

## Article

# Design Principles of Large Cation Incorporation in Halide Perovskites

Heesoo Park <sup>1</sup> , Syam Kumar <sup>2</sup> , Sanjay Chawla <sup>3</sup> and Fedwa El-Mellouhi <sup>1,\*</sup> 

<sup>1</sup> Qatar Environment and Energy Research Institute, Hamad Bin Khalifa University, Doha P.O. Box 34110, Qatar; hpark@hbku.edu.qa

<sup>2</sup> Department of Health and Nutritional Science, Institute of Technology Sligo, Ash Lane, F91 YW50 Sligo, Ireland; Kumar.Syam@itsligo.ie

<sup>3</sup> Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha P.O. Box 34110, Qatar; schawla@hbku.edu.qa

\* Correspondence: felmellouhi@hbku.edu.qa

**Abstract:** Perovskites have stood out as excellent photoactive materials with high efficiencies and stabilities, achieved via cation mixing techniques. Overcoming challenges to the stabilization of Perovskite solar cells calls for the development of design principles of large cation incorporation in halide perovskite to accelerate the discovery of optimal stable compositions. Large fluorinated organic cations incorporation is an attractive method for enhancing the intrinsic stability of halide perovskites due to their high dipole moment and moisture-resistant nature. However, a fluorinated cation has a larger ionic size than its non-fluorinated counterpart, falling within the upper boundary of the mixed-cation incorporation. Here, we report on the intrinsic stability of mixed Methylammonium (MA) lead halides at different concentrations of large cation incorporation, namely, ethylammonium (EA;  $[\text{CH}_3\text{CH}_2\text{NH}_3]^+$ ) and 2-fluoroethylammonium (FEA;  $[\text{CH}_2\text{FCH}_2\text{NH}_3]^+$ ). Density functional theory (DFT) calculations of the enthalpy of the mixing and analysis of the perovskite structural features enable us to narrow down the compositional search domain for EA and FEA cations around concentrations that preserve the perovskite structure while pointing towards the maximal stability. This work paves the way to developing design principles of a large cation mixture guided by data analysis of DFT data. Finally, we present the automated search of the minimum enthalpy of mixing by implementing Bayesian optimization over the compositional search domain. We introduce and validate an automated workflow designed to accelerate the compositional search, enabling researchers to cut down the computational expense and bias to search for optimal compositions.

**Keywords:** halide perovskites; photocatalyst; density functional theory; large cation; fluorinated cation; octahedral deformation; mixed cation; non-covalent interaction; Bayesian optimization



**Citation:** Park, H.; Kumar, S.; Chawla, S.; El-Mellouhi, F. Design Principles of Large Cation Incorporation in Halide Perovskites. *Molecules* **2021**, *26*, 6184. <https://doi.org/10.3390/molecules26206184>

Academic Editors: Federico Bella and Shijing Sun

Received: 2 September 2021

Accepted: 24 September 2021

Published: 13 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The majority of the world's population is moving towards green energy technologies, out of which solar energy is considered the most suitable form to harness. Over the last two decades, the family of materials known as Halide Perovskites have stood out as being excellent photoactive materials, thanks to their exceptional formability and bandgap tunability [1].

The perovskite family of materials offers a huge opportunity for the compositional and structural tuning that opens the possibility of discovering high performance materials for energy conversion. Halide perovskites ( $ABX_3$ ) consist of metal atoms or organic molecules at the *A*-site; Ge, Pb, and Sn at the *B*-site; and I, Br, and Cl at the *X*-site [2,3]. Methylammonium lead halide ( $\text{CH}_3\text{NH}_3\text{PbI}_3$ ) is the mainstay of the perovskite family due to its bandgap of 1.5 eV and its facile fabrication by low-temperature solution processing. The baseline compound for numerous mixed halide perovskites reported in the past few years has been optimized to obtain better stabilities and photoconversion efficiencies [4].

The use of perovskites and their derivatives in the field of photovoltaics has been intensely researched ever since its discovery, and a vast number of articles and reviews are already available, starting from their fundamentals up to photovoltaic module physics [5–7].

Several of the materials that belong to this family have shown power conversion efficiencies (PCE) up to 25.5% [8] due to an exceptional combination of properties suitable for photoconversion such as (i) bandgaps, (ii) charge carrier mobilities and diffusion lengths, (iii) lower surface recombination rates, (iv) high optical absorption, (v) structural defect tolerance and (vi) defect characteristics. However, halide perovskites exhibit long-term performance constraints [9–11] such as degradation under thermal treatment, anomalous photocurrent with voltage (J-V) hysteresis and susceptibility to hydrolysis. These constraints stand as the bottleneck against the large-scale deployment of these materials in the photovoltaics market [12].

In addition to photovoltaic applications, Sunghak et al. recently demonstrated the splitting of the aqueous HI solution into H<sub>2</sub> and I<sub>3</sub><sup>−</sup> with methylammonium lead halide perovskite under solar irradiation [13]. Halide perovskites have been reported for visible-light water-splitting and organic pollutant degradation such as CO<sub>2</sub> reduction, which can be collectively classified under photocatalysis [14–17]. Their suitable bandgap and band edge positions can activate a wide range of redox reactions since they can be tuned via compositional optimization [17]. The relative positions of the conduction band (CB) generally determine the reduction ability for H<sub>2</sub> generation and CO<sub>2</sub> reduction; while those of the valence band (VB) are critical for water oxidation, allowing H<sub>2</sub>O splitting. In addition to the above suitable light absorption properties, halide perovskites possess properties that enhance their photocatalytic activity, such as charge migration/stability, and the combination of their ferroelectric and piezoelectric effects [3,5,18].

Despite the perovskites stability issues under visible light irradiation and moisture conditions during photocatalytic device operation [18], there have been recent reports on water-stable halide perovskites that are useful for H<sub>2</sub> production from direct water splitting. Ju et al. showed a single crystal of [(CH<sub>3</sub>)<sub>2</sub>NH<sub>2</sub>]<sub>2</sub>SnI<sub>3</sub> is stable in pure water for up to 16 hours and does not dissolve [19]. Pisanu et al. uncovered that the bromide analogue ([CH<sub>3</sub>)<sub>2</sub>NH<sub>2</sub>]<sub>2</sub>SnBr<sub>3</sub> has a similar stability while generating H<sub>2</sub> from deionized water without dissolving [20]. Although this material formed an opalescent solution in the deionized water, the authors could recover the pristine phase perovskite as soon as the material was dried. Alternatively, as proposed by Ding et al., the nanocrystalline (NC) glass formation can improve the stability of the inorganic counterparts such as CsPbBr<sub>3</sub> [21]. The formation of NC glass solves the rapid materials deactivation in water by fostering the encapsulation effect.

In APbI<sub>3</sub> perovskites, the contribution to band edge states originates from the PbI<sub>6</sub> octahedra, where the Pb–I bond lengths and angles determine the orbital overlap that modulates the positions of the CB and VB [4]. For instance, the replacement of an MA cation with a larger FA cation results in the reduction of bandgap by 0.05 eV while the replacement with a smaller Cs ion increases the bandgap by 0.16 eV [3,22]. Hence, A-site cations contribute very weakly to the band edges but strongly influence the perovskites phase stability due to their size and dynamics [23,24]. A common approach to remaining within the bandgap range of pristine lead halide perovskites has been the partial incorporation of the cations. Interestingly, there are recent reports that demonstrate a promising route of the commercialization of halide perovskite solar cells light absorbers based on formamidinium (FA) and a mixture of other cations [25–27]. In addition to MA and FA, there are other commonly used monovalent cations such as ethylammonium (EA), guanidinium (GUA), hydrazinium (HZ), and hydroxylammonium (HA) [28,29]. Furthermore, sulfonium-based cations, such as trimethylsulfonium and trimethylsulfoxonium, have shown the formation of lower dimensional perovskites resulting in a stability increase, as their ionic size is bulkier than MA [30,31].

A mixture of the above cations were found to positively affect the dynamics of the organic molecule, thus offering the possibility of stabilizing the perovskite structure due to entropic effects and by minimizing the internal energy [32].

Alongside these aliphatic organic cations used in perovskites, Huang et al. reported the successful use of a halogenated-methylammonium cation by synthesizing the halide perovskite,  $\text{CH}_2\text{FNH}_3\text{PbBr}_3$  [33]. Interestingly, the resulting crystal structure of the synthesized perovskite material preserved a 3-dimensional crystal structure (3D) while having a slightly larger bandgap (2.3 eV) than  $\text{MAPbBr}_3$  (2.2 eV) with the enhanced absorption coefficient and smaller exciton binding energies compared to MA-based compounds. Interestingly, due to the high electronegativity of fluorine, fluorinated cations can introduce strong ionic and intermolecular bonding within the perovskite material. They are also helpful for passivating surfaces, thereby suppressing their chemical reactivity in the presence of moisture or contact with water [34]. The lattice distortions induced by fluorinated cations, which are usually even larger than MA, result in a broad and robust photoluminescence spectrum upon excitations of samples with light sources at frequencies below the bandgap [35]. Furthermore, recent theoretical exploration suggested that partial substitution of MA with Fluorinated cations induces not only enhanced intrinsic stability but could also boost the device stability under applied bias during operation due to the suppression of ionic migration [24,35,36].

Fluorinated organic cation's partial or total substitution is increasingly being reported as an efficient way to tune the properties of the perovskite and enhance their stability [34,37] by strengthening some of the weak hydrogen bonds between the organic cations and the surrounding  $[\text{PbI}_3]$  framework [36].

Nevertheless, it is important to find the critical balance between the ionic size and concentrations of large cation incorporation. The larger cations in a mixed-cation perovskite can suppress the hysteresis and aging of the device while reducing the bandgap [38]. Besides affecting the size of the unit cell, the organic cation also helps to exert a screening effect of the moving ionic charges [39]. The Pb-I-Pb angles are severely affected by the strong steric and electrostatic interactions between the organic cation and the Pb-I lattice [29]. Furthermore, Garoufalidis et al. pointed out the importance of the orientational disorder of the molecular cation by looking into their impact on the perovskite nanoparticles, as the orientational disorder influences the structural and electronic properties [40,41]. Though the fine-tuning of the parent perovskite stands as a never-ending endeavor, a thorough study of their structure–property relation seems to be missing in majority of the reports available.

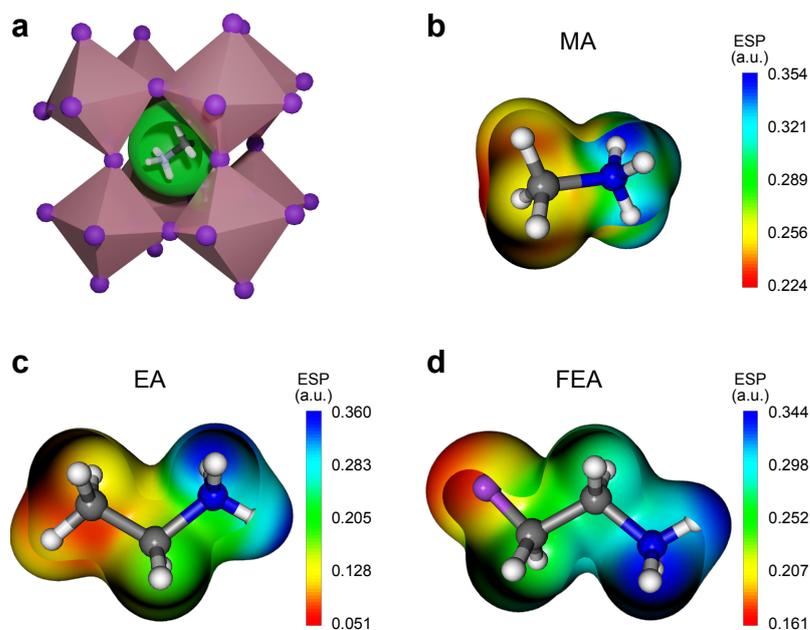
In our previous reports, the data-driven and machine-learning-aided investigations we discussed the perovskites' structural deformation induced by the various mixed-cation compositions of  $\text{MA}_{1-x}\text{A}_x\text{PbI}_3$ , where  $A$  is the additive cation [42–46]. We reported that upon mixing  $A$ -site cations in  $\text{MAPbI}_3$  perovskite framework, cations with a larger size than MA behave differently in stabilizing the framework than smaller cations. This suggests that different cations exert on the  $[\text{PbI}_3]$  inorganic network a variety of cage distortion and stress mechanisms. In this work, we ought to deepen this understanding when large cations are incorporated in the perovskites by comparing the behaviors of ethylammonium (EA;  $[\text{CH}_3\text{CH}_2\text{NH}_3]^+$ ) and 2-fluoroethylammonium (FEA;  $[\text{CH}_2\text{FCH}_2\text{NH}_3]^+$ ) cations and then search for interesting compositions that help minimize the enthalpy of mixing. This is achieved by employing manual DFT calculations and automated workflows guided by Bayesian optimization (BO) to explore the energy landscape of these mixed-cation perovskites at different concentrations and molecular orientations, taking enthalpic effects into account.

## 2. Results and Discussion

### 2.1. Ionic Size

A hybrid organic–inorganic perovskite,  $\text{APbI}_3$ , is composed of a positively charged molecular cation ( $A^+$ ) accommodated within a  $[\text{PbI}_3]^-$  network where the molecule plays

a role in charge balance. The molecular size, shape, and dipole moment determine the perovskite's intrinsic stability as well as the three dimensional (3D) connectivity of the  $\text{PbI}_6$  octahedral network via the electrostatic interaction as well as non-covalent interactions. As illustrated in Figure 1a, changes in the molecular occupancy of the cavity benefit from the flexibility of the  $\text{PbI}_6$  octahedra and their ability to elongate, tilt and rotate hence allowing the preservation of the 3D connected network, which is desirable for optoelectronic and catalytic applications.



**Figure 1.** (a) Crystal structure of a  $\text{APbI}_3$  perovskite. An organic cation is surrounded by the connected  $\text{PbI}_6$  octahedra and is accommodated within the cavity shown as green sphere. Electrostatic potential (ESP) maps of organic cations: (b) methylammonium, MA; (c) ethylammonium, EA; and (d) 2-fluoroethylammonium, FEA. The ESP potential is mapped in Hartree atomic units (a.u.) for a density isosurface of  $\rho = 0.01 \text{ e}\cdot\text{Bohr}^{-3}$ .

Figure 1b–d displays Electrostatic potential (ESP) maps of the MA, EA, and FEA organic cations considered in the present work. ESP analysis is a useful metric for analyzing the electrostatic properties of a given cation and its suitability to occupy the cavity within the perovskite framework. The ESP maps reveal that all three singly positive charged molecular cations are electropositive around all the constituent atoms.

As a common feature in all three cations, the most electron-deficient sites are located within the amine ( $-\text{NH}_3$ ) groups. Thus, once any of the three molecules are hosted within the perovskite structure cavity, the electron-deficient amine groups would exert the strongest ionic interactions on the surrounding  $[\text{PbI}_3]^-$  network. Focusing on the  $-\text{CH}_3$  groups of the three molecules, one might note that the electropositivity is significantly different around the terminal C atoms. EA and FEA molecules that have larger radii and stronger dipole moments than MA (see Table 1) have a less positive  $-\text{CH}_3$  and  $-\text{CH}_2\text{F}$  terminal group as compared to the  $-\text{CH}_3$  group in MA. This feature suggests that the electrostatic attraction of the terminal C group with the  $[\text{PbI}_3]^-$  network is weaker in the case of the larger molecules like EA and FEA compared to MA. However, this ESP analysis does not account for non-covalent interactions, such as hydrogen and halogen bonding, that are known to play an essential role in stabilizing the hybrid perovskites structure [36].

**Table 1.** The effective ionic radii and dipole moment of the used organic cations. And the tolerance factor (TF) in APbI<sub>3</sub> is calculated according to  $\tau = (r_A + r_I) / (\sqrt{2}(r_{Pb} + r_I))$ , where  $r_i$  is the ionic radius of the corresponding  $i$  ion.

Cation	Formula	Ionic Radius (pm)	Dipole Moment (Debye)	$\tau$
MA	[CH <sub>3</sub> NH <sub>3</sub> ] <sup>+</sup>	217	2.20	0.91
EA	[CH <sub>3</sub> CH <sub>2</sub> NH <sub>3</sub> ] <sup>+</sup>	274	3.85	1.03
FEA	[CH <sub>2</sub> FCH <sub>2</sub> NH <sub>3</sub> ] <sup>+</sup>	338	9.17	1.16

The complete substitution of MA with significantly larger cations EA and FEA, to form EAPbI<sub>3</sub> and FEAPbI<sub>3</sub>, led to a Goldschmidt tolerance factor ( $\tau$ ) exceeding 1 (see Table 1), which is indicative that the 2D hexagonal phase might be favored over the highly symmetric 3D cubic phase.

On the other hand, the partial substitution of MA with a larger cation at different concentrations brings several advantages for the stability and efficiency of the devices [34,47] while preserving the 3D connectivity of the octahedral network [46]. For example, the large dipole moment of molecular cation in mixed-cation halide perovskites [46] is a desired property that is known to enhance the charge carrier separation when the incident light generates photoexcited electron-hole pairs. In addition, the alignment of these molecules was found to be favorable for the hole transfer at the interfaces of the perovskite light absorbers [6,48].

The empirical cubic perovskite stability indicators, such as the tolerance factor and its enhanced versions, have been recently challenged and questioned after the successful synthesis of mixed-cation halide perovskites ( $A_{1-x}A'_x\text{PbI}_3$ ) that crystallize in the 3D pseudocubic perovskite structures. Our group recently demonstrated that the enthalpy of mixing plays a significant role in enhancing the intrinsic stability of mixed-cation halide perovskites compared to the conventional single cation MAPbI<sub>3</sub> material [46]. The enthalpy of mixing APbI<sub>3</sub> and A'PbI<sub>3</sub> ( $\Delta H_{\text{mix}}$ ) is calculated according to:

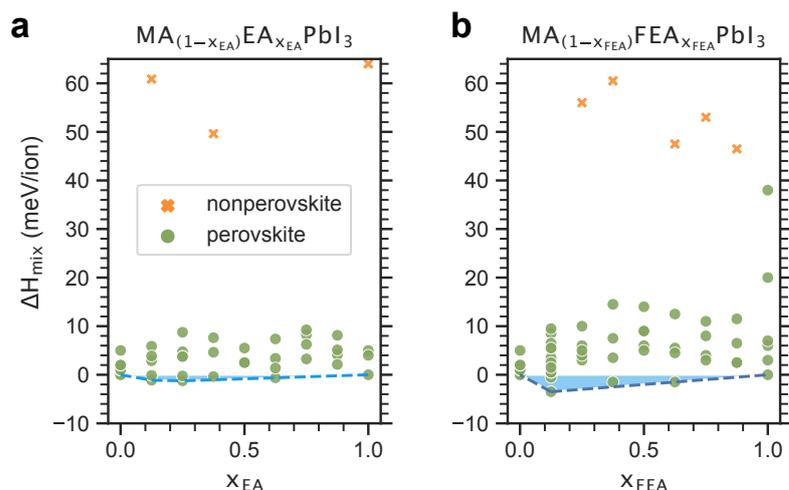
$$\begin{aligned} \Delta H_{\text{mix}}(A_{1-x}A'_x\text{PbI}_3) \\ = \Delta H(A_{1-x}A'_x\text{PbI}_3) - \{(1-x)\Delta H(A_{1-x}\text{PbI}_3) + x\Delta H(A'_x\text{PbI}_3)\}, \end{aligned} \quad (1)$$

where  $\Delta H$  is the enthalpy of formation for a given chemical formula of perovskite obtained from DFT calculations.

Although empirical models indicate that  $\tau > 1$  due to the large ionic size of EA and FEA, it remains remarkably relevant to search for promising compositions by exploring the energy landscape of these mixed-cation perovskites at different cation mixture concentrations and orientations taking enthalpic and entropic effects into account.

## 2.2. Enthalpy of Cation Mixing

Here, we report the enthalpies of mixing,  $\Delta H_{\text{mix}}$ , upon the substitution of MA with EA and with FEA within a perovskite framework, as shown in Figure 2. The resulting mixed-cation perovskites have the chemical formula  $\text{MA}_{1-x_{\text{EA}}}\text{EA}_{x_{\text{EA}}}\text{PbI}_3$  and  $\text{MA}_{1-x_{\text{FEA}}}\text{FEA}_{x_{\text{FEA}}}\text{PbI}_3$ , where  $0 \leq x_{\text{EA}} \leq 1$  and  $0 \leq x_{\text{FEA}} \leq 1$ , respectively. Among the fully optimized supercells, few non-perovskite structures have emerged where the connectivity between and within the [PbI<sub>6</sub>] octahedra was broken. One can note that, across the full compositional range, the non-perovskites represent the minority of three and five structures among 46 relaxed supercells of  $\text{MA}_{1-x_{\text{EA}}}\text{EA}_{x_{\text{EA}}}\text{PbI}_3$  and 61 of  $\text{MA}_{1-x_{\text{FEA}}}\text{FEA}_{x_{\text{FEA}}}\text{PbI}_3$ , respectively. Thus, in the remainder of the article, we exclude these non-perovskites from the dataset and investigate only the perovskite structures.



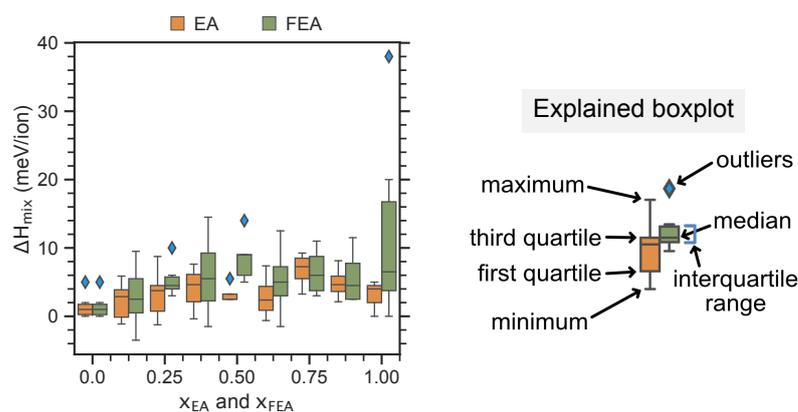
**Figure 2.** Enthalpies of mixing,  $\Delta H_{\text{mix}}$  (meV/ion), as a function of the large molecule concentration  $x_{\text{EA}}$  and  $x_{\text{FEA}}$  for (a)  $\text{MA}_{1-x_{\text{EA}}}\text{EA}_{x_{\text{EA}}}\text{PbI}_3$  and (b)  $\text{MA}_{1-x_{\text{FEA}}}\text{FEA}_{x_{\text{FEA}}}\text{PbI}_3$ . The orange crosses and green circles distinguish between non-perovskite and perovskite structure, respectively. The blue shaded areas denote the resulting region of enhanced stability with respect to the parent  $\text{MAPbI}_3$ , where  $\Delta H_{\text{mix}} < 0$  meV/ion.

Structurally, it is well accepted that the unequal molecular sizes typically induce some local distortion in the  $[\text{PbI}_6]$  octahedral network that often results in a positive value of  $\Delta H_{\text{mix}}$ . Interestingly, in the case of EA and FEA, one might notice a negative  $\Delta H_{\text{mix}}$  region as marked by the blue shaded region, where the extensive sampling of initial molecular orientations led to the exploration of the energy landscape.

Cation segregation is a common problem often encountered during the substitution of MA with other cations, such as Cs, K, Rb, FA, and GUA. Cation mixture relaxes the lattice contraction, thereby balancing the lattice strain and homogenizing alloyed organic-inorganic perovskites [49–51]. Our discovery of the perovskite structures representing minima at which  $\Delta H_{\text{mix}} < 0$  meV/ion implies the relaxation of the pure perovskites by the cation mixing. Subsequently, the negative value of  $\Delta H_{\text{mix}}$  suggests a suppressed EA and FEA cation segregation within the perovskite structures at the corresponding concentrations.

Contrary to the empirical structural model predictions, most mixed-cation compositions,  $\text{MA}_{1-x_{\text{EA}}}\text{EA}_{x_{\text{EA}}}\text{PbI}_3$  and  $\text{MA}_{1-x_{\text{FEA}}}\text{FEA}_{x_{\text{FEA}}}\text{PbI}_3$ , lead to the formation of stabilized 3D perovskite structures, despite the considerable molecular size difference between MA and its substituents, namely EA and FEA. Among the computed FEA-MA mixed perovskites compositions, the chemical formula  $\text{MA}_{0.875}\text{FEA}_{0.125}\text{PbI}_3$  turns out to have the lowest  $\Delta H_{\text{mix}}$  of  $-4$  meV/ion. This negative value implies that the perovskite  $[\text{PbI}_3]$  network would be favorable to accommodating this given concentration of FEA molecules, despite their bulky size of 338 pm. It is worth noting that the optimal concentration of the large cation replacement of MA should increase the intrinsic stability without compromising the desired structural and electronic properties. Consequently, extensive analysis and discussion of the bandgap and structural properties of the mixed cations are provided in the supplementary information.

Figure 3 displays the distribution of the  $\Delta H_{\text{mix}}$  computed data as a function of the concentration of the substituted molecules EA and FEA, represented by using the box-and-whisker plots (boxplots). It can be noticed that for EA incorporation, most  $\Delta H_{\text{mix}}$  values lay within an interquartile range of less than 5 meV/ion at most  $x_{\text{EA}}$  values. Interestingly, for  $\Delta H_{\text{mix}}$  calculated for FEA incorporation, most  $\Delta H_{\text{mix}}$  values lay within an interquartile range of less than 10 meV/ion at most  $x_{\text{FEA}}$  values except at  $x_{\text{FEA}} = 1$ , which spreads the most widely.



**Figure 3.**  $\Delta H_{\text{mix}}$  distribution in the dataset as a function of the concentration of the substituted molecule. We use Seaborn's boxplot [52] to display the properties and their distribution. The outliers are marked by the blue diamonds. The different parts of a boxplot are explained in the right panel showing the five-number summary of a set of data: minimum, lower quartile, median, upper quartile, and maximum.

### 2.3. Cation Alignment

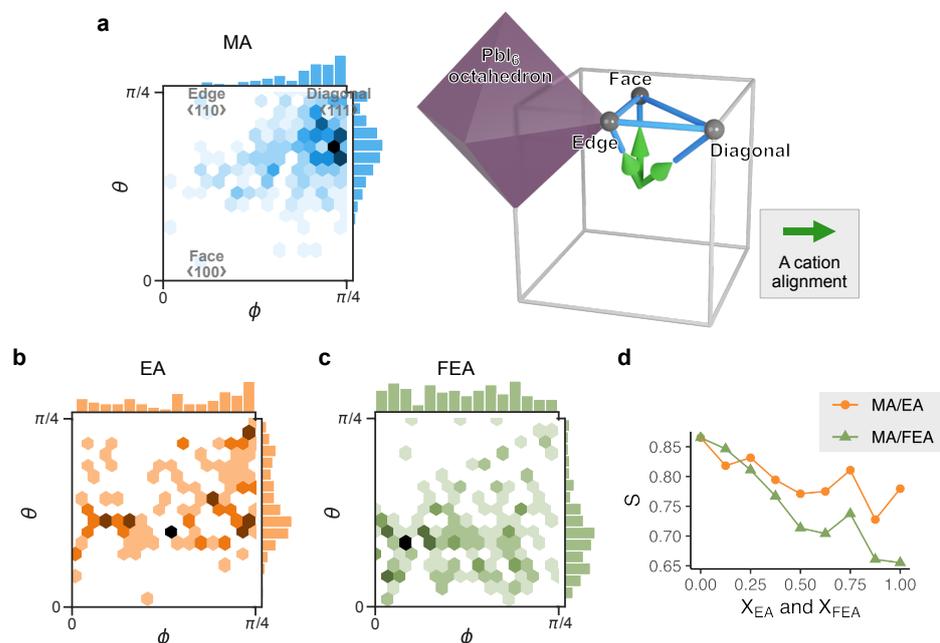
Figure 4a illustrates the MA alignment distribution in the entire dataset of computed perovskite crystal structures within a 48-fold symmetry group (Oh). In the perovskites, MA cations prefer to align in parallel to the diagonal direction ( $\langle 111 \rangle$ ;  $\theta = 4/\pi$  and  $\phi = 4/\pi$ ). This preferential orientation originates from the tendency to form hydrogen bonds between MA and the  $\text{I}^-$  anions. Meanwhile, the ionic radii of both EA and FEA cations exceed that of MA, hence stretching the cavity of the  $[\text{PbI}_3]$  network. As shown in Figure 4b,c for EA and FEA, respectively, these cations align toward the faces of the cube ( $\theta = 0$ ,  $\phi = 0$ ) by avoiding the short-range repulsion between cation and  $\text{I}^-$ . EA and FEA have a preference to align further toward the faces mainly because of their larger ionic radius if compared to MA. This cation size effect results in the average direction deviating from the diagonal direction as the EA and FEA concentrations increase (Figure 4d). The cations' direction subsequently influences the formation energies by deforming the  $[\text{PbI}_3]$  network and affecting the non-covalent interactions, such as the hydrogen bond and the halogen–halogen bond.

Hence, the structural stability of fully substituted  $\text{FEAPbI}_3$  perovskite crystal structures appears to be strongly dependent on the FEA molecular alignments. This impact is much stronger than the case with MA cations in  $\text{MAPbI}_3$  [53] which is due to the larger molecular size and dipole moment.

Phase segregation into either a 3D/2D or a 3D/1D mixture [31,47,54], when  $\tau > 1$ , implies that large-cation incorporation affects the restricted region while retaining the individual features within the inorganic network. Furthermore, empirically, a pure perovskite compound turns out a hexagonal phase favorably when  $\tau > 1$ . As we consider the enthalpy change within the cubic phase, the above stability estimation by mixing cations lacks the phase transition assessment. However, the comparisons of varying molecular alignments demonstrate its importance within the organic network, stabilizing the compound.

The search of DFT calculated structures giving the minimal enthalpy of mixing for a given concentration of substituted large molecular cations requires extensive sampling of randomized initial orientations within the perovskite supercell, making energy landscape exploration challenging. Detailed analysis of the current dataset points towards some promising large cation concentrations and trends in the cation alignments that would enable a guided search for the optimal concentration of large cation replacement of MA, without compromising the desired structural and electronic properties. The outcomes of the above analysis could be utilized by researchers to manually focus their compositional search on a narrower set of concentrations and molecular alignments. However, this approach might introduce bias into the search space. Furthermore, relevant concentrations

could be missed as it is highly dependent on the researcher's intuition and background. The following section reports our findings from Bayesian optimization (BO) to guide the optimal molecular concentrations and alignments.



**Figure 4.** Density map of molecular alignment vectors projected on spherical coordinates represented within 48-fold symmetry group (Oh) for (a) MA with the illustration of the representative directions, (b) EA, and (c) FEA. Here, the alignment vectors of a cation are constructed by connecting the terminal C, N, F atoms. The darker color indicates the highest incidence as also shown from the histograms constructed by binning the spherical angles. Face, edge and diagonal refer to the alignment of the molecules with respect to the perovskite cage, and correspond to the  $\langle 100 \rangle$  ( $\theta = 0$  and  $\phi = 0$ ),  $\langle 110 \rangle$  ( $\theta = \pi/4$  and  $\phi = 0$ ) and  $\langle 111 \rangle$  ( $\theta = \pi/4$  and  $\phi = \pi/4$ ) directions. (d) Orientation order parameter  $S = \frac{3}{2} \langle \cos^2 \alpha \rangle - \frac{1}{2}$  as a function of the substituted large cation concentration, where  $\alpha$  is the angle between  $\langle 111 \rangle$  and the molecular directions.

#### 2.4. The Search for Optimal Concentration of Large Cation Replacement of MA: Bayesian Optimization Aided Design

Although DFT is a compelling method for investigating materials' structural and electronic properties, its principal limitation is the computational expense incurred for extensively sampling all possible combinations of cation concentrations and orientations within a large supercell in a high-throughput scheme.

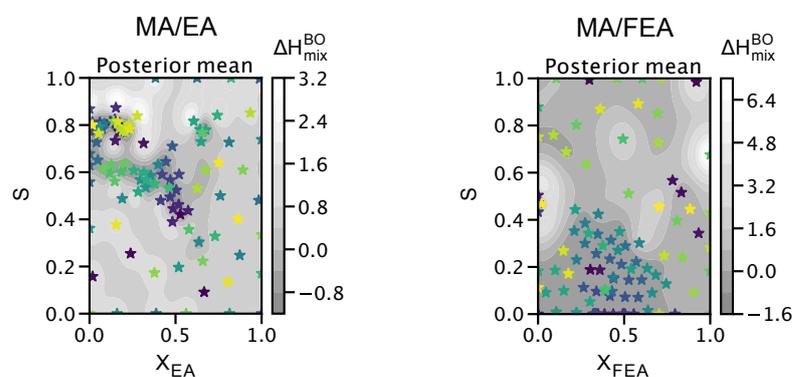
By increasing the supercell size, it is burdensome to empirically identify the minima of  $\Delta H_{\text{mix}}$  with confidence because not only does the DFT computation cost grow, but also the count of mixed-cation combinations to compute enlarges dramatically. For example, consider  $\text{MA}_{0.5}\text{FEA}_{0.5}\text{PbI}_3$  within  $2 \times 2 \times 2$  supercell. Without any symmetry, the number of mixed-cation initial configurations is  ${}^8C_4 = 70$ . On the other hand, for a  $3 \times 3 \times 3$  supercell, the number of possible initial configurations reaches  ${}^{27}C_{13}$  (and  ${}^{27}C_{14}$ ) = 20,058,300, making their exploration using DFT calculations prohibitively expensive.

We explore the applicability and efficiency of an automated workflow based on Bayesian Optimization (BO) in finding the structures that have minimal enthalpy of mixing, taking into account the large cation concentration and the cation alignment. Our initial dataset of mixed large-cations was constructed using computationally expensive DFT calculations for the function  $\Delta H_{\text{mix}}$ . A manual inspection of the dataset suggests that there are regions in the two-dimensional search  $\mathcal{X}$  space (*molecular concentration and alignment*) where  $\Delta H_{\text{mix}}$  achieves low values pointing toward the minimum. These are regions where the optimal *concentration* might exist and where large cation replacement of the MA cation

can be carried out without compromising the desired structural and electronic properties of the resulting perovskite. However, instead of proceeding manually, we can use BO to guide and automate our search towards the optimal configuration, of which  $\Delta H_{\text{mix}}$  is the lowest, in a computationally efficient manner. BO is ideally suited for this task because (i) each data point is obtained using a computationally expensive DFT calculation, and (ii) we do not have access to the underlying  $\Delta H_{\text{mix}}$  function but only its values at the expensively computed points. BO can incrementally guide our search towards the minimum of  $\Delta H_{\text{mix}}$  despite the absence of the gradient in the observation by developing a computationally inexpensive surrogate function to suggest the following configuration in the search space for which the DFT calculation should be invoked.

In our particular scenario, the search for the minimum is carried out within the 2D-space domain  $\mathcal{X}$  composed of (concentration, alignment) while the function  $f$  we aim to optimize is  $\Delta H_{\text{mix}}$  where each evaluation requires an expensive DFT computation. Full BO computational details are given in the methodology section.

Figure 5 presents the exploration and exploitation of 100 evaluations of  $\arg \min_{(x,S) \in \mathcal{D}} \Delta H_{\text{mix}}^{\text{BO}}(x,S)$  using Bayesian optimization (BO) with the posterior mean ( $\mu(x)$ , see Section 3.2), applied on the DFT-calculated dataset ( $\mathcal{D}$ ). The BO searched for the lowest  $\Delta H_{\text{mix}}$  values as a function of the concentration and molecular alignment. Comparing both panels in Figure 5, we observe that the MA/EA perovskite points are clustered in different  $S$  regions to the MA/FEA perovskite ones, whereas both MA/EA and MA/FEA data points are clustered within the concentration domain,  $0 < x < 0.5$ . Notably, the cluster of the searched points is located at the points where  $\Delta H_{\text{mix}}$  values are relatively low, while the BO model searches for the global minimum with  $x$  and  $S$ . The rest of the region is explored sparsely as BO determines that the probability of finding the minima there is low. This is a manifestation of the exploration–exploitation trade-off. BO will always have a non-zero probability of exploring regions of high uncertainty with the possibility of finding a minimum in those regions and this property allows BO to escape settling for local minima.



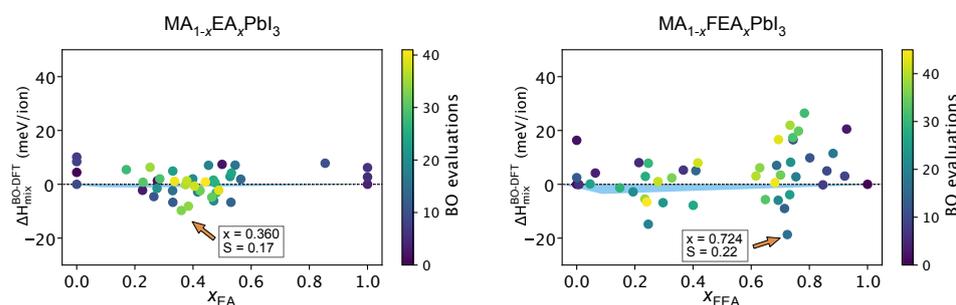
**Figure 5.** Exploration and exploitation of 100 evaluations of  $\arg \min_{(x,S) \in \mathcal{D}} \Delta H_{\text{mix}}^{\text{BO}}(x,S)$  (in meV/ion), where  $\mathcal{D}$  denotes the domain from the dataset. As the 100 cycles of the BO evaluation proceed, the points are marked gradually in order of sampling on a color scale from dark blue to yellow stars.

### 2.5. DFT Workflow Automation Aided by Bayesian Optimization

In the above DFT-calculation sections, we had manually selected the mixed-cation concentration when preparing the initial structures of DFT input in a sequential decision way. The number of initial configurations also limited us to conducting DFT calculations at each concentration. In this section, a BO model implemented in the workflow plays the role of the researcher who prepares the initial structures balancing exploration and exploitation.

As shown in Figure 6, we present the results of our automated BO-guided DFT workflow that searches for the minimum  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  across the cations' concentrations in  $\text{MA}_{1-x}\text{EA}_x\text{PbI}_3$  and  $\text{MA}_{1-x}\text{FEA}_x\text{PbI}_3$  actively and autonomously. The Rocketsled [55] framework has been adopted to build our automated workflow that searches for the

minimum  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  by exploiting and exploring the 2D space domain of large cation concentrations and molecular alignments.



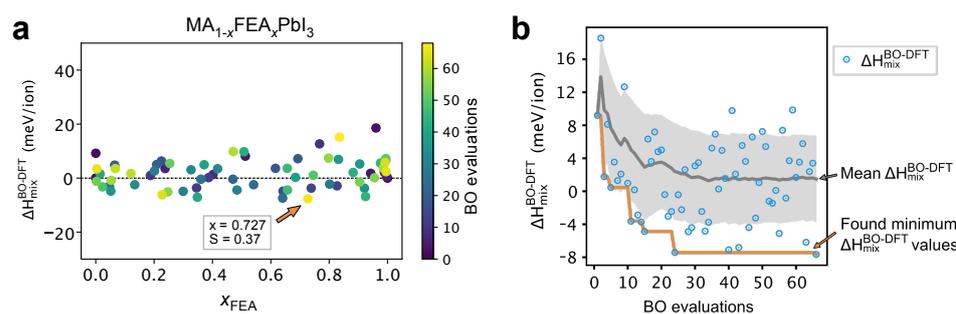
**Figure 6.** Exploration and exploitation of  $\text{argmin} \Delta H_{\text{mix}}^{\text{BO-DFT}}$  (in meV/ion), as the post-analyzer, processes the data during the DFT workflow automation, where the next best cation's concentration and orientation are selected by BO. As the BO search proceeds, the points are marked gradually in order of sampling on a color scale from dark blue to yellow circles. The configuration giving the minimum enthalpy found using the BO aided automated DFT workflow is pointed to by the orange arrow. The blue shaded areas denote the DFT results where  $\Delta H_{\text{mix}} < 0$  meV/ion, conducted manually.

As we consider the molecular orientation, the BO model efficiently explored the search space to identify an additional minimum for  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  compared to the manual procedure. A comparison of Figure 6 with Figure 2 shows that, using BO, we were able to obtain a *lower* minimum  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  value of  $-20$  meV compared to  $-4$  meV from the manually-prepared DFT calculation. This enhanced energy landscape exploration is mainly attributed to extensive molecular orientation sampling carried out automatically by the BO procedure.

In the manually-prepared DFT results of the  $\text{MA}_{1-x}\text{EA}_x\text{PbI}_3$  perovskites, where the initial molecular orientations were given randomly, the minimum was located at  $x = 0.250$ . Interestingly, the BO-guided automated search workflow located another lower minimum of  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  at around  $x = 0.375$  by exploiting and exploring molecular orientations during the preparation of input perovskite structures.

Meanwhile,  $\text{MA}_{1-x}\text{FEA}_x\text{PbI}_3$  perovskites show the minimum  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  at the  $x = 0.724$ ; which corresponds to  $x = 0.725$  in the  $(2 \times 2 \times 2)$ -supercell perovskites. Although the BO model also exploited the lower concentration at around  $x = 0.250$ , it also explored the higher concentration. As a result of the BO method, we found that the global minimum is around  $x = 0.725$ . It was unlikely that this *global* minimum could be located using manual inspection as it is not consistent with the common belief of researchers that large cations do not fit well within the perovskite structure.

Subsequently, we have extended the use of automated DFT workflows aided by Bayesian optimization to explore other interesting concentrations of  $\text{MA}_{1-x}\text{FEA}_x\text{PbI}_3$   $3 \times 3 \times 3$  supercells. As shown in Figure 7a, the BO model was able to locate a potential minimum of  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  with a large cation concentration and molecular alignment. We obtain the minimum point at  $x = 0.727$  similarly to the outcomes of the BO-guided DFT workflow in  $2 \times 2 \times 2$  supercells. The only distinction is that the minimum values are quantitatively higher across the FEA concentration ( $x_{\text{FEA}}$ ) in  $3 \times 3 \times 3$  supercells than those in  $2 \times 2 \times 2$  supercells mainly because the high degrees of freedom tend to lower the total energy of pure perovskites' structures. Figure 7b shows the convergence of the automated BO-guided DFT workflow searches in finding a potential global minimum as a function of the number of BO evaluations. The BO optimizer was able to locate the minimum mixing enthalpy around the 23rd iteration; that minimum continues to sustain until the 65th iteration.



**Figure 7.** (a) Exploration and exploitation of  $\text{argmin} \Delta H_{\text{mix}}^{\text{BO-DFT}}$  (in meV/ion) for  $3 \times 3 \times 3$   $\text{MA}_{1-x}\text{FEA}_x\text{PbI}_3$  supercell, as the post-analyzer processes the data during the DFT workflow automation and the next best cation's concentration and orientation are selected by BO. As the BO evaluation progresses, the points are marked gradually in order of sampling on a color scale from dark blue to yellow circles. The configuration giving the minimum enthalpy found using the BO aided automated DFT workflow is pointed to by the orange arrow. (b) The number of BO evaluations and found  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  values of the BO for the  $3 \times 3 \times 3$  supercell. The blue circles are the DFT calculated  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  as the BO evaluation progresses. The orange line delineates the found minimum at each BO iteration  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  and the gray line illustrates the mean values.

Hence, we have demonstrated the efficacy of DFT workflow automation aided by Bayesian optimization to accelerate the compositional search.

### 3. Computational Methods

#### 3.1. Dft Calculation Methods

We performed the DFT calculations of the  $(2 \times 2 \times 2)$ -supercell perovskites, using the Vienna Ab-initio Simulations Package (VASP) [56–58]. The projector-augmented wave [59,60] method described the wavefunction of the valence electrons at the Perdew–Burke–Ernzerhof (PBE) [61,62] functional level. The plane–wave cutoff was set at 520 eV for all the calculations. Besides, to make up for the GGA's inadequate dispersion interactions, we include the Tkatchenko–Scheffler van der Waals correction [63]. We optimized the structures until the energies and forces were converged within  $10^{-7}$  eV/atom and  $0.01$  eV/Å, respectively. Moreover, the calculated energies of the structures were obtained with the Monkhorst–Pack scheme [64]  $4 \times 4 \times 4$  k-point grid.

While we varied the content of the organic cations in a perovskite framework, we placed methylammonium (MA), ethylammonium (EA), and 2-fluoroethylammonium cations (FEA) in the vacant organic-cation sites. Their locations were undefined when we sampled the mixed-cation perovskites. Accordingly, in an  $\text{MA}_{1-x}\text{A}_x\text{PbI}_3$  perovskite (where  $A = \text{EA}$  and  $\text{FEA}$ ; and  $0 < x < 1$ ), the MA and A cations were distributed in random locations among the A-cation sites, in the cavity of the  $[\text{PbI}_3]^-$  network. Besides, the orientation of each cation was rotated differently in the initial coordinates. To manage the workflow, we customized FireWorks [65], and Pymatgen [66] tools were used when we set the perovskites and analyzed the optimized coordinates. Accordingly, we optimized and sampled 46  $\text{MA}_{1-x_{\text{EA}}}\text{EA}_{x_{\text{EA}}}\text{PbI}_3$  and 61  $\text{MA}_{1-x_{\text{FEA}}}\text{FEA}_{x_{\text{FEA}}}\text{PbI}_3$  compounds, where  $0 \leq x_{\text{EA}} \leq 1$  and  $0 \leq x_{\text{FEA}} \leq 1$ , respectively.

When we calculated the cations' dipole moment and electrostatic potential (ESP) in the gas phase, GAMESS 2020 R2 [67] was performed with B3LYP/6-311+G(d,p). Using these optimized coordinates, we obtained the effective ionic radii of the organic cations according to the method suggested by Kieslich et al. [68].

#### 3.2. Bayesian Optimization Methodology

Bayesian Optimization (BO) addresses the problem of finding a global optimization of a function whose form is not available but is accessible through (possibly noisy) function evaluation [69]. Thus, if  $f : \mathcal{X} \rightarrow \mathcal{Y}$ , which needs to be optimized, then  $f$  is accessible through a set of tuples  $\{(x_i, f(x_i))\}_{i=1}^n$ . Furthermore, it is assumed that function calls to  $f$

are computationally expensive, and thus a small number of evaluations should be carried out to arrive at the optimal.

In BO, one begins with a prior distribution on the functional forms of  $f$  and, after each evaluation, the prior is updated using the Bayes rule to arrive at a posterior distribution. The class of prior distributions is called the surrogate model, which is assumed to close under the application of a Bayesian update: if  $\mathcal{F}$  is the surrogate model and  $g \in \mathcal{F}$ , then after the function evaluation, the new function  $g' \in \mathcal{F}$ . Since the assumption is that the evaluation of the (unknown) function is computationally expensive, the selection of  $x$ , the next evaluation element, is guided by an acquisition function,  $a$ , which is derived from the latest posterior function  $g'$  and is relatively (computationally) cheaper to evaluate. After  $a$  proposes the next point to evaluate, the function call  $f$  is made and that in turn updates the posterior distribution on  $\mathcal{F}$ . This process is repeated until a computational budget is exhausted or  $a'$  proposes an evaluation point close to the one that has already been computed. A key characteristic of the acquisition function  $a$  is that it is designed to trade-off exploration vs exploitation. Exploration is designed to propose evaluation points for which the model has high uncertainty about the function value of  $f$ , while exploitation does the converse, that is, suggests points where there is a higher certainty about the value of  $f$ .

The most common surrogate model in BO is a Gaussian Processes (GP), which is a generalization of a multivariate Gaussian (Normal) distribution to functions [70]. A GP is defined by a mean function  $\mu(x)$  and a covariance function  $k$  and is denoted as  $\mathcal{GP}(\mu(x), k)$ . Problem specific information and constraints can be incorporated in  $k$ . A common covariance function is the Gaussian kernel  $k(x, y) = \exp\left(-\frac{\|x-y\|^2}{h}\right)$ , where  $\|x-y\|^2$  is the Euclidean norm and  $h$  is a length-scale parameter. Generalizations of the Gaussian kernel, including the Matern kernel and anisotropic kernels, have been used.

For the acquisition function  $a$ , several choices are available. A common choice is the *Expected Improvement*. Suppose  $x^*$  is the minimal value of  $f$  observed so far. Then define a function  $a$  as:

$$a(x) = \mathbb{E}_f(\max(0, x^* - f(x))).$$

The next evaluation point is then  $\arg \max_x a(x)$ , that is, a point is selected where the expected improvement is maximized.

Our constructed BO models guide the cations' concentrations and their initial coordinates in the DFT workflow automation. When we set the search domains, in addition to concentration variable  $x$ , we took the  $\Delta a_{\max}$  as the feature domain to prepare the molecular alignments in the DFT input. Because the molecular alignment relaxes during the ionic optimization calculations, the BO model sets the initial molecular alignments around the  $\langle 111 \rangle$  direction instead of the relaxed structures' orientation order parameter ( $S$ ). Here, we put  $\Delta a_{\max}$ , the maximum angle between the  $\langle 111 \rangle$  direction and molecular direction, to launch the DFT geometry relaxation tasks. Then, the workflow relocates the cations by allowing them in this molecular orientation range within the 48-fold symmetry group, preparing the input coordinates. For example, when  $\Delta a_{\max} = 0$ , all the cations are aligned in the  $\langle 111 \rangle$  direction in the perovskite. Besides, when  $\Delta a_{\max} = \pi/8$ , the cations are oriented randomly. Subsequently, the DFT relaxation task returns the estimated  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$  once the DFT geometry relaxation is completed. The input preparation of initial structures to be explored as a function of input with the concentrations and molecular alignments of the organic cations ( $x$  and  $\Delta a_{\max}$ ) in  $2 \times 2 \times 2$  and  $3 \times 3 \times 3$  supercells. The expected improvement (EI) acquisition function enabled us to sample the points to be explored and exploited. The  $\Delta H_{\text{mix}}$  values are predicted by the Gaussian Process (GP) regressor. Meanwhile, when we carried out the BO-guided workflow automation to assess  $\Delta H_{\text{mix}}^{\text{BO-DFT}}$ , the workflow sets up a  $2 \times 2 \times 2$ -grid mesh of the  $\mathbf{k}$ -points for the  $2 \times 2 \times 2$  supercell and a gamma-centered  $1 \times 1 \times 1$ -grid mesh for the  $3 \times 3 \times 3$  supercell during the VASP calculations.

#### 4. Conclusions

Large molecular cation incorporation in halide perovskites has recently emerged as an appealing strategy for enhancing materials and device stability. We evaluate the impact of large-sized cations' incorporation within the perovskite framework by mixing MA with two representative cations, EA and FEA. The analysis of the DFT calculations shows the importance of extensive sampling in the cation-concentration and orientation domains. Furthermore, the entropy-favored configurations hinder finding the global minimum. While the cation orientation order correlates with the local strain and instability, it suggests that orderly oriented cations can incorporate large cations at a given higher concentration.

Allowing for the effects at different concentrations and molecular orientations, our Bayesian optimization (BO)-guided workflow automation explored the energy landscape of these mixed-cation perovskites. The BO-guided workflow efficiently identifies the minimum enthalpy of mixing compared to the manual procedure. The enhanced energy landscape exploration is mainly attributed to extensive molecular orientation sampling aiming at the global minimum. This approach has the potential to reduce the DFT computational expense associated with searching for optimal compositions and might lead to the establishment of design principles for large molecular cation incorporation in halide perovskites expandable to other classes of materials.

**Supplementary Materials:** Supplementary Materials are available online. Figure S1: (a) Bandgap EA- and FEA-substituted perovskites (b) Comparison of unit cell volume increase; Figure S2: Correlation of octahedral deformation on the EA and FEA concentration in a perovskite. We identified the octahedral deformations: (a) quadratic elongation  $\lambda$  and (b) angle variance  $\sigma^2$ ; Figure S3: Correlation of octahedral tilt  $\theta^2$  on the E and FE concentration in a perovskite. The cartoon sketches the mechanistic recovery of the octahedral tilt; Figure S4: NCI isosurfaces ( $s = 0.6$ ) around the FEA in  $\text{MA}_{0.875}\text{FEA}_{0.125}\text{PbI}_3$ . Only the isosurface around the FEA cation is illustrated in yellow for clarity; Figure S5: Enhanced non-covalent interactions with FEA. Reduced density gradient plots as a function of sign  $(\lambda_2)\rho$  are compared in the range  $-0.04 < \rho < 0.04$ .  $\text{MAPbI}_3$ ,  $\text{MA}_{0.875}\text{EA}_{0.125}\text{PbI}_3$  and  $\text{MA}_{0.875}\text{FEA}_{0.125}\text{PbI}_3$  compounds are marked in orange, green, and blue, respectively. The gray box denotes the broaden range of non-covalent interactions in  $\text{MA}_{0.875}\text{FEA}_{0.125}\text{PbI}_3$ .

**Author Contributions:** Conceptualization, F.E.-M., H.P. and S.C.; Data curation, H.P.; Formal analysis, H.P., F.E.-M. and S.C.; Funding acquisition, F.E.-M. and S.C.; Investigation, H.P., S.K., S.C. and F.E.-M.; Methodology, H.P., S.C. and F.E.-M.; Project administration, F.E.-M.; Software, H.P.; Supervision, F.E.-M.; Validation, F.E.-M., H.P. and S.C.; Visualization, H.P.; Writing—original draft, H.P., S.K., S.C. and F.E.-M.; Writing—review & editing, H.P., S.K., S.C. and F.E.-M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This publication was made possible from the funding received for the Project, An Artificial Intelligence Platform for Accelerating Materials Discovery (AIPAM), awarded by the Hamad Bin Khalifa University Vice President Office. The findings herein reflect the work and are solely the responsibility of the author[s]. FE acknowledges support from the Qatar Environment and Energy Research Institute through the project code NOMDEE.

**Data Availability Statement:** The raw/processed data and recipe/workflow codes can be obtained by contacting the authors.

**Acknowledgments:** Computational resources have been provided by the research computing team at Texas A&M University at Qatar.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Jeon, N.J.; Noh, J.H.; Yang, W.S.; Kim, Y.C.; Ryu, S.; Seo, J.; Seok, S.I. Compositional engineering of perovskite materials for high-performance solar cells. *Nature* **2015**, *517*, 476. [[CrossRef](#)]
2. Yin, W.J.; Yang, J.H.; Kang, J.; Yan, Y.; Wei, S.H. Halide perovskite materials for solar cells: A theoretical review. *J. Mater. Chem. A* **2015**, *3*, 8926–8942. [[CrossRef](#)]
3. Walsh, A. Principles of chemical bonding and band gap engineering in hybrid organic–inorganic halide perovskites. *J. Phys. Chem. C* **2015**, *119*, 5755–5760. [[CrossRef](#)] [[PubMed](#)]

4. An, Y.; Hidalgo, J.; Perini, C.A.R.; Castro-Méndez, A.F.; Vagott, J.N.; Bairley, K.; Wang, S.; Li, X.; Correa-Baena, J.P. Structural Stability of Formamidinium-and Cesium-Based Halide Perovskites. *ACS Energy Lett.* **2021**, *6*, 1942–1969. [CrossRef]
5. Kanhere, P.; Chen, Z. A review on visible light active perovskite-based photocatalysts. *Molecules* **2014**, *19*, 19995–20022. [CrossRef]
6. Frost, J.M.; Butler, K.T.; Brivio, F.; Hendon, C.H.; Van Schilfgaarde, M.; Walsh, A. Atomistic origins of high-performance in hybrid halide perovskite solar cells. *Nano Lett.* **2014**, *14*, 2584–2590. [CrossRef]
7. Park, N.G. Perovskite solar cells: An emerging photovoltaic technology. *Mater. Today* **2015**, *18*, 65–72. [CrossRef]
8. Best Research-Cell Efficiency Chart. Available online: <https://www.nrel.gov/pv/cell-efficiency.html> (accessed on 25 May 2021).
9. Green, M.A.; Dunlop, E.D.; Hohl-Ebinger, J.; Yoshita, M.; Kopidakis, N.; Hao, X. Solar cell efficiency tables (version 56). *Prog. Photovoltaics Res. Appl.* **2020**, *28*, 629–638. [CrossRef]
10. Yin, W.J.; Shi, T.; Yan, Y. Unique properties of halide perovskites as possible origins of the superior solar cell performance. *Adv. Mater.* **2014**, *26*, 4653–4658. [CrossRef]
11. Chen, Q.; De Marco, N.; Yang, Y.M.; Song, T.B.; Chen, C.C.; Zhao, H.; Hong, Z.; Zhou, H.; Yang, Y. Under the spotlight: The organic–inorganic hybrid halide perovskite for optoelectronic applications. *Nano Today* **2015**, *10*, 355–396. [CrossRef]
12. Song, Z.; Waththage, S.C.; Phillips, A.B.; Heben, M.J. Pathways toward high-performance perovskite solar cells: Review of recent advances in organo-metal halide perovskites for photovoltaic applications. *J. Photonics Energy* **2016**, *6*, 022001. [CrossRef]
13. Sunghak, P.; Je Chang, W.; Woo Lee, C.; Sangbaek, P.; Hyo-Yong, A.; Tae Nam, K. Photocatalytic hydrogen generation from hydriodic acid using methylammonium lead iodide in dynamic equilibrium with aqueous solution. *Nat. Energy* **2016**, *2*, 16185. [CrossRef]
14. Bresolin, B.M.; Park, Y.; Bahnemann, D.W. Recent Progresses on Metal Halide Perovskite-Based Material as Potential Photocatalyst. *Catalysts* **2020**, *10*, 709. [CrossRef]
15. Zhao, Y.; Wang, L.; Song, T.; Mudryi, A.; Li, Y.; Chen, Q. Recent Progress in Designing Halide-Perovskite-Based System for the Photocatalytic Applications. *Front. Chem.* **2021**, *8*, 1178. [CrossRef]
16. Romani, L.; Malavasi, L. Solar-Driven Hydrogen Generation by Metal Halide Perovskites: Materials, Approaches, and Mechanistic View. *ACS Omega* **2020**, *5*, 25511–25519. [CrossRef]
17. Huang, H.; Pradhan, B.; Hofkens, J.; Roeffaers, M.B.J.; Steele, J.A. Solar-Driven Metal Halide Perovskite Photocatalysis: Design, Stability, and Performance. *ACS Energy Lett.* **2020**, *5*, 1107–1123. [CrossRef]
18. Chen, P.; Ong, W.J.; Shi, Z.; Zhao, X.; Li, N. Pb-Based Halide Perovskites: Recent Advances in Photo (electro) catalytic Applications and Looking Beyond. *Adv. Funct. Mater.* **2020**, *30*, 1909667. [CrossRef]
19. Ju, D.; Zheng, X.; Liu, J.; Chen, Y.; Zhang, J.; Cao, B.; Xiao, H.; Mohammed, O.F.; Bakr, O.M.; Tao, X. Reversible Band Gap Narrowing of Sn-Based Hybrid Perovskite Single Crystal with Excellent Phase Stability. *Angew. Chem. Int. Ed.* **2018**, *57*, 14868–14872. doi: 10.1002/anie.201810481. [CrossRef] [PubMed]
20. Pisanu, A.; Speltini, A.; Quadrelli, P.; Drera, G.; Sangaletti, L.; Malavasi, L. Enhanced air-stability of Sn-based hybrid perovskites induced by dimethylammonium (DMA): Synthesis, characterization, aging and hydrogen photogeneration of the MA<sub>1-x</sub>DMA<sub>x</sub>SnBr<sub>3</sub> system. *J. Mater. Chem. C* **2019**, *7*, 7020–7026. [CrossRef]
21. Ding, L.; Shen, C.; Zhao, Y.; Chen, Y.; Yuan, L.; Yang, H.; Liang, X.; Xiang, W.; Li, L. CsPbBr<sub>3</sub> nanocrystals glass facilitated with Zn ions for photocatalytic hydrogen production via H<sub>2</sub>O splitting. *Mol. Catal.* **2020**, *483*, 110764. doi: 10.1016/j.mcat.2020.110764. [CrossRef]
22. Koh, T.M.; Fu, K.; Fang, Y.; Chen, S.; Sum, T.C.; Mathews, N.; Mhaisalkar, S.G.; Boix, P.P.; Baikie, T. Formamidinium-containing metal-halide: An alternative material for near-IR absorption perovskite solar cells. *J. Phys. Chem. C* **2014**, *118*, 16458–16462. [CrossRef]
23. Mozur, E.M.; Neilson, J.R. Cation Dynamics in Hybrid Halide Perovskites. *Annu. Rev. Mater. Res.* **2021**, *51*, 269–291. [CrossRef]
24. Fu, Y.; Jiang, X.; Li, X.; Traoré, B.; Spanopoulos, I.; Katan, C.; Even, J.; Kanatzidis, M.G.; Harel, E. Cation Engineering in Two-Dimensional Ruddlesden–Popper Lead Iodide Perovskites with Mixed Large A-Site Cations in the Cages. *J. Am. Chem. Soc.* **2020**, *142*, 4008–4021. [CrossRef] [PubMed]
25. Liu, Y.; Wu, Z.; Dou, Y.; Zhang, J.; Bu, T.; Zhang, K.; Fang, D.; Ku, Z.; Huang, F.; Cheng, Y.B.; et al. Formamidinium-Based Perovskite Solar Cells with Enhanced Moisture Stability and Performance via Confined Pressure Annealing. *J. Phys. Chem. C* **2020**, *124*, 12249–12258. [CrossRef]
26. Hui, W.; Chao, L.; Lu, H.; Xia, F.; Wei, Q.; Su, Z.; Niu, T.; Tao, L.; Du, B.; Li, D.; et al. Stabilizing black-phase formamidinium perovskite formation at room temperature and high humidity. *Science* **2021**, *371*, 1359–1364. [CrossRef]
27. Park, B.W.; Kwon, H.W.; Lee, Y.; Lee, D.Y.; Kim, M.G.; Kim, G.; Kim, K.J.; Kim, Y.K.; Im, J.; Shin, T.J.; et al. Stabilization of formamidinium lead triiodide  $\alpha$ -phase with isopropylammonium chloride for perovskite solar cells. *Nat. Energy* **2021**, *6*, 419–428. [CrossRef]
28. Ono, L.K.; Juarez-Perez, E.J.; Qi, Y. Progress on perovskite materials and solar cells with mixed cations and halide anions. *ACS Appl. Mater. Interfaces* **2017**, *9*, 30197–30246. [CrossRef]
29. Maheshwari, S.; Patwardhan, S.; Schatz, G.C.; Renaud, N.; Grozema, F.C. The effect of the magnitude and direction of the dipoles of organic cations on the electronic structure of hybrid halide perovskites. *Phys. Chem. Chem. Phys.* **2019**, *21*, 16564–16572. [CrossRef]
30. Rahman, M.M.; Ahmed, A.; Ge, C.y.; Singh, R.; Yoo, K.; Sandhu, S.; Kim, S.; Lee, J.J. Trimethylsulfonium lead triiodide (TMSPI<sub>3</sub>) for moisture-stable perovskite solar cells. *Sustain. Energy Fuels* **2021**, *5*, 4327–4335. [CrossRef]

31. Parashar, M.; Singh, R.; Yoo, K.; Lee, J.J. Formation of 1-D/3-D Fused Perovskite for Efficient and Moisture Stable Solar Cells. *ACS Appl. Energy Mater.* **2021**, *4*, 2751–2760. [[CrossRef](#)]
32. Yi, C.; Luo, J.; Meloni, S.; Boziki, A.; Ashari-Astani, N.; Grätzel, C.; Zakeeruddin, S.M.; Röthlisberger, U.; Grätzel, M. Entropic stabilization of mixed A-cation ABX<sub>3</sub> metal halide perovskites for high performance perovskite solar cells. *Energy Environ. Sci.* **2016**, *9*, 656–662. [[CrossRef](#)]
33. Huang, S.; Huang, P.; Wang, L.; Han, J.; Chen, Y.; Zhong, H. Halogenated-Methylammonium Based 3D Halide Perovskites. *Adv. Mater.* **2019**, *31*, 1903830. doi: 10.1002/adma.201903830. [[CrossRef](#)]
34. Lefler, B.M.; May, S.J.; Fafarman, A.T. Role of fluoride and fluorocarbons in enhanced stability and performance of halide perovskites for photovoltaics. *Phys. Rev. Mater.* **2020**, *4*, 120301. [[CrossRef](#)]
35. Lermer, C.; Birkhold, S.T.; Moudrakovski, I.L.; Mayer, P.; Schoop, L.M.; Schmidt-Mende, L.; Lotsch, B.V. Toward Fluorinated Spacers for MAPI-Derived Hybrid Perovskites: Synthesis, Characterization, and Phase Transitions of (FC<sub>2</sub>H<sub>4</sub>NH<sub>3</sub>)<sub>2</sub>PbCl<sub>4</sub>. *Chem. Mater.* **2016**, *28*, 6560–6566. [[CrossRef](#)]
36. El-Mellouhi, F.; Rashkeev, S.N.; Marzouk, A.; Kabalan, L.; Belaidi, A.; Merzougui, B.; Tabet, N.; Alharbi, F.H. Intrinsic stability enhancement and ionic migration reduction by fluorinated cations incorporated in hybrid lead halide perovskites. *J. Mater. Chem. C* **2019**, *7*, 5299–5306. [[CrossRef](#)]
37. Chu, L.L.; Zhang, T.; Gao, Y.F.; Zhang, W.Y.; Shi, P.P.; Ye, Q.; Fu, D.W. Fluorine Substitution in Ethylamine Triggers Second Harmonic Generation in Noncentrosymmetric Crystalline [NH<sub>3</sub>CH<sub>2</sub>CH<sub>2</sub>F]<sub>3</sub>BiCl<sub>6</sub>. *Chem. Mater.* **2020**, *32*, 6968–6974. [[CrossRef](#)]
38. Haruyama, J.; Sodeyama, K.; Han, L.; Tateyama, Y. First-principles study of ion diffusion in perovskite solar cell sensitizers. *J. Am. Chem. Soc.* **2015**, *137*, 10048–10051. [[CrossRef](#)]
39. Tong, C.J.; Geng, W.; Prezhd, O.V.; Liu, L.M. Role of Methylammonium Orientation in Ion Diffusion and Current–Voltage Hysteresis in the CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3</sub> Perovskite. *ACS Energy Lett.* **2017**, *2*, 1997–2004. [[CrossRef](#)]
40. Garoufalidis, C.S.; Galanakis, I.; Zeng, Z.; Hayrapetyan, D.B.; Baskoutas, S. Structural and Electronic Properties of Small Perovskite Nanoparticles of the Form ABX<sub>3</sub> (A = MA, DEA, FA, GA, B = Pb, Sn, X = Cl, Br, I). *Electron. Mater.* **2021**, *2*, 382–393. [[CrossRef](#)]
41. Koliogiorgos, A.; Garoufalidis, C.S.; Galanakis, I.; Baskoutas, S. Electronic and Optical Properties of Ultrasmall ABX<sub>3</sub> (A = Cs, CH<sub>3</sub>NH<sub>3</sub>/B = Ge, Pb, Sn, Ca, Sr/X = Cl, Br, I) Perovskite Quantum Dots. *ACS Omega* **2018**, *3*, 18917–18924. [[CrossRef](#)] [[PubMed](#)]
42. Park, H.; Mall, R.; Alharbi, F.H.; Sanvito, S.; Tabet, N.; Bensmail, H.; El-Mellouhi, F. Exploring new approaches towards the formability of mixed-ion perovskites by DFT and machine learning. *Phys. Chem. Chem. Phys.* **2019**, *21*, 1078–1088. [[CrossRef](#)]
43. Park, H.; Mall, R.; Alharbi, F.H.; Sanvito, S.; Tabet, N.; Bensmail, H.; El-Mellouhi, F. Learn-and-Match Molecular Cations for Perovskites. *J. Phys. Chem. A* **2019**, *123*, 7323–7334. [[CrossRef](#)]
44. Ali, A.; Park, H.; Mall, R.; Aïssa, B.; Sanvito, S.; Bensmail, H.; Belaidi, A.; El-Mellouhi, F. Machine Learning Accelerated Recovery of the Cubic Structure in Mixed-Cation Perovskite Thin Films. *Chem. Mater.* **2020**, *32*, 2998–3006. [[CrossRef](#)]
45. Park, H.; Mall, R.; Ali, A.; Sanvito, S.; Bensmail, H.; El-Mellouhi, F. Importance of structural deformation features in the prediction of hybrid perovskite bandgaps. *Comput. Mater. Sci.* **2020**, *184*, 109858. [[CrossRef](#)]
46. Park, H.; Ali, A.; Mall, R.; Bensmail, H.; Sanvito, S.; El-Mellouhi, F. Data-driven enhancement of cubic phase stability in mixed-cation perovskites. *Mach. Learn. Sci. Technol.* **2021**, *2*, 025030. [[CrossRef](#)]
47. Jodlowski, A.D.; Roldán-Carmona, C.; Grancini, G.; Salado, M.; Ralaiarisoa, M.; Ahmad, S.; Koch, N.; Camacho, L.; Miguel, G.D.; Nazeeruddin, M.K. Large guanidinium cation mixed with methylammonium in lead iodide perovskites for 19% efficient solar cells. *Nat. Energy* **2017**, *2*, 972–979. [[CrossRef](#)]
48. Zhao, S.; Xie, J.; Cheng, G.; Xiang, Y.; Zhu, H.; Guo, W.; Wang, H.; Qin, M.; Lu, X.; Qu, J.; et al. General Nondestructive Passivation by 4-Fluoroaniline for Perovskite Solar Cells with Improved Performance and Stability. *Small* **2018**, *14*, 1803350. doi: 10.1002/smll.201803350. [[CrossRef](#)] [[PubMed](#)]
49. Zheng, X.; Wu, C.; Jha, S.K.; Li, Z.; Zhu, K.; Priya, S. Improved Phase Stability of Formamidinium Lead Triiodide Perovskite by Strain Relaxation. *ACS Energy Lett.* **2016**, *1*, 1014–1020. [[CrossRef](#)]
50. Correa-Baena, J.P.; Luo, Y.; Brenner, T.M.; Snaider, J.; Sun, S.; Li, X.; Jensen, M.A.; Hartono, N.T.P.; Nienhaus, L.; Wieghold, S.; et al. Homogenized halides and alkali cation segregation in alloyed organic-inorganic perovskites. *Science* **2019**, *363*, 627–631. [[CrossRef](#)] [[PubMed](#)]
51. Knight, A.J.; Borchert, J.; Oliver, R.D.J.; Patel, J.B.; Radaelli, P.G.; Snaith, H.J.; Johnston, M.B.; Herz, L.M. Halide Segregation in Mixed-Halide Perovskites: Influence of A-Site Cations. *ACS Energy Lett.* **2021**, *6*, 799–808. [[CrossRef](#)]
52. Waskom, M.L. seaborn: Statistical data visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [[CrossRef](#)]
53. Mozur, E.M.; Maughan, A.E.; Cheng, Y.; Huq, A.; Jalarvo, N.; Daemen, L.L.; Neilson, J.R. Orientational Glass Formation in Substituted Hybrid Perovskites. *Chem. Mater.* **2017**, *29*, 10168–10177. [[CrossRef](#)]
54. Sahayaraj, S.; Radicchi, E.; Ziółek, M.; Ścigaj, M.; Tamulewicz-Szwajkowska, M.; Serafińczuk, J.; De Angelis, F.; Wojciechowski, K. Combination of a large cation and coordinating additive improves carrier transport properties in quasi-2D perovskite solar cells. *J. Mater. Chem. A* **2021**, *9*, 9175–9190. [[CrossRef](#)]
55. Dunn, A.; Brenneck, J.; Jain, A. Rocketsled: A software library for optimizing high-throughput computational searches. *J. Phys. Mater.* **2019**, *2*, 034002. [[CrossRef](#)]
56. Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **1993**, *47*, 558–561. [[CrossRef](#)] [[PubMed](#)]

57. Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169–11186. [[CrossRef](#)]
58. Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15–50. [[CrossRef](#)]
59. Blöchl, P.E. Projector augmented-wave method. *Phys. Rev. B* **1994**, *50*, 17953. [[CrossRef](#)]
60. Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59*, 1758. [[CrossRef](#)]
61. Perdew, J.P.; Burke, K.; Wang, Y. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B* **1996**, *54*, 16533–16539. [[CrossRef](#)]
62. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [[CrossRef](#)]
63. Tkatchenko, A.; Scheffler, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102*, 073005. [[CrossRef](#)]
64. Monkhorst, H.J.; Pack, J.D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **1976**, *13*, 5188. [[CrossRef](#)]
65. Jain, A.; Ong, S.P.; Chen, W.; Medasani, B.; Qu, X.; Kocher, M.; Brafman, M.; Petretto, G.; Rignanese, G.M.; Hautier, G.; et al. FireWorks: A dynamic workflow system designed for high-throughput applications. *Concur. Comp.-Pract. E.* **2015**, *27*, 5037–5059. [[CrossRef](#)]
66. Ong, S.P.; Richards, W.D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V.L.; Persson, K.A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **2013**, *68*, 314–319. [[CrossRef](#)]
67. Barca, G.M.J.; Bertoni, C.; Carrington, L.; Datta, D.; De Silva, N.; Deustua, J.E.; Fedorov, D.G.; Gour, J.R.; Gunina, A.O.; Guidez, E.; et al. Recent developments in the general atomic and molecular electronic structure system. *J. Chem. Phys.* **2020**, *152*, 154102. [[CrossRef](#)]
68. 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. [[CrossRef](#)]
69. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **2016**, *104*, 148–175. [[CrossRef](#)]
70. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.