Determination of stable proton configurations by black-box optimization using an Ising machine

Jianbo Lin[1,*], *Tomofumi Tada*[2,*], *Ai Koizumi*[1], *Masato Sumita*[3], *Koji Tsuda*[4,1,3], *and Ryo Tamura*[1,3,4,*]

[1]Center for Basic Research on Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

[2]Kyushu University Platform of Inter/Transdisciplinary Energy Research, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan.

[3]RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

[4]Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwa-no-ha, Kashiwa, Chiba 277-8561, Japan

KEYWORDS Ising machine, black-box optimization, density functional theory calculations, proton configuration, stable configuration

ABSTRACT

Stable proton configurations in solid-state materials are a prerequisite for the theoretical microscopic investigation of solid-state proton-conductive materials. However, a large number of initial atomistic configurations should be considered to find stable proton configurations, and relaxation calculations using the density functional theory approach are required for each initial configuration. Consequently, the determination of stable configurations is a difficult and time-consuming task. Furthermore, when the size of the simulation cells or the number of doped atoms increases, the number of initial configurations leads to a combinatorial explosion, rendering the computation infeasible. In this study, black-box optimization was combined with an Ising machine and density functional calculations to perform an efficient search for stable proton configurations. Scandium-doped barium zirconate, a typical high-proton conductive oxide, was selected as the model system. The Ising machine was able to rapidly select the initial atomistic configuration, ultimately leading to stable proton configurations after subsequent relaxation calculations. This optimization strategy should be able to solve various issues related to configuration optimization in solid-state materials, thereby promoting novel scientific discoveries.

INTRODUCTION

Protonic ceramic fuel cells (PCFCs) are expected to exhibit high power generation efficiencies. As previously reported, doping is a key process in achieving high proton conductivities in PCFC oxide electrolytes.[1–6] More specifically, doping with low-valence elements creates oxygen vacancies that maintain electrical neutrality, and these vacancies can subsequently be filled with hydroxy groups to introduce protons into the system. Previously, the relationship between the conductivity and the doping level has been studied experimentally.[7,8] However, to understand the mechanism of proton conductivity from a microscopic viewpoint, atomistic simulations based on molecular dynamics are essential,[9,10] and stable atomistic configurations are required to commence these simulations. At present, determination of the stable atomistic configurations of doped atoms and protons is a difficult and time-consuming task since the initial configurations of the doping atoms and protons can change depending on the size of the simulation cell and the number of doped atoms. In addition, relaxation calculations using the density functional theory (DFT) approach are required for each initial configuration, thereby rendering it impossible to perform relaxation calculations for all possible initial configurations and to find stable structures using such approaches.

Using machine learning to optimize atomistic configurations, Ju et al.[11] demonstrated that the configurations of silicon and germanium atoms in a lattice could be rapidly explored by means of Bayesian optimization to maximize or minimize phonon transport. Bayesian optimization is a method for optimizing black-box functions,[12] and is characterized by the use of a Gaussian process as a surrogate model.[13,14] Following the work of Ju et al., many structure search problems aimed at atomistic configurations have been solved using Bayesian optimization.[15–19] However, Bayesian optimization has a problem in that when the number of candidates causes a combinatorial

explosion, the time required for the selection of promising candidates increases exponentially because acquisition function calculations are required for all candidates. To overcome this limitation, Kitai et al. proposed a black-box optimization (BBO) algorithm using a quantum annealer known as the factorization machines with quantum annealing (FMQA) algorithm.[20] In the FMQA algorithm, the factorization machine (FM)[21] model is adopted as a surrogate model. Because the FM model is expressed by the Ising model, the lower-energy states of the FM can be effectively obtained using quantum annealers and Ising machines.[22] Indeed, this algorithm has been widely used to solve BBO problems in materials science and chemistry.[23–30] For example, the FMQA algorithm was adopted to optimize the atomic configurations of the magnesium and germanium ions in a magnetic tunnel junction structure,[31] and its efficiency was confirmed for small problems. Furthermore, its application in crystal structure prediction has also been considered.[32,33]

In this study, the FMQA algorithm combined with the Vienna Ab Initio Simulation Package (VASP) is employed to effectively search for stable proton configurations in PCFCs (**Figure 1**). In each optimization cycle, the FM model is trained to predict the total energy after the relaxation calculation from the initial proton configuration. By solving the FM model using an Ising machine, an initial proton configuration that leads to stable proton configurations after the relaxation calculation is suggested. For the suggested configuration, the total energy is calculated by relaxation calculations using VASP, and the amount of training data for the FM model is increased. Subsequently, in each cycle, a relaxation calculation must be performed using VASP to acquire the total energy with the required calculation accuracy. The optimization task is therefore to determine the stable proton configurations using as few cycles as possible. For the purpose of this study, scandium-doped barium zirconate,[34] which is a representative electrolyte candidate for

PCFCs, is employed as a model system to perform the stable proton configuration searches by combining FMQA and VASP. In addition, a graphics processing unit-based Ising machine is used to solve the FM model.

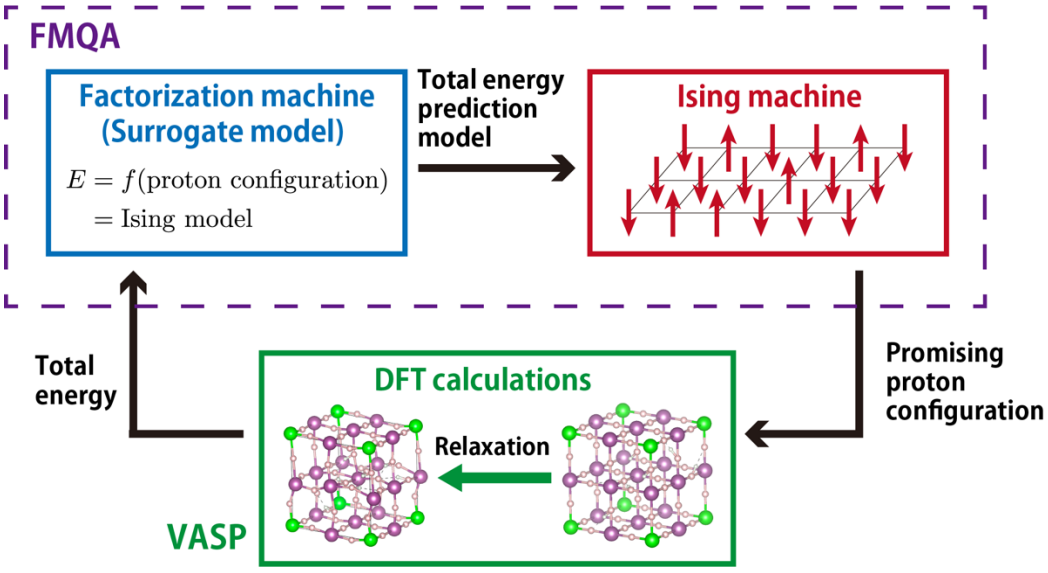

Figure 1 Optimization cycle for the BBO based on use of the FMQA algorithm and VASP to obtain a stable proton configuration. The atomistic configurations were drawn using VESTA.[35]

METHODS

Target material

As a model system for proton configuration optimization, scandium-doped barium zirconate was employed according to the crystal structure shown in **Figure 2(a)**. The mother compound was the $BaZrO_3$ perovskite, and fast proton conduction has previously been observed in its doped materials,[4,5,36] wherein a high proton conductivity was achieved by substituting zirconium with scandium.[34] Under these conditions, oxygen vacancies are generated, and hydroxy groups can be introduced into the system in the presence of water vapor to compensate for these vacancies. In the current study, eight initial proton positions are considered at each oxygen atom, as determined by DFT structural optimizations for several configurations, including those of the protons. Each proton position was directed to the second nearest oxygen atom to form a hydrogen bond (see Figure 2(b)). Thus, these proton positions can be used as the initial proton configurations for the relaxation calculations. A 2 × 2 × 2 supercell model was adopted to search for a stable configuration. Based on the number of scandium ions (x) in the supercell, the composition of the model was expressed as $Ba_8Zr_{8-x}Sc_xO_{24}H_x$, corresponding to the condition that all oxygen vacancies are compensated for by hydroxy groups; therefore, the number of introduced protons is the same as the number of scandium ions. When the doped positions of the scandium ions are fixed, the number of proton configurations is ${}_{24}C_x \times 8^x$, without considering symmetry. Thus, when the value of x increases, many proton configurations must be considered. Of course, when symmetry is considered, the number of independent initial proton configurations is reduced; however, a combinatorial explosion of the initial configurations occurs when larger systems are considered.

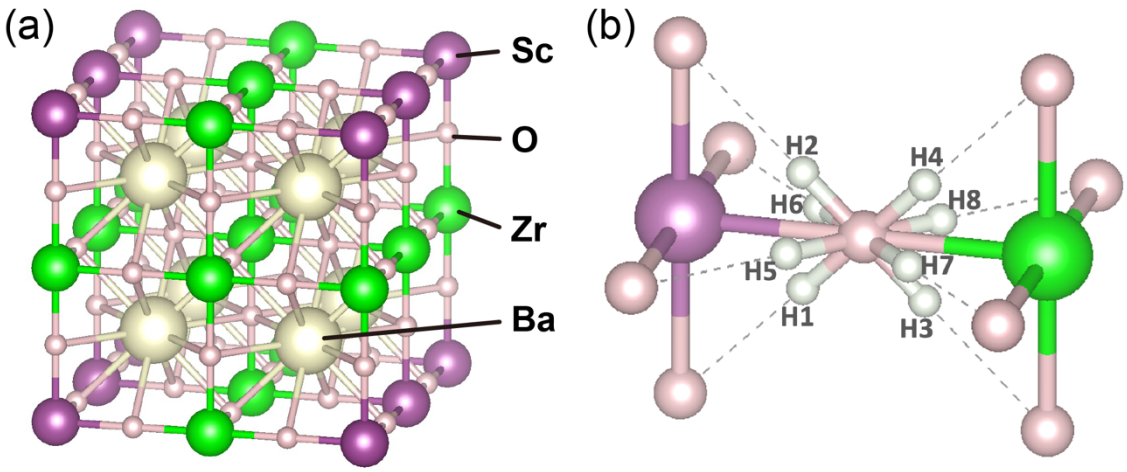

Figure 2. (a) Atomistic structure of the scandium-doped barium zirconate. (b) Eight stable proton positions. The atomistic configurations were drawn using VESTA.[35]

Structural relaxation based on DFT calculations

Based on one of the initial proton configurations, structural relaxation calculations were performed using VASP.[37–39] To effectively obtain accurate stable configurations, multiple-step structure relaxation is often performed.[40,41] Thus, for this study, three total energies were used as metrics to determine a stable configuration. The first is the total energy of the initial proton configuration without the relaxation calculation, which is expressed as E_0 and is known as the no-relaxation total energy. The second is the total energy with normal accuracy in the structural relaxation (E_1), and the third is the highly accurate total energy after structural relaxation (E_2). The detailed setup of the VASP calculations employed for multiple-step structural relaxation to obtain the total energy is summarized in Supplementary Note A.

Optimization of the proton configuration using the FMQA algorithm

In quantum annealers and Ising machines, only 0/1 variables can be used; thus, it is necessary to encode atomistic configurations into a bit string composed of 0/1 variables. Oxygen ions can take the form of protonated (i.e., hydroxy) or non-protonated species. For protonated oxygen ions, a proton is placed at any one of eight possible positions in the initial state for structural relaxation (Figure 2(b)). Thus, nine bits were used to express the position of a proton at each oxygen atom site, and a one-hot constraint was considered to ensure that no more than two protons were attached to the same oxygen. In total, the proton configuration can be expressed as 24 × 9 = 216 bits, while the bit string can be expressed as $\mathbf{q} = \{q_{ni}\}_{n=1,\ldots,24,i=1,\ldots,9}$ where $q_{ni} = 0$ or 1. The indices n and i represent the indices of the oxygen ions and their proton positions, respectively. For the proton position, $i = 1, \ldots, 8$ refers to the positions shown in Figure 2(b), while $i = 9$ indicates the absence of a proton. The optimization cycle employed to obtain a stable proton configuration (i.e., the lowest total energy) using the FMQA algorithm therefore involved four key steps. Firstly, M proton configurations were generated as the initial training dataset from ${}_{24}C_x \times 8^x$ configurations, which were used as the initial configurations for structural relaxation. Relaxation calculations were performed using VASP and the total energies were obtained. These data were used as the training data for the FM model, as shown in Figure 1. In step 2, the FM model was used as a surrogate model to predict the total energy from the initial proton configuration and was trained when bit string $\mathbf{q}$ was input. The values of the total energy predicted by the FM are as follows:

$$E(\mathbf{q}) = \sum_{n=1}^{24}\sum_{i=1}^{9} w_{ni}^{*} q_{ni} + \sum_{n=1}^{24}\sum_{i=1}^{9}\sum_{m=1}^{24}\sum_{j=1}^{9}\sum_{k=1}^{K} v_{nik}^{*} v_{mjk}^{*} q_{ni} q_{mj}\,, \qquad (1)$$

where the hyperparameter of K is fixed at a default value of 8.[21] The trained parameters are denoted w_{ni}^{*} and v_{nik}^{*}. In the third step, the FM is solved by the Ising machine under the following constraints:

$$\sum_{n=1}^{24} q_{n9} = 24 - x, \tag{2}$$

$$\sum_{i=1}^{9} q_{ni} = 1, \forall n. \tag{3}$$

The former constraint determines the number of hydrogen atoms, whereas the latter ensures that at most, one hydrogen atom is attached to one oxygen atom. For the purpose of this study, Fixstars Amplify AE[42] was used as the Ising machine. A single solution for bit string $\mathbf{q}^{*}$ that minimizes the FM model defined by Eq. (1) was selected using the Fixstars Amplify AE. If the selected $\mathbf{q}^{*}$ was already included in the training data, then the bit string was randomly generated under the constraints of Eqs. (2) and (3). Finally, in step four, DFT relaxation calculations were conducted using the initial proton configuration expressed by the selected $\mathbf{q}^{*}$, and the total energy was obtained. The number of training data points was increased to $M + 1$. Steps 2–4 were subsequently repeated as desired.

The optimization task was defined as finding a superior initial state for structural relaxation, denoted by bit string $\mathbf{q}$, which will lead to the minimum total energy after the relaxation calculations, using as few cycles as possible. Note that in the initial configurations of ${}_{24}C_x \times 8^x$, multiple configurations can be considered to have the same symmetry. Thus, from a symmetry consideration, in a case where the configuration suggested by $\mathbf{q}^{*}$ in the third step had already been

investigated in the previous cycles, the calculated total energy would have been used without the relaxation calculation, thereby reducing the computation time. However, for the ease of implementation, symmetry was not considered in the current study.

RESULTS AND DISCUSSION

Optimization performance for stable proton configurations

Initially, the optimization performance of the system was investigated using a fixed scandium ion configuration. A case was considered in which the number of scandium ions was two, i.e., $x = 2$. In this case, by considering the symmetry, three possible configurations exist for the scandium ions for a 2 × 2 × 2 supercell. By assigning each zirconium ion position in $BaZrO_3$ with a number between 1 and 8 (see **Figure 3(a)**), the scandium ion configurations were denoted as Sc@(1,2) (Figure 2(a)), Sc@(1,6), and Sc@(1,8), respectively. For each case, the number of possible proton configurations as the initial states for the relaxation calculations was defined as ${}_{24}C_x \times 8^x = 17{,}664$, without considering symmetry. For all configurations, DFT relaxation calculations were performed in advance using a Fugaku supercomputer, and the total energies of E_0, E_1, and E_2 were obtained with different accuracies. Thus, for $x = 2$, the optimal solutions for the stable proton configurations were known for the various relaxation steps, and these solutions were used to estimate the efficiency of the FMQA algorithm in finding stable proton configurations. In addition, during each optimization cycle, the total energy of the selected proton configuration can be accessed from the previously calculated total energies. Consequently, it is not necessary to perform an additional DFT calculation to estimate the efficiency.

To verify the BBO performance using the FMQA algorithm for exploring stable proton configurations, optimization was performed for cases where the total energies of E_0, E_1, and E_2 were targeted. The number of initial training data points was fixed at $M = 10$, and optimization runs were performed until the optimal solution was obtained. Ten runs were performed independently by changing the initial training data selections, and Figure 3(b) shows box plots of

the number of cycles required to obtain the most stable proton configuration for Sc@(1,2). Here, the most stable configurations are defined as those with total energies ranging from the lowest total energy up to 0.005 eV. For comparison, five random samplings were conducted, starting from each initial training dataset used in the FMQA optimizations; thus, 50 runs were performed. Figure S1 summarizes the results for obtained Sc@(1,6) and Sc@(1,8), wherein the most stable proton configurations were identified within 150 BBO cycles using the FMQA algorithm. Consequently, it was found that the FMQA outperformed the random sampling approach. Notably, when random sampling is employed, the number of cycles required to find stable proton configurations for E_0 is larger than those required for E_1 and E_2. This can be attributed to the fact that the number of most stable configurations for E_0 is smaller than those for E_1 and E_2. Furthermore, Figures 3(c) and 3(d) show the relationships among E_0, E_1, and E_2 for the same initial proton configurations. It can be seen that the order between E_0 and E_2 is largely interchanged when the values of E_2 are small, indicating that find stable proton configurations using E_0, i.e., without the use of a relaxation calculation, are not the true stable proton configurations using E_2. However, the order is comparable between E_1 and E_2, indicating that optimization using E_1 is sufficient for finding stable proton configurations in the present problem setting.

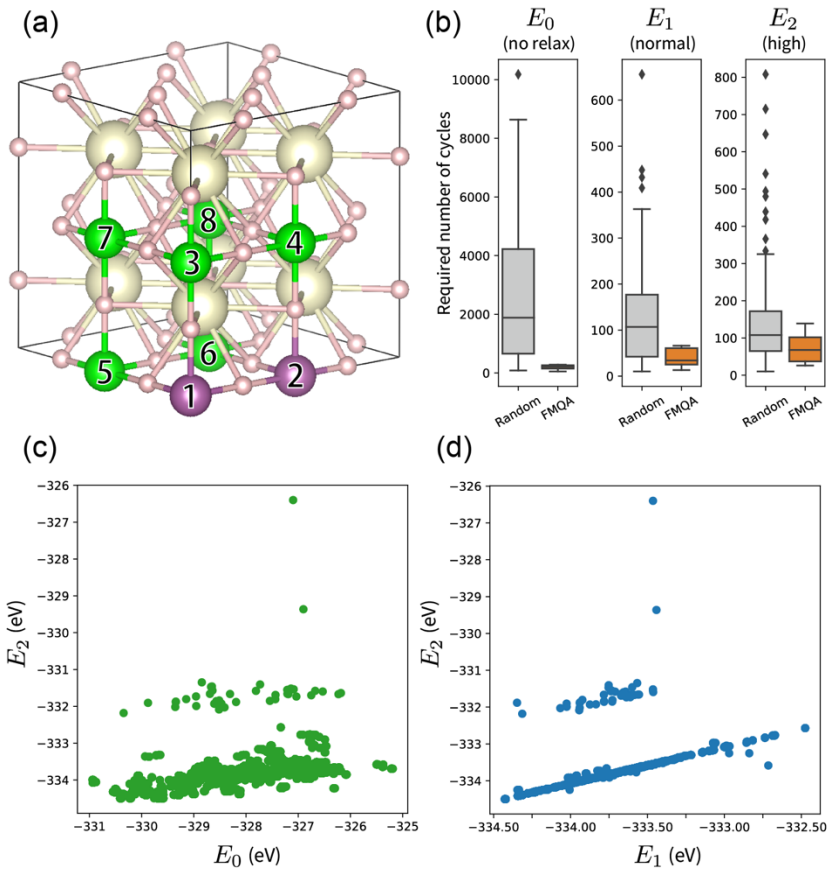

Figure 3. (a) Eight zirconium ion positions in $BaZrO_3$ and the configuration of the two scandium ions in the Sc@(1,2) structure. (b) Number of cycles required to obtain the most stable proton configurations from the FMQA algorithm and random sampling using the total energies E_0 (no relaxation), E_1 (relaxation calculation with normal accuracy), and E_2 (relaxation calculation with high accuracy) for the Sc@(1,2) structure. The horizontal line in the box represents the median value (50th percentile data), while the top and bottom areas represent the 75th and 25th percentiles, respectively. The points denote the outliers, while the error bars represent the maximum and minimum values without outliers. (c) Relationship between E_0 and E_2 for the same initial state. (d) Relationship between E_1 and E_2 for the same initial state. The atomistic configurations were drawn using VESTA.[35]

Subsequently, to examine the optimization performance for more complex problems, the case of $x = 5$ was considered, and the results are shown in **Figure 4**. In this case, three possible configurations of scandium ions exist, which were denoted as Sc@(3,4,6,7,8) (Figure 4(a)), Sc@(1,3,4,6,8), and Sc@(1,2,4,6,7). When the scandium ion configuration was fixed, the number of initial candidate structures for the relaxation calculations was ${}_{24}C_5 \times 8^5 = 1{,}392{,}771{,}072$ without considering symmetry, and so it was not possible to calculate the total energies of all configurations. Consequently, the number of iteration cycles of the BBO runs was fixed at 200, including the generation of ten initial training datasets. Three independent runs were performed by altering the initial training data selections. Figures 4(a), 4(b), and 4(c) show the minimum total energies over the various cycles for Sc@(3,4,6,7,8) when E_0, E_1, and E_2 were targeted for optimization. In all cases, use of the FMQA approach generated more stable structures over fewer cycles compared to those required for random sampling. For the stable configurations determined by FMQA optimizations using E_0 and E_1, configurations were changed by performing further relaxations. More specifically, starting from the top five configurations with small E_0 and E_1 values for each independent FMQA optimization run, further relaxation calculations were performed, which have the same accuracy as those performed using E_2. The total energies obtained by the highly accurate relaxation calculations are summarized in Figure 4 (d). As a result, it was confirmed that stable configurations with a value equivalent to the total energy obtained from the FMQA optimizations performed using E_2 could be obtained from the optimizations using E_1. This implies that by performing FMQA optimization using E_1, followed by further relaxation for a small number of stable configurations, the computation time can be reduced to obtain stable proton configurations.

However, as observed for $x = 2$, the discovery of stable configurations was also meaningless using E_0.

Figure S2 summarizes the results obtained for Sc@(1,3,4,6,8) and Sc@(1,2,4,6,7) when the FMQA optimization was performed using E_1. Further relaxation calculations were also performed on the identified proton configurations, and the total energies obtained with the same accuracy as that of E_2were compared. Consequently, the Sc@(3,4,6,7,8) structure was defined as possessing the lowest total energy.

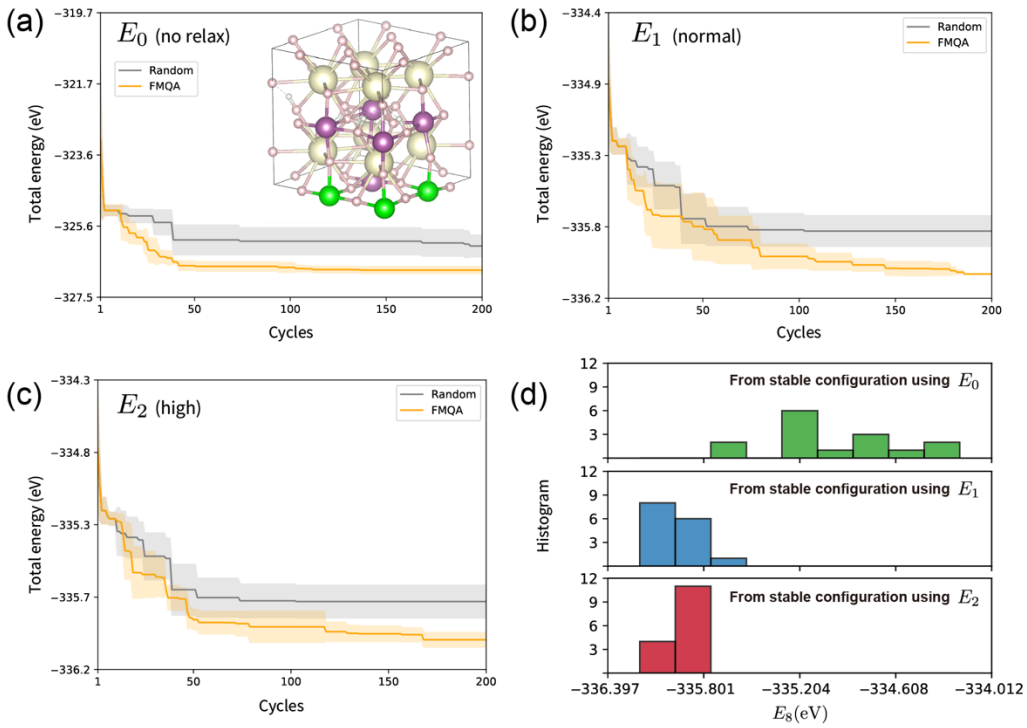

Figure 4. Minimum total energies depending on the number of iteration cycles when (a) E_0, (b) E_1, and (c) E_2 were targeted as the total energies for Sc@(3,4,6,7,8). Three independent runs were performed using the FMQA and by random sampling. The means and deviations are indicated by lines and shaded areas. The inset of (a) shows the configuration the scandium ions in Sc@(3,4,6,7,8). (d) Histogram of the total energies with the high accuracies (E_2), as determined

by performing further relaxation calculations based on the configurations found by FMQA optimization using E_0 and E_1. The total energy of the proton configuration found by FMQA optimization using E_2 is also plotted. The atomistic configurations were drawn using VESTA.[35]

Discussion regarding the stable proton configurations

When the number of Sc ions (i.e., x) is equal to seven or eight, only one configuration exists wherein the scandium ions can be substituted for symmetry. However, the number of initial proton configurations for the relaxation calculation drastically increases under these conditions, reaching 10^{11} and 10^{12} for x = 7 and 8, respectively, without symmetry. In these cases, owing to the large number of candidates, conventional BBO methods such as Bayesian optimization cannot be performed. Thus, stable proton configuration searches were performed by FMQA optimization using E_1. More specifically, three independent FMQA optimizations were carried out, and further relaxation calculations were performed for the top 15 configurations. The results of the optimization cycle using E_1 are shown in Figures S2(c) and S2(d), and the most stable configurations are summarized in Figure 5, wherein the numbers of scandium ions (x) are 2, 5, 7, and 8.

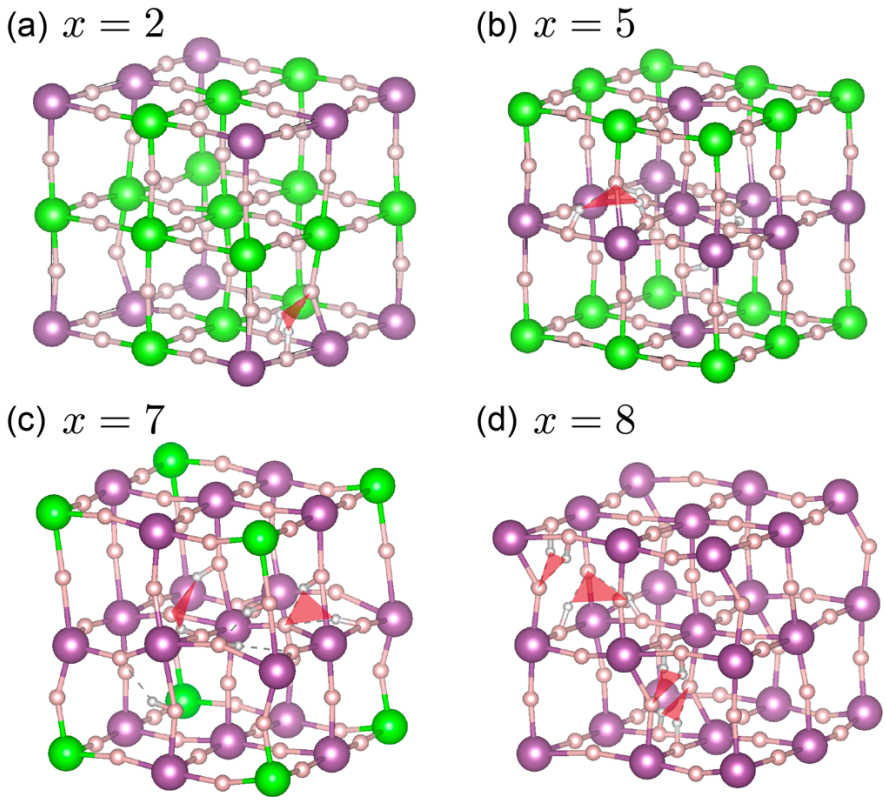

Figure 5. The most stable proton configurations found for (a) 2, (b) 5, (c) 7, and (d) 8, scandium ions. When $x = 2$, the optimal configuration is shown. The red triangles indicate pairs of hydrogen bonds. Barium ions are omitted for clarity. The atomistic configurations were drawn using VESTA.[35]

From the obtained configurations, it was possible to discuss the characteristics of the stable proton configurations. Firstly, in many cases, hydrogen atoms attach to the oxygen sites between the scandium ions, forming Sc–O–Sc species. This indicates that the position of hydrogen is more stable on the oxygen atom of Sc–O–Sc than on Sc–O–Zr or Zr–O–Zr. This stabilization of protons near the scandium ions was consistent with the results of previous ab initio calculations[43,44,36] and nuclear magnetic resonance experiments.[45] Subsequently, hydrogen bond pairs appear, as denoted

by the red triangles in Figure 5. Figure 6 compares the most stable configuration containing a pair of hydrogen bonds and a metastable configuration containing two independent hydrogen bonds when $x = 2$. The lengths of the hydrogen bonds were calculated to be 1.89 and 1.81 Å in the most stable case, which are longer than those calculated for the metastable case, i.e., 1.73 and 1.67 Å. These results indicate that the metastable state has a lower energy (i.e., it is more stable) if the length of the hydrogen bonds is considered. In contrast, lattice deformation was found to increase the total energy, and in the most stable state, the lattice deformation was reduced by the formation of a pair of hydrogen bonds. This property of forming a hydrogen bond pair in a stable proton configuration was observed even when the number of scandium ions was increased (see Figure 5). The contributions of lattice deformation and the formation of a pair of hydrogen bonds for total energy depend on simulation conditions. Future work will therefore examine whether the characteristics of the stable structures observed in this study are universal for simulation conditions using the FMQA algorithm.

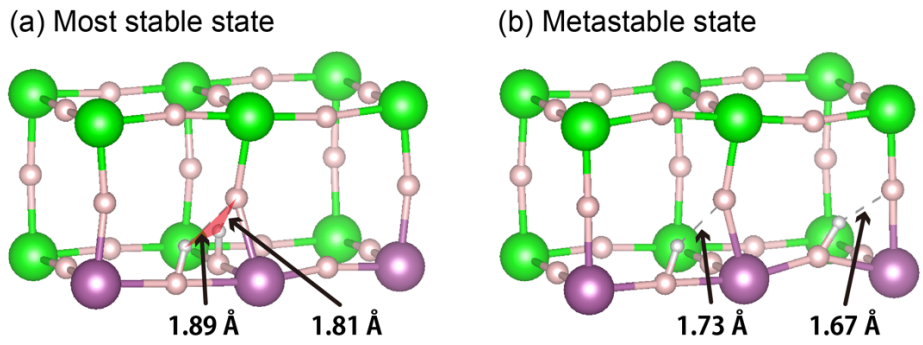

Figure 6. Hydrogen bonds in the (a) most stable and (b) metastable configurations when $x = 2$. The atomistic configurations were drawn using VESTA.[35]

Simultaneous optimization of the doping and proton configurations

It is also possible to optimize the doping configuration of scandium ions simultaneously with the proton configuration. The doping configuration should also be treated as a bit string, and so the bits for the scandium/zirconium sites were prepared as 1 for scandium ions and 0 for zirconium ions. In the 2 × 2 × 2 supercell model, eight bits were used to represent the doping configuration, which was defined by $\{q_l\}_{l=1,\ldots,8}$. In other words, when simultaneously considering the doping and proton configurations, in addition to the 216-bit, $\{q_{ni}\}_{n=1,\ldots,24,i=1,\ldots,9}$ was also employed for doping. Therefore, a 224-bit FM was utilized as a surrogate model to predict the total energy. When solving the trained FM with the Ising machine, in addition to the constraints defined in Eqs. (2) and (3), the following constraint is required:

$$\sum_{l=1}^{8} q_l = x. \tag{4}$$

However, in the present study, symmetry was observed in the doping positions. For example, for $x = 2$, one scandium ion can be fixed at index 1, and the remaining scandium ions can be placed at one of the three sites (i.e., 2, 6, or 8) indicated in Fig. 3(a). Therefore, instead of using Eq. (4), the following constraints can be used to narrow the search space, wherein only three bits are considered (i.e., q_2, q_6, and q_8):

$$q_1 = 1, \tag{5}$$

$$\sum_{l=2,6,8} q_l = 1. \tag{6}$$

In the case where $x = 2$, the doping and proton configurations were simultaneously optimized by BBO using FMQA, and the results are presented in Figure 7 for searching with the total energies E_0, E_1, and E_2. The results of random sampling are also shown for comparison, and the optimization performance of FMQA was observed to outperform that of random sampling.

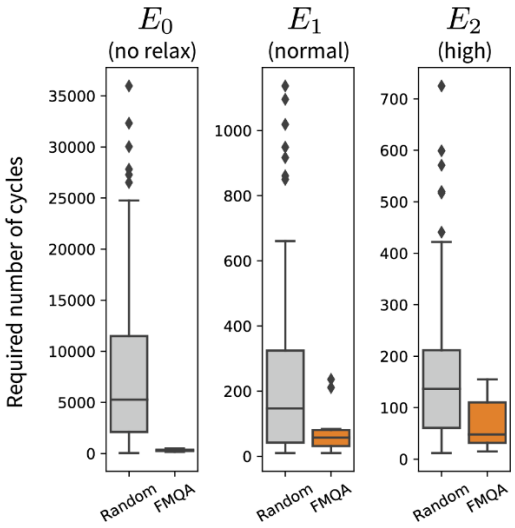

Figure 7. Number of cycles required when the most stable proton configuration was identified by FMQA optimization and random sampling using the total energies E_0 (no relaxation), E_1 (relaxation calculation with normal accuracy), and E_2 (relaxation calculation with high accuracy) for optimization of the doped scandium ion and proton configurations.

CONCLUSIONS

In this study, black-box optimization (BBO) calculations were performed to determine stable proton configurations using the factorization machines with quantum annealing (FMQA) algorithm combined with the Vienna Ab Initio Simulation Package (VASP). A graphics processing unit-based Ising machine was used to select a promising initial state for the relaxation calculations. As a model system for proton conductors, scandium-doped barium zirconate was used, wherein the number of scandium ions (x) was set as 2, 5, 7, or 8 in the 2 × 2 × 2 supercell model. For $x = 2$, the number of possible proton configurations was 52,992, and so it was possible to calculate the total energies for all cases using VASP, and the most stable configuration was identified. In this case, the optimization performance of FMQA algorithm always outperformed that of random sampling for finding the most stable configurations. Subsequently, using the FMQA algorithm, the proton configurations were optimized for x = 5, 7, and 8, wherein the numbers of initial configurations were in the order of 10^9, 10^{11}, and 10^{12}, respectively. Notably, such vast number of possible configurations are difficult to treat using conventional BBO methods, such as Bayesian optimization, because of the computation times required. Moreover, the optimization results identified the formation of hydrogen bond pairs in the stable proton configurations, regardless of the number of scandium ions. Notably, simultaneous optimization of the doped scandium ions and proton configurations was achieved. It is therefore expected that this optimization strategy based on a combination of Ising machines and density functional theory calculations will be able to solve various issues related to configuration optimization in solid-state materials, and promote novel scientific discoveries.

ASSOCIATED CONTENT

Supporting Information. Supporting Information is provided.

AUTHOR INFORMATION

Corresponding Author

Jianbo Lin (linjb86@gmail.com)

Tomofumi Tada (tada.tomofumi.054@m.kyushu-u.ac.jp)

Ryo Tamura (tamura.ryo@nims.go.jp).

Author Contributions

T.T. and R.T. conceived and designed the study. J. L. and A. K. conducted the investigations and calculations. R.T., M.S., and K.T. devised the original idea for the optimization method. All members contributed to manuscript preparation.

Funding Sources

This study was supported by a project subsidized by the Core Research for Evolutional Science and Technology (CREST) (grant number JPMJCR2234) and KAKENHI 21H01008. The computational DFT work was partly supported by the MEXT Program: Data Creation and Utilization Type Material Research and Development Project (grant number JPMXP1122683430) and the MEXT Program for Promoting Researches on the Supercomputer Fugaku (Data-Driven

Research Methods Development and Materials Innovation Led by Computational Materials Science; grant number JPMXP1020230327).

Notes

The authors declare no competing financial interests. Data supporting the findings of this study are available from the corresponding author upon request.

ACKNOWLEDGMENT

The authors would like to thank Yoshihiro Yamazaki and Shusuke Kasamatsu for their fruitful discussions. The computational resources of the Fugaku supercomputer were provided by the RIKEN Center for Computational Science (Project ID: hp230212). The computations performed in this study were performed on an Ising machine in Fixstars Amplify. We would like to thank Editage (www.editage.jp) for English language editing.

REFERENCES

(1) Iwahara, H.; Esaka, T.; Uchida, H.; Maeda, N. Proton Conduction in Sintered Oxides and Its Application to Steam Electrolysis for Hydrogen Production. *Solid State Ionics* **1981**, *3–4*, 359–363. https://doi.org/10.1016/0167-2738(81)90113-2.

(2) Norby, T. Solid-State Protonic Conductors: Principles, Properties, Progress and Prospects. *Solid State Ionics* **1999**, *125* (1), 1–11. https://doi.org/10.1016/S0167-2738(99)00152-6.

(3) Kreuer, K. D. Proton-Conducting Oxides. *Annual Review of Materials Research* **2003**, *33* (Volume 33, 2003), 333–359. https://doi.org/10.1146/annurev.matsci.33.022802.091825.

(4) Yamazaki, Y.; Hernandez-Sanchez, R.; Haile, S. M. High Total Proton Conductivity in Large-Grained Yttrium-Doped Barium Zirconate. *Chem. Mater.* **2009**, *21* (13), 2755–2762. https://doi.org/10.1021/cm900208w.

(5) Hyodo, J.; Kitabayashi, K.; Hoshino, K.; Okuyama, Y.; Yamazaki, Y. Fast and Stable Proton Conduction in Heavily Scandium-Doped Polycrystalline Barium Zirconate at Intermediate Temperatures. *Advanced Energy Materials* **2020**, *10* (25), 2000213. https://doi.org/10.1002/aenm.202000213.

(6) Cao, J.; Ji, Y.; Shao, Z. Perovskites for Protonic Ceramic Fuel Cells: A Review. *Energy Environ. Sci.* **2022**, *15* (6), 2200–2232. https://doi.org/10.1039/D2EE00132B.

(7) Yamazaki, Y.; Babilo, P.; Haile, S. M. Defect Chemistry of Yttrium-Doped Barium Zirconate: A Thermodynamic Analysis of Water Uptake. *Chem. Mater.* **2008**, *20* (20), 6352–6357. https://doi.org/10.1021/cm800843s.

(8) Yamazaki, Y.; Yang, C.-K.; Haile, S. M. Unraveling the Defect Chemistry and Proton Uptake of Yttrium-Doped Barium Zirconate. *Scripta Materialia* **2011**, *65* (2), 102–107. https://doi.org/10.1016/j.scriptamat.2010.12.034.

(9) Kitamura, N.; Akola, J.; Kohara, S.; Fujimoto, K.; Idemoto, Y. Proton Distribution and Dynamics in Y- and Zn-Doped BaZrO3. *J. Phys. Chem. C* **2014**, *118* (33), 18846–18852. https://doi.org/10.1021/jp502455v.

(10) Hossain, M.; Khan, M.; Akash, T.; Hashizume, K. Molecular Dynamics Simulation For Barium Zirconate Proton Conducting Oxide Energy Materials: A Mini Review. *International Exchange and Innovation Conference on Engineering & Science(IEICES)* **2021**, *7*, 1. https://doi.org/10.5109/4738550.

(11) Ju, S.; Shiga, T.; Feng, L.; Hou, Z.; Tsuda, K.; Shiomi, J. Designing Nanostructures for Phonon Transport via Bayesian Optimization. *Phys. Rev. X* **2017**, *7* (2), 021024. https://doi.org/10.1103/PhysRevX.7.021024.

(12) Terayama, K.; Sumita, M.; Tamura, R.; Tsuda, K. Black-Box Optimization for Automated Discovery. *Acc. Chem. Res.* **2021**, *54* (6), 1334–1346. https://doi.org/10.1021/acs.accounts.0c00713.

(13) Ueno, T.; Rhone, T. D.; Hou, Z.; Mizoguchi, T.; Tsuda, K. COMBO: An Efficient Bayesian Optimization Library for Materials Science. *Materials Discovery* **2016**, *4*, 18–21. https://doi.org/10.1016/j.md.2016.04.001.

(14) Motoyama, Y.; Tamura, R.; Yoshimi, K.; Terayama, K.; Ueno, T.; Tsuda, K. Bayesian Optimization Package: PHYSBO. *Computer Physics Communications* **2022**, *278*, 108405. https://doi.org/10.1016/j.cpc.2022.108405.

(15) Garijo del Río, E.; Mortensen, J. J.; Jacobsen, K. W. Local Bayesian Optimizer for Atomic Structures. *Phys. Rev. B* **2019**, *100* (10), 104103. https://doi.org/10.1103/PhysRevB.100.104103.

(16) Todorović, M.; Gutmann, M. U.; Corander, J.; Rinke, P. Bayesian Inference of Atomistic Structure in Functional Materials. *npj Comput Mater* **2019**, *5* (1), 1–7. https://doi.org/10.1038/s41524-019-0175-2.

(17) Yan, J.; Wei, Han; Xie, Han; Gu, X.; Bao, H. Seeking for Low Thermal Conductivity Atomic Configurations in SiGe Alloys with Bayesian Optimization. *ES Energy & Environment* **2020**, *Volume 8 (June 2020)* (4), 56–64.

(18) Ono, S. Optimization of Configurations of Atomic Species on Two-Dimensional Hexagonal Lattices for Copper-Based Systems. *AIP Advances* **2022**, *12* (8), 085313. https://doi.org/10.1063/5.0098517.

(19) Kusaba, A.; Kangawa, Y.; Kuboyama, T.; Oshiyama, A. Exploration of a Large-Scale Reconstructed Structure on GaN(0001) Surface by Bayesian Optimization. *Applied Physics Letters* **2022**, *120* (2), 021602. https://doi.org/10.1063/5.0078660.

(20) Kitai, K.; Guo, J.; Ju, S.; Tanaka, S.; Tsuda, K.; Shiomi, J.; Tamura, R. Designing Metamaterials with Quantum Annealing and Factorization Machines. *Phys. Rev. Res.* **2020**, *2* (1), 013319. https://doi.org/10.1103/PhysRevResearch.2.013319.

(21) Rendle, S. Factorization Machines with Libfm. *ACM Transactions on Intelligent Systems and Technology (TIST)* **2012**, *3* (3), 57.

(22) Mohseni, N.; McMahon, P. L.; Byrnes, T. Ising Machines as Hardware Solvers of Combinatorial Optimization Problems. *Nat Rev Phys* **2022**, *4* (6), 363–379. https://doi.org/10.1038/s42254-022-00440-8.

(23) Wilson, B. A.; Kudyshev, Z. A.; Kildishev, A. V.; Kais, S.; Shalaev, V. M.; Boltasseva, A. Machine Learning Framework for Quantum Sampling of Highly Constrained, Continuous Optimization Problems. *Applied Physics Reviews* **2021**, *8* (4), 041418. https://doi.org/10.1063/5.0060481.

(24) Kim, S.; Shang, W.; Moon, S.; Pastega, T.; Lee, E.; Luo, T. High-Performance Transparent Radiative Cooler Designed by Quantum Computing. *ACS Energy Lett.* **2022**, 4134–4141. https://doi.org/10.1021/acsenergylett.2c01969.

(25) Matsumori, T.; Taki, M.; Kadowaki, T. Application of QUBO Solver Using Black-Box Optimization to Structural Design for Resonance Avoidance. *Sci Rep* **2022**, *12* (1), 12143. https://doi.org/10.1038/s41598-022-16149-8.

(26) Inoue, T.; Seki, Y.; Tanaka, S.; Togawa, N.; Ishizaki, K.; Noda, S. Towards Optimization of Photonic-Crystal Surface-Emitting Lasers via Quantum Annealing. *Opt. Express, OE* **2022**, *30* (24), 43503–43512. https://doi.org/10.1364/OE.476839.

(27) Urushihara, M.; Karube, M.; Yamaguchi, K.; Tamura, R. Optimization of Core–Shell Nanoparticles Using a Combination of Machine Learning and Ising Machine. *Advanced Photonics Research* **2023**, *4* (12), 2300226. https://doi.org/10.1002/adpr.202300226.

(28) Mao, Z.; Matsuda, Y.; Tamura, R.; Tsuda, K. Chemical Design with GPU-Based Ising Machines. *Digital Discovery* **2023**, *2* (4), 1098–1103. https://doi.org/10.1039/D3DD00047H.

(29) Tučs, A.; Berenger, F.; Yumoto, A.; Tamura, R.; Uzawa, T.; Tsuda, K. Quantum Annealing Designs Nonhemolytic Antimicrobial Peptides in a Discrete Latent Space. *ACS Med. Chem. Lett.* **2023**, *14* (5), 577–582. https://doi.org/10.1021/acsmedchemlett.2c00487.

(30) Kim, S.; Jung, S.; Bobbitt, A.; Lee, E.; Luo, T. Wide-Angle Spectral Filter for Energy-Saving Windows Designed by Quantum Annealing-Enhanced Active Learning. *Cell Reports Physical Science* **2024**, *5* (3), 101847. https://doi.org/10.1016/j.xcrp.2024.101847.

(31) Nawa, K.; Suzuki, T.; Masuda, K.; Tanaka, S.; Miura, Y. Quantum Annealing Optimization Method for the Design of Barrier Materials in Magnetic Tunnel Junctions. *Phys. Rev. Appl.* **2023**, *20* (2), 024044. https://doi.org/10.1103/PhysRevApplied.20.024044.

(32) Couzinie, Y.; Seki, Y.; Nishiya, Y.; Nishi, H.; Kosugi, T.; Tanaka, S.; Matsushita, Y. Machine Learning Supported Annealing for Prediction of Grand Canonical Crystal Structures. arXiv August 7, 2024. https://doi.org/10.48550/arXiv.2408.03556.

(33) Luo, T.; Xu, Z.; Shang, W.; Kim, S.; Lee, E. QALO: Quantum Annealing-Assisted Lattice Optimization. August 27, 2024. https://doi.org/10.21203/rs.3.rs-4518513/v1.

(34) Hoshino, K.; Kasamatsu, S.; Hyodo, J.; Yamamoto, K.; Setoyama, H.; Okajima, T.; Yamazaki, Y. Probing Local Environments of Oxygen Vacancies Responsible for Hydration in Sc-Doped Barium Zirconates at Elevated Temperatures: In Situ X-Ray Absorption Spectroscopy, Thermogravimetry, and Active Learning Ab Initio Replica Exchange Monte Carlo Simulations. *Chem. Mater.* **2023**, *35* (6), 2289–2301. https://doi.org/10.1021/acs.chemmater.2c02116.

(35) Momma, K.; Izumi, F. VESTA 3 for Three-Dimensional Visualization of Crystal, Volumetric and Morphology Data. *J Appl Cryst* **2011**, *44* (6), 1272–1276. https://doi.org/10.1107/S0021889811038970.

(36) Vera, C. Y. R.; Ding, H.; Peterson, D.; Gibbons, W. T.; Zhou, M.; Ding, D. A Mini-Review on Proton Conduction of BaZrO3-Based Perovskite Electrolytes. *J. Phys. Energy* **2021**, *3* (3), 032019. https://doi.org/10.1088/2515-7655/ac12ab.

(37) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47* (1), 558–561. https://doi.org/10.1103/PhysRevB.47.558.

(38) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169–11186. https://doi.org/10.1103/PhysRevB.54.11169.

(39) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Computational Materials Science* **1996**, *6* (1), 15–50. https://doi.org/10.1016/0927-0256(96)00008-0.

(40) Peressi, M.; Baldereschi, A. Chapter 2 - Ab Initio Studies of Structural and Electronic Properties. In *Characterization of Semiconductor Heterostructures and Nanostructures (Second Edition)*; Lamberti, C., Agostini, G., Eds.; Elsevier: Oxford, 2013; pp 21–73. https://doi.org/10.1016/B978-0-444-59551-5.00002-9.

(41) Musielewicz, J.; Wang, X.; Tian, T.; Ulissi, Z. FINETUNA: Fine-Tuning Accelerated Molecular Simulations. *Mach. Learn.: Sci. Technol.* **2022**, *3* (3), 03LT01. https://doi.org/10.1088/2632-2153/ac8fe0.

(42) *Fixstars Amplify*. https://amplify.fixstars.com/en/.

(43) Björketun, M. E.; Sundell, P. G.; Wahnström, G. Structure and Thermodynamic Stability of Hydrogen Interstitials in BaZrO3 Perovskite Oxide from Density Functional Calculations. *Faraday Discuss.* **2006**, *134* (0), 247–265. https://doi.org/10.1039/B602081J.

(44) Yamazaki, Y.; Kuwabara, A.; Hyodo, J.; Okuyama, Y.; Fisher, C. A. J.; Haile, S. M. Oxygen Affinity: The Missing Link Enabling Prediction of Proton Conductivities in Doped Barium Zirconates. *Chem. Mater.* **2020**, *32* (17), 7292–7300. https://doi.org/10.1021/acs.chemmater.0c01869.

(45) Buannic, L.; Sperrin, L.; Dervişoğlu, R.; Blanc, F.; Grey, C. P. Proton Distribution in Sc-Doped BaZrO3: A Solid State NMR and First Principle Calculations Analysis. *Phys. Chem. Chem. Phys.* **2018**, *20* (6), 4317–4328. https://doi.org/10.1039/C7CP08523K.