Discovery of Lead-Free Perovskites for High-Performance Solar Cells via Machine Learning: Ultrabroadband Absorption, Low Radiative Combination, and Enhanced Thermal Conductivities

Xia Cai, Yiming Zhang, Zejiao Shi, Ying Chen, Yujie Xia, Anran Yu, Yuanfeng Xu, Fengxian Xie, Hezhu Shao, Heyuan Zhu, Desheng Fu, Yiqiang Zhan,* and Hao Zhang*

Exploring lead-free candidates and improving efficiency and stability remain the obstacle of hybrid organic-inorganic perovskite-based devices commercialization. Traditional trial-and-error methods seriously restrict the discovery especially for large search space, complex crystal structure and multi-objective properties. Here, the authors propose a multi-step and multi-stage screening scheme to accelerate the discovery of hybrid organic-inorganic perovskites $A_2BB'X_6$ from a large number of candidates through combining machine learning with high-throughput calculations for pursuing excellent efficiency and thermal stability in solar cells. Followed by a series of screenings, the structure-property relationships mapping $A_2BB'X_6$ properties are built and the predictions are close to reported experimental results. Successfully, four experimental-feasibly candidates with good stability, high Debye temperature and suitable band gap are screened out and further verified by density-functional theory calculations, in which the predicted efficiency for three lead-free candidates (($CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$ and $(C_2NH_6)_2AgInBr_6$) achieves 20.6%, 19.9% and 27.6% due to ultrabroadband absorption region ranging from UVC to IRC with excitonic radiative combination rates as low as 10 ps, large or intermediate polarons form with properties similar to $CH_3NH_3PbI_3$ and the calculated thermal conductivities are 5.04, 4.39 and 5.16 $Wm^{-1}K^{-1}$, respectively, with Debye temperatures larger than 500 K, beneficial for suppression of both nonradiative combination and heat-induced degradation.

1. Introduction

Hybrid organic–inorganic perovskites (HOIPs) represented by $CH_3NH_3PbX_3$ as promising next-generation photovoltaic materials have attracted great attention in recent years and have revolutionized solar cells since in 2009 Kojima et al. first successfully reported with a power conversion efficiency (PCE) of 3.8%.[1] The PCE of HOIPs-based photovoltaic system has increased to 25.5% in only 10 years.[2] In addition to further improving PCE of HOIPs, many efforts have been done, such as 2D/3D mixed dimensional perovskite,[3–5] charge transport layer modification,[6–8] interlayer insertion,[9,10] and post encapsulation[11] to restrain HOIPs decomposing into Pb^{2+} composites. However, due to the limited impacts of these works on material intrinsic stability, Pb^{2+} eventually poisons the earth after a long period time scale.[12–14] Pb substitution is a widely accepted concept to wipe out the toxic issue from HOIPs. Another concern is the intrinsic instability of perovskite. Industry market standard requires solar cells to work steady over 25 years in ambient.[15]

X. Cai, Y. Zhang, Z. Shi, Y. Chen, Y. Xia, A. Yu, F. Xie, H. Zhu, Y. Zhan, H. Zhang
School of Information Science and Technology
Fudan University
Shanghai 200433, China
E-mail: yqzhan@fudan.edu.cn; zhangh@fudan.edu.cn

X. Cai, Z. Shi, A. Yu, Y. Zhan
Center of Micro-Nano System
School of Information Science and Technology
Fudan University
Shanghai 200433, China

Y. Zhang, Y. Xia, H. Zhu, H. Zhang
Key Laboratory of Micro and Nano Photonic Structures (MOE) and Department of Optical Science and Engineering
Fudan University
Shanghai 200433, China

Y. Xu
School of Science
Shandong Jianzhu University
Jinan, Shandong 250101, China

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
SCIENCE**
Open Access

www.advancedscience.com

Light- and heat-induced degradation have been integrally studied and proved to be two major issues in HOIPs.[16–18] For example, $CH_3NH_3PbI_3$ can be easily degraded due to the light-induced reaction, that is, $CH_3NH_3PbI_3 \overset{h\nu}{\rightleftarrows} PbI_2 + CH_3NH_2 \uparrow + HI\uparrow$,[19] and heat-induced degradation due to the thermal conductivity as low as 0.5 $Wm^{-1}K^{-1}$ at room temperature,[20] respectively. After baking at over 150 ^oC, HOIPs easily go through an endothermic reaction and decompose into its components like PbI_2 and HI.[21,22] The general solutions to thermal issue in device are to dissipate accumulated heat efficiently by adding heat-dissipation layers, while the complexity of device fabrication process increases.[23,24] Generally, in materials, Debye temperature $\theta_D \propto n(\frac{\rho}{M})^{1/3}v_m$, in which n, ρ, M, and v_m are the number of atoms per formula unit, the crystal structure's density, molar mass, and average sound velocity, respectively, plays a key role in proxy for structure rigidity and directly relates to the thermal conductivity. If perovskite material possesses a high Debye temperature, generally leading to high thermal conductivity, which is beneficial to rapidly dissipate heat and conducive to the thermal stability of devices.[25,26] High Debye temperature may decrease the non-radiative recombination in solar cells as well, due to the high phonon energies. Therefore, HOIPs with the merits of lead free and high Debye temperature should be explored to pursuit the goal of excellent efficiency and thermal stability in perovskite solar cells. The double-perovskite $A_2BB'X_6$ named HOIDP with A the molecular cation, B/B' the metal cation, and X the anionic bridging ligand, has emerged as a new class promising lead-free HOIPs, due to diverse electronic structures and multiple material selections. However, currently, most of the works on lead-free $A_2BB'X_6$ focus on inorganic materials, for example, $Cs_2AgBiBr_6$, which mainly suffers from the following problems: 1) its absorption is usually below 650 nm[27] which prevents the charge generation; 2) low thermal conductivity of $A_2BB'X_6$ leads to high working temperature of devices; 3) complex structures make it difficult to maintain the lattice stability; 4) the formation of small polarons or self-trapped excitons limits mobilities, thus suppressing diffusion/drift lengths; 5) multiple material options lead trial-and-error method work labor-intensive and time-consuming. Therefore, the search for new $A_2BB'X_6$ HOIDPs with excellent performance and good stability has become an urgent task.

Compared with the traditional trial-and-error approach to discover new materials based on scientists' physical and chemical intuition, the emergence of advanced techniques, such as high-throughput computation based on density-functional theory (DFT) method, has greatly accelerated the search process.[28–32] However, the inherent complexity of materials severely hinders its efficiency due to the large-scale real-world chemical space. Fortunately, the rapid advances in Materials Genome Project[33] and artificial intelligence technologies offer exciting hope for this dilemma.[34–49] Unlike first-principles approaches, machine learning (ML) technique can be used smoothly to rapidly predict one or more target properties without relying on numerically solving complex systems of quantum mechanical equations, once suitable material database is established and efficient model is selected, which requires several orders of magnitude fewer computational resources than traditional method. Recently, ML technique has made significant progress in the design of rational materials, such as photovoltaic materials,[34–39] catalysts,[40–42] lithium batteries,[43] and so on. Notably, many materials predicted by ML technique have been synthesized experimentally and have shown excellent properties. These successful attempts[44–49] have demonstrated that intelligent ML technique bypassing intense DFT calculations or experimental trials can enable low-cost, fast, and highly accurate prediction of target material properties, greatly accelerating material discovery.

In this work, we develop a multi-step material screening scheme to accelerate the discovery of new lead-free HOIDPs with enhanced thermal stability for high-performance solar cells, by combining high-throughput DFT calculations with ML technique. The perovskite stability, band gap, and Debye temperature are selected as three target properties to step-wise screen the chemical space. As the fundamental step for global search of possible HOIDPs candidates, a database containing 180 038 electrically neutral compounds is firstly screened from a full chemical space of elemental combinations in the periodic table based on 32 organic cations, and then the condition of structure stability is used to screen out candidates with unstable structure. Third, different ML models are built for multi-goal and multi-stage screen with high accuracy and the importance of relevant features on learning goal is analyzed. Based on ML-predicted results some orthorhombic-like HOIDPs candidates for light-harvesting applications stand out, and Br-based and environment-friendly candidates are selected for further DFT validation. Finally, four lead-free HOIDPs with enhanced thermal stability are picked out as promising solar cell materials with appropriate band gap and high Debye temperature.

2. Results

2.1. Material Screening Framework

In this work, we propose to use the multi-step material screening approach to discover novel hybrid organic–inorganic double perovskites for photovoltaic application with good stability, high Debye temperature, and suitable band gap. The workflow of screening processes is schematically illustrated in **Figure 1**a. In the first screening process of our material-search scheme, the HOIDP compounds must satisfy the charge neutrality criterion, and then tolerance factor and octahedral factor are used to evaluate the structural stability of perovskite candidates. A large chemical space for HOIDP candidates is generated subsequently.

The procedure of machine-learning screening for the generated large chemical space includes three parts: first, a regression

H. Shao
College of Electrical and Electronic Engineering
Wenzhou University
Wenzhou 325035, China
D. Fu
Department of Electronics & Materials Sciences
Faculty of Engineering, & Department of Optoelectronics and Nanostructure Science
Graduate School of Science and Technology
Shizuoka University
Hamamatsu 432-8561, Japan
H. Zhang
Yiwu Research Institute of Fudan University
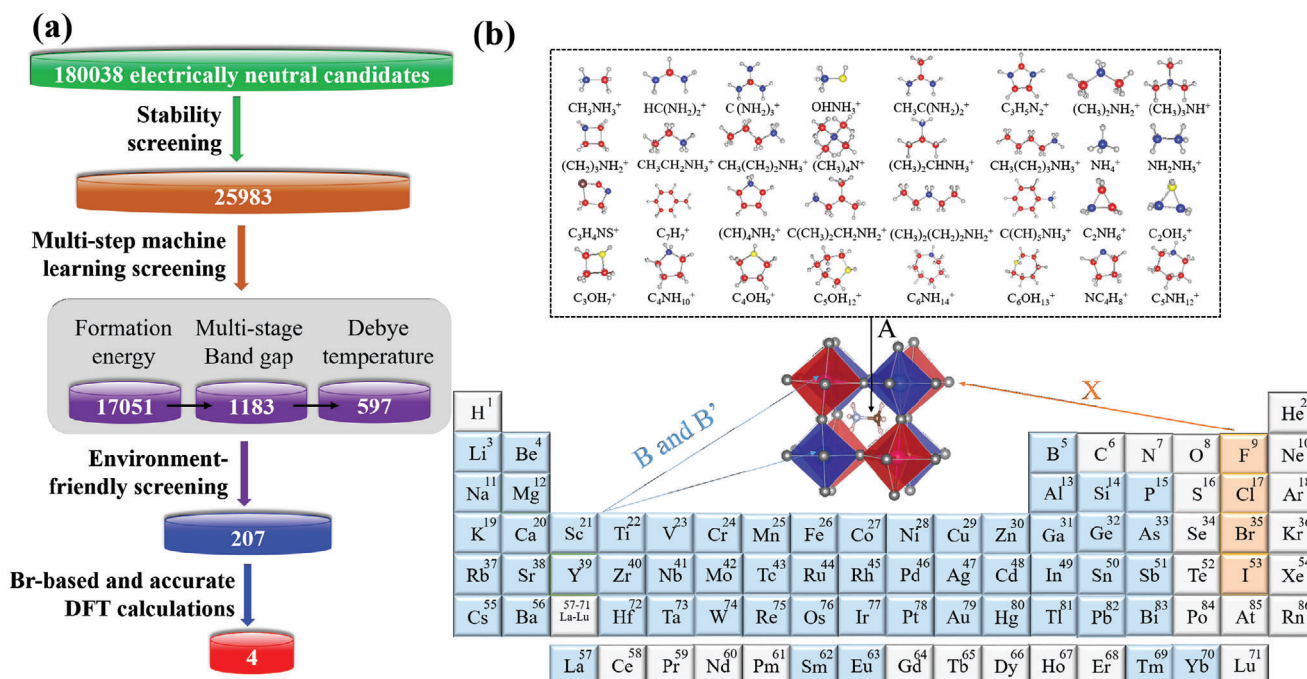Chengbei Road, Yiwu City, Zhejiang 322000, China

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
SCIENCE**
Open Access

www.advancedscience.com

Figure 1. The left chart for the process of discovering novel HOIDPs according to the combination of ML and DFT calculation for photovoltaic application and the right chart for the composition and structure of perovskites in prediction set. Here, the combinations of 32 monovalent organic molecular cations for A site, 9 monovalent, 49 divalent, and 35 trivalent B/B′-site cations, and 4 X-site anions cross the periodic table produce the set of unexplored HOIDPs candidates. In the condition of charge neutrality, 180038 initial candidates are obtained. Then through the stability condition and ML method, 597 HOIDPs suitable for solar cells are chosen. Finally, electronic and other properties of these chosen candidates are further verified by DFT calculations, and four ideal HOIDPs with good quality are finally selected.

model is built to obtain the average formation energy per atom of HOIDPs, in order to judge the chemical stability of each compound. Second, three regression models are built to predict the band gap of HOIDPs to ensure the accuracy of the screening, which can help identify candidates with the appropriate band gap. In the parallel step, another regression model is built to pick the candidates with high Debye temperature to select candidates with large thermal conductivity and good thermal stability. A much smaller set of HOIDP candidates is generated after the machine-learning screening based on the formation-energy, band-gap and Debye-temperature models. Finally, the accuracy of the predicted small set is further verified by DFT calculation, including computing the electronic property, formation energy, Debye temperature and other properties for screening candidates.

2.2. Dataset and Modeling: High-Throughout Calculation of Debye Temperature

Generally, in order to train a good ML model, a high-quality training dataset is required. Data diversity is important for the prediction accuracy of ML models. In addition to inert gas and radioactive elements, the elements utilized in our training and test sets cover almost the entire periodic table, which should be reliable enough to realize accurate ML model. The input dataset of this work composed of 4456 HOIDPs is obtained from high-throughput DFT calculations.[31,32] All the selected compounds possess the double perovskite structures with the chemical

formula of $A_2BB'X_6$. As shown in Figure 1b, the halogen atoms X (X=F, Cl, Br, I) occupy the vertices of the corner-sharing BX_6 or $B'X_6$ octahedron. The elements of B and B′ can be same or different, and the number of different B-site and B′-site cations in our dataset is 20. 16 kinds of monovalent A^+-site cations are filled in the cavity formed by the adjacent octahedrons. In order to define the feature vector of each $A_2BB'X_6$ compound, 95 features based on the main characteristics of elements listed in Table S1, Supporting Information, are utilized to represent, which can be obtained directly from the periodic table to facilitate ML prediction.

The calculated formation energies of 4456 HOIDPs used as the dataset for building formation-energy ML model are shown in Figure S1a, Supporting Information, which plots the relation between Goldschmidt tolerance factor (T_f) and formation energy. The T_f is defined by $\frac{R_A+R_X}{\sqrt{2}(R_{\bar{B}}+R_X)}$, with R_A and R_X representing the ionic radius of A cation and X anion, respectively, and $R_{\bar{B}}$ representing the average ionic radius of the B and B′ cations. Effective radii[50] and Shannon's ionic radii[51] are adopted for molecular ions and atomic ions, respectively. As shown in Figure S1a, Supporting Information, most of the chemically stable HOIDPs in the dataset with the formation energies less than 0 eV atom^{-1} possess the T_f values ranging from 0.80 to 1.25, which is consistent with the chemical tuition for perovskites that the appropriate tolerance factor T_f should be in the range between 0.8 and 1.2.[52–55]

The dependence of T_f on Perdew–Burke–Ernzerhof (PBE) band gap of 4456 HOIDPs are shown in Figure S1b, Supporting

Information, in which all the compounds can be divided into three categories: metals (zero band gap, 3533), semiconductors (band gap between 0 and 3.5 eV, 776), and insulators (band gap lager than 3.5 eV, 147). We extract 425 compounds with direct band gap and verify the band gap by PBE-DFT calculation, to constitute the dataset for further modeling. Considering that PBE method usually underestimates the band gap of semiconductor, we also build the band-gap ML model based on the reported 663 compounds with accurate Heyd–Scuseria–Erenzhof (HSE) band gap. Involving all the 95 features for the PBE/HSE band-gap training dataset does not mean characterizing perovskites well, since too many features sometimes increase the complexity of ML model (called "curse of dimension"), leading to over fitting. Thus, we also implement the process of feature selection to simplify the inputs of model.

As we know, Debye temperature (Θ_D) is the most reliable proxy for structural rigidity, and the materials with higher Θ_D tend to have higher-energy phonon modes that decrease the probability of non-radiative relaxation,[25,26] which is beneficial for the enhancement of quantum coherence and device performances in HOIDP-based solar cells. Generally, above Θ_D, all phonons in materials are excited. According to the slack model,[56] the lattice thermal conductivity in materials can be estimated by $\kappa_L = A \times \frac{\bar{M}\Theta_D^3\delta}{\gamma^2 n^{2/3} T}$, where $\bar{M}$ is the average mass per atom in the crystal, δ^3 is the volume per atom, γ is the Grüneisen parameter, n is the number of atoms in the primitive unit cell, and A is a constant determined by γ, that is, $A = \frac{2.43 \cdot 10^{-8}}{1 - 0.514/\gamma + 0.228/\gamma^2}$. The Grüneisen parameter (γ) can been calculated by $\gamma = \frac{3}{2}\left(\frac{1+v_p}{2-3v_p}\right)$,[57] where $v_p = \frac{1 - 2(v_T/v_L)^2}{2 - 2(v_T/v_L)^2}$ according to longitudinal (v_L) and transverse (v_T) sound velocity found in Equation (17). Therefore, large Debye temperature leads to high κ_L. Due to computational intensity, we only calculate Debye temperature for the 425 HOIDP compounds with direct band gap.

The ML models for formation energies per atom, band gap and Debye temperature built by seven ML algorithms are shown in Figure S2, Supporting Information, in which black dots are the values of determination coefficient (R^2), red dots represent mean square error (MSE), and blue dots are mean absolute error (MAE). The higher R^2, the smaller MSE, the smaller MAE, and the better performance of the trained model. As shown in Figure S2a– d, Supporting Information, among the seven regression algorithms, the GBR algorithm performs best among these models, which gives R^2/MSE/MAE of 0.990/0.021/0.013 eV atom^{-1}, 0.920/0.307/0.241 eV, 0.870/0.384/0.278 eV, and 0.990/16.741/11.737 K for models of formation energy, PBE band gap, HSE band gap, and Debye temperature, respectively.

2.3. Basic Screening

Herein, the common valence of element is considered to use to calculate the charge neutrality of each candidate. In this way, we first screen out a large number of non-neutral compounds from the possible candidate compounds. Monovalent A$^+$-site organic cation, divalent B^{2+}-site and B$'^{2+}$-site cations and monovalent X$^-$-site halogen anion or monovalent A$^+$-site organic cation, monovalent B$^+$-site cation, triva-

lent B$'^{3+}$-site cation, and monovalent X$^-$-site halogen anion are collected to generate possible HOIDP candidates. Here, 32 different kinds of monovalent organic molecular cations including CH$_3$NH$_3^+$, HC(NH$_2$)$_2^+$, C(NH$_2$)$_3^+$, OHNH$_3^+$, CH$_3$C(NH$_2$)$_2^+$, C$_3$H$_5$N$_2^+$, (CH$_3$)$_2$NH$_2^+$, (CH$_3$)$_3$NH$^+$, (CH$_2$)$_3$NH$^+$, CH$_3$CH$_2$NH$_3^+$, CH$_3$CH$_2$CH$_2$NH$_3^+$, (CH$_3$)$_4$N$^+$, (CH$_3$)$_2$CHNH$_3^+$, CH$_3$(CH$_2$)$_3$NH$_3^+$, NH$_4^+$, NH$_2$NH$_3^+$, C$_3$H$_4$NS$^+$, C$_7$H$_7^+$, (CH)$_4$NH$_2^+$, C(CH$_3$)$_2$CH$_2$NH$_2^+$, (CH$_3$)$_2$(CH$_2$)$_2$NH$_2^+$, C(CH)$_5$NH$_3^+$, C$_2$NH$_6^+$, C$_2$OH$_5^+$, C$_3$OH$_7^+$, C$_4$NH$_{10}^+$, C$_4$OH$_9^+$, C$_5$OH$_{12}^+$, C$_6$NH$_{14}^+$, C$_6$OH$_{13}^+$, NC$_4$H$_8^+$, C$_5$NH$_{12}^+$, are used in A$^+$ site, which are shown in the top of Figure 1b. Meanwhile, 9 monovalent B$^+$-site cations (Li, Na, K, Rb, Cs, Cu, Ag, Au, Tl), 49 divalent B^{2+}-site and B$'^{2+}$-site cations (Be, Mg, Sr, Ba, La, Ti, Zr, Hf, V, Nb, Ta, Cr, Mo, W, Mn, Tc, Re, Fe, Ru, Os, Co, Rh, Ir, Ni, Pd, Pt, Cu, Ag, Zn, Cd, Hg, Si, Ge, Sn, Pb, Ca, Eu, Tm, Yb, Sm, P, Sc, As, Y, In, Sb, Au, Tl, Bi) and 35 trivalent B$'^{3+}$ site cations (Sc, Y, La, Ti, Zr, Hf, V, Nb, Ta, Cr, Mo, W, Mn, Tc, Re, Fe, Ru, Os, Co, Rh, Ir, Ni, Pd, Pt, Cu, Ag, Au, B, Al, Ga, In, Tl, As, Sb, Bi) are employed. The list of atoms at B-, B'-, and X-site are shown in the periodic table of Figure 1b. According to the rule of charge neutrality, 180 038 electrically neutral HOIDPs are firstly generated.

Then, we conduct further screening regarding the structural formability of the firstly generated candidates, based on T_f and the octahedral factor (O_f)[58] given by $\frac{R_B}{R_X}$. As mentioned above, to form a stable perovskite structure, generally, T_f and O_f have values of $0.8 \approx 1.2$ and $0.4 \approx 0.7$, respectively.[52–55] Thus, 25983 compounds with structural formability are selected from the 180 038 charge neutral candidates, which can be used to conduct further prediction with ML method.

2.4. Model Inference

2.4.1. *Chemical Stability: Formation Energy*

The comparison of predicted formation energies with those in test dataset for GBR model is shown in **Figure 2**a, which are basically coincident. The inset shows the percent error between predicted $\triangle H^{ML}$ and $\triangle H^{DFT}$ of 95% compounds is less than 10%. The SHAP (Shapley Additive exPlanations) method is used to further interpret our model,[59] which is a generalized metric for feature importance and employ the game-theory-based Shapley values to obtain the contribution of each feature to the model's output. The top 15 features are displayed in Figure 2b, which are listed by the ranking importance, that is, the heat of formation of X site (H_X^{fh}), the electronegativity of X site (χ_x), molar mass (M), the Mendeleev's number of B site (N_B^{men}), and the first ionization energy of B and B' sites (E_B^{ip} and $E_{B'}^{ip}$).

Overall, as shown in Figure 2b, instead of the traditional chemical tuition of T_f and O_f for structural formability of perovskites, the features from B, B' and X sites, that is, H_X^{fh}, χ_x, M, N_B^{men}, E_B^{ip}, and $E_{B'}^{ip}$, rank top six in determining the formation energy, revealing the dominant role of the BX$_6$ and B'X$_6$ octahedrons in the chemical stabilities of A$_2$BB'X$_6$ perovskites. The heat of formation H_X^{fh} and electronegativity χ_x of X-site halogen atom play the key role in $\triangle H^{ML}$, which is consistent with the widely-accepted point that the structural stability of HOIDPs is roughly affected by the geometric criteria T_f and $B - X$ bonding
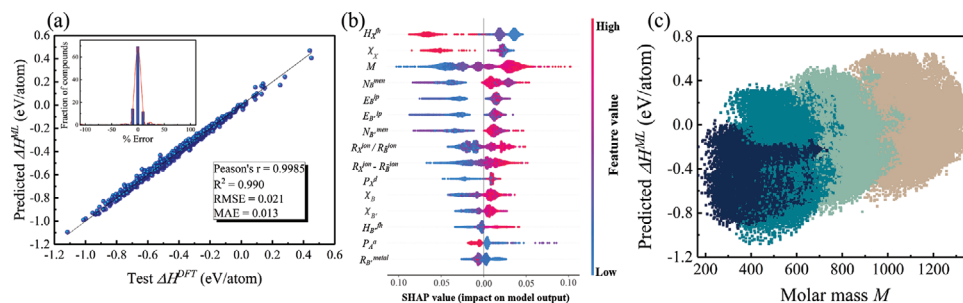
Figure 2. The chart results from formation-energy ML model. a) Actual formation energy per atom $\triangle H^{DFT}$ in the test set and predicted formation energy per atom $\triangle H^{ML}$. Pearson coefficient (r), Coefficient of determination (R^2), mean squared error (MSE), and mean absolute error (MAE) are computed to estimate the prediction errors. The ideal line shown as red line is about ideal prediction result and the fit line shown as dark line is about actual prediction result. These two lines basically coincide. The inset is the fraction of compounds according to their percent error between predicted $\triangle H^{ML}$ and $\triangle H^{DFT}$. The red curve shows the trend. b) The feature importance ranking produced from the gradient boosting regression and SHAP library, showing the elemental properties in descending order of importance. All samples in the dataset are presented and a point in the graph is corresponding to a sample. The x-axis labeled as the SHAP value represents the impact of features on formation energy. The red and blue color indicate high and low values of a given feature, respectively. c) The prediction of formation energy for all electrically neutral candidates with molar mass of compound.

strength,[60] and the latter is usually determined by the difference of electronegativity and distance between bond partners. The X-resolved dependence of $\triangle H^{ML}$ on M are shown in Figure 2c, which reveals that $\triangle H^{ML}$ of X-based HOIDPs possess the relationship of $\triangle H_F^{ML} < \triangle H_{Cl}^{ML} < \triangle H_{Br}^{ML} < \triangle H_I^{ML}$, confirming that stronger electronegativity of halogen atom at X site or lighter molar mass M leads to lower formation energy. Based on the model of formation energy, we can obtain the formation energy of each candidate and reduce the number of candidates to 17 051, by excluding compounds with formation energies larger than -0.2 eV atom^{-1}.

2.4.2. Electronic and Optical Properties: Band Gap

The electronic and optical properties of semiconducting perovskite solar cells are fundamental to understand or further manipulate their interactions with light, and the production, transport, collection, and recombination of photo-generated non-equilibrium carriers. For large absorption coefficients the HOIDPs are required to possess direct band gap, since the optical transition of electrons for indirect-bandgap semiconductors inevitably involves phonons to meet the conservation of momentums, which leads to very weak absorption. Furthermore, the $p \rightarrow p$ transition for electrons from valence band (VB) to conduction band (CB) in direct-bandgap semiconductors is highly beneficial for large absorption coefficient. As we know now, the VB of HOIP, for example, $CH_3NH_3PbX_3$ ($X = F/Cl/Br/I$), is usually formed by the antibonding states of $X - p$ and $B/B' - s$ atomic orbitals, while the CB is formed by the antibonding states of $X - p$ and $B/B' - p$ atomic orbitals,[61] as shown in **Figure 3**a. Since both VB and CB possess p −orbitals, the absorption coefficients for directly semiconducting HOIPs are generally large, beneficial for large PCE.

Here, we only consider HOIDPs with direct band gap. The GBR model for PBE band gap shown in Figure 3b reveals a good match between the predicted band gap and PBE-calculated ones in the test set. By using GBR model to record the feature importance and combining "last-place elimination" algorithm,[35] the feature selection based on importance is achieved, as shown in

Figure 3c. Each lavender dot is R^2 value of GBR model during each selection process, each cyan dot represents MSE value, and each khaki dot denotes MAE value. Obviously, R^2 abruptly decreases, and MSE and MAE abruptly increase when the number of feature is smaller than 32. So we select these 32 features as important features for PBE band-gap model, as shown in Figure S3, Supporting Information.

The top 10 important features are $R_X^{ion} - R_B^{ion}$, χ_B, $\chi_{B'}$, $R_X^{ion}/R_B^{ion} (= 1/O_f)$, M_A, M, O_f, N_X^{proton}, N_A, and P_A, which manifest the key role of B/B'- and X-sites in the formation of band gap, consistent with the aforementioned discussions. Since the VB and CB of HOIDPs are dominantly attributed to the atomic orbitals of B/B' and X atoms, large energy-level differences ΔE between B/B' − s/p orbitals and X−s/p orbitals lead to large band gap, as shown in Figure 3a, Supporting Information. The s/p orbital energies for the halogen elements of F/Cl/Br/I can be calculated by the Rydberg model, that is, $E = \frac{-Z_{eff}^2}{n^2} 13.6 eV$, with Z_{eff} and n representing the effective nuclear charge and principal quantum number. The values of the orbital exponents ζ for neutral halogen atoms, defined as $\zeta = Z_{eff}/n$, are obtained as listed in Table S2, Supporting Information, then the s/p orbital energies can be calculated as $E_{Cl}^{3p} = -56.52 eV > E_{Br}^{4p} = -69.28 eV > E_I^{5p} = -73.33 eV > E_{Cl}^{3s} = -75.49 eV > E_F^{2p} = -88.43 eV > E_F^{2s} = -89.41 eV > E_{Br}^{4s} = -94.66 eV > E_I^{5s} = -97.74 eV$. When $F \rightarrow Cl \rightarrow Br \rightarrow I$, R_X^{ion} and N_X^{proton} increase, as well as the absolute values of the energy-level $E_X^{s/p}$ of X atoms. When the Mendeleev's number of B/B' atoms $N_{B/B'}^{men}$ decreases, generally R_B^{ion} decreases and $\chi_{B/B'}$ increases, and the absolute values of energy $E_{B/B'}^{s/p}$ of B/B' − s/p orbitals decrease. The X-resolved dependences of band gap on $R_X^{ion} - R_B^{ion}$ and M are shown in Figure 3d,e, respectively, which reveal that, as $F \rightarrow Cl \rightarrow Br \rightarrow I$, $R_X^{ion} - R_B^{ion}$ and M increase, and the band gap reduces, resulted from the decrease of energy difference ΔE between B/B' − s/p and X−s/p atomic orbitals. The X-resolved dependence of band gap on octahedral factor O_f is shown in Figure 3f, which reveals that, as O_f increases, indicating enhanced octahederal distortions in HOIDPs, band gap increase, which is due to the reduction of overlap between B/B' − s/p orbitals and X−s/p orbitals by enhanced octahederal
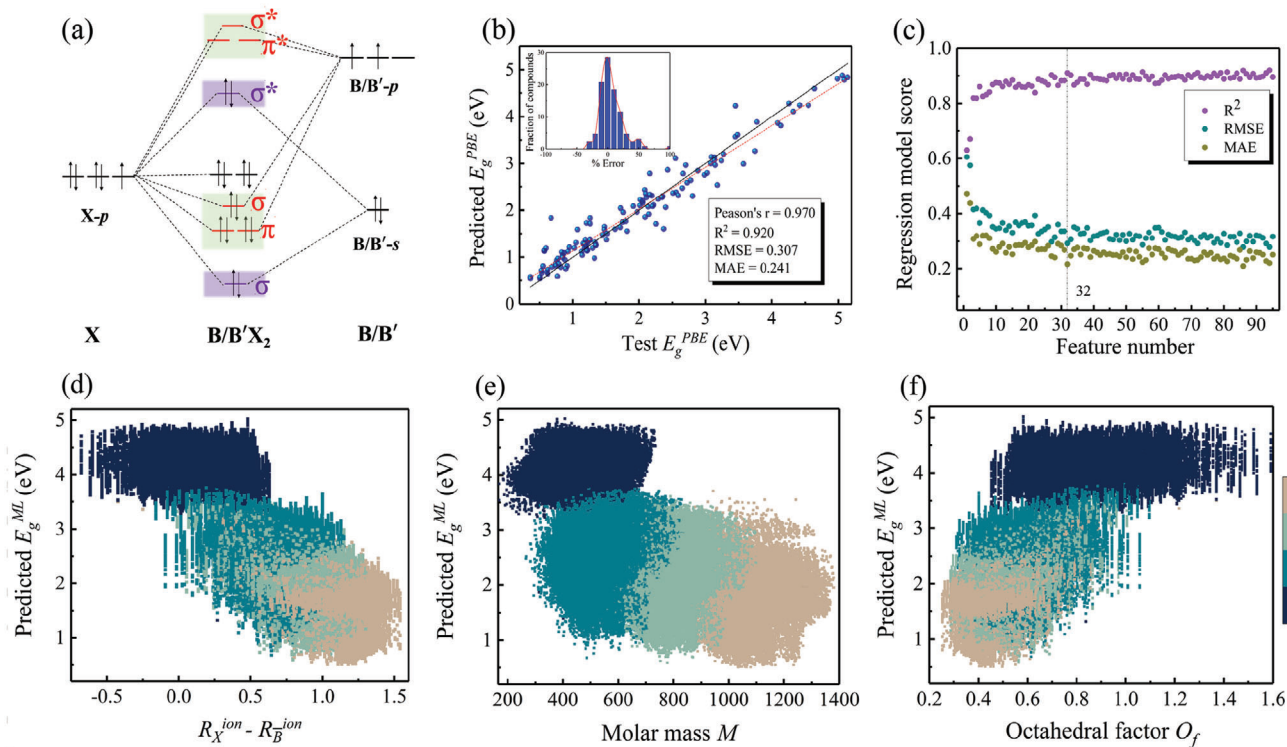
Figure 3. a) Schematic illustration of the band diagram for HOIDPs. The results of GBR model for band gap E_g^{PBE} b) the actual value of band gap E_g^{PBE} by DFT in the test set and the predicted value E_g^{PBE} by ML. Pearson coefficient (r), coefficient of determination (R^2), mean squared error (MSE), and mean absolute error (MAE) are calculated to explain the training performance of the built model. The ideal line is shown as red line and the fit line is shown as dark line. The inset is the fraction of compounds according to their percent error between predicted E_g^{ML} and actual E_g^{DFT}. c) The values of R^2, MSE, and MAE of GBR model with the number of selected features. Relationship visualization of the prediction of band gap for all electrically neutral HOIDP candidates with d) the difference of ion radius between X and B/B′ sites $R_X^{ion} - R_{\bar{B}}^{ion}$, e) molar mass M, and f) octahedral factor O_f, respectively. Different colors represents different X-site halogen elements.

distortion, responsible for widening band gap. Similar analysis can be conducted on the influences from χ_B and $\chi_{B'}$.

To further analyze the correlation among the selected 32 features, the Pearson correlation coefficient is calculated, which can produce the positive and negative correlation between one pair of features, and the results are shown in the inset of Figure S3, Supporting Information. If the correlation value between two features exceeds 0.8, the features with lower importance will be deleted in order to further reduce the redundancy of features. As shown in Figure S4, Supporting Information, through feature pruning process, the total number of features is decreased to 14 and the top 10 important features before pruning is decreased to 6 ($R_X^{ion} - R_{\bar{B}}^{ion}$, χ_B, $\chi_{B'}$, M_A, M, O_f).

In addition, as mentioned above, PBE calculations always underestimate the band gap for semiconductors, therefore, we also select HSE-calculation results to obtain precise band gap for semiconducting HOIDPs and build ML model for HSE band gap. The related prediction performance is shown in Supporting Information, and the calculated R^2, MSE, and MAE are 0.870, 0.384, 0.278, respectively, as shown in Figure S5a, Supporting Information. Compared with E_g^{PBE} model, the top ten important features for ML-models of HSE-bandgaps are M, $R_X^{ion} - R_{\bar{B}}^{ion}$, the Mendeleev's number of B′ and B, that is, $N_{B'}^{men}$ and N_B^{men},

$R_X^{ion} + R_{\bar{B}}^{ion}$, $\chi_{B'}$, χ_B, $R_X^{ion}/R_{\bar{B}}^{ion}$, P_A and O_f as shown in Figure S5b, Supporting Information, which are nearly identical with those for E_g^{PBE} models.

We apply these ML-bandgap models to predict the band gap of 17051 HOIDP candidates screened on basis of formation energy, and 1183 candidates with band gap ranging from 0.6 to 2.2 eV and from 1.1 to 3.0 eV, respectively, for E_g^{PBE} and E_g^{HSE} beneficial for light harvesting and PCE of PSCs are generated.

2.4.3. Thermal Properties: Debye Temperature

Although currently $CH_3NH_3PbX_3$-based PSCs are extensively used for photovoltaics, they not only suffer from the structural instabilities in ambient environment and toxicity due to lead, but also suffer from low lattice thermal conductivity κ_L due to low Debye temperatures Θ_D around 200 K,[62] which may lead to decomposition of the crystals due to heat accumulation. Materials with low Θ_D tend to have low-energy phonon modes, and enhance phonon-phonon scattering processes since over Θ_D all phonons are excited, leading to the decrease of κ_L according to Slack model, and the increase of the probability of non-radiative relaxation time, the latter of which directly decreases PCEs of PSCs. Therefore, the materials with high Θ_D are highly expected

Table 1. Twelve selected HOIDPs with relevant properties, where c and p represent calculated and predicted results, respectively, and $p - E_g^{PBE1}$ and E_g^{PBE2} represent the predicted results from the PBE band gap models of 4456 and 425 HOIDPs, respectively.

Formula	T_f	O_f	$c - \Delta H$	$c - \Theta_D$	$c - E_g^{PBE}/E_g^{HSE}$	$c - \kappa_L$	Direct/Indirect	$p - \Delta H$	$p - \Theta_D$	$p - E_g^{PBE1}/E_g^{PBE2}/E_g^{HSE}$
$(CH_3NH_3)_2AgAlBr_6$	1.02	0.54	-0.42	544.9	1.57 / 2.94	5.49	Direct	-0.41	525.3	0.67 / 1.73 / 2.99
$(CH_3NH_3)_2AgGaBr_6$	1.00	0.56	-0.37	518.3	0.93 / 1.97	5.04	indirect	-0.36	511.0	0.86 / 1.60 / 2.48
$(CH_3NH_3)_2AgInBr_6$	0.97	0.61	-0.38	496.9	0.51 / 1.70	4.39	Direct	-0.36	502.5	0.43 / 0.73 / 2.12
$(HC(NH_2)_2)_2AgGaBr_6$	1.11	0.56	-0.45	521.4	2.35 / 3.84	4.63	Direct	-0.40	525.6	0.65 / 2.59 / 2.94
$(C_2H_6N)_2AgTiBr_6$	1.02	0.56	-0.49	572.9	0.00 /2.49	5.73	Indirect	-0.43	554.9	0.93 / 1.82 / 2.52
$(C_2H_6N)_2AgAlBr_6$	1.03	0.54	-0.47	577.4	1.85 / 3.29	5.74	Indirect	-0.39	555.3	1.61 / 1.78 / 2.97
$(C_2H_6N)_2AgInBr_6$	0.99	0.61	-0.44	535.5	1.13 / 2.48	5.16	Indirect	-0.31	515.0	1.24 / 1.60 / 2.47
$(C_2OH_5)_2AgAlBr_6$	1.04	0.54	-0.99	545.1	1.86 / 3.20	5.37	Indirect	-0.70	523.7	1.42 / 1.68 / 2.93
$(C_2H_6N)_2TiTiBr_6$	1.13	0.41	-0.75	628.6	0.00 / 1.67	6.96	Indirect	-0.62	591.2	0.93 / 1.80 / 2.01
$(C_2H_6N)_2TiMnBr_6$	1.12	0.43	-0.35	618.4	0.00 / 2.62	6.53	Indirect	-0.31	586.6	0.93 / 1.78 / 2.40
$(C_2H_6N)_2TiZnBr_6$	1.10	0.45	-0.48	581.2	0.00 / 1.70	5.77	Indirect	-0.41	571.6	0.85 / 1.59 / 1.94
$(C_2H_6N)_2TiGeBr_6$	1.15	0.39	-0.45	578.9	0.00 / 1.02	5.59	Indirect	-0.27	560.8	0.82 / 1.57 / 1.87

for the design of high-efficiency and highly stable PSCs. According to the formulae for Θ_D shown as Equation (15), Θ_D value is related to the bond strength in crystals.

Based on the high-throughput calculation of Debye temperatures for 425 HOIDPs, GBR model predicts Θ_D^{ML} well matched with the calculated Θ_D^{DFT} as shown in Figure S7a, Supporting Information. Based on the feature-importance selection process as shown in Figure S6b, Supporting Information, 28 important features are selected as shown in Figure S7b, Supporting Information. The top ten important features are χ_X, $R_X^{ion} + R_{\bar{B}}^{ion}$, E_X^{ip}, M, R_X^{ion}, T_f, N_X^{proton}, E_X^{pa}, N_X^{period}, and R_X^{s+p}, which manifest the important role of corner-sharing B/B'X$_6$ octahedra in determining Θ_D, and five of them are sharing with those for the ML model of formation energy. We also calculate the Pearson correlation value between Debye temperatures and formation energies, and the calculated result is 0.42, indicating the high correlation between Θ_D and ΔH, which is due to the similar underlying mechanism of $B/B' - X$ bond strength for Θ_D and ΔH.

The X-resolved dependence of Θ_D on χ_X, M, T_f, and O_f is shown in Figure S8a– d, Supporting Information, respectively. As shown in Figure S8a,b, Supporting Information, as $F \rightarrow Cl \rightarrow Br \rightarrow I$, the value of χ_X decreases, and as a result $B/B' - X$ bonds weaken, as well M increases, leading to the decrease of Θ_D. Similar tendency regarding X-site contribution to Θ_D can be observed in Figure S8c,d, Supporting Information. Different from those in case of ΔH, the Goldsmith tolerance factor T_f, which indicates the deviation of HOIDP structures from cubic crystal, also plays an important role in determining Θ_D, as shown in Figure S8c, Supporting Information, which reveals that, as T_f increases, Θ_D increases. According to the ML model for Θ_D, we select 597 HOIDPs from 1183 candidates, whose Debye temperature is higher than 500 K (much larger than Θ_D of $CH_3NH_3PbX_3$ crystal).

3. Discussions

It is worth to mention that three compounds ($(CH_3NH_3)_2AgSbI_6$, $(CH_3NH_3)_2AgBiBr_6$, and $(CH_3NH_3)_2TlBiBr_6$) from 1183 HOIDP

candidates screened from the criteria of formation energy and band gap have been experimentally synthesized, and the predicted HSE-bandgaps for them are quite close to the experimentally observed ones, as shown in Table S3, Supporting Information, validating the precision of our ML models. After screening the HOIDP candidates with the chemical space of 180 038 by formation energy less than -0.2 eV atom^{-1}, PBE/HSE band gap ranging from 0.6/1.1 to 2.2/3.0 eV and Debye temperatures larger than 500 K, 597 HOIDP candidates are subsequently selected, which are probable to possess chemical stability, appropriate band gap and good thermal transport property. Furthermore, considering selecting non-toxic elements and accessible Br-based HOIDPs in experiments,[63] 12 rhombohedral Br-based HOIDPs are finally selected for further detailed investigations on their electronic, optical, photovoltaic, and thermal properties by DFT calculation.

Since different elements occupy the vertices of the regular corner-sharing BX$_6$ and B'X$_6$ octahedra and the lone-pair electrons of the B/B'-site ions leads to the coordination symmetry breaking, all these 12 undetected HOIDPs have different degrees of distorted BX$_6$/B'X$_6$ octahedra, although they possess typical perovskite structures as listed in Table S4, Supporting Information. The calculated octahedral factors O_f for these 12 HOIDP candidates are listed in **Table 1**, which reveals a larger O_f for Ag-based HOIDPs compared to Ti-based HOIDPs. The calculated formation energies for them are smaller than -0.35 eV atom^{-1}, most of which are well smaller than those of $CH_3NH_3PbI_3$ (-0.39 eV atom^{-1}[32]), confirming their chemical stabilities.

The PBE/HSE calculations of band gap for these 12 HOIDP candidates are listed in Table 1 as well, which are in good agreement with ML-predicted ones for Ag-based HOIDP candidates, but PBE calculations give zero band gap for Ti-based HOIDP candidates. However, HSE calculations give band gap well matched with those predicted by HSE ML-model for all these 12 HOIDPs. By using the criterion of HSE band gap ranging from 1.1 to 3.0 eV and excluding Ti-based HOIDPs, four HOIDPs of $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ are selected in the end as shown in
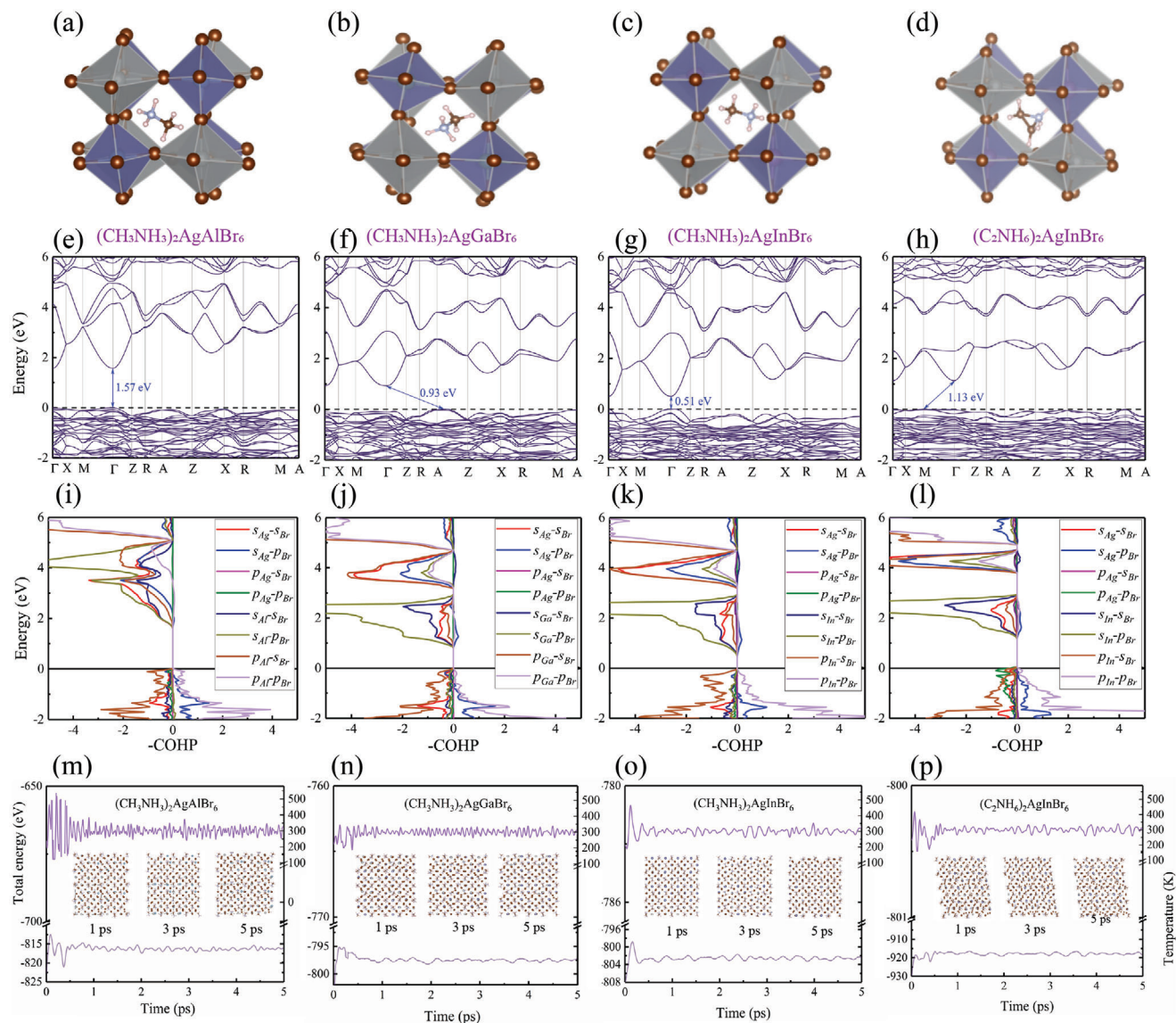
ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED
SCIENCE
Open Access
www.advancedscience.com

Figure 4. a–d) DFT optimized crystal structure, e–h) the calculated electronic band structures, i–l) orbital-resolved-pCOHP, m–p) the variation of total energy and crystal structure during 5 ps AIMD simulations at room temperature for selected four HOIDPs $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$, respectively.

Figure 4a–d, respectively. The PBE-calculated bandstructures for these four HOIDPs are shown in Figure 4e–h, respectively, which reveals that, $(CH_3NH_3)_2AgAlBr_6$ and $(CH_3NH_3)_2AgInBr_6$ possess direct band gap, and the HSE-calculations of band gap give 2.94 and 1.70 eV, respectively, well consistent with ML-predicted ones as listed in Table 1. Since the difference of the band gap for these two HOIDPs is large, therefore the band gap for $(CH_3NH_3)_2AgAl_xIn_{1-x}Br_6$ can be tunable by manipulating the Al ratio according to the bowing effect.[64] In addition, the hole or electron effective mass of these 12 HOIDP candidates in $x/y/z$ direction are different as listed in Table S5, Supporting Information, and the average of hole effective mass is usually higher electron's except $(CH_3NH_3)_2AgGaBr_6$, $(C_2NH_6)_2TiTiBr_6$, and $(C_2NH_6)_2TiZnBr_6$, which indicates the hole extraction in transport layer is more difficult and employing excellent hole

transport layer will lead to high photovoltaic performance in the actual PSC design.

To further reveal the chemical bonding states for VB and CB in $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$, we also use the crystal orbital Hamilton population (COHP) method proposed by Bloechl et al.[65] to calculate their Hamiltonian elements with positive/negative values indicating the bonding/antibonding states between chosen atomic orbitals. The calculated orbital-resolved -pCOHP for these four HOIDPs are shown in Figure 4i–l, which reveals that the CBs in all these HOIDPs are dominantly formed by the antibonding states of $B' - s$ and $Br - p$, $B' - s$ and $Br - s$, and $B - s$ and $Br - s$, with $B = Ag$ and $B' = Al/Ga/In$, and the VBs for these four HOIDPs are dominantly formed by the antibonding states of $B' - p$ and $Br - s$. Similar to the case in $CH_3NH_3PbI_3$ men-

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
SCIENCE**
Open Access

www.advancedscience.com

tioned above, the absorption coefficients of these four HOIDPs will be beneficial from the p orbitals involved in their CBs and VBs. In addition, we also calculate the electron localization function (ELF), which is position-dependent with the value ranging from 0.0 to 1.0. The value of unity for ELF reflects maximum probability to find the localized electrons, ELF = 0.5 corresponds to the electron-gas-like pair behavior for electrons, and the value less than 0.5 of ELF means electrons are emptied from this position. The calculated ELFs for these four HOIDPs are shown in Figure S9, Supporting Information, which reveals that electrons are emptied from the B/B' metal-atom sites and accumulated at the X halogen-atom sites, indicating that $B/B' - X$ bonds are essentially ionic.

We also conduct AIMD simulations to investigate the thermal stability of these four HOIDPs at 300 K for the total simulation time of 5 ps with a time interval of 1 fs, as shown in Figure 4m–p, which reveal that during simulation time, their time-dependent total energies fluctuate in a narrow range and the crystal structures remain integrated, indicating their thermal stabilities at 300 K. Furthermore, for the results of time-dependent evolution of XRD simulations, as shown in Figure S10, Supporting Information, no shift diffraction peaks appear, which further confirms that these selected HOIDPs have no change in structure and keep their structural integrity in the process of simulations. In addition, we calculate the general tolerance factor proposed by Filip et al.[66] for these four HOIDPs to check the limits of stretch, octahedral, tilt, and chemistry shown in Table S6, Supporting Information, which indicates that $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ meet the stability criteria for double perovskites. Furthermore, the H_2O adsorption energy of four selected HOIDPs estimated by $E_{ads} = E_{slab+H_2O} - E_{slab} - E_{H_2O}$, with E_{slab+H_2O}, E_{slab}, and E_{H_2O} the total energies of H_2O adsorbed perovskite slab, perovskite slab, and the isolated H_2O molecule, respectively, is calculated to check the stabilities against H_2O. The H_2O molecule is adsorbed along the z-direction of perovskite slab with the vacuum thickness of 18 $\overset{\circ}{A}$. Comparing with $CH_3NH_3PbI_3$ whose H_2O-adsorption energy is -0.48 eV, $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, and $(CH_3NH_3)_2AgInBr_6$ show better environmental stability against H_2O with E_{ads} of -0.32, -0.42, and -0.04 eV, respectively. H_2O-adsorption energy for $(C_2NH_6)_2AgInBr_6$ is -1.87 eV.

For accurate calculation of device performance of PSCs, for example, PCE, we also calculate the quasi-particles (QP) band structures for excited states based on many-body perturbation theory (MBPT) and obtain the optical absorption by considering many-body electron–hole interactions. We use one-shot G_0W_0 method to calculate the QP states based on the calculated Kohn–Sham (KS) orbitals implemented in Quantum Espresso software package,[67,68] and on top of QP states, the electron–hole excitation states for these four HOIDPs can be calculated using the G_0W_0+BSE method combined with the Tamm–Dancoff approximation method, which are implemented in Yambo.[69] By the summation of respective contribution from a large number of electron-hole pairs according to Equation (9), the imaginary part of the dielectric constant ϵ_2 can be obtained, and the absorption can be subsequently calculated as well.

The calculated QP band gap for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and

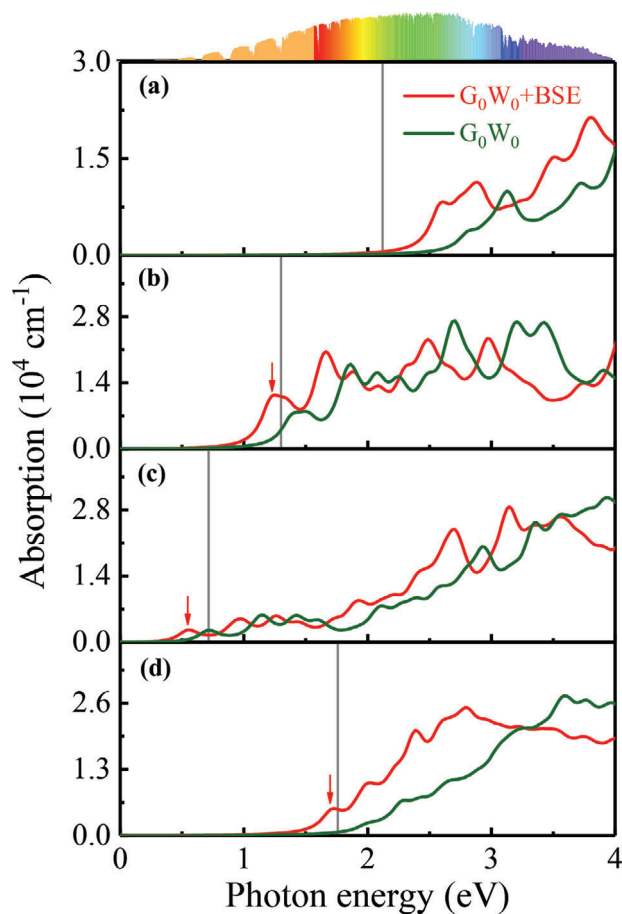**Figure 5.** Energy-dependent optical absorption calculated by G_0W_0+BSE of selected four HOIDPs a) $(CH_3NH_3)_2AgAlBr_6$, b) $(CH_3NH_3)_2AgGaBr_6$, c) $(CH_3NH_3)_2AgInBr_6$, and d) $(C_2NH_6)_2AgInBr_6$. The gray line indicates the QP bandgap.

$(C_2NH_6)_2AgInBr_6$ are 2.12, 1.30, 0.71, and 1.76 eV, respectively. As shown in **Figure 5**, the calculated absorption with and without considerations of electron–hole interactions for the selected four HOIDPs, are plotted by red and green lines, respectively. Obviously, the contribution from electron–hole interactions influences the optical absorption significantly, and no excitonic absorption peak can be observed in $(CH_3NH_3)_2AgAlBr_6$ crystal, as shown in Figure 5a. The absorption spectra for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ cover the ultra-broadband region ranging from 180/163/180/143 nm to 496/1127/2480/775 nm, respectively. Notably, the absorption spectra for all the four HOIDPs cover the major part of the solar-energy spectrum as shown in Figure 5, especially for $(CH_3NH_3)_2AgGaBr_6$ and $(C_2NH_6)_2AgInBr_6$, indicating their abilities to efficiently harvesting the solar energy by photovoltaic effects similar to $CH_3NH_3PbI_3$. The excitonic peaks within the QP bandgap for $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ locate at 1.22, 0.54, 1.70 eV, respectively, denoted by red arrows as shown in Figure 5, and the corresponding excitonic binding energies defined as the energy difference between the excitonic eigenenergy and QP bandgap, are 0.08,
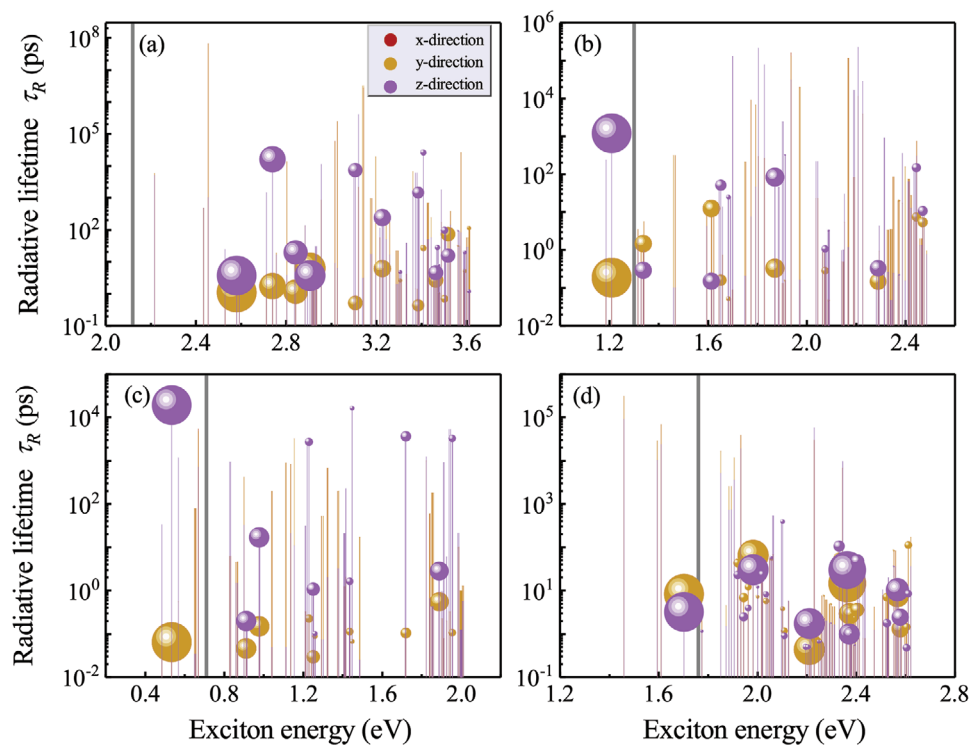
Figure 6. Exciton lifetimes induced by many-body interactions for selected HOIDPs. The gray line is QP band gap and the circle radius shows the contribution from exciton to excitonic absorption peak.

0.17, and 0.06 eV, respectively, which are comparable to 0.05 eV for $CH_3NH_3PbI_3$ crystal.[70] Among them, the relatively large binding energy of 0.17 eV in $(CH_3NH_3)_2AgInBr_6$ is observed, implying the potential applications for light-emission devices. The vanishing excitonic absorption peaks or small excitonic binding energies for these four HOIDPs imply easy dissociation of photogenerated excitons, which is beneficial for high-performance PSC devices. The calculated energy-dependent optical absorption of $CH_3NH_3PbI_3$ is shown in Figure S12, Supporting Information, for comparison. In addition, we also conduct the analysis on the selection rule for optical transition based on the group-theory argument for four HOIDP candidates which is provided in Tables S7-S10, Supporting Information. Actually, for the four candidates, the transitions of valence-band electrons at some high-symmetric points in the Brillouin zone are allowed, and some are forbidden, which means that, not all the optical transition between valence bands and conduction bands are forbidden restricted by the parity-induced dipole-forbidden rule. Although the optical transition between VBM and CBM is forbidden in some double perovskites, those transitions around the band edge are still allowed due to the C_1 symmetry, thus the optical properties can be still significantly large.

As we know, the creation of the tail states in the band edge of materials generates serious increase of open-circuit voltage losses V_{loss}, which can be characterized quantitatively through the evaluation of absorption tail, expressed by $\alpha = \alpha_0 + e^{\left(\frac{E}{E_u}\right)}$, where α is the absorption coefficient, E is the photon energy, and E_u is the Urbach energy.[71] The Urbach energy is presumed as the width of the tail of localized defect states in the band gap. It has

already been reported that E_u shows a direct correlation with V_{loss} and the smaller E_u (i.e., sharper absorption edge) is beneficial to suppressing V_{loss}. Based on the absorption curves calculated by the precise G_0W_0+BSE method for the four selected candidates, we calculate their Urbach energies E_u by fitting the reciprocal of the slope of $ln(\alpha)$, which gives 58.0, 18.3, 18.1, and 28.9 meV for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$, respectively, as shown in Figure S13, Supporting Information, in comparison to the case for $CH_3NH_3PbI_3$ as shown in Figure S14, Supporting Information, giving E_u of 21.9 meV consistent with the experimental result of 11.3 meV.[72]

The calculated exciton-resolved radiative combination rates τ_R at room temperature according to Equation (12) for these four HOIDPs are shown in **Figure 6**, in which the circle radius is proportional to the oscillator strength for the corresponding exciton. The stronger oscillator strength leads to larger contribution from the exciton to the excitonic absorption peak. As shown in Figure 6, τ_R is isotopic irrespective of x- or y-polarization but z-polarization is different, and the τ_R values for excitons with large oscillator strength range from 10 ps to 10 ns, comparable to 14 ns for $CH_3NH_3PbI_3$ and 37 ns for $CH_3NH_3PbBr_3$.[73]

According to Equation (20), the calculated short-circuit currents J_{sc} for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ are 7.0, 23.4, 53.2, 20.8 mAcm^{-2}. Then, according to Equation (18), the calculated PCEs for these four HOIDPs are 11.4%, 20.6%, 19.9%, and 27.6%, respectively. The relatively smaller PCE for $(CH_3NH_3)_2AgAlBr_6$ compared to the rest is probable due to the

smaller overlap of the absorption spectrum and solar-energy spectrum as shown in Figure 5a.

In hybrid perovskites, since the materials are generally soft and polar, charged carriers are generally dressed in phonons resulted from large dielectric electron-phonon couplings, forming the so-called polarons. Recent experimental and theoretical studies have shown that the coupling between polar longitudinal optical (LO) phonon and carriers, that is, the Fröhlich interaction, dominates the charge carrier dynamics in hybrid perovskites at room temperatures, and the low-frequency LO phonon vibrations, namely, X-B-X bond bending and stretching motions, contribute dominantly to the Fröhlich interaction.[74] The Fröhlich coupling strength is defined as, $\alpha_F = \frac{e^2}{\hbar}(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0})\sqrt{\frac{m}{2\hbar\omega}}$ where ϵ_∞ and ϵ_0 are the high-frequency and static dielectric constant, and ω is the LO-phonon frequency. In the weak coupling regime, that is, $\alpha_F \leqslant 2$, large polarons form. We calculate the electron and hole α_F for the four selected HOIDP candidates and the results are listed in Table S11, Supporting Information, which shows that, their hole α_F (3.35, 5.54, 2.80, and 2.71 for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$, respectively) are larger than $CH_3NH_3PbI_3$ of 2.67,[75] and their electron α_F (1.48, 4.10, 2.32, and 1.60 for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$, respectively) are comparable to $CH_3NH_3PbI_3$ of 2.39,[75] revealing the formation and existence of large or intermediate polarons in the four selected HOIDP candidates, similar to $CH_3NH_3PbI_3$. To further investigate polaron behaviors in the predicted HOIDP candidates, we also calculate their polaron masses and polaron radii, and the results are listed in Table S11, Supporting Information.

Furthermore, we calculate the polaron mobility by solving the low-field temperature-dependent mobility equation with the FHIP, Kadanoff and Hellwarth methods, implemented in the *PolaronMobility.jl* code developed by J. M. Frost.[76] The results are listed in Table S11, Supporting Information, which reveal that the electron polaron mobilities for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ are 90, 8, 53, 81 $cm^2V^{-1}s^{-1}$, respectively, which are smaller than that of $CH_3NH_3PbI_3$ of 197 $cm^2V^{-1}s^{-1}$ but comparable to that of double-perovskite $Cs_2AgBiBr_6$ of 34 $cm^2V^{-1}s^{-1}$ calculated by the Kadanoff method.[77] The hole polaron mobilities for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ are 5, 2, 27, 12 $cm^2V^{-1}s^{-1}$, respectively, which are similarly smaller than that of $CH_3NH_3PbI_3$ of 133 $cm^2V^{-1}s^{-1}$, but comparable to that of double-perovskite $Cs_2AgBiBr_6$ of 42 $cm^2V^{-1}s^{-1}$.[77] In addition, the carrier mobility induced by deformation potential scattering due to long-wavelength acoustic phonons implemented by the ab initio scattering and transport AMSET code[78] for the four HOIDP candidates is also calculated and the results are shown in Table S11, Supporting Information, which reveals that, in comparison to inorganic $Cs_2AgBiBr_6$, for the four selected candidates, the deformation potential scattering limited carrier mobilities are larger by two to three orders than those limited by Fröhlich coupling, especially for electrons. Therefore, the Fröhlich coupling dominates in the four selected HOIDP candidates.

As listed in Table 1, the calculated Debye temperatures for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ are 544.9, 518.3, 496.9, and 535.5 K, respectively, much larger than those in $CH_3NH_3PbI_3$ crystal, which is probable due to the stiffer properties and larger phonon velocities of these four HOIDPs, since they have larger optical phonon frequencies resulted from smaller densities as listed in Table S5, Supporting Information. The corresponding calculated Grueneisen parameters γ are 1.28, 1.28, 1.33, and 1.31, respectively, which are larger than 1.17 experimentally reported in $CH_3NH_3PbI_3$ crystal,[62] implying enhanced anharmonic interactions among phonons in these four HOIDPs. According to the aforementioned Slack model, the calculated thermal conductivities κ_L at 300 K for $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$ are 5.49, 5.04, 4.39, and 5.16 $Wm^{-1}K^{-1}$, which are much larger than experimentally reported 0.5 $Wm^{-1}K^{-1}$ at room temperature in $CH_3NH_3PbI_3$ crystal,[20] indicating their good thermal transport properties beneficial for heat dissipation. Therefore, although the anharmonic interactions in these four HOIDPs are larger than $CH_3NH_3PbI_3$ crystal, the much larger Debye temperature resulted from larger stiffness overrides the anharmonic interactions, leading to the larger thermal conductivities by an order in $(CH_3NH_3)_2AgAlBr_6$, $(CH_3NH_3)_2AgGaBr_6$, $(CH_3NH_3)_2AgInBr_6$, and $(C_2NH_6)_2AgInBr_6$.

4. Conclusion

By combining high-throughput DFT calculation with ML technology, we developed a multi-step material screening scheme to accelerate the discovery of hybrid organic-inorganic double perovskite material $A_2BB'X_6$. 597 stable HOIDP solar cell materials with suitable band gap and high Debye temperature were successfully selected from 180038 compounds. After that, based on the conditions of non-toxicity and experimental accessibility, 12 candidate materials were selected, and four of them pass the verification of DFT with excellent properties. More importantly, with the help of ML technology, the prediction results of our HSE band gap model can achieve the level of results of experimental preparation, and the multi-stage scheme of band gap prediction could ensure the validity of final candidates. The structure property relationship under different targets is established by full use of input data, which provides a further understanding of material properties. After DFT verification, the screened-out candidates possess excellent properties.

As a new generation of material design strategy, ML driven scheme can achieve high-precision material discovery in a "cheap" way without deep physical and chemical knowledge, just based on existing data and appropriate algorithms. Meanwhile, ML technology can capture the structure property relationship hidden in the data, which provides a unique way to understand the complex material properties, so it can help researchers jump out of the framework of known knowledge and find more suitable descriptors. In addition, multi-objective screening can effectively improve the screening efficiency, and the number of targets is not limited. However, the limitations of ML-based material design method also exist, such as the prediction accuracy depends on the reliability of dataset, which can be reflected from our two PBE band gap models. It is full of challenges to obtain the dataset with consistent data samples but also to ensure the diversity of samples. We choose the mandatory way that the predicted values

of the two models both need to meet the requirement at the expense of the number of final candidates. A more flexible solution could be explored in the future research.

5. Methodology

5.1. Gradient Boosted Regression

Gradient boosted regression (GBR) is one of the tree-based machine learning algorithms, which is based on the gradient boosting method and realized in the open-source Scikit-learn package.[79] It is a single strong learner $f(x)$ through combining many weak learners step by step using the gradient descent algorithm. It means that GBR usually find the gradient of $f(x)$ which can minimize the average value of loss function to achieve the improvement of the accuracy of regression results. First, the model is initialized with a constant and the specific expression is as follows:

$$f_0(x) = \underset{c}{argmin} \sum_{i=1}^{m} L(\gamma_i, c) \tag{1}$$

here x and γ_i are the input and output values in the training set, c is the step size of gradient learning, and m is the number of data samples. Basic learner $h_j(x)$ needs to fit the negative gradient $\tilde{\gamma}_j(x_i)$ of the loss function $L(\gamma_i, f(x_i))$ for n ($j < n$) times

$$\tilde{\gamma}_j(x_i) = - \left[\frac{\partial L(\gamma_i, f(x_i))}{\partial f(x_i)} \right]_{f(x) = f_{j-1}(x)} \tag{2}$$

The next optimization function is solved to get each c, and the equation is as follows:

$$c_j = \underset{c}{argmin} \sum_{i=1}^{m} L(\gamma_i, f_{j-1}(x_i) + ch_j(x_i)) \tag{3}$$

After that, the model is obtained with the following form:

$$f_j(x) = f_{j-1}(x) + c_j h_j(x) \tag{4}$$

Therefore, the final model $f_n(x)$ is applied to predict.

5.2. Hyper-Parameters Selection

In ML algorithm, in order to improve the efficiency and generalization performance of model, the hyper-parameters of each ML algorithm should be selected before starting the formal training process. The optimization of hyper-parameters is very important and this can be completed through a cross-validated grid search or a randomized search over the parameter setting. For GBR algorithm, the related hyper-parameters are loss function, learning rate, the number of boosting stages to perform, maximum depth of the individual regression estimators, the minimum number of samples required to be at a leaf node, and the number of features to consider when looking for the best split. When many hyper-parameters exists together, the speed of traditional way of optimizing hyper-parameters will be slow. A global

search algorithm based on the simulated annealing algorithm is used to solve this problem, which is an open-source python module called hyperopt.[80] Through this module, the speed of hyper-parameters search can be greatly improved. This search method is also applied to other algorithms involved in this article. The prediction results of our GBR model are the average values through ten cross-validations after 100 iterations.

5.3. DFT Calculations

The calculations of material properties are performed using the Vienna ab initio simulation package (VASP), which is based on DFT.[81] The pseudopotential used to describe the interaction between valence electrons and core charges is based on the projector-augmented wave (PAW) method. We use the generalized gradient approximation (GGA) in parametrization of Perdew–Burke–Ernzerhof (PBE) version to describe the exchange-correlation functional, with a kinetic energy cutoff of 600 eV. The K points in the Brillouin zone are sampled with the smallest k-mesh interval of 0.2 under the Monkhorst-Pack scheme. The convergence is set as 1×10^{-5} eV in total energy for two consecutive electronic steps and 0.01 eV Å^{-1} for maximum Hellman–Feynman force in the crystal. To eliminate the well-known error in PBE calculation, we also use the Heyd–Scuseria–Erzenhof (HSE) hybrid functional method to improve the accuracy of the calculation of band gap for semiconductors. The hole and electron effective mass tensors, $\mathbf{m}_h^*$ and $\mathbf{m}_e^*$ are expressed as the double partial differential of the energy band $E(\mathbf{k})$ at the valence band maximum (VBM) and conduction band minimum (CBM), respectively:[82]

$$\frac{1}{m_{h,i,j}^*} = \frac{1}{\hbar^2} \frac{\partial^2 E(\mathbf{k})}{\partial k_i \partial k_j} \bigg|_{\mathbf{k} = VBM} \quad (i, j = x, y, z) \tag{5}$$

$$\frac{1}{m_{e,i,j}^*} = \frac{1}{\hbar^2} \frac{\partial^2 E(\mathbf{k})}{\partial k_i \partial k_j} \bigg|_{\mathbf{k} = CBM} \tag{6}$$

And the averages of diagonal terms in each tensor are used as the hole and electron effective mass, respectively.

In order to investigate the thermodynamic stability of candidates, we also carry out the ab initio molecular dynamics simulation (AIMD) in canonical ensemble to simulate the evolution of crystals at room temperature using the NOSè–Hoover method,[83] where the NVT set is adopted and $2 \times 1 \times 2$ supercell is constructed. A total of 5-ps AIMD simulation is implemented with the time interval of 1 fs.

To investigate the optical properties of the candidates, we imply the G_0W_0+BSE method implemented by Yambo code to calculate the absorption spectra of those candidates.[84,85] The Kohn sham states are obtained using $5 \times 5 \times 5$ Monkhorst-Pack mesh on top of 200 empty bands. The quasiparticle (QP) states containing eight highest valance bands and eight lowest conduction bands are used in the Bethe–Salpeter equation kernel written as

$$\left(E_{c\mathbf{k}}^{QP} - E_{v\mathbf{k}}^{QP} \right) A_{vc\mathbf{k}}^s + \sum_{v'c'\mathbf{k}'} \left\langle vc\mathbf{k} \left| K^{eh} \right| v'c'\mathbf{k}' \right\rangle A_{v'c'\mathbf{k}'}^s = \Omega_S A_{vc\mathbf{k}}^s \tag{7}$$

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
SCIENCE**
Open Access

www.advancedscience.com

where A_{vck}^s is the amplitude of the Sth excitonic wave function, Ω_S is the corresponding excitonic eigenenergy, and K^{eh} is the electron–hole interaction kernel. On the basis of the Tamm–Dancoff approximation, the exciton wave function in real space can be written as

$$\Psi(\mathbf{r}_e, \mathbf{r}_h) = \sum_{\mathbf{k}, v, c} A_{vck}^s \psi_{\mathbf{k}, c}(\mathbf{r}_e) \psi_{\mathbf{k}, v}^*(\mathbf{r}_h) \tag{8}$$

The imaginary part of the dielectric function ϵ_2 can be produced by summarization over all the exciton eigenstates,

$$\epsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_S |e \times \langle 0|v|S\rangle|^2 \delta(\omega - \Omega_S) \tag{9}$$

where $e \times \langle 0|v|S\rangle$, and v is velocity operator of the incident photons along polarization direction. The real part of the dielectric function ϵ_1 can be obtained by the well-known Kramers–Kronig relation as follows:

$$\epsilon_1(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\epsilon_2(\omega')\omega'}{\omega'^2 - \omega^2 + i\eta} d\omega' \tag{10}$$

Based on the calculated frequency-dependent complex dielectric function, the absorption coefficient $\alpha(\omega)$ can be calculated by as follows:

$$\alpha(\omega) = \frac{\sqrt{2}\omega}{c} \left\{ \left[\epsilon_1^2(\omega) + \epsilon_2^2(\omega) \right]^{1/2} - \epsilon_1(\omega) \right\}^{1/2} \tag{11}$$

The radiative decay rate of a given exciton state S at temperature T for bulk crystal under the assumption that the exciton momentum has a thermal equilibrium distribution, can be calculated as follows:[86]

$$\gamma_S(T) = \tau_S^{-1} = \frac{8\sqrt{\pi\epsilon}e^2\hbar p_S^2}{3\epsilon_0 m^2 V E_S(0)^2} \left(\frac{E_S(0)^2}{2M_S c^2 k_B T} \right)^{3/2} \tag{12}$$

where the exciton energy $E_S(0)$ and the transition dipole $\mathbf{p}_S$ (and $p_S^2 = |\mathbf{p}_S|^2$) are produced by solving the BSE, and M_S denotes the exciton mass defined by $M_S = m_e^* + m_h^*$.

To investigate the coupling of carriers to acoustic phonons, we calculate the carrier mobility induced by deformation potential scattering implemented by the ab initio scattering and transport AMSET package, in which the electron-acoustic phonon scattering matrix element is given by[78]

$$g_{nm}^{ADP} = \left\langle m\mathbf{k} + \mathbf{q} | \mathbf{S_q} : (\mathbf{D}_{nk} + v_{nk} \bigotimes v_{nk}) | n\mathbf{k} \right\rangle \tag{13}$$

where $|n\mathbf{k}\rangle$ and $|m\mathbf{k} + \mathbf{q}\rangle$ denote the initial and final states of electrons, respectively, $\mathbf{S_q}$ is the strain associated with an acoustic phonon $\mathbf{q}$, v_{nk} is the group velocity for electron $|n\mathbf{k}\rangle$, and $\mathbf{D}_{nk}$ is the second rank deformation potential tensor.

5.4. Formation Energy and Debye Temperature

For HOIDPs with the formulae of $A_2BB'X_6$ under investigations here, the average formation energy per atom ($\triangle H$) is defined as

$$\triangle H = \left(E_{A_2BB'X_6} - 2E_A - E_B - E_{B'} - 3E_{X_2} - E_{H_2} \right)/N \tag{14}$$

where N is the number of atoms in the unit cell of HOIDPs compounds; $E_{A_2BB'X_6}$, E_A, E_B, $E_{B'}$, E_{X_2}, and E_{H_2} are the total energies of HOIDPs unit cell, the isolated neutral organic molecule A, a single B atom and a single B' atom in the corresponding elemental crystals, the isolated X_2, and H_2 molecules, respectively. For the group of tetramethylammonium cation ($C_4H_{12}N^+$), the energy of neutral trimethylamine (C_3H_9N) is used for E_A, and the energies of the molecules C_2H_6 is used instead of E_{H_2}.

To evaluate the thermal stabilities and thermal conductivities, we also investigate Debye temperature Θ_D for $A_2BB'X_6$-HOIDPs, which can be estimated on the basis of the results of the average sound velocity (v_m), as follows:

$$\Theta_D = \frac{h}{k_B} \left[\frac{3n}{4\pi} \left(\frac{N_A\rho}{M} \right) \right]^{\frac{1}{3}} v_m \tag{15}$$

where h is Planck constant, k_B is Boltzmann constant, n is the number of atoms per formula unit, N_A is Avogadro constant, ρ is the crystal structure's density, M is the molar mass, and v_m is the average sound velocity. In a poly-crystalline material, v_m can be approximated as

$$v_m = \left[\frac{1}{3} \left(\frac{2}{v_T^3} + \frac{1}{v_L^3} \right) \right]^{-\frac{1}{3}} \tag{16}$$

here, v_L and v_T are the longitudinal and transverse sound velocity, respectively, which can be calculated from the bulk (B) and shear (G) moduli according to

$$v_L = \left(\frac{B + \frac{4G}{3}}{\rho} \right)^{\frac{1}{2}}, v_T = \left(\frac{G}{\rho} \right)^{\frac{1}{2}} \tag{17}$$

5.5. Device Performances for Photovoltaic PSCs

The key parameter to characterize the photovoltaic devices is PCE, also called as spectroscopic limited maximum efficiency, that, η,[87] which is defined as the ratio of the maximum output power density (P_{max}) and the total incident solar energy density (P_{in}), as follows,

$$\eta = \frac{P_{max}}{P_{in}} = \frac{max\left\{ (J_{sc} - J_0(e^{eV/k_BT} - 1))V \right\}_V}{\int_0^\infty EI_{sun}(E)dE} \tag{18}$$

where P_{max} is obtained by numerically maximizing the product of current density J and voltage V. We assume that the solar cell is illuminated under the photon flux I_{sun} and can be approximated as an ideal diode. At temperature T, the current density J and voltage V follow

$$J = J_{sc} - J_0(e^{eV/k_BT} - 1) \tag{19}$$

where J_{sc} is the short-circuit current density, and

$$J_{sc} = e \int_0^\infty \alpha(E) I_{sun}(E)dE \tag{20}$$

where $\alpha(E)$ is the energy-dependent absorption, e is the elementary charge, and I_{sun} is the AM 1.5 G solar spectrum.[88]

The reverse saturation current J_0 includes the radiative current J_0^r and nonradiative current J_0^{nr},

$$J_0 = J_0^r + J_0^{nr} = \frac{J_0^r}{f_r} \tag{21}$$

Here, $J_0^r = e\pi \int_0^\infty \alpha(E) I_{bb}(E, T) dE$ and the fraction of radiative recombination current f_r is given by $f_r = e^{(E_g - E_g^{da}/k_B T)}$, where E_g is the electronic band gap, E_g^{da} is the direct-allowed optical band gap, and I_{bb} is the black-body spectrum at temperature T.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

[1] A. Kojima, K. Teshima, Y. Shirai, T. Miyasaka, *J. Am. Chem. Soc.* **2009**, *131*, 6050.

[2] NREL, Best research-cell efficiencies, https://www.nrel.gov/pv/cell-efficiency.html.

[3] Y. Jiang, J. Yuan, Y. Ni, J. Yang, Y. Wang, T. Jiu, M. Yuan, J. Chen, *Joule* **2018**, *2*, 1356.

[4] Y. Yan, S. Yu, A. Honarfar, T. Pullerits, K. Zheng, Z. Liang, *Adv. Sci.* **2019**, *6*, 1900548.

[5] Y. Hu, Z. Zhang, A. Mei, Y. Jiang, X. Hou, Q. Wang, K. Du, Y. Rong, Y. Zhou, G. Xu, H. Han, *Adv. Mater.* **2018**, *30*, 1705786.

[6] M. M. Tavakoli, P. Yadav, R. Tavakoli, J. Kong, *Adv. Energy Mater.* **2018**, *8*, 1800794.

[7] X. Zhang, Z. Shi, H. Lu, X. Li, H. Wang, S. Yuan, F. Liu, Y. Pan, Z. Weng, H. Zhang, X. Cui, Q. Liu, C. Li, A. A. Umar, J. Wang, X.-C. Hang, Z. Sun, Y. Zhan, *J. Mater. Chem. A* **2019**, *7*, 22323.

[8] K. Liu, S. Chen, J. Wu, H. Zhang, M. Qin, X. Lu, Y. Tu, Q. Meng, X. Zhan, *Energy Environ. Sci.* **2018**, *11*, 3463.

[9] D. B. Kim, S. Lee, C. H. Jang, J. H. Park, A.-y. Lee, M. H. Song, *Adv. Mater. Interfaces* **2020**, *7*, 1902158.

[10] H. Wang, Z. Zhang, J. V. Milić, L. Tan, Z. Wang, R. Chen, X. Jing, C. Yi, Y. Ding, Y. Li, Y. Zhao, X. Zhang, A. Hagfeldt, M. Grätzel, J. Luo, *Adv. Energy Mater.* **2021**, *11*, 2101082.

[11] Y. Cheng, L. Ding, *Energy Environ. Sci.* **2021**, *14*, 3233.

[12] A. Wakamiya, M. Endo, T. Sasamori, N. Tokitoh, Y. Ogomi, S. Hayase, Y. Murata, *Chem. Lett.* **2014**, *43*, 711.

[13] J. M. Frost, K. T. Butler, F. Brivio, C. H. Hendon, M. van Schilfgaarde, A. Walsh, *Nano Lett.* **2014**, *14*, 2584.

[14] G. Niu, W. Li, F. Meng, L. Wang, H. Dong, Y. Qiu, *J. Mater. Chem. A* **2014**, *2*, 705.

[15] S. Chen, K. R. Choudhury, J. Subbiah, C. M. Amb, J. R. Reynolds, F. So, *Adv. Energy Mater.* **2011**, *1*, 963.

[16] T. Leijtens, G. E. Eperon, N. K. Noel, S. N. Habisreutinger, A. Petrozza, H. J. Snaith, *Adv. Energy Mater.* **2015**, *5*, 1500963.

[17] T. A. Berhe, W.-N. Su, C.-H. Chen, C.-J. Pan, J.-H. Cheng, H.-M. Chen, M.-C. Tsai, L.-Y. Chen, A. A. Dubale, B.-J. Hwang, *Energy Environ. Sci.* **2016**, *9*, 323.

[18] N. H. Tiep, Z. Ku, H. J. Fan, *Adv. Energy Mater.* **2016**, *6*, 1501420.

[19] Z. Wang, Z. Shi, T. Li, Y. Chen, W. Huang, *Angew. Chem. Int. Ed.* **2017**, *56*, 1190.

[20] T. Haeger, R. Heiderhoff, T. Riedl, *J. Mater. Chem. C* **2020**, *8*, 14289.

[21] J. H. Heo, S. H. Im, J. H. Noh, T. N. Mandal, C.-S. Lim, J. A. Chang, Y. H. Lee, H.-j. Kim, A. Sarkar, M. K. Nazeeruddin, M. Grätzel, S. I. Seok, *Nat. Photonics* **2013**, *7*, 486.

[22] G. Divitini, S. Cacovich, F. Matteocci, L. Ciná, A. Di Carlo, C. Ducati, *Nat. Energy* **2016**, *1*, 15012.

[23] J. Zhao, K. O. Brinkmann, T. Hu, N. Pourdavoud, T. Becker, T. Gahlmann, R. Heiderhoff, A. Polywka, P. Görrn, Y. Chen, B. Cheng, T. Riedl, *Adv. Energy Mater.* **2017**, *7*, 1602599.

[24] S. Sandell, J. Maire, E. Chávez-Ángel, C. M. Sotomayor Torres, H. Kristiansen, Z. Zhang, J. He, *Nanomater.* **2020**, *10*, 4.

[25] K. A. Denault, J. Brgoch, S. D. Kloß, M. W. Gaultois, J. Siewenie, K. Page, R. Seshadri, *ACS Appl. Mater. Interfaces* **2015**, *7*, 7264.

[26] Y. Zhuo, A. M. Tehrani, A. O. Oliynyk, A. C. Duke, J. Brgoch, *Nat. Commun.* **2018**, *9*, 1.

[27] F. Ji, Y. Huang, F. Wang, L. Kobera, F. Xie, J. Klarbring, S. Abbrent, J. Brus, C. Yin, S. I. Simak, I. A. Abrikosov, I. A. Buyanova, W. M. Chen, F. Gao, *Adv. Funct. Mater.* **2020**, *30*, 2005521.

[28] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, O. Levy, *Nat. Mater.* **2013**, *12*, 191.

[29] S. Chakraborty, W. Xie, N. Mathews, M. Sherburne, R. Ahuja, M. Asta, S. G. Mhaisalkar, *ACS Energy Lett.* **2017**, *2*, 837.

[30] Y. Li, K. Yang, *Energy Environ. Sci.* **2019**, *12*, 2233.

[31] T. Nakajima, K. Sawada, *J. Phys. Chem. Lett.* **2017**, *8*, 4826.

[32] C. Kim, T. D. Huan, S. Krishnan, R. Ramprasad, *Sci. Data* **2017**, *4*, 170057.

[33] Materials genome initiative for global competitiveness, http://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf, **2011**.

[34] D. P. Tabor, L. M. Roch, S. K. Saikin, C. Kreisbeck, D. Sheberla, J. H. Montoya, S. Dwaraknath, M. Aykol, C. Ortiz, H. Tribukait, C. Amador-Bedolla, C. J. Brabec, B. Maruyama, K. A. Persson, A. Aspuru-Guzik, *Nat. Rev. Mater.* **2018**, *3*, 5.

[35] S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li, J. Wang, *Nat. Commun.* **2018**, *9*, 3405.

[36] Z. Li, Q. Xu, Q. Sun, Z. Hou, W. J. Yin, *Adv. Funct. Mater.* **2019**, *29*, 1807280.

[37] T. Wu, J. Wang, *Nano Energy* **2019**, *66*, 104070.

[38] S. Kanno, Y. Imamura, M. Hada, *Phys. Rev. Materials* **2019**, *3*, 075403.

[39] S. Sun, N. T. Hartono, Z. D. Ren, F. Oviedo, A. M. Buscemi, M. Layurova, D. X. Chen, T. Ogunfunmi, J. Thapa, S. Ramasamy, C. Settens, B. L. DeCost, A. G. Kusne, Z. Liu, S. I. Tian, I. M. Peters, J.-P. Correa-Baena, T. Buonassisi, *Joule* **2019**, *3*, 1437.

[40] K. Tran, Z. W. Ulissi, *Nat. Catal.* **2018**, *1*, 696.

[41] S. Lin, H. Xu, Y. Wang, X. C. Zeng, Z. Chen, *J. Mater. Chem. A* **2020**, *8*, 5663.

[42] G. H. Gu, C. Choi, Y. Lee, A. B. Situmorang, J. Noh, Y.-H. Kim, Y. Jung, *Adv. Mater.* **2020**, *32*, 1907865.

[43] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, R. D. Braatz, *Nat. Energy* **2019**, *4*, 383.

[44] R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams, A. Aspuru-Guzik, *Nat. Mater.* **2016**, *15*, 1120.

[45] A. Mansouri Tehrani, A. O. Oliynyk, M. Parry, Z. Rizvi, S. Couper, F. Lin, L. Miyagi, T. D. Sparks, J. Brgoch, *J. Am. Chem. Soc.* **2018**, *140*, 9844.

[46] R. Yuan, Z. Liu, P. V. Balachandran, D. Xue, Y. Zhou, X. Ding, J. Sun, D. Xue, T. Lookman, *Adv. Mater.* **2018**, *30*, 1702884.

[47] P. V. Balachandran, B. Kowalski, A. Sehirlioglu, T. Lookman, *Nat. Commun.* **2018**, *9*, 1668.

[48] M. Askerka, Z. Li, M. Lempen, Y. Liu, A. Johnston, M. I. Saidaminov, Z. Zajacz, E. H. Sargent, *J. Am. Chem. Soc.* **2019**, *141*, 3682.

[49] J. Kirman, A. Johnston, D. A. Kuntz, M. Askerka, Y. Gao, P. Todorović, D. Ma, G. G. Privé, E. H. Sargent, *Matter* **2020**, *2*, 938.

[50] G. Kieslich, S. Sun, A. K. Cheetham, *Chem. Sci.* **2015**, *6*, 3430.

[51] R. D. Shannon, *Acta Cryst. A* **1976**, *32*, 751.

[52] S. Burger, M. Ehrenreich, G. Kieslich, *J. Mater. Chem. A* **2018**, *6*, 21785.

[53] M. Becker, T. Klüner, M. Wark, *Dalton Trans.* **2017**, *46*, 3500.

[54] W. Travis, E. Glover, H. Bronstein, D. Scanlon, R. Palgrave, *Chem. Sci.* **2016**, *7*, 4548.

[55] T. M. Brenner, D. A. Egger, L. Kronik, G. Hodes, D. Cahen, *Nat. Rev. Mater.* **2016**, *1*, 15007.

[56] G. A. Slack, *J. Phys. Chem. Solids* **1973**, *34*, 321.

[57] D. S. Sanditov, V. N. Belomestnykh, *Tech. Phys.* **2011**, *56*, 1619.

[58] L. Feng, L. Jiang, M. Zhu, H. Liu, X. Zhou, C. Li, *J. Phys. Chem. Solids* **2008**, *69*, 967.

[59] S. M. Lundberg, S.-I. Lee, in *Adv. Neural Inf. Process. Syst.*, vol. 30, MIT Press, Cambridge, MA **2017**, pp. 4765–4774.

[60] G. P. Nagabhushana, R. Shivaramaiah, A. Navrotsky, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 7717.

[61] S. Gu, R. Lin, Q. Han, Y. Gao, H. Tan, J. Zhu, *Adv. Mater.* **2020**, *32*, 1907392.

[62] C. Ge, M. Hu, P. Wu, Q. Tan, Z. Chen, Y. Wang, J. Shi, J. Feng, *J. Phys. Chem. C* **2018**, *122*, 15973.

[63] J. Huang, Y. Yuan, Y. Shao, Y. Yan, *Nat. Rev. Mater.* **2017**, *2*, 17042.

[64] J. A. Van Vechten, T. K. Bergstresser, *Phys. Rev. B* **1970**, *1*, 3351.

[65] R. Dronskowski, P. Bloechl, *J. Phys. Chem. C* **1993**, *97*, 8617.

[66] M. R. Filip, F. Giustino, *Proc. Natl. Acad. Sci.* **2018**, *115*, 5397.

[67] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, *J. Phys.: Condens. Matter* **2009**, *21*, 395502.

[68] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, *J. Phys.: Condens. Matter* **2017**, *29*, 465901.

[69] A. Marini, C. Hogan, M. Grüning, D. Varsano, *Comput. Phys. Commun.* **2009**, *180*, 1392.

[70] M. Zhong, W. Zeng, H. Tang, L. X. Wang, F. S. Liu, B. Tang, Q. J. Liu, *Sol. Energy* **2019**, *190*, 617.

[71] M. Nishiwaki, K. Nagaya, M. Kato, S. Fujimoto, H. Tampo, T. Miyadera, M. Chikamatsu, H. Shibata, H. Fujiwara, *Phys. Rev. Materials* **2018**, *2*, 085404.

[72] V. M. Caselli, Z. Wei, M. M. Ackermans, E. M. Hutter, B. Ehrler, T. J. Savenije, *ACS Energy Lett.* **2020**, *5*, 3821.

[73] Y. Yang, M. Yang, Z. Li, R. Crisp, K. Zhu, M. C. Beard, *J. Phys. Chem. Lett.* **2015**, *6*, 4688.

[74] M. Sendner, P. K. Nayak, D. A. Egger, S. Beck, C. Müller, B. Epding, W. Kowalsky, L. Kronik, H. J. Snaith, A. Pucci, R. Lovrinčić, *Mater. Horiz.* **2016**, *3*, 613.

[75] K. Miyata, T. L. Atallah, X. Y. Zhu, *Sci. Adv.* **2017**, *3*, e1701469.

[76] J. M. Frost, *Phys. Rev. B* **2017**, *96*, 195202.

[77] B. Wu, W. Ning, Q. Xu, M. Manjappa, M. Feng, S. Ye, J. Fu, S. Lie, T. Yin, F. Wang, T. W. Goh, P. C. Harikesh, Y. K. E. Tay, Z. X. Shen, F. Huang, R. Singh, G. Zhou, F. Gao, T. C. Sum, *Sci. Adv.* **2021**, *7*, 8.

[78] A. M. Ganose, J. Park, A. Faghaninia, R. Woods-Robinson, K. A. Persson, A. Jain, *Nat. Commun.* **2021**, *12*, 2222.

[79] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. VanderPlas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *J. Mach. Learn. Res.* **2011**, *12*, 2825.

[80] J. Bergstra, D. Yamins, D. D. Cox, in *Proc. of the 12th Python in Science Conference*, SciPy, Austin, TX **2013**, pp. 13–19.

[81] G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169.

[82] N. W. Ashcroft, N. D. Mermin, *Solid State Physics*, Holt, Rinehart and Winston, New York **1976**.

[83] G. J. Martyna, M. L. Klein, M. Tuckerman, *J. Chem. Phys.* **1992**, *97*, 2635.

[84] B. Hou, Y. Zhang, H. Zhang, H. Shao, C. Ma, X. Zhang, Y. Chen, K. Xu, G. Ni, H. Zhu, *J. Phys. Chem. Lett.* **2020**, *11*, 3116.

[85] B. Peng, H. Zhang, H. Shao, K. Xu, G. Ni, L. Wu, J. Li, H. Lu, Q. Jin, H. Zhu, *ACS Photonics* **2018**, *5*, 4081.

[86] H.-Y. Chen, V. A. Jhalani, M. Palummo, M. Bernardi, *Phys. Rev. B* **2019**, *100*, 075135.

[87] K. Choudhary, M. Bercx, J. Jiang, R. Pachter, D. Lamoen, F. Tavazza, *Chem. Mater.* **2019**, *31*, 5900.

[88] D. G. Collins, W. G. Blättner, M. B. Wells, H. G. Horak, *Appl. Opt.* **1972**, *11*, 2684.