Taylor & Francis
Taylor & Francis Group

Science and Technology of Advanced Materials: Methods

ISSN: 2766-0400 (Online) Journal homepage: www.tandfonline.com/journals/tstm20

Periodic table-based compositional descriptors for accelerating electrochemical material discovery: Li-ion conductors and oxygen evolution electrocatalysts

Yen-Ju Wu, Yibin Xu, Lei Fang, Wenqin Peng, Ken Sakaushi, Meiqi Zhang, Masao Arai & Yukinori Koyama

To cite this article: Yen-Ju Wu, Yibin Xu, Lei Fang, Wenqin Peng, Ken Sakaushi, Meiqi Zhang, Masao Arai & Yukinori Koyama (30 May 2025): Periodic table-based compositional descriptors for accelerating electrochemical material discovery: Li-ion conductors and oxygen evolution electrocatalysts, Science and Technology of Advanced Materials: Methods, DOI: 10.1080/27660400.2025.2513218

To link to this article: https://doi.org/10.1080/27660400.2025.2513218

Publisher: Taylor & Francis & The Author(s). Published by National Institute for Materials Science in partnership with Taylor & Francis Group
Journal: *Science and Technology of Advanced Materials: Methods*
DOI: 10.1080/27660400.2025.2513218

Periodic table-based compositional descriptors for accelerating electrochemical material discovery: Li-ion conductors and oxygen evolution electrocatalysts

Yen-Ju Wu[1], Yibin Xu[1,*], Lei Fang[1], Wenqin Peng[2], Ken Sakaushi[2], Meiqi Zhang[1], Masao Arai[1], Yukinori Koyama[1]

[1] Center for Basic Research on Materials, National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

[2] Research Center for Energy and Environmental Materials, National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

*Corresponding author: Yibin Xu, XU.Yibin@nims.go.jp

Keywords: solid electrolyte, ionic conductor, electrocatalyst, energy storage, descriptor

The discovery of high-performance electrochemical materials is essential for sustainable energy technologies, yet conventional methods rely on trial-and-error experiments, time-consuming computations, and detailed structural data. To address these challenges, we introduce a periodic table-based compositional descriptor that requires only chemical formulas, enabling efficient material discovery with reversible design. We applied this approach to two key applications: fast Li-ion conductors for solid-state electrolytes and platinum-group metal (PGM)-free oxygen evolution reaction (OER) electrocatalysts. Our model identified both known and new Li-ion conductors,

including anti-fluorite structures with high ionic conductivity at 600–700 K—significantly lower than traditional compounds like Li_2S and Li_2Se. For OER electrocatalysts, we predicted $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide, which exhibited experimentally validated performance comparable to RuO_2 but at a lower overpotential. The periodic descriptor offers a scalable and efficient framework for accelerating the discovery of green energy materials, including next-generation batteries and hydrogen production, contributing to carbon neutrality.

1. Introduction

The urgent need for sustainable energy solutions has sparked a global pursuit of advanced electrochemical materials that can power next-generation technologies, such as rechargeable batteries and green hydrogen production.[1-5] Ionic conductors and electrocatalysts are fundamental to these technologies, determining the efficiency and scalability of key applications. Despite extensive efforts, the discovery of high-performance materials has largely depended on traditional methods, such as first-principle computational simulations and trial-and-error experimental processes, which are often time-consuming and hinder rapid material innovation.[6-9]

Recent advancements in data-driven approaches for materials discovery have aimed to overcome these limitations by predicting promising materials through various metrics, such as composition and structural properties. For instance, Sendek et al. combined compositional, structural, and electronic data for computational screening,[10,11] while Muy et al. used lattice dynamics descriptors to identify ionic conductors.[12] Other approaches, like those of Zhang et al. used unsupervised learning to discover ionic conductors by analyzing crystal structures, anion frameworks, and

modified XRD patterns,[13] while Kajita et al. predicted ionic conductors by exploring ensemble structures derived from atomic configurations.[14] Similarly, various data-driven techniques have been applied to identify high-performance electrocatalysts.[15-18] However, a significant bottleneck persists: the scarcity of sufficient structural data for new materials, which restricts the training data available for machine learning models in materials science.

Tools like XenonPy[19] and Matminer[20] generate compositional descriptors from chemical formulas, sidestepping the need for detailed structural information. They capture the chemical properties of materials by taking weighted averages, variances, and other statistical measures of characteristics, such as atomic weight, atomic radius, and electronegativity for each element. However, their complexity introduces challenges, particularly in achieving reversible design—a process that enables deducing a material's chemical formula directly from its descriptors. Additionally, when descriptor dimensions are high, the search space for exploring the entire chemical composition grows exponentially, making optimization beyond 10-20 dimensions difficult. This limitation underscores the need for a more well-designed approach for descriptor generation, one that balances low dimensionality with predictive accuracy.

In response to these challenges, we propose an advanced approach that uses periodic descriptors derived directly from the periodic table. The periodic descriptors require only the chemical formula of materials, bypassing the need for detailed structural information and enabling exploration across a broader spectrum of unknown materials. Furthermore, our approach facilitates reversible design, allowing for a direct translation between descriptors and chemical compositions.

Our analysis emphasizes the predictive performance, simplicity, and practical

applicability of the periodic descriptors in material science research. This paper outlines our methodology, starting with the introduction of periodic descriptors, followed by the evaluation of their predictive performance using five distinct machine learning models. We present a comparative analysis to illustrate the superiority of our approach over established benchmarks.

Furthermore, we demonstrate two model studies for discovering fast Li-ion conductors and oxygen evolution electrocatalysts using our proposed periodic descriptors, as illustrated in **Figure 1**. In the case of ionic conductors, our predictive framework identified several candidates, including anti-fluorite structures that exhibit superionic conductivity behavior around 600-700 K, as confirmed by simulation. For electrocatalysts, our model identified $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide as a high-performance, PGM-free oxygen evolution electrocatalyst, demonstrating electrochemical properties comparable to the PGM benchmark RuO_2. The **Model training** section describes the consistent model training method used for comparative analysis and the two model studies. By summarizing these two successful model studies, we highlight the versatility and effectiveness of periodic descriptors in accelerating the discovery and optimization of materials for electrochemical materials and beyond.

2. Periodic descriptor design

Periodic descriptors are derived from chemical formulas using a two-step process involving group coefficients and group properties. Each of the 17 groups in the periodic table, excluding Group 18 (which does not form compounds), is represented by these two components, resulting in 34 periodic descriptors, as shown in **Figure 2**.

(i) Group Coefficient (*Cg*)

Let the chemical formula consist of n elements, each with a stoichiometric coefficient a_i belonging to group g. The group coefficient for group g is defined as: $C_g = \frac{\sum_{i \in g} a_i}{\sum_{i=1}^{n} a_i}$ (Eq. 1)

This represents the fraction of atoms in the formula that belong to group g. For example, in $Li_3PO_{3.56}N_{0.44}$, the total number of atoms is 8. The group coefficient for Group 1 (Li) is 3/8 = 0.375, and for Group 15 (P and N), it is (1 + 0.44)/8 = 0.18.

(ii) Group Property ($P_g^{(p)}$)

Let p_i be the atomic property (e.g., atomic number, atomic radius) for element i. The group property $P_g^{(p)}$ is calculated as:

$$P_g^{(p)} = \frac{\sum_{i \in g} a_i \cdot p_i}{\sum_{i \in g} a_i} \quad \text{(Eq. 2)}$$

This represents the weighted average of atomic property p for elements in group g, where weights are based on their stoichiometric coefficients. For instance, for Group 15 in $Li_3PO_{3.56}N_{0.44}$, the atomic number-based group property is (1 × 15 + 0.44 × 7)/1.44 = 12.56.

In cases where the sum of coefficients in a periodic table group is zero, the average value of properties for that group becomes undefined. For machine learning prediction models, it is necessary to express these values numerically, so we fill these undefined values with zero in our models.

Our periodic descriptors have low dimensionality, and the independence of each group makes them straightforward to use. The reverse transformation from descriptors to chemical compositions is also relatively simple. Additionally, these descriptors

effectively capture substitutions of chemically similar elements within the same group by reflecting changes in average property values. This ensures that even when materials have elements within a group, the periodic descriptors can still identify them. For instance, consider $Li_3P(S_{3.8}Se_{0.2})$ and $Li_3P(S_{3.6}Se_{0.4})$, which substitutes S in Li_3PS_4 with Se. The coefficient sum in Group 16 (S and Se) remains the same, while the periodic descriptors can identify the material based on the specific coefficient of the two elements and their atomic properties. Therefore, even low-dimensional descriptors can sufficiently represent compositional information.

The periodic descriptors provide a comprehensive representation of both the composition and atomic properties, adhering to the principles of simplicity and reversibility. Further modifications can be made to the descriptors to meet specific requirements by incorporating additional groupings or using different atomic properties. Possible grouping methods include separating Lanthanides and Actinides from Group 3, as the large number of elements within this group can complicate identification.

To analyze the impact of different atomic properties on prediction performance, we conducted a comparison using a range of properties, including Pauling electronegativity, neutron cross-section, atomic number, atomic radius, atomic weight, first ionization energy, and electron affinity. Using these periodic descriptors, we trained models using ionic conductivity data, as detailed in Model Study 1. The training followed the method outlined in the Model Training section. The prediction performance is presented in **Figure 3**, showing similar R^2 values for most properties, with electron affinity yielding a relatively lower R^2. The selection of properties may need to be adjusted according to various targets. In the subsequent analysis, atomic radius is used.

2.1 Comparative analysis

We evaluate our periodic descriptors in comparison with established compositional descriptor generators, namely XenonPy and Matminer. Our comparative framework employed five machine learning models to predict ionic conductivity, using the same dataset across all methods to ensure fairness in evaluation. The training data used is described in Model Study 1, and the method detailed in the section of Model Training. The results of the comparative analysis for predicting ionic conductivities are summarized in **Table 1**, showing the mean test R^2 values for each descriptor type across all machine learning models.

In addition to performance, our descriptors offer two important advantages. First, they exhibit enhanced generalizability across chemical space. While the training data includes only 55 elements (**Table 2**), all 17 periodic groups are represented in the dataset (**Table 3**). This enables the model to learn from chemically similar elements within the same group, even when specific elements are absent. Second, the compact and chemically structured nature of periodic descriptors enables inverse design, allowing one to search in descriptor space and translate optimized values back into chemical formulas. In contrast, high-dimensional descriptors often lack reversibility and are impractical for compositional optimization due to the complexity of the feature space.

The comparison demonstrates the efficiency of our periodic descriptors offer not only a reduced number of descriptors but also maintain or surpass the predictive performance of more complex sets. This efficiency signifies a substantial reduction in computational overhead and a more interpretable model, without sacrificing accuracy. This balance of

simplicity and effectiveness shows the potential of the periodic descriptors to accelerate the discovery of electrochemical materials.

2.2 Descriptor coverage and uniqueness across chemical systems

We address the question of why periodic descriptors based solely on chemical composition are effective for training machine learning models, particularly in scenarios where structural information is limited or unavailable. To support this, we provide an analysis of descriptor combinations across different chemical systems and evaluate the prediction performance of models with and without structural descriptors.

The analysis was conducted using the AtomWork-Adv. (AWA) database [21], developed by NIMS, which provides data on crystal structures, phase diagrams, and properties of inorganic materials derived from experimental reports. **Table 4** presents the number of unique substances categorized by chemical systems (unary, binary, ternary, and ≥ quaternary), and summarizes the ratios of unique descriptor combinations to the total number of substances. These combinations include chemical formula alone (referred to as "compound") and chemical formula combined with structural descriptors, such as Bravais lattice, space group, and Pearson symbol.

As shown in **Table 4**, the descriptor uniqueness ratio increases with the complexity of the chemical system. This trend reflects a significant observation: in high-element systems (e.g., ≥ quaternary), the structural type of a material is often uniquely determined by its chemical formula. For instance, 90% of substances in total exhibit a unique structural type once their chemical formula is known. For unary systems, the ratio increases from 18% (compound) to 93% (compound + space group + Pearson symbol), and for binary systems, it increases from 74% to 95%. This trend reflects that

structural ambiguity is more pronounced in low-element systems. For higher-element systems (e.g., ≥ quaternary), the ratio reaches 89% for compounds and 99% when structural descriptors are included, reflecting the strong correlation between chemical formulas and structural types in these systems.

To investigate the role of structural descriptors further, we evaluated the prediction performance of machine learning models for ionic conductivity using periodic descriptors alone and periodic descriptors augmented with structural descriptors (space group and Pearson symbol). The details of the training dataset are provided in **Model Study 1**. The results indicate no significant improvement in prediction accuracy with the inclusion of structural descriptors as shown in **Table 1**. The mean R^2 values for ionic conductivity predictions remain comparable between models trained with and without structural descriptors. Notably, we use space group and Pearson symbol as structural descriptors instead of more detailed features such as atomic distances or coordination numbers. This ensures applicability to high-element system materials, for which detailed structural information is often unavailable. These findings align with the observation that quaternary and higher systems, which dominate the training dataset, already exhibit strong correlations between chemical formulas and structural types, as reflected in **Table 4**.

3. Model study 1: fast Li-ion conductor

The dataset used for model training comprises 887 materials, including 588 chemical formulas and 208 chemical systems. During preprocessing, materials classified as composites were excluded. Additionally, any chemical formulas containing undefined variables (e.g., “x”, “y”) or not conforming to standard compositional formats were

removed to ensure data consistency. The frequency of elements in the resulting training set is summarized in **Table 2**. The data points represent ionic conductivity fitted to 300 K, derived from experimental measurements across a range of temperatures using the Battery Material Database, AtomWork-Battery (AWB)[22]. For cases where the temperature range does not include 300 K, the values were extrapolated. The distribution of ionic conductivity and structure types are shown in **Figure 4**. Most samples fall into categories such as garnet, NASICON, perovskite, and LISICON.

This database provides comprehensive information on the composition, structure, and properties of battery materials, as well as the performance metrics of batteries constructed from these materials. Further details about AWB are available in our related publication[23]. Using periodic descriptors based on atomic radius, the model achieved a mean test R^2 of 0.723, as shown in Table 1. Feature importance analysis using ensemble tree models revealed that descriptors corresponding to Group 12 and Group 16 were consistently ranked as highly influential (Figure S1). The prominence of Group 16 is chemically reasonable, as it comprises anionic species (e.g., O, S, Se) known to play a critical role in Li-ion transport. Previous studies have shown that larger anions can enhance ionic conductivity by increasing accessible diffusion volume and reducing electrostatic interactions with Li-ion.[24] Furthermore, anion substitutional disorder has been linked to improved ionic mobility in several solid electrolytes.[25,26] The high feature importance observed for Group 12 may be related to the high variability in ionic conductivity among samples containing Group 12 elements. As shown in Table 3, Group 12 (33 samples) exhibits a notably large standard deviation compared to other groups, suggesting that changes in the composition of this group may correlate to substantial variation in ionic conductivity. However, given the limited number and

variability of samples, the scientific significance of Group 12 requires further analysis. The parity plot for the Random Forest model is provided in **Figure 5**, demonstrating strong alignment with observed values across all cross-validation folds. Parity plots of other models are shown in **Figure S2**.

Following model training, we applied the models to predict Li-containing materials from AWA [21], aiming to explore potential candidates. The AWA database comprises 7,705 Li-containing materials spanning 4,285 chemical systems, with 241 materials overlapping with the AWB database. We ranked predicted ionic conductors from highest to lowest based on their predicted ionic conductivities. We then focused on the overlap among the top-1000 predictions across the five models (GPR(rbf), GPR(rq), NGB, XGB, and RF), considering materials identified by five models as viable fast ionic conductor candidates. This selection criterion led to the identification of 28 materials. The elements present in the 28 materials, along with the proportion of materials containing each element, are illustrated in **Figure 6 (a)**.

Among these 28 candidates, **Figure 6 (b)** revealed that: 21.4% were already included in our training data, validating the models' ability to recognize known conductors. 35.7% matched with verified ionic conductors documented in other scientific literature. Notably, several candidates exhibit experimentally measured high ionic conductivities, as summarized in **Table 5**. For instance, $Li_{6.5}Si_{0.5}S_5Sb_{0.5}I$ (No. 6) stands out with a reported conductivity of 14.8 mS/cm, similarly, $Li_{10}P_2S_{11.2}Ge_{0.6}Se_{0.8}Sn_{0.4}$ (No. 5) approximates an ionic conductivity of 2.7 mS/cm. These two categories comprise 57.1% of the predictions, demonstrating the models' predictive performance and the effectiveness of the periodic descriptors. The remaining 42.9% of the predictions, which have not yet been reported as ionic conductors, represent promising candidates for high

ionic conductivity.

3.1 Validation of predicted new ionic conductors through simulation

Among the new findings in **Table 5**, well-known structure types such as thio-LISICON and argyrodite are identified as fast Li-ion conductors. Notably, we also identified anti-fluorite structures, a class not traditionally associated with high ionic conductivity. We analyzed the antifluorite of $Li_{1.73}Ge_{0.27}As_{0.82}Se_{0.18}$ (No. 17 in **Table 5**), modifying the coefficient slightly to $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ for simulation. Utilizing molecular dynamics (MD) simulations with machine-learned force fields trained to replicate first-principles results, we examined the ionic conductivity of this material across various temperatures. The results in **Figure 7** reveal a temperature-dependent transition in ionic diffusion for the anti-fluorite structure $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ occurring around 800–900 K. When extrapolating ionic conductivity using data points above 900 K, a high value of 4.88 mS/cm at 300 K is obtained. However, extrapolation based on simulated data from the lower temperature range of 575–800 K yields a much lower ionic conductivity of 8.09×10^{-6} mS/cm. As shown in **Figure 8**, compared to conventional anti-fluorites like Li_2S and Li_2Se[27,28]—which require temperatures exceeding 1000 K to achieve similar conductivity levels—the anti-fluorite structure $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ demonstrates ionic conductivity exceeding 10 mS/cm at lower temperatures of 600–700 K and lower transition temperature. Similar temperature-dependent transitions have been observed in ionic conductors like Li_3YCl_6, and the transition temperature was further reduced by substituting Cl with Br[29]. This suggests that further improvements in room-temperature ionic conductivity could be achieved by strategically lowering the transition temperature through compositional tuning.

These findings also show a limitation of the periodic descriptor approach. While effective in predicting trends based on chemical formulas, the method cannot account for changes resulting from temperature-dependent transitions within the same composition. The observed temperature-dependent transitions across all simulated anti-fluorite structures highlight their potential for achieving high ionic conductivity. Further exploration of these findings will be presented in a forthcoming publication.

4. Model study 2: OER electrocatalyst

For efficient water electrolysis, platinum-group metals like platinum, ruthenium, and iridium are typically employed as electrocatalysts in water electrolysers[30]. However, the limited availability and high cost of these metals pose challenges for scaling up sustainable hydrogen production. This model study aims to assess the predictive performance of our periodic descriptors and identify PGM-free OER electrocatalysts.

To evaluate electrocatalysts, several target properties, including overpotential, current density, and Tafel slope, are generally considered. After comparing these target properties in terms of data availability, data quality, and prediction accuracy, we selected overpotential at a current density of 1 mA/cm² as our target property. Our machine learning models were trained on a systematically curated experimental dataset comprising 120 metal oxides and prepared using a consistent synthetic protocol. All compositions are quaternary oxides consisting of four metal elements (excluding oxygen). The dataset includes 11 different metal elements, with Ni present in all 120 samples. The remaining elements—Mo, Sc, Ta, Fe, Zn, Zr, Mn, Co, Cu, and Ti—each appear in 36 samples. The distribution of overpotential at 1 mA/cm^2 is shown in **Figure 9 (**a), where the majority of samples fall between 400–800 mV. A small subset of 37

entries exhibits overpotentials reported as 1000 mV, which reflects an upper measurement limit rather than exact values. These entries were excluded from the histogram, and their imprecision may have contributed to the reduced model performance. Using periodic descriptors based on atomic radius, the GPR model with an rbf kernel achieved the highest predictive performance among the tested models, with a test R^2 of 0.635 (**Figure 9**(b)). The remaining model results are presented in **Figure S**3, where the performance trends are generally consistent but slightly affected by samples with extreme overpotentials near 1000 mV.

To explore potential electrocatalyst candidates, we employed a dataset comprising compounds made from 18 PGM-free elements: Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Y, Zr, Nb, Mo, Ag, Hf, Ta, and W, as shown in **Figure 10 (a)**. These elements contribute to more cost-effective and sustainable electrocatalysts, thereby supporting green hydrogen production. Notably, only the metal compositions were considered when generating the search space for metal oxides. This approach ensured consistency, as all compounds underwent the same experimental oxidation process, allowing us to focus solely on the influence of metal compositions on electrocatalyst performance. We generated 1,358,816 compounds, ranging from binary system ($E_{1x_1}E_{2x_2}$) to quinary systems ($E_{1x_1}E_{2x_2}E_{3x_3}E_{4x_4}E_{5x_5}$), where E_i represents each element and x_i denotes its coefficient. The coefficient, x_i, ranged from 0.1 to 1 in steps of 0.1, ensuring that the sum of all x_i equaled 1.

We ranked the predicted electrocatalysts from lowest to highest overpotential, as lower overpotential corresponds to more efficient water electrolysis. We focused on the overlap among the top-3000 predictions across the five models (GPR(rbf), GPR(rq),

NGB, XGB, and RF), identifying 20 candidates consistently predicted by all models as potential electrocatalysts.

4.1 Validation of predicted new electrocatalysts through experiments

Among the predicted candidates, $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide demonstrated promising linear sweep voltammetry (LSV) curves when compared against two benchmark electrocatalysts in **Figure 10 (b)**. The first benchmark, RuO_2, is a PGM-based oxide known for its high activity. The second benchmark, a PGM-free $MnNiFeZn_{0.5}Ag$ oxide, was previously predicted and validated for high electrochemical performance in OER[15].

As shown in **Figure 10 (b)**, the predicted candidate $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide exhibits a steeper slope than the benchmarks. The predicted overpotential at 1 mA/cm^2 for $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide was 335.38 mV, which closely aligns with the experimental value of 353.40 mV. Additionally, the potential required to achieve a current density of 40 mA/cm^2 is approximately 1.9 V vs. the reversible hydrogen electrode (RHE) for the PGM-based benchmark and 2.0 V vs. RHE for the PGM-free benchmark. In contrast, our candidate achieves a lower potential of about 1.8 V vs RHE, indicating improved water electrolysis efficiency.

This experimental validation shows that $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide is a highly efficient electrocatalyst for water electrolysis, achieving a lower overpotential than the standard PGM-based and PGM-free electrocatalysts. The consistent results between the predictions and experiments underscore the reliability and accuracy of our model with periodic descriptors for identifying high-performance electrocatalysts.

5. Conclusion

In this study, we developed an advanced approach for material design using periodic descriptors, focusing on applications in fast Li-ion conductors for solid electrolytes and electrocatalysts for water electrolysis. The periodic descriptors, characterized by their simplicity and reversibility, provide a new perspective for discovering materials essential to green energy applications.

We demonstrated through a comprehensive dataset and five distinct machine learning models that the periodic descriptors outperform established benchmarks such as XenonPy and Matminer. For Fast Li-ion conductors, our predictive framework identified 28 potential candidates, with 56.1% correlating with known ionic conductors in training data or verified by other literature. Notably, 42.9% of these candidates represent new findings, offering significant opportunities for further research and development. Simulation validation revealed that the anti-fluorite structures exhibit superionic conductivity behavior at 600-700 K, significantly lower than conventional anti-fluorite compounds like Li_2S and Li_2Se.

For electrocatalysts, our model identified $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide as a promising PGM-free candidate with a current density comparable to the benchmark PGM electrocatalyst of RuO_2. The experimental validation confirmed that this candidate achieves efficient water electrolysis at lower overpotentials, highlighting the practical applicability of the periodic descriptors.

The periodic descriptors demonstrate robust predictive power and can serve as design guidelines for developing electrochemical materials with enhanced properties. Our approach offers an efficient method for exploring and designing solid electrolytes,

electrocatalysts, and other inorganic materials, contributing to the advancement of renewable energy solutions.

6. Methods

6.1 Model training

The datasets comprise periodic descriptors and targeted properties. The periodic descriptors incorporate coefficients and atomic radius for both comparative analysis and the two model studies. The datasets are divided into 80% for training and 20% for testing, ensuring ample data for model training while reserving a portion for evaluation. Hyperparameters are optimized using a grid search method. A 5-fold cross-validation is applied to partition the training data into subsets, allowing for robust validation of each model's performance.

Five machine learning algorithms from scikit-learn are used: Gaussian Process Regression with Radial Basis Function (GPR(rbf)), Gaussian Process Regression with Rational Quadratic kernel (GPR(rq)), NGBoost (NGB), XGBoost (XGB), and Random Forest (RF). The Support Vector Regression (SVR) is used in Comparative analysis. The feature importance rankings from RF and XGB models trained on the Li-ion conductor dataset are extracted. Model performance is evaluated using the R^2 score, which measures how well the model captures data variance, with a score closer to 1 indicating better predictive accuracy.

6.2 Simulation process for antifluorite

The CIF file for Li_2S, Li_2Se, and the target material $Li_{1.73}Ge_{0.27}As_{0.82}Se_{0.18}$ was sourced from AtomWork-Adv. (AWA) [21]. The composition is slightly adjusted to

$Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ to generate 4×4×4 and 2×2×2 supercell structure models. For each supercell size, atomic arrangements of both anion and cation sites in $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ were enumerated to obtain all symmetry-inequivalent configurations. From these, 20 configurations with the lowest Ewald electrostatic energies were selected for subsequent DFT calculations. All computations were performed using the Vienna Ab initio Simulation Package (VASP)[31] with the projector augmented wave (PAW) method[32]. The generalized gradient approximation (GGA) with the Perdew–Burke–Ernzerhof (PBE) functional was used for electron exchange-correlation energy[33]. A plane-wave energy cutoff of 500 eV and a 3×3×3 Γ-centered k-mesh were employed. Convergence criteria were set to 1×10^{-6} eV for energy and 0.01 eV/Å for forces. The lowest energy structure from DFT was then used to train a Machine learning force field (MLFF) via on-the-fly ab initio molecular dynamics (AIMD) simulations within VASP[34]. Subsequently, machine learning molecular dynamics (MLMD) simulations were conducted in an NVT ensemble, using the MLFF with a 1 fs time step and a Nosé–Hoover thermostat[35] over 100 ps.

6.3 Experimental process for electrocatalysts

Electrocatalysts were synthesized according to our previously reported procedure[15]. First, stock solutions of metal salts were prepared with ethanol (99.5%, Fujifilm) as solvent. Then a solution was obtained by mixing the stock solutions at a desired ratio (e.g. Fe : Co : Cu : Ag : W=1 : 1 : 1 : 1 : 6 for $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$). The mixed solution was deposited on a Ti substrate to form a thin film and then it was heated at 350 ℃ for 10 min. The deposition and heating process was repeated for 10 times. Finally, the film was calcined at 450℃ for 1 h to obtain an oxide layer on the Ti

substrate. 0.1M KOH solution was prepared by mixing KOH (99.99%, Sigma-Aldrich) and Milli-Q water (18.2 MΩ・cm; TOC < 5 ppb), and used as the electrolyte in this study.

The Electrochemical measurements were performed on a multichannel potentio/galvanostat (VMP-3, Bio-Logic, France). A three-electrode cell was set up, with the as-prepared Thin film electrode as the working electrode, a platinum wire as the counter electrode and an Ag/AgCl electrode as the reference electrode. The linear sweep voltammetry (LSV) method was conducted in 0.1 M KOH aqueous solution at room temperature (298 K) with a scan rate of 1 mV/s.

Acknowledgments

This work was supported in part by the 'Advanced Battery Collaboration' project of COI-NEXT Grant Number JPMJPF2016 of the Japan Science and Technology Agency (JST), the Materials Open Platform for All Solid-State Battery, the Research Network and Facility Services Division in the National Institute for Materials Science (NIMS), and the MEXT Program: Data Creation and Utilization-Type Material Research and Development Project Grant Number JPMXP1122712807. The authors are indebted to the National Institute for Material Science.

References

[1] Zheng F, Kotobuki M, Song S, et al. Review on solid electrolytes for all-solid-state lithium-ion batteries. Journal of Power Sources. 2018;389:198-213. doi: 10.1016/j.jpowsour.2018.04.022

[2] Zhao W, Yi J, He P, et al. Solid-State Electrolytes for Lithium-Ion Batteries: Fundamentals, Challenges and Perspectives. Electrochemical Energy Reviews. 2019;2(4):574-605. doi: 10.1007/s41918-019-00048-0

[3] Li Q, Chen J, Fan L, et al. Progress in electrolytes for rechargeable Li-based batteries and beyond. Green Energy & Environment. 2016;1(1):18-42. doi: 10.1016/j.gee.2016.04.006

[4] Gao X, Chen Y, Wang Y, et al. Next-Generation Green Hydrogen: Progress and Perspective from Electricity, Catalyst to Electrolyte in Electrocatalytic Water Splitting. Nanomicro Lett. 2024 Jul 5;16(1):237. doi: 10.1007/s40820-024-01424-2

[5] Zainal BS, Ker PJ, Mohamed H, et al. Recent advancement and assessment of green hydrogen production technologies. Renewable and Sustainable Energy Reviews. 2024;189. doi: 10.1016/j.rser.2023.113941

[6] Xiao Y, Jun K, Wang Y, et al. Lithium Oxide Superionic Conductors Inspired by Garnet and NASICON Structures. Advanced Energy Materials. 2021;11(37). doi: 10.1002/aenm.202101437

[7] Jalem R, Yamamoto Y, Shiiba H, et al. Concerted Migration Mechanism in the Li Ion Dynamics of Garnet-Type Li7La3Zr2012. Chemistry of Materials. 2013;25(3):425-430. doi: 10.1021/cm303542x

[8] Liao X, Lu R, Xia L, et al. Density Functional Theory for Electrocatalysis. Energy & Environmental Materials. 2021;5(1):157-185. doi: 10.1002/eem2.12204

[9] Schmidt J, Marques MRG, Botti S, et al. Recent advances and applications of machine learning in solid-state materials science. npj Computational Materials. 2019;5(1). doi: 10.1038/s41524-019-0221-0

[10] Sendek AD, Yang Q, Cubuk ED, et al. Holistic computational structure screening of more than 12 000 candidates for solid lithium-ion conductor materials. Energy & Environmental Science. 2017;10(1):306-320. doi: 10.1039/c6ee02697d

[11] Sendek AD, Cubuk ED, Antoniuk ER, et al. Machine Learning-Assisted Discovery of Solid Li-Ion Conducting Materials. Chemistry of Materials. 2018;31(2):342-352. doi: 10.1021/acs.chemmater.8b03272

[12] Muy S, Voss J, Schlem R, et al. High-Throughput Screening of Solid-State Li-Ion Conductors Using Lattice-Dynamics Descriptors. iScience. 2019 Jun 28;16:270-282. doi: 10.1016/j.isci.2019.05.036

[13] Zhang Y, He X, Chen Z, et al. Unsupervised discovery of solid-state lithium ion conductors. Nat Commun. 2019 Nov 20;10(1):5260. doi: 10.1038/s41467-019-13214-1

[14] Kajita S, Ohba N, Suzumura A, et al. Discovery of superionic conductors by ensemble-scope descriptor. NPG Asia Materials. 2020;12(1). doi:

10.1038/s41427-020-0211-1

[15] Sakaushi K, Hoisang W, Tamura R. Human-Machine Collaboration for Accelerated Discovery of Promising Oxygen Evolution Electrocatalysts with On-Demand Elements. ACS Cent Sci. 2023 Dec 27;9(12):2216-2224. doi: 10.1021/acscentsci.3c01009

[16] Zhong M, Tran K, Min Y, et al. Accelerated discovery of CO(2) electrocatalysts using active machine learning. Nature. 2020 May;581(7807):178-183. doi: 10.1038/s41586-020-2242-8

[17] Steinmann SN, Wang Q, Seh ZW. How machine learning can accelerate electrocatalysis discovery and optimization. Mater Horiz. 2023 Feb 6;10(2):393-406. doi: 10.1039/d2mh01279k

[18] Esterhuizen JA, Mathur A, Goldsmith BR, et al. High-Performance Iridium-Molybdenum Oxide Electrocatalysts for Water Oxidation in Acid: Bayesian Optimization Discovery and Experimental Testing. J Am Chem Soc. 2024 Feb 28;146(8):5511-5522. doi: 10.1021/jacs.3c13491

[19] XenonPy: yoshida-lab; 2022 [cited 2023 2023/10/12]; XenonPy is a Python library that implements a comprehensive set of machine learning tools for materials informatics.]. Available from: https://xenonpy.readthedocs.io/en/latest/index.html

[20] Ward L, Dunn A, Faghaninia A, et al. Matminer: An open source toolkit for materials data mining. Comp Mater Sci. 2018 Sep;152:60-69. doi: 10.1016/j.commatsci.2018.05.018

[21] AtomWork-Adv.: NIMS; 2018 [cited 2023 2023/3/20]. Available from: https://atomwork-adv.nims.go.jp/

[22] AtomWork-Battery 2020 [cited 2023 2023/3/20]. Available from: https://www.nims.go.jp/cbrm/en/about/pc8qa60000000032.html

[23] Xu Y, Wu Y-J, Li H, et al. A comprehensive data network for data-driven study of battery materials. Science and Technology of Advanced Materials. 2024;25(1). doi: 10.1080/14686996.2024.2403328

[24] Bachman JC, Muy S, Grimaud A, et al. Inorganic Solid-State Electrolytes for Lithium Batteries: Mechanisms and Properties Governing Ion Conduction. Chem Rev. 2016 Jan 13;116(1):140-62. doi: 10.1021/acs.chemrev.5b00563

[25] Dyre JC, Maass P, Roling B, et al. Fundamental questions relating to ion conduction in disordered solids. Reports on Progress in Physics. 2009;72(4). doi: 10.1088/0034-4885/72/4/046501

[26] Morgan BJ. Mechanistic Origin of Superionic Lithium Diffusion in Anion-

Disordered Li(6)PS(5) X Argyrodites. Chem Mater. 2021 Mar 23;33(6):2004-2018. doi: 10.1021/acs.chemmater.0c03738

[27] Gupta MK, Singh B, Goel P, et al. Lithium diffusion in Li2X (X=O, S, and Se): Ab initio simulations and inelastic neutron scattering measurements. Physical Review B. 2019;99(22). doi: 10.1103/PhysRevB.99.224304

[28] Liu F, Wang L, Zhang Z, et al. A Mixed Lithium - Ion Conductive Li2S/Li2Se Protection Layer for Stable Lithium Metal Anode. Advanced Functional Materials. 2020;30(23). doi: 10.1002/adfm.202001607

[29] Liu Z, Chien PH, Wang S, et al. Tuning collective anion motion enables superionic conductivity in solid-state halide electrolytes. Nat Chem. 2024 Oct;16(10):1584-1591. doi: 10.1038/s41557-024-01634-6

[30] Hoisang W, Sakaushi K. Key criteria for next-generation dimensionally stable electrodes towards large-scale green hydrogen production by water electrolysis. Current Opinion in Electrochemistry. 2022;36. doi: 10.1016/j.coelec.2022.101136

[31] Kresse G, Furthmuller J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. Phys Rev B Condens Matter. 1996 Oct 15;54(16):11169-11186. doi: 10.1103/physrevb.54.11169

[32] Blochl PE. Projector augmented-wave method. Phys Rev B Condens Matter. 1994 Dec 15;50(24):17953-17979. doi: 10.1103/physrevb.50.17953

[33] Perdew JP, Burke K, Ernzerhof M. Generalized Gradient Approximation Made Simple. Phys Rev Lett. 1996 Oct 28;77(18):3865-3868. doi: 10.1103/PhysRevLett.77.3865

[34] Jinnouchi R, Karsai F, Kresse G. On-the-fly machine learning force field generation: Application to melting points. Physical Review B. 2019;100(1). doi: 10.1103/PhysRevB.100.014105

[35] Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. Phys Rev A Gen Phys. 1985 Mar;31(3):1695-1697. doi: 10.1103/physreva.31.1695

[36] Aono H, bin Idris MA, Sadaoka Y. Ionic conductivity and crystal structure for the Li3-2xCr2-xTax(PO4)3 system. Solid State Ionics. 2004 Jan;166(1-2):53-59. doi: 10.1016/j.ssi.2003.11.005

[37] Sun Y, Yan WN, An L, et al. A facile strategy to improve the electrochemical stability of a lithium ion conducting Li10GeP2S12 solid electrolyte. Solid State Ionics. 2017 Mar;301:59-63. doi: 10.1016/j.ssi.2017.01.014

[38] Morales M, Mestres L, Dlouhá M, et al. Crystal structure of the mixed conductors phases, Li0.5-3xLa0.5+x+yTi1-3yM3yO3 (M=Mn, Cr) with x=0.133 and

y=0.20. Journal of Materials Chemistry. 1998;8(12):2691-2694. doi: 10.1039/a804764b

[39] Mori D, Sugimoto K, Matsuda Y, et al. Synthesis, Structure and Ionic Conductivity of Garnet Like Lithium Ion Conductor Li6.25+xGa0.25La3-xSrxZr2012. J Electrochem Soc. 2018 Dec 13;166(3):A5168-A5173. doi: 10.1149/2.0171903jes

[40] Yang K, Dong J, Zhang L, et al. Dual Doping: An Effective Method to Enhance the Electrochemical Properties of Li10GeP2S12 - Based Solid Electrolytes. Journal of the American Ceramic Society. 2015 Dec;98(12):3831-3835. doi: 10.1111/jace.13800

[41] Zhou L, Assoud A, Zhang Q, et al. New Family of Argyrodite Thioantimonate Lithium Superionic Conductors. J Am Chem Soc. 2019 Dec 4;141(48):19002-19013. doi: 10.1021/jacs.9b08357

[42] Yang K, Dong JY, Zhang L, et al. Dual Doping: An Effective Method to Enhance the Electrochemical Properties of Li10GeP2S12-Based Solid Electrolytes. Journal of the American Ceramic Society. 2015 Dec;98(12):3831-3835. doi: 10.1111/jace.13800

[43] Xie H, Li Y, Han J, et al. Li6La3SnMO12(M = Sb, Nb, Ta), a Family of Lithium Garnets with High Li-Ion Conductivity. J Electrochem Soc. 2012;159(8):A1148-A1151. doi: 10.1149/2.009208jes

[44] Stramare S, Weppner W. Structure and conductivity of B-site substituted (Li,La)TiO3. Materials Science and Engineering: B. 2004;113(1):85-90. doi: 10.1016/j.mseb.2004.07.003

[45] Li YJ, Cao ML, Zhou CX, et al. A novel synthesis of gadolinium-doped Li3V2(PO4)3/C with excellent rate capacity and cyclability. Rsc Advances. 2016;6(34):28624-28632. doi: 10.1039/c6ra03109a

[46] Šalkus T, Barre M, Kežionis A, et al. Ionic conductivity of Li1.3Al0.3-xScxTi1.7(PO4)3 (x=0, 0.1, 0.15, 0.2, 0.3) solid electrolytes prepared by Pechini process. Solid State Ionics. 2012;225:615-619. doi: 10.1016/j.ssi.2012.03.045

[47] Zhao E, Guo Y, Xu G, et al. High ionic conductivity Y doped Li1.3Al0.3Ti1.7(PO4)3 solid electrolyte. Journal of Alloys and Compounds. 2019;782:384-391. doi: 10.1016/j.jallcom.2018.12.183

[48] Kwon O, Hirayama M, Suzuki K, et al. Synthesis, structure, and conduction mechanism of the lithium superionic conductor Li10+δGe1+δP2-δS12. Journal of Materials Chemistry A. 2015;3(1):438-446. doi: 10.1039/c4ta05231e

[49] Kato Y, Saito R, Sakano M, et al. Synthesis, structure and lithium ionic conductivity of solid solutions of Li10(Ge1-M)P2S12 (M = Si, Sn). Journal of Power Sources. 2014;271:60-64. doi: 10.1016/j.jpowsour.2014.07.159

[50] Bron P, Dehnen S, Roling B. Li10Si0.3Sn0.7P2S12 - A low-cost and low-grain-boundary-resistance lithium superionic conductor. Journal of Power Sources. 2016;329:530-535. doi: 10.1016/j.jpowsour.2016.08.115

[51] Iwasaki R, Hori S, Kanno R, et al. Weak Anisotropic Lithium-Ion Conductivity in Single Crystals of Li10GeP2S12. Chemistry of Materials. 2019;31(10):3694-3699. doi: 10.1021/acs.chemmater.9b00420

Table 1 Prediction performance of Li-ion conductors with different descriptor sets. The table presents the mean test R^2 values for each descriptor set across the five machine learning models: Gaussian Process Regression with Radial Basis Function (GPR(rbf)), Gaussian Process Regression with Rational Quadratic Kernel (GPR(rq)), NGBoost (NGB), XGBoost (XGB), Random Forest (RF), and Support Vector Regression (SVR).

Descriptor sets	Number of descriptors	GPR(rbf)	GPR(rq)	NGB	XGB	RF	SVR	Mean test R^2
Periodic descriptors	34	0.686	0.718	0.727	0.731	0.743	0.735	0.723
Periodic descriptors + space group + Pearson symbol	36	0.665	0.664	0.715	0.742	0.761	0.678	0.704
XenonPy	290	0.557	0.677	0.752	0.749	0.733	0.175	0.607
Matminer	132	-0.303	0.591	0.788	0.755	0.762	0.567	0.527

Table 2 Element frequency in the training data set of ionic conductors.

Element	**Frequency**	**Element**	**Frequency**
Li	850	In	16
O	736	Hf	16
La	373	F	16
Zr	311	Ce	14

P	287	V	12
Al	221	Sn	12
S	138	Te	11
Ti	125	C	11
Ta	125	Co	11
Nb	101	Cu	10
Si	83	N	10
Ge	77	Ni	10
Sr	61	Cr	9
B	55	Sm	9
Ba	48	Mo	7
H	36	Bi	7
Na	35	As	7
Zn	32	Sb	6
Cl	28	Eu	6
Sc	27	Se	6
Mg	26	I	3
Br	25	U	2
W	23	Mn	2
Fe	22	Nd	1
Ca	22	Ru	1
Ga	22	Gd	1
Y	22	Cd	1
K	21		

Table 3 The mean, standard deviation (STD), and count of Log_{10}(ionic conductivity σ_S/m) for each periodic group (Group) in the training data set of ionic conductors.

Group	**Mean Log_{10} (σ_S/m)**	**STD Log_{10} (σ_S/m)**	**Count**
1	-2.244	1.609	871
2	-2.511	1.806	157

3	-2.322	1.083	433
4	-2.230	1.020	448
5	-2.080	1.007	238
6	-3.209	1.951	39
7	-3.935	1.366	2
8	-4.283	1.942	23
9	-4.364	1.250	11
10	-3.805	0.066	10
11	-3.215	1.154	10
12	-8.448	5.198	33
13	-2.209	1.400	312
14	-1.729	1.298	182
15	-2.065	1.963	313
16	-2.248	1.615	853
17	-1.573	0.994	70

Table 4 Number of substances and descriptor uniqueness ratios for different descriptor combinations across chemical systems. The ratios indicate the fraction of substances uniquely described by each descriptor combination.

	unary	unary (%)	binary	binary (%)	ternary	ternary (%)	≥ quaternary	≥ quaternary (%)	total	total (%)
substances	648	100%	15264	100%	57993	100%	84363	100%	158268	100%
compound	116	18%	11248	74%	51853	89%	78743	93%	141960	90%
compound + Bravais lattice	486	75%	13752	90%	56422	97%	83226	99%	153886	97%
compound + space group	544	84%	14098	92%	56987	98%	83527	99%	155156	98%
compound + Pearson symbol	595	92%	14368	94%	57144	99%	83739	99%	155846	98%
compound + space group + Pearson symbol	601	93%	14507	95%	57393	99%	83894	99%	156395	99%

Table 5 Predicted candidate materials for potential fast Li-ion conductors. This table lists the 28 materials identified by our models using periodic descriptors as potential fast ionic conductors, consistently predicted across all five machine learning models. It includes each candidate's chemical formula, structure type, and experimentally

measured ionic conductivity (mS/cm). Notes indicate whether the material is part of our training data ("In training data"), verified by other studies ("verified candidates"), or represents new discoveries ("NEW findings").

No.	Material	Structure type	Ionic conductivity (mS/cm)	Note
1	$Li_{2.2}O_{12}P_3Cr_{1.6}Ta_{0.4}$		0.024	verified candidates[36]
2	$Li_{9.4}P_2S_{12}GeBa_{0.3}$	thio-LISICON	0.7	verified candidates[37]
3	$Li_{0.1}O_3Ti_{0.4}Mn_{0.6}La_{0.833}$	perovskite	1.1	verified candidates[38]
4	$Li_{6.35}O_{12}Ga_{0.25}Sr_{0.1}Zr_2La_{2.9}$	garnet family	1.3	verified candidates[39]
5	$Li_{10}P_2S_{11.2}Ge_{0.6}Se_{0.8}Sn_{0.4}$	thio-LISICON	2.7	verified candidates[40]
6	$Li_{6.5}Si_{0.5}S_5Sb_{0.5}I$	argyrodite family	14.8	verified candidates[41]
7	$Li_{10}P_2S_{10}Se_2Sn$	thio-LISICON	2	verified candidates[42]
8	$Li_6O_{12}SnSbLa_3$	garnet family	0.022	verified candidates[43]
9	$Li_{0.19}O_3Ti_{0.9}Ni_{0.1}La_{0.67}$	perovskite		verified candidates[44]
10	$Li_3O_{12}P_3V_{1.95}Gd_{0.05}$		cathode	verified candidates[45]
11	$Li_{1.3}O_{12}Al_{0.25}P_3Sc_{0.05}Ti_{1.7}$	Nasicon family	0.18	In training data[46]
12	$Li_{1.3}O_{12}Al_{0.2}P_3Ti_{1.7}Y_{0.1}$	Nasicon family	0.78	In training data[47]
13	$Li_{10.5}P_{1.5}S_{12}Ge_{1.5}$	thio-LISICON	14.3	In training data[48]

14	$Li_{10}P_2S_{12}Ge_{0.7}Sn_{0.3}$	thio-LISICON	7.67	In training data[49]
15	$Li_{10}P_2S_{12}Sn$	thio-LISICON	5.78	In training data[50]
16	$Li_{10}P_2S_{12}Ge$	thio-LISICON	27.8	In training data[51]
17	$Li_{1.73}Ge_{0.27}As_{0.82}Se_{0.18}$	fluorite family		NEW findings
18	$Li_{1.86}P_{0.71}S_{0.29}Cr_{0.14}$	fluorite family		NEW findings
19	$Li_{1.86}Cr_{0.14}As_{0.72}Se_{0.28}$	fluorite family		NEW findings
20	$Li_{1.73}Si_{0.27}As_{0.82}Se_{0.18}$	fluorite family		NEW findings
21	$Li_{10}P_2S_{12}Ge_{0.67}Sn_{0.33}$	thio-LISICON		NEW findings
22	$Li_{2.58}In_{0.42}EuBi_2$			NEW findings
23	Li_8N_2Te			NEW findings
24	Li_6AsSe_5I	argyrodite family		NEW findings
25	$Li_{6.69}PCl_{0.31}Se_{5.69}$	argyrodite family		NEW findings
26	$Li_{6.25}PSe_{5.25}I_{0.75}$	argyrodite family		NEW findings
27	Li_7PSe_6	argyrodite family		NEW findings
28	$Li_{6.5}PSe_{5.5}Br_{0.5}$	argyrodite family		NEW findings

Figure 1 Workflow of two model studies: fast Li-ion conductor and OER electrocatalyst. Model Study 1 trains on ionic conductivity data for 887 materials from the AtomWork-Battery database[22] and screens 7,705 Li-containing compounds, with

top-ranked candidates validated through simulations. Model Study 2 trains on overpotential data for 120 materials and screens 1,358,816 compounds, with top-ranked candidates validated experimentally. Both studies use periodic descriptors derived from the periodic table.

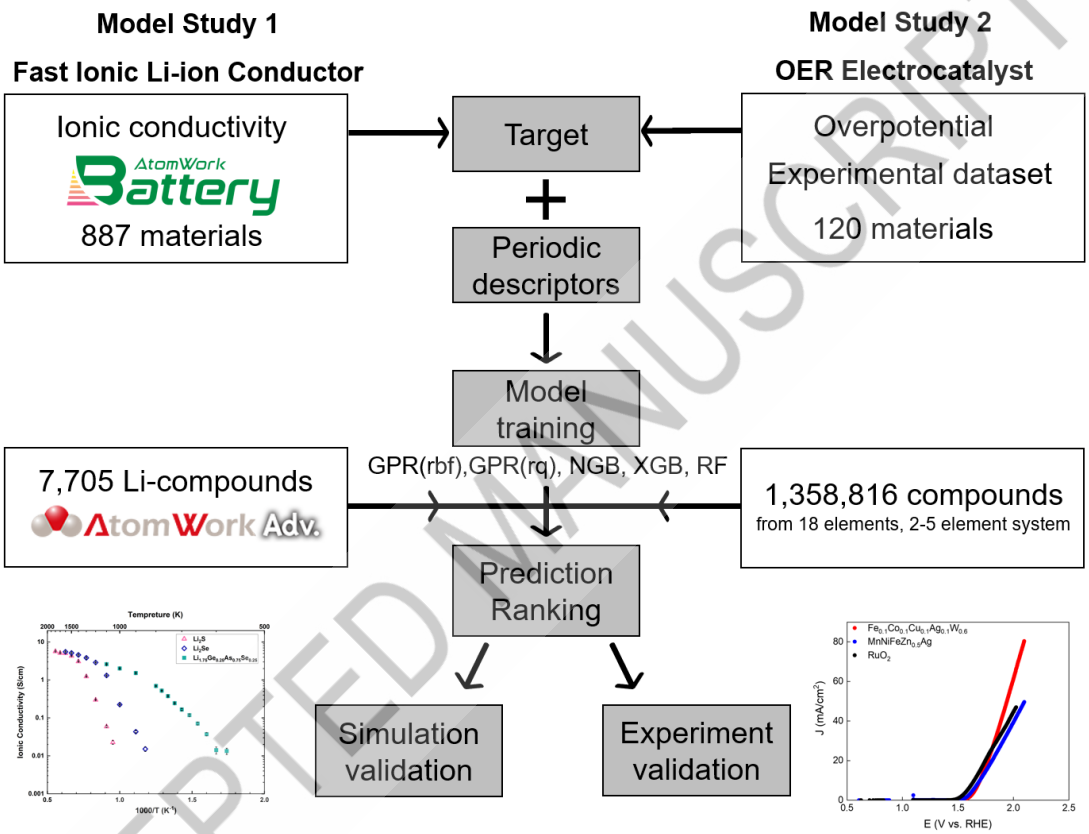


Figure 2 Generation of periodic descriptors from the periodic table, demonstrated using the chemical formula $Li_3PO_{3.56}N_{0.44}$. Each group in the periodic table is assigned two

values: a group coefficient and a group property. The group coefficient is calculated as the ratio of the total number of atoms in a group to the total number of atoms in the formula. The group property is determined as the weighted average of the atomic properties of the elements in the group, weighted by their coefficients. For this example, the atomic property used is the atomic number.

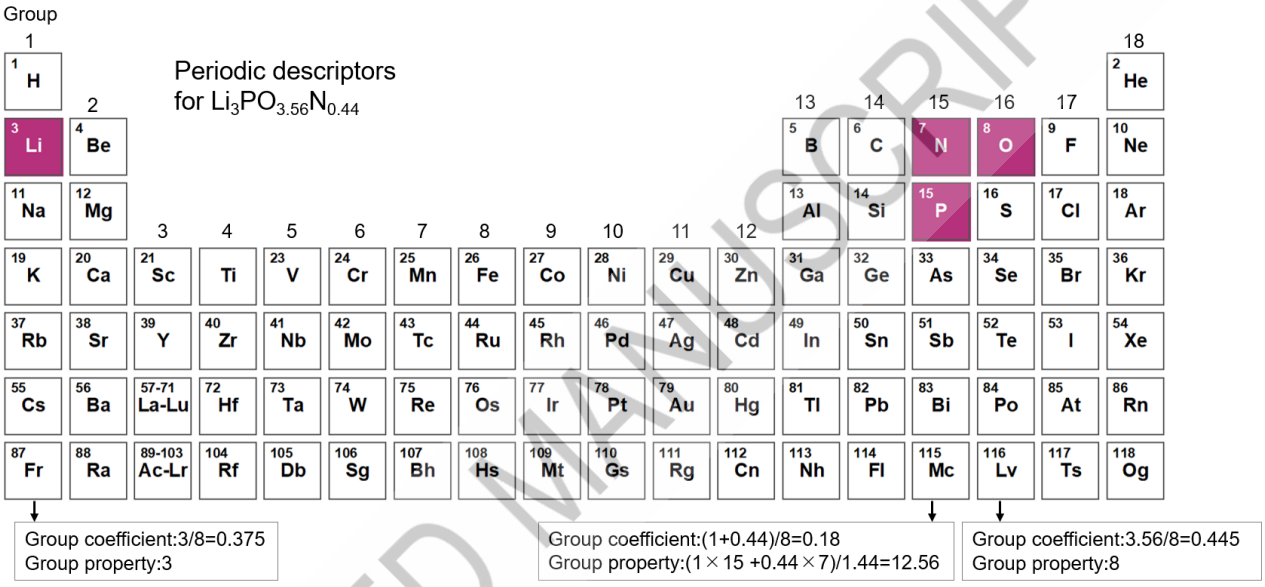


Figure 3 Prediction performance of periodic descriptors using different atomic properties for ionic conductivities. R^2 is the average with the standard deviation of the five models.

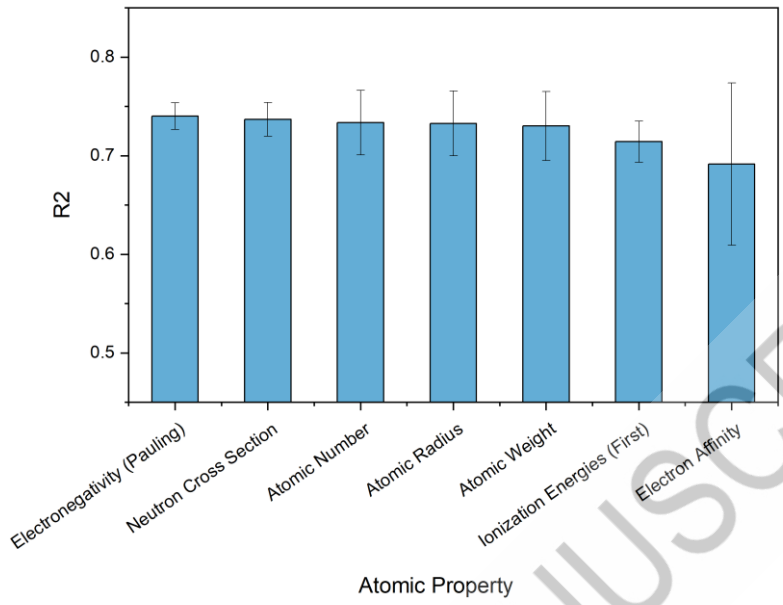


Figure 4 (a) Histogram of $\log_{10}$(ionic conductivity) in the Li-ion conductor dataset. (b) Structure type distribution, highlighting the prevalence of garnet, NASICON, and other structure types.

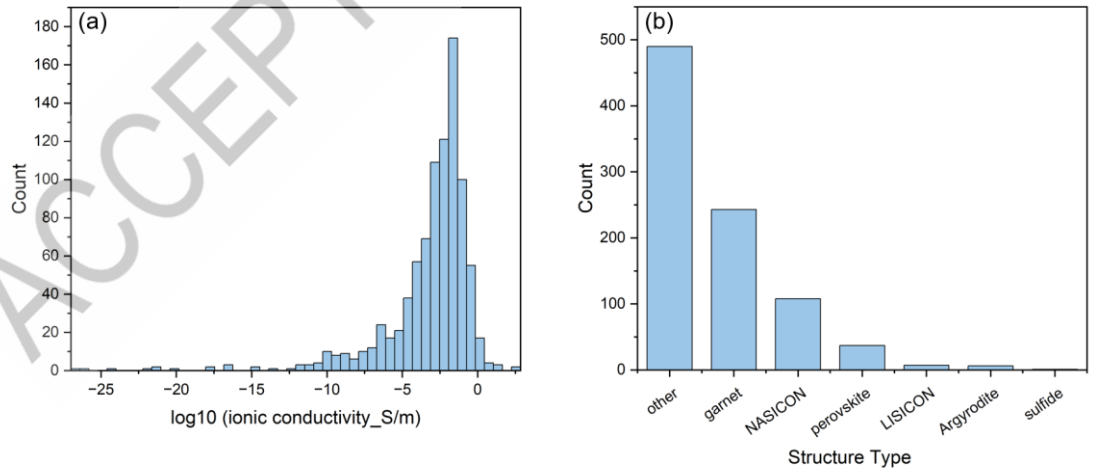


Figure 5 Parity plot of predicted vs. observed $\log_{10}$(ionic conductivity_S/m) using the

Random Forest model, with a cross-validated R^2 of 0.743. The results show reliable agreement across cross-validation folds.

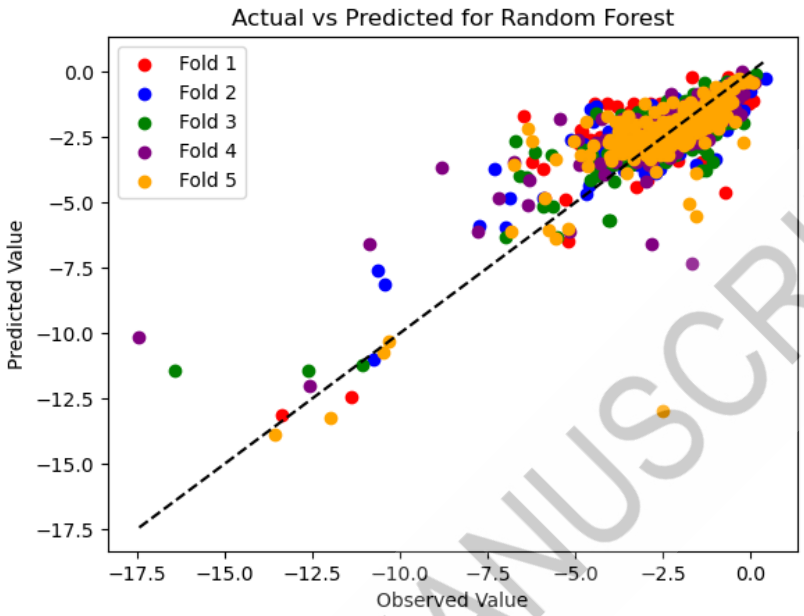


Figure 6 (a) Occurrence of elements in the 28 candidate materials. (b) Validation of the 28 candidate materials. The material numbers (percentages) are shown under the categories.

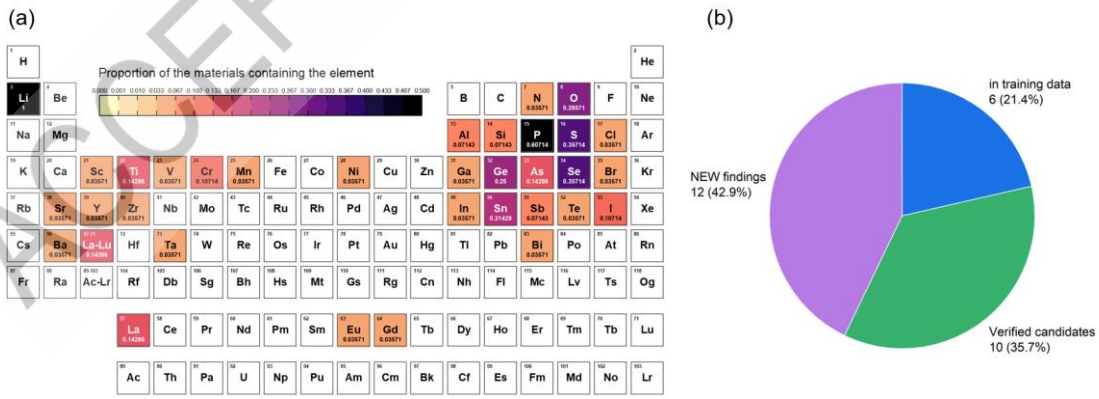


Figure 7 Ionic conductivity extrapolated to 300 K for $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ (4×4×4

and 2×2×2 supercells). Extrapolation using simulated data points above 900 K predicts a high conductivity of 4.88 mS/cm, whereas using data from 575 K to 800 K results in a lower extrapolated conductivity of 8.09 × 10^{-6} mS/cm. The star symbols indicate the extrapolated conductivity points, illustrating the impact of temperature-dependent transitions on ionic conductivity predictions.

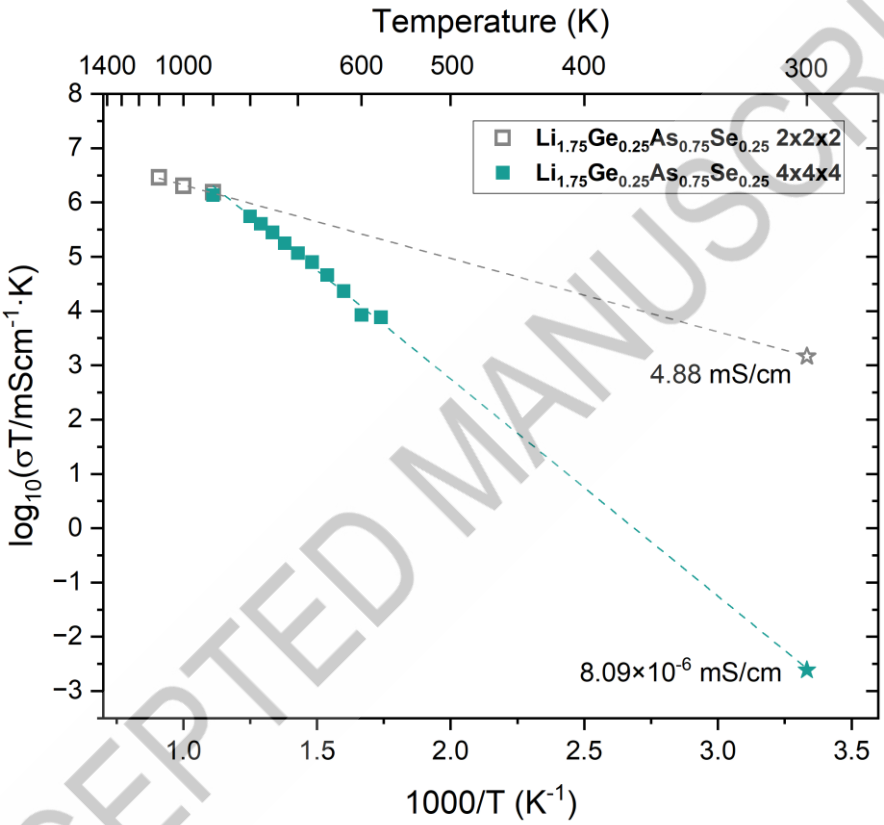


Figure 8 Simulated ionic conductivities of anti-fluorites (4×4×4 and 2×2×2 supercells).

The data compares $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ with Li_2S and Li_2Se. The high ionic conductivity of $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ at lower temperatures highlights its potential for applications at moderate temperatures. The black line for $Li_{1.75}Ge_{0.25}As_{0.75}Se_{0.25}$ indicates the slope differences for easier visualization.

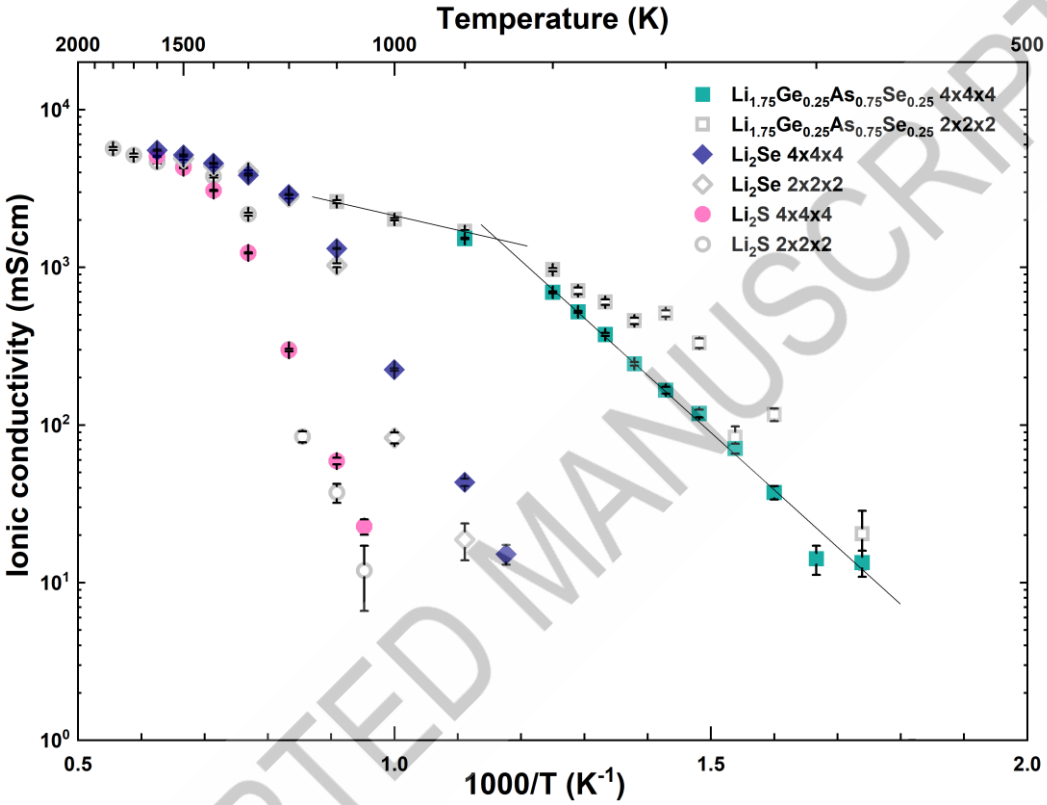


Figure 9 (a) Histogram of overpotential values for the electrocatalyst dataset. The

majority of values are centered between 400 and 800 mV. A subset of 37 samples with overpotentials recorded as 1000 mV are excluded from the plot. (b) Parity plot of observed vs. predicted overpotential (mV) using GPR (rbf) model. The model achieved the highest test R^2 of 0.635 among all models.

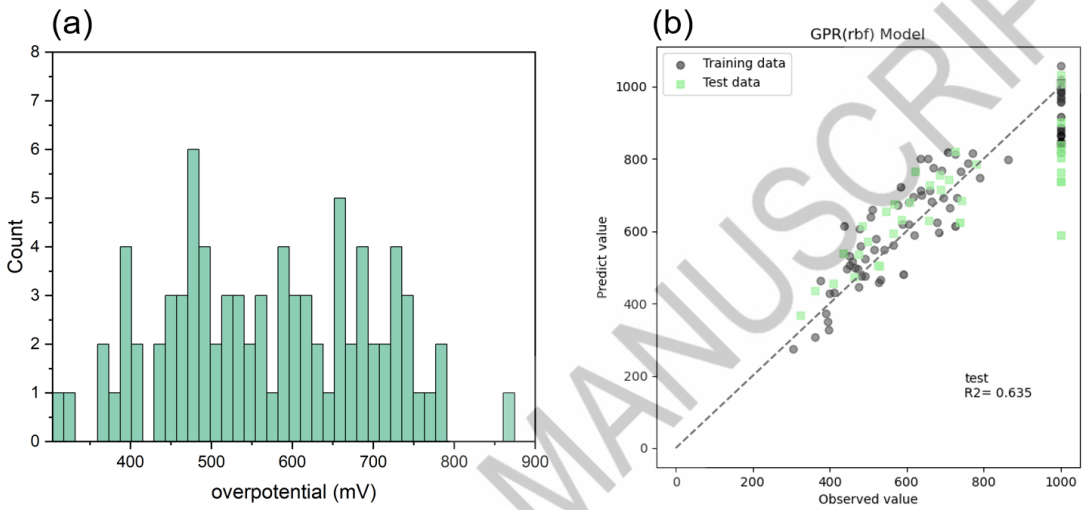
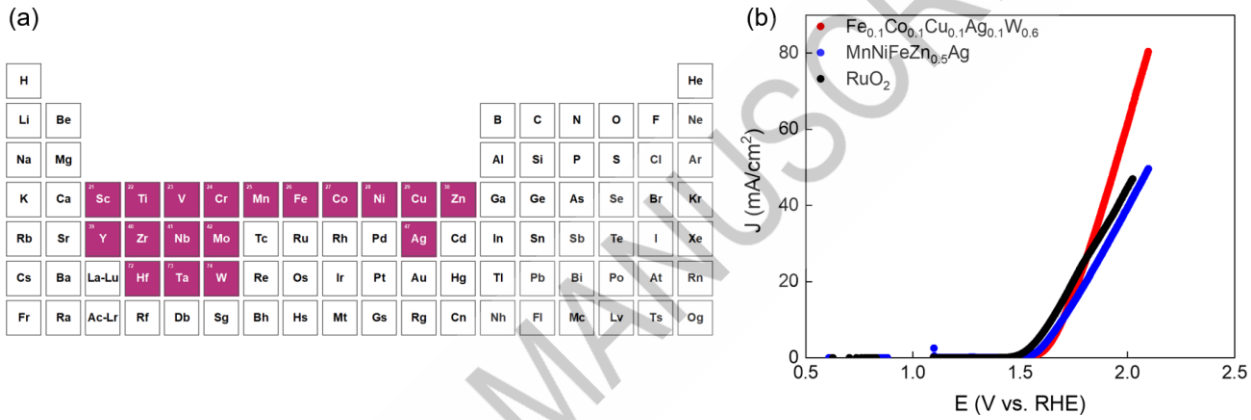
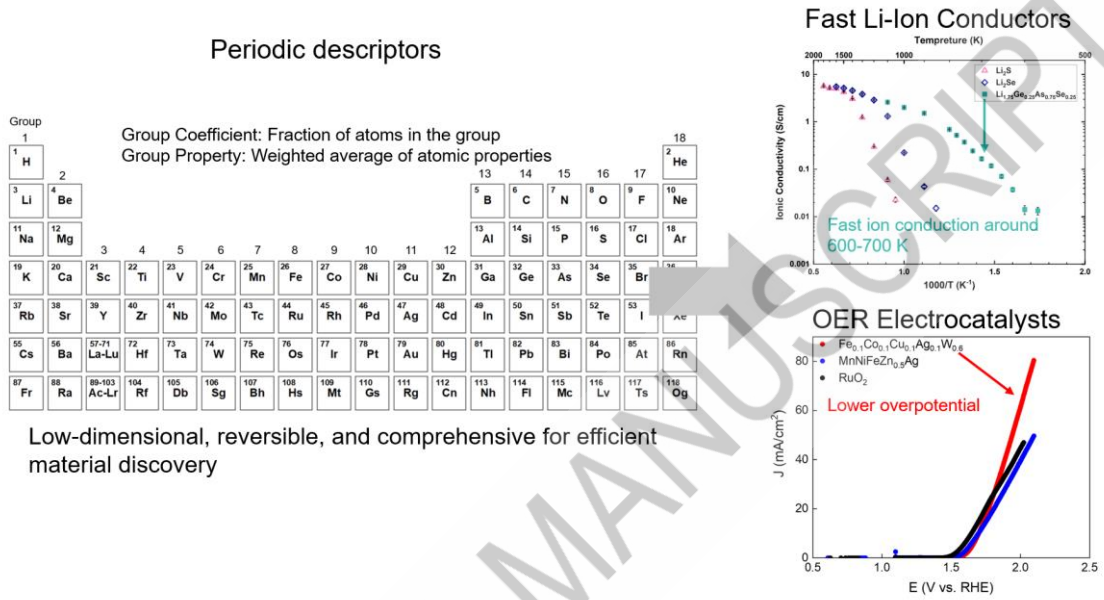


Figure 10 (a) The 18 elements used for generating materials for OER electrocatalysts prediction, highlighted on the periodic table. (b)The current density against voltage of the electrocatalysts. The predicted candidate, $Fe_{0.1}Co_{0.1}Cu_{0.1}Ag_{0.1}W_{0.6}$ oxide, shows a lower operational voltage at the current density of 40 mA/cm^2 compared to benchmarks such as RuO_2 and $MnNiFeZn_{0.5}Ag$ oxide.



Graphical abstract

Periodic descriptors simplify material discovery, validated in fast Li-ion conductors and OER electrocatalysts, accelerating progress toward electrochemical materials.



Impact Statement:

This work introduces a periodic table-based compositional descriptor that enables efficient, structure-independent material discovery, accelerating the identification of high-performance Li-ion conductors and PGM-free electrocatalysts, driving innovation in sustainable energy applications.

Supplementary

Periodic Table-Based Compositional Descriptors for Accelerating Electrochemical Material Discovery: Li-Ion Conductors and Oxygen Evolution Electrocatalysts

Yen-Ju Wu[1], Yibin Xu[1,*], Lei Fang[1], Wenqin Peng[2], Ken Sakaushi[2], Meiqi Zhang[1], Masao Arai[1], Yukinori Koyama[1]

[1] Center for Basic Research on Materials, National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

[2] Research Center for Energy and Environmental Materials, National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

To provide insight into the contribution of each periodic descriptor, we extracted feature importance rankings from Random Forest and XGBoost models trained on the Li-ion conductor dataset. The plots (Figure S1) display the top 20 descriptors as ranked by each model. Both models identified descriptors associated with Group 12 (e.g., coef_12, prop_12) and Group 16 (e.g., prop_16, coef_16) as highly important, suggesting that certain element groups are more relevant for predicting ionic conductivity within the compositional space studied. This further validates the chemical interpretability of the proposed periodic descriptor framework.

Figure S1 Top 20 feature importance rankings derived from (a) Random Forest and (b) XGBoost models for the Li-ion conductor dataset. Group 12 and Group 16 descriptors were repeatedly identified as dominant contributors.

Figure S1(a) Feature importance for Random Forest

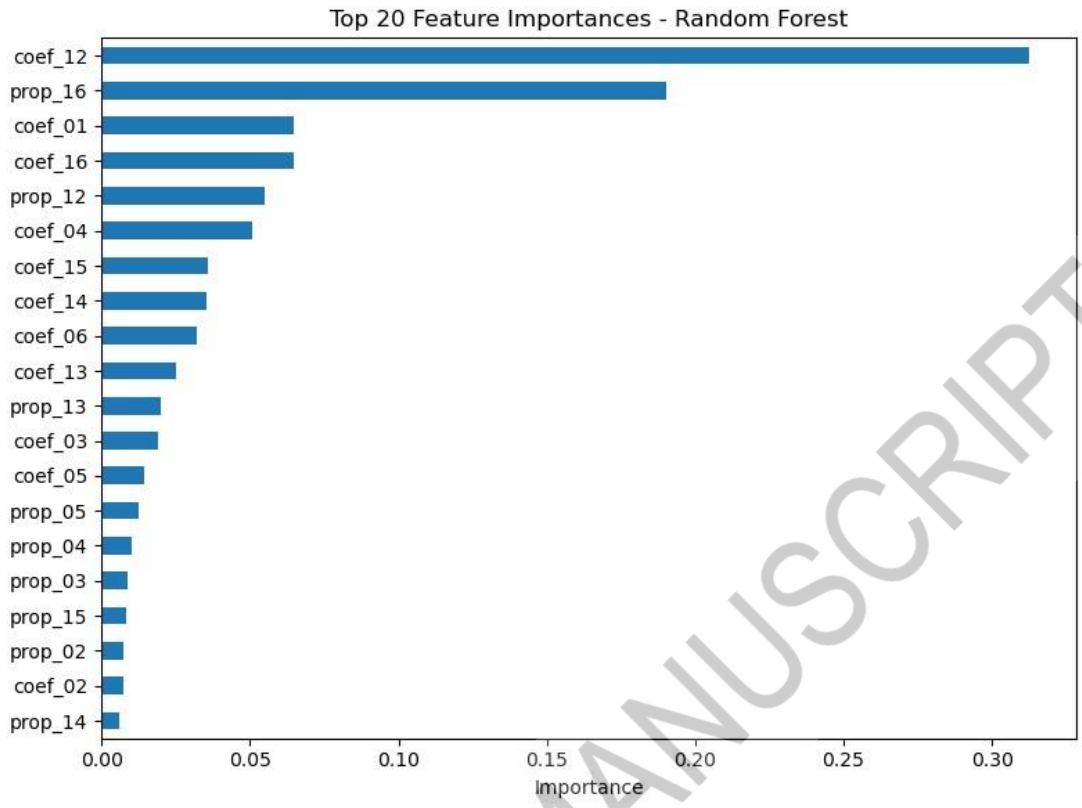


Figure S1(b) Feature importance for XGBoost (XGB)

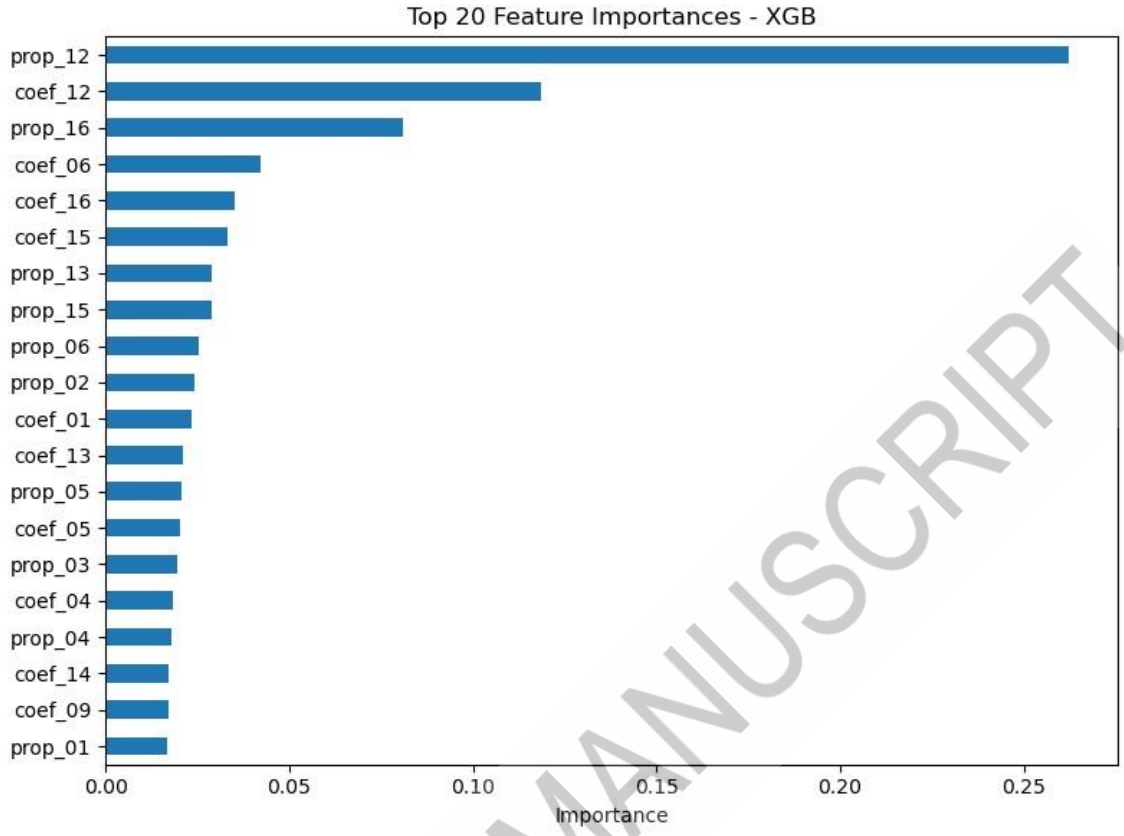
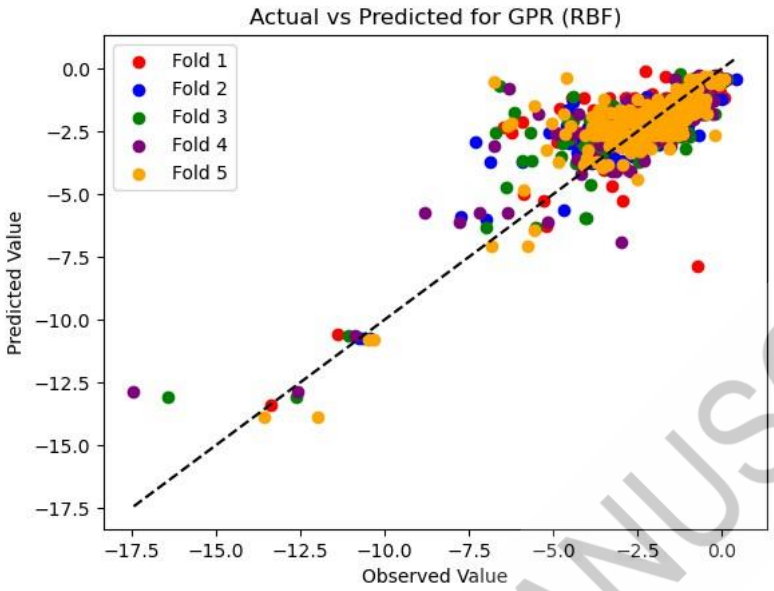
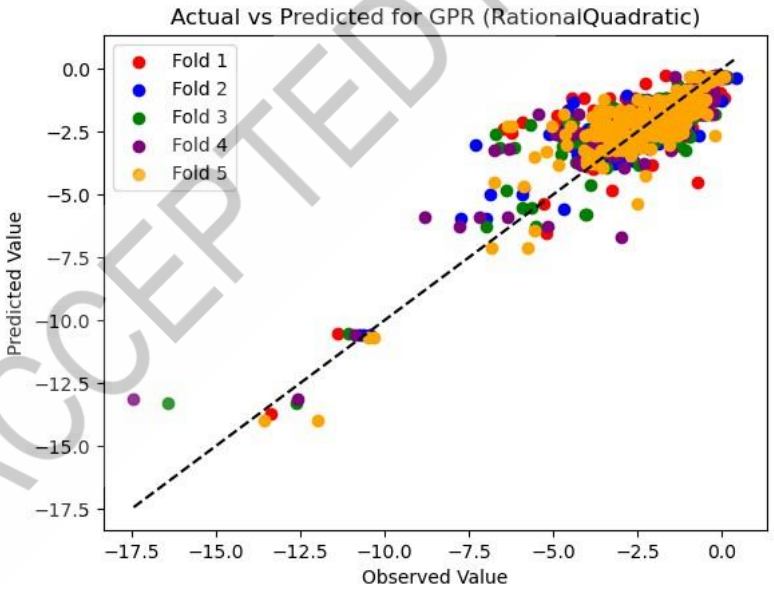


Figure S2 Parity plots for Li-ion conductivity prediction of $\log_{10}$(ionic conductivity_S/m) using five-fold cross-validation across:

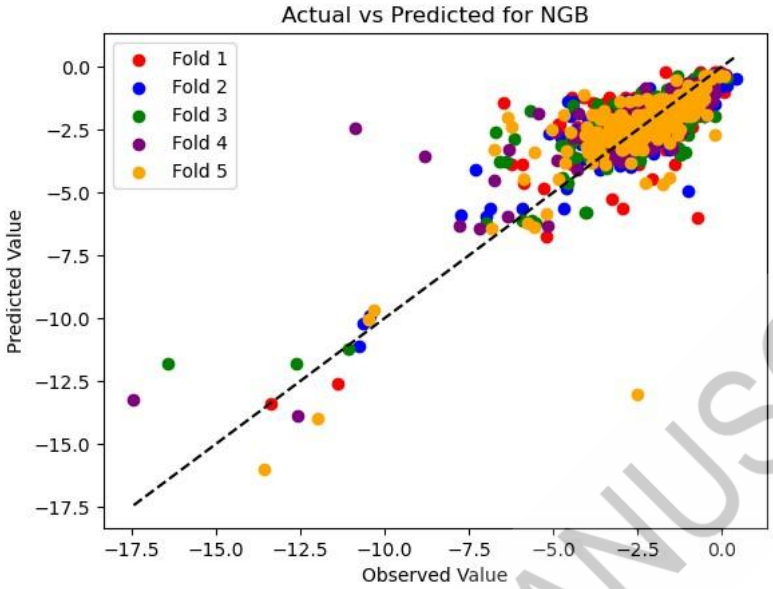
(a) GPR (RBF), R^2 = 0.686



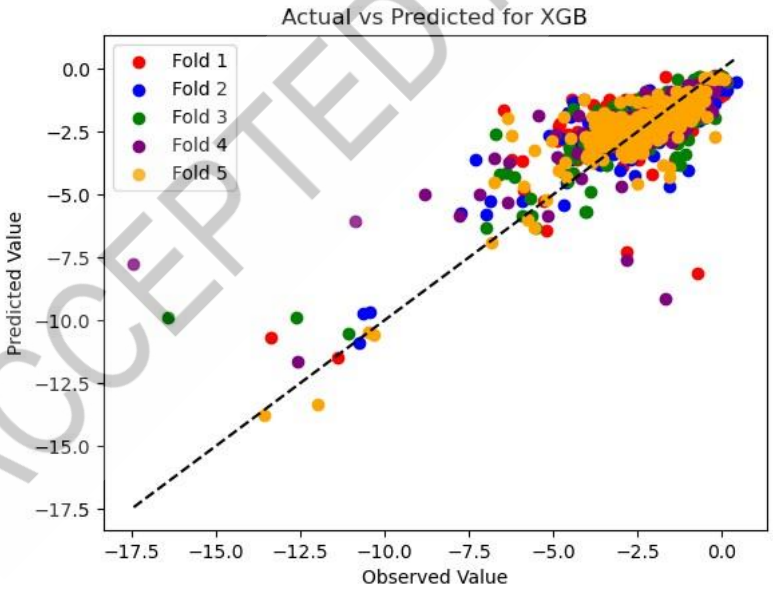
(b) GPR (Rational Quadratic), R^2 = 0.718



(c) NGBoost (NGB), R^2 = 0.727



(d) XGBoost (XGB), R^2 = 0.731



(e) SVR, $R^2 = 0.735$

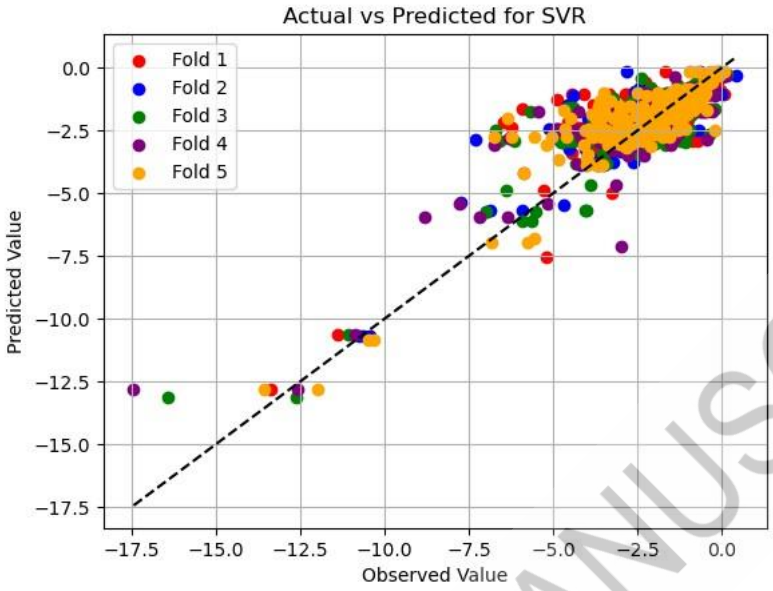


Figure S3 Parity plots for overpotential (mV) prediction using: (a) GPR(rq), R^2 = 0.632, (b) NGB, R^2 = 0.610, (c) XGB, R^2 = 0.621, and (d) Random Forest (RF), R^2 = 0.579