A Physical Data Fusion Approach to Optimize Compositional Stability of Halide Perovskites

Shijing Sun[1], Armi Tiihonen[1], Felipe Oviedo[1], Zhe Liu[1], Janak Thapa[1], Noor Titan P. Hartono[1], Anuj Goyal[2], Clio Batali[1], Alex Encinas[1], Jason J. Yoo[1], Ruipeng Li[3], Zekun Ren[4], Moungi G. Bawendi[1], Vladan Stevanovic[2], John Fisher III[1], Tonio Buonassisi[1,4]

[1] Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Colorado School of Mines, Golden, CO 80401, USA
[3] Brookhaven National Laboratory, Upton, NY 11970, USA
[4] Singapore-MIT Alliance for Research and Technology, 138602, Singapore

Abstract

Improving the heat–moisture–light stability of organic-inorganic perovskites, a widely studied semiconductor material class, is a critical challenge. Compositional search within multinary perovskites employing brute force synthesis followed by environmental tests are prohibitively expensive in large chemical spaces. To identify the most stable multi-cation lead iodide perovskites containing Cs, formamidinium (FA) and methylammonium (MA), we fuse results from density functional theory (DFT) calculations and *in situ* thin-film degradation test within an end-to-end machine learning (ML) algorithm to inform the compositional optimization of $Cs_xMA_yFA_{1-x-y}PbI_3$. We integrate phase thermodynamics modelling as a *probabilistic constraint* in a Bayesian optimization (BO) loop, which effectively guides the experimental search while considering both structural and environmental stability. After three optimization rounds and only sampling 1.8% of the compositional space, we identify thin-film compositions centred at $Cs_{0.17}MA_{0.03}FA_{0.80}PbI_3$ that achieve a 3x delay in macroscopic degradation onset under elevated temperature, humidity, and light compared with the more complex state-of-the-art $Cs_{0.05}(MA_{0.17}FA_{0.83})_{0.95}Pb(I_{0.83}Br_{0.17})_3$. We find up to 8% of MA can be incorporated into the perovskite structure before stability is significantly compromised. Cs is beneficial at low concentrations, however, beyond 17% is found to contribute to reduced stability**.** Synchrotron-based grazing-incidence wide-angle X-ray scattering (GIWAXS) further validates that the interplay of chemical decomposition and phase separation governs the non-linear instability landscape of this compositional space. We reveal the detrimental role of the δ-$CsPbI_3$ minority phase in accelerating degradation and it can be kinetically suppressed by co-optimising Cs and MA content, providing insights into simplifying perovskite compositions for further environmental stability enhancement. Our approach realizes the effectiveness of ML-enabled data fusion in achieving a holistic, efficient, and physics-informed experimentation for multinary systems, potentially generalisable to materials search in the vast structural and alloyed spaces beyond halide perovskites.

Keywords: compositional materials search, perovskite stability, machine learning, Bayesian optimization.

Main

The environmental instability of halide perovskite materials limits their usage in real-world applications such as solar cells, light emitters, lasers, and photodetectors.[1] Compositional engineering is to date one of the most effective methods to improve perovskites' stability in the presence of heat, humidity and light without sacrificing optoelectronic performance.[2] This fact has led to intensive research within the combinatorial space of $A_xB_yC_{1-x-y}Pb(I_zBr_{1-z})_3$.[3] However, only a fraction of this total compositional space has been experimentally explored, in part due to the large state space (an estimated 5 x 10^5 compositions, with 1% compositional steps). The paucity of resulting degradation data inhibits generalisation of mechanisms across this state space, requiring each compositional search to start their experimental investigations *ab initio*.[4] This challenge is similar to those faced by other materials communities, including heterogeneous catalysts, alloyed battery electrodes, and high-entropy metal alloys for structural and magnetic materials.[5–7] The halide perovskite field and several others require new tools to experimentally navigate these vast composition spaces efficiently to locate optima and to extract generalisable design principles.[8–10]

Recent *in situ* experiments and first-principle calculations independently reveal two thermodynamic origins of instability in $Cs_xMA_yFA_{1-x-y}PbI_3$ multi-cation perovskites.[11,12] The first relates to the chemical instability of organic cations, in particular the more volatile MA cations, which promotes decomposition into PbI_2 with multiple degradation pathways activated in the presence of water, oxygen or light.[11,13] The second arises from the phase instability of multi-cation α-perovskites relative to their single-cation perovskite constituents. The resultant phase separation is reported detrimental to device performance when the black FA- and Cs-based α-perovskites transform to their non-photoactive but more stable polymorphs, δ-$APbI_3$ (*A* = FA or Cs).[14] *A*-site cations play an important role in determining the environmental stability of halide perovskites.[15] There is a widespread need to understand the fundamental composition-dependent behaviour under elevated temperature, humidity, and illumination, however, merging computational and experimental insights on selective compositions into a generalizable optimisation policy over the entire chemical space remains a challenge.[16] State-of-the-art two-step approaches of directly combining theoretical screening prior to shortlisted synthesis as a hard constraint are limited by the inefficiencies from: 1) high-performing theoretical calculations for organic–inorganic systems are often too sparse to guide experimentation, and 2) the discrepancies between the calculation assumptions and the experiments at non-thermodynamic equilibria decreases search accuracy.[17–19] The lack of physics-informed and iterative materials search hinders the ultimate goal of designing perovskite compositions for enhanced environmental stability.

Here, we introduce a data fusion approach to combine the DFT-modelled Gibbs free energy of mixing (ΔG_{mix}) in this study and from literature,[12] and the experimentally quantified time-resolved degradation into a closed-loop machine learning (ML) framework. We apply this framework to optimise state-of-the-art solution-processed thin films suffering from severe heat and moisture-induced degradation within the quasi-ternary space of $Cs_xMA_yFA_{1-x-y}PbI_3$.[3] Under accelerated environmental stress tests with elevated temperature, humidity in air, we identify compositions overperforming the $MAPbI_3$ starting-point by 17x and our state-of-the-art reference composition of $Cs_{0.05}MA_{0.17}FA_{0.83}Pb(I_{0.83}Br_{0.17})_3$ by 3x within three optimisation rounds. The data fusion approach allows us to interpret both compositional and phase contributions in multi-cation perovskites to their environmental (in)stability. DFT here serves as principled guidance in the search algorithm to reduce experiments on compositions forming stable minority phases rather

than the desired multi-cation α-perovskites. With synchrotron-based GIWAXS, we further identify the competing role of MA and Cs in two degradation mechanisms, as MA accelerates chemical decomposition and suppresses phase separation, whereas Cs suppresses the former mechanism but accelerates the latter. In optimised compositions, phase separation which take places over a longer timeframe than chemical decomposition. The observed composition- and time-dependent nature of the two degradation mechanisms explains the non-linearity in the instability landscape and validates the importance of achieving a holistic view of the full compositional space. Our findings highlight the detrimental effects of δ-$CsPbI_3$ minority phase leading to accelerated degradation of α-phase perovskites, which shed light on degradation control of multinary perovskites through a machine-learning driven simultaneous optimization of the three *A*-site cations.

Closed-loop Experimentation Platform Directed by Physics-informed Bayesian Optimization.

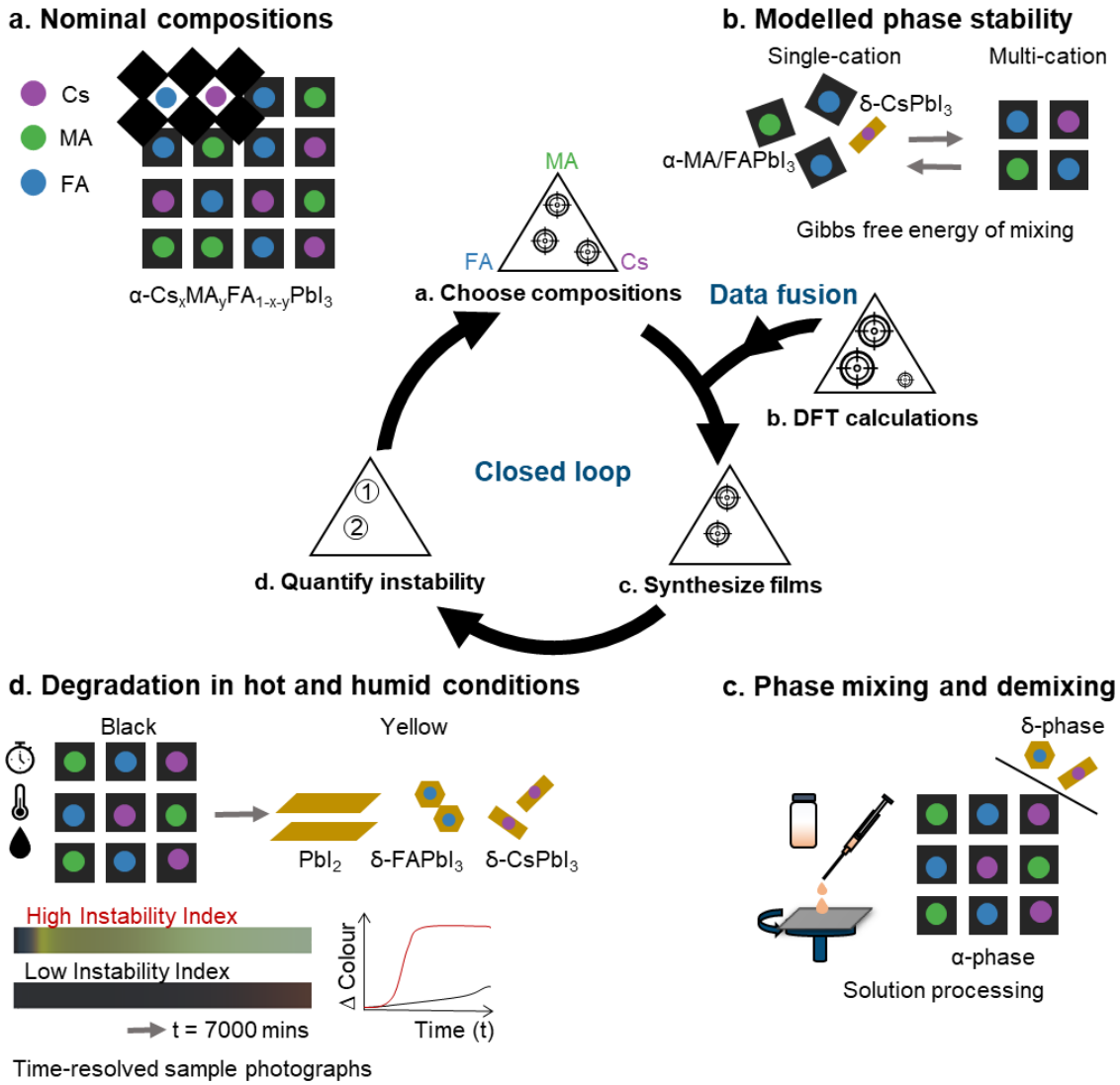


Figure 1 Each optimisation round consists of three steps of 'composition selection', 'film synthesis', and 'instability quantification' and a fourth step of theoretical incorporation into the closed-loop workflow by a data fusion approach. a-b. Composition selection: the selection of nominal compositions of multi-cation perovskites is driven by machine learning algorithms with DFT-modelled ΔG_{mix} incorporated by data fusion. c. Film synthesis: tuning A-site cations in lead iodide perovskites forms $Cs_xMA_yFA_{1-x-y}PbI_3$, a space of 5151 compositions (estimated experimental resolution 1%). Thin-film samples are spin-coated in series using precursor solutions of nominal compositions. d. Instability quantification: we perform accelerated degradation tests with *in situ* optical monitoring, enabling 28 thin-film samples being degraded in parallel. Near-black photoactive perovskite films turn yellow over time due to the emergence of high-bandgap degradation products. Quantified optical changes over time, recorded in R (red), G (green), and B (blue) channels, are used as a proxy to evaluate the chemical instability of the samples under 85% relative humidity (RH), 85°C sample temperature, and 0.15 Sun visible only illumination. Two representative sample photographs and curves of the total (R + G + B) area-averaged value as a function of time is illustrated in d.

To direct the compositional search in a $Cs_xMA_yFA_{1-x-y}PbI_3$ five-element space, we construct a physics-informed batch Bayesian optimization (BO) framework through data fusion (Fig. 1). BO explores efficiently vast variable spaces in a 'closed-loop' fashion, whereby the outcome of one experimental round informs the next without human intervention.[20] BO is an established ML method[21] that recently successfully directed experimentation in several materials systems including the self-driving deposition of organic hole-transport layers for solar cells[5] and the search for inorganic oxide perovskites for piezoelectrics.[6] As a key algorithm contribution, we fuse ΔG_{mix} as a *probabilistic constraint* of the BO acquisition function in the 'composition selection' step, providing additional information on phase stability to effectively identify the composition suggestions of multi-cation perovskites that are thermodynamically stable relative to their single-cation counterparts (Fig. 1a-b). In this study, we define a figure of merit for optimising stability that we call the '*Instability Index (I_c)'*. The goal of each optimization round, which consists of three steps of 'composition selection', 'film synthesis' and 'instability quantification', is to minimise this value. Our batch BO algorithm makes use of a surrogate ML model, Gaussian process regression (GP), to estimate the value and uncertainty of I_c in non-explored regions of the compositional space.[22] An 'acquisition function', *expected improvement* $EI(\Theta)$ (see algorithm details in supplementary information), takes the estimated mean and uncertainty and suggests promising compositions for the next experimental round, balancing the exploitation of the most stable regions minimising mean I_c, and the exploration of high-uncertainty regions within the compositional space. Within each experimentation round (one batch), 28 spin-coated thin-film samples (Fig. 1c) are examined *in situ* in parallel using an environmental chamber under 85 RH% and 85°C in air (chamber schematics see Fig. S1).

Photoactive α-perovskite phases within $Cs_xMA_yFA_{1-x-y}PbI_3$ exhibit a bandgap of ~1.5 eV, whereas their main degradation products under hot and humid conditions, PbI_2 (2.27 eV),[23] δ-$CsPbI_3$ (2.82 eV)[24] or δ-$FAPbI_3$ (2.43 eV)[25] show deteriorated photophysical properties (Fig. S2).[26–28] We hence employ a colour-based metric as a proxy to capture the macroscopic evolution of the high-bandgap, non-perovskite phases through the optical change of the near-black perovskite films.[29] 0.15 Sun visible only illumination is applied to enable automatic image capture by an RGB camera (~200 µm resolution). We define *Instability Index (I_c)* (Eq. 1) as the integrated colour change of an unencapsulated perovskite film under severe environmental stresses measured every five minutes throughout a 7000-minute (116 hours 40 mins) stability test (Fig. 1d)[30,31]:

$$I_c\,(\Theta) = \sum_{c=\{R,G,B\}} \int_{t=0}^{t=7000\text{min}} |c(t,\Theta) - c(0,\Theta)|\, dt, \quad (1)$$

where composition $\Theta = (x, y, 1-x-y)$, t is time and c are area-averaged, colour-calibrated red, green and blue pixel values of the sample (see supplementary information for colour calibration).

The cut-off time was determined based on the observed divergence between the most and least stable compositions (Fig. S3). Our closed-loop and iterative workflow enable the systematic optimization of multi-cation perovskites against degradation by varying the nominal compositions, Θ, within $Cs_xMA_yFA_{1-x-y}PbI_3$ (x, y limit to two decimal places) (See Table S1-2 for experimental summary).

Data fusion Approach: Incorporation of Phase Thermodynamics into Iterative Composition Selection

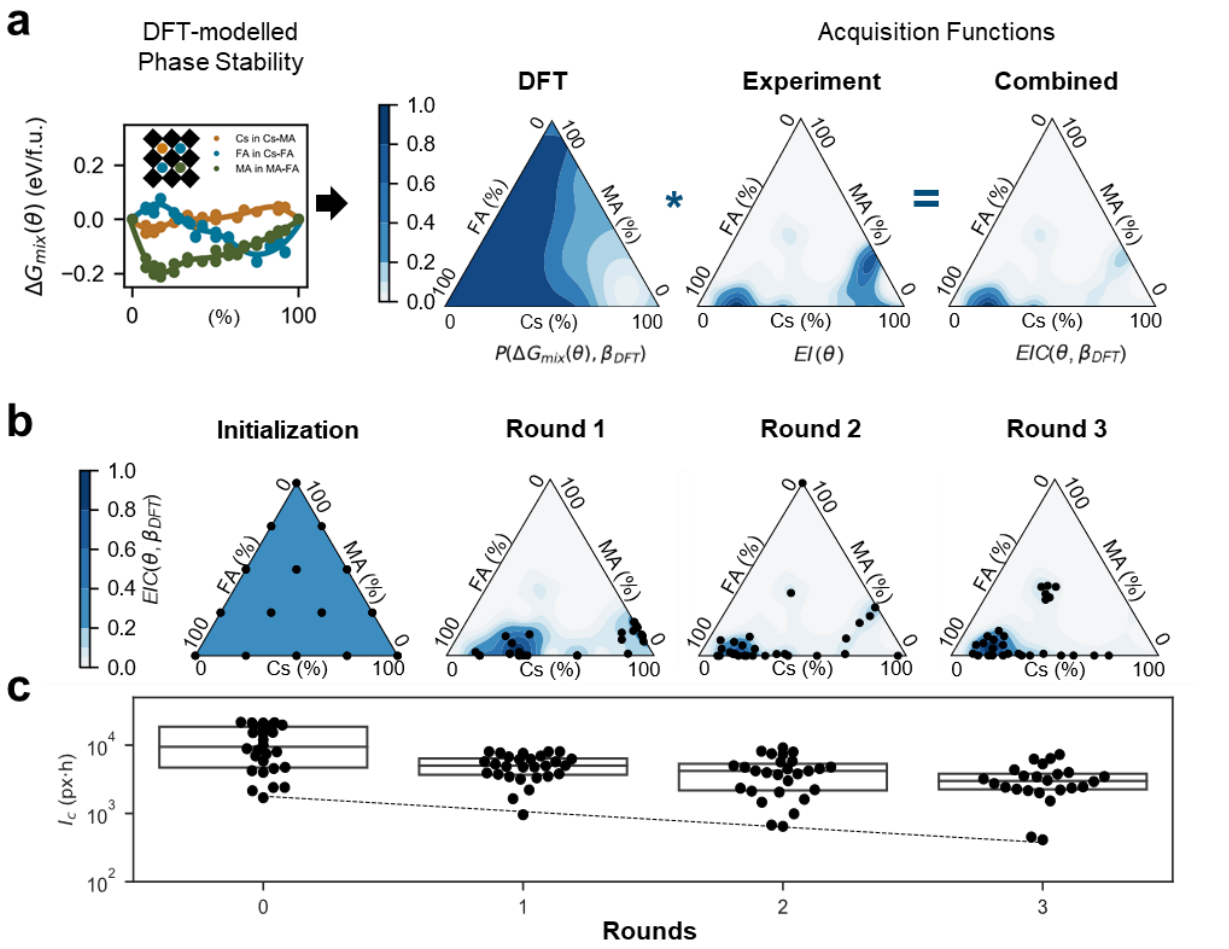


Figure 2 a. DFT-modelled Gibbs Free Energy of mixing, ΔG_{mix}, of 47 binary compositions of CsMA, CsFA, and MAFA α-perovskites relative to end members of δ-$CsPbI_3$, α-$MAPbI_3$, and α-$FAPbI_3$, are presented. We regress these phase thermodynamics data using a Gaussian process model with a radial basis function kernel, defining a Gibbs free-energy model, $\Delta G_{mix}(\Theta)$, where Θ is a composition in the ternary space. $\Delta G_{mix}(\Theta)$ is then transformed into a probabilistic constraint, $P(\Delta G_{mix}(\Theta), \beta_{DFT})$ that models the cumulative probability of phase mixing at above 300 K according to ΔG_{mix}. The data fusion parameter β_{DFT} is calibrated according to the estimated uncertainty of DFT calculations. Multiplying the probabilistic constraint model with the acquisition function of the batch Bayesian optimization gives a DFT-weighted acquisition function. b. Starting from equally-spaced 15 compositions in Round 0, 28 compositions are synthesized and undergo degradation tests following the DFT-weighted acquisition function in each optimization round. The acquisition functions are indicated in blue. See Fig. S4 for uncertainty landscapes of each round. c. Experimentally measured instability indices of 112 samples over 4 experimental rounds. The black boxes indicate the mean and standard deviation of each round. The unit of the *Instability Index*, I_c is pixels*hours

(px.h). The dashed line indicating the most stable compositions in each experimental round is for eye guidance only.

Due to the polymorphic nature, identical perovskite compositions crystallised into different phases can exhibit diverse degradation behaviours, making it essential to evaluate phase stabilities in any perovskite composition optimisation.[1] The end members of the compositional space in this study consist of the cubic α-FA/MAPbI$_3$ perovskites and the non-perovskite δ-CsPbI$_3$ at the synthesis temperature.[14] We observe Cs-rich compositions (> 20% Cs) of aimed at α-Cs$_x$MA$_y$FA$_{1-x-y}$PbI$_3$ contain minority phases of the non-perovskite, but more stable δ-CsPbI$_3$ at the room temperature (Fig. S10). Such phase demixing during synthesis leads to minority phases within thin-film samples prior to degradation tests and are, therefore, not captured in I_c. Schelhas *et al.* recently demonstrated the use of DFT calculations to predict the phase demixing tendency between α-Cs$_x$MA$_y$A$_{1-x-y}$PbI$_3$ (G_{mix}) and their single-cation perovskite polymorphs *A*PbI$_3$ (*A* = Cs, MA, or FA) (G_0) at a given temperature.[12] We herein fuse the composition-dependent change in Gibbs free energy of mixing ΔG_{mix} as a constraint into the experimental optimisation loop (Figs. 2a,b). This approach allows the α- and δ-phase relative stability to be considered in the composition selection, thus enabling us to reduce sampling in regions with high probability of minority phase formation.

Data fusion refers to a set of techniques where ML is used to map two or more datasets coming from related but distinct distributions. In our case, we relate the theoretical $\Delta G_{mix}(\Theta)$ and the experimental $I_c\,(\Theta)$. The two data streams account for distinct mechanisms of modelled thermodynamic phase instability and measured macroscopic thermal-moisture instability, respectively. Hence, it is inadequate to combine both datasets as equivalent or include DFT directly as a prior following state-of-the-art model-free BO.[32,33] We herein define a data-fused *probabilistic constraint* approach according to Eq. (2):

$$P(\Delta G_{mix}(\Theta), \beta_{DFT}) = \frac{1}{1 + e^{-\Delta G_{mix}(\Theta)/\beta_{DFT}}}, \qquad (2)$$

where $P(\Delta G_{mix}(\Theta), \beta_{DFT})$ is a logistic cumulative distribution function (CDF) modelling the phase mixing probability. β_{DFT} is a data fusion parameter calibrated according to ΔG_{mix} calculations to control the smoothness of the boundaries from stable to unstable compositions, forming a soft compositional boundary presented in Fig. 2a. (See Fig. S6-7 for algorithm details).

Given the computational cost and complexity of DFT calculations on organic–inorganic hybrid systems, Schelhas *et al*. mostly computed ΔG_{mix} around the edges of the α-Cs$_x$MA$_y$FA$_{1-x-y}$PbI$_3$ phase diagram. To overcome this data-sparsity challenge, we first regress 85 DFT-modelled ΔG_{mix} values on 47 single-cation and binary alloyed compositions (3 single-cation, 29 MAFA and CsFA compositions from Schelhas *et al.*[14] 12 CsMA compositions computed for the present work using the same method) over the quasi-ternary phase space by using an auxiliary GP model that defines $\Delta G_{mix}(\Theta)$ (Fig. 2a). Figure 2b visualises the probability of phase mixing $P(\Delta G_{mix}(\Theta), \beta_{DFT}) \in [0,1]$ as defined by Eq. (3), where low values suggest phase instability ($\Delta G_{mix} >> 0$) and high values suggest phase stability ($\Delta G_{mix} << 0$).

Our work is inspired by the *unknown constraint* BO proposed by Gelbart *et al*.[34] By developing a probabilistic constraint model $P(\Delta G_{mix}(\Theta), \beta_{DFT})$ instead of applying a hard-constraint boundary, we are able to discount regions predicted by DFT to go through phase demixing rather than completely exclude any unfavourable regions. This approach accounts for the inherent uncertainty in DFT predictions and chemical accuracy, and data scarcity through the use of the soft compositional boundary to model the stability threshold (see supporting information for β_{DFT}

calibration). The proposed algorithm allows us to seamlessly adapt DFT into the experimental optimizations loop, thereby achieving a physics-informed and sample-efficient search without being limited by the unknown exact phase boundaries across a vast compositional space.

To integrate the *probabilistic constraint* into the BO formulation, we weigh the acquisition function with the value of $P(\Delta G_{mix}(\Theta), \beta_{DFT})$ and obtain a DFT-weighted BO acquisition function, $EIC\ (\Theta)$, as illustrated in Fig. 2b. Results of our first experimental round indicate two potential optima in Cs-poor and Cs-rich regions, respectively, based on I_c. The DFT-weighted algorithm effectively reduces sampling in energetically unfavourable Cs-rich regions despite low I_c: the subsequent optimization rounds converge to stable nominal compositions with a high probability of stable α-perovskite films among Cs-poor regions (Fig. 2a). Comparisons of optimisations with and without DFT-weighting using a teacher-student model are shown in Fig. S8-9, which validates without data fusion, the model-free BO algorithm continues to suggest sampling in Cs-rich regions despite of their phase instability.

Figures. 2c demonstrates that batch BO sequentially identifies the most stable regions over four experimental rounds of synthesis and degradation tests. Fig. 2d reveals a rapid decrease in experimentally quantified I_c from Rounds 0–3. The search converges after three optimization rounds (see Fig. S5 for convergence conditions) to an optimal composition region centred at $Cs_{0.17}MA_{0.03}FA_{0.80}PbI_3$ and bounded by 8-29% Cs, <14 % MA and 68-92 % FA. The identification of the global optimum lying within an FA-rich, Cs and MA-poor region is consistent with the reports that FA-rich perovskites show superior environmental stability compared with their MA-rich counterparts and the less volatile Cs is expected to enhance the heat and moisture resistance.[35] Interestingly, we found a local optimum near $Cs_{0.26}MA_{0.36}FA_{0.38}PbI_3$, which emerged in Round 1. We intentionally sample four additional compositions and verify the existence of local optima in Round 3. We define the compositional space as the discretized quasi-ternary phase space subdivided by the minimum achievable experimental resolution (1% composition). This yields 5151 possible singular, binary and ternary cation compositions, 1.8% of which were sampled experimentally, converging to the optimal region. Three additional degradation rounds of seven representative compositions were performed, to validate the instability trend with structural and optical characterisation shown in the supplementary information. (See Table S3 and Fig. S13 for bandgap measurements via UV-Vis spectroscopy and Fig. 12 for minority phase identification via grazing incidence X-ray diffraction)

Composition-dependent and Time-resolved Instability Landscape

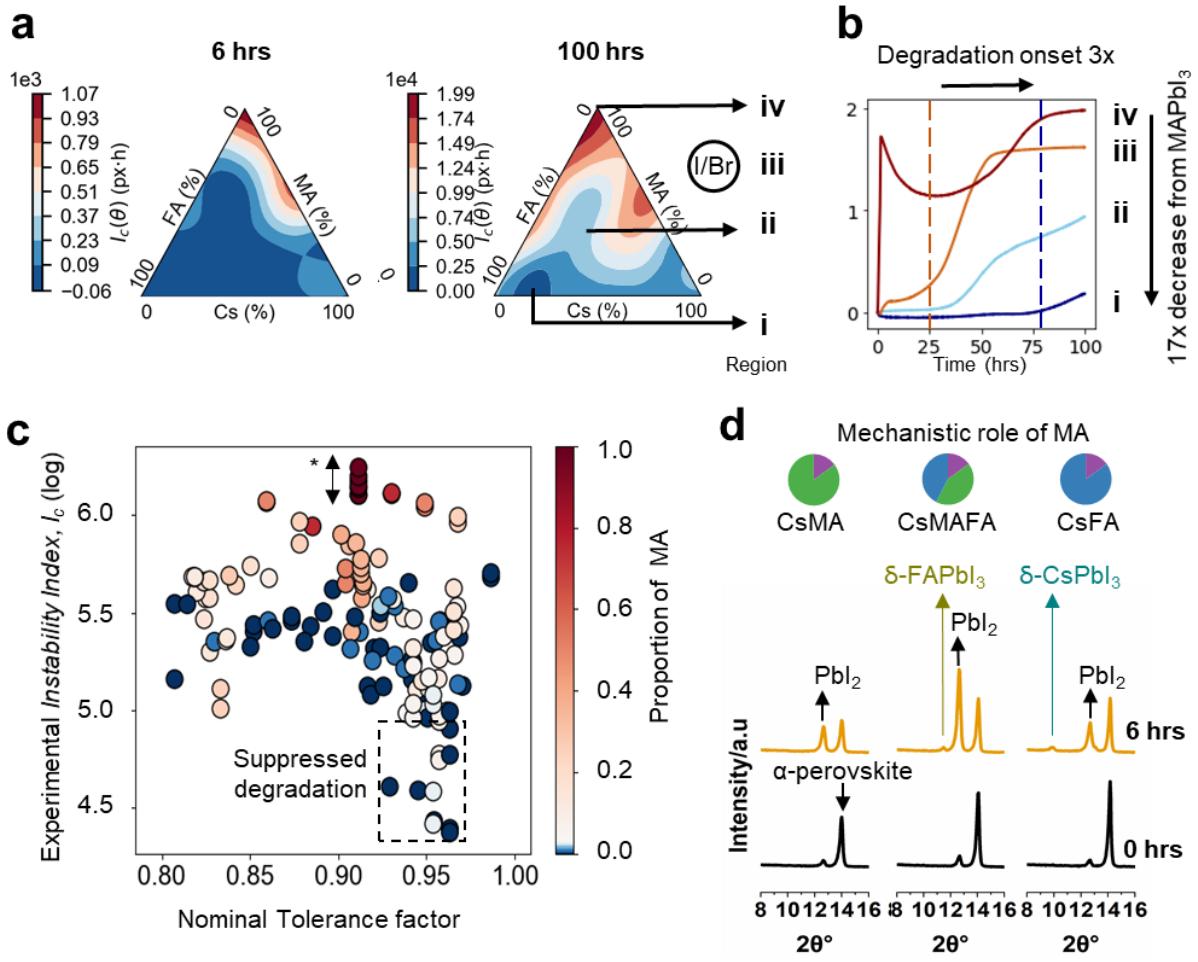


Figure 3 a. Evolution of *Instability landscape,* $I_c(\Theta)$*:* the Gaussian process (GP) regressed the posterior mean of experimentally measured *Instability Index*, I_c, of the algorithm-chosen compositions mapped to the full compositional space. The *Instability Index* is the integrated optical change of an unencapsulated perovskite film measured every five minutes under under 85 RH% and 85°C and 0.15 Sun illumination in air. The two instability landscape triangles are obtained using degradation profiles of 112 experimentally measured samples after 6- and 100-hour degradation tests respectively. Four distinct compositional regions are identified at 100 hours within $Cs_xMA_yFA_{01-x-y}PbI_3$: the most stable (i), local optimum (ii), and the least stable (iv), as well as the I/Br mixed anion compositional space which is not optimized in this study (iii). Regions i – iv are labelled in ascending order of measured I_c. b. Relative optical changes measured as a function of degradation time, showing the onset of degradation for representative compositions in regions i – iv, $Cs_{0.17}MA_{0.03}FA_{0.80}PbI_3$, $Cs_{0.26}MA_{0.36}FA_{0.38}PbI_3$, $Cs_{0.05}(MA_{0.17}FA_{0.83})_{0.95}Pb(I_{0.83}Br_{0.17})_3$, and $MAPbI_3$ respectively. We reveal the ML-optimized composition in Region iii stay macroscopically stable for 3 times longer than $Cs_{0.05}(MA_{0.17}FA_{0.83})_{0.95}Pb(I_{0.83}Br_{0.17})_3$. c. Measured I_c from Round 0 – Round 3 over full degradation runs of 7000 mins (116 hours 40 mins) as a function of the proportion of MA in the composition nominal tolerance factor, which is the Goldschmidt's tolerance factor based on nominal compositions, over the entire $Cs_xMA_{1-x}FA_{1-x-y}PbI_3$ space. *Experimental uncertainty of measured I_c across batches in a control composition, $MAPbI_3$, across multiple batches (see Appendix for full sample list). The composition range within the dashed line box shows comparable final degradation products after a full degradation run. d. Laboratory grazing incidence X-ray diffraction measurements on binary and ternary alloyed compositions with fixed Cs proportions to evaluate the effect of MA. Over-stoichiometric precursors with excess PbI_2 were

added following the high-performing perovskite semiconductor recipe in ref.[35] CsFA: 15%Cs 85%FA. CsMA: 15%Cs 85%MA. CsMAFA: 15%Cs 42.5%MA 42.5%FA.

The overall instability landscape for $Cs_xMA_{1-x}FA_{1-x-y}PbI_3$ after three optimization rounds on I_c over full degradation rounds of 7000 minutes identify three distinct compositional regions with descending I_c (see Fig. S4 for iterative landscape estimation and uncertainty). Figure 3a presents the evolution of GP posterior mean of, I_c, i.e. I_c (Θ) as a function of degradation time. Fast degradations of MA-rich compositions are evident after six hours of degradation tests (Region (iv),) while two additional regions, representing several compositions in local optima (Region (ii)) and the global optima (Region (i)), are sequentially distinctive after 100 hours of degradation tests (see Fig. S3 for time-resolved instability landscape and uncertainty). Quantitative optical change analysis of representative compositions from the three regions, the ML-optimum $Cs_{0.17}MA_{0.03}FA_{0.80}PbI_3$ (i), the ML-local optimum $Cs_{0.26}MA_{0.36}FA_{0.38}PbI_3$ (ii) and $MAPbI_3$ (iv), reveal a 17x reduction in I_c from composition (iv) to (i). To evaluate the stability improvement in this study, we further compare the optimised five-element (i) with a six-element reference composition, $Cs_{0.05}MA_{0.17}FA_{0.83}Pb(I_{0.83}Br_{0.17})_3$ (iii), which is outside the design space of this study and has been extensively employed in state-of-the-art perovskite devices (see Fig. S14 for morphological characterisation).[35] We found that (i) yields a 3.5x lower I_c than (iii). The postponed degradation onset is reflected by the 3x delay in the onset of sharp optical change as per Fig. 3b. Our finding provides insight into the limitation of more complex alloying methods to improve the environmental stability of hybrid perovskites The improvement in stability under severe environmental stresses does not scale with the number of elements incorporated into the perovskite structure, where the entropic stabilizations through cationic and/or anionic mixing[36] can be offset by the multiple degradation pathways under long exposure of environmental stress of heat, moisture and illumination as seen in this study.

We found the composition–stability relationships within $Cs_xMA_yFA_{1-x-y}PbI_3$ are non-linear. As we increase the Cs content, degradations are first suppressed (up to 26% Cs) and then increased. Figure 3c visualises experimentally measured I_c as a function of Goldschmidt's tolerance factor (TF) based on the average ionic radius of A-site cations of nominal compositions. TF is empirical guidance that has been widely applied to estimate the intrinsic structural stability of hybrid perovskites.[37,38] During optimization Rounds 1–3, an increasing number of compositions within a TF of 0.93–0.97 are suggested by BO (Fig. S18), indicating high stability of compositions with a TF around 0.95. This value is lower than TF in an ideal cubic structure (TF = 1), attributing to the incorporation of small-radius and non-volatile Cs into the α-lattice to improve moisture and heat resistance. By achieving this holistic view shown in Fig. 3c over the entire compositional space, we should note that tolerance factor optimization alone is a necessary but not sufficient requirement for achieving high stability. The observation of limited stability improvement with Cs addition can be explained by the observation that more than 20% Cs in nominal compositions form minority phases during synthesis (annealing temperature of 403 K) as observed in grazing incidence X-ray diffraction (GIXRD) (Fig. S10), Cs crystalised in δ-$CsPbI_3$ does not subsequently contribute to the stability improvement of the α-perovskite.

Interestingly, several MA-containing compositions (up to 8%) show comparable $I_c$to their CsFA binary cation counterparts as per Fig. 3c. In addition to the ML optimum, top-performing compositions include $Cs_{0.13}MA_{0.08}FA_{0.79}PbI_3$ and $Cs_{0.13}FA_{0.87}PbI_3$. To understand the distinctive role of MA in the degradation dynamics of multi-cation perovskites behind the optimisation results, we fix the Cs content (x = 0.15) and compare the structural changes *via* GIXRD of α-CsMA, CsFA

and CsMAFA with an equal amount of MA and FA (y = (1-x)/2 = 0.425) after an initial six hours of degradation (Fig. 3d). Two composition-dependent degradation mechanisms are observed. While all three samples show increased PbI_2 content (evident of chemical decomposition to precursors), the emergence of δ-$CsPbI_3$ and δ-$FAPbI_3$ in CsFA and CsMAFA films, respectively, indicates additional mechanisms of phase separation during degradation tests. Both increases in PbI_2 and two high-bandgap δ-phases contribute to the overall yellowing of the films throughout the degradation run. This finding from empirical observations validates the recent theoretical prediction that MA is more effective than Cs in stabilising the α-phase alloyed perovskties, driven by lower enthalpy of mixing on addition of MA.[12,14] Carefully tuned MA contribute to the suppressed degradation by controlling phase separations in hot and humid conditions.

Composition- and Phase-dependent Degradation Kinetics

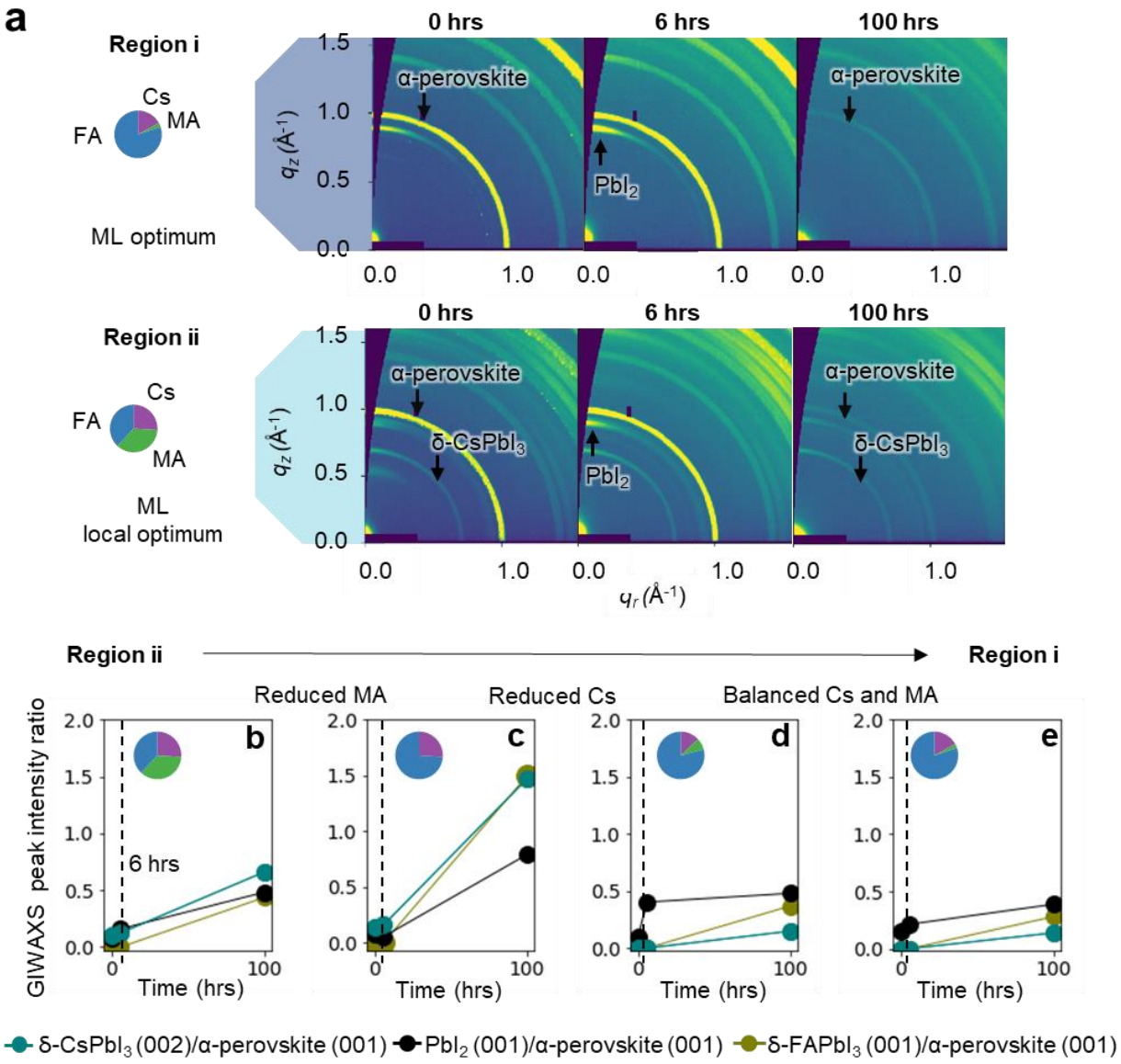


Figure 4 a. GIWAXS images of the ML optimum, $Cs_{0.17}MA_{0.03}FA_{0.80}PbI_3$ in Region i and $Cs_{0.26}MA_{0.36}FA_{0.38}PbI_3$ in Region ii after 0, 6 and 100 hours of degradation tests. b. GIWAXS peak intensity ratios of the non-perovskite phases and PbI_2 relative to the perovskite phase as an estimate for the extent of degradation. From left to right: $Cs_{0.26}MA_{0.36}FA_{0.38}PbI_3$, $Cs_{0.26}FA_{0.74}PbI_3$, $Cs_{0.13}MA_{0.8}FA_{0.79}PbI_3$, and $Cs_{0.17}MA_{0.03}FA_{0.8}PbI_3$.

We seek to determine the structural characteristics that lead to divergences in degradation kinetics. As shown in Fig. 4a, *ex situ* GIWAXS after 0, 6 and 100 hours of degradation reveal that the ML-optimum, (i) $Cs_{0.17}MA_{0.03}FA_{0.80}PbI_3$ in the most stable region, Region i, and the ML-local optimum, (ii) $Cs_{0.26}MA_{0.36}FA_{0.38}PbI_3$ in the second most stable region, Region ii, show comparable lattice parameters of α-perovskites prior to degradation tests, however (ii) contains δ-$CsPbI_3$ minority phase at 0 hours. Interestingly, only a slightly faster increase in PbI_2 is observed in (ii) in a 6-hour degradation run. After 100 hours, (ii) exhibits a significantly bigger loss of intensity of α-perovskites and crystallinity. We also observe lattice shrinkage after 100 hours for both compositions, which may due to a higher proportion of the small-radius Cs is left in the α-perovskites after degradation as the more volatile and larger radii MA and FA leave the cubic structure (Fig. S17).

The data fusion approach effectively discounts sampling in the regions with high probabilities of minority phase formation. To evaluate the role of δ-minority phases in the degradation kinetics of the co-existing α-perovskites of interest, we use the scattering of PbI_2(001), δ-$CsPbI_3$(002), $FAPbI_3$(001) and α-perovskite (001) as indicators of corresponding phases and the ratios of intensity between the peaks to describe the chemical decomposition and phase separation, respectively. We found that samples with δ-$CsPbI_3$ minority phase at initial states (Figs. 4b, c) undergo faster chemical decomposition over the 100-hour degradation tests compared with samples with low Cs content and no pre-(degradation test)-crystalised δ-$CsPbI_3$ (Figs. 4d,e). Nevertheless, despite no δ-$CsPbI_3$ seen at 0 hours or 6 hours in $Cs_{0.13}MA_{0.08}FA_{0.79}PbI_3$ (Fig. 4d) and the ML-optimum $Cs_{0.17}MA_{0.03}FA_{0.80}PbI_3$ (Fig. 4e), δ-$CsPbI_3$ formed under elevated temperature and humidity conditions are visible after 100 hours. These results indicate that in hot and humid conditions, δ-phases are both a degradation product of phase separation and an accelerator for chemical decomposition. Tuning the MA/Cs ratios is an effective strategy for modulating the δ-minority phase formation during degradation. Comparing Figs. 4b and c4, we observe unfavourable phase separation of δ-$FAPbI_3$ emerged between 6 -and 100 hours of degradation if we completely replace MA with the less volatile FA. GIWAXS images of compositions listed in Fig.S15-16 illustrates that phase separation, indicated by additional minority phase formation, takes place in a longer time scale compared with chemical decomposition since an increase in PbI_2 is observed within the initial six hours. Overall, the compositional optimisation strategy from Region (ii) to Region (i) can be interpreted as follows: 1) reducing MA for suppressed chemical decomposition, 2) reducing Cs for limiting minority phase formation, and 3) balancing MA, FA, and Cs for restraining phase separation in hot and humid conditions.

Conclusions

In summary, we develop a closed-loop optimization strategy for $Cs_xMA_yFA_{1-x-y}PbI_3$ multi-cation perovskites against heat and moisture-induced degradation by introducing an iterative and physics-informed Bayesian optimization framework. We identify an FA-rich and Cs-poor region centred at $Cs_{0.17}MA_{0.03}FA_{0.8}PbI_3$ exhibiting enhanced environmental stability while sampling only 1.8% of the compositional space, achieving superior search efficiency to brute-force screening. Our study demonstrates the power of data fusion to allow material search over vast and sparsely-sampled compositional spaces, where the DFT-modelled thermodynamics serves as a *probabilistic constraint* and provides principled guidance to experimentation. This approach

successfully fuses multiple data sources into a single search algorithm and can be utilised to include any other experimental or theoretical constraints with non-negligible uncertainty into the materials design strategy.

We apply this physics-informed optimization framework to achieve a holistic understanding of the fundamental composition- and phase- dependent alloy behaviour within the quasi-ternary $Cs_xMA_yFA_{1-x-y}PbI_3$ space. We perform *in situ* optical assessment and *ex situ* synchrotron-based GIWAXS, and reveal that the interplay between the minority phase formation led by thermodynamic structural instability and the chemical decompositions led by thermo-moisture instability govern the overall degradation profile of multi-cation perovskites in the initial 100 hours under elevated temperature and high humidity environments. As a consequence of competing roles of cations in different degradation mechanisms, a composition window of up to 8% of addition of the least chemically stable cation, MA, contributes to kinetically suppressed degradation, whereas the most chemically stable cation in this design space, Cs, contributes to accelerated degradation even in the most macroscopically-stable candidates of both CsFA and CsMAFA compositions, which the potentially limits the benefits of Cs as a perovskite stabilizing agent.. This finding highlights the detrimental effects of minority phase formation, which occurs over a longer timeframe under the 85°C/85%RH degradation tests than chemical decomposition, and hence are easily underestimated during compositional engineering. We further suggest several optimized iodide perovskites, including $Cs_{0.13}FA_{0.87}PbI_3$ and $Cs_{0.13}MA_{0.08}FA_{0.79}PbI_3$, that show superior stability than the state-of-the-art iodide-bromide mixed bromide perovskite ($Cs_{0.05}MA_{0.17}FA_{0.83}Pb(I_{0.83}Br_{0.17})_3$) under elevated temperature and humidity, providing insights into simplifying perovskite compositions for stability.

Our data fusion and probabilistic constraint method can be generalized to a wide range of solid-state material systems, merging complementary experiments and theory into an end-to-end close-loop ML framework to identify new materials with a holistic view of the entire design space, as well as to achieve simultaneous structural and chemical optimization in a vast compositional space.

Data and Code Availability

The machine learning framework and camera-based in situ degradation monitoring platform details are avalible in Supplementary information. The codes and the datasets used for Bayesian optimization are available in GitHub repository https://github.com/PV-Lab/SPProC. Lists of samples, XRD characterization, and thin-film degradation results are shown in Appendix available at https://github.com/PV-Lab/SPProC.

Acknowledgments

The authors thank members of MIT Photovoltaics Research Laboratory, in particularly Jim Serdy for setting up and maintaining the environmental chamber for accelerated aging tests; Dr. I. Marius Peters for fruitful discussions on the initial design of experiments, Antonio Buscemi and Isaac Metcalf for assistance in XRD measurements, Isaac Siyu Tian for initial image processing, and Shreyaa Raghavan for user-friendly Python package development. SS thanks Dr. Masafumi Fukuto and Dr. Kevin Yager from Brookhaven National Laboratory for beamtime assistance. High-throughput GIWAXS images were collected at BM-11, NSLS-II. The authors further thank Prof. Joshua Schrier (Fordham university) for discussion on perovskite material discovery using

machine learning techniques, and SS thanks Felice Frankel for valuable insights on scientific graphics and visualization.

This study is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. HR001118C0036. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA. SS, NTPH, TB thank TOTAL S.A., US National Science Foundation grant CBET-1605547, and the Skoltech NGP program for financial support. AT was supported by Alfred Kordelin Foundation and Svenska Tekniska Vetenskaps-akademien i Finland. FO was supported by the U.S. Department of Energy under Photovoltaic Research and Development program under Award DE-EE0007535. This research used computational resources sponsored by the DOE Office of Energy Efficiency and Renewable Energy and located at National Renewable Energy Laboratory. JJY and MB were supported by the Institute for Soldier Nanotechnology (ISN) grant W911NF-13-D-0001, the National Aeronautics and Space Administration (NASA) grant NNX16AM70H, and the Eni-MIT Alliance Solar Frontiers Center.

Author Contributions

SS and TB conceived the project. SS designed the experiments and led the laboratory and synchrotron-based measurements. AT developed colour and chamber environment calibration. AT and LZ processed the camera data. FO and JF designed the Bayesian optimization algorithms and instability index, and FO, AT, and LZ wrote the Bayesian optimization implementation. JT synthesized all thin-films samples supervised by SS. AE built the automatic humidity control supervised by SS and AT. CB and AT developed the camera control program and visualized degradation data. AG performed DFT calculations. FO, AT, LZ, and JF developed the incorporation of DFT into the Bayesian optimization loop, and AT implemented the Gaussian process model and probability distribution on phase stability. JT, AT, NTPH performed the degradation tests. JT and NTPH performed laboratory structural and morphological characterization. CB performed optical characterization supervised by SS. SS, JT, NTPH, and JY performed structural characterization using synchrotron radiation. RL assisted the measurements at beamline BM-11, NSLS-II. ZR performed the teacher-student model for DFT and non-DFT optimization comparison. MB, VS, JF and TB supervised the project. SS, AT, FO and TB wrote the manuscript with contributions from all co-authors. All authors discussed the results and inferred the implications of the project.

References

1. Boyd, C. C., Cheacharoen, R., Leijtens, T. & McGehee, M. D. Understanding Degradation Mechanisms and Improving Stability of Perovskite Photovoltaics. *Chemical Reviews* vol. 119 3418–3451 (2019).

2. Jeon, N. J. *et al.* Compositional engineering of perovskite materials for high-performance solar cells. *Nature* **517**, 476–480 (2015).

3. Saidaminov, M. I. *et al.* Multi-cation perovskites prevent carrier reflection from grain surfaces. *Nat. Mater.* **19**, 412–418 (2020).

4. Correa-Baena, J.-P. *et al.* Promises and challenges of perovskite solar cells. *Science.* **358**, 739–744 (2017).

5. Rohr, B. *et al.* Benchmarking the acceleration of materials discovery by sequential learning. *Chem. Sci.* **11**, 2696–2706 (2020).

6. Zhang, W., Mao, J., Li, S., Chen, Z. & Guo, Z. Phosphorus-Based Alloy Materials for Advanced Potassium-Ion Battery Anode. *J. Am. Chem. Soc.* **139**, 3316–3319 (2017).

7. George, E. P., Raabe, D. & Ritchie, R. O. High-entropy alloys. *Nature Reviews Materials* vol. 4 515–534 (2019).

8. Zhong, M. *et al.* Accelerated discovery of CO2 electrocatalysts using active machine learning. *Nature* **581**, 178–183 (2020).

9. Talapatra, A. *et al.* Experiment design frameworks for accelerated discovery of targeted materials across scales. *Frontiers in Materials* vol. 6 82 (2019).

10. Bai, Y. *et al.* Accelerated Discovery of Organic Polymer Photocatalysts for Hydrogen Evolution from Water through the Integration of Experiment and Theory. *J. Am. Chem. Soc.* **141**, 9063–9071 (2019).

11. Smecca, E. *et al.* Stability of solution-processed MAPbI3 and FAPbI3 layers. *Phys. Chem. Chem. Phys.* **18**, 13413–13422 (2016).

12. Schelhas, L. T. *et al.* Insights into operational stability and processing of halide perovskite active layers. *Energy Environ. Sci.* **12**, 1341–1348 (2019).

13. Tan, W., Bowring, A. R., Meng, A. C., McGehee, M. D. & McIntyre, P. C. Thermal Stability of Mixed Cation Metal Halide Perovskites in Air. *ACS Appl. Mater. Interfaces* **10**, 5485–5491 (2018).

14. Kim, S., Eom, T., Ha, Y.-S., Hong, K.-H. & Kim, H. Thermodynamics of Multicomponent Perovskites: A Guide to Highly Efficient and Stable Solar Cell Materials. *Chem. Mater.* **32**, 4265–4272 (2020).

15. Poorkazem, K. & Kelly, T. L. Compositional Engineering to Improve the Stability of Lead Halide Perovskites: A Comparative Study of Cationic and Anionic Dopants. *ACS Appl. Energy Mater.* **1**, 181–190 (2018).

16. Saliba, M. Polyelemental, Multicomponent Perovskite Semiconductor Libraries through Combinatorial Screening. *Adv. Energy Mater.* **9**, 1803754 (2019).

17. Li, Z., Xu, Q., Sun, Q., Hou, Z. & Yin, W.-J. *Stability Engineering of Halide Perovskite via Machine Learning*. https://arxiv.org/ftp/arxiv/papers/1803/1803.06042.pdf.

18. Curtarolo, S. *et al.* The High-throughput Highway to Computational Materials Design. *Nat. Mater.* **12**, 191–201 (2013).

19. Jesper Jacobsson, T. *et al.* Exploration of the compositional space for mixed lead halogen perovskites for high efficiency solar cells. *Energy Environ. Sci.* **9**, 1706–1724 (2016).

20. Packwood, D. *Bayesian Optimization for Materials Science*. vol. 3 (Springer, 2017).

21. Häse, F., Roch, L. M., Kreisbeck, C. & Aspuru-Guzik, A. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **4**, 1134–1145 (2018).

22. SheffieldML/GPyOpt: Gaussian Process Optimization using GPy. https://github.com/SheffieldML/GPyOpt.

23. Zhu, X. H., Wei, Z. R., Jin, Y. R. & Xiang, A. P. Growth and characterization of a PbI2 single crystal used for gamma ray detectors. *Cryst. Res. Technol.* **42**, 456–459 (2007).

24. Hu, Y. *et al.* Bismuth Incorporation Stabilized α-CsPbI3 for Fully Inorganic Perovskite Solar Cells. *ACS Energy Lett.* **2**, 2219–2227 (2017).

25. Masi, S. *et al.* Chemi-Structural Stabilization of Formamidinium Lead Iodide Perovskite by Using Embedded Quantum Dots. *ACS Energy Lett.* 418–427 (2020) doi:10.1021/acsenergylett.9b02450.

26. Giorgi, G., Fujisawa, J. I., Segawa, H. & Yamashita, K. Cation role in structural and electronic properties of 3D organic-inorganic halide perovskites: A DFT analysis. *J. Phys. Chem. C* **118**, 12176–12183 (2014).

27. Hashmi, S. G. *et al.* Long term stability of air processed inkjet infiltrated carbon-based printed perovskite solar cells under intense ultra-violet light soaking. *J. Mater. Chem. A* **5**, 4797–4802 (2017).

28. Hashmi, S. G. *et al.* Air Processed Inkjet Infiltrated Carbon Based Printed Perovskite Solar Cells with High Stability and Reproducibility. *Adv. Mater. Technol.* **2**, 1600183 (2017).

29. Stoddard, R. J. *et al.* Forecasting the Decay of Hybrid Perovskite Performance Using Optical Transmittance or Reflected Dark-Field Imaging. *ACS Energy Lett.* **5**, 946–954 (2020).

30. Tiihonen, A. *et al.* Critical analysis on the quality of stability studies of perovskite and dye solar cells. *Energy Environ. Sci.* **11**, 730–738 (2018).

31. Menesatti, P. *et al.* RGB color calibration for quantitative image analysis: The '3D Thin-Plate Spline' warping approach. *Sensors (Switzerland)* **12**, 7063–7079 (2012).

32. Herbol, H. C., Poloczek, M. & Clancy, P. Cost-effective materials discovery: Bayesian optimization across multiple information sources. *Mater. Horizons* (2020) doi:10.1039/D0MH00062K.

33. Doan, H. A. *et al.* Quantum Chemistry-Informed Active Learning to Accelerate the Design and Discovery of Sustainable Energy Storage Materials. *Chem. Mater.* (2020) doi:10.1021/acs.chemmater.0c00768.

34. Gelbart, M. A., Snoek, J. & Adams, R. P. Bayesian optimization with unknown constraints. in *Uncertainty in Artificial Intelligence - Proceedings of the 30th Conference, UAI 2014* 250–259 (2014).

35. Saliba, M. *et al.* Cesium-containing triple cation perovskite solar cells: improved stability, reproducibility and high efficiency. *Energy Environ. Sci.* **9**, 1989–1997 (2016).

36. Senocrate, A., Kim, G. Y., Grätzel, M. & Maier, J. Thermochemical Stability of Hybrid Halide Perovskites. *ACS Energy Letters* vol. 4 2859–2870 (2019).

37. Kieslich, G., Sun, S. & Cheetham, A. K. Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog. *Chem. Sci.* **5**, 4712–4715 (2014).

38. Li, Z. *et al.* Stabilizing Perovskite Structures by Tuning Tolerance Factor: Formation of Formamidinium and Cesium Lead Iodide Solid-State Alloys. *Chem. Mater.* **28**, 284–292 (2016).

Supplementary Information

A Physical Data Fusion Approach to Optimize Compositional Stability of Halide Perovskites

Shijing Sun[1], Armi Tiihonen[1], Felipe Oviedo[1], Zhe Liu[1], Janak Thapa[1], Noor Titan P. Hartono[1], Anuj Goyal[2], Clio Batali[1], Alex Encinas[1], Jason J. Yoo[1], Ruipeng Li[3], Zekun Ren[4], Moungi G. Bawendi[1], Vladan Stevanovic[2], John Fisher III[1], Tonio Buonassisi[1,4]

[1] Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2] Metallurgical and Materials Department, Colorado School of Mines, CO 80401, USA

[3] Brookhaven National Laboratory, Upton, NY 11970, USA

[4] Singapore-MIT Alliance for Research and Technology, 138602, Singapore

Contents

1. Synthesis

Perovskite precursor solutions were spin-coated on UV-ozone glass substrates. Glass microscope slides (VWR) cut into square pieces were cleaned with sonication in 2% Hellmanex-DI water mix, DI water, and IPA respectively. We perform perovskite synthesis with over-stoichiometric PbI_2 in the molar ratio of 1.09 (PbI_2) to 1 (halide salt of CsI, MAI, and FAI). Lead (II) iodide stock solution was prepared in 9:1 N,N-dimethylformamide (Sigma-Aldrich) to dimethyl sulfoxide (Sigma-Aldrich) solvent. The perovskite precursor solution was prepared by mixing individual stock solutions following the ratios of Cs, FA, and MA suggested by the machine learning algorithm. The films were annealed at 403 K for 20 minutes using the central part of a hot plate in the glovebox. The spin-coating program follows a 2-step approach: 1000 rpm for 10 seconds and acceleration of 200 rpm/s, with a subsequent 6000 rpm for 30 seconds and acceleration of 2000 rpm/s. Chlorobenzene (Sigma-Aldrich) antisolvent in the quantity of 150 µL was dropped 5 seconds at the beginning of the second step of spin-coating. The annealed samples were cooled to room temperature before transferring to the degradation test. In total, 202 films were synthesised, and 196 of which were degraded in this study over 7 degradation rounds, including one repeated round, as shown in the main text. Of the 202 films synthesized across all rounds, 99 films were distinctly unique in composition, 95 of which were characterized by laboratory grazing incidence X-ray diffraction with 7 compositions were further characterized after 6 and 100 hours of degradation tests, respectively, using grazing-incidence wide-angle scattering.

Amongst the final selected seven compositions for advanced characterization, the bromine-containing composition, $Cs_{0.05}(FA_{0.83}MA_{0.17})_{0.95}Pb(I_{0.87}Br_{0.13})_3$, following a different recipe from Ref.[1] The halide perovskite precursor solutions for the composition were prepared by mixing FAI and CsI salts with PbI_2 solution, separately, and MAI with $PbBr_2$ solution, using the same method and molarity ratios as described above. All prepared precursors solutions were mixed together in the volume ratios listed for FA, MA and Cs.

Full list of samples and synthetic parameters are shown in Appendix Table A1 at (MIT PVLab Github repository, https://github.com/PV-Lab/SPProC).

Table S 1 Summary of sample synthesis and degradation tests.

Synthesis Round	No. of samples synthesized	No. of compositions	Full or partial degradation round	Structure Characterization	Sampling strategies	No. of triple cation compositions	No. of single or double cation composition
0	30	15	Full	XRD	Equal-spaced sampling	3	12
1	28	21	Full	XRD	ML suggests with DFT	12	16
1b	28	28	Full	XRD	ML suggests without DFT	10	18
2	28	28	Full	XRD	ML suggests with DFT, 1 manually added (two compositions are within 1% differences)	18	10
3	32	32	Full	XRD	ML suggests with DFT, 4 added manually (validation of local optimum)	21	11
4	28	7	Full	XRD	Selected 7 representative compositions, validation round	4	3
5	14	7	6 hours	GIWAXS	Samples degraded up to a time point for synchrotron measurement.	4	3
6	14	7	100 hours	GIWAXS	Samples degraded up to a time point for synchrotron measurement.	4	3

Table S2 List of 7 compositions selected for advanced structural, optical and morphological characterization.

Stability Region	Composition ID	Cs%	MA%	FA%	Ionic Radius (Å)	Tolerance Factor	Reason for in-depth characterization
I	a	13	0	87	241.820	0.963	Top binary cations Round 0 -3
I	b	13	8	79	238.940	0.957	Top binary cations Round 0 -3
Ii	c	26	74	0	204.000	0.884	Reference composition with no MA
Ii	d	26	38	36	216.960	0.911	ML local optimum
I	e	17	3	80	237.300	0.954	ML optimum
iii	f	5	17	83	242.886	0.969	Reference composition, state of the art
iv	g	0	100	0	206.150	0.891	Reference composition with no FA or Cs

2. Degradation Tests

2.1 The Setup

In total 196 individual thin-films within 7 experimental rounds were degraded in this study, including one initial round, three optimization rounds, and three rounds for validation and advanced characterization. Humidity, temperature, and visible light illumination level were controlled at 85±2°C, 85±5%, and 0.15±0.01Sun, respectively, within the house-built environmental chamber, enabling automated photographing of the samples and tracking of the chamber temperature and humidity. Samples were photographed in every 5 min during the aging test, and additionally the humidity and temperature of the aging chamber were tracked automatically. The illumination conditions remaining stable during the aging tests was confirmed by following a printed 28-patch reference colour chart that had been placed into the picture area. After the aging test, the samples were stored in a glovebox until further characterization.

The sample holder was made of graphite for maximum heat transfer, painted with medium grey colour for preventing over- or under saturating of the pictures taken, and designed to minimize reflections from the samples to the camera. Sample holder was heated from the bottom and its temperature was set to 85°C with SOLO 4824 temperature controller. The temperature was controlled using a thermocouple wired into the sample holder.

The chamber was illuminated using Advanced Illumination DL097 LED lamp and Advanced Illumination ICS 2.0 LED controller. The lamp provided even visible light only illumination. In this study, the illumination acts foremost as a necessary lighting of the samples for the camera, and also a degradation factor.Air humidity in the chamber was controlled and monitored using a control program running on Arduino, Si7021 humidity sensor, and water evaporation unit utilizing a fan. The stability of the humidity sensor during the aging tests was followed and the chamber conditions were tracked using EasyLog EL-USB-2 humidity-temperature logger. Both humidity sensor and logger were located close to the samples but with distance of centimetres from any surface possibly condensing water.

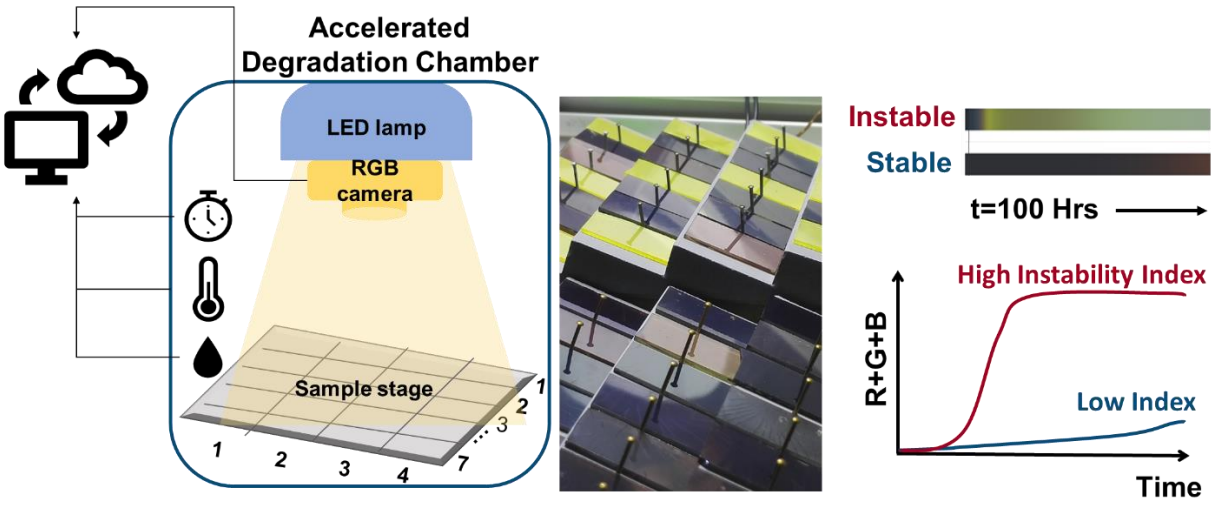


Figure S1 Schematics of the environmental chamber with in situ optical monitoring.

2.2 The Camera-based *in situ* Optical Imaging

2.2.1 Camera Integration and Colour Calibration

Samples were photographed using ThorLabs DCC1645C camera (with a removed IR filter to maximize data collection bandwidth of the camera) and ThorLabs MVL6WA lens, providing a 3-band approximation of the colour of the sample. The photographs of the samples were taken in BMP format in every 5 minutes using a control program implemented using LabVIEW. The photographs were saved automatically directly to Dropbox folder for further analysis. The photographs of the samples were automatically sliced into pictures of individual samples and colour calibrated to ensure reproducible and repeatable collection of quantitative colour data.

he data was used for determining the instability index for each sample as shown in the main article. The variations in the *Instability Index* (I_c) values arise from several sources: the variations between the sample replicates, variations in the environmental conditions between the sample holder locations, and in how camera sees the colour of the samples (e.g., reflections decrease the accuracy of colour determination). Experimental uncertainty is highly dependent on the composition and varies between 5000 px*min 225000 px*min. The estimated error for our optimized region of interest is 9000 px*min (with mean of approximately 30000 px*min). On average 7284 pixels were captured per thin film samples of 2.54 cm x 1.27 cm, leading to an estimated resolution of 210 µm.

Colour calibration by transforming the sample colours to a stable reference colour space ensures the colours are comparable even when the pictures are taken under different illumination conditions or with different camera-lens setups. This was done utilizing an X-Rite Colour Checker Passport with 28 reference colour patches that was photographed at the beginning of each aging test in the aging chamber. During the aging tests, camera settings were defined to be as illuminated as possible without oversaturating the white colour patch in Xrite colour chart, ensuring the colour calibration procedure succeeding without distortions.

The sample colours were transformed into a larger L*a*b colour space and clour warped by 3D thin-plate spline that has been shown to be among the most accurate colour warping methods for colour calibration.[2] Distortion between the colours of the reference colour chart in real and reference colour spaces is defined as:

$$D = \frac{\begin{bmatrix} V \\ O(4,3) \end{bmatrix}}{\begin{bmatrix} K & P \\ P^T & O(4,4) \end{bmatrix}},$$

where *O(n,m)* is a zero matrix with shape *m* x *n*, *V* represents the colours of the reference colour chart in the reference space (obtained from the colour chart manufacturer), matrix *P* represents the colours of the reference colour chart in the original space, and matrix *K* is the distortion between the colour patches in the reference and real colour spaces . The definitions for these matrices are:

$$V = \begin{bmatrix} 1 & x'_1 & y'_1 & z'_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x'_N & y'_N & z'_N \end{bmatrix},$$

$$P = \begin{bmatrix} 1 & x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_N & y_N & z_N \end{bmatrix},$$

$$K = \begin{bmatrix} 0 & \dots & U(r_{1N}) \\ U(r_{21}) & \dots & U(r_{2N}) \\ \vdots & \vdots & \vdots \\ U(r_{N1}) & \dots & 0 \end{bmatrix},$$

where *N* is the number of colour patches (in our reference colour chart *N*=24) and *U* is defined as:

$$U(r_{ij}) = 2r_{ij}^2 \, log(r_{ij} + 10^{-20})$$

(with a constant value added for numerical stability) and the Euclidian distance between the two colours is:

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

where x_i, y_i, z_i and x_j, y_j, z_j are the three colour components of each colour (in this study the components are *L*, *a*, and *b* in Lab colour space) in the real colour space and reference colour space, respectively.

Applying the same distortion *D* to the samples completes the colour warping:

$$\begin{bmatrix} V_s \\ O(4,3) \end{bmatrix} = \begin{bmatrix} K_s & P_s \\ P_s^T & O(4,4) \end{bmatrix} D,$$

where matrices V_s and P_s represent the colours of the samples in the reference and original space, respectively, and matrix K_s represents the distortion between the colours of the samples and

reference colour chart. The shape distortion between the colours of the samples and the reference colour chart patches in the real space is:

$$K_s = \begin{bmatrix} 0 & \dots & U(r_{s_1,c_N}) \\ U(r_{s_2,c_1}) & \dots & U(r_{s_2,c_N}) \\ \vdots & \vdots & \vdots \\ U(r_{s_M,c_1}) & \dots & 0 \end{bmatrix},$$

where $s_i$refer to sample colours, *M* is the number of samples, and c_i refer to reference colour chart colours. Matrices V_s and P_s that represent colours of the samples in the reference and real colour spaces, respectively, are:

$$V = \begin{bmatrix} 1 & x_1' & y_1' & z_1' \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_M' & y_M' & z_M' \end{bmatrix},$$

$$P = \begin{bmatrix} 1 & x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_M & y_M & z_M \end{bmatrix}.$$

After colour calibration, the sample colours were transformed back to RGB space.

2.2.2 Quantified Colour Change Over Test Duration

The decomposed and calibrated RGB values were determined for each degrading sample. This was evaluated across the film surface, and therefore resulted in a spatial distribution of colouration; this was evaluated by determining the mean, 5% threshold, and 95% threshold for each colour channel within each film. These colour channels were evaluated separately, and also accumulated to yield an instability measure at each time point. This accumulation consisted of the sum of the R, G, and B mean values at a time point, normalized by the initial sum at zero minutes of degradation such that each sample had an instability measure of zero at the initial time recorded. The mean RGB was likewise visually represented as colour bars, relating the normalized average colour evaluated as a DI to an absolute recorded average colour over time.

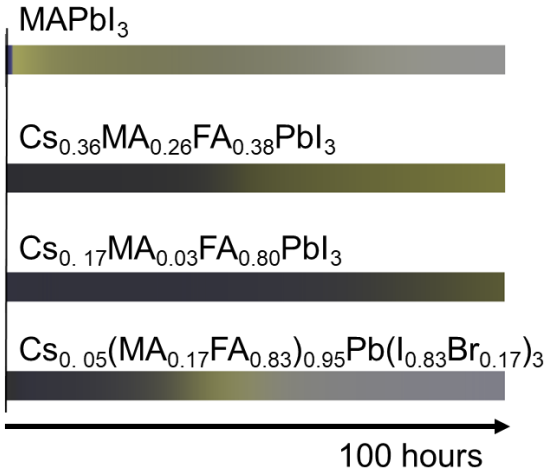


Figure S2 In situ monitored colour evolution of four representative compositions during the degradation test, quantified R+G+B colour changes of the compositions are shown in Figure 3b in the main article.

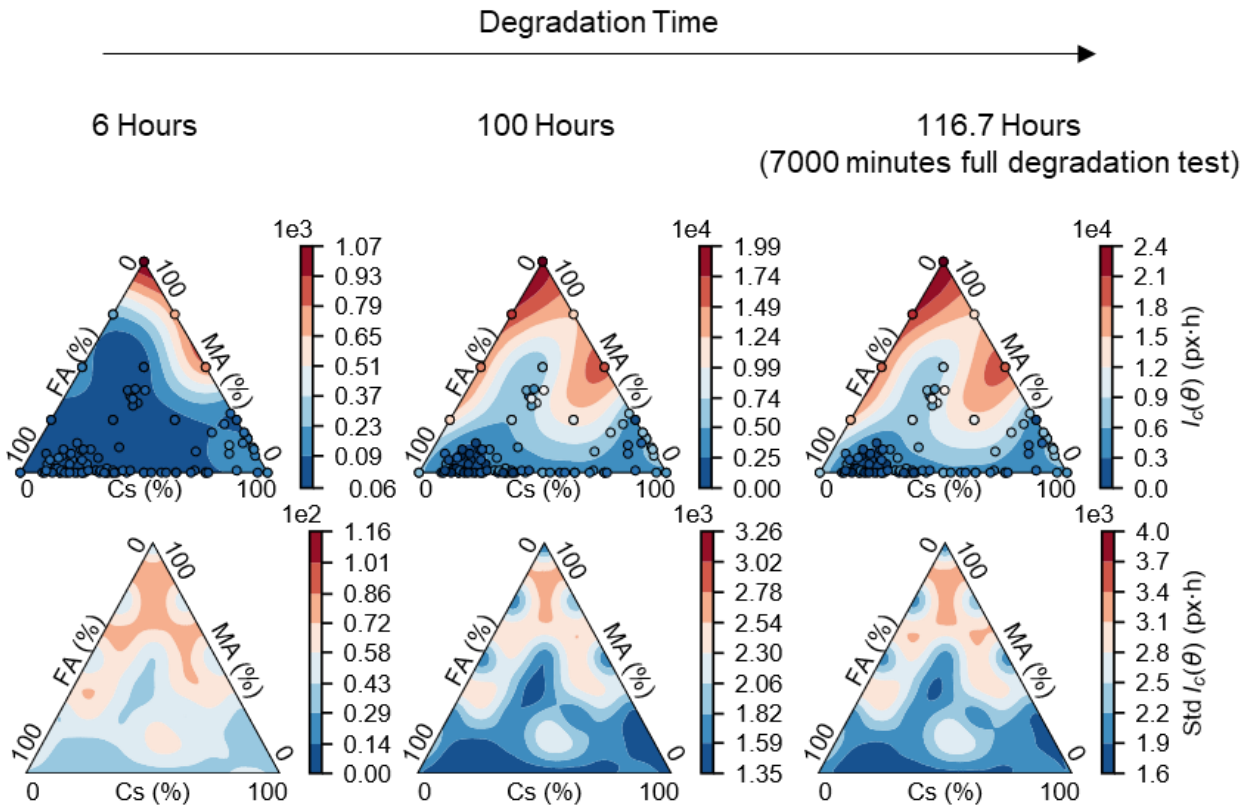


Figure S3 Integrated colour change over the first 6h (left column), 100h (middle column), and 7000min (right column) of the aging test. 7000min result corresponds to the instability index. The resulting integrated colour changes (first row) and standard deviations (second row) are modelled using samples from rounds 0-3 and shown as contours. Each sample is shown as a dot with a colour corresponding to its experimentally measured integrated colour change over the integration time.

Figure S3 illustrates the integrated colour changes and the uncertainties over the full compositional space using a GP model for three different integration (degradation test) periods. The results illustrate that after 6h of degradation samples that do not have FA and have high MA content have clearly changed their colour, whereas other samples have gone through small, uniform colour changes. By 100h of degradation, the samples have already gone through most of the colour change seen at the final stage after 7000min, and the stable regions have separated from the rest of the space. The instability landscape derived from the colour change is consistent with the bandgap change trends (Fig. S13), degradation products and structural changes measured by synchrotron-based GIWAXS at the same timepoints. GIWXAS (Fig. S15) validates that mixed-cation perovskites have gone through minor degradation after 6 hours, leading to colour changes, and the more stable regions are not distinguishable from less stable ones. After 100 hours of degradation tests, the differences between the stable regions are visible through colour changes.

The *Instability Index* (I_c) in this study is defined as the integrated colour change over the 7000 min aging test duration (main article Eq. 1). Full list of samples and measured instability index are shown in Appendix Table A2.

3. Machine Learning Framework

3.1 Physics-constrained Bayesian Optimization (BO)

Bayesian optimization is used to find Θ^*, the perovskite composition with smallest instability index I_c, using few samples as possible, according to:

$$\Theta^* = argmin_{\Theta}\, \mathbb{E}[I_c(\Theta)] \quad \text{s.t. } P(\Delta G_{mix}(\Theta), \beta_{DFT}) > 0$$

where $\Theta^* = arg_{\Theta} min\, \mathrm{E}[I_c(\Theta)]$ is a BO framework optimising the I_c, and $I_c(\Theta)$ is a noisy, black-box function that we evaluate by making samples at different Θ compositions. We propose to use the probabilistic constraints from DFT-computations given by $P(\Delta G_{mix}(\Theta), \beta_{DFT})$. The probabilistic constraint models the probability of phase according to Gibbs Free Energy $\Delta G_{mix}(\Theta)$ and data-fusion factor β_{DFT}.

Our chosen acquisition function is *expected improvement* $\mathrm{EI}(\Theta)$,

$$\mathrm{EI}(\Theta) = (\mu_n(\Theta) - \tau)\Phi\left(\frac{\mu_n(\Theta) - \tau}{\sigma_n(\Theta)}\right) + \sigma_n(\Theta)\, \phi\left(\frac{\mu_n(\Theta) - \tau}{\sigma_n(\Theta)}\right)$$

where Φ is the standard normal cumulative distribution, $\mu_n(\Theta)$ is the mean of the GP model's posterior, τ is an incumbent best point, σ_n is the variance of the GP model's posterior, and ϕ is the standard normal probability distribution. By maximizing $\mathrm{EI}(\Theta)$, the most promising location Θ^* for the next experimental round is determined. To suggest more than the one promising composition Θ per round, we employ the local penalization algorithm to resample $\mathrm{EI}(\Theta)$, as described in [3].

To include the physical constraints in the Bayesian optimization algorithm, the acquisition function is weighted according to the probabilistic model $P(\Delta G_{mix}(\Theta), \beta_{DFT})$ following the method in ref.[4]:

$$\mathrm{EIC}(\Theta) = \mathrm{EI}(\Theta)\, P(\Delta G_{mix}(\Theta), \beta_{DFT})$$

These discounts the probability of phase separation coming from DFT, allowing to consider first-principles calculations into the experimental optimization loop. Figure S4 presents various rounds of Bayesian optimization, included the GP models mean and standard deviation, along with the DFT-weighted and unweighted acquisition function. The effectiveness of the constraint to reduce the search space and avoid non-stable optima.

In Round 0, a grid of 28 evenly-spaced sample compositions (15 compositions) was aged, and I_c was quantified for each composition. In each subsequent round, we use Bayesian optimization to efficiently sample compositions and iteratively approach the global optimum. The Bayesian optimization algorithm is implemented based on the GPyOpt optimization package.[5] In the Bayesian optimization setting, a surrogate machine learning model, Gaussian Process (GP) regression, is used to approximate the mean and uncertainty of $I_c(\Theta)$ in non-sampled regions of the compositional space. For our applications, we use a GP model with a Matérn 5/2 kernel, with kernel parameters chosen to optimize the maximum likelihood. The model is updated in each experimental round. Once the model is fitted, an *acquisition function* is maximized to suggest

locations in the compositional space with high chance of leading to an optimum, either due to a low expected instability index or a high uncertainty at a certain composition (*i.e.* more chances to produce an optimal composition). Our Bayesian optimization algorithm uses two convergence criteria: the modelled mean stabilized, and the instability index of the most stable compositions found reached the experimental uncertainty of our test setup. The convergence of the instability index towards the optimum is illustrated in Figure S5 and Figure S6.

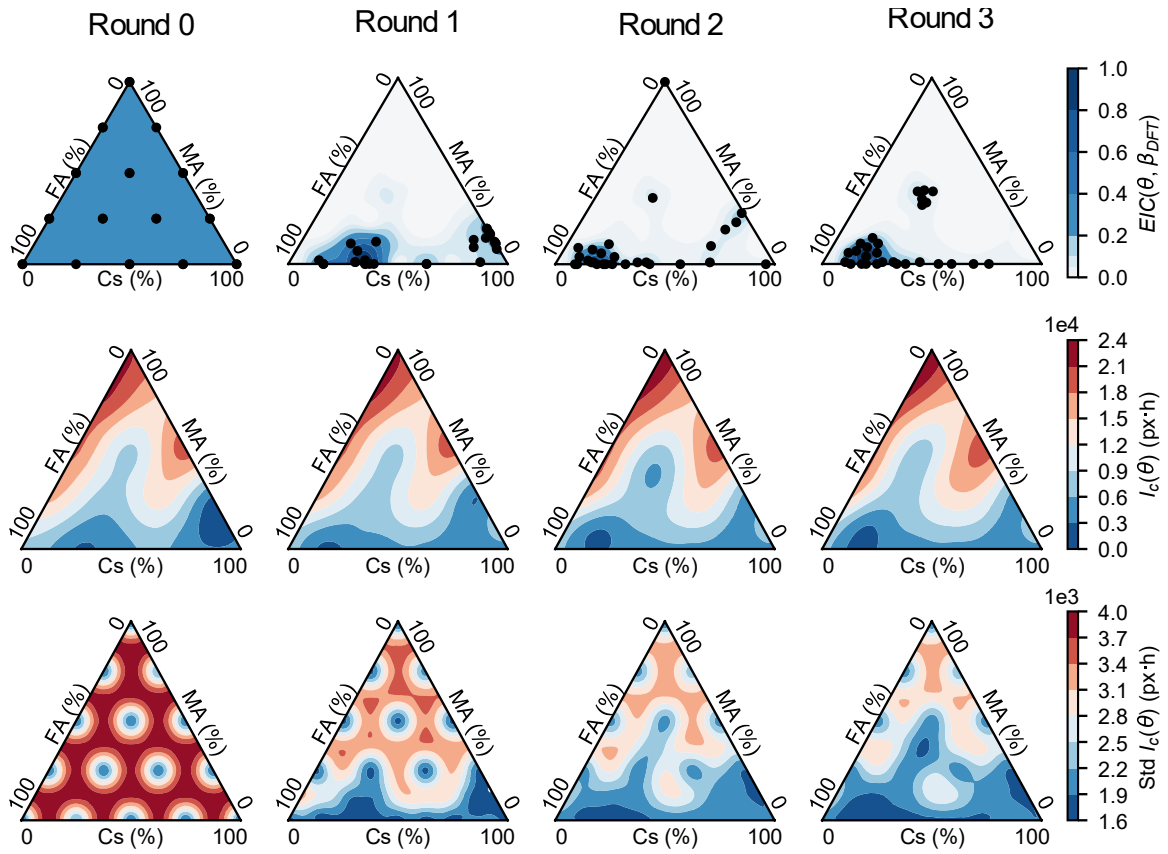


Figure S4 DFT-weighted acquisition function, posterior mean, and uncertainty landscapes over three optimization rounds.

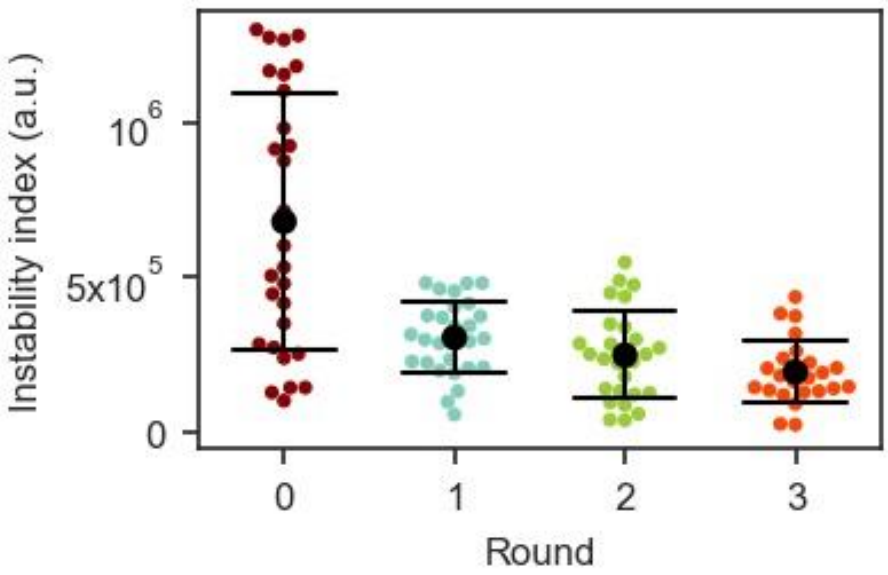


Figure S5 Instability index in logarithm scale for each degradation round with mean and standard deviation.

3.2 DFT Calculations, Data Fusion and Probabilistic Constraint

3.2.1 Gibbs Free Energy Calculations

Alloy structures are created using the pseudo-cubic as the starting structure for the pure compositions (obtained from Ref.[6]) with random substitution at the A-site. Special quasi-random structure (SQS) method,[7] as implemented in the ATAT package, [8,9], is used to obtain structures for various A-site alloy compositions. Two different supercell sizes 96 and 144-atoms are used and for each A-site composition with multiple (two or three) structures, varying in the orientation of the MA and FA molecules, are considered. All the DFT relaxed modelled structure files of binary and ternary A-site alloys are provided with the supplementary information.

DFT calculations are performed within the projected augmented wave (PAW) method[10] as implemented in the VASP code.[11] The Perdew Burke Ernzerhof (PBE) exchange correlation functional[12] is used with GGA and spin-orbit coupling (SOC) is included in the total energy calculations. Plane wave cutoff of 340 eV, and a Monkhorst-Pack k-point sampling scheme[13] is used. All degrees of freedom (cell shape, volume, and ionic positions) are relaxed in DFT calculations. Following the relaxations, A-site alloy structures at various compositions are found to retain the overall cubic symmetry.

To model the phase stability of mixed A-site halide perovskites, we compute the Gibbs free energy of mixing ($\Delta G_{mix} = \Delta H_{mix} - T\Delta S_{mix}$) of these materials as a function of the A-site composition. The modeled ΔG_{mix} has two components, (1) the enthalpy of mixing (ΔH_{mix}), and the entropy of mixing (ΔS_{mix}). The enthalpy of mixing is calculated from DFT by taking the difference between the total energy of the mixed A-site halide perovskite with respect to the total energy of the constituent, or pure, phase. The temperature dependence ($T\Delta S_{mix}$) to Gibbs free energy is incorporated by considering the entropic contributions associated with the configuration and rotations degrees of freedom, as discussed in detail in Ref.10. Further details of thermodynamic modeling can be

found in the Ref 5. As benchmarked in Ref.[14], the reproducibility and precision in our DFT total energy calculations is very high. The variability in the computed value of Gibbs free energy at a specific composition is between 5 – 20 meV/unit, and it comes from the varying orientation of the FA and MA molecules between the multiple structures, considered in our simulations.

The key sources of uncertainty in this study relating to incorporating DFT calculations into the experimental loops consist of three aspects: 1) the intrinsic precision limit of the DFT calculations leads to small differences over different computational runs of the same organic-inorganic structure (~ 0.025 $(\mathrm{eV/f.u.})$), and 2) the uncertainty of the regression model that map the sparse DFT data over a continuous compositional space, and 3) the differences between simulation environment and experimental conditions due to kinetic factors.

3.2.2 Regression Model for Gibbs Free Energy

Calculating ΔG_{mix} for ternary compounds is computationally expensive and infeasible over the whole composition space; therefore, ΔG_{mix} is modelled using an auxiliary GP regression model trained on 91 DFT-computed ΔG_{mix} from 47 binary perovskite compositions.

Based on the DFT data, the $\Delta G_{mix}(\Theta)$ is approximated over the ternary composition space, as shown in Figure S6. The auxiliary GP model uses a radial basis function (RBF) kernel with initialized length scale of $0.003\ (\mathrm{eV/f.u.})$ and variance of $0.025\ (\mathrm{eV/f.u.})^2$. The kernel, initial length scale and variance values were determined based on grid search, and optimized according to maximum likelihood.

The model was validated using leave-one-out cross validation, which resulted in the mean squared error $\mathrm{MSE} = 0.0004\ (\mathrm{eV/f.u.})^2$. Additionally, the final model was tested against three datapoints that are out of the validation dataset (ternary triple cation phase stability calculations in Ref.[15]), resulting in $\mathrm{MSE} = 0.00015\ (\mathrm{eV/f.u.})^2$. Both MSE values are low in comparison to the $\Delta \mathrm{G_{mix}}$ of the training set, which shows that the model regresses $\Delta \mathrm{G_{mix}}$ successfully for ternary compositions.

3.2.3 Definition of Phase Stability Constraint

Once $\Delta \mathrm{G_{mix}}$ is modelled, the probabilistic model $P(\Delta G_{mix}(\Theta), \beta_{DFT})$ can be computed. The inherent bias and precision of DFT calculations justifies the probabilistic treatment of the constraint.

The regressed $\Delta G_{mix}(\Theta)$ is integrated as probabilistic constraint according to a probabilistic model given by the logistic cumulative distribution function:

$$P(\Delta G_{mix}(\Theta), \beta_{DFT}) = \frac{1}{1 + e^{-\Delta G_{mix}(\Theta)/\beta_{DFT}}}$$

Referring to the constraint value as $\Phi = P(\Delta G_{mix}(\Theta), \beta_{DFT})$, we formulate the data fusion process as maximizing the likelihood $\mathcal{L}(\beta_{DFT;}\Phi, \Delta G_{mix})$ such that:

$$\beta^*_{DFT} = arg_{\beta_{DFT}} \mathcal{L}(\beta_{DFT;}\Phi, \Delta G_{mix})$$

Using a Bernoulli likelihood, this definition is equivalent to fitting a logistic regression model via maximum likelihood, with ΔG_{mix} as the independent variable and the probability of phase de-mixing as the dependent variable. In this sense, one could estimate a certain critical energy $\Delta G_{mix-critic}$ above which the crystalline structure is unstable and will decompose into its constituent phases. In the context of convex hull stability calculations, this value is often

considered to be around -0.025 eV/f.u. We hence choose a cumulative probability of $P\left(-0.025eV/f.u,\beta^*{}_{DFT}\right)$=0.7 and $P\left(-0.05eV/f.u,\beta^*{}_{DFT}\right)$=0.9. This assumption defines a smooth limit for phase stability in the compositional space, taking into account the inherent uncertainty of first-principles calculations. Our choice of probabilistic model is common in machine learning literature, due to the simplicity and expressivity of logistic models.[16]

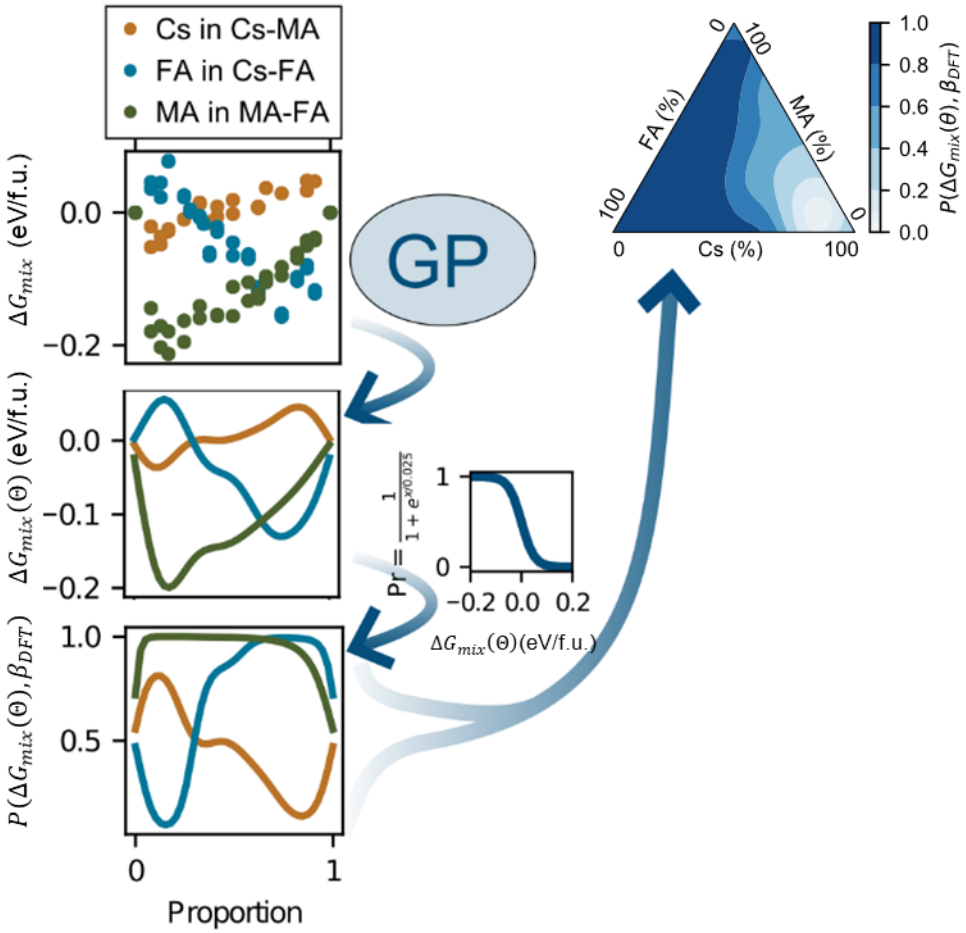


Figure S6 Construction of Gibbs free energy model using DFT-calculated Gibbs free energy of mixing.

3.2.4 Data Fusion Improves Search Effectiveness

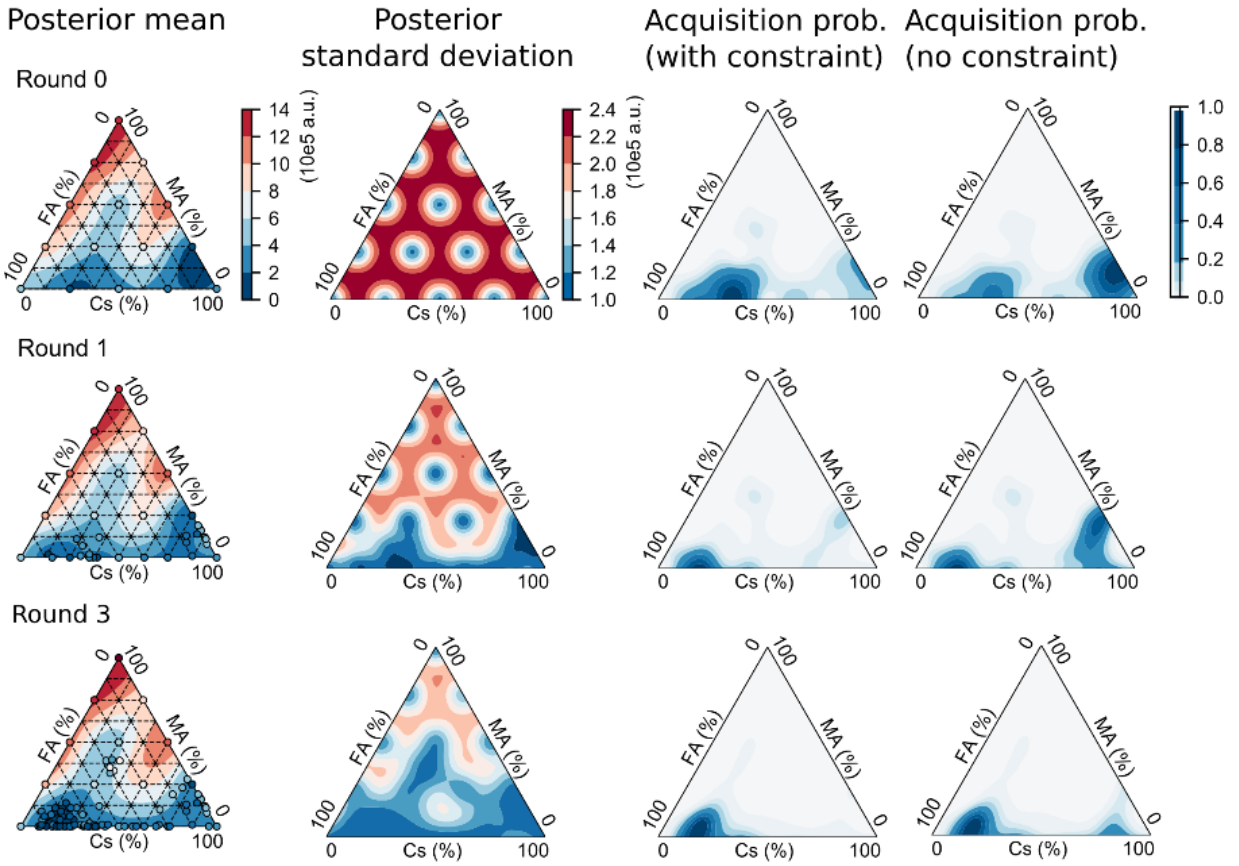


Figure S7 DFT-constrained acquisition function with samples of each round shown as black dots (the first row) and the resulting modelled instability index (the second row) with standard deviation (the third row) shown for each optimization round.

We use teacher-student model to compare the performance of Bayesian optimization fused with DFT data (BO_DFT) and standard Bayesian Optimization (BO_baseline).[17] Firstly, we train a GP surrogate model as the "ground-truth" model and use it to calculate the instability index (metric) of newly suggested experimental conditions from both BO_DFT and BO_baseline. We feed the results of the initial run (run 0) to the above two different Bayesian optimization methods and suggest new conditions iteratively from run 1 to run 4, in order to simulate our experimental optimisation rounds. The teacher-student model allows comparison of those two methods without performing the actual time-consuming degradation experiments.

Figure S8 shows the loss evolution of the student model with batch size of 28. The median loss decreases for both and BO_DFT. BO_baseline has lower instability metric than BO_DFT because BO_baseline allocates most of points in the lower right corner of the phase diagram (Figure S9). This region has phase separation from the XRD measurement (Figure S11) despite minimal colour change of films in this region. BO_DFT avoids exploring that area with DFT data fusion and finds the local minimum at lower left corner of the phase diagram.

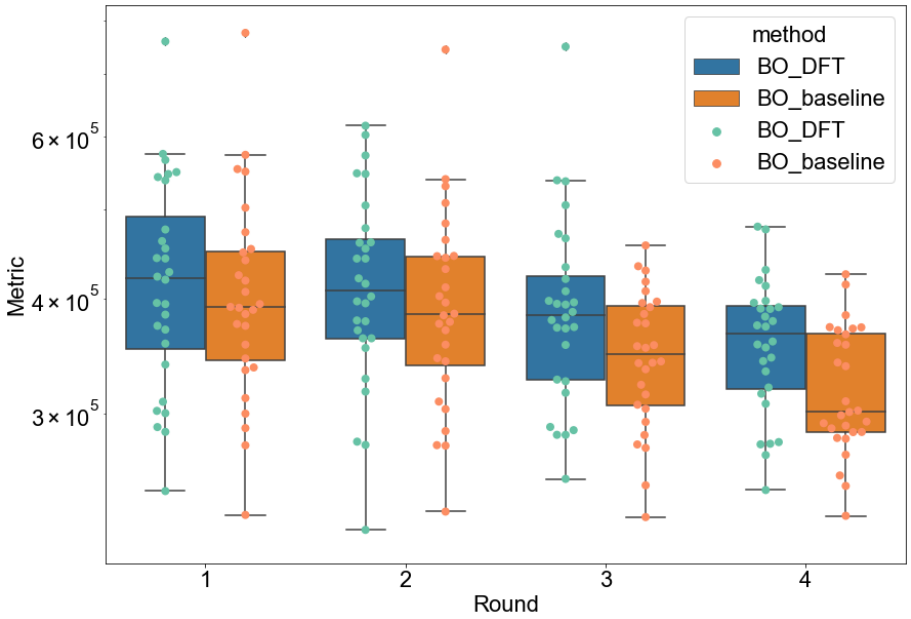


Figure S8 Loss evolution of student model, batch size=28

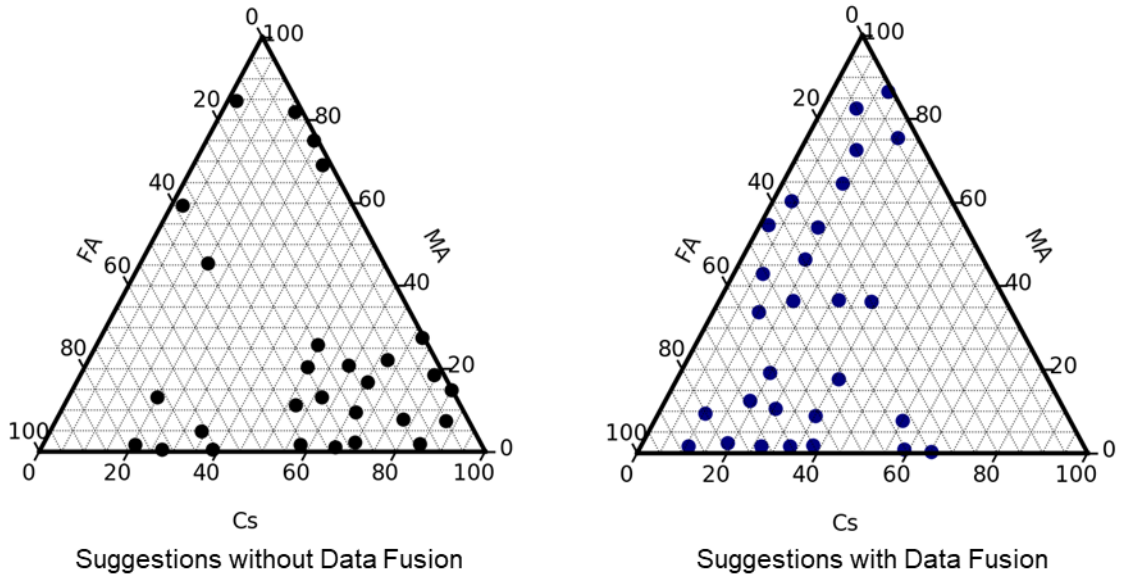


Figure S9 Suggested points of student model for Round 4, batch size=28

4. Laboratory Characterization

4.1 X-ray Diffraction and Phase Analysis

Grazing incidence X-ray diffraction (incident angle of 1°) were performed on 94 compositions from Round 0-3 samples using Rigaku SmartLab with Cu-*K*α sources on the as-synthesized thin films to understand the crystal structures and to examine the formation of minority phases.

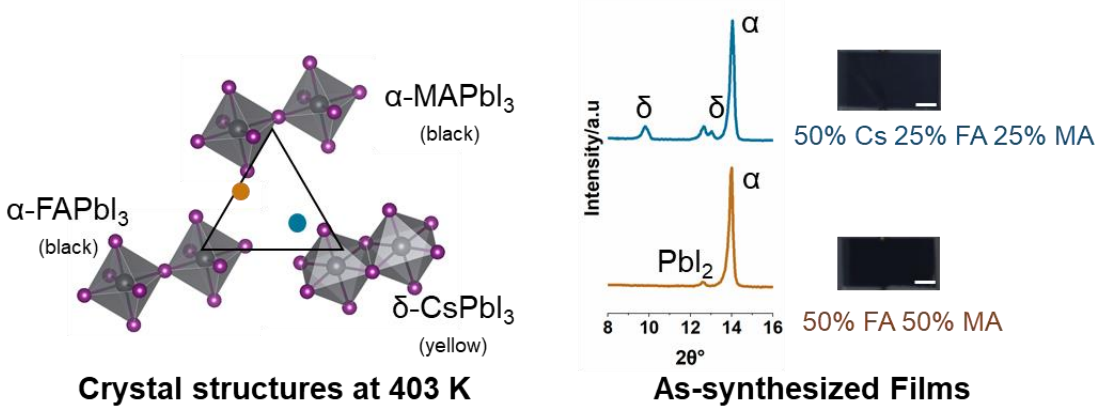


Figure S10 Crystal structures of the end members of the $Cs_xMA_yFA_{1-x-y}PbI_3$ phase space at room temperature. Minority phases are detected via X-ray diffraction measurement on as-synthesized thin film samples. Overstochemetric lead iodide was added in the synthesis following reference,[1] leading to PbI_2 crystallization in as-synthesized films.

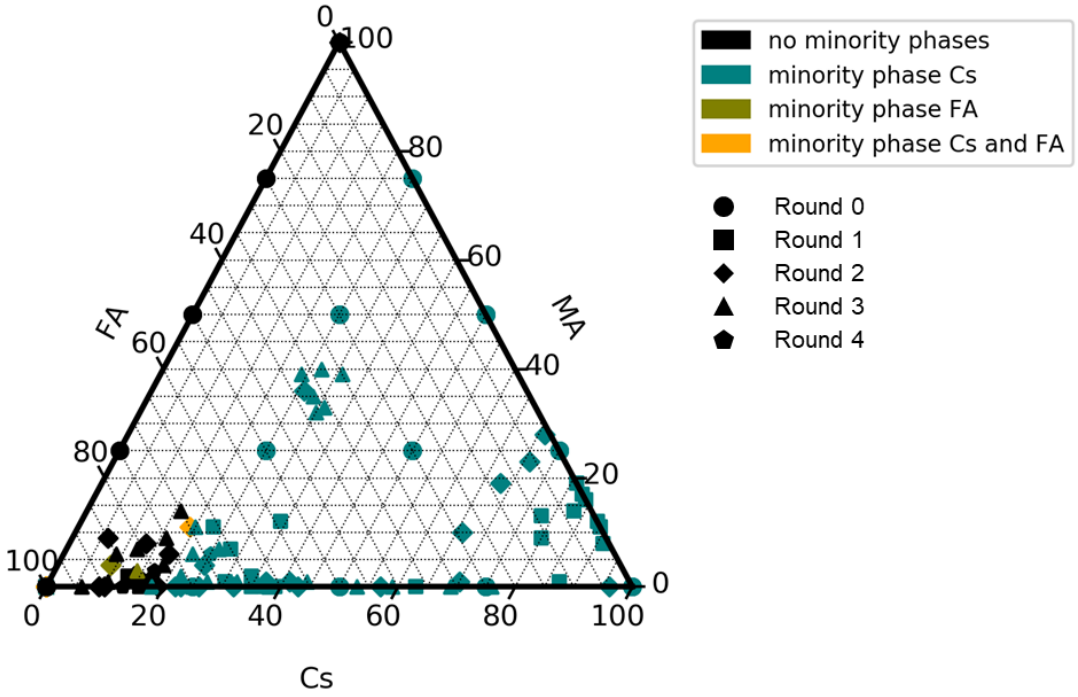
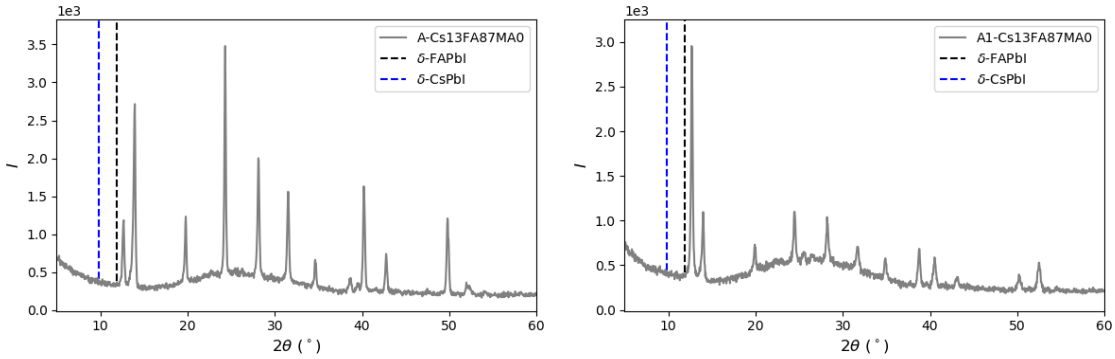
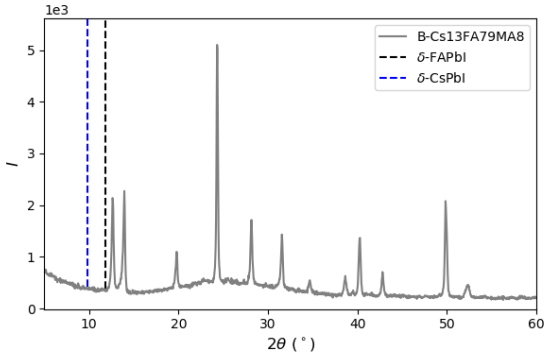


Figure S11 Phase diagrams constructed with as synthesized thin films. No minority phases indicate single-phase α-perovskite, minority phase Cs indicates the existence of δ-$CsPbI_3$, minority phase FA indicates the existence of δ-$FAPbI_3$ crystalized at 403 K thin-film annealing temperature followed by cooling to 300 K room temperature in N_2 glovebox prior to degradation tests.

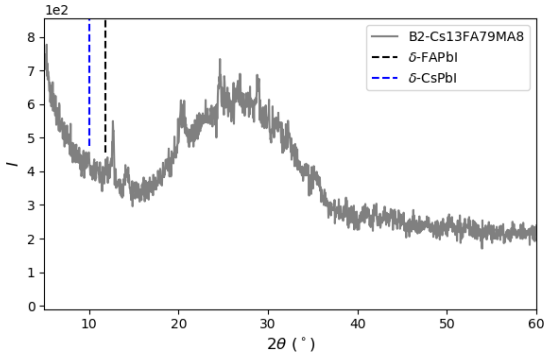


a – 0 hours

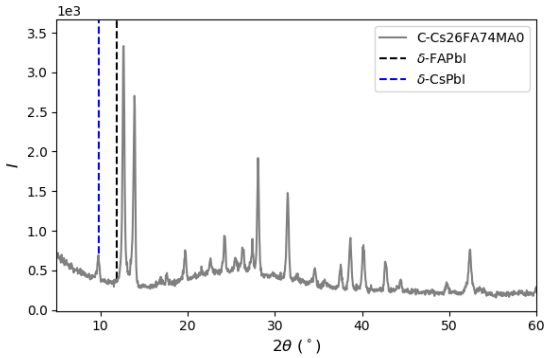
Degraded (full degradation run)



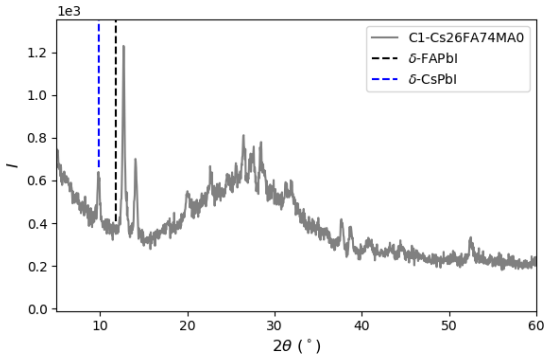
b – 0 minutes



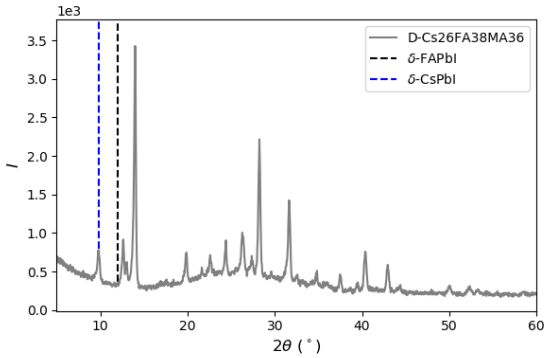
Degraded



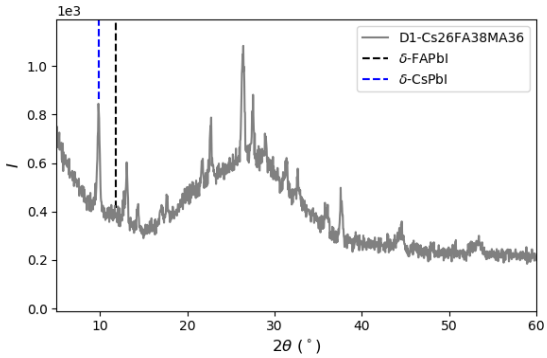
c – 0 minutes



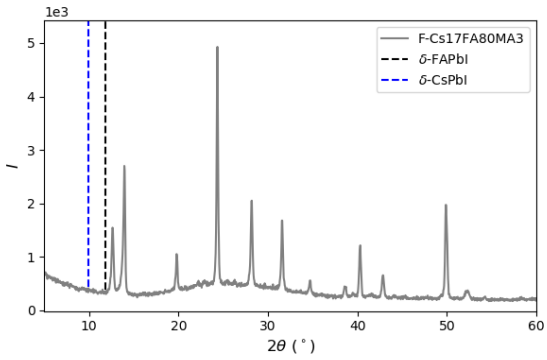
Degraded

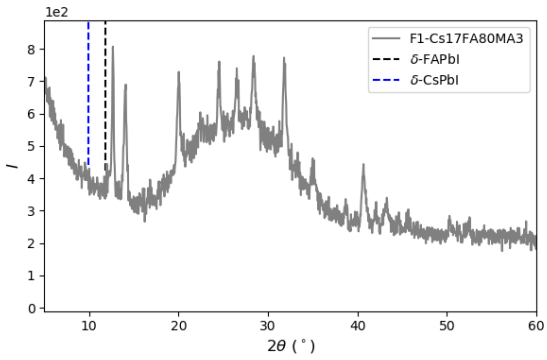


d – 0 minutes



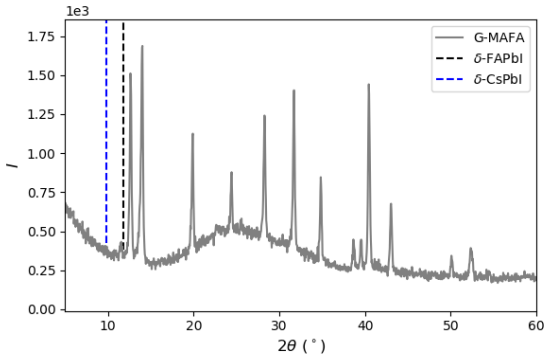
Degraded

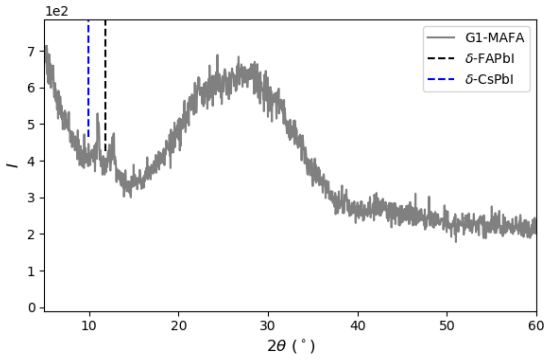




e – 0 minutes

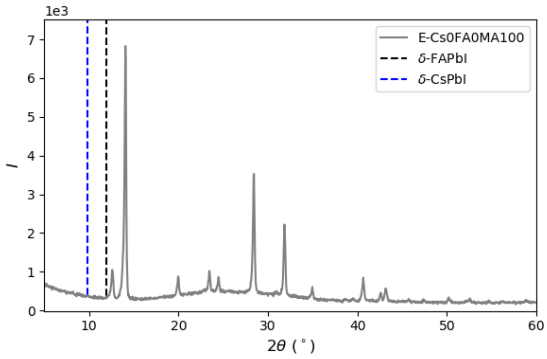
Degraded

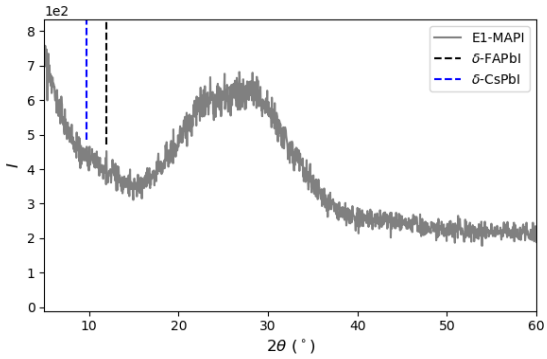




f – 0 minutes

Degraded





g – 0 minutes

Degraded

Figure S12 X-ray diffraction of 7 distinct compositions (a-g) from Round 4 degradation run, showing structural changes before and after a full degradation run. The labelled minority phases are according to: δ-CsPbI3: ICSD number: 250744; α-perovskite: ICSD: 250735, and δ-FAPbI3: ICSD: 230491

4.2 UV-Visible Spectroscopy

The absorptance for the films was calculated based on transmission and reflection measurements done using Perkin-Elmer Lambda 950 UV/Vis Spectrophotometer (Perkin-Elmer). Bandgaps were calculated using Tauc methods assuming direct bandgaps.

Table S3 Bandgap changes during degradation tests..

	Bandgap (eV)*						
Time (hours)	a	B	c	d	e	f	g
0	1.48	1.5	1.48	1.5	1.49	1.58	1.55
6	1.48	1.5	1.49	1.5	1.51	1.58	2.28
100	1.61	1.64	2.61	1.67	1.61	2.64	2.6

*Estimated error of ±0.03 eV.

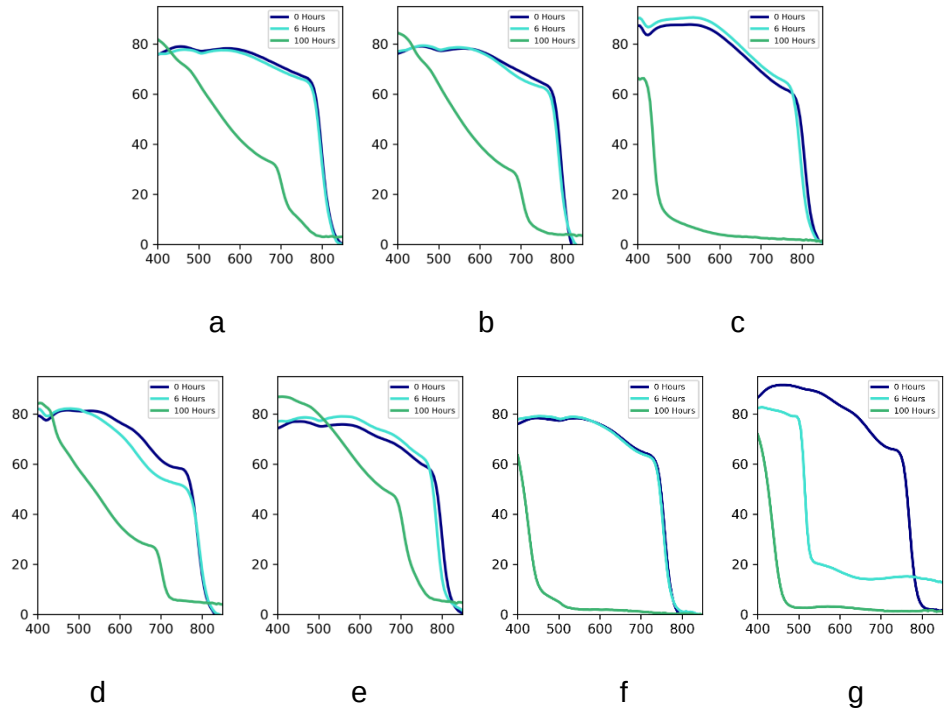


Figure S13 Absorption spectra of composition a-g, dark blue: as synthesized, cyan: after degradation for 6 hours, green: after degradation for 100 hours.

.

4.3 Scanning Electron Microscopy (SEM)

The film morphology was investigated using a ZEISS Ultra-55 field-emission scanning electron microscope (FESEM, ZEISS), with InLens detector and 3.00 kV EHT gun. The grain sizes were counted using ImageJ within the area of ~0.72 μm^2. Various compositions show different distribution of grain sizes. The grain sizes of composition with high Cs ($Cs_{0.26}FA_{0.74}PbI_3$), with Cs > 20%, are mostly between 300-500 nm and some of them reach up to 1200 nm, indicating the presence of δ-phase. The grain sizes of $MAPbI_3$ are mostly between 200-400 nm, and reach up

to 1000 nm. The grain sizes of the rest of the compositions, which have low Cs (< 20%) are mostly between 200-400 nm.

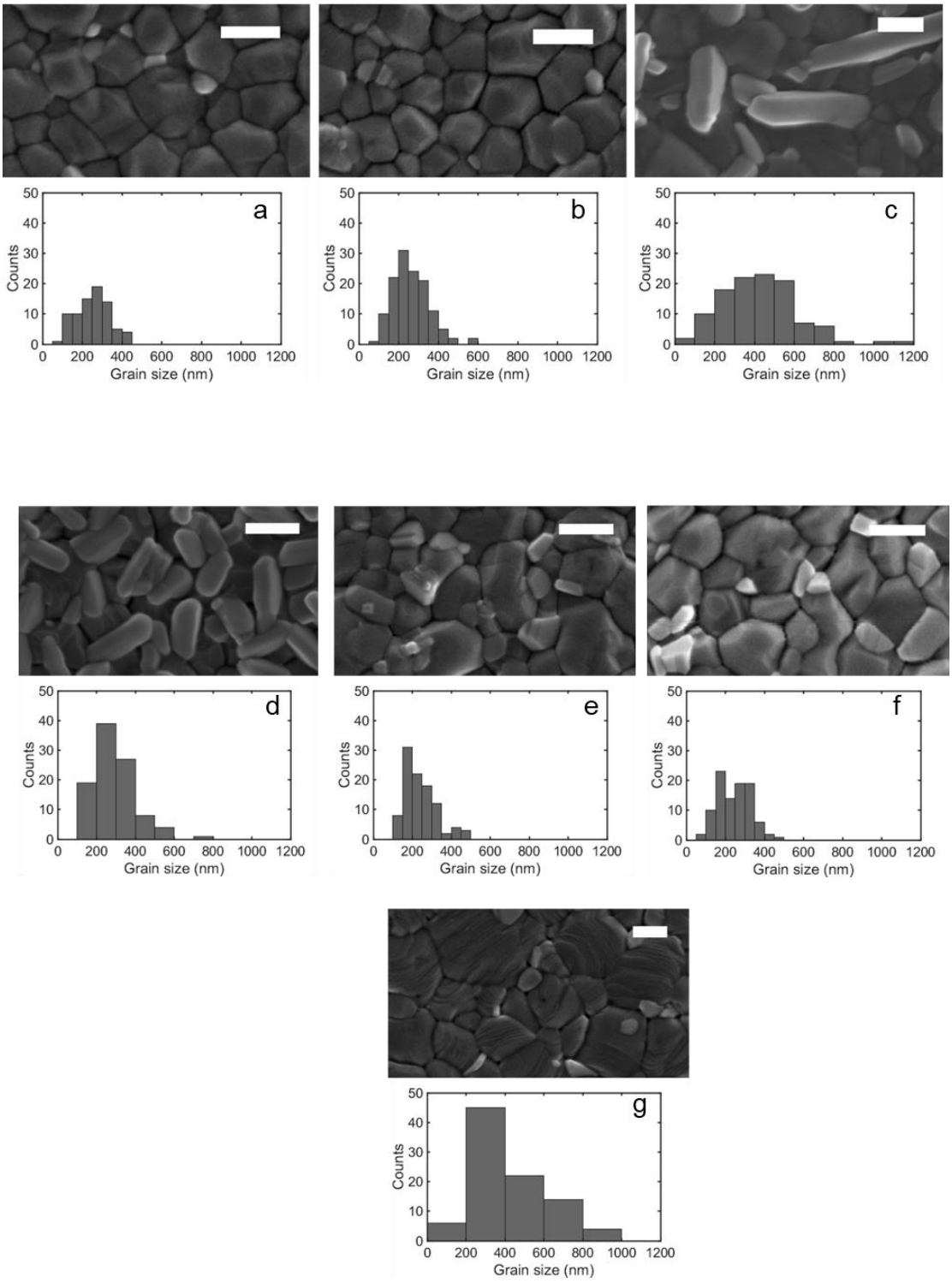


Figure S14 SEM images and grain size count for as-synthesized composition a-g in Round 4. The scale bar is 300 nm. Rod-like δ-$CsPbI_3$ minority phases are visible in c and d.

5. Synchrotron Characterization

5.1 GIWAXS Measurements

Grazing-incidence wide-angle x-ray scattering (GIWAXS) measurements were taken at beamline 11-BM (CMS) at the National Synchrotron Light Source II (NSLS-II) of Brookhaven National Laboratory. The x-ray beam with the energy of 13.5 keV shone on the perovskite films in the grazing incident geometry. The data presented in the study was taken at incident angle θ = 0.2 which probes the bulk structure of the films. The scattering spectra were collected with the exposure time of 30 seconds by an area detector (DECTRIS Pilatus 800K) placed 257 mm away from the sample. The data analysis was performed by using custom-made software SciAnalysis.[18,19]

Table S4 Sample list for GIWAXS examination. The Composition ID follows the ID in Table S1.

Region	Composition ID	Cs%	MA%	FA%	Duration of Degradation (Hours)
I	A	13	0	87	0, 6, 100
I	B	13	8	79	0, 6, 100
Ii	C	26	74	0	0, 6, 100
Ii	D	26	38	36	0, 6, 100
I	E	17	3	80	0, 6, 100
Iii	F	5	17	83	0, 6, 100

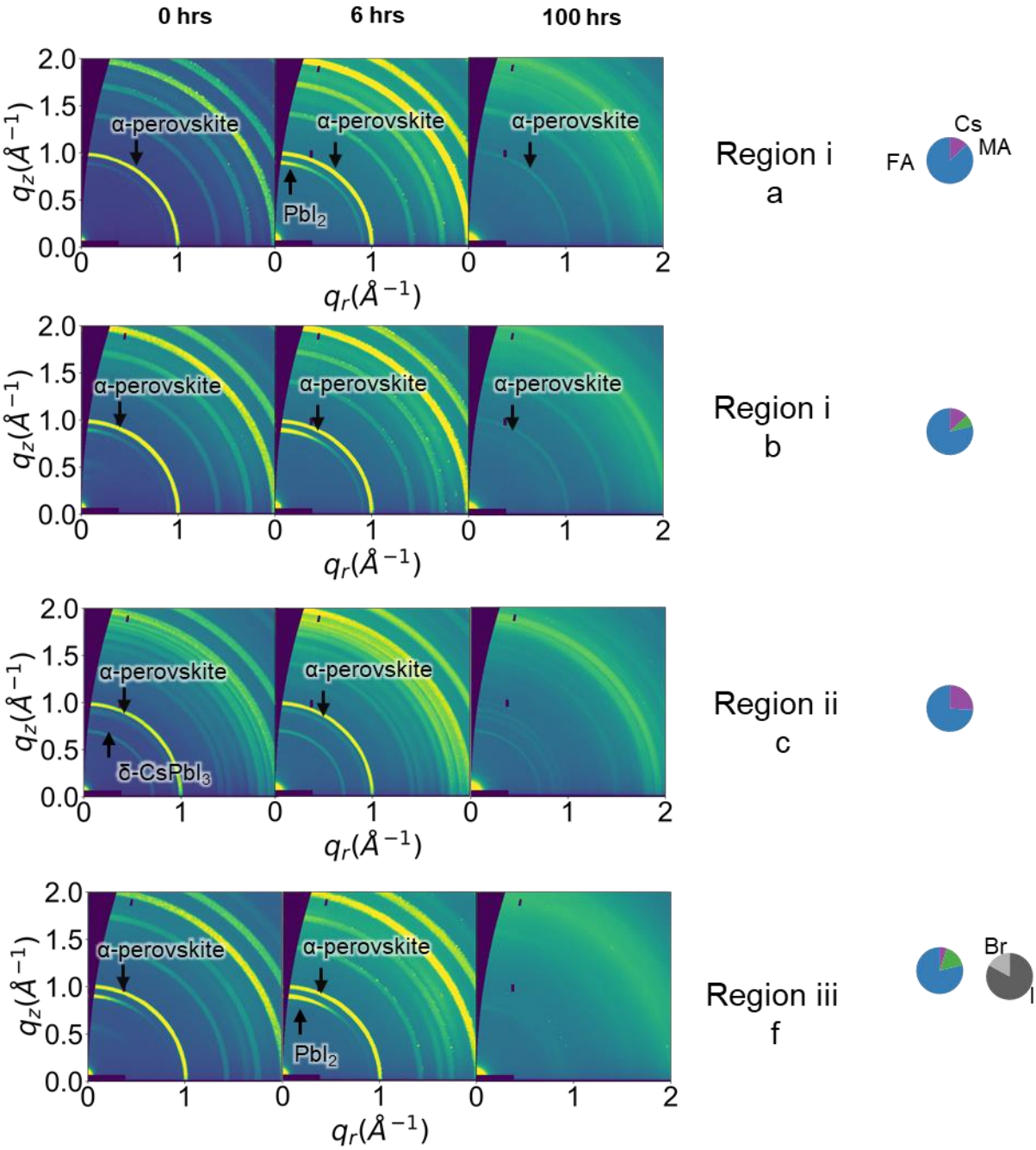


Figure S15 GIWAXS images with data acquisition at incident angle of θ = 0.2° for composition a, b, c, and f after 0, 6 (Round 5), and, 100 hours (Round 6) of degradation tests. GIWAXS images of d and e are shown in main text Figure 4a.

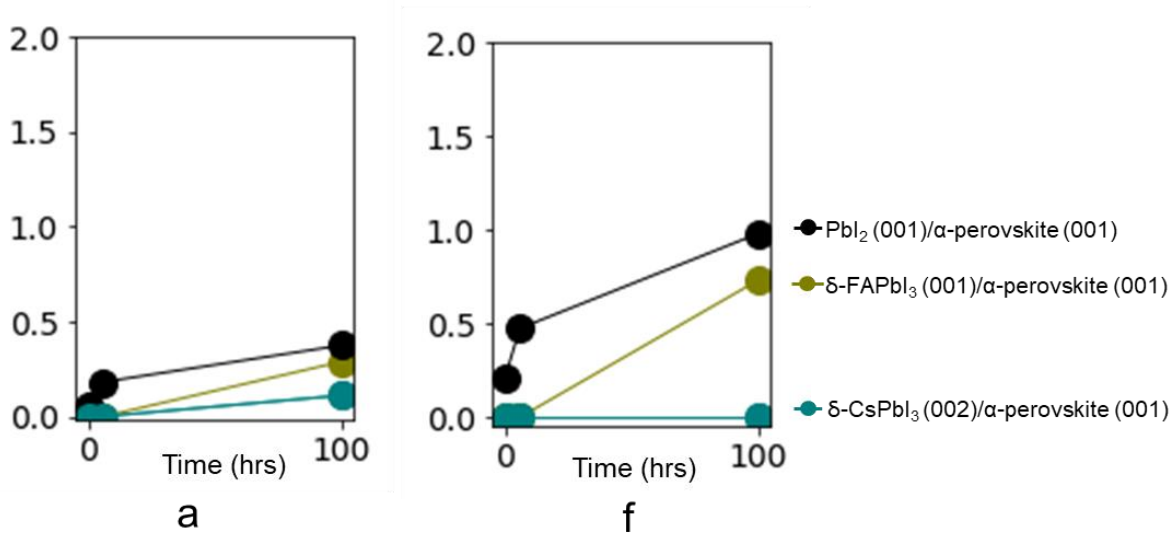


Figure S16 Relative peak insensitive of GIWAXS data (circular average) for composition a and f after 0, 6 (Round 5), and, 100 hours (Round 6) of degradation tests. Figures of b-e are shown in main text Figure 4 b-e.

5.2 Data Analysis

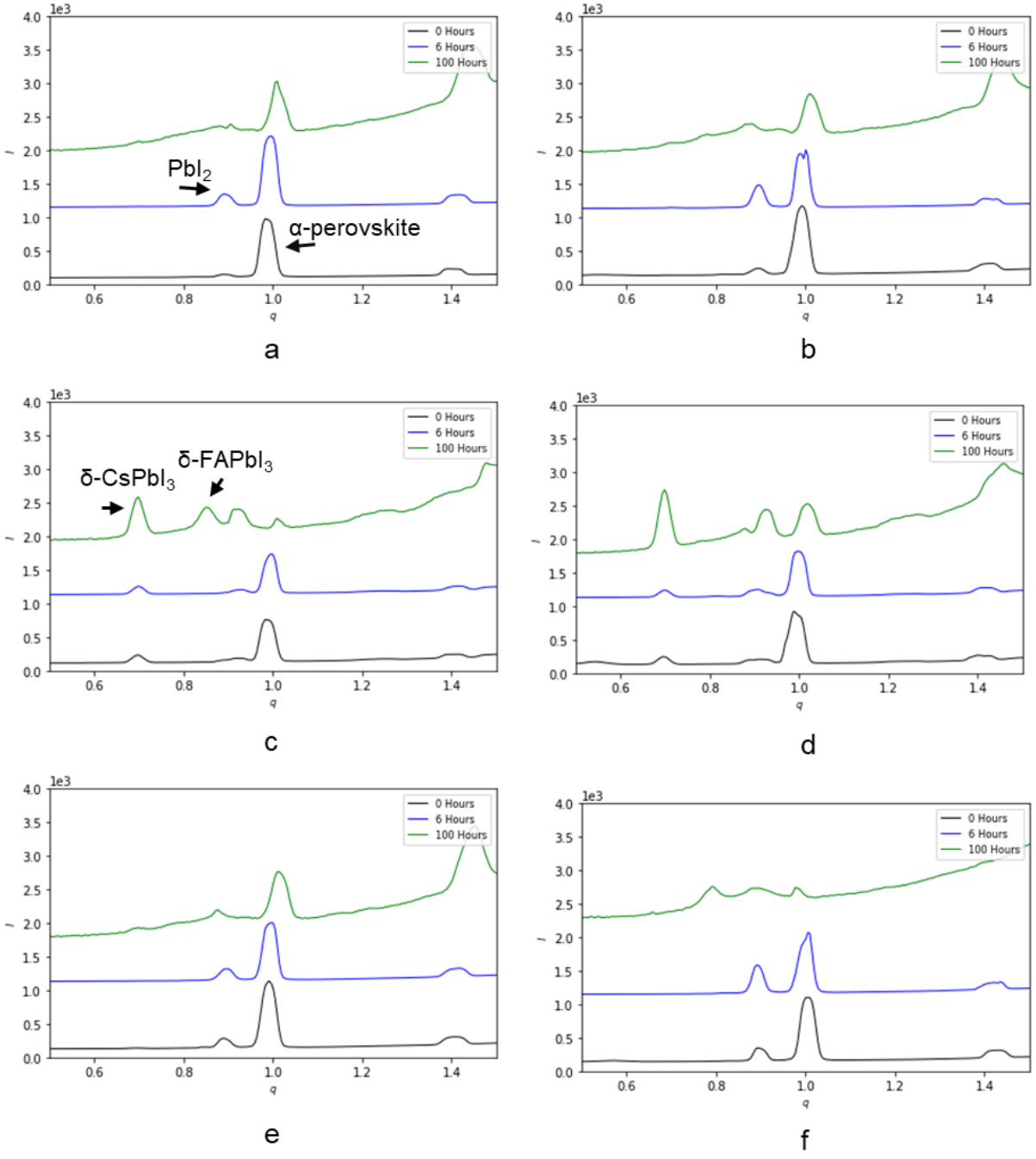


Figure S17 Integrated peak intensities for compositions a, b, c, d, e, and f after 0 ,6, and 100 hours of degradation tests respectively. We observe the emergence of PbI_2, δ-$CsPbI_3$ and δ-$FAPbI_3$ during the course of degradation test. We observe a lattice shrinkage for all the iodide-based perovskites, including both CsFA, and CsMAFA perovskites.

Table S5 Peak extraction from GIWAXS images

0 Hours		Peak Positions (*q*) and Intensities (*I*) (Circular Average)										Background-subtracted				
	Background	Perovskite (001)		CsPbI3 (002)		FAPbI3 (001)		PbI2 (001)		CsPbI3 (102)		Perovskite	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)
	I	*Q*	*I*	*Q*	*I*	*q*	*I*	*Q*	*I*	*q*	*I*	*I*	*I*	*I*	*I*	*I*
A	97.288	0.982	976.143	0.707	102.043	0.840	109.921	0.891	145.630			878.855		12.633	48.342	
B	136.883	0.991	1170.167	0.707	135.709	0.845	147.407	0.895	236.317			1033.284		10.524	99.434	
C	115.909	0.987	759.443	0.707	200.604	0.845	128.772	0.900	163.261	0.932	184.984	643.534	84.695	12.863	47.352	69.075
D	132.751	0.991	911.080	0.707	209.716	0.849	138.121	0.895	196.309	0.987	204.799	778.329	76.965	5.370	63.558	72.049
E	128.968	0.991	1137.794	0.707	138.190	0.849	155.693	0.891	284.678			1008.826		26.725	155.709	
F	145.045	1.005	1102.730	0.707	144.083	0.845	154.400	0.895	345.547			957.685		9.355	200.502	
6 Hours		Peak Positions (*q*) and Intensities (*I*) (Circular Average)										Background-subtracted				
	Background	Perovskite (001)		CsPbI3 (002)		FAPbI3 (001)		PbI2 (001)		CsPbI3 (102)		Perovskite	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)
	I	*q*	*I*	*Q*	*I*	*q*	*I*	*Q*	*I*	*q*	*I*	*I*	*I*	*I*	*I*	*I*
A	152.982	0.996	1215.836	0.707	161.419	0.845	169.747	0.891	346.907			1062.854	8.437	16.765	193.925	
B	132.009	1.001	1006.776	0.707	145.289	0.845	149.510	0.895	484.918			874.767	13.280	17.501	352.909	
C	136.276	0.996	742.240	0.707	232.101	0.845	149.960	0.900	168.679	0.927	206.762	605.963	95.825	13.684	32.403	70.486
D	132.035	0.996	822.620	0.707	217.400	0.845	145.112	0.900	242.712	0.927	205.790	690.585	85.365	13.077	110.677	73.755
E	133.970	0.996	1006.303	0.707	141.311	0.845	150.138	0.895	320.102			872.333	7.341	16.168	186.132	
F	149.877	1.005	1070.753	0.707	154.771	0.845	169.093	0.891	587.316			920.876	4.894	19.216	437.439	
100 Hours		Peak Positions (*q*) and Intensities (*I*) (Circular Average)										Background-subtracted				
	Background	Perovskite (001)		CsPbI3 (002)		FAPbI3 (001)		PbI2 (001)		CsPbI3 (102)		Perovskite	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)
	I	*q*	*I*	*Q*	*I*	*q*	*I*	*Q*	*I*	*q*	*I*	*I*	*I*	*I*	*I*	*I*
A	200.577	1.010	303.289	0.707	211.940	0.845	230.260	0.905	239.050	0.923	231.936	102.712	11.364	29.684	38.473	31.359
B	197.898	1.010	284.630	0.703	210.813	0.845	229.940	0.881	239.511	0.941	232.187	86.732	12.916	32.042	41.614	34.289
C	194.306	1.010	227.055	0.707	242.386	0.849	243.216	0.900	220.161	0.927	239.739	32.749	48.080	48.910	25.855	45.433
E	261.700	1.014	387.500	0.707	344.843	0.849	316.700	0.900	322.400	0.923	366.800	125.800	83.143	55.000	60.700	105.100
E	179.400	1.014	276.388	0.707	192.784	0.845	206.514	0.881	217.042	0.927	208.668	96.988	13.384	27.114	37.642	29.268
F	229.370	0.977	274.638	0.707	239.175	0.845	262.505	0.895	273.765	0.927	266.559	45.268	9.806	33.136	44.396	37.189

Table S6 Peak intensity analysis

0 Hours	Signal/Noise Ratio (If > 1.05 threshold) *					Peak Intensity Ratios (if > 0.05 threshold) **			
	Perovskite (001)	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)
	Relative to background	Relative to background	Relative to background	Relative to background	Relative to background	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite
A	10.034		1.130	1.497				0.055	
B	8.549		1.077	1.726				0.096	
C	6.552	1.731	1.111	1.409	1.596	0.132		0.074	0.107
D	6.863	1.580		1.479	1.543	0.099		0.082	0.093
E	8.822	1.072	1.207	2.207				0.154	
F	7.603		1.064	2.382				0.209	
6 Hours	Signal/Noise Ratio (If > 0.05 threshold)					Peak Intensity Ratios (if > 0.05 threshold)			
	Perovskite (001)	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)
	Relative to background	Relative to background	Relative to background	Relative to background	Relative to background	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite
A	7.948	1.055	1.110	2.268				0.182	
B	7.627	1.101	1.133	3.673				0.403	
C	5.447	1.703	1.100	1.238	1.517	0.158		0.053	0.116
D	6.230	1.647	1.099	1.838	1.559	0.124		0.160	0.107
E	7.511	1.055	1.121	2.389				0.213	
F	7.144		1.128	3.919				0.475	
100 Hours	Signal/Noise Ratio (If > 0.05 threshold)					Peak Intensity Ratios (if > 0.05 threshold)			
	Perovskite (001)	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)	CsPbI3 (002)	FAPbI3 (001)	PbI2 (001)	CsPbI3 (102)
	Relative to background	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite	Relative to Perovskite
A	1.512	1.057	1.148	1.192	1.156	0.111	0.289	0.375	0.305
B	1.438	1.065	1.162	1.210	1.173	0.149	0.369	0.480	0.395
C	1.169	1.247	1.252	1.133	1.234	1.468	1.493	0.789	1.387
E	1.481	1.318	1.210	1.232	1.402	0.661	0.437	0.483	0.835
E	1.541	1.075	1.151	1.210	1.163	0.138	0.280	0.388	0.302
F	1.197		1.144	1.194	1.162		0.732	0.981	0.822

** Account for systematic error in background subtraction due to amorphous substrates and degradation products (5%)*

*** Account for human peak extraction error (5*

6. Additional Information

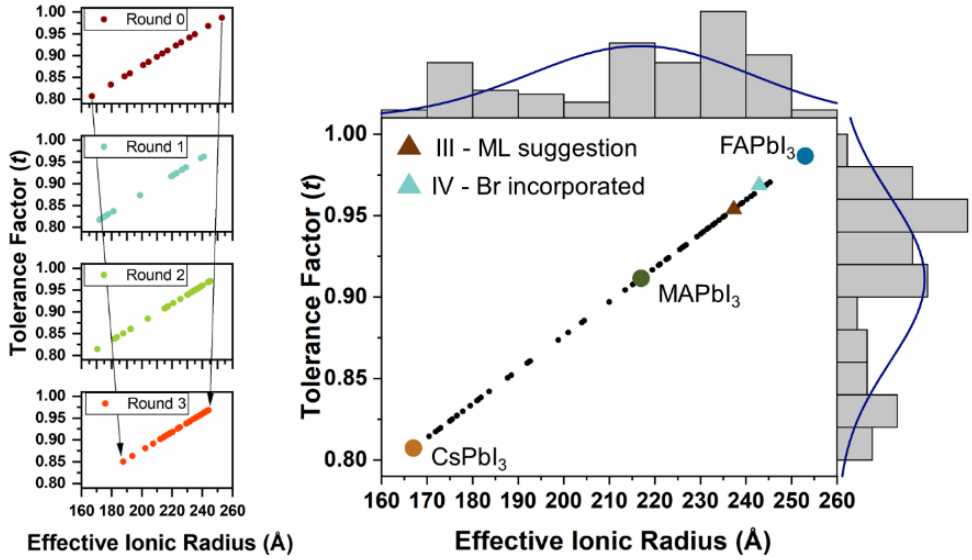


Figure S18 Relationships between effective ionic radius and calculated Goldschmidt's tolerance factor for compositions sampled in Round 0 -3.

7. Data Integrity

The synthesis conditions were calculated using Google Sheets, recorded using a laboratory notebook by the experimenter, then transcribed to a summary Google Sheets. Accelerated aging test data (camera image time series) were automatically pushed to Dropbox, and the Bayesian optimization codes were written using Python and gave the subsequent round 's composition suggestions saved to a spreadsheet, which were then given to the experimenter, to be fabricated. Raw XRD, UV-Vis, SEM and GIWAXS data were processed using their own software packages with individual file format, labelling conventions, and stored on different local computers. Most of the data analysis was done using Python and MATLAB. Metadata, which linked different files containing synthesis conditions, calibrated aging test data, GIWAXS, XRD, and UV-Vis data, were created on an ad-hoc basis on the 7 compositions, as samples were deemed of high scientific significance.

8. References

1. Correa-Baena, J.-P. *et al.* Homogenized Halides and Alkali Cation Segregation in Alloyed Organic-inorganic Perovskites. *Science* **363**, 627–631 (2019).
2. Menesatti, P. *et al.* RGB color calibration for quantitative image analysis: The '3D Thin-Plate Spline' warping approach. *Sensors* **12**, 7063–7079 (2012).

3. González, J., Dai, Z., Hennig, P. & Lawrence, N. *Batch Bayesian Optimization via Local Penalization*. http://sheffieldml.github.io/GPyOpt/.

4. Gelbart, M. A., Snoek, J. & Adams, R. P. Bayesian optimization with unknown constraints. in *Uncertainty in Artificial Intelligence - Proceedings of the 30th Conference, UAI 2014* 250–259 (2014).

5. SheffieldML/GPyOpt: Gaussian Process Optimization using GPy. https://github.com/SheffieldML/GPyOpt.

6. Stoumpos, C. C., Malliakas, C. D. & Kanatzidis, M. G. Semiconducting Tin and Lead Iodide Perovskites with Organic Cations: Phase Transitions, High Mobilities, and Near-infrared Photoluminescent Properties. *Inorg. Chem.* **52**, 9019–9038 (2013).

7. Zunger, A., Wei, S. H., Ferreira, L. G. & Bernard, J. E. Special quasirandom structures. *Phys. Rev. Lett.* **65**, 353–356 (1990).

8. Van de Walle, A., Asta, M. & Ceder, G. The alloy theoretic automated toolkit: A user guide. *Calphad Comput. Coupling Phase Diagrams Thermochem.* **26**, 539–553 (2002).

9. Van De Walle, A. *et al.* Efficient stochastic generation of special quasirandom structures. *Calphad Comput. Coupling Phase Diagrams Thermochem.* **42**, 13–18 (2013).

10. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).

11. 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).

12. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* vol. 77 3865–3868 (1996).

13. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zon integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).

14. Lejaeghere, K. *et al.* Reproducibility in density functional theory calculations of solids. *Science,* **351**, (2016).

15. Schelhas, L. T. *et al.* Insights into operational stability and processing of halide perovskite active layers. *Energy Environ. Sci.* **12**, 1341–1348 (2019).

16. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning*. (MIT press Cambridge, 2016).

17. Sollich, P. Can Gaussian process regression be made robust against model mismatch? in *International Workshop on Deterministic and Statistical Methods in Machine Learning* vol. 3635 LNAI 199–210 (Springer, Berlin, Heidelberg, 2004).

18. https://github.com/CFN-softbio/SciAnalysis.

19. https://scistreams.readthedocs.io/en/latest/index.html.