

Unveiling stable and efficient antiperovskite semiconductors via high-throughput computation and interpretable machine learning

Hao Qu(瞿浩), Tao Hu(胡涛), Mingjun Li(李明军), Jiangyu Yang(杨江渝), Yunyi Zhou(周云逸), Shichang Li(李世长), Dengfeng Li(李登峰), Gang Tang(唐刚), and Chunbao Feng(冯春宝)
Citation: Chin. Phys. B, 2026, 35 (**4**): 046102. DOI: 10.1088/1674-1056/ae24e7
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Physical properties of Cr_2S_3 at high pressure

Lun Xiong(熊伦), Mingquan Jiang(江明全), Jinxia Zhu(竹锦霞), Lin Xia(夏林), Hao Wang(王毫), Shenghan Zhang(张升瀚), Pengfei Tang(汤鹏飞), Zhiqiang Chen(陈志强), Sheng Jiang(蒋升), and Hongliang Dong(董洪亮)
Chin. Phys. B, 2026, 35 (**3**): 036103. DOI: 10.1088/1674-1056/ae1cb3

Physical properties of high-pressure synthesized $Al_{65}Cu_{20}Fe_{15}$ quasicrystal

Yibo Liu(刘一博), Changzeng Fan(范长增), Zhefeng Xu(许哲峰), Ruidong Fu(付瑞东), Feng Ke(柯峰), Lin Wang(王霖), Bin Wen(温斌), Lifeng Zhang(张立峰), Marek Mihalkovič, and Bo Xu(徐波)
Chin. Phys. B, 2025, 34 (**9**): 096103. DOI: 10.1088/1674-1056/add4fb

Layer-dependent structural stability and electronic properties of $CrPS_4$ under high pressure

Jian Zhu(朱健), Dengman Feng(冯登满), Liangyu Wang(王亮予), Liang Li(李亮), Fangfei Li(李芳菲), Qiang Zhou(周强), and Yalan Yan(闫雅兰)
Chin. Phys. B, 2025, 34 (**6**): 066102. DOI: 10.1088/1674-1056/adc65f

General-purpose moment tensor potential for Ga-In liquid alloys towards large-scale molecular dynamics with *ab initio* accuracy

Kai-Jie Zhao(赵凯杰) and Zhi-Gong Song(宋智功)
Chin. Phys. B, 2025, 34 (**6**): 066101. DOI: 10.1088/1674-1056/adc661

Wafer-scale 30° twisted bilayer graphene epitaxially grown on $Cu_{0.75}Ni_{0.25}$ (111)

Peng-Cheng Ma(马鹏程), Ao Zhang(张翱), Hong-Run Zhen(甄洪润), Zhi-Cheng Jiang(江志诚), Yi-Chen Yang(杨逸尘), Jian-Yang Ding(丁建阳), Zheng-Tai Liu(刘正太), Ji-Shan Liu(刘吉山), Da-Wei Shen(沈大伟), Qing-Kai Yu(于庆凯), Feng Liu(刘丰), Xue-Fu Zhang(张学富), and Zhong-Hao Liu(刘中灏)
Chin. Phys. B, 2024, 33 (**6**): 066101. DOI: 10.1088/1674-1056/ad2d53

SPECIAL TOPIC — Artificial intelligence and smart materials innovation: From fundamentals to applications

Unveiling stable and efficient antiperovskite semiconductors via high-throughput computation and interpretable machine learning

Hao Qu(瞿浩)[1,†], Tao Hu(胡涛)[1,†], Mingjun Li(李明军)[1], Jiangyu Yang(杨江渝)[1], Yunyi Zhou(周云逸)[1], Shichang Li(李世长)[1], Dengfeng Li(李登峰)[1,‡], Gang Tang(唐刚)[3,§], and Chunbao Feng(冯春宝)[1,2,¶]

[1] *School of Electronic Science and Engineering, Chongqing University of Posts and Telecommunications, Chongqing* 400065, *China*

[2] *Chongqing Key Laboratory of Dedicated Quantum Computing and Quantum Artificial Intelligence, Chongqing University of Posts and Telecommunications, Chongqing* 400065, *China*

[3] *School of Interdisciplinary Science, Beijing Institute of Technology, Beijing* 100081, *China*

(Received 6 August 2025; revised manuscript received 11 November 2025; accepted manuscript online 27 November 2025)

Nitride antiperovskites have recently been theoretically identified as promising optoelectronic materials, yet their chemical space remains largely unexplored. Here, we employ a high-throughput first-principles screening workflow to systematically investigate the X_3BA antiperovskite family. Six candidates that exhibit both structural and dynamical stability together with desirable bandgaps are identified. Electronic-structure calculations reveal that the alkaline-earth-based compounds (e.g., Ca_3AsSb, Sr_3AsSb, Ba_3AsSb) not only possess suitable direct bandgaps and strong optical absorption, but also exhibit favorable ambipolar carrier mobilities and low exciton binding energies (< 45 meV). Notably, Sr_3AsSb and Ba_3AsSb are predicted to achieve theoretical maximum power-conversion efficiencies of 28.1% and 29.4%, respectively. Finally, an interpretable machine-learning model demonstrates that the electronegativity of the A-site anion is the single most influential descriptor governing bandgap trends across the chemical space. This work establishes a data-driven design heuristic and provides a predictive framework for the accelerated discovery of efficient and stable antiperovskite-based optoelectronic materials.

Keywords: antiperovskite, physical properties, first-principles calculations, interpretable machine learning

PACS: 61.50.Ah, 71.15.Mb, 63.20.dk, 07.05.Mh

DOI: 10.1088/1674-1056/ae24e7 **CSTR:** 32038.14.CPB.ae24e7

1. Introduction

In recent years, halide perovskite semiconductors have attracted tremendous research interest in optoelectronics due to their tunable direct bandgaps,[1–3] high optical absorption coefficients,[4] and long carrier diffusion lengths and lifetimes.[5,6] Since the first report of a modest 3.8% power-conversion efficiency in 2009,[7] perovskite solar cells have undergone remarkable progress, with efficiencies now approaching 27%.[8] Their low-temperature solution processability, mechanical flexibility, and seamless heterojunction integration further highlight their potential for next-generation cost-effective optoelectronic technologies.[9]

Despite these advances, perovskite materials remain intrinsically prone to hydrolytic, oxidative, and thermal degradation, which continues to limit their long-term stability and device reliability.[10–13] To address these challenges, researchers have explored diverse engineering strategies. Beyond compositional tuning, external stimuli have emerged as powerful tools for modulating physical properties.[14] These methods, together with structural–compositional approaches such as two-dimensional/three-dimensional heterostructure design,[15,16] interface passivation,[17] and defect control,[18] have shown promise in significantly improving the stability of perovskites under moisture-, thermal-, and light-induced stress. More recently, an innovative ion-type inversion strategy has attracted growing attention, in which the cation and anion sites of perovskite structures are interchanged to yield antiperovskites.[19–21] These antiperovskites retain the cubic symmetry and three-dimensional framework characteristic of typical perovskites, while exhibiting superior compositional tunability and rich structural polymorphism.[22,23] These features offer promising opportunities to design antiperovskite candidates with optimized optoelectronic performance through rational compositional engineering.

In this work, we employed high-throughput first-principles screening to explore elemental combinations across a wide range of antiperovskite compounds. Key descriptors, including tolerance factors, electronic bandgaps, phase stability (energy differences), and phonon spectra, were systematically evaluated, leading to the identification of six struc-

[†]These authors contributed equally to this work.

[‡]Corresponding author. E-mail: lidf@cqupt.edu.cn

[§]Corresponding author. E-mail: gtang@bit.edu.cn

[¶]Corresponding author. E-mail: fengcb@cqupt.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

turally robust candidates: Cs_3OCl, Cs_3SI, Rb_3ClTe, Ca_3AsSb, Sr_3AsSb, and Ba_3AsSb. These materials were further characterized in terms of their electronic band structures, carrier mobilities, exciton binding energies, optical absorption spectra, and theoretical maximum photovoltaic efficiencies. Moreover, interpretable machine-learning analysis revealed that the electronegativity of the A-site anion serves as the dominant factor controlling the bandgap in cubic X_3BA antiperovskites.

2. Computation details

All calculations were carried out using the Vienna *ab initio* simulation package (VASP) with the projector-augmented-wave (PAW) method.[24,25] The Perdew–Burke–Ernzerhof (PBE) exchange–correlation functional within the generalized-gradient approximation (GGA)[26] was employed for geometry optimizations. A plane-wave cutoff energy of 550 eV and a k-point mesh density of $2\pi \times 0.02$ Å^{-1} were used. Structural convergence was achieved when the total energy variation was less than 10^{-6} eV and the residual atomic forces were below 0.02 eV·Å^{-1}. To obtain more accurate electronic structures and bandgaps, the Heyd–Scuseria–Ernzerhof hybrid functional (HSE06)[27] was adopted. Spin–orbit coupling (SOC) was included to account for relativistic effects, which are particularly pronounced in heavy-element systems.[28] Phonon dispersions were computed using density-functional perturbation theory (DFPT) with $2 \times 2 \times 2$ supercells, and the resulting data were post-processed using the SUMO package.[29] To assess dynamical stability at finite temperatures and include anharmonic effects, the phonon spectrum at 300 K was additionally calculated using the temperature-dependent effective potential (TDEP) code.[30] The required force–displacement data for this method were generated from *ab initio* molecular dynamics (AIMD) simulations carried out in the canonical (NVT) ensemble. Carrier mobility calculations were performed at 300 K using the *ab initio* scattering and transport code (AMSET),[31] with a $31 \times 23 \times 33$ k-point grid employed for all computations. This code explicitly incorporates three major scattering processes within the Boltzmann transport framework: acoustic deformation potential (ADP) scattering from lattice vibrations; polar optical phonon (POP) scattering driven by long-range electrostatic interactions; and ionized impurity (IMP) scattering induced by charged defects.[31] Electronic and optical properties were further analyzed with VASPKIT.[32] The exciton binding energy (E_b) was evaluated using the Wannier–Mott model

$$E_b = Ry\frac{\mu^*}{\varepsilon_r^2}, \tag{1}$$

where $Ry = 13.606$ eV is the atomic Rydberg energy, and the reduced exciton mass μ^* is defined as $1/\mu^* = 1/m_h + 1/m_e$. The optical absorption coefficient $\alpha(\omega)$ was derived from the frequency-dependent dielectric function using the expression

$$\alpha(\omega) = \frac{\sqrt{2}\omega}{c}\left(\sqrt{\varepsilon_1^2+\varepsilon_2^2}-\varepsilon_1\right)^{\frac{1}{2}}, \tag{2}$$

where $\varepsilon_1(\omega)$ and $\varepsilon_2(\omega)$ denote the real and imaginary parts of the dielectric function, respectively, and ω is the optical frequency. Data processing and machine-learning workflows were carried out in Python.[33] Gradient-boosting models were constructed using the XGBoost library,[34] and model interpretability was obtained using the SHapley additive exPlanations (SHAP) framework.[35]

3. Results and discussion

3.1. High-throughput screening workflow

We focus on antiperovskite compounds with the general formula X_3BA, obtained by interchanging the anionic and cationic species at the X, B, and A sites relative to conventional perovskites.[21] The elemental selection criteria used in the screening process are presented in Fig. 1(a). For the X site, alkali metals with a +1 oxidation state (e.g., Li, Na) and alkaline-earth metals with a +2 oxidation state (e.g., Mg, Ca) were considered. The choice of the B- and A-site elements was guided by compositions commonly reported for antiperovskite compounds,[36–41] ensuring charge neutrality and structural feasibility of the resulting compounds.

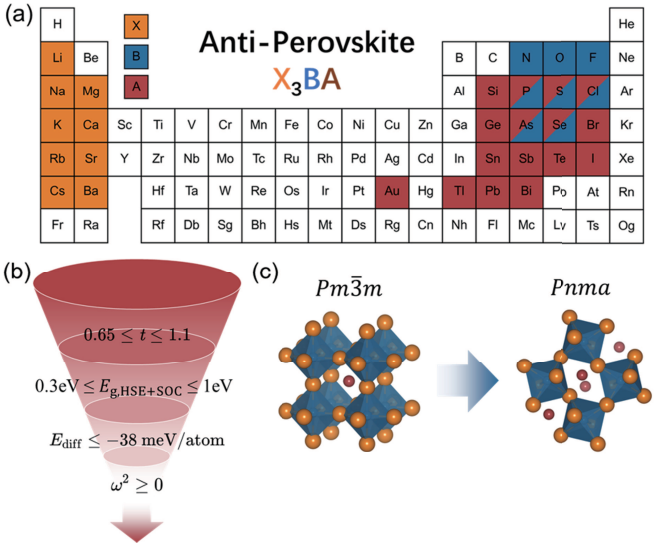


Fig. 1. (a) Periodic table locations of the X-, B-, and A-site elements considered in the X_3BA antiperovskite chemical space. (b) Schematic overview of the high-throughput screening workflow, including tolerance-factor filtering, bandgap evaluation, dynamical-stability assessment, and electronic-structure analysis. (c) Illustration of the structural transformation from the ideal high-symmetry cubic phase (*Pm*-3*m*) to a thermodynamically favored low-symmetry distorted phase.

A high-throughput computational workflow was then established to identify antiperovskite phases that are both structurally and dynamically stable and possess potentially suitable bandgaps for optoelectronic applications, as illustrated in

Fig. 1(b). As the first structural-stability filter, we applied the Goldschmidt tolerance factor t,[42] defined as $t = \frac{r_A + r_X}{\sqrt{2}(r_B + r_X)}$, where r denotes the effective Shannon ionic radius of each species.[43] Owing to its simplicity and predictive utility, the tolerance factor has been widely adopted as a stability descriptor in perovskite systems;[42,44,45] its applicability has also been extended to antiperovskites.[45] In this work, compounds satisfying the empirical stability window of $0.65 \leq t \leq 1.1$[46] were retained. Subsequently, bandgaps were screened based on the Shockley–Queisser (SQ) efficiency limit, which indicates that high-performance photovoltaic absorbers typically exhibit bandgaps in the range of 1.1–1.55 eV.[47] To mitigate the well-known bandgap underestimation of PBE, PBE + SOC calculations were employed, and a preliminary selection window of 0.3–1.0 eV was adopted to ensure that the HSE06-corrected bandgaps would reasonably fall within the SQ optimal range.

To further evaluate phase stability, we defined $E_{\text{diff}} = E_{Pm\text{-}3m} - E$ as the energy difference between the ideal cubic phase ($E_{Pm\text{-}3m}$) and distorted structural variants (E). Structures with $E_{\text{diff}} < -38$ meV/atom were retained (see Figs. S1–S3). This threshold is consistent with widely used heuristics in high-throughput materials screening, since many experimentally synthesizable metastable compounds are found within similar energy windows above the convex hull.[48,49] Finally, phonon dispersion calculations were performed to confirm dynamical stability (Fig. 1(c)). Six dynamically stable antiperovskite compounds were identified: Ca_3AsSb, Sr_3AsSb, Ba_3AsSb, Cs_3OCl, Cs_3SI, and Rb_3ClTe. All six adopt the orthorhombic *Pnma* space group. Notably, the small imaginary modes observed in the phonon spectrum of Cs_3OCl near the Γ and U points are attributed to numerical artifacts inherent to density functional theory (DFT) calculations rather than any intrinsic dynamical instability (see Fig. S4).[50] To further validate the stability of these materials at room temperature, we performed anharmonic phonon calculations at 300 K for the representative candidate Ba_3AsSb. As shown in Fig. S5, the resulting spectrum exhibits no imaginary frequencies, confirming its dynamical stability at typical photovoltaic operating temperatures and thus strengthening the reliability of our screening results.

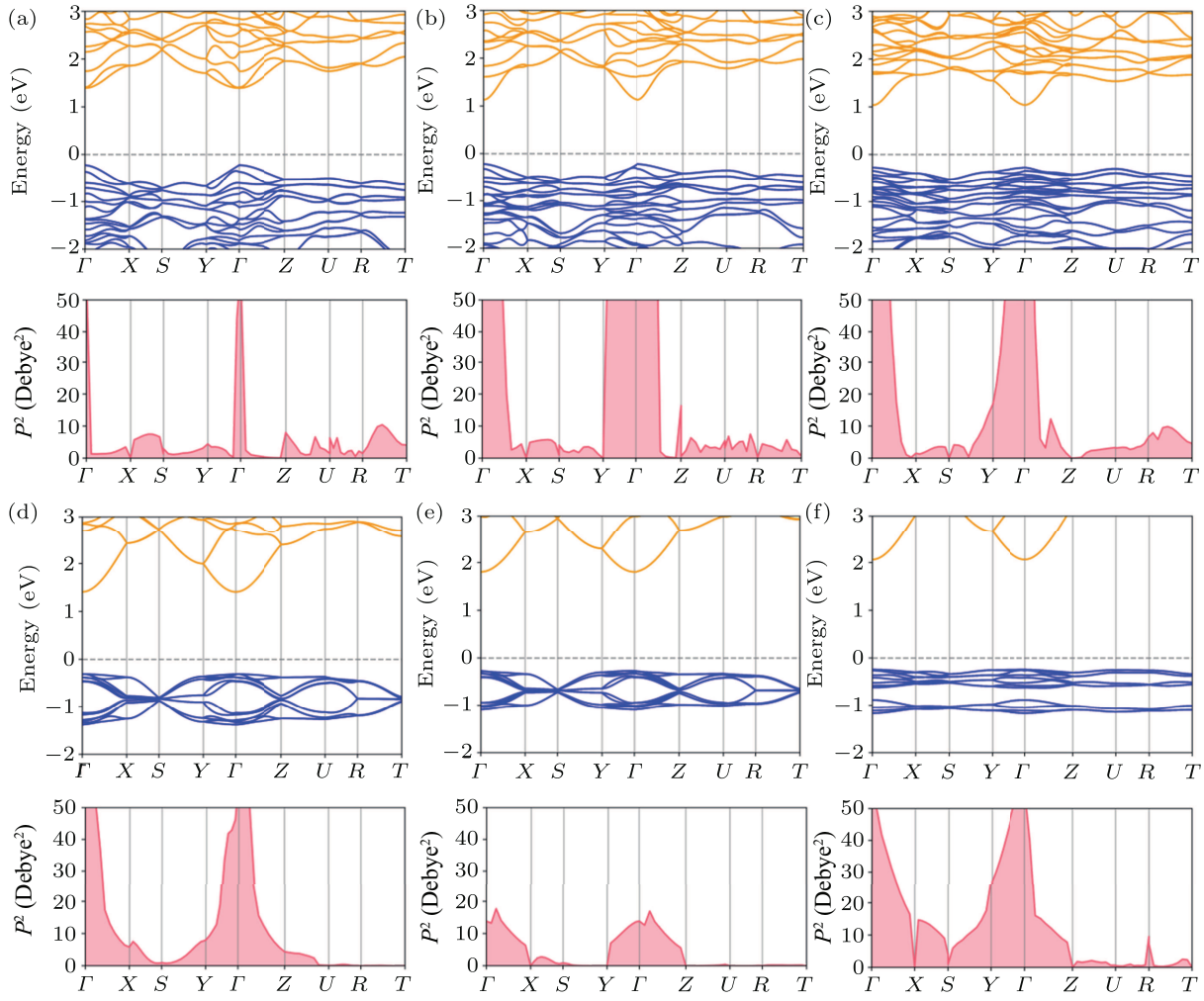


Fig. 2. Band structures (upper panel) and distributions of transition dipole moments (lower panel) for six antiperovskite materials, calculated using HSE06 + SOC. The materials shown correspond to (a) Ca_3AsSb, (b) Sr_3AsSb, (c) Ba_3AsSb, (d) Cs_3OCl, (e) Cs_3SI, and (f) Rb_3ClTe.

3.2. Electronic band structures and density of states

To assess the electronic properties of the six selected antiperovskite compounds, their band structures were calculated using the HSE06 + SOC method. The corresponding bandgap values are summarized in Table 1. Overall, the three compounds containing alkaline-earth metals at the X site exhibit smaller bandgaps than those incorporating alkali-metal cations. In particular, Sr_3AsSb and Ba_3AsSb show bandgaps of 1.351 eV and 1.315 eV, respectively, which lie within the optimal photovoltaic window suggested by the Shockley–Queisser limit (1.1–1.55 eV). These results indicate that both materials may be promising candidates for photovoltaic applications.

Figure 2 shows the band structures and the corresponding transition-dipole matrix elements for the six compounds under study, all of which are found to be direct-gap semiconductors with both the valence-band maximum (VBM) and the conduction-band minimum (CBM) located at the Γ point. A systematic trend observed across our dataset is that alkali-metal-based compounds consistently exhibit wider bandgaps than their alkaline-earth-based counterparts. The origin of this phenomenon can be rationalized by the interplay between ionicity and covalency, which is governed by two key factors according to established orbital-interaction theory:[36–38,41,50–52] (i) the initial energy separation (Δ) between the frontier orbitals, as dictated by their electronegativity difference, and (ii) the covalent hybridization strength (V). In the alkali-metal-based systems, the large electronegativity difference results in a substantial Δ, which suppresses covalent hybridization and leads to an ionicity-driven wide bandgap, where a deeply stabilized VBM and a high-lying localized CBM dominate. Conversely, the smaller electronegativity difference in the alkaline-earth-based systems allows for strong covalent hybridization, particularly the p–d hybridization between cation d and anion p orbitals, which significantly lowers the energy of the antibonding CBM and results in a narrower, covalency-driven bandgap.

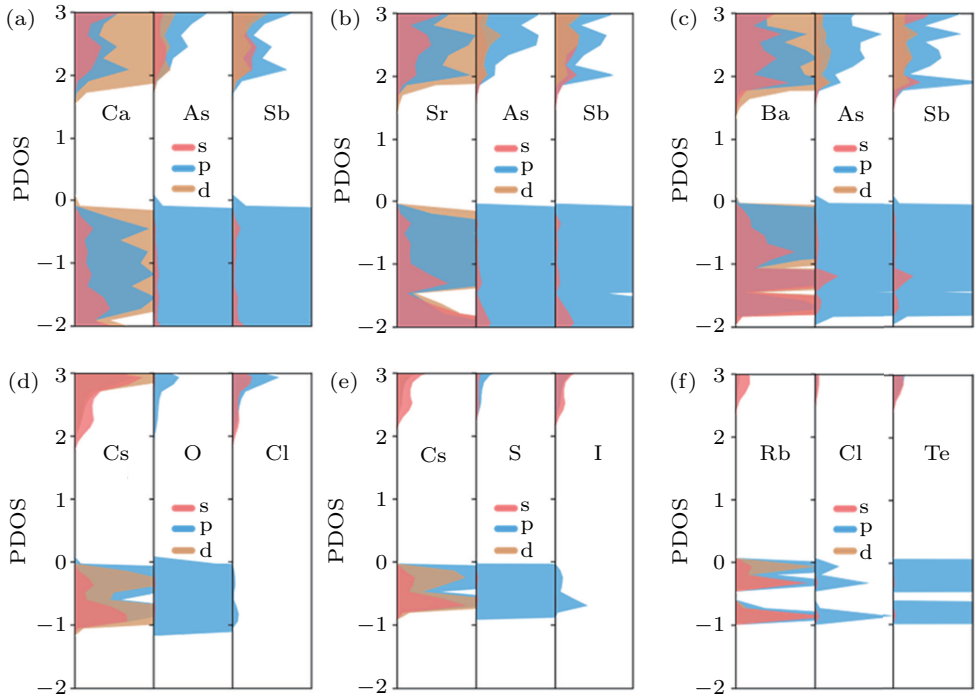


Fig. 3. Projected density of states (PDOS) of (a) Ca_3AsSb, (b) Sr_3AsSb, (c) Ba_3AsSb, (d) Cs_3OCl, (e) Cs_3SI, and (f) Rb_3ClTe.

The PDOS results in Fig. 3 further support this mechanism. In the alkali-metal systems, the PDOS exhibits sharp, localized, atomic-like peaks at the band edges, while the alkaline-earth systems display broad, highly overlapping continuous states, which are clear indicators of weak and strong hybridization, respectively. Notably, an interesting chemical trend also emerges within the alkaline-earth family. The bandgap of Ca_3AsSb (1.619 eV) is significantly larger than those of Sr_3AsSb and Ba_3AsSb. Our projected band-structure analysis suggests that this is primarily due to the progressive lowering of the cation d-orbital energies from Ca-3d to Sr-4d to Ba-5d, which enhances the p–d hybridization in Sr_3AsSb and Ba_3AsSb. This stronger hybridization lowers their CBM, ultimately resulting in a reduced bandgap.

3.3. Carrier mobility and exciton binding energy

We further evaluated the carrier-transport properties by performing first-principles calculations of both electron (n-type) and hole (p-type) mobilities in the six candidate materials. As shown in Fig. 4, all compounds exhibit strong po-

tential for efficient charge transport. In the intrinsic or lightly doped regime (carrier concentration $< 10^{16}$ cm^{-3}), electron mobilities generally exceed 100 cm^2·V^{-1}·s^{-1}, with particularly high values observed in the alkaline-earth-based systems. In this regime, mobility is predominantly limited by polar optical phonon (POP) scattering, a characteristic feature of polar semiconductors in which conduction electrons strongly couple to macroscopic polarization fields associated with longitudinal optical phonons.[53,54] These Fröhlich interactions lead to polaron formation, thereby reducing carrier mobility. As the carrier concentration increases into the heavily doped regime ($> 10^{18}$ cm^{-3}), ionized-impurity IMP scattering becomes the dominant mechanism, resulting in a pronounced decrease in mobility.

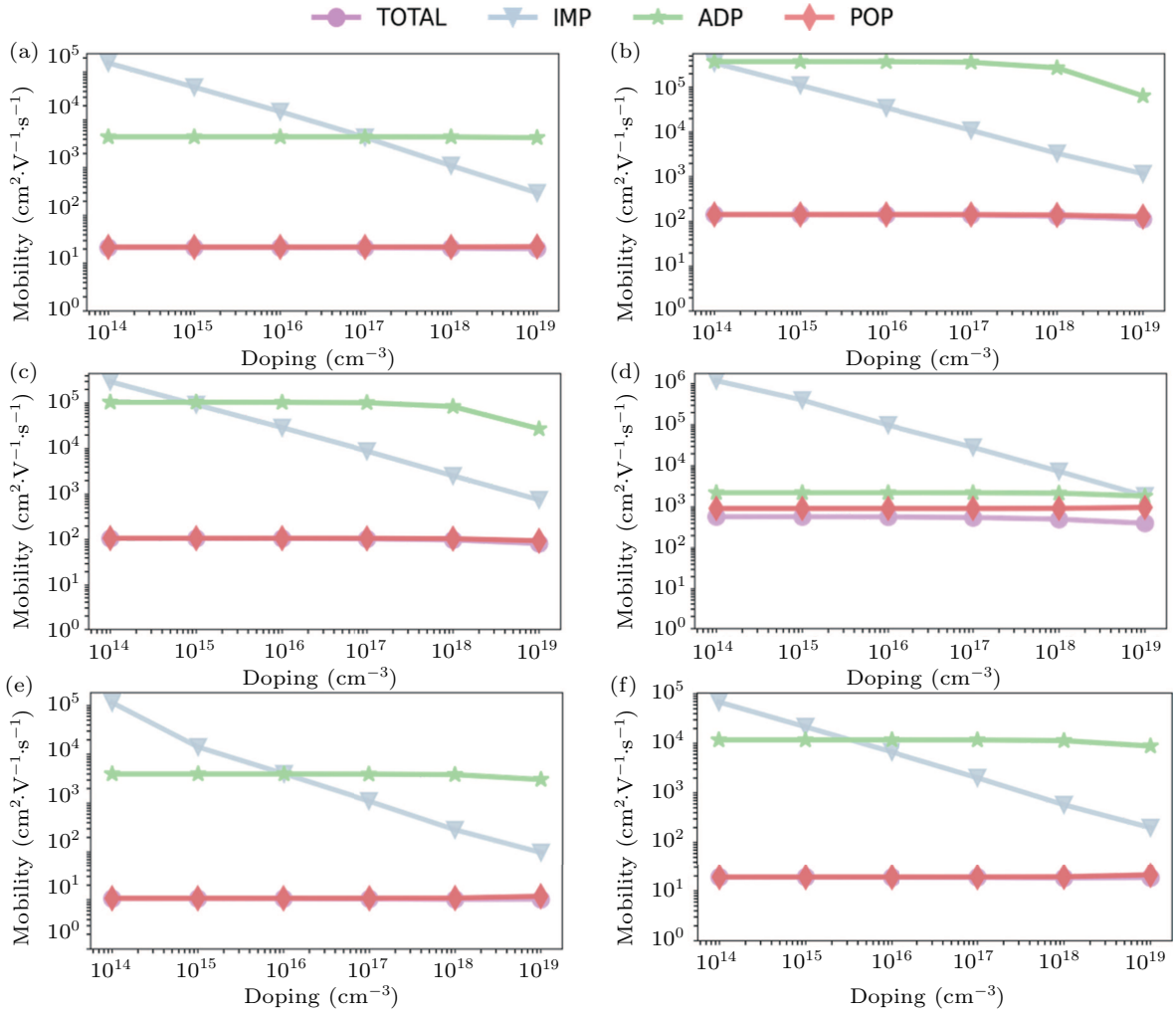


Fig. 4. Carrier mobility as a function of doping concentration for n-type antiperovskites at 300 K for the six materials calculated based on DFT, including (a) Ca_3AsSb, (b) Sr_3AsSb, (c) Ba_3AsSb, (d) Cs_3OCl, (e) Cs_3SI, and (f) Rb_3ClTe. The purple, light blue, green, and red curves represent the total mobility, IMP scattering, ADP scattering, and POP scattering, respectively.

In contrast to the consistently high electron mobilities, the hole-transport properties (Fig. 5) exhibit pronounced material-dependent behavior and a substantially broader performance spread, reflecting a more intricate set of governing physical mechanisms. The alkaline-earth-based compounds (Figs. 5(a)–5(c)) maintain relatively high hole mobilities, with values comparable to their electron counterparts. Consistent with the n-type behavior, POP scattering remains the dominant mobility-limiting mechanism, and the mobility decreases systematically with increasing hole effective mass from Ca_3AsSb to Ba_3AsSb (see Table 1). However, the alkali-metal-based materials (Figs. 5(d)–5(f)) display markedly different behavior. For Cs_3SI and Rb_3ClTe, the hole mobilities are extremely low (< 1 cm^2·V^{-1}·s^{-1}), indicative of poor p-type conductivity. This originates from their exceptionally flat valence-band maxima, which result in very large hole effective masses ($m_h = 1.160\ m_0$ and $3.680\ m_0$, respectively).

The most notable outlier is Cs_3OCl (Fig. 5(d)). Despite having one of the largest hole effective masses among all candidates ($m_h \approx 3.034\ m_0$), it unexpectedly maintains a moderate mobility of ~ 10 cm^2·V^{-1}·s^{-1}. Analysis of the scattering contributions elucidates the origin of this behavior: in Cs_3OCl, the dominant mobility-limiting mechanism shifts from POP scattering to ADP scattering, with POP scattering strongly suppressed. This suppression is attributed to its exceptionally large static dielectric constant ($\varepsilon_{st} \approx 82$, see Table 1),

which provides substantial dielectric screening and significantly weakens the long-range Coulomb interactions between charge carriers and longitudinal optical phonons. As a result, the POP scattering strength is reduced by nearly an order of magnitude, effectively mitigating the detrimental influence of the large effective mass and preserving a non-negligible level of hole mobility. These observations highlight a potential design principle: enhancing dielectric screening may serve as an effective strategy to counteract large carrier effective masses in antiperovskites.

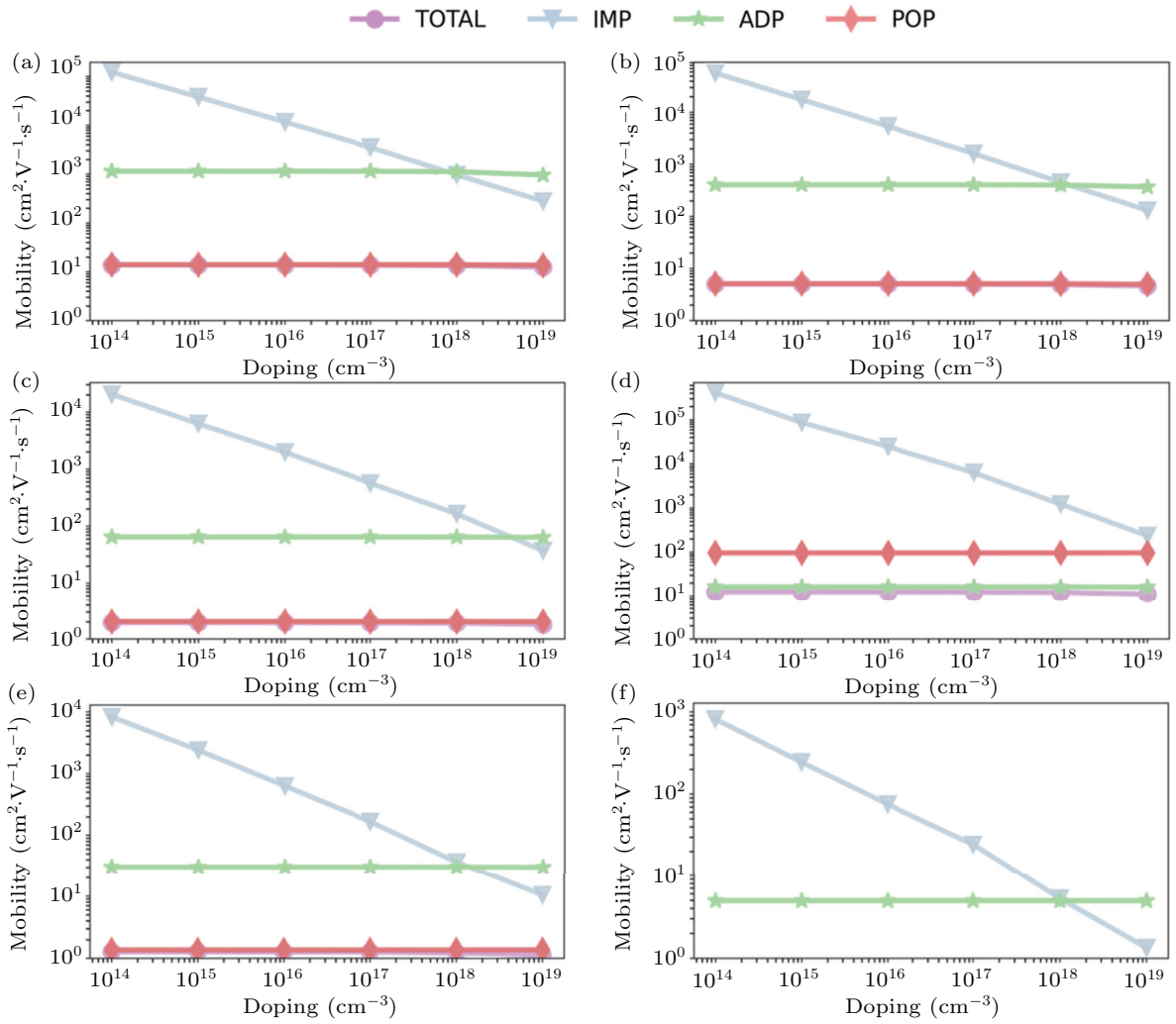


Fig. 5. Hole mobility as a function of doping concentration for p-type antiperovskites at 300 K for the six materials calculated based on DFT, including (a) Ca_3AsSb, (b) Sr_3AsSb, (c) Ba_3AsSb, (d) Cs_3OCl, (e) Cs_3SI, and (f) Rb_3ClTe. The purple, light blue, green, and red curves represent the total mobility, IMP scattering, ADP scattering, and POP scattering, respectively.

Table 1. Static dielectric constants (ε_{st}^{poly}) including both electronic ($\varepsilon_{\infty}^{poly}$) and ionic contributions (ε_{ion}^{poly}), bandgaps calculated using HSE06 + SOC, average hole and electron effective masses, exciton binding energies (E_b) estimated by the Wannier–Mott model, and maximum predicted photovoltaic efficiencies (η_{max}, %) of 2 μm thin films computed via the spectroscopic limited maximum efficiency (SLME) method for six antiperovskite materials.

Compounds	a (Å)	b (Å)	c (Å)	$\alpha=\beta=\gamma$ (°)	∠B–X–B (°)	ε_{ion}^{poly}	$\varepsilon_{\infty}^{poly}$	ε_{st}^{poly}	E_g	m_h	m_e	E_b (meV)	η_{max}
Ca_3AsSb	7.980	10.733	7.541	90	142.828	9.175	9.439	18.614	1.619	0.449	0.646	42	24.0
Sr_3AsSb	8.473	11.373	7.949	90	141.610	8.204	9.008	17.212	1.351	0.499	0.163	22	28.1
Ba_3AsSb	8.991	12.014	8.362	90	140.497	9.595	9.220	18.815	1.315	0.838	0.200	27	29.4
Cs_3OCl	8.346	11.441	7.900	90	148.055	78.790	3.188	81.978	1.717	3.034	0.381	433	21.8
Cs_3SI	9.531	13.009	9.084	90	146.469	5.003	3.020	8.024	2.075	1.160	0.410	456	17.0
Rb_3ClTe	9.296	12.003	8.349	90	141.212	3.922	3.331	7.253	2.310	3.680	0.325	374	13.8

The exciton binding energies for all six antiperovskites, evaluated via the Wannier–Mott model based on the calculated effective masses and static dielectric constants, are summarized in Table 1. The results reveal that, owing to their relatively large hole effective masses, the alkali-metal-based antiperovskites generally exhibit higher exciton binding energies compared to their alkaline-earth-metal-based counterparts. Notably, Sr_3AsSb and Ba_3AsSb possess optimal

bandgaps alongside exciton binding energies of 22 meV and 27 meV, respectively. These relatively low binding energies imply facilitated exciton dissociation, which is advantageous for improving charge-carrier transport efficiency in photovoltaic applications.

3.4. Photovoltaic properties

In addition to suitable bandgaps and charge-carrier effective masses, strong visible-light absorption is an indispensable property for solar-cell absorber materials. Using the HSE06 + SOC method, the absorption coefficients of six antiperovskite materials were computed, with results presented in Fig. 6(a). The absorption onset closely aligns with the bandgap values of the respective materials. As photon energy increases, the absorption coefficient rapidly rises, reaching approximately 10^4 cm^{-1} within a range of ~ 0.44 eV. At this point, the absorption growth of alkali-metal-based antiperovskites begins to plateau, whereas that of alkaline-earth-metal-based materials continues to increase, ultimately exceeding 10^5 cm^{-1} within ~ 0.37 eV before leveling off. Notably, the alkaline-earth-metal-based antiperovskites demonstrate significantly superior absorption performance compared to their alkali-metal-based counterparts, primarily due to their more favorable direct bandgaps. Furthermore, the strong optical transitions at the Γ point between the valence-band maximum and conduction-band minimum, characterized by transition-dipole moments $P^2 > 50$ Debye^2, further facilitate effective photon excitation and enhance their light-absorption capabilities.

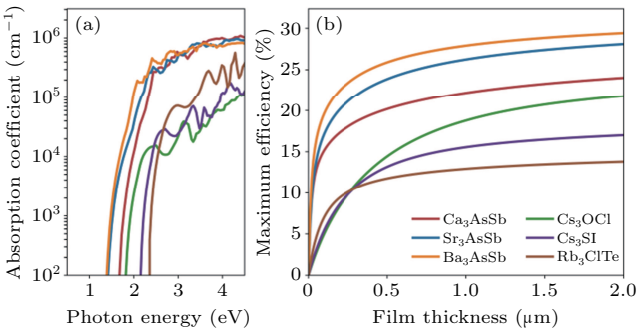


Fig. 6. (a) Absorption spectra of the six materials calculated using the HSE06 + SOC method; (b) SLME as a function of thin-film thickness for the six candidate antiperovskites.

The photovoltaic performance of the materials was ultimately assessed using the SLME method.[55] By integrating optical absorption properties with bandgap characteristics, SLME offers a systematic and comprehensive theoretical approach for the screening and evaluation of emerging photovoltaic materials. Figure 6(b) illustrates the variation of SLME as a function of thin-film thickness for the six antiperovskite compounds. Notably, at a film thickness of 2 μm, Sr_3AsSb and Ba_3AsSb exhibit SLME values of 28.1% and 29.4%, respectively, indicating their promising potential as absorber materials for solar cells. This is primarily attributed to their optimal bandgaps and excellent visible-light absorption capabilities.

3.5. Key bandgap descriptors identified by interpretable machine learning

To identify the key factors governing the bandgaps (E_g) of X_3BA-type antiperovskites, we developed and trained an interpretable machine-learning model using PBE + SOC bandgaps computed for 114 compounds obtained from our high-throughput screening. It is noted that the present screening set comprises only cubic X_3BA antiperovskites; mixed-halide, chalcogenide, and vacancy-ordered variants were not considered due to the limited availability of reliable structural and electronic data in existing databases. A consistency check was performed by comparing PBE + SOC and HSE06 + SOC bandgaps for a representative subset of materials (Fig. S6), which confirmed that both approaches yield similar trends. This agreement supports the use of PBE + SOC data for subsequent interpretability analysis to infer bandgap-driving features. To represent the compositional characteristics of each material, atomic descriptors including electronegativity, covalent radius, and polarizability were assigned to each crystallographic site. Statistical metrics such as extrema, standard deviation, and variance were then computed, yielding a 24-dimensional feature vector per compound. Pearson correlation analysis was carried out to select the ten features most strongly correlated with the bandgap for input into the model. As shown in Fig. 7(a), electronegativity-related descriptors exhibit the strongest correlation with E_g, consistent with chemical intuition and prior theoretical understanding.

Subsequently, the robustness of the selected descriptors was further evaluated using uniform manifold approximation and projection (UMAP) for dimensionality reduction.[56] As shown in Fig. 7(b), compounds with similar E_g values form well-separated clusters in the UMAP embedding space. Density-based clustering via the DBSCAN algorithm[57] reveals seven major clusters, each containing more than ten samples (highlighted in gray). The consistent color distribution within individual clusters indicates that materials grouped together exhibit comparable bandgaps, thereby validating the discriminative capability and physical relevance of the selected descriptors for capturing bandgap trends in antiperovskites.

In our preliminary model assessment, we compared several representative regression algorithms, including linear regression, random forest, and XGBoost. XGBoost was ultimately selected due to its superior performance in five-fold cross-validation, consistently achieving the highest coefficient of determination (R^2) and the lowest mean absolute error (MAE). The model hyperparameters were then optimized using Bayesian optimization, with the R^2 on the training set used

as the evaluation metric, which led to a clear improvement in performance. A comparison of the parity plots before and after tuning is shown in Fig. S7, and the corresponding numerical results are listed in Table S1. Model evaluation was conducted using a standard 80/20 train-test split. As shown in Fig. 8(a), the parity plot between the predicted and calculated E_g values reveals that the data points for both the training and test sets lie close to the ideal diagonal line. The optimized model achieves an R^2 of 0.97, demonstrating its strong predictive capability and indicating that XGBoost effectively captures the underlying correlations between the selected descriptors and the bandgap.

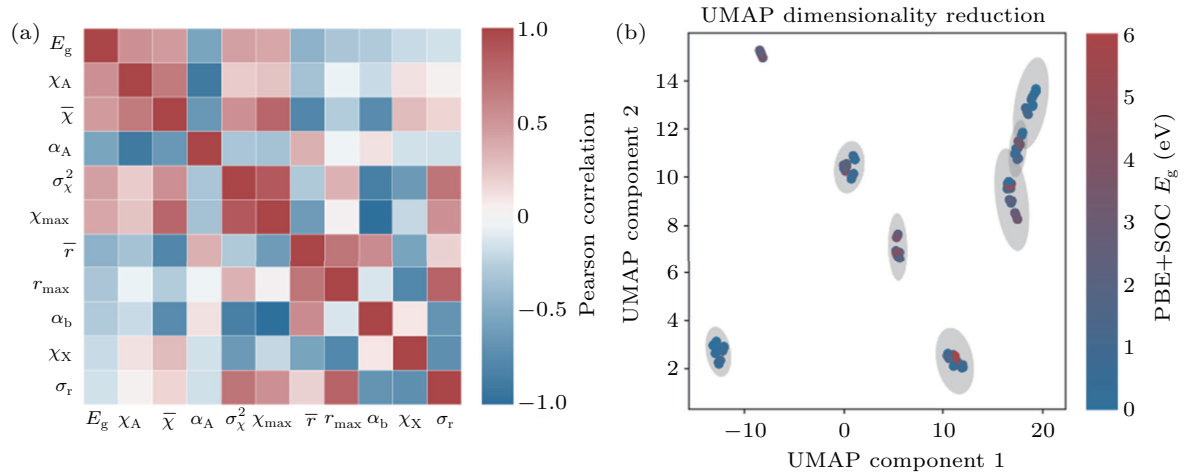


Fig. 7. (a) Pearson correlation matrix between the selected features and E_g; (b) dimensionality reduction of the selected features using the UMAP method followed by clustering analysis based on the DBSCAN algorithm[56] (gray regions indicate clusters). The color of each data point corresponds to the magnitude of E_g.

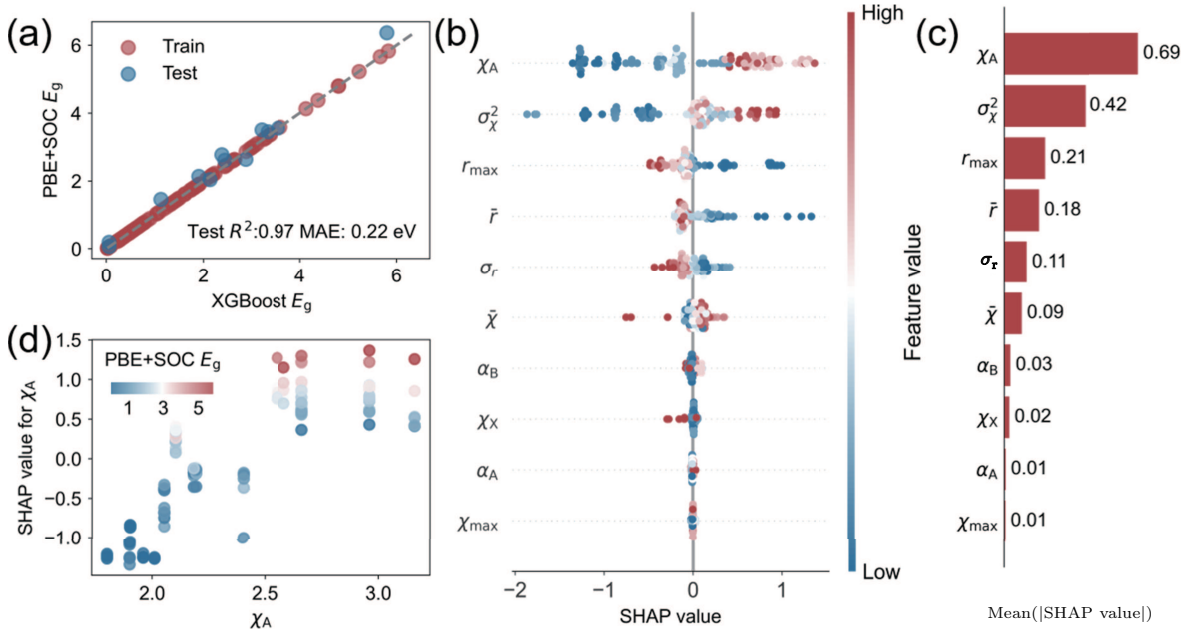


Fig. 8. (a) Parity plot comparing PBE + SOC-calculated E_g with XGBoost model predictions; (b) SHAP beeswarm plot illustrating the contribution of selected features across samples; (c) bar chart of feature importance based on the mean absolute SHAP values; (d) scatter plot of the A-site ion electronegativity (χ_A) versus its SHAP contribution, with point colors representing the corresponding E_g.

To further interpret the model's decision-making process, SHAP values were employed to evaluate the contribution of each feature to the model output. As shown in Figs. 8(b) and 8(c), the SHAP beeswarm plot and the feature importance bar chart based on mean absolute SHAP values are presented. Notably, the feature importance ranking obtained from SHAP analysis is highly consistent with the previous Pearson correlation results, further validating the rationality of the feature selection approach. From Figs. 8(b) and 8(c), it is evident that the electronegativity of the A-site ion (χ_A) contributes most significantly to the prediction of PBE + SOC E_g. The color gradient from blue to red in the SHAP beeswarm plot indicates a positive correlation between χ_A and its contribution to the E_g prediction, which is further confirmed by the scatter plot of χ_A versus SHAP values in Fig. 8(d). Moreover, among the top five most important features identified by the model, four

are related to electronegativity. This highlights the decisive role of electronegativity in determining E_g, which is consistent with its fundamental influence on chemical bond strength and electronic structure.[58] These findings further validate the reliability of our model interpretability analysis.

4. Conclusion

In conclusion, we have performed a systematic high-throughput screening of X_3BA-type antiperovskite materials based on first-principles calculations. Six dynamically stable compounds were identified, including three alkali metal–based (Cs_3OCl, Cs_3SI, and Rb_3ClTe) and three alkaline-earth metal–based systems (Ca_3AsSb, Sr_3AsSb, and Ba_3AsSb). The HSE06 + SOC results indicate that the alkaline-earth metal–based compounds possess narrower bandgaps E_g due to the lower electronegativity of their constituent elements, while exhibiting strong optical absorption with coefficients exceeding 10^5 cm^{-1}, which can be attributed to their direct bandgaps and large transition dipole moments (P^2). Moreover, Sr_3AsSb and Ba_3AsSb exhibit theoretical maximum photovoltaic efficiencies of 28.1% and 29.4%, respectively, comparable to $CH_3NH_3PbI_3$ perovskites, suggesting their great potential as efficient photovoltaic absorbers. Interpretable machine learning analysis further reveals that the electronegativity of the A-site ion serves as the dominant descriptor governing the PBE + SOC E_g in cubic antiperovskites. Although this work focuses on 114 cubic compounds, future studies will expand the dataset to over 200 systems, including mixed halides, chalcogenides, and vacancy-ordered antiperovskites, to reduce chemical bias and test the universality of the descriptors. To assess the experimental feasibility, we calculated the decomposition enthalpies (ΔH) relative to binary precursors. The negative ΔH values for Sr_3AsSb (-2.003 eV/f.u.) and Ba_3AsSb (-2.121 eV/f.u.) indicate favorable thermodynamic stability (see Table S2), with magnitudes significantly smaller than those of typical perovskite oxides and halides.[58,59] Such characteristics suggest that these compounds can be synthesized via conventional solid-state reactions under inert conditions. Overall, this work provides both fundamental insights into the structure–property relationships of antiperovskites and concrete design strategies for developing stable, high-performance, and lead-free optoelectronic materials.

Acknowledgments

Project supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202200632), the National Natural Science Foundation of China (Grant No. 12574089), the Beijing Institute of Technology Research Fund Program for Young Scholars (Grant No. XSQD-202222008), the Beijing National Laboratory for Condensed Matter Physics (Grant No. 2023BNLCMPKF003), and the Guangdong Key Laboratory of Electronic Functional Materials and Devices Open Fund (Grant No. EFMD2023004M).

References

[1] Kim J Y, Lee J W, Jung H S, Shin H and Park N G 2020 *Chem. Rev.* **120** 7867
[2] Tiwari A, Satpute N S, Mehare C M and Dhoble S J 2021 *J. Alloys Compd.* **850** 156827
[3] Straus D B and Cava R J 2022 *ACS Appl. Mater. Interfaces* **14** 34884
[4] Dong H, Ran C, Gao L, Li M, Xia Y and Huang W 2023 *eLight* **3** 3
[5] Yang X, Fu Y, Su R, Zheng Y, Zhang Y, Yang W, Yu M, Chen P, Wang Y, Wu J, Luo D, Tu Y, Zhao L, Gong Q and Zhu R 2020 *Adv. Mater.* **32** 2002585
[6] Jiang J, Sun X, Chen X, Wang B, Chen Z, Hu Y, Guo Y, Zhang L, Ma Y, Gao L, Zheng F, Jin L, Chen M, Ma Z, Zhou Y, Padture N P, Beach K, Terrones H, Shi Y, Gall D, Lu T M, Wertz E, Feng J and Shi J 2019 *Nat. Commun.* **10** 4145
[7] Kojima A, Teshima K, Shirai Y and Miyasaka T 2009 *J. Am. Chem. Soc.* **131** 6050
[8] Green M A, Dunlop E D, Yoshita M, Kopidakis N, Bothe K, Siefer G, Hao X and Jiang J Y 2025 *Prog. Photovoltaics Res. Appl.* **33** 3
[9] Wolf C and Lee T W 2018 *Mater. Today Energy* **7** 199
[10] Leijtens T, Eperon G E, Noel N K, Habisreutinger S N, Petrozza A and Snaith H J 2015 *Adv. Energy Mater.* **5** 1500963
[11] Wang H, Li Q, Zhu Y, Sui X, Fan X, Lin M, Shi Y, Zheng Y, Yuan H, Zhou Y, Jin H, Yang H G, Hou Y and Yang S 2025 *Energy Environ. Sci.* **18** 2254
[12] Sun X, Shi W, Liu T, Cheng J, Wang X, Xu P, Zhang W, Zhao X and Guo W 2025 *Science* **388** 957
[13] Zhang Y, Song Q, Liu G, Chen Y, Guo Z, Li N, Niu X, Qiu Z, Zhou W, Huang Z, Zhu C, Zai H, Ma S, Bai Y, Chen Q, Huang W, Zhao Q and Zhou H 2023 *Nat. Photonics* **17** 1066
[14] Chowdhury T A, Bin Zafar M A, Sajjad-Ul Islam M, Shahinuzzaman M, Islam M A and Khandaker M U 2023 *RSC Adv.* **13** 1787
[15] Tan S, Shih M C, Lu Y, Choi S G, Dong Y, Lee J H, Yavuz I, Larson B W, Park S Y, Kodalle T, Zhang R, Grotevent M J, Lin Y K, Zhu H, Bulović V, Sutter-Fella C M, Park N G, Beard M C, Lee J W, Zhu K and Bawendi M G 2025 *Science* **388** 639
[16] Liu R, Lan C, Zeng M, Zheng Z, Zheng X, Guo R, Guo J, Yang S, Wang Z and Li X 2025 *Adv. Mater.* 2504321
[17] Deng C, Yang Y, Wu J, Tan L, Liu F, Du Y, Chen Q, Chen X, Sun L, Sun W, Lin J, Xie Y, Lan Z, Bai Y and Abate A 2025 *ACS Energy Lett.* 3132
[18] Duan L and Uddin A 2022 *Mater. Chem. Front.* **6** 400
[19] Yang H, Wang W, Wang Y, Hu L, Yang S and Jiao S 2025 *Adv. Mater.* **37** 2500031
[20] Zhong H, Feng C, Wang H, Han D, Yu G, Xiong W, Li Y, Yang M, Tang G and Yuan S 2021 *ACS Appl. Mater. Interfaces* **13** 48516
[21] Liu S, Biju V P, Qi Y, Chen W and Liu Z 2023 *NPG Asia Mater.* **15** 27
[22] Gao S, Broux T, Fujii S, Tassel C, Yamamoto K, Xiao Y, Oikawa I, Takamura H, Ubukata H, Watanabe Y, Fujii K, Yashima M, Kuwabara A, Uchimoto Y and Kageyama H 2021 *Nat. Commun.* **12** 201
[23] Tang G, Liu X, Wang S, Hu T, Feng C, Zhu C, Zhu B and Hong J 2024 *Mater. Horiz.* **11** 5320
[24] Kresse G and Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15
[25] Kresse G and Joubert D 1999 *Phys. Rev. B* **59** 1758
[26] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[27] Heyd J, Peralta J E, Scuseria G E and Martin R L 2005 *J. Chem. Phys.* **123** 174101
[28] Oba F and Kumagai Y 2018 *Appl. Phys. Express* **11** 060101
[29] Ganose A M, Jackson A J and Scanlon D O 2018 *J. Open Source Softw.* **3** 717
[30] Knoop F, Shulumba N, Castellano A, Batista J P A, Farris R, Verstraete M J, Heine M, Broido D, Kim D S, Klarbring J, Abrikosov I A, Simak S I and Hellman O 2024 *J. Open Source Softw.* **9** 6150
[31] Ganose A M, Park J, Faghaninia A, Woods-Robinson R, Persson K A and Jain A 2021 *Nat. Commun.* **12** 2222
[32] Wang V, Xu N, Liu J C, Tang G and Geng W T 2021 *Comput. Phys. Commun.* **267** 108033
[33] Raschka S, Patterson J and Nolet C 2020 *Information* **11** 193

[34] Chen T and Guestrin C 2016 *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA, pp. 785–794

[35] Lundberg S M, Erion G, Chen H, DeGrave A, Prutkin J M, Nair B, Katz R, Himmelfarb J, Bansal N and Lee S I 2020 *Nat. Mach. Intell.* **2** 56

[36] Nuss J, Muhle C, Hayama K, Abdolazimi V and Takagi H 2015 *Acta Cryst. B* **71** 300

[37] Goh W F and Pickett W E 2018 *Phys. Rev. B* **97** 035202

[38] Ahiavi E, Dawson J A, Kudu U, Courty M, Islam M S, Clemens O, Masquelier C and Famprikis T 2020 *J. Power Sources* **471** 228489

[39] Duran-Pinilla J M, Romero A H and Garcia-Castro A C 2022 *Phys. Rev. Mater.* **6** 125003

[40] Li X, Zhang Y, Kang W, Yan Z, Shen Y and Huo J 2023 *Comput. Mater. Sci.* **225** 112188

[41] Lin S, Yue J, Ren W, Shen C and Zhang H 2024 *J. Mater. Chem. A* **12** 19567

[42] Dawson J A, Famprikis T and Johnston K E 2021 *J. Mater. Chem. A* **9** 18746

[43] Shannon R 1976 *Acta Crystallographica Section A* **32** 751

[44] Gebhardt J and Rappe A M 2017 *ACS Energy Lett.* **2** 2681

[45] Mochizuki Y, Sung H J, Takahashi A, Kumagai Y and Oba F 2020 *Phys. Rev. Mater.* **4** 044601

[46] Travis W, Glover E N K, Bronstein H, Scanlon D O and Palgrave R G 2016 *Chem. Sci.* **7** 4548

[47] Shockley W and Queisser H 1961 *J. Appl. Phys.* **32** 510

[48] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G and Persson K A 2013 *APL Mater.* **1** 011002

[49] Sun W, Dacek S T, Ong S P, Hautier G, Jain A, Richards W D, Gamst A C, Persson K A and Ceder G 2016 *Sci. Adv.* **2** e1600225

[50] Rahim W, Skelton J M, Savory C N, Evans I R, Evans J S O, Walsh A and Scanlon D O 2020 *Chem. Sci.* **11** 7904

[51] Goldschmidt V M 1926 *Naturwissenschaften* **14** 477

[52] Naskar A, Khanal R and Choudhury S 2021 *Materials* **14** 1032

[53] Filippetti A, Mattoni A, Caddeo C, Saba M I and Delugas P 2016 *Phys. Chem. Chem. Phys.* **18** 15352

[54] Zhao T, Gibson Q D, Daniels L M, Slater B and Corà F 2020 *J. Mater. Chem. A* **8** 25245

[55] Blank B, Kirchartz T, Lany S and Rau U 2017 *Phys. Rev. Appl.* **8** 024032

[56] McInnes L, Healy J and Melville J 2018 arXiv:1802.03426

[57] Schubert E, Sander J, Ester M, Kriegel H P and Xu X 2017 *ACM Trans. Database Syst.* **42** Article 19

[58] Castelli I E, García-Lastra J M, Thygesen K S and Jacobsen K W 2014 *APL Mater.* **2** 081514

[59] Sun Q and Yin W J 2017 *J. Am. Chem. Soc.* **139** 14905