Journal of Materials Chemistry A

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: J. Mater. Chem. A, 2025, 13, 35447

Machine-learning-driven modelling of amorphous and polycrystalline $BaZrS_3$

Laura-Bianca Paşca, [a] Yuanbin Liu, [a] Andy S. Anker, [ab] Ludmilla Steier [a] and Volker L. Deringer *[a]

The chalcogenide perovskite material $BaZrS_3$ is of growing interest for emerging thin-film photovoltaics. Here we show how machine-learning-driven modelling can be used to describe the material's amorphous precursor as well as polycrystalline structures with complex grain boundaries. Using a bespoke machine-learned interatomic potential (MLIP) model for $BaZrS_3$, we study the atomic-scale structure of the amorphous phase, quantify grain-boundary formation energies, and create realistic-scale polycrystalline structural models which can be compared to experimental data. Beyond $BaZrS_3$, our work exemplifies the increasingly central role of MLIPs in materials chemistry and marks a step towards realistic device-scale simulations of materials that are gaining momentum in the fields of photovoltaics and photocatalysis.

Received 4th June 2025
Accepted 1st September 2025

DOI: 10.1039/d5ta04536c

rsc.li/materials-a

Introduction

In the search for new, sustainable photoabsorbers, sulfide-based chalcogenide perovskite materials have emerged as attractive lead-free candidates.[1] However, while oxide and halide perovskites have defined much of the progress in photovoltaics and related fields, chalcogenide perovskites have only more recently begun to be explored. Among the latter, $BaZrS_3$ presents optical absorption matching or even surpassing those of halide perovskites and GaAs,[2] competitive charge carrier lifetimes, and improved stability to environmental factors compared to other perovskite materials.[3–5] Thin films of $BaZrS_3$ can be synthesised from earth-abundant and non-toxic elements: by sulfidation of Ba–Zr–O precursors[4,6,7] or by directly depositing sulfide species using pulsed laser deposition,[8] molecular beam epitaxy,[9] or sputtering.[3] Most methods involve the deposition of amorphous precursors that require temperatures of $\approx$ 900 °C to crystallise. Their growth and subsequent crystallisation has been followed experimentally using X-ray diffraction (XRD) or X-ray spectroscopy techniques.[10–12]

Given the rapidly growing interest in $BaZrS_3$, computational methods are increasingly used to complement experimental studies of this material. Density-functional theory (DFT) and phonon computations were employed to map out the thermodynamic conditions under which $BaZrS_3$ films might form and which surface termination is expected to be the most stable.[13,14] To reach beyond the system-size limits of DFT-based methods, machine-learned interatomic potentials (MLIPs) have now been applied to many functional materials,[15–17] including halide perovskites.[18–21] The chalcogenide alternatives, *viz.* $BaZrS_3$ and homologous compounds, were recently studied in a comprehensive work using ML-accelerated molecular dynamics (MD).[22] These studies have typically focused on the crystalline material[20,22] and the formation of other phases, such as the binary crystals or 2D Ruddlesden–Popper structures.[13] To validate ML-accelerated MD, Kayastha *et al.* compared simulated XRD patterns for MD-generated $BaZrS_3$ structures with experimental XRD patterns.[23] However, these simulations also were focused on ordered unit cells, corresponding to single-crystalline samples.

This limitation is more generally a current research challenge in modelling perovskite solar-cell materials: experimentally synthesised materials are usually polycrystalline, and fully realistic simulations would therefore need to involve structural models representing individual grains, with sizes typically on the order of tens or hundreds of nanometres.[10,12] A single-crystalline structural model will therefore likely not suffice to fully understand the structure and properties of $BaZrS_3$ films. We have recently reported very-large-scale atomistic models of functional materials, including phase-change materials for data storage[30] and amorphous silicon which is relevant to solar cells.[31,32] It would seem highly beneficial to achieve this type of realism for chemically complex, perovskite-type photoabsorber materials as well.

Here, we introduce a machine-learned interatomic potential (MLIP) model for ordered and disordered forms of $BaZrS_3$, based on the atomic cluster expansion (ACE) framework.[33,34] For training, we employ a combination of *de novo*[35,36] and domain-

[a]Inorganic Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QR, UK. E-mail: volker.deringer@chem.ox.ac.uk
[b]Department of Energy Conversion and Storage, Technical University of Denmark, Kgs. Lyngby 2800, Denmark

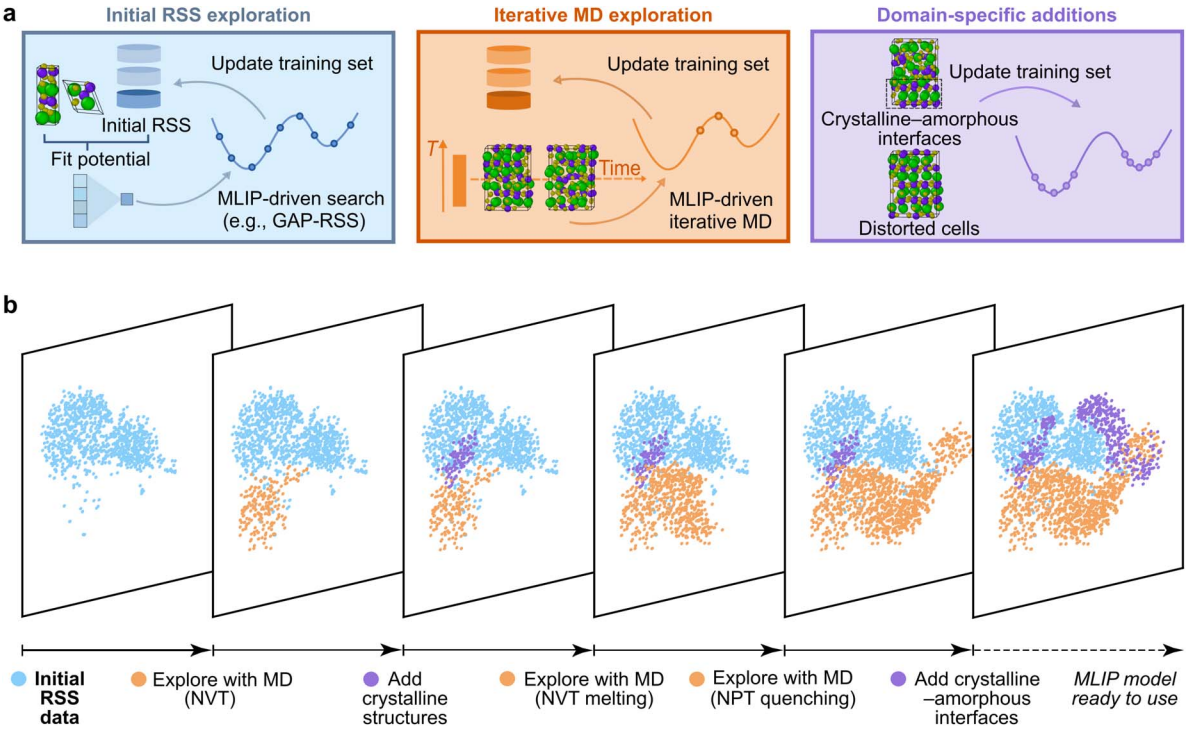

Fig. 1 A machine-learned interatomic potential for crystalline and amorphous $BaZrS_3$. (a) Schematics of the different approaches used in training dataset construction, showing examples of the different configuration types sampled. Structural images were created using OVITO.[24] (b) Evolution of the training dataset, visualised in a style similar to ref. 27. Each slice provides a two-dimensional representation (using the UMAP algorithm[28]) of the relevant configurational space, showing the reference training dataset of the ML potential, distributed based on the structures' average atomic-environment similarity. The latter is quantified using the SOAP kernel similarity metric[29] with a cut-off radius of 5 Å and smoothness of 0.75 Å (see SI for more details): the distances between points in this two-dimensional space therefore reflect the structural (dis-) similarity between entries of the training dataset.

specific iterative training (Fig. 1), aiming for the final dataset to capture the structural complexities of $BaZrS_3$. We show how ML-driven simulations can describe three scenarios relevant to experimental studies: (i) the amorphous precursor; (ii) large-scale grain boundaries; and (iii) polycrystalline $BaZrS_3$ structures. This way, ML-driven simulations can corroborate experimental observations regarding the atomistic structure of this material and provide insights that experiments on their own can not. Beyond their application to $BaZrS_3$, we expect that ML-driven approaches for simulating polycrystalline structures—from precursors to individual grains—can more broadly accelerate computational studies of diverse polycrystalline solar-cell materials.

Methodology

The choice of data used to train MLIP models is now a central consideration in the field,[37] and different approaches to dataset building have been discussed in the literature.[38,39] Here, we begin with random structure searching (RSS)[40,41] using an iterative protocol similar to ref. 35 and 36, whereby an initial MLIP is trained on randomised structures and then used to sequentially drive the RSS exploration (Fig. 1a). It was previously shown that RSS can sample the complex atomistic environments relevant to grain boundaries and interfaces by generating diverse starting structures.[42] High-temperature structures obtained from ML-driven MD, as well as crystalline–amorphous interfaces were subsequently added (SI). As shown in a series of "structure maps" in Fig. 1b, the structures generated at different temperatures using iterative MD (shown as orange points in Fig. 1b) sample different regions of configurational space compared to the initial RSS dataset (light blue). Furthermore, our "domain-specific" additions to the dataset (purple), such as small-scale structural models representing crystalline–amorphous interfaces, include structures in-between disordered, high-temperature snapshots from MD and crystalline structures. Domain-specific training data such as interface structures have been used before to help describe crystallisation processes in Ge–Sb–Te memory materials, for example.[30]

We used different MLIP fitting approaches as part of the development of the final model. Initially, the Gaussian Approximation Potential (GAP)[43] framework was used because of its data efficiency: it allowed us to generate a stable initial potential with relatively few training data points (90 initial RSS structures, with a further 899 structures obtained from *de novo* GAP-RSS exploration).[35,36] Once a larger dataset had been built by iterative training, the computationally efficient ACE framework as implemented in `pacemaker` was used to fit a faster MLIP model to that dataset.[34,44]

The final ACE model was obtained by iterative training until it could reliably generate a structural model for the amorphous phase using an MD melt–quench protocol in the NPT ensemble

(see SI for further details). Two model versions were fitted to the final dataset: the first by filtering out structures with high DFT energies (>1 eV per atom), indicative of very close contacts between atoms or high-energy RSS structures, and the second using the full reference dataset (SI). The first version was used to generate all the quantitative data presented, as it achieved good accuracy on the structures relevant in the present study, showing an energy root-mean-square error (RMSE) of 13.9 meV per atom relative to DFT results using the PBEsol functional.[45] The second version, which includes higher-energy dimers and random structures, is less accurate (energy RMSE: 23.1 meV per atom), but it did not fail when handling structures with closer contacts between atoms, and therefore it was used to relax the polycrystalline structures with randomly-oriented grains. (We consider a simulation to have "failed" if, during the MD run, atoms come closer to each other than 1 Å, collapsing the structure, or if atoms are lost during the simulation.) Details of numerical errors are provided in Fig. S1 and S2 in the SI. As an additional test, we computed the phonon dispersion curves for crystalline $BaZrS_3$ (Fig. S3).

Results and discussion

We describe the computational modelling of $BaZrS_3$ in the same sequence as would be relevant to experimental synthesis and characterisation. First, we use the MLIP model to simulate the amorphous phase, corresponding to precursor phases that are deposited in experiments.[3,4] Second, we validate the model for grain boundaries, which need to be accurately described so that the model can be applied to polycrystalline samples. With both of these aspects available, we finally apply the model to simulating structures with different grain sizes, providing a direct connection to experimental scattering data.

Amorphous $BaZrS_3$

The amorphous structure simulated using the MLIP model is shown in Fig. 2a. The structure was obtained using the melt–quench protocol described in the SI. The effect of changing the quench rate and the starting configuration is detailed in Tables S3 and S4, respectively. In our structural model of amorphous $BaZrS_3$, many of the Zr atoms still have a (defective) octahedral coordination by S, similar to the crystalline structure; however, the ZrS_x polyhedra lack long-range order. The geometry of the different Zr coordination environments, which also present undercoordinated and a few overcoordinated Zr atoms, and those of the BaS_x polyhedra, are shown in the histogram plots in Fig. 2b and c, respectively. We show examples of the coordination polyhedra for each coordination number (CN). A wider distribution of CNs is observed for Ba, and in this case, also, we observe a pronounced undercoordination of the cation, which has an expected CN of 8 in crystalline $BaZrS_3$ (due to the orthorhombic distortion of its perovskite-like structure, and compared to CN = 12 for the cubic archetype).

The local ordering of S atoms around the A- and B-site cations can be observed in the radial distribution function (RDF) of the quenched amorphous structure (Fig. 2d, left), which shows well-defined peaks in the short-range region, at distances roughly below 4 Å. These two main peaks correspond to the Zr–S and Ba–S interatomic distances, as confirmed by the partial RDF plots (Fig. 2d, right). The RDF data obtained from atomistic simulations can be compared with experimental

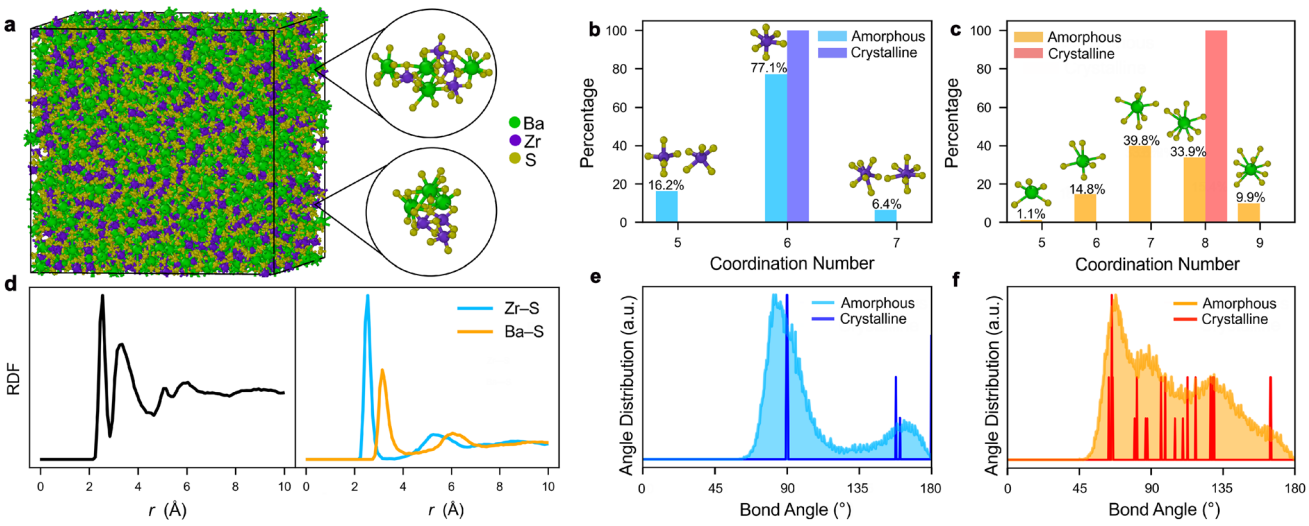

Fig. 2 Amorphous $BaZrS_3$. (a) The atomistic structure of the simulated amorphous phase (10 240 atoms) and manually-chosen close-ups showing the range of geometries and connectivity types between the coordination polyhedra of the A- and B-site cations. (b) Histogram plot showing the distribution of coordination numbers around Zr atoms in the amorphous and crystalline phases, respectively. (Note that environments with CN <0.5% are not visualised in the histogram, but Table S3 presents full details of the coordination numbers.) (c) Same but for Ba atoms. (d) Total and partial RDFs for the Zr–S and Ba–S interatomic distances in the 10 240-atom simulated amorphous phase. (e) ADF plot showing the S–Zr–S bond-angle distribution in the amorphous phase, compared to the crystalline structure. (f) Same but for S–Ba–S bond angles. The radial cut-off used for calculating coordination numbers and ADF plots was set to 3.1 Å in the case of the Zr environments and to 3.8 Å in the case of Ba.

extended X-ray absorption fine structure spectroscopy (EXAFS) results,[10] which probe the local structure around Zr atoms in the amorphous phase. Although the techniques are different, both the EXAFS data and our simulations qualitatively indicate under-coordinated Zr environments in amorphous $BaZrS_3$ compared to its crystalline counterpart. Specifically, analysis of the EXAFS data yielded a Zr–S bond length of 2.593 Å and coordination number (CN) of 5.2 (see ref. 10 for details), while our simulations yield an average bond length of 2.575 Å and CN of 5.9 (determined using a 3.1 Å cut-off in OVITO).[24] A more in-depth comparison between the experimental and simulated values of the bond length and CN of Zr is provided in Table S5. In addition to the differences in methodology, the experimental values may be influenced by factors such as defects in the amorphous precursor material, for example sulfur vacancies, which were not taken into account in our current model. While modelling defects is beyond the scope of the current work, datasets such as those provided by the *ab initio* study of the defect landscape of $BaZrS_3$ by Desai *et al.*,[25] as well as methodologies such as those developed by Mosquera-Lois *et al.*,[26] could be used to expand the training of MLIP models for $BaZrS_3$ further.

The presence of B-site cation halide fragments in our simulated structure, some maintaining a similar octahedral geometry to the crystalline phase, is in agreement with experimentally reported structural details of the amorphous phases derived from other materials adopting the perovskite structure.[10,46,47] Preserved local bonding units of TiO_6 connected in a random network *via* apex-, edge-, and face-sharing octahedra have also been observed in the amorphous phases of $BaTiO_3$ (ref. 48 and 49) and $SrTiO_3$.[50]

The expected bulk density of amorphous $BaZrO_3$ phases was reported to be in the range of 82–84% of the crystalline density (ref. 51). In our ML-driven NPT simulation of amorphous $BaZrS_3$, the observed density was 3.94 g cm^{-3}, which is roughly 92.5% of the crystalline density, in qualitative agreement with the lower density observed in the related oxide compound. The average bond length of the 6-fold-coordinated Zr atoms in the amorphous phase, with a value of 2.58 Å, is similar to the expected bond length of 2.55–2.56 Å in the ZrS_6 octahedra of the crystalline phase. In the case of the 5- and 7-coordinated Zr–S environments, the bonds are slightly compressed or elongated compared to the crystalline phase (2.50 Å and 2.66 Å, respectively). The Ba–S bonds vary in length within a similar range to that observed in the crystalline phase, around 3.0–3.4 Å; however, as also observed in the partial RDF peak, the coordination can vary more than in the case of the Zr environment.

The relationship between the geometry of the cation environments in the amorphous and crystalline forms of $BaZrS_3$ can be observed in the angle distribution function (ADF) plots for the S–Zr–S and S–Ba–S bond angles (Fig. 2e and f, respectively). The main ADF peaks in the case of Zr are clearly distributed around the values expected for the octahedral crystalline environment, specifically 90° angles between equatorial and axial Zr–S bonds, 180° between the axial Zr–S bonds, and about 150° for the bonds connecting the octahedra in the orthorhombic crystal structure. The distribution is harder to assess for the S–Ba–S bond angles, both due to the wider range of bond angles observed for higher coordination geometries, and due to the greater variety of CNs present in the amorphous phase. Given the undercoordination of the cations compared to the crystalline phase, there is an observed increase in close S–S contacts in the amorphous structure (Fig. S4).

We note that structural properties of the amorphous phase are of interest not only to provide insight into the atomic environments found in the as-deposited precursor to the polycrystalline material, but additionally to reveal possible structure–property relationships in amorphous or surface-amorphised perovskites which have shown good performance as electro- or photoelectrocatalysts in the case of oxide perovskite materials.[52] The presence of dangling bonds from undercoordinated atoms has been suggested as a possible reason for the efficiency of perovskites with amorphised surfaces during electrochemical processes.[52] Further experimental work could determine whether electrocatalytic surface reconstruction occurs in $BaZrS_3$ and therefore whether this could explain the performance of the material in electrocatalysed oxygen or hydrogen evolution reactions, where the reactant binds to undercoordinated Zr sites.[53,54] In the long run, knowledge of the coordination environments in the amorphous precursor phase might aid in the development of higher-quality crystallised materials with fewer defects.

Grain boundaries

Crucially, the ML potential also allows the modelling of interfaces, such as those observed at grain boundaries (GBs). To date, the atomistic structure of the boundary region and the associated GB formation energies in $BaZrS_3$ have remained unexplored. Therefore, we use the misorientation angles predicted by coincidence site lattice (CSL) theory applied to an orthorhombic system[56] to construct models of the anticipated stable structures for higher misorientation angles around the (001) axis, in the lattice plane defined by the *a* and *b* axes. Additionally, we test the potential's performance on low-angle GBs, choosing misorientation angles up to 10°. The MLIP was used to relax the positions of atoms in the boundary region to obtain structures with GB energies in good agreement with DFT results (Fig. 3a). The GB formation energy, E_{GBf}, was calculated as

$$E_{GBf} = \frac{E_{GB} - E_{bulk}}{2A},$$

where E_{GB} is the energy of the modelled system containing two identical grain boundaries, E_{bulk} is the energy of the same supercell without grain boundaries (the bulk single-crystal system), and A is the area of the grain-boundary region.

For most GB systems studied, the MLIP's prediction has a minimum accuracy of 0.10 J m^{-2} relative to DFT values. This is within 0.07 J m^{-2} of the errors obtained in a study by Ito *et al.* for an ML potential specialised on grain-boundary structures.[57] The maximum error is obtained for the GB with a misorientation angle of 1.5°, which is 0.18 J m^{-2} off the DFT value. Overall, the potential is also able to capture the relative stabilities of the different GB systems, with the exception of the GB with an angle

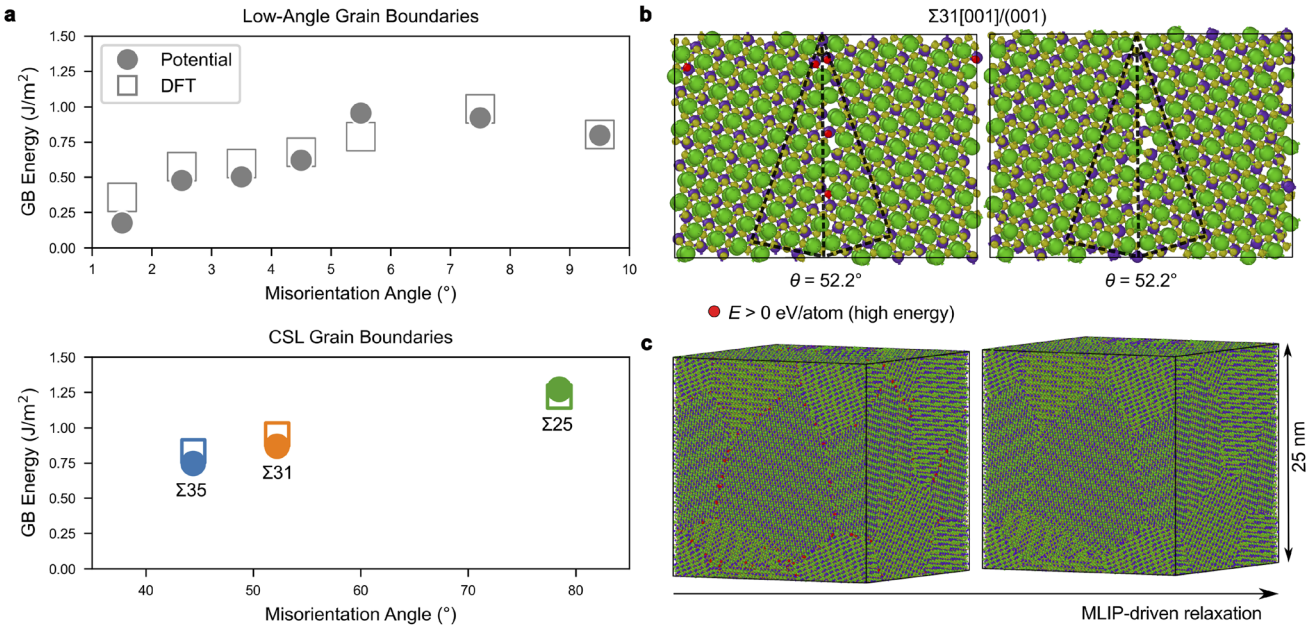

Fig. 3 Grain boundaries. (a) The grain-boundary (GB) energy prediction for relaxed GB structures of $BaZrS_3$ using the MLIP model compared with the single-point energies predicted by DFT at the PBEsol level of theory. (b) Optimisation of the Σ31[001]/(001) GB geometry, where high-energy atoms with energies higher than 1 eV per atom (red) are relaxed to a lower-energy structure using the ML potential. (c) Relaxation of a 615 214-atom 3D polycrystalline structure with six randomly-oriented grains.

of 5.5°, where the energy prediction incorrectly identifies it as higher in energy than its relative ground-truth value. These errors could be addressed by adding the structures of interest to the dataset or changing the weightings of relevant configuration types such that the potential is more specialised on the region of interest (Fig. S5). As noted, it has been previously shown that ML potentials can be specifically trained on different GB structures.[57,58] However, such targeted training is beyond the scope of the current work.

An example of the successful relaxation of the expected CSL Σ31[001] per (001) GB system is shown in Fig. 3b. The relaxation of an approximately 1000-atom GB structure, at the limit of what is achievable using DFT methods, is achieved within seconds using the MLIP, while systems of up to hundreds of thousands of atoms can be relaxed on the timescale of minutes on a 128-core CPU compute node.

Polycrystalline structures

As discussed above, MLIPs allow the relaxation of much larger and more realistic systems which are inaccessible to *ab initio* methods. Going beyond a simple grain-boundary system, we created 3D polycrystalline models of up to 600 000 atoms (Fig. 3c). The polycrystalline unit cell with 6 randomly-oriented grains was generated by Voronoi tessellation in Atomsk[59] and relaxed with the potential trained on the full dataset, to avoid unphysical close contacts between atoms (see SI for details). The extrapolation grade based on the D-optimality algorithm is an established method for measuring the uncertainty of an ACE ML potential in a particular region of the configurational space being modelled.[57,60] We found that the MLIP is able to relax the polycrystalline unit cell, resulting in a structure with extrapolation grades up to a maximum value of 2, which is considered accurate, as described in ref. 60 (see Fig. S6). This result suggests that the potential can predict the atomistic structure with relatively high certainty, including in the boundary region. Given the potential's performance for modelling the amorphous phase, its transferability to the polycrystalline system is unsurprising, as we found that in the polycrystalline relaxed structure, the coordination around the Zr atom is similar to that found in the amorphous phase, but with a shorter Zr–S bond length ($CN_{Zr} = 5.87$ and $d_{Zr-S} = 2.54$ Å; see also Fig. S7). Thus, the undercoordinated Zr environment suggests that the relaxed grain boundaries are S-deficient (Fig. S8).

While the potential can efficiently relax systems exceeding 600 000 atoms, calculating scattering patterns of discrete structures remains computationally demanding,[61] and so we focused our study on polycrystalline unit cells with a side length of 20 nm (≈310 000 atoms). Fig. 4 illustrates the effect that polycrystallinity has on the simulated XRD pattern and pair distribution function (PDF) of $BaZrS_3$, which were generated using the DebyeCalculator package (ref. 61). Going from a single-crystal model shown in Fig. 4a to a polycrystalline model with 8 grains and 50 grains in Fig. 4b and c, respectively, the peaks in the simulated XRD patterns broaden—an effect primarily attributable to size variation: the size of the grains decreases when their number increases in a simulation box of constant size, as expected from applying the Scherrer equation in XRD experiments. This broadening causes the smaller, secondary peaks observed in the crystalline pattern to gradually become unresolved for the systems with a larger number of grains; the effect is detailed in Fig. S9 and S10, which compare

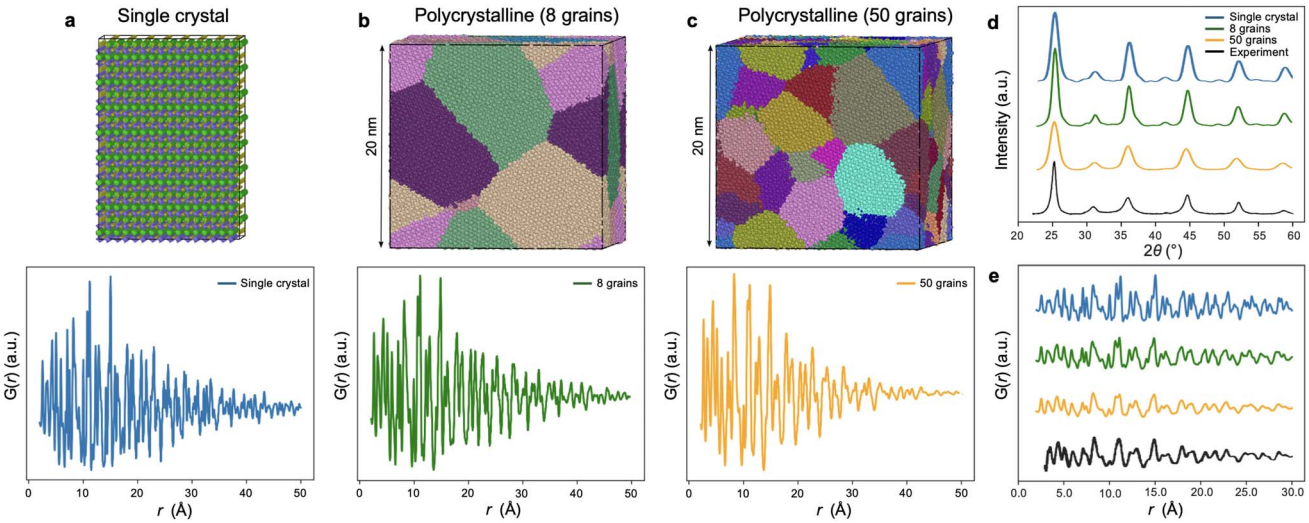

Fig. 4 Effect of grain size on the X-ray diffraction (XRD) patterns and pair distribution functions (PDFs) of crystalline $BaZrS_3$. (a) A single-crystal 10 240-atom model (top) and its simulated PDF [$G(r)$; bottom]. (b) A polycrystalline model with a simulation box side length of 20 nm containing 8 grains (313 449 atoms). (c) A polycrystalline model with a simulation box side length of 20 nm containing 50 grains (312 370 atoms). (d and e) Comparison between the XRD patterns and PDFs, respectively, for the simulated models and the experimental data from ref. 55. For simulating the scattering data, we employed the DebyeCalculator default parameters, except for the intensity data shown in panel (d) where Q_{min} and Q_{max} were set to 1.61 $Å^{-1}$ and 5.25 $Å^{-1}$, respectively, assuming the use of K-α1 radiation for the conversion to 2θ.

the simulated partial XRD patterns of single-crystal and polycrystalline $BaZrS_3$ models, showing the effect of increasing the number of grains even further.

In the PDF, size effects manifest as a more rapid dampening of the signal at high r values in the model with a large number of grains relative to the polycrystalline model with fewer, larger grains and to that of the single-crystal model. Additionally, the local structure of each model varies slightly, presumably due to differences in the grain-boundary fraction. In Fig. 4d and e, the simulated data are compared to experimental data from ref. 55. The experimental XRD and PDF patterns are described best by the polycrystalline 50-grains model, consistent with the nanoparticulate nature of the material described in ref. 55, which presents a solution-phase synthesis of plate-like, aggregated $BaZrS_3$ nanoparticles. Similar results were also reported in ref. 62: a low-temperature synthesis of $BaZrS_3$ nanoparticles with grain sizes of 3–5 nm. A related trend in the experimental XRD patterns can be observed when samples are progressively annealed at higher temperatures, whereby the crystal size increases.[3]

Conclusions and outlook

Our study has introduced an ML-based interatomic potential model for disordered and polycrystalline $BaZrS_3$ and shown how it can be used to simulate multiple types of structures that are relevant for this emerging functional material. We addressed questions related to the structural and physical properties of $BaZrS_3$ across different stages of the experiment, from the deposition of the amorphous precursor phase to the study of polycrystalline systems with different grain sizes. These simulations can be directly compared with experimental results, such as XRD patterns, PDFs, or local bonding information obtained from EXAFS. Furthermore, the potential is capable of relaxing grain-boundary structures and estimating their formation energies at modest computational cost for large-scale structural models.

The present study is an early step towards the realistic modelling of thin-film photovoltaic materials. In the future, as an extension to this work, the modelling of surface structures, or amorphised surfaces, could provide further insights into the material's functionality. More generally, our work shows how polycrystalline perovskite materials can now be modelled with atomistic machine learning, having in view the realism that has already been achieved for other technologically relevant systems.[30,63] The present study thus lays the groundwork for such studies involving more complex photovoltaic materials, including mixed-cation and anion perovskites, with hybrid organic–inorganic components, and for starting to approach their modelling at the length scale of real devices.

Conflicts of interest

There are no conflicts to declare.

Data availability

Data supporting this work, including MLIP parameter files, structural models, and analysis code, are available at **https://github.com/BiancaPasca/polycrystalline-BaZrS3**.

Supplementary information: details about training data, model hyperparameters, and simulation protocols. See DOI: **https://doi.org/10.1039/d5ta04536c**.

Acknowledgements

We thank D. F. Thomas du Toit and Y. Zhou for helpful discussions. L.-B.P. acknowledges funding from the EPSRC Centre for Doctoral Training in Inorganic Chemistry for Future Manufacturing (OxICFM), EP/S023828/1. This work was supported by UK Research and Innovation [grant numbers EP/X016188/1 and EP/Z000343/1]. The work presented in this article is supported by Novo Nordisk Foundation grant NNF23OC0081359. We are grateful for computational support from the UK national high performance computing service, ARCHER2, for which access was obtained *via* the UKCP consortium and funded by EPSRC grant ref. EP/X035891/1 (see also ref. 64).

References

1 T. Gupta, D. Ghoshal, A. Yoshimura, S. Basu, P. K. Chow, A. S. Lakhnot, J. Pandey, J. M. Warrender, H. Efstathiadis, A. Soni, E. Osei-Agyemang, G. Balasubramanian, S. Zhang, S.-F. Shi, T.-M. Lu, V. Meunier and N. Koratkar, *Adv. Funct. Mater.*, 2020, **30**, 2001387.
2 Y. Nishigaki, T. Nagai, M. Nishiwaki, T. Aizawa, M. Kozawa, K. Hanzawa, Y. Kato, H. Sai, H. Hiramatsu, H. Hosono and H. Fujiwara, *Sol. RRL*, 2020, **4**, 1900555.
3 C. Comparotto, A. Davydova, T. Ericson, L. Riekehr, M. V. Moro, T. Kubart and J. Scragg, *ACS Appl. Energy Mater.*, 2020, **3**, 2762–2770.
4 Z. Yu, X. Wei, Y. Zheng, H. Hui, M. Bian, S. Dhole, J.-H. Seo, Y.-Y. Sun, Q. Jia, S. Zhang, S. Yang and H. Zeng, *Nano Energy*, 2021, **85**, 105959.
5 N. Thakur, K. A. Aly, M. Mohery, M. M. Ebrahium, P. Kumar and P. Sharma, *J. Alloys Compd.*, 2023, **957**, 170457.
6 X. Wei, H. Hui, C. Zhao, C. Deng, M. Han, Z. Yu, A. Sheng, P. Roy, A. Chen, J. Lin, D. F. Watson, Y.-Y. Sun, T. Thomay, S. Yang, Q. Jia, S. Zhang and H. Zeng, *Nano Energy*, 2020, **68**, 104317.
7 J. A. Márquez, M. Rusu, H. Hempel, I. Y. Ahmet, M. Kölbach, I. Simsek, L. Choubrac, G. Gurieva, R. Gunder, S. Schorr and T. Unold, *J. Phys. Chem. Lett.*, 2021, **12**, 2148–2153.
8 M. Surendran, H. Chen, B. Zhao, A. S. Thind, S. Singh, T. Orvis, H. Zhao, J.-K. Han, H. Htoon, M. Kawasaki, R. Mishra and J. Ravichandran, *Chem. Mater.*, 2021, **33**, 7457–7464.
9 I. Sadeghi, K. Ye, M. Xu, Y. Li, J. M. LeBeau and R. Jaramillo, *Adv. Funct. Mater.*, 2021, **31**, 2105563.
10 S. Mukherjee, S. Riva, C. Comparotto, F. O. L. Johansson, G. J. Man, D. Phuyal, K. A. Simonov, J. Just, K. Klementiev, S. M. Butorin, J. J. S. Scragg and H. Rensmo, *ACS Appl. Energy Mater.*, 2023, **6**, 11642–11653.
11 S. P. Ramanandan, A. Giunto, E. Z. Stutz, B. Reynier, I. T. F. M. Lefevre, M. Rusu, S. Schorr, T. Unold, A. F. I. Morral, J. A. Márquez and M. Dimitrievska, *J. Phys.: Energy*, 2023, **5**, 014013.
12 S. Riva, S. Mukherjee, S. M. Butorin, C. Comparotto, G. Aggarwal, E. Johannesson, M. Abdel-Hafiez, J. Scragg and H. Rensmo, *ACS Appl. Mater. Interfaces*, 2024, **16**, 40210–40221.
13 P. Kayastha, G. Longo and L. D. Whalley, *ACS Appl. Energy Mater.*, 2024, **7**, 11326–11333.
14 E. Osei-Agyemang and K. Rajan, *J. Phys. Chem. C*, 2023, **127**, 17146–17156.
15 J. Behler, *Angew. Chem., Int. Ed.*, 2017, **56**, 12828–12840.
16 V. L. Deringer, M. A. Caro and G. Csányi, *Adv. Mater.*, 2019, **31**, 1902765.
17 P. Friederich, F. Häse, J. Proppe and A. Aspuru-Guzik, *Nat. Mater.*, 2021, **20**, 750–761.
18 R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse and M. Bokdam, *Phys. Rev. Lett.*, 2019, **122**, 225701.
19 T. Braeckevelt, R. Goeminne, S. Vandenhaute, S. Borgmans, T. Verstraelen, J. A. Steele, M. B. J. Roeffaers, J. Hofkens, S. M. J. Rogge and V. Van Speybroeck, *Chem. Mater.*, 2022, **34**, 8561–8576.
20 E. Fransson, J. Wiktor and P. Erhart, *J. Phys. Chem. C*, 2023, **127**, 13773–13781.
21 W. J. Baldwin, X. Liang, J. Klarbring, M. Dubajic, D. Dell'Angelo, C. Sutton, C. Caddeo, S. D. Stranks, A. Mattoni, A. Walsh and G. Csányi, *Small*, 2024, **20**, 2303565.
22 N. Jaykhedkar, R. Bystrický, M. Sýkora and T. Bučko, *Inorg. Chem.*, 2023, **62**, 12480–12492.
23 P. Kayastha, E. Fransson, P. Erhart and L. Whalley, *J. Phys. Chem. Lett.*, 2025, **16**, 2064–2071.
24 A. Stukowski, *Modell. Simul. Mater. Sci. Eng.*, 2009, **18**, 015012.
25 R. Desai, S. Agarwal, K. C. Vincent, A. Strachan, R. Agrawal and A. Mannodi-Kanakkithodi, *J. Phys. Chem. C*, 2025, **129**, 7967–7976.
26 I. Mosquera-Lois, S. R. Kavanagh, A. Walsh and D. O. Scanlon, *npj Comput. Mater.*, 2023, **9**, 25.
27 Z. El-Machachi, D. Frantzov, A. Nijamudheen, T. Zarrouk, M. A. Caro and V. L. Deringer, *Angew. Chem., Int. Ed.*, 2024, **63**, e202410088.
28 L. McInnes, J. Healy, N. Saul and L. Großberger, *J. Open Source Softw.*, 2018, **3**, 861.
29 A. P. Bartók, R. Kondor and G. Csányi, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**, 184115.
30 Y. Zhou, W. Zhang, E. Ma and V. L. Deringer, *Nat. Electron.*, 2023, **6**, 746–754.
31 J. D. Morrow, C. Ugwumadu, D. A. Drabold, S. R. Elliott, A. L. Goodwin and V. L. Deringer, *Angew. Chem., Int. Ed.*, 2024, **63**, e202403842.
32 L. A. M. Rosset, D. A. Drabold and V. L. Deringer, *Nat. Commun.*, 2025, **16**, 2360.
33 R. Drautz, *Phys. Rev. B*, 2019, **99**, 014104.
34 Y. Lysogorskiy, C. van der Oord, A. Bochkarev, S. Menon, M. Rinaldi, T. Hammerschmidt, M. Mrovec, A. Thompson, G. Csányi, C. Ortner and R. Drautz, *npj Comput. Mater.*, 2021, **7**, 97.
35 V. L. Deringer, C. J. Pickard and G. Csányi, *Phys. Rev. Lett.*, 2018, **120**, 156001.
36 N. Bernstein, G. Csányi and V. L. Deringer, *npj Comput. Mater.*, 2019, **5**, 99.

37 C. Ben Mahmoud, J. L. A. Gardner and V. L. Deringer, *Nat. Comput. Sci.*, 2024, **4**, 384–387.
38 V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti and G. Csányi, *Chem. Rev.*, 2021, **121**, 10073–10141.
39 M. Kulichenko, B. Nebgen, N. Lubbers, J. S. Smith, K. Barros, A. E. A. Allen, A. Habib, E. Shinkle, N. Fedik, Y. W. Li, R. A. Messerly and S. Tretiak, *Chem. Rev.*, 2024, **124**, 13681–13714.
40 C. J. Pickard and R. J. Needs, *Phys. Rev. Lett.*, 2006, **97**, 045504.
41 C. J. Pickard and R. J. Needs, *J. Phys.: Condens. Matter*, 2011, **23**, 053201.
42 G. Schusteritsch and C. J. Pickard, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2014, **90**, 035424.
43 A. P. Bartók, M. C. Payne, R. Kondor and G. Csányi, *Phys. Rev. Lett.*, 2010, **104**, 136403.
44 Y. Zhou, D. F. Thomas du Toit, S. R. Elliott, W. Zhang and V. L. Deringer, Full-cycle device-scale simulations of memory materials with a tailored atomic-cluster-expansion potential, *arXiv*, 2025, preprint, arXiv:2502.08393, DOI: **10.48550/arXiv.2502.08393**.
45 J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, *Phys. Rev. Lett.*, 2008, **100**, 136406.
46 E. Wachtel and I. Lubomirsky, *Adv. Mater.*, 2010, **22**, 2485–2493.
47 S. A. Rigter, X. L. Quinn, R. E. Kumar, D. P. Fenning, P. Massonnet, S. R. Ellis, R. M. A. Heeren, K. L. Svane, A. Walsh and E. C. Garnett, *Adv. Funct. Mater.*, 2021, **31**, 2010330.
48 A. I. Frenkel, Y. Feldman, V. Lyahovitskaya, E. Wachtel and I. Lubomirsky, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2005, **71**, 024116.
49 D. Ehre, H. Cohen, V. Lyahovitskaya, A. Tagantsev and I. Lubomirsky, *Adv. Funct. Mater.*, 2007, **17**, 1204–1208.
50 A. I. Frenkel, D. Ehre, V. Lyahovitskaya, L. Kanner, E. Wachtel and I. Lubomirsky, *Phys. Rev. Lett.*, 2007, **99**, 215502.
51 V. Lyahovitskaya, Y. Feldman, I. Zon, E. Wachtel, K. Gartsman, A. K. Tagantsev and I. Lubomirsky, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2005, **71**, 094205.
52 H. Sun, X. Xu, G. Chen and Z. Shao, *Carbon Energy*, 2024, **6**, 595.
53 N. Yasmin, M. Safdar, G. Ali, H. M. Khan and M. Mirza, *J. Phys. Chem. Solids*, 2023, **172**, 111056.
54 N. Humphrey, A. Tsung, S. Singh, A. Irshad, B. Zhao, S. Narayan, J. Ravichandran and S. Mallikarjun Sharada, *ChemPhysChem*, 2024, **25**, e202300953.
55 D. Zilevu, O. O. Parks and S. E. Creutz, *Chem. Commun.*, 2022, **58**, 10512–10515.
56 A. H. King and A. Singh, *J. Appl. Phys.*, 1993, **74**, 4627–4630.
57 K. Ito, T. Yokoi, K. Hyodo and H. Mori, *npj Comput. Mater.*, 2024, **10**, 255.
58 T. Yokoi, K. Adachi, S. Iwase and K. Matsunaga, *Phys. Chem. Chem. Phys.*, 2022, **24**, 1620–1629.
59 P. Hirel, *Comput. Phys. Commun.*, 2015, **197**, 212–219.
60 Y. Lysogorskiy, A. Bochkarev, M. Mrovec and R. Drautz, *Phys. Rev. Mater.*, 2023, **7**, 043801.
61 F. L. Johansen, A. S. Anker, U. Friis-Jensen, E. B. Dam, K. M. ø Jensen and R. Selvan, *J. Open Source Softw.*, 2024, **9**, 6024.
62 R. Yang, A. D. Jess, C. Fai and C. J. Hages, *J. Am. Chem. Soc.*, 2022, **144**, 15928–15931.
63 B. W. J. Chen and M. Mavrikakis, *Nat. Chem. Eng.*, 2025, **2**, 181–197.
64 G. Beckett, J. Beech-Brandt, K. Leach, Z. Payne, A. Simpson, L. Smith, A. Turner and A. Whiting, *ARCHER2 Service Description*, Zenodo, 2024, DOI: **10.5281/zenodo.14507040**.