

# <sup>1</sup> Dimensional Control over Metal Halide Perovskite Crystallization Guided by Active Learning

<sup>3</sup> Zhi Li, Philip W. Nega, Mansoor Ani Najeeb Nellikkal, Chaochao Dun, Matthias Zeller, Jeffrey J. Urban,  
<sup>4</sup> Wissam A. Saidi, Joshua Schrier, Alexander J. Norquist, and Emory M. Chan\*



Cite This: <https://doi.org/10.1021/acs.chemmater.1c03564>



Read Online

ACCESS |

Metrics & More

Article Recommendations

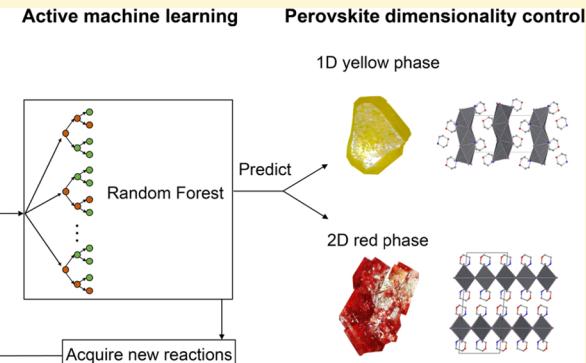
Supporting Information

**ABSTRACT:** Metal halide perovskite (MHP) derivatives, a promising class of optoelectronic materials, have been synthesized with a range of dimensionalities that govern their optoelectronic properties and determine their applications. We demonstrate a data-driven approach combining active learning and high-throughput experimentation to discover, control, and understand the formation of phases with different dimensionalities in the morpholinium (morph) lead iodide system. Using a robot-assisted workflow, we synthesized and characterized two novel MHP derivatives that have distinct optical properties: a one-dimensional (1D) morphPbI<sub>3</sub> phase ( $[C_4H_{10}NO]_1[PbI_3]$ ) and a two-dimensional (2D) (morph)<sub>2</sub>PbI<sub>4</sub> phase ( $[C_4H_{10}NO]_2[PbI_4]$ ). To efficiently acquire the data needed to construct a machine learning (ML) model of the reaction conditions where the 1D and 2D phases are formed, data acquisition was guided by a diverse-mini-batch-sampling active learning algorithm, using prediction confidence as a stopping criterion. Querying the ML model uncovered the reaction parameters that have the most significant effects on dimensionality control. Based on these insights, we discuss possible reaction schemes that may selectively promote the formation of morph-Pb-I phases with different dimensionalities. The data-driven approach presented here, including the use of additives to manipulate dimensionality, will be valuable for controlling the crystallization of a range of materials over large reaction-composition spaces.

## 1. INTRODUCTION

Metal halide perovskite (MHP) derivatives are emerging optoelectronic materials with tunable physical properties<sup>1–4</sup> and applications in photovoltaic<sup>5,6</sup> and ferroelectric devices,<sup>7</sup> light-emitting diodes,<sup>8,9</sup> and lasers.<sup>9,10</sup> MHP derivatives have been synthesized with diverse crystal structures, with their metal halide frameworks exhibiting connectivities across one,<sup>11</sup> two,<sup>12</sup> or three dimensions.<sup>13</sup> MHP dimensionality (1D, 2D, and 3D) is a critical material parameter because it governs physical properties such as optical absorption, luminescence wavelength, charge transport, and stability.<sup>14–16</sup>

One strategy for controlling the dimensionality of MHP derivatives is tuning the inorganic or organoammonium A-cations that lie between metal halide  $[BX_6]^{4-}$  octahedra, for example, in the canonical ABX<sub>3</sub> perovskite unit cell. When A-cations are small, with effective radii of 1.7–2.6 Å (e.g., methylammonium),<sup>17</sup> three-dimensional (3D) MHP derivatives form. In contrast, larger cations (>2.6 Å) often give rise to lower-dimensional MHP derivatives.<sup>18,19</sup> When the cross-sectional area of the cations is larger than 40 Å<sup>2</sup>, MHP derivatives almost exclusively exhibit zero-dimensional (0D) or one-dimensional (1D) connectivities rather than two-dimensional (2D), owing to steric hindrance.<sup>20</sup> For example,

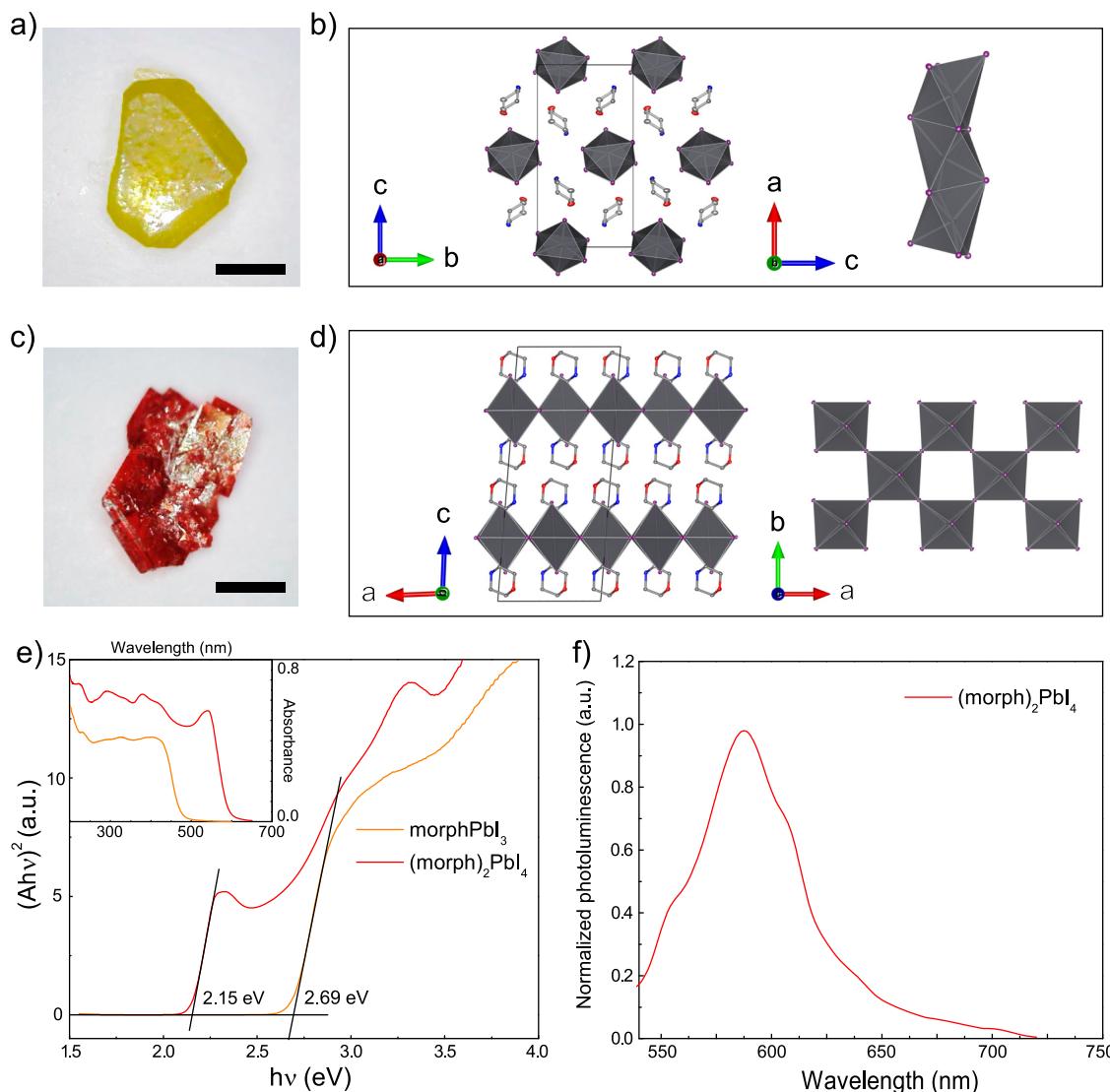


heterocyclic organoammonium cations with six or fewer ring members are able to form 2D MHP derivatives with lead iodide, while seven-member and larger rings rarely form in 2D.<sup>20–23</sup> Despite this limited ability to predict the dimensionalities of heterocyclic organoammonium lead halides, controlling dimensionalities for MHP derivatives incorporating a wide range of cations is still a major challenge.

One factor that complicates the prediction of MHP dimensionality is that different phases in an A–B–X reaction system (e.g., the 3D FAPbBr<sub>3</sub> and 2D FA<sub>2</sub>PbBr<sub>4</sub>, where FA = formamidinium) may form under different reaction conditions (FA:Pb ratio), given a reaction of A-cations, metal (B<sup>2+</sup>) cations, and halide anions (X<sup>-</sup>). Controlling dimensionality is a practical challenge for device applications—mixed phases<sup>24</sup> in MHP thin films can lead to low device performance.<sup>25</sup> Developing design rules for synthesizing MHPs with control-

Received: October 14, 2021

Revised: December 27, 2021



**Figure 1.** (a) Optical micrograph of a large crystal of the 1D “yellow” phase (scale bar: 1 mm). (b) Packing of  $[PbI_{6/2}]^-$  chain structures in  $\text{morphPbI}_3$  (1D yellow phase). (c) Optical micrograph of a large crystal of the 2D “red” phase (scale bar: 1 mm). (d) Packing of  $[PbI_{2/1}I_{4/2}]^{2-}$  layer structures in  $(\text{morph})_2\text{PbI}_4$  (2D red phase). For (b) and (d), dark gray polyhedra represent  $[PbI_6]$  octahedra, while purple, light gray, red, and blue atoms correspond to iodine, carbon, oxygen, and nitrogen. Hydrogen atoms have been removed for clarity. (e) Diffuse reflectance spectra of ground powders of  $\text{morphPbI}_3$  and  $(\text{morph})_2\text{PbI}_4$  in Tauc plots and absorbance units (in the inset). (f) PL spectra of  $(\text{morph})_2\text{PbI}_4$  ( $\lambda_{\text{ex}} = 470 \text{ nm}$ ).

lable dimensionalities will drive increases in the performance of real-world MHP devices. Furthermore, understanding the kinetic and thermodynamic factors that influence dimensionality in a single cation-metal-halide (CMH) system will allow us to design functional materials and their reaction pathways. Nevertheless, it is still unclear how to control the dimensionalities of MHP derivatives by tuning the reaction parameters in a single CMH system. Multiple crystalline phases in a single CMH system are typically discovered through trial and error. This time- and resource-intensive approach is particularly inefficient for crystallizing new MHP phases because it can require simultaneous optimization of a large number of experimental variables in high-dimensional experimental parameter space. Insufficient sampling of reaction spaces risks missing rare MHP phases and precludes a comprehensive understanding of the formation of different phases.

Machine learning (ML) and high-throughput experimentation (HTE) have been recently leveraged to accelerate material

discovery and design,<sup>26,27</sup> inspiring our efforts to apply these tools to the crystallization of MHP derivatives.<sup>28</sup> However, even with HTE, it can be impractical to perform a large number of experiments needed to train common ML models. To overcome this challenge, researchers have utilized active learning (AL),<sup>29,30</sup> a sequential learning method in which an ML model is iteratively refined over repeated cycles of experimentation through an algorithm that selects new experiments based on the performance of the most recent model. An efficient AL sampling algorithm increases the learning speed of ML models and reduces the number of experiments needed. This efficiency is beneficial for constructing material phase diagrams,<sup>31–33</sup> AL has been used to accelerate the acquisition of phase and composition diagrams of multicomponent materials, including ferroelectric ceramics,<sup>34</sup> piezoelectric materials,<sup>35</sup> phase-change materials,<sup>36</sup> catalysts,<sup>37</sup> and MHP thin films.<sup>38</sup> In these workflows, samples have been typically labeled using data acquired from

99 simulations,<sup>31–33</sup> existing datasets,<sup>37</sup> or high-throughput  
100 characterization of existing material libraries.<sup>36–38</sup>

101 Controlling the dimensionality of MHP derivatives requires  
102 the analogous task of mapping the reaction conditions that  
103 produce specific phases in different regions of synthetic  
104 composition space. We hypothesized that an approach  
105 combining AL and HTE would be advantageous for building  
106 ML models that, given reaction conditions as inputs, predict  
107 the phase and dimensionality of MHP products. Solution-  
108 phase crystallization of materials presents a stringent test for  
109 AL because each reaction tends to be more costly than a  
110 simulation or measurement and can involve large numbers of  
111 reaction parameters, reagents, and additives. These syntheses  
112 are often governed by complex reaction networks<sup>39</sup> and  
113 stochastic processes (e.g., nucleation) that can lead to noisier  
114 and less predictable outcomes than observed with simulation  
115 and characterization. These problems are particularly acute  
116 when mapping phases across more than three dimensions.  
117 When exploring such high-dimensional space, determining  
118 when to stop AL experimentation is challenging<sup>40</sup> (and an  
119 ongoing subject of research<sup>41</sup>) because one cannot determine  
120 the actual accuracy of a model without synthesizing a large test  
121 set that is representative of the entire experimental space. To  
122 realize the potential of AL-guided material synthesis, there is a  
123 strong need for robust AL workflows with clear stopping  
124 criteria and tolerance for noisy, high-dimensional data.

125 In this work, we used AL + HTE to discover, control, and  
126 understand the formation of MHP derivatives with different  
127 dimensionalities in the morpholinium lead iodide (morph-Pb-  
128 I) system. We focused on this reaction system because morph<sup>+</sup>  
129 is a six-membered heterocyclic organoammonium that should  
130 theoretically form a 2D MHP derivative with PbI<sub>2</sub>.<sup>20–22</sup>  
131 However, only 1D structures have been observed when morph  
132 is combined with different metal cations (e.g., Pb<sup>2+</sup>, Sb<sup>4+</sup>) and  
133 halides.<sup>42–49</sup> Using HTE, we successfully synthesized the 2D  
134 MHP derivative in the morph-Pb-I system. We adopted and  
135 modified an AL method to train ML classification models to  
136 predict the dimensionalities of phases formed in this chemical  
137 system. We established and validated a stopping criterion,  
138 based on the model prediction confidence, for terminating  
139 experimentation when the ML model was not improving  
140 significantly. We used a predictive ML model to uncover the  
141 reaction parameters that have the most significant effects on  
142 dimensionality control. These insights, combined with density  
143 functional theory calculations, allowed us to formulate a  
144 plausible reaction scheme that rationalized the formation of  
145 MHP derivatives with different dimensionalities in the morph-  
146 Pb-I system.

## 2. RESULTS AND DISCUSSION

147 **2.1. Benchtop Synthesis of Morpholinium Lead**  
148 **Iodide.** A range of synthetic methods is used to grow MHP  
149 derivative crystals, such as seeded crystal growth,<sup>50</sup> slow  
150 evaporation,<sup>51</sup> and inverse temperature crystallization.<sup>52</sup> In this  
151 work, to synthesize morpholinium lead iodide (morph-Pb-I)  
152 structures across a range of dimensionalities, we used  
153 antisolvent vapor-assisted crystallization (ASVC), a straightfor-  
154 ward, room-temperature approach known to produce high-  
155 quality MHP crystals suitable for structure determination using  
156 single-crystal X-ray diffraction (sXRD).<sup>53</sup> In initial experi-  
157 ments, we manually performed ASVC reactions by exposing a  
158 solution of morpholinium iodide and lead iodide (1:1  
159 morph:Pb mole ratio) in dimethylformamide (DMF) to

160 saturated vapor of dichloromethane (DCM), the antisolvent  
161 (the scheme illustrated in Figure S1). These benchtop  
162 syntheses yielded yellow crystals (Figure 1a). Structural  
163 determination based on sXRD confirmed a new MHP  
164 derivative—morphPbI<sub>3</sub>, crystallizing in the orthorhombic  
165 space group *P*2<sub>1</sub>2<sub>1</sub> (No. 19). Full crystallographic details of  
166 this new phase are given in Table S1; bond length and angles  
167 are listed in Tables S2a and S2b, respectively. In morphPbI<sub>3</sub>,  
168 [PbI<sub>6</sub>]<sup>4-</sup> units are arranged in 1D chains of face-sharing  
169 octahedra (Figure 1b). The powder XRD (pXRD) pattern of  
170 morphPbI<sub>3</sub> matches the pXRD pattern simulated for the  
171 sXRD-derived crystal structure (Figure S2a). The Tauc  
172 analysis of the absorption spectrum of the ground powder of  
173 morphPbI<sub>3</sub> indicates a direct band gap of 2.69 eV (Figure 1e).  
174 No photoluminescence (PL) is detected for this compound.  
175 Despite the large band gap and the absence of PL, the 1D  
176 yellow phase could have possible applications in second-order  
177 nonlinear optics<sup>54</sup> and piezoelectric devices<sup>55</sup> owing to its  
178 noncentrosymmetric space group (only ~18% of inorganic  
179 crystal structures reported are noncentrosymmetric).<sup>56</sup>

180 **2.2. Robot-Accelerated ASVC Perovskite Workflow.** Although we only observed a single 1D morph-Pb-I structure with our isolated benchtop syntheses, 2D derivatives have been predicted theoretically.<sup>20</sup> To more comprehensively search for morph-Pb-I phases of different dimensionalities (especially 2D), we developed a robot-accelerated perovskite workflow based on high-throughput (HT) ASVC to explore a much larger reaction-composition space (see Figure S3 for the workflow). Similar to our previous robot-assisted perovskite investigation and discovery workflow,<sup>28</sup> our HT-ASVC workflow utilizes a liquid-handling robot to dispense perovskite precursor solutions into reaction vials. We designed a custom, multiwell ASVC microplate (see Figure S3; CAD file available<sup>57</sup>) that allowed our liquid-handling robot to prepare 24 parallel ASVC reactions on the 500 μL scale. Additional details for the HT-ASVC process are described in the Supporting Information. After crystallization, we photographed the reaction vials and recorded the morphologies and colors of the solid products. Then, we characterized the solid products using HT optical microscopy, absorption and PL spectroscopy, and pXRD (HT characterization workflow illustrated in Figure S4). Representative optical micrographs, absorption/PL spectra, and pXRD patterns are shown in Figures S5–S8.

181 **2.3. Primary Screening of 3D Reaction-Composition Space.** Using our robotic workflow, we performed primary screening of the reaction-composition space of PbI<sub>2</sub> concentration ([Pb]), morphI concentration ([morph]), and formic acid concentration ([FAH]). We chose this reaction space because our previous HTE work has demonstrated that tuning the analogous concentrations in different ammonium lead halide systems resulted in the successful synthesis of 19 MHP derivatives from 45 A-cation candidates.<sup>28</sup> We used the Kennard-Stone (KS) algorithm<sup>58,59</sup> to uniformly sample 48 primary reactions from grid points generated in the allowed reaction-composition space (concentration constraints shown in Table S4). KS and grid-point generation algorithms are implemented in the ESCALATE software pipeline we developed to manage HT experiments and capture data.<sup>60</sup>

182 Unlike benchtop syntheses, primary HT-ASVC screening produced both yellow solids and red solids with an unknown structure. Structural determination based on sXRD confirmed that the red crystals (Figure 1c) are a new MHP derivative with formula (morph)<sub>2</sub>PbI<sub>4</sub> in the monoclinic space group C2/

(No. 15, see Tables S1, S3a and S3b for details). In this crystal structure,  $[PbI_6]^{4-}$  octahedral units are corner-connected to form 2D layers (Figure 1d). Therefore, we successfully demonstrated the efficacy of the robotic workflow in the discovery of a new 2D phase in the morph-Pb-I system, which was found neither in our benchtop synthesis nor experimentally reported in the literature. Compared to the 1D yellow phase, the 2D red phase has distinct optical properties. The absorption spectrum of the 2D phase indicates a direct band gap of 2.15 eV, which is 0.54 eV lower than that of the 1D phase (Figure 1e). Under 470 nm excitation, the 2D phase exhibits a broad emission peak at 588 nm (Figure 1f), which could be utilized in the application of light-emitting diodes.<sup>61</sup> Unlike the 1D phase, the 2D phase has a centrosymmetric crystal structure.

Visualizing the distribution of reaction outcomes across the [morph]–[Pb]–[FAH] composition space (Figure 2) illus-

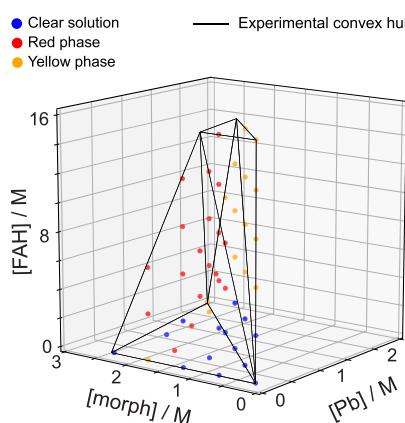
trend in the fabrication of MHP devices is to use mixed solvents and additives to modify the crystallinity and morphology of MHP thin films, which can improve device performance.<sup>62</sup> However, the effect of these solvents and additives on the dimensionality of MHP products is still unclear. To understand such effects, we simultaneously incorporated four additional solvents and additives into our HT-ASVC reactions. These included three common solvents for MHP syntheses: DMF, dimethyl sulfoxide (DMSO), and  $\gamma$ -butyrolactone (GBL). These solvents were selected because of their distinct physical properties, such as their polarity<sup>63</sup> and their affinity for coordinating metal ions<sup>62</sup> and accepting hydrogen bonds.<sup>64</sup> We also included water as an additive because our previous study demonstrated that water content in perovskite precursor solutions affects the crystallinity of MHP single crystals and thin films.<sup>65</sup> Unlike DMSO or DMF, water and formic acid are both hydrogen bond donors and acceptors.<sup>66</sup> When coexisting in solution, DMSO (or DMF) and water (or formic acid) are likely to form hydrogen bonds. Our new reaction-composition space was thus composed of six reaction parameters: [morph]; [Pb]; the volume fraction of DMSO,  $V_{f,DMSO}$ ; the volume fraction of GBL,  $V_{f,GBL}$ ; [FAH]; and the concentration of water,  $[H_2O]$ . A seventh parameter,  $V_{f,DMF}$ , was not included in this reaction space because it is calculated as  $1 - V_{f,DMSO} - V_{f,GBL}$ . The modified robotic-synthesis procedure is described in the Supporting Information, and the constraints of all six parameters are given in Table S5. In the constrained six-dimensional (6D) space, we generated a pool of 469,326 possible reaction compositions located on a fixed grid.

To rapidly characterize the outcomes of reactions performed in this extended 6D space, we acquired absorption spectra on products with a multifunction plate reader. Automated scripts classified reaction products with absorption edges  $\geq 2.3$  eV as the “yellow phase” and the reactions with absorption edges  $\leq 2.15$  eV as the “red phase” (see Figure S6a). If no solids formed, we labeled the reaction as a “clear solution.” Representative pXRD patterns (Figures S7 and S8) verified that the diffraction peaks of reaction products correspond to the phases predicted using absorption edges.

## 2.5. Exploring the Role of MHP Additives Using AL.

To understand how combinations of additives contribute to the dimensionality of crystals in the morph-Pb-I system, we sought to train an ML model to predict the phase and dimensionality (i.e., 1D yellow phase or 2D red phase) for each combination of reagents in our 6D reaction pool. To train such a model efficiently in such high-dimensional space, we developed an uncertainty-based AL method to perform repeated cycles of HT-ASVC microplate reactions that iteratively refine the ML model. For each cycle in our method, an AL algorithm selects the next batch of reactions to perform by identifying the regions of the 6D reaction-composition space, where the ML model has the highest prediction uncertainty.<sup>29,67</sup>

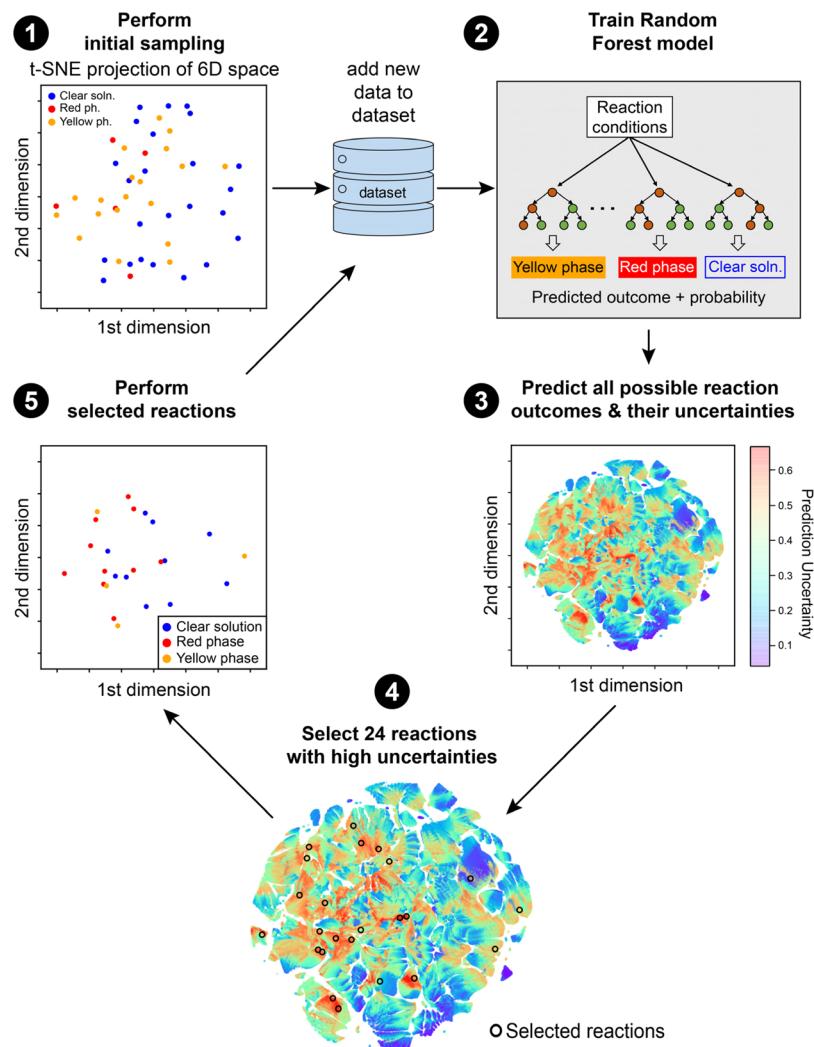
To initiate AL, we performed a uniform sampling of 48 reactions using the KS algorithm. Because visualizing 6D data graphically is difficult, we projected reaction outcomes onto a lower-dimensional space using t-distributed stochastic neighborhood embedding (t-SNE)<sup>68</sup> (Step 1 in Figure 3). In this initial sampling, we observed all three classes of reaction outcomes. We tested Random Forest (RF) models and Pearson VII universal function Kernel-based supporting vector



**Figure 2.** Convex hull of the allowed reaction-composition space (black lines) and the primary screening experiments (colored circles) contained within it, as a function of [Pb], [morph], and [FAH]. The blue circles indicate a clear solution with no crystals. The red and yellow circles indicate reaction outcomes of 2D (red) and 1D (yellow) phases, respectively.

rates how reaction conditions determine the crystallization of the 1D and 2D (red and yellow) phases. For each HT-AVSC experiment in this space, we assigned a reaction outcome from one of the three classes: (1) clear solution, (2) red phase, and (3) yellow phase (see “Product Scoring Rubric Based on Human Inspection” in the Supporting Information). For mixtures of yellow and red products, we labeled the reaction outcomes based on the major product. In general, clear solutions (no solids) were observed below 6 M FAH. Above 6 M FAH, the yellow phase formed at lower morph:Pb ratios ( $\sim 1$ ), while the red phase was more likely to form at higher morph:Pb. This dependence on the morph-to-Pb ratio is due to the different chemical stoichiometries of the two phases (morph:Pb = 1 in the yellow phase and 2 in the red phase). Tuning the reactant ratio per target-compound stoichiometry has been utilized previously in several CMH systems, such as FA-Pb-Br<sup>24</sup>. Therefore, not only did primary HT-ASVC screening of the morph-Pb-I system identify a 2D phase, but the resulting dataset also provided guidance for controlling the dimensionality of morph-Pb-I in 3D reaction-composition space.

**A. Modified Workflow To Screen Additives in Six-Dimensional Reaction-Composition Space.** A prevailing



**Figure 3.** Illustration of the diverse-mini-batch-sampling AL loop. Uniformly sampled seed reactions (Step 1) are used to train an RF model (2), which is then used to calculate prediction uncertainties (3) for the entire pool of potential reactions previously generated in 6D experimental space. In subsequent cycles of AL, reactions are selected using a diverse-mini-batch-sampling algorithm that prioritizes reaction conditions where model predictions have high uncertainty (4). These reactions are performed using a synthesis robot (5), and the data are used to retrain the model and perform additional cycles of active learning (2–5). All AL-sampled reactions (colored scattered circles) and prediction uncertainty distributions (colored maps) are projected onto 2D space using t-distributed stochastic neighborhood embedding (t-SNE).

325 machine (SVM\_PUFK) models, optimizing their hyper-  
 326 parameters on the initial dataset using fivefold cross-validation  
 327 (CV). We selected the RF model as the basis for our AL  
 328 algorithm because it exhibited the highest CV accuracy of  $0.80 \pm 0.09$  (Table S6). The RF model, trained by the initial dataset  
 329 in Step 2, is used to predict reaction outcomes (yellow, red,  
 330 or no crystals) and calculate the prediction probabilities  $P$  for  
 331 each member of the reaction pool.  $P$  is defined as the  
 332 probability of the reaction outcome predicted to have the  
 333 highest likelihood of forming (e.g.,  $P = 0.6$  if prediction  
 334 likelihoods are 0.2 for the clear solution, 0.6 for the red phase,  
 335 and 0.2 for the yellow phase). Then, the prediction uncertainty  
 336 ( $U$ ) for each point is calculated as  $1 - P$  (Step 3 in Figure  
 337 3). The distribution of  $U$  shows that the regions with high  
 338 uncertainty ( $U > 0.5$ ) are found where multiple compounds  
 339 form under similar conditions.  
 340 Our AL algorithm is designed to generate the next batch of  
 341 24 robot reactions based on the points in reaction-composition  
 342 space with highest  $U$ . Simply selecting the 24 reactions with

343 the highest  $U$ , however, would result in reactions with very  
 344 similar conditions. To avoid such over-sampling in a small  
 345 region, we implemented the diverse-mini-batch-sampling  
 346 algorithm,<sup>69</sup> which divides high-uncertainty reactions ( $U >$   
 347 0.5) into 24 mini batches using  $k$ -mean clustering,<sup>70</sup> weighted  
 348 by  $U$ . Then, the centers of mass of the 24 mini batches are  
 349 selected as the set of reactions to perform in the next AL cycle.  
 350 As shown in Step 4 in Figure 3, the selected reactions have  
 351 diverse reaction conditions and are located in regions of high  
 352  $U$ . After this batch of reactions is performed and characterized  
 353 in Step 5, the reaction outcomes are collected and added to  
 354 the dataset. The updated dataset is then used to retrain the RF  
 355 model, at which point, the next AL cycle (2–5)  
 356 commences.

357 A critical outstanding question in AL-guided materials  
 358 synthesis is when to terminate the AL loop. In principle, the  
 359 RF model performance should be evaluated during AL to  
 360 terminate the AL process when the model ceases to improve,  
 361 but such evaluation is practically challenging because a large  
 362

363 test set is not available to evaluate the RF model. To determine  
 364 the stopping point for our AL runs, we monitored the average  
 365 uncertainty (AU) and prediction confidence (PC) of the RF  
 366 model after each AL iteration.<sup>71</sup> AU and PC are defined in eqs  
 367 1 and 2, respectively:

$$368 \quad AU = \frac{\sum_{k=1}^N U_k}{N} \quad (1)$$

$$369 \quad PC = \frac{\sum_{k=1}^N (P_k - P'_k)}{N} \quad (2)$$

370 Here,  $N = 469,326$  is the number of potential reactions in  
 371 the reaction pool;  $U_k$  is the prediction uncertainty for the  $k$ th  
 372 reaction in the reaction pool;  $P_k$  and  $P'_k$  are the prediction  
 373 probabilities of the most likely class and second most likely  
 374 class for the  $k$ th reaction. A rising PC (decreasing AU) over AL  
 375 cycles indicates that AL is still improving the RF model. A  
 376 decreasing PC (increasing AU) suggests that the RF model has  
 377 ceased to improve, and AL should be stopped. Heuristically,  
 378 PC usually increases at the beginning of AL and then decreases,  
 379 which indicates that ML models were often improved by AL in  
 380 the first several cycles and then remained little changed.<sup>71,72</sup> In  
 381 general, adding new data to the training set does not reduce  
 382 model performance, so decreasing PC implies a lack of  
 383 improvement rather than a reduction in prediction quality.

384 After the first AL iteration of 24 reactions, the PC of our RF  
 385 model increased from 0.41 to 0.45 and AU decreased from  
 386 0.34 to 0.32 (Figure 4a), which indicates an improvement of  
 387 the RF model. Surprisingly, the second AL iteration reduced  
 388 PC and increased AU, suggesting that experimentation should  
 389 stop, as the model is no longer improving with the added data  
 390 points. To confirm the downward trend of PC (and upward  
 391 trend of AU), we performed three more AL iterations after the

392 second iteration. PC continued to decrease, so we ceased the 392  
 393 AL experiments after the fifth iteration (the practical stopping 393  
 394 point). The dataset from AL iterations 2–5 (shown in Figure 394  
 395 S9) is included in the training set because expanding a training 395  
 396 set rarely reduces the prediction accuracy for the test set. 396

397 To justify the practical stopping point, one could, in 397  
 398 principle, calculate the prediction accuracy of the RF model 398  
 399 (trained by initial sampling +5 AL iterations) over a large test 399  
 400 set representative of the overall reaction pool. However, the 400  
 401 time and labor required to collect this test set defeat the 401  
 402 purpose of AL,<sup>71</sup> which is to minimize the number of 402  
 403 experiments. Here, we performed three analyses to support 403  
 404 the conclusion that we have reached a reasonable stopping 404  
 405 point after five AL iterations.

406 A good ML model should show modest variance during the 406  
 407 last few AL iterations before the practical stopping point (from 407  
 408 the second to fifth iteration),<sup>73</sup> shown as small changes in its 408  
 409 CV accuracy when including the last few AL runs in the 409  
 410 training set. It is worth noting that because the training set has 410  
 411 been collected, we are not limited to using RF models in the 411  
 412 search for the best ML model. Our first analysis calculated the 412  
 413 CV accuracies of multiple ML models after different AL 413  
 414 iterations (Table S7 and Figure S10). After the 5th AL 414  
 415 iteration, the Gaussian process (GP) model shows the highest 415  
 416 CV accuracy of  $0.78 \pm 0.04$  (Figure 4b) and close to the 416  
 417 theoretical limit of 0.78 (calculated by overfitting the dataset 417  
 418 with deep neural networks, shown in Figure S11).<sup>74</sup> The CV 418  
 419 accuracy is well above the random-classification (control) 419  
 420 accuracy of 0.34 and the majority-class vote accuracy of 0.41. 420  
 421 As the most accurate ML model, the GP model shows only a 421  
 422 slight variation in the CV accuracy from the second to the fifth 422  
 423 AL iteration, which indicates an insignificant variance in the 423  
 424 GP model and suggests that the AL experiments have reached 424  
 425 the stopping point.

426 If the most accurate model, the GP model (trained on the 426  
 427 dataset after the 5th AL iteration), can accurately predict 427  
 428 unseen reaction conditions, AL has likely reached its stopping 428  
 429 point. Therefore, in our second analysis, we performed 24 429  
 430 reactions located far from the tested reactions (initial sampling 430  
 431 +5 AL runs) in the reaction-composition space (Figure S12a). 431  
 432 The 24 reactions were selected using the KS algorithm. The 432  
 433 GP model, trained on the tested reactions, shows a prediction 433  
 434 accuracy of 0.92 for the unseen 24 reactions (Figure S12b), 434  
 435 which leaves little room for improvement and indicates that AL 435  
 436 has reached its stopping point.

437 Testing the AL algorithm on a synthetic dataset (with labels) 437  
 438 allows us to monitor the model prediction accuracy for the 438  
 439 whole dataset and to investigate whether the stopping criterion 439  
 440 based on prediction confidence is reasonable. As the final 440  
 441 analysis, we ran the AL algorithm on a synthetic dataset with a 441  
 442 similar structure to our experimental dataset (see Figure S13 442  
 443 and the Supporting Information for details). After AL 443  
 444 commences, the RF model's PC increases until the fourth 444  
 445 AL cycle, after which PC decreases (Figure S14a). Thus, the 445  
 446 fourth iteration is considered as the theoretical stopping point. 446  
 447 Like our AL experiment, the decreasing trend can be 447  
 448 confirmed with a few additional iterations. Meanwhile, the 448  
 449 prediction accuracy for the whole dataset rapidly increases 449  
 450 from 0.48 (after initial sampling) to 0.82 (after four AL 450  
 451 iterations) (Figure S14b). After the 4th cycle, the accuracy 451  
 452 plateaus, eventually stabilizing at 0.85 after 100 iterations. 452  
 453 Therefore, the model performance has only a 4% improvement 453  
 454 with additional 96 AL iterations after the theoretical stopping 454

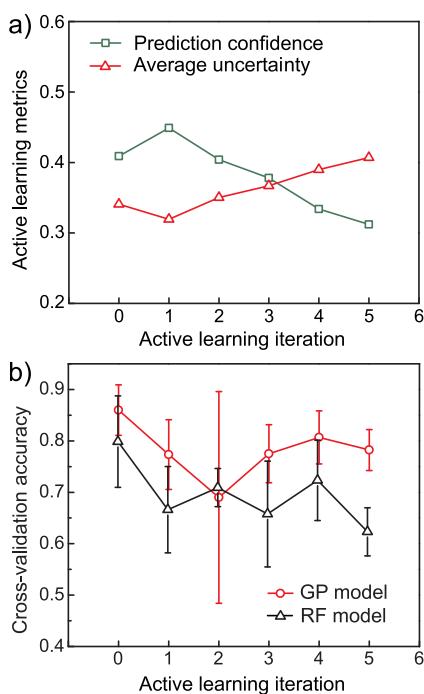


Figure 4. (a) Prediction confidence and average uncertainty of the RF model in each iteration of AL. (b) CV accuracies of the Gaussian process and RF models on the dataset collected after each iteration of AL (error bars represent the standard deviation across fivefold CV).

455 point, confirming that the stopping criterion based on  
456 prediction confidence is reasonable. Clear stopping criteria,  
457 such as that demonstrated here, will benefit AL-assisted  
458 materials and chemistry research.

459 To summarize, we determined the AL stopping point based  
460 on the changes in prediction confidence and performed three  
461 tractable analyses to support the stopping point. Using diverse-  
462 mini-batch AL and a stopping criterion based on prediction  
463 confidence, we only needed to explore at most 0.035% of the  
464 reaction pool to successfully build and confirm the stopping  
465 point of ML models that accurately predict the formation of  
466 the 1D and 2D MHP derivatives in the morph-Pb-I system.

## 467 2.6. Importance of Features and Their Effect on Dimensionality.

468 A predictive ML model can be used to understand the physicochemical process of morpholinium lead  
469 iodide crystallization. To uncover the reaction parameters that  
470 have the most significant influence on morph-Pb-I dimension-  
471 ality, we performed a permutation-feature-importance anal-  
472 ysis<sup>75</sup> (see details in the Supporting Information) on the RF  
473 model trained by the portion of the dataset that includes only  
474 yellow- and red-phase outcomes. Feature-importance analysis  
475 revealed that [Pb], [morph], [FAH], and [H<sub>2</sub>O] are important  
476 for controlling the formation of 1D and 2D phases (feature  
477 importance > 0.1), while the compositions of solvents (i.e.,  
478 V<sub>f,DMSO</sub> and V<sub>f,GBL</sub>) are much less important (feature  
479 importance < 0.05) and can be ignored (Figure 5a). To  
480 validate the feature downselection, we retrained the RF model  
481 with the yellow/red dataset and used only [Pb], [morph],  
482 [FAH], and [H<sub>2</sub>O] as features. Similar CV accuracies were  
483 observed compared to the model trained by the full set of  
484 features (Figure 5b), suggesting that the feature selection is  
485 effective. This analysis is corroborated by the visually  
486 distinguishable boundary between yellow and red phases in  
487 the parameter space of ln([morph]/[Pb]), [FAH], and [H<sub>2</sub>O]  
488 (Figure 5c).

489 To determine whether the important features have positive  
490 or negative effects on the formation of the yellow and red  
491 phases, we developed a data-driven approach that combines  
492 logistic regression (LR) modeling and statistical hypothesis  
493 testing. Using a “crystals only” dataset containing only  
494 outcomes that produced solids, we performed LR using only  
495 [Pb], [morph], [FAH], and [H<sub>2</sub>O] as inputs, and with the  
496 outcomes labeled as “1” for the yellow phase and “0” for the  
497 red phase. Because the LR model demonstrated a reasonable  
498 CV accuracy of  $0.81 \pm 0.07$ , the sign of the slope can be  
499 treated as the direction of the correlation between the  
500 corresponding feature and yellow phase formation. To test  
501 the hypotheses with statistical significance, we used boot-  
502 strapping<sup>76</sup> to sample the “crystals only” dataset 1000 times.  
503 We fit the 1000 samples into the LR model and obtained a  
504 distribution of slopes for each important feature (Figure S16a).  
505 The slopes of [morph], [Pb], [FAH], and [H<sub>2</sub>O] are  $-2.33 \pm$   
506 0.24,  $2.37 \pm 0.16$ ,  $1.09 \pm 0.24$ , and  $1.51 \pm 0.22$ , respectively.  
507 Based on this LR analysis, we hypothesized that when solid is  
508 formed, [morph] has a negative effect on yellow phase  
509 formation while [Pb], [FAH], and [H<sub>2</sub>O] have positive effects.  
510 To further validate the correlations between reagent  
511 concentrations and crystal phases, we performed one-tailed  
512 statistical hypothesis testing.<sup>77</sup> We defined one null hypothesis  
513 for each feature: for the yellow phase formation, [morph] is  
514 hypothesized to have positive or no effect while [Pb], [FAH],  
515 and [H<sub>2</sub>O] are hypothesized to have negative or no effect.  
516 Given the mean values and standard errors of the slopes, we

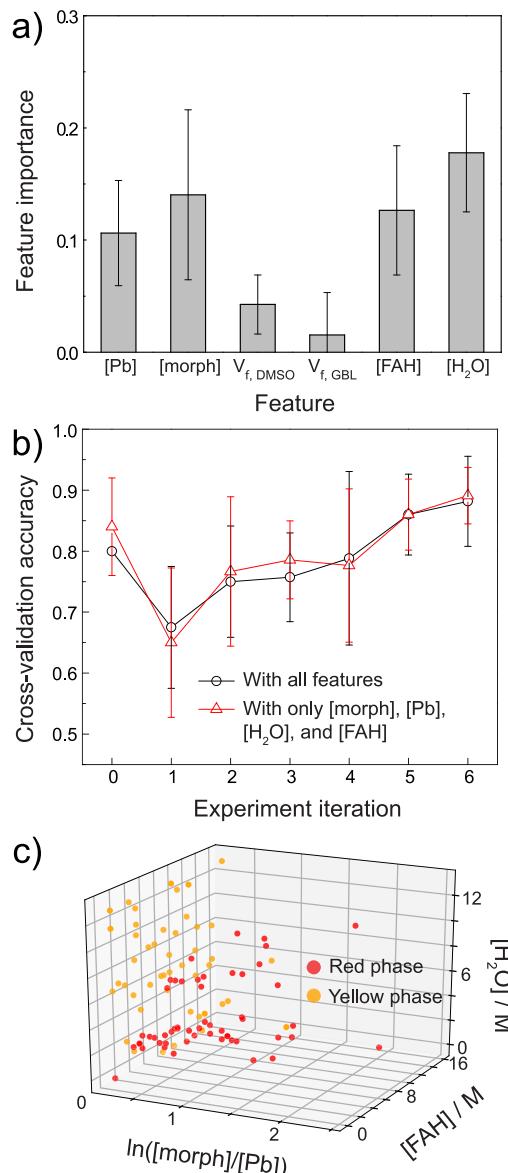


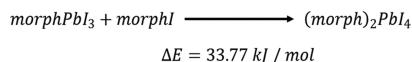
Figure 5. (a) Permutation-feature-importance of the RF model trained by the yellow and red phase reactions. (b) CV accuracies of the RF model on predicting yellow phase vs red phase, with all features (black line) and only [morph], [Pb], [H<sub>2</sub>O], and [FAH] (red line). (c) Outcomes of the yellow and red phase reactions as a function of ln([morph]/[Pb]), [FAH], and [H<sub>2</sub>O].

518 rejected all four null hypotheses with a confidence level of 99%.  
519 Therefore, all four features have hypothesized effects on the  
520 yellow phase formation. Using the same approach, we  
521 discovered that when solid is formed, [morph] has a positive  
522 effect on red phase formation while [Pb], [FAH], and [H<sub>2</sub>O]  
523 have negative effects.

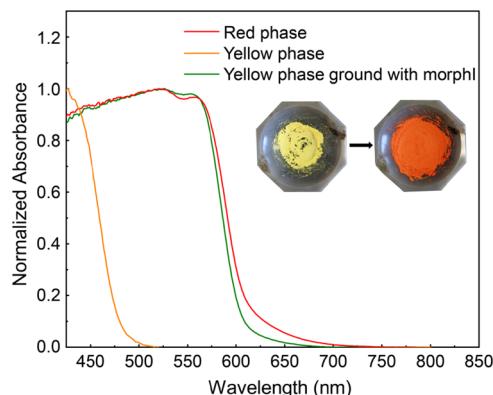
524 **2.7. Understanding the Underlying Physicochemical Process of Dimensionality Control.** The effects of [Pb] and  
525 [morph] on the dimensionality of morph-Pb-I perovskite  
526 derivatives can be largely explained by the chemical  
527 stoichiometries of the red and yellow phases. The phys-  
528 icochemical process through which additives (i.e., water and  
529 formic acid) influence the dimensionality is still unclear. To  
530 understand this physicochemical process, we studied both the  
531 thermodynamics and kinetics of the ASVC reaction. First, we  
532 investigated whether the reaction is under thermodynamic  
533

s<sub>34</sub> control. We calculated the total energy of the yellow phase and s<sub>35</sub> red phase using density functional theory (DFT).<sup>78</sup> DFT s<sub>36</sub> calculations show that the yellow phase is slightly more stable s<sub>37</sub> than the red phase. However, the formation energy difference s<sub>38</sub> between these two phases is negligible ( $\Delta E_{\text{form}} = 6.8 \text{ kJ}$  per s<sub>39</sub> mole of Pb) and within the typical intrinsic error of DFT s<sub>40</sub> ( $\sim 0.1 \text{ eV}$  or  $\sim 10 \text{ kJ/mol}$ ).<sup>79</sup> Thermogravimetric analysis s<sub>41</sub> (TGA) shows that the thermal decomposition temperatures s<sub>42</sub> for the yellow and red phases are close (240 and 200 °C, s<sub>43</sub> respectively, shown in Figure S17), which agrees with DFT s<sub>44</sub> results. DFT calculations also show that only a small amount of s<sub>45</sub> energy (33.77 kJ/mol) is needed to convert the yellow phase s<sub>46</sub> to the red phase in the solid state (Scheme 1). Theoretically,

**Scheme 1. Solid-State Conversion of the Yellow Phase to Red Phase**



s<sub>47</sub> this conversion can be achieved by mechanochemical grinding, s<sub>48</sub> which provides an energy of 95–112 kJ/mol.<sup>80</sup> Ultraviolet– s<sub>49</sub> visible (UV-vis) absorption spectra show that at room s<sub>50</sub> temperature, the yellow phase can be converted to the red s<sub>51</sub> phase in the solid state with one equivalent morphI and s<sub>52</sub> manual grinding (Figure 6). The room-temperature synthesis

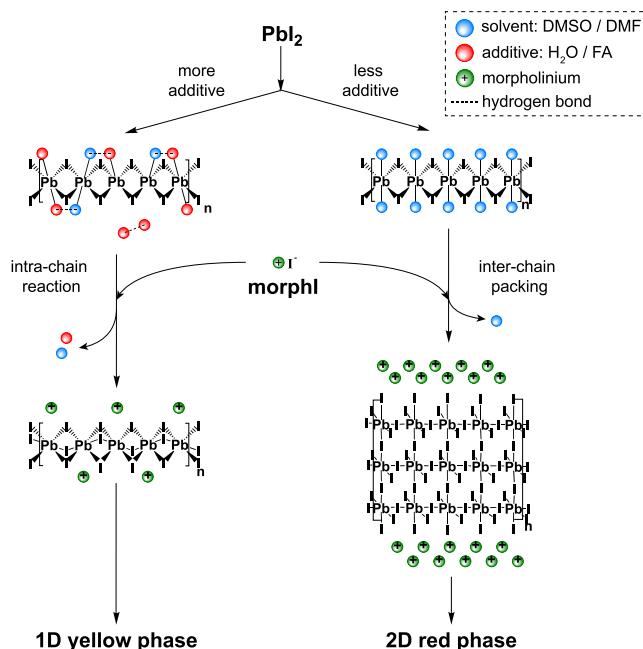


**Figure 6.** UV-vis absorption spectra of the red phase, yellow phase, and yellow phase ground with additional morphI. The inset photographs are powders of the yellow phase (left) and the yellow phase ground with morphI (right).

s<sub>53</sub> of ASVC and the small energy difference between the 1D and s<sub>54</sub> 2D phases suggest that the formation of morph-Pb-I phases s<sub>55</sub> with varying dimensionalities is not under thermodynamic s<sub>56</sub> control.

s<sub>57</sub> The small energetic difference between the 1D and 2D s<sub>58</sub> phases provides the opportunity for additives such as H<sub>2</sub>O and s<sub>59</sub> formic acid to exert influence on the dimensionality of the final s<sub>60</sub> crystals. LR correlates the use of H<sub>2</sub>O and FA with the s<sub>61</sub> formation of the 1D yellow phase, but further exploration of s<sub>62</sub> the mechanistic origins of such phase selectivity is needed. s<sub>63</sub> Experimental and computational investigations of perovskite s<sub>64</sub> formation pathways have identified key prenucleation inter- s<sub>65</sub> mediates in the form of 1D chains of PbI<sub>2</sub> coordinated with s<sub>66</sub> solvent and additive (e.g., H<sub>2</sub>O) molecules.<sup>81–83</sup> These 1D s<sub>67</sub> PbI<sub>2</sub> oligomers, shown in Scheme 2, may be converted into the s<sub>68</sub> 1D lead iodide frameworks in yellow-phase morphPbI<sub>3</sub> or may s<sub>69</sub> stack into the 2D lead iodide frameworks of the red phase.

**Scheme 2. Possible Reaction Scheme for the Formation of the 1D Yellow Phase and 2D Red Phase**



Because intramolecular hydrogen bonding in proteins and polymers has been observed to alter their coordination with metal ions,<sup>84</sup> intrachain H-bonding induced by protic additives may play a role in the phase selectivity imparted by additives in perovskite crystallization. H-bonding between additive and solvent molecules coordinated to adjacent Pb<sup>2+</sup> ions on the same 1D PbI<sub>2</sub> chain could geometrically strain metal–ligand bonds (Scheme 2), promoting the dissociation of H-bonded ligand pairs.<sup>82</sup> Faster dissociation of ligands on neighboring sites could promote the formation of the 1D yellow phase by allowing faster conversion of the 1D prenucleation complexes to the 1D lead iodide frameworks of the morphPbI<sub>3</sub> phase.<sup>81</sup> Meanwhile, slower dissociation could make 1D PbI<sub>2</sub> intermediates more likely to pack with other 1D chains to form the 2D lead iodide framework of the red (morph)<sub>2</sub>PbI<sub>4</sub> phase.<sup>84</sup> A more detailed mechanism following this reasoning is provided in Scheme S1 in the Supporting Information. If validated experimentally, the proposed scheme suggests that H<sub>2</sub>O and FA additives selectively promote the formation of the 1D morphPbI<sub>3</sub> phase by modulating the reactivity of PbI<sub>2</sub> intermediates via intrachain hydrogen bonding.

### 3. CONCLUSIONS

Using a robotic workflow based on ASVC, we synthesized two novel MHP derivatives [1D morphPbI<sub>3</sub> and 2D (morph)<sub>2</sub>PbI<sub>4</sub>] with distinct optical properties. Although the existence (but not crystal structures) of 2D MHP derivatives based on morpholinium has been postulated theoretically, the synthesis and characterization of a 2D derivative has not been reported until this work. We demonstrated the efficacy of the KS sampling algorithm + robotic workflow in finding rare MHP derivatives. Using the uncertainty-based AL method with decreased prediction confidence as a stopping criterion, we sampled only 0.035% of the reaction-composition space to build a predictive ML model to classify the reaction conditions where 1D and 2D phases are formed. By analyzing the feature importance of the predictive ML model, we elucidated

[Pb], [morph], [FAH], and [H<sub>2</sub>O] have significant influence on the dimensionality control in the morph-Pb-I system. Using these data, along with DFT calculations, thermogravimetric measurements, and mechanochemistry observations, we explored the mechanistic origins of the selective formation of the 1D and 2D phases. In one possible scheme, water and formic acid may accelerate the formation of the 1D phase via intrachain hydrogen bonding, which could be observed with other A-cations and MHP systems. Our strategy of using additives to control dimensionality has the potential to be applied in many other CMH systems. With the AL stopping criterion developed and tested in this work, the AL + HTE approach will be valuable for any material research that benefits from predicting and controlling different phases/compounds in a vast reaction-composition space.

## 4. METHODS

**4.1. Materials.** Lead iodide (PbI<sub>2</sub>) (99%), formic acid (FAH) ( $\geq 95\%$ ), dimethylformamide (DMF) (99.8%), dimethyl sulfoxide (DMSO) ( $\geq 99.5\%$ ), and dichloromethane (DCM) ( $\geq 99.8\%$ ) were purchased from Sigma Aldrich Chemicals.  $\gamma$ -Butyrolactone (GBL) ( $\geq 98\%$ ) was purchased from Spectrum Chemical. Morpholinium iodide (morphI) (98%) was purchased from GreatCell Solar.

**4.2. Robotic Workflow.** Our HT-ASVC workflow utilizes a Hamilton Microlab NIMBUS4 liquid-handling robot equipped with four independent micropipettors. Stock solutions of PbI<sub>2</sub>-morphI mixture solution and morphI solution were prepared based on experimental data entry files generated by ESCALATE<sup>60</sup> using the solubility data. After stock solution preparation, all reaction components (i.e., the stock solutions, pure solvents, and additives) were placed in programmatically designated locations on NIMBUS operation deck. A synthetic flow chart describing stock solution preparation and robotic procedures is shown in Figure S3 in the Supporting Information. A customized robot-compatible crystallization block, containing 24 pairs of wells, was placed on the Hamilton heater and shaker (HHS). For one pair of wells in the block, perovskite stock solution occupied one well, and the other well contained the antisolvent (DCM). The 8 × 43 mm (diameter × height) glass scintillation vials were used as reaction and antisolvent vessels. The vials were maintained at 75 °C during the addition of the stock solutions for dissolution. Formic acid was added to each reaction vial, followed by 15 min of shaking to avoid premature precipitation of PbI<sub>2</sub>, morphI, or perovskite. After vortexing, the crystallization block was cooled to room temperature before DCM was added to the antisolvent wells in the crystallization block. After DCM addition, we manually sealed the block with a metal cap and stored the block at 20 °C without disturbance for 16 h. Additional details are given in the “Robotic Workflow” section in the Supporting Information.

**4.3. Characterization.** Powder XRD (pXRD) measurements were performed on a Bruker AXS D8 Discover GADDS X-ray diffractometer with a Vantec-500 area detector, and it is operated at 35 kV/40 mA with a Co K $\alpha$  radiation source with a wavelength of 1.79 Å. UV-vis absorption spectra were collected using an Agilent Cary-5000 UV-vis-near-infrared (NIR) spectrophotometer with an internal diffuse reflectance accessory. PL spectra were measured using a Horiba Jobin Yvon Fluorolog-3 spectrofluorometer and collected from 530 to 720 nm with 1 nm wavelength steps and 0.01 s integration time per step. TGA was performed using TA Instruments QSS00 TGA-MS. The sample weight change was measured from room temperature to 450 °C with a ramp rate of 10 °C/min under nitrogen. HT pXRD measurements were performed on glass slides in a customized sample holder on the same Bruker X-ray diffractometer. The sample locations are programmatically defined in pXRD software. HT UV-vis-NIR absorption spectra were collected using a custom-built reflection-mode UV-vis-NIR absorption spectrometer, which has a motorized XY stage to enable automated measurement. The spectra were measured from 350 to 2500 nm and averaged over 100

acquisitions (each acquisition takes 1 s). HT PL spectra were collected using a Biotek Synergy 4 UV-vis absorption/fluorescence microplate reader. The PL spectra were measured from 540 to 720 nm with 450 nm excitation. HT optical micrographs were acquired using a Biotek Cytation5 Cell Imaging Multimode Reader with a 4× objective lens.

**4.4. Reaction Outcome Scoring.** All AL reactions from the modified workflow were scored using the UV-vis spectra of the products. The reactions from primary screening were scored by human inspection into three classes. Clear solution: no solid observed in the solutions. Red phase: red crystals or powder are the major solid products. Yellow phase: yellow crystals or powder are the major solid products.

**4.5. Software and ML.** Our custom-developed pipeline software, ESCALATE,<sup>60</sup> was used to specify experimental parameters in robot readable files, provide instructions for human operators, and capture experiment results and observations. All algorithms in this work were written in Python 3.6 in Jupyter notebooks using the following libraries: Numpy 1.18.0, Pandas 0.22.0, Scipy 1.3.0, Matplotlib 3.1.0, Scikit-learn 0.21.3, and modal 0.3.5. We used a “Stratified Shuffle Split” method from Scikit-learn to generate training/testing datasets for CV of ML models. In the case of fivefold CV, there are five different train/test splits on the dataset: in each split, 80% of the data were used to train the ML model, while 20% of the data were reserved for testing. The testing sets were randomly drawn from whole datasets in a stratified style (i.e., testing sets have the same percentage of samples of each target class as the whole datasets). Before each drawing, the datasets were shuffled, so the testing datasets are not necessarily exclusive between splits. CV accuracies were calculated by averaging the prediction accuracies of five different train/test splits created by CV on the dataset. Our experimental results were interpreted as either two-class or three-class classification problems. In the two-class case, “yellow phase” outcome was considered as a “positive” result while “red phase” outcome was considered as a “negative” result. The “clear solution” outcomes were excluded. In the three-class case, we calculated the overall prediction accuracy by averaging the prediction accuracies from all possible one-vs-all classifications [e.g., yellow phase (positive) vs nonyellow phase (negative)].

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.chemmater.1c03564>.

Detailed materials and methods; crystal scoring rubric; single crystal structure refinement and CIF files; additional XRD patterns, absorption, and emission spectra; state spaces and metrics for experimental and simulated active learning runs; TGA data; experimental datasets; and Python code used for analysis and experiment generation (PDF)

Crystallographic information of morphPbI<sub>3</sub> (CIF)

Crystallographic information of (morph)<sub>2</sub>PbI<sub>4</sub> (CIF)

Python code used for analysis and experiment generation (ZIP)

## AUTHOR INFORMATION

### Corresponding Author

Emory M. Chan — Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; [orcid.org/0000-0002-5655-0146](https://orcid.org/0000-0002-5655-0146); Email: EMChan@lbl.gov

### Authors

Zhi Li — Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

734 Philip W. Nega — Molecular Foundry, Lawrence Berkeley  
735 National Laboratory, Berkeley, California 94720, United  
736 States  
737 Mansoor Ani Najeeb Nellikkal — Department of Chemistry,  
738 Haverford College, Haverford, Pennsylvania 19041, United  
739 States; [0000-0002-8258-0613](https://orcid.org/0000-0002-8258-0613)  
740 Chaochao Dun — Molecular Foundry, Lawrence Berkeley  
741 National Laboratory, Berkeley, California 94720, United  
742 States; [0000-0002-3215-6478](https://orcid.org/0000-0002-3215-6478)  
743 Matthias Zeller — Department of Chemistry, Purdue  
744 University, West Lafayette, Indianapolis 47907, United  
745 States; [0000-0002-3305-852X](https://orcid.org/0000-0002-3305-852X)  
746 Jeffrey J. Urban — Molecular Foundry, Lawrence Berkeley  
747 National Laboratory, Berkeley, California 94720, United  
748 States; [0000-0002-6520-830X](https://orcid.org/0000-0002-6520-830X)  
749 Wissam A. Saidi — Department of Mechanical Engineering and  
750 Materials Science, University of Pittsburgh, Pittsburgh,  
751 Pennsylvania 15261, United States; [0000-0001-6714-4832](https://orcid.org/0000-0001-6714-4832)  
753 Joshua Schrier — Department of Chemistry, Fordham  
754 University, New York 10458, United States; [0000-0002-2071-1657](https://orcid.org/0000-0002-2071-1657)  
756 Alexander J. Norquist — Department of Chemistry, Haverford  
757 College, Haverford, Pennsylvania 19041, United States

758 Complete contact information is available at:  
759 <https://pubs.acs.org/10.1021/acs.chemmater.1c03564>

## 760 Notes

761 The authors declare no competing financial interest.  
762 The CIF files of morphPbI<sub>3</sub> and (morph)<sub>2</sub>PbI<sub>4</sub> crystal  
763 structures have been submitted to the Cambridge Crystallo-  
764 graphic Data Centre under Deposition Numbers 2110021 and  
765 2110022, respectively.

## 766 ■ ACKNOWLEDGMENTS

767 This study is based upon work supported by the Defense  
768 Advanced Research Projects Agency (DARPA) under Contract  
769 No. HR001118C0036. Any opinions, findings, and conclusions  
770 or recommendations expressed in this material are those of the  
771 authors and do not necessarily reflect the views of DARPA.  
772 Work at the Molecular Foundry was supported by the Office of  
773 Science, Office of Basic Energy Sciences, of the U.S.  
774 Department of Energy under Contract No. DE-AC02-  
775 05CH11231. J.S. acknowledges the Henry Dreyfus Teacher-  
776 Scholar Award (TH-14-010) and the resources of the  
777 MERCURY consortium (<http://mercuryconsortium.org/>)  
778 under NSF Grant No. CNS-2018427. M.Z. acknowledges the  
779 National Science Foundation, Major Research Instrumentation  
780 Program under Grant No. CHE 1625543, for funding for the  
781 single crystal X-ray diffractometer. W.A.S. acknowledges the  
782 support from the U. S. National Science Foundation (Award  
783 No. CSSI-2003808). Also, we are grateful for computing time  
784 provided in part by the Pittsburgh Center for Research  
785 Computing (CRC) resources at the University of Pittsburgh.

## 786 ■ REFERENCES

- 787 (1) Herz, L. M. Charge-Carrier Mobilities in Metal Halide  
788 Perovskites: Fundamental Mechanisms and Limits. *ACS Energy Lett.*  
789 2017, 2, 1539–1548.  
790 (2) Zhao, Y.; Zhu, K. Organic–Inorganic Hybrid Lead Halide  
791 Perovskites for Optoelectronic and Electronic Applications. *Chem.*  
792 *Soc. Rev.* 2016, 45, 655–689.

- (3) Takahashi, Y.; Obara, R.; Nakagawa, K.; Nakano, M.; Tokita, J.; Inabe, T. Tunable Charge Transport in Soluble Organic–Inorganic Hybrid Semiconductors. *Chem. Mater.* 2007, 19, 6312–6316.
- (4) Leveillee, J.; Katan, C.; Even, J.; Ghosh, D.; Nie, W.; Mohite, A. D.; Tretiak, S.; Schleife, A.; Neukirch, A. J. Tuning Electronic Structure in Layered Hybrid Perovskites with Organic Spacer Substitution. *Nano Lett.* 2019, 19, 8732–8740.
- (5) Correa-Baena, J.-P.; Abate, A.; Saliba, M.; Tress, W.; Jacobsson, T. J.; Grätzel, M.; Hagfeldt, A. The Rapid Evolution of Highly Efficient Perovskite Solar Cells. *Energy Environ. Sci.* 2017, 10, 710–727.
- (6) Correa-Baena, J.-P.; Saliba, M.; Buonassisi, T.; Grätzel, M.; Abate, A.; Tress, W.; Hagfeldt, A. Promises and Challenges of Perovskite Solar Cells. *Science* 2017, 358, 739–744.
- (7) Yang, D.; Luo, L.; Gao, Y.; Chen, S.; Zeng, X. C. Rational Design of One-Dimensional Hybrid Organic–Inorganic Perovskites with Room-Temperature Ferroelectricity and Strong Piezoelectricity. *Mater. Horiz.* 2019, 6, 1463–1473.
- (8) Wang, H.-C.; Bao, Z.; Tsai, H.-Y.; Tang, A.-C.; Liu, R.-S. Perovskite Quantum Dots and Their Application in Light-Emitting Diodes. *Small* 2018, 14, 1702433–1702455.
- (9) Veldhuis, S. A.; Boix, P. P.; Yantara, N.; Li, M.; Sum, T. C.; Mathews, N.; Mhaisalkar, S. G. Perovskite Materials for Light-Emitting Diodes and Lasers. *Adv. Mater.* 2016, 28, 6804–6834.
- (10) Kao, T. S.; Hong, Y.-H.; Hong, K.-B.; Lu, T.-C. Perovskite Random Lasers: A Tunable Coherent Light Source for Emerging Applications. *Nanotechnology* 2021, 32, 282001.
- (11) Seth, C.; Jana, D.; Jindal, V.; Khushalani, D.; Ghosh, S. One-Dimensional Behavior of Imidazolium Lead Iodide. *J. Phys. Chem. C* 2019, 123, 16449–16455.
- (12) Stoumpos, C. C.; Cao, D. H.; Clark, D. J.; Young, J.; Rondinelli, J. M.; Jang, J. I.; Hupp, J. T.; Kanatzidis, M. G. Ruddlesden–Popper Hybrid Lead Iodide Perovskite 2D Homologous Semiconductors. *Chem. Mater.* 2016, 28, 2852–2867.
- (13) Poglitsch, A.; Weber, D. Dynamic Disorder in Methylammoniumtrihalogenoplumbates (II) Observed by Millimeter-wave Spectroscopy. *J. Chem. Phys.* 1987, 87, 6373–6378.
- (14) Gao, P.; Bin Mohd Yusoff, A. R.; Nazeeruddin, M. K. Dimensionality Engineering of Hybrid Halide Perovskite Light Absorbers. *Nat. Commun.* 2018, 9, 5028.
- (15) Hong, K.; Le, Q. V.; Kim, S. Y.; Jang, H. W. Low-Dimensional Halide Perovskites: Review and Issues. *J. Mater. Chem. C* 2018, 6, 2189–2209.
- (16) Yusoff, A. R. B. M.; Nazeeruddin, M. K. Low-Dimensional Perovskites: From Synthesis to Stability in Perovskite Solar Cells. *Adv. Energy Mater.* 2018, 8, No. 1702073.
- (17) Kieslich, G.; Sun, S.; Cheetham, A. K. Solid-State Principles Applied to Organic–Inorganic Perovskites: New Tricks for an Old Dog. *Chem. Sci.* 2014, 5, 4712–4715.
- (18) Billing, D. G.; Lemmerer, A. Synthesis, Characterization and Phase Transitions in the Inorganic–Organic Layered Perovskite-Type Hybrids [(CnH2n+1NH3)2PbI4], n = 4, 5 and 6. *Acta Crystallogr. B* 2007, 63, 735–747.
- (19) Billing, D. G.; Lemmerer, A. Synthesis, Characterization and Phase Transitions of the Inorganic–Organic Layered Perovskite-Type Hybrids [(CnH2n+1NH3)2PbI4] (n = 12, 14, 16 and 18). *New J. Chem.* 2008, 32, 1736–1746.
- (20) Saparov, B.; Mitzi, D. B. Organic–Inorganic Perovskites: Structural Versatility for Functional Materials Design. *Chem. Rev.* 2016, 116, 4558–4596.
- (21) Billing, D. G.; Lemmerer, A. Inorganic–Organic Hybrid Materials Incorporating Primary Cyclic Ammonium Cations: The Lead Iodide Series. *CrystEngComm* 2007, 9, 236–244.
- (22) Billing, D. G.; Lemmerer, A. Inorganic–Organic Hybrid Materials Incorporating Primary Cyclic Ammonium Cations: The Lead Bromide and Chloride Series. *CrystEngComm* 2009, 11, 1549–1562.
- (23) Chen, X.-G.; Song, X.-J.; Zhang, Z.-X.; Zhang, H.-Y.; Pan, Q.; Yao, J.; You, Y.-M.; Xiong, R.-G. Confinement-Driven Ferroelectricity

- 862 in a Two-Dimensional Hybrid Lead Iodide Perovskite. *J. Am. Chem.*  
863 *Soc.* **2020**, *142*, 10212.
- 864 (24) Fateev, S. A.; Petrov, A. A.; Marchenko, E. I.; Zubavichus, Y. V.;  
865 Khrustalev, V. N.; Petrov, A. V.; Aksenov, S. M.; Goodilin, E. A.;  
866 Tarasov, A. B. FA2PbBr4: Synthesis, Structure, and Unusual Optical  
867 Properties of Two Polymorphs of Formamidinium-Based Layered  
868 (110) Hybrid Perovskite. *Chem. Mater.* **2021**, *33*, 1900–1907.
- 869 (25) Huang, J.; Yuan, Y.; Shao, Y.; Yan, Y. Understanding the  
870 Physical Properties of Hybrid Perovskites for Photovoltaic Applica-  
871 tions. *Nat. Rev. Mater.* **2017**, *2*, 17042.
- 872 (26) Stach, E.; DeCost, B.; Kusne, A. G.; Hattrick-Simpers, J.;  
873 Brown, K. A.; Reyes, K. G.; Schrier, J.; Billinge, S.; Buonassisi, T.;  
874 Foster, I.; Gomes, C. P.; Gregoire, J. M.; Mehta, A.; Montoya, J.;  
875 Olivetti, E.; Park, C.; Rotenberg, E.; Saikin, S. K.; Smullin, S.; Staney,  
876 V.; Maruyama, B. Autonomous Experimentation Systems for  
877 Materials Development: A Community Perspective. *Matter* **2021**, *4*,  
878 2702–2726.
- 879 (27) Liu, Y.; Niu, C.; Wang, Z.; Gan, Y.; Zhu, Y.; Sun, S.; Shen, T.  
880 Machine Learning in Materials Genome Initiative: A Review. *J. Mater.*  
881 *Sci. Technol.* **2020**, *57*, 113.
- 882 (28) Li, Z.; Najeeb, M. A.; Alves, L.; Sherman, A. Z.; Shekar, V.;  
883 Cruz Parrilla, P.; Pendleton, I. M.; Wang, W.; Nega, P. W.; Zeller, M.;  
884 Schrier, J.; Norquist, A. J.; Chan, E. M. Robot-Accelerated Perovskite  
885 Investigation and Discovery. *Chem. Mater.* **2020**, *32*, 5650–5663.
- 886 (29) Lookman, T.; Balachandran, P. V.; Xue, D.; Yuan, R. Active  
887 Learning in Materials Science with Emphasis on Adaptive Sampling  
888 Using Uncertainties for Targeted Design. *npj Comput. Mater.* **2019**, *5*,  
889 1–17.
- 890 (30) del Rosario, Z.; Rupp, M.; Kim, Y.; Antono, E.; Ling, J.  
891 Assessing the Frontier: Active Learning, Model Accuracy, and Multi-  
892 Objective Candidate Discovery and Optimization. *J. Chem. Phys.*  
893 **2020**, *153*, No. 024112.
- 894 (31) Dai, C.; Glotzer, S. C. Efficient Phase Diagram Sampling by  
895 Active Learning. *J. Phys. Chem. B* **2020**, *124*, 1275–1284.
- 896 (32) Zhao, S.; Cai, T.; Zhang, L.; Li, W.; Lin, J. Autonomous  
897 Construction of Phase Diagrams of Block Copolymers by Theory-  
898 Assisted Active Machine Learning. *ACS Macro Lett.* **2021**, *10*, 598–  
899 602.
- 900 (33) Tian, Y.; Yuan, R.; Xue, D.; Zhou, Y.; Wang, Y.; Ding, X.; Sun,  
901 J.; Lookman, T. Determining Multi-Component Phase Diagrams with  
902 Desired Characteristics Using Active Learning. *Adv. Sci.* **2021**, *8*,  
903 No. 2003165.
- 904 (34) Balachandran, P. V.; Kowalski, B.; Sehirlioglu, A.; Lookman, T.  
905 Experimental Search for High-Temperature Ferroelectric Perovskites  
906 Guided by Two-Step Machine Learning. *Nat. Commun.* **2018**, *9*, 1668.
- 907 (35) Xue, D.; Balachandran, P. V.; Yuan, R.; Hu, T.; Qian, X.;  
908 Dougherty, E. R.; Lookman, T. Accelerated Search for BaTiO3-Based  
909 Piezoelectrics with Vertical Morphotropic Phase Boundary Using  
910 Bayesian Learning. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 13301–  
911 13306.
- 912 (36) Kusne, A. G.; Yu, H.; Wu, C.; Zhang, H.; Hattrick-Simpers, J.;  
913 DeCost, B.; Sarker, S.; Oses, C.; Toher, C.; Curtarolo, S.; Davydov, A.;  
914 V.; Agarwal, R.; Bendersky, L. A.; Li, M.; Mehta, A.; Takeuchi, I. On-  
915 the-Fly Closed-Loop Materials Discovery via Bayesian Active  
916 Learning. *Nat. Commun.* **2020**, *11*, 5966.
- 917 (37) Rohr, B.; Stein, H. S.; Guevarra, D.; Wang, Y.; Haber, J. A.;  
918 Aykol, M.; Suram, S. K.; Gregoire, J. M. Benchmarking the  
919 Acceleration of Materials Discovery by Sequential Learning. *Chem.*  
920 *Sci.* **2020**, *11*, 2696–2706.
- 921 (38) Sun, S.; Tiilinen, A.; Oviedo, F.; Liu, Z.; Thapa, J.; Zhao, Y.;  
922 Hartono, N. T. P.; Goyal, A.; Heumueller, T.; Batali, C.; Encinas, A.;  
923 Yoo, J. J.; Li, R.; Ren, Z.; Peters, I. M.; Brabec, C. J.; Bawendi, M. G.;  
924 Stevanovic, V.; Fisher, J.; Buonassisi, T. A Data Fusion Approach to  
925 Optimize Compositional Stability of Halide Perovskites. *Matter* **2021**,  
926 *4*, 1305–1322.
- 927 (39) Dahl, J. C.; Wang, X.; Huang, X.; Chan, E. M.; Alivisatos, A. P.  
928 Elucidating the Weakly Reversible Cs-Pb-Br Perovskite Nanocrystal  
929 Reaction Network with High-Throughput Maps and Transformations.  
930 *J. Am. Chem. Soc.* **2020**, *142*, 11915–11926.
- 931 (40) Hino, H. Active Learning: Problem Settings and Recent  
932 Developments, 2020. arXiv: 2012.04225 (Computer Science and  
Machine Learning). <https://arxiv.org/abs/2012.04225>.
- 933 (41) Ueno, T.; Ishibashi, H.; Hino, H.; Ono, K. Automated  
934 Stopping Criterion for Spectral Measurements with Active Learning.  
*npj Comput. Mater.* **2021**, *7*, 1–9.
- 935 (42) Ramalingam, K.; Rajaraman, T. Metal – Organic Hybrids of  
937 Tin(IV) with Tuneable Band Gap: Synthesis, Spectral, Single Crystal  
938 X-Ray Structural, BVS and CSM Analysis of Morpholinium  
939 Hexahalostannate(IV). *J. Mol. Struct.* **2020**, *1218*, No. 128489.
- 940 (43) Xiong, K.; Liu, W.; Teat, S. J.; An, L.; Wang, H.; Emge, T. J.; Li,  
J. New Hybrid Lead Iodides: From One-Dimensional Chain to Two-  
942 Dimensional Layered Perovskite Structure. *J. Solid State Chem.* **2015**,  
943 *230*, 143–148.
- 944 (44) Owczarek, M.; Jakubas, R.; Pietraszko, A.; Medycki, W.; Baran,  
J. Investigation of Structure–Properties Relationship in a Novel  
946 Family of Halogenoantimonates(III) and Halogenobismuthates(III)  
947 with Morpholinium Cation: [NH2(C2H4)2O]MX4. Crystal Struc-  
948 ture, Phase Transitions and Dynamics of Molecules. *Dalton Trans.* **2013**,  
949 *42*, 15069–15079.
- 950 (45) Owczarek, M.; Szklarz, P.; Jakubas, R.; Miniewicz, A.  
951 [NH2(C2H4)2O]MX5: A New Family of Morpholinium Nonlinear  
952 Optical Materials among Halogenoantimonate(III) and  
953 Halogenobismuthate(III) Compounds. Structural Characterization,  
954 Dielectric and Piezoelectric Properties. *Dalton Trans.* **2012**, *41*,  
955 7285–7294.
- 956 (46) Kawasaki, T.; Takahashi, M.; Ohhara, T.; Tanaka, I.; Kusaka,  
K.; Hosoya, T.; Yamada, T.; Kurihara, K. Structure of Morpholinium  
958 Tribromoplumbate C4H8ONH2PbBr3 Studied Using Single-Crystal  
959 Neutron Diffraction. *J. Phys. Soc. Jpn.* **2012**, *81*, No. 094602.
- 960 (47) Li, H.-H.; Chen, Z.-R.; Cheng, L.-C.; Wang, Y.-J.; Feng, M.;  
961 Wang, M. Hybrid Polymeric Iodoplumbates Constructed from  
962 Morpholine and Its Derivatives: Structures and Properties. *Dalton*  
963 *Trans.* **2010**, *39*, 11000–11007.
- 964 (48) Li, H.-H.; Chen, Z.-R.; Cheng, L.-C.; Liu, J.-B.; Chen, X.-B.; Li,  
J.-Q. A New Hybrid Optical Semiconductor Based on Polymeric  
966 Iodoplumbate Co-Templated by Both Organic Cation and Polyiodide  
967 Anion. *Cryst. Growth Des.* **2008**, *8*, 4355–4358.
- 968 (49) Corradi, A. B.; Bruni, S.; Cariati, F.; Ferrari, A. M.; Saccani, A.;  
969 Sandrolini, F.; Sgarabotto, P. Organic/Inorganic Composite Materi-  
970 als: Synthesis and Properties of One-Dimensional Polymeric  
971 Haloplumbate(II) Systems. *Inorg. Chim. Acta* **1997**, *254*, 137–143.
- 972 (50) Yue, H.-L.; Sung, H.-H.; Chen, F.-C. Seeded Space-Limited  
973 Crystallization of CH3NH3PbI3 Single-Crystal Plates for Perovskite  
974 Solar Cells. *Adv. Electron. Mater.* **2018**, *4*, No. 1700655.
- 975 (51) Shi, C.; Ye, L.; Gong, Z.-X.; Ma, J.-J.; Wang, Q.-W.; Jiang, J.-Y.;  
976 Hua, M.-M.; Wang, C.-F.; Yu, H.; Zhang, Y.; Ye, H.-Y. Two-  
977 Dimensional Organic–Inorganic Hybrid Rare-Earth Double Perov-  
978 skite Ferroelectrics. *J. Am. Chem. Soc.* **2019**, *142*, 545–551.
- 979 (52) Saidaminov, M. I.; Abdelhady, A. L.; Murali, B.; Alarousu, E.;  
980 Burlakov, V. M.; Peng, W.; Dursun, I.; Wang, L.; He, Y.; Maculan, G.;  
981 Gorilev, A.; Wu, T.; Mohammed, O. F.; Bakr, O. M. High-Quality  
982 Bulk Hybrid Perovskite Single Crystals within Minutes by Inverse  
983 Temperature Crystallization. *Nat. Commun.* **2015**, *6*, 7586–7592.
- 984 (53) Shi, D.; Adinolfi, V.; Comin, R.; Yuan, M.; Alarousu, E.; Buin,  
985 A.; Chen, Y.; Hoogland, S.; Rothenberger, A.; Katsiev, K.; Losovyj, Y.;  
986 Zhang, X.; Dowben, P. A.; Mohammed, O. F.; Sargent, E. H.; Bakr, O.  
987 M. Low Trap-State Density and Long Carrier Diffusion in Organolead  
988 Trihalide Perovskite Single Crystals. *Science* **2015**, *347*, 519–522.
- 989 (54) Ferrando, A.; Martínez Pastor, J. P.; Suárez, I. Toward Metal  
990 Halide Perovskite Nonlinear Photonics. *J. Phys. Chem. Lett.* **2018**, *9*,  
991 5612–5623.
- 992 (55) Park, H.; Ha, C.; Lee, J.-H. Advances in Piezoelectric Halide  
993 Perovskites for Energy Harvesting Applications. *J. Mater. Chem. A*  
994 **2020**, *8*, 24353–24367.
- 995 (56) Newnham, R. *Properties of Materials: Anisotropy, Symmetry,*  
996 *Structure*; Oxford University Press: New York, 2005.
- 997 (57) Li, Z. multi-well crystallization block by ZhiLBNL. <https://www.thingiverse.com/thing:4927375> (accessed September 20, 2021).

- 1000 (58) Kennard, R. W.; Stone, L. A. Computer Aided Design of  
1001 Experiments. *Technometrics* **1969**, *11*, 137–148.
- 1002 (59) Xie, J. Karoka/Kennard-Stone-Algorithm. [https://github.com/  
1003 karoka/Kennard-Stone-Algorithm](https://github.com/karoka/Kennard-Stone-Algorithm) (accessed September 20, 2021).
- 1004 (60) Pendleton, I. M.; Cattabriga, G.; Li, Z.; Najeeb, M. A.; Friedler,  
1005 S. A.; Norquist, A. J.; Chan, E. M.; Schrier, J. Experiment  
1006 Specification, Capture and Laboratory Automation Technology  
1007 (ESCALATE): A Software Pipeline for Automated Chemical  
1008 Experimentation and Data Management. *MRS Commun.* **2019**, *9*,  
1009 846–859.
- 1010 (61) Yangui, A.; Garrot, D.; Lauret, J. S.; Lusson, A.; Bouchez, G.;  
1011 Deleporte, E.; Pillet, S.; Bendeif, E. E.; Castro, M.; Triki, S.; Abid, Y.;  
1012 Boukheddaden, K. Optical Investigation of Broadband White-Light  
1013 Emission in Self-Assembled Organic–Inorganic Perovskite  
1014 (C<sub>6</sub>H<sub>11</sub>NH<sub>3</sub>)<sub>2</sub>PbBr<sub>4</sub>. *J. Phys. Chem. C* **2015**, *119*, 23638–23647.
- 1015 (62) Zhou, Y.; Game, O. S.; Pang, S.; Padture, N. P. Microstructures  
1016 of Organometal Trihalide Perovskites for Solar Cells: Their Evolution  
1017 from Solutions and Characterization. *J. Phys. Chem. Lett.* **2015**, *6*,  
1018 4827–4839.
- 1019 (63) Gu, E.; Tang, X.; Langner, S.; Duchstein, P.; Zhao, Y.; Levchuk,  
1020 I.; Kalancha, V.; Stubhan, T.; Hauch, J.; Egelhaaf, H. J.; Zahn, D.;  
1021 Osvet, A.; Brabec, C. J. Robot-Based High-Throughput Screening of  
1022 Antisolvents for Lead Halide Perovskites. *Joule* **2020**, *4*, 1806–1822.
- 1023 (64) Islam, T.; Sarker, M. Z. I.; Uddin, A. H.; Yunus, K. B.; Reddy  
1024 Prasad, D. M.; Mia, M. A. R.; Ferdosh, S. Kamlet Taft Parameters: A  
1025 Tool to Alternate the Usage of Hazardous Solvent in Pharmaceutical  
1026 and Chemical Manufacturing/Synthesis – A Gateway towards Green  
1027 Technology. *Anal. Chem. Lett.* **2020**, *10*, 550–561.
- 1028 (65) Nega, P. W.; Li, Z.; Ghosh, V.; Thapa, J.; Sun, S.; Hartono, N.  
1029 T. P.; Nelliikkal, M. A. N.; Norquist, A. J.; Buonassisi, T.; Chan, E. M.;  
1030 Schrier, J. Using Automated Serendipity to Discover How Trace  
1031 Water Promotes and Inhibits Lead Halide Perovskite Crystal  
1032 Formation. *Appl. Phys. Lett.* **2021**, *119*, No. 041903.
- 1033 (66) Kamlet-Taft solvent parameters. [http://www.stenutz.eu/chem/  
1034 solv26.php](http://www.stenutz.eu/chem/solv26.php) (accessed May 17, 2021).
- 1035 (67) Settles, B. *Active Learning Literature Survey*. University of  
1036 Wisconsin–Madison, 2010, Vol. 67.
- 1037 (68) Hinton, G.; Roweis, S. Stochastic Neighbor Embedding.  
1038 *NIPS'02: Proceedings of the 15th International Conference on Neural  
1039 Information Processing Systems*; MIT Press: Cambridge, MA, United  
1040 States, 2002; pp 857–864.
- 1041 (69) Zhdanov, F. Diverse Mini-Batch Active Learning, 2019. arXiv:  
1042 1901.05954 (Computer Science and Machine Learning). <https://arxiv.org/abs/1901.05954>.
- 1044 (70) MacQueen, J. Some Methods for Classification and Analysis of  
1045 Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium  
1046 on Mathematical Statistics and Probability, Volume 1: Statistics*;  
1047 University of California Press, 1967; pp 281–297.
- 1048 (71) Vlachos, A. A Stopping Criterion for Active Learning. *Comput.  
1049 Speech Lang.* **2008**, *22*, 295–312.
- 1050 (72) Zhu, J.; Wang, H.; Hovy, E.; Ma, M. Confidence-Based  
1051 Stopping Criteria for Active Learning for Data Annotation. *ACM  
1052 Trans. Speech Lang. Process.* **2010**, *6*, 1–24.
- 1053 (73) Yang, Y.; Loog, M. A Variance Maximization Criterion for  
1054 Active Learning. *Pattern Recognit.* **2018**, *78*, 358–370.
- 1055 (74) Cortes, C.; Jackel, L. D.; Chiang, W.-P. Limits on Learning  
1056 Machine Accuracy Imposed by Data Quality. *KDD'95: Proceedings of  
1057 the First International Conference on Knowledge Discovery and Data  
1058 Mining*; AAAI Press, 1995; pp 57–62.
- 1059 (75) 4.2. Permutation feature importance — scikit-learn 0.24.2  
1060 documentation. [https://scikit-learn.org/stable/modules/  
1061 permutation\\_importance.html#outline-of-the-permutation-importance-algorithm](https://scikit-learn.org/stable/modules/permutation_importance.html#outline-of-the-permutation-importance-algorithm) (accessed June 3, 2021).
- 1063 (76) Efron, B.; Tibshirani, R. Bootstrap Methods for Standard  
1064 Errors, Confidence Intervals, and Other Measures of Statistical  
1065 Accuracy. *Stat. Sci.* **1986**, *1*, 54–75.
- 1066 (77) Lehmann, E. L.; Romano, J. P. *Testing Statistical Hypotheses*, 3rd  
1067 ed.; Springer Texts in Statistics; Springer-Verlag: New York, 2005,  
1068 DOI: [10.1007/0-387-27605-X](https://doi.org/10.1007/0-387-27605-X).
- 1078 (78) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; 1069  
Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the 1070  
Density-Gradient Expansion for Exchange in Solids and Surfaces. 1071  
*Phys. Rev. Lett.* **2008**, *100*, No. 136406. 1072
- 1079 (79) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; 1073  
Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials 1074  
Database (OQMD): Assessing the Accuracy of DFT Formation 1075  
Energies. *npj Comput. Mater.* **2015**, *1*, 1–15. 1076
- 1080 (80) Do, J.-L.; Friščić, T. Mechanochemistry: A Force of Synthesis. 1077  
*ACS Cent. Sci.* **2017**, *3*, 13–19. 1078
- 1081 (81) Guo, Y.; Shoyama, K.; Sato, W.; Matsuo, Y.; Inoue, K.; Harano, 1079  
K.; Liu, C.; Tanaka, H.; Nakamura, E. Chemical Pathways Connecting 1080  
Lead(II) Iodide and Perovskite via Polymeric Plumbate(II) Fiber. *J.* 1081  
*Am. Chem. Soc.* **2015**, *137*, 15907–15914. 1082
- 1082 (82) Zhang, K.; Wang, Z.; Wang, G.; Wang, J.; Li, Y.; Qian, W.; 1083  
Zheng, S.; Xiao, S.; Yang, S. A Prenucleation Strategy for Ambient 1084  
Fabrication of Perovskite Solar Cells with High Device Performance 1085  
Uniformity. *Nat. Commun.* **2020**, *11*, 1006. 1086
- 1083 (83) Xiao, S.; Zhang, K.; Zheng, S.; Yang, S. Good or Evil: What Is 1087  
the Role of Water in Crystallization of Organometal Halide 1088  
Perovskites? *Nanoscale Horiz.* **2020**, *5*, 1147–1154. 1089
- 1089 (84) Natale, D.; Mareque-Rivas, J. C. The Combination of 1090  
Transition Metal Ions and Hydrogen-Bonding Interactions. *Chem.* 1091  
*Commun.* **2008**, *4*, 425–437. 1092