

Generative Inverse Design of Solid-State Fluoride-Ion Electrolyte Candidates in Data-Sparse Chemical Space

Joohyeon Park

Gachon University

Seonwoo Choi

Gachon University

Myeongjae Heo

Sogang University

Heonjae Jeong

`h.jeong@sogang.ac.kr`

Sogang University

Article

Keywords:

Posted Date: October 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11026639/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Fluoride-ion batteries (FIBs) are constrained by the scarcity of solid-state electrolytes and the limited diversity of existing materials databases. Most computational searches start from predefined chemical systems or known fluorite- and tysonite-type prototypes. Here, we apply a diffusion-based generative foundation model to inverse-design fluoride-ion electrolyte candidates without prescribing cation identities, stoichiometry, space group, or parent framework, conditioning generation only on thermodynamic stability. Under a common screening framework, a compositionally general Alex-MP-20 model explores a broader elemental space but suffers heavy attrition during fluoride-specific filtering, whereas a fluoride-restricted MP-F model raises the per-sample yield of stable fluoride candidates roughly 28-fold. Hierarchical screening and first-principles validation identify five Ni-based computational leads: KNi_2F_5 and NaNi_2F_5 best combine band alignment, elastic stability, and interfacial reaction energies, while $\text{CsNaNi}_2\text{F}_6$ shows the weakest driving force for interfacial decomposition. Compositionally and structurally unconstrained generation, coupled with physics-based screening, can thus open underexplored regions of data-sparse chemical space.

Introduction

The development of next-generation batteries is fundamentally constrained by the lack of materials that can simultaneously deliver high energy density, intrinsic safety, and long-term scalability.^{1–3} Although lithium-ion batteries remain the dominant commercial technology, their reliance on organic liquid electrolytes introduces critical safety risks, most notably thermal runaway associated with lithium dendrite formation.⁴ In parallel, the growing dependence on geographically concentrated resources such as Li, Mn, and Co has intensified supply-chain vulnerabilities.⁵ While solid-state electrolytes have been proposed as a partial solution to safety concerns,⁶ most ceramic systems still rely on lithium as the charge carrier, leaving resource scarcity largely unaddressed.⁷ As an alternative chemistry, fluoride-ion batteries (FIBs) have attracted increasing attention.⁸ By employing fluoride ions as charge carriers, FIBs enable high-voltage electrochemistry and multi-electron conversion reactions, yielding theoretical volumetric energy densities approaching $\sim 5000 \text{ Wh L}^{-1}$.⁹ Moreover, FIB architectures are intrinsically compatible with inorganic solid electrolytes, offering substantial safety advantages through the elimination of flammable liquid components.¹⁰ Despite these promises, progress in FIB development has been severely limited by the scarcity of fast fluoride-ion conductors that are simultaneously electronically insulating.¹¹ To date, research efforts have largely focused on a narrow set of binary fluorides and solid-solution derivatives, such as LaF_3 - and PbSnF_4 -based systems,^{10,12–14} thereby restricting exploration to limited compositional and structural spaces. The discovery of new inorganic solid-state materials requires navigating a high-dimensional and strongly non-linear relationship among composition, crystal structure, and functional properties. Conventional forward modeling approaches, including density functional theory (DFT), high-throughput screening, and evolutionary structure searches, have proven effective for evaluating known or hypothesized candidates,^{15,16} but inherently

confine exploration to pre-defined chemical compositions and structural prototypes. As a result, these methods struggle to access sparsely populated regions of chemical space, such as those relevant to fluoride-ion conductors.

Inverse design strategies enabled by machine learning-based generative models offer a powerful alternative, in which candidate structures satisfying predefined stability or property constraints are directly generated.¹⁷ Recent advances in generative materials modeling, particularly diffusion-based approaches, have demonstrated the ability to generate diverse and physically plausible crystal structures by learning joint probability distributions over compositions and atomic configurations.¹⁸ To be effective in practical materials discovery, however, such models must combine broad chemical generality with explicit control over thermodynamic stability. MatterGen, a diffusion-based crystal generative foundation model, represents a state-of-the-art framework that satisfies these requirements.¹⁹ Trained on large-scale inorganic DFT datasets derived from the Materials Project²⁰ and Alexandria,^{21,22} MatterGen supports stability-conditioned generation through the energy above convex hull metric (E_{hull}), substantially reducing the computational burden of downstream first-principles screening. While this approach has recently proven successful in the relatively data-rich lithium-ion battery domain,²³ solid-state FIB electrolytes occupy a markedly different regime, characterized by sparse data coverage and a limited number of known structural archetypes, predominantly fluorite- and tysonite-type phases.^{10,24} FIB electrolytes therefore constitute a stringent and physically meaningful testbed for evaluating the extrapolative capability of generative inverse design in data-sparse chemical spaces.

In this work, we investigate whether generative inverse design can identify solid-state fluoride-ion electrolyte candidates without prescribing a specific chemical system or relying on established structural prototypes. At the generation stage, atom types, stoichiometry, and space group are not fixed. Thermodynamic stability is used as the sole explicit property condition. Fluoride relevance is introduced through the choice of training domain and post-generation chemical screening rather than through predefined candidate formulas. This strategy enables compositionally and structurally unconstrained exploration beyond conventional fluorite- and tysonite-based design spaces. By comparing a compositionally general model with a fluoride-domain-aligned model, we further examine how training-data distribution governs generative efficiency in a sparsely represented functional chemical space. Generated structures are then subjected to hierarchical high-throughput screening and first-principles validation of thermodynamic, electronic, mechanical, and interfacial properties. To the best of our knowledge, this study is the first application of compositionally and structurally unconstrained generative crystal design to solid-state FIB electrolyte candidates.

Results

Generation and hierarchical screening workflow

To examine how the generality of the training data governs generative behavior, we deliberately varied the compositional scope of the training corpora and performed generation and downstream

screening independently for a compositionally general Alexandria-plus-Materials Project corpus (Alex-MP-20) and a fluoride-restricted Materials Project subset (MP-F). Generated structures were first screened using E_{hull} and electronic-insulation criteria, followed by fluoride-chemistry, elemental-safety, charge-neutrality, and structural filters. First-principles DFT calculations were then used to evaluate thermodynamic and elastic stability. Fig. 1 summarizes this workflow together with the retained crystal structures. Because no cation set, chemical system, stoichiometry, space group, or structural prototype was prescribed at the sampling stage, the workflow is chemistry-domain constrained but compositionally and structurally unconstrained. The complete workflow yields one candidate from Alex-MP-20 and four from MP-F.

The Alex-MP-20 and MP-F models generated 108,000 and 21,600 structures, respectively. Applying a MatterSim-predicted $E_{\text{hull}} < 0.1 \text{ eV atom}^{-1}$ stability criterion retained 76,344 Alex-MP-20 and 13,235 MP-F structures. Additionally, fluoride-only screening was performed for Alex-MP-20, substantially reducing the candidate pool to 2,356 structures. An ALIGNN-predicted band gap criterion of $E_g \geq 3.5 \text{ eV}$ reduced the candidate pools to 672 and 1,259 structures, respectively. Excluding compounds containing noble gases, radioactive elements, or Hg left 670 and 1,227 structures, which were further reduced to 438 and 877 structures through oxidation-state screening. DFT calculations converged for 404 Alex-MP-20 and 801 MP-F structures. Subsequent DFT-based screening sequentially applied an $E_{\text{hull}} < 0.02 \text{ eV atom}^{-1}$ criterion, exclusion of O- and H-containing compounds, and an HSE06-calculated $E_g \geq 3.5 \text{ eV}$ criterion, reducing the Alex-MP-20 and MP-F candidate pools to 14 and 57 structures, then to 8 and 28, and finally to 5 and 13 structures, respectively. Finally, mechanical-stability screening based on the calculated elastic properties yielded one final candidate from Alex-MP-20 and four from MP-F.

Elemental and stability landscapes of the generated structures

Elemental occurrence in the generated fluoride-containing materials depends strongly on the training corpus (Fig. 2), which is compositionally broad for Alex-MP-20 (Fig. 2a) and fluoride-restricted for MP-F (Fig. 2b). The heatmaps report elemental occurrence counts for generated candidates that satisfy the thermodynamic stability criterion of MatterSim, a machine learning force field (MLFF) trained on large-scale first-principles data ($E_{\text{hull}} < 0.1 \text{ eV atom}^{-1}$). The model trained on Alex-MP-20 exhibits a broadly distributed elemental landscape spanning alkali and alkaline-earth metals, transition metals, post-transition metals, and lanthanides. This wide dispersion reflects the compositionally unfiltered nature of the Alex-MP-20 training corpus, which preserves extensive compositional diversity even after stability-based filtering. In contrast, the MP-F trained model yields a markedly localized elemental distribution, with counts strongly concentrated among a limited set of elements, including Li, O, V, F, and Mn. This localization directly reflects the intrinsic compositional bias imposed by the fluoride-restricted training dataset. Despite both models being evaluated under identical stability criteria, these results demonstrate that generative outcomes are strongly governed by the compositional priors encoded in the training data, underscoring the critical role of dataset design in generative materials discovery.

Stability-conditioned generation retained 76,344 of 108,000 Alex-MP-20 samples (70.7%) and 13,235 of 21,600 MP-F samples (61.3%) below 0.1 eV atom^{-1} (Figs. 3a and 3b). Imposing fluoride relevance then separates the two models sharply. Only 2,356 of the stable Alex-MP-20 structures (3.09%) contain fluorine, whereas the MP-F model generates fluorides by construction and retains all 13,235. Normalized to the number of samples drawn, domain-aligned training therefore raised the yield of stable fluoride candidates from 218 to 6,127 per 10,000 samples, a factor of approximately 28. The much smaller number of fluoride-containing candidates entering electronic screening for Alex-MP-20 is apparent from the total counts in the band gap (E_g) histograms of Figs. 3c and 3d. The normalized E_g distributions from the two datasets show comparable profiles, indicating that the discrepancy in raw candidate counts is governed primarily by sampling budget and fluoride-specific filtering rather than a systematic shift in predicted electronic structure. Because the initial sampling sizes differ by a factor of five, absolute counts are not interpreted as a direct measure of model performance. Instead, the comparison focuses on per-sample retention under the tested settings. The compositionally general model also produced the higher fraction of thermodynamically stable structures per sample (70.7% versus 61.3%), indicating that narrowing the training domain carries a modest cost in stability conditioning. These counts include duplicate structures, which a post hoc StructureMatcher analysis identified in 1.5% of Alex-MP-20 and 32.5% of MP-F samples (Methods). If duplicates are distributed uniformly across stable and unstable outputs, the per-sample fractions and the yield ratio are unaffected; any concentration of duplicates among stable fluorides would instead make the 28-fold ratio an upper bound.

Parity plots of MatterSim-predicted against DFT-calculated E_{hull} for the Alex-MP-20 (Fig. 4a) and MP-F (Fig. 4b) candidate sets reveal a consistent offset. Across both datasets, MatterSim systematically predicts lower E_{hull} values than DFT, resulting in a consistent downward deviation from the ideal parity line. Quantitatively, for Alex-MP-20, the root-mean-square error (RMSE) and mean absolute error (MAE) are both $0.06 \text{ eV atom}^{-1}$. For the MP-F dataset, both metrics are slightly larger, at $0.07 \text{ eV atom}^{-1}$. It should be noted that strict quantitative agreement with DFT is not required at this stage, as the primary role of MatterSim is to serve as a computationally efficient pre-screening tool. Because MatterSim systematically underestimates E_{hull} , it acts as a permissive pre-filter that minimizes false negatives, ensuring that stable candidates are not discarded before DFT validation.

Retained candidates and their magnetic ordering

Five candidate materials satisfy the complete screening criteria (Fig. 5): one originates from the Alex-MP-20 model and four from the MP-F model. All candidates were fully relaxed using DFT after generative sampling and hierarchical high-throughput filtering, and their structural and compositional information is provided in Table S1. Under the sampling budgets used and the same downstream filtering criteria, the fluoride-domain-aligned model yields a greater number of retained candidates, whereas the general model accesses a broader elemental space. This result supports a trade-off between chemical breadth and domain alignment in generative exploration of sparse materials spaces, while motivating normalized, equal-budget comparisons for quantitative model assessment. Normalized to the number of

stable fluoride structures entering downstream screening, the two models showed similar observed final-candidate retention rates: 0.042% for Alex-MP-20 (1 of 2,356) and 0.030% for MP-F (4 of 13,235). The advantage of domain alignment is therefore concentrated at the fluoride-retention step, where it is largely definitional, and reduces the sampling cost of reaching viable candidates rather than improving the intrinsic quality of the candidates generated.

The magnetic states of the five Ni-containing candidates were evaluated by comparing non-spin-polarized (nonmagnetic, NM), ferromagnetic (FM), and antiferromagnetic (AFM) configurations using spin-polarized DFT+U calculations. The FM and AFM configurations are nearly degenerate for LiNi_2F_5 , KNi_2F_5 , and NaNi_2F_5 (within 0.05 eV per formula unit), whereas $\text{CsNaNi}_2\text{F}_6$ and NiF_2 show a clearer preference for AFM ordering (0.19 eV per formula unit; Table S3). The calculated average absolute magnetic moments of Ni atoms and total energies per formula unit for the NM, FM, and AFM configurations are summarized in Tables S2 and S3, respectively. The lowest-energy configuration is reported for each candidate; the band-edge, elastic, and interfacial descriptors discussed below were obtained from the non-spin-polarized screening-stage calculations.

Surface-referenced band alignment

Surface-referenced band alignment was examined as an electronic descriptor relevant to oxidative and reductive stability. Conventional Perdew-Burke-Ernzerhof (PBE) functionals are known to underestimate E_g .^{25,26} HSE06 calculations were therefore performed to determine band-edge positions. Of the thirty-six candidates that reached the hybrid-functional stage, only eighteen exhibited HSE06-calculated E_g above 3.5 eV. Subsequent elastic-tensor analysis excluded thirteen mechanically unstable compounds, leaving LiNi_2F_5 , $\text{CsNaNi}_2\text{F}_6$, NiF_2 , KNi_2F_5 , and NaNi_2F_5 . In particular, NiF_2 shares its composition with the known tetragonal ($\text{P4}_2/\text{mm}$) and orthorhombic (Pnnm) phases reported in the Materials Project. The generated structure is nevertheless a distinct polymorph: across the entire symmetry-tolerance range examined, from 0.10 down to 0.001 Å, it is assigned to $\text{C2}/\text{m}$ or a lower-symmetry subgroup and never converges to either reported space group (Table S4). It is retained as a screening outcome but is not prioritized further, as it exhibits the lowest Young's and shear moduli of the retained set (Table S5).

Slab models were constructed using (001) surfaces for a consistent vacuum reference. Fig. 6 compares the conduction band minimum (CBM) and valence band maximum (VBM) of these candidates and PbSnF_4 with the F_2/F^- redox level (-7.33 eV)^{27,28} and a hypothetical 4 V anode level. The band edges of KNi_2F_5 and NaNi_2F_5 span the selected reference window, prioritizing them for further electrochemical and transport validation in future work. This band-alignment analysis does not include thermodynamic decomposition potentials or interphase formation and therefore should not be interpreted as a complete electrochemical stability window.

Interfacial reactivity with representative electrodes

Thermodynamic interfacial reactivity with representative electrode materials was evaluated to further prioritize the inverse-designed candidates. Reaction energies were computed using the procedure described in the Methods, and the resulting values for all electrolyte-electrode pairs are listed as bar charts in Fig. S1. Values closer to zero indicate a smaller thermodynamic driving force for interfacial decomposition at 0 K, but do not by themselves establish kinetic or operational interface stability.

Fig. 7 summarizes, as a heatmap, the interfacial reaction energies between the five retained candidates and representative cathode materials (BiF_3 , FeF_3 , and CuF_2) and anode materials (Sn, Pb, and Ce).^{8,29} Reaction energies near zero correspond to a smaller thermodynamic decomposition driving force, whereas increasingly negative values indicate a stronger tendency toward interfacial reaction at 0 K. Clear electrode-dependent trends are observed. BiF_3 produces reaction energies closer to zero for several candidates than FeF_3 , for which the more negative values indicate a greater thermodynamic tendency toward interfacial decomposition. Sn and Pb show intermediate reaction energies with comparatively limited variation across the candidate set, whereas Ce produces strongly negative values for nearly all candidates. LiNi_2F_5 , $\text{CsNaNi}_2\text{F}_6$, KNi_2F_5 , and NaNi_2F_5 therefore emerge as the most favorable leads within the present electrode set because their reaction energies against BiF_3 and Sn are the least negative in the set. All pairs nevertheless show substantial 0 K driving forces (more negative than approximately $-0.7 \text{ eV atom}^{-1}$), so interphase formation is expected with every electrode examined and the ranking, rather than the absolute magnitude, is the relevant output of this analysis. These results indicate relative differences in thermodynamic driving force for decomposition, rather than confirmed long-term interfacial stability.

Discussion

In this work, we evaluate whether generative inverse design can navigate the highly data-sparse chemical space of solid-state FIB electrolyte candidates without fixing a specific composition, cation set, stoichiometry, space group, or structural prototype. MatterGen is applied under a minimal property-conditioning framework in which thermodynamic stability is the sole explicit generation target. Fluoride relevance is introduced through training-domain selection and downstream chemical screening, yielding an approach that is constrained to the functional chemistry of interest but remains compositionally and structurally unconstrained. This distinction enables exploration beyond the fluorite- and tysonite-derived families that dominate conventional FIB electrolyte searches.

A central finding is that the generated candidate space is strongly shaped by the compositional distribution of the training data. The compositionally general Alex-MP-20 model samples a broader elemental landscape, but many outputs are removed by fluoride-specific post-generation screening. The fluoride-focused MP-F model produces a more chemically aligned candidate pool and, under the sampling budgets used, contributes four of the five final candidates. Normalized to sampling effort, this corresponds to an approximately 28-fold increase in the yield of stable fluoride candidates, from 218 to 6,127 per 10,000 samples. These results indicate that domain alignment can improve downstream

retention in a sparse functional chemical space. The effect is, however, concentrated at the fluoride-retention step and is to a large extent definitional. Once conditioned on being a stable fluoride, the two models yield final candidates at similar observed rates. Because the two sampling budgets, source corpora, pretraining data (607,683 versus 13,414 structures), and duplicate fractions differ, the comparison is interpreted as evidence of a domain-alignment effect under the present workflow rather than as a universal quantitative ranking of the two models.

Through a multi-stage screening pipeline integrating MatterSim-based stability evaluation, E_g prediction, compositional and charge-neutrality filtering, and first-principles DFT validation, we identify a small set of fluoride-only solid electrolyte candidates that satisfy essential thermodynamic, electronic, and mechanical criteria. Notably, the retained candidates span triclinic, monoclinic, and tetragonal crystal systems and remain distinguishable across the full range of symmetry tolerances examined (Tables S4 and S6-S9), indicating that meaningful structural diversity can emerge even under strictly fluoride-only compositional constraints. This observation highlights the capacity of generative models, when coupled with physically motivated screening, to uncover nontrivial structural motifs in sparsely populated chemical spaces.

All five retained candidates exhibit lower energies in spin-polarized configurations than in the corresponding nonmagnetic states (Table S3), indicating that magnetic ordering is energetically relevant for this Ni-based candidate set. The FM and AFM configurations are relatively close in energy, with LiNi_2F_5 slightly favoring FM ordering and the remaining candidates favoring AFM ordering within the configurations examined. These results highlight the importance of accounting for magnetic ordering when evaluating the relative stability and electronic properties of Ni-containing fluoride candidates. Because the screening-stage energies, elastic constants, and band-edge positions were obtained without spin polarization, and because the E_{hull} values were referenced to Materials Project hull entries without a GGA/GGA + U compatibility correction, these descriptors should be regarded as approximate. Non-spin-polarized energies overestimate the total energy of Ni^{2+} compounds, so the E_{hull} criterion was conservative for the retained candidates, but candidates rejected at that stage may include false negatives. A fully spin-polarized, Hubbard-corrected re-evaluation of the screening pipeline is a natural refinement of the present workflow.

The multi-criteria screening results provide a basis for prioritizing the five finalists as computational leads. KNi_2F_5 and NaNi_2F_5 are prioritized because their band edges span the selected F_2/F^- -to-anode reference window (Fig. 6) and they exhibit comparatively small interfacial decomposition driving forces against BiF_3 cathodes and Sn anodes (Fig. 7). KNi_2F_5 also exhibits the highest stiffness of the candidate set (Hill-averaged bulk and Young's moduli of 85.2 and 93.8 GPa, respectively; Table S5). NaNi_2F_5 , by contrast, shows the largest Voigt-Reuss spread of the set (17.8 GPa in bulk modulus and 29.2 GPa in Young's modulus; Table S5), indicating pronounced elastic anisotropy. It is prioritized on the basis of its band alignment and interfacial energetics rather than its mechanical isotropy. Although its band edges do not span the full selected window, $\text{CsNaNi}_2\text{F}_6$ shows the least negative interfacial reaction energies

across the electrode set and a nearly isotropic elastic response (a Voigt-Reuss difference of 0.15 GPa in bulk modulus; Table S5), making it a strong secondary candidate. The consistently negative reaction energies at Ce interfaces indicate that Ce metal is unfavorable for the present Ni-based fluoride family, while LiNi_2F_5 is less attractive from a resource-diversification perspective because it contains lithium. More broadly, all five retained candidates are Ni-based, which reflects the prevalence of Ni-containing fluorides in the training corpora rather than an explicit design choice. Because nickel is itself a supply-constrained element, incorporating elemental abundance and criticality as explicit screening criteria is a natural extension of the present workflow.

The principal advance of this study is not only the identification of several fluoride compounds, but the demonstration that physically plausible candidates can emerge without prescribing their elemental combinations, stoichiometries, or structural prototypes. Coupling compositionally and structurally unconstrained generation with hierarchical high-throughput screening and first-principles validation provides a systematic route for exploring functional materials domains in which both training data and established design rules are limited. Generative modeling is therefore positioned as an exploratory front end that expands the candidate space before targeted physics-based calculations and experiments.

The present workflow establishes thermodynamic, electronic, elastic, and interfacial plausibility rather than directly confirming fluoride-ion conductivity. The five retained materials should therefore be regarded as prioritized computational leads, not validated solid electrolytes. Direct evaluation of native and defect-mediated fluoride migration, including defect formation energies, migration barriers, molecular-dynamics simulations, and experimental impedance measurements, remains necessary. This limitation defines the next validation stage while preserving the central methodological conclusion: compositionally and structurally unconstrained generation can open underexplored regions of fluoride chemical and structural space.

Methods

Stability-Conditioned Generative Modeling

MatterGen was employed to generate fluoride-containing candidate materials using a parallelized sampling workflow based on two independent source datasets: Materials Project (v2022.10.28)²⁰ and Alexandria (Materials Cloud Archive).^{21,22} By default, MatterGen generates a diverse collection of inorganic crystal structures with up to 20 atoms per primitive cell. To compare generation efficiency across datasets, two different training datasets were used. For the MP-derived dataset, a fluoride-specific subset (MP-F) was constructed by retaining only compounds containing fluorine. After excluding radioactive elements and noble-gas-containing compounds present in the Materials Project dataset, 13,414 structures were retained for model training. No additional restriction on the number of atoms per cell or E_{hull} was applied to the MP-F dataset, as imposing these criteria would have substantially reduced the amount of training data available. The model was first pretrained on this MP-F dataset and subsequently fine-tuned using E_{hull} as a conditioning property. In contrast, the model pretrained on the

Alexandria-plus-MP dataset (Alex-MP-20) was used without additional compositional filtering. The Alex-MP-20 dataset encompasses both fluorinated and non-fluorinated compounds, representing a compositionally general chemical space. The training set was further restricted to structures containing no more than 20 atoms per unit cell and having E_{hull} of less than 0.1 eV atom⁻¹. The released Alex-MP-20 dataset comprised 607,683 training and 67,521 validation structures. The pretrained MatterGen model was fine-tuned using E_{hull} as a conditioning property.

Candidate generation was conditioned exclusively on thermodynamic stability to prioritize physical realizability and synthetic accessibility. Specifically, structures were generated under an E_{hull} conditioning target of 0.05 eV atom⁻¹, ensuring that proposed candidates possessed favorable thermodynamic formation energies.³⁰ Sampling was performed using 1,000 diffusion steps with a diffusion guidance factor of 2.0. Each generation run produced 2,160 structures in 10 batches of 216. This run was repeated 50 times for the Alex-MP-20 model and 10 times for the MP-F model, yielding 108,000 and 21,600 generated structures, respectively. The repetitions were carried out independently, so a given structure may be sampled in more than one run. A duplicate analysis with the StructureMatcher module in pymatgen identified 1,628 duplicate candidates among the 108,000 structures generated by the Alex-MP-20 model and 7,020 duplicate structures among the 21,600 structures generated by the MP-F model. This analysis was performed post hoc; duplicates were not removed, and all counts reported in the Results refer to the full generated sets. The implications for the reported yields are discussed in the Results. At the sampling stage, no cation set, chemical system, stoichiometry, space group, or structural prototype was prescribed. Fluoride relevance was introduced through the MP-F training domain or, for the Alex-MP-20 model, through post-generation retention of F-containing structures.

Hierarchical High-Throughput Screening

Following generation, a multi-stage filtering pipeline was applied to identify stable, electronically insulating, and practically tractable fluoride-ion solid electrolyte candidates. First, generated structures were evaluated using the SUN (Stable, Unique, and Novel) criterion implemented in the MatterGen evaluation pipeline.¹⁹ Stability was assessed with MatterSim, an MLFF trained on large-scale first-principles data and used here to efficiently evaluate relative energetic stability across a wide range of inorganic compositions.³¹ Candidates were required to satisfy a stability threshold of $E_{\text{hull}} \leq 0.1$ eV atom⁻¹. Uniqueness and novelty were assessed with the same pipeline, which applies pymatgen StructureMatcher within the generated set (uniqueness) and against the Alex-MP reference dataset of 845,997 DFT-relaxed entries (novelty). Default settings were used for both models (Itol = 0.2, stol = 0.3, angle_tol = 5°, with primitive-cell reduction and volume scaling enabled), so the resulting statistics are directly comparable between the two. These statistics were used to characterize the generated pools; structures were not additionally removed on the basis of uniqueness or novelty. Duplicate structures across generation runs were identified post hoc with StructureMatcher, as described above, but were not removed. Second, electronic insulation was assessed using E_g as a screening criterion. The ALIGNN model³² was employed to rapidly estimate E_g directly from atomic structures, bypassing

computationally intensive first-principles calculations. To reduce the likelihood of electronic leakage, candidates were required to have E_g values larger than 3.5 eV consistent with the band gaps of established fluoride-ion solid electrolytes.³³ Third, compositional filters were applied to exclude elements incompatible with practical solid electrolyte chemistry. Candidates containing (i) noble gases, (ii) radioactive elements, or (iii) mercury were removed due to chemical inertness, regulatory constraints, or toxicity concerns.³⁴ Finally, formal charge neutrality was applied as a basic chemical-plausibility criterion. A candidate was retained if its total formal charge could sum to zero under a chemically reasonable combination of oxidation states. This filter removes compositions lacking a plausible formal oxidation-state assignment but does not by itself establish thermodynamic or kinetic stability. For transition-metal-containing compounds, which may exhibit variable valency, structures were retained provided that charge neutrality could be achieved under any chemically reasonable combination of oxidation states.

Validation by First-Principles Calculations

Candidates passing the post-generation filtering pipeline were further validated using first-principles calculations. Density functional theory (DFT) simulations were performed using the Vienna Ab initio Simulation Package (VASP).^{35,36} The Projector Augmented Wave (PAW) pseudopotential³⁷ and the Perdew-Burke-Ernzerhof (PBE)³⁸ Generalized Gradient Approximation (GGA) exchange-correlation functional were used. The plane-wave cutoff value was set to 520 eV. Monkhorst-Pack meshes were generated with a reciprocal-space sampling density of $2\pi \times 0.03 \text{ \AA}^{-1}$.³⁹ The lattice parameters and atomic positions were fully relaxed until the residual force on each atom fell below $0.01 \text{ eV \AA}^{-1}$. Elastic constants were subsequently computed from the relaxed structures. To mitigate E_g underestimation inherent in the PBE functional, electronic structure calculations were performed using the HSE06 hybrid functional.⁴⁰ Candidates with HSE06-calculated E_g below 3.5 eV were excluded. All screening-stage calculations, including structural relaxations, E_{hull} evaluation, elastic constants, and HSE06 band-edge positions, were performed without spin polarization and without a Hubbard correction, and E_{hull} was evaluated against Materials Project hull entries without a GGA/GGA+U compatibility correction. For the five retained candidates, the magnetic ordering was subsequently examined with spin-polarized PBE+ U_{eff} calculations using an effective U_{eff} ($U - J$) of 6.2 eV for the Ni 3d states in the Dudarev formulation, consistent with the Materials Project convention (Tables S2 and S3); these calculations establish the preferred magnetic ordering and were not propagated to the screening-stage descriptors. Following the structural relaxations, stability analysis was performed. Thermodynamic stability was evaluated using E_{hull} , and only candidates with $E_{\text{hull}} < 0.02 \text{ eV atom}^{-1}$ were retained.³⁰ To avoid the formation of oxide or hydroxide phases that can deteriorate electrolyte performance, an additional screening criterion was applied to exclude candidates containing oxygen or hydrogen. Such secondary phases also strengthen metal-oxygen bonding and reduce fluoride-ion mobility, thereby lowering ionic conductivity.¹¹ In addition, elastic properties obtained from DFT calculations were used as a mechanical screening criterion. The elastic moduli were obtained using the Voigt-Reuss-Hill averaging scheme.⁴¹

Elastic stability was assessed using the Born criterion, requiring the elastic matrix (C_{ij}) to be positive definite.⁴² Candidates violating this condition were discarded, since a non-positive-definite C_{ij} implies a lattice that is unstable against long-wavelength deformation. For the remaining structures, the Hill averages B_H and G_H , defined as the arithmetic means of the Voigt and Reuss bounds, were adopted as the representative bulk and shear moduli. Space-group symmetries were determined using the spglib library as implemented in pymatgen.^{43,44} To ensure consistency between the assigned symmetry and the relaxed lattice geometry, an adaptive symmetry-tolerance scheme was used. The symmetry tolerance (symprec) was set to 0.10 Å and iteratively reduced in increments of 0.01 Å, down to a minimum of 0.001 Å. For each structure, the final space group was assigned using the largest symprec value at which the symmetry-consistent cell reproduced the relaxed lattice metrics. The selected symprec values for LiNi_2F_5 , $\text{CsNaNi}_2\text{F}_6$, NiF_2 , KNi_2F_5 , and NaNi_2F_5 were 0.001, 0.1, 0.01, 0.01, and 0.02 Å, respectively. An angular tolerance of 5° was applied throughout the analysis. This choice assigns the highest symmetry compatible with the relaxed geometry while avoiding over-symmetrization of the generated cells. To confirm that the assignments are not artifacts of this choice, space groups were additionally evaluated across the full tolerance range from 0.10 to 0.001 Å for every retained candidate (Tables S4 and S6-S9). Electrochemical stability was further evaluated using three-layer (001) slab models constructed from the relaxed bulk structures with a 15 Å vacuum region introduced along the z direction. For each candidate, the surface termination showing the most favorable band-edge alignment with respect to the F_2/F^- redox level was selected for the electrochemical stability analysis. Accordingly, F-, Cs-, Ni-, K-, and Na-terminated surfaces were selected for LiNi_2F_5 , $\text{CsNaNi}_2\text{F}_6$, NiF_2 , KNi_2F_5 , and NaNi_2F_5 , respectively. The planar-averaged electrostatic potential along the z direction was calculated to align the HSE06 band-edge positions relative to the vacuum level. A dipole correction was not applied because the non-orthogonal lattice geometry of the slab cells is incompatible with the surface-normal dipole correction scheme. The vacuum-referenced band edges in Fig. 6 are therefore reported with this limitation in mind. The Dudarev formulation of the Hubbard correction was used.⁴⁵

Thermodynamic Interfacial Reactivity Analysis

Interfacial reactivity between the electrolyte and electrode was evaluated in terms of the thermodynamic reaction driving force obtained from a convex-hull analysis. For each electrode-electrolyte pair, a reference convex hull was constructed with pymatgen⁴³ from Materials Project entries lying within the combined chemical system. Pseudo-binary compositions are given by

$$C(x) = xA + (1-x)B, 0 \leq x \leq 1 \text{ (Eq. 1),}$$

where $C(x)$ is the composition of the electrolyte-electrode mixture, A and B are the compositions of the electrolyte and electrode, respectively, and x is the molar fraction of the electrolyte. For each composition $C(x)$, the reaction energy was computed as the energy difference between $C(x)$ and the equilibrium phases of the reference phase diagram at that composition, normalized per atom of the

mixture. The most negative reaction energy over the scanned composition range was taken as the measure of the interfacial reactivity. All calculations were performed at 0 K.

Declarations

Code Availability

The custom scripts and configuration files used for structure generation, hierarchical high-throughput screening, and post-processing are available at GitHub URL: <https://github.com/Jeong-Research-Group/FIB.git>. MatterGen, MatterSim, ALIGNN, and pymatgen are publicly available from their respective repositories.

Data Availability

The relaxed crystal structures of the final fluoride-ion electrolyte candidates and the source data underlying the figures are available at GitHub URL: <https://github.com/Jeong-Research-Group/FIB.git>. Additional data are available from the corresponding author upon reasonable request.

Author Contributions

J.P. and S.C. developed the generative-screening methodology, carried out the calculations, wrote the software, curated and analyzed the data, validated the results, prepared the figures, and drafted the manuscript; they contributed equally to this work. M.H. contributed to data curation, software development, and validation. H.J. conceived the project, designed the computational workflow, administered and supervised the work, and drafted the manuscript. All authors reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Acknowledgements

This work was supported by the Sogang University Research Grant of 202510019.01. We gratefully acknowledge the use of high-performance computing (HPC) resources, specifically the Nurion and Neuron clusters, provided by the KISTI R&D Innovation-Support Program at the Korea Institute of Science and Technology Information (KISTI).

ORCID

Joohyeon Park: <https://orcid.org/0009-0001-5779-6463>

Seonwoo Choi: <https://orcid.org/0009-0004-7930-701X>

References

1. Larcher, D. & Tarascon, J. M. Towards greener and more sustainable batteries for electrical energy storage. *Nat. Chem.* **7**, 19–29 (2015).
2. Armand, M. & Tarascon, J.-M. Building better batteries. *Nature* **451**, 652–657 (2008).
3. Dunn, B., Kamath, H. & Tarascon, J.-M. Electrical energy storage for the grid: a battery of choices. *Science* **334**, 928–935 (2011).
4. Xu, X. *et al.* Dendrite-accelerated thermal runaway mechanisms of lithium metal pouch batteries. *SusMat* **2**, 435–444 (2022).
5. Ren, H. *et al.* Vulnerability to geopolitical disruptions of the global electric vehicle lithium-ion battery supply chain network. *Comput. Ind. Eng.* **188**, 109919 (2024).
6. Manthiram, A., Yu, X. & Wang, S. Lithium battery chemistries enabled by solid-state electrolytes. *Nat. Rev. Mater.* **2**, 16103 (2017).
7. Janek, J. & Zeier, W. G. Challenges in speeding up solid-state battery development. *Nat. Energy* **8**, 230–240 (2023).
8. Nowroozi, M. A. *et al.* Fluoride ion batteries—past, present, and future. *J. Mater. Chem. A* **9**, 5980–6012 (2021).
9. Xiao, A. W., Galatolo, G. & Pasta, M. The case for fluoride-ion batteries. *Joule* **5**, 2823–2844 (2021).
10. Rongeat, C., Anji Reddy, M., Witter, R. & Fichtner, M. Solid Electrolytes for Fluoride Ion Batteries: Ionic Conductivity in Polycrystalline Tysonite-Type Fluorides. *ACS Appl. Mater. Interfaces* **6**, 2103–2110 (2014).
11. Zhang, M. *et al.* Recent progress, challenges and prospects of electrolytes for fluoride-ion batteries. *Energy Rev.* **3**, 100083 (2024).
12. Ji, Q. *et al.* Mechanochemical synthesis, microstructure and electrochemical properties of solid electrolytes with stabilized fluorite-type structure in the $\text{PbF}_2\text{-SrF}_2\text{-KF}$ system for solid-state fluoride-ion batteries. *Ceram. Int.* **49**, 16901–16908 (2023).
13. Xiong, L. *et al.* Exploring efficient solid electrolyte based on Nd doped BaSnF_4 for fluoride-ion batteries at atomic scale. *J. Power Sources* **518**, 230718 (2022).
14. Murakami, M. *et al.* High Anionic Conductive Form of $\text{Pb}_x\text{Sn}_{2-x}\text{F}_4$. *Chem. Mater.* **31**, 7704–7710 (2019).
15. Curtarolo, S. *et al.* The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).
16. Oganov, A. R., Pickard, C. J., Zhu, Q. & Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **4**, 331–348 (2019).

17. Franceschetti, A. & Zunger, A. The inverse band-structure problem of finding an atomic configuration with given electronic properties. *Nature* **402**, 60–63 (1999).
18. Park, H., Li, Z. & Walsh, A. Has generative artificial intelligence solved inverse materials design? *Matter* **7**, 2355–2367 (2024).
19. Zeni, C. *et al.* A generative model for inorganic materials design. *Nature* **639**, 624–632 (2025).
20. Jain, A. *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
21. Schmidt, J. *et al.* Machine-Learning-Assisted Determination of the Global Zero-Temperature Phase Diagram of Materials. *Adv. Mater.* **35**, 2210788 (2023).
22. Schmidt, J., Wang, H. C., Cerqueira, T. F. T., Botti, S. & Marques, M. A. L. A dataset of 175k stable and metastable materials calculated with the PBEsol and SCAN functionals. *Sci. Data* **9**, 64 (2022).
23. Parida, C., Roy, D., Lastra, J. M. G. & Bhowmik, A. Mining Chemical Space with Generative Models for Battery Materials. *Batter. Supercaps* **9**, e202500309 (2026).
24. Mohammad, I., Chable, J., Witter, R., Fichtner, M. & Reddy, M. A. Synthesis of Fast Fluoride-Ion-Conductive Fluorite-Type $\text{Ba}_{1-x}\text{Sb}_x\text{F}_{2+x}$ ($0.1 \leq x \leq 0.4$): A Potential Solid Electrolyte for Fluoride-Ion Batteries. *ACS Appl. Mater. Interfaces* **10**, 17249–17256 (2018).
25. Perdew, J. P. & Levy, M. Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities. *Phys. Rev. Lett.* **51**, 1884–1887 (1983).
26. Sham, L. J. & Schlüter, M. Density-Functional Theory of the Energy Gap. *Phys. Rev. Lett.* **51**, 1888–1891 (1983).
27. Bratsch, S. G. Standard Electrode Potentials and Temperature Coefficients in Water at 298.15 K. *J. Phys. Chem. Ref. Data* **18**, 1–21 (1989).
28. Trasatti, S. The absolute electrode potential: an explanatory note (Recommendations 1986). *Pure Appl. Chem.* **58**, 955–966 (1986).
29. Liu, J. *et al.* Unveiling kinetic characteristics of the interface between solid electrolyte and cathode for all-solid-state fluoride ion batteries towards a wide-temperature application. *J. Power Sources* **625**, 235736 (2025).
30. Sun, W. *et al.* The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).
31. Yang, H. *et al.* MatterSim: A Deep Learning Atomistic Model Across Elements, Temperatures and Pressures. Preprint at <https://arxiv.org/abs/2405.04967> (2024).
32. Choudhary, K. & DeCost, B. Atomistic Line Graph Neural Network for improved materials property predictions. *npj Comput. Mater.* **7**, 185 (2021).
33. Meena, M., Money, B. K. & Bhide, A. Physicochemical characterization of fluoride solid electrolyte PbSnF_4 synthesized via sonochemical reaction. *J. Fluor. Chem.* **285–286**, 110450 (2025).
34. Schoenung, J. M., Ogunseitan, O. A. & Heine, L. G. Sustainability-informed materials selection, design, discovery, and development. *MRS Bull.* **50**, 1225–1242 (2025).

35. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
36. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
37. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
38. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
39. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
40. Heyd, J., Peralta, J. E., Scuseria, G. E. & Martin, R. L. Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional. *J. Chem. Phys.* **123**, 174101 (2005).
41. Chung, D. H. & Buessem, W. R. The Voigt-Reuss-Hill Approximation and Elastic Moduli of Polycrystalline MgO, CaF₂, β -ZnS, ZnSe, and CdTe. *J. Appl. Phys.* **38**, 2535–2540 (1967).
42. Mouhat, F. & Coudert, F.-X. Necessary and sufficient elastic stability conditions in various crystal systems. *Phys. Rev. B* **90**, 224104 (2014).
43. Ong, S. P. et al. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
44. Togo, A., Shinohara, K. & Tanaka, I. Spglib: a software library for crystal symmetry search. *Sci. Technol. Adv. Mater. Methods* **4**, 2384822 (2024).
45. Dudarev, S. L., Botton, G. A., Savrasov, S. Y., Humphreys, C. J. & Sutton, A. P. Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study. *Phys. Rev. B* **57**, 1505–1509 (1998).

Figures

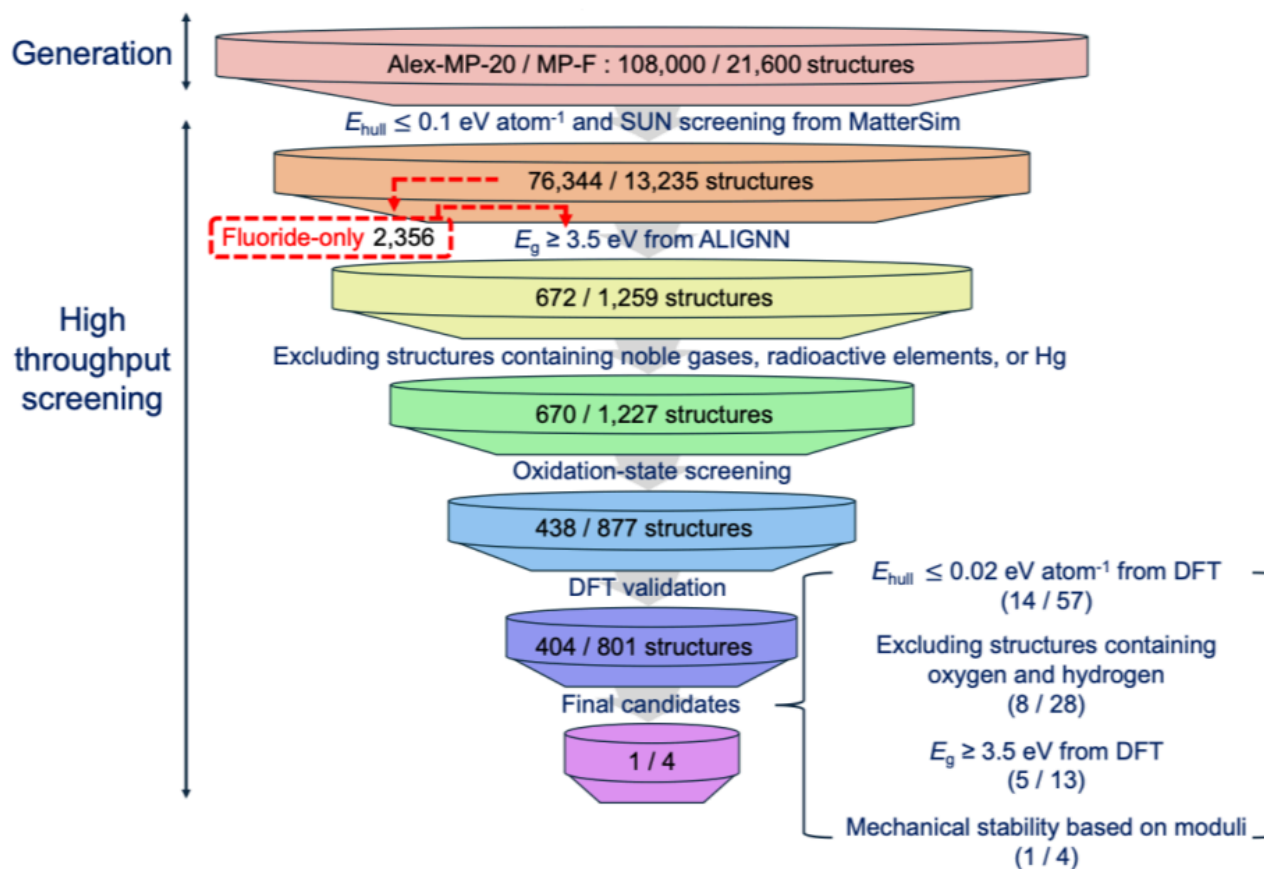


Figure 1

Crystal-structure generation and hierarchical high-throughput screening workflow, showing the progressive reduction from the generated structures to the five retained candidates for the Alex-MP-20 and MP-F models.

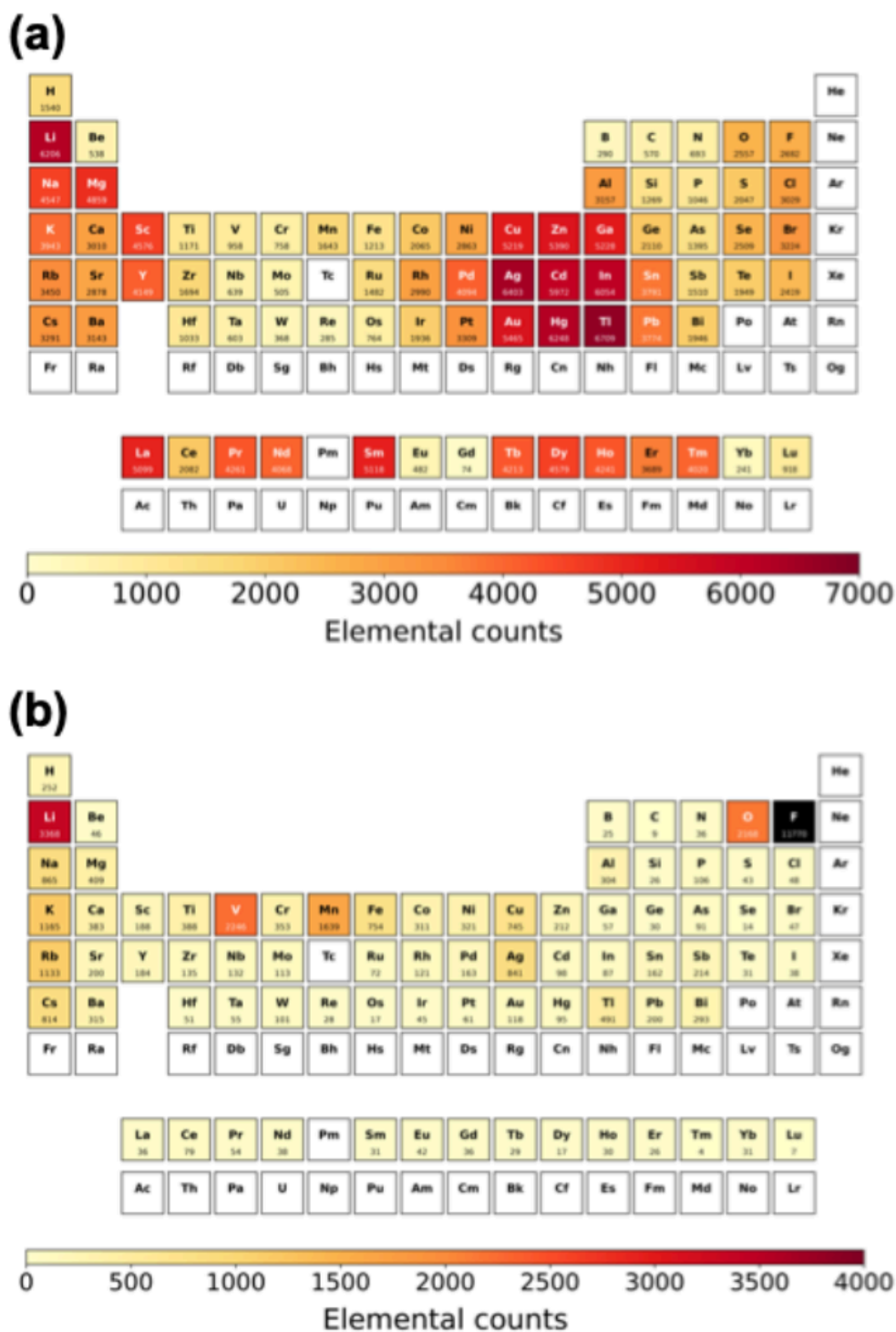


Figure 2

Periodic-table heatmaps of elemental occurrence counts in the stable ($E_{\text{hull}} \leq 0.1 \text{ eV atom}^{-1}$) fluoride-containing structures generated by models trained on (a) the compositionally general Alex-MP-20 dataset and (b) the fluoride-restricted MP-F dataset.

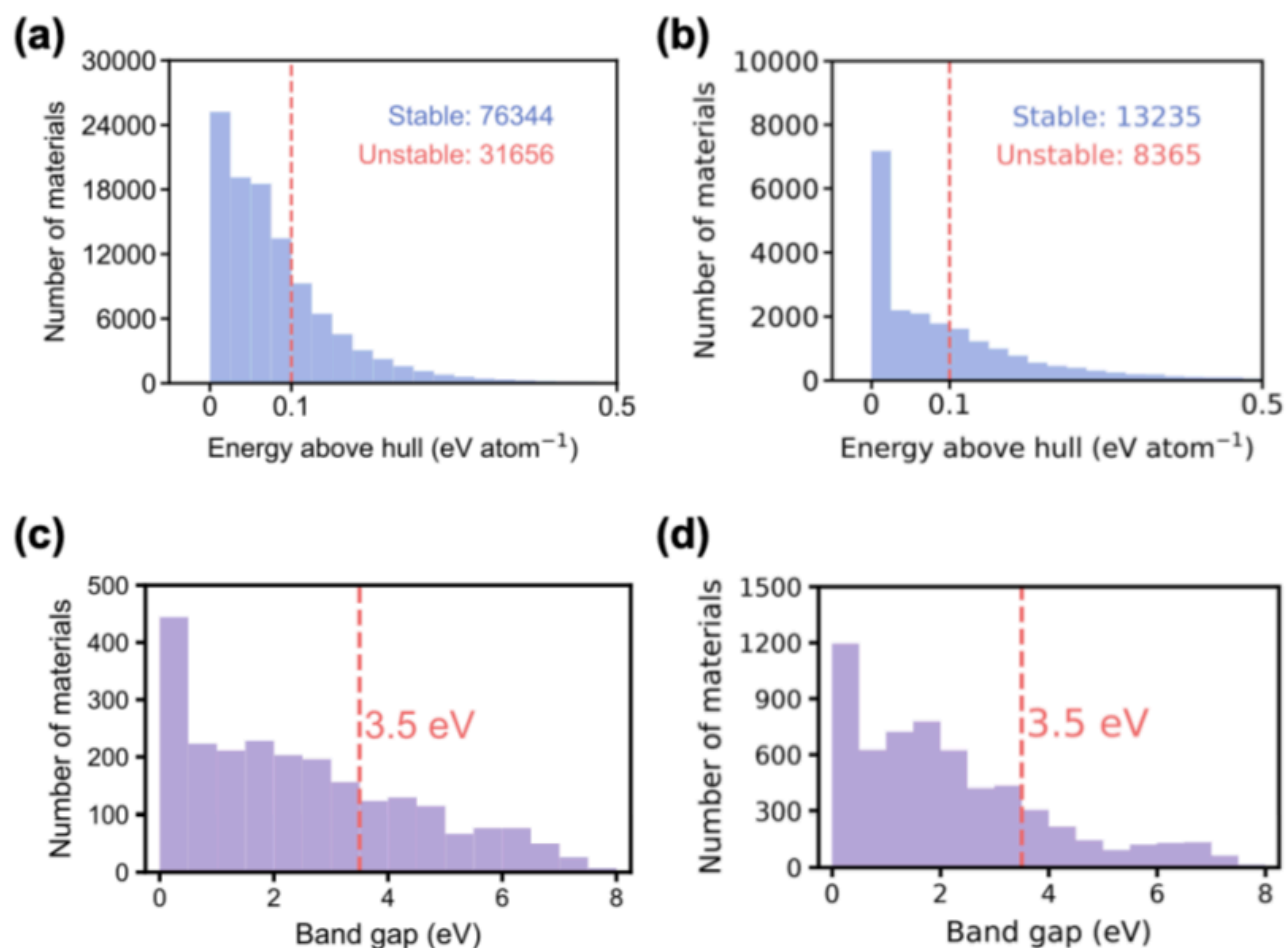


Figure 3

Histograms of MatterSim-predicted E_{hull} for all structures generated by models trained on (a) the Alex-MP-20 dataset and (b) the MP-F dataset. The dashed vertical line at 0.1 eV atom^{-1} indicates the threshold for stability. Histograms of ALIGNN-predicted E_g for the stable fluoride-containing structures from (c) the Alex-MP-20 model and (d) the MP-F model. The red dashed line at 3.5 eV indicates the threshold for electronic insulation suitable for electrolyte applications.

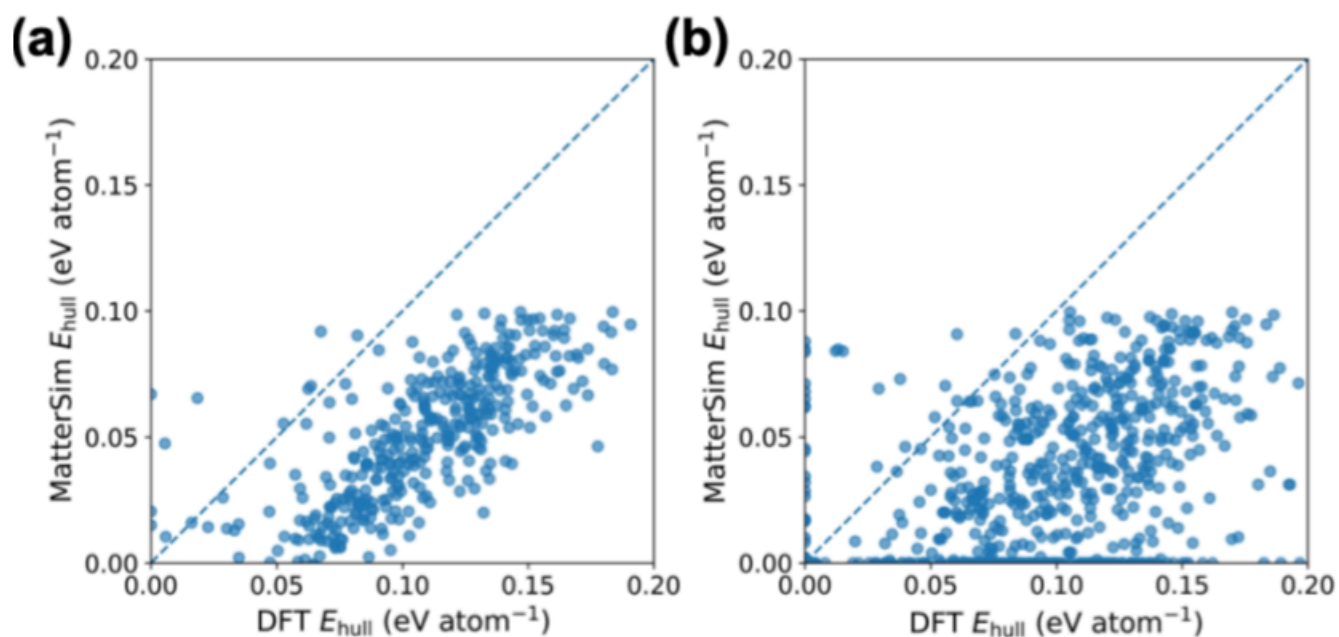


Figure 4

Parity plots of MatterSim-predicted E_{hull} against DFT-calculated values for the (a) Alex-MP-20 and (b) MP-F candidate sets. The dashed line represents perfect agreement ($y = x$).

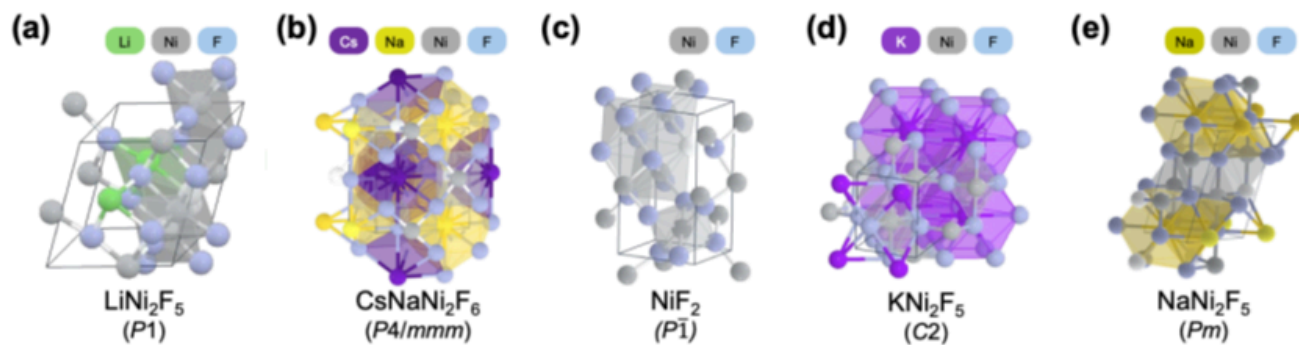


Figure 5

DFT-relaxed crystal structures of the five retained FIB electrolyte candidates: (a) LiNi_2F_5 from the Alex-MP-20 model, and (b) $\text{CsNaNi}_2\text{F}_6$, (c) NiF_2 , (d) KNi_2F_5 , and (e) NaNi_2F_5 from the MP-F model.

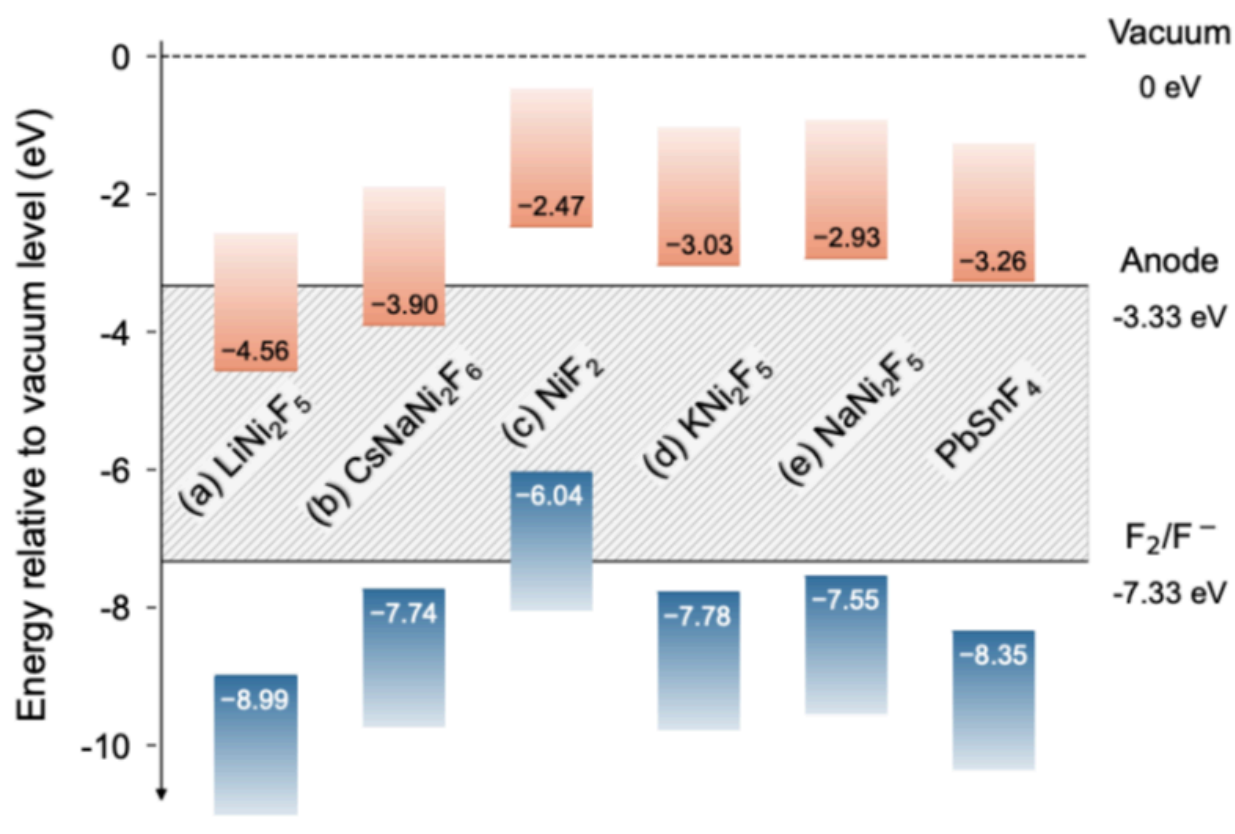


Figure 6

Surface-referenced band-edge alignment of the five retained FIB electrolyte candidates and the PbSnF_4 reference relative to the vacuum level (0 eV), from HSE06 calculations on (001) slab models. Orange and blue bars represent the CBM and VBM, respectively. The F_2/F^- redox level is located at -7.33 eV, and the hypothetical 4 V anode level at -3.33 eV. The comparison is an electronic band-alignment descriptor and not a complete thermodynamic electrochemical stability window.

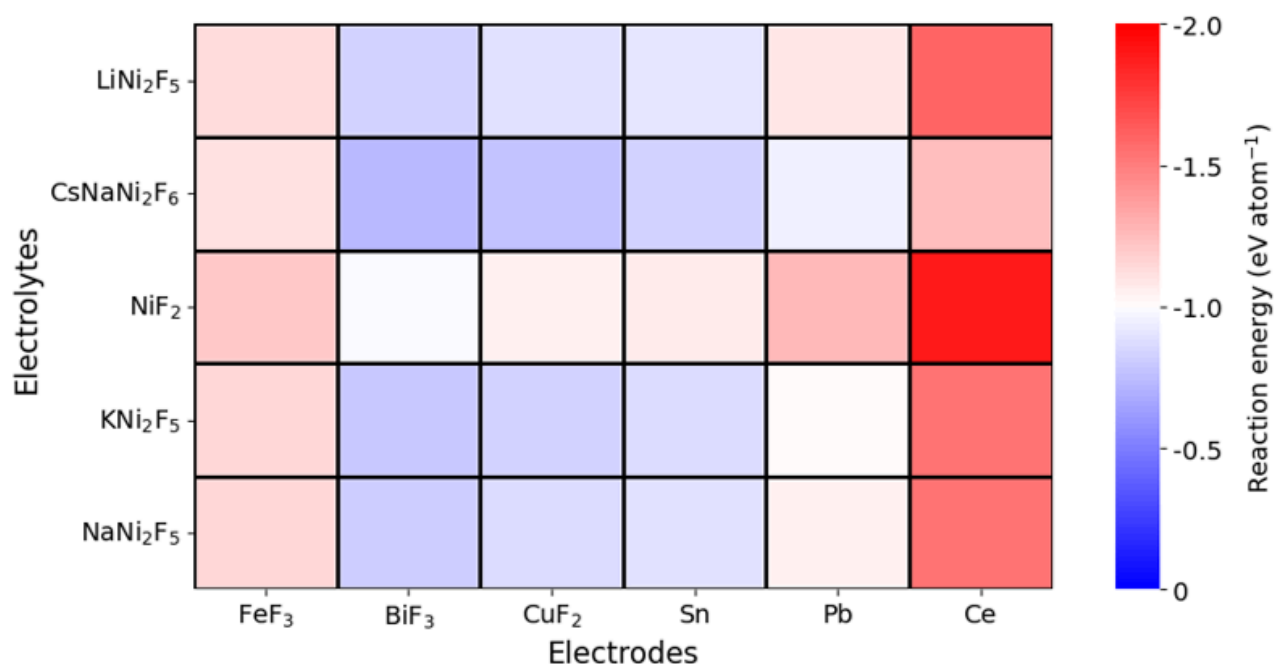


Figure 7

Heatmap of interfacial reaction energies (eV atom⁻¹) between the five retained FIB electrolyte candidates and six representative electrode materials (BiF₃, FeF₃, CuF₂, Sn, Pb, and Ce). Less negative values (blue) indicate a smaller 0 K thermodynamic driving force for decomposition, whereas increasingly negative values (red) indicate a stronger driving force. The individual values are shown as bar charts in Fig. S1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SMnpjComputMaterJeong.docx](#)