

# **Supplementary Information:**

## **Accelerating the discovery of disordered multi-component solid-state electrolytes using machine learning interatomic potentials**

Yanhao Deng,<sup>†</sup> Yan Li,<sup>†</sup> Gopalakrishnan Sai Gautam,<sup>‡</sup> Bonan Zhu,<sup>\*,¶,§</sup> and Zeyu

Deng<sup>\*,†</sup>

*†Department of Materials Science and Engineering, National University of Singapore, 9  
Engineering Drive 1, 117575, Singapore*

*‡Department of Materials Engineering, Indian Institute of Science, Bengaluru, Karnataka,  
560012, India*

*¶School of Aerospace Engineering, Beijing Institute of Technology, Beijing, 100081 China.*

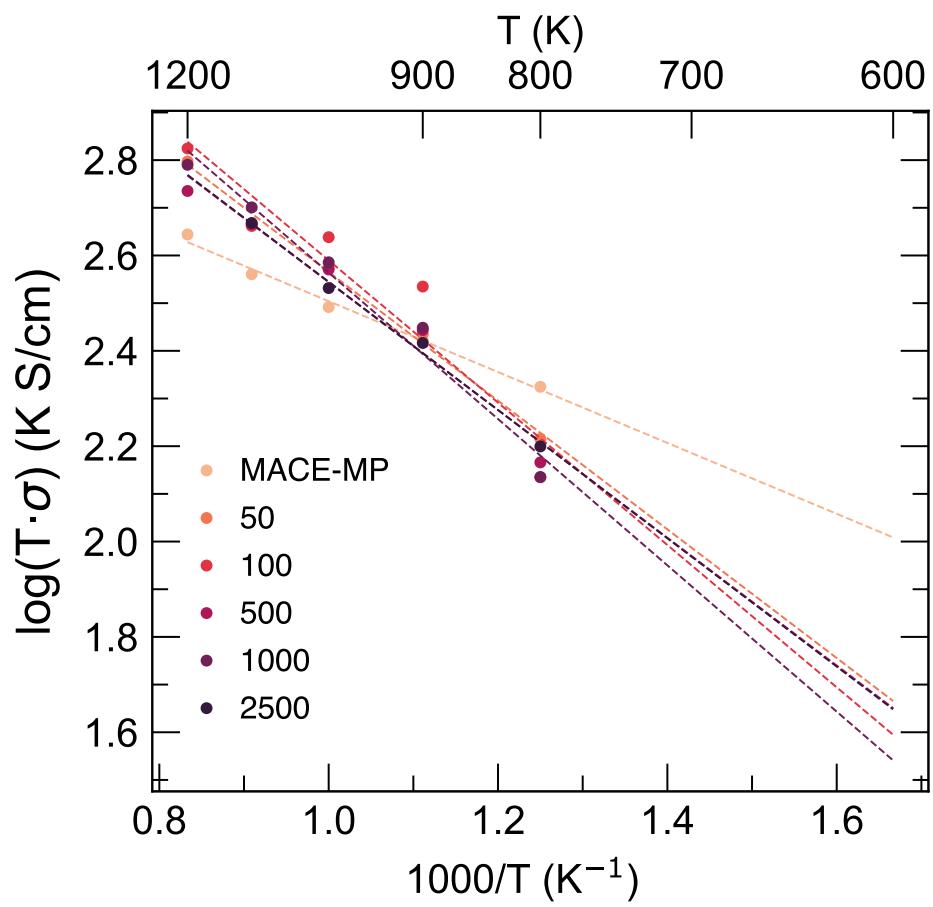
*§State Key Laboratory of Environment Characteristics and Effects for Near-space, Beijing,  
100081 China*

E-mail: bzhu@bit.edu.cn; msedz@nus.edu.sg

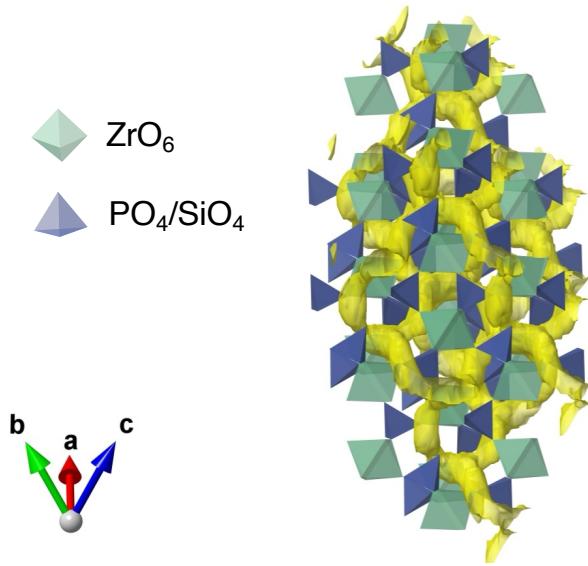
# 1 NaSICON

**Table S1:** Ionic conductivity at 673 K and activation energy of  $\text{Na}_3\text{Zr}_2\text{Si}_2\text{PO}_{12}$ , comparing results obtained from the MACE-MP model and fine-tuned models (trained on different sizes of data) with experimental measurements.

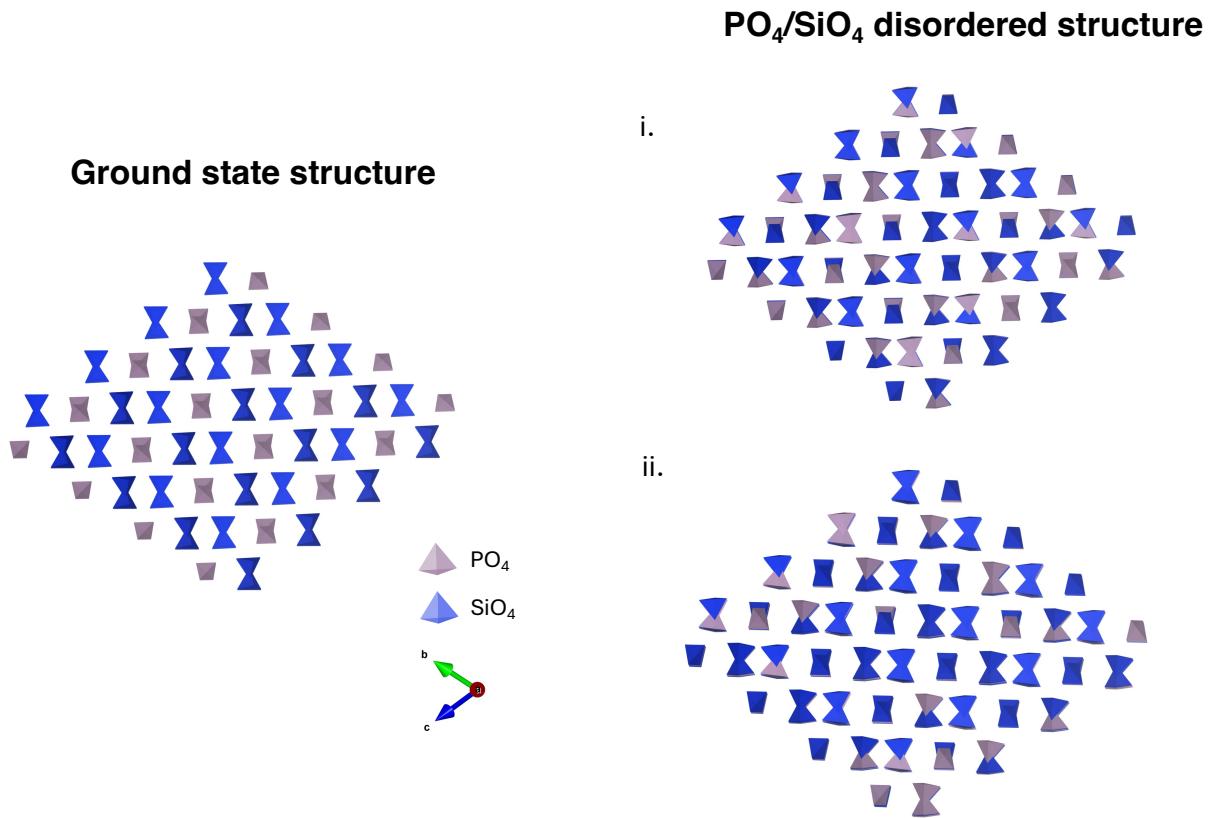
Source	$\sigma_{300^\circ\text{C}}$ (S/cm)	$E_a$ (eV)
MACE-MP	0.170	$0.15 \pm 0.01$
50	0.077	$0.27 \pm 0.01$
100	0.065	$0.30 \pm 0.06$
Fine-tuning	500	0.074
	1000	$0.27 \pm 0.03$
	2500	0.058
	2500	$0.30 \pm 0.03$
experiment <sup>1</sup>	0.080	$0.29 \pm 0.01$
experiment <sup>2</sup>	0.20	0.29
experiment <sup>2</sup>	0.14	0.28



**Fig. S1:** Arrhenius plots of Na-ion diffusion in  $\text{Na}_3\text{Zr}_2\text{Si}_2\text{PO}_{12}$  calculated using fine-tuned models (varying dataset sizes) and the MACE-MP model.

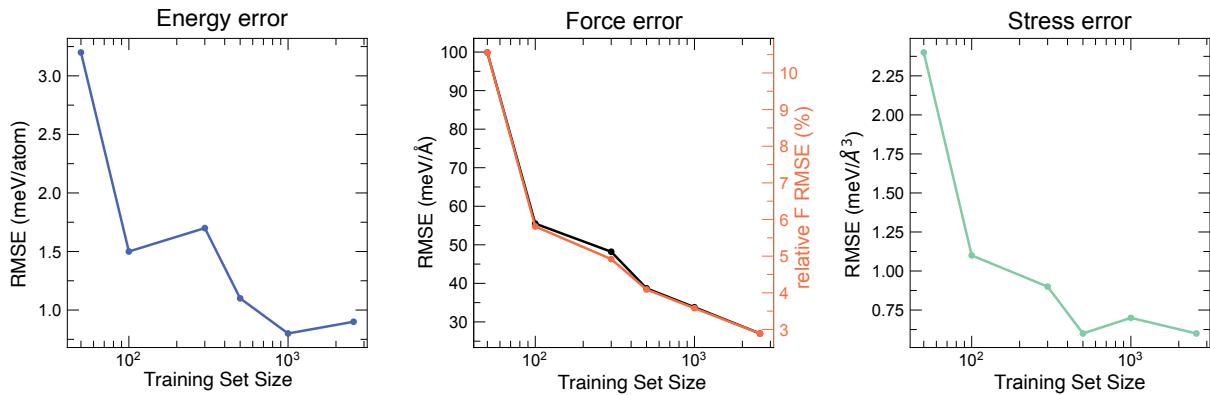


**Fig. S2:** The probability density of Na ions in  $\text{Na}_3\text{Zr}_2\text{Si}_2\text{PO}_{12}$ , analyzed from MLIP-MD trajectory at 800 K using GEMDAT. The isosurface level is set to  $10^{-4} \text{ \AA}^{-1}$ . For better visualization purposes, the Na sites are not plotted.

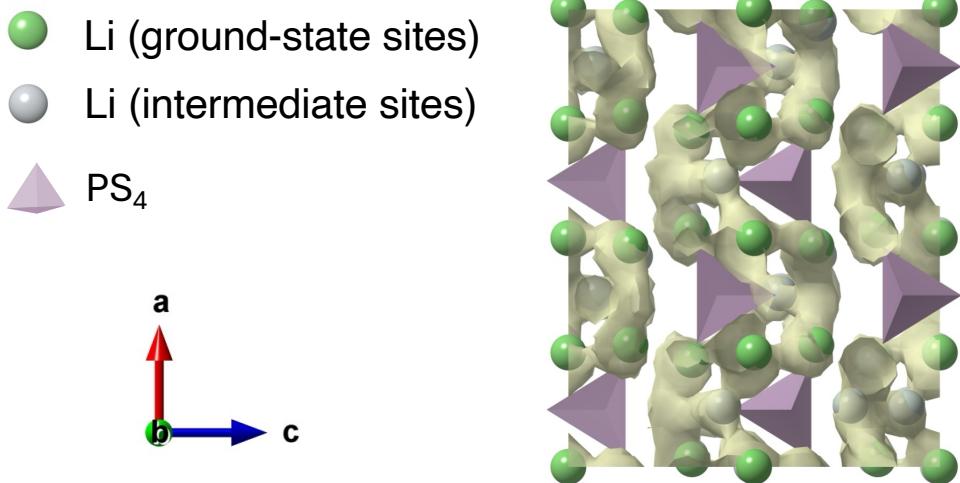


**Fig. S3:** The arrangements of PO<sub>4</sub>/SiO<sub>4</sub> polyhedrals in a 4 × 4 × 4 supercell of Na<sub>3</sub>Zr<sub>2</sub>Si<sub>2</sub>PO<sub>12</sub>. The atoms are not plotted for better visualization. Left: the ground state structure used to run MD in this work. Right: two examples of disordered structures generated by Monte Carlo.

## 2 LPGSO



**Fig. S4:** Convergence test on the training datasets for the Li<sub>3+x</sub>P<sub>1-x</sub>Ge<sub>x</sub>S<sub>4-4x</sub>O<sub>4x</sub> system.

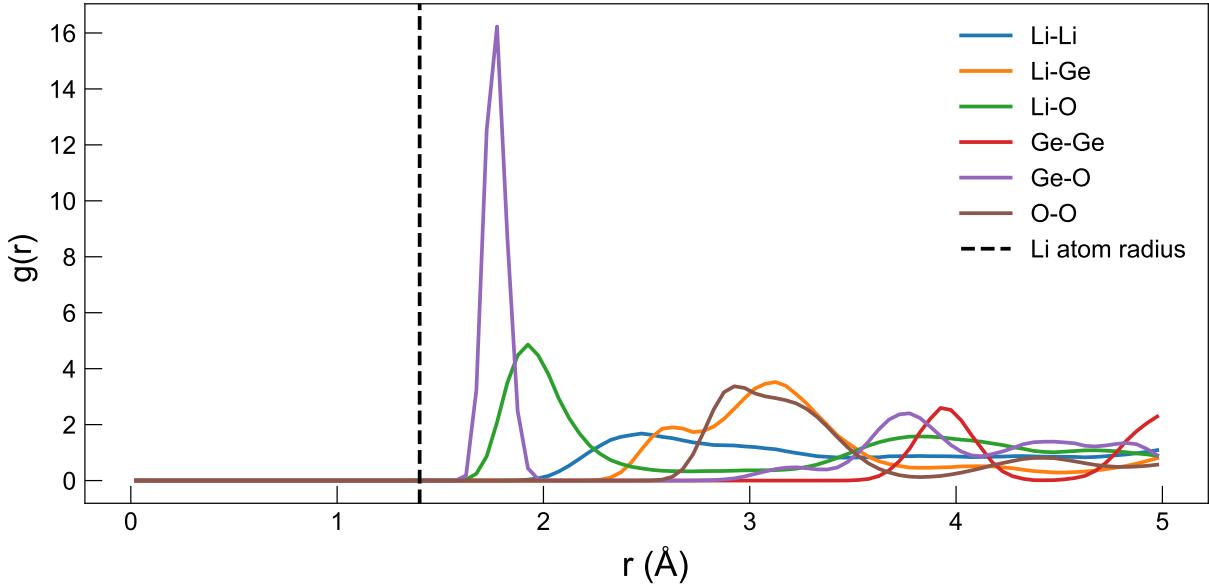


**Fig. S5:** The probability density of Li ions in  $\text{Li}_3\text{PS}_4$ , analyzed from MLIP-MD trajectory at 800 K using GEMDAT. The isosurface level is set to  $10^{-4} \text{ \AA}^{-1}$ . The green spheres are the ground-state sites. The white spheres are the intermediate sites analyzed using Gemdat.

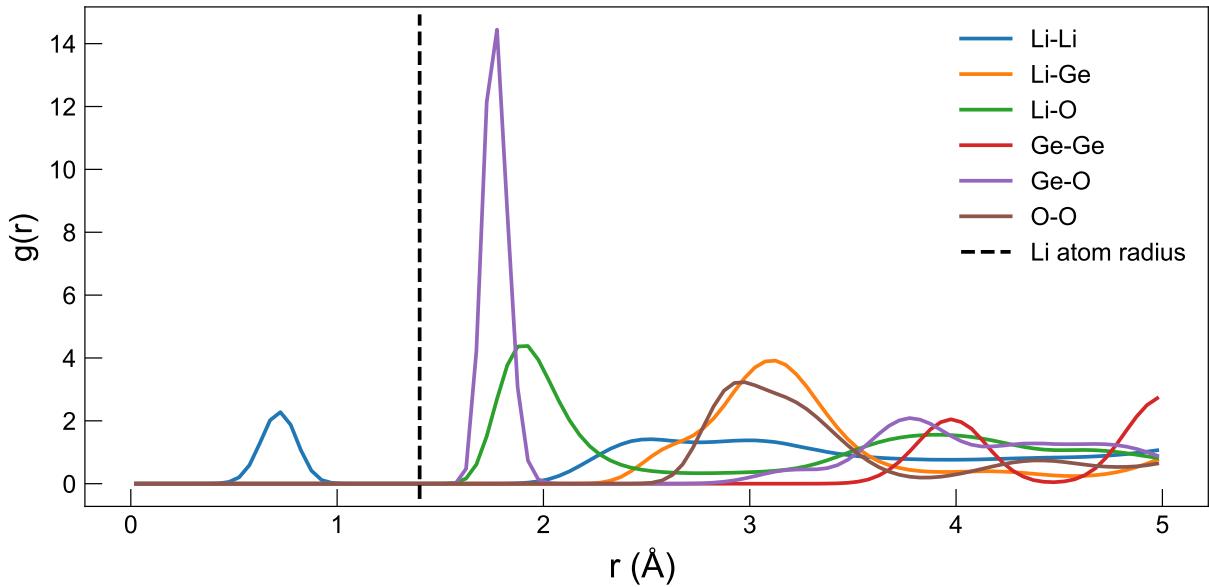
### 3 Comparison between fine-tuning and training from scratch

**Table S2:** Comparison of validation performance between fine-tuning model and the model trained from scratch for the NaSICON and LPGSO systems. The model trained from scratch is based on MACE architecture, using a similar hyperparameter setting for fine-tuning.

Material System	Training Method	RMSE (E/meV/atom)	RMSE (F/meV/ $\text{\AA}$ )
NaSICON	Fine-tuning	0.3	14.4
	Training from scratch	1.2	19.2
LPGSO	Fine-tuning	0.6	36.5
	Training from scratch	0.9	56.5



**Fig. S6:** The pair distribution function  $g(r)$  from correct MD trajectories of  $\text{Li}_4\text{GeO}_4$  at 800 K, simulated by fine-tuned models. The dash line (at 1.4 Å) indicates the atomic radius of Li.



**Fig. S7:** The pair distribution function  $g(r)$  from incorrect MD trajectories of  $\text{Li}_4\text{GeO}_4$  at 800 K, simulated by models trained from scratch. The dash line (at 1.4 Å) indicates the atomic radius of Li.

We compared the performance of fine-tuning and training from scratch for both NaSICON and LPGSO systems. The model trained from scratch adopts the MACE architecture and

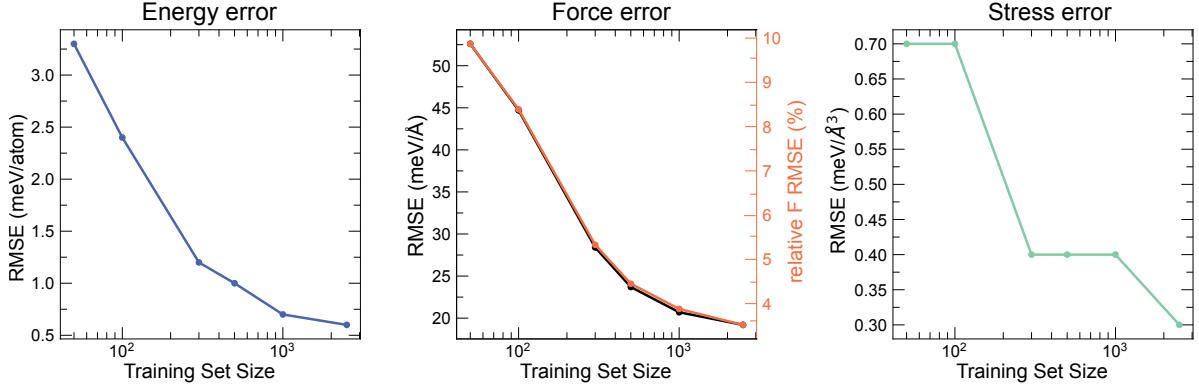
uses hyperparameters consistent with those employed for fine-tuning. Both models were trained on the same dataset with 2,500 data.

While overall validation errors are comparable across methods (Table S2), the fine-tuned model displays considerably reduced energy and force errors, suggesting improved predictive accuracy, for both the NASICON and LPGSO systems. However, in terms of kinetic properties, the model trained from scratch exhibits significant limitations due to its insufficient training data. During MD simulations, this model often leads to unphysical diverging energies and an exploded system, as it encounters atomic configurations outside its training data distribution.

Further evidence for the unreliability of the model trained from scratch is provided by its inaccurate representation of the pair distribution function in LPGSO system, as illustrated in Fig. S7. This unphysical behavior is characterized by abnormal short Li–Li distances, peaking at approximately 0.7 Å. Consequently, such unrealistic Li–Li distances lead to the deviated diffusion pathways of Li ions and meaningless diffusivities.

As a comparison, the MD simulation with a fine-tuned model gives a more reasonable pair distribution function in Fig. S6, where no Li–Li pairs appear with distances smaller than 1 Å. Fine-tuning benefits from prior knowledge embedded in the uMLIPs, allowing it to generalize better and remain stable even in extrapolative regimes, and to work with training data from different functionals.

## 4 LIYBC



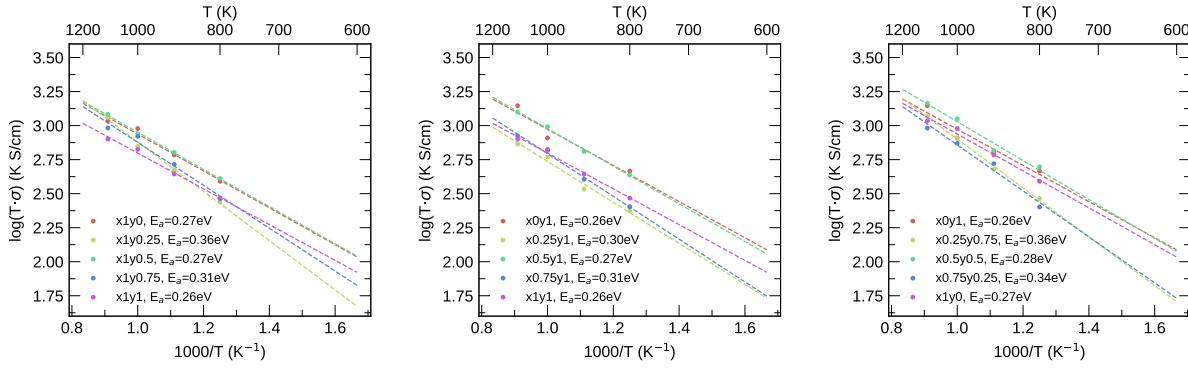
**Fig. S8:** Convergence test on the training datasets for  $\text{Li}_3\text{In}_x\text{Y}_{1-x}\text{Br}_{6y}\text{Cl}_{6-6y}$ .

**Table S3:** Energy above hull ( $E_{\text{hull}}$ ) and activation energy ( $E_a$ ) for  $\text{Li}_3\text{In}_x\text{Y}_{1-x}\text{Br}_{6y}\text{Cl}_{6-6y}$ , as determined by the simulation using the fine-tuned model. “stable” indicates that the composition has a configuration on the convex hull.

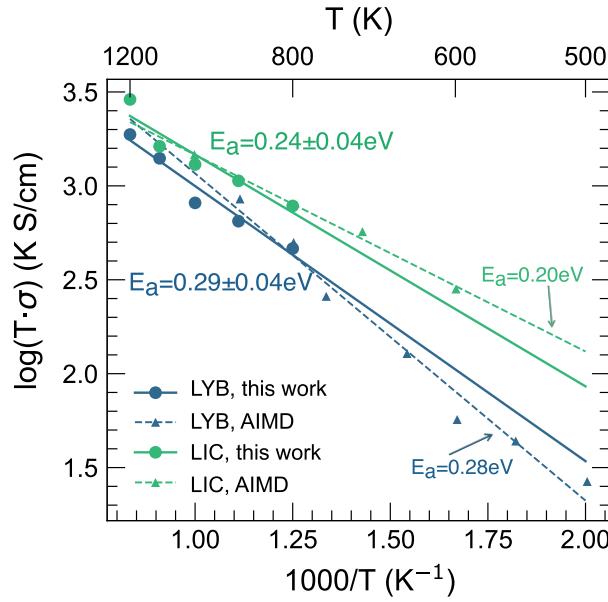
$(x, y)$	Composition	$E_{\text{hull}}$ (meV/atom)	$E_a$ (eV)
(1,1)	$\text{Li}_3\text{InBr}_6$	stable	0.33
(1,0.75)	$\text{Li}_3\text{InBr}_4\text{Cl}_2$	1	0.31
(1,0.5)	$\text{Li}_3\text{InBr}_3\text{Cl}_3$	1	0.23
(1,0.25)	$\text{Li}_3\text{InBr}_2\text{Cl}_4$	1	0.37
(1,0)	$\text{Li}_3\text{InCl}_6$	stable	0.24
(0.75,0.25)	$\text{Li}_3\text{In}_{0.75}\text{Y}_{0.25}\text{Br}_2\text{Cl}_4$	3	0.32
(0.5,0.5)	$\text{Li}_3\text{In}_{0.5}\text{Y}_{0.5}\text{Br}_3\text{Cl}_3$	stable	0.31
(0.25,0.75)	$\text{Li}_3\text{In}_{0.25}\text{Y}_{0.75}\text{Br}_4\text{Cl}_2$	2	0.34
(0,1)	$\text{Li}_3\text{YBr}_6$	stable	0.29
(0.25,1)	$\text{Li}_3\text{In}_{0.25}\text{Y}_{0.75}\text{Br}_6$	1	0.30
(0.5,1)	$\text{Li}_3\text{In}_{0.5}\text{Y}_{0.5}\text{Br}_6$	stable	0.28
(0.75,1)	$\text{Li}_3\text{In}_{0.75}\text{Y}_{0.25}\text{Br}_6$	1	0.26

**Table S4:** Lattice parameters of crystal structures for each composition in  $\text{Li}_3\text{In}_x\text{Y}_{1-x}\text{Br}_{6y}\text{Cl}_{6-6y}$ , optimized by the fine-tuned model.

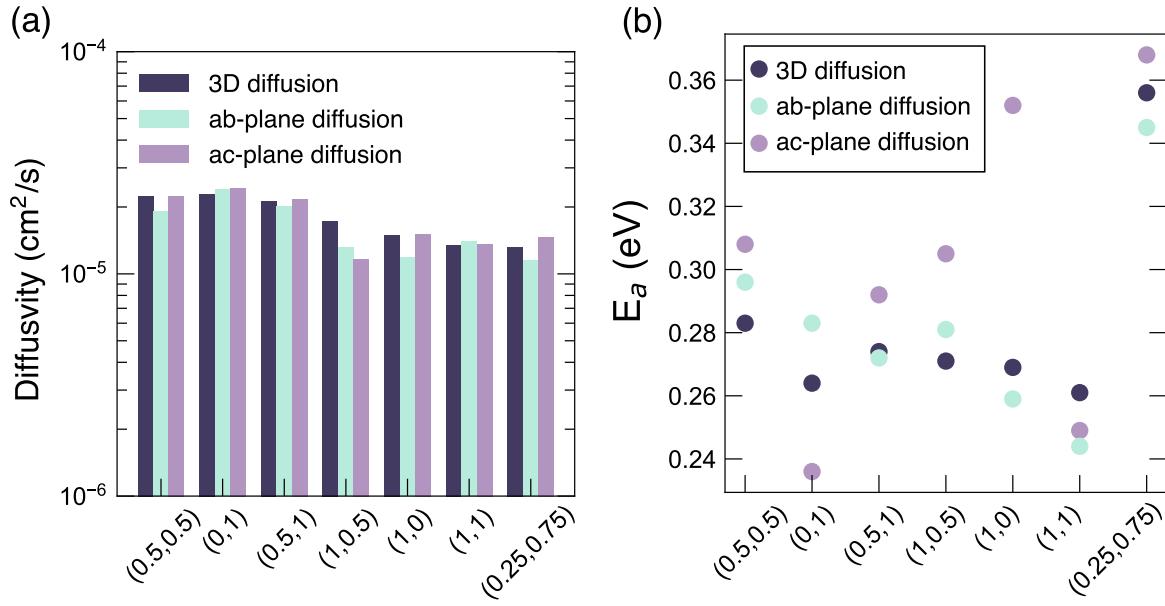
$(x, y)$	Composition	a (Å)	b (Å)	c (Å)	$\alpha$ (°)	$\beta$ (°)	$\gamma$ (°)
(1,1)	$\text{Li}_{12}\text{In}_4\text{Br}_{24}$	6.777	11.716	13.612	90.0	108.3	90.0
(1,0.75)	$\text{Li}_6\text{In}_2\text{Br}_8\text{Cl}_4$	6.755	6.708	11.556	90.0	90.3	109.4
(1,0.5)	$\text{Li}_6\text{In}_2\text{Br}_6\text{Cl}_6$	6.641	6.601	11.344	90.0	90.0	109.1
(1,0.25)	$\text{Li}_6\text{In}_2\text{Br}_4\text{Cl}_8$	6.570	6.528	11.170	90.1	90.7	109.9
(1,0)	$\text{Li}_6\text{In}_2\text{Cl}_{12}$	6.419	11.028	6.423	90.0	109.2	90.0
(0.75,0.25)	$\text{Li}_9\text{In}_3\text{YBr}_6\text{Cl}_{12}$	6.808	11.760	13.081	90.0	79.7	90.0
(0.5,0.5)	$\text{Li}_6\text{InYBr}_6\text{Cl}_6$	6.665	6.676	11.418	90.0	90.0	109.9
(0.25,0.75)	$\text{Li}_{12}\text{InY}_3\text{Br}_{18}\text{Cl}_6$	6.800	11.713	12.864	89.5	80.0	89.8
(0,1)	$\text{Li}_6\text{Y}_2\text{Br}_{12}$	6.892	11.909	6.845	90.0	109.8	90.0
(0.25,1)	$\text{Li}_4\text{InY}_3\text{Br}_{24}$	6.865	11.860	13.068	90.0	80.0	90.0
(0.5,1)	$\text{Li}_{12}\text{In}_2\text{Y}_2\text{Br}_{24}$	6.845	11.787	13.053	90.0	79.7	90.0
(0.75,1)	$\text{Li}_{12}\text{In}_3\text{YBr}_{24}$	6.808	11.760	13.081	90.0	79.7	90.0



**Fig. S9:** Arrhenius plots for the compositions in  $\text{Li}_3\text{In}_x\text{Y}_{1-x}\text{Br}_{6y}\text{Cl}_{6-6y}$ . Left:  $\text{Li}_3\text{InBr}_{6y}\text{Cl}_{6-6y}$ ; Middle:  $\text{Li}_3\text{In}_x\text{Y}_{1-x}\text{Br}_6$ ; Right:  $\text{Li}_3\text{In}_x\text{Y}_{1-x}\text{Br}_{6y}\text{Cl}_{6-6y}$  with  $x + y = 1$ .



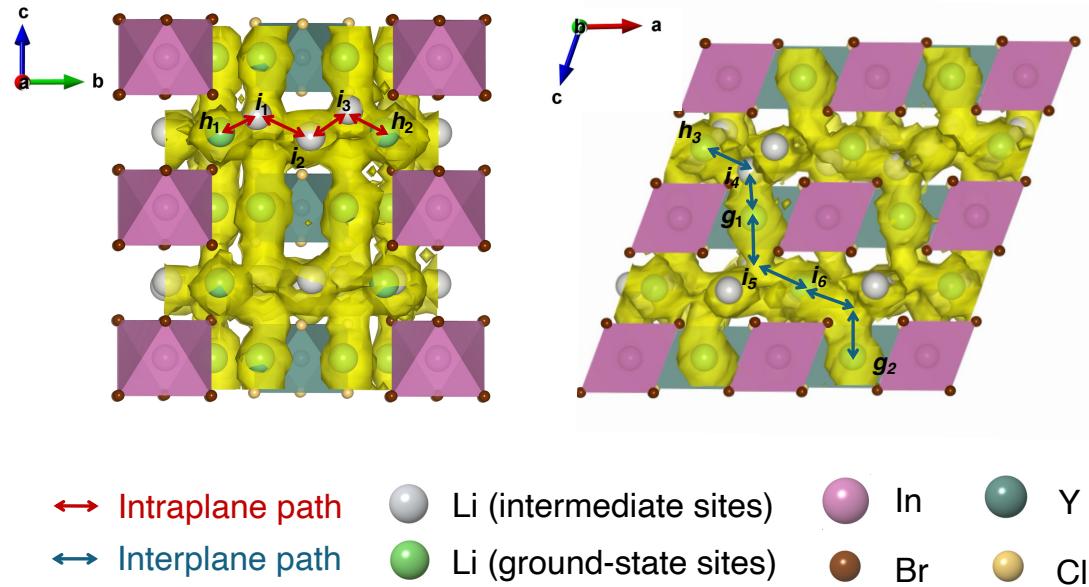
**Fig. S10:** Arrhenius plots of  $\text{Li}_3\text{InCl}_6$  and  $\text{Li}_3\text{YBr}_6$ , compared with AIMD results.<sup>3,4</sup>



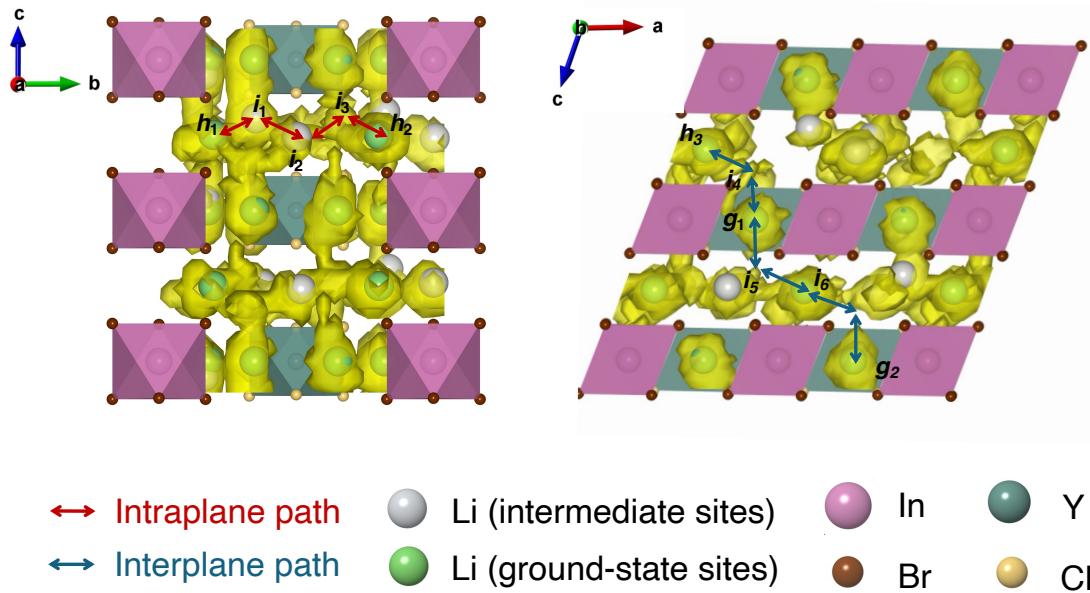
**Fig. S11:** (a) Computed total diffusivity, and diffusivities within  $ab$ - and  $ac$ -planes, and (b) the corresponding activation energies for various compositions of  $\text{Li}_3\text{In}_x\text{Y}_{1-x}\text{Br}_6\text{Cl}_{6-y}$ , labeled as  $(x, y)$ .

**Table S5:** The diffusivity  $D$  and activation energy  $E_a$  projected in ab-plane and ac-plane for each composition in LIYBC system.

$(x, y)$	Composition	ab-plane		ac-plane	
		$D$ (cm $^2$ /s)	$E_a$ (eV)	$D$ (cm $^2$ /s)	$E_a$ (eV)
(1,1)	$\text{Li}_3\text{InBr}_6$	1.41e-05	$0.249 \pm 0.043$	1.36e-05	$0.244 \pm 0.012$
(1,0.75)	$\text{Li}_3\text{InBr}_4\text{Cl}_2$	1.18e-05	$0.292 \pm 0.047$	1.33e-05	$0.335 \pm 0.051$
(1,0.5)	$\text{Li}_3\text{InBr}_3\text{Cl}_3$	1.31e-05	$0.305 \pm 0.037$	1.16e-05	$0.281 \pm 0.041$
(1,0.25)	$\text{Li}_3\text{InBr}_2\text{Cl}_4$	1.04e-05	$0.402 \pm 0.044$	1.11e-05	$0.323 \pm 0.017$
(1,0)	$\text{Li}_3\text{InCl}_6$	1.19e-05	$0.352 \pm 0.044$	1.51e-05	$0.259 \pm 0.053$
(0.75,0.25)	$\text{Li}_3\text{In}_{0.75}\text{Y}_{0.25}\text{Br}_2\text{Cl}_4$	9.16e-06	$0.374 \pm 0.001$	1.04e-05	$0.296 \pm 0.068$
(0.5,0.5)	$\text{Li}_3\text{In}_{0.5}\text{Y}_{0.5}\text{Br}_3\text{Cl}_3$	1.91e-05	$0.308 \pm 0.023$	2.24e-05	$0.296 \pm 0.054$
(0.25,0.75)	$\text{Li}_3\text{In}_{0.25}\text{Y}_{0.75}\text{Br}_4\text{Cl}_2$	1.15e-05	$0.368 \pm 0.045$	1.46e-05	$0.345 \pm 0.002$
(0,1)	$\text{Li}_3\text{YBr}_6$	2.40e-05	$0.236 \pm 0.040$	2.42e-05	$0.283 \pm 0.081$
(0.25,1)	$\text{Li}_3\text{In}_{0.25}\text{Y}_{0.75}\text{Br}_6$	9.81e-06	$0.295 \pm 0.034$	1.20e-05	$0.288 \pm 0.015$
(0.5,1)	$\text{Li}_3\text{In}_{0.5}\text{Y}_{0.5}\text{Br}_6$	2.02e-05	$0.292 \pm 0.004$	2.17e-05	$0.272 \pm 0.033$
(0.75,1)	$\text{Li}_3\text{In}_{0.75}\text{Y}_{0.25}\text{Br}_6$	1.28e-05	$0.272 \pm 0.026$	1.12e-05	$0.331 \pm 0.008$

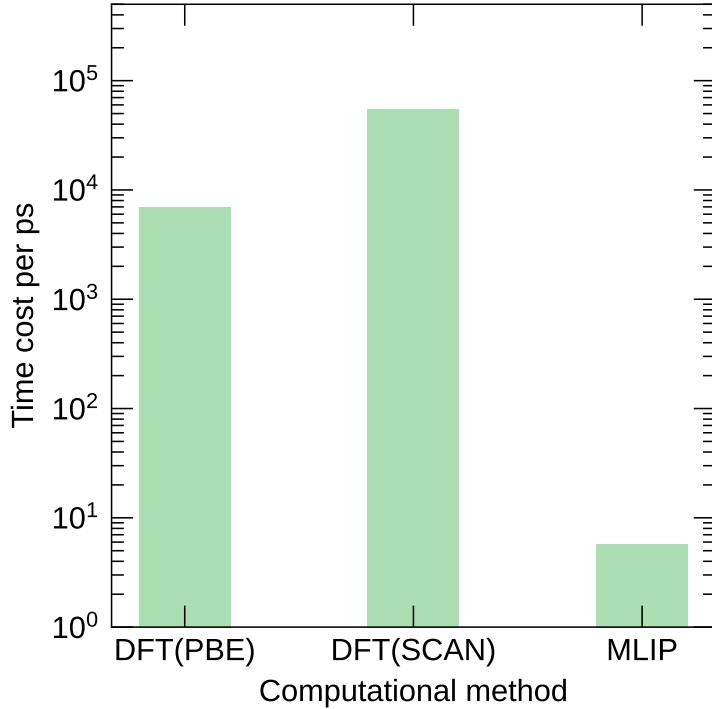


**Fig. S12:** The probability density of Li ions in  $\text{Li}_3\text{In}_{0.5}\text{Y}_{0.5}\text{Br}_3\text{Cl}_3$ , analyzed from a 150 ps-AIMD trajectory at 800 K using GEMDAT. PBE functional is used. The isosurface level is set to  $10^{-4}$  Å $^{-1}$ . The left panel shows the view along the  $a$ -axis, and the right panel shows the view along the  $b$ -axis. Green spheres represent host Li sites in the ground-state structure, while white spheres indicate natural interstitial sites identified by GEMDAT, which correspond to high-probability Li-ion positions during the MD simulations. Two distinct migration paths are identified from the analysis, indicated by the arrows.



**Fig. S13:** The probability density of Li ions in  $\text{Li}_3\text{In}_{0.5}\text{Y}_{0.5}\text{Br}_3\text{Cl}_3$ , analyzed from a 11 ps-AIMD trajectory at 800 K using GEMDAT. SCAN functional is used. The isosurface level is set to  $10^{-4} \text{ \AA}^{-1}$ . The left panel shows the view along the  $a$ -axis, and the right panel shows the view along the  $b$ -axis. Green spheres represent host Li sites in the ground-state structure, while white spheres indicate natural interstitial sites identified by GEMDAT, which correspond to high-probability Li-ion positions during the MD simulations. Two distinct migration paths are identified from the analysis, indicated by the arrows.

## 5 Discussion



**Fig. S14:** A comparison of the computational cost for each ps between AIMD and MLIP-MD. All AIMD simulations were performed on CPUs using two nodes (each equipped with two of Intel 8452Y) with 72 CPU cores, while MLIP simulations were conducted on NVIDIA A40 GPU.

From Figure S14, AIMD simulations require 3–4 orders of magnitude more computational time than MLIP-MD. Moreover, as the number of atoms increases, the CPU-based AIMD times grow approximately linearly, whereas the GPU-based MLIP times do not scale strictly linearly. The A40 GPU has not yet reached its performance limit, and the H100 GPU appears capable of handling even larger systems without a significant increase in time per picosecond.

## References

- (1) Goodenough, J. B.; Hong, H. Y.-P.; Kafalas, J. A. Fast Na+-ion transport in skeleton structures. *Materials Research Bulletin* **1976**, 11, 203–220.
- (2) Deng, Z.; Mishra, T. P.; Mahayoni, E.; Ma, Q.; Tieu, A. J. K.; Guillon, O.; Chotard, J.-N.; Seznec, V.; Cheetham, A. K.; Masquelier, C.; Gautam, G. S.; Canepa, P. Fundamental investigations on the sodium-ion transport properties of mixed polyanion solid-state battery electrolytes. *Nature Communications* **2022**, 13, 4470, Publisher: Nature Publishing Group.
- (3) Li, X.; Liang, J.; Yang, X.; Adair, K. R.; Wang, C.; Zhao, F.; Sun, X. Progress and perspectives on halide lithium conductors for all-solid-state lithium batteries. *Energy & Environmental Science* **2020**, 13, 1429–1461, Publisher: The Royal Society of Chemistry.
- (4) Park, D.; Park, H.; Lee, Y.; Kim, S.-O.; Jung, H.-G.; Chung, K. Y.; Shim, J. H.; Yu, S. Theoretical Design of Lithium Chloride Superionic Conductors for All-Solid-State High-Voltage Lithium-Ion Batteries. *ACS Applied Materials & Interfaces* **2020**, 12, 34806–34814.