Supporting Information: Computing chemical potentials with machine-learning-accelerated simulations to accurately predict thermodynamic properties of molten salts

Luke D. Gibson[†*], Rajni Chahal[‡], Vyacheslav S. Bryantsev[‡*]

[†] Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA
[‡] Chemical Science Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA
[*] Email: gibsonld@ornl.gov
[*] Email: bryantsevv@ornl.gov

Machine Learning Force Field (MLFF) Training Details

The liquid-phase LiCl MLFF was trained on-the-fly in VASP with a system comprising 50 LiCl ion pairs in the isothermal-isobaric ensemble (NpT) at 1 bar with the temperature linearly ramped from 900-1400 K over 100 ps. Another 50 ps of on-the-fly training was performed at 1400 K. The final model ultimately included 1,909 frames, from which 2,299 (Li) and 1,544 (Cl) local reference configurations were selected to construct the MLFF. The list of model parameters for the liquid MLFF are provided in **Table S1**. For the liquid phase test set, the mean absolute error (MAE) in energies was 0.59 meV/atom, the root-mean-squared error (RMSE) in forces was 0.077 eV/Å, and the RMSE in stresses was 0.21 kbar (**Figure S1(a)-(c)**). The Li-Cl pairwise radial distribution function predicted by the MLFF is compared against ab initio molecular dynamics (AIMD) and experiment in **Figure 1** of the main text. While the MLIPs developed herein do not consider long-range interactions, previous studies[1, 2] have shown that MLIPs with cutoffs ≤ 8 Å are still able to accurately reproduce several experimental observables of molten salt systems, such as density, heat capacity, thermal conductivity, and viscosity. Consistent with this trend, we note that the accuracy of the MLFF did not appreciably change when increasing the angular and radial cutoff parameters beyond 4.0 and 6.5 Å, respectively. For instance, while testing various hyperparameters, the MAE in energy and RMSE in forces and stress from a model with 4.0 and 6.5 Å cutoffs were 0.63 meV/atom, 0.074 eV/Å, and 0.19 kbar, respectively; while keeping all other parameters the same, increasing the cutoffs to 5.5 and 7.5 Å yielded errors of 0.49 eV/Å, 0.067 eV/Å, and 0.16 kbar, respectively.

The solid-phase LiCl MLFF was trained in a similar manner to the liquid-phase model, but with 108 LiCl ion pairs (3×3×3 supercell of the conventional cell) and with the temperature linearly ramped from 700-900 K over 50 ps. The final model ultimately included 889 frames, from which 3,021 (Li) and 1,986 (Cl) local reference configurations were selected to construct the MLFF. The list of model parameters for the solid MLFF are also provided in **Table S1**. For the solid phase test set, the mean absolute error (MAE) in energies was 1.05 meV/atom, the root-mean-squared error (RMSE) in forces was 0.036 eV/Å, and the RMSE in stresses was 0.16 kbar (**Figure S1(d)-(f)**).

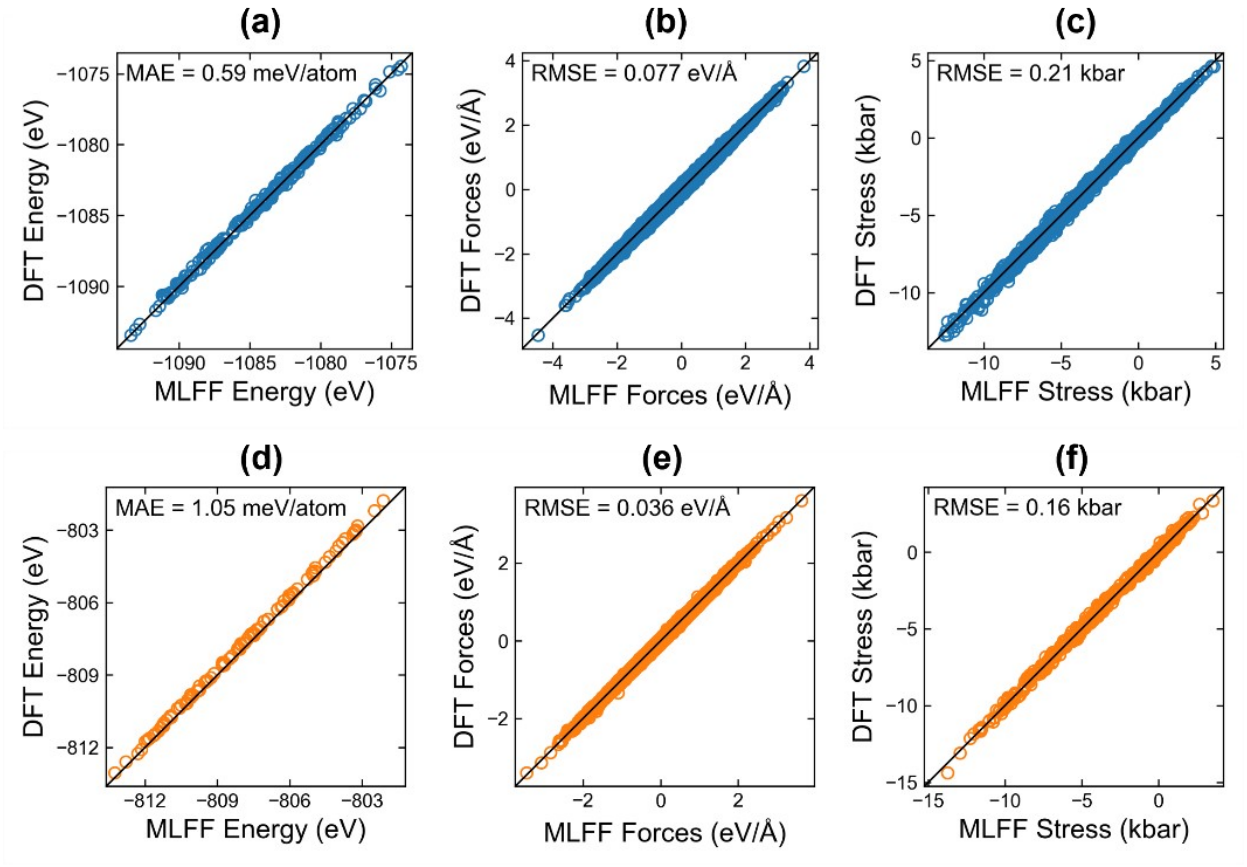

Figure S1: Test set benchmarking of MLFF predictions against DFT for (a, d) energies, (b, e) forces, and (c, f) stresses. (a-c) Liquid phase MLFF and DFT values for 200 uncorrelated configurations sampled from AIMD simulations with 150 LiCl at 893, 943, 993, and 1043 K (50 configurations per temperature). (d-f) Solid phase MLFF and DFT values for 100 uncorrelated configurations sampled from an AIMD NPT simulation with 108 LiCl at 1 bar and temperatures from 700 to 900 K.

Table S1: Model parameters for liquid- and solid-phase LiCl MLFFs. Parameter descriptions can be found in ref [3].

Liquid Phase		**Solid Phase**	
$\beta^{(2)}$	0.5	$\beta^{(2)}$	0.1
$r_{cut}^{(2)}$	6.5 Å	$r_{cut}^{(2)}$	6.5 Å
$N_{basis}^{(2)}$	8	$N_{basis}^{(2)}$	8
$\sigma^{(2)}$	0.5 Å	$\sigma^{(2)}$	0.3 Å
$\beta^{(3)}$	0.5	$\beta^{(3)}$	0.9
$r_{cut}^{(3)}$	4.0 Å	$r_{cut}^{(3)}$	4.0 Å
$N_{basis}^{(3)}$	6	$N_{basis}^{(3)}$	6
$\sigma^{(3)}$	0.5 Å	$\sigma^{(3)}$	0.3 Å
$\zeta^{(3)}$	4	$\zeta^{(3)}$	4

Neural Network Interatomic Potential (NNIP) Description and Training Details

Deep-Pot-Smooth Edition (DeepPot-SE) potential[4] contained inside DeePMD-kit (DP-kit) package (version 2.1.1)[5] was used for training of NNIP interatomic potential on VASP data. From previous investigations of molten salts,[2, 6] DeepPot-SE has been shown to fit a smooth and continuously differentiable potential energy surface. To train the NNIP for molten LiCl, the training dataset comprised 66,082 LiCl (150 pairs) configurations at relaxed volume at 943 K, in addition to ~285,577 configurations with varied densities. Specifically, among varied densities configurations, 50,860 configurations were 7.5% expanded, 53,341 at 7.5% compression, 41,072 at 15% expansion, 36,627 at 15% compression, 41,099 at 25% expansion, and 62,578 configurations were obtained from 25% compression of volumes relative to the experimental densities at 943 K. The inclusion of expanded and compressed configurations from the same system was essential to ensure that the NNIP remains stable and is capable of density predictions.[2, 6] During the training, the entire dataset was shuffled and split with 80% and 20% representing the training and validation sets, respectively. During training, the DeepPot-SE model maps the local environment of each atom within a cut-off (here, 8 Å) to a per-atom energy, such that the sum of atomic energies corresponds to the reference DFT energy obtained from VASP. Thereupon, the gradients of the predicted energies from NNIP are used to compute the atomic forces and both the reference energies and forces are included to evaluate the loss function to be minimized during the training of the NNIP. Here, the smooth cutoff and hard cutoff radius of 2 Å and 8 Å is chosen and the embedding network and fitting network sizes are chosen to be [25, 50, 100] and [240, 240, 240], respectively. For the loss function, the tunable pre-factors for energies and forces were chosen as 0.002, 1000, 1, 1 for p_e^{start}, p_f^{start}, p_e^{limit}, and p_f^{limit}, respectively. Previously, these hyperparameters showed success in accurately predicting the structure and transport properties of a molten salt system.[6] Seed value of 1 is chosen for assigning random values to weights and biases of the neural network when starting the NNIP training. After 800,000 training steps, the MAE in energies was 0.27 meV/atom and the RMSE in forces was 0.035 eV/Å (**Figure S2(a,b)**). For the validation set (20% of dataset not used for training), the MAE in energies was 0.26 meV/atom and the RMSE in forces was 0.035 eV/Å (**Figure S2(c,d)**).

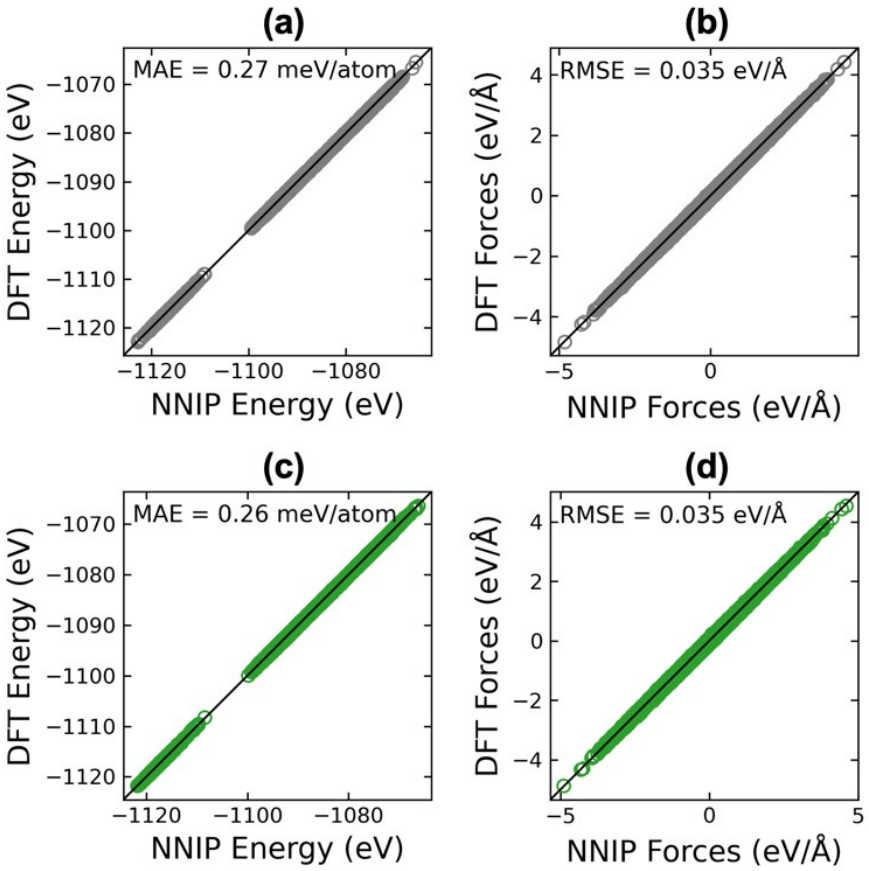

Figure S2: Comparison of DFT and NNIP predictions in (a, c) energies and (b, d) forces for liquid phase structures from the (a, b) training set (80% of initial dataset) and (c, d) validation set (20% of initial dataset).

MLFF Liquid Stability

To assess the stability of our MLFF trained on liquid structures, we performed 2 ns of MD in the NVT ensemble at 943 K with 150 LiCl pairs and analyzed the last 1 ns by computing g(r) from 1 ps segments of the trajectory, which are compared against g(r) from AIMD in **Figure S3(a)**. Following the recommendations of a recent benchmarking study,[7] we also computed Wright's factor and the Jensen-Shannon divergence for the 1,000 g(r) samples (respectively shown in **Figure S3(b)** and **(c)**), which demonstrate the ability of the MLFF to continuously maintain the correct liquid structure.

Wright's Factor, R_χ

$$R_\chi = \frac{\sum \left(g(r) - g_{ref}(r)\right)^2}{\sum \left(g_{ref}(r)\right)^2} \#(S1)$$

Jensen-Shannon Divergence (JSD)

$$JSD\left(g(r)||g_{ref}(r)\right) = \frac{1}{2}\left[KL(g(r)||\hat{g}(r)) + KL\left(g_{ref}(r)||\hat{g}(r)\right)\right] \#(S2)$$

$$\hat{g}(r) = \frac{1}{2}\left[g(r) + g_{ref}(r)\right]$$

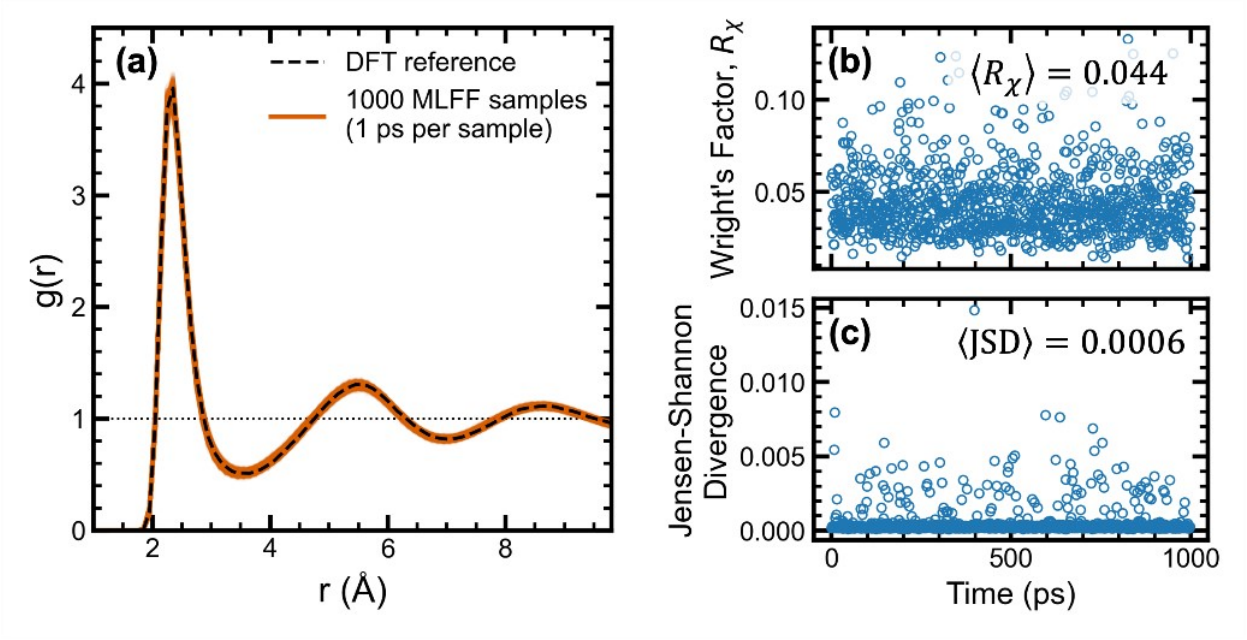

Figure S3: Demonstrating stability of liquid phase MLFF. (a) Comparison of AIMD-computed g(r) against 1,000 samples computed from 1 ps segments (MLFF-MD). (b) Distribution of values for Wright's factor computed on 1,000 sampled g(r) functions. (c) Distribution of values for the Jensen-Shannon divergence of the 1,000 sampled g(r) functions with the AIMD-computed g(r).

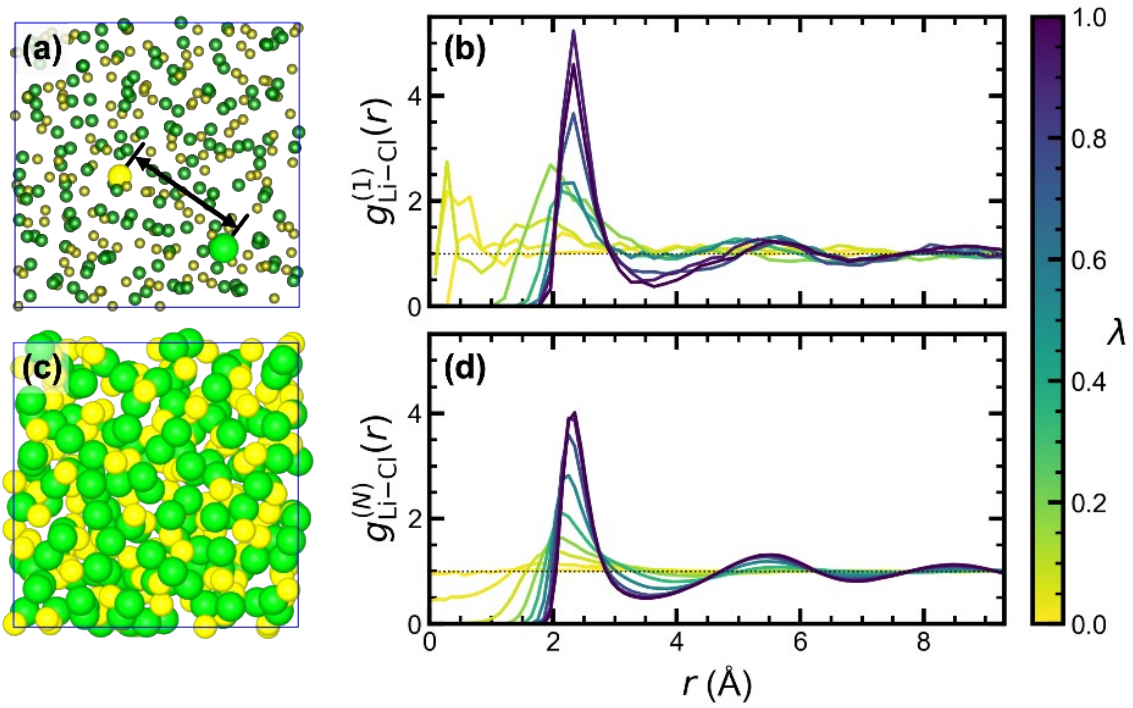

Figure S4: Li-Cl distance distributions of transmuted ions for (a, b) $\Delta N = 1$ and (c, d) $\Delta N = 1$ thermodynamic pathways. Each distribution in (b) and (d) are constructed from the TI window corresponding to λ in the color bar.

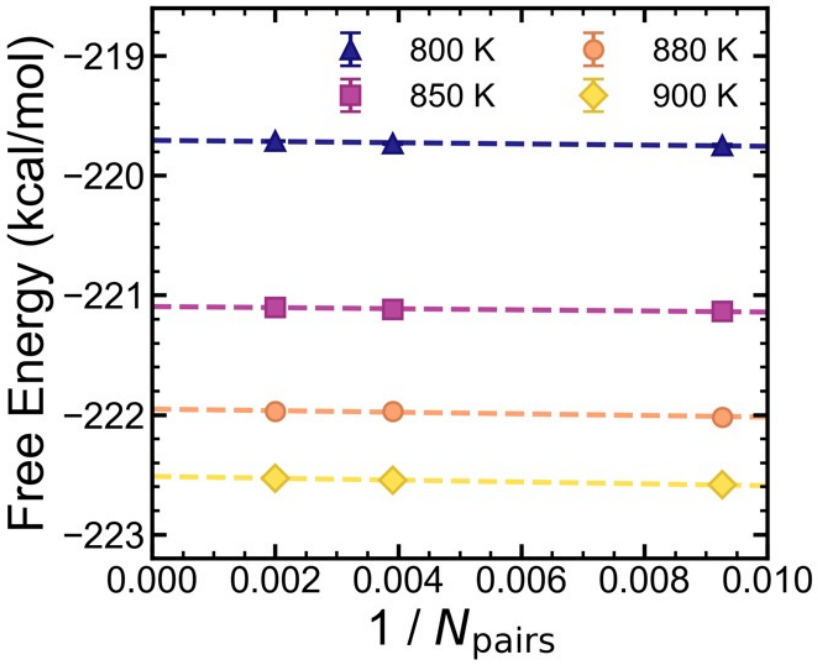

Figure S5: System size scaling of Einstein crystals at each temperature.

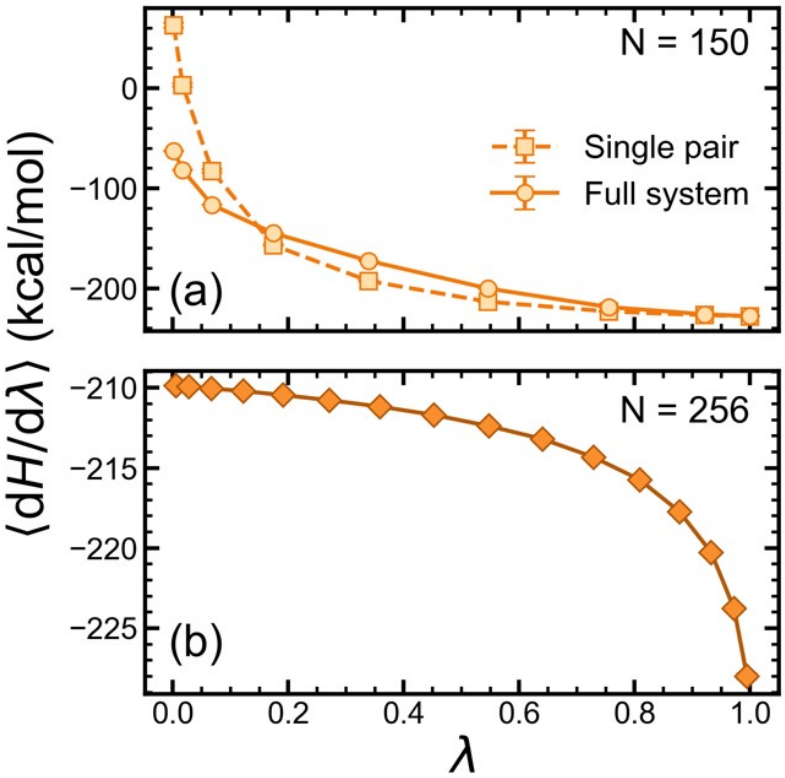

Figure S6: Thermodynamic integration for (a) liquid and (b) solid LiCl phases. Two different approaches are shown for the liquid phase in (a), where either a single LiCl pair or all ions are transformed into ideal gas particles. Thermodynamic integration for the solid phase in (b) involves the transformation from an Einstein crystal ($\lambda = 0$) to the real crystal ($\lambda = 1$).

Table S2: Free energy components from Einstein crystal[8-11] calculations. All free energies are provided in kcal/mol.

T (K)	N_{LiCl}	F_{EC}^{COM}/N_{LiCl}	$\frac{\left(\Delta F_{EC}^{COM} + \Delta F_{RC}^{COM}\right)}{N_{LiCl}}$	$\Delta F_{EC\rightarrow RC}^{COM}/N_{LiCl}$	$\mu_{LiCl_{(s)}}$
800	108	-6.858	-0.172	-212.72 ± 0.01	-219.75 ± 0.01
	256	-6.858	-0.081	-212.794 ± 0.007	-219.732 ± 0.007
	500	-6.858	-0.045	-212.809 ± 0.005	-219.711 ± 0.005
850	108	-7.901	-0.182	-213.05 ± 0.01	-221.13 ± 0.01
	256	-7.901	-0.085	-213.131 ± 0.007	-221.117 ± 0.007
	500	-7.901	-0.047	-213.153 ± 0.005	-221.100 ± 0.005
880	108	-8.543	-0.187	-213.29 ± 0.01	-222.02 ± 0.01
	256	-8.543	-0.088	-213.335 ± 0.008	-221.967 ± 0.008
	500	-8.543	-0.049	-213.360 ± 0.006	-221.952 ± 0.006
900	108	-8.979	-0.191	-213.41 ± 0.01	-222.58 ± 0.01
	256	-8.979	-0.090	-213.782 ± 0.008	-222.545 ± 0.008
	500	-8.979	-0.050	-213.499 ± 0.006	-222.527 ± 0.006

Table S3: Free energy components of solid lithium chloride crystal from method in ref [12]. All free energies are provided in kcal/mol.

T (K)	N_{LiCl}	Perfect Crystal	Harmonic Approximation	Anharmonic Contributions	$\mu_{LiCl_{(s)}}$
700	108	-207.124	-10.947	1.15 ± 0.03	-216.92 ± 0.03
	256	-207.124	-10.633	0.79 ± 0.02	-216.97 ± 0.02
750	108	-206.926	-12.679	1.32 ± 0.03	-218.28 ± 0.03
	256	-206.926	-12.506	1.11 ± 0.02	-218.32 ± 0.02
800	108	-206.756	-14.444	1.54 ± 0.03	-219.66 ± 0.03
	256	-206.756	-14.465	1.51 ± 0.01	-219.71 ± 0.01

Implementations of the Thermodynamic Integration Method in VASP

We note that TI calculations can be performed in VASP using two different implementations when using MLFFs. To utilize the coupling parameter, λ, in **Equation (18)** in the main text, the ML_LCOUPLE keyword must be enabled, the selected subset of atoms must be provided to ML_ICOUPLE, and only then can λ be specified via the ML_RCOUPLE keyword. In this way, TI is performed as described in ref [13], where the integrated quantity, $\partial H(\lambda)/\partial\lambda$, is analytically computed from the derivative of **Equation (18)** in the main text,

$$\frac{\partial H(\lambda)}{\partial \lambda} = \sum_{i \notin M} \frac{\partial U_i(\lambda)}{\partial \lambda} + \sum_{i \in M} U_i(\lambda) + \lambda \sum_{i \in M} \frac{\partial U_i(\lambda)}{\partial \lambda}. \#(S3)$$

However, when operating on all atoms in the system, TI can also be performed by simply setting the SCALEE keyword to λ. In this way, MD simulations can be performed using a hybrid Hamiltonian that couples the real system to either an ideal gas (for liquids) or harmonic crystal (for solids) via $H(\lambda) = \lambda H_{real} + (1-\lambda)H_{ref}$, where H_{ref} is either H_{ideal} or $H_{harmonic}$, as described in ref [12]. Differentiating this Hamiltonian with respect to λ instead becomes,

$$\frac{\partial H(\lambda)}{\partial \lambda} = U_{real} - U_{ref}. \#(S4)$$

In the case of H_{ideal}, to accurately compute chemical potentials, the electronic energy of N isolated Li^+ and N isolated Cl^- ions must be removed from **Equation (S4)** before integration. Harmonic crystal calculations were performed for the 3×3×3 and 4×4×4 supercells by sampling the entire Brillouin zone using the MedeA version 3.7.0 software.[14] Using MLFF-based calculations, we confirmed for several solid- and liquid-phase systems that both approaches yield consistent free energy values. This approach is also compatible with DFT-based calculations and was employed to compute the DFT-based chemical potentials in **Table 2** of the main text.

Uncertainty Quantification

The uncertainties of all reported chemical potentials and free energies are represented by 95% confidence intervals (CI) estimated from bootstrapping. The integrand, $\langle \partial H(\lambda)/\partial\lambda \rangle$, for each TI window was resampled 2,000 times via bootstrapping to construct error bars for plots of

$\langle \partial H(\lambda)/\partial \lambda \rangle$ vs. λ (e.g., **Figure 2c** or **Figure 4** in the main text). These 2,000 samples at each λ were then used to estimate the uncertainty in the resulting free energies. The uncertainty in free energies of solid LiCl extrapolated to infinite system sizes were estimated from the variance of the intercept of the fitted line. Lastly, the uncertainty in the melting point prediction was estimated by propagating the variances in the fitted linear parameters to liquid- and solid-phase free energies in **Figure 5** of the main text.

References

(1) Fayfar, S.; Chahal, R.; Williams, H.; Gardner, D. N.; Zheng, G.; Sprouster, D.; Neuefeind, J. C.; Olds, D.; Hwang, A.; Mcfarlane, J.; et al. Complex Structure of Molten FLiBe (2LiF-BeF2) Examined by Experimental Neutron Scattering, X-Ray Scattering, and Deep-Neural-Network Based Molecular Dynamics. *PRX Energy* **2024**, *3* (1).

(2) Rodriguez, A.; Lam, S.; Hu, M. Thermodynamic and Transport Properties of LiF and FLiBe Molten Salts with Deep Learning Potentials. *ACS Appl Mater Interfaces* **2021**, *13* (46), 55367-55379.

(3) Jinnouchi, R.; Karsai, F.; Kresse, G. On-the-fly machine learning force field generation: Application to melting points. *Phys. Rev. B* **2019**, *100* (1).

(4) Zhang, L. F.; Han, J. Q.; Wang, H.; Saidi, W. A.; Car, R.; E, W. N. End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems. *Adv Neur In* **2018**, *31*.

(5) Wang, H.; Zhang, L. F.; Han, J. Q.; E, W. N. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178-184.

(6) Chahal, R.; Roy, S.; Brehm, M.; Banerjee, S.; Bryantsev, V.; Lam, S. T. Transferable Deep Learning Potential Reveals Intermediate-Range Ordering Effects in LiF-NaF-ZrF4 Molten Salt. *JACS Au* **2022**, *2* (12), 2693-2702.

(7) Bihani, V.; Mannan, S.; Pratiush, U.; Du, T.; Chen, Z. M.; Miret, S.; Micoulaut, M.; Smedskjaer, M. M.; Ranu, S.; Krishnan, N. M. A. EGraFFBench: evaluation of equivariant graph neural network force fields for atomistic simulations. *Digit Discov* **2024**.

(8) Anwar, J.; Frenkel, D.; Noro, M. G. Calculation of the melting point of NaCl by molecular simulation. *J. Chem. Phys.* **2003**, *118* (2), 728-735.

(9) Khanna, V.; Anwar, J.; Frenkel, D.; Doherty, M. F.; Peters, B. Free energies of crystals computed using Einstein crystal with fixed center of mass and differing spring constants. *J. Chem. Phys.* **2021**, *154* (16).

(10) Frenkel, D.; Ladd, A. J. C. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres. *J. Chem. Phys.* **1984**, *81* (7), 3188-3193.

(11) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Elsevier Science, 2001.

(12) Dorner, F.; Sukurma, Z.; Dellago, C.; Kresse, G. Melting Si: Beyond Density Functional Theory. *Phys. Rev. Lett.* **2018**, *121* (19).

(13) Jinnouchi, R.; Karsai, F.; Kresse, G. Making free-energy calculations routine: Combining first principles with machine learning. *Phys. Rev. B* **2020**, *101* (6).

(14) *MedeA*; Materials Design, Inc.: San Diego, CA, USA, 2021.