Literature-Grounded LLMs for Predicting High Ionic Conductivity Solid-State Electrolytes

Sunggi An[a,*], Kunik Jang[b,*], Junyoung Choi[b], Gunwook Nam[b], Jang Wook Choi[b,c], Yousung Jung[b,d,e,*]

[*]Equal contribution [a]Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291, Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea [b]Department of Chemical and Biological Engineering, and Institute of Chemical Processes, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea [c]Seoul National University Energy Initiative (SNUEI), Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea [d]Institute of Chemical Processes, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea [e]Institute of Engineering Research, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea

Correspondence to: Yousung Jung yousung.jung@snu.ac.kr

1. Introduction

All-solid-state batteries (ASSBs) require solid electrolytes (SEs) that deliver both high room-temperature Li^+conductivity and robust stability.[1] However, discovering such materials remains challenging and slow. This is largely because ionic conductivity is governed by both composition and crystal structure, while high-quality, experimentally labeled datasets that capture these factors are still scarce.

Recent large language models (LLMs) provide a practical opportunity to incorporate literature-grounded structural descriptions (even when CIF-quality labels are scarce) into property prediction.[2,3] Herein, we present a fine-tuned, literature-grounded LLM for ionic conductivity prediction and integrate it with high-throughput virtual screening (HTVS) and machine-learned interatomic potentials (MLIP)-based MD validation to identify promising SE candidates from Materials Project (MP) (**Figure 1**).

2. Methods

Literature dataset construction:
We curated 874 experimental entries from two prior papers,[4,5] linking compositions/ICSD IDs to source papers and extracting structure-related text from full articles.

Structural text descriptors:
We extracted structural information and categorized (basic crystal descriptors; cell/coordination; experimental notes) the Structural description from the literature to test which

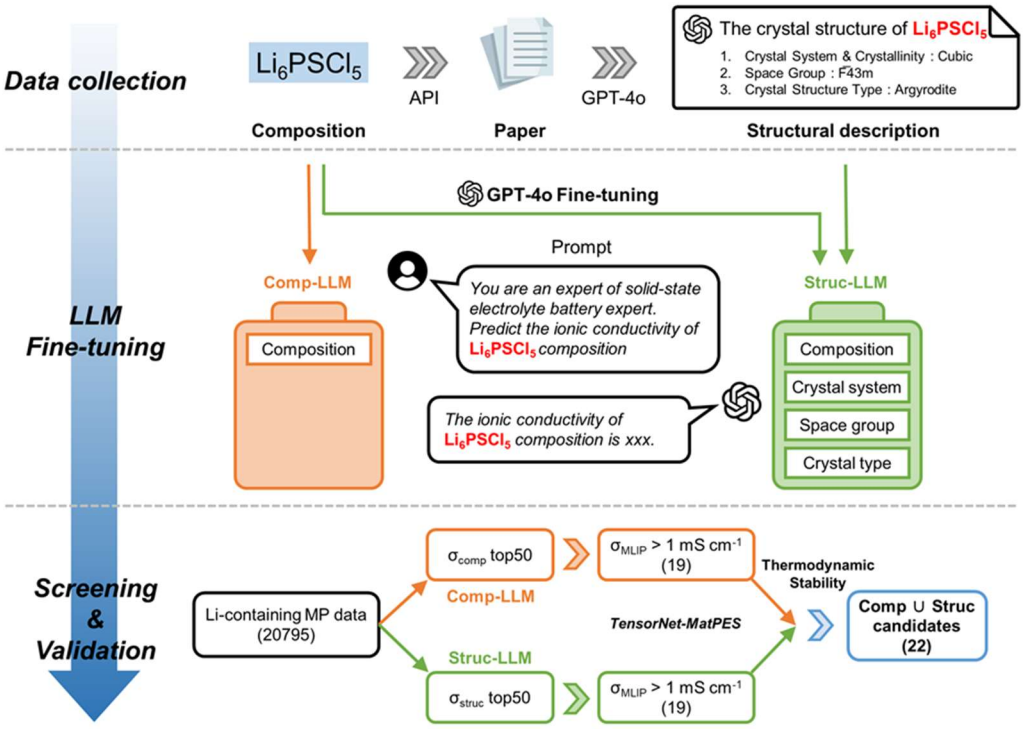


Figure 1. Schematic overview of this work, from data collection and model training to HTVS.

granularity best supports conductivity prediction.

LLM fine-tuning and screening pipeline:
We fine-tuned a pretrained GPT-4o to predict log ionic conductivity, comparing Comp-LLM (composition-only) vs. Struc-LLM (composition + structural descriptors), using an 8:1:1 split and 5-fold ensembles; MP candidates were converted to text using RoboCrystallographer.[6]

Ionic conductivity validation via MLIP calculation:
Top candidates were validated by solid-state MD using TensorNet-MatPES,[7] and down-selected using MLIP calculated conductivity.

3. Results and Discussion

Regression performance and the role of structural descriptors

Both Comp-LLM and Struc-LLM exceeded a graph-based ML baseline.[8] Adding structure improved generalization with ensemble MAE and R^2 of 0.704 and 0.721 for Comp-LLM and 0.689 and 0.744 for Struc-LLM. Basic crystal descriptors delivered the most reliable gains, whereas detailed cell and coordination or heterogeneous experimental narratives reduced accuracy, suggesting that noisy text can obscure transport signals.

HTVS results from Materials Project candidates

Applying the models to 20,795 Li-containing Materials Project entries[9], we selected the top 50 for each model and obtained 70 unique candidates after removing duplication (**Figure 2**). All exceeded the practical threshold of 0.1 mS/cm in predicted conductivity, indicating broad SE potential and enabling screening without standardized structure labels.

MLIP based MD validation of ranked candidates

Solid-state MD using TensorNet-MatPES[7], refined the candidate set and showed that LLM-ranked groups generally tracked MLIP-calculated conductivity trends. Prior benchmarking reported strong MLIP to AIMD correlation near R2 0.80, supporting MLIP MD as a scalable validation stage and revealing overestimation versus experimental conductivities in checks.

Complementarity of two LLM views and final candidates

Struc-LLM selected higher-conductivity candidates on average and the overlap between models performed best. Mean MLIP conductivities were 1.44 for Comp-LLM, 1.68 for Struc-LLM, and 1.77 mS cm^{-1} for the overlap. Applying conductivity and stability criteria yielded 22 final candidates meeting MLIP and E_{hull} thresholds.

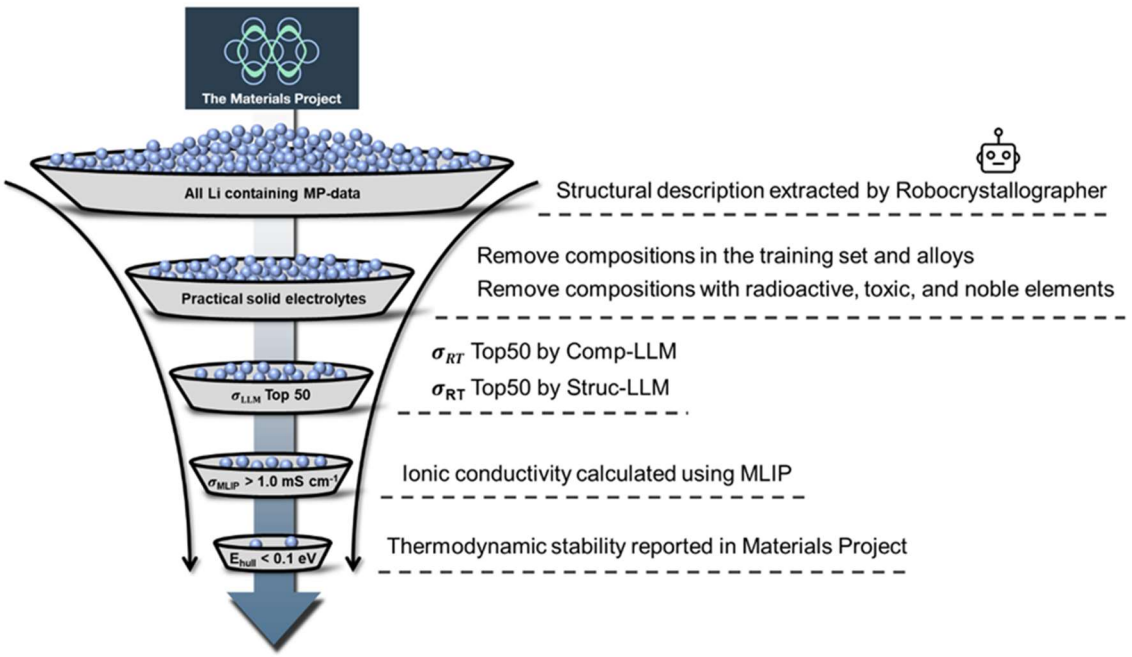


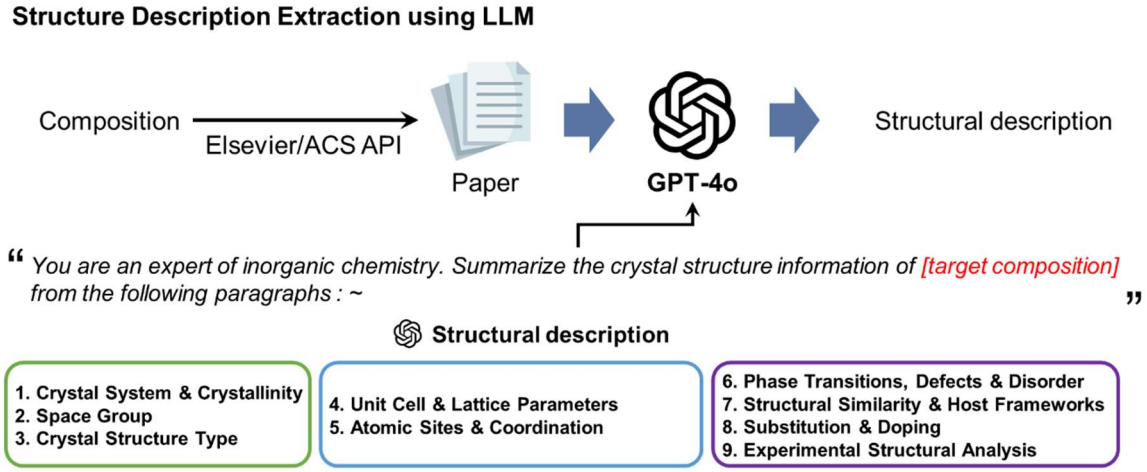
Figure 2. High-throughput virtual screening (HTVS) procedure for sampling promising solid electrolytes (SEs).

References

[1] Janek, J.; Zeier, W. G. Challenges in speeding up solid-state battery development. Nature Energy 2023, 8 (3), 230-240, DOI: 10.1038/s41560-023-01208-9
[2] Zhang, Q.; Ding, K.; Lv, T.; Wang, X.; Yin, Q.; Zhang, Y.; Yu, J.; Wang, Y.; Li, X.; Xiang, Z. Scientific large language models: A survey on biological & chemical domains. ACM Comput. Surv. 2025, 57 (6), 1-38, DOI: 10.1145/3715318
[3] Abir, A. R.; Toki Tahmid, M.; Bayzid, M. S. BioLLMNet: enhancing RNA-interaction prediction with a specialized cross-LLM transformation network. Brief. Bioinform. 2025, 26 (5), bbaf549, DOI: 10.1093/bib/bbaf549
[4] Laskowski, F. A.; McHaffie, D. B.; See, K. A. Identification of potential solid-state Li-ion conductors with semi-supervised learning. Energy Environ. Sci. 2023, 16 (3), 1264-1276, DOI: 10.1039/d2ee03499a
[5] Hargreaves, C. J.; Gaultois, M. W.; Daniels, L. M. A database of experimentally measured lithium solid electrolyte conductivities evaluated with machine learning. npj Computational Mater. 9. 2023, DOI: 10.1038/s41524-022-00951-z
[6] Ganose, A. M.; Jain, A. Robocrystallographer: automated crystal structure text descriptions and analysis. MRS Communications 2019, 9 (3), 874-881, DOI: 10.1557/mrc.2019.94
[7] Kaplan, A. D.; Liu, R.; Qi, J.; Ko, T. W.; Deng, B.; Riebesell, J.; Ceder, G.; Persson, K. A.; Ong, S. P. A foundational potential energy surface dataset for materials. arXiv preprint arXiv:2503.04070 2025, DOI: 10.48550/arXiv.2503.04070
[8] Wang, S.; Gong, S.; Böger, T.; Newnham, J. A.; Vivona, D.; Sokseiha, M.; Gordiz, K.; Aggarwal, A.; Zhu, T.; Zeier, W. G. Multimodal Machine Learning for Materials Science: Discovery of Novel Li-Ion Solid Electrolytes. Chem. Mater. 2024, 36 (23), 11541-11550, DOI: 10.1021/acs.chemmater.4c02257
[9] Horton, M. K.; Huck, P.; Yang, R. X.; Munro, J. M.; Dwaraknath, S.; Ganose, A. M.; Kingsbury, R. S.; Wen, M.; Shen, J. X.; Mathis, T. S. Accelerated data-driven materials science with the Materials Project. Nat. Mater. 2025, 24 (10), 1-11, DOI: 10.1038/s41563-025-02272-0

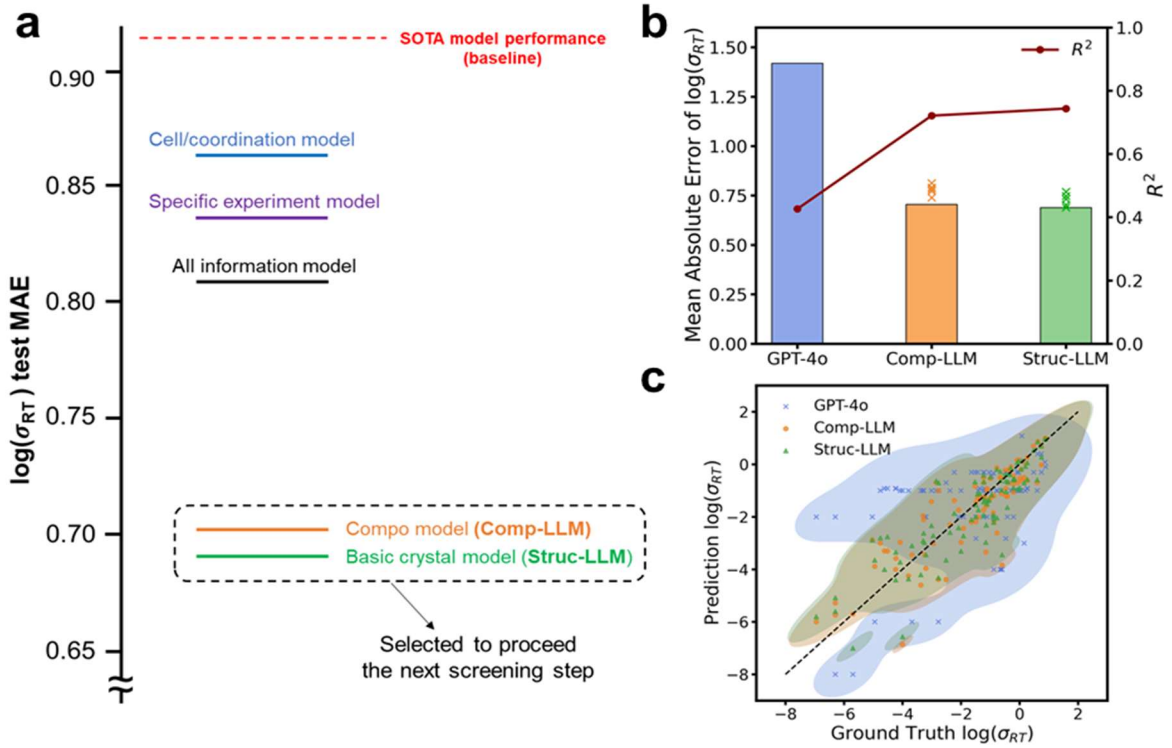
Appendix A. Extraction process of structural description from the bibliographic information using LLM.

3 Categories of 9 structural descriptions are described in three boxes. The green box includes basic information, the blue box includes structural information, and the purple box includes experimental information.

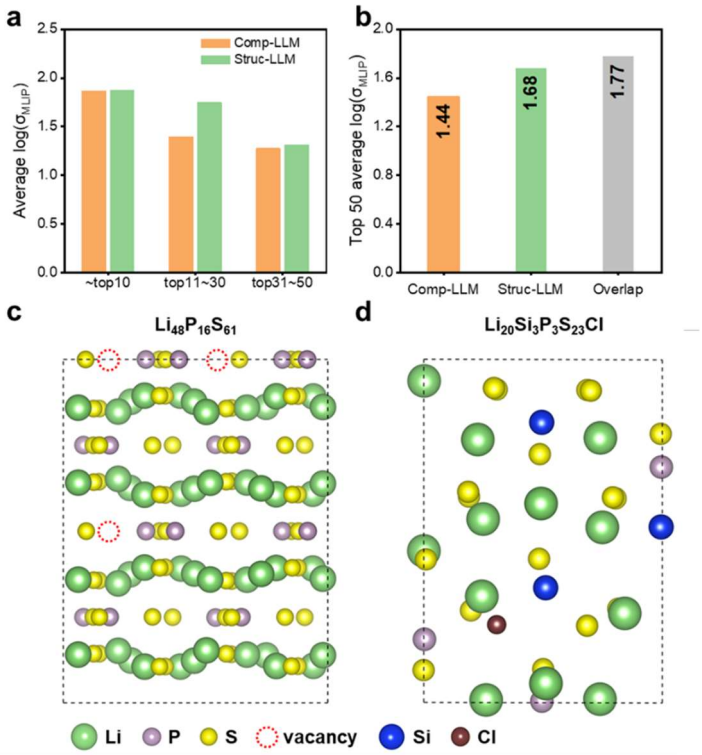


Appendix B. Performance of our prediction model.

(a) Performance comparison between prior work and each categorized information involved models. (b) Performance of ensemble prediction of Comp-LLM and Struc-LLM between non-finetuned predictions. (c) Parity plots of each prediction, the dashed line is the y=x line.



Appendix C. The result of HTVS.



(a) Average $\log(\sigma_{MLIP})$ values of the top 10, 11-30, and 31-50 entries for each model. (b) Average $\log(\sigma_{MLIP})$ values of the top 50 entries for each model and overlapped data. Representative structures of the final candidate materials: (c) $Li_{48}P_{16}S_{61}$ (mp-1001069), (d) $Li_{20}Si_3P_3S_{23}Cl$ (mp-1097035).