salt simulations, here allowing for 1024 atom simulations over ~ 10 ns timescales; this is in stark contrast to typical AIMD simulations using ~ 200 atoms for 10s of picoseconds and represents a conservative four order of magnitude speedup that will improve further for larger systems. The increased system sizes, improved sampling, and near-DFT accuracy of GAP MD leads to excellent agreement with experimental characterizations of densities, structure factors, pair distribution functions, self-diffusion constants, and ionic conductivities for the case of molten LiCl studied here. The developed workflow is extensible to any molten salt chemistry and provides hope for the *in silico* screening of molten salt structure and dynamics in experimentally inaccessible regions of chemical and thermodynamic space.

Acknowledgement

This material is based upon work supported by Laboratory Directed Research and Development (LDRD-2020-0226, LDRD-CLS-1-630) funding from Argonne National Laboratory, provided by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-06CH11357. This research was supported in part by the Exascale Computing Project (17-SC-20-SC) of the U.S. Department of Energy (DOE), by DOE’s Advanced Scientific Research Office (ASCR) under contract DE-AC02-06CH11357. We gratefully acknowledge the computing resources provided on Bebop; a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory. HEXRD measurements were made on beamline 6-ID-D at the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Argonne National Laboratory’s work was supported by the U.S. Department of Energy, Office of Science, under contract DE-AC02-06CH11357. G. S. would like to thank Dr. Anand Narayanan Krishnamoorthy for fruitful discussions on ionic liquids. G. S. would like to thank Samuel Tovey for implementing the Green-Kubo ionic conductivity for the NaCl study. J. G. would like to thank Dr. Zhi-Gang Mei for discussions on AIMD simulations.

Supporting Information Available

RIM Configuration Generation

A 64 atom LiCl box (32 Li^+ , 32 Cl^-) is randomly packed at the experimental density and used for BMHTF RIM simulations.^{16,17} Simulations are performed using LAMMPS.⁴⁰ The initial structure is melted at 3000 K and then equilibrated to a target temperature of 1800 K. Equilibrations are performed in an NVT ensemble using the Nose-Hoover thermostat.^{42,43} Our goal is sample configurations at an elevated temperature (\gg melting temperature) so

as to gain a diverse sampling of melted configurations. Two different simulations are used for sampling at 1800 K: a) simulations at the fixed experimental density b) simulations using a deformation of the cubic box over a density interval that corresponds to $\pm 40\%$ of reference experimental density. Both of the samplings are performed over 12.5 ns with an NVT ensemble and a time step of 0.5 fs. From each of the independent runs 50,000 configurations are drawn and passed to the active learning algorithm for clustering and selection.

DFT

Dispersion-corrected DFT single point calculations are performed using the Vienna Ab initio Simulation Package.⁶³ The Perdew–Burke–Ernzerhof exchange-correlation functional with the “D3” dispersion correction and projector-augmented wave method are employed.^{33–35,64} A large plane wave cutoff of 700 eV with an electronic convergence criterion of 10^{-5} eV is used. A Γ -centered $1 \times 1 \times 1$ k-mesh is used for reciprocal sampling.

Active learning

The pseudocode for the clustering based active learning,²⁸ described below, bears overlap with the previous work of Sivaraman *et al.*:³⁹

Initialization: trajectory, distance measure, target accuracy.

1. Configurations of the input trajectory are featurized by using distance matrices.
2. Perform unsupervised clustering based on HDBSCAN algorithm to obtain uncorrelated clusters. The distance measure of root mean square deviation of atomic positions is used.
3. Training and test configurations are sequentially drawn from the clustering performed on large pool of RIM simulation configuration.
4. Perform single point DFT configuration

5. Perform hyperparameter tuning using Bayesian optimization (BO) as implemented in the GPyOpt library⁶⁵ to get the best GAP model for a chosen set of training configurations as validated against independent test configurations.
6. Draw more samples from the clustered configurations if the GAP target accuracy of < 5 meV/atom has not been achieved and repeat [3-5].
7. Exit if the target accuracy is achieved. For the accuracy we use Mean Absolute Error in GAP predicted energy with respect to DFT in the unit of meV/atom.

The number of training configurations drawn at each iteration depends on the outcome of unsupervised clustering, and hence adjusts automatically with the input chemical system of interest.²⁸ A target accuracy of 5 meV/atom was achieved after four data iterations with fewer than 600 training configurations (Table 2). Validation on independent test configurations is shown in Figure 1b). The detailed code implementation, including the list of SOAP hyperparameters used for Bayesian optimization, range of Bayesian optimization parameter search, clustering criterion, and usage notes are available elsewhere.²⁸

Table 2: Data iteration

| Iteration | Number of configurations |
|--------------|--------------------------|
| 1 | 152 |
| 2 | 107 |
| 3 | 142 |
| 4 | 183 |
| Total | 584 |

GAP fitting

The total energy in a GAP model is expanded as a sum of two body (2B) and many body (MB) terms of local atomic energies (ϵ) as follows:.³⁸

$$E = \delta^{(2B)} \sum_{i < j} \epsilon^{(2B)}(r_{ij}) + \delta^{(MB)} \sum_q \epsilon^{(MB)}(q) \quad (1)$$

The first term describes the pair-wise atomic distance used in the squared exponential kernel. The second term is the rotationally invariant power spectrum vector q that enters the SOAP kernel. The δ s are the scaling parameters for the linear combination. The GAP methodology is described elsewhere.³⁸ The machine learning fit is performed using sparse Gaussian process regression as implemented in the GAP code.³⁶ The hyperparameters are further discussed in the next subsection.

GAP hyperparamters

Table 3: GAP training Parameters

| Parameter | SOAP | Two body |
|------------------------------|------------|----------|
| Cutoff radius (Å) | 6.7 | 6.7 |
| Smooth cutoff transition (Å) | 1.0 | 1.0 |
| sparse method | cur points | uniform |
| Sparse points | 500 | 100 |
| (n_{max}, l_{max}) | (5,5) | - |
| Kernel exponent | 4 | - |
| GAP Version | 1548461341 | |

R-factor

R-factor gives the quantitative disagreement between the experiments and simulation using the relation:⁵⁰

$$R_X = \left(\frac{\sum_i [T_{exp}(r_i) - T_{model}(r_i)]^2}{\sum_i T_{exp}^2(r_i)} \right)^{\frac{1}{2}}, \quad (2)$$

where $T_{exp}(r)$ and $T_{model}(r)$ are the total correlation function ($T(r) = r\rho pr + D(r)$) from experiments and simulations respectively.

Binding affinities of cation-anion pairs

Binding affinities between two species are computed from the pair PDF ($g_{ij}(r)$) using the relation:⁶⁶

$$\Delta F(r) = -k_B T \log(g_{ij}(r)), \quad (3)$$

where k_B is the Boltzmann constant, and T is the temperature.

Table 4: Binding affinity

| Temperature (K) | Li-Cl (Kcal/mol) |
|-----------------|------------------|
| 900 K | -3.31 |
| 1000 K | -3.82 |
| 1100 K | -4.36 |
| 1200 K | -4.96 |

Coordination number

Coordination numbers of atomic pairs in molten LiCl at various temperatures are calculated by integrating the partial PDF, $g_{\alpha\beta}$, over the radius r to the first minimum, r_{min} , as defined in

$$N_{\alpha\beta} = 4\pi\rho_\beta \int_0^{r_{min}} g_{\alpha\beta}(r)r^2 dr, \quad (4)$$

with results listed in Table 5

Table 5: Coordination numbers

| Temperature | Li-Li | Li-Cl | Cl-Cl |
|-------------|-------|-------|-------|
| 900 K | 13.12 | 4.52 | 12.92 |
| 1000 K | 12.74 | 4.46 | 12.64 |
| 1100 K | 12.37 | 4.39 | 12.33 |
| 1200 K | 11.95 | 4.27 | 11.95 |

Self-diffusion coefficient

The self-diffusion coefficient, D , is computed by fitting the slope of mean square displacement vs. time given by Einstein's relation,¹

$$D_\alpha = \frac{1}{6} \lim_{t \rightarrow \infty} \frac{\partial}{\partial t} \langle [r_{i,\alpha}(t) - r_{i,\alpha}(0)]^2 \rangle, \quad (5)$$

where $r_{i,\alpha}$ is the i^{th} ion of species $\alpha = Li^+, Cl^-$.

Ionic conductivity

The Nernst-Einstein relation can be used to estimate the ionic conductivity of a 1:1 symmetric electrolyte at low concentration using the self-diffusion coefficients:⁶⁷

$$\sigma_{N-E} = q^2 \rho \beta (D_{Li^+} + D_{Cl^-}), \quad (6)$$

where β is the inverse temperature, q is the ionic charge and ρ is the number density.

However, the Nernst-Einstein relation shown in Equation 6 does not account for cross correlation between ions.⁶¹ The Green-Kubo relation computed using the current autocorrelation function includes the cross-correlation terms, thereby allowing computation of higher accuracy ionic conductivity.^{68,69}

$$\sigma_{G-C} = \frac{\beta}{3V} \int_{t=0}^{t=\infty} \langle J(0) \cdot J(t) \rangle dt, \quad (7)$$

where V is the volume, $j(t) = \sum_{i=1}^N q_i \cdot v_i(t)$ is the current with the q_i being the ionic charge and $v_i(t)$ velocity at time t . The benchmark of the code implementation is discussed elsewhere.²¹

Angle distribution function

The angular distribution functions of the Cl-Li-Cl bonds in molten LiCl at various temperatures are obtained from the trajectory of the simulated cell and are displayed in Figure 5.

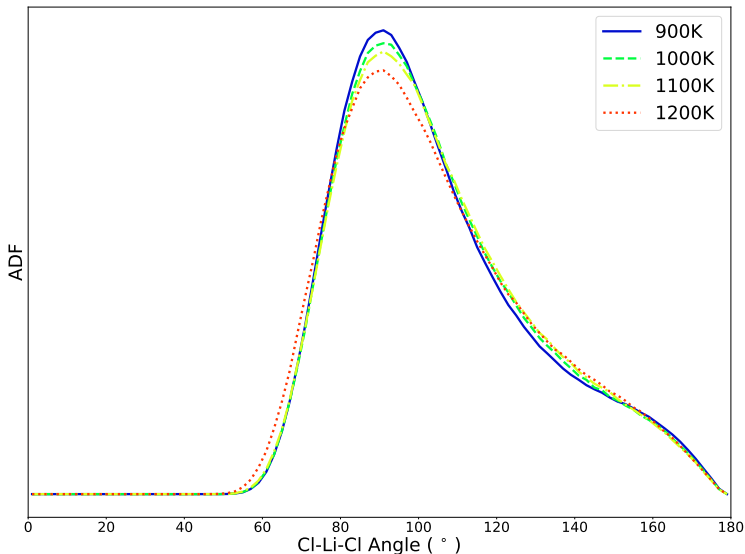


Figure 5: Angle distribution function computed from GAP MD simulations

References

- (1) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press, 2017.
- (2) Carter, E. A. Challenges in modeling materials properties without experimental input. *Science* **2008**, *321*, 800–803.
- (3) Stone, A. J. Intermolecular potentials. *Science* **2008**, *321*, 787–789.
- (4) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical Physics* **2016**, *145*, 170901.
- (5) Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R.

- Machine learning of accurate energy-conserving molecular force fields. *Science Advances* **2017**, *3*, e1603015.
- (6) Huan, T. D.; Batra, R.; Chapman, J.; Krishnan, S.; Chen, L.; Ramprasad, R. A universal strategy for the creation of machine learning-based atomistic force fields. *NPJ Computational Materials* **2017**, *3*, 1–8.
 - (7) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost. *Chemical Science* **2017**, *8*, 3192–3203.
 - (8) Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A. Performance and cost assessment of machine learning interatomic potentials. *The Journal of Physical Chemistry A* **2020**, *124*, 731–745.
 - (9) Jia, W.; Wang, H.; Chen, M.; Lu, D.; Liu, J.; Lin, L.; Car, R.; Zhang, L. Pushing the limit of molecular dynamics with ab initio accuracy to 100 million atoms with machine learning. *arXiv preprint arXiv:2005.00223* **2020**,
 - (10) Thompson, A. P.; Swiler, L. P.; Trott, C. R.; Foiles, S. M.; Tucker, G. J. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *Journal of Computational Physics* **2015**, *285*, 316–330.
 - (11) Dupont, J. From molten salts to ionic liquids: A “nano” journey. *Accounts of Chemical Research* **2011**, *44*, 1223–1231.
 - (12) Kim, H.; Boysen, D. A.; Newhouse, J. M.; Spatocco, B. L.; Chung, B.; Burke, P. J.; Bradwell, D. J.; Jiang, K.; Tomaszowska, A. A.; Wang, K.; Wei, W.; Ortiz, L. A.; Barriga, S. A.; Poizeau, S. M.; Sadoway, D. R. Liquid metal batteries: Past, present, and future. *Chemical Reviews* **2013**, *113*, 2075–2099.

- (13) Guo, J.; Hoyt, N.; Williamson, M. Multielectrode array sensors to enable long-duration corrosion monitoring and control of concentrating solar power systems. *Journal of Electroanalytical Chemistry* **2021**, 115064.
- (14) Mehos, M.; Turchi, C.; Vidal, J.; Wagner, M.; Ma, Z.; Ho, C.; Kolb, W.; Andraka, C.; Kruizenga, A. *Concentrating solar power Gen3 demonstration roadmap*; 2017; National Renewable Energy Lab. (NREL), Golden, CO (United States).
- (15) Serp, J.; Allibert, M.; Beneš, O.; Delpech, S.; Feynberg, O.; Ghetta, V.; Heuer, D.; Holcomb, D.; Ignatiev, V.; Kloosterman, J. L.; Luzz, L.; Merle-Lucotte, E.; Uhlíř, J.; Yoshioka, R.; Zhimin, D. The molten salt reactor (MSR) in generation IV: Overview and perspectives. *Progress in Nuclear Energy* **2014**, 77, 308–319.
- (16) Fumi, F.; Tosi, M. Ionic sizes and born repulsive parameters in the NaCl-type alkali halides—I: The Huggins-Mayer and Pauling forms. *Journal of Physics and Chemistry of Solids* **1964**, 25, 31–43.
- (17) Tosi, M.; Fumi, F. Ionic sizes and born repulsive parameters in the NaCl-type alkali halides—II: The generalized Huggins-Mayer form. *Journal of Physics and Chemistry of Solids* **1964**, 25, 45–52.
- (18) Salanne, M.; Madden, P. A. Polarization effects in ionic solids and melts. *Molecular Physics* **2011**, 109, 2299–2315.
- (19) DeFever, R. S.; Wang, H.; Zhang, Y.; Maginn, E. J. Melting points of alkali chlorides evaluated for a polarizable and non-polarizable model. *The Journal of Chemical Physics* **2020**, 153, 011101.
- (20) Salanne, M.; Rotenberg, B.; Jahn, S.; Vuilleumier, R.; Simon, C.; Madden, P. A. Including many-body effects in models for ionic liquids. *Theoretical Chemistry Accounts* **2012**, 131, 1–16.

- (21) Tovey, S.; Narayanan Krishnamoorthy, A.; Sivaraman, G.; Guo, J.; Benmore, C.; Heuer, A.; Holm, C. DFT accurate interatomic potential for molten NaCl from machine learning. *The Journal of Physical Chemistry C* **2020**, *124*, 25760–25768.
- (22) Jinnouchi, R.; Miwa, K.; Karsai, F.; Kresse, G.; Asahi, R. On-the-fly active learning of interatomic potentials for large-scale atomistic simulations. *The Journal of Physical Chemistry Letters* **2020**, *11*, 6946–6955, PMID: 32787192.
- (23) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. Less is more: Sampling chemical space with active learning. *The Journal of Chemical Physics* **2018**, *148*, 241733.
- (24) Gubaev, K.; Podryabinkin, E. V.; Shapeev, A. V. Machine learning of molecular properties: Locality and active learning. *The Journal of Chemical Physics* **2018**, *148*, 241727.
- (25) Shao, Y.; Knijff, L.; Dietrich, F. M.; Hermansson, K.; Zhang, C. Modelling bulk electrolytes and electrolyte interfaces with atomistic machine learning. *Batteries & Supercaps*
- (26) Settles, B. *Active learning literature survey*; 2009.
- (27) Snoek, J.; Larochelle, H.; Adams, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems*. 2012.
- (28) Sivaraman, G.; Krishnamoorthy, A. N.; Baur, M.; Holm, C.; Stan, M.; Csányi, G.; Benmore, C.; Vázquez-Mayagoitia, Á. Machine-learned interatomic potentials by active learning: amorphous and liquid hafnium dioxide. *npj Computational Materials* **2020**, *6*, 1–8.
- (29) Bernstein, N.; Csányi, G.; Deringer, V. L. De novo exploration and self-guided learning of potential-energy surfaces. *npj Computational Materials* **2019**, *5*.

- (30) Born, M.; Mayer, J. E. Zur gittertheorie der ionenkristalle. *Zeitschrift für Physik* **1932**, *75*, 1–18.
- (31) Huggins, M. L.; Mayer, J. E. Interatomic distances in crystals of the alkali halides. *The Journal of Chemical Physics* **1933**, *1*, 643–646.
- (32) McInnes, L.; Healy, J.; Astels, S. HDBSCAN: Hierarchical density based clustering. *Journal of Open Source Software* **2017**, *2*, 205.
- (33) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Physical Review Letters* **1996**, *77*, 3865.
- (34) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010**, *132*, 154104.
- (35) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011**, *32*, 1456–1465.
- (36) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Physical Review Letters* **2010**, *104*, 136403.
- (37) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Physical Review B* **2013**, *87*, 184115.
- (38) Deringer, V. L.; Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Physical Review B* **2017**, *95*, 094203.
- (39) Sivaraman, G.; Gallington, L.; Krishnamoorthy, A. N.; Stan, M.; Csanyi, G.; Vazquez-Mayagoitia, A.; Benmore, C. J. Experimentally driven automated machine-learned interatomic potential for a refractory oxide. *Physical Review Letters* **2021**,

- (40) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics* **1995**, *117*, 1–19.
- (41) Janz, G. J. *Molten Salts Handbook*; Elsevier, 2013.
- (42) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984**, *81*, 511–519.
- (43) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985**, *31*, 1695.
- (44) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **1981**, *52*, 7182–7190.
- (45) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994**, *101*, 4177–4189.
- (46) Shinoda, W.; Shiga, M.; Mikami, M. Rapid estimation of elastic constants by molecular dynamics simulation under constant stress. *Physical Review B* **2004**, *69*, 134103.
- (47) Guo, J.; Merwin, A.; Benmore, C. J.; Mei, Z.-G.; Hoyt, N. C.; Williamson, M. A. Fluid structure of molten LiCl–Li solutions. *The Journal of Physical Chemistry B* **2019**, *123*, 10036–10043.
- (48) Faber, T.; Ziman, J. A theory of the electrical properties of liquid metals: III. The resistivity of binary alloys. *Philosophical Magazine* **1965**, *11*, 153–173.
- (49) Keen, D. A. A comparison of various commonly used correlation functions for describing total scattering. *Journal of Applied Crystallography* **2001**, *34*, 172–177.
- (50) Wright, A. C. Neutron scattering from vitreous silica. V. The structure of vitreous silica: What have we learned from 60 years of diffraction studies? *Journal of non-crystalline solids* **1994**, *179*, 84–115.

- (51) Deringer, V. L.; Bernstein, N.; Bartók, A. P.; Cliffe, M. J.; Kerber, R. N.; Marbella, L. E.; Grey, C. P.; Elliott, S. R.; Csányi, G. Realistic atomistic structure of amorphous silicon from machine-learning-driven molecular dynamics. *The journal of physical chemistry letters* **2018**, *9*, 2879–2885.
- (52) Walz, M.-M.; Van der Spoel, D. Molten alkali halides—temperature dependence of structure, dynamics and thermodynamics. *Physical Chemistry Chemical Physics* **2019**, *21*, 18516–18524.
- (53) McGreevy, R.; Howe, M. The structure of molten LiCl. *Journal of Physics: Condensed Matter* **1989**, *1*, 9957–9962.
- (54) McGreevy, R.; Pusztai, L. The structure of molten salts. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* **1990**, *430*, 241–261.
- (55) Mosyagin, I.; Gambino, D.; Sangiovanni, D. G.; Abrikosov, I. A.; Caffrey, N. M. Effect of dispersion corrections on ab initio predictions of graphite and diamond properties under pressure. *Physical Review B* **2018**, *98*, 174103.
- (56) Janz, G. J.; Bansal, N. Molten salts data: Diffusion coefficients in single and multi-component salt systems. *Journal of Physical and Chemical Reference Data* **1982**, *11*, 505–693.
- (57) Lenke, R.; Uebelhack, W.; Klemm, A. Selbstdiffusion in geschmolzenem LiCl/Self-diffusion in molten LiCl. *Zeitschrift für Naturforschung A* **1973**, *28*, 881–884.
- (58) Janz, G. J. Thermodynamic and transport properties for molten salts: correlation equations for critically evaluated density, surface tension, electrical conductance, and viscosity data. *Journal of Physical and Chemical Reference Data* **1988**, *17*.
- (59) Merwin, A.; Williamson, M. A.; Willit, J. L.; Chidambaram, D. Metallic lithium and

- the reduction of actinide oxides. *Journal of the Electrochemical Society* **2017**, *164*, H5236.
- (60) Williamson, M.; Willit, J. Pyroprocessing flowsheets for recycling used nuclear fuel. *Nuclear Engineering and Technology* **2011**, *43*, 329–334.
- (61) Shao, Y.; Hellström, M.; Yllö, A.; Mindemark, J.; Hermansson, K.; Behler, J.; Zhang, C. Temperature effects on the ionic conductivity in concentrated alkaline electrolyte solutions. *Physical Chemistry Chemical Physics* **2020**, *22*, 10426–10430.
- (62) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Physical Review Letters* **2015**, *115*, 036402.
- (63) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B* **1996**, *54*, 11169.
- (64) Blöchl, P. E. Projector augmented-wave method. *Physical Review B* **1994**, *50*, 17953.
- (65) González, J.; Dai, Z. GPyOpt: A Bayesian optimization framework in Python. 2016.
- (66) Smiatek, J.; Wohlfarth, A.; Holm, C. The solvation and ion condensation properties for sulfonated polyelectrolytes in different solvents—a computational study. *New Journal of Physics* **2014**, *16*, 025001.
- (67) Michalowsky, J.; Zeman, J.; Holm, C.; Smiatek, J. A polarizable MARTINI model for monovalent ions in aqueous solution. *The Journal of Chemical Physics* **2018**, *149*, 163319.
- (68) Gillan, M. The molecular dynamics calculation of transport coefficients. *Physica Scripta* **1991**, *1991*, 362.
- (69) Walz, M.-M.; van der Spoel, D. Microscopic origins of conductivity in molten salts unraveled by computer simulations. *Communications Chemistry* **2021**, *4*, 1–10.

Graphical TOC Entry

