Electronic Supplementary Material (ESI) for ChemComm.

Supplementary Information: How fast do defects migrate in halide perovskites: insights from on-the-fly machine-learned force fields

Mike Pols, Victor Brouwers, Sofía Calero, and Shuxia Tao*

Materials Simulation & Modelling, Department of Applied Physics, Eindhoven University of Technology, 5600 MB, Eindhoven, The Netherlands

E-mail: s.x.tao@tue.nl

Contents

Supporting Notes

1 Computational methodology

1.1 Gaussian process models

To obtain a potential energy surface (PES) using sparse Gaussian process (SGP) regression, the total energy E of the system is decomposed into M local energies $\varepsilon(\rho_i)$,[1] so that

$$E = \sum_{i}^{M} \varepsilon(\rho_i) = \sum_{i}^{M} \sum_{s \in S}^{n_s} \alpha_s k(\boldsymbol{d}_s, \boldsymbol{d}_i). \tag{S1}$$

ρ_i describes the atom-centered local environment as a function of atomic species and interatomic distances within a cutoff radius R_{cut}. We use the multielement atomic cluster expansion[2] to describe the local environments with descriptors $\boldsymbol{d}$ that are invariant under permutations, translations, and rotations. To evaluate forces and stresses, the gradients of the descriptor are calculated with respect to the Cartesian coordinates of each neighbour in the local environment.[1] Both the sparse weights α_s, collected in the sparse vector $\boldsymbol{\alpha}_s$, and the kernel function $k(\boldsymbol{d}_s, \boldsymbol{d}_i)$ must be optimized to obtain a correct description of the PES. The descriptors $\boldsymbol{d}_i$ are compared with the n_s reference descriptors $\boldsymbol{d}_s$ in the sparse data set S, which is obtained from ab-initio calculations. The kernel function is used as a similarity measure as

$$k(\boldsymbol{d}, \boldsymbol{d'}) = \gamma^2 \left(\frac{\boldsymbol{d} \cdot \boldsymbol{d'}}{|\boldsymbol{d}||\boldsymbol{d'}|} \right)^{\xi}. \tag{S2}$$

Here, γ quantifies the noise in local energy and the integer ξ determines the sharpness or correlation length of the kernel. We set $\xi = 2$. The sparse vector $\boldsymbol{\alpha}_s$ and the associated variances $\mathcal{V}[\varepsilon_*]$ of the predictions are computed using Bayes' theorem[1] as

$$\boldsymbol{\alpha}_S = \left[K_{SF} (\sigma_{\text{n}}^2 I)^{-1} K_{FS} + K_{SS} \right] K_{SF} \boldsymbol{y}, \tag{S3}$$

$$\mathcal{V}[\varepsilon] = k(\boldsymbol{d}, \boldsymbol{d}) - \boldsymbol{k}(\boldsymbol{d})^T (K + \sigma_n^2 I)^{-1} \boldsymbol{k}(\boldsymbol{d}). \tag{S4}$$

The sparse vector $\boldsymbol{\alpha}_s$ is determined from aggregated kernels K_{AB} where A and B can denote either the full data set F or the sparse data set S. A benefit is that for smaller sparse data sets S, the loop in eqn. S1 is performed across fewer data points, making its evaluation less demanding. $\sigma_{\mathrm{n}}^2 I$ is a diagonal matrix that quantifies the expected noise in the forces (σ_{F}), energies (σ_E), and stresses (σ_{S}). $\boldsymbol{y}$ collects the data of the reference ab-initio calculations. The variance is used as an estimate of the Bayesian error during the on-the-fly training. Whenever the estimated error exceeds a threshold, a reference calculation is performed of which the full structure is added to the full data set F, and only the local environments with an uncertainty above a sparsification threshold are added to the sparse set S. The SGP model is retrained by optimizing a log marginal likelihood function in the deterministic training conditional (DTC) approximation[1,3]

$$\begin{aligned}\mathcal{L} = &-\frac{1}{2}\boldsymbol{y}^T\left(K_{SF}K_{SS}^{-1}K_{FS}+\sigma_{\mathrm{n}}^2 I\right)^{-1}\boldsymbol{y}\\ &-\frac{1}{2}\log\left|K_{SF}K_{SS}^{-1}K_{FS}+\sigma_{\mathrm{n}}^2 I\right|\\ &-\frac{n}{2}\log 2\pi,\end{aligned} \tag{S5}$$

where the first term rewards the quality of the fit, the second term punishes model complexity, and the third term is a normalization constant with n the total number of labels in the training set. As such, the log marginal likelihood results in a trade-off between model accuracy and complexity to prevent model overfitting. Using the L-BFGS algorithm[4,5] the log marginal likelihood is optimized with the hyperparameters (γ, σ_{F}, σ_{E}, and σ_{S}), as a result of which the SGP model is continuously improved until the on-the-fly training ends, resulting in a machine-learned force field (MLFF). In line with earlier works in literature,[1,6] we mapped the final MLFFs onto much faster polynomial models, that are quadratic in the descriptor ($\xi_{\mathrm{map}=2}$), for use in large-scale dynamical simulations."

1.2 Density functional theory calculations

The ab-initio density functional theory (DFT) calculations used to train the MLFF were performed with the projector augmented wave (PAW) method[7] as implemented in the Vienna Ab-Initio Simulation Package (VASP).[8–10] The Perdew, Burke and Ernzerhof (PBE) functional was used to model the exchange-correlation interaction between the electrons.[11] The outermost electrons of Cs ($5s^2 5p^6 6s^1$); Pb ($6s^2 6p^2$), I ($5s^2 5p^5$) and Br ($4s^2 4p^5$) were treated as valence electrons and an energy cutoff of 300 eV was used for the plane wave basis set. Long-range dispersive interactions were explicitly accounted for by the DFT-D3(BJ) dispersion correction.[12] In each ionic step the self-consistency cycle was repeated until a convergence of 1×10^{-2} meV was reached in the energy. The Brillouin zones of the structures were sampled using Γ-centered k-points meshes.[13] For small perovskite structures (less than 20 formula units), the reciprocal space was sampled using a $2 \times 2 \times 2$ k-mesh. Larger perovskite structures (more than 20 formula units), were sampled with only the Γ-point ($1 \times 1 \times 1$ k-mesh).

1.3 Large-scale molecular dynamics simulations

Large-scale molecular dynamics (MD) simulations were run with the mapped models in LAMMPS.[14] Depending on the material property of interest, simulations were run in an NVT-ensemble (e.g. defect migration) and NpT-ensemble (e.g. phase transition). In all simulations, the temperature was maintained at the desired target temperature using a Nosé-Hoover chains (NHC) thermostat.[15,16] For the NpT-ensemble, lattice fluctuations were accounted for by maintaining an atmospheric pressure (1 atm) using a Martyna-Tobias-Klein (MTK) barostat.[17] As a result of the relatively large mass ($>$ 79.904 a.m.u) of all atomic species involved, a simulation time step Δt in the range of 1-3 fs was used. Details of specific simulations can be found in the respective sections of the Supporting Information.

2 Phase transition models

2.1 On-the-fly machine learning hyperparameters

The on-the-fly MLFF training routine requires a certain set of parameters to be defined prior to the training cycles. These so-called hyperparameters used during the on-the-fly training of the MLFF models in this work are shown in Table S1. We note that γ, σ_{F}, σ_{E}, and σ_{S} are kernel hyperparameters are changed during the log marginal likelihood optimization in the on-the-fly training.

Table S1: Hyperparameters of the on-the-fly MLFF training.

Hyperparameter	Value
N_{rad}	8
l_{max}	3
ξ	2
R_{cut}	6.0 Å
γ	2.0 eV
σ_{E}	1.0 meV/atom
σ_{F}	1.0 eV Å^{-1}
σ_{S}	1.0×10^8 Pa

2.2 Phase transition model training

We trained the MLFF models for the different bulk phases of inorganic $CsPbX_3$ perovskites. For the cubic, tetragonal and orthorhombic phases we used $2 \times 2 \times 2$, $\sqrt{2} \times \sqrt{2} \times 2$ and $2 \times 2 \times 1$ supercells for the on-the-fly MLFF training (Figure S1). All simulations were done in an NpT-ensemble with a simulation time step of $\Delta t = 3$ fs at atmospheric pressure. In line with the training procedure used by Jinnouchi et al.[18], we started the training from the high-temperature cubic phase, so that during the initial training iteration a large portion of the phase-space would be sampled as a result of the increased dynamics at elevated temperatures. The MLFF was subsequently trained against the low-temperature phases (tetragonal and orthorhombic) of the inorganic perovskites, for which the thresholds (σ_{th} and $\sigma_{\mathrm{th,s}}$) were decreased to allow for the incorporation of new data into the training set. A

complete overview of the training cycles in the on-the-fly MLFF training is given in Table S2.

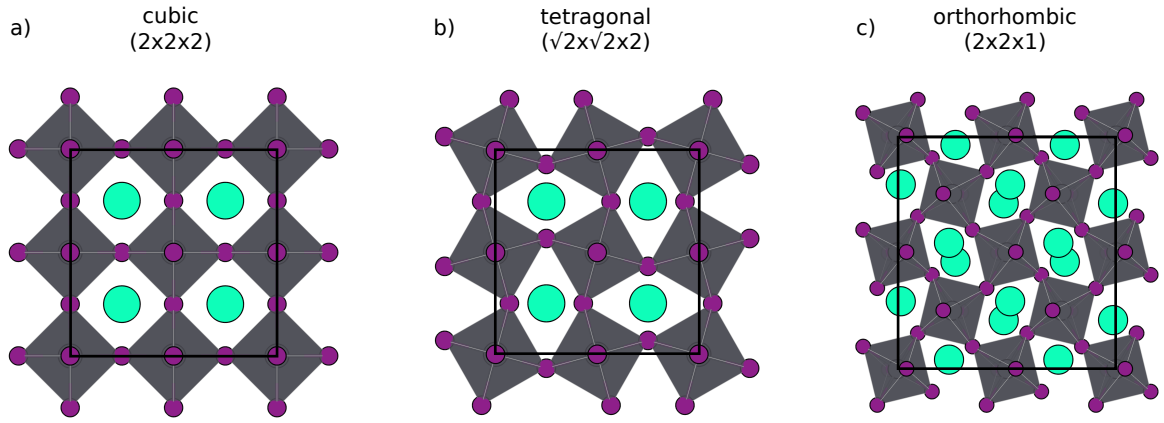

Figure S1: Structural models used during the training of the phase transition models for $CsPbX_3$ perovskites. (a) $2 \times 2 \times 2$ cubic $CsPbI_3$. (b) $\sqrt{2} \times \sqrt{2} \times 2$ tetragonal $CsPbI_3$. (c) $2 \times 2 \times 1$ orthorhombic $CsPbI_3$. Unit cells are indicated in black.

Table S2: On-the-fly machine learning training cycles for inorganic halide perovskite ($CsPbX_3$) MLFFs for bulk phase transitions between the black phases.

Material	Iteration	Phase	T (K)	N_{steps} (-)	σ_{th} (eV/Å)	$\sigma_{\text{th,s}}$ (eV/Å)
$CsPbI_3$	1	Cubic	600	2500	0.050	0.040
	2	Tetragonal	470	2500	0.050	0.040
	3		510	2500		
	4	Orthorhombic	200	1500	0.030	0.025
	5		275	1000		
	6		350	2500		
$CsPbBr_3$	1	Cubic	450	2500	0.050	0.035
	2	Tetragonal	370	5000	0.030	0.025
	3	Orthorhombic	200	5000	0.025	0.015
	4		250	5000		
	5		300	5000		

The details of the resulting MLFF models are shown in Table S3. This overview includes the number of training steps N_{train}, the number of local environments N_{local} and the number of sparse local environments N_{sparse} per atomic species.

Table S3: Details of the MLFF models trained against the different bulk phases of inorganic $CsPbX_3$ perovskites.

Perovskite	N_{train} (-)	N_{local} (-)	N_{sparse} (-)
$CsPbI_3$	87	4240	163 (Cs) 62 (Pb) 472 (I)
$CsPbBr_3$	82	4200	173 (Cs) 74 (Pb) 509 (I)

2.3 Phase transition model validation

The MLFF trained against the various bulk phases of $CsPbI_3$ was validated using large-scale molecular dynamics simulations of a model system with 40 atoms (8 formula units). The used model system is shown in Figure S1a. In the simulation, the perovskite system was kept at 700 K and atmospheric pressure (1 atm) in an NpT-ensemble. The simulation was run with a time step $\Delta t = 2$ fs for a total length of 1 ns. Since 50 equally spaced frames were extracted from the simulation, this resulted in a spacing between the frames of 20 ps. The resulting total energies and atomic force components of these frames were compared against ab-initio calculations.

2.4 Bulk moduli of $CsPbX_3$

The reference data for the equations of state was generated with the same settings in VASP as detailed in SI Note 1.2. The reciprocal space of the cubic, tetragonal, and orthorhombic phase of $CsPbI_3$ was sampled with a $10 \times 10 \times 10$, $7 \times 7 \times 10$ and $7 \times 7 \times 5$ k-mesh, respectively. The unit cell volumes are changed with isotropic strain, during which ionic positions are allowed to relax. The bulk moduli are obtained by fitting the Murnaghan equation of state[19] to the data for each of the $CsPbI_3$ phases. We note that the bulk moduli from the MLFF model (Table S4) show a good agreement with the bulk moduli obtained from ab-initio DFT calculations and other predictions for $CsPbI_3$ in literature,[20–23] highlighting

an overall accurate description of the PES of the perovskite.

Table S4: The bulk moduli of the different black phases of $CsPbI_3$ as determined using from ab-initio DFT calculations and the $CsPbI_3$ MLFF model with the Murnaghan equation of state.

Material	Phase	$B_{0,\mathrm{MLFF}}$ (GPa)	$B_{0,\mathrm{DFT}}$ (GPa)
$CsPbI_3$	Cubic	18.2	16.7
	Tetragonal	18.3	14.2
	Orthorhombic	16.0	13.7

2.5 Octahedral inversion barrier

The octahedral inversion barrier of the tetragonal phase of $CsPbI_3$ is calculated with similar VASP settings as detailed in SI Note 1.2 using a $7 \times 7 \times 10$ k-mesh to sample the reciprocal space. Five intermediate geometries were used to determine the inversion barrier of the octahedra using the Climbing Image Nudged Elastic Band (CI-NEB) method[24,25] using a force convergence criterion of $3 \times 10^{-2}\,\mathrm{eV\,\AA^{-1}}$.

3 Phase transition analysis

3.1 Phase transition simulations

We used large-scale molecular dynamics simulations with the MLFFs to simulate the phase transitions of the two inorganic perovskites. A $3 \times 3 \times 3$ orthorhombic supercell, with 540 atoms (108 formal units) was used as the model system. At each target temperature, the simulations were carried out in an NpT-ensemble. Each simulation run made use of a time step $\Delta t = 1\,\mathrm{fs}$ for a total duration of 400 ps, where the first 50 ps of each trajectory were discarded as equilibration stage. A temperature spacing of 50 K was used, a smaller spacing of 25 K was used close to the experimentally reported phase transition temperatures.[26,27] Lattice vectors were obtained by a fitting of Gaussian distributions to the fluctuations in the lattice vectors, the complete procedure of which was outlined in earlier works in literature.[18,28]

3.2 Thermal expansion of $CsPbX_3$

To assess the macroscopic behaviour of the pure bulk perovskites, we compute their volumetric expansion coefficient α_V, as

$$\alpha_\mathrm{V} = \frac{1}{V_0}\frac{\partial V}{\partial T} \tag{S6}$$

with V_0 the reference volume at T= 300 K. An overview of the volumetric expansion coefficients for the $CsPbX_3$ inorganic perovskites is shown in Table S5. We highlight the good overall agreement between experiments and the MLFF models.

Table S5: The volumetric expansion coefficients α_V of the black phases of $CsPbI_3$ and $CsPbBr_3$ as determined with the MLFF models and through experiments.

Material	$\alpha_{\mathrm{V,MLFF}}$ ($\cdot 10^{-4}$ K^{-1})	$\alpha_{\mathrm{V,exp}}$ ($\cdot 10^{-4}$ K^{-1})
$CsPbI_3$	1.35(0.02)	1.53[26], 1.18[29]
$CsPbBr_3$	1.28(0.04)	1.2[27]

3.3 Ionic motion of halide ions

To probe the typical motion of the halide ions and the effect that temperature has on this motion, we analyze the trajectories used for the construction of the phase diagrams of the perovskites. We employ the root mean squared displacement (RMSD) of the halide ions with respect to their average position as a measure for their typical motion in the lattice. For a fair comparison between $CsPbI_3$ and $CsPbBr_3$, we scale the root mean square displacements with the cubic lattice vectors obtained from ab-initio calculations ($a_{CsPbI_3} = 6.29\,\text{Å}$ and $a_{CsPbBr_3} = 5.92\,\text{Å}$). The extent of this halide movement along with its temperature dependence is shown in Figure S2. We note that at elevated temperatures ($> 500\,\text{K}$) iodide ions exhibit more ionic motion than bromide ions in as evidenced by the higher values of the plateaus in Figure S2 for $CsPbI_3$ (0.185) than for $CsPbBr_3$ (0.165).

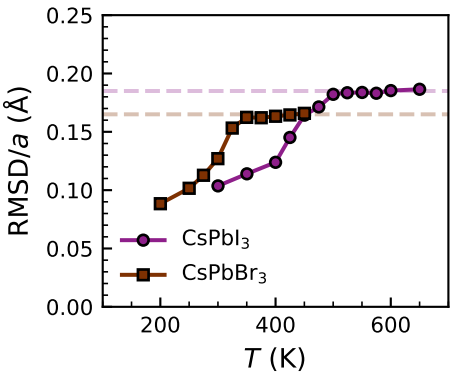

Figure S2: Degree of ionic motion in inorganic perovskite lattices ($CsPbI_3$ and $CsPbBr_3$), quantified by the root mean square displacement (RMSD) of the halide ions scaled by the cubic lattice vectors.

4 Model transferability

4.1 Defect model training

The MLFF model explicitly trained against the migration of a V_I in $CsPbI_3$ was trained against a cubic supercell with one point defect. We made use of a $2 \times 2 \times 2$ cubic supercell with one V_I defect (39 atoms; 8 formula units with one V_I). During the on-the-fly training in an NpT-ensemble, the temperature and pressure were maintained at 600 K and 1 atm, respectively, with a simulation time step of $\Delta t = 3$ fs. The threshold and sparse threshold hyperparameters were set to $\sigma_{\text{th}} = 0.050\,\text{eV}\,\text{Å}^{-1}$ and $\sigma_{\text{th,s}} = 0.035\,\text{eV}\,\text{Å}^{-1}$. A total of 10000 on-the-fly training steps (30 ps) were needed to end up with a MLFF with mean absolute errors (MAEs) in the force components below the error threshold of $0.050\,\text{eV}\,\text{Å}^{-1}$. Notably, the vacancy does not move during training.

4.2 Defect validation simulations

For the MLFF models, we validated their performance for the description of a bulk system with an iodine vacancy (V_I). In these simulations, we used a $4 \times 4 \times 4$ cubic supercell with 319 atoms (64 formula units with one V_I) as the model system. In the simulations in an NpT-ensemble, the temperature was kept to 750 K at an atmospheric pressure (1 atm). The simulation was run with a time step of $\Delta t = 2$ fs for a total duration of 2.5 ns. We extracted a total of 100 frames from the simulate and only evaluate the frames in which we identify the V_I defect, this resulted in a total of 59 analyzed frames for which the force components could be compared. An overview of the accuracy of the models is shown in Table S6.

Table S6: R^2 values of the force components for a $CsPbI_3$ system with one V_I defect for different MLFF models. A spatial decomposition of the R^2 values is performed into full structure, local defect region and residual bulk region. A chemical decomposition of the R^2 values is performed to the different chemical species: Cs, Pb and I.

System	Model	Environment	All	Cs	Pb	I
Iodine vacancy (V_I)	Phase transition	Full	0.911	0.908	0.912	0.911
		Bulk	0.914	0.930	0.912	0.913
		Defect	0.843	0.895	0.822	0.857
	Vacancy trained	Full	0.932	0.930	0.932	0.932
		Bulk	0.933	0.951	0.927	0.934
		Defect	0.904	0.941	0.888	0.915

4.3 Defect environment recognition

The identification of the iodine vacancy (V_I) was done based on the local coordination of Pb species. In the pristine bulk of $CsPbI_3$ each Pb species is surrounded by six I atoms, which in the presence of a vacancy reduces to five I atoms. We employ a spherical cutoff distance of $r_{\text{cut}}^{\text{Pb-I}} = 4.5\,\text{Å}$ to determine the I-coordination for each Pb atom. Whenever we identify two neighbouring five-coordinated Pb atoms in our system, we identify the position between those two Pb particles as the V_I defect. A spherical cutoff of $r_{\text{cut}}^{\text{pair}} = 7.5\,\text{Å}$ is used to check if any undercoordinated Pb atoms are close to each other. A schematic overview of the defect recognition in bulk $CsPbI_3$ is shown in Figure S3a. Finally, we consider all atoms that are within a radius of $r_{\text{cut}}^{\text{def}} = 4.5\,\text{Å}$ from the identified defect, to be in the so-called 'defect environment', all atoms further than this cutoff distance are considered in a 'bulk environment'. The distinction between these two environments is illustrated in Figure S3b.

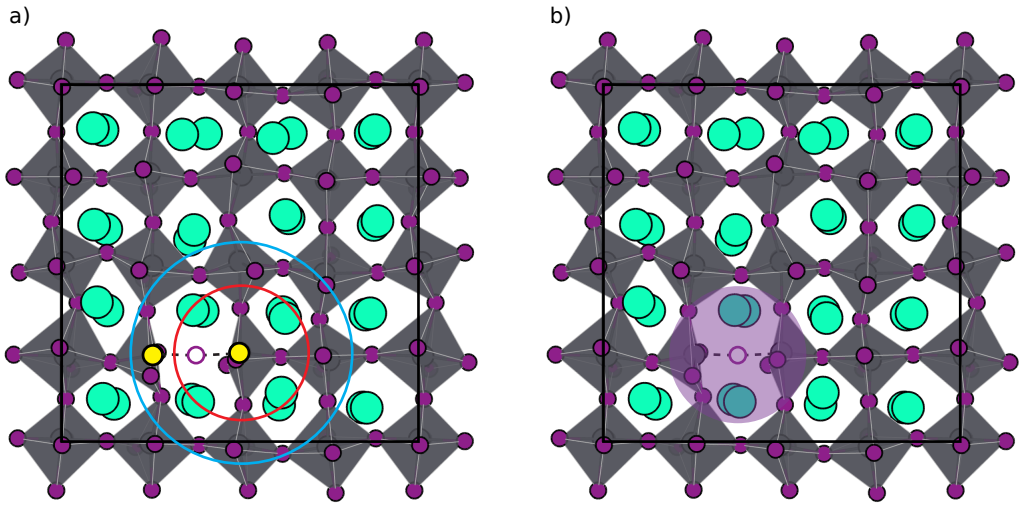

Figure S3: Schematic overview of the recognition of a V_I defect in bulk $CsPbI_3$. (a) Two undercoordinated Pb atoms are highlighted by the yellow circles and the V_I defect is indicated with the purple circle. The red circle highlight the $r_{\text{cut}}^{\text{Pb-I}}$ cutoff distance, used to check which Pb atoms have fewer than six I atoms coordinating it. The blue circle illustrates the $r_{\text{cut}}^{\text{pair}}$ cutoff distance, which is to check if two undercoordinated Pb atoms are close to each other. (b) Distinction between the region of atoms far away from ('bulk environment') and close to ('defect environment') the point defect, with the purple shading highlighting the are close to the defect.

5 Defect migration models and analysis

5.1 Defect migration model training

The different MLFF models were trained in a similar way to that aimed at describing the migration of V_I in $CsPbI_3$ (see SI Note 5). Pure $2 \times 2 \times 2$ cubic supercells either with one vacancy (39 atoms; 8 formula units with one V_X) or one interstitial (41 atoms; 8 formula units with one I_X) point defect were used. For all systems, the on-the-fly training was done in an NpT-ensemble with similar simulation conditions as in the previous run ($T = 600\,\mathrm{K}$, $p = 1\,\mathrm{atm}$, $\Delta t = 3\,\mathrm{fs}$), with identical hyperparameter thresholds (σ_{th}= $0.050\,\mathrm{eV\,\AA^{-1}}$ and $\sigma_{\mathrm{th,s}}$= $0.035\,\mathrm{eV\,\AA^{-1}}$). Since we observe that halide interstitials are considerably more mobile than halide vacancies, we only require 5000 steps (15 ps) for the training of interstitials, opposed to the 10000 steps (30 ps) used for vacancies to obtain MLFFs with MAEs below $0.050\,\mathrm{eV\,\AA^{-1}}$. Notably, interstitials move during the MLFF training, in contrast to vacancies that do not. A complete overview of the obtained models is shown in Table S7.

Table S7: Obtained models from on-the-fly MLFF training on $CsPbI_3$ defects with one V_I or I_I defect and $CsPbBr_3$ with one V_{Br} or I_{Br} defect.

Perovskite	Defect	N_{train} (-)	N_{local} (-)	N_{sparse} (-)
$CsPbI_3$	V_I	101	3939	197 (Cs) 83 (Pb) 514 (I)
	I_I	103	4223	165 (Cs) 70 (Pb) 699 (I)
$CsPbBr_3$	V_{Br}	106	4134	171 (Cs) 111 (Pb) 493 (I)
	I_{Br}	81	3321	158 (Cs) 91 (Pb) 674 (I)

5.2 Defect migration simulations

The defect migration of point defects was studied through large-scale molecular dynamics simulations with MLFFs for perovskite systems with point defects. A $4 \times 4 \times 4$ orthorhombic supercell with 1280 atoms (256 formula units) was used as the model system. Defect were created by either removing (V_X) or adding (I_X) two halide ions from the perovskite. Each model system contained two point defects, resulting in a defect concentration of $3.3 \times 10^{19}\,\mathrm{cm}^{-3}$ for $CsPbI_3$ and $3.9 \times 10^{19}\,\mathrm{cm}^{-3}$ for $CsPbBr_3$, respectively. To prevent a high local concentration a spacing between the defects of at least 20 Å was used. The molecular dynamics simulations were run with a time step of $\Delta t = 2\,\mathrm{fs}$. The model systems were equilibrated to their target temperature and atmospheric pressure (1 atm) in an NpT-ensemble. The length of the equilibration stage was set to $t_{\mathrm{eq}}^{\mathrm{I_X}} = 100\,\mathrm{ps}$ for the systems with I_X defects and $t_{\mathrm{eq}}^{\mathrm{V_X}} = 60\,\mathrm{ps}$ for systems with V_X defects. Longer equilibration times were used for interstitial point defects as a result of their increased mobility in the perovskite lattice. After equilibration, we take the volume, pressure and velocities of the final equilibration frame and use those as a starting point for the 2.5 ns production runs in an NVT-ensemble at the desired target temperature.

5.3 Defect migration analysis

To analyze the migration behaviour of point defects in inorganic halide perovskites we use a method that is similar to previous works.[23,28] We employ the Einstein relation for diffusion[30] to characterize the motion of atomic species with their self-diffusion coefficient D as

$$D = \frac{1}{2d} \lim_{t \to \infty} \frac{d}{dt} \mathrm{MSD}(t) \tag{S7}$$

in which d is the dimensionality of the system ($d = 3$ for bulk perovskites) and MSD (mean square displacement) is the average displacement of a species at time t. If we are interested in the diffusion of point defects in inorganic perovskites, the MSD of these point defects is

defined as

$$\mathrm{MSD}\left(t\right) = \left\langle \frac{1}{N_{\mathrm{def}}} \sum_{i=1}^{N_{\mathrm{X}}} |\vec{r}_i\left(t\right) - \vec{r}_i\left(0\right)|^2 \right\rangle_{t_0} \tag{S8}$$

where N_{def} is the number of point defects, N_{X} the number of halide ions, $\vec{r}_i\left(t\right)$ describes the position of halide ion i after a time t, $\vec{r}_i\left(0\right)$ the starting position of halide ion i and $\langle \cdots \rangle_{t_0}$ showing the window-averaging over different time origins t_0. To speed up the computation of we employ a fast Fourier transform algorithm for the computation of the MSD.[31]

The self-diffusion coefficients obtained with equation S7 and S8 are inherently noisy for shorter trajectories. As such, we repeated every simulation three times in total and averaged the corresponding self-diffusion coefficients for each temperature. Furthermore, we determined the self-diffusion coefficients of systems with halide vacancies at more intermediate temperatures, due to the lower mobility of such defects. Finally, we excluded the simulations of $\mathrm{V_{Br}}$ in $\mathrm{CsPbBr_3}$ at 550 K, since we did not observe any defect migration in our simulations.

References

(1) Vandermause, J.; Xie, Y.; Lim, J. S.; Owen, C. J.; Kozinsky, B. Active Learning of Reactive Bayesian Force Fields Applied to Heterogeneous Catalysis Dynamics of H/Pt. *Nature Communications* **2022**, *13*, 5183.

(2) Drautz, R. Atomic Cluster Expansion for Accurate and Transferable Interatomic Potentials. *Physical Review B* **2019**, *99*, 014104.

(3) Quiñonero-Candela, J.; Rasmussen, C. E. A Unifying View of Sparse Approximate Gaussian Process Regression. *Journal of Machine Learning Research* **2005**, *6*, 1939–1959.

(4) Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal of Scientific Computing* **1995**, *16*, 1190–1208.

(5) Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization. *ACM Transactions on Mathematical Software* **1997**, *23*, 550–560.

(6) Xie, Y.; Vandermause, J.; Sun, L.; Cepellotti, A.; Kozinsky, B. Bayesian Force Fields from Active Learning for Simulation of Inter-Dimensional Transformation of Stanene. *npj Computational Materials* **2021**, *7*, 1–10.

(7) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59*, 1758–1775.

(8) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal–Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49*, 14251–14269.

(9) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Computational Materials Science* **1996**, *6*, 15–50.

(10) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(11) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(12) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *Journal of Computational Chemistry* **2011**, *32*, 1456–1465.

(13) Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-zone Integrations. *Phys. Rev. B* **1976**, *13*, 5188–5192.

(14) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Computer Physics Communications* **2022**, *271*, 108171.

(15) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Molecular Physics* **1984**, *52*, 255–268.

(16) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover Chains: The Canonical Ensemble via Continuous Dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2643.

(17) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

(18) Jinnouchi, R.; Lahnsteiner, J.; Karsai, F.; Kresse, G.; Bokdam, M. Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference. *Phys. Rev. Lett.* **2019**, *122*, 225701.

(19) Murnaghan, F. D. The Compressibility of Media under Extreme Pressures. *Proceedings of the National Academy of Sciences* **1944**, *30*, 244–247.

(20) Roknuzzaman, M.; Ostrikov, K. K.; Wang, H.; Du, A.; Tesfamichael, T. Towards Lead-Free Perovskite Photovoltaics and Optoelectronics by Ab-Initio Simulations. *Scientific Reports* **2017**, *7*, 14025.

(21) Jong, U.-G.; Yu, C.-J.; Kye, Y.-H.; Kim, Y.-S.; Kim, C.-H.; Ri, S.-G. A First-Principles Study on the Chemical Stability of Inorganic Perovskite Solid Solutions Cs1-xRbxPbI3 at Finite Temperature and Pressure. *Journal of Materials Chemistry A* **2018**, *6*, 17994–18002.

(22) Yu, C.-J.; Ko, U.-H.; Hwang, S.-G.; Kim, Y.-S.; Jong, U.-G.; Kye, Y.-H.; Ri, C.-H. First-Principles Study on Material Properties and Stability of Inorganic Halide Perovskite Solid Solutions $\mathrm{CsPb}{({\mathrm{I}}_{1\ensuremath{-}x}{\mathrm{Br}}_{x})}_{3}$. *Physical Review Materials* **2020**, *4*, 045402.

(23) Balestra, S. R. G.; Vicent-Luna, J. M.; Calero, S.; Tao, S.; Anta, J. A. Efficient Modelling of Ion Structure and Dynamics in Inorganic Metal Halide Perovskites. *Journal of Materials Chemistry A* **2020**, *8*, 11824–11836.

(24) Henkelman, G.; Jónsson, H. Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points. *J. Chem. Phys.* **2000**, *113*, 9978–9985.

(25) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113*, 9901–9904.

(26) Marronnier, A.; Roma, G.; Boyer-Richard, S.; Pedesseau, L.; Jancu, J.-M.; Bonnassieux, Y.; Katan, C.; Stoumpos, C. C.; Kanatzidis, M. G.; Even, J. Anharmonicity and Disorder in the Black Phases of Cesium Lead Iodide Used for Stable Inorganic Perovskite Solar Cells. *ACS Nano* **2018**, *12*, 3477–3486.

(27) Stoumpos, C. C.; Malliakas, C. D.; Peters, J. A.; Liu, Z.; Sebastian, M.; Im, J.; Chasapis, T. C.; Wibowo, A. C.; Chung, D. Y.; Freeman, A. J.; Wessels, B. W.; Kanatzidis, M. G. Crystal Growth of the Perovskite Semiconductor CsPbBr3: A New Material for High-Energy Radiation Detection. *Crystal Growth & Design* **2013**, *13*, 2722–2727.

(28) Pols, M.; Vicent-Luna, J. M.; Filot, I.; van Duin, A. C. T.; Tao, S. Atomistic Insights Into the Degradation of Inorganic Halide Perovskite CsPbI3: A Reactive Force Field Molecular Dynamics Study. *The Journal of Physical Chemistry Letters* **2021**, *12*, 5519–5525.

(29) Trots, D. M.; Myagkota, S. V. High-Temperature Structural Evolution of Caesium and Rubidium Triiodoplumbates. *Journal of Physics and Chemistry of Solids* **2008**, *69*, 2520–2526.

(30) Einstein, A. Über Die von Der Molekularkinetischen Theorie Der Wärme Geforderte Bewegung von in Ruhenden Flüssigkeiten Suspendierten Teilchen. *Annalen der Physik* **1905**, *322*, 549–560.

(31) de Buyl, P. Tidynamics: A Tiny Package to Compute the Dynamics of Stochastic and Molecular Simulations. *Journal of Open Source Software* **2018**, *3*, 877.