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

Abstract

Achieving high-performance inverted perovskite solar cells (PSCs) still remains a significant challenge, necessitating innovative approaches in materials selection and manufacturing technique optimization of perovskites. In this work, we unveil a paradigm shift in PSCs optimization. Through a judicious selection from a repertoire of 60 perovskite variants, we identified a composition with exemplary optical, thermal and electrical stability. Employing Bayesian machine learning, we navigated a labyrinth of over 1 billion process conditions, culminating in a record-breaking efficiency within a mere 80 iterations. Finally, the integration of bespoke in-situ polymerized ionic molecules allowed us to further augment performance of inverted PSCs, reaching an unparalleled power conversion efficiency of 25.76% (certified at 25.21%). The PSCs retained 94% of the initial efficiency after continuous operation in nitrogen atmosphere at 65 °C for 1,920 hours. This work not only redefines the benchmarks for PSCs but also illuminates the path forward for photovoltaic innovations.

Introduction

The field of photovoltaics has undergone a paradigm shift with the advent of perovskite solar cells (PSCs), which have emerged as frontrunner materials due to their exceptional optoelectronic properties, ease of fabrication, and cost-effectiveness[1-6]. Since their initial demonstration in 2009, PSCs have experienced an unprecedented surge in power conversion efficiencies (PCEs), improving from an initial 3.8% to over 26% within more than a decade, rivaling and even surpassing traditional silicon-based solar cells[7-9].

Despite this impressive progress, the journey towards achieving optimal performance and long-term stability in PSCs is still fraught with challenges. The intrinsic instability of perovskite materials, coupled with their susceptibility to environmental factors such as humidity, temperature, and oxygen, poses significant hurdles to their commercialization[10, 11]. Ion

migration within the perovskite lattice, leading to material degradation and device failure, further complicates the stability issue[12-15].

The vast and complex parameter space of perovskite fabrication, encompassing variations in composition, morphology, and processing conditions, adds another layer of complexity to the optimization process[16, 17]. Traditional experimental methodologies, reliant on iterative trial-and-error, are rendered impractical given the enormity of the parameter space, with many trillions of possible process conditions[18-20]. This underscores the importance of adopting innovative strategies to accelerate the identification of optimal material combinations and process parameters.

In response to these challenges, this work adopts a holistic and innovative approach, integrating advanced machine learning techniques with molecular engineering to navigate the labyrinth of perovskite optimization. We started our approach from a comprehensive library of 60 perovskite compositions and meticulously identified a variant that exhibited unparalleled optical, thermal and electrical stability. Utilizing Bayesian machine learning (ML), we then explored a high-dimensional parameter perovskite film preparation space of over 1 billion film formation process conditions and achieved a device efficiency exceeding 25% within just 80 iterations. This result underscores the efficiency of ML in process optimization.

To further augment device efficiency and stability, we have then introduced the *in situ polymerization* of ionic polymer into perovskite as a grain boundary modulator. The polymerizable ionic molecule, which includes a constellation of functional groups, is shown to provides a robust protection and passivation to the perovskite film, which improve the efficiency and tolerance against light, heat and moisture of PSCs.[21, 22] By integrating the machine learning and molecular engineering approach, the inverted PSCs culminated in a PCE of 25.75%, featuring a high open circuit voltage (V_{OC}) of 1.19 V, with a certified efficiency of 25.21%. The unencapsulated PSCs retained 94% of the initial efficiency after 1,920 hours at maximum power point (MPP) tracking under continuous 1-sun illumination at 65 °C in N_2 atmosphere, setting a new benchmark in PSC stability. This work not only redefines the performance benchmarks for PSCs but also provides a comprehensive framework for the design and optimization of future photovoltaic devices, showcasing the transformative impact of

integrating machine learning and molecular engineering in the field.

Components screening

As depicted in **Figure 1a**, a systematic screening process was conducted across 60 compositions to identify the most stable perovskite formulations. Perovskite structures, characterized by an ABX_3 framework, typically involve a monovalent cation "A", lead (Pb) as "B", and a halogen element as "X". Our approach for customizing perovskite compositions involved introducing various cations (FA, MA, Cs, Rb, K) at the "A" site and controlled halide mixtures (Br, Cl) at the "X" site (up to 15%), leading to design of 60 distinct perovskite formulations, as documented in **Supplementary Table 1**. Perovskite films were prepared by the vacuum flash-assisted growth technique (VFGT)[23, 24]. Recent advancements in VFGT have shown promising results in not only matching the thermal stability offered by spin-coating but also in potentially outperforming it, particularly noted in the stabilization of the black phase, which is crucial for the optimal performance of perovskite films. **Supplementary Figure 1** displays the perovskite aspect with contrasting yellow and black phases among the 60 compositions, highlighting the impact of precursor composition on film phase purity. The perovskite films with various compositions were subjected to photo-thermal aging in a controlled glove box environment, enduring heat conditions of 50 °C under a nitrogen atmosphere, complemented by solar irradiance (AM1.5 illumination, $100\,mW\,cm^{-2}$). Regular UV-vis spectra measurements between the 600-850 nm range provided us insights into the optical degradation of the films. The statistical data on the T_{90}^A lifetimes for all 60 compositions, along with the decay curves of six subset formulations, are presented in **Figure 1b and Supplementary Figure 2**. Notably, the employment of MACl as additive (ranging from 5 to 50%) positively influenced the optical stability of the films, with a highest point observed at a 50% addition level,[25, 26] as shown in **Supplementary Figure 3**. Among these compositions, numbers 14 ($FAPbI_3$+50% MACl), 35 ($FAPbI_3$+3% CsI+7% MAI), and 57 (($MA_{0.02}Cs_{0.05}FA_{0.93}$)$Pb(I_{0.97}Br_{0.01}Cl_{0.02})_3$ + 50% MACl) stood out for their exceptional thermal stability, with T_{90}^A lifetimes exceeding 2,000 hours.

However, photothermal stability is only one aspect of the overall robustness of the devices.

Operational stability, measured through MPP tracking until the device efficiency drops to 90% of its initial value (T_{90}^P lifetime), is equally crucial. These assessments were conducted under environmental conditions of 30 °C and 40-50% relative humidity (RH). The corresponding T_{90}^P data for different components under MPP testing and the decay of *J-V* curves over time are detailed in **Figure 1c. Supplementary Figure 4** provides data for some compositions, offering a clear visual representation of the decay of *J-V* curves over time. This figure effectively illustrates the differences in stability among the compositions and demonstrates the method used to calculate the time to 90% degradation. Specifically, components 13 (FAPbI$_3$+30% MACl), 14 (FAPbI$_3$+50% MACl), 57 (MA$_{0.02}$Cs$_{0.05}$FA$_{0.93}$)Pb(I$_{0.97}$Br$_{0.01}$Cl$_{0.02}$)$_3$ + 50% MACl), and 58 ((Cs$_{0.05}$MA$_{0.15}$FA$_{0.80}$)PbI$_3$) exhibited exceptional operational stability, making them strong candidates for further investigation. The culmination of these photo-thermal and operational stability experiments led to the selection of composition 57 for in-depth analysis. This composition not only demonstrated robust light-absorption characteristics but also stood out for its exceptional stability in solar cells, making it a prime candidate to advance the study of PSC lifetimes.
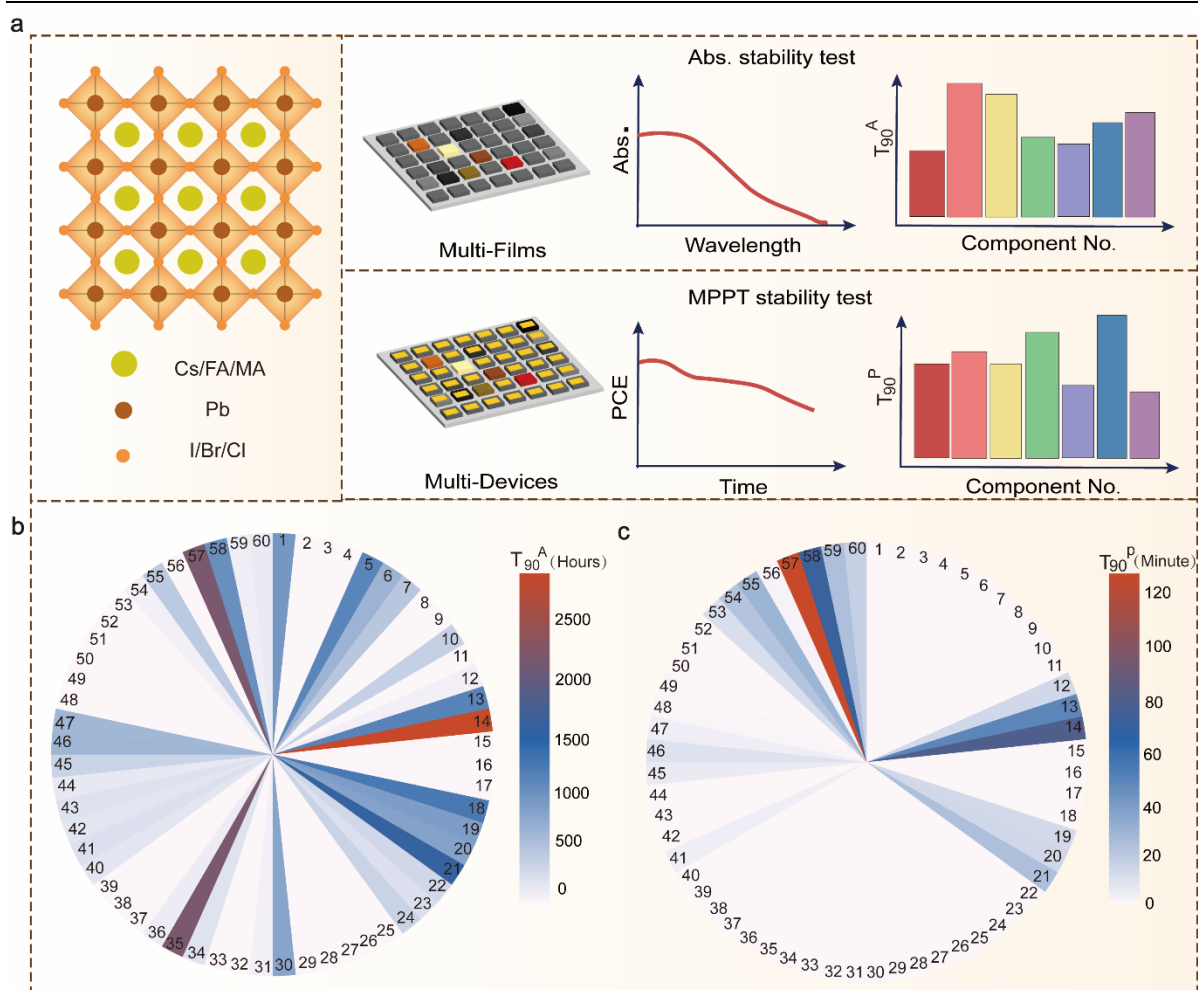
Fig. 1 Stability analysis and screening of 60 multi-component perovskites. a, Schematic diagram of the screening of the perovskite components stability. The multi-component perovskite film was subjected to photo-thermal aging, and some of the prepared devices were subjected to photo-thermal electrical aging. Characterization methods include UV-vis absorption spectroscopy (Abs.) and current-voltage curve changes with time. **b,** Colormap representation of the T_{90}^A lifetime of 60 different perovskite samples subjected to aging at 50 °C under a nitrogen atmosphere and illuminated with 100 mW cm^{-2} of metal halide light. **c,** Colormap representation of the T_{90}^P efficiency lifetime of 60 perovskite devices continuously tested under 30 °C, 40-50% RH and 100 mW cm^{-2} metal halide lamp illumination using MPP tracking.

Machine learning for film formation parameters optimization

Next, we applied Bayesian optimization (BO) within a sequential machine learning paradigm to optimize the perovskite film formation VFGT parameters to maximize the PCE of perovskite thin-film photovoltaic solar cells. These parameters are provided in the experimental

section (**Supplementary Material**). The optimized parameters resulted in a PCE under test conditions of 25.32%, a short-circuit current density (J_{SC}) of 25.69 mAcm^{-2}, a V_{OC} of 1.182 V and a fill factor (FF) of 83.3%. **Figure 2a** illustrates the methodology used. To systematically navigate the complex experimental parameter space, several rounds of experimentation were performed, with the aim of gradually learning the relationship between this space and the device efficiency, while identifying with increasing confidence and accuracy regions of the space resulting in peak PCE values.

In a preliminary step, a search space, optimizing experimental scope and feasibility, was defined. Initially, 45 process conditions were tested to validate this search space, revealing a PCE ranging from 5.6% to 21.9%, confirming the potential of the selected space to significantly modulate PCE **(Supplementary Figure 5,** the range of the considered search space is shown in **Supplementary Table 2)**. Subsequently, the BO campaign was launched. To match the capacity of the thermal evaporation process, batches of 20 samples were processed simultaneously, optimizing cost-effectiveness by balancing measurement time against model feedback frequency. The initial manual sampling was informed by the preliminary screening, while aiming for wide coverage across process variables and minimizing solvent use (**Supplementary Figure 6)**. Upon fabrication of multiple devices under the same conditions, this initial sampling showed significant PCE variability, particularly at lower efficiencies, with an average PCE spread of 3.2%. This variability – which can be attributed to measurement and sampling errors – was mitigated in our approach by taking the mean PCE across four devices per condition[27]. With this approach, an initial mean PCE of 20.1% was obtained for this batch **(Supplementary Figure 7)**.

The initial sample served as the basis for constructing a surrogate model, predictive of the objective function. For a comprehensive explanation of this process, encompassing the relevant mathematical formulas and methodologies, readers are directed to the detailed discussion in the Supplementary Information, specifically within the 'Machine Learning Methodology' section. We then employed a compounded acquisition function to identify new promising conditions for experimental validation. In this compounded acquisition function, the surrogate model was combined with probabilistic constraints, respectively corresponding to qualitative film quality

data, as well as the individual PCE measurements for the 45 preliminary devices, fabricated to validate the scope of the search space. The goal of these probabilistic constraints was to augment the acquisition function based on the surrogate model with secondary, related – though not directly transferable – information sources. With each batch, the BO algorithm's search space became more targeted, increasingly focusing on higher PCE regions **(Supplementary Figures 8 and 9)**. The second batch revealed a fine-tuning of the search range, which became narrower as more data was gathered **(Supplementary Figure 9)**. The steady progress and subsequent diminishing returns are graphically depicted in **Supplementary Figures 10 and 11**. By the third batch, the mean PCE had increased to 25.1%, with a peak at 25.3%, suggesting a plateau in the PCE improvement trajectory, and surpassing published records.[23] This advancement demonstrates the efficacy of our methodical optimization strategy in pushing the boundaries of PSC performance **(Supplementary Figure 12 and 13).** The gradual refinement of the surrogate model as more data is gathered can be visualized through contour plots, visualizing the evolution of the predicted objective function across two-dimensional cross-sections of the process variable space **(Supplementary Figures 8, 11 and 14)**. It is noteworthy that these contour plots did not show significant changes post the second batch, suggesting that convergence has indeed been reached. Finally, **Figures 2c-f** and **Supplementary Figure 14** present the last set of contour plots from the final optimization round, indicating only minor adjustments, and solidifying our decision to conclude the campaign. Based on the final set of contour plots, a refined range of the various process parameters, resulting in maximized PCE values, can be delineated **(Supplementary Table 3)**. The process condition resulting in the best PCE measured corresponded to the following parameter combination: DMF/DMSO/NMP solvent ratio at 85/15/0, perovskite precursor concentration at 1.4 M, annealing temperature set to 135°C, vacuum pressure maintained at 20 Pa, and vacuum application duration fixed at 21 seconds. Overall, our integrated approach, blending machine learning with empirical insights, achieved remarkable efficiency enhancements and presents a robust model for optimizing material properties in complex multidimensional spaces.
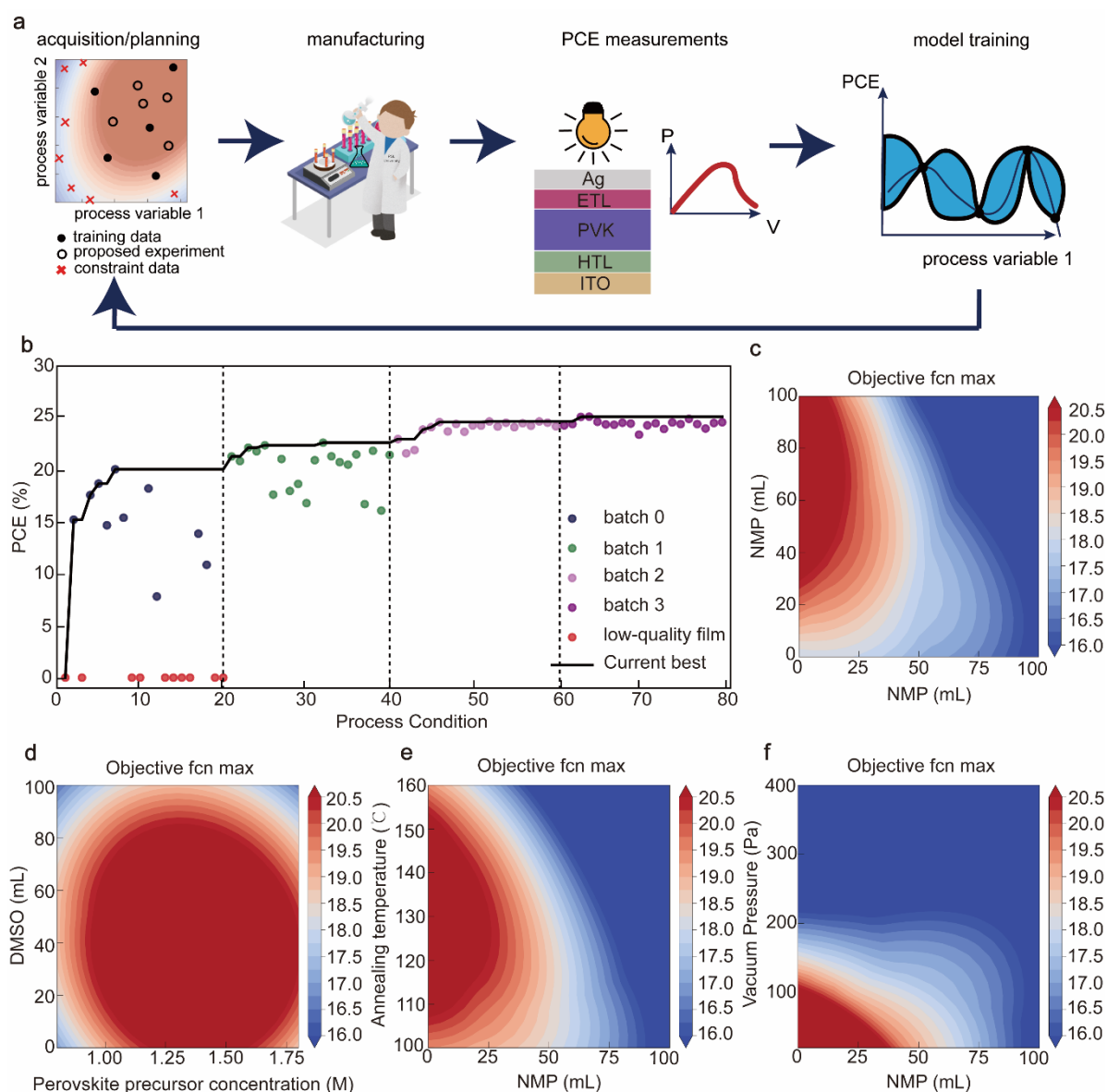
Fig. 2 Schematic of sequential learning optimization of PSCs with probabilistic constraints. a, Planning, manufacturing, measuring and model training & prediction workflow. This workflow is typically iterated until the target efficiency is achieved or the maximum experimental budget is reached. **b,** The acquisition process for the final sampling round in this study. Plot of the predicted and measured PCE values for both the initial batch and the first selection of acquired points; the best PCE value measured so far is denoted by the bold black curve. **c-f,** Sample of the 2D contour plots for the surrogate objective function: (c) DMF vs. NMP volumes, (d) DMSO vs. perovskite precursor concentration, (e) annealing temperature vs. NMP volume, (f) vacuum pressure vs. NMP volume. Maximum surrogate objective function values across a sample of process conditions at every grid point are visualized.

Layer engineering by a multifunctional polymeric compound

The ionic characteristics of perovskite crystal enable molecular defect passivation approaches through the interaction between functional groups and defect sites. Further, a functional molecule, cryloyloxyethyltrimethylammonium tetrakis(pentafluorophenyl)borate (AETA-BCF, **Figure 3a**), that was meticulously crafted to include a constellation of functional groups was introduced into perovskite films, to simultaneously enhance the efficiency and stability of inverted PSCs. Central to its design is the polymerizable acryloyl moiety (C=C), which upon thermally induced polymerization, initiates linearly polymerized, engendering a meshwork within the perovskite matrix. Fourier Transform Infrared Spectroscopy (FTIR) analysis before and after polymerization furnishes additional verification of the in-situ linearly polymerization process. Compared to the un-polymerized perovskite films, the polymerized samples exhibited significant reductions in the vibrational characteristics of C=C, C-O, C=O, and CH_2=CH groups, as evidenced by v=CH (3102 cm^{-1}), v=CH_2 (3055 cm^{-1}), and γ=CH_2 (870 cm^{-1})[28]. These changes indicate that some molecules were successfully linearly polymerized, confirming the effectiveness of our linearly polymerized strategy (**Supplementary Figure 15**). Photographs in **Supplementary Figure 16** display AETA-BCF in both powder and solution form before (a) and after (b) the polymerization process. As observed, the solutions dissolved in DMF are clear and transparent before polymerization. After polymerization, the solution becomes a turbid suspension, reflecting the formation of a polymer network and decreased solubility. Transmission electron microscopy (TEM) and nuclear magnetic resonance provides unequivocal evidence of in-situ linearly polymerized within the perovskite, delineating the amorphous and crystalline domains **(Supplementary Figures 17-19)**. The perovskite crystal lattice is discerned with pronounced clarity, with the polymer matrix enveloping the grain boundaries—affirming the successful polymerization and stabilization conferred by the novel compound. To verify the interaction between AETA-BCF and the perovskite structure and to confirm the optimized composition of the doped perovskite films, X-ray Photoelectron Spectroscopy (XPS) analysis was conducted. The observed shifts in the binding energies of the

Pb 4f peak effectively demonstrate the interaction between AETA-BCF molecules and the perovskite. Additionally, the presence of C=O, C-N, C-O, and C-C peaks in the C 1s spectrum confirms the successful incorporation of AETA-BCF into the perovskite matrix (Supplementary Figure 20). Our examination, illustrated in **Figure 3a**, reveals the transformative impact of AETA-BCF on the film formation and crystallization processes in perovskite structures. The inherent uncontrolled conditions in perovskite synthesis often yield a significant quantity of uncoordinated Pb^{2+} ions and atomic defects involving lead and iodine. However, the incorporation of AETA-BCF, with its unique C=O functional group acting as a Lewis base, facilitates the formation of coordination bonds with these uncoordinated Pb^{2+} ions, effectively stabilizing them. Moreover, the presence of amine and fluorine atoms in AETA-BCF not only allows for hydrogen bond interactions with the perovskite structure but also imparts enhanced hydrophobic properties. This multifaceted interaction between AETA-BCF and the perovskite matrix plays a crucial role in defect passivation, promoting superior crystallization and hydrophobicity. Consequently, this leads to the formation of high-quality perovskite films characterized by lower defect densities and improved stability.

Our DFT calculations, utilizing $FAPbI_3$ as a simplified model, delved into the intricate interactions between the chosen molecule and the perovskite structure. We investigated three prevalent defect models: lead vacancy (V_{Pb}), iodine vacancy (V_I), and $Pb_{I\text{-antisite}}$, with the initial defect models presented in **Figures 3b-d**. The differential charge density, portrayed in **Figures 3e-f**, elucidates the charge transfer dynamics between the molecule and the uncoordinated Pb^{2+}, I^-, and FA^+ ions. This interaction fosters electron cloud delocalization and the formation of coordination bonds, theoretically validating our approach to mitigate defect formation. Further insights are gleaned from **Supplementary Figures 21-24**, where we showcase the optimized defect state density (DOS) and electronic structure in comparison to the original models. We observed that surface defects on the perovskite, particularly those involving lead and iodine vacancies and antisite defects, introduce new hybridized states. These emergent trap states subtly modulate the conduction and valence band edges, potentially leading to charge capture. Crucially, the charge transfer facilitated by the passivation molecule to the perovskite surface

defects effectively neutralizes gap states. This phenomenon is pivotal in creating perovskite films with a remarkably low defect density, a cornerstone for enhancing the overall performance and stability of perovskite-based devices. The Kelvin Probe Force Microscopy (KPFM) analysis demonstrated a more uniform surface potential across the perovskite films treated with AETA-BCF, compared to the untreated samples (**Supplementary Figure 25**). This uniformity indicates a reduction in local charge traps and defects, corroborating the decreased density of defect states observed in our DFT studies. Our scanning electron microscope (SEM) analysis revealed that perovskite films, both with and without molecular passivation, displayed pronounced morphological changes. Optimized films, as shown in **Supplementary Figure 26**, exhibited larger grains and improved cross-sectional morphology. Complementing this, time-resolved photoluminescence (TRPL) data (**Supplementary Figure 27**) indicated a reduced lifetime in the passivated perovskite, suggesting lower recombination rates and more efficient charge dynamics. To further elucidate the specific role of molecular optimization in suppressing electron-hole recombination, additional Time-Resolved Photoluminescence (TRPL) tests were conducted on a Glass/perovskite structure. The TRPL data (**Supplementary Figures 28**) indicate that the passivated perovskite films exhibit significantly enhanced lifetimes, suggesting reduced recombination rates and more efficient charge dynamics.

This polymer improves the mechanical elasticity of the perovskite layer, mitigating the disintegration of the perovskite structure as evidenced by post-heating structural analysis (**Figure 3a**). The molecule's quaternary ammonium segment ($-N^+(CH_3)_3$) effectively passivates intrinsic defect states that compromise the perovskite's integrity. Concurrently, the ester functionality (C=O) is anticipated to engage in robust coordination interactions, forming robust Pb-O bonds that enhance molecular adhesion to the perovskite surface.[29] The hydrophobic character of the molecule is predominantly imparted by the pentafluorophenyl moieties appended to the borate nucleus.[30] These fluorinated aromatic groups are renowned for their hydrophobicity, proffering moisture resistance that is vital to the longevity and operational stability of perovskite-based devices.[31] The deliberate integration of these groups aims at conferring a hydrophobic sheath over the perovskite surface, thereby curbing moisture ingress,

a well-known cause of perovskite degradation.[32] Furthermore, the voluminous nature of the pentafluorophenyl groups introduces spatial bulk around the borate core, averting tight packing of molecules. This strategic steric arrangement is projected to diminish non-radiative recombination by thwarting the agglomeration of molecules at the grain boundaries, potentially prolonging charge carrier lifetimes and thereby amplifying device efficiency.
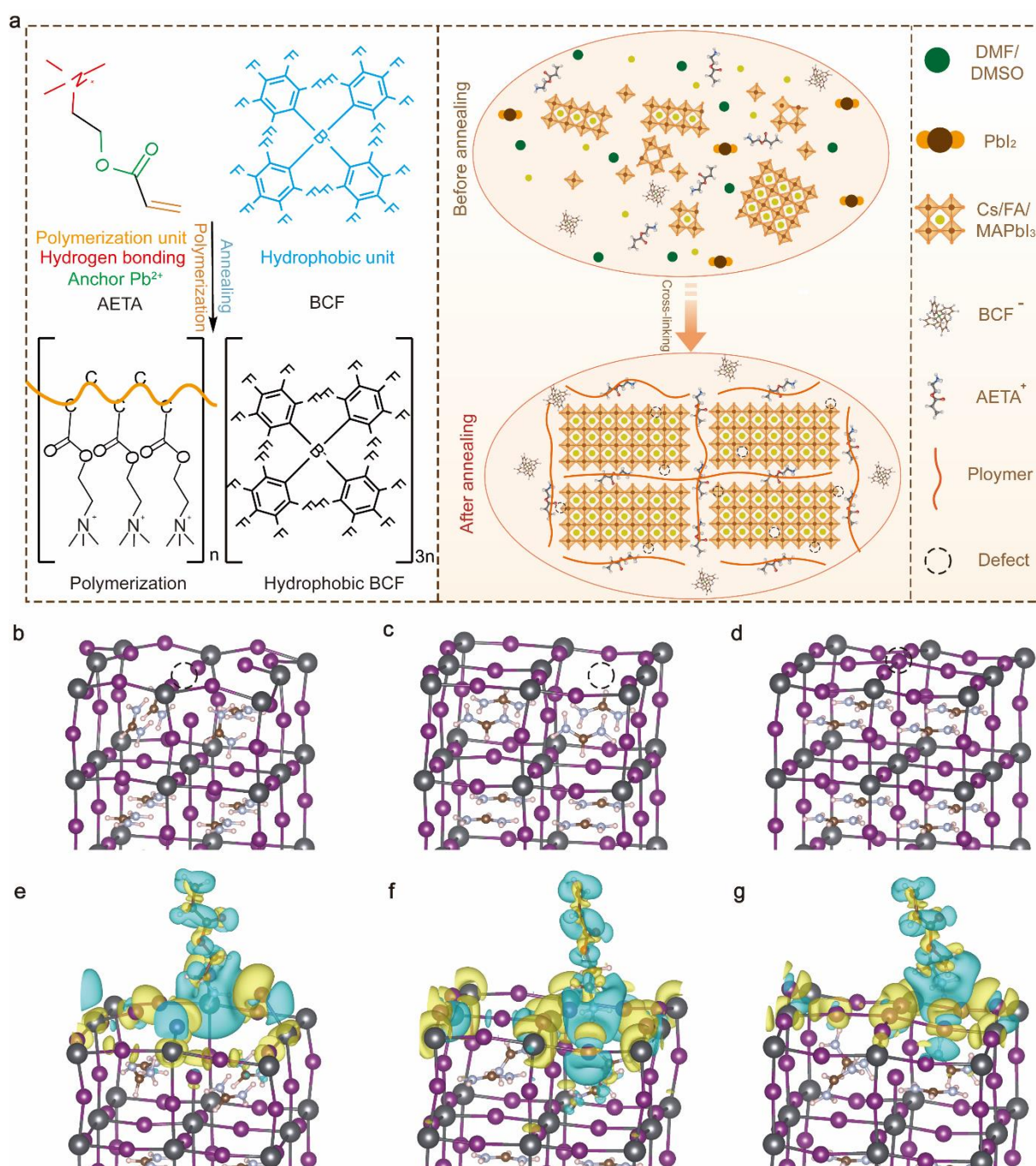
Fig. 3 Molecular engineering of the perovskite. a, Comparative analysis of chemical structures and defect passivation mechanisms. Left: depicts the chemical structure of AETA-BCF and the process of in situ polymerization under heating conditions, illustrating the molecular arrangement and polymerization dynamics. Right: schematic representation of the distribution of defects in perovskite materials, alongside the mechanism of defect passivation by organic polymer molecules, highlighting the interaction between the perovskite crystal and the organic passivators. DFT modeling of AETA-perovskite interactions and common defects impact. Panels (b-g) illustrate the influence of AETA interactions with perovskite on the formation of common defects: Pb vacancy (V_{Pb}), I vacancy (V_I), and PbI antisite ($Pb_{I\text{-antisite}}$) defects, and black circle represents the location of the defect. Panels (b), (c), and (d) present side views of the original atomic structures of unmodified perovskite with defects: V_{Pb}, V_I, and $Pb_{I\text{-antisite}}$, respectively. Panels (e), (f), and (g) show side views of the charge density differences and atomic structures of the corresponding three defects after AETA optimization. The brown, blue, pink, black, and purple spheres represent C, N, H, Pb, and I atoms, respectively. Blue regions indicate decreased charge density, whereas yellow regions denote increased charge density.

Solar cells characteristics and stability

To investigate the performance of our optimized perovskite layers, we fabricated devices with the ITO/NiOx/MeO-2PACz/perovskite/PCBM/BCP/Ag structure. We optimized the concentration of the AETA-BCF additive, finding that 1.5 mg/mL yielded devices with optimal performance, as shown in **Supplementary Figure 29**. The control devices, optimized using machine learning, exhibited a reverse-scan PCE of 25.32%, with a J_{SC} of 25.69 mAcm^{-2}, an open-circuit voltage (V_{OC}) of 1.182 V, and a fill factor (FF) of 83.3%. Forward-scan measurements showed a PCE of 24.93%, J_{sc} of 25.70 mAcm^{-2}, a V_{OC} of 1.181 V, and a FF of 82.1%, indicating a significant hysteresis effect, as displayed in **Supplementary Figure 30**. In contrast, the AETA-BCF-modified devices demonstrated superior performance, with the best target device exhibiting a reverse-scan PCE of 25.76%, a J_{SC} of 25.71 mAcm^{-2}, a V_{OC} of 1.192 V, a FF of 84.1%, and a forward-scan PCE of 25.62%, of 25.70 mAcm^{-2}, a V_{OC} of 1.191V, a FF of 83.7%, showing negligible hysteresis (**Figure 4a**). The stabilized PCE was 25.60%, as shown in **Supplementary Figure 31. Figure 4b** displays the Gaussian fitted statistical distribution of PCE for both control and target devices, with an average PCE of 25.05% for

control devices and 25.60% for target devices, further confirming the effectiveness of AETA-BCF. Our best device, optimized based on AETA-BCF, was sent to the China National Photovoltaic Product Quality Inspection & Testing Center for certification, confirming its performance with a PCE of 25.21% (V_{OC} = 1.191V, FF = 82.32%, and J_{SC} = 25.71 mA cm^{-2}), as shown in **Supplementary Figure 32**. The external quantum efficiency (EQE) spectra (**Supplementary Figure 33**) resulted in an integrated J_{SC} of 25.43 mAcm^{-2}. These values are marginally lower than those derived from *J-V* measurements, yet they exhibit good alignment with the current values recorded during JV testing. The addition of AETA-BCF showed a significant improvement in PCE, correlating with enhanced V_{OC} and FF, as well as J_{SC}, as indicated by the transient photocurrent (TPC) results (**Supplementary Figure 34**). The target devices exhibited faster decay compared to control devices, indicating more efficient charge extraction, which accounts for the increased V_{OC} and FF. The increase in J_{SC} is attributed to suppressed recombinations and effective charge extraction.

The operational stability of the high-efficiency inverted PSCs under aging conditions was evaluated using the International Summit on Organic Photovoltaic Stability ISOS-L-2 protocol. To enhance device stability, we used copper as the electrode instead of the original silver electrode for aging tests. Therefore, both control and target devices experienced a three percent drop in efficiency. The long-term operational stability of the encapsulated control and target PSCs was assessed by MPP tracking under constant simulated AM1.5G illumination (100 mW cm^{-2}). The AETA-BCF optimized devices demonstrated robustness in two distinct aspects of stability. Firstly, they exhibited excellent photo-thermal and -operational stability, which refers to their ability to retain performance under prolonged exposure to high temperatures, as evidenced by maintaining about 94% of the initial efficiency (22.2% to 20.87%) after 1,920 hours at 65 °C in N$_2$ atmosphere. This 94% stability, assessed through *J-V* curve testing every 20 hours, encapsulates both photo-thermal endurance and operational resilience, reflecting the device's robust performance in real-world conditions. Secondly, their operational stability, which pertains to the durability of the device under normal working conditions, was also remarkable. This was showcased in **Figure 4c**, where the devices were protected under an N$_2$

atmosphere, simulating the stability of the devices during regular operation. In comparison, the control devices, after rigorous component screening, also showed relatively good thermal and operational stability, maintaining about 85% of the initial efficiency (22.1% to 18.7%) after 1,920 hours of aging at 65 °C in N_2 atmosphere, as shown in **Figure 4d**. The stability of the best devices is illustrated in **Figure 4e**. The stability in the J-V curves is consistent with the stable output before and after aging at the MPP. This characteristic of operational stability and thermal stability throughout the aging process represents a significant improvement over the current state-of-the-art technologies. In contrast to other organic material-based devices where FF significantly decreases, the FF in devices prepared by this process remained almost constant during aging. As shown in **Figure 4f**, the control device's FF and PCE output exhibited severe fluctuations compared to the target devices. Water contact angle measurements were utilized to assess the moisture stability of perovskite films after adding AETA-BCF and polymerization. The results demonstrated a significant enhancement in hydrophobicity: the contact angle of the perovskite films increased from 43.45° to 60.77° after polymerization (**Supplementary Figure 35**). To further investigate the air stability of our devices, PSCs were exposed to ambient air and their performance was assessed. The results demonstrated that, compared to control PSCs without additive, the target PSCs retained 97% of their initial efficiency after 420 hours of exposure, whereas the efficiency of control PSCs declined to below 93%. This significant outcome highlights the effectiveness of our strategy to enhance the air stability of PSCs (Supplementary Figure 36). These data indicate that the combination of component screening, machine learning, and molecular engineering represents a solid step towards the optimizing of the solar cell performances while reaching high stability.
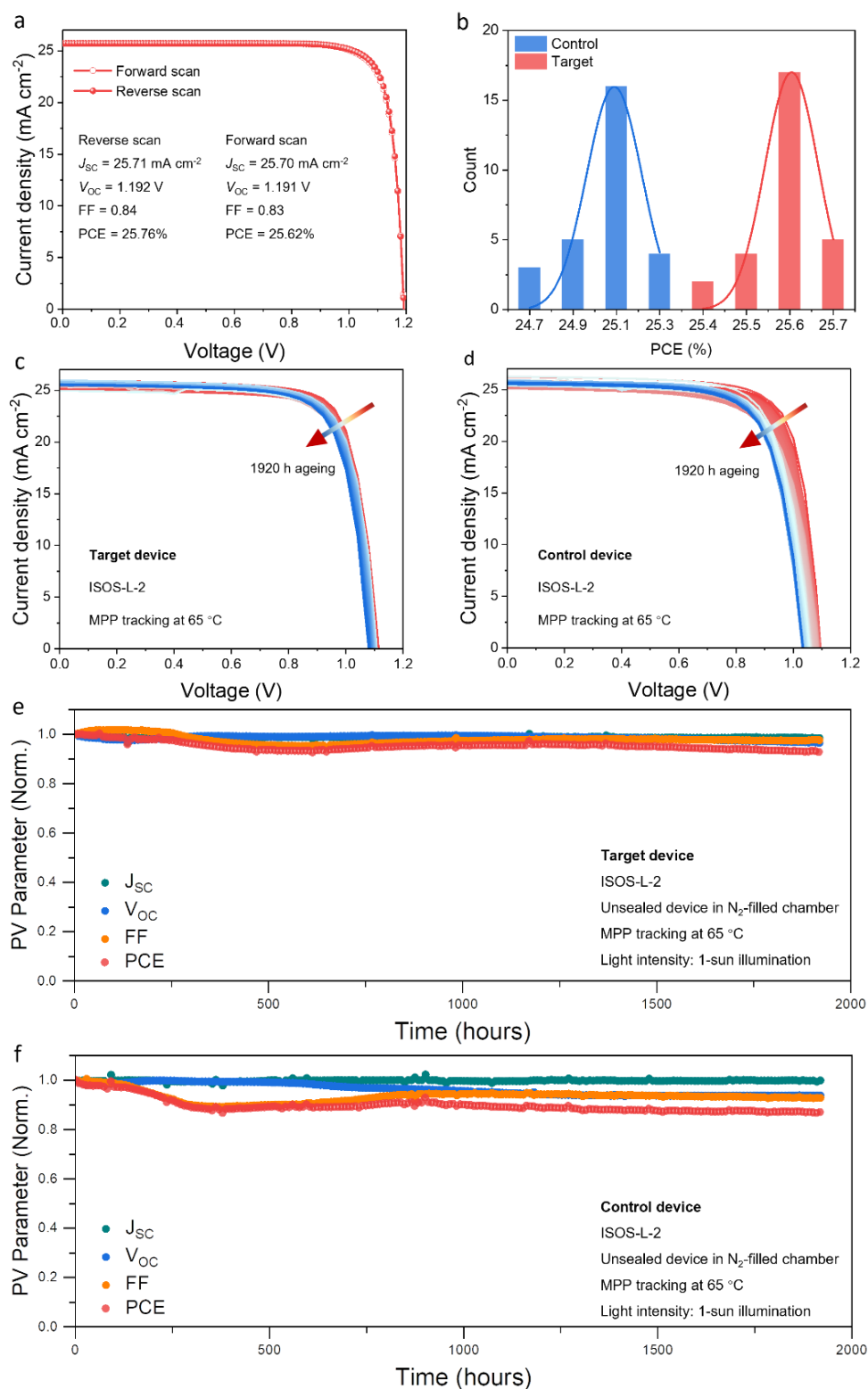
Fig. 4 Photovoltaic performances of PSCs. a, J–V curves of the best-performing target device. b, Histograms showing the device efficiencies of 28 cells per type, fitted with Gaussian distributions (solid lines). c-d, MPP tracking J–V curves for control and target PSCs aged at 65 °C with continuous light illumination (100 mW cm^{-2}) for 1,920 hours in N_2. e-f. The long-term stability for PSCs tracking of

control and target devices, data points originated from *J-V* scan.

Conclusions

In summary, we have successfully unveiled a paradigm shift to develop efficient and stable inverted PSCs. Our approach combined component screening, machine learning and molecular engineering. Through a judicious selection from a repertoire of 60 perovskite variants, we have identified a composition with exemplary optical, thermal and electrical stability. Employing Bayesian machine learning, we have navigated a labyrinth of over 1 billion process conditions, culminating in a record-breaking efficiency within a mere 80 iterations. Finally, the integration of bespoke in-situ linearly polymerized ionic molecules allowed us to further augment device performance, reaching a power conversion efficiency of 25.76% (certified at 25.21%). Additionally, through the synergistic effect of selecting stable components and incorporating hydrophobic molecules, the devices retained 94% of their initial efficiency after operating for 1,920 hours under continuous 1-sun illumination and MPP tracking at 65 °C in N_2 atmosphere. Finally, this work describes an innovative strategy that can be successfully implemented to considerably accelerate the development and optimization of photovoltaic and, beyond, other optoelectronic devices.

Acknowledgements

Author contributions

B.Z and T.P conceived the idea and designed the project. B.Z, and J.L fabricated the devices. H.Z. and H.Y performed MPPT characterization. Z.W., C.J. and J.L. performed TRPL, SEM

and FTIR characterization. T.S. and B.Z performed the ML work. B.Z., T.P and T.S wrote the manuscript. D.Z., Z.W., C.J. and J.L. improved the manuscript. All authors discussed the results and commented on the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information. The online version contains supplementary material available at XX.

Data availability

The main data supporting the findings of this study are available within the Article and its Supplementary Information.

Code availability

All code and data used during the Bayesian optimization campaign is available in a GitHub repository at https://github.com/Boxue2023/Perovskite.

References

1. H. Wei, D. DeSantis, W. Wei, Y. Deng, D. Guo, T. J. Savenije, L. Cao and J. Huang, *Nat Mater*, 2017, **16,** 826-833.

2. J. Burschka, N. Pellet, S.-J. Moon, R. Humphry-Baker, P. Gao, M. K. Nazeeruddin and M. Graetzel, *Nature*, 2013, **499**, 316-319.

3. N. J. Jeon, J. H. Noh, W. S. Yang, Y. C. Kim, S. Ryu, J. Seo and S. I. Seok, *Nature*, 2015, **517**, 476-480.

4. H. Tsai, W. Nie, J.-C. Blancon, C. C. Stoumpos, R. Asadpour, B. Harutyunyan, A. J. Neukirch, R. Verduzco, J. J. Crochet, S. Tretiak, L. Pedesseau, J. Even, M. A. Alam, G. Gupta, J. Lou, P. M. Ajayan, M. J. Bedzyk, M. G. Kanatzidis and A. D. Mohite, *Nature*, 2016, **536**, 312-316.

5. N. Arora, M. I. Dar, A. Hinderhofer, N. Pellet, F. Schreiber, S. M. Zakeeruddin and M. Grätzel, *358, 762-771*.

6. W. S. Yang, B.-W. Park, E. H. Jung, N. J. Jeon, Y. C. Kim, D. U. Lee, S. S. Shin, J. Seo, E. K. Kim, J. H. Noh and S. I. Seok, *Science*, 2017, **356**, 1376-1379.

7.   H. Li, C. Zhang, C. Gong, D. Zhang, H. Zhang, Q. Zhuang, X. Yu, S. Gong, X. Chen, J. Yang, X. Li, R. Li, J. Li, J. Zhou, H. Yang, Q. Lin, J. Chu, M. Grätzel, J. Chen and Z. Zang, *Nat. Energy*, **8, 946-955** 2023.

8.   Q. Tan, Z. Li, G. Luo, X. Zhang, B. Che, G. Chen, H. Gao, D. He, G. Ma, J. Wang, J. Xiu, H. Yi, T. Chen and Z. He, *Nature*, 2023, **620**, 545-551.

9.   S. Zhang, F. Ye, X. Wang, R. Chen, H. Zhang, L. Zhan, X. Jiang, Y. Li, X. Ji, S. Liu, M. Yu, F. Yu, Y. Zhang, R. Wu, Z. Liu, Z. Ning, D. Neher, L. Han, Y. Lin, H. Tian, W. Chen, M. Stolterfoht, L. Zhang, W. H. Zhu and Y. Wu, *Science*, 2023, **380**, 404-409.

10. M. Z. Liu, M. B. Johnston and H. J. Snaith, *Nature*, 2013, **501,** 395-398.

11. E. Bi, H. Chen, F. Xie, Y. Wu, W. Chen, Y. Su, A. Islam, M. Grätzel, X. Yang and L. Han, *Nature Communications*, 2017, **8**, 15330.

12. J. L. Minns, P. Zajdel, D. Chernyshov, W. van Beek and M. A. Green, *Nature Communications*, 2017, **8**, 15152.

13. D. Zheng, T. Zhu and T. Pauporté, *Solar RRL*, 2021, **5**, 2100010.

14. Y. Zhao, I. Yavuz, M. Wang, M. H. Weber, M. Xu, J. H. Lee, S. Tan, T. Huang, D. Meng, R. Wang, J. Xue, S. J. Lee, S. H. Bae, A. Zhang, S. G. Choi, Y. Yin, J. Liu, T. H. Han, Y. Shi, H. Ma, W. Yang, Q. Xing, Y. Zhou, P. Shi, S. Wang, E. Zhang, J. Bian, X. Pan, N. G. Park, J. W. Lee and Y. Yang, *Nat Mater*, 2022, **21**, 1396-1402.

15. J. Luo, B. Liu, H. Yin, X. Zhou, M. Wu, H. Shi, J. Zhang, J. Elia, K. Zhang, J. Wu, Z. Xie, C. Liu, J. Yuan, Z. Wan, T. Heumueller, L. Luer, E. Spiecker, N. Li, C. Jia, C. J. Brabec and Y. Zhao, *Nat Commun*, 2024, **15**, 2002.

16. D. Zheng, F. Raffin, P. Volovitch and T. Pauporte, *Nat Commun*, 2022, **13**, 6655.

17. N.-G. Park and K. Zhu, *Nature Reviews Materials*, 2020, **5**, 333-350.

18. M. Ahmadi, M. Ziatdinov, Y. Zhou, E. A. Lass and S. V. Kalinin, *Joule*, 2021, **5**, 2797-2822.

19. Q. Tao, P. Xu, M. Li and W. Lu, *npj Computational Materials*, 2021, **7**, 23.

20. S. Sun, N. T. P. Hartono, Z. D. Ren, F. Oviedo, A. M. Buscemi, M. Layurova, D. X. Chen, T. Ogunfunmi, J. Thapa, S. Ramasamy, C. Settens, B. L. DeCost, A. G. Kusne, Z. Liu, S. I. P. Tian, I. M. Peters, J.-P. Correa-Baena and T. Buonassisi, *Joule*, 2019, **3**, 1437-1451.

21. L. Luo, H. Zeng, Z. Wang, M. Li, S. You, B. Chen, A. Maxwell, Q. An, L. Cui, D. Luo, J. Hu, S. Li, X. Cai, W. Li, L. Li, R. Guo, R. Huang, W. Liang, Z.-H. Lu, L. Mai, Y. Rong, E. H. Sargent and X. Li, *Nat. Energy*, 2023, **8**, 294-303.

22. J. Zhang, X. Niu, C. Peng, H. Jiang, L. Yu, H. Zhou and Z. Zhou, *Angew Chem Int Ed Engl*, 2023, e202314106.

23. D. Gao, R. Li, X. Chen, C. Chen, C. Wang, B. Zhang, M. Li, X. Shang, X. Yu, S. Gong, T. Pauporte, H. Yang, L. Ding, J. Tang and J. Chen, *Adv Mater*, 2023, **35**, e2301028.

24. X. Li, D. Bi, C. Yi, J.-D. Décoppet, J. Luo, S. M. Zakeeruddin, A. Hagfeldt and M. Grätzel, *Science*, 2016, **353**, 58-62.

25. D. Zheng, F. Chen, M. N. Rager, L. Gollino, B. Zhang and T. Pauporté, *Advanced Materials Interfaces*, 2022, **9**, 2201436.

26. T. Zhu, D. Zheng, M.-N. Rager and T. Pauporté, *Solar RRL*, 2020, **4**, 2000348.

27. J. K. T. Byron Kratochvil, *Analytical Chemistry*, 1981, **53**, 924A-938A.

28. J. Zhang, X. Niu, C. Peng, H. Jiang, L. Yu, H. Zhou and Z. Zhou, *Angew Chem Int Ed Engl*, 2023, **62**, e202314106.

29.   T. W. Yanbo Wang, Julien Barbaud, Weiyu Kong, Danyu Cui, Han Chen, Xudong Yang, Liyuan Han, *Science*, 2019, **365**, 687-691.

30.   S. S. Reddy, V. M. Arivunithi, V. G. Sree, H. Kwon, J. Park, Y.-C. Kang, H. Zhu, Y.-Y. Noh and S.-H. Jin, *Nano Energy*, 2019, **58**, 284-292.

31.   J. Wang, M. A. Uddin, B. Chen, X. Ying, Z. Ni, Y. Zhou, M. Li, M. Wang, Z. Yu and J. Huang, *Advanced Energy Materials*, 2023, **13**.

32.   H. Wang, Z. Zhang, J. V. Milić, L. Tan, Z. Wang, R. Chen, X. Jing, C. Yi, Y. Ding, Y. Li, Y. Zhao, X. Zhang, A. Hagfeldt, M. Grätzel and J. Luo, *Advanced Energy Materials*, 2021, **11**, 2101082.