Generative Pipeline for Discovering Solid-State Battery Materials with Universal Atomistic Potentials

Anonymous Authors[1]

Abstract

We introduce a computational pipeline for discovering protective coating materials for Li-ion solid-state batteries, built on universal machine learning interatomic potentials (uMLIPs) that play the role of foundation models in atomistic simulation. The pipeline integrates three stages: (i) generation of candidate crystal structures using WyFormer, a symmetry-aware transformer based on Wyckoff representations; (ii) high-throughput evaluation of thermodynamic stability and electrochemical compatibility with prominent solid electrolyte and electrode materials via uMLIP-driven structure optimization; and (iii) screening of Li-ion transport using descriptors derived from uMLIP-predicted potential energy surfaces. Top candidates are further validated using machine learning-accelerated molecular dynamics, to be followed by higher-accuracy *ab initio* calculations. This framework provides a practical approach for exploring large chemical spaces and discovering functional materials by combining generative models with foundational atomistic potentials.

1. Introduction

The discovery of functional crystalline materials is increasingly formulated as a high-dimensional combinatorial optimization problem over an intractably large chemical and structural space (Matsokin et al., 2025). In applications such as solid-state electrolytes for solid-state metal-ion (e.g., Li- or Na-ion) batteries, viable candidates must simultaneously satisfy multiple criteria, including electronic insulation, fast ionic transport, and thermodynamic and electrochemical stability (Janek & Zeier, 2023). The need to balance these criteria significantly constrains the pool of suitable materials and makes their identification nontrivial.

Standard high-throughput discovery pipelines rely on enumerating candidates from existing materials databases, such as the Materials Project (MP) (Jain et al., 2013), followed by quantum-mechanical (QM) evaluation and filtering (Xiao et al., 2019). While effective in early stages of materials exploration, these approaches are fundamentally limited by the finite coverage of known databases and the prohibitive cost of QM calculations, leading to saturation in discovered property improvements.

Recent machine learning approaches offer a path beyond enumeration through (i) crystal structure generative models (Jiao et al., 2023; Miller et al., 2024; Kazeev et al., 2025; Cao et al., 2025) that expand the candidate space beyond existing databases, and (ii) universal machine learning interatomic potentials (uMLIPs) (Deng et al., 2023; Batatia et al., 2025) that provide scalable approximations to QM evaluation. However, it remains unclear whether these components can support reliable end-to-end discovery under realistic physical constraints, and how they interact within practical screening pipelines.

Here, we evaluate the WyFormer (Kazeev et al., 2025) crystal structure generative model within a surrogate-accelerated materials discovery pipeline for solid-state battery applications, as depicted in Figure 1. Specifically, we integrate the model with uMLIPs as a substitute for QM evaluation and assess its ability to generate physically viable candidate structures for protective coating materials stabilizing electrode | solid-electrolyte interfaces in Li-ion solid-state batteries.

2. Related work

2.1. Crystal structure generation

Generative models for inorganic crystals aim to learn the distribution over stable, periodic atomic arrangements and then sample novel structures to serve as candidate materials for high-throughput screening. Crystal generation has

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the AI for Science Workshop at International Conference on Machine Learning (ICML). Do not distribute.

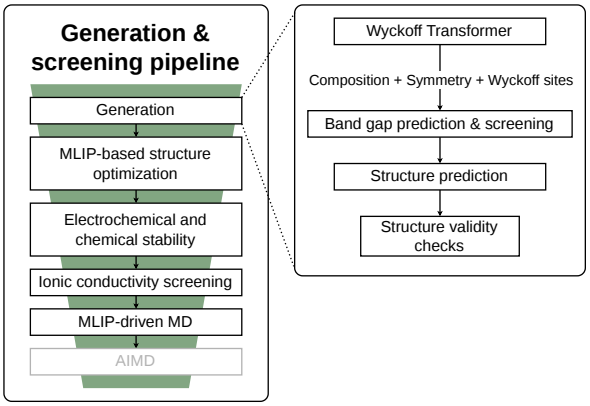


Figure 1. Schematic diagram of the developed screening pipeline. The ab initio molecular dynamics (AIMD) step shown in light gray block was not performed in this work.

unique constraints: outputs must be periodic (defined by a lattice plus a unit cell of atoms), respect crystallographic symmetries (the 230 space groups), and remain chemically plausible in terms of stoichiometry, charge balance, and bond lengths. The models are typically trained on datasets produced with DFT, and employ advanced deep learning architectures, such as equivariant diffusion (Jiao et al., 2023), flow-matching (Miller et al., 2024), and transformers (Cao et al., 2025).

A generative model at the start of a material design pipeline introduces an exploration/exploitation tradeoff between the novelty and plausibility of the generated structures. Here, we chose WyFormer (Kazeev et al., 2025), a leading model in terms exploration power, with uniqueness and novelty $\sim 90\%$ for a 10 000 sample (Okhotin et al., 2025).

2.2. Universal machine learning interatomic potentials

Machine learning interatomic potentials (MLIPs) have become essential tools in computational materials science, enabling large-scale atomistic simulations with near ab initio accuracy at a dramatically reduced computational cost compared to traditional QM methods. Pioneered by Behler and Parrinello in 2007 through their neural network approach using atom-centered symmetry functions (Behler & Parrinello, 2007), the field has advanced rapidly to modern equivariant architectures (Jacobs et al., 2025). These leverage message-passing graph neural networks that respect the physical symmetries of atomic systems (rotational, translational, and permutational invariance/equivariance), allowing for greater data efficiency and accuracy. Training MLIPs on data covering diverse chemical systems and structural environments enables transferable predictions across compositionally varied materials, making these potentials universal (uMLIPs). In this work, we employ MACE (Batatia et al., 2025), SevenNet (Park et al., 2024), and TensorNet (Simeon & De Fabritiis, 2023) – the latter trained on the MatPES (Kaplan et al., 2025) dataset – throughout our solid electrolyte discovery pipeline; the specific application of each model is detailed in Section 3.

2.3. Protective coatings for solid-state batteries

In solid-state Li-ion batteries the flammable liquid electrolyte is replaced with an inorganic solid, improving safety and energy density of the device. However, reactive electrode–electrolyte interfaces lead to degradation and increased resistance, limiting performance (Janek & Zeier, 2023).

Protective coatings – thin Li-ion–conductive interlayers – physically separate electrode and electrolyte materials while maintaining ionic transport (Zhou et al., 2025). Identifying suitable coatings is challenging due to requirements for high Li-ion conductivity, electronic insulation, and electrochemical stability. Atomistic modeling, particularly density functional theory (DFT), is widely used to screen candidate coating materials for cathodes and anodes by evaluating interfacial stability and transport properties (Xiao et al., 2019; Liu et al., 2022).

2.4. Band gap regression

Band gap (E_g) is frequently used as a proxy for electronic conductivity (Xiao et al., 2019) in high-throughout screenings. Band gap values are available in the MP database, and are commonly predicted with machine learning models (Bechtel et al., 2025). MP uses PBE functional for its DFT computation, which underestimates the band gap (Perdew & Levy, 1983). This is naturally inherited by machine learning models trained on this database. While band gap underestimation may cause our screening procedure to discard some of the otherwise eligible insulating materials (false negatives), it increases the robustness of filtering out electronically conductive materials, therefore reducing the risk of processing them in the computationally expensive subsequent stages.

2.5. Surrogate models for ionic mobility

Machine learning has become increasingly powerful for predicting ionic mobility and conductivity in crystalline solids, addressing the computational bottleneck of calculating migration barriers and diffusion coefficients through expensive first-principles methods. Early fundamental work (Wang et al., 2015) linked crystal structure and ionic transport through anion packing

geometry which was shown to determine lithium-ion migration pathways. Subsequently, lattice-dynamics-based descriptors (Muy et al., 2019) and local structural and topological descriptors for lithium environment (Sendek et al., 2019) have been used as effective inputs for ML-based screening tools. Semi-supervised ML techniques were later applied to identify powerful descriptor representations in ionic conductivity screening (Laskowski et al., 2023), revealing the smooth overlap of atomic positions (SOAP) descriptor to be most strongly correlated with experimental ionic conductivity.

With the development of uMLIPs, traditional screening techniques like Nudged Elastic Band (NEB) with underlying quantum mechanical computations replaced by machine learning potentials have become applicable in high-throughout searches (Lian et al., 2025; Dembitskiy et al., 2025). Furthermore, treating uMLIPs as fundamental models allows to extract new powerful descriptors from the potential energy surface (PES), which was shown to be both fast and effective in screening for high ionic conductivity (Maevskiy et al., 2025). In this work, we rely on these PES descriptors, as detailed in Section 3.4.

3. Proposed pipeline

We employ crystal structure generative models and uMLIPs to discover new stable protective coatings for Li-ion solid-state batteries. In Section 3.1, we formulate the problem addressed in this work, and in Section 3.2, we define the corresponding search space. Sections 3.3 and 3.4 describe the computational methodology used at each stage of the screening pipeline following structure generation.

The overall workflow is illustrated in Figure 1.

3.1. Problem setup: Protective coatings selection

As a model system, we consider a solid-state battery utilizing $Li_{10}GeP_2S_{12}$ (LGPS) as the solid electrolyte, $LiCoO_2$ as the cathode, and metallic Li as the anode. LGPS exhibits high room-temperature ionic conductivity exceeding 10^{-3} S/cm (Mo et al., 2012) and is considered one of the most promising solid electrolytes. The use of a $LiCoO_2$ cathode together with a metallic Li anode enables high energy density.

However, LGPS has a narrow electrochemical stability window, leading to decomposition at both Li | LGPS and $LiCoO_2$ | LGPS interfaces, which results in performance degradation and potential short-circuiting (Han et al., 2016). To mitigate these effects, protective coatings – stabilizing interlayers between the electrolyte and electrodes – are introduced.

To serve as a stabilizing interlayer, for the anode side, a protective coating must be stable at 0.0 V $vs.$ Li/Li^+, and its oxidation potential must exceed the reduction limit of LGPS. For the cathode side, a protective coating must be stable at 4.5 V vs. Li/Li^+ (upper operating potential for $LiCoO_2$), and its reduction limit must lie below the oxidation limit of LGPS. In addition, a protective coating must be electronically insulating, and no chemical reactions should occur at the interfaces. Additionally, owing to their much smaller thickness compared to solid electrolytes, protective coatings can operate at lower ionic conductivity than bulk solid electrolytes.

3.2. Crystal structure generation constraints

WyFormer (Kazeev et al., 2025) is a symmetry-constrained generative model that represents crystals in a coordinate-free Wyckoff space. Each structure is encoded as a space group (S.G.) followed by an unordered set of Wyckoff-position tokens, where each token contains the chemical element, site symmetry, and enumeration. In this work, we trained WyFormer on Alex-MP-20 and performed inference in Chemical System eXploration (CSX) mode, where required and allowed elements are specified by the user to constrain generation.

We targeted ternary chemical systems that must contain Li. The allowed chemical space was restricted to H, Li, Be, B, C, N, O, F, Mg, Al, Si, P, S, Cl, Ca, Sc, Ti, Cu, Zn, Ga, Ge, As, Se, Br, Sr, Y, Zr, Nb, Ag, Cd, In, Sn, Sb, Te, I, Ba, La, Hf, Ta, W, Pb, and Bi. Alkali elements other than Li were excluded to avoid pseudo-binary systems arising from partial substitution of Li by other alkali ions in the generated structures. Transition metals (except for Ti) were also excluded to focus the search on electrochemically passive compounds and to avoid systems that may exhibit electronic conductivity through polaronic transport, which is undesirable for insulating coating applications.

To remove candidates with potential electronic conductivity, we used a WyFormer band gap (E_g) prediction module trained on PBE band gap values and crystal structures from the MP database. The model achieves a mean absolute error of 0.23 eV. Because WyFormer operates directly on Wyckoff representations, band gaps can be predicted immediately after generating Wyckoff representations, without first constructing full atomic structures, ensuring fast operation. Structures with predicted band gaps below 0.5 eV are rejected. While the band gap serves only as a proxy and does not guarantee low electronic conductivity, it provides an efficient and computationally inexpensive way to eliminate materials with potential for significant electronic conduction (Xiao et al., 2019). Finally, novelty was evaluated using SWORD fingerprints (Huang et al., 2026), with generated structures compared against the LeMat-Bulk database (Siron et al.,

2025) and the Inorganic Crystal Structure Database (ICSD) (Zagorac et al., 2019).

3.3. Thermodynamic calculations

Following the E_g prediction, the crystal structures are optimized using the MACE-MPA-0 (Batatia et al., 2025) uMLIP for further evaluation of phase stability via computation of the energy above the convex hull (E_{hull}), electrochemical stability windows (reduction and oxidation limits), and chemical compatibility with LGPS and $LiCoO_2$.

The MACE-MPA-0 uMLIP was used to compute thermodynamic properties of candidate materials. The model is trained on 3.5 million crystal structure configurations drawn from the combined MPTrj (Deng et al., 2023) and sAlex (Barroso-Luque et al., 2024) datasets, spanning 89 chemical elements. The training data are generated at the PBE level of theory for systems without transition metals, oxygen, or fluorine, and at the PBE+U level for systems containing transition metals and oxygen or fluorine.

The inclusion of Hubbard U corrections improves agreement between DFT-computed formation energies and experimental measurements (Hautier et al., 2012). The MP's energy correction scheme (Wang et al., 2021) was used to further enhance consistency with reference data.

For phase diagram construction, crystal structures were obtained from the MP database and subsequently re-optimized, including both cell shape and atomic positions, using the selected uMLIP with a force convergence criterion of 0.05 eV/Å. The mean absolute difference between MP total energies and uMLIP-derived values is 0.04 eV/atom.

The electrochemical stability windows and chemical mixing reaction energies were calculated following the methodologies proposed by Ong et al. (Ong et al., 2008) and Richards et al. (Richards et al., 2016), respectively, as implemented in `Pymatgen` (Ong et al., 2013).

3.4. Ionic conductivity screening

Candidate structures passing our stability criteria are further evaluated using an ionic conductivity screening step adapted from the methodology by Maevskiy et al. (2025). This approach relies on descriptors extracted from PES computed using uMLIP. We employ two complementary computational strategies to capture diverse conductivity mechanisms. In the first approach, the original structure is frozen and a single interstitial Li atom is positioned at a set of locations from a regular 3D grid to generate a 3D PES map; we then select candidates based on the Ξ descriptor from (Maevskiy et al., 2025), which favors structures exhibiting large regions with energy below 0.5 eV. For this first approach, SevenNet-0 uMLIP (Park et al., 2024) is used for all energy evaluations. For the second approach, all Li atoms are removed from the structure and the resulting framework is frozen; a single Li atom is subsequently added and translated to generate the PES. Given the substantial domain shift between our dataset and the ICSD-derived training data of (Maevskiy et al., 2025), we employ a percolation barrier as an intuitive descriptor for this second approach, rather than relying on the transferred descriptor. For this second approach, MACE-MPA-0 (Batatia et al., 2025) is used to evaluate the energies. The two selections are combined through logical OR, allowing structures favorable under either mechanism to be retained.

Structures passing the ionic conductivity screening are evaluated through NVT molecular dynamics simulations using TensorNet-MatPES-PBE-v2025.1-PES (Kaplan et al., 2025; Ko et al., 2025; Simeon & De Fabritiis, 2023). For each candidate, a supercell is constructed containing at least 100 atoms and accommodating a 10 Å diameter sphere. Atomic positions are first relaxed to a maximum force threshold of 0.005 eV/Å, followed by a 10 ps thermalization phase using a Berendsen thermostat with time constant $\tau = 1$ ps and 100 ps of production dynamics with a Nose-Hoover thermostat with Nose time constant of 25 fs. Li-ion diffusion coefficients are extracted by partitioning the trajectory into 15 ps windows and fitting the mean squared displacement slope in the 3–15 ps interval for each window; the final diffusion coefficient and standard error are estimated from the ensemble of window-based fits. This procedure is adopted from (Kahle et al., 2020), which is shown to avoid the overestimation of the diffusion coefficient due to vibrational dynamics for the slow diffusion cases. The procedure was shown to be applicable for diffusion coefficient values $D > 10^{-4}$ Å^2/ps, which agrees with our conductivity-based selection explained in Section 4.

4. Results

The selection criteria applied at each screening stage, together with the number of candidates passing each filter, are summarized in Table 1. We adopt relatively tolerant thermodynamic criteria to account for discrepancies between uMLIP and DFT predictions.

The results for uMLIP-driven MD runs are shown in Figure 2. This step was performed on 466 candidates passing the descriptor-based ionic conductivity screening stage.

To filter out remaining unstable structures, we evaluate the diffusion coefficient D not only for Li-ion but for each

Table 1. Selection criteria applied at each stage of the pipeline along with numbers of passing candidates. The smallest interatomic distance within a structure is denoted with d_{min}. AC, CC, and SE stand for anode coating, cathode coating, and solid electrolyte, respectively. E_{red} and E_{ox} stand for the reduction and oxidation limits, respectively. The quantities $\Xi^{\text{interstitial}}$ and $E^{\text{empty}}_{\text{mig.}}$ are the Ξ descriptor from (Maevskiy et al., 2025) and migration barrier, respectively, as explained in Section 3.4. The standard error of the mean is denoted as *s.e.m.*

Filter	Criterion	No. of candidates passing
Band gap	$E_{\text{g}} > 0.5$ eV	101,789
Subsampling for structure reconstruction	–	50,000
Validity, uniqueness & novelty	$d_{\text{min}} > 0.7$ Å uniqueness novel vs. LeMat-Bulk & ICSD	43,311
Stability	$E_{\text{hull}} < 0.2$ eV/atom	23,607
Electrochemical stability & chemical mixing	$\text{E}_{\text{red}}(\text{AC}) < 0.3$ V *vs.* Li/Li$^+$ $\text{E}_{\text{ox}}(\text{AC}) > \text{E}_{\text{red}}(\text{SE})$ $\text{E}_{\text{red}}(\text{CC}) < \text{E}_{\text{ox}}(\text{SE})$ $\text{E}_{\text{ox}}(\text{CC}) > 4.5$ V *vs.* Li/Li$^+$ $\text{E}_{\text{rxn}} >$ -0.2 eV/atom	1,661
Ionic conductivity (descriptor-based)	$\Xi^{\text{interstitial}} > 0.8$ (OR) $E^{\text{empty}}_{\text{mig.}} < 0.5$ eV	466
Ionic conductivity (uMLIP-driven MD)	$\sigma^{300\text{ K}}_{\text{Li}} - 1 \cdot \textit{s.e.m.} > 10^{-6}$ S / cm $\text{MSD(100 ps)}^{1000\text{ K}}_{\text{framework}} < 2$ Å^2 $\max_t \text{MSD}(t)^{500\text{ K}}_{\text{Li}} > 4$ Å^2	29

species and check the value of D at $T = 1000$ K for the slowest non-Li atom (denoted as $D^{1000\text{ K}}_{\text{framework}}$ in Figure 2). We require that $D^{1000\text{ K}}_{\text{framework}}$ of a stable structure translates to mean squared displacement (MSD) of a full 100 ps trajectory no larger than 2 Å^2. Next, structures in which the maximum MSD of Li within 100 ps at $T = 500$ K was below 4 Å^2 were omitted.

Finally, to select potentially conductive structures, similarly to Hwang *et al.* (Hwang et al., 2025), we extrapolate Li-ion conductivity from $T = 500$ K to $T = 300$ K assuming relatively pessimistic Li-ion activation barrier, E_a, of 0.5 eV in the Arrhenius law $\left(\sigma \sim \frac{1}{T}\exp(-E_a/kT)\right)$. We select structures with such extrapolated conductivity above 0.001 mS/cm, which is sufficient for a protective coating.

As a result, 29 out of 527 365 element-constrained generated ternary candidates pass all filters in our pipeline. Tables 2 and 3 summarize details for the selected anode and cathode coating candidates, respectively.

Almost all candidate cathode coatings are halides (predominantly fluorides), which is consistent with previous literature reports (Liu et al., 2019; Zhou et al., 2025) showing wide electrochemical stability, electronic insulation, and compatibility of such coatings with transition metal oxide cathode materials.

As for anode coatings, most of the candidates that passed all screening filters are polymorphs of $LiCaI_3$. The top candidate for ionic conductivity is Li_3SBr (S.G. $P2_1/c$), with $\sigma^{300\text{ K}}_{\text{Li}} = 0.6$ mS/cm. However, the generated structure has a relatively high E_{hull} of 0.11 eV/atom and shows atypical polyhedral connectivity and an unusual coordination environment for Li ions (Figure 4), suggesting that this structure is unlikely to be synthesizable.

Manual verification of the generated structures reveals several likely implausible entries in the final candidate list (Figure 4). For example, Li_2SO_4 (S.G. $I4_1/amd$), with E_{hull} of 0.16 eV/atom, exhibits an unusual square-planar coordination of Li ions, each surrounded by four oxygen atoms. Similarly, Li_3MgF_5 (S.G. $C2/c$, not present in the tables due to low Li-ion MSD) contains Li-ions in an atypical planar triangular environment formed by neighboring fluorine ions (Figure 4). Beyond local coordination anomalies, this composition may also reflect deeper inconsistencies: in ternary ionic systems such as Li–Mg–F, structures with positive E_{hull} values

Table 2. Generated crystal structures passing screening filters for protective anode coatings. Li-ion conductivity value $\sigma_{\text{Li}}^{300\text{ K}}$ was obtained by extrapolating $\sigma_{\text{Li}}^{500\text{ K}}$, assuming activation barrier $E_a = 0.5$ eV.

Formula	Space group	E_{hull}, eV / atom	E_{red}, V *vs.* Li/Li$^+$	E_{ox}, V *vs.* Li/Li$^+$	E_{rxn}, eV / atom	$\sigma_{\text{Li}}^{300\text{ K}}$, mS / cm
Li_3SBr	$P2_1/c$	0.11	0.0	2.14	0.0	0.62 ± 0.12
Li_2CaI_4	$Imma$	0.06	0.01	2.84	-0.0	0.21 ± 0.05
Li_4CaBr_6	$C2/m$	0.03	0.06	3.66	-0.01	0.133 ± 0.034
$LiCaI_3$	$P\bar{1}$	0.02	0.01	2.84	-0.01	0.114 ± 0.022
Li_4SCl_2	$P2_1/c$	0.06	0.0	2.14	0.0	0.10 ± 0.08
$LiCaI_3$	$C2/c$	0.04	0.01	2.84	-0.01	0.044 ± 0.025
$LiCaI_3$	$C2/m$	0.03	0.01	2.84	-0.01	0.029 ± 0.015
$LiCaI_3$	$C2/c$	0.12	0.01	2.84	-0.01	0.0177 ± 0.0031

Table 3. Generated crystal structures passing screening filters for protective cathode coatings. Li-ion conductivity value $\sigma_{\text{Li}}^{300\text{ K}}$ was obtained by extrapolating $\sigma_{\text{Li}}^{500\text{ K}}$, assuming activation barrier $E_a = 0.5$ eV.

Formula	Space group	E_{hull}, eV / atom	E_{red}, V *vs.* Li/Li$^+$	E_{ox}, V *vs.* Li/Li$^+$	E_{rxn}, eV / atom	$\sigma_{\text{Li}}^{300\text{ K}}$, mS / cm
Li_3AlF_6	$P312$	0.06	1.06	6.44	-0.01	0.68 ± 0.09
Li_3YCl_6	$P312$	0.06	0.64	4.28	0.0	0.585 ± 0.033
Li_3AlF_6	$P\bar{1}$	0.04	1.06	6.44	-0.01	0.52 ± 0.15
Li_2MgCl_4	$C2/c$	0.05	0.87	4.28	0.0	0.46 ± 0.10
Li_2HfCl_6	$P4_2/mnm$	0.02	1.36	4.28	-0.05	0.41 ± 0.10
$LiLaCl_4$	$I4_1/amd$	0.02	0.44	4.28	0.0	0.40 ± 0.06
Li_3AlF_6	$C2/m$	0.04	1.06	6.44	-0.01	0.35 ± 0.07
Li_2SO_4	$I4_1/amd$	0.16	1.63	4.61	0.0	0.33 ± 0.12
Li_2ZrCl_6	$C2/c$	0.01	1.79	4.28	-0.02	0.28 ± 0.04
Li_3GaCl_6	$R3$	0.05	2.14	4.28	0.0	0.28 ± 0.04
Li_2ZrCl_6	$C2$	0.0	1.79	4.28	-0.02	0.27 ± 0.04
$LiYCl_4$	Pc	0.02	0.64	4.28	0.0	0.19 ± 0.04
$LiMgCl_3$	$C2/m$	0.01	0.87	4.28	0.0	0.17 ± 0.04
Li_5AlF_8	$P1$	0.05	1.06	6.34	-0.01	0.15 ± 0.05
Li_2ZrCl_6	$C2/m$	0.0	1.7	4.32	-0.01	0.13 ± 0.04
Li_3GaF_6	$C2/c$	0.04	2.26	6.51	0.0	0.106 ± 0.021
Li_2ZrCl_6	$C2/c$	0.02	1.79	4.28	-0.02	0.10 ± 0.04
Li_2MgCl_4	$P2/m$	0.01	0.87	4.28	0.0	0.10 ± 0.04
Li_2SiF_6	$P312$	0.02	1.85	6.61	-0.01	0.049 ± 0.029
Li_2ZrCl_6	$P312$	0.02	1.79	4.28	-0.02	0.048 ± 0.026
$LiLaCl_4$	$I4_1/a$	0.01	0.44	4.28	0.0	0.021 ± 0.015

whose compositions lie on a binary tie-line between stable phases (e.g., LiF and MgF_2) are expected to decompose into phase-separated mixtures of those binary endpoints. Similarly, the generated $LiLaCl_4$ phase ($I4_1/amd$) shown in Figure 3 has a relatively low E_{hull} of 0.02 eV/atom; however, no solid solutions or intermediate phases are known in the LiCl–$LaCl_3$ system (Shujian et al., 1998), suggesting that this candidate is unlikely to be synthesizable, at least via a solid-solution route.

5. Limitations and future work

The occurrence of implausible coordination environments in the generated candidates highlights an inherent limitation of symmetry-constrained generative models such as WyFormer. Because the model operates in a coordinate-free Wyckoff representation, structures are primarily determined by space group symmetry (e.g., $I4_1/amd$) and site assignments rather than explicit chemical preferences. As a result, configurations that are fully consistent with symmetry – such as the square-planar LiO_4 environment observed in Li_2SO_4 – can emerge even when they are atypical for the underlying chemistry. While subsequent uMLIP relaxation (or even phonon calculations) may confirm these structures as locally stable minima, they can still be thermodynamically unfavorable (e.g., elevated energy above the convex hull) and chemically unlikely due to the absence of constraints on coordination environments, bond valence, or electrostatics in the generative prior. This illustrates a broader mismatch between symmetry-driven sampling and chemically realistic structure formation, motivating the inclusion of additional chemical filters or physically informed priors in post-processing.

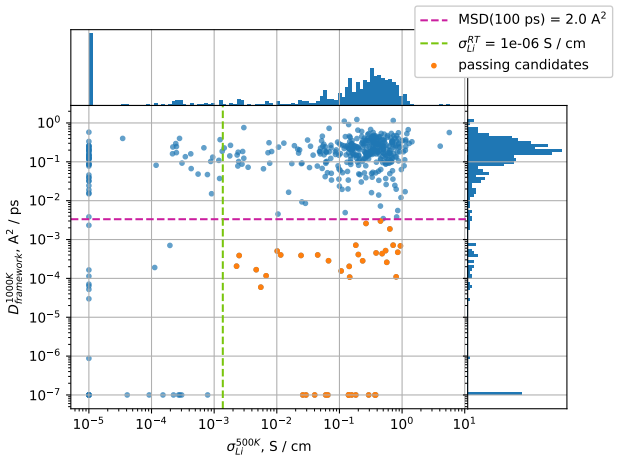


Figure 2. Conductivity values extracted from the uMLIP-driven MD at $T = 500$ K ($\sigma_{\mathrm{Li}}^{500\ \mathrm{K}}$, horizontal axis) versus diffusion coefficient for the slowest non-lithium species extracted from MD at $T = 1000$ K (vertical axis). Framework stability and ionic conductivity thresholds for the final selection step are shown with dashed lines.

We therefore emphasize that candidates identified during screening must be further validated with DFT beyond E_{hull}, including verification of well-relaxed minima, dynamical stability (e.g., absence of imaginary phonon modes), and ionic mobility. This also presents a size-quality trade-off for the selection of the training dataset for the generative model. We used Alex-MP, majority of the structures in it were only screened by E_{hull}. A potential alternative, MP-20, contains only experimentally-observed structures, but is an order of magnitude smaller.

In addition, statistics of coordination environments in the generated structures can be used to assess whether local geometries are typical (George et al., 2020), helping to flag potentially "unrealistic" candidates. Therefore, the development of crystal-chemistry analysis tools for validating the results of generative models would be beneficial.

Most structures passing the PES descriptor-based filter were subsequently classified as unstable in high-temperature ($T = 1000$ K) MD simulations using TensorNet uMLIP, based on a conservative threshold of 2 Å^2 for the mean squared displacement of the framework. While this criterion is practical, it may misclassify cases where species other than Li are mobile but the overall structural motif is preserved. Phonon calculations may therefore provide a more reliable alternative for identifying unstable structures.

For application in solid-state batteries, further DFT evaluation of interfacial adhesion and charge transfer is required to ensure stable, low-resistance electrode–electrolyte interfaces (Burov et al., 2024).

6. Conclusion

Our sequential high-throughput workflow effectively filtered a massive initial search space of 527 365 element-constrained generated ternary structures. The novelty of the structures was evaluated against the largest computational metastable materials (LeMat-Bulk) and experimental (ICSD) databases, taking into account disordered structures. We systematically evaluated these candidates against the stringent requirements for protective interlayers: electronic insulation, thermodynamic phase stability, electrochemical and chemical compatibility with the $Li_{10}GeP_2S_{12}$ solid electrolyte and adjacent electrodes, and sufficient Li-ion transport. Ultimately, this multi-stage approach, combining potential energy surface descriptors with uMLIP-driven molecular dynamics simulations, narrowed the vast candidate pool down to 29 structures, subject to further *ab initio* evaluation.

Our study reveals important limitations inherent to symmetry-driven generative modeling: structures that are thermodynamically competitive and locally stable according to machine learning metrics can nevertheless exhibit chemically implausible coordination environments. We emphasize that the 29 identified candidates should be validated through density functional theory calculations of dynamical stability, interfacial properties, and charge-transfer mechanisms before experimental synthesis is pursued. Future work should incorporate chemical plausibility filters into the post-processing stage, and further tune generative models with feedback from stability analysis to reduce the false-positive rate and enhance the chemical realism of discovered structures.

Data availability

The data supporting this study will be made available upon acceptance of the manuscript.

References

Barroso-Luque, L., Shuaibi, M., Fu, X., Wood, B. M., Dzamba, M., Gao, M., Rizvi, A., Zitnick, C. L., and Ulissi, Z. W. Open materials 2024 (omat24) inorganic materials dataset and models. *arXiv preprint arXiv:2410.12771*, 2024.

Batatia, I., Benner, P., Chiang, Y., Elena, A. M., Kovács, D. P., Riebesell, J., Advincula, X. R., Asta, M., Avaylon, M., Baldwin, W. J., et al. A foundation model for atomistic materials chemistry. *The Journal of chemical*

physics, 163(18), 2025.

Bechtel, T., Speckhard, D. T., Godwin, J., and Draxl, C. Band-gap regression with architecture-optimized message-passing neural networks. *Chemistry of Materials*, 37(4):1358–1369, 2025.

Behler, J. and Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical review letters*, 98(14):146401, 2007.

Burov, A. S., Boev, A. O., Abakumov, A. M., and Aksyonov, D. A. Mechanism of li+ charge transfer at li/li 7 la 3 zr 2 o 12 interfaces: A density functional theory study. *Physical Review B*, 109(4):045305, 2024.

Cao, Z., Luo, X., Lv, J., and Wang, L. Space group informed transformer for crystalline materials generation. *Science Bulletin*, 2025.

Dembitskiy, A. D., Humonen, I. S., Eremin, R. A., Aksyonov, D. A., Fedotov, S. S., and Budennyy, S. A. Benchmarking machine learning models for predicting lithium ion migration. *npj Computational Materials*, 11 (1):131, 2025.

Deng, B., Zhong, P., Jun, K., Riebesell, J., Han, K., Bartel, C. J., and Ceder, G. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, 2023.

George, J., Waroquiers, D., Di Stefano, D., Petretto, G., Rignanese, G.-M., and Hautier, G. The limited predictive power of the pauling rules. *Angewandte Chemie*, 132(19):7639–7645, 2020.

Han, F., Zhu, Y., He, X., Mo, Y., and Wang, C. Electrochemical stability of li10gep2s12 and li7la3zr2o12 solid electrolytes. *Advanced Energy Materials*, 6(8):1501590, 2016.

Hautier, G., Ong, S. P., Jain, A., Moore, C. J., and Ceder, G. Accuracy of density functional theory in predicting formation energies of ternary oxides from binary oxides and its implication on phase stability. *Physical Review B—Condensed Matter and Materials Physics*, 85(15): 155208, 2012.

Huang, Y., Nong, W., Yamazaki, S., Petersen, M. H., Wang, J., Zhu, R., and Hippalgaonkar, K. SWORD: Symmetry and Wyckoff-sequence of Ordered and Disordered crystals, 2026. URL https://arxiv.org/abs/2604.17994.

Hwang, S., Lee, J., Han, S., Kang, Y., and Kang, S. Discovery of oxide li-conducting electrolytes in uncharted chemical space via topology-constrained crystal structure prediction. *arXiv preprint arXiv:2509.25763*, 2025.

Jacobs, R., Morgan, D., Attarian, S., Meng, J., Shen, C., Wu, Z., Xie, C. Y., Yang, J. H., Artrith, N., Blaiszik, B., Ceder, G., Choudhary, K., Csanyi, G., Cubuk, E. D., Deng, B., Drautz, R., Fu, X., Godwin, J., Honavar, V., Isayev, O., Johansson, A., Kozinsky, B., Martiniani, S., Ong, S. P., Poltavsky, I., Schmidt, K., Takamoto, S., Thompson, A. P., Westermayr, J., and Wood, B. M. A practical guide to machine learning interatomic potentials — status and future. *Current Opinion in Solid State and Materials Science*, 35:101214, 2025. ISSN 1359-0286. doi: 10.1016/j.cossms.2025.101214.

Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.

Janek, J. and Zeier, W. G. Challenges in speeding up solid-state battery development. *Nature Energy*, 8(3):230–240, 2023.

Jiao, R., Huang, W., Lin, P., Han, J., Chen, P., Lu, Y., and Liu, Y. Crystal structure prediction by joint equivariant diffusion. *Advances in Neural Information Processing Systems*, 36:17464–17497, 2023.

Kahle, L., Marcolongo, A., and Marzari, N. High-throughput computational screening for solid-state li-ion conductors. *Energy Environ. Sci.*, 13:928–948, 2020. doi: 10.1039/C9EE02457C.

Kaplan, A. D., Liu, R., Qi, J., Ko, T. W., Deng, B., Riebesell, J., Ceder, G., Persson, K. A., and Ong, S. P. A foundational potential energy surface dataset for materials. *arXiv preprint arXiv:2503.04070*, 2025.

Kazeev, N., Nong, W., Romanov, I., Zhu, R., Ustyuzhanin, A., Yamazaki, S., and Hippalgaonkar, K. Wyckoff transformer: Generation of symmetric crystals. *arXiv preprint arXiv:2503.02407*, 2025.

Ko, T. W., Deng, B., Nassar, M., Barroso-Luque, L., Liu, R., Qi, J., Thakur, A. C., Mishra, A. R., Liu, E., Ceder, G., et al. Materials Graph Library (MatGL), an open-source graph deep learning library for materials science and chemistry. *npj Computational Materials*, 11(1):253, 2025.

Laskowski, F. A., McHaffie, D. B., and See, K. A. Identification of potential solid-state li-ion conductors with semi-supervised learning. *Energy & Environmental Science*, 16(3):1264–1276, 2023. doi: 10.1039/D2EE03499A.

Lian, J., Fu, X., Gong, X., Xiao, R., and Li, H. High-throughput neb for li-ion conductor discovery via fine-tuned chgnet potential. *J. Mater. Chem. A*, 13:34918–34926, 2025. doi: 10.1039/D5TA05355B.

Liu, B., Wang, D., Avdeev, M., Shi, S., Yang, J., and Zhang, W. High-throughput computational screening of li-containing fluorides for battery cathode coatings. *ACS Sustainable Chemistry & Engineering*, 8(2):948–957, 2019.

Liu, C., Früchtl, H., Irvine, J. T., and Bühl, M. Computational screening of anode coatings for garnet-type solid-state batteries. *Batteries & Supercaps*, 5(4): e202100357, 2022.

Maevskiy, A., Carvalho, A., Sataev, E., Turchyna, V., Noori, K., Rodin, A., Castro Neto, A. H., and Ustyuzhanin, A. Predicting ionic conductivity in solids from the machine-learned potential energy landscape. *Phys. Rev. Res.*, 7:023167, May 2025. doi: 10.1103/PhysRevResearch.7.023167.

Matsokin, N. A., Eremin, R. A., Kuznetsova, A. A., Humonen, I. S., Krautsou, A. V., Lazarev, V. D., Vassilyeva, Y. Z., Pak, A. Y., Budennyy, S. A., Kvashnin, A. G., et al. Discovery of chemically modified higher tungsten boride by means of hybrid gnn/dft approach. *npj Computational Materials*, 11(1):163, 2025.

Miller, B. K., Chen, R. T., Sriram, A., and Wood, B. M. Flowmm: Generating materials with riemannian flow matching. *arXiv preprint arXiv:2406.04713*, 2024.

Mo, Y., Ong, S. P., and Ceder, G. First principles study of the li10gep2s12 lithium super ionic conductor material. *Chemistry of Materials*, 24(1):15–17, 2012.

Muy, S., Voss, J., Schlem, R., Koerver, R., Sedlmaier, S. J., Maglia, F., Lamp, P., Zeier, W. G., and Shao-Horn, Y. High-throughput screening of solid-state Li-ion conductors using lattice-dynamics descriptors. *iScience*, 16:270–282, Jun 2019. doi: 10.1016/j.isci.2019.05.036.

Okhotin, A., Nakhodnov, M., Kazeev, N., Ustyuzhanin, A. E., and Vetrov, D. Miad: Mirage atom diffusion for de novo crystal generation. *arXiv preprint arXiv:2511.14426*, 2025.

Ong, S. P., Wang, L., Kang, B., and Ceder, G. Li- fe- p- o2 phase diagram from first principles calculations. *Chemistry of Materials*, 20(5):1798–1807, 2008.

Ong, S. P., Richards, W. D., Jain, A., Hautier, G., Kocher, M., Cholia, S., Gunter, D., Chevrier, V. L., Persson, K. A., and Ceder, G. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.

Park, Y., Kim, J., Hwang, S., and Han, S. Scalable parallel algorithm for graph neural network interatomic potentials in molecular dynamics simulations. *J. Chem. Theory Comput.*, 20(11):4857–4868, 2024. doi: 10.1021/acs.jctc.4c00190.

Perdew, J. P. and Levy, M. Physical content of the exact kohn-sham orbital energies: band gaps and derivative discontinuities. *Physical Review Letters*, 51(20):1884, 1983.

Richards, W. D., Miara, L. J., Wang, Y., Kim, J. C., and Ceder, G. Interface stability in solid-state batteries. *Chemistry of Materials*, 28(1):266–273, 2016.

Sendek, A. D., Cubuk, E. D., Antoniuk, E. R., Cheon, G., Cui, Y., and Reed, E. J. Machine learning-assisted discovery of solid li-ion conducting materials. *Chemistry of Materials*, 31(2):342–352, Jan 2019. ISSN 0897-4756. doi: 10.1021/acs.chemmater.8b03272.

Shujian, T., Shixiang, L., and Chaogui, Z. Investigation on the phase diagram of the ternary system lacl3–nacl–licl. *Journal of alloys and compounds*, 279(2):127–131, 1998.

Simeon, G. and De Fabritiis, G. Tensornet: Cartesian tensor representations for efficient learning of molecular potentials. *Advances in Neural Information Processing Systems*, 36:37334–37353, 2023.

Siron, M., Djafar, I., Ramlaoui, A., du Fayette, E., Rossello, A., Fako, E., McDermott, M., Therrien, F., Barroso-Luque, L., Cipcigan, F., Schwaller, P., Wolf, T., and Duval, A. LeMat-Bulk: aggregating, and de-duplicating quantum chemistry materials databases, 2025. URL https://arxiv.org/abs/2511.05178.

Wang, A., Kingsbury, R., McDermott, M., Horton, M., Jain, A., Ong, S. P., Dwaraknath, S., and Persson, K. A. A framework for quantifying uncertainty in dft energy corrections. *Scientific reports*, 11(1):15496, 2021.

Wang, Y., Richards, W. D., Ong, S. P., Miara, L. J., Kim, J. C., Mo, Y., and Ceder, G. Design principles for solid-state lithium superionic conductors. *Nature Materials*, 14(10):1026–1031, Oct 2015. doi: 10.1038/nmat4369.

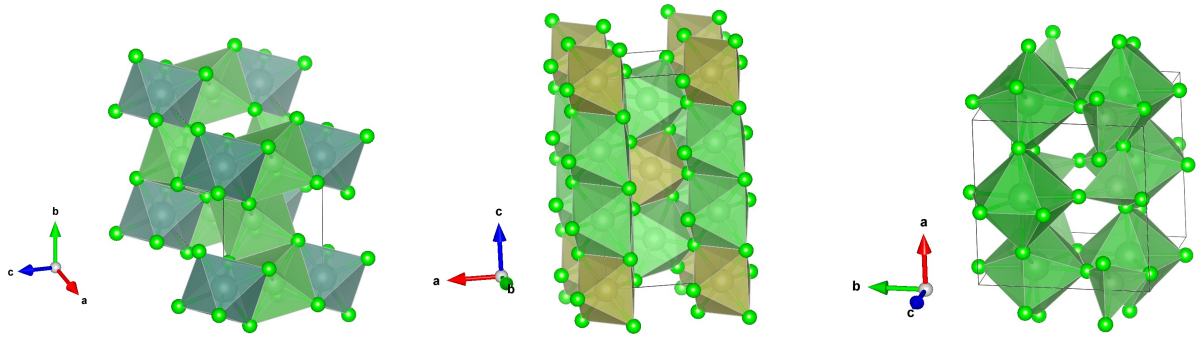
Xiao, Y., Miara, L. J., Wang, Y., and Ceder, G. Computational screening of cathode coatings for solid-state batteries. *Joule*, 3(5):1252–1275, 2019.

Zagorac, D., Müller, H., Ruehl, S., Zagorac, J., and Rehme, S. Recent developments in the inorganic crystal structure database: theoretical crystal structure data and related features. *Applied Crystallography*, 52(5):918–925, 2019.

Zhou, X., Chang, C.-Y., Yu, D., Zhang, K., Li, Z., Jiang, S.-K., Zhu, Y., Xia, Y., Hwang, B. J., and Wang, Y. Li2zrf6 protective layer enabled high-voltage licoo2 positive electrode in sulfide all-solid-state batteries. *Nature Communications*, 16(1):112, 2025.

A. Examples of generated structures

Figure 3 shows examples of structures passing all the selection criteria that showed high ionic conductivity in uMLIP-driven MD simulations. Additionally, we present structures in Figure 4 that, while also passing all the selection criteria, demonstrate unusual Li-ion coordination environments.

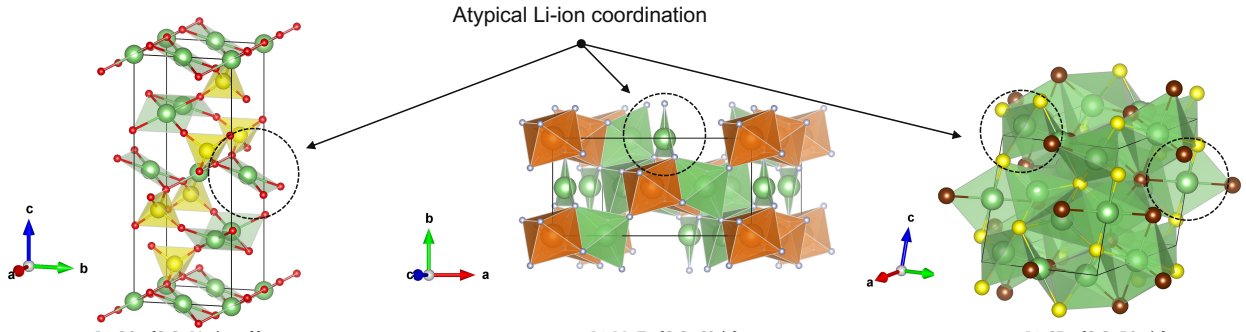


Li_3YCl_6 (S.G. $P312$)
E_{hull} = 0.06 eV/atom
$\sigma_{Li}^{300\,K} = 0.585 \pm 0.033$ mS/cm

Li_2HfCl_6 (S.G. $P4_2/mnm$)
E_{hull} = 0.02 eV/atom
$\sigma_{Li}^{300\,K} = 0.41 \pm 0.10$ mS/cm

$LiLaCl_4$ (S.G. $I4_1/amd$)
E_{hull} = 0.02 eV/atom
$\sigma_{Li}^{300\,K} = 0.40 \pm 0.06$ mS/cm

Figure 3. Examples of generated structures with high estimated Li-ion conductivity.



Li_2SO_4 (S.G. $I4_1/amd$)
E_{hull} = 0.16 eV/atom
$\sigma_{Li}^{300\,K} = 0.33 \pm 0.12$ mS/cm

Li_3MgF_5 (S.G. $C2/c$)
E_{hull} = 0.04 eV/atom

Li_3SBr (S.G. $P2_1/c$)
E_{hull} = 0.11 eV/atom
$\sigma_{Li}^{300\,K} = 0.62 \pm 0.12$ mS/cm

Figure 4. Examples of generated structures that pass thermodynamic screening filters but exhibit highly atypical Li-ion coordination environments.