

Supplementary Information

Molecular Dynamics (MD) Simulations

The MD workflow for generating pre-training datasets is shown in Figure S1. Using empirical force fields such as OPLS-AA [1], various molecular systems can be simulated efficiently to generate structural data.

Different solutions of ions and electrolytes were used to build the electrolyte dataset. As shown in Figure S2 and Figure S3, we implemented MD simulations with OPLS-AA force field on 16 different combinations. The components include $[\text{Na}]^+$ or $[\text{Li}]^+$ cations, $[\text{PF}_6]^-$ or $[\text{Tf}_2\text{N}]^-$ anions, and DME or 1:1 EC+DMC solvents.

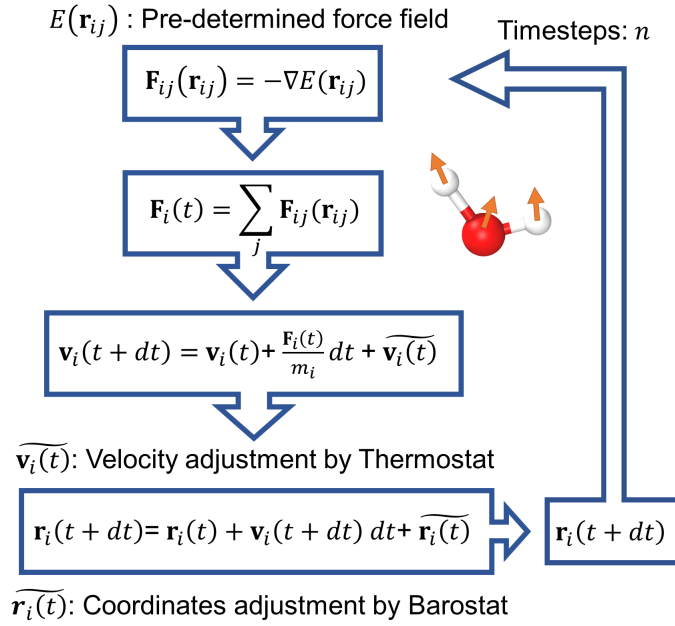


Fig. S1 A schematic representation of MD workflow that generates pre-training datasets. The potential energies as well as interatomic forces are calculated by a pre-determined force field. Then the coordinates of atoms are iteratively changed according to Newton's law of motion.

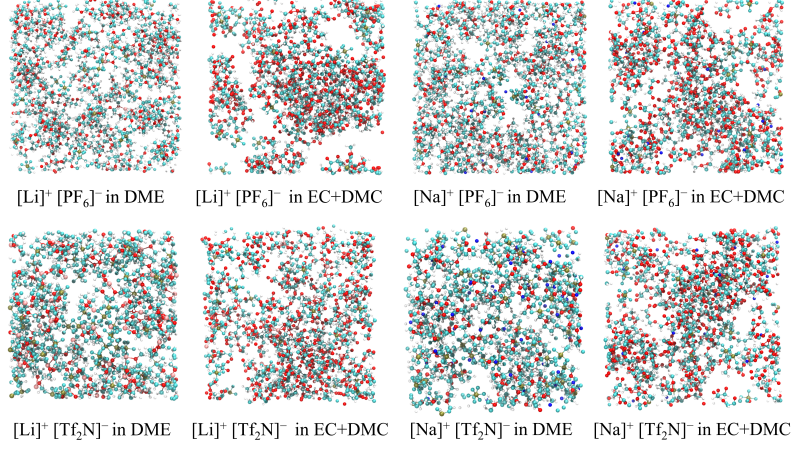


Fig. S2 Snapshots of MD simulations on electrolyte solutions: 1M [Li]⁺ [PF₆]⁻, [Na]⁺ [PF₆]⁻, [Li]⁺ [Tf₂N]⁻, or [Na]⁺ [Tf₂N]⁻ Electrolyte with DME or 1:1 EC+DMC solvent.

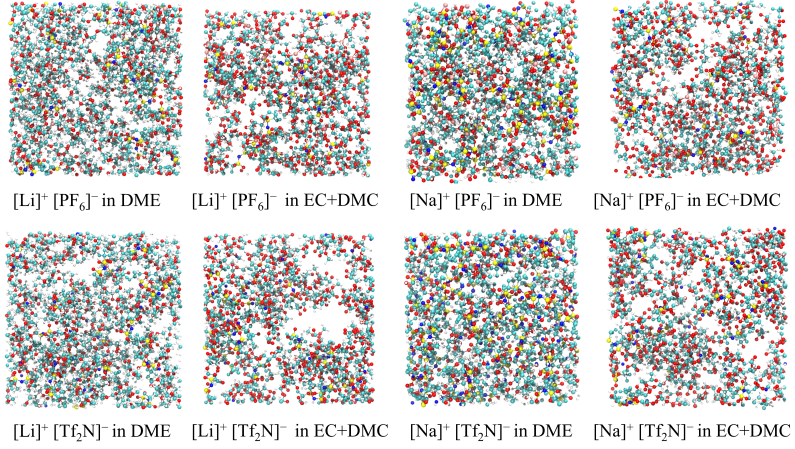


Fig. S3 Snapshots of MD simulations on electrolyte solutions: 4M [Li]⁺ [PF₆]⁻, [Na]⁺ [PF₆]⁻, [Li]⁺ [Tf₂N]⁻, or [Na]⁺ [Tf₂N]⁻ Electrolyte with DME or 1:1 EC+DMC solvent.

Table S1 Ablation study on MD17 dataset. MAE of forces (F) are shown in the unit of [kcal/mol/Å]. T1 indicates the masked feature reconstruction task, T2 indicates the noise prediction task, and GPIP indicates all the 3 pre-training tasks are involved. Results of all baseline models are directly taken or adapted (if the unit varies) from the original papers. The weights of force over energy in loss functions are set to 100 in line with the original papers of the different baselines for fair comparisons.

		SchNet [2]	SchNet-T2	SchNet-GPIP	SphereNet [3]	SphereNet-T1	SphereNet-T2	SphereNet-T1+T2	SphereNet-GPIP
Aspirin.	F	1.35	0.98	0.85	0.430	0.394	0.383	0.332	0.301
Benzene.	F	0.31	0.29	0.27	0.178	0.172	0.164	0.159	0.155
Ethanol.	F	0.39	0.34	0.27	0.208	0.164	0.157	0.154	0.150
Malonaldehyde.	F	0.66	0.60	0.53	0.340	0.307	0.303	0.292	0.261
Naphthalene.	F	0.58	0.50	0.45	0.178	0.145	0.119	0.095	0.093
Salicylic acid.	F	0.85	0.67	0.58	0.360	0.313	0.227	0.208	0.204
Toluene	F	0.57	0.47	0.41	0.155	0.130	0.102	0.098	0.090
Uracil	F	0.56	0.50	0.48	0.267	0.243	0.221	0.233	0.195
std. MAE	F	2.38	1.37	1.08	0.97	0.85	0.78	0.79	0.70

Table S2 Ablation study on Electrolyte dataset. MAE of energy and force are shown in the unit of [eV] and [eV/Å], respectively. T1 indicates the masked feature reconstruction task, T2 indicates the noise prediction task, T3 indicates the contrastive learning task, and GPIP indicates all the 3 pre-training tasks are involved.

	SchNet	SchNet-T1	SchNet-T2	SchNet-T3	SchNet-T1+T2	SchNet-GPIP
Energy	0.713	0.708	0.597	0.721	0.529	0.516
Force	0.161	0.161	0.155	0.159	0.140	0.135

Ablation Study

We perform several ablation studies using SchNet and SphereNet to study the effect of each task in GPIP. The tasks are masked feature reconstruction (T1), noise prediction (T2), and contrastive learning with 3D network (T3). The results on MD17 dataset and electrolyte solution dataset are shown in Table S1 and Table S2, respectively. The errors continue to decrease as more pre-training tasks are involved, which shows that the three pretext tasks are complementary and can jointly improve the performance of GNNs for MLIPs.

Hyperparameters

For fine-tuning on MD17 and electrolyte solution datasets, we adopt the learning rate adjustment strategy as [3]. We first warmup on the initial learning rate and then use two learning rate strategies (ReduceLROnPlateau and StepLR) for training. The learning rate is decayed by the decay ratio every fixed epoch represented as step size. For all experiments except liquid water, some hyperparameters are fixed values, and some are tuned by grid search, which are provided in Table S3, Table S4, and Table S5. We adopt the hyperparameters as in [4] for liquid water.

Table S3 Values/search space for hyperparameters on MD17.

Hyperparameters	Values/search space
Interaction blocks (SchNet)	6
Interaction block - distance LB2 intermediate size	8
Interaction block - angle LB2 intermediate size	8
Interaction block - torsion LB2 intermediate size	8
# of interaction blocks (SphereNet)	4
# of RBFs N	6
# of spherical harmonics L	3
Cutoff distance	5.0, 6.0, 10.0
Batch size	1, 2, 4, 16, 32
Initial learning rate	1e-3, 5e-4, 1e-3
Learning rate strategy	StepLR
Learning rate decay ratio	0.4, 0.5, 0.6
Learning rate step size	50, 100, 200
Max # of Epochs	500, 1000, 2000

Table S4 Values/search space for hyperparameters on ISO17.

Hyperparameters	Values/search space
Interaction blocks (SchNet)	5
Max chirality type	5
Max num_neighbors type	32
Cutoff distance	5.0, 6.0
Interaction blocks (EGNN)	5
Batch size	32, 64, 128, 256
Initial learning rate	2e-3, 2e-4, 2e-5
Minimum learning rate	1e-7
Learning rate strategy	StepLR
Learning rate cosine decay ratio	0.0, 0.1, 0.2
Learning rate step size	50, 100, 200
Learning rate warmup factor	0.7
Patience	0.3
Max # of Epochs	50, 80, 100, 130

Table S5 Values/search space for hyperparameters on Electrolyte.

Hyperparameters	Values/search space
Interaction blocks (SchNet)	5, 6
Max chirality type	5
Max num_neighbors type	32
Cutoff distance	5.0, 6.0, 10.0
Interaction blocks (EGNN)	3, 4 , 5
Interaction blocks (PaiNN)	3, 4 , 5
Batch size	1, 2, 4, 8
Initial learning rate	1e-3, 5e-4, 1e-3
Learning rate strategy	StepLR
Learning rate decay ratio	0.4, 0.5, 0.6
Learning rate step size	50, 100, 200
Learning rate warmup factor	0.7
Patience	0.3
Max # of Epochs	500, 1000, 2000

References

- [1] Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., Jorgensen, W.L.: Evaluation and reparametrization of the opls-aa force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B* **105**(28), 6474–6487 (2001)
- [2] Schütt, K., Kindermans, P.-J., Sauceda Felix, H.E., Chmiela, S., Tkatchenko, A., Müller, K.-R.: Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems* **30** (2017)
- [3] Liu, Y., Wang, L., Liu, M., Lin, Y., Zhang, X., Oztekin, B., Ji, S.: Spherical message passing for 3d molecular graphs. In: *International Conference on Learning Representations (ICLR)* (2022)
- [4] Fu, X., Wu, Z., Wang, W., Xie, T., Keten, S., Gomez-Bombarelli, R., Jaakkola, T.: Forces are not enough: Benchmark and critical evaluation for machine learning force fields with molecular simulations. *arXiv preprint arXiv:2210.07237* (2022)