

ARTICLE OPEN



Anion order in oxysulfide perovskites: origins and implications

Ghanshyam Pilania¹✉, Ayana Ghosh², Steven T. Hartman³, Rohan Mishra³, Christopher R. Stanek¹ and Blas P. Uberuaga¹

Heteroanionic oxysulfide perovskite compounds represent an emerging class of new materials allowing for a wide range of tunability in the electronic structure that could lead to a diverse spectrum of novel and improved functionalities. Unlike cation ordered double perovskites—where the origins and design rules of various experimentally observed cation orderings are well known and understood—anion ordering in heteroanionic perovskites remains a largely uncharted territory. In this contribution, we present and discuss insights that have emerged from our first-principles-based electronic structure analysis of a prototypical anion-ordered $\text{SrHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ oxysulfide chemistry, studied in all possible anion configurations allowed within a finite size supercell. We demonstrate that the preferred anion ordering is always an all-*cis* arrangement of anions around an HfO_3S_3 octahedron. As a general finding beyond the specific chemistry, the origins of this ordering tendency are traced back to a combined stabilization effect stemming from electronic, elastic, and electrostatic contributions. These qualitative notions are also quantified using state-of-the-art machine learning models. We further study the relative stability of the identified ordering as a function of A (Ca, Sr, Ba) and B (Ti, Zr, Hf) site chemistries and probe chemistry-dependent trends in the electronic structure and functionality of the material. Most remarkably, we find that the identified ground-state anion ordering breaks the inversion symmetry to create a family of oxysulfide ferroelectrics with a macroscopic polarization $>30 \mu\text{C}/\text{cm}^2$, exhibiting a significant promise for electronic materials applications.

npj Computational Materials (2020)6:71; <https://doi.org/10.1038/s41524-020-0338-1>

INTRODUCTION

Perovskite compounds—with a generic formula ABX_3 —represent a versatile class of materials exhibiting tremendous synthetic flexibility. For instance, $>90\%$ of the metal ions in the Periodic Table are known to be stable in a perovskite-type structure^{1,2}. Moreover, the anion sublattice X can be populated by several members of halide (X=F, Cl, Br, and I), chalcogenide (X=O, S, Se) or pnictide (X=N) families. A further possibility of partial substitutions at A, B, and X sites accompanied with a rich set of possible cation- and/or anion-orderings make this class a truly staggering collection of compounds. This diverse perovskite composition and ordering naturally leads to a rich portfolio of chemical, electronic and magnetic properties for applications ranging from catalysis to electronic devices and from sensors to energy storage^{3–6}.

Mixed-anion compounds (i.e., containing more than one anion species) add new dimensions to control and tune stability as well as different aspects of functionality, owing to the different elemental characteristics of the anions, such as charge, ionic radii, electronegativity, and polarizability, as well as their relative arrangement in the crystal lattice⁷. Although mixed cation perovskites have been widely studied^{8–13}, mixing on the anion sublattice in perovskites remains largely unexplored, with the exception of a few recent studies^{14–22}.

Little is yet known about the structure and properties of perovskite oxysulfides, despite the potential advantages of mixing these two anions. Although metal oxides are a mainstay of many devices owing to their stability and good dielectric properties, the increased covalency of metal-sulfide bonds offers the chance to narrow the bandgap and improve electronic transport properties. Indeed, sulfide perovskites are under active investigation as environment-friendly solar energy harvesters. Recent studies focused on chalcogenides have shown that ABX_3 compounds

within this class—with A=Ca, Sr, or Ba; B=Ti, Zr, or Hf; X=S or Se—are stable in a distorted perovskite structure with the bandgaps ranging from 0.3 eV to 2.3 eV^{23–26}. Many of these semiconducting compounds possess direct bandgaps with high optical absorption coefficients—comparable to those of traditional optoelectronic semiconductors such as GaAs—with large valence and conduction band dispersions, suggesting good carrier mobility²³. For some ABX_3 sulfides, competing non-perovskite phases have also been reported^{24,27–30}. More recently, a wide range of ABS_3 perovskites and non-perovskites were screened for relevant properties, such as thermodynamic stability, light absorption, charge mobility, and defect tolerance. The study identified LaYS_3 as a promising candidate³¹. As a first step in the exploration of perovskite oxysulfides, Perera et al¹⁸ recently synthesized a series of $\text{BaZr}(\text{O}_{1-x}\text{S}_x)_3$ oxysulfides over the wide range of $x \in [0, 0.54]$ to show that all the compounds within this composition range are stable in a perovskite crystal structure with a systematic bandgap variation of >1 eV.

In general, the properties of perovskites are sensitive to small structural changes, such as rotations, distortions, and tilting of the metal-anion octahedra. In the case of oxysulfide perovskites (or more broadly in heteroanionic perovskites), the B-site octahedral coordination is controlled by two or more anions. Different anion configurations around a given metal ion could, in principle, lead to different levels of crystal field strength, bonding characteristics, and local polarization, which in turn dictate the ground-state structure and functionalities. In spite of this potential for rich structure–property relationships, investigations of anion order in mixed chalcogenide perovskites are missing.

Here, we report a first-principles investigation of anion order in $\text{AB}(\text{O}_{0.5}\text{S}_{0.5})_3$ perovskites, with A=Ca, Sr, Ba and B=Ti, Zr, Hf. We first analyze all 246 symmetry-unique configurations accessible within a 20-atom supercell of $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$, and distill down

¹Materials Science and Technology Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. ²Materials Science & Engineering, University of Connecticut, Storrs, CT 06269, USA. ³Institute of Materials Science & Engineering and Department of Mechanical Engineering & Materials Science, Washington University in St. Louis, St. Louis, MO 63130, USA. ✉email: gpilania@lanl.gov

various factors that control their relative energetics and bandgap variations. Specifically, we find that a *cis*-arrangement of O and S around the BX_6 octahedron is highly preferred over a *trans*-arrangement. We show that a combination of electronic, local strain, and electrostatic effects largely dictates the anion ordering in this class of compounds. The bandgap is controlled by the crystal field splitting imposed by the surrounding anion in the BX_6 octahedron, which primarily changes the position of the Hf *d*-states. We then generalize this insight by exploring stability trends for the larger set of $AB(O_{0.5}S_{0.5})_3$ oxysulfides with $A \in \{\text{Ca, Sr, Ba}\}$ and $B \in \{\text{Ti, Zr, Hf}\}$ to demonstrate that the trends can be rationalized using geometrical notions and ionic radii of the constituent ions. In particular, our computations predict that zirconates of Ca and Sr are thermodynamically stable and, therefore, should be amenable to experimental synthesis. At last, we demonstrate that the asymmetric and stable *cis*-arrangement of O and S anions in the perovskite structures can support ferroelectricity.

RESULTS

Bulk CaHfO_3 , CaHfS_3 , and $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ chemistries

Both CaHfO_3 and CaHfS_3 are known to be stable in a GdFeO_3 -type crystal structure with a $Pnma$ space group and can be represented within a 20-atom orthorhombic supercell. Using density-functional theory (DFT) calculations, we computed the lattice parameters and bandgaps for these two chemistries, and we compare them with past studies from the literature in Table 1. Going to equimolar mixed-anion $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ chemistry, we simply average the experimentally reported lattice parameters for CaHfO_3 and CaHfS_3 in the same space group as a starting point for further geometry optimization using DFT computations. Our theoretical investigation starts with an explicit enumeration of all possible configurations spanned by different relative arrangements of S and O atoms within a 20-atom supercell of $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$. As both the oxide and sulfide parent perovskite chemistries stabilize in a $Pnma$ space group, it is fair to assume that this size of supercell is enough to enumerate all low energy anion orderings and octahedral distortion patterns possible for the equimolar mixed oxysulfide chemistry and going to larger cells will not reveal any new orderings. To enumerate the symmetry-unique configurations exhaustively for the target composition, we rely on an algorithm for generating derivative structures^{32,33}, using the implementation made available in `enumlib`³⁴. Enumeration of all possible

symmetry-unique configurations for the equimolar mixed-anion composition $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ within this supercell leads to a total of 246 possible candidate structures. Further details of our DFT computations are provided in the Methods section towards the end of the manuscript.

Figure 1 captures the DFT-computed relative energetics and Kohn-Sham bandgaps for the entire set of configurations. The relative energy is defined by the enthalpy of mixing with respect to the two end-point compositions (i.e., CaHfO_3 and CaHfS_3 perovskites) in units of eV per $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ 5-atom formula unit. It is interesting to note that the different anion orderings have a remarkable effect on the stability and bandgap. For a fixed composition, just by changing the relative arrangement of S and O atoms on the anion sublattice results in a large variation in the mixing enthalpies (17.6–1.135 eV) and bandgaps (1.52–2.75 eV). Note that, for this particular chemistry, it turns out that all of the mixing energies are positive, such that it is thermodynamically favorable (at least at 0 K, i.e., considering enthalpic contributions alone) for the oxysulfide to phase separate into the respective oxide and sulfide components. We later show that by systematically changing A and B-site chemistries, this mixing enthalpy can be made negative. However, first we are interested in understanding configuration-dependent relative trends across the $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ data set. With this aim, natural next questions are: “What different factors are responsible for the observed trends and variations in the energetics and electronic structure?” and “Is it possible to develop a more intuitive understanding of the same at a microscopic level?”

Relative stability of anion-ordered configurations

In the quest of addressing the above questions, we first focus on the mixing enthalpies. A visualization of the relaxed geometries for the entire set of 246 configurations reveals that the most-stable arrangement is a *fac*-type HfO_3S_3 octahedron, presented by configurations shown in Fig. 1a, b (i.e., all S–Hf–S and O–Hf–O bonds are *cis* with a 90° interaction), followed by a *mer*-type HfO_3S_3 octahedron (each anion type forms a “T” shape having two *cis* interactions, and one *trans*). The least-stable configuration is mixed $\text{HfO}_4\text{S}_2/\text{HfO}_2\text{S}_4$ octahedra with the maximum number of *trans* interactions (*cf.*, configurations shown in Fig. 1c, d). A deeper analysis reveals that there are at least three different factors—namely, electronic, elastic strain, and electrostatic effects—that are responsible for energetically favoring the *cis*-configurations and penalizing the *trans*-configurations.

Electronic effects are known to play an important role in stabilizing a *cis*-configuration for strongly bonded ligands in octahedral complexes containing high valence d^0 transition metal ions. For instance, $\text{Mo}(\text{NR})_2^{2+}$ (with $R = \text{alkyl or aryl group}$) or MoO_2^{2+} complexes with a *cis*-configuration of ligands tend to maximize the metal(d_n)–ligand(p_n) covalency^{35,36}. In addition to octahedral complexes, covalency effects are known to favor the formation of *cis*-octahedra over the *trans*-arrangement in a number of oxynitride perovskites^{20,37}. Similar effects also play a role in stabilization of configurations with a *cis*-arrangement of anions around the transition metal ion octahedra in oxysulfides. In fact, a partial density-of-states (DOS) analysis of representative *cis*- versus *trans*-configurations in the BX_6 octahedra (shown in Fig. 2a) results in very different valence band DOS for the *p*-states of O and S—with O *p*-states pushed significantly lower in the *cis*-configuration as compared with the *trans*-configuration, as depicted in Fig. 2b—confirming that the electronic effects indeed have an important role in the relative stabilization of the anion *cis*-configurations.

The role of elastic strain can be understood on the basis of simple and intuitive arguments by considering different arrangements of metal-anion-metal chains reached by the *cis*- and *trans*-anion configurations. As graphically illustrated in Fig. 2c, whereas a *cis*-configuration of the O and S ions in an octahedral

Table 1. Lattice parameters and bandgaps of the end-point compositions, CaHfO_3 and CaHfS_3 perovskites, forming the mixed $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ equimolar oxysulfide chemistries studied here.

ABX ₃ chemistry	Space group	Lattice parameters			Bandgap (eV)
		<i>a</i> (Å)	<i>b</i> (Å)	<i>c</i> (Å)	
CaHfS_3	<i>Pnma</i>	7.015	9.582	6.543	1.51^a
	<i>Pnma</i>	6.983 ^b	9.537 ^b	6.518 ^b	2.65 ^c
CaHfO_3	<i>Pnma</i>	5.767	8.036	5.594	4.49^a
	<i>Pnma</i>	5.732 ^d	7.984 ^d	5.568 ^d	6.75 ^e , 6.09 ^f

Entries computed in this work are shown in a bold font. Past experimental results and bandgaps computed at different levels of theory are also presented for comparison.

^aGGA-PBE Bandgaps.

^bRef. 67.

^cRef. 31.

^dRef. 2.

^eExperimental bandgap from ref. 68.

^fBandgap using the hybrid Heyd–Scuseria–Ernzerhof functional (HSE06) from ref. 39.

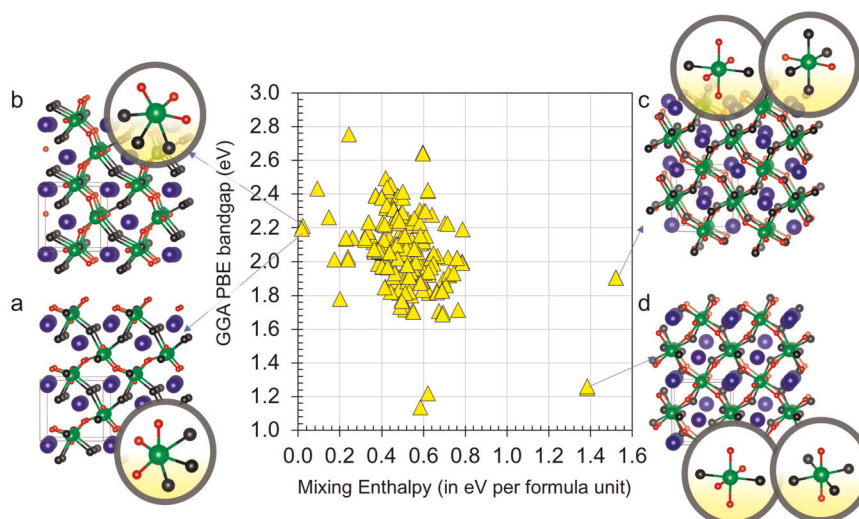


Fig. 1 Anion ordering dependent energetics and bandgaps. Configuration-dependent mixing enthalpies and bandgaps of 246 different anion-ordered $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ configurations possible within a 20-atom $\sqrt{2} \times \sqrt{2} \times 2$ pseudo-cubic perovskite supercell. Representative octahedral configurations of anions around Hf are highlighted in a few selected most-stable **a**, **b** and least-stable **c**, **d** supercell geometries.

environment leads to equi-length zigzag chains of only one type (i.e., $-\text{O}-\text{Hf}-\text{S}-\text{Hf}-\text{O}-$) in all the three directions within a crystal, a *trans*-configuration can result in both $-\text{O}-\text{Hf}-\text{O}-\text{Hf}-\text{O}-$ and $-\text{S}-\text{Hf}-\text{S}-\text{Hf}-\text{S}-$ chains running parallel to each other in a given crystal direction (cf. Fig. 2d). In the latter case, the two different chains will have different "relaxed" lengths, owing to the significantly different ionic radii of the O and S anions (1.40 and 1.84 Å, respectively)³⁸. Therefore, to be able to accommodate these chains along a given lattice vector, the crystal periodicity would require the longer $-\text{S}-\text{Hf}-\text{S}-\text{Hf}-\text{S}-$ chains to be compressed and the shorter $-\text{O}-\text{Hf}-\text{O}-\text{Hf}-\text{O}-$ chains to be stretched. As a result, the *trans*-configurations lead to an elastic strain along one or more directions in the mixed-anion crystal, and the effect eventually leads to the relative stabilization of the *cis*-configurations.

In brief, alluded to above, in oxysulfides the two anionic species—the O^{2-} and S^{2-} ions—have a large size mismatch and therefore the elastic strain effects naturally become important. However, the two anions share the same nominal charge states and, at least for a fixed anion sublattice, rearranging the anions at a given composition cannot alter the net electrostatic energy of the system. An Ewald summation based on the nominal charge states of the constituents for a perfect (unrelaxed) perovskite crystal results in a constant electrostatic energy for the entire set of configurations. This might lead to an erroneous conclusion that, unlike oxynitrides (where the two anion species have different nominal charge states: O^{2-} and N^{3-}), for oxysulfides containing isovalent O^{2-} and S^{2-} anions the electrostatic effects are not important. However, our analysis shows that relative energetics of the different configurations is directly governed by lowering of the overall electrostatic energy during geometry relaxation, as discussed below.

To understand the role played by electrostatics in stabilizing certain anion orderings in a mixed-anion oxysulfide, it is instructive to look at the A- and/or B-site cation ordered perovskites where such effects are well understood^{9,10,12}. Figure 2e–g show local coordination environments of an anion X in different A- and B-site cation ordered configurations of a double perovskite. In each case, the anion is coordinated by two B-site cations and four A-site cations. In case of a B-site rocksalt-ordered perovskite $\text{A}_2[\text{B}^1\text{B}^2]\text{X}_6$, the two different B-site cations are on opposite sides from one another, as shown in Fig. 2e. As a result of the asymmetric local configuration caused by the B-site sublattice,

the anion may easily shift towards the smaller cation and away from the larger cation to not only relieve lattice strains arising from the size mismatch, but also minimize the overall electrostatic energy of the system during the course of relaxation. Figure 2f, g compare local anion configurations for two different A-site orderings, namely rocksalt and layer orderings, in a $[\text{A}^1\text{A}^2]\text{B}_2\text{X}_6$ -type double perovskite. Unlike the B-site ordered case, the rocksalt ordering of the A^1 and A^2 cations creates a symmetric local configuration for the anion whereas the layer ordering breaks this symmetry, allowing for a larger degree of relaxation for the anion. Similar arguments can also be used to rationalize energetics of simultaneous ordering at the A- and B-sites of a $[\text{A}^1\text{A}^2][\text{B}^1\text{B}^2]\text{X}_6$ double perovskite. Indeed, for ordered perovskite systems where electrostatics has a dominant role (i.e., compounds that demonstrate large size and charge disparity with respect to the ordered cation pair at a given cation sublattice), a rocksalt ordering at the B-sublattice is often accompanied by a layer ordering at the A-sublattice.

The aforementioned symmetry arguments can be readily extended to an anion-ordered perovskite. Although *cis*-ordering of the anions leads to a highly asymmetric octahedron, a *trans*-ordered octahedron is symmetric. As a result, an oxysulfide configuration with *cis*-ordering of the O^{2-} and S^{2-} anions around the B-site cations can lower its strain and electrostatic energies more easily than those with the *trans*-ordering of the anions. If the local-symmetry arguments above are valid, relative energetics of mixing enthalpies presented in Fig. 1 must exhibit a strong correlation with the corresponding lowering of the overall electrostatic energy during the course of geometry relaxation. To confirm this, we computed the Madelung energy for each DFT-optimized configuration with respect to a common unrelaxed reference state. Indeed, the results presented in Fig. 2h exhibit a strong correlation between the two quantities, indicating the validity of the arguments discussed above to understand and rationalize the relative energetics of anion ordering in oxysulfide perovskites.

Configuration-dependent trends in the bandgap

After rationalizing the trends in the mixing enthalpy as a function of different anion configurations, next we focus our attention on the bandgap. As can be seen from Fig. 1, there is no obvious correlation between the plotted mixing enthalpies and bandgaps for the different configurations. In fact, for different configurations

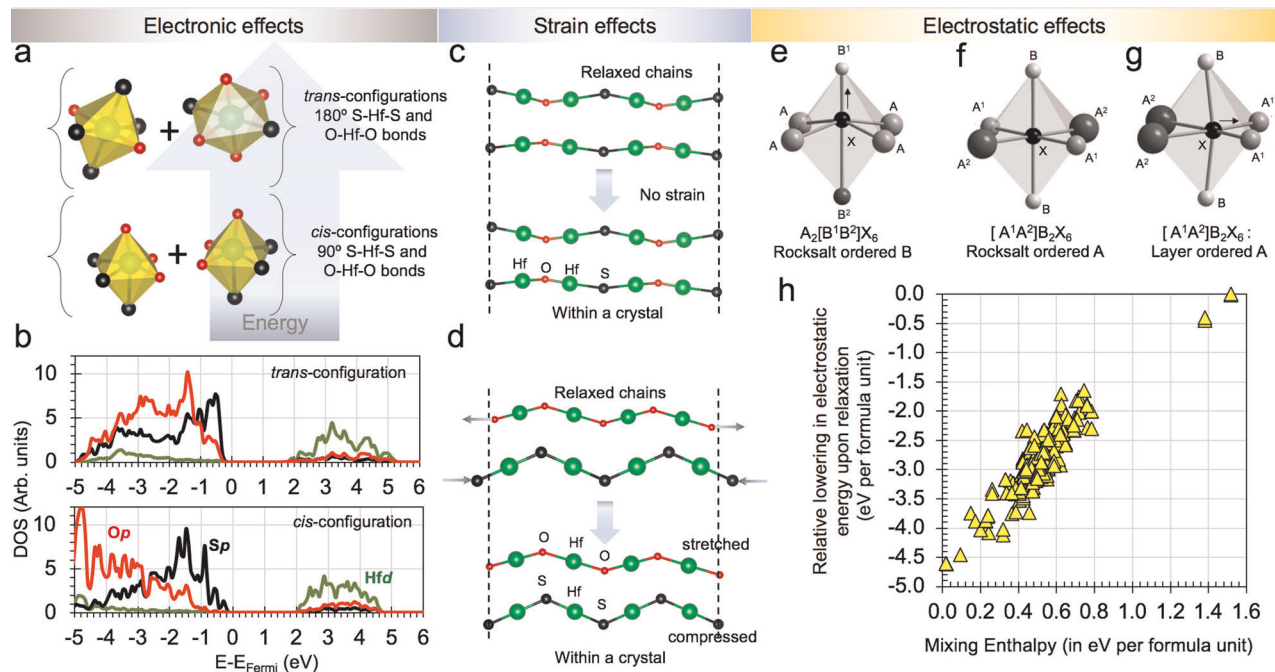


Fig. 2 Factors dictating anion ordering in oxysulfides. A graphical representation of the electronic, elastic strain, and electrostatic effects that stabilize *cis*-ordering while penalizing the *trans*-order in oxysulfide perovskites. **a** A schematic representing electronic stabilization of *cis*-configuration for strongly bonded ligands in octahedral complexes containing d^0 transition metal ions, as compared with the *trans*-configuration. **b** Partial density-of-states (DOS) plots showing that the *cis*-configured octahedra in **a** have oxygen p -states at lower energies. The *cis*- and *trans*-configurations of anions also lead to different anion-orderings in anion-metal-anion chains. For instance, in a *cis*-configured $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ structure **c** all the chains have the same $-\text{S}-\text{Hf}-\text{O}-\text{Hf}-\text{S}-\text{Hf}-\text{O}-$ arrangement, as shown in top panel, which can fit in a crystal without introducing any additional lattice strain (bottom panel). However, the *trans*-configuration shown in **d** has mismatched chain lengths, as O and S have significantly different radii (top panel), which have to be either stretched or compressed to fit in the periodic boundary conditions posed by a crystal lattice (bottom panel). **e-g** Representative anion-site local environments in rocksalt-ordered $\text{A}_2\text{B}^1\text{B}^2\text{X}_6$ - and $\text{A}^1\text{A}^2\text{B}_2\text{X}_6$ -type, and layered-ordered $\text{A}^1\text{A}^2\text{B}_2\text{X}_6$ -type double perovskites, respectively. **h** The mixing enthalpy of the oxysulfide configurations, which combines the effects of electronic effects, strain relaxation and local-symmetry-breaking, is well-correlated with the increase of electrostatic cohesive energy during relaxation.

with a fixed value of mixing enthalpy, the DFT-computed bandgap can vary up to >1.5 eV. To systematically understand the factors responsible for this variation, in Fig. 3a–d, we plot orbital-decomposed DOS for four different configurations, all having a mixing enthalpy value of 0.60 ± 0.02 eV per formula unit, with significantly different bandgaps. In particular, the DOS associated with Ca s -states, Hf d -states and O and S p -states are shown in each panel. In agreement with previous reports in literature^{39,40}, the valence band maximum and conduction band minimum in this class of ABX_3 compounds are primarily formed by anion p -states and B-site transition metal d -states. From the results plotted in Fig. 3, it can be seen that the bandgap variation across different configurations is largely controlled by the extent of lowering of the B-site transition metal ion d -states, relative to the valence band maximum. This lowering for a particular B-site ion is mainly governed by the crystal field imposed by a specific arrangement of different anions surrounding the octahedrally coordinated ion. The overall bandgap is then a collective effect of these local d -band lowerings at the individual sites. Therefore, the bandgap trends in the target chemical space are expected to be determined by not only local arrangements of anion species around the B-site octahedra but also global arrangement of these octahedral motifs with respect to each other in a crystal.

Configurational machine learning

Although the analysis presented thus far provides a physics-based qualitative understanding of relative energetics and bandgap variations as a function of anion configurations, next we focus on learning these trends in a quantitative manner using machine

learning (ML). Further, a ML-based analysis allows us to quantify the impact of local relaxation effects on the two properties, by training the model on either the relaxed or unrelaxed geometries. Regardless of specific aims, materials ML strategies generally consist of two distinct steps. The first step is to come up with a physically meaningful numerical representation for each material in the data set. As a result of this step, each input material gets reduced to a string of numbers (frequently referred to as a fingerprint or descriptor). This is an enormously important step, requiring significant expertise and knowledge of the materials class and the application, i.e., “domain expertise”. The second step establishes a mapping between the fingerprinted input and the target property, and is entirely numerical in nature, largely devoid of the need for domain knowledge.

For the problem at hand, as the relative arrangement of O^{2-} and S^{2-} species in different systems is deemed most relevant, we naturally resort to fingerprinting schemes that explicitly account for atomic configurations. Specifically, we employ (1) bag of bonds (BoB) fingerprint based on a sine Coulomb matrix (SCM) and (2) simulated X-ray diffraction (XRD) powder pattern spectra. The SCM represents a simple extension of the molecular Coulomb matrix to periodic crystals⁴¹. For the XRD spectra based fingerprint, we used a one-dimensional array representing powder diffraction pattern of a crystal structure⁴². A fixed 2θ range (from 0° to 127°) with $\text{Cu K}\alpha$ radiation was consistently used for the entire set of configurations and a Gaussian broadening of 0.05 was used for smoothing the XRD spectra. To map the configurations on to the target properties, namely the mixing enthalpy and the bandgap, we employ kernel ridge regression (KRR) as a statistical learning algorithm^{43–45}. A detailed description of both the

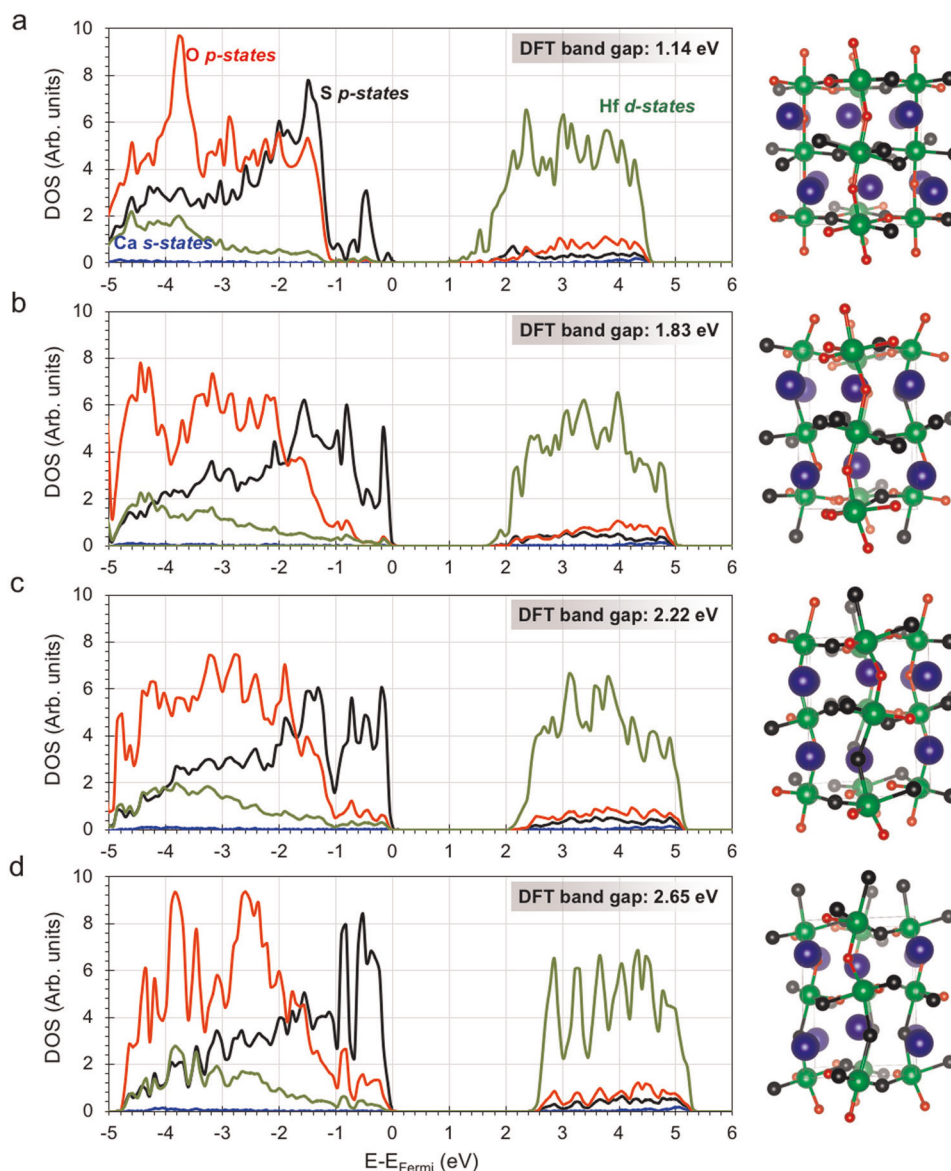


Fig. 3 Anion ordering dependent electronic structure. Orbital-decomposed partial DOS for four different $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ configurations **a–d**, selected to have a similar range of the mixing enthalpies (within 0.60 ± 0.02 eV per formula unit), but with significantly different GGA-PBE bandgaps. Ca *s*-states, O and S *p*-states, and Hf *d*-states are shown in blue, red, black, and green, respectively.

employed descriptors and the KRR ML model is provided in Supplementary Discussion and Supplementary Figure 1. For the SCM-based BoB and XRD powder spectra descriptors implementations available in the `matminer` library⁴² were used and KRR implementation with a Gaussian kernel available in `scikit-learn` was used for ML⁴⁶.

Results of our ML model's training and prediction performance for the two targeted properties are presented in Table 2. In each case, the two configurational descriptors (SCM-based BoB and XRD powder spectra) are computed on either the DFT-relaxed ground-state geometry or the unrelaxed geometry where both lattice parameters and the internal atomic coordinates were fixed to ideal GdFeO_3 -type structure for all different configurations. A set of 90% of randomly selected data was used for the KRR ML models' training and the remaining 10% unseen data set was employed to test prediction performance of the trained models. To avoid overfitting, the optimal KRR model hyperparameters were selected during model training stage via minimizing a fivefold cross-validation root mean squared error (RMSE). To

account for slight variability with respect to different training test data splits, in Table 2 the reported mean absolute errors (MAEs) and RMSEs are averaged over five different training/test splits.

It can be seen from the results presented in Table 2 that while reasonable estimates of the two properties on unseen data are possible using either of the two descriptors, descriptors computed on top of the relaxed geometries result in significantly better estimates of the target properties. This is not entirely unexpected and points towards a sensitive dependence of the target properties on the details of local structural relaxations. Although the ML models built on the descriptors computed using the relaxed geometries are more accurate, the models employing unrelaxed geometries are practically more useful as they are able to provide reasonably accurate trends of the properties without requiring an actual DFT computation. Furthermore, owing to systematic trends in the two properties across different A and B-site chemistries (as discussed in detail in the following section), it is expected that such ML models can easily be extended beyond the specific $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ chemistry employed here, with

Table 2. Performance of KRR ML models built on different sets of configurational fingerprints or descriptors in predicting the computed mixing enthalpies and DFT-PBE bandgaps of various $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ configurations considered in the present study.

Target property	Fingerprint scheme employed	Performance with DFT relaxed geometry Avg. MAE (RMSE) in eV		Performance with DFT unrelaxed geometry Avg. MAE (RMSE) in eV	
		Training set	Test set	Training set	Test set
		Mixing enthalpy	BoB with SCM	0.016 (0.026)	0.047 (0.082)
	XRD spectra	0.011 (0.015)	0.034 (0.058)	0.066 (0.114)	0.084 (0.148)
DFT bandgap	BoB with SCM	0.013 (0.017)	0.082 (0.126)	0.074 (0.098)	0.137 (0.191)
	XRD spectra	0.027 (0.035)	0.093 (0.131)	0.116 (0.164)	0.128 (0.174)

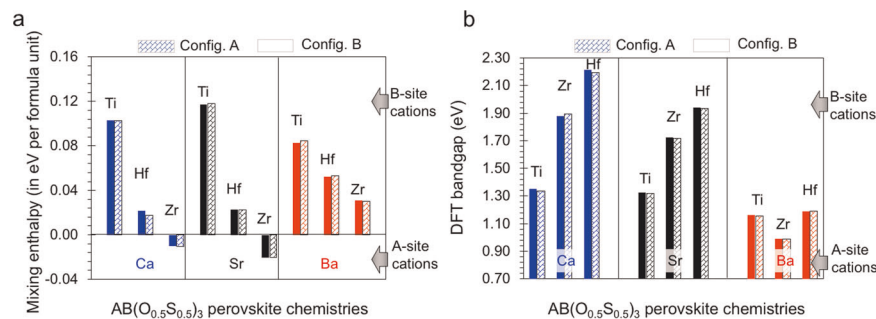


Fig. 4 Relative chemical trends across chemistries. Systematic trends in the computed mixing enthalpies and DFT bandgaps in $\text{AB}(\text{O}_{0.5}\text{S}_{0.5})_3$ chemistries with $A \in \{\text{Ca}, \text{Sr}, \text{Ba}\}$ and $B \in \{\text{Ti}, \text{Zr}, \text{Hf}\}$. Each chemistry is studied in two different *cis*-ordered configurations (both with a $Pmc2_1$ space group) identified as the lowest-energy configurations for $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ from within the entire set of 246 configurations studied and shown in Figs. 1a, b.

exceedingly lower amounts of required training data as new chemistries are added.

General trends across chemistries

After developing an understanding of the trends in energetics and bandgaps as a function of anion configuration in $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ perovskite, next we focus on how these trends might be affected in different (but systematically varying) cation chemistries. In particular, we consider nine different chemistries with $A \in \{\text{Ca}, \text{Sr}, \text{Ba}\}$ and $B \in \{\text{Ti}, \text{Zr}, \text{Hf}\}$. Each of these nine chemistries were studied in two different structures, which are close in energy and exhibit the two lowest enthalpies of mixing for $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ within the entire set of 246 configurations studied. Although the relaxed geometries for all the structures are provided as Supplementary Data 1, we note that both the structures show a *cis*-configuration of S and O atoms around HfO_6 octahedra and are closely related to each other. The only difference lies in the direction in which the zigzag -S-Hf-S-Hf-S- *cis*-bonded chains run. Henceforth these two configurations are referred to as Configs. A and B. A space group determination using FINDSYM⁴⁷, with a tolerance of 0.01 Å for the internal atomic positions as well as the lattice, resulted in the $Pmc2_1$ space group for both the structures considered in the entire set of chemistries with $A \in \{\text{Ca}, \text{Sr}, \text{Ba}\}$ and $B \in \{\text{Ti}, \text{Zr}, \text{Hf}\}$.

The computed mixing enthalpies and generalized gradient approximation (GGA)-PBE bandgaps across this chemical space for the two configurations are presented in Fig. 4a, b, respectively. Note that for a given chemistry, both configurations always show very similar mixing enthalpies and bandgaps. Further, it can be seen from the figure that there are systematic trends going from one cation chemistry to the other, on both the A- and B-site sublattices, which suggests that the analysis made thus far for the specific $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ chemistry is likely to extrapolate to the other related cation chemistries.

Besides the systematic variation of the two properties across chemistries, Figs. 4a, b reveal several important points. First, the ordering of chemical trends in the two properties differ with respect to the B-site chemistries. Although the mixing enthalpies decrease in an order of $\text{Ti} > \text{Hf} > \text{Zr}$, the bandgaps generally increase in an order of $\text{Ti} < \text{Zr} < \text{Hf}$, irrespective of the A-site chemistry. In other words, zirconates and hafnates have different ordering for the two properties. This can be rationalized by noting the fact that, for different B-site cation chemistries, the size effects dictate the variation in mixing enthalpies, whereas the electronegativity of the cation plays a dominant role in determining the bandgap. Although the B cation electronegativities follow the expected trend and decrease going down the Periodic Table from Ti to Hf, the Shannon's ionic radii (for a VI-fold coordinated tetravalent oxidation state) of the cations increase in an order of $\text{Ti} (0.605 \text{ \AA}) < \text{Hf} (0.71 \text{ \AA}) < \text{Zr} (0.72 \text{ \AA})$. The different ionic radii ordering is due to the well-known Lanthanide contraction effect—as a result of the filling of the 4f shell in Hf, an increased nuclear charge leads to a much stronger electrostatic interaction between the nuclei and the electron cloud resulting in a slightly smaller ionic radii for Hf (0.71 Å) as compared with that of Zr (0.72 Å). Second, the DFT-computed bandgap for $\text{BaTi}(\text{O}_{0.5}\text{S}_{0.5})_3$ is anomalously large, as can be seen from Fig. 4b. A closer look at the relaxed structure reveals that, despite being dynamically stable (i.e., no soft mode phonon instabilities), a combination of a large A-site cation and a small B-site cation leads to extremely stretched Ti-S bonds in this system. As a consequence of this local strain, the bandgap further opens up, breaking the common trend exhibited by the remaining chemistries. This local strain also leads to a relatively higher positive enthalpy of mixing for $\text{BaTi}(\text{O}_{0.5}\text{S}_{0.5})_3$. Based on our results, only zirconates of Ca and Sr are predicted to be thermodynamically stable based on the mixing enthalpies. Finally, we note that we have also confirmed that the chemical

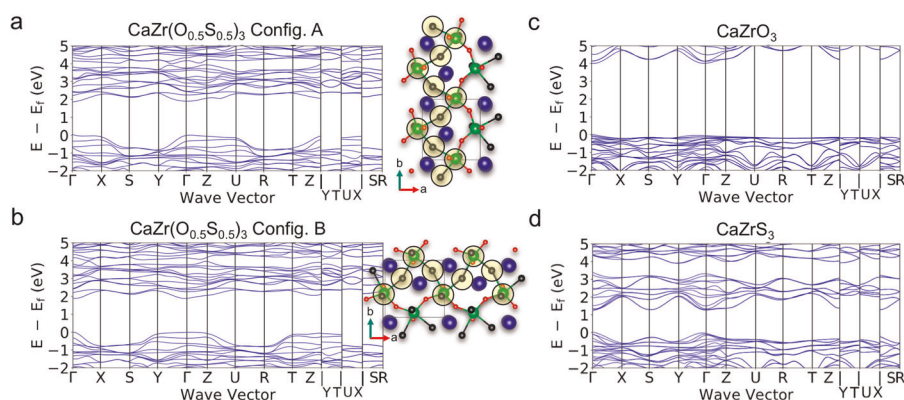


Fig. 5 Electronic band structures of homo and heteroanion chemistries. DFT-computed band structures for **a** $\text{CaZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ config. A (space group $Pmc2_1$), **b** $\text{CaZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ config. B (space group $Pmc2_1$), **c** CaZrO_3 (space group $Pnma$), and **d** CaZrS_3 (space group $Pnma$). The two configurations for the mixed heteroanion chemistry differ in the manner the $-\text{S}-\text{Zr}-\text{S}-\text{Zr}-$ zigzag chains run along the supercell, as identified by black circles in the first two panels.

trends in mixing enthalpies and GGA-PBE bandgaps identified in Figs. 4a, b, respectively, using the GGA-PBE functional do not change when using higher level meta-GGA and hybrid functionals. In particular, for a selected set of cation orderings, we recomputed the mixing enthalpies and GGA-PBE bandgaps using Strongly Constrained and Appropriately Normed (SCAN)⁴⁸ and hybrid Heyd–Scuseria–Ernzerhof HSE06 functionals⁴⁹, respectively, for each of the nine chemistries shown in Fig. 4. Our results, presented in the Supplementary Discussion, Supplementary Figs. 2 and 3 indicate that these functionals possibly improve the quantitative nature of our results, albeit at a significantly higher computational cost, but do not change any of the physical trends or conclusions.

Comparing the electronic band structure of the two configurations (configs. A and B) in each of these stable chemistries reveals that details of anisotropic distributions of anions are reflected in the band structures. For instance, as shown in Figs. 5a, b for $\text{CaZr}(\text{O}_{0.5}\text{S}_{0.5})_3$, config. A showcases continuous zigzag $-\text{S}-\text{Zr}-\text{S}-\text{Zr}-$ chains running along the y axis, whereas in config. B, these chains run along the x axis (for comparison, the band structures of CaZrO_3 and CaZrS_3 are shown in Figs. 5c, d, respectively). Given that the anion p -states largely contribute to the valence band maximum, along the axis of anion ordering (i.e., along the $\Gamma-X$ and $\Gamma-Y$ paths), the valence band dispersion resembles that of the pure sulfide perovskite, whereas in the other two directions, oxygen blocks electron hopping and the bands formed present details that are characteristic of the oxide. The band structure of $\text{SrZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ exhibits a qualitatively similar behavior.

To understand why only Ca and Sr zirconates exhibit negative mixing enthalpy and to further investigate thermodynamic formability in these oxysulfide chemistries, we resort to an empirical analysis based on classical structure factors. The problem of perovskite formability has traditionally been addressed using a pair of key geometric factors, namely the Goldschmidt tolerance factor (t_f) and the octahedral factor (o_f). Although the former directly captures the ability of a given ABX_3 -type perovskite chemistry to achieve a packing arrangement of different atoms in alternate AX and BX_2 {001} family of planes as close to that of an ideal cubic perovskite crystal structure, the latter is only concerned with the stability window of the B-site octahedral coordination. The conventional t_f and o_f are defined as: $(r_A + r_X)/\sqrt{2}(r_B + r_X)$ and r_B/r_X , respectively, where r_i represents Shannon's coordination-dependent ionic radius for the atom i . Based on simple geometrical arguments, t_f should be as close to unity as possible and o_f should fall within $[\sqrt{2}-1, \sqrt{3}-1]$ to form a stable cubic perovskite. Extensions of t_f and o_f to a mixed-anion chemistry $\text{AB}(\text{O}_{0.5}\text{S}_{0.5})_3$ by taking compositional averages of the

anion ionic radii as $[r_A + 0.5(r_O + r_S)]/[\sqrt{2}(r_B + 0.5(r_O + r_S))]$ and $r_B/0.5(r_O + r_S)$, however, fail to capture the stability trends presented in Fig. 4a. In particular, such a definition of a modified t_f results in a metric which is near the formability borderline (on the lower side) for $\text{CaZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ and $\text{SrZr}(\text{O}_{0.5}\text{S}_{0.5})_3$, whereas $\text{BaZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ is predicted to lie in a highly favorable region of the structure map.

A closer look at the *cis*-configured $\text{AB}(\text{O}_{0.5}\text{S}_{0.5})_3$ structure, shown in Fig. 6a, readily points out the problem associated with the employed definition of the compositionally averaged modified t_f . From Fig. 6a, it can be seen that for the $\text{B}(\text{O}_{0.5}\text{S}_{0.5})_2$ {001} planes, the alternating arrangement of O and S does justify use of an averaged anion ionic radius in both the o_f and denominator of the t_f . However, the {100} planes with A-site cations exhibit atomic layers containing either 100% O or 100% S as anionic species. Therefore, given the much larger atomic size of S as compared with O, only AS planes can be close packed. This insight indicates that the modified tolerance factor should rather be defined as $(r_A + r_S)/\sqrt{2}[r_B + 0.5(r_O + r_S)]$ for the present case. Using this revised modified t_f and an anion-averaged o_f , a structure map presented in Fig. 6b correctly explains the DFT-computed relative stabilities across the considered set of chemistries. In particular, titanates are unstable because of a much smaller B-site cation radius (average $o_f < 0.414$) and Ba-containing chemistries are unfavorable because A-site radius is larger than desired (modified $t_f > 1.04$). The figure also shows that the energetics of the oxysulfides is very sensitive to the size of B-site cations (as compared with that of the A-site cations) and a very small size difference between the Zr and Hf is actually sufficient to switch the trends of thermodynamic stabilities between some of these chemistries. At last, we note that, in addition to being helpful in rationalization of stability trends, the structure map presented in Fig. 6b can, in principle, be used to screen other potentially stable equimolar *cis*-ordered oxysulfide chemistries. If a given chemistry falls near and above the highlighted region (shown by a dashed rectangle) of stable zirconates, the geometric arguments favor stability in a perovskite crystal structure, though many other factors may also need to be addressed before any firm conclusions can be made.

At last, we note that based on the computed mixing enthalpies both $\text{SrZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ and $\text{CaZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ are predicted to be thermodynamically stable. Since synthesis of $\text{BaZr}(\text{O}_{0.5}\text{S}_{0.5})_3$ has already been demonstrated in the past¹⁸, we anticipate that experimental realization of relatively stable (although thermodynamically metastable)⁵⁰ $\text{CaHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ and $\text{SrHf}(\text{O}_{0.5}\text{S}_{0.5})_3$ should also be highly likely.

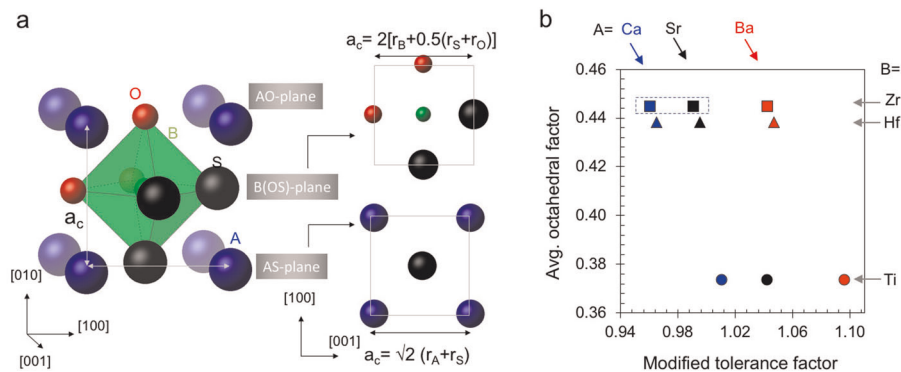


Fig. 6 Geometric factors influencing formability in a *cis*-ordered $AB(O_{0.5}S_{0.5})_3$ perovskite oxysulfide. **a** A schematic showing local atomic arrangements in a *cis*-ordered $AB(O_{0.5}S_{0.5})_3$ perovskite oxysulfide. The AS and BOS close-packed planes along the [010] direction are explicitly shown. Note that alternate {001} AX atomic planes feature either all-oxide or all-sulfide stackings. Owing to the much larger size of the S atom, only AS planes can be close-packed, forming the basis for a modified tolerance factor t_f proposed in this work. **b** The proposed modified tolerance factor t_f and anion-averaged octahedral factor o_f readily cluster the two stable zirconate chemistries (shown within a dashed rectangle) and allow us to rationalize the DFT-computed trends in relative energetics. Note that the Cartesian axes in the figure refer to a five-atom cubic perovskite unit cell.

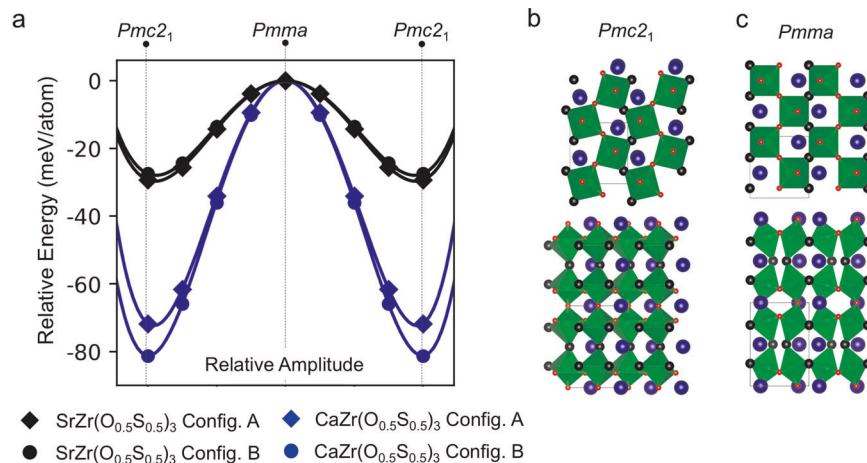


Fig. 7 Anion ordering induced ferroelectric polarization. **a** The DFT-computed barrier for ferroelectric switching in the two thermodynamically stable zirconate chemistries. The results for both config. A and config. B are shown in each case. **b, c** present the top and side views of the $Pmc2_1$ ground state and $Pmma$ saddle point geometric structures in the config. A.

Anion-order-driven ferroelectricity

This work not only identifies and analyzes the factors that control anion ordering in perovskite oxysulfides, but also has considerable implications towards functionality exhibited by heteroanionic perovskites. For instance, considering the two stable zirconate chemistries shown in Fig. 4a, we note that while the two end members $AZrO_3$ and $AZrS_3$ (with $A=Ca$ and Sr) are both centrosymmetric with a $Pmma$ space group, the most-stable *cis*-type anion ordering breaks mirror symmetry along the c axis owing to the alternating AO and AS planes, lowering the symmetry to a polar $Pmc2_1$ space group. Therefore, the *cis*-ordering could also potentially lead to the phenomena of hybrid improper ferroelectricity^{51–54}, if the preferred anion ordering can manifest in large single domains. To examine the potential ferroelectricity in these polar materials, we compute the macroscopic polarization and the associated switching barrier using the Berry phase approach⁵⁵ made available within the framework of the modern theory of polarization^{55–57}. As a non-polar reference, we take the centrosymmetric $Pmma$ structure obtained by undoing the octahedral rotation around the c axis using PSEUDO⁵⁸, and we relax this with constrained symmetry. To fit the polarization branch⁵⁵, we use ISODISTORT⁵⁹ to interpolate three intermediate images along the switching pathway between

$Pmma$ and $Pmc2_1$. The calculated polarizations and switching barriers for the two stable zirconate chemistries are shown in Fig. 7a and the structural details of the $Pmc2_1$ and $Pmma$ phases are depicted in Figs. 7b, c, respectively. We find that both chemistries are predicted to be ferroelectric with a computed macroscopic polarization of $>30 \mu C/cm^2$. Further, quantitative details for our computations are provided in Table 3.

Although the ferroelectric mechanism here is similar to that of the known hybrid improper ferroelectrics^{52–54}, here we note an important distinction. In a $Pmma$ ABX_3 -type perovskite, the A-site cations in each {001} plane along the unique axis of tilting displace in alternating directions. The perovskite is non-polar because these dipoles cancel exactly. However, typically when symmetry is lowered to the $Pmc2_1$ space group by breaking the mirror symmetry along the c axis (for instance in a double perovskite containing two A-site cations arranged in a layered-type ordering)⁵⁴, one of the dipoles becomes larger than the other, resulting in a relatively small net polarization. However, in these ordered oxysulfides, as a consequence of the *cis*-type anion ordering, all the A-site cations get displace in the same direction, resulting in a significantly larger macroscopic polarization than that of typical improper ferroelectrics reported in the literature⁵⁴. Specifically, the cations in the AS layers behave as expected, whereas those in the AO layer reverse direction. This anion-order-driven polarization is

Table 3. The DFT-computed macroscopic polarizations and ferroelectric switching barriers in the two thermodynamically stable zirconate chemistries.

Composition configuration	Space group	Polarization $\mu\text{C}/\text{cm}^2$	Polarization direction	Well-depth meV/atom
<i>CaZr(O_{0.5}S_{0.5})₃</i>				
Config. A	<i>Pmc2₁</i>	36	$\pm[010]$	71.8
Config. B	<i>Pmc2₁</i>	35	$\pm[100]$	81.2
<i>SrZr(O_{0.5}S_{0.5})₃</i>				
Config. A	<i>Pmc2₁</i>	32	$\pm[010]$	29.4
Config. B	<i>Pmc2₁</i>	32	$\pm[100]$	27.6

The results for both the config. A and config. B are shown for each of the chemistries. Net polarization direction refers to the Cartesian axes in a GdFeO₃-type *Pnma* crystal structure.

predicted to result in a macroscopic polarization that is larger than that of the well-known ferroelectrics such as BaTiO₃⁶⁰ and KNbO₃⁶¹.

DISCUSSION

Oxysulfide perovskite compounds represent an emerging class of materials with a potential use in a diverse range of technological applications, including solar energy harvesting, catalysis, and sensing. In contrast to heterocationic double perovskite compounds, where the underlying design rules and cation ordering tendencies are relatively well understood and documented, the factors that govern the anion ordering phenomena within this class of heteroanionic compounds remain poorly understood. As an important step in this direction, the present manuscript employed a selected set of first-principles-based DFT computations to first study a wide range of anion orderings in a prototypical equimolar oxysulfide perovskite CaHf(O_{0.5}S_{0.5})₃ chemistry. Subsequently, the most-stable anion ordering was studied and compared in a range of different AB(O_{0.5}S_{0.5})₃ chemistries with A ∈ {Ca, Sr, Ba} and B ∈ {Ti, Zr, Hf}. The salient findings of our work can be summarized as follows:

- AB(O_{0.5}S_{0.5})₃ oxysulfide perovskites with a B-site cation in a *d*⁰ electronic configuration exhibit a strong preference for a *cis*-type anion ordering. The driving force for this specific ordering tendency stems from a combined interplay of electronic, elastic, and electrostatic factors.
- Descriptors based on either BoB-SCM or simulated XRD spectra are shown to be effective in learning the anion - configuration-dependent variations in the mixing enthalpies and the bandgaps, with both descriptors leading to a comparable performance.
- Heteroanionic *cis*-ordered AB(O_{0.5}S_{0.5})₃ chemistries with A ∈ {Ca, Sr, Ba} and B ∈ {Ti, Zr, Hf} exhibit well-behaved trends in both mixing enthalpies and bandgaps. While the bandgap variations are well aligned with trends in the metallic nature of the A- and B-site cations in the Periodic Table, mixing enthalpies exhibit a relatively more complex behavior.
- The trends in mixing enthalpies of these oxysulfides (computed relative to the respective pure oxide and sulfide chemistries) can be understood by using the compositionally averaged octahedral factor and a modification of the classical tolerance factor that accounts for presence of AS planes, as the larger sulfide ions are the limiting factor in forming a close-packed structure.
- Based on the computed mixing enthalpies, we predict a relative stability ordering of the following oxysulfide

perovskites as: SrZr(O_{0.5}S_{0.5})₃ > CaZr(O_{0.5}S_{0.5})₃ > CaHf(O_{0.5}S_{0.5})₃ > SrHf(O_{0.5}S_{0.5})₃ > BaZr(O_{0.5}S_{0.5})₃, with a negative enthalpy of mixing for the first two chemistries. Given that a number of BaZr(O_{1-x}S_x)₃ perovskites spanning over the wide range of $x \in [0, 0.54]$ have already been synthesized and reported¹⁸, all of the above chemistries should be amenable to experimental synthesis. In particular, zirconates of Ca and Sr should exhibit long term stability.

- The strongly favored *cis*-type anion ordering breaks inversion symmetry locally, and can potentially lead to a large macroscopic polarization >30 $\mu\text{C}/\text{cm}^2$, if large anion-ordered domains are sustained.

METHODS

Details of DFT computations

The DFT calculations were performed using the Vienna Ab initio Simulation Package (VASP)^{62,63} employing the Perdew, Burke, and Ernzerhof (PBE)⁶⁴ GGA exchange-correlation functional. The electronic wave functions were expanded in plane waves up to a cutoff energy of 500 eV. We used potentials based on the projector-augmented wave method⁶⁵ explicitly including the following valence electronic configurations for different elemental species; Ca: 3s²3p⁶4s², Sr: 4s²4p⁶5s², Ba: 5s²5p⁶6s², Zr: 4p⁶4d²5s², Hf: 5p⁶5d²6s², O: 2s²2p⁴, and S: 3s²3p⁴. A Gamma-centered 5 × 5 × 3 Monkhorst–Pack *k*-point mesh⁶⁶ was used for Brillouin-zone integrations for an orthorhombic supercell with *Pnma* space group containing 20 atoms (i.e., a $\sqrt{2} \times \sqrt{2} \times 2$ pseudo-cubic supercell). Spin-unpolarized calculations were employed as none of the simulated systems contained any unpaired electrons. To obtain a geometry optimized equilibrium structure, atomic positions and the supercell lattice parameters were fully relaxed using the conjugate-gradient method until all the Hellmann–Feynman forces and the stress components were >0.05 V/Å and 1.0×10^{-2} GPa, respectively.

DATA AVAILABILITY

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

CODE AVAILABILITY

The VASP code used in this study is a commercial electronic structure modeling software, available from <https://www.vasp.at>.

Received: 29 November 2019; Accepted: 11 May 2020;

Published online: 03 June 2020

REFERENCES

1. Mitchell, R. H. *Perovskites: modern and ancient*; Almaz Press Thunder Bay, Vol. 7 (2002).
2. Muller, O. & Roy, R. *The major ternary structural families*; Springer (1974).
3. Kudo, A. & Miseki, Y. Heterogeneous photocatalyst materials for water splitting. *Chem. Soc. Rev.* **38**, 253–278 (2009).
4. NavarroYerga, R. M. et al. Water splitting on semiconductor catalysts under visible-light irradiation. *ChemSusChem* **2**, 471–485 (2009).
5. Scott, J. F. Applications of modern ferroelectrics. *science* **315**, 954–959 (2007).
6. Tejuca, L. G., Fierro, J. L. *Properties and applications of perovskite-type oxides*; CRC Press, 1992
7. Kageyama, H. et al. Expanding frontiers in materials chemistry and physics with multiple anions. *Nat. Commun.* **9**, 772 (2018).
8. Thind, A. S. et al. KBaTeBiO₆: a lead-free, inorganic double-perovskite semiconductor for photovoltaic applications. *Chem. Mater.* **13**, 4769–4778 (2019).
9. Vasala, S. & Karppinen, M. A2B'B'O₆ perovskites: a review. *Prog. Solid State Chem.* **43**, 1–36 (2015).
10. King, G. & Woodward, P. M. Cation ordering in perovskites. *J. Mater. Chem.* **20**, 5785–5796 (2010).
11. Anderson, M. T., Greenwood, K. B., Taylor, G. A. & Poeppelmeier, K. R. B-cation arrangements in double perovskites. *Prog. Solid State Chem.* **22**, 197–233 (1993).

12. Knapp, M. C. & Woodward, P. M. A-site cation ordering in AA'BB'O₆ perovskites. *J. Solid State Chem.* **179**, 1076–1085 (2006).
13. Howard, C. J., Kennedy, B. J. & Woodward, P. M. Ordered double perovskites—a group-theoretical analysis. *Acta Crystallogr. Sect.* **59**, 463–471 (2003).
14. Laurita, G. et al. Covalency-driven structural evolution in the polar pyrochlore series Cd₂Nb₂O_{7-x} S x. *Chem. Mater.* **31**, 7626–7637 (2019).
15. Talanov, M. V., Shirokov, V. B. & Talanov, V. M. Anion order in perovskites: a group-theoretical analysis. *Acta Crystallogr. Sect.* **72**, 222–235 (2016).
16. Wang, Q. et al. Oxsulfide photocatalyst for visible-light-driven overall water splitting. *Nat. Mater.* **18**, 827–832 (2019).
17. Charles, N., Saballos, R. J. & Rondinelli, J. M. Structural diversity from anion order in heteroanionic materials. *Chem. Mater.* **30**, 3528–3537 (2018).
18. Perera, S. et al. Chalcogenide perovskites—an emerging class of ionic semiconductors. *Nano Energy* **22**, 129–135 (2016).
19. Camp, P. J., Fuertes, A. & Attfield, J. P. Subextensive entropies and open order in perovskite oxynitrides. *J. Am. Chem. Soc.* **134**, 6762–6766 (2012).
20. Yang, M. et al. Anion order in perovskite oxynitrides. *Nat. Chem.* **3**, 47 (2011).
21. Kaneko, M., Fujii, M., Hisatomi, T., Yamashita, K. & Domen, K. Regression model for stabilization energies associated with anion ordering in perovskite-type oxynitrides. *J. Energy Chem.* **36**, 7–14 (2019).
22. Hartman, S. T., Cho, S. B. & Mishra, R. Multiferroism in iron-based oxyfluoride perovskites. *Inorg. Chem.* **57**, 10616–10624 (2018).
23. Sun, Y.-Y., Agiorgousis, M. L., Zhang, P. & Zhang, S. Chalcogenide perovskites for photovoltaics. *Nano Lett.* **15**, 581–585 (2015).
24. Lee, C.-S., Kleinke, K. M. & Kleinke, H. Synthesis, structure, and electronic and physical properties of the two SrZrS₃ modifications. *Solid State Sci.* **7**, 1049–1054 (2005).
25. Wang, Y., Sato, N. & Fujino, T. Synthesis of BaZrS₃ by short time reaction at lower temperatures. *J. Alloy Compd.* **327**, 104–112 (2001).
26. Wang, Y., Sato, N., Yamada, K. & Fujino, T. Synthesis of BaZrS₃ in the presence of excess sulfur. *J. Alloy Compd.* **311**, 214–223 (2000).
27. Clearfield, A. The synthesis and crystal structures of some alkaline earth titanium and zirconium sulfides. *Acta Crystallogr.* **16**, 135–142 (1963).
28. Takeda, Y., Kanamura, F., Shimada, M. & Koizumi, M. The crystal structure of BaNiO₃. *Acta Crystallogr. Sect. B* **32**, 2464–2466 (1976).
29. Brasseur, H. & Pauling, L. The crystal structure of ammonium cadmium chloride, NH₄CdCl₃. *J. Am. Chem. Soc.* **60**, 2886–2890 (1938).
30. Tranchitella, L. J., Chen, B.-H., Fettinger, J. C. & Eichhorn, B. W. Structural evolutions in the Sr_{1-x}BaxZrSe₃Series. *J. Solid State Chem.* **130**, 20–27 (1997).
31. Kuhar, K. et al. Sulfide perovskites for solar energy conversion applications: computational screening and synthesis of the selected compound LaY₃. *Energy Environ. Sci.* **10**, 2579–2593 (2017).
32. Hart, G. L. & Forcade, R. W. Algorithm for generating derivative structures. *Phys. Rev. B* **77**, 224115 (2008).
33. Hart, G. L., Nelson, L. J. & Forcade, R. W. Generating derivative structures at a fixed concentration. *Comput. Mater. Sci.* **59**, 101–107 (2012).
34. Derivative structure enumeration library. Contribute to msg-byu/enumlib development by creating an account on GitHub. <https://github.com/msg-byu/enumlib> (2018).
35. Tatsumi, K. & Hoffmann, R. Bent cis d₀ MoO₂₂. vs. linear trans d₀f₀ UO₂₂: a significant role for nonvalence 6p orbitals in uranyl. *Inorg. Chem.* **19**, 2656–2658 (1980).
36. Barrie, P., Coffey, T. A., Forster, G. D. & Hogarth, G. Bent vs. linear imido ligation at the octahedral molybdenum (VI) dithiocarbamate stabilised centre. *J. Chem. Soc. Dalton Transac.* 4519–4528 (1999).
37. Wolff, H. & Dronskowski, R. First-principles and molecular-dynamics study of structure and bonding in perovskite-type oxynitrides ABO₂N (A=Ca, Sr, Ba; B=Ta, Nb). *J. Comput. Chem.* **29**, 2260–2267 (2008).
38. Shannon, R. D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr. Sect. A* **32**, 751–767 (1976).
39. Pilania, G., Yadav, S. K., Nikl, M., Uberuaga, B. P. & Stanek, C. R. Role of multiple charge states of ce in the scintillation of AB₃O₃ perovskites. *Phys. Rev. Applied* **10**, 024026 (2018).
40. Hoat, D. M., Silva, J. R. & Blas, A. M. First principles study of structural, electronic and optical properties of perovskites CaZrO₃ and CaHfO₃ in cubic phase. *Solid State Commun.* **275**, 29–34 (2018).
41. Faber, F., Lindmaa, A., von Lilienfeld, O. A. & Armiento, R. Crystal structure representations for machine learning models of formation energies. *Int. J. Quant. Chem.* **115**, 1094–1101 (2015).
42. Ward, L. et al. Matminer: an open source toolkit for materials data mining. *Comput. Mater. Sci.* **152**, 60–69 (2018).
43. Friedman, J., Hastie, T. & Tibshirani, R. *The elements of statistical learning*; Springer series in statistics New York, 1 (2001).
44. Pilania, G., Wang, C., Jiang, X., Rajasekaran, S. & Ramprasad, R. Accelerating materials property predictions using machine learning. *Sci. Rep.* **3**, 2810 (2013).
45. Pilania, G. et al. Machine learning bandgaps of double perovskites. *Sci. Rep.* **6**, 19375 (2016).
46. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
47. Stokes, H. T. & Hatch, D. M. FINDSYM: program for identifying the space-group symmetry of a crystal. *J. Appl. Crystallogr.* **38**, 237–238 (2005).
48. Sun, J., Ruzsinszky, A. & Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 036402 (2015).
49. Heyd, J., Scuseria, G. E. & Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. **118**, 8207–8215 (2003).
50. Sun, W. et al. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).
51. Levanyuk, A. P. & Sannikov, D. G. Improper ferroelectrics. *Sov. Phys. Uspekhi* **17**, 199 (1974).
52. Bousquet, E. et al. Improper ferroelectricity in perovskite oxide artificial superlattices. *Nature* **452**, 732 (2008).
53. Benedek, N. A. & Fennie, C. J. Hybrid improper ferroelectricity: a mechanism for controllable polarization-magnetization coupling. *Phys. Rev. Lett.* **106**, 107204 (2011).
54. Pilania, G. & Lookman, T. Electronic structure and biaxial strain in RbHgF₃ perovskite and hybrid improper ferroelectricity in (Na, Rb) Hg₂F₆ and (K, Rb) Hg₂F₆ superlattices. *Phys. Rev.* **90**, 115121 (2014).
55. King-Smith, R. D. & Vanderbilt, D. Theory of polarization of crystalline solids. *Phys. Rev. B* **47**, 1651 (1993).
56. Rabe, K., Ahn, C. H. & Triscone, J.-M. Physics of Ferroelectrics: A Modern Perspective. *Topics in Applied Physics* (Springer, 2007).
57. Resta, R. Modern theory of polarization in ferroelectrics. *Ferroelectrics* **151**, 49–58 (1994).
58. Capillas, C. et al. A new computer tool at the Bilbao Crystallographic Server to detect and characterize pseudosymmetry. *Z. Kristallogr. Cryst. Mater.* **226**, 186–196 (2011).
59. Stokes, H. T., Hatch, D. M., Campbell, B. J. & Tanner, D. E. ISODISPLACE: a web-based tool for exploring structural distortions. *J. Appl. Crystallogr.* **39**, 607–614 (2006).
60. Cudney, R. S. et al. Photorefractive and domain gratings in barium titanate. *Appl. Phys. Lett.* **63**, 3399–3401 (1993).
61. Triebwasser, S. Behavior of ferroelectric KNbO₃ in the vicinity of the cubic-tetragonal transition. *Phys. Rev.* **101**, 993 (1956).
62. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
63. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758 (1999).
64. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
65. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
66. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188 (1976).
67. Lelieveld, R. & Ijdo, D. J. W. Sulphides with the GdFeO₃ structure. *Acta Crystallogr. Sect. B* **36**, 2223–2226 (1980).
68. Dorenbos, P. The electronic structure of lanthanide doped compounds with 3d, 4d, 5d, or 6d conduction band states. *J. Lumin.* **151**, 224–228 (2014).

ACKNOWLEDGEMENTS

G.P., C.R.S., and B.P.U. gratefully acknowledge support from the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project # 20190043DR. Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of US Department of Energy (Contract No. 89233218CNA000001). S.T.H. and R.M. acknowledge support from the National Science Foundation through DMR-1806147. Computational support for this work was provided by LANL's high-performance computing clusters.

AUTHOR CONTRIBUTIONS

G.P. designed with inputs from R.M., C.R.S., and B.P.U. A.G. and S.T.H. performed the DFT computations and help analyze the data. All authors discussed the results and commented on the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at <https://doi.org/10.1038/s41524-020-0338-1>.

Correspondence and requests for materials should be addressed to G.P.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2020