A Dynamic Database of Solid-State Electrolyte (DDSE) Picturing All-Solid-State Batteries

Fangling Yang\textsuperscript{a,b}, Egon Campos dos Santos\textsuperscript{a,*,} Xue Jia\textsuperscript{a}, Ryuhei Sato\textsuperscript{a}, Kazuaki Kisu\textsuperscript{c}, Hashimoto Yusuke\textsuperscript{d}, Shin-ichi Orimo,\textsuperscript{a,c} and Hao Li\textsuperscript{a,*}

\textsuperscript{a} Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, Sendai 980-8577, Japan

\textsuperscript{b} Key Laboratory of Radiation Physics and Technology, Ministry of Education, College of Physics, Sichuan University, Chengdu 610065, People’s Republic of China

\textsuperscript{c} Institute for Materials Research (IMR), Tohoku University, Sendai 980-8577, Japan

\textsuperscript{d} Tohoku Forum for Creativity, Tohoku University, Sendai 980-8577, Japan

* Corresponding Authors:

egoncs@gmail.com (E. C. S.)

li.hao.b8@tohoku.ac.jp (H.L.)
Abstract

All-solid-state batteries (ASSBs) are a class of safer and higher-energy-density materials compared to conventional devices, from which solid-state electrolytes (SSEs) are their essential components. To date, investigations to search for high ion-conducting solid-state electrolytes have attracted broad concern. However, obtaining SSEs with high ionic conductivity is challenging due to the complex structural information and the less-explored structure-performance relationship. To provide a solution to these challenges, developing a database containing typical SSEs from available experimental reports would be a new avenue to understand the structure-performance relationships and find out new design guidelines for reasonable SSEs. Herein, a dynamic experimental database containing >600 materials was developed in a wide range of temperatures (132.40 to 1261.60 K), including mono- and divalent cations (e.g., Li\(^+\), Na\(^+\), K\(^+\), Ag\(^+\), Ca\(^{2+}\), Mg\(^{2+}\), and Zn\(^{2+}\)) and various types of anions (e.g., halide, hydride, sulfide, and oxide). Data-mining was conducted to explore the relationships among different variates (e.g., transport ion, composition, activation energy, and conductivity). Overall, we expect that this database can provide essential guidelines for the design and development of high-performance SSEs in ASSB applications. This database is dynamically updated, which can be accessed via our open-source online system.

**Keywords:** solid-state electrolyte (SSE); all-solid-state battery (ASSB); ionic conductivity; dynamic database; machine learning.

**Graphical Abstract**
Introduction

As an energy storage devise, battery has been widely employed in portable electric devices, electric vehicles, and harvesting energy from renewable energy sources.\textsuperscript{1-2} However, a key problem of some conventional batteries is that they employ organic solvents as the battery electrolytes, which would lead to potential leakage and explosion.\textsuperscript{3-4} With the increasing demands for energy storage, developing safer and lower-cost batteries are urgently needed.\textsuperscript{5-7}

In the recent decade, all-solid-state batteries (ASSBs) have received considerable attention due to their excellent security, energy density, and thermal stability.\textsuperscript{8-11} Due to these benefits, it can be a promising alternative to circumvent these issues by employing solid-state electrolytes (SSEs).\textsuperscript{12-14} However, one of the key challenges of developing industry-level SSEs is the limited ionic conductivity (i.e., sluggish cation diffusion kinetics) at room temperature.\textsuperscript{15-16} With the recent intensive development of SSEs, ionic conductivities reaching the magnitudes of 10\textsuperscript{-4} - 10\textsuperscript{-2} S/cm were reported by some literatures,\textsuperscript{17-19} which are close to those of some liquid electrolytes.\textsuperscript{20-24} Therefore, developing high ionic conductivity SSEs is feasible and promising to improve battery techniques. So far, plenty of sulfide-based electrolytes have shown high ionic conductivities at room temperature, such as Li\textsubscript{2}S–P\textsubscript{2}S\textsubscript{5} and Li\textsubscript{2}S–P\textsubscript{2}S\textsubscript{5}–MS\textsubscript{x}.\textsuperscript{18, 25-28} However, one shortcoming of these materials is their high sensitivity to H\textsubscript{2}O and air.\textsuperscript{13, 29} Oxide-based solid electrolytes, such as garnet,\textsuperscript{30-32} LISICON (Li\textsuperscript{+} super ionic conductor),\textsuperscript{33-34} NASICON (Na\textsuperscript{+} super ionic conductor),\textsuperscript{35-36} and perovskite-type materials,\textsuperscript{37-38} generally have higher stability but lower conductivity than sulfide-based electrolytes.\textsuperscript{39-41} Therefore, it is particularly important to search for novel SSEs with promising conductivity and stability under operating conditions. However, the high complexity of the materials space makes the exploration and screening of high-performance SSEs difficult - conventional experimental methods and theoretical computations for exhaustive material search are time- and resource-consuming.\textsuperscript{42-44}

To date, machine learning (ML) techniques can extract useful information from available databases using different algorithms and provide predictions.\textsuperscript{45-48} In the recent decade, more and more successful examples certified that ML would be a powerful tool for seeking suitable functional materials with minimum time consumption.\textsuperscript{49-52} However, the accuracy of a ML model depends significantly on the quality of training data, while a state-of-the-art dataset of ionic conductivity with diverse conducting
cations, was rarely reported. Without a comprehensive, dynamically updated database, it would be hard to explore the structure-performance relationship of SSEs, which would hamper the design of new SSEs with promising performance.

To address the challenges above, a SSEs database (namely, the Dynamic Database of Solid-State Electrolyte, DDSE) was developed with the ionic conductivity of 678 typical materials from available experimental reports (from 291 resources). By searching literatures containing the keywords related to solid-state electrolyte and the dynamic data mining from typical experimental literatures reported to date, this database collects the SSEs with various mono- and divalent cations (e.g., Li\(^+\), Na\(^+\), K\(^+\), Ag\(^+\), Ca\(^{2+}\), Mg\(^{2+}\), and Zn\(^{2+}\)) and advanced anions (e.g., halide, hydride, sulfide, and oxide) from experimental reports. To testify the quality of the database, the correlations among various features in the database were investigated via data mining. Based on the current database, high accuracies in predicting conductivity can be obtained by ML, indicating a promising potential to research alternative materials through the database. We expect that this comprehensive database can contribute to the future screening and design of high-performance SSEs.

**Results and Discussion**

As of May 2023, a database with 5,430 performance records (from 678 reported materials) was built, containing the solid-state ionic conductivity (\(\sigma\)) of the materials with various chemical compositions and the temperature ranging from 132.40 to 1261.60 K, with the activation energies (\(E_a\)) fitted from the Arrhenius plot of \(\ln(\sigma T)\) vs. \(1000/T\). In the database, there are seven (i.e., Li\(^+\), Na\(^+\), K\(^+\), Ag\(^+\), Ca\(^{2+}\), Mg\(^{2+}\), and Zn\(^{2+}\)) species of conducting cations involving mono- and divalent materials. In this database, each composition was classified according to its structural characteristics. **Figure 1** shows the distribution of materials between the year of 1978 and 2023, demonstrating the increasing interest in researching SSEs over the past four decades.
Ionic conductivity is a crucial parameter to evaluate the performance of SSEs because it is essential to produce ASSBs with high power/rate performance. The distribution of ionic conductivity in this database is shown in Figure 2. It can be seen that more than half of the materials in this database are monovalent due to the larger research population in monovalent electrolytes in the materials community. The log_{10}(\sigma) values of monovalent electrolytes fall between -12 and 1 log_{10}(S/cm). For divalent materials, the values concentrate on the range from -12 to -1 log_{10}(S/cm). The materials with the log_{10}(\sigma) focusing on -5 to -2 log_{10}(S/cm) comprise >60% of all materials in the database. Note that to ensure the quality of the database, those literature with ambiguous values or identified incorrect results were excluded.
Figure 2. Distribution of conductivity for mono- and divalent electrolyte materials in the DDSE database. These conductivities have been measured in the temperature ranging from 132.40 to 1261.60 K.

To obtain more information from the database, each complex was assigned with a label based on the structure of materials (material class). It is known that temperature can play a key role in determining the ionic conductivity. Therefore, the trends among $\log_{10}(\sigma)$ ($x$-axis), classification ($y$-axis), and temperature ($z$-axis) are shown in Figure 3, where the valence of the cations are distinguished by different colors. The distribution of $\log_{10}(\sigma)$ for each material class was displayed on the $xy$-plane.

From the statistical analysis, the monovalent compositions can distribute on all the material classes and many of them present excellent performance, suggesting that monovalent cation is compatible well with different anions. In divalent materials, the closo- (i.e., cage-like) and hydride-type materials have attracted intense concern especially in the recent decade,\textsuperscript{53-55} demonstrating that these types of electrolytes could be promising in ASSB applications. Compared to monovalent SSEs, divalent SSEs
generally have a lower mobility. This is partly due to the stronger electrostatic interactions between the cation and anion in divalent SSEs. Noted that several divalent materials show better performance than monovalent complexes near room temperature, suggesting that obtaining divalent high-performance SSEs could be promising. Currently, divalent materials are less reported in the experimental literature, preventing us from obtaining more information for a larger-scale understanding. Therefore, in future works, it would be necessary to expand the search in the unexplored space of divalent materials by experiments and predict novel divalent materials with high conductivity using appropriate regression models based on available data.

**Figure 3.** Distribution of the materials in the DDSE database as a function of material class, ionic conductivity, and temperature. The gray and cyan points in $xy$ and $xz$ planes represent the mono- and divalent materials, respectively.

Projecting **Figure 3** into the $xz$-plane ($\log_{10}(\sigma)$ and temperature), the relationship between temperature and $\log_{10}(\sigma)$ can be analyzed. A large number of materials with the $\log_{10}(\sigma)$ values between -5 and -2 $\log_{10}(\text{S/cm})$ can be found in the temperature from 300 to 400 K. Meanwhile, high conductivity can also be found between 400 and 600 K for some monovalent materials like the NASICON type. For divalent materials, the $\log_{10}(\sigma)$ values between 300 and 400 K are more dispersive than monovalent
complexes, ranging from $-12$ to $-3 \log_{10}(S/cm)$. And some of the divalent materials like phosphates could present an excellent value of $\log_{10}(\sigma)$ over 700 K. Closo- and hydride- type materials can present high conductivity under 400 K.

In the DDSE, the distribution of $\log_{10}(\sigma)$ for each conducting cation is shown in Figure 4. The number of Li-based materials is the highest among all the materials, with the $\log_{10}(\sigma)$ values focusing on ranging from $-4$ to $-2 \log_{10}(S/cm)$. Similar to Li-based materials, Na-based materials generally have the $\log_{10}(\sigma)$ values between $-3$ and $-1 \log_{10}(S/cm)$. The distributions with long tails are observed for Ca-, Mg-, and Zn-based electrolyte materials, which generally have lower $\log_{10}(\sigma)$ because of the strong Coulomb interaction between divalent cations and the counter-anions. However, divalent cations present higher theoretical volumetric capacity than monovalent metals due to the larger charge per ion. Moreover, plenty of materials with the $\log_{10}(S/cm)$ values between $-8$ and $-3 \log_{10}(S/cm)$ can be found from the database, indicating their promising performance as SSEs.

Figure 4. Performance of the cation conductivity [in the unit of $\log_{10}(\sigma)$] of the SSE materials summarized based on the DDSE database.
A webpage of the DDSE (https://ssbed-ssbed-database-gui-ntj1tz.streamlit.app/; DOI: doi.org/10.50974/00137195) was developed via Streamlit, with its interface shown in Figure 5. This online interface can visualize the relationships among various materials’ properties (e.g., conductivity vs. conducting ions, classification vs. $E_a$, conductivity vs. temperature, and so on) in the database, and users can choose the interested ranges on the plot. More valuable information can be explored according to users’ input on the left panel. Meanwhile, the DDSE can display the properties of a separate compound and the corresponding literature’s source.

Figure 5 Example of the online interface of the latest version of DDSE. The left panel interacts with the users and the right panel shows the corresponding outputs based on the users’ input.

To test the potential applications of ML based on the DDSE, we structured a series of features to describe these materials and used an advanced ML method, the gradient boosting regression (GBR) in
the scikit-learn package\textsuperscript{,57} to model the ionic conductivity based on the current version of the database (training set: testing set = 4:1). All features considered as the ML inputs were acquired using the following equations:

\begin{align}
\bar{P} &= \sum_{i=1}^{n} m_i P_i \\

P_{\text{mean}} &= \frac{1}{n} \sum_{i=1}^{n} N_i P_i \\

P_{\text{var}} &= \frac{1}{n} \sum_{i=1}^{n} (N_i P_i - P_{\text{mean}})^2 \\

\delta P &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} m_i (1 - P_i/\bar{P})^2}
\end{align}

where $m_i$, $P_i$, and $N_i$ are the percentage of atomic mole, elemental property, and the number of the $i^{\text{th}}$ element, respectively. The molar mean and mismatch values of the elemental properties of compounds were described by $\bar{P}$ and $\delta P$,\textsuperscript{58-59} respectively. $P_{\text{var}}$ and $P_{\text{mean}}$ represent the variance and mean value of the elemental properties, respectively.

As a result, after modeling with various combinations of inputs, the top eight most important features (with the Pearson correlation coefficient\textsuperscript{60} lower than 0.9) were acquired to present the ML model, including the Period Number (Mismatch) (\textit{i.e.}, the mismatch value of the period number among the elements in the composition through Eq. 4\textsuperscript{59}; note: the mismatch value of other features can be obtained by the identical equation with different atomic properties), T/Kelvin (temperature with the Kelvin scale), Electron Affinity (Molar Mean) (\textit{i.e.}, molar mean value of electron affinity calculated by Eq. 1), Atomic Number (Mismatch) (\textit{i.e.}, the mismatch value of the atomic number), Dipole Polarizability (Ratio $m/n$) (\textit{i.e.}, the total dipole polarizability of metal divided by that of non-metal), Electronegativity (Pauling) (Var) (\textit{i.e.}, the variance value of Pauling electronegativity among the atoms in the composition), vdW Radius (Mismatch) (\textit{i.e.}, the mismatch value of van der Waals radius), and Mendeleev Number (Mismatch) (\textit{i.e.}, the mismatch value of Mendeleev number\textsuperscript{61}). As shown in Figure 6, a high-level predictive accuracy can be achieved (based on 10-fold cross-validation) using the developed model for monovalent materials, with an $R^2$ value of 0.92 (Figure 6a-b). The optimized hyperparameters were fixed to perform training for other material groups. The divalent compounds can reach an $R^2$ of 0.86 (Figure 6e-f). We also used this ML model with the same hyperparameters for
regression analysis on all materials, including both the mono- and divalent materials, and obtained an \( R^2 \) of 0.91 (Figure 6i-j). Based on the above results, the importance of eight features was qualitatively analyzed via the SHapley Additive exPlanations (SHAP) package.\(^6\) The differences in the feature contribution are shown in Figure 6 (c, d, g, h, k, l), suggesting that these factors could play the key roles in determining the ionic conductivities in SSEs. These preliminary ML analyses show that building a database of SSEs is particularly important for analyzing the structure-performance relationships of SSEs and provide new predictive insights into this type of essential materials. In the next-step study, we will further focus on the precise ML modeling based on various state-of-the-art methods with optimized hyperparameters.

**Figure 6** ML modeling with the GBR algorithm and feature analysis based on the data of DDSE. (a, e, i) The results of training set of the model for (a) monovalent, (b) divalent, and (c) all materials in the database. (b, f, j) The results of testing set based on (b) monovalent, (b) divalent, and (c) all materials in the database. (c, g, k) The derived feature importance based on the data of (c) monovalent, (g), divalent, and (k) all materials. (d, h, l) The calculated SHAP values (i.e., the impact on model output) of the features based on the data of (d) monovalent, (h) divalent, and (l) all data.

**Conclusion**

In summary, motivated by the current lack of comprehensive understandings of the structure-performance relationships of SSEs and the growing demand for searching high-performance SSE
materials, we performed data mining from the available typical experimental literatures published during the past four decades. The acquired data include the SSE materials consisting of mono- and divalent cation components (e.g., Li⁺, Na⁺, K⁺, Ag⁺, Ca²⁺, Mg²⁺, and Zn²⁺) and various types of anions (e.g., halide, hydride, sulfide, and oxide). A dynamically updated database (namely, the DDSE) with a user-friendly online interface was developed. The performances (i.e., ionic conductivity and ionic diffusion activation energy) of SSE materials in a wide range of temperatures were collected. Based on the statistical analysis from this database, SSEs with mono- and divalent cations present different distributions of conductivity, with the log_{10}(σ) values of monovalent concentrating on the range of -5 to -2 log_{10}(S/cm). Divalent materials with high conductivity have gradually reported in recent years, suggesting an extensive space for researching excellent materials with a higher energy density. The ML and feature analyses indicate that our database provides an opportunity to explore diverse SSE materials and to better picture the structure-performance relationships. We expect this dynamic database can accelerate the discovery of novel materials with improved electrochemical performance and guide the design of optimal solid-state electrolytes.

Data Availability

The webpage of DDSE is: https://hdl.handle.net/10097/00137195. The registered DOI for the database is: https://doi.org/10.50974/00137195, which will be automatically activated upon the publication of this article.

Acknowledgement

This project was supported by the Ensemble Grant for Early Career Researchers 2022 and the 2023 Ensemble Continuation Grant of Tohoku University, the Hirose Foundation, the Iwatani Naoji Foundation, and the AIMR Fusion Research Grant. It was also supported by JSPS KAKENHI Nos. JP23K13599, JP23K13703, JP22H01803, and JP18H05513. We acknowledge the Center for Computational Materials Science, Institute for Materials Research, Tohoku University for the use of MASAMUNE-IMR (Nos. 202212-SCKXX0204 and 202208-SCKXX-0212) and the Institute for Solid State Physics (ISSP) at the University of Tokyo for the use of their supercomputers. F.Y. acknowledges the China Scholarship Council (CSC) fund to pursue studies in Japan.
References

15. Yu, C.; Ganapathy, S.; De Klerk, N. J.; Roslon, I.; van Eck, E. R.; Kentgens, A. P.; Wagemaker,
M., Unravelling Li-Ion Transport from Picoseconds to Seconds: Bulk Versus Interfaces in an Argyrodite Li6ps5cl–Li2s All-Solid-State Li-Ion Battery. *Journal of the American Chemical Society* **2016**, *138*, 11192-11201.


26. Trevey, J. E.; Jung, Y. S.; Lee, S.-H., High Lithium Ion Conducting Li2s–Ges2–P2s5 Glass–


37. Rao, R. P.; Adams, S., Studies of Lithium Argyrodite Solid Electrolytes for All-Solid-State
38. Catti, M., Short-Range Order and Li$^+$ Ion Diffusion Mechanisms in Li$_{5}$La$_{9}$□$_{2}$TiO$_{3}$$_{16}$ (Llto). *Solid State Ionics* 2011, 183, 1-6.


60. Pearson, K., Contributions to the Mathematical Theory of Evolution. Philosophical Transactions
of the Royal Society of London. A 1894, 185, 71-110.
