Anion Ordering and Phase Stability Govern Optical Band Gaps in $BaZrS_{3x}Se_{3-3x}$

Erik Fransson,[1] Michael Xu,[2] Prakriti Kayastha,[3,4] Kevin Ye,[2] Ida Sadeghi,[2]
Rafael Jaramillo,[2] James M. LeBeau,[2] Lucy Whalley,[3] and Paul Erhart[1,*]

[1]*Department of Physics and Astronomy, Chalmers University of Technology, SE-41296, Gothenburg, Sweden*
[2]*Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America*
[3]*Department of Engineering, Physics and Mathematics, Northumbria University, Newcastle upon Tyne, NE1 8QH, United Kingdom*
[4]*Department of Chemistry, University College London, London WC1E 6BT, United Kingdom*

(Dated: April 15, 2026)

Chalcogenide perovskites have emerged as promising lead-free materials for photovoltaic and thermoelectric applications. Among them, $BaZrS_3$ has attracted particular attention due to its thermal and chemical stability, favorable optoelectronic properties, and low thermal conductivity. Here, we combine molecular dynamics and Monte Carlo simulations based on a machine-learned interatomic potential with scanning transmission electron microscopy to investigate mixing thermodynamics and phase stability in the $BaZrS_{3x}Se_{3-3x}$ system. We identify an unusual ordered structure that persists at room temperature, most prominently at 33 % S, where S and Se atoms form alternating layers within the crystal. Free-energy calculations yield the temperature–composition phase diagram, including a non-perovskite δ phase in the Se-rich limit and a perovskite phase in the S-rich limit, separated by a broad two-phase region. Analysis of the dielectric function and the absorption coefficient demonstrates that composition, crystal structure, and anion ordering jointly control the optical band gap. Selenium alloying enables tuning between approximately 1.6 and 1.9 eV, while anion ordering within a given composition reduces the gap by about 0.12 eV. Lastly, variations between structural polymorphs give rise to band gap differences of up to 0.4 eV.

I. INTRODUCTION

Perovskites with the general formula ABX_3 constitute a versatile class of materials for optoelectronic applications. Hybrid lead–halide perovskites, in particular, have enabled record-breaking single-junction and tandem photovoltaic efficiencies [1]. However, their long-term chemical instability and the presence of toxic lead remain significant barriers to large-scale deployment. These limitations have motivated the search for alternative perovskite chemistries that retain favorable optoelectronic properties while improving chemical robustness and sustainability.

Chalcogenide perovskites (X = S, Se) provide one such platform. These materials combine strong optical absorption with improved environmental stability and are composed of earth-abundant, non-toxic elements [2–7]. Among them, $BaZrS_3$ has emerged as a model system, exhibiting strong band-to-band photoluminescence and favorable defect physics [8–11]. However, its band gap of 1.9 eV [12, 13] lies above the ideal range for single-junction photovoltaics, as the Shockley–Queisser limit places the maximum theoretical efficiency at 1.3 eV to 1.4 eV. Band gap engineering through alloying is therefore a natural strategy to enhance its photovoltaic potential. Partial substitution on the B-site with Ti [14, 15] or on the X-site with Se [16] reduces the band gap into the optimal range [13, 16–19]. An additional benefit is that Se incorporation is predicted to weaken electron–phonon coupling, potentially suppressing non-radiative recombination in $BaZrS_{3x}Se_{3-3x}$ alloys [18]. While alloying thus improves optoelectronic properties, it also modifies bonding and lattice energetics, raising fundamental questions about phase stability and anion ordering.

The structural stability of the perovskite sulfide end member is well established. At ambient conditions, $BaZrS_3$ adopts the corner-sharing perovskite structure with orthorhombic *Pnma* symmetry [20]. Upon heating, it undergoes displacive phase transitions to tetragonal and cubic polymorphs [21–24].

In contrast, there are conflicting structures reported for the perovskite selenide end member. An early experimental attempt to synthesize stoichiometric $BaZrSe_3$ via a solid-state reaction yielded off-stoichiometric, needle-like structures composed of face-sharing octahedra, rather than the corner-sharing perovskite network [25]. More recently, stoichiometric $BaZrSe_3$ in a closely related face-sharing hexagonal phase has been reported for nanocrystalline samples prepared by hot injection [26]. First-principles studies constrained to the 1:1:3 stoichiometry instead predict a different edge-sharing polymorph with *Pnma* symmetry [27, 28]. Se-rich corner-sharing perovskite phases, including stoichiometric $BaZrSe_3$, have been realized using non-equilibrium methods [16, 19]. Although these perovskite phases are likely metastable, they exhibit long-term stability at ambient conditions. For the mixed-anion series $BaZrS_{3x}Se_{3-3x}$, powder synthesis studies indicate phase segregation above approximately 40 % Se [9], also implying limited stability of the perovskite framework a high Se content under near-equilibrium conditions.

* erhart@chalmers.se

Collectively, the available evidence suggests that the perovskite form of $BaZrSe_3$ is not the equilibrium ground state under ambient conditions, but that it is kinetically stable. However, a unified thermodynamic description capable of reconciling equilibrium phase segregation, metastable perovskite formation, and possible anion ordering across composition and temperature has not yet been established. Developing such a description is essential for the optimization of mixed-anion chalcogenide perovskites and their optical properties.

Here, we establish the temperature–composition phase behavior of the mixed $BaZrS_{3x}Se_{3-3x}$ system combining atomic scale modeling and scanning transmission electron microscopy (STEM) analysis. Building on our previously developed dataset for $BaZrS_3$ [24], we train a machine-learned interatomic potential (MLIP) with density-functional theory (DFT) data. This approach enables large-scale sampling of the configurational and vibrational degrees of freedom while retaining close to hybrid-functional accuracy. Using molecular dynamics (MD) and Monte Carlo molecular dynamics (MCMD) simulations, we explore anion configurations and competing structural motifs over a broad range of temperatures and compositions, keeping the Ba:Zr:(S+Se) ratio fixed at 1:1:3. We predict the emergence of an unusual trans-ordered S/Se arrangement at 33 % S content, which we confirm experimentally by STEM. Furthermore, we construct a temperature–composition phase diagram that rationalizes the experimentally observed stability limits of the perovskite phase and clarifies the thermodynamic origins of metastability in Se-rich compositions. Together, these results provide quantitative thermodynamic guidelines for stabilizing Se-rich perovskite phases and tuning anion order to optimize band gap and phase stability in lead-free chalcogenide absorbers. Lastly, we analyze the dielectric function as well as the absorption coefficient, and demonstrate that composition, crystal structure, and anion ordering jointly control the optical band gap.

II. RESULTS AND DISCUSSION

A. Mixing energies and ordered structures

The MLIP model developed in this work is based on the neuroevolution potential (NEP) framework [30–35] (Supp. Note 1) and accurately reproduces the small energy differences predicted by DFT calculations [36] using the HSE06 hybrid functional [37] (Supp. Note 2) between the perovskite, needle-like (δ $Pnma$), and hexagonal ($P6_3/mmc$) phases for the end members $BaZrS_3$ and $BaZrSe_3$ (Supp. Note 4 and Table S1). In addition, the MLIP yields mixing energies in quantitative agreement with the corresponding DFT results across the full composition range (Supp. Note 4; Figure S2 and Figure S3), demonstrating the reliability of the model for thermodynamic analysis. We note that throughout this study the mixing energy is defined relative to the pure needle-like end members (Supp. Note 7).

The relative 0 K stability of the phases within the 1:1:3 stoichiometry was further assessed using simulated annealing in supercells containing several hundred atoms, which are computationally inaccessible at the DFT level (Supp. Note 4). This analysis identifies the δ needle-like phase as the lowest-energy structure across nearly the entire composition range up to approximately 95 % sulfur (Figure 1). These results indicate a two-phase coexistence region separating the perovskite and needle-like phases at low temperature, which we analyze in more detail below (Sect. II C).

For both the perovskite and needle-like structures, the mixing energy exhibits a pronounced local minimum near $x = 1/3$, corresponding to layered anion ordering patterns commensurate with the respective primitive unit cells (Figure 1b,c). In the perovskite case, this trans-type ordering is analogous to arrangements previously identified in mixed-halide perovskites [38]. However, such layered ordering appears to be uncommon in oxychalcogenide perovskites; oxysulfide and oxynitride perovskites typically favor cis-configurations, in which identical anions occupy adjacent sites rather than opposite vertices of the octahedron [39–42]. Studies of oxysulfides further demonstrate that anion ordering can strongly influence electronic properties, underscoring the importance of understanding the origin of the ordered structures identified here [40] (see Sect. II D).

We propose that layered ordering is energetically favored in $BaZrS_{3x}Se_{3-3x}$ because it enables octahedral tilting to adapt to the local chemical environment. In the orthorhombic $Pnma$ structure with tilt pattern $a^-a^-c^+$, layered ordering leads to the c^+ rotation acting predominantly on Se anions, while the a^- tilts involve both S and Se. Consistent with this interpretation, the relaxed Zr–Se–Zr bond angles are smaller than the corresponding Zr–S–Zr angles (155° versus 156°), reflecting a slightly larger tilt amplitude associated with the larger ionic radius of Se. Although layered ordering can induce significant interlayer strain in oxysulfide and oxynitride systems with large ionic-radius mismatch [40], the smaller size difference between S and Se likely mitigates such strain here, stabilizing the observed trans-type ordering.

B. Order-disorder transitions

To assess the finite-temperature ordering behavior of $BaZrS_{3x}Se_{3-3x}$, we performed heating simulations starting from the ordered ground states identified at 0 K (Supp. Note 5 and Supp. Note 6) [31, 43–46]. We focus on the composition $x = 1/3$, which exhibits the strongest ordering tendency in the mixing energies. The temperature evolution of the potential energy, heat capacity, and Warren–Cowley SRO parameter [47] reveals an order–disorder transition near 210 K for both the δ and γ phases (Figure 1d–g). The transition appears broad, consistent with a continuous loss of long-range order rather than a

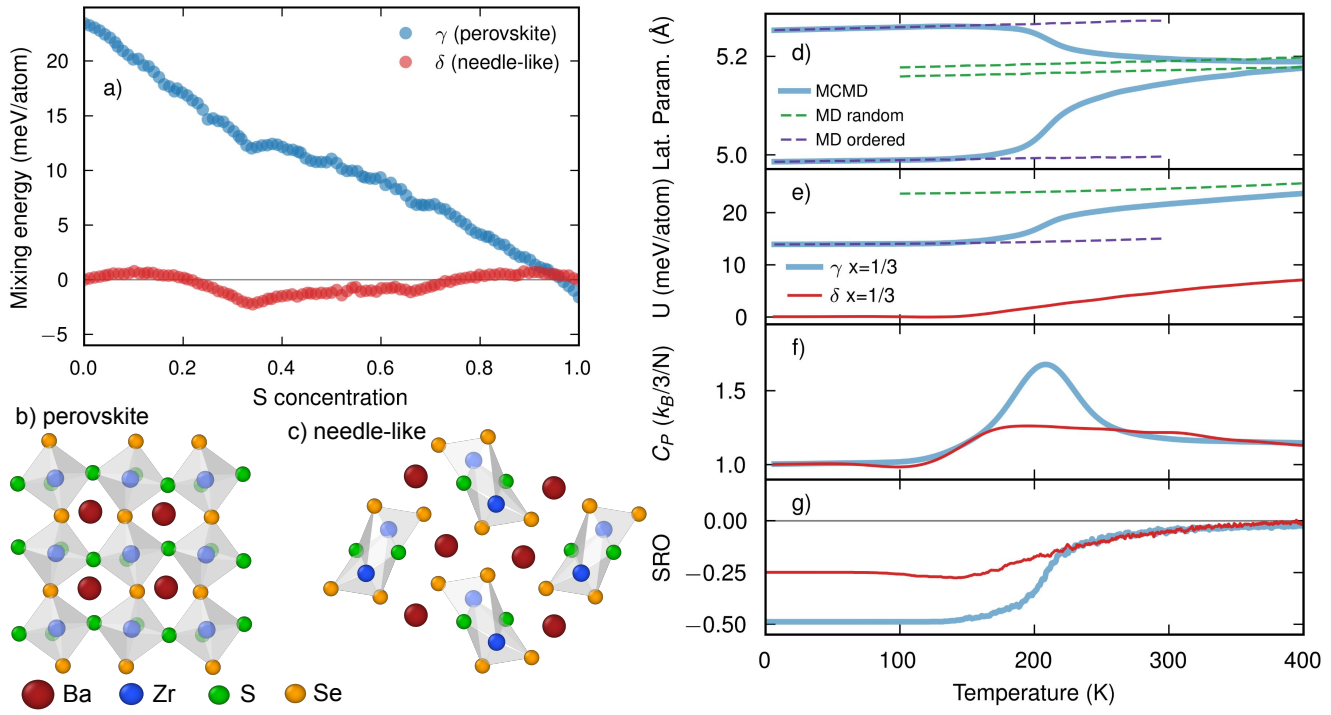

FIG. 1. **Mixing energies and order-disorder transition from simulations.** (a) Comparison of the mixing energies of the perovskite (γ) and needle-like (δ) phases obtained from MCMD cooling simulations. The structures in (b) and (c) display the ordered structures at $x = 1/3$ with Ba atoms shown in red, Zr atoms in blue, S atoms in green, and Se atoms in orange. The Zr-centered coordination polyhedra are shown in gray. Structures were visualized using OVITO [29]. (d–g) Thermodynamic properties from heating simulations at $x = 1/3$ starting from the ordered state. (d) Lattice parameters referred to the primitive cubic cell for the perovskite (γ) phase using both MCMD and MD with fully ordered or fully randomized occupations. (e) Change in potential energy after removal of the Dulong–Petit contribution ($\frac{3}{2}k_BT$), (f) heat capacity C_P, and (g) nearest-neighbor Warren–Cowley SRO parameter.

first-order discontinuity. Accordingly, no sharp jump in the energy is observed.

To demonstrate the impact of ordering, we additionally performed MD simulations of the perovskite phase with fully ordered and fully randomized anion occupations (Figure 1d,e). In these simulations, the system remains trapped in the imposed configurations, preventing equilibration between ordered and disordered states. In contrast, the combined MCMD approach enables exchange of S and Se on the X-site sublattice and produces a smooth crossover from the ordered to the disordered regime as temperature increases. Above the transition temperature, the SRO parameter decreases continuously toward zero, reflecting progressive randomization of S and Se on the X-site sublattice. Similar order–disorder behavior is observed across the composition range, but the transition is particularly pronounced at $x = 1/3$, where ideal commensurate ordering patterns are possible.

To provide experimental evidence for this unusual anion ordering and its associated order–disorder transition, we carried out STEM characterization of $BaZrS_{3x}Se_{3-3x}$ thin-film cross sections with compositions near $x = 1/3$ (Supp. Note 9–11) [12, 19]. Atomic-number-sensitive HAADF imaging along the pseudocubic ⟨100⟩ direction reveals an atom-column arrangement consistent with the perovskite structure of $BaZrS_{3x}Se_{3-3x}$ (Figure 2a,b). Significant intensity variations are observed at the anion columns, with columns along projected ⟨110⟩ directions exhibiting alternating low and high intensity. To quantify these variations, peak intensities extracted from two-dimensional Gaussian fits were mapped onto the projected structure (Figure 2c), revealing spatial patterns consistent with layered anion ordering.

To enable a direct comparison between experiment and simulation, we evaluated the Moran's I nearest-neighbor autocorrelation statistic for both experimental images and simulated STEM images generated from MCMD snapshots (Figure 2d,e; see Eq. (8)) [48]. As a function of increasing anion neighbor shell along ⟨110⟩ directions, the order parameters extracted from simulated images exhibit an order–disorder transition consistent with that obtained directly from the atomistic simulations. The experimentally determined order parameters at approximately 300 K fall between the simulated values at 200 and 300 K, indicating partial ordering at room temperature. These results confirm that the anion ordering predicted by simulations is observed experimentally at comparable

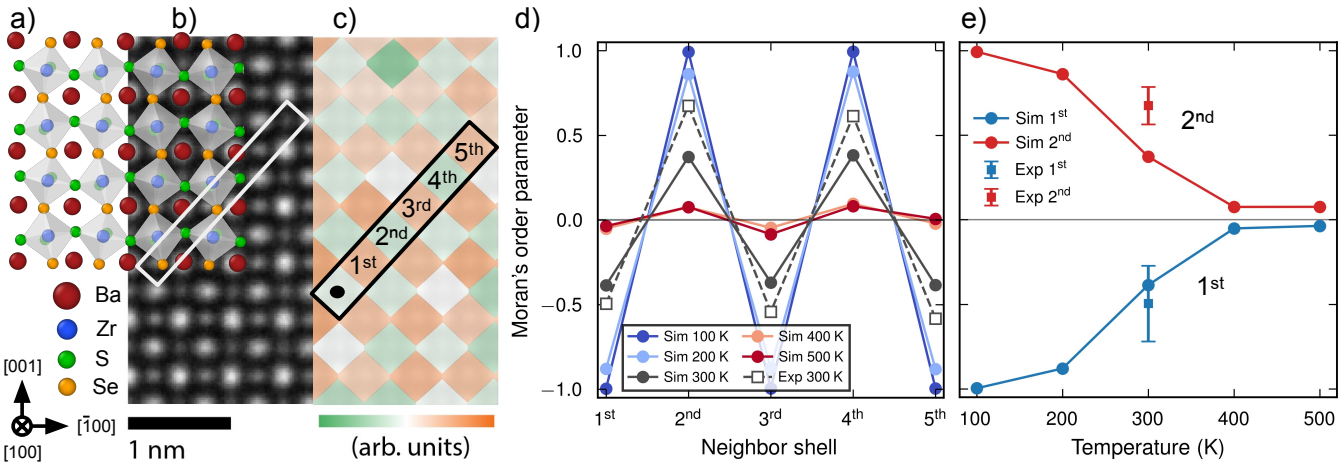

FIG. 2. **Experimental observation of ordering.** (a) Atomistic structure of the ideal ordered perovskite at x=1/3. (b) STEM HAADF image of a sample of $BaZrS_{3x}Se_{3-3x}$ with $x \approx 0.26$ along the pseudocubic ⟨100⟩ zone axis, overlaid with the ideal atomic structure. (c) Anion atom column intensities corresponding to the X-sites extracted from the HAADF images (where orange and green markers scale with higher (Se) and lower (S) image intensity, respectively). (d) Moran's I order parameter from experiment at approximately 300 K (Exp) and from images based on the atomistic simulations (Sim) as a function of increasing pseudocubic ⟨110⟩ neighbor shell. (e) Moran's order parameter as a function of temperature for the first and second neighbor shell.

temperature and composition.

It is important to note that the perovskite γ phase is metastable with respect to the needle-like δ phase within the 1:1:3 stoichiometry, implying the presence of significant kinetic barriers between these structures. The persistence of nonzero SRO above the nominal transition temperature, including at room temperature, suggests that local anion correlations can survive even after long-range order is lost.

C. Phase diagram

We now examine the temperature–composition phase behavior of $BaZrS_{3x}Se_{3-3x}$. We first determine the transitions among the perovskite polymorphs from MCMD heating simulations (Figure S4)which reveal successive transitions from the orthorhombic γ ($Pnma$) to the tetragonal β ($I4/mcm$) and finally to the cubic α ($Pm\bar{3}m$) phase across the composition range. These transitions are identified through a combined analysis of lattice parameters, heat capacities, and symmetry-adapted octahedral tilt mode projections [49] as functions of temperature.

In contrast, the transition between the needle-like δ phase and the perovskite γ phase is not directly accessible from MD or MCMD simulations because of the large nucleation barrier separating these structures. We therefore determine this boundary from free energy calculations (Supp. Note 6).

We begin by computing the Gibbs free energy of the pure $BaZrSe_3$ end member, $G(x = 0, T)$, using thermodynamic integration to an Einstein crystal reference (Figure 3a). The δ phase is stable up to approximately 870 K, above which the γ phase becomes thermodynamically favored. Although the γ phase is higher in energy by about 24 meV/atom at 0 K (Table S1), it possesses a significantly larger vibrational entropy, which stabilizes it at elevated temperature. This enhanced entropy arises from a greater density of low-frequency phonon modes in the γ phase (Figure 3a and Figure S5).

Using $G(x = 0, T)$ as a reference, the mixing free energy $\Delta G_{\mathrm{mix}}(x, T)$ is obtained across the full composition range via free energy integration (Supp. Note 7). At 300 K, a two-phase region is predicted between approximately $x = 0.2$ and $x = 0.95$, separating the γ and δ phases (Figure 3b). At 500 K, this coexistence region narrows and shifts asymmetrically to approximately $x = 0.25$ and $x = 0.65$ (Figure 3c).

Free energy calculations performed over a dense grid of temperatures then yield the full temperature–composition phase diagram of $BaZrS_{3x}Se_{3-3x}$ (Figure 3d). On the Se-rich side, the needle-like δ phase is thermodynamically stable, whereas on the S-rich side the perovskite phase is favored, with a two-phase coexistence region at intermediate compositions. At higher temperatures, the orthorhombic γ phase is stabilized across the entire composition range and subsequently transforms to the tetragonal β and cubic α polymorphs upon further heating.

When synthesized via a conventional solid-state reaction at 700 K [9], Se substitution up to approximately 40 % in the sulfide perovskite has been reported without observation of non-perovskite precipitates at room temperature. Our calculations predict that only about 9 % Se is thermodynamically soluble in the perovskite

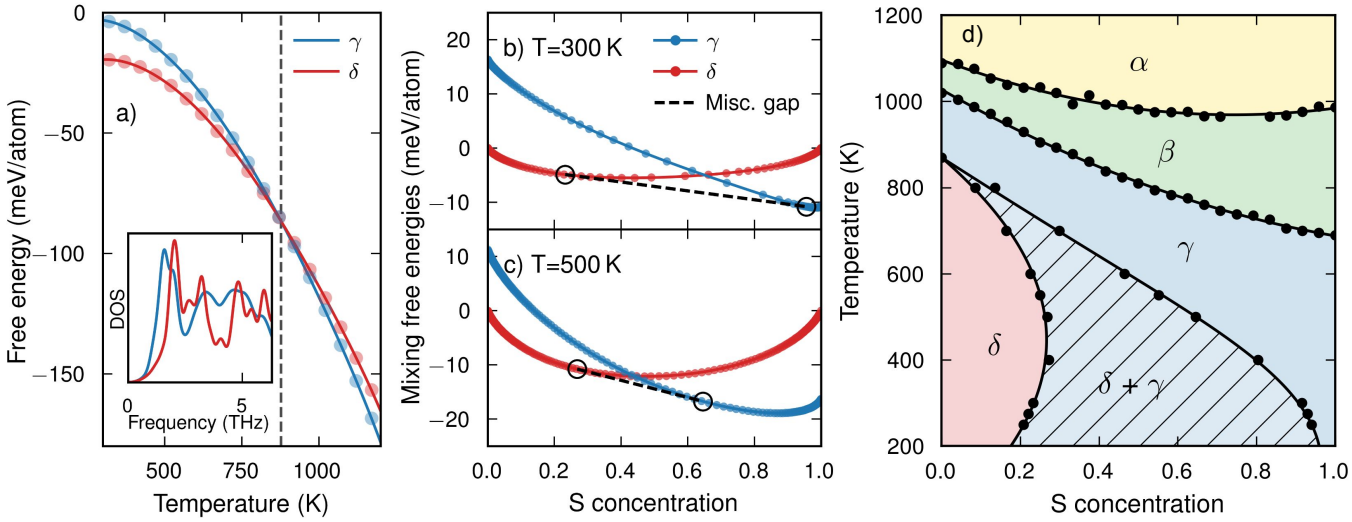

FIG. 3. **Free energies and phase diagram.** (a) Free energies of the pure Se (x=0) γ and δ phases. The inset shows the vibrational density of states obtained from MD at 300 K via the mass-weighted velocity autocorrelation function. Mixing free energies for γ and δ phases at (b) 300 K and (c) 500 K. The dashed lines indicate the miscibility gap with end points being marked with open circles. (c) Predicted phase diagram for the $BaZrS_{3x}Se_{3-3x}$ system. Here, α ($Pm\bar{3}m$), β ($I4/mcm$), and γ ($Pnma$) refer to the three perovskite phases, while δ refers to the needle-like phase ($Pnma$). The striped $\delta+\gamma$ region indicates the two-phase region.

phase at 300 K. However, the calculated solubility limit increases to approximately 20 % at 400 K and to 40 % at 500 K, bringing the theoretical predictions closer to experimental observations under elevated-temperature growth conditions. Furthermore, the computed free energy differences are small, and modest uncertainties in the underlying DFT reference data, the MLIP model, or the sampling can shift phase boundaries. Indeed, first-principles predictions of phase transition temperatures in similar systems can commonly differ from experiment by 100 K to 200 K depending on the underlying exchange-correlation functional [24, 50]. Finally, we note that the persistence of the single-phase perovskite at room temperature indicates kinetic limitations that suppress nucleation of the thermodynamically favored δ phase at low temperature.

D. Optoelectronic Properties

Finally, to understand how the electronic and optical properties of the material are affected by changes in composition, phase, and atomic ordering, we compute the dielectric function and the absorption coefficient. All calculations are performed at the PBE level, followed by two corrections: a band gap correction derived from HSE06 ($\alpha = 0.25$) calculations and a spin-orbit coupling (SOC) correction. The latter decreases the band gap by up to 0.17 eV for Se-containing compounds (see Supp. Note 3 for further details).

The absorption coefficients of the perovskite and needle-like phases at 300 K are calculated for all compositions (Figure 4a,b). For both phases, the onset of absorption shifts monotonically to higher energies with increasing sulfur concentration. Notably, the needle phase exhibits an extended but weak absorption tail, in contrast to the sharper onset observed for the perovskite phase. A similar difference between the two phases has previously been reported for $SrZrS_3$ [51].

From the absorption coefficients, we extract the optical band gap via a Tauc analysis (Figure 4c; see Supp. Note 3 and Figure S6 for details). For both phases, the Tauc gap increases monotonically with increasing sulfur content, in agreement with previous experimental and theoretical studies [16–18]. The predictions are in very good agreement with experimental data from Ref. 16, particularly for the end members.

Previous DFT studies have reported that the needle phase exhibits a lower *fundamental* band gap than the perovskite phase [52]. Consistently, the band gaps obtained directly from the electronic eigenstates are smaller for the needle phase than for the perovskite phase (Table S2). In contrast, the Tauc analysis shows that the *optical* gap of the needle phase is notably larger. This discrepancy arises from the extended but weak absorption tail observed in the needle phase (Figure 4b), highlighting that the fundamental and optical gaps can differ significantly.

Finally, we demonstrate that the anion ordering identified in this work (Figure 2) has a significant impact on the optical properties. To this end, we perform three types of simulations at 33 % S and temperatures of 100, 200, and 300 K: MD with fully ordered anion occupation, MD with random occupation, and fully equilibrated MCMD.

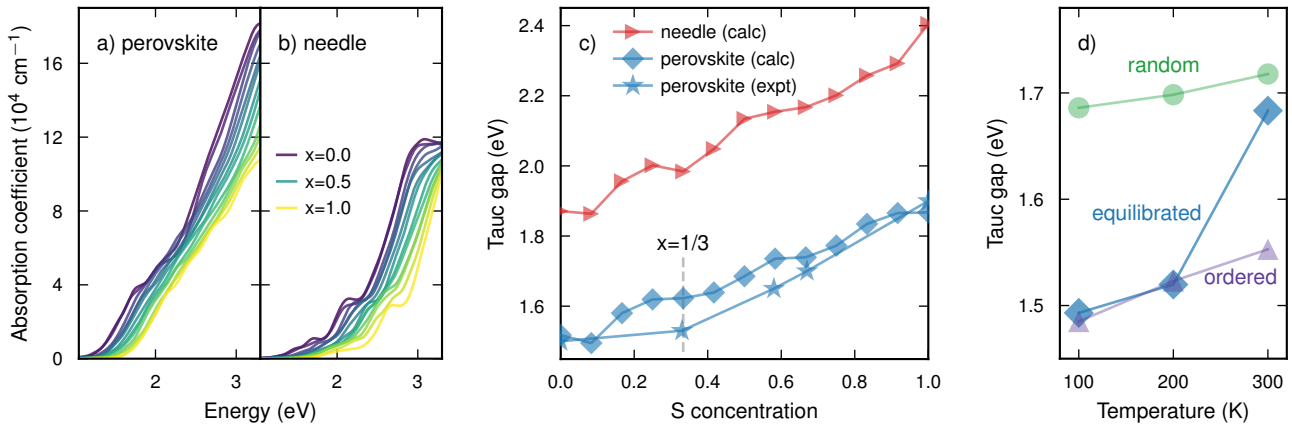

FIG. 4. **Optical properties.** a, b) Absorption coefficients of the perovskite and needle-like phases, respectively, at 300 K for varying sulfur concentrations using equilibrated MCMD snapshots. c) Tauc gaps obtained from the data in (a,b) (calculated), together with experimental data from Ref. 16. d) Tauc gaps at 33 % sulfur concentration obtained from three different simulations: MD with random occupation, MD with ordered occupation, and fully equilibrated MCMD, at 100, 200, and 300 K.

Analysis of the optical properties from MD simulations with fixed occupations (ordered or random) reveals a consistent difference of 0.16 eV to 0.19 eV in the Tauc gap, independent of temperature. In both cases, the Tauc gap increases with temperature, opposite to typical semiconductor behavior but consistent with trends reported for other perovskites [53]. The observed difference is therefore attributable to the ordering of S and Se atoms in the system. The fully equilibrated structures follow the Tauc gap of the ordered configuration up to 200 K. Between 200 K and 300 K, however, the Tauc gap increases rapidly toward the value of the random configuration. This behavior is consistent with the order–disorder transition identified above, occurring between 200 and 300 K.

III. CONCLUSIONS

By combining a MLIP with large-scale MD simulations, we have revealed ordering phenomena in the $BaZrS_{3x}Se_{3-3x}$ perovskite system. The simulations predict the emergence of a trans-ordered configuration with alternating layers of S and Se atoms, which is particularly pronounced at $x = 1/3$. This ordering is supported by STEM measurements and corresponding image simulations. In addition, we performed free-energy calculations based on MCMD simulations to construct the full phase diagram, considering all relevant ABX_3 polymorphs. Our results show that while the perovskite phase is thermodynamically stable only at high sulfur content, a wide two-phase region exists, and the perovskite phase remains dynamically stable across the entire composition range. Finally, we demonstrate that composition, crystal phase, and anion ordering all strongly influence the optoelectronic properties of this prototypical chalcogenide perovskite. In particular, the order–disorder transition occurring near room temperature leads to a shift in the Tauc gap of approximately 0.16 eV to 0.19 eV. This demonstrates that a detailed understanding of the underlying thermodynamic properties, such as phase stability and ordering, is essential for the rational design of materials with tailored properties.

DATA AVAILABILITY STATEMENT

The NEP model and DFT data generated in this study are openly available via Zenodo at `https://doi.org/10.5281/zenodo.19483979`.

ACKNOWLEDGMENTS

This work has been supported by the Swedish Research Council (Nos. 2020-04935 and 2021-05072) as well as the Turing Scheme and the UK Engineering and Physical Sciences Research Council (EPSRC) CDT in Renewable Energy Northeast Universities (ReNU) via Grant EP/S023836/1. The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at C3SE, partially funded by the Swedish Research Council through grant agreement no. 2022-06725, the Berzelius resource provided by the Knut and Alice Wallenberg Foundation at NSC, and the ARCHER2 UK National Supercomputing Service through the HEC Materials Chemistry Consortium (EPSRC EP/X035859), and the UK Materials and Molecular Modelling Hub (EPSRC EP/T022213, EP/W032260, EP/P020194). The authors acknowledge support from the National Science Foundation (NSF) under grant no. 1751736, "CAREER: Fundamentals of Complex Chalcogenide Electronic Materials".

A portion of this project was funded by the Skolkovo Institute of Science and Technology as part of the MIT-Skoltech Next Generation Program. K. Y. acknowledges support by the NSF Graduate Research Fellowship, grant no. 1745302. M. X. and J. M. L acknowledge support for this work from the Air Force Office of Scientific Research (FA9550-20-0066) and the MIT Research Support Committee. This work made use of the MIT.nano Characterization facilities.

[1] NREL Best Research-Cell Efficiency Chart, `https://www.nrel.gov/pv/cell-efficiency.html`.

[2] K. V. Sopiha, C. Comparotto, J. A. Márquez, and J. J. Scragg, Chalcogenide perovskites: Tantalizing prospects, challenging materials, Advanced Optical Materials **10**, 2101704 (2022).

[3] R. Jaramillo and J. Ravichandran, In praise and in search of highly-polarizable semiconductors: Technological promise and discovery strategies, APL Materials **7**, 100902 (2019).

[4] S. Niu, D. Sarkar, K. Williams, Y. Zhou, Y. Li, E. Bianco, H. Huyan, S. B. Cronin, M. E. McConney, R. Haiges, *et al.*, Optimal bandgap in a 2D Ruddlesden–Popper perovskite chalcogenide for single-junction solar cells, Chemistry of Materials **30**, 4882 (2018).

[5] P. Kayastha, D. Tiwari, A. Holland, O. S. Hutter, K. Durose, L. D. Whalley, and G. Longo, High-temperature equilibrium of 3D and 2D chalcogenide perovskites, Solar RRL **7**, 2201078 (2023).

[6] E. Osei-Agyemang, N. Koratkar, and G. Balasubramanian, Examining the electron transport in chalcogenide perovskite $BaZrS_3$, Journal of Materials Chemistry C **9**, 3892 (2021).

[7] Y. Wu, Y. Chen, Z. Fang, Y. Ding, Q. Li, K. Xue, H. Shao, H. Zhang, and L. Zhou, Ultralow lattice thermal transport and considerable wave-like phonon tunneling in chalcogenide perovskite $BaZrS_3$, The Journal of Physical Chemistry Letters **14**, 11465 (2023).

[8] R. S. Nielsen, Á. L. Álvarez, Y. Tomm, G. Gurieva, A. Ortega-Guerrero, J. Breternitz, L. Bastonero, N. Marzari, C. A. Pignedoli, S. Schorr, *et al.*, $BaZrS_3$ lights up: The interplay of electrons, photons, and phonons in strongly luminescent single crystals, Advanced Optical Materials **13**, e00915 (2025).

[9] Y. Nishigaki, T. Nagai, M. Nishiwaki, T. Aizawa, M. Kozawa, K. Hanzawa, Y. Kato, H. Sai, H. Hiramatsu, H. Hosono, *et al.*, Extraordinary strong band-edge absorption in distorted chalcogenide perovskites, Solar RRL **4**, 1900555 (2020).

[10] Z. Yuan, D. Dahliah, R. Claes, A. Pike, D. P. Fenning, G.-M. Rignanese, and G. Hautier, Assessing carrier mobility, dopability, and defect tolerance in the chalcogenide perovskite $BaZrS_3$, PRX Energy **3**, 033008 (2024).

[11] Y. Mualem, D. Cahen, Y. Rakita, and H. N. Barad, Using bond characteristics to predict self-healing properties in materials for sustainable energy conversion, ChemRxiv 10.26434/chemrxiv-2025-8r4l2-v2 (2025).

[12] I. Sadeghi, K. Ye, M. Xu, Y. Li, J. M. LeBeau, and R. Jaramillo, Making $BaZrS_3$ Chalcogenide Perovskite Thin Films by Molecular Beam Epitaxy, Advanced Functional Materials **31**, 2105563 (2021).

[13] S. Agarwal, K. C. Vincent, and R. Agrawal, From synthesis to application: a review of $BaZrS_3$ chalcogenide perovskites, Nanoscale **17**, 4250 (2025).

[14] X. Wei, H. Hui, S. Perera, A. Sheng, D. F. Watson, Y.-Y. Sun, Q. Jia, S. Zhang, and H. Zeng, Ti-alloying of $BaZrS_3$ chalcogenide perovskite for photovoltaics, ACS Omega **5**, 18579 (2020).

[15] D. Zilevu, K. M. Miller, N. Arrykova, A. B. Locke, and S. E. Creutz, Solution-phase synthesis of alloyed $BaZr_{1-x}Ti_xS_3$ perovskite and non-perovskite nanomaterials, Nanoscale **16**, 17126 (2024).

[16] I. Sadeghi, J. Van Sambeek, T. Simonian, M. Xu, K. Ye, T. Cai, V. Nicolosi, J. M. LeBeau, and R. Jaramillo, Expanding the perovskite periodic table to include chalcogenide alloys with tunable band gap spanning 1.5–1.9 eV, Advanced Functional Materials **33**, 2304575 (2023).

[17] W. Meng, B. Saparov, F. Hong, J. Wang, D. B. Mitzi, and Y. Yan, Alloying and defect control within chalcogenide perovskites for optimized photovoltaic application, Chemistry of Materials **28**, 821–829 (2016).

[18] Q.-q. Li, L. Yan, W. Chu, J. He, H. Luo, T. Frauenheim, S. Tretiak, and L. Zhou, Control of polaronic behavior and carrier lifetimes via metal and anion alloying in chalcogenide perovskites, The Journal of Physical Chemistry Letters **13**, 4955–4962 (2022).

[19] K. Ye, I. Sadeghi, M. Xu, J. Van Sambeek, T. Cai, J. Dong, R. Kothari, J. M. LeBeau, and R. Jaramillo, A processing route to chalcogenide perovskites alloys with tunable band gap via anion exchange, Advanced Functional Materials **34**, 2405135 (2024).

[20] A. Jess, R. Yang, and C. J. Hages, On the phase stability of chalcogenide perovskites, Chemistry of Materials **34**, 6894–6901 (2022).

[21] N. Jaykhedkar, R. Bystrický, M. Sýkora, and T. Bučko, How the temperature and composition govern the structure and band gap of Zr-based chalcogenide perovskites: Insights from ML accelerated AIMD, Inorganic Chemistry **62**, 12480 (2023).

[22] R. Bystrický, S. K. Tiwari, P. Hutár, and M. Sýkora, Thermal stability of chalcogenide perovskites, Inorganic Chemistry **63**, 12826 (2024).

[23] A. Jaiswal, K. A. Sakharov, Y. Lekina, K. Kamonsuangkasem, Y. Tomm, F. Wei, and T. J. White, High-temperature polymorphism and band-gap evolution in $BaZrS_3$, Inorganic Chemistry **63**, 24157 (2024).

[24] P. Kayastha, E. Fransson, P. Erhart, and L. Whalley, Octahedral tilt-driven phase transitions in $BaZrS_3$ chalcogenide perovskite, The Journal of Physical Chemistry Letters **16**, 2064 (2025).

[25] L. J. Tranchitella, J. C. Fettinger, P. K. Dorhout, P. M. Van Calcar, and B. W. Eichhorn, Commensurate columnar composite compounds: Synthesis and structure of $Ba_{15}Zr_{14}Se_{42}$ and $Sr_{21}Ti_{19}Se_{57}$, Journal of the American Chemical Society **120**, 7639–7640 (1998).

[26] S. Agarwal, S. Rodriguez Perilla, M. Rios Marques, D. C. Hayes, K. C. Vincent, and R. Agrawal, Hexagonal ABX_3 nanocrystals: rod-shaped $BaNbS_3$ and $BaTaS_3$; $BaTiSe_3$,

$BaZrSe_3$, and other selenide derivatives for optoelectronic applications, Nanoscale Advances **7**, 6864 (2025).

[27] M. Ong, D. M. Guzman, Q. Campbell, I. Dabo, and R. A. Jishi, $BaZrSe_3$: Ab initio study of anion substitution for bandgap tuning in a chalcogenide material, Journal of Applied Physics **125**, 235702 (2019).

[28] D. Liu, H. Zeng, H. Peng, and R. Sa, Computational study of the fundamental properties of Zr-based chalcogenide perovskites for optoelectronics, Physical Chemistry Chemical Physics **25**, 13755–13765 (2023).

[29] A. Stukowski, Visualization and analysis of atomistic simulation data with ovito–the open visualization tool, Modelling and Simulation in Materials Science and Engineering **18**, 015012 (2009).

[30] Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, A. J. Gabourie, J. Liu, H. Dong, J. Wu, Y. Chen, Z. Zhong, J. Sun, P. Erhart, Y. Su, and T. Ala-Nissila, GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations, The Journal of Chemical Physics **157**, 114801 (2022).

[31] K. Xu, H. Bu, S. Pan, E. Lindgren, Y. Wu, Y. Wang, J. Liu, K. Song, B. Xu, Y. Li, T. Hainer, L. Svensson, J. Wiktor, R. Zhao, H. Huang, C. Qian, S. Zhang, Z. Zeng, B. Zhang, B. Tang, Y. Xiao, Z. Yan, J. Shi, Z. Liang, J. Wang, T. Liang, S. Cao, Y. Wang, P. Ying, N. Xu, C. Chen, Y. Zhang, Z. Chen, X. Wu, W. Jiang, E. Berger, Y. Li, S. Chen, A. J. Gabourie, H. Dong, S. Xiong, N. Wei, Y. Chen, J. Xu, F. Ding, Z. Sun, T. Ala-Nissila, A. Harju, J. Zheng, P. Guan, P. Erhart, J. Sun, W. Ouyang, Y. Su, and Z. Fan, GPUMD 4.0: A high-performance molecular dynamics package for versatile materials simulations with machine-learned potentials, Materials Genome Engineering Advances **3**, e70028 (2025).

[32] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, The atomic simulation environment—a python library for working with atoms, Journal of Physics: Condensed Matter **29**, 273002 (2017).

[33] F. Eriksson, E. Fransson, and P. Erhart, The hiphive package for the extraction of high-order force constants by machine learning, Advanced Theory and Simulations **2**, 1800184 (2019).

[34] M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod, and P. Erhart, ICET – A python library for constructing and sampling alloy cluster expansions, Advanced Theory and Simulations **2**, 1900015 (2019).

[35] G. L. W. Hart and R. W. Forcade, Algorithm for generating derivative structures, Physical Review B **77**, 224115 (2008).

[36] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Ab initio molecular simulations with numeric atom-centered orbitals, Computer Physics Communications **180**, 2175 (2009).

[37] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, Influence of the exchange screening parameter on the performance of screened hybrid functionals, J. Chem. Phys. **125**, 224106 (2006).

[38] J. S. Bechtel and A. Van der Ven, First-principles thermodynamics study of phase stability in inorganic halide perovskite solid solutions, Physical Review Materials. **2**, 045401 (2018).

[39] G. Pilania, A. Ghosh, S. T. Hartman, R. Mishra, C. R. Stanek, and B. P. Uberuaga, Anion order in oxysulfide perovskites: origins and implications, npj Computational Materials **6**, 71 (2020).

[40] H. Shang, D. West, H. Hui, H. Zeng, and S. Zhang, Spontaneous ordering in ultra-large anion size-mismatched $BaZr(S_{1-x}O_x)_3$, Applied Physics Letters **126**, 232111 (2025).

[41] S. H. Porter, Z. Huang, and P. M. Woodward, Study of anion order/disorder in $RTaN_2O$ (R = La, Ce, Pr) perovskite nitride oxides, Crystal Growth & Design **14**, 117 (2014).

[42] M. Yang, J. Oró-Solé, J. A. Rodgers, A. B. Jorge, A. Fuertes, and J. P. Attfield, Anion order in perovskite oxynitrides, Nature Chemistry **3**, 47 (2011).

[43] R. Freitas, M. Asta, and M. de Koning, Nonequilibrium free-energy calculation of solids using LAMMPS, Computational Materials Science **112**, 333 (2016).

[44] J. G. Kirkwood, Statistical Mechanics of Fluid Mixtures, The Journal of Chemical Physics **3**, 300 (1935).

[45] N. L. Allan, G. D. Barrera, M. Y. Lavrentiev, I. T. Todorov, and J. A. Purton, Ab initio calculation of phase diagrams of ceramics and minerals, Journal of Materials Chemistry **11**, 63–68 (2001).

[46] B. Sadigh, P. Erhart, A. Stukowski, A. Caro, E. Martinez, and L. Zepeda-Ruiz, Scalable parallel monte carlo algorithm for atomistic simulations of precipitation in alloys, Physical Review B **85**, 184203 (2012).

[47] J. M. Cowley, An approximate theory of order in alloys, Physical Review **77**, 669 (1950).

[48] M. Xu, A. Kumar, and J. M. LeBeau, Correlating local chemical and structural order using Geographic Information Systems-based spatial statistics, Ultramicroscopy **243**, 113642 (2023).

[49] E. Fransson, P. Rosander, F. Eriksson, J. M. Rahm, T. Tadano, and P. Erhart, Limits of the phonon quasiparticle picture at the cubic-to-tetragonal phase transition in halide perovskites, Communications Physics **6**, 173 (2023).

[50] E. Fransson, J. Wiktor, and P. Erhart, Phase transitions in inorganic halide perovskites from machine-learned potentials, The Journal of Physical Chemistry C **127**, 13773 (2023).

[51] S. Niu, H. Huyan, Y. Liu, M. Yeung, K. Ye, L. Blankemeier, T. Orvis, D. Sarkar, D. J. Singh, R. Kapadia, and J. Ravichandran, Bandgap control via structural and chemical tuning of transition metal perovskite chalcogenides, Advanced Materials **29**, 1604733 (2016).

[52] Y.-Y. Sun, M. L. Agiorgousis, P. Zhang, and S. Zhang, Chalcogenide perovskites for photovoltaics, Nano Letters **15**, 581 (2015).

[53] G. Mannino, I. Deretzis, E. Smecca, A. La Magna, A. Alberti, D. Ceratti, and D. Cahen, Temperature-dependent optical band gap in $CsPbBr_3$, $MAPbBr_3$, and $FAPbBr_3$ single crystals, The Journal of Physical Chemistry Letters **11**, 2490 (2020).

Supporting Information

Anion Ordering and Phase Stability Govern Optical Band Gaps in $BaZrS_{3x}Se_{3-3x}$

Erik Fransson,[1] Michael Xu,[2] Prakriti Kayastha,[3,4] Kevin Ye,[2] Ida Sadeghi,[2]
Rafael Jaramillo,[2] James M. LeBeau,[2] Lucy Whalley,[3] and Paul Erhart[1,*]

[1] Department of Physics, Chalmers University of Technology, SE-41296, Gothenburg, Sweden
[2] Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America
[3] Department of Mathematics, Physics and Electrical Engineering, Northumbria University, Newcastle upon Tyne, NE1 8QH, United Kingdom
[4] Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom
[*] erhart@chalmers.se

Contents

Supplementary Notes

Supplementary Note 1: Neuroevolution potential construction

A machine-learned interatomic potential (MLIP) was constructed using the neuroevolution potential (NEP) framework as implemented in the GPUMD package (1, 2). The NEP formalism was chosen for its high computational efficiency relative to other state-of-the-art MLIP approaches, which is essential for the extensive configurational and thermodynamic sampling performed here. Its local descriptor-based architecture additionally enables efficient local Monte Carlo updates, providing a substantial computational advantage for hybrid Monte Carlo molecular dynamics (MCMD) simulations (Note 7).

Training structures were generated from the relevant structural prototypes, including corner-sharing perovskites ($Pnma$, $I4/mcm$, $Pm\bar{3}m$), a needle-like edge-sharing phase ($Pnma$), and the hexagonal phase ($P6_3/mmc$) using the ASE (3) and HIPHIVE (4) packages. Mixed-anion configurations were generated by enumeration using ICET (5) following the algorithm of Hart and Forcade (6), complemented by randomly occupied supercells. The training set was iteratively refined by adding structures generated via molecular dynamics (MD) and MCMD simulations in the NPT ensemble between 0 and 1200 K. About 100 selected structures from the $BaZrS_3$ training dataset of Ref. 7 were also incorporated.

The final model achieves for the train (test) set a energy root-mean-square error (RMSE) of 0.0034 eV/atom (0.0049 eV/atom), a force RMSE of 0.0835 eV/Å (0.0843 eV/Å), and a stress RMSE of 0.2058 GPa (0.2093 GPa).

The ASE (3) and CALORINE (8) packages were used to prepare the training structures, set up MD simulations and post-process the results. Random displacements were generated using the HIPHIVE package (4).

Supplementary Note 2: Density functional theory calculations

All density-functional theory (DFT) calculations were performed with the FHI-AIMS code (9) using the HSE06 exchange–correlation functional(10), which has been previously found to yield structural parameters and phase transitions in good agreement with experiment (7). All computational parameters were identical to those reported in Ref. 7. A Monkhorst–Pack k-point mesh with a minimum k-spacing of 0.2/Å was employed. The *light* basis set was used throughout. Stress tensors were evaluated for small and medium-sized cells included in the training set, but were not computed for structures containing more than approximately 50 atoms due to memory constraints.

Supplementary Note 3: Dielectric function calculations

Dielectric functions were calculated using VASP(11) with the PBE functional (12), using projector-augmented wave (13, 14) setups. We use an plane wave energy cutoff of 520 eV automatically generated **k**-point grids with a maximum spacing of 0.2/Å.

The dielectric function $\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega)$ is converted to the absorption coefficient via

$$\alpha(\omega) = \frac{\sqrt{2}\,\omega}{c}\sqrt{\sqrt{\epsilon_1(\omega)^2 + \epsilon_2(\omega)^2} - \epsilon_1(\omega)}. \tag{1}$$

The Tauc gap is subsequently extracted from the absorption coefficient by plotting $(\alpha(\omega)\cdot\hbar\omega)^2$ as a function of photon energy and performing a linear fit to the onset region, excluding the low-energy Urbach tail (Figure S6). The optical band gap (Tauc gap) is taken as the zero-energy intercept of this linear fit.

For each composition and temperature, 5 to 10 snapshots extracted from NVT MCMD and MD simulations are used, with lattice parameters fixed to those obtained from the corresponding NPT runs. The dielectric functions are averaged over these snapshots for each composition and phase.

To correct for the band gap underestimation of PBE, we compute HSE06 band gaps for a representative set of snapshots using FHI-AIMS, obtaining a consistent PBE-to-HSE06 shift of 0.63 eV, which is applied as a rigid scissors correction to all dielectric functions.

Additionally, we assess the role of spin-orbit coupling (SOC) by computing HSE06+SOC band gaps for the fully relaxed pure S and Se end-members in both the perovskite and needle (both $Pnma$) phases; results are provided in Table S2. SOC reduces the band gap in both cases, with a larger effect for Se than for S, consistent with the expected trend for heavier chalcogenides (15).

Supplementary Note 4: Ground state structures

For the sulfide, the perovskite γ ($Pnma$) phase is identified as the ground state, in agreement with experiment, with the non-perovskite δ phase lying 2 meV/atom higher in energy. The hexagonal $P6_3/mmc$ phase is substantially higher in energy, by 72 meV/atom. For the selenide, the non-perovskite δ phase is identified as the ground state, with the perovskite γ phase 24 meV/atom higher in energy (see Table S1). Again, the hexagonal $P6_3/mmc$ phase is much higher in energy at 33 meV/atom, and remains energetically unfavorable across the full composition range. Accordingly, this hexagonal phase is excluded from further thermodynamic analysis.

Supplementary Note 5: Mixing energies

Mixing energies at 0 K were computed for relaxed structures generated by enumeration of symmetry-inequivalent configurations based on the parent perovskite and δ needle-like phases, as well as from cooling MCMD simulations (Figure S2 and Figure S3). The cooling simulations were performed from 1200 K to 0 K over 40 ns using supercells containing approximately 300 to 400 atoms. For nearly all compositions, the cooling simulations yield structures with energies equal to or lower than those obtained from enumeration. The relaxed enumerated structures themselves remain in quantitative agreement with direct DFT calculations, confirming the reliability of the NEP model.

Supplementary Note 6: Molecular dynamics and free energy calculations

All MD simulations employed a timestep of 1 fs. Constant-pressure simulations were performed in the NPT ensemble using the stochastic cell rescaling barostat with a pressure target of 0 GPa. Equilibrium lattice parameters at each temperature and composition were obtained by averaging over heating and cooling runs conducted at a rate of -1 K/ns.

The Helmholtz free energies of the end members ($x = 0$ and $x = 1$) were computed using thermodynamic integration (TI) with respect to an Einstein crystal reference (16), as implemented in GPUMD (2). The interpolated Hamiltonian was defined as

$$H(\lambda) = (1-\lambda)H_{\mathrm{Ein}} + \lambda H_{\mathrm{NEP}}, \tag{2}$$

where λ is the Kirkwood coupling parameter (17). The free energy difference between the NEP system and the Einstein crystal reference is given by

$$G_{\mathrm{NEP}} - G_{\mathrm{Ein}} = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda, \tag{3}$$

where $\langle \cdots \rangle_\lambda$ denotes an ensemble average evaluated with the Hamiltonian $H(\lambda)$. TI simulations were performed for 0.1 ns using a spring constant of 4 eV/Å^2 and a supercell containing 23 040 atoms.

Supplementary Note 7: Monte Carlo–molecular dynamics

To explicitly sample both vibrational and chemical configurational degrees of freedom in the mixed system, we performed MCMD simulations (18–20), in which MD steps are interleaved with Monte Carlo (MC) trial moves. After every 100 fs of MD, N_{X} MC trial moves were attempted, where N_{X} denotes the number of X-site atoms in the simulation cell. All MD segments employed a timestep of 1 fs.

Configurational sampling was carried out using flip trial moves in the semi-grand canonical (SGC) ensemble, in which the chemical species of a randomly selected X-site atom is changed from S to Se or vice versa. For a binary sublattice, we define the concentration of species A as $x_{\mathrm{A}} = N_{\mathrm{A}}/N$, where N_{A} is the number of A atoms on the X-site sublattice and N is the total number of X-site atoms. Within the canonical ensemble, the chemical potential is given by

$$\mu_{\mathrm{A}} = \left(\frac{\partial G}{\partial N_{\mathrm{A}}} \right)_{T,p,N_{i\neq \mathrm{A}}}, \tag{4}$$

which can equivalently be expressed in terms of the free energy per X-site atom, $g = F/N$, as

$$\mu_{\mathrm{A}} = \frac{\partial g}{\partial x_{\mathrm{A}}}. \tag{5}$$

The concentration-dependent free energy $f(x_{\mathrm{A}}, T)$ can therefore be obtained by thermodynamic integration along the composition axis,

$$f(x_{\mathrm{A}}, T) = f(0, T) + \int_0^{x_{\mathrm{A}}} \frac{\partial g}{\partial x'_{\mathrm{A}}} \mathrm{d}x'_{\mathrm{A}}, \tag{6}$$

where the relation between μ_A and x_A was sampled via the MCMD simulations.

Supplementary Note 8: Free energy and phase diagram construction

For each structural phase, the Helmholtz free energy $F(x,T)$ is determined as described above, where here and in the remainder of this paper x denotes the S concentration on the X-site sublattice, while the Ba and Zr sublattices remain fixed. We define the mixing free energy of a given phase as

$$\Delta G_{\mathrm{mix}}(x,T) = F(x,T) - \left[xG_\delta(1,T) + (1-x)G_\delta(0,T)\right], \tag{7}$$

where $G_\delta(0,T)$ and $G_\delta(1,T)$ are the free energies of the pure Se and pure S needle-like δ phases, respectively. The needle-like phase thus serves as a common reference state for all compositions.

Part of the temperature–composition phase diagram is constructed by mapping the convex hull of the free energy as a function of composition at each temperature. Phase coexistence regions are identified from common-tangent constructions, while single-phase stability corresponds to compositions lying on the convex envelope.

In addition to compositional phase separation, the perovskite phase exhibits a continuous tetragonal-to-cubic ($Pm\bar{3}m$) transition. Because this second-order transition does not involve a free-energy crossing between distinct phases, its location cannot be determined from convex-hull analysis. Instead, it is identified from cooling simulations initiated in the cubic phase and monitoring changes in structural and thermodynamic observables, following the procedure of Refs. 7, 21.

Supplementary Note 9: Order parameters

To distinguish the cubic, tetragonal, and orthorhombic perovskite polymorphs observed in the MD and MCMD simulations, we project the instantaneous atomic displacements onto normal modes obtained for the ideal cubic structure corresponding to octahedral tilting distortions (21). Specifically, we evaluate the projection onto the lowest-frequency phonon modes at the M and R points of the Brillouin zone, which correspond to in-phase and out-of-phase octahedral tilts, respectively. The relative amplitudes of these mode projections allow us to identify the tilt patterns characteristic of the cubic, tetragonal, and orthorhombic phases, as discussed in detail in Ref. 7.

Chemical short-range order (SRO) on the X-site sublattice is quantified using the Warren–Cowley parameter (22). For an atom i of type A on the X sublattice, the first-shell parameter is defined as

$$\alpha_i = 1 - \frac{Z_{\mathrm{B}}}{Z_{\mathrm{tot}} c_{\mathrm{B}}}, \tag{8}$$

where Z_{B} is the number of first-nearest-neighbor B atoms on the X sublattice, Z_{tot} is the total number of first-nearest neighbors, and c_{B} is the overall concentration of species B. The SRO parameter is averaged over all X-site atoms and over equilibrated snapshots of the simulation trajectory. Only first-nearest neighbors on the X sublattice are considered. For a random distribution, $\alpha = 0$; $\alpha > 0$ indicates a tendency toward clustering or phase separation, while $\alpha < 0$ signifies preferential unlike-neighbor bonding and chemical ordering.

Supplementary Note 10: Film growth

Thin films were grown in a custom-built chalcogenide molecular beam epitaxy (MBE) system (Mantis Deposition M500). Detailed deposition procedures are described in Refs. 23, 24. Films were deposited on 10 mm × 10 mm × 0.5 mm $(001)_{\mathrm{PC}}$-oriented $LaAlO_3$ substrates (MTI Corporation). Here, PC denotes the pseudocubic setting, and the $(001)_{\mathrm{PC}}$ orientation corresponds to the (012) family of reflections in the rhombohedral basis.

Epitaxial $BaZrS_3$ films were first grown and subsequently subjected to selenizing anneals for anion exchange. Prior to growth, substrates were outgassed in the MBE chamber at 1000 °C under flowing H_2S. $BaZrS_3$ films were deposited at 900 °C to 1000 °C using elemental Ba and Zr sources under a 0.8 SCCM flow of H_2S. Alloying with selenium was achieved by annealing the epitaxial $BaZrS_3$ films at 800 °C for 60 min under a mixed chalcogen flow of 0.2 SCCM H_2S and 0.2 SCCM H_2Se.

Cross-sectional compositions were quantified by scanning transmission electron microscopy (STEM) energy-dispersive X-ray spectroscopy. Elemental mapping confirms a uniform distribution of Ba, Zr, S, and Se across the film thickness, yielding an overall composition of $BaZrS_{3x}Se_{3-3x}$ with $x = 0.27$ (23).

Supplementary Note 11: Scanning transmission electron microscopy

Samples for STEM were prepared by non-aqueous mechanical polishing followed by thinning to electron transparency using single-sector argon-ion milling (Fischione 1051 TEM Mill) (25). Atomic-resolution high-angle

annular dark-field (HAADF) image series were acquired on a probe aberration-corrected Thermo Fisher Scientific Themis Z S/TEM operated at 200 kV. A probe semi-convergence angle of 18.9 mrad, a detector collection range of 65 mrad to 200 mrad, and a dwell time of 1 µs were used. Image series were aligned and corrected for distortions arising from linear drift using the revolving STEM method (26).

Supplementary Note 12: Scanning transmission electron microscopy image analysis

STEM image simulations were performed using a custom Python implementation of the multislice method (27, 28). For both simulated and experimental HAADF images, peak intensities of anion atomic columns were extracted by fitting two-dimensional Gaussian functions to the projected columns.

To quantify spatial ordering and directly compare simulations with experiment, Moran's I spatial autocorrelation parameter was evaluated for the anion peak intensities (29). For a given feature A (here, the anion column intensity) over N atomic columns, the autocorrelation metric is defined as

$$M^I = \frac{N}{W} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}(A_i - \bar{A})(A_j - \bar{A})}{\sum_{i=1}^{N}(A_i - \bar{A})^2}, \tag{9}$$

with

$$W = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}, \tag{10}$$

$\bar{A}$ is the mean value of A, and w_{ij} is a weighting factor defining the neighbor relationship between atomic columns i and j. In this work, neighbors are defined along the projected pseudocubic $\langle 110 \rangle$ directions. This metric provides a quantitative measure of spatial correlations in anion intensity variations associated with chemical ordering.

Supplementary Figures

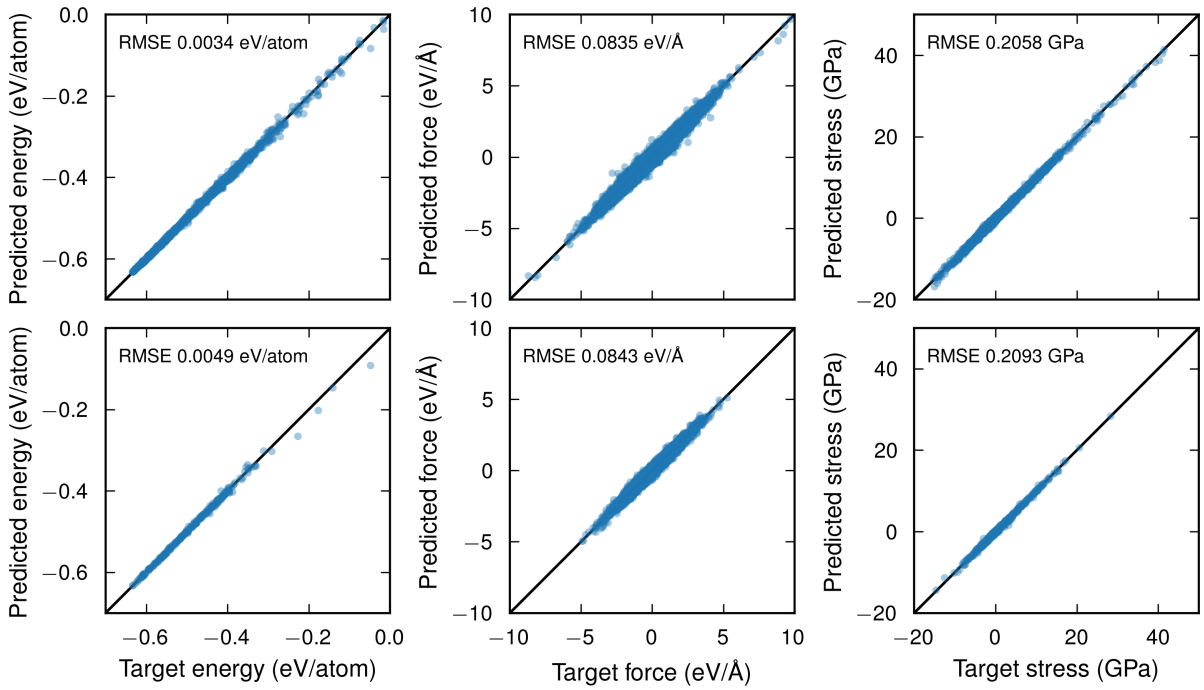

Figure S1: Parity plots for the MLIP compared to the reference DFT data. The top row corresponds to the training data the and botto tom row the test data.

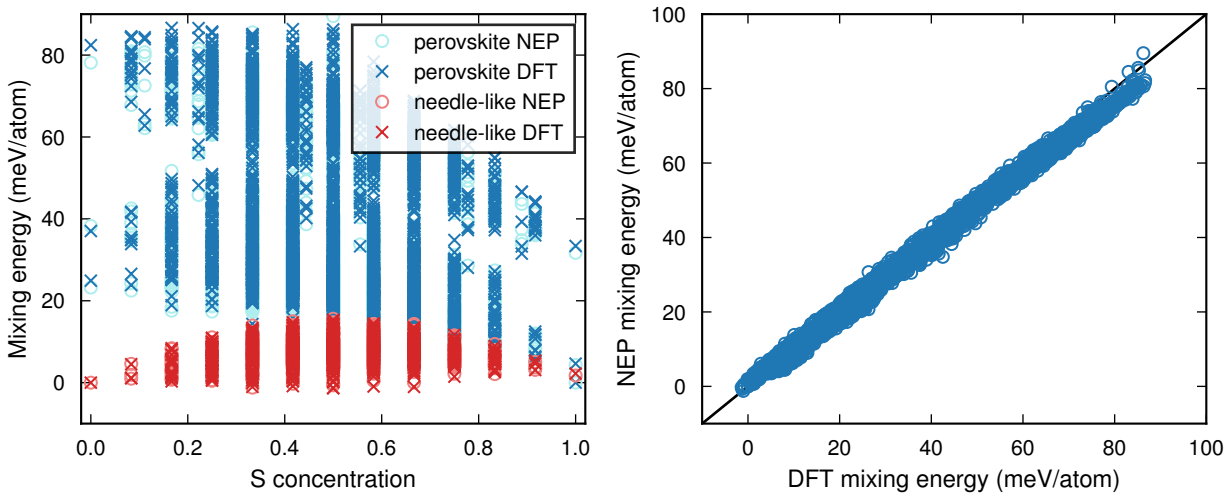

Figure S2: Mixing energies from MLIP and DFT. Left: Mixing energies based on enumerated structures for the perovskite and needle (δ) phase calculated with DFT and NEP (left). Right: Parity plot for the mixing energies.

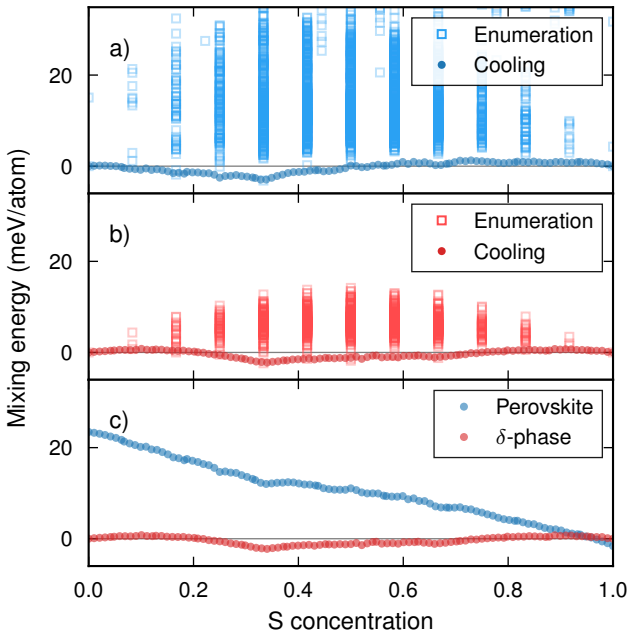

Figure S3: Mixing energy of the (a) perovskite and (b) needle-like (δ) phase as a function of composition. Structures labeled "enumeration" were obtained from systematic combinatorial sampling followed by relaxation, whereas structures labeled "cooling" were obtained from MCMD cooling simulations. The mixing energies are referenced to respective end members of the same structure to facilitate comparison within the respective structure family. (c) Comparison of mixing energies of the perovskite (γ) and needle-like (δ) phases obtained from cooling simulations using a common energy reference.

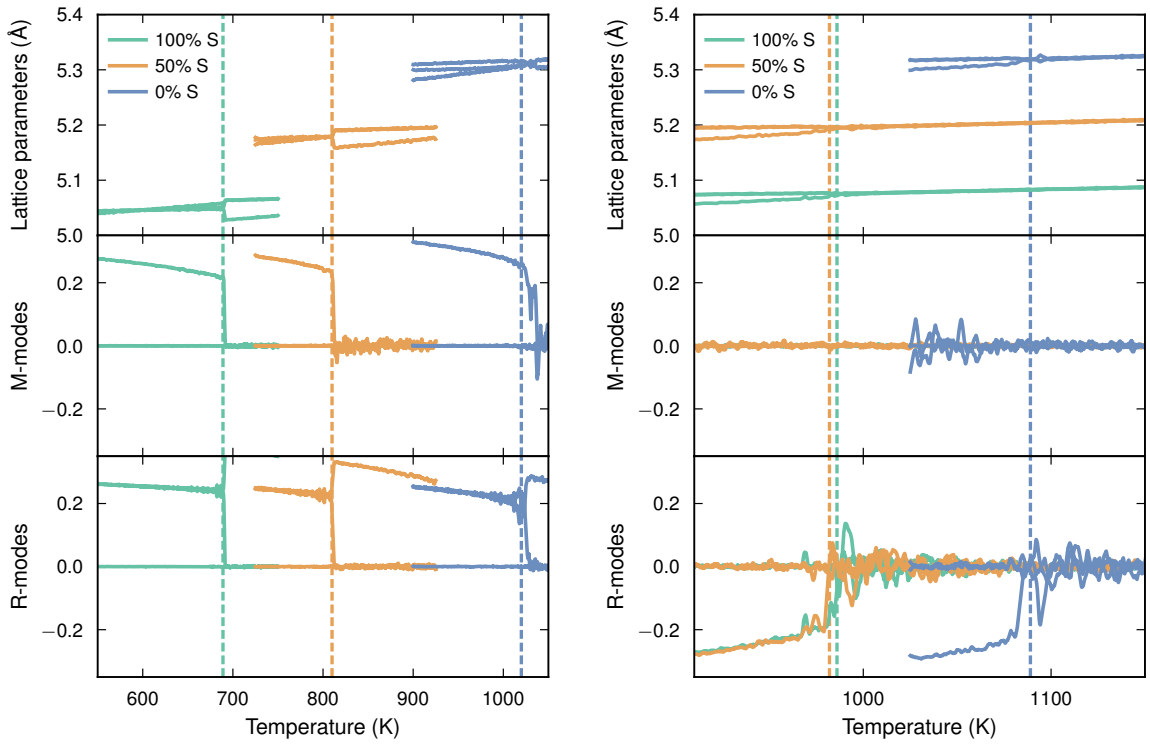

Figure S4: Phase transitions observed in MCMD simulations. Left panels shows the observation of orthorhombic to tetragonal transition (transition temperature indicated by dashed vertical lines), and right panels shows the tetragonal to cubic transition.

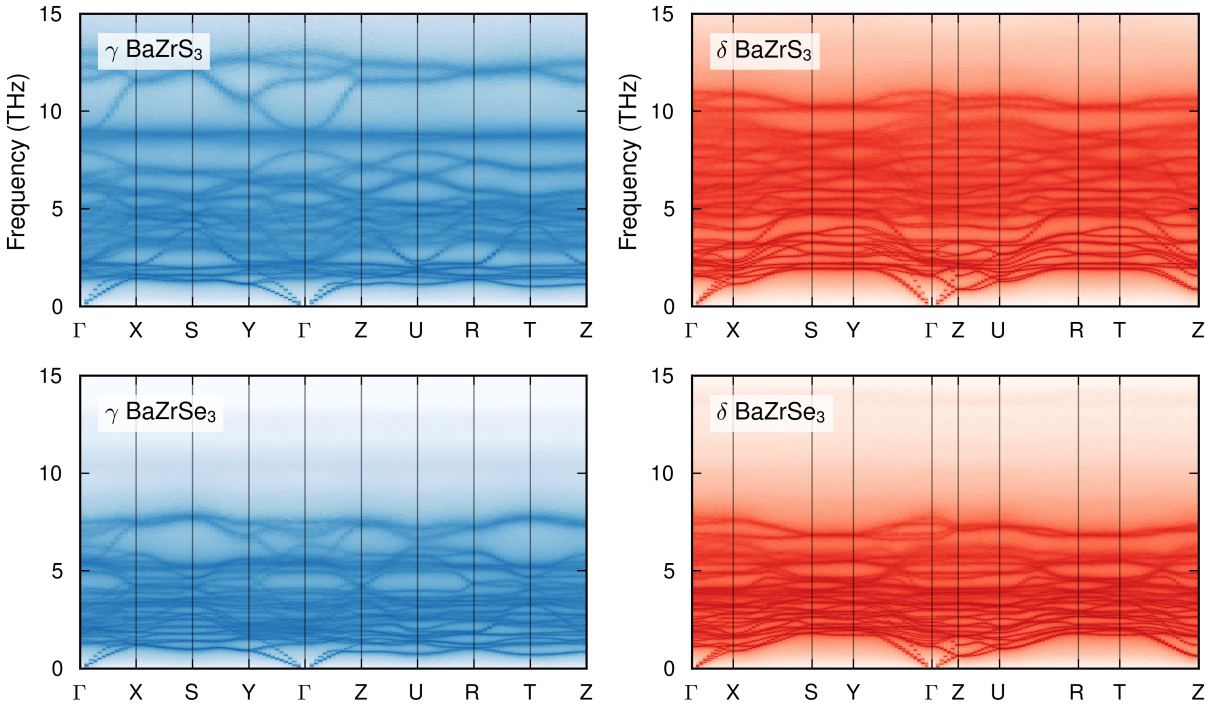

Figure S5: Phonon dispersions for δ and γ phases of $BaZrS_3$ and $BaZrSe_3$ at 300 K obtained from the spectral energy density (30) computed with DYNASOR (31, 32).

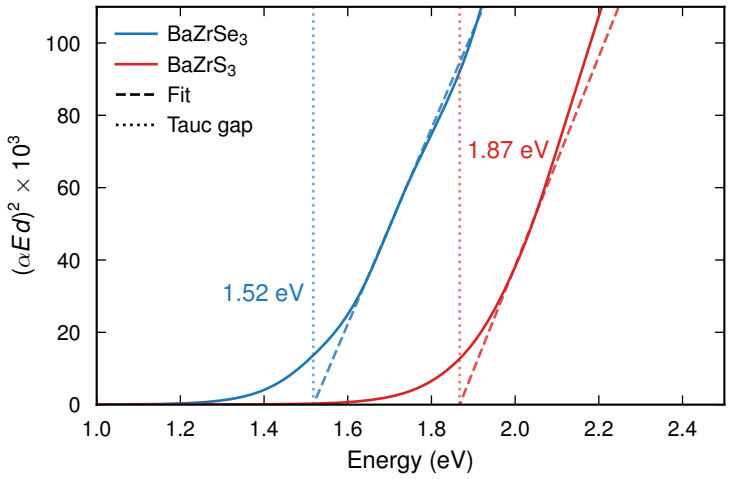

Figure S6: Fitting of the Tauc gap for pure $BaZrS_3$ and $BaZrSe_3$ at 300 K.

Supplementary Tables

Table S1: Comparison of the relaxed energy differences (0 K) for the relevant pure phases in units of meV/atom. The octahedral tilt patterns of the perovskite phases are described using Glazer notation.

		$BaZrSe_3$		$BaZrS_3$	
		NEP	DFT	NEP	DFT
Non-perovskite phases					
δ $Pnma$		0.0	0.0	0.0	0.0
$P6_3/mmc$		32.8	32.5	74.2	71.3
Perovskite phases					
γ $Pnma$	$a^-a^-c^+$	23.4	24.4	−1.6	−2.3
β $I4/mcm$	$a^0a^0c^-$	38.5	37.0	2.7	2.4
α $Pm\bar{3}m$	$a^0a^0a^0$	78.5	82.5	29.7	31.3

Table S2: Band gaps and Tauc gaps for $BaZrSe_3$ and $BaZrS_3$ in different phases calculated with HSE06 $\alpha = 0.25$. The values for the relaxed structures are calculated both without and with spin-orbit coupling (values given in parenthesis).

	Fundamental band gap (eV) relaxed	Fundamental band gap (eV) MD 300 K	Tauc gap (eV) MD 300 K
$BaZrSe_3$ perovskite	1.35 (1.18)	1.35	1.64
$BaZrS_3$ perovskite	1.81 (1.76)	1.66	1.91
$BaZrSe_3$ needle	0.98 (0.90)	1.04	
$BaZrS_3$ needle	1.54 (1.53)	1.57	

Supplementary References

[1] Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, A. J. Gabourie, J. Liu, H. Dong, J. Wu, Y. Chen, Z. Zhong, J. Sun, P. Erhart, Y. Su, and T. Ala-Nissila, *GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations,* The Journal of Chemical Physics **157**, 114801 (2022). doi:10.1063/5.0106617.

[2] K. Xu, H. Bu, S. Pan, E. Lindgren, Y. Wu, Y. Wang, J. Liu, K. Song, B. Xu, Y. Li, T. Hainer, L. Svensson, J. Wiktor, R. Zhao, H. Huang, C. Qian, S. Zhang, Z. Zeng, B. Zhang, B. Tang, Y. Xiao, Z. Yan, J. Shi, Z. Liang, J. Wang, T. Liang, S. Cao, Y. Wang, P. Ying, N. Xu, C. Chen, Y. Zhang, Z. Chen, X. Wu, W. Jiang, E. Berger, Y. Li, S. Chen, A. J. Gabourie, H. Dong, S. Xiong, N. Wei, Y. Chen, J. Xu, F. Ding, Z. Sun, T. Ala-Nissila, A. Harju, J. Zheng, P. Guan, P. Erhart, J. Sun, W. Ouyang, Y. Su, and Z. Fan, *GPUMD 4.0: A high-performance molecular dynamics package for versatile materials simulations with machine-learned potentials,* Materials Genome Engineering Advances **3**, e70028 (2025). doi:10.1002/mgea.70028.

[3] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, *The atomic simulation environment—a Python library for working with atoms,* Journal of Physics: Condensed Matter **29**, 273002 (2017). doi:10.1088/1361-648X/aa680e.

[4] F. Eriksson, E. Fransson, and P. Erhart, *The Hiphive Package for the Extraction of High-Order Force Constants by Machine Learning,* Advanced Theory and Simulations **2**, 1800184 (2019). doi:10.1002/adts.201800184.

[5] M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod, and P. Erhart, *ICET – A Python Library for Constructing and Sampling Alloy Cluster Expansions,* Advanced Theory and Simulations **2**, 1900015 (2019). doi:https://doi.org/10.1002/adts.201900015.

[6] G. L. W. Hart and R. W. Forcade, *Algorithm for generating derivative structures,* Physical Review B **77**, 224115 (2008). doi:10.1103/PhysRevB.77.224115.

[7] P. Kayastha, E. Fransson, P. Erhart, and L. Whalley, *Octahedral Tilt-Driven Phase Transitions in $BaZrS_3$ Chalcogenide Perovskite,* The Journal of Physical Chemistry Letters **16**, 2064 (2025). doi:10.1021/acs.jpclett.4c03517.

[8] E. Lindgren, M. Rahm, E. Fransson, F. Eriksson, N. Österbacka, Z. Fan, and P. Erhart, *calorine: A Python package for constructing and sampling neuroevolution potential models,* Journal of Open Source Software **9**, 6264 (2024). doi:10.21105/joss.06264.

[9] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Ab initio molecular simulations with numeric atom-centered orbitals,* Computer Physics Communications **180**, 2175 (2009). doi:10.1016/j.cpc.2009.06.022.

[10] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, *Influence of the exchange screening parameter on the performance of screened hybrid functionals,* J. Chem. Phys. **125**, 224106 (2006). doi:10.1063/1.2404663.

[11] G. Kresse and J. Furthmüller, *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set,* Computational Materials Science **6**, 15 (1996). doi:10.1016/0927-0256(96)00008-0.

[12] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple,* Physical review letters **77**, 3865 (1996).

[13] P. E. Blöchl, *Projector augmented-wave method,* Physical Review B **50**, 17953 (1994). doi:10.1103/PhysRevB.50.17953.

[14] G. Kresse and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method,* Physical Review B **59**, 1758 (1999). doi:10.1103/PhysRevB.59.1758.

[15] M. Moustafa, A. Ghafari, A. Paulheim, C. Janowitz, and R. Manzke, *Spin orbit splitting in the valence bands of ZrSxSe2-x: Angle resolved photoemission and density functional theory,* Journal of Electron Spectroscopy and Related Phenomena **189**, 35 (2013). 12th International Conference on Electronic Spectroscopy and Structure (ICESS-12). doi:10.1016/j.elspec.2012.12.010.

[16] R. Freitas, M. Asta, and M. de Koning, *Nonequilibrium Free-Energy Calculation of Solids Using LAMMPS,* Computational Materials Science **112**, 333 (2016). doi:10.1016/j.commatsci.2015.10.050.

[17] J. G. Kirkwood, *Statistical Mechanics of Fluid Mixtures,* The Journal of Chemical Physics **3**, 300 (1935). doi:10.1063/1.1749657.

[18] N. L. Allan, G. D. Barrera, M. Y. Lavrentiev, I. T. Todorov, and J. A. Purton, *Ab initio calculation of phase diagrams of ceramics and minerals,* Journal of Materials Chemistry **11**, 63–68 (2001). doi:10.1039/b002951n.

[19] M. Widom, W. P. Huhn, S. Maiti, and W. Steurer, *Hybrid Monte Carlo/Molecular Dynamics Simulation of a Refractory Metal High Entropy Alloy,* Metallurgical and Materials Transactions A **45**, 196–200 (2013). doi:10.1007/s11661-013-2000-8.

[20] J. M. Rahm and P. Erhart, *Understanding Chemical Ordering in Bimetallic Nanoparticles from Atomic-Scale Simulations: The Competition between Bulk, Surface, and Strain,* The Journal of Physical Chemistry C **122**, 28439–28445 (2018). doi:10.1021/acs.jpcc.8b10874.

[21] E. Fransson, P. Rosander, F. Eriksson, J. M. Rahm, T. Tadano, and P. Erhart, *Limits of the phonon quasiparticle picture at the cubic-to-tetragonal phase transition in halide perovskites,* Communications Physics **6**, 173 (2023). doi:10.1038/s42005-023-01297-8.

[22] J. M. Cowley, *An Approximate Theory of Order in Alloys,* Physical Review **77**, 669 (1950). doi:10.1103/PhysRev.77.669.

[23] K. Ye, I. Sadeghi, M. Xu, J. Van Sambeek, T. Cai, J. Dong, R. Kothari, J. M. LeBeau, and R. Jaramillo, *A Processing Route to Chalcogenide Perovskites Alloys with Tunable Band Gap via Anion Exchange,* Advanced Functional Materials **34**, 2405135 (2024). doi:10.1002/adfm.202405135.

[24] I. Sadeghi, K. Ye, M. Xu, Y. Li, J. M. LeBeau, and R. Jaramillo, *Making $BaZrS_3$ Chalcogenide Perovskite Thin Films by Molecular Beam Epitaxy,* Advanced Functional Materials **31**, 2105563 (2021). doi:10.1002/adfm.202105563.

[25] L. Dieterle, B. Butz, and E. Müller, *Optimized Ar^+-ion milling procedure for TEM cross-section sample preparation,* Ultramicroscopy **111**, 1636 (2011). doi:10.1016/j.ultramic.2011.08.014.

[26] X. Sang and J. M. LeBeau, *Revolving scanning transmission electron microscopy: correcting sample drift distortion without prior knowledge,* Ultramicroscopy **138**, 28 (2014). doi:10.1016/j.ultramic.2013.12.004.

[27] J. M. Cowley and A. F. Moodie, *The scattering of electrons by atoms and crystals. I. A new theoretical approach,* Acta Crystallographica **10**, 609 (1957). doi:10.1107/S0365110X57002194.

[28] E. J. Kirkland, *Advanced computing in electron microscopy* (New York, NY: Springer, 2010). doi:10.1007/978-1-4419-6533-2.

[29] M. Xu, A. Kumar, and J. M. LeBeau, *Correlating local chemical and structural order using Geographic Information Systems-based spatial statistics,* Ultramicroscopy **243**, 113642 (2023). doi:10.1016/j.ultramic.2022.113642.

[30] J. A. Thomas, J. E. Turney, R. M. Iutzi, C. H. Amon, and A. J. H. McGaughey, *Predicting phonon dispersion relations and lifetimes from the spectral energy density,* Physical Review B **81**, 081411 (2010). doi:10.1103/PhysRevB.81.081411.

[31] E. Fransson, M. Slabanja, P. Erhart, and G. Wahnström, *dynasor—A Tool for Extracting Dynamical Structure Factors and Current Correlation Functions from Molecular Dynamics Simulations,* Advanced Theory and Simulations **4**, 2000240 (2021). doi:https://doi.org/10.1002/adts.202000240.

[32] E. Berger, E. Fransson, F. Eriksson, E. Lindgren, G. Wahnström, T. H. Rod, and P. Erhart, *Dynasor 2: From simulation to experiment through correlation functions,* Computer Physics Communications **316**, 109759 (2025). doi:https://doi.org/10.1016/j.cpc.2025.109759.