Journal of Materials Chemistry A

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d4ta00721b

Machine-learning assisted high-throughput discovery of solid-state electrolytes for Li-ion batteries†

Xingyu Guo, *[a] Zhenbin Wang,[bc] Ji-Hui Yang[a] and Xin-Gao Gong*[a]

The development of high-performance solid-state lithium-ion batteries (LIBs) requires designing solid-state electrolytes (SEs) with high ionic conductivity and excellent electrochemical stability. Here, we report 130 novel materials as promising SEs, identified through large-scale high-throughput calculations. These calculations employed a universal machine-learning interatomic potential (ML-IAP) as a surrogate model for density functional theory (DFT). In calculating the Li^+ conductivity of well-known conductors, we found that the universal ML-IAP tends to underestimate the activation energy for Li^+ conductivity by about 150 meV compared to DFT calculations. To identify practical SEs, we parameterized screening criteria for key material properties including synthesizability, interfacial stability, and ionic conductivity. We then established a tiered workflow for the accurate and efficient assessment of potential SEs. Additionally, we developed a tree-based ML model to unravel the relationship between structure, chemistry, and conductivity. Our findings indicate that features such as maximum packing efficiency, volume per atom, packing fraction, and differences in electronegativity are crucial in influencing Li^+ conduction. This work not only introduces novel SEs vital for the advancement of solid-state LIBs but also showcases the potential of ML in material innovation.

Received 31st January 2024
Accepted 26th March 2024

DOI: 10.1039/d4ta00721b

rsc.li/materials-a

Introduction

All solid-state lithium-ion batteries (LIBs) are widely recognized as a revolutionary technology for electrochemical energy storage. Compared to commercialized LIBs, solid-state LIBs, featuring nonflammable solid electrolytes (SEs), offer enhanced safety and cost-effectiveness, while also holding the promise of achieving higher energy density by utilizing Li metal anodes or high-energy cathodes.[1] Over the past few decades, significant progress has been achieved in the development of novel SEs with fast ionic conductivity, broader electrochemical stability, and excellent electrode compatibility. One of the potential contenders for SEs are the state-of-the-art fast ionic conductors such as $Li_{10}MP_2S_{12}$ (M = Si, Ge, Sn)[2–6] and lithium argyrodites $Li_6(P,Sb)S_5X$ (X = halogen).[7,8] They exhibit high ionic conductivities (1–10 mS cm^{-1}) that are comparable to commercial liquid electrolytes in LIBs. The garnet-type $Li_7La_3Zr_2O_{12}$ (LLZO)[9] and perovskite $La_{2/3-x}Li_{3x}TiO_3$ (LLTO)[10] oxides are proved to have a decent Li^+ conductivity (<1 mS cm^{-1}) with good electrochemical stability within a wide voltage range of ~3 V. Recently, halide SEs (Li_3MX_6, M = metal and X = halogen) are also reported to exhibit excellent electrochemical stabilities with cathode materials and high ionic conductivities (~1 mS cm^{-1}).[11–14]

Despite these tremendous achievements, there are few SEs satisfying all the requirements for practical applications. For instance, sulfide SEs tend to decompose at the SE/electrode interfaces and form detrimental interphases, resulting in either high interfacial impedance[15] or growth of lithium dendrites.[16] Moreover, oxide SEs commonly exhibit high stiffness, making it challenging to achieve conformal contact with the electrodes and accommodate volume changes during cycling. The incompatibility between electrodes and SEs poses risks of crack formation and internal short circuits during battery operations.[1] Therefore, there is a pressing need to develop novel SEs with superior performance for the commercialization of solid-state LIBs.

High-throughput (HT) density functional theory (DFT) calculations and machine learning (ML) predictions have recently become invaluable tools for the discovery of novel

[a]Department of Physics, Key Laboratory of Computational Physical Sciences (Ministry of Education), Institute of Computational Physical Sciences, State Key Laboratory of Surface Physics, Fudan University, Shanghai 200433, China. E-mail: xguo@fudan.edu.cn; xggong@fudan.edu.cn

[b]Department of Materials Science and Engineering, City University of Hong Kong, Hong Kong 999077, China

[c]School of Energy and Environment, City University of Hong Kong, Hong Kong 999077, China

† Electronic supplementary information (ESI) available: Li^+ conductivity of known lithium-ion conductors; benchmark of electronic band gaps; statistics of screening process; tree-based classification model on Li^+ conductivities. Structure prototypes of the 130 identified promising candidate materials. Calculated properties of identified promising solid-state electrolytes ($\sigma_{300\ K}$ < 10 mS cm^{-1}). See DOI: https://doi.org/10.1039/d4ta00721b

materials.[17–19] The HT DFT calculations typically begin with the assessment of a series of thermodynamic properties. Evaluation of the kinetic property (*e.g.*, diffusivity) is in general limited to a few promising candidates due to its high computational costs. In contrast, ML can rapidly screen the properties of tens of thousands of materials at a relatively low cost.[20–25] But the reliability of its prediction is considerably influenced by the volume and quality of available data, which remains one of the biggest challenges in AI-driven materials design.

In this work, using universal machine-learning interatomic potential (ML-IAP) as a surrogate model for DFT, we conducted comprehensive HT calculations of 740 000 lithium compounds to identify novel SEs. To streamline the identification process, we established screening criteria of key material properties essential for high-performance SEs. Utilizing this approach, we successfully pinpointed 130 materials as promising SEs. Furthermore, we developed a tree-based ML model to disentangle factors governing Li^+ conductivities in SEs. Based on insights gained, we propose strategies to address challenges impeding the development of high-performance solid-state LIBs.

Results

Benchmark of Li^+ conductivity

Conductivity is the most important metric for designing fast Li^+ conductors as SEs. Although previous study[26] has demonstrated the high accuracy of universal ML-IAP in calculating materials' thermodynamic properties, the effectiveness of this approach in evaluating materials' conductivity remains unclear. According to the benchmark studies[19] of *ab initio* molecular dynamic (MD) simulations against experimental data, a quick estimation of the activation energy (E_a) of Li^+ conductivity can be accessed by the mean square displacements (MSDs) of short-time (50 ps) MD simulations at 800 K and 1200 K. It has been found that when the calculated $MSD_{800\ K} > 5\ Å^2$ and estimated $E_a < 400$ meV (derivation of E_a is provided in Methods), the given material is identified as a superionic Li^+ conductor.[19] Building upon this finding, we examine the uncertainty of universal ML-IAP in calculating $MSD_{800\ K}$ and E_a with respect to DFT calculations for assessing Li^+ conductivity.

We first evaluate the uncertainty in MSDs calculated using the universal ML-IAP compared to those obtained from DFT. In this work, an $MSD_{800\ K}$ exceeding 5 $Å^2$ is indicative of a material displaying baseline conductivities. Fig. 1(a) presents the calculated MSDs at 800 K for 18 well-known fast and slow Li^+ conductors, as outlined in Table S1.† These MSDs are derived from 50 ps MD simulations using both the universal ML-IAP and DFT. We observe that all known superionic Li^+ conductors, characterized with experimental σ_{Li^+} higher than 1 mS cm^{-1}, exhibit universal ML-IAP and DFT calculated $MSD_{800\ K}$ exceeding ~10 $Å^2$. In contrast, slower Li^+ conductors, such as LLZO, LLTO, $LiFePO_4$, NCM and $LiCoO_2$, show calculated $MSD_{800\ K}$ less than 5 $Å^2$ using both methods. This finding suggests limited harmonic oscillations of Li^+ ions within their equilibrium positions, rather than diffusive behavior. It should be noted that only pristine materials were used to calculate Li^+ conductivity in this work. Potential enhancements due to occupational disorder or defects are not included, as their computational cost is prohibitive.[27] This exclusion could lead to an underestimation of the calculated Li^+ conductivities in SEs, such as garnet-type oxides and anti-perovskites.

The E_a can be estimated by the ratio of the MSD computed at 1200 K and 800 K. Fig. 1(b) shows the calculated $MSD_{1200\ K}$ *versus* $MSD_{800\ K}$ for nine Li^+ conductors, which possess baseline conductivities ($MSD_{800\ K} > 5\ Å^2$) at 800 K. It is observed that the universal ML-IAP generally predicts larger MSDs at both temperatures, resulting in an average E_a about 150 meV lower than those by DFT. The ratios of $MSD_{1200\ K}$ to $MSD_{800\ K}$ computed using the universal ML-IAP are less discriminative as indicators of Li^+ conductivities compared to those derived from DFT. For instance, the SEs of the LGPS family, *i.e.* $Li_{10}GeP_2S_{12}$,

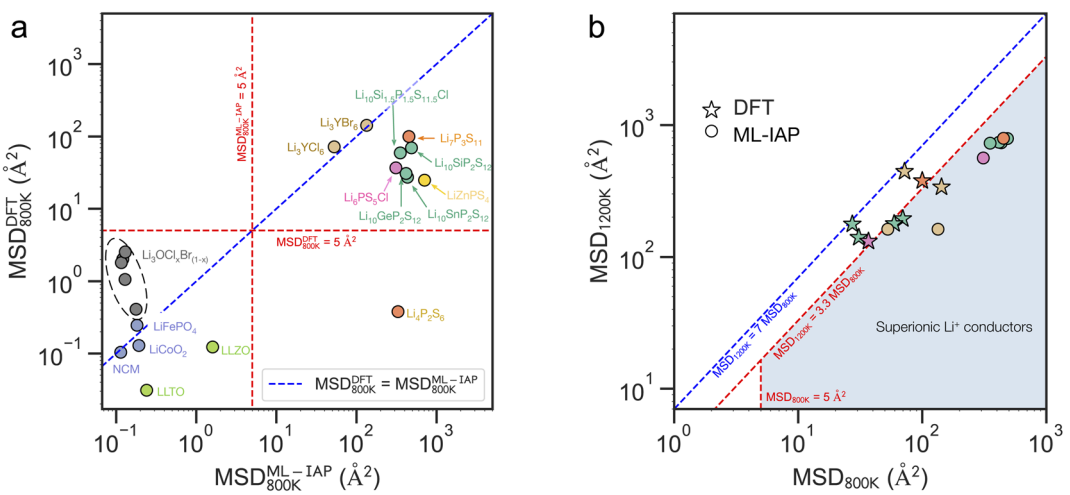

Fig. 1 (a) Comparison of mean square displacements (MSDs) calculated using machine-learning interatomic potential (ML-IAP) and density functional theory (DFT) at 800 K. The dashed red line denotes the $MSD_{800\ K} = 5\ Å^2$. The dashed blue line represents the $MSD_{800\ K}$ calculated by ML-IAP, which equals that of DFT. (b) Comparison of ML-IAP and DFT calculated ratio of MSDs at 1200 K and 800 K.

$Li_{10}SnP_2S_{12}$, and $Li_{10}SiP_2S_{12}$, have experimentally measured Li^+ conductivities of 12, 4, and 2.3 mS cm^{-1}, respectively. Their DFT-computed ratios of $MSD_{1200\ K}$ to $MSD_{800\ K}$ amount to 6.48, 4.55, and 2.80, in contrast to 1.69, 1.78, and 1.63, respectively, as calculated using universal ML-IAP. These ratios predicted by the universal ML-IAP are markedly more uniform compared to the diverse values derived from DFT, indicating a reduced sensitivity of the universal ML-IAP in evaluating E_a for materials with varying Li^+ conductivities. Consequently, we propose the criteria of $MSD_{800\ K} > 5\ Å^2$ and $MSD_{1200\ K}/MSD_{800\ K} < 3.3$ for fast evaluations of Li^+ conductivities.

Parameterization of high-throughput screening

A SE should meet a series of requirements for practical applications, which include fast ionic conduction, broad electrochemical windows, and good compatibility with electrodes (Fig. 2(a)). In response to these needs, we developed a high-throughput screening framework that considers multiple properties for designing promising SEs, as shown in Fig. 2(b). A crucial task in HT calculations is the parameterization of screening criteria, which directly influences the accuracy of predictions. Here, we present the rationale behind the designed screening criteria and

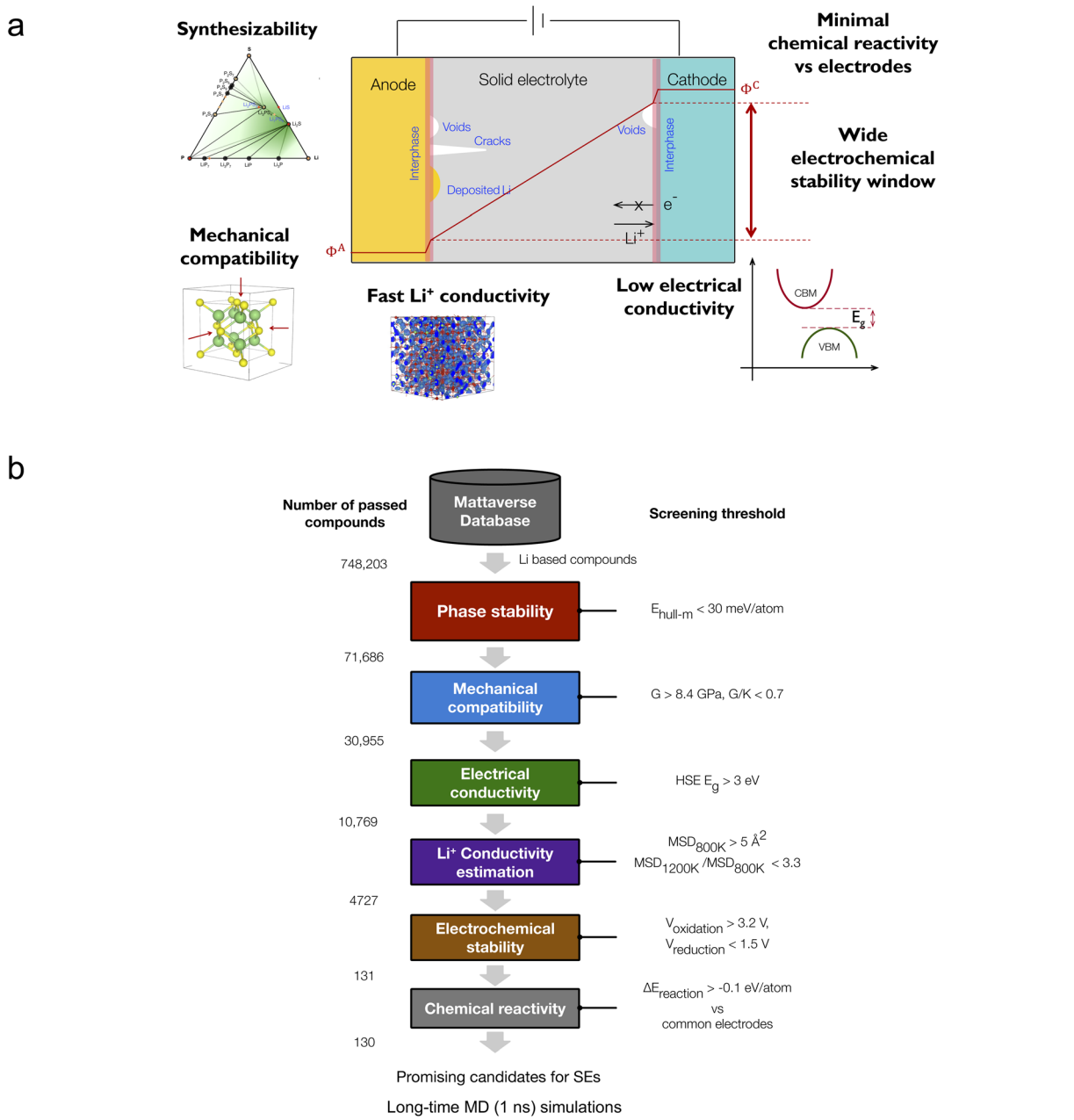

Fig. 2 (a) Schematic diagram illustrating material properties essential for the design of high-performance solid electrolytes (SEs). CBM and VBM represents conduction band minimum (CBM) and valence band maximum (VBM), respectively. E_g denotes the band gap. (b) Designed flowchart for the identification of promising SEs for solid-state lithium-ion batteries.

establish their values based on DFT calculations and experimental data.

Phase stability. The first criterion that new material discovery for technological applications must be satisfied is the phase stability or synthesizability. In line with previous work,[26] we use the energy distance ($E_{\text{hull-m}}$) to the Materials Project convex hull to estimate the new material's phase stability. According to previous high-throughput computational efforts,[19,28,29] a conservative threshold of 30 meV per atom was adopted for $E_{\text{hull-m}}$. The smaller the value of $E_{\text{hull-m}}$, the more stable the predicted material is, increasing the likelihood of successful synthesis in experiments.

Mechanical compatibility. An ideal SE should be soft and ductile such that it can deform and maintain good conformal contact with the electrodes, which is essential for long-term operations of solid-state LIBs. But it also necessitates a certain extent of hardness to inhabit the growth of lithium dendrites when coupling with Li metal anodes. The mechanical properties of a material are assessed by its calculated shear modulus (G) and bulk modulus (K) and the ductility is measured by Pugh's ratio (G/K). Previous studies suggest that (i) the buffer layer between SE and electrodes with a shear modulus (G) twice that of Li metal (4.2 GPa) can effectively suppress the growth of lithium dendrites.[30] (ii) Most common SEs exhibit Pugh's ratios (G/K) smaller than 0.7.[31] Based on these discoveries, we adopted criteria of $G > 8.4$ GPa and $G/K < 0.7$ to search for SEs with good mechanical compatibility.

Electronic conductivity. Electronic conductivity in SEs has been identified as a contributing factor to the degradation of solid-state LIBs. For example, electron conduction (e^-) can promote the nucleation and growth of Li dendrites at the interface of electrodes and SEs, ultimately leading to short-circuits and battery failures.[32,33] Here, we utilized the Heyd–Scuseria–Ernzerhof (HSE) band gaps (E_g) to approximate the electronic conductivity. It is important to note that while materials with small band gaps are highly likely to exhibit certain electronic conductivity, materials with wide band gaps may also exhibit conductivity due to defects introduced during synthesis and other processes.[16] Therefore, a screening criterion of HSE $E_g > 3$ eV was determined, based on the observation that most known SEs have a ML-predicted HSE E_g larger than 3 eV (Table S2†).

Li^+ conductivity. While ML-IAP allows us to explore Li^+ conductivity with significantly reduced computational costs, conducting calculations for tens of thousands of materials remains a computational challenge. To address this, we developed a two-step strategy for identifying fast Li^+ conductors. According to our benchmark results, a material is preliminarily classified as a fast ionic conductor if its $MSD_{800\ K} > 5$ Å^2 and $MSD_{1200\ K}/MSD_{800\ K} < 3.3$. In the next phase, long-time (1 ns) molecular dynamics (MD) simulations at 300 K were performed for the most promising candidates to calculate their Li^+ conductivity.

Electrochemical stability. The electrochemical stability of SEs relates to their ability to resist redox reactions when subjected to a range of voltages during operation. The voltage range, relative to Li metal, where an SE maintains stability against decomposition due to either lithium consumption or release, is defined as the electrochemical stability window.[15] The upper and lower boundaries of this window are defined as the oxidation and reduction limits, respectively. SEs with wide electrochemical stability windows are highly sought after for high-performance applications. Common anodes, such as Li metal and $Li_4Ti_5O_{12}$ (LTO), operate at potentials around 0 V and 1.5 V *vs.* Li/Li^+, respectively. Commercial cathodes, like $LiFePO_4$ (LFPO) and $LiNi_xMn_yCo_{1-x-y}O_2$ (NCM), typically operate within a potential range of 3.2 to 4.5 V *vs.* Li/Li^+. A material with a reduction limit below 1.5 V and an oxidation limit above 3.2 V is predicted to be a promising SE.

Chemical reactivity with electrodes. Besides redox reactions, chemical reactions between SEs and electrodes should be minimized. Previous DFT studies[34] have indicated that when the reaction energy (ΔE_{rxn}) between SEs and electrodes is ≥ -0.1 eV per atom, the SE would have a robust chemical reactivity with electrodes. Hence we applied this criterion to assess the stability of SEs in conjunction with common electrodes.

Screening analysis

The initial set of candidate SEs was constructed by selecting all compounds containing lithium from the Matterverse database,[26] which is the world's largest database of inorganic materials and contains ~33 million newly calculated structures. During this selection process, compounds containing actinide and lanthanide elements, with the exception of La and Lu, were excluded due to large errors in their DFT calculated properties.[35,36] The initial pool of candidates comprises 748 203 unique materials. To facilitate further analysis, these candidates are categorized into seven groups according to their anion chemistries, which plays an important role in determining the electrochemical properties of SEs.[15,34] These compounds are grouped into oxides, sulfides, fluorides, chlorides, bromides, alloys, and other compounds. In general, the number of materials passing through each filter decreases sequentially, indicating the effectiveness of the designed screening workflow (Fig. S1†).

Out of 748 203 unique Li-compounds, 71 686 materials exhibit an $E_{\text{hull-m}}$ lower than 30 meV per atom, indicating a higher potential for facile synthesis. This selection includes 13 097 oxides, 38 735 sulfides, and 19 773 halides, which comprise 3382 fluorides, 6767 chlorides, and 8624 bromides, in addition to 556 alloys and 448 compounds in other categories. Notably, compared to previous HT DFT calculations or ML predictions replying on databases like the Materials Project and Inorganic Crystal Structure Database,[17,19,29,34,37] there is a significant increase, at least tenfold, in the pool of potential candidates, which is especially noticeable in the sulfides and halides categories (Table S3†). This substantial expansion of the materials library offers great potential for exploring novel SEs previously uncharted in earlier studies.

Fig. 3(a) shows the distribution of calculated shear and bulk moduli for compounds with $E_{\text{hull-m}}$ lower than 30 meV per atom. We find that the majority of oxides, fluorides, and alloys

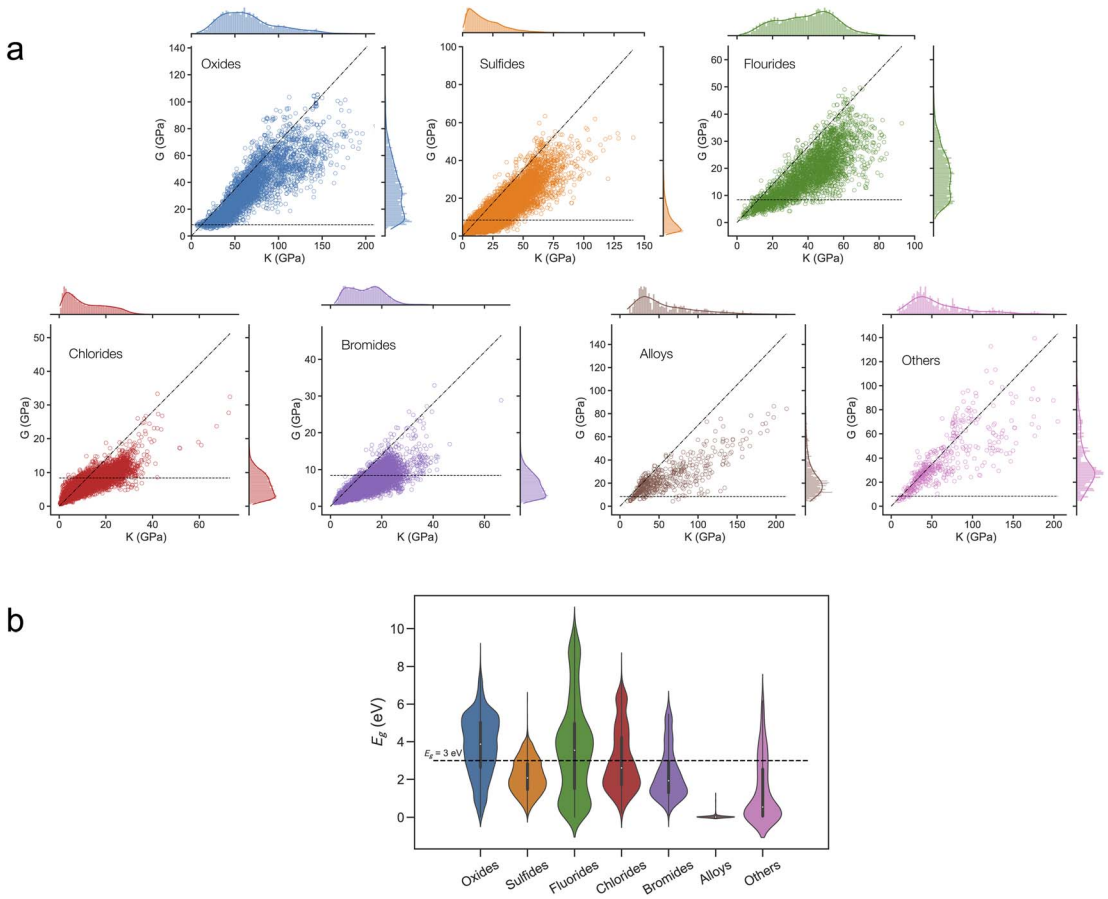

Fig. 3 (a) Calculated shear modulus (G) *vs.* bulk modulus (K) of candidate materials grouped by their anion chemistries. The dashed and dot-dashed lines stand for $G = 8.4$ GPa and $G/K = 0.7$, respectively. (b) Calculated HSE band gaps (E_g) for candidate materials categorized by their anion chemistries.

exhibit favorable mechanical compatibility. Oxides, in particular, tend to be harder than other material types (with higher G values). However, they also exhibit higher brittleness, indicated by their large Pugh's ratio (high G/B value). Approximately 25% of oxide candidates are eliminated at this stage due to their high risk of developing cracks and experiencing high-resistance contact at electrode/SE interfaces during battery operations. Fluorides are softer with G ranges from 1.4 to 50 GPa and less brittle, which are suitable for SE materials. Only around 30% of sulfides and 20% of chlorides and bromides meet the criteria for mechanical compatibility. Compared to oxides, sulfides and halides generally exhibit smaller G and K values. More specifically, a notable portion of sulfides, chlorides, and bromides demonstrates significant softness and compressibility (G < 20 GPa, K < 40 GPa), indicating an increased risk of failing to suppress lithium dendrite growth during cycling. As a result, these materials were excluded in this screening phase.

Fig. 3(b) presents ML-predicted HSE band gaps for candidate materials categorized by their anion chemistries. We observe that the majority of sulfides tend to have band gaps below 3.0 eV, suggesting a higher likelihood of electronic conduction and making them less suitable as SEs. In contrast, oxides and halides typically have wider band gaps, resulting in lower electronic conductivities compared to sulfides. Notably, a significant number of fluorides display exceptionally large band gaps, often exceeding 6 eV. Alloys and most other material types are typically known for their good electronic conductivity, characterized by band gaps significantly smaller than 3.0 eV. Consequently, these materials were excluded in the screening process for electronic insulators.

Next, the Li^+ conductivities were screened using our established benchmark criteria of $MSD_{800\ K} > 5\ Å^2$ and $MSD_{1200\ K}/MSD_{800\ K} < 3.3$. A total of 4727 compounds meets these criteria, indicating potential for superionic conductivities. Among these, sulfides emerge as the most promising candidates for superionic Li^+ conduction, with 1550 sulfide compounds passing the conductivity criteria. Indeed, most experimentally reported superionic Li^+ conductors to date are sulfides.[4–7] Additionally, 1151 oxides, 282 chlorides and 207 bromides passed the conductivity screening, indicating their potential for novel SE design in various chemical compositions. Fig. 4 shows the distribution of calculated $MSD_{1200\ K}$ *vs.* $MSD_{800\ K}$ for the 10 769 unique compounds. It is evident that many sulfides, chlorides, and bromides exhibit high baseline conductivity at 800 K ($MSD_{800\ K} > 5\ Å^2$), with half of them showing low E_a ($MSD_{1200\ K}/MSD_{800\ K} < 3.3$). In contrast, a considerable number of

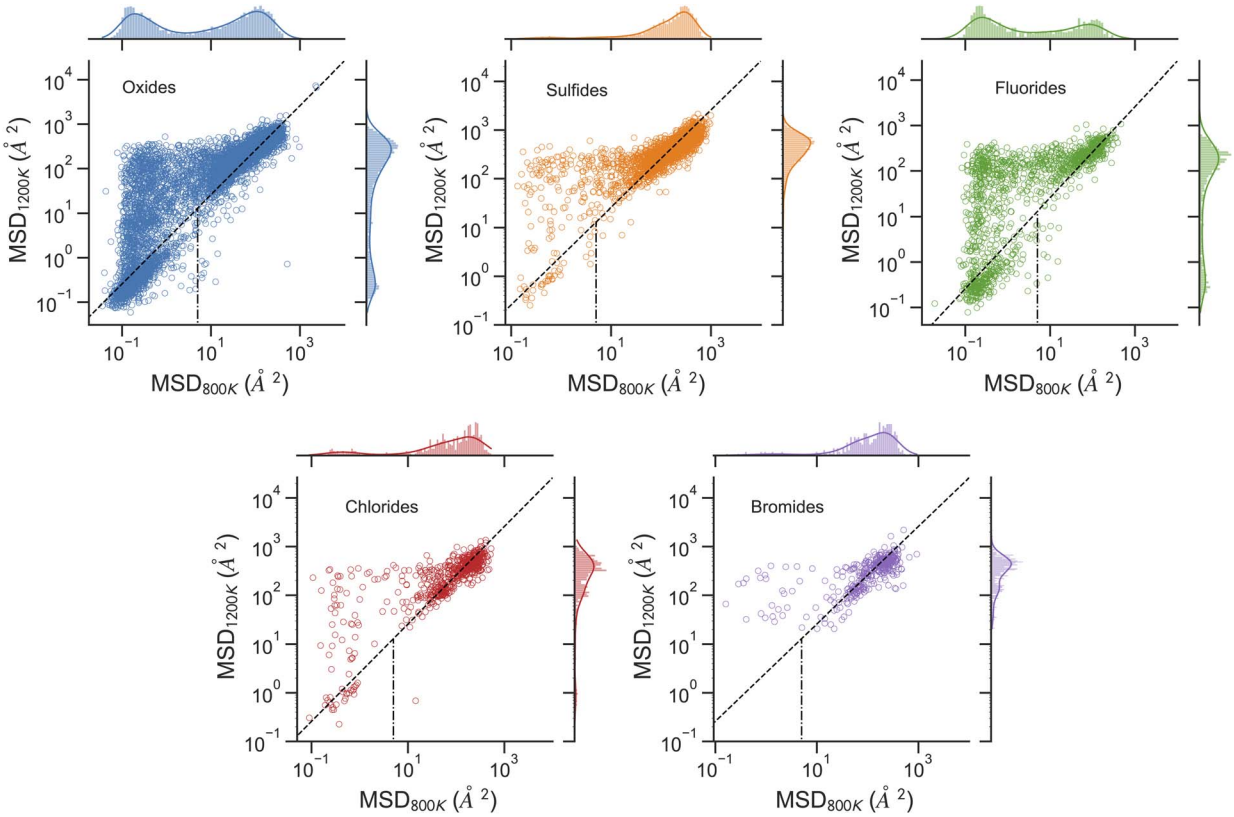

Fig. 4 Distribution of the calculated mean squared displacement (MSD) at 1200 K ($MSD_{1200\ K}$) *vs.* 800 K ($MSD_{800\ K}$) for various materials including oxides, sulfides, fluorides, chlorides and bromides. The dashed line represents the ratio $\frac{MSD_{1200\ K}}{MSD_{800\ K}} = 3.3$, while the dashed-dotted lines correspond to $MSD_{800\ K} = 5$ Å^2. Materials with $\frac{MSD_{1200\ K}}{MSD_{800\ K}} < 3.3$ and $MSD_{800\ K} > 5.0$ Å^2 are predicted to be fast Li^+ conductors.

compounds in oxides and fluorides exhibit small $MSD_{800\ K}$ less than 5 Å^2 and high $MSD_{1200\ K}/MSD_{800\ K}$ ratios, indicative of high E_a and, consequently, a sluggish Li^+ conductivity.

Only 131 compounds exhibit favorable electrochemical stability within a voltage range spanning from 1.5 to 3.2 V *vs.* Li/Li^+. Fig. 5(a) illustrates the distribution of calculated oxidation and reduction limits for candidates that passed the conductivity screening. Among all types of compounds, sulfides generally have the narrowest electrochemical stability window, with oxidation limit below 3 V and reduction limit above 1.5 V *vs.* Li/Li^+. Thus, sulfides are less likely to be thermodynamically stable when paired with common electrodes. In contrast, oxide SEs tend to have oxidation limits aligning with the working potentials (3 to 4.5 V *vs.* Li/Li^+) of typical cathodes, but their reduction limits often surpass 2 V *vs.* Li/Li^+, rendering them unstable against common anodes (Li metal and LTO anodes). Fluorides, however, demonstrate stability at exceptionally high potentials, with most fluoride SEs exhibiting oxidation limits ranging from 4.5 to 6.9 V and reduction limits between 3.0 and 5.1 V *vs.* Li/Li^+. Chlorides and bromides generally exhibit optimal electrochemical stabilities. Most chloride SEs, and about half of the bromide SEs, remain stable at potentials exceeding 3.2 V *vs.* Li/Li^+, with some even showing stability down to 0 V. This suggests their potential to enhance the utilization of Li metal, potentially leading to a significant improvement in the energy density. Overall, only 20 oxides, 5 sulfides, 7 fluorides, 51 chlorides, and 48 bromides exhibit electrochemical stability that spans the voltage range of 1.5 to 3.3 V *vs.* Li/Li^+. Furthermore, we analyzed the chemical reactivity of the electrochemical stable SEs in contact with common electrodes. As depicted in Fig. 5(b), 35 SEs remain stable in contact with Li metal anodes, exhibiting $\Delta E_{rxn} > -0.1$ eV per atom. Most SEs also exhibit minimal reactivity with LTO anodes, indicating their potential for long-term stable function in solid-state LIBs. Additionally, most SEs are compatible with typical cathodes, such as LFPO, LCO and LMO. While only 38 SEs demonstrate good stability in contact with NCM111 electrodes.

Ultimately, we identified 130 unique compounds that satisfy all the aforementioned criteria, making them as promising candidates for SEs. A structural analysis reveals that these 130 new materials are classified into 68 structure prototypes (Table S4†), which, to the best of our knowledge, have not been previously studied. To further confirm their Li^+ conductivity, we conducted MD simulations over a long duration (1 ns) at 300 K. Particularly, these simulations reveal that 28 of these materials exhibit superionic Li^+ conductivities exceeding 10 mS cm^{-1} at 300 K, as detailed in Table 1. Notably, 26 of these materials, representing 93%, are halides. Halides are an emerging class of

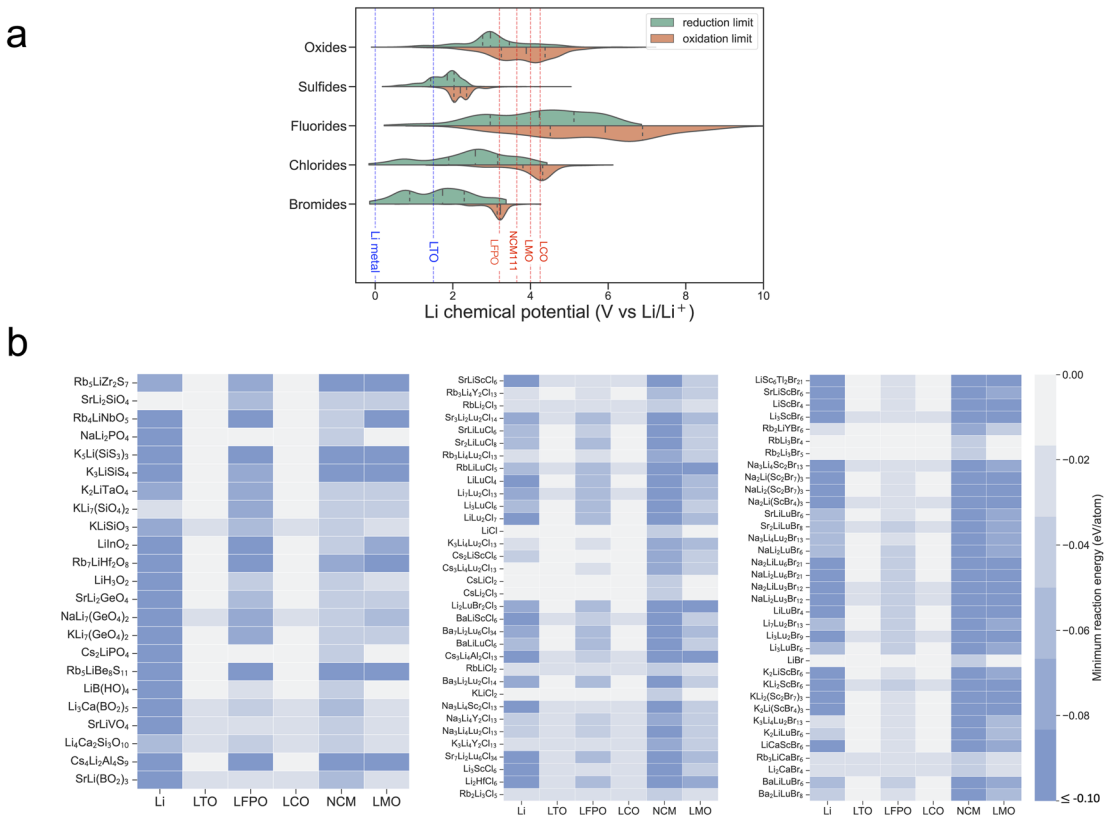

Fig. 5 (a) Distribution of the calculated oxidation and reduction limits for solid electrolytes (SEs) that meet the conductivity screening criteria. (b) Calculated minimum reaction energies between SEs and common electrode materials, such as Li metal, $Li_4Ti_5O_{12}$ (LTO), $LiFePO_4$ (LFPO), $LiCoO_2$ (LCO), $Li[Ni_xCo_yMn_z]O_2$ (NCM), $LiMnO_2$ (LMO) in experiment.

materials currently being explored as SEs in experimental studies.[11–14] These findings offer great promise for the discovery of novel SEs through experimental validation. The calculated phase stability, mechanical properties, and electrochemical properties of these 28 materials, along with the others, are presented in Tables 1 and S5,† respectively, to aid in experimental validations.

Structure–composition–conductivity relationship

Elucidating the relationship between structure, composition and conductivity is essential for designing better SEs. To unravel this relationship, we developed a ML tree-based classification model. Fig. 6(a) shows the permutation importance of the top 10 features in our model. Notably, features such as maximum packing efficiency, packing fraction, volume per atom, and mean difference in electronegativity are the most significant contributors to the predicted Li^+ conductivities. Fig. 6(b)–(d) presents the partial dependence plots, which illustrates the marginal effects of these four features on the predicted probability of materials being fast Li^+ conductors.

The maximum packing efficiency refers to the largest space that an atom occupies within its local environment in a crystal structure.[38] As shown in Fig. 6(b), the average predicted probability of Li^+ conduction decreases with increasing maximum packing efficiency. This implies that larger space for Li^+, *i.e.*, more extensive diffusing tunnels, are crucial for rapid Li^+ diffusion. While maximum packing efficiency measures the local arrangement within a structure, density features reflect the overall arrangement of atoms and void spaces. Fig. 6(c) reveals that the probability of Li^+ conduction diminishes as the packing fraction increases. The packing fraction represents the portion of space occupied by atoms in a structure; a higher packing fraction results in fewer void spaces for Li^+ hopping, thereby reducing connectivities of Li^+ diffusing tunnels. Conversely, as illustrated in Fig. 6(d), the Li^+ conduction probability increases with volume expansion. The rise in Li^+ conductivity stabilizes when the volume per atom reaches ~17 $Å^3$ per atom. Fig. 6(e) shows that Li^+ conductivity decreases with increasing electronegativity difference, likely because large electronegativity difference correspond to stronger interactions between Li^+ and neighboring polyhedral anions. Consequently, structures with large electronegativity difference tend to restrict Li^+ motion, reducing the likehood of fast Li^+ conduction. In summary, three primary factors influence Li^+ diffusion in SEs: (1) the size of local Li^+ diffusion channels; (2) the connectivity of these channels; (3) the electrostatic interactions between Li^+ and the polyhedral anions.

Discussion

Compared to conventional HT DFT computations, universal ML-IAPs provide reasonably accurate results with significantly

Table 1 Calculated properties of 28 identified promising solid electrolytes with Li^+ conductivities higher than 10 mS cm^{-1} at 300 K. These properties include the composition of candidate material, phase stability ($E_{hull\text{-}m}$), shear modulus (G), Pugh's ratio (G/K), hybrid functional band gap (E_g), electrochemical stability window, Li^+ conductivity at 300 K and material's ID in the Matterverse database. The data is sorted by the calculated $E_{hull\text{-}m}$

Material	$E_{hull\text{-}m}$ (meV per atom)	G (GPa)	G/K	E_g (eV)	Electrochemical window (V)	$\sigma_{300\ K}$ (mS cm^{-1})	Material ID
$K_5Li(SiS_3)_3$	−174	8.60	0.38	2.90	1.26–5.01	10.9	mv-15353762
$K_2Li(ScBr_4)_3$	2	8.54	0.41	3.06	0.92–3.23	17.5	mv-6707348
$LiLuBr_4$	4	9.18	0.46	3.82	0.72–3.23	26.0	mv-31587649
$RbLi_3Br_4$	5	9.56	0.51	4.33	0.01–3.23	133.5	mv-22499633
Cs_2LiPO_4	8	13.87	0.44	4.02	0.76–3.96	20.4	mv-24746431
$NaLi_2Lu_6Br_{21}$	8	9.21	0.56	3.65	0.72–3.23	89.3	mv-7301111
$Ba_7Li_2Lu_6Cl_{34}$	9	9.58	0.39	4.81	0.74–4.26	37.9	mv-16242402
$CsLiCl_2$	10	8.50	0.46	4.93	0.01–4.29	12.7	mv-31595624
$Ba_3Li_2Lu_2Cl_{14}$	13	9.04	0.40	4.83	0.74–4.26	54.9	mv-3243154
$NaLi_2LuBr_6$	14	8.92	0.54	4.07	0.72–3.23	248.4	mv-28300644
$Li_7Lu_2Br_{13}$	14	10.86	0.54	3.77	0.72–3.23	55.1	mv-15824306
$LiScBr_4$	15	8.68	0.38	3.06	0.92–3.23	91.4	mv-31583403
Li_2HfClF_5	15	9.12	0.50	6.08	1.18–4.26	31.4	mv-22716067
$Rb_2Li_3Br_5$	16	8.41	0.49	4.19	0.01–3.23	485.0	mv-8407962
$KLiCl_2$	16	10.98	0.51	4.88	0.01–4.26	237.7	mv-31595564
$SrLiScBr_6$	17	10.27	0.46	3.11	0.92–3.23	45.9	mv-8065059
$RbLiCl_2$	17	10.41	0.55	4.87	0.01–4.26	152.0	mv-31397981
$Li_7Lu_2Cl_{13}$	18	12.59	0.47	4.85	0.74–4.26	14.6	mv-15824305
$K_3Li_4Y_2Cl_{13}$	19	9.61	0.44	4.65	0.64–4.26	12.70	mv-15827581
$Rb_2Li_3Br_5$	19	9.46	0.54	4.22	0.01–3.23	101.2	mv-7977990
$Sr_7Li_2Lu_6Cl_{34}$	23	10.37	0.42	4.51	0.74–4.26	16.7	mv-16242397
Li_2CaBr_4	24	9.02	0.45	4.27	0.07–3.23	418.8	mv-20014595
$Rb_2Li_3Cl_5$	25	10.91	0.52	4.69	0.01–4.26	159.4	mv-8407950
$Rb_3LiCaBr_6$	26	9.18	0.59	4.25	0.01–3.23	18.2	mv-16033288
$Sr_3Li_2Lu_2Cl_{14}$	28	9.85	0.41	4.53	0.74–4.26	50.4	mv-3242974
$Na_2Li(ScBr_4)_3$	29	9.94	0.56	3.15	0.92–3.23	93.3	mv-6704684
KLi_2ScBr_6	29	9.90	0.46	3.11	0.92–3.23	337.7	mv-28459714
$Li_3Lu_2Br_9$	30	8.77	0.48	3.96	0.72–3.23	28.7	mv-15922265

lower computational costs, enabling assessments of material properties that were previously unattainable. However, there is still room for improvement. Firstly, our benchmark results indicate that the universal ML-IAP (*i.e.*, M3GNET) tends to systematically underestimate the activation energy of Li^+ by about 150 meV compared to DFT. Consequently, this leads to an overestimation of predicted Li^+ conductivity by two orders of magnitude. The discrepancy arises likely because the ML-IAP was trained predominantly on structures near their ground states,[26] while Li^+ diffusion involves transitioning through high-energy metastable states not sufficiently represented in the training dataset, leading to inaccuracies in predicting Li^+ conductivities. Secondly, the current framework of M3GNet, based on message-passing deep graph neural networks, faces challenges in parallelizing simulations. This limitation may restrict its application to large and complex material systems, such as heterogeneous interfaces. While preparing the manuscript, we noticed that a preprint paper[39] reported the discovery of new Li-ion SEs using M3GNet. Our work distinguishes itself by assessing the uncertainty of the M3GNet model in predicting Li^+ conductivity and by developing a multi-property screening workflow with well-benchmarked criteria. These efforts are conducive to accurately identifying novel SEs with potential for high performance. Additionally, none of the promising materials we identified have been reported previously.

A major hurdle in SE design is establishing and maintaining stable interfaces. Our findings suggest that ~97% of predicted superionic conductors are not directly viable as SE candidates, mainly due to their limited electrochemical stability windows, which could lead to detrimental oxidation and reduction reactions at the interface, compromising LIB reliability. To address this, replacing sulfur elements with halides could broaden the electrochemical stability window, especially at cathodes. Experimentally, the formation of beneficial passivation layers at interfaces, as seen with LiPON decomposing into Li_2O, Li_3N, and Li_3PO_4 at the Li metal/LiPON interface, offers another solution.[40,41] These decomposition products are ionically conductive but electronically insulating, forming beneficial interphases that enhance battery cycling stability. Therefore, eliminating elements that decompose into electronically conductive compounds could promote the formation of durable passivation layers. Additionally, using different SEs that are stable at the anode and cathode interfaces respectively could be a practical approach to overcome the stability issues faced by SEs in LIBs.[15]

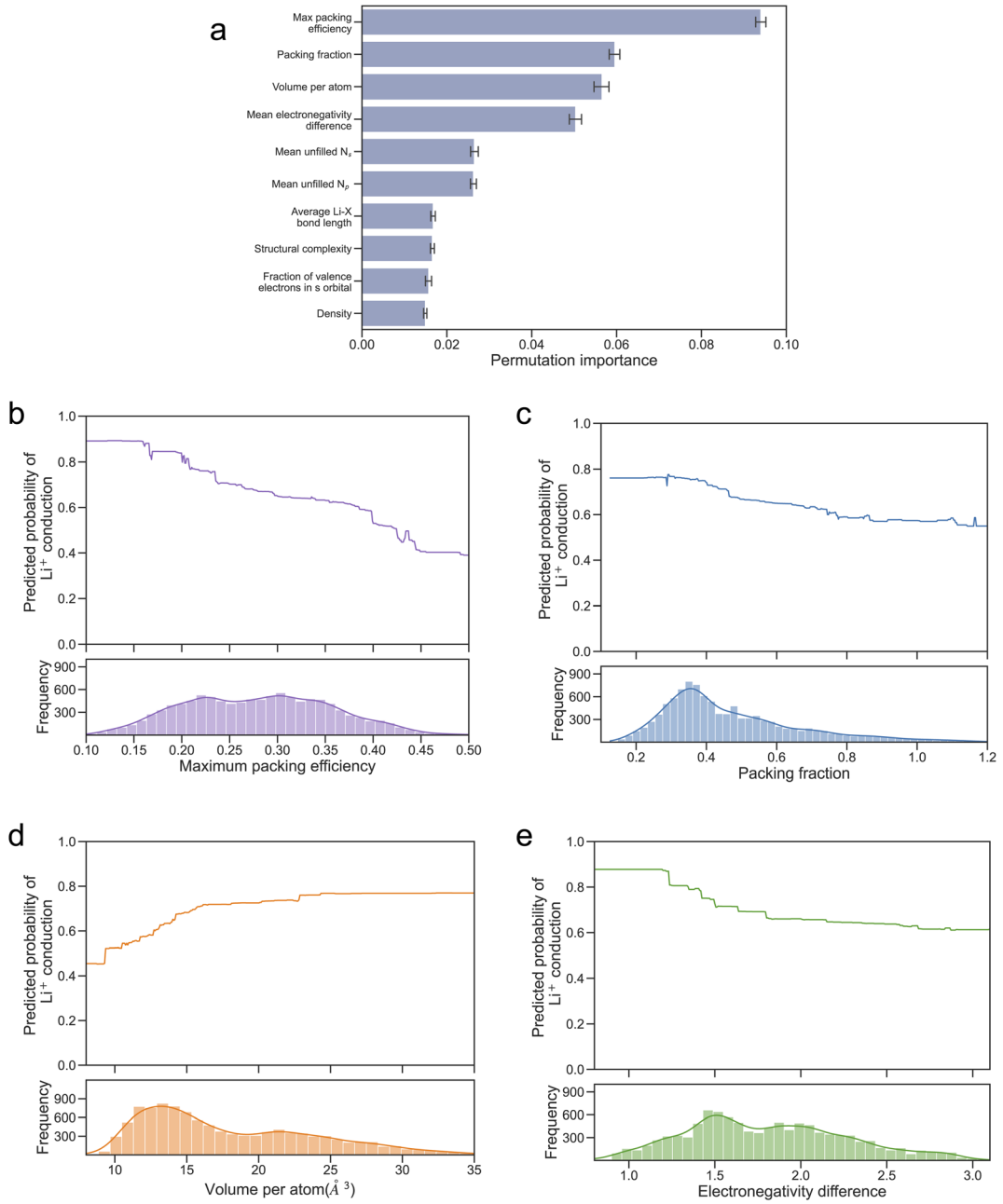

Fig. 6 (a) Permutation feature importance of developed machine learning model for identifying fast and slow Li^+ conductors. Top ten features scoring the highest permutation importance were presented. Partial dependence plots illustrating the influence of different features, including (b) maximum packing efficiency, (c) packing fraction, (d) volume per atom, and (e) electronegativity difference, on the classification of Li^+ diffusion. The accompanying histogram displays the frequency distribution of the datasets according to the feature values.

Conclusion

To summarize, we have performed machine-learning aided high-throughput calculations on over 740 000 lithium-based compounds to identify promising SEs for solid-state LIBs. By benchmarking the properties calculated by universal ML-IAP against experimental and DFT data, we established a tiered-properties HT workflow. This workflow efficiently and effectively pinpointed 130 novel SE materials, which are projected to possess properties on par with or surpassing those of the current state-of-the-art SEs. Furthermore, we developed a tree-based ML model to decipher the behavior of Li^+ within SEs. This model has successfully identified the key factors influencing Li^+ diffusion in SEs. In addition, we proposed potential strategies to address the challenges in developing high-performance solid-state LIBs.

Methods

The structure optimizations, energy calculations and molecular dynamics simulations were carried out using the universal ML-

IAP as implemented in M3GNet.[26] A pre-trained model based on the Materials Project database[42] of 2021.2.8 version was used. The initial crystal structures were retrieved from the Matterverse database website.[43] All crystal structure manipulations, file generation and data analysis were performed using the Python Materials Genomics (pymatgen) package.[44]

Phase stability

The phase stability of a compound was estimated by evaluating the distance of its energy relative to the linear combination of energies of stable phases in the 0 K phase diagram. This distance, denoted as the energy above hull (E_{hull}), provides an indication of materials' phase stability. The lower E_{hull}, the higher the phase stability.[45] The 0 K phase diagram is constructed by considering the energetics of all relevant materials from the Materials Project version 2021.2.8 (ref. 42) with the Materials Project 2020 energy correction scheme. The energy above hull calculated by this scheme is denoted as $E_{\text{hull-m}}$, which is in line with the definition in Chen and Ong,[26] a material is deemed more stable than all known materials in the Materials Project if its calculated $E_{\text{hull-m}}$ has a non-positive value.

Electronic property

The hybrid functional Heyd–Scuseria–Ernzerhof (HSE) band gap of a material was predicted using the multi-fidelity graph network models.[46] The multi-fidelity models were trained on datasets comprising multi-fidelity experimental and computed band gaps for both ordered and disordered materials. Using four-fidelity models, the average MAE of the predicted HSE band gaps is ~0.31 eV. The four-fidelity model was used to compute the HSE band gap in this work.

Li^+ conductivity

The Li^+ conductivity at a given temperature, T_0 can be estimated *via* the Nernst–Einstein relation:

$$\sigma_{T_0} = (\rho z^2 F^2/RT) \times D_{T_0} \quad (1)$$

where ρ is the molar density, F is Faraday constant, z is the charge of the mobile ions, R is the gas constant and T is temperature. Assuming that the Li^+ diffusivity in solid electrolytes follows an Arrhenius relationship, the diffusivity (D_{T_0}) at temperature T_0 is given by,

$$D_{T_0} = D_0\, e^{\frac{-E_a}{kT_0}} \quad (2)$$

where E_a is the activation energy, k is Boltzmann's constant and D_0 is the diffusivity at $T \rightarrow \infty$.

From molecular dynamic (MD) simulations, the self-diffusivity (D_{T_0}) of Li^+ can be obtained by:

$$D_{T_0} = \frac{\langle \Delta r(t) \rangle^2}{2dt} = \frac{\text{MSD}}{2dt} \quad (3)$$

where d is the dimensionality factor that equals three for bulk crystal structure, and $\langle \Delta r(t) \rangle^2$ is the average Li^+ mean square displacement over a time duration t at a given temperature T_0.

The MD simulations based on universal ML-IAP were performed in the NVT ensemble with Berendsen thermostat[47] using ASE package.[48] A time step of 1 fs was adopted. The *ab initio* molecular dynamic (AIMD) simulations were performed in the NVT ensemble with Nose–Hoover thermostat. Perdew–Burke–Ernzerhof functional[49] was used for all AIMD simulations. A plane-wave energy cutoff of 400 eV, a γ-centered $1 \times 1 \times 1$ k-point mesh, and a time step of 2 fs was used. For both methods, the simulation cell parameters were fixed at their fully relaxed values at ground state. The model was then heated to the target temperatures with a gradient of 0.25 K fs^{-1} and equilibrated for at least 20 ps. A supercell size with at least 9 Å along each lattice direction was generated for MD simulations.

Electrochemical stability

The electrochemical stability was assessed using the grand potential phase diagram.[40] In this approach, Li is considered as the primary mobile species and the solid electrolyte/electrode interface is assumed to be an open system with respect to Li. The thermodynamic potential ϕ thus can be constructed by a Legendre transformation of the Gibbs free energy with respect to the Li chemical potential:

$$\phi = G - \frac{\partial G}{\partial N_{\text{Li}}} N_{\text{Li}} \approx E_{\text{ML-IAP}} - \mu_{\text{Li}} N_{\text{Li}} \quad (4)$$

where G is the Gibbs free energy, N_{Li} is the number of lithium atoms in the open system. The Gibbs free energy was approximated by ML-IAP calculated total energy ($E_{\text{ML-IAP}}$) due to the negligible entropy effect of solids. The partial derivative of G with respect to N_{Li} is the chemical potential of Li. The phase equilibrium at the anode and charged cathode can be approximated as the SE's composition at the high $\mu_{\text{Li}} = \mu_{\text{Li}^0}$ and low $\mu_{\text{Li}} = \mu_{\text{Li}^0} - 10$ eV (μ_{Li^0} is the chemical potential of Li metal), respectively.

Electrolytes/electrodes interface

To assess the reactivity between solid electrolytes and electrodes, the point at which the driving force of the reaction reaches its maximum was identified by determining the fraction of SE and electrode materials. The thermodynamic reactivity is calculated based on the phase diagram and energetics of the combined chemical space of solid electrolytes and electrodes. By minimizing the reaction energy ($\Delta E[c_{\text{SE}}, c_{\text{E}}]$)

$$\Delta E[c_{\text{SE}}, c_{\text{E}}] = E_{\text{pd}}[xc_{\text{SE}} + (1-x)c_{\text{E}}] - xE[c_{\text{SE}}] - (1-x)E[c_{\text{E}}] \quad (5)$$

over x, where c_{SE} and c_{E} are the composition of solid electrolyte and electrode, respectively. E_{pd} is the lowest energy combination of the reaction products at composition $xc_{\text{SE}} + (1-x)c_{\text{E}}$.

Machine learning modeling

The eXtreme Gradient Boosting (XGBoost) algorithm was used to create the classification models as implemented in XGBoost and the scikit-learn package.[50,51] The dataset used to train the machine learning model for predicting diffusion properties consists of 10 769 unique materials. 87 features were created for

each material using matminer package.[52] These features covers the descriptions of composition properties, crystal structure and crystal sites. To avoid redundancy and overfitting, we calculated the Pearson correlation coefficient ($\rho = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$, X and Y represents values of the two descriptors that are being compared) between each pairs of the descriptors, as shown in Fig. S2.† For pairs of descriptors that are highly correlated ($|\rho| > 0.75$), the descriptor results in better performance of the machine learning model is selected. As a result, 28 independent descriptors were selected to develop the model.

Five-fold cross-validation was used for model fitting and hyper-parameter optimization. During the optimization process, we performed a grid search to identify optimal values for parameters directly related to the model performance. These parameters include number of trees in the ensemble ($n_{\mathrm{estimator}}$), regularization parameter γ, the maximum depth of each tree (d) and the learning rate (η). For all other parameters, the default values were used. The top-1 accuracy and Jaccard index were used to measure the performances of classifiers and the optimized model has a CV score of 0.88. The final model was trained with $n_{\mathrm{estimator}} = 1800$, $\gamma = 0.2$, $\eta = 0.01$ and $d = 6$, and yield an accuracy score of 0.975 and a Jaccard score of 0.962.

Data availability

Further information and data requests should be directed to Xingyu Guo.

Author contributions

Conceptualization, X. G.; methodology, X. G.; investigation, X. G. and Z. W.; writing – original draft, X. G.; writing – review and editing, Z. W., J.-H. Y. and X.-G. G.; supervision, X.-G. G.; funding acquisition, X.-G. G.

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

This work was partially supported by the China National Key R&D Program (2022YFA1404603), the National Natural Science Foundation of China (Grants No. 11974078, No. 12188101, and No. 11991061) and the Guangdong Major Project of the Basic and Applied Basic Research (Future functional materials under extreme conditions–2021B0301030005).

References

1 J. Janek and W. G. Zeier, Challenges in speeding up solid-state battery development, *Nat. Energy*, 2023, 1–11.
2 N. Kamaya, K. Homma, Y. Yamakawa, M. Hirayama, R. Kanno, M. Yonemura, T. Kamiyama, Y. Kato, S. Hama and K. Kawamoto, A lithium superionic conductor, *Nat. Mater.*, 2011, **10**, 682–686.
3 S. P. Ong, Y. Mo, W. D. Richards, L. Miara, H. S. Lee and G. Ceder, Phase stability, electrochemical stability and ionic conductivity of the Li 10±1 MP 2 X 12 (M= Ge, Si, Sn, Al or P, and X= O, S or Se) family of superionic conductors, *Energy Environ. Sci.*, 2013, **6**, 148–156.
4 P. Bron, S. Johansson, K. Zick, J. Schmedt auf der Günne, S. Dehnen and B. Roling, Li10SnP2S12: an affordable lithium superionic conductor, *J. Am. Chem. Soc.*, 2013, **135**, 15694–15697.
5 A. Kuhn, O. Gerbig, C. Zhu, F. Falkenberg, J. Maier and B. V. Lotsch, A new ultrafast superionic Li-conductor: ion dynamics in Li11Si2PS12 and comparison with other tetragonal LGPS-type electrolytes, *Phys. Chem. Chem. Phys.*, 2014, **16**, 14669–14674.
6 J. M. Whiteley, J. H. Woo, E. Hu, K.-W. Nam and S.-H. Lee, Empowering the lithium metal battery through a silicon-based superionic conductor, *J. Electrochem. Soc.*, 2014, **161**, A1812.
7 S. Boulineau, M. Courty, J.-M. Tarascon and V. Viallet, Mechanochemical synthesis of Li-argyrodite Li6PS5X (X= Cl, Br, I) as sulfur-based solid electrolytes for all solid state batteries application, *Solid State Ionics*, 2012, **221**, 1–5.
8 P. R. Rayavarapu, N. Sharma, V. K. Peterson and S. Adams, Variation in structure and Li+-ion migration in argyrodite-type Li 6 PS 5 X (X= Cl, Br, I) solid electrolytes, *J. Solid State Electrochem.*, 2012, **16**, 1807–1813.
9 R. Murugan, V. Thangadurai and W. Weppner, Fast lithium ion conduction in garnet-type Li7La3Zr2O12, *Angew. Chem., Int. Ed.*, 2007, **46**, 7778–7781.
10 A. Belous, G. Novitskaya, S. Polyanetskaya and Y. I. Gornikov, Investigation into complex oxides of La 2/3-x Li 3x TiO 3 composition, *Izv. Akad. Nauk SSSR, Neorg. Mater.*, 1987, **23**, 470–472.
11 C. Wang, J. Liang, J. T. Kim and X. Sun, Prospects of halide-based all-solid-state batteries: From material design to practical application, *Sci. Adv.*, 2022, **8**, eadc9516.
12 T. Asano, A. Sakai, S. Ouchi, M. Sakaida, A. Miyazaki and S. Hasegawa, Solid halide electrolytes with high lithium-ion conductivity for application in 4 V class bulk-type all-solid-state batteries, *Adv. Mater.*, 2018, **30**, 1803075.
13 X. Li, J. Liang, J. Luo, M. N. Banis, C. Wang, W. Li, S. Deng, C. Yu, F. Zhao, Y. Hu, *et al.*, Air-stable Li 3 InCl 6 electrolyte with high voltage compatibility for all-solid-state batteries, *Energy Environ. Sci.*, 2019, **12**, 2665–2671.
14 L. M. Riegger, R. Schlem, J. Sann, W. G. Zeier and J. Janek, Lithium-Metal Anode Instability of the Superionic Halide Solid Electrolytes and the Implications for Solid-State Batteries, *Angew. Chem., Int. Ed.*, 2021, **60**, 6718–6723.
15 W. D. Richards, L. J. Miara, Y. Wang, J. C. Kim and G. Ceder, Interface stability in solid-state batteries, *Chem. Mater.*, 2016, **28**, 266–273.
16 P. Gorai, T. Famprikis, B. Singh, V. Stevanovic and P. Canepa, Devil is in the defects: electronic conductivity in solid electrolytes, *Chem. Mater.*, 2021, **33**, 7484–7498.
17 G. Hautier, A. Jain, H. Chen, C. Moore, S. P. Ong and G. Ceder, Novel mixed polyanions lithium-ion battery

cathode materials predicted by high-throughput ab initio computations, *J. Mater. Chem.*, 2011, **21**, 17147–17153.
18 H. Chen, G. Hautier, A. Jain, C. Moore, B. Kang, R. Doe, L. Wu, Y. Zhu, Y. Tang and G. Ceder, Carbonophosphates: a new family of cathode materials for Li-ion batteries identified computationally, *Chem. Mater.*, 2012, **24**, 2009–2016.
19 Z. Zhu, I.-H. Chu and S. P. Ong, Li3Y (PS4) 2 and Li5PS4Cl2: new lithium superionic conductors predicted from silver thiophosphates using efficiently tiered ab initio molecular dynamics simulations, *Chem. Mater.*, 2017, **29**, 2474–2484.
20 E. D. Cubuk, A. D. Sendek and E. J. Reed, Screening billions of candidates for solid lithium-ion conductors: A transfer learning approach for small data, *J. Chem. Phys.*, 2019, **150**, 214701.
21 A. D. Sendek, E. D. Cubuk, E. R. Antoniuk, G. Cheon, Y. Cui and E. J. Reed, Machine learning-assisted discovery of solid Li-ion conducting materials, *Chem. Mater.*, 2018, **31**, 342–352.
22 A. D. Sendek, Q. Yang, E. D. Cubuk, K.-A. N. Duerloo, Y. Cui and E. J. Reed, Holistic computational structure screening of more than 12000 candidates for solid lithium-ion conductor materials, *Energy Environ. Sci.*, 2017, **10**, 306–320.
23 S. J. Honrao, X. Yang, B. Radhakrishnan, S. Kuwata, H. Komatsu, A. Ohma, M. Sierhuis and J. W. Lawson, Discovery of novel Li SSE and anode coatings using interpretable machine learning and high-throughput multi-property screening, *Sci. Rep.*, 2021, **11**, 16484.
24 C. Lv, X. Zhou, L. Zhong, C. Yan, M. Srinivasan, Z. W. Seh, C. Liu, H. Pan, S. Li, Y. Wen, *et al.*, Machine learning: an advanced platform for materials development and state prediction in lithium-ion batteries, *Adv. Mater.*, 2022, **34**, 2101474.
25 Y. Liu, B. Guo, X. Zou, Y. Li and S. Shi, Machine learning assisted materials design and discovery for rechargeable batteries, *Energy Storage Mater.*, 2020, **31**, 434–450.
26 C. Chen and S. P. Ong, A universal graph deep learning interatomic potential for the periodic table, *Nat. Comput. Sci.*, 2022, **2**, 718–728.
27 Z. Deng, B. Radhakrishnan and S. P. Ong, Rational composition optimization of the lithium-rich Li3OCl1–x Br x anti-perovskite superionic conductors, *Chem. Mater.*, 2015, **27**, 3749–3755.
28 Z. Wang, J. Ha, Y. H. Kim, W. B. Im, J. McKittrick and S. P. Ong, Mining unexplored chemistries for phosphors for high-color-quality white-light-emitting diodes, *Joule*, 2018, **2**, 914–926.
29 G. Hautier, A. Jain, S. P. Ong, B. Kang, C. Moore, R. Doe and G. Ceder, Phosphates as lithium-ion battery cathodes: an evaluation based on high-throughput ab initio calculations, *Chem. Mater.*, 2011, **23**, 3495–3508.
30 C. Monroe and J. Newman, The impact of elastic deformation on deposition kinetics at lithium/polymer interfaces, *J. Electrochem. Soc.*, 2005, **152**, A396.
31 Z. Deng, Z. Wang, I.-H. Chu, J. Luo and S. P. Ong, Elastic properties of alkali superionic conductor electrolytes from first principles calculations, *J. Electrochem. Soc.*, 2015, **163**, A67.
32 J. Janek and W. G. Zeier, A solid future for battery development, *Nat. Energy*, 2016, **1**, 1–4.
33 T. Famprikis, P. Canepa, J. A. Dawson, M. S. Islam and C. Masquelier, Fundamentals of inorganic solid-state electrolytes for batteries, *Nat. Mater.*, 2019, **18**, 1278–1291.
34 Y. Xiao, L. J. Miara, Y. Wang and G. Ceder, Computational screening of cathode coatings for solid-state batteries, *Joule*, 2019, **3**, 1252–1275.
35 V. I. Anisimov, J. Zaanen and O. K. Andersen, Band theory and Mott insulators: Hubbard U instead of Stoner I, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1991, **44**, 943.
36 P. R. Kent and G. Kotliar, Toward a predictive theory of correlated materials, *Science*, 2018, **361**, 348–354.
37 S. Kirklin, B. Meredig and C. Wolverton, High-throughput computational screening of new Li-ion battery anode materials, *Adv. Energy Mater.*, 2013, **3**, 252–262.
38 L. Ward, R. Liu, A. Krishna, V. I. Hegde, A. Agrawal, A. Choudhary and C. Wolverton, Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations, *Phys. Rev. B*, 2017, **96**, 024104.
39 C. Chen, D. T. Nguyen, S. J. Lee, N. A. Baker, A. S. Karakoti, L. Lauw, C. Owen, K. T. Mueller, B. A. Bilodeau, V. Murugesan, *et al.*, Accelerating computational materials discovery with artificial intelligence and cloud high-performance computing: from large-scale screening to experimental validation, *arXiv*, 2024, preprint, arXiv:2401.04070, DOI: **10.48550/arXiv.2401.04070**.
40 Y. Zhu, X. He and Y. Mo, Origin of outstanding stability in the lithium solid electrolyte materials: insights from thermodynamic analyses based on first-principles calculations, *ACS Appl. Mater. Interfaces*, 2015, **7**, 23685–23693.
41 D. Cheng, T. A. Wynn, X. Wang, S. Wang, M. Zhang, R. Shimizu, S. Bai, H. Nguyen, C. Fang, M.-c. Kim, *et al.*, Unveiling the stable nature of the solid electrolyte interphase between lithium metal and LiPON via cryogenic electron microscopy, *Joule*, 2020, **4**, 2484–2500.
42 A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, *et al.*, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater.*, 2013, **1**, 011002.
43 **https://matterverse.ai/**.
44 S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson and G. Ceder, Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis, *Comput. Mater. Sci.*, 2013, **68**, 314–319.
45 S. P. Ong, L. Wang, B. Kang and G. Ceder, Li- Fe- P- O2 phase diagram from first principles calculations, *Chem. Mater.*, 2008, **20**, 1798–1807.
46 C. Chen, Y. Zuo, W. Ye, X. Li and S. P. Ong, Learning properties of ordered and disordered materials from multi-fidelity data, *Nat. Comput. Sci.*, 2021, **1**, 46–53.
47 H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

48 A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Du lak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, *et al.*, The atomic simulation environment—a Python library for working with atoms, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.

49 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

50 T. Chen and C. Guestrin, Xgboost: a scalable tree boosting system, *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp 785–794.

51 F. Pedregosa, *et al.*, Scikit-learn: Machine Learning in Python, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.

52 L. Ward, A. Dunn, A. Faghaninia, N. E. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, *et al.*, Matminer: An open source toolkit for materials data mining, *Comput. Mater. Sci.*, 2018, **152**, 60–69.