Acta Materialia 292 (2025) 120999

Contents lists available at ScienceDirect

Acta Materialia

journal homepage: www.elsevier.com/locate/actamat

Ruddlesden-Popper tolerance factor: An indicator predicting stability of 2D Ruddlesden-Popper phases

Hyo Gyeong Shin [a], Eun Ho Kim [a], Jaeseon Kim [a], Hyo Kim [a], Donghwa Lee [a,b,*]

[a] Department of Materials Science and Engineering (MSE), Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea
[b] Institute for Convergence Research and Education in Advanced Technology (I_CREATE), Yonsei University, Incheon 21983, Republic of Korea

ARTICLE INFO

Keywords:
Machine learning
Descriptor
Two-dimensional
Perovskites
Density functional theory

ABSTRACT

Two-dimensional Ruddlesden-Popper (RP) phases receive the focus of extensive research because of their unique optical and electrical properties. Accurate prediction of stable RP phases can expedite finding new RP compositions with improved properties for practical applications. However, most attempts are limited to finding new RP phases by employing time-consuming computational approaches. Although descriptors such as cationic radius ratio or Goldschmidt tolerance factor can be used alternatively, they have shown limitation in predicting stable RP phases. In this study, thus we develop a novel RP tolerance factor (t_{RP}) derived through machine learning based process, which exhibits high accuracy in classifying RP and non-RP phases. The t_{RP} is a simple form based solely on ionic radii and incorporates two meaningful parameters: inverse cationic radius ratio and inverse octahedral factor. Additionally, t_{RP} shows a linear correlation with first-principles density functional theory formation energies, allowing us to determine the relative stability of RP phases beyond mere classification. By utilizing t_{RP} as a descriptor, we propose the new compounds that showing potential for RP phases, and we expect it will pave the way for the discovery of novel RP phases for future optoelectronic applications.

1. Introduction

Ruddlesden-Popper (RP) phases, known as half-shifting stacks of two-dimensional (2D) perovskite layers, have emerged as a promising material class due to their superior environmental stability and versatile optical and electrical properties [1–5]. These properties include superconductivity, tunable bandgap, high carrier mobility and ferroelectricity. The unique combination of these properties makes RP phases as advantageous candidates across diverse applications [6–10]. For instance, RP phases offer tunable bandgaps and strong photoluminescence boosting the performance of light-emitting diodes and lasers [6,7]. In the realm of solar cells, both high power conversion efficiency (PCE) and high stability is achieved by using RP phases as capping layers [8,9]. Given the superior properties of RP phases and their potential in various applications, there is a growing interest in exploring new RP phases that exhibit improved stability and desirable properties.

Recently, various computational approaches have been employed to expediate the discovery of new RP phases. For example, Zhang et al. utilized machine learning (ML) techniques to identify factors influencing the formation of RP phases and validated their findings through the synthesis of new RP phases [11]. Hu et al. combined high-throughput density functional theory (DFT) calculations with ML approaches to identify new RP phases suitable for photovoltaic applications [12]. Meftahi et al. developed new RP phase films for solar cells using the high-throughput fabrication and optimization strategy aided by ML methods [13].

However, these studies require extensive time and computational resources, limiting their accessibility. In contrast, descriptors offer an easier way to predict the stability of perovskite materials without the need for intensive computational calculations. One common descriptor used to predict the stability of single-layer RP phases (A_2BX_4) is the cationic radius ratio (c), composed of the ionic radii of A and B ions (r_A, r_B) [14].

$$c = \frac{r_A}{r_B}$$

While it provides a straightforward means of assessing stability based on A and B ions, neglects the impact of X ions. As an alternative, the Goldschmidt tolerance factor (t) incorporating the impact of X ions, has been utilized to predict the stability of RP phases [3,15]. The t is defined as:

* Corresponding author.
E-mail address: donghwa96@postech.ac.kr (D. Lee).

https://doi.org/10.1016/j.actamat.2025.120999
Received 1 January 2025; Received in revised form 18 March 2025; Accepted 28 March 2025
Available online 31 March 2025

$$t = \frac{r_A + r_B}{\sqrt{2}(r_B + r_X)}$$

where r_A, r_B, and r_X are the ionic radii of A, B, and X ion, respectively. Despite its broad application and easy accessibility, recent studies have mentioned certain limitations of the t [16,17]. The generally accepted ranges for the t is not directly applicable to RP phases and result in poor accuracy, especially when determining the stability of RP phases containing heavier halides (Table S3) [18–21].

Therefore, accurately predicting the stability of compounds as RP phases remains a challenging task necessitating the development of more refined descriptors. In this study, we introduce the RP tolerance factor (t_{RP}), a groundbreaking predictive descriptor derived from a sophisticated ML-based methodology. The t_{RP} achieves exceptional classification accuracy in differentiating RP from non-RP phases by integrating the inverse form of two parameters: c and octahedral factor (μ). This unique integration allows t_{RP} to accurately forecast stable RP phases, offering a simple yet effective tool for stability prediction. Our comparisons of t_{RP} with DFT results demonstrate that t_{RP} not only classifies RP phases but also assesses their degree of stability, enabling the identification and proposal of new promising RP compounds. We believe that t_{RP} can be used as a general metric for finding innovative RP materials with applications in optoelectronics and beyond.

2. Methods

2.1. Construction of datasets

The dataset, comprising 930 compounds, is constructed from experimental sources: 854 from the Handbook of Inorganic Substances, 8 from the 2D Perovskites Database, and 68 from the Inorganic Crystal Structure Database (ICSD) and various research papers [22–25]. For the RP phases, K_2NiF_4-type compounds are considered for A_2BX_4, $Sr_3Ti_2O_7$-type for $A_3B_2X_7$, and $Sr_4Ti_3O_{10}$-type for $A_4B_3X_{10}$ compounds. For non-RP phases, we identify the most prevalent compound types sharing the same stoichiometry as their RP counterparts; this includes six types for A_2BX_4 (spinel, olivine, K_2SO_4, Th_3P_4, high-Tc, and CaV_2O_4), four types for $A_3B_2X_7$ ($SrTb_2Fe_2O_7$, $La_3Mn_{0.5}SiS_7$, La_3CuSiS_7, and $Ca_2MgSi_2O_7$) and four types for $A_4B_3X_{10}$ ($K_2Sb_2Sn_3S_{10}$, $Cs_4Mg_3F_{10}$, $Ca_4Ti_3O_{10}$, and $Ba_4Zr_3S_{10}$). In the case of organic-inorganic hybrid compounds, the 2D Perovskite Database is used, which contains organic-inorganic 2D perovskite materials [25]. Within this database, we select only compounds whose organic cation sizes were known [26]. As a result, we obtained 8 organic-inorganic RP hybrid perovskites containing methylammonium ($CH_3NH_3^+$, MA^+), ethylammonium ($CH_3CH_2NH_3^+$, EA^+), and ammonium (NH_4^+) as organic cations. Within the scope of the t_{RP} extraction process, among the 544 all-inorganic compounds, 476 sourced from the Handbook of Inorganic Substances are utilized as the training set, and the remaining compounds allocate as the test set.

2.2. SISSO classification framework

In this study, sure independence screening and sparsifying operator (SISSO) classification framework with stratified k-fold (5-fold) cross-validation is implemented to extract common descriptors [27]. Features and labels for compounds in each fold are randomly split into a training set and a test set, and the training set serves as the input set of SISSO. In each fold of the cross-validation, SISSO does not use the test set for generating or refining descriptors. Each SISSO is trained on a distinct training set. Then, a combination of algebraic operations considering three levels of complexity (Φ_1, Φ_2 and Φ_3) along with a maximum feature complexity of 10 and using the operator set $H \equiv \{I, +, -, \times, /, \exp, \log, \sqrt{}, ^{-1}, ^2\}$, is repeatedly executed to construct a descriptor space (feature space) in each fold. The descriptor space increases approximately from 6×10^1 at Φ_1, to 3×10^3 at Φ_2, and up to 1×10^7 at Φ_3. Out of the vast descriptor space, sure independence screening (SIS) scores the descriptors based on the correlation magnitude between the feature and the target and selects only top ranked as much as the subspace size (100,000). According to the SISSO process, the selected descriptors go through a sparsifying operator (SO) that finally selects the best nD ($n > 1$) descriptors. However, in this study, only the SIS step of SISSO is performed by setting the output dimension of SISSO up to 1D descriptors. Then, 100,000 descriptors are extracted as the output of SISSO from each fold. Of the total 5 (5-fold) × 100,000, common descriptors to all folds, which are descriptor candidates, are finally obtained.

2.3. DFT calculations

DFT calculations are performed to obtain the formation energy (E_{form}) of compounds with a RP phase (K_2NiF_4-type) backbone structure. The DFT calculations are implemented using the Vienna Ab-initio Simulation Package (VASP) code [28,29]. The projector augmented wave (PAW) method with the Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) for the exchange and correlation potentials is utilized [30–32]. At the Γ-point, the Monkhorst-Pack k-point sampling of 5 × 5 × 5 grid with energy cutoff of 520 eV is used [33]. For the fully relaxed structure, the convergence criterion of energy and force are 10^{-6} eV and 0.001 eV/ Å, respectively.

3. Results and discussion

3.1. Extraction of t_{RP} through a 3-step process

Our dataset consists of 930 compounds with RP phase stoichiometry ($A_{n+1}B_nX_{3n+1}$) from $n = 1$ to $n = 3$, where n represents the thickness of the perovskite layer. Of which, the most extensively studied are the all-inorganic A_2BX_4 compounds (Fig. 1a). Therefore, to extract the t_{RP}, we focus on the 544 all-inorganic A_2BX_4 compounds, while the remaining compounds are used for evaluating the derived t_{RP}. Among 544 A_2BX_4 compounds, 195 are in the RP phases and 349 are in the non-RP phases. The A-site in these compounds includes 41 elements, spanning the lanthanide series, alkali metals, alkaline earth metals, and a small number of transition metals. The B-site includes 43 elements, predominantly composed of transition metals. The X-site comprises chalcogens (O, S, Se, and Te) and halogens (F, Cl, Br, and I) (Fig. 1b). For more information, see the supplementary materials.

To extract the t_{RP}, a 3-step process is designed using the dataset of A_2BX_4 compounds (Fig. 2). In the first step, the dataset along with atomic and ionic features serve as inputs (Table S1). The features include Shannon ionic radii (r_A, r_B, and r_X), oxidation states (n_A, n_B, and n_X), electronegativities (x_A, x_B, and x_X), valence electrons (v_A, v_B, and v_X), and ionization energies (i_A, i_B, and i_X) [34]. Important features among these are identified via SHaply Additive exPlanations (SHAP), which ranks features according to their importances [35]. In this analysis, feature importances are quantified based on how each feature influences the decision to classify a material as forming the RP phase. Then, the r_A, r_B, r_X and x_X are selected as the most important features (Fig. S1). Second step involves SISSO classification framework with 5-fold cross-validation applied to the datasets and selected features [27]. To minimize the risk of data leakage, we carefully set aside 68 data points out of a total 544 as a completely independent test set, while the remaining 476 data points are used for 5-fold cross-validation process. SISSO outputs the top 100,000 descriptors for each fold. In our 5-fold cross-validation, the 100,000 descriptors outputted by SISSO are not all equal. Hence, we extract common descriptors present in all fold outputs (Fig. S2). Third step considers three factors: classification accuracy, complexity, and unit to identify the best simple unitless descriptor among the common ones (Table S2). First of all, a random forest classifier is used to assess each descriptor's classification accuracy

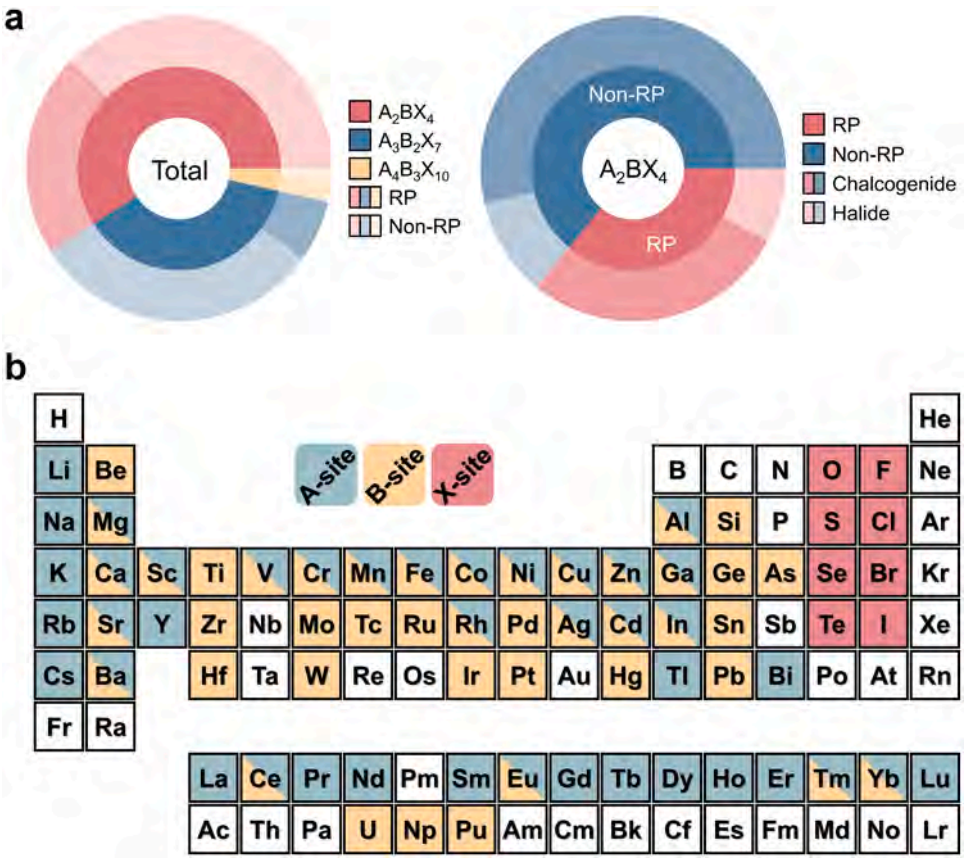

Fig. 1. Overview of dataset and elemental composition. **(a)**, On the right, the datasets of RP and non-RP phases $n = 1$ (A_2BX_4), $n = 2$ ($A_3B_2X_7$) and $n = 3$ ($A_4B_3X_{10}$) are illustrated. The darker colors represent RP phases and lighter colors represent non-RP phases as shown in the color legend. On the left, the A_2BX_4 compounds are displayed. The pink color represents RP phases and blue color represents non-RP phases. Darker shades denote chalcogenides, while lighter shades represent halides. **(b)**, Distribution of elemental composition across the A, B, and X sites within the A_2BX_4 dataset is depicted in blue, yellow, and pink, respectively.

[36,37]. For this evaluation, we employ two measures: accuracy on the independent 68-point test set and the average accuracy on the test set within the 5-fold cross-validation. By aligning the descriptors based on their independent test accuracy and then confirming with the average cross-validation test accuracy, we ensure a robust performance evaluation. Then, descriptor analysis is conducted to evaluate the complexity of each descriptor by considering the number of features and mathematical symbols and to identify the unit of descriptors. As a result, we extract t_{RP}, the best simple descriptor that demonstrates high classification accuracy, simplicity, and unitless.

3.2. High accuracy and two key parameters defining t_{RP}

The t_{RP} is expressed as follows:

$$t_{RP} = \left(\frac{r_B}{r_A}\right)^2 + \sqrt{\frac{r_X}{r_B}}$$

where r_A, r_B, and r_X are the ionic radii of A, B, and X ion, respectively. We take a closer look at t_{RP} and find that it contains two important terms: inverse cationic radius ratio ($1/c$, $c = r_A/r_B$) and the inverse octahedral factor ($1/\mu$, $\mu = r_B/r_X$). The term c evaluates the relative sizes of A-site and B-site ions. The term μ evaluates whether a B-site cation can enter an octahedron outlined by X-site anions. We note here that the ionic radii depend on the oxidation states. Since chalcogenides and halides have different oxidation states of cations, the distribution of ionic radii constituting the RP compounds are also different. Accordingly, halides and chalcogenides exhibit different distributions of t_{RP} values. Therefore, we establish separate t_{RP} criteria to classify RP and non-RP phases for chalcogenides and halides. The criteria for t_{RP} are established by analyzing the decision criteria ranges derived from a random forest classifier. The optimal criterion that gives the highest classification accuracy is then selected. The t_{RP} criterion is identified as 1.83 for chalcogenides and 1.72 for halides. Compounds with t_{RP} values below the criteria are classified as stable RP phases, while those with values above are classified as non-RP phases.

Fig. 3a illustrates the distribution on RP and non-RP phases according to t_{RP} values of chalcogenides (left) and halides (right). The t_{RP} achieves an overall classification accuracy of 90 % for both chalcogenides and halides. Specifically, the classification accuracy in the RP phases is 93 % for chalcogenides and 81 % for halides. The lower accuracy observed in the halides compared to the chalcogenides is due to the fact that halide compounds is more flexible for the formation of RP phases with B-site cations. In the non-RP phases, the classification accuracy is 88 % for chalcogenides and 96 % for halides. While t and c do not clearly distinguish between RP and non-RP phases, t_{RP} shows much higher accuracy on distinguishing between RP and non-RP phases (Fig. S3).

Fig. 3b illustrates the distribution of RP and non-RP phases according to the two parameters ($1/c$, $1/\mu$) incorporated within t_{RP}. The inverse relationship between $1/c$ and $1/\mu$ is revealed in this figure. In the case of RP phases, $1/c$ is mainly distributed in the range of 0.4 to 0.6 and $1/\mu$ is mainly distributed in the range of 1.3 to 2.6. The majority RP phases are discernibly present in these ranges satisfying the proposed t_{RP} criteria. The presence of halide RP phases with low $1/\mu$ values (yellow star) represents that the halide octahedral cage is more flexible, allowing large B-site cations to enter as well. In contrast, non-RP phases are widely distributed in area where $1/c$ or $1/\mu$ values are high individually or the sum of the two values is high. This indicates that the formation of

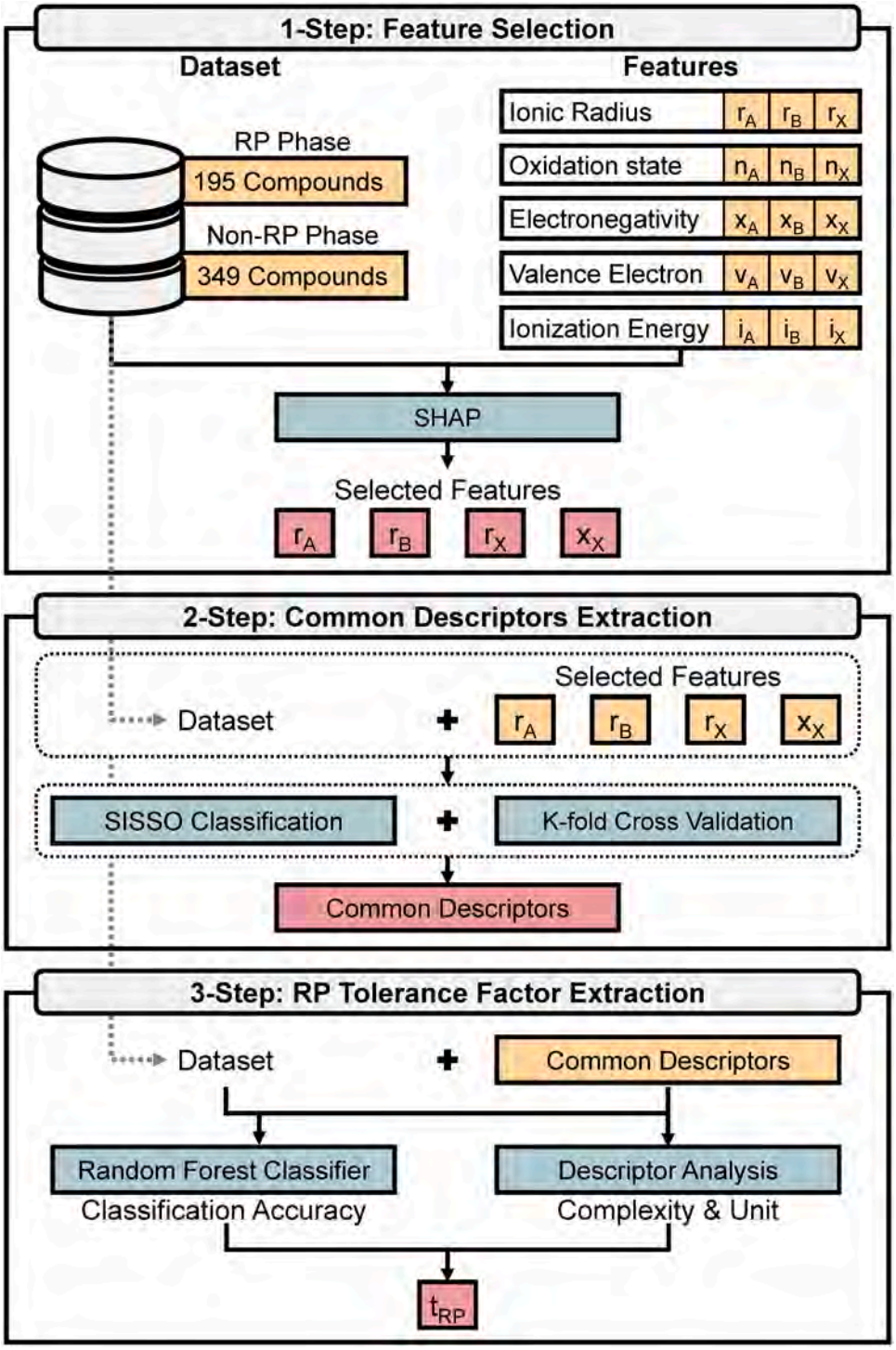

Fig. 2. A 3-step process to extract the best simple descriptor. Each step consists of the inputs shown in yellow boxes, the methods shown in blue boxes, and the outputs shown in the red boxes. In the 1st step, datasets along with their ionic and atomic features are considered. SHAP analysis is performed on these inputs to select important features. In the 2nd step, the datasets and selected features are served as inputs. The SISSO algorithm coupled with k-fold cross-validation is then used to derive common descriptors. In the 3rd step, the dataset and common descriptors are used as inputs. Then, random forest classifier and descriptor analysis are conducted to extract the t_{RP}.

RP phases cannot be clearly distinguished by using $1/c$ or $1/\mu$ separately and the sum of two values, should be small as what we proposed in t_{RP} criteria. Therefore, the synergy of $1/c$ and $1/\mu$ is important to estimate the formation of RP phases and t_{RP} is a good indicator for RP phase formation.

3.3. Correlation between t_{RP} and E_{form}

To investigate the ability of t_{RP} for evaluating the degree of stability of the RP phases beyond simple classification, we further compare it with E_{form}, which is a generally accepted criterion judging stability of materials. Fig. 4 shows the correlation between DFT-calculated E_{form} and t_{RP}. The E_{form} is expressed in eV/f.u (where f.u stands for formula unit)

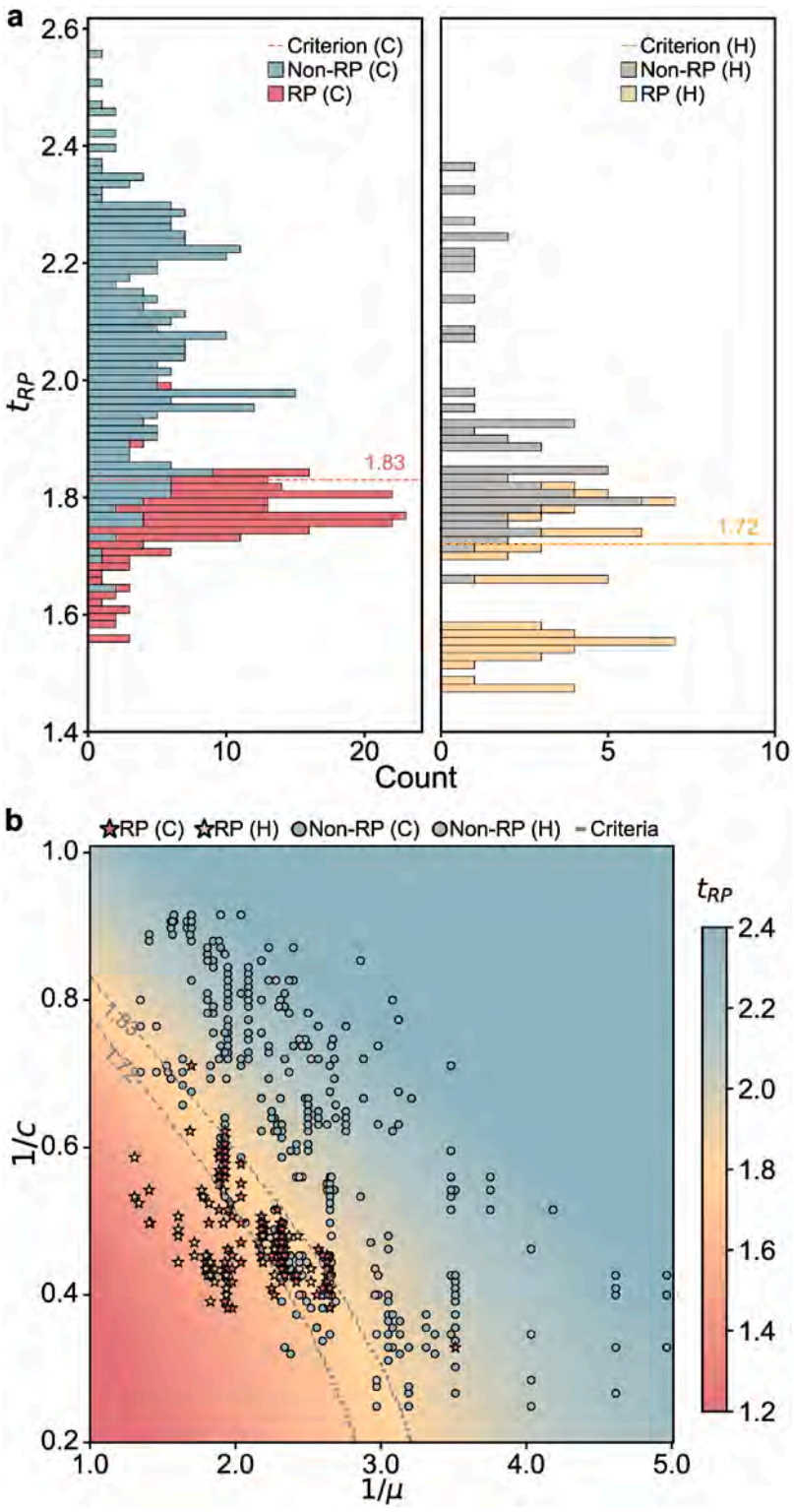

Fig. 3. Classification of A_2BX_4 compounds through t_{RP}. **(a)**, On the left, a stacked format is used to display the RP and non-RP phases for chalcogenides (C), which appear in red and blue, respectively. On the right, a stacked format is used to display the RP and non-RP phase for halides (H), which appear in yellow and gray, respectively. The dotted lines represent the t_{RP} criteria. Compounds below the t_{RP} criteria are stable RP phases, whereas those above are non-RP phases. **(b)**, Displays the distribution of A_2BX_4 compounds in relation to the $1/c$ and $1/\mu$. The color bar shows t_{RP}. values. Red and yellow stars represent RP phases in C and H, respectively. Blue and gray circles represent non-RP phases in C and H, respectively. The gray dotted line indicates t_{RP} criteria in C and H.

and quantifies the energy required to form a compound from its precursors. The E_{form} is calculated as follows:

$$E_{form} = E_{bulk}^{total} - \left(aE_{A_kX_l}^{tot} + bE_{B_mX_n}^{tot}\right)$$

where E_{bulk}^{total} is the total energy of the A_2BX_4 compound with the RP phase backbone structure, and a and b are coefficients multiplied by the precursors to maintain the A_2BX_4 stoichiometry. The chosen precursors are binary compounds (A_kX_l and B_mX_n) commonly used in 2D perovskite

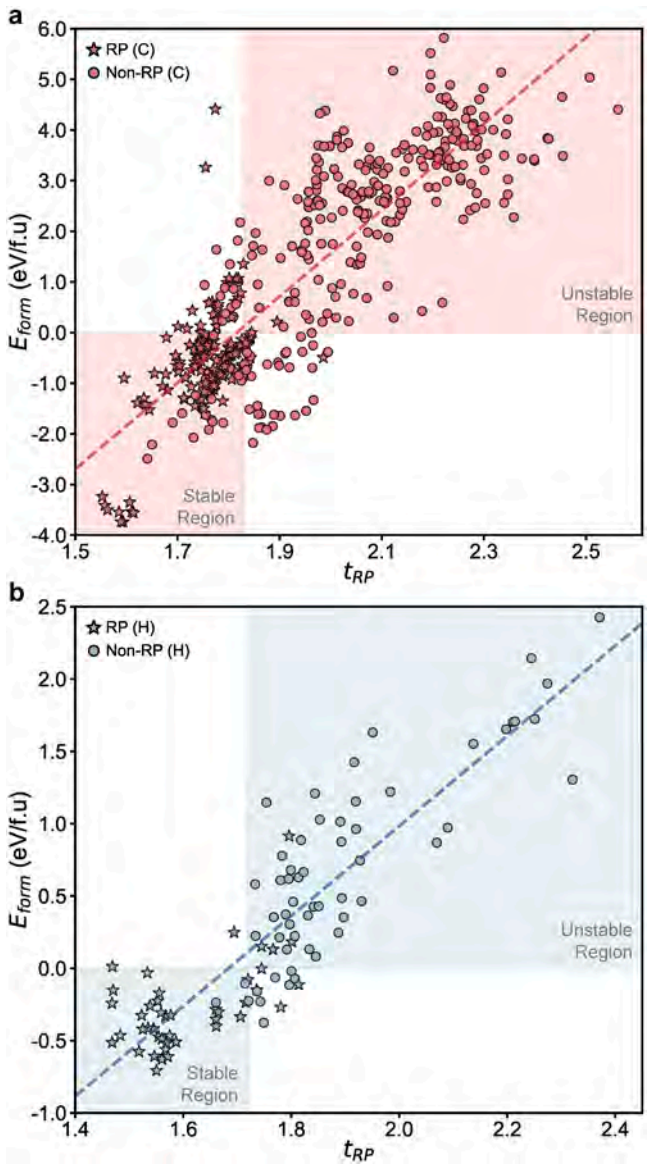

Fig. 4. Correlation between E_{form} and t_{RP}. The correlation of E_{form} and t_{RP} is shown, **(a)**, for chalcogenides (C) and **(b)**, for halides (H). Stars represent RP phases, while circles indicate non-RP phases. Dotted lines depict the line of best fit, along with their correlation coefficients.

synthesis, taking into account the oxidation states of the ions involved [22,38,39]. The subscripts k, l, m and n are determined by these oxidation states. $E^{tot}_{A_kX_l}$ and $E^{tot}_{B_mX_n}$ represent the total energies of these precursors.

The positive linear relationship between E_{form} and t_{RP} is shown in Fig. 4 with a correlation coefficient 0.85 for chalcogenides and 0.89 for halides. This suggests that as the t_{RP} value decreases below its criterion the E_{form} value also decreases and the possibility of forming a stable RP phases increase. So, our study clearly shows that t_{RP} can be an indicator representing degree of stability to the RP phase. In Fig. 4a and 4b, the color regions correspond to where both E_{form} and t_{RP} consistently classify compounds as either stable RP phases (lower color region) or unstable RP phases (upper color region). The white regions denote areas where the classifications based on E_{form} and t_{RP} are inconsistent.

We further investigate the white region to understand the origin of the inconsistency. In the upper white region of chalcogenides (Fig. 4a), a significant portion (37 out of 41) of these compounds contain lanthanide elements, leading misclassification of E_{form} or t_{RP}. This indicates that DFT calculations are limited in predicting E_{form} of lanthanide compounds and t_{RP} also has difficulty classifying lanthanide compounds [40,41]. In the lower white region for chalcogenides (Fig. 4a) and halides (Fig. 4b), inconsistencies are also observed. For the RP phases in the lower white region, t_{RP} incorrectly predicts them as non-RP phases by slightly exceeding the criteria. This indicates that t_{RP} criteria should be used in caution since it is difficult to accurately predict the stability of RP phase near t_{RP} criteria. For the non-RP phases in the lower white region, DFT calculations incorrectly predict the negative E_{form} for them. This indicates that some compositions are difficult to accurately predict E_{form} as E_{form} is highly dependent on precursor selection.

We further examine the lower color region to determine the possibility of forming RP phases among non-RP phases. In this region, 21 chalcogenide and 3 halide non-RP phases are classified as RP phases from both E_{form} and t_{RP}. Especially, Ba_2GeO_4, $EuNaGeO_4$, $Sr_2NiO_2Cl_2$, $NaNdGeO_4$ and Cs_2CdBr_4 show lower E_{form} in the RP-phase backbone structures than their original structures (Fig. S4). This suggests that they can adopt the RP phases if proper synthesis method is applied. Thus, we encourage material scientists to try synthesizing them into new RP compounds.

3.4. Expansion of t_{RP} to higher n-value RP phases

Lastly, we expand our study to assess the viability of applying t_{RP} not only to A_2BX_4 ($n = 1$) compounds, but also to $A_3B_2X_7$ ($n = 2$) and $A_4B_3X_{10}$ ($n = 3$) compounds. A total of 930 compounds with different n-values are evaluated using t_{RP}. The t_{RP} criteria for each n-value are identified by analyzing decision criteria ranges (Fig. 5). For chalcogenides, the t_{RP} criteria are identified as 1.83 ($n = 1$), 1.82 ($n = 2$), and 1.79 ($n = 3$), and the classification accuracy according to the t_{RP} criteria is 89 % for all cases. The decrease in t_{RP} criteria with higher n-value indicates that the probability of forming a stable RP phase decreases as the n-value increases, consistent with findings from previous studies [42,43]. This represents that although the n-value is not directly included in t_{RP}, t_{RP} can still be applied to higher n-value RP phases with different t_{RP} criteria. In the case of halides, a criterion of $n = 1$ is uniformly applied across all n-values because data for non-RP phases are not available for $n = 2$ and $n = 3$. Our results show that most RP phases fall within the expected range, yielding an 87 % classification accuracy. Thus, we believe that t_{RP} can be applied to classify higher n-value RP phases using different t_{RP} criteria.

4. Conclusion

In this study, we have employed a systematic 3-step process that integrates SHAP, SISSO with k-fold cross-validation, and random forest classifier with descriptor analysis to develop a simple and accurate descriptor, t_{RP}. It is designed to distinguish between RP and non-RP phases with high classification accuracy by effectively integrating two relevant parameters (c and μ) based solely on ionic radii. A crucial aspect of t_{RP} is its linear correlation with E_{form}, indicating its ability to identify compounds with potential as new RP phases. Furthermore, t_{RP} exhibits the capability to assess the stability of high n-value RP phases.

Evaluating stability of new compositional materials is a fundamental step in discovering new materials [44–46]. For material scientists, quickly and reliably assessing stability is a crucial yet challenging task. Our study clearly shows that t_{RP} developed through a systematic 3-step process can efficiently and effectively predict the RP phase stability of compounds. Although methods such as backward elimination can be used as a valid approach for feature selection, we find that our simple and systematic 3-step process sufficiently yields a robust and meaningful descriptor, t_{RP}. Given that our 3-step process can be applied to develop tolerance factor for various phases, our study can be extended to search for novel materials in different fields. In summary, our study provides not only a straightforward t_{RP} that enables rapid design and discovery of

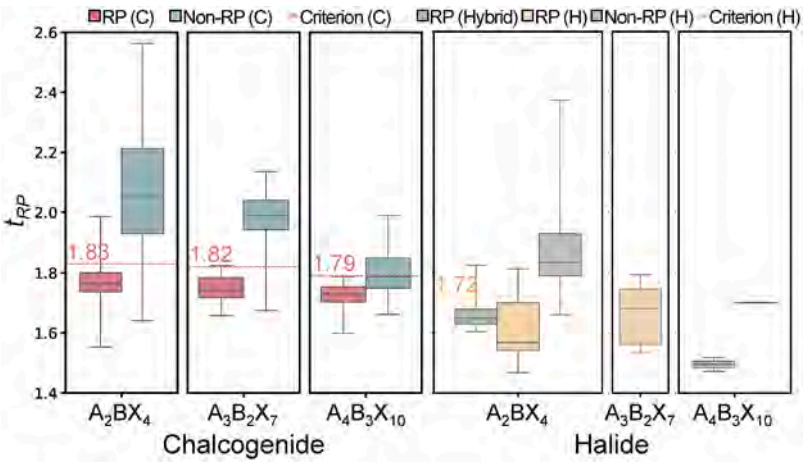

Fig. 5. Distribution of RP and non-RP phases in A_2BX_4 ($n = 1$), $A_3B_2X_7$ ($n = 2$), and $A_4B_3X_{10}$ ($n = 3$) compounds according to t_{RP}. The three figures on the left show chalcogenides (C), with the RP phases in red and the non-RP phases in blue. The red lines show the criteria of C. The three figures on the right show halides (H), with the hybrid RP phases in green, the RP phases in yellow, and the non-RP phases in gray. The yellow lines show the criteria of H. Compounds below the t_{RP} criteria are stable RP phases, whereas those above are non-RP phases. The lines in each box represent the median value, and the lines outside each box represent the minimum and maximum values.

novel RP phases but also a systematic 3-step process that can be used to develop tolerance factors for estimating the stability of new compositional materials, potentially revolutionizing materials discovery in various applications.

Code availability

The code related to the t_{RP} descriptor and its application can be found at https://github.com/CNMD-POSTECH/tRP.

CRediT authorship contribution statement

Hyo Gyeong Shin: Writing – review & editing, Writing – original draft, Visualization, Validation, Data curation, Conceptualization. **Eun Ho Kim:** Validation. **Jaeseon Kim:** Formal analysis. **Hyo Kim:** Investigation. **Donghwa Lee:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning No. NRF-2020R1A6C101A202 and NRF-2024M3A7C2045166 and NRF-2021M3I3A1084940 and RS-2023-00257666 and RS-2024-00446683 and RS-2024-00450836.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actamat.2025.120999.

References

[1] H. Kim, J.-K. Kim, J. Kwon, J. Kim, H.-W.J. Kim, S. Ha, K. Kim, W. Lee, J. Kim, G. Y. Cho, Quantum spin nematic phase in a square-lattice iridate, Nature 625 (7994) (2024) 264–269.
[2] H. Sun, M. Huo, X. Hu, J. Li, Z. Liu, Y. Han, L. Tang, Z. Mao, P. Yang, B. Wang, Signatures of superconductivity near 80 K in a nickelate under high pressure, Nature 621 (7979) (2023) 493–498.
[3] W. Yi, Z. Liu, L. Zhang, X. Chen, Y. Huang, S. Ma, Q. Li, X. Zhao, J. Wang, H. Zhou, S. Jin, La2SrSc2O7: a-site cation disorder induces ferroelectricity in Ruddlesden–Popper layered perovskite oxide, J. Am. Chem. Soc. 146 (2024) 4570–4581.
[4] Machine learning As-sisted synthetic acceleration of Ruddlesden–Popper and Dion–Jacobson 2D lead halide perovskites, Acta Mater 245 (2023) 118638.
[5] S.S. Nikitin, A.D. Kiverin, V.M. Kovrugin, et al. defect chemistry and charge transport in LaSr3Fe3− xMoxO10− δ, Acta Mater. 203 (2021) 116489.
[6] B. Zhao, M. Vasilopoulou, A. Fakharuddin, F. Gao, A.R.b. Mohd Yusoff, R. H. Friend, D. Di, Light management for perovskite light-emitting diodes, Nat. Nanotechnol. 18 (9) (2023) 981–992.
[7] Y. Li, H. Zhou, M. Xia, H. Shen, T. Wang, H. Gao, X. Sheng, Y. Han, Z. Chen, L. Dou, Phase-pure 2D tin halide perovskite thin flakes for stable lasing, Sci. Adv. 9 (32) (2023) eadh0517.
[8] Y.-W. Jang, S. Lee, K.M. Yeom, K. Jeong, K. Choi, M. Choi, J.H. Noh, Intact 2D/3D halide junction perovskite solar cells via solid-phase in-plane growth, Nature Energy 6 (1) (2021) 63–71.
[9] Y. Luo, K. Liu, L. Yang, W. Feng, L. Zheng, L. Shen, Y. Jin, Z. Fang, P. Song, W. Tian, Dissolved-Cl2 triggered redox reaction enables high-performance perovskite solar cells, Nat. Commun. 14 (1) (2023) 3738.
[10] L. Wang, T. Liu, T. Wu, J. Lu, Strain-retardant coherent perovskite phase stabilized Ni-rich cathode, Nature 611 (7934) (2022) 61–67.
[11] Z.-Z. Zhang, T.-M. Guo, Z.-G. Li, F.-F. Gao, W. Li, F. Wei, X.-H. Bu, Machine learning assisted synthetic acceleration of Ruddlesden-Popper and Dion-Jacobson 2D lead halide perovskites, Acta Mater. 245 (2023) 118638.
[12] J. Hu, C. Wang, Q. Li, R. Sa, P. Gao, Accelerated design of photovoltaic Ruddlesden–Popper perovskite Ca6Sn4S14− xOx using machine learning, APL Mater. 8 (11) (2020).
[13] N. Meftahi, M.A. Surmiak, S.O. Fürer, K.J. Rietwyk, J. Lu, S.R. Raga, C. Evans, M. Michalska, H. Deng, D.P. McMeekin, Machine learning enhanced high-throughput fabrication and optimization of quasi-2D Ruddlesden–Popper Perovskite solar cells, Adv. Energy Mater. 13 (38) (2023) 2203859.
[14] D. Ganguli, Cationic radius ratio and formation of K2NiF4-type compounds, J. Solid State Chem. 30 (3) (1979) 353–356.
[15] V.M. Goldschmidt, Die gesetze der krystallochemie, Naturwissenschaften 14 (21) (1926) 477–485.
[16] C.J. Bartel, C. Sutton, B.R. Goldsmith, R. Ouyang, C.B. Musgrave, L.M. Ghiringhelli, M. Scheffler, New tolerance factor to predict the stability of perovskite oxides and halides, Sci. Adv. 5 (2) (2019) eaav0693.
[17] J. Feng, X. Wang, J. Li, H. Liang, W. Wen, E. Alvianto, C.-W. Qiu, R. Su, Y. Hou, Resonant perovskite solar cells with extended band edge, Nat. Commun. 14 (1) (2023) 5392.
[18] Y. Fu, M.P. Hautzinger, Z. Luo, F. Wang, D. Pan, M.M. Aristov, I.A. Guzei, A. Pan, X. Zhu, S. Jin, Incorporating large A cations into lead iodide perovskite cages: relaxed goldschmidt tolerance factor and impact on exciton–phonon interaction, ACS Cent. Sci. 5 (8) (2019) 1377–1386.
[19] S. Ramos-Terron, A.D. Jodlowski, C. Verdugo-Escamilla, L. Camacho, G. de Miguel, Relaxing the goldschmidt tolerance factor: sizable incorporation of the guanidinium cation into a two-dimensional ruddlesden–popper perovskite, Chem. Mater. 32 (9) (2020) 4024–4037.
[20] G. Grancini, M.K. Nazeeruddin, Dimensional tailoring of hybrid perovskites for photovoltaics, Nature Rev. Mater. 4 (1) (2019) 4–22.
[21] Y. Shao, W. Gao, H. Yan, R. Li, I. Abdelwahab, X. Chi, L. Rogée, L. Zhuang, W. Fu, S.P. Lau, Unlocking surface octahedral tilt in two-dimensional Ruddlesden-Popper perovskites, Nat. Commun. 13 (1) (2022) 138.

[22] J. Li, Q. Yu, Y. He, C.C. Stoumpos, G. Niu, G.G. Trimarchi, H. Guo, G. Dong, D. Wang, L. Wang, Cs2PbI2Cl2, all-inorganic two-dimensional Ruddlesden–Popper mixed halide perovskite with optoelectronic response, J. Am. Chem. Soc. 140 (35) (2018) 11085–11090.
[23] G. Bergerhoff, R. Hundt, R. Sievers, I. Brown, The inorganic crystal structure data base, J. Chem. Inf. Comput. Sci. 23 (2) (1983) 66–69.
[24] P. Villars, K. Cenzual, R. Gladyshevskii, Handbook, Walter de Gruyter GmbH & Co KG, 2014.
[25] E.I. Marchenko, S.A. Fateev, A.A. Petrov, V.V. Korolev, A. Mitrofanov, A.V. Petrov, E.A. Goodilin, A.B. Tarasov, Database of two-dimensional hybrid perovskite materials: open-access collection of crystal structures, band gaps, and atomic partial charges predicted by machine learning, Chem. Mater. 32 (17) (2020) 7383–7388.
[26] G. Kieslich, S. Sun, A.K. Cheetham, Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog, Chem. Sci. 5 (12) (2014) 4712–4715.
[27] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, L.M. Ghiringhelli, SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates, Phys. Rev. Mater. 2 (8) (2018) 083802.
[28] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B 54 (16) (1996) 11169.
[29] M. Fuchs, M. Scheffler, Ab initio pseudopotentials for electronic structure calculations of poly -atomic systems using density-functional theory, Comput. Phys. Commun. 119 (1) (1999) 67–98.
[30] P.E. Blöchl, Projector augmented-wave method, Phys. Rev. B 50 (24) (1994) 17953.
[31] K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (18) (1996) 3865.
[32] J.P. Perdew, K. Burke, Y. Wang, Generalized gradient approximation for the exchange-correlation hole of a many-electron system, Phys. Rev. B 54 (23) (1996) 16533.
[33] H.J. Monkhorst, J.D. Pack, Special points for Brillouin -zone integrations, Phys. Rev. B 13 (12) (1976) 5188.
[34] R.D. Shannon, Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, Acta Crystallograph. Sect. A 32 (5) (1976) 751–767.
[35] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees, Nature Mach. Intell. 2 (1) (2020) 56–67.
[36] A. Liaw, M. Wiener, Classification and regression by randomForest, Rnews 2 (3) (2002) 18–22.
[37] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[38] L. Dou, A.B. Wong, Y. Yu, M. Lai, N. Kornienko, S.W. Eaton, A. Fu, C.G. Bischak, J. Ma, T. Ding, Atomically thin two-dimensional organic-inorganic hybrid perovskites, Science 349 (6255) (2015) 1518–1521.
[39] Q. Zhang, L. Chu, F. Zhou, W. Ji, G. Eda, Excitonic properties of chemically synthesized 2D organic–inorganic hybrid perovskite nanosheets, Adv. Mater. 30 (18) (2018) 1704055.
[40] V.A. Basiuk, O.V. Prezhdo, E.V. Basiuk, Thermal smearing in DFT calculations: how small is really small? A case of La and Lu atoms adsorbed on graphene, Mater. Today Commun. 25 (2020) 101595.
[41] M.G. Taylor, D.J. Burrill, J. Janssen, E.R. Batista, D. Perez, P. Yang, Architector for high-throughput cross-periodic Table 3D complex building, Nat. Commun. 14 (1) (2023) 2786.
[42] C.M.M. Soe, G. Nagabhushana, R. Shivaramaiah, H. Tsai, W. Nie, J.-C. Blancon, F. Melkonyan, D.H. Cao, B. Traoré, L. Pedesseau, Structural and thermodynamic limits of layer thickness in 2D halide perovskites, Proc. Natl. Acad. Sci. 116 (1) (2019) 58–66.
[43] M.R. Barone, M. Jeong, N. Parker, J. Sun, D.A. Tenne, K. Lee, D.G. Schlom, Synthesis of metastable Ruddlesden–Popper titanates, (ATiO3) nAO, with n≥ 20 by molecular-beam epitaxy, APL Mater. 10 (9) (2022).
[44] A. Merchant, S. Batzner, S.S. Schoenholz, M. Aykol, G. Cheon, E.D. Cubuk, Scaling deep learning for materials discovery, Nature 624 (7990) (2023) 80–85.
[45] N.J. Szymanski, B. Rendy, Y. Fei, R.E. Kumar, T. He, D. Milsted, M.J. McDermott, M. Gallant, E.D. Cubuk, A. Merchant, An autonomous laboratory for the accelerated synthesis of novel materials, Nature 624 (7990) (2023) 86–91.
[46] H. Meng, Y. Wang, Y. Zhang, et al., Formation ability descriptors for high-entropy diborides established through high throughput experiments and machine learning, Acta Mater. 256 (2023) 119132.