Supplementary Information

Combining Component Screening, Machine Learning and Molecular Engineering for the Design of High-Performance Inverted Perovskite Solar Cells

Boxue Zhang,[1] Huaibiao Zeng,[2] Haomiao Yin,[2] Daming Zheng,[1] Zhongquan Wan,[2] Chunyang Jia,[2] Thijs Stuyver,[3]* Junsheng Luo,[2]* Thierry Pauporté[1]*

[1] *Institut de Recherche de Chimie Paris (IRCP) UMR8247, Chimie ParisTech, PSL University CNRS, 11 rue P. et M. Curie, Paris F-75005, France*

[2] *National Key Laboratory of Electronic Films and Integrated Devices, School of Integrated Circuit Science and Engineering, University of Electronic Science and Technology of China, 611731 Chengdu, P. R. China*

[3] *Chimie ParisTech, Université PSL, CNRS, Institute of Chemistry for Life and Health Sciences, 75005 Paris, France*

*E-mails: thierry.pauporte@chimieparistech.psl.eu; luojs@uestc.edu.cn; thijs.stuyver@chimieparistech.psl.eu

Table of Contents

Methods

Materials

Formamidinium Iodide (FAI), Methylammonium Iodide (MAI), Formamidinium Bromide (FABr), Methylammonium Bromide (MABr), and Methylammonium Chloride (MACl) were purchased from Greatcell Solar (Australia). Lead Iodide (PbI_2), Lead Bromide ($PbBr_2$), and Lead Chloride ($PbCl_2$), the [2-(3,6-Dimethoxy-9H-carbazol-9-yl)ethyl]phosphonic Acid (MeO-2PACz) were sourced from Tokyo Chemical Industry (TCI, Japan). For the electron transport layer, Bathocuproine (BCP) and [6,6]-phenyl-C61-butyric acid methyl ester ($PC_{61}BM$) were acquired from Xi'an Polymer Light Technology Corporation (China). Additional materials such as Cesium Iodide (CsI), Rubidium Iodide (RbI), and Rubidium Chloride (RbCl), nickel nitrate hexahydrate (($Ni(NO_3)_2·6H_2O$), sodium hydroxide (NaOH), dimethylformamide (DMF), dimethyl sulfoxide (DMSO), isopropanol (IPA), 1-Methyl-2-pyrrolidinone (NMP) and chlorobenzene (CB) were purchased from Sigma-Aldrich.

Multi-component film and devices preparation

In our comprehensive exploration of perovskite materials, we meticulously prepared a series of multi-component thin films, utilizing a 1.2M solution of seven distinct perovskite precursors: $FAPbI_3$, $CsPbI_3$, $FAPbBr_3$, $MAPbI_3$, $MAPbCl_3$, $RbPbI_3$, CsI, RbI, MAI, and MACl. First, 1.2381 g of FAI and 3.3192 g of PbI_2 were dissolved in 6 mL of a DMF and DMSO with a 4:1 volume ratio. Subsequently, separate solutions were created: 0.0749 g of FABr and 0.2202 g of $PbBr_2$, 0.0953 g of MAI and 0.2766 g of PbI_2, 0.0671 g of MABr and 0.2202 g of $PbBr_2$, and 0.0402 g of MACl and 0.1688 g of $PbCl_2$, each dissolved in 0.5 mL of the same DMF:DMSO solvent. 0.1247 g of CsI and 0.2212 g of PbI_2, 0.1019 g of RbI and 0.2212 g of PbI_2, each dissolved in 0.4 mL of the DMF:DMSO solvent. Separately, 0.1247 g of CsI, 0.1274 g of RbI, 0.0953 g of MAI, and 0.0402 g of MACl were each dissolved in 0.5 mL of DMSO. These solutions were then combined in various ratios to form a diverse range of perovskite compositions, leading to the creation of 60 different perovskite variants for the study **(Supplementary Table 1)**. The conducting FTO substrates were subjected to a UV ozone treatment for 15 minutes, a crucial step to modify the surface wettability and promote better

adhesion of the perovskite material. Following substrate preparation, 40 microliters of the mixed perovskite solution, corresponding to 60 different component variations as outlined in Table 1, were carefully drop-casted onto the glass substrate. These films were then spin-coated at 500 rpm for 6 seconds, followed by 3000 rpm for 20 seconds, ensuring a uniform distribution of the material onto the substrate. To eliminate any potential influences from these solvents and preparation methods, we solely relied on spin-coating followed by annealing to fabricate the perovskite films. They were then annealed at 100°C for 1 hour, to obtain excellent optoelectronic properties of perovskite films, and these films were used for optical attenuation testing.

To fabricate the multi-component perovskite solar cells, we initiated the process with a meticulous preparation of the ITO (indium tin oxide) coated glass substrates. The substrates were laser etched to define the device area, followed by an extensive ultrasonic cleaning regimen using glass cleaner, deionized water, acetone, ethanol, and isopropyl alcohol, each for a duration of 15 minutes. This cleaning process was crucial to remove any residues and ensure a pristine substrate surface. The ITO-coated substrates were then subjected to ultraviolet ozone (UVO) treatment for 30 minutes, a step implemented to enhance the wettability of the substrate and promote better adhesion of subsequent layers. Next, we prepared a 30 mg ml^{-1} aqueous ink of NiO$_x$ nanoparticles (NPs) dispersed in deionized water. This NiOx NP ink was spin-coated onto the UVO-treated ITO glass at 4,000 rpm for 30 seconds, forming a uniform NiO$_x$ layer. The deposited NiO$_x$ film was subsequently annealed at 150 °C for 15 minutes in the atmosphere to improve its crystallinity and electrical properties. Following the NiOx layer preparation, 50 microliters of each of the 60 different component perovskite solutions were applied onto the NiO$_x$/ITO substrate. This step was performed meticulously to ensure uniform coverage and prepare the devices for electrical testing. The perovskite films were spin-coated at 4000 rpm for 6 s, followed by a vacuum flash-assisted crystallization step and a final annealing at 150 °C for 20 minutes. To complete the device architecture, a 20 mg mL^{-1} solution of PC$_{61}$BM in CB was spin-coated onto the perovskite film at 1500 rpm for 30 seconds. This was followed by the deposition of a BCP layer, prepared by dissolving 2 mg of BCP in 1 mL of IPA, and filtering through a PTFE filter prior to use. The BCP layer was spin-coated onto the PC$_{61}$BM layer at 4000 rpm for 30 seconds. The device fabrication was finalized with the thermal evaporation of

100 nm thick Ag electrodes, establishing the electrical contacts necessary for device operation and testing. All the above steps were performed in a N_2 filled glove box.

Solar cells fabrication

In the fabrication of our perovskite solar cells, we initiated the process by the deposition of a NiO_x nanocrystal layer. A 30 mg.mL^{-1} NiOx solution was spin-coated onto the ITO substrate at 4,000 rpm for 30 seconds in ambient air, without any subsequent post-processing. The relative humidity (RH) was maintained at a constant level of 40% to ensure controlled environmental conditions. The substrate was then immediately transferred to a glove box to maintain an inert atmosphere. The NiO_x nanoparticles were synthesized via a hydrolysis reaction of nickel nitrate, as detailed in a previous work.[1] Following the NiO_x layer, a 0.3 mg mL^{-1} solution of Meo-2PACz in ethanol was spin-coated onto the NiO_x film at 3,000 rpm for 30 seconds. The substrate was then annealed at 100°C for 10 minutes to ensure proper film formation and adherence. For the perovskite layer, we selected the No. 57 perovskite precursor (**Supplementary Table S1**), identified as having the best optical and electrical stability from our composition screening. Different concentrations of AETA-BCF (0 mg/ml, 1 mg/ml, 1.5 mg/ml and 2 mg/ml) were dissolved in the perovskite precursor. An 80 μL drop of the 1.5 M perovskite precursor solution was applied to the NiOx-Meo-2PACz substrate, followed by spin-coating at 4,000 rpm for 6 seconds. The film then underwent a vacuum-assisted crystallization process, immediately followed by annealing on a hot plate at 150°C for 20 minutes. Next, a 20 mg ml^{-1} solution of PC61BM in chlorobenzene (CB) was spin-coated onto the treated perovskite film at 1,500 rpm for 30 seconds. A supersaturated solution of BCP was prepared by dissolving 2 mg of BCP in 1 mL of IPA, followed by 5 minutes of agitation and filtration through an organic filter. This BCP solution was then spin-coated onto the $PC_{61}BM$ layer at 4,000 rpm for 30 seconds. The device fabrication was completed with the thermal evaporation of 100 nm thick Ag electrodes, establishing the necessary electrical contacts for device operation. For stability testing, devices were fabricated using Cu as the electrode material. A 100 nm thick layer of copper was thermally evaporated to complete the device manufacturing.

Device test

Current-voltage (*J-V*) characteristic measurements under controlled conditions. The

measurements were carried out using a Keithley 2400 sourcemeter, with the solar cells illuminated by a standard solar simulator. The light intensity was meticulously calibrated to 100 mW cm^{-2}, ensuring consistency and comparability of the results (calibrated using a standard silicon solar cell) in an atmospheric environment, with the humidity maintained between 30-50%. For the forward scan, the voltage was varied from -0.2V to 1.2V at a scanning rate of 0.01 mV s^{-1}. Similarly, for the reverse scan, the voltage was varied from 1.2V to -0.2V at the same scanning rate of 0.01 mV s^{-1}. The effective area of the solar cells was determined using an aperture mask, ensuring precision in the measurement of the photocurrent. The devices were measured with a 0.1 cm^2 aperture mask.

Measurement and Characterization

Transient photocurrent spectroscopy (TPC) was conducted using ENLITECH's PD-RS system. The perovskite absorbance of the films was measured with a Cary 5000 UV-Vis-NIR spectrophotometer, using an FTO sample as the baseline. Fourier-transform infrared (FTIR) spectra were obtained in reflectance mode using the FTIR-850. The Agilent 5500 scanning probe microscope was employed for Kelvin Probe Force Microscopy (KPFM) in tapping mode. Time-resolved photoluminescence (TRPL) was measured with a Horiba Fluorolog-3 time-correlated single photon counting system. Scanning electron microscopy (SEM) determinations were conducted using a JEOL JSM7610F SEM. Transmission electron microscopy (TEM) images were acquired with a JEM-2100F (JEOL, Japan).

Computational Details

All first-principles calculations, were performed with the Vienna Ab initio Simulation Package (VASP),[2] making use of the Projector Augmented-Wave (PAW) method in combination with the PBE exchange-correlation functional.[3,4] Our simulations utilized a $2 \times 2 \times 1$ gamma-centered Monkhorst-Pack electronic k-point grid with a plane-wave cutoff energy of 400 eV. The calculations were conducted with a 20Å vacuum space to preclude any aperiodic interactions between the $2 \times 2 \times 1$ supercells. Computations were performed allowing full relaxation of all structures until the total energy of the atoms reached a convergence threshold of 1.0×10^{-5} eV/atom. Starting FAPbI$_3$ crystal cell structure were constructed using Materials

studio.[5] The files associated with the DFT calculations performed in VASP have been deposited on FigShare (10.6084/m9.figshare.24782304).

Machine learning Methodology

To optimize the fabrication process of vacuum-assisted perovskite solar cells, we employed Gaussian regression Bayesian optimization, a robust machine learning technique renowned for its efficacy in high-dimensional optimization problems.[6,7] As input variables and target output, we meticulously selected seven critical process variables based on their significant impact on the perovskite film formation and the overall device performance. These variables included: NMP volume (ML), DMF volume (ML), DMSO volume (ML), perovskite precursor concentration (M, the concentration of the raw materials used in the synthesis of perovskites), annealing temperature (°C), vacuum pressure (Pa), and vacuum pressure holding time (s) (**Supplementary Table 2**). The target output for our optimization model was the Power Conversion Efficiency (PCE) of the solar cell, a direct indicator of device performance. The optimization ranges for each of the input variables were carefully chosen based on a comprehensive review of previous work and expert consideration – and validated through a preliminary sampling (cf. main text) – ensuring a balance between experimental feasibility and the breadth of the parameter space. These ranges are detailed in **Supplementary Table 1.** Given the selected sampling interval, a full grid sampling approach would result in approximately 1,342,863,522 unique process conditions, highlighting the necessity of an efficient optimization strategy.

In BO for process optimization, one attempts to find the specific process condition, x_i, for which the black-box objective function, $f(x)$, mapping the condition search space (A), to the target property (PCE), is maximal,[8]

$$x_i = \operatorname*{argmax}_{x \in A} f(x) \tag{1}$$

Such a BO campaign typically starts by selecting an initial sample of process conditions, after which the output property of interest is determined for each of those. Subsequently, a

surrogate model for the true objective function $\hat{f}(x)$ – the prior – is trained based on the observations made, and an acquisition function is constructed, which calculates the utility of performing an additional measurement, i.e., evaluating the objective function, at the point x. Utility may be measured in a number of ways: the predicted PCE value, the amount of information this new point will provide the surrogate model, the likelihood this point will improve upon the current maximum, etc.[9] Once new condition combinations have been acquired, the surrogate model is updated – resulting in the posterior – and acquisition can be repeated. This updating process can be performed in an iterative manner until significant improvements in PCE values are no longer observed from batch to batch.

For the construction of the main surrogate model, Gaussian Process regression with a matern52 kernel was selected.[10] Matern52 is a flexible kernel, enabling it to treat potential discontinuities in the data better. Additionally, the kernel was selected in its anisotropic form, so that each of the kernel parameters could be tuned independently for every input variable, facilitating the assignment of a stronger impact/relevance on the predicted output variable, i.e., the measured efficiency, by some variables compared to others. It should be noted that the surrogate model, evaluated at point x, yields both a mean prediction for that point, $\hat{f}(x)$, as well as a standard deviation $\sigma(x)$, which quantifies the uncertainty of the model about its prediction at this point.

To inform the selection of process condition combinations to acquire in subsequent BO batches, a compounded acquisition function was selected. The utility function of upper confidence bound (UCB) was selected as the basis,[11]

$$acq_{UCB}(x;\beta) = \hat{f}(x) + \beta * \sigma(x) \tag{2}$$

where β is the parameter that adjusts the relative weight of prediction uncertainty $\sigma(x)$ over the prediction mean value $\hat{f}(x)$. In general terms of balancing exploration and exploitation, higher β values lead to more exploration, while lower β values lead to more exploitation. In this work, we set β equal to 1, in line with previous work.[12]

While it was our goal to focus primarily on the (accurate) efficiencies measured as part of the BO campaign, we also did not want to lose the information from the preliminary sample and the visual defects data altogether during acquisition. As such, two probabilistic constraints were defined, in line with previous work by Buonassisi and co-workers.[13] The first of those probabilistic constraints, $Pr[C_1(x)]$, corresponds to the probability that the thin film will be produced without visual defects. This film constraint is defined based on a latent constraint function $g_{constr,1}(x)$,[14]

$$Pr[C_1(x)] = Pr[g_{constr,1}(x) \geq 0)] \qquad (3)$$

where $g_{constr,1}(x)$ is a secondary surrogate model in its own right, which is constructed with the help of Gaussian Processes as well, based on the process conditions evaluated so far. In the training data of this Gaussian process model, $\{(x_i; y_i)\}$, y_i is assigned the value +0.5 if the fabricated device did not have visually detectable defects, and -0.5 if the device did contain such defects. As a result of this convention, Eq. 3 evaluates to 0 for all training points which yielded devices with visual defects, and to 1 for all training points resulting in devices without defects. For all other points in the search space, a probability – to have a device without defects – is interpolated based on the constructed surrogate model.

The second probabilistic constraint, $Pr[C_2(x)]$, has been defined with the aim of penalizing regions of the search space around points, evaluated during preliminary sampling, which yielded subpar PCE values. This constraint is defined based on a latent constraint function $g_{constr,2}(x)$,[14]

$$Pr[C_2(x)] = Pr[g_{constr,2}(x) \geq 0)] \qquad (4)$$

where $g_{constr,2}(x)$ is another secondary surrogate model, which has again been constructed with the help of Gaussian Processes. The training points $\{(x_i; y_i)\}$ for $g_{constr,2}(x)$ are derived

from the process conditions and corresponding PCE values for all the points in the preliminary sample, but then with the PCE values rescaled as follows,

$$\{(x_i; y_i)\} \rightarrow \{(x_i; y_i - mean(\{y_i\}))\} \tag{5}$$

where $mean(\{y_i\})$ is the mean of all the PCE values encountered in the preliminary sample. Based on these definitions, it can indeed be inferred that Eq. 4 will evaluate to 0 for all points in the preliminary sample that yielded a PCE below average, whereas it will evaluate to 1 for points that yielded a PCE above average during preliminary sampling. For all other points in the search space, a probability to yield an above average PCE is interpolated based on the constructed surrogate model.

The two constraints were then combined with acq_{UCB} to yield the final acquisition function as follows,

$$acq_{full}(x) = acq_{UCB}(x) * (0.5 + 0.5 * Pr[C_1(x)]) * (0.8 + 0.2 * Pr[C_2(x)]) \tag{6}$$

Note that the individual constraints have been softened, i.e., rescaled, in Eq. 6, to make the acquisition process more conservative with respect to the film and preliminary screening data. The weighting factors introduced to this end were chosen to make the current data 5 times more important than the preliminary screening data, and the qualitative film quality information is considered half as important as the information of the actual efficiency measurements. By introducing these constraints in the acquisition function, one can expect to sample fewer points in the region where film quality is poor, or where low device efficiency was obtained in the preliminary experiments, throughout the BO campaign.

As a final note, it should be mentioned that, since the purpose of the probabilistic constraints is mainly to guide the model away from regions in the search space with a low probability of success, radial basis function kernels were used to model surrogates $g_{constr,1}(x)$ and $g_{constr,2}(x)$. Radial basis functions are less flexible than matern52 kernels, and

consequently, their smoother/more rigid behavior intuitively feels more appropriate to crudely define unfavorable regions of the search space.

To construct 2D contour plots of the surrogate objective function during a given iteration of the Bayesian optimization procedure, we started by sampling 1000 random points across the 7-dimensional search space. Subsequently, we modified the values for the two selected dimensions (x1 and x2) iteratively, so that the complete grid of potential (x1, x2)-values is covered. This results in 1000 simulated samples for every combination of (x1, x2): for the 5 non-selected dimensions, the values are fully randomized, so that they constitute a representative sample of the full population of possible process conditions. Finally, the 1000 samples in every (x1, x2) point is passed through the surrogate function, obtaining the corresponding predictions. Finally, we can construct our contour plots, where we either visualize in every grid point the maximum, mean or minimum of the respective 1000 predictions.

The surrogate models and acquisition functions were implemented, based on previous work by Buonassisi and co-workers,[13] with the help of Emukit[15] and GPy.[16] All the code associated with this project has been made available through a GitHub repository https://github.com/Boxue2023/Perovskite.

Synthesis of Acryloyloxyethyltrimethylammonium tetrakis(pentafluorophenyl)borate

(S1)
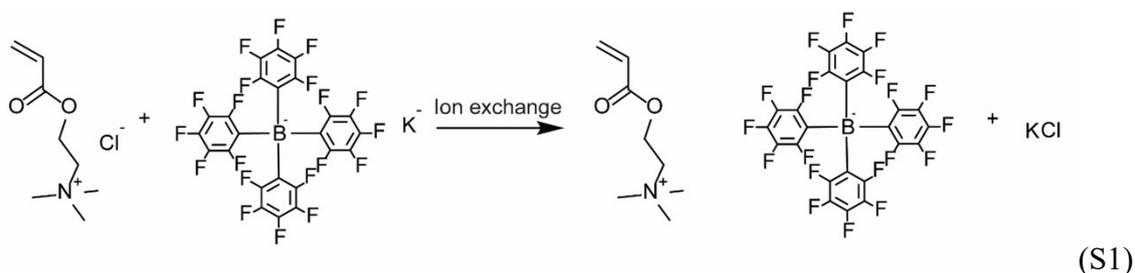
In a 100 mL Schlenk flask, equipped with a magnetic stir bar, acryloyloxyethyltrimethyl ammonium chloride (AETAC) (1.00 g, 4.58 mmol) was dissolved in 30 mL of anhydrous tetrahydrofuran (THF). The solution was cooled to -78°C in a dry ice/acetone bath under an inert nitrogen atmosphere. Meanwhile, potassium tetrakis(pentafluorophenyl)borate (2.15 g,

4.58 mmol) was separately prepared in a glovebox to prevent moisture and oxygen contamination, and then suspended in 20 mL of anhydrous THF.

The potassium tetrakis(pentafluorophenyl)borate suspension was added dropwise to the AETAC solution over a period of 30 minutes, ensuring the reaction temperature did not rise above -60°C. After the addition, the reaction was allowed to slowly warm to 0°C over a period of 2 hours, followed by stirring at room temperature for an additional 16 hours, allowing for complete reaction (S1) as evidenced by thin-layer chromatography (TLC) analysis using a mixture of hexanes and ethyl acetate (1:1) as the mobile phase.

The reaction was then quenched by the slow addition of saturated aqueous ammonium chloride solution. The mixture was extracted three times with dichloromethane (DCM). The combined organic extracts were washed with brine, dried over anhydrous sodium sulfate, and concentrated under reduced pressure. Purification of the residue by flash column chromatography (silica gel, using a gradient of 0-10% ethyl acetate in hexanes as eluent) afforded the title compound as a white solid.

The structure of the synthesized compound was confirmed by ^{1}H NMR and ^{19}F NMR spectroscopy. The ^{1}H NMR (400 MHz, DMSO-d6) spectrum exhibited characteristic peaks at δ 6.38 (d, J = 17.3 Hz, 1H), 6.29 – 6.13 (m, 1H), 6.02 (d, J = 7.3 Hz, 1H), 4.55 (s, 2H), 3.69 (s, 2H), and 3.13 (s, 9H). The ^{19}F NMR (376 MHz, DMSO-d6) spectrum showed peaks at δ -132.35 (s, 8F), -161.36 (s, 4F), and -165.96 (s, 8F), confirming the presence of the pentafluorophenyl and tetrakis(pentafluorophenyl)borate moieties.

Supplementary Fig. 1 Device and film photographs of the 60 components prepared through the vacuum-assisted crystal.
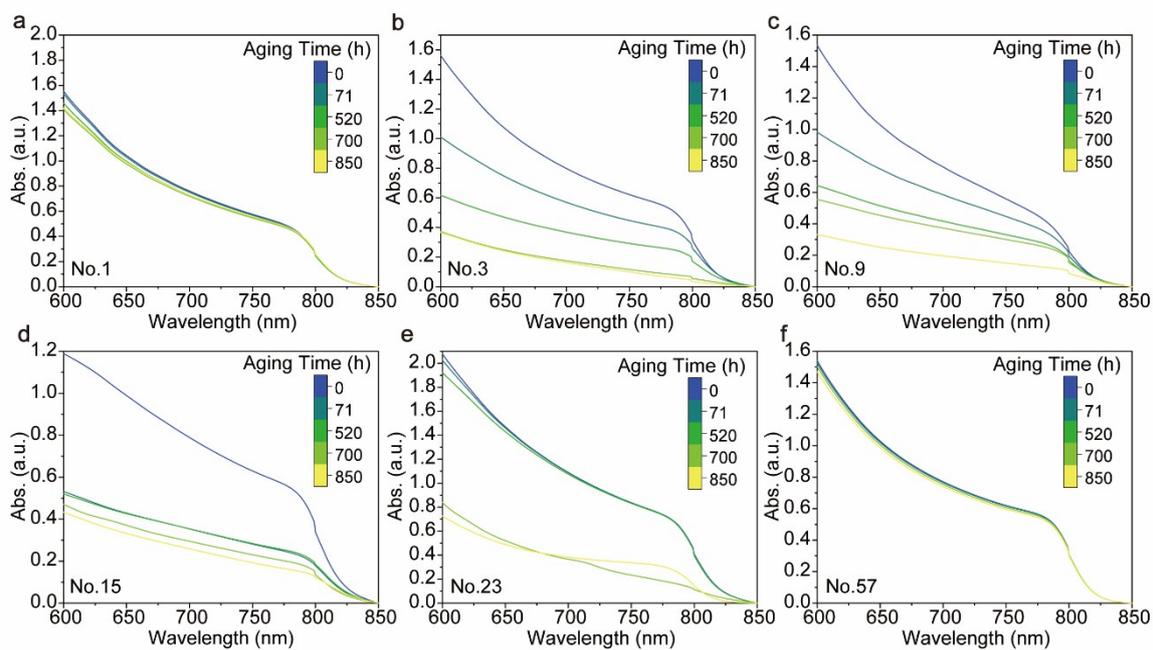
Supplementary Fig. 2 Visible/NIR Absorbance spectrum evolution of some component films during photothermal in N_2 atmosphere aging (hours).
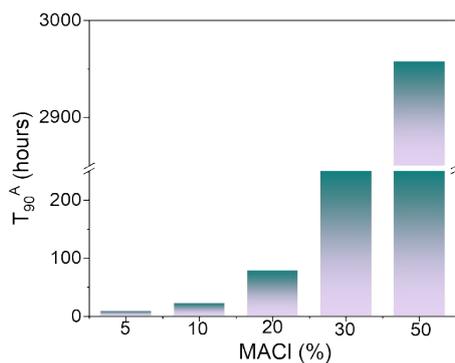
Supplementary Fig. 3 Diagram between the amount of MACl additive and the photothermal aging lifetime T_{90}^A.
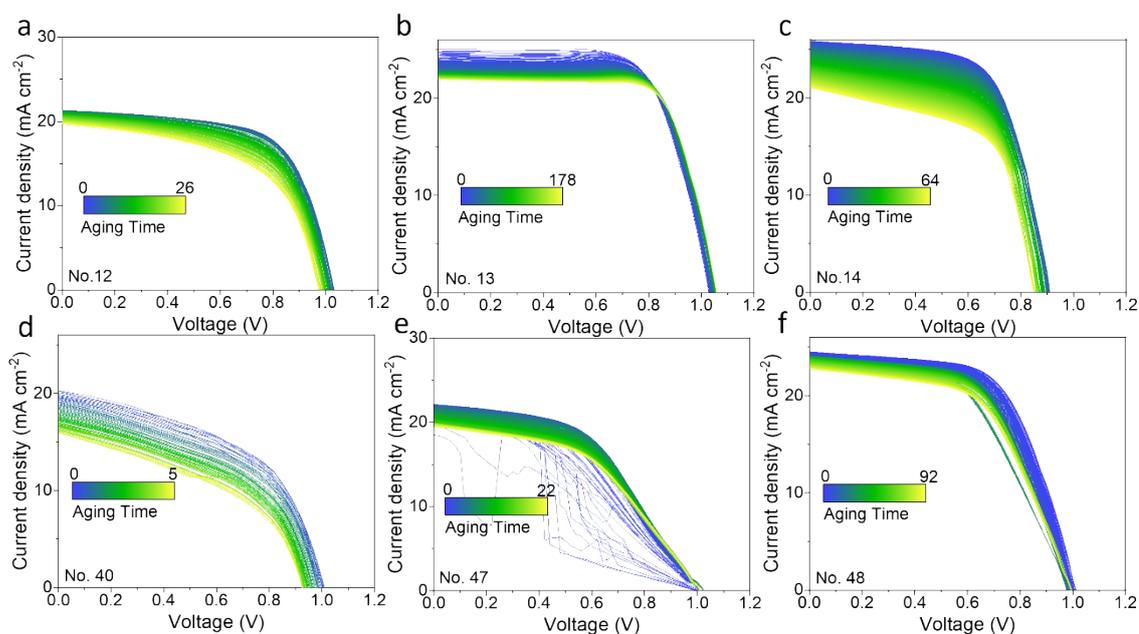
Supplementary Fig. 4 In-situ *J-V* performance (Reverse scan) evolution of some components upon aging time (hours).
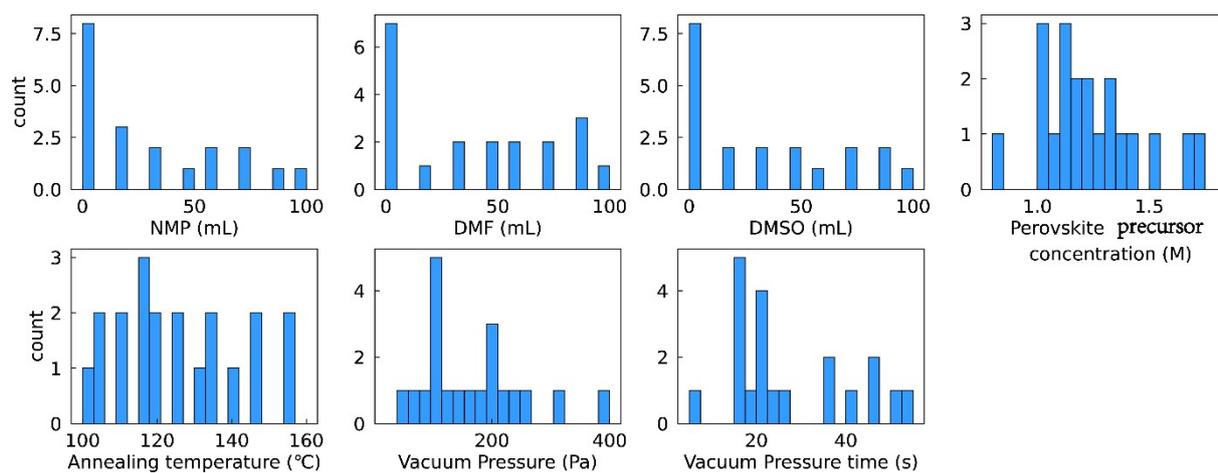
Supplementary Fig. 5 Bar plots visualizing the distribution of the initial samples across the search space for each of the process variables.
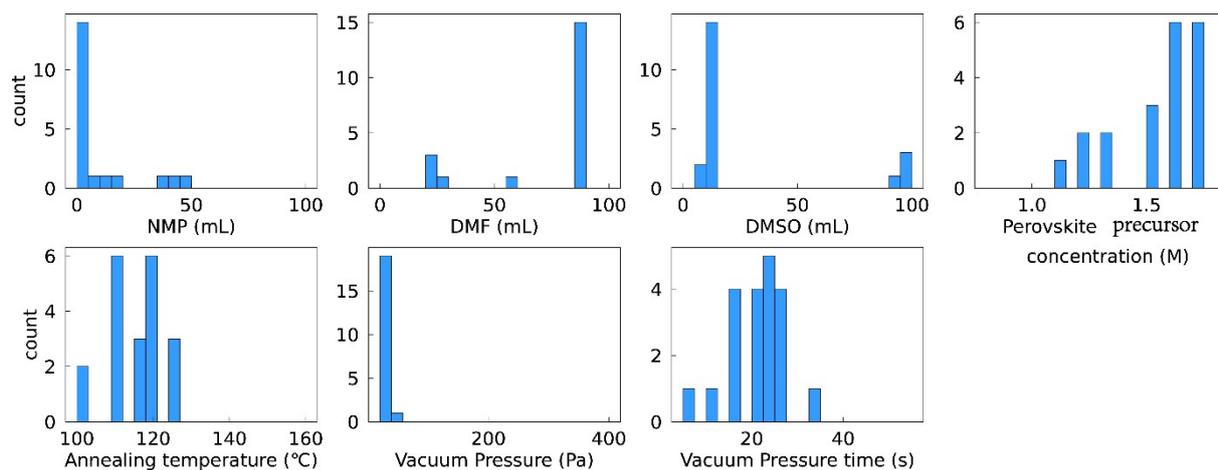
Supplementary Fig. 6 Bar plots visualizing the distribution of the first batch of experiments, selected by the BO algorithm (through construction of the compounded acquisition function), across the search space for each of the process variables.
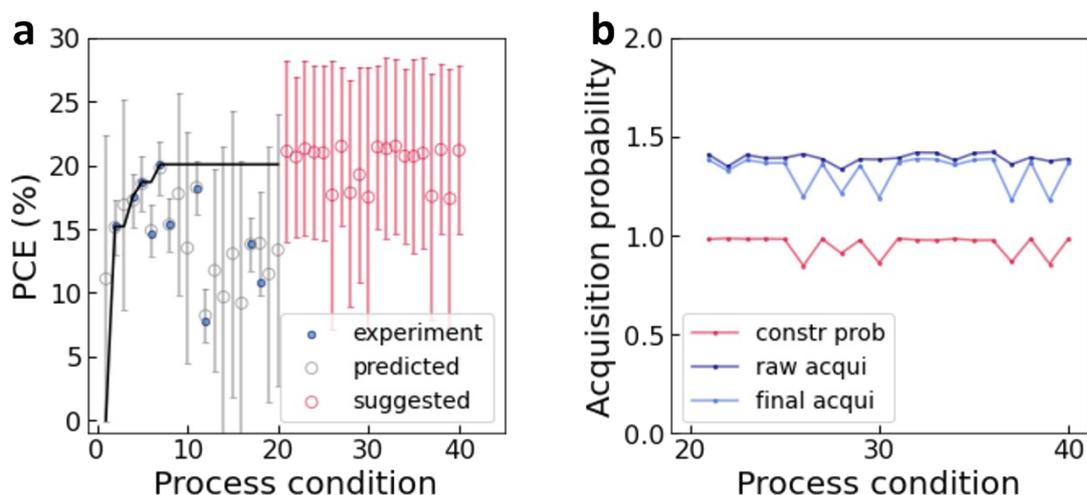
Supplementary Fig. 7 The acquisition process for the first sampling round. Plot of the predicted – and measured – PCE values for both the initial batch and the first selection of acquired points; (a) the best PCE value measured so far is denoted by the bold black curve (left). (b) Graphical summary of the acquisition function values across this first selection round, where 'raw acqui' corresponds to the upper confidence bound acquisition function, 'constr prob', corresponds to the product of the two constraint probabilities and 'final acqui' corresponds to the product between the raw acquisition function and the constraint probabilities. Note that the acquisition function values have been scaled by a factor 20 to enable a visualization on the same scale as the constraint probabilities.
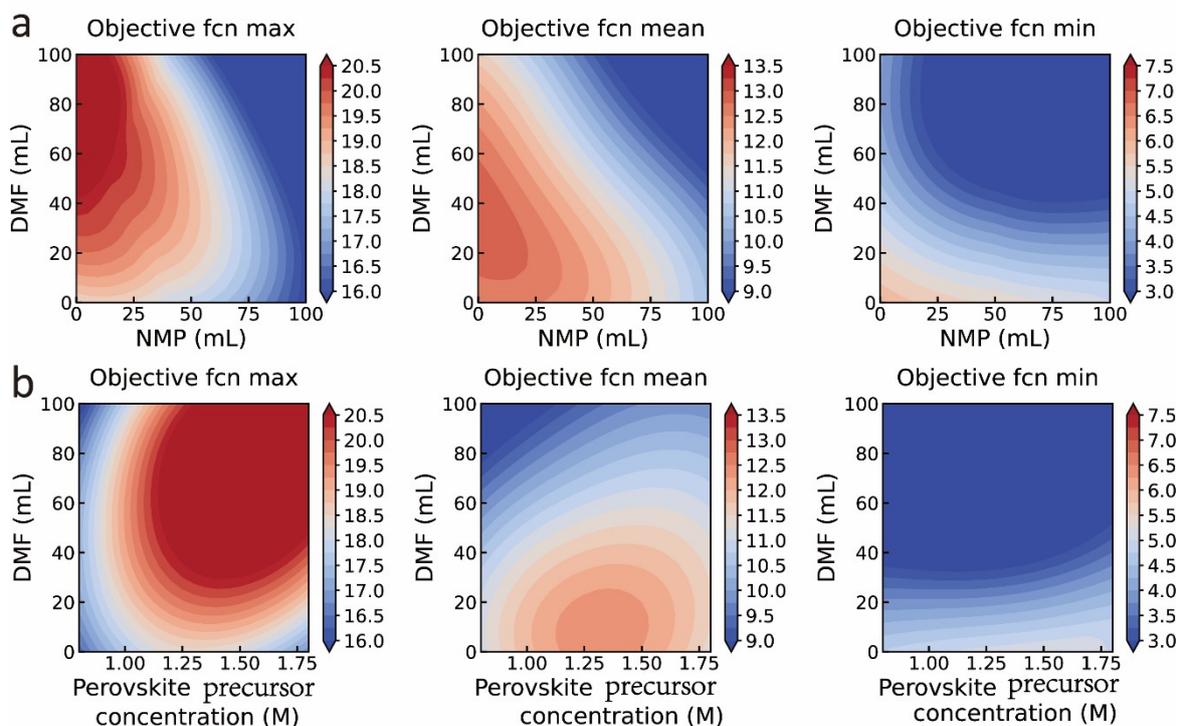
Supplementary Fig. 8 Sample of the 2D contour plots for the surrogate objective function: (a) DMF vs. NMP volumes, (b) Perovskite precursor concentration vs. DMF volume. These plots were

constructed by randomly sampling 1000 points in the search space, after which the values for the two selected parameters are iteratively modified. In each iteration, a distribution of process conditions corresponding to a point in the 2D space spanned by the two selected parameters is sampled in this manner. Evaluating the objective function for these distributions in every point, contour plots in 2D space can be generated, visualizing respectively the maximum (left), mean (center) and minimum (left) of the process condition distributions.
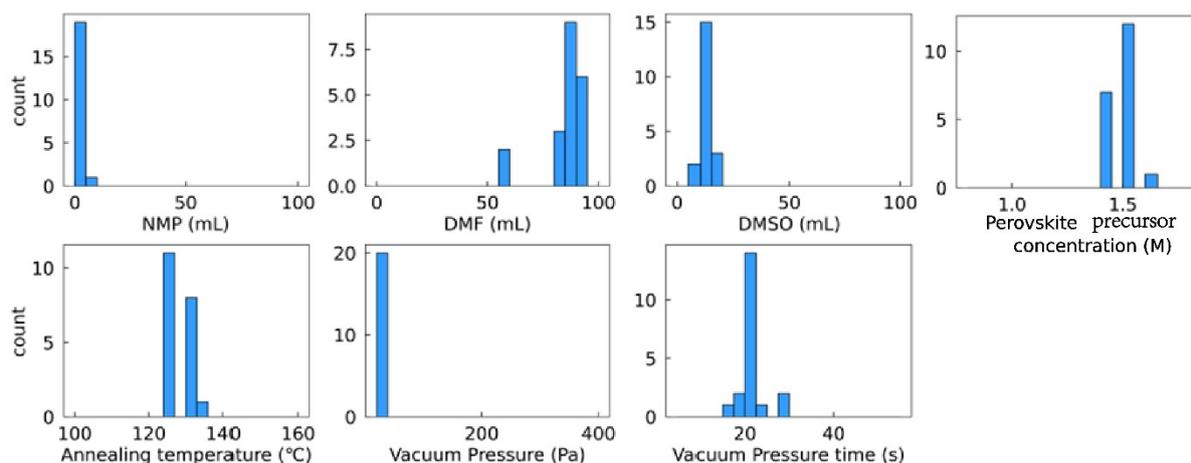
Supplementary Fig. 9 Bar plots visualizing the distribution of the second batch of experiments, selected by the BO algorithm, across the search space for each of the process variables.
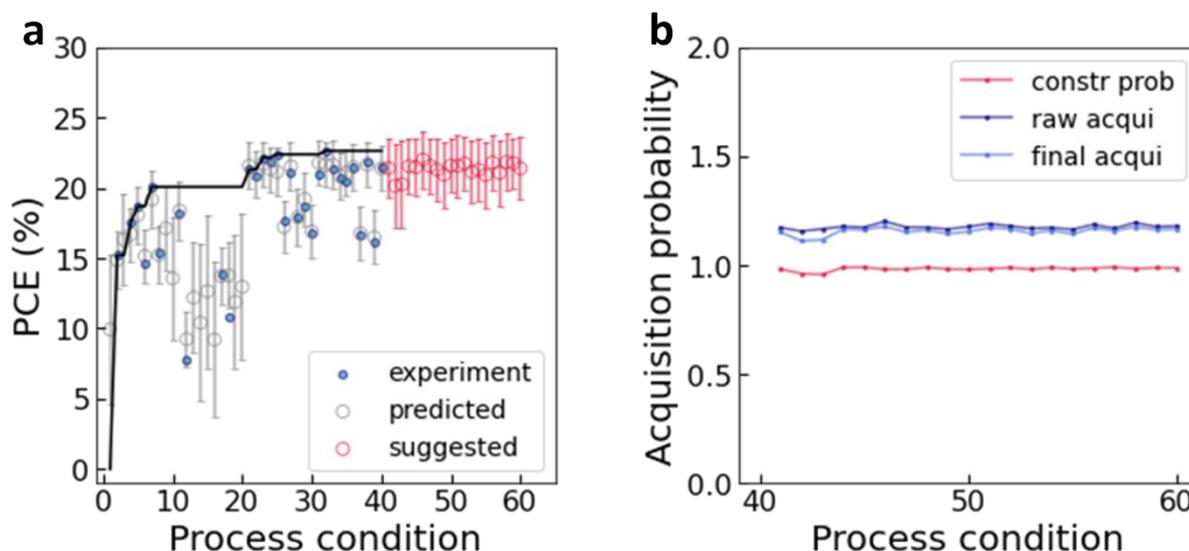
Supplementary Fig. 10 The acquisition process for the second sampling round. Plot of the predicted – and measured – PCE values for both the initial batch and the first selection of acquired points; (a) the best PCE value measured so far is denoted by the bold black curve (left). (b) Graphical summary of the acquisition function values across this second selection round, where 'raw acqui' corresponds to the

upper confidence bound acquisition function, 'constr prob', corresponds to the product of the two constraint probabilities and 'final acqui' corresponds to the product between the raw acquisition function and the constraint probabilities (right). Note that the acquisition function values have been scaled by a factor 20 to enable a visualization on the same scale as the constraint probabilities.
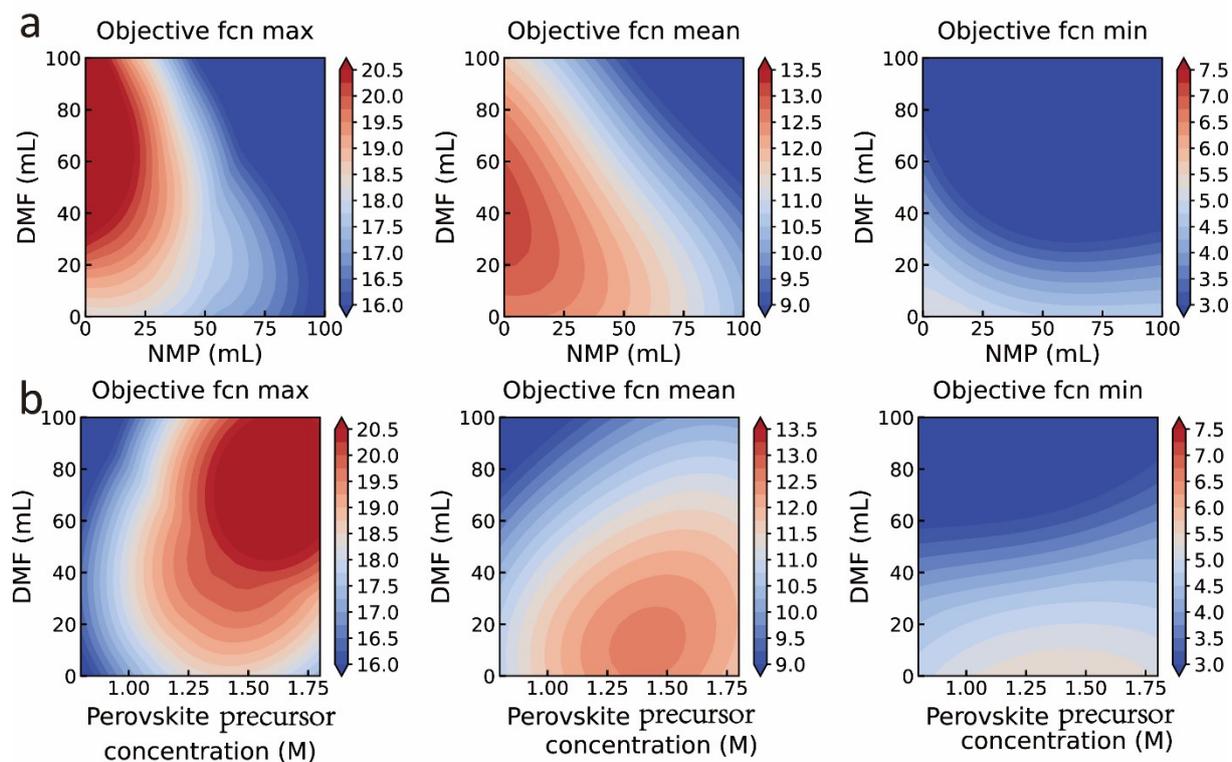
Supplementary Fig. 11 Sample of the 2D contour plots for the surrogate objective function: (a) DMF vs. NMP volumes, (b) Perovskite precursor concentration vs. DMF volume. Maximum (left), mean (middle) and minimum (right) surrogate objective function values across the sample of process conditions at every grid point are visualized.
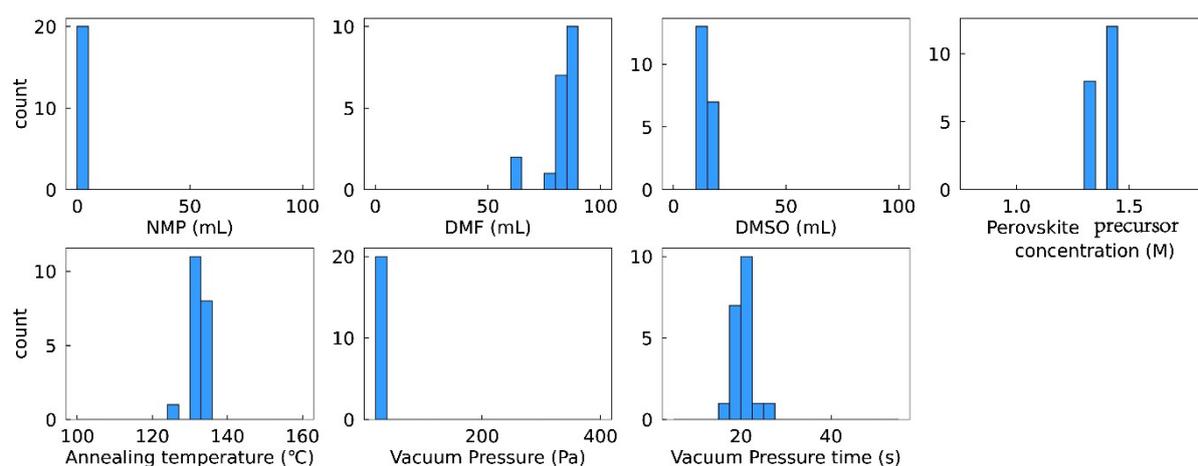
Supplementary Fig. 12 Bar plots visualizing the distribution of the third batch of experiments, selected by the BO algorithm, across the search space for each of the process variables.
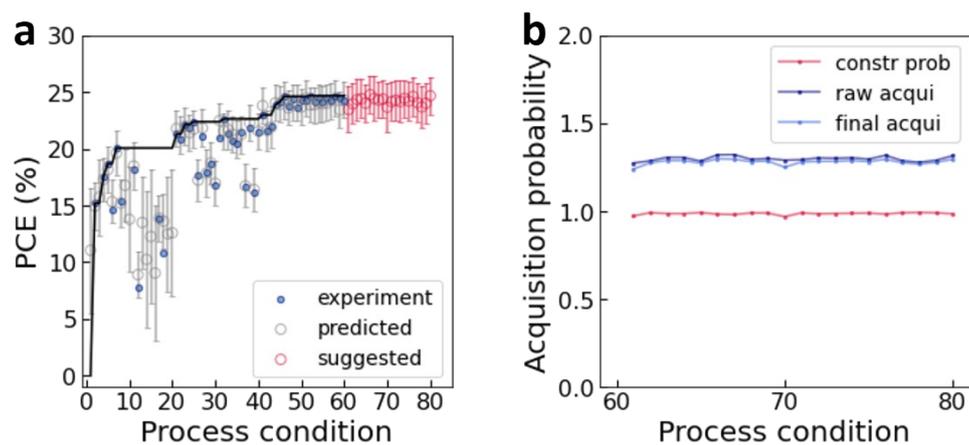
Supplementary Fig. 13 The acquisition process for the third sampling round. Plot of the predicted – and measured – PCE values for both the initial batch and the first selection of acquired points; (a) the best PCE value measured so far is denoted by the bold black curve. (b) Graphical summary of the acquisition function values across this third selection round.
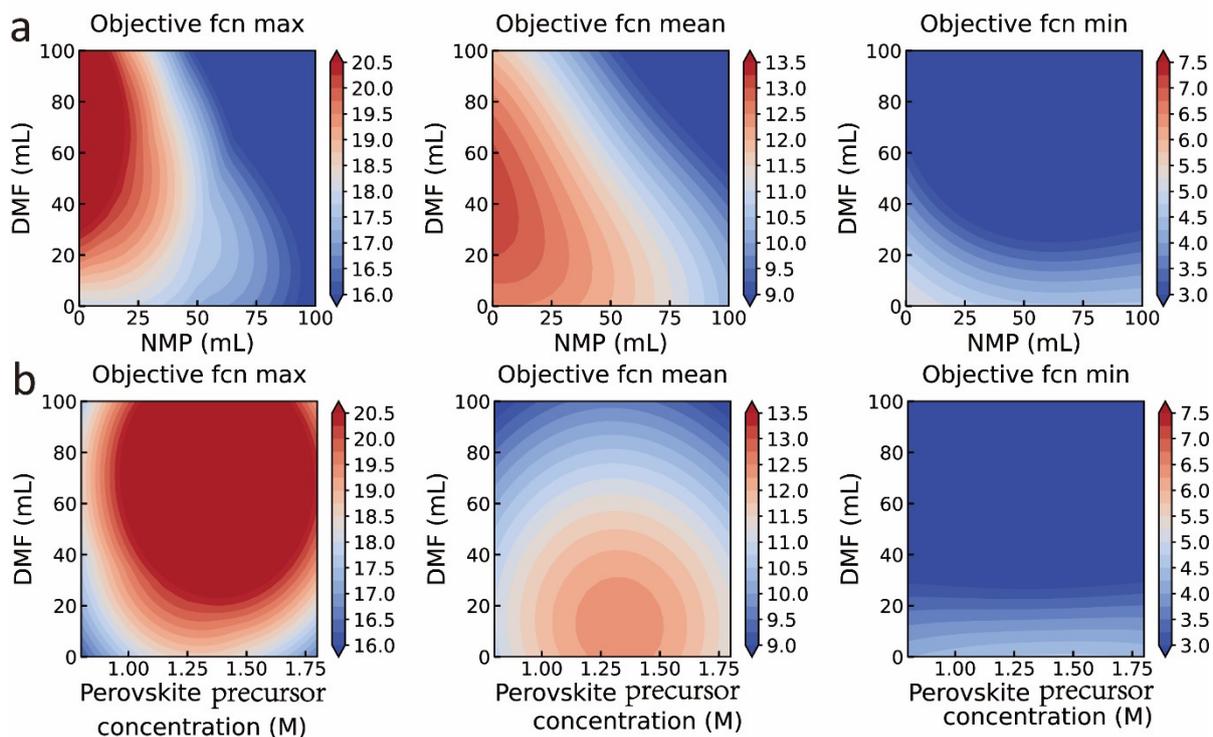
Supplementary Fig. 14 Sample of the 2D contour plots for the surrogate objective function: (a) DMF vs. NMP volumes, (b) Perovskite precursor concentration vs. DMF volume. Maximum (left), mean (middle) and minimum (right) surrogate objective function values across the sample of process conditions at every grid point are visualized.
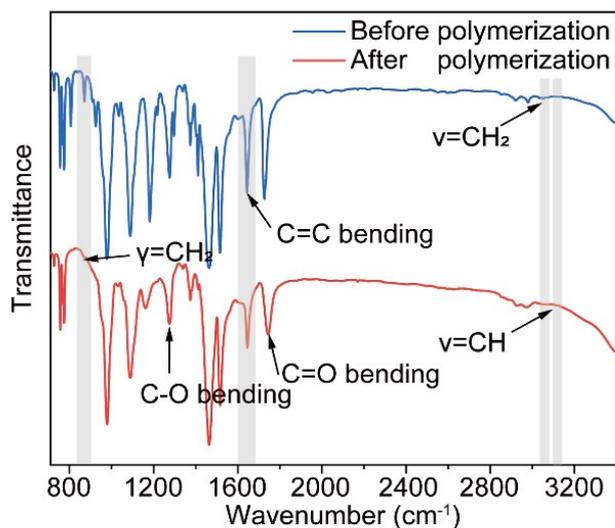
Supplementary Fig. 15 FTIR spectra of AETA-BCF before and after linear polymerization.
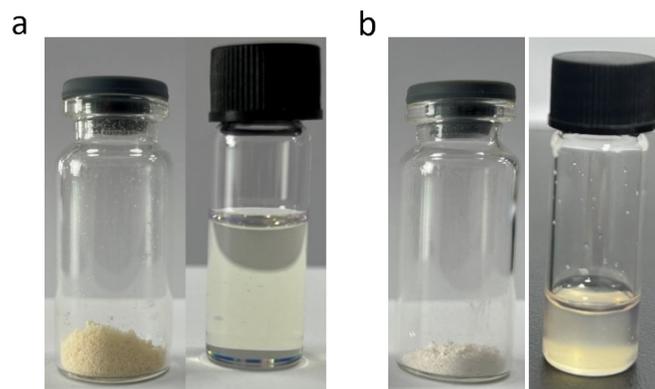
Supplementary Fig. 16 Photographs of AETA-BCF in both powder and solution forms before (a) and after (b) the polymerization. The powder samples before and after cross-linking were obtained through heat treatment.
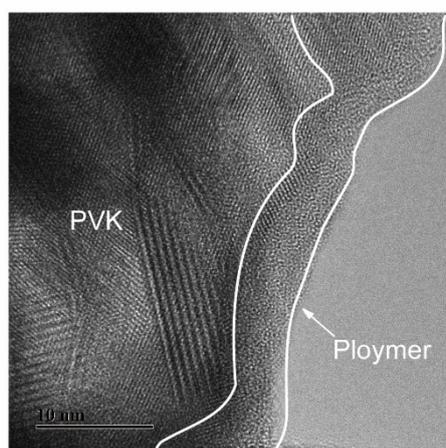
Supplementary Fig. 17 TEM morphology after AETA-BCF linear polymerization in perovskite.
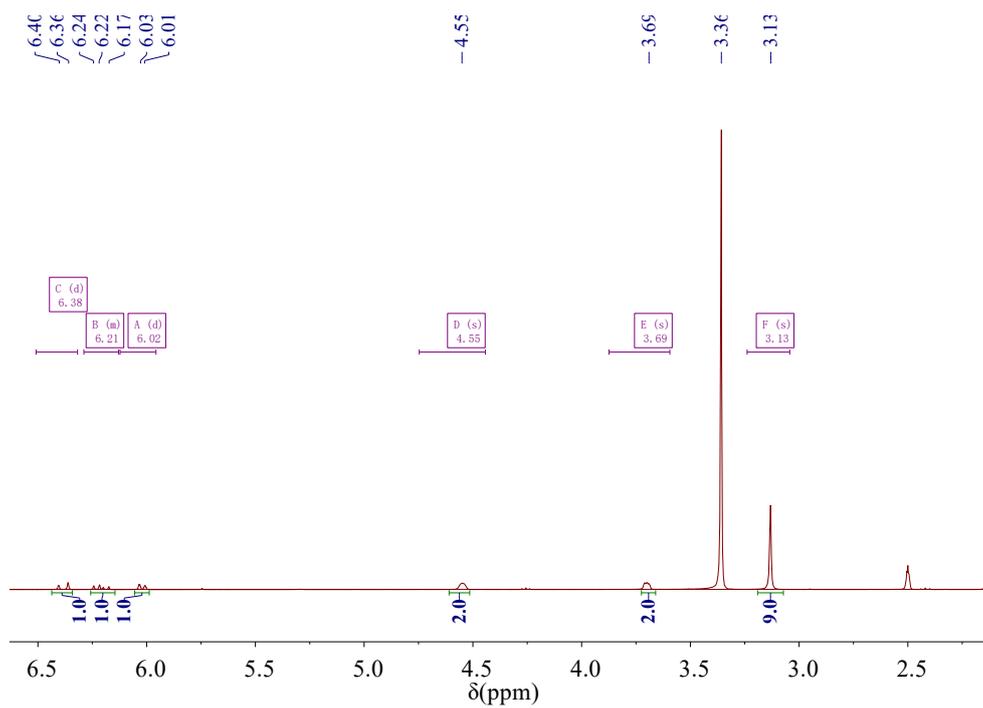
Supplementary Fig. 18 1H NMR spectrum of AETA-BCF.
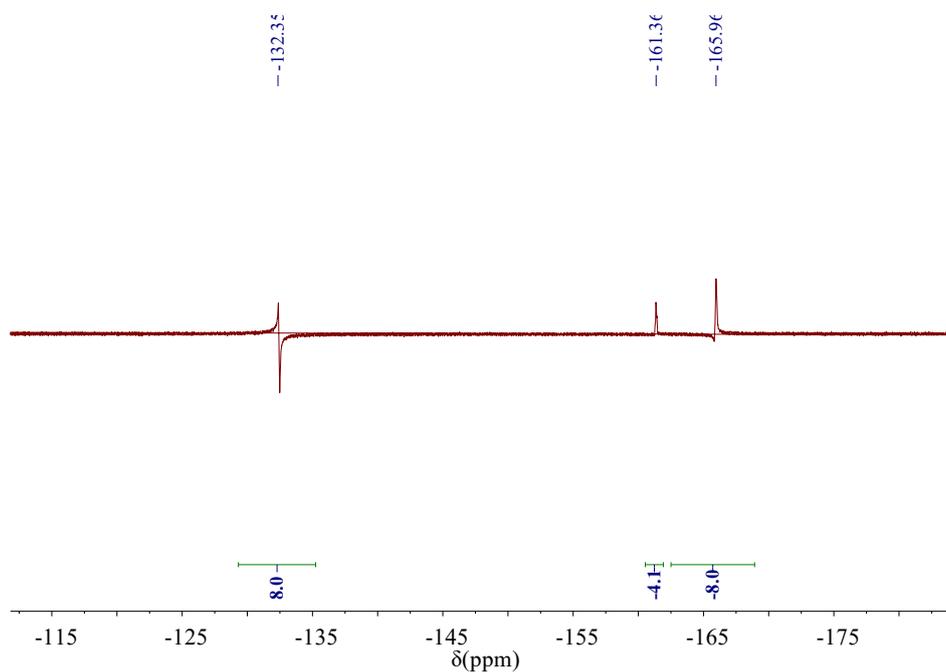
Supplementary Fig. 19 ^{19}F NMR spectrum of AETA-BCF.
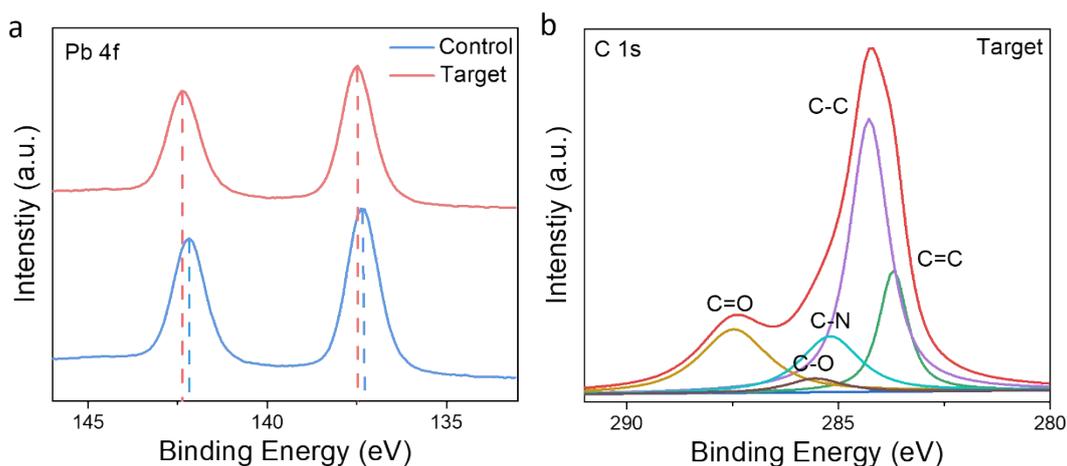
Supplementary Fig. 20 XPS spectra of (a) Pb 4f and (b) C 1s.
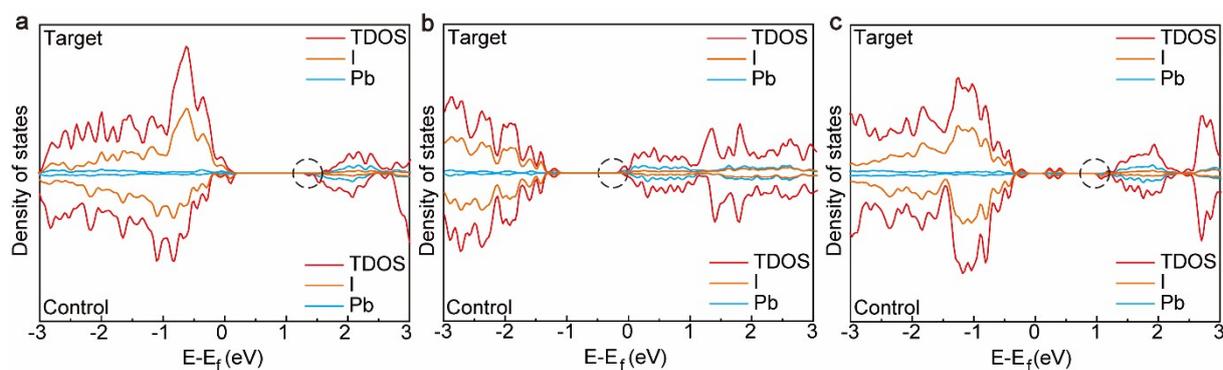
Supplementary Fig. 21 DOS of (a) V_{Pb}, (b) V_I and (c) $Pb_{I-antisite}$ before and after AETA-BCF treatment.
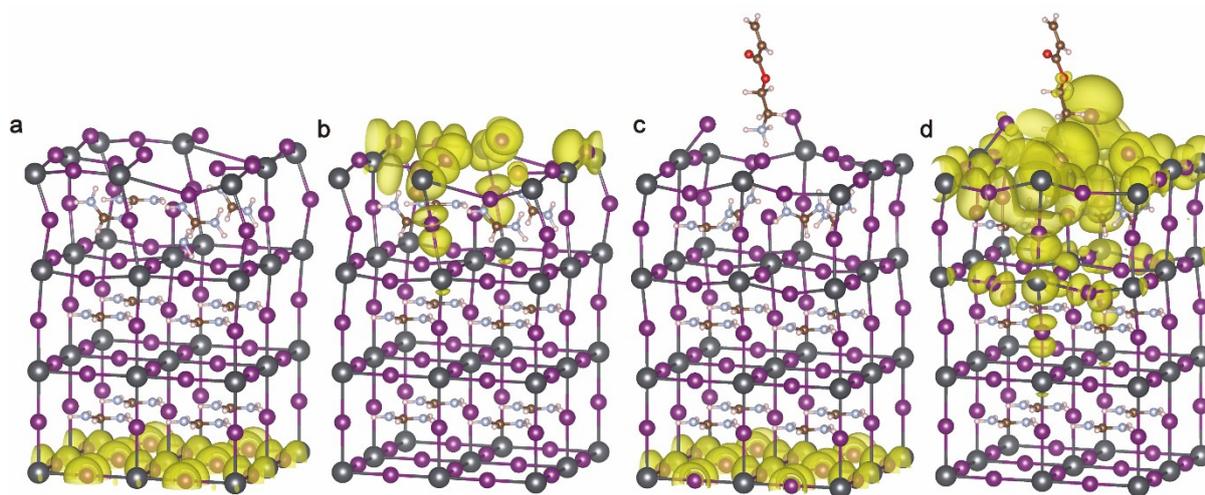
Supplementary Fig. 22 Iso-surface representations of perovskite surfaces with V_{Pb} defects, pre- and post-molecular passivation. Panels (a) and (c) show the highest occupied valence band for untreated and passivated surfaces, respectively. Panels (b) and (d) illustrate the lowest unoccupied conduction band for the same conditions. These images highlight the electronic structure changes due to V_{Pb} and the

impact of molecular passivation.
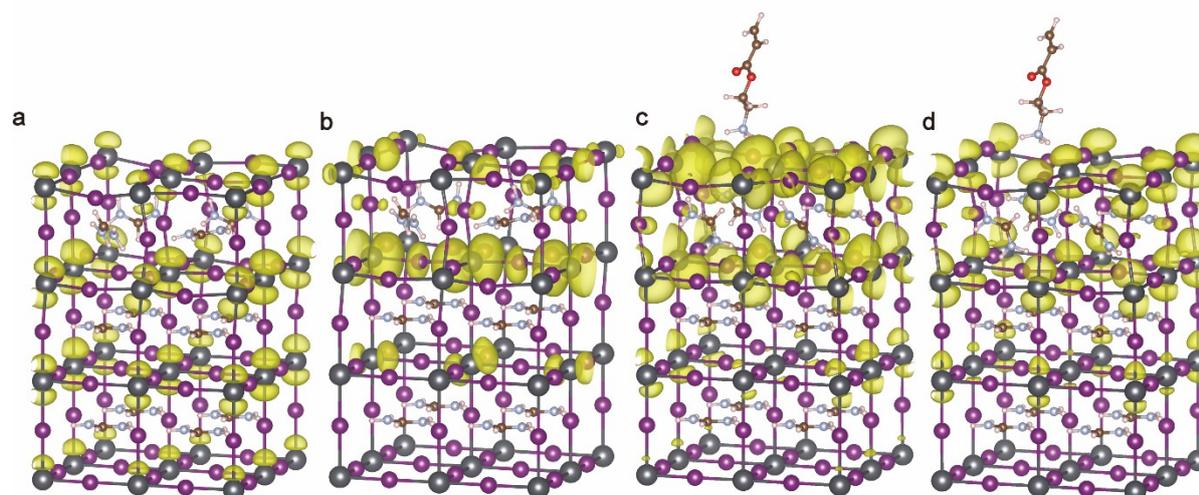
Supplementary Fig. 23 Iso-surface representations of perovskite surfaces with V_I defects, pre- and post-molecular passivation. Panels (a) and (c) show the highest occupied valence band for untreated and passivated surfaces, respectively. Panels (b) and (d) illustrate the lowest unoccupied conduction band for the same conditions. These images highlight the electronic structure changes due to V_I and the impact of molecular passivation.
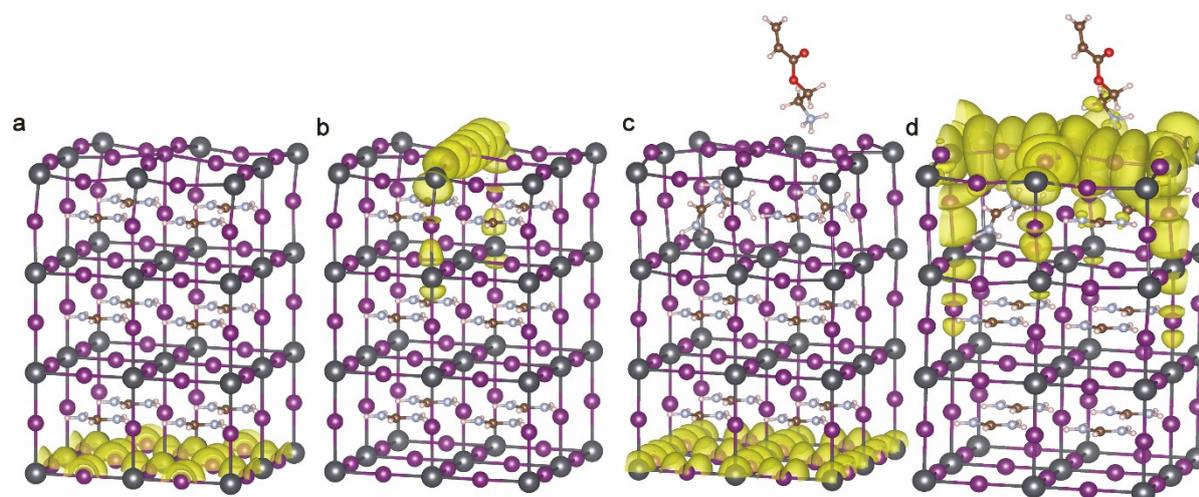
Supplementary Fig. 24 Iso-surface representations of perovskite surfaces with $Pb_{I\text{-antisite}}$ defects, pre- and post-molecular passivation. Panels (a) and (c) show the highest occupied valence band for untreated and passivated surfaces, respectively. Panels (b) and (d) illustrate the lowest unoccupied conduction band for the same conditions. These images highlight the electronic structure changes due to $Pb_{I\text{-antisite}}$ and the impact of molecular passivation.
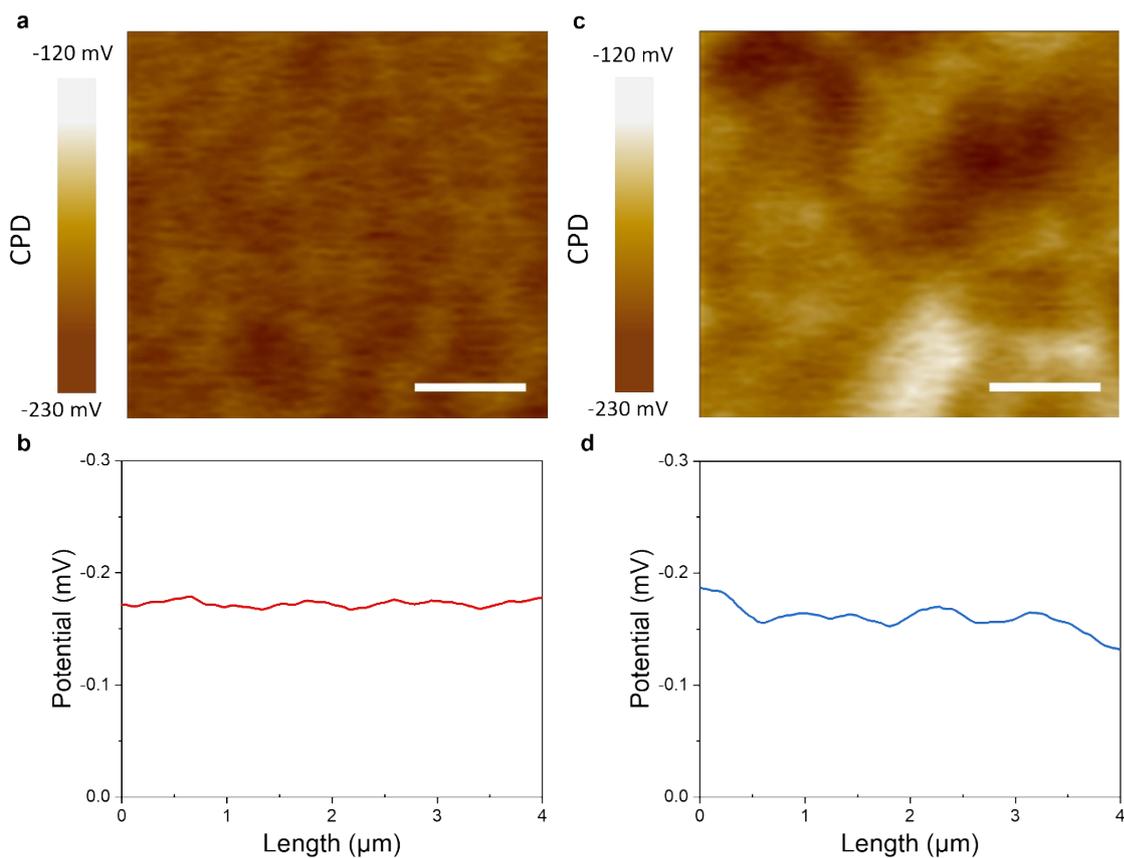
Supplementary Fig. 25 Kelvin Probe Force Microscopy (KPFM) and Surface Potential Maps of Thin Films. Panel (a) displays the KPFM image of the thin film after AETA-BCF optimization, with a scale bar of 1000 nm. Panel (b) shows the corresponding baseline surface potential map of this optimized film, and its average surface potential was found to be -0.173 mV with a fluctuation range of -0.01 mV. Panel (c) presents the KPFM image of the thin film before AETA-BCF optimization, also scaled at 1000 nm. Panel (d) illustrates the corresponding baseline surface potential map of the unoptimized film, and its average surface potential was found to be -0.162 mV with a fluctuation range of -0.55 mV. These images highlight the surface potential changes induced by the AETA-BCF optimization process in the thin films.
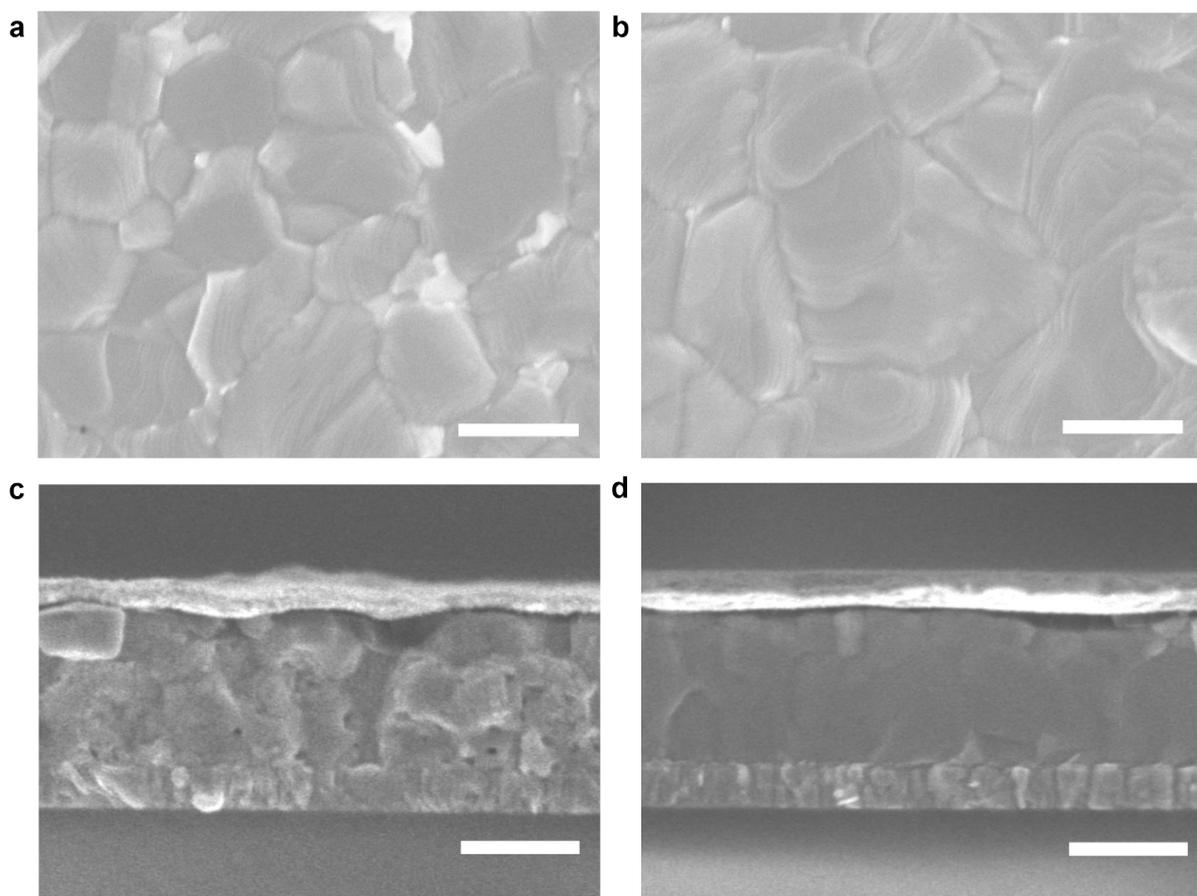
Supplementary Fig. 26 SEM Images of Control and Target Thin Films. Panels (a) and (b) present the SEM (Scanning Electron Microscopy) images of the control thin film and the target thin film, respectively, each with a scale bar of 1000 nm. Panels (c) and (d) show the cross-sectional views of the same films, with a scale bar of 500 nm for close-up detail.
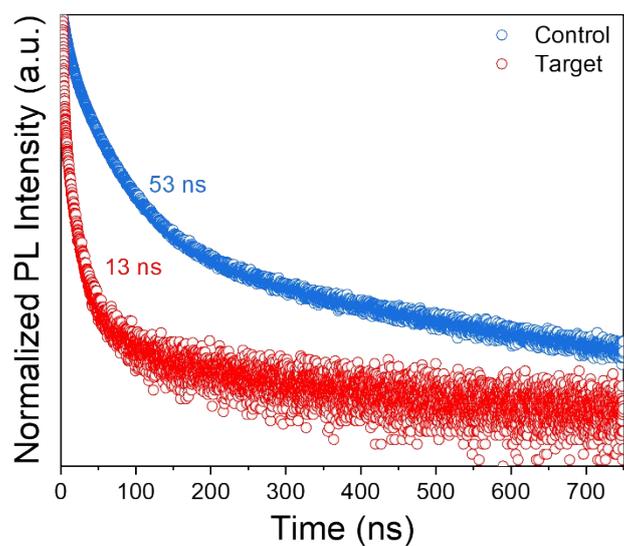
Supplementary Fig. 27 Time-Resolved Photoluminescence (TRPL) Spectra of Control and Target Thin Films. This figure presents the TRPL spectra for both the control and target thin films, with a sample structure of Glass/NiOx/MeO-2PACz/Perovskite.
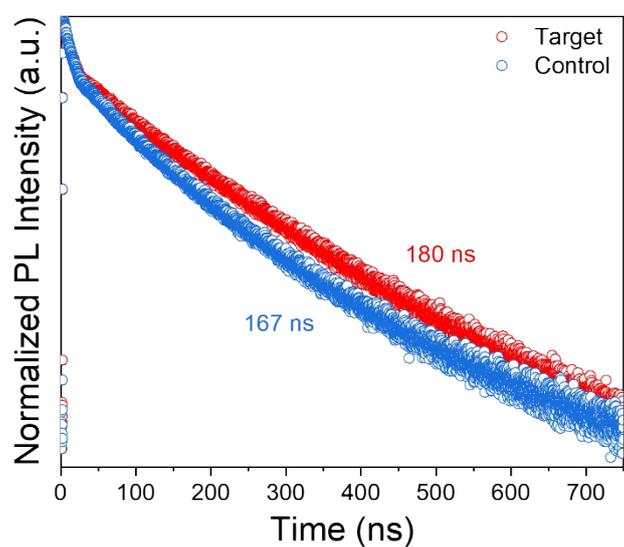
Supplementary Fig. 28 Time-Resolved Photoluminescence (TRPL) Spectra of Control and Target Thin Films. This figure presents the TRPL spectra for both the control and target thin films, with a sample structure of Glass/Perovskite.
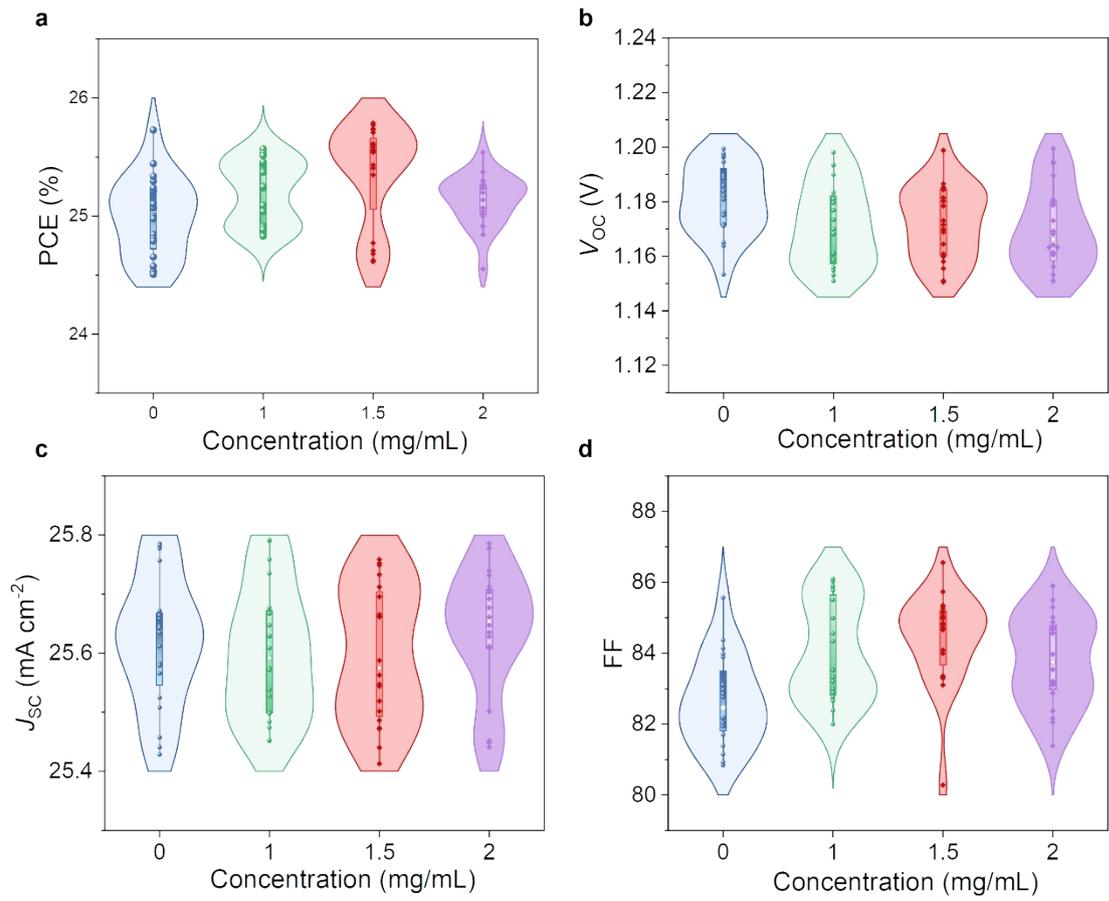
Supplementary Fig. 29 Performance Variation with AETA-BCF Concentrations. Statistical graphs depicting how different concentrations of AETA-BCF affect PCE, V_{OC}, J_{SC}, and FF in photovoltaic devices.
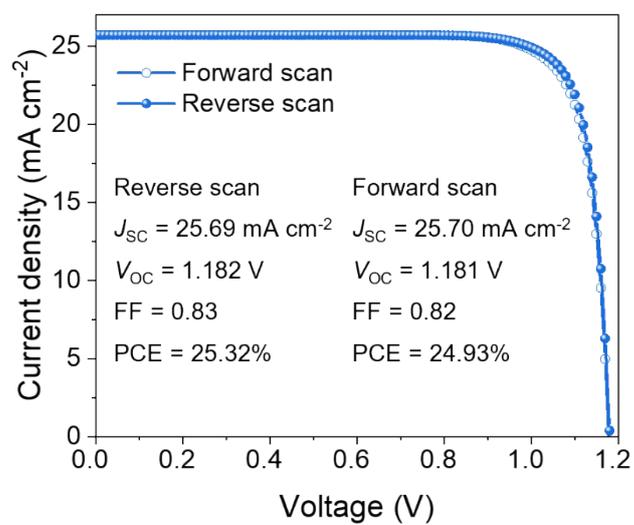
Supplementary Fig. 30 The J–V curves of the best-performing control device.
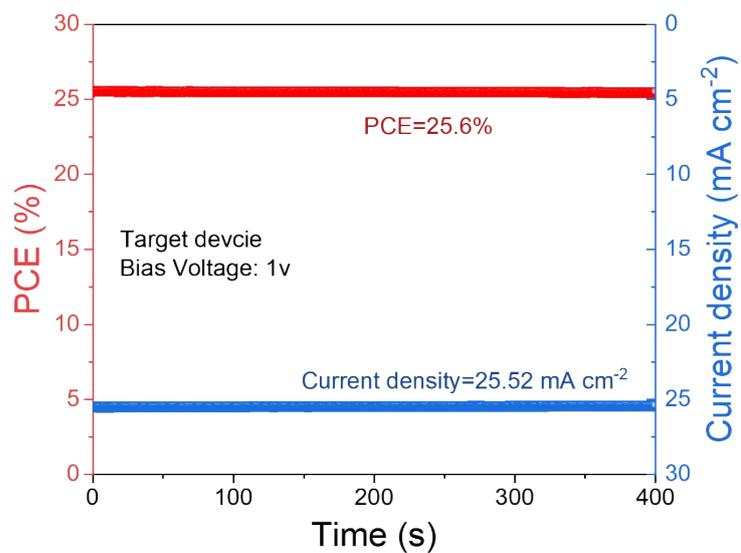
Supplementary Fig. 31 Stability of Power Output for Target Device at 1.0v Bias.

Chengdu Institute of Product Quality Inspection Co., Ltd.
National Photovoltaic Product Quality Inspection &Testing Center
TEST REPORT

Product Name	Perovskite solar cells	Trade Mark	/
Manufacture Date	/	Model /Type	2cm×2cm
Sample No.	AGXB123W00309	Sample Grade	/
Sample Quantity	One piece	Sample State	/
Delivery Date	19/09/2023	Sample Delivered personnel	Junsheng Luo/Boxue Zhang
Commission unit	Junsheng Luo, University of Electronic Science and Technology of China	Manufacturer	School of Integrated Circuit Science and Engineering, University of Electronic Science and Technology of China/ Chimie ParisTech, PSL Research University
Commission unit address	University of Electronic Science and Technology of China, No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, P. R. China	Manufacturer Address	University of Electronic Science and Technology of China, No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, P. R. China/ Chimie ParisTech, PSL Research University, 11 rue P. et M. Curie, F-75005 Paris, France
Commission unit Zip code	611731	Manufacturer Zip code	611731/75005
Commission unit Tel.	18215558952	Manufacturer Tel.	18215558952/15104486052
Center Address	No. 355, 2nd Tengfei Road, Southwest Airport Economic Development Zone, Chengdu, Sichuan, P. R. China.	Measurement Date	19/09/2023
Methods	IEC 60904-1:2020 Photovoltaic devices-Part 1: Measurement of Photovoltaic Current-Voltage Characteristics.		
Test conclusion	This column blank.		
Remarks	The mask area is provided by the Commission unit: 0.0911cm^2.		

(Special chapter for test report)

Issue Date 22/09/2023

Approved by	Reviewed by	Measured by

Chengdu Institute of Product Quality Inspection Co., Ltd.
National Photovoltaic Product Quality Inspection &Testing Center
TEST REPORT

Test Report No. AGXB123W00309   Page 2 of 3

Test Results（Forward scanning）:

No.	Test item(s)	Unit	Results
1	Current-voltage characteristics measurement	---	---
1.1	Open-circuit voltage, V_{OC}	V	1.185
1.2	Short-circuit current, I_{SC}	mA	2.342
1.3	Maximum-power, P_{max}	mW	2.267
1.4	Maximum-power voltage, V_{p-max}	V	1.030
1.5	Maximum-power current, I_{p-max}	mA	2.201
1.6	Fill factor, FF	%	81.69
1.7	Conversion efficiency, η	%	24.88

Current-voltage characteristics
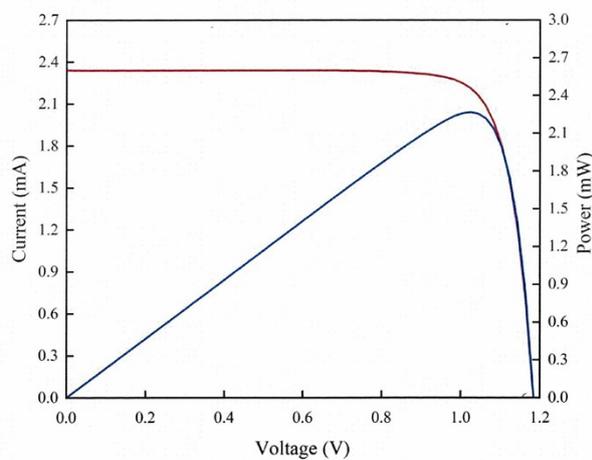
Remark: Sample was tested under the irradiation with a steady-state class calibrated AAA solar simulator (AM1.5-G 1000.0 W/m² basted on mono-Si reference cell) at 25 ± 1 °C. Designated area defined by thin metal aperture mask.

Chengdu Institute of Product Quality Inspection Co., Ltd.
National Photovoltaic Product Quality Inspection &Testing Center
TEST REPORT

Test Report No. AGXB123W00309   Page 3 of 3

Test Results（Reverse scanning）:

No.	Test item(s)	Unit	Results
1	Current-voltage characteristics measurement	---	---
1.1	Open-circuit voltage, V_{OC}	V	1.191
1.2	Short-circuit current, I_{SC}	mA	2.343
1.3	Maximum-power, P_{max}	mW	2.297
1.4	Maximum-power voltage, V_{p-max}	V	1.038
1.5	Maximum-power current, I_{p-max}	mA	2.213
1.6	Fill factor, FF	%	82.32
1.7	Conversion efficiency, η	%	25.21

Current-voltage characteristics under STC

Remark: Sample was tested under the irradiation with of steady state class calibrated AAA solar simulator (AM1.5-G 1000.0 W/m² basted on mono-Si reference cell) at 25±1 °C. Designated area defined by thin metal aperture mask.

————Blank————
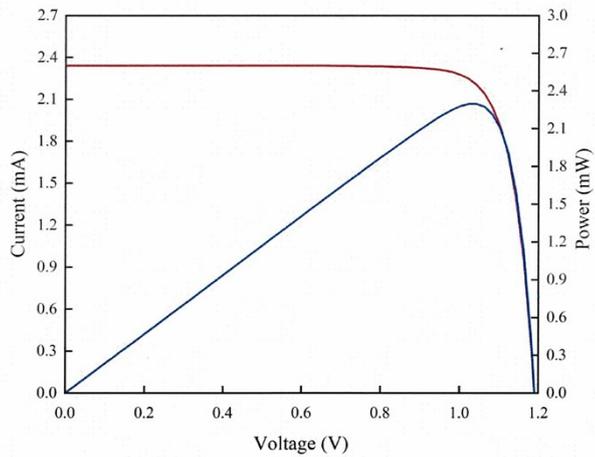
Supplementary Fig. 32 Efficiency Certification. This figure presents the certified efficiency report obtained from the National Photovoltaic Quality Inspection and Testing Center, validating the performance metrics of our photovoltaic device.
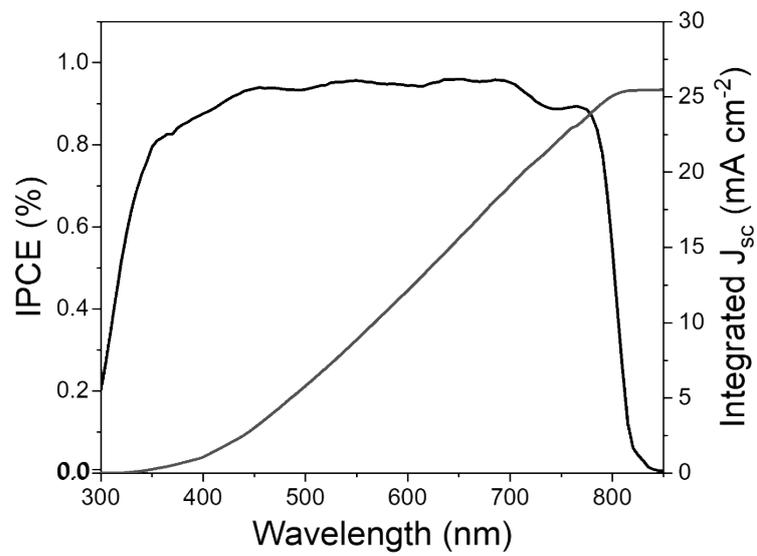
Supplementary Fig. 33 IPCE spectrum of target device with AETA-BCF passivation, the corresponding integrated current is 25.43 mA cm^{-2}.
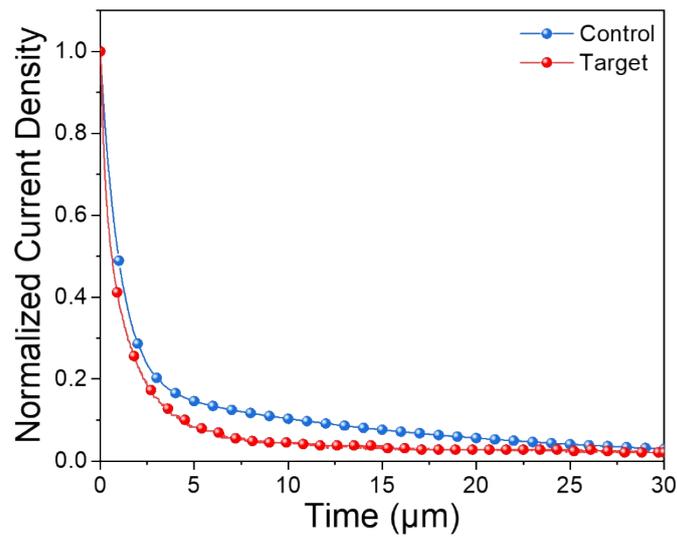
Supplementary Fig. 34 Transient photocurrent curves for the control and target devices.
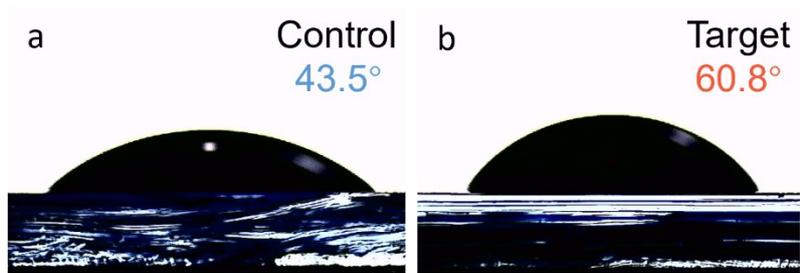
Supplementary Fig. 35 The water contact angle measurement for PSCs of control and target film.
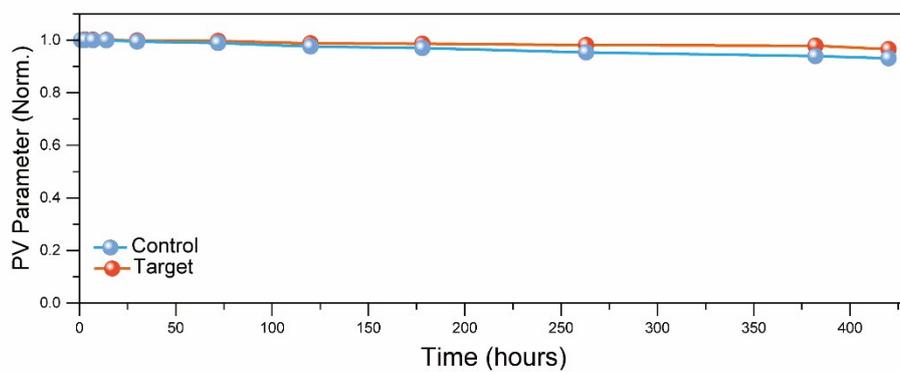
Supplementary Fig. 36 The long-term stability for PSCs of control and target devices for the encapsulated under AM 1.5 illumination in ambient air.

Supplementary Table 1. Compilation of 60 Perovskite Compositions Used in This Study. This table provides a detailed list of the 60 perovskite compositions explored in our research, outlining their specific formulations. This percentage refers to the molar concentration ratio of MACl added to the perovskite precursor solution. *For detailed solution preparation, solution concentrations, and their respective volume ratios, please refer to the Experimental Section. The 60 different $FAPbI_3$-based perovskite components are classified and arranged based on the types and ratios of additives used. The components are grouped as follows: (1) $FAPbI_3$ with Single Additives: These compositions include $FAPbI_3$ with varying concentrations of a single additive such as RbI, CsI, MAI, or MACl. They are arranged in ascending order of additive concentration. The corresponding components are highlighted in red. (2) $FAPbI_3$ with Dual Additives: These compositions involve combinations of two additives. They are categorized based on the types of additives and their relative concentrations. The corresponding components are highlighted in orange. (3) $FAPbI_3$ with Multiple Additives: These compositions contain three or more additives. The corresponding components are highlighted in green. (4) Mixed Cation and Halide Perovskites: These include compositions where multiple cations (Rb, Cs, MA) and halides (I, Br, Cl) are mixed with $FAPbI_3$. The corresponding components are highlighted in blue.

$FAPbI_3$-based Composites and corresponding recipes
serial Number	Target Composites	Recipes*	Serial number	Target composition	Recipe
1	$FAPbI_3$	200 µL $FAPbI_3$	31	$FAPbI_3$+3%RbI+7%MAI	6µL RbI + 14µL MAI + 200 µL $FAPbI_3$
2	$FAPbI_3$+3%RbI	6µL RbI + 200 µL $FAPbI_3$	32	$FAPbI_3$+7%RbI+3%MAI	14µL RbI + 6µL MAI + 200 µL $FAPbI_3$
3	$FAPbI_3$+7%RbI	14µL RbI + 200 µL $FAPbI_3$	33	$FAPbI_3$+5%CsI+2%MAI+3%RbI+50% MACl	10µL CsI + 4µL MAI + 6 RbI + 200 µL $FAPbI_3$+100µL MACl
4	$FAPbI_3$+15%RbI	30µL RbI + 200 µL $FAPbI_3$	34	$FAPbI_3$+3%CsI+7%MAI	5µL CsI + 14µL MAI + 200 µL $FAPbI_3$
5	$FAPbI_3$+3%CsI	6µL CsI + 200 µL $FAPbI_3$	35	$FAPbI_3$+7%CsI+3%MAI	14µL CsI + 6µL MAI + 200 µL $FAPbI_3$
6	$FAPbI_3$+7%CsI	14µL CsI + 200 µL $FAPbI_3$	36	$FAPbI_3$+5%CsI+2%MAI+3%RbI	10µL CsI + 4µL MAI + 6µL RbI + 200 µL $FAPbI_3$
7	$FAPbI_3$+15%CsI	30µL CsI + 200 µL $FAPbI_3$	37	$FAPbI_3$+5%RbI+10%MAI	10µL RbI + 20µL MAI + 200 µL $FAPbI_3$
8	$FAPbI_3$+3%MAI	6µL MAI + 200 µL $FAPbI_3$	38	$FAPbI_3$+10%RbI+5%MAI	20µL RbI + 10µL MAI + 200 µL $FAPbI_3$
9	$FAPbI_3$+7%MAI	14µL MAI + 200 µL $FAPbI_3$	39	$FAPbI_3$+10%RbI+5%MACl	20µL RbI + 10µL MACl + 200 µL $FAPbI_3$
10	$FAPbI_3$+15%MAI	30µL MAI + 200 µL	40	$FAPbI_3$+7%RbI+3%CsI+5%MAI	14µL RbI + 6µL CsI +

No.	Composition	Recipe	No.	Composition	Recipe
	$FAPbI_3$				$10\mu L$ MAI + $200\ \mu L$ $FAPbI_3$
11	$FAPbI_3$+10%MACl	$20\mu L$ MACl + $200\ \mu L$ $FAPbI_3$	41	$FAPbI_3$+5%CsI+10% MAI	$10\mu L$ CsI + $20\mu L$ MAI + $200\ \mu L$ $FAPbI_3$
12	$FAPbI_3$+20%MACl	$40\mu L$ MACl + $200\ \mu L$ $FAPbI_3$	42	$FAPbI_3$+10%CsI+5% MAI	$20\mu L$ CsI + $10\mu L$ MAI + $200\ \mu L$ $FAPbI_3$
13	$FAPbI_3$+30%MACl	$60\mu L$ MACl + $200\ \mu L$ $FAPbI_3$	43	$FAPbI_3$+10%CsI+5% MACl	$20\mu L$ CsI + $10\mu L$ MACl + $200\ \mu L$ $FAPbI_3$
14	$FAPbI_3$+50%MACl	$100\mu L$ MACl + $200\ \mu L$ $FAPbI_3$	44	$FAPbI_3$+3%RbI+7%CsI+5%MAI	$6\mu L$ RbI + $14\mu L$ CsI + $10\mu L$ MAI + $200\ \mu L$ $FAPbI_3$
15	$(Rb_{0.03}FA_{0.97})PbI_3$	$6\mu L$ $RbPbI_3$ + $200\ \mu L$ $FAPbI_3$	45	$FAPbI_3$+5%CsI+15%MAI	$10\mu L$ CsI + $30\mu L$ MAI + $200\ \mu L$ $FAPbI_3$
16	$(Rb_{0.07}FA_{0.93})PbI_3$	$14\mu L$ $RbPbI_3$ + $200\ \mu L$ $FAPbI_3$	46	$FAPbI_3$+15%CsI+5%MAI	$30\mu L$ CsI + $10\mu L$ MAI + $200\ \mu L$ $FAPbI_3$
17	$(Rb_{0.15}FA_{0.85})PbI_3$	$30\mu L$ $RbPbI_3$ + $200\ \mu L$ $FAPbI_3$	47	$FAPbI_3$+5%CsI+15%MACl	$10\mu L$ CsI + $30\mu L$ MACl + $200\ \mu L$ $FAPbI_3$
18	$(Cs_{0.03}FA_{0.97})PbI_3$	$6\mu L$ $CsPbI_3$ + $200\ \mu L$ $FAPbI_3$	48	$FAPbI_3$+5%RbI+10%CsI+5%MAI	$10\mu L$ RbI + $20\mu L$ CsI + $10\mu L$ MAI + $200\ \mu L$ $FAPbI_3$
19	$(Cs_{0.7}FA_{0.93})PbI_3$	$14\mu L$ $CsPbI_3$ + $200\ \mu L$ $FAPbI_3$	49	$(Rb_{0.03}MA_{0.07}FA_{0.9})PbI_3$	$6\mu L$ $RbPbI_3$ + $14\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$
20	$(Cs_{0.15}FA_{0.85})PbI_3$	$30\mu L$ $CsPbI_3$ + $200\ \mu L$ $FAPbI_3$	50	$(Rb_{0.07}MA_{0.03}FA_{0.9})PbI_3$	$14\mu L$ $RbPbI_3$ + $6\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$
21	$MAPbI_3$	$200\ \mu L$ $MAPbI_3$	51	$(Rb_{0.05}Cs_{0.03}MA_{0.02}FA_{0.9})PbI_3$	$10\mu L$ $RbPbI_3$ + $6\mu L$ $CsPbI_3$ + $4\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$
22	$(MA_{0.03}FA_{0.97})PbI_3$	$6\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$	52	$(Rb_{0.03}MA_{0.07}FA_{0.90})Pb(I_{0.95}Br_{0.03}Cl_{0.02})_3$	$6\mu L$ $RbPbI_3$ + $14\mu L$ $MAPbCl_3$ + $9\mu L$ **$FAPbBr_3$** + $200\ \mu L$ $FAPbI_3$
23	$(MA_{0.07}FA_{0.93})PbI_3$	$14\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$	53	$(Cs_{0.05}MA_{0.10}FA_{0.85})PbI_3$	$10\mu L$ $CsPbI_3$ + $20\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$
24	$(MA_{0.15}FA_{0.85})PbI_3$	$30\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$	54	$(Cs_{0.10}MA_{0.05}FA_{0.85})PbI_3$	$20\mu L$ $CsPbI_3$ + $10\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$
25	$(MA_{0.03}FA_{0.97})PbI_{2.97}Cl_{0.03}$	$6\mu L$ $MAPbCl_3$ + $200\ \mu L$ $FAPbI_3$	55	$(Rb_{0.03}Cs_{0.10}MA_{0.02}FA_{0.85})PbI_3$	$6\mu L$ $RbPbI_3$ + $20\mu L$ $CsPbI_3$ + $4\mu L$ $MAPbI_3$ + $200\ \mu L$ $FAPbI_3$
26	$(MA_{0.07}FA_{0.93})PbI_{2.93}Cl_{0.07}$	$14\mu L$ $MAPbCl_3$ + $200\ \mu L$ $FAPbI_3$	56	$(Rb_{0.03}MA_{0.07}FA_{0.90})Pb(I_{0.95}Br_{0.03}Cl_{0.02})_3$ + 30%MACl	$6\mu L$ $RbPbI_3$ + $14\mu L$ $MAPbCl_3$ + $9\mu L$ $FAPbBr_3$ + $200\ \mu L$ $FAPbI_3$ + $60\mu L$ MACl
27	$(MA_{0.15}FA_{0.85})PbI_{2.85}Cl_{0.15}$	$30\mu L$ $MAPbCl_3$ + $200\ \mu L$ $FAPbI_3$	57	$(MA_{0.02}Cs_{0.05}FA_{0.93})Pb(I_{0.97}Br_{0.01}Cl_{0.02})_3$ + 50%MACl	$4\mu L$ $MAPbCl_3$ + $10\mu L$ $CsPbI_3$ + $2\mu L$ $FAPbBr_3$ + $200\ \mu L$ $FAPbI_3$ + $100\mu L$ MACl

28	$(MA_{0.10}FA_{0.90})Pb(I_{0.95}Br_{0.02}Cl_{0.03})_3$	20µL MAPbCl$_3$ + 12µL **FAPbBr$_3$** + 200 µL FAPbI$_3$	58	$(Cs_{0.05}MA_{0.15}FA_{0.80})PbI_3$	10µL CsPbI$_3$ + 30µL MAPbI$_3$ + 200 µL FAPbI$_3$
29	$(MA_{0.10}FA_{0.90})Pb(I_{0.90}Br_{0.07}Cl_{0.03})_3$	20µL MAPbCl$_3$ + 41µL **FAPbBr$_3$** + 200 µL FAPbI$_3$	59	$(Cs_{0.15}MA_{0.05}FA_{0.80})PbI_3$	30µL CsPbI$_3$ + 10µL MAPbI$_3$ + 200 µL FAPbI$_3$
30	$(MA_{0.20}FA_{0.80})Pb(I_{0.90}Br_{0.03}Cl_{0.07})_3$	41µL MAPbCl$_3$ + 18µL **FAPbBr$_3$** + 200 µL FAPbI$_3$	60	$(Rb_{0.05}Cs_{0.05}MA_{0.10}FA_{0.80})PbI_3$	10µL RbPbI$_3$ + 10µL CsPbI$_3$ + 20µL MAPbI$_3$ + 200 µL FAPbI$_3$

Supplementary Table 2: An overview of the search space considered in this work.

Process variable	Total range (interval)
Volume NMP (ml)	0 – 100 (5)
Volume DMF (ml)	0 – 100 (5)
Volume DMSO (ml)	0 – 100 (5)
Perovskite precursor concentration (M)	0.8 – 1.8 (0.1)
Annealing temperature (°C)	100 – 160 (5)
Vacuum pressure (Pa)	20 – 400 (10)
Vacuum pressure time (s)	0 – 55 (2)

Supplementary Table 3: The region in the multidimensional parameter space where our surrogate model, trained with the help of Bayesian optimization, indicateds the highest likelihood of achieving an optimal device performance.

Process variable	Best range (interval)
Volume NMP (ml)	0 – 20 (5)
Volume DMF (ml)	60 – 100 (5)
Volume DMSO (ml)	20 – 60 (5)
Perovskite precursor concentration (M)	1.3 – 1.7 (0.1)
Annealing temperature (°C)	120 – 140 (5)
Vacuum pressure (Pa)	20 – 60 (10)
Vacuum pressure time (s)	20 – 40 (2)

Supplementary References:

1   H. Chen, S. Teale, B. Chen, Y. Hou, L. Grater, T. Zhu, K. Bertens, S. M. Park, H. R. Atapattu, Y. Gao, M. Wei, A. K. Johnston, Q. Zhou, K. Xu, D. Yu, C. Han, T. Cui, E. H. Jung, C. Zhou, W. Zhou, A. H. Proppe, S. Hoogland, F. Laquai, T. Filleter, K. R. Graham, Z. Ning and E. H. Sargent. Quantum-size-tuned heterostructures enable efficient and stable inverted perovskite solar cells. *Nature Photonics* **16**, 352-358, doi:10.1038/s41566-022-00985-1 (2022).

2   G. Kresse av, J. F. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* **6**, 15-50, doi: 10.1016/0927-0256(96)00008-0 (1996).

3   Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **77**, 3865-3868, doi:10.1103/PhysRevLett.77.3865 (1996).

4   Guangyu Suna, Jenö Kürtia,b,Péter Rajczyb, Miklos Kertesza,,Jürgen Hafnerc, Georg Kressec. Performance of the Vienna ab initio simulation package (VASP) in chemical applications. *Journal of Molecular Structure (Theochem)* **624**, 37-45, doi: 10.1016/S0166-1280(02)00733-9 (2002).

5   Meunier, M. & Robertson, S. Materials Studio 20th anniversary. *Molecular Simulation* **47**, 537-539, doi:10.1080/08927022.2021.1892093 (2021).

6   B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams and A. G. Doyle. Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **590**, 89-96, doi:10.1038/s41586-021-03213-y (2021).

7   J. A. G. Torres, S. H. Lau, P. Anchuri, J. M. Stevens, J. E. Tabora, J. Li, A. Borovika, R. P. Adams and A. G. Doyle. A Multi-Objective Active Learning Platform and Web App for Reaction Optimization. *Journal of the American Chemical Society* **144**, 19999-20007, doi:10.1021/jacs.2c08592 (2022).

8   Frazier, P. I. A tutorial on Bayesian optimization. *arXiv preprint* **1807.02811,** doi: 10.48550/arXiv.1807.02811 (2018).

9   Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE* **104**, 148-175, doi:10.1109/jproc.2015.2494218 (2016).

10   Genton, M. G. Classes of Kernels for Machine Learning: A Statistics Perspective. *Journal of Machine Learning Research* **2**, 299-312, doi: 10.5555/944790.944815 (2001).

11   Srinivas, N., Krause, A., Kakade, S. M. & Seeger, M. W. Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting. *IEEE Transactions on Information Theory* **58**, 3250-3265, doi:10.1109/TIT.2011.2182033 (2012).

12   B. Rohr, H. S. Stein, D. Guevarra, Y. Wang, J. A. Haber, M. Aykol, S. K. Suram and J. M. Gregoire. Benchmarking the acceleration of materials discovery by sequential learning. *Chemical Science* **11**, 2696-2706, doi:10.1039/C9SC05999G (2020).

13   Z. Liu, N. Rolston, A. C. Flick, T. W. Colburn, Z. Ren, R. H. Dauskardt and T. Buonassisi. Machine learning with knowledge constraints for process optimization of open-air perovskite solar cell manufacturing. *Joule* **6**, 834-849, doi:10.1016/j.joule.2022.03.003 (2022).

14   Michael A. Gelbart, J. S., Ryan P. Adams. Bayesian Optimization with Unknown

Constraints. *arXiv preprint*, arXiv:**1403.5607**, doi: 10.48550/arXiv.1403.5607 (2014).

15   Andrei Paleyes, M. P., Maren Mahsereci, Cliff McCollum, Neil Lawrence. Emulation of physical processes with Emukit. *In Second Workshop on Machine Learning and the Physical Sciences* **NeurIPS,** arXiv:**2110.13293,** doi:10.48550/arXiv.2110.13293 (2019).

16   GPy: A Gaussian process framework in Python. Available at: \url{http://github.com/SheffieldML/GPy}. (since 2012).