Computational Materials Science 173 (2020) 109415

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Cuboctahedral stability in Titanium halide perovskites *via* machine learning

Hisham A. Maddah[a,b], Vikas Berry[a], Sanjay K. Behura[a,*]

[a] Department of Chemical Engineering, University of Illinois at Chicago, 945 W Taylor St., 60607, Engineering Innovation Building, Chicago, IL, United States
[b] Department of Chemical Engineering, King Abdulaziz University, Rabigh 21911, Saudi Arabia

ARTICLE INFO

Keywords:
Perovskites
Structural formability
Crystal stability
Ionic radius
Statistical analysis
Machine learning

ABSTRACT

Titanium (Ti)-based halide perovskites with A_2TiX_6 and $ATiX_3$ stoichiometry can be promising alternatives to lead-based perovskites due to their high cuboctahedral stability and reduced internal stresses. It is important to study compositional stability and crystal chemistry controlling the formation of BX_6 octahedral periodic arrays in Ti-based perovskites. Here, we investigate on the formability of Ti-based mixed organic-inorganic and all-inorganic perovskite crystals *via* establishing a library of tolerance (t) and octahedral (μ) stability factors using random sampling of ionic radii. Structural mapping and machine learning analysis are performed for twelve Ti-perovskites selected by mixing/matching cations with halide anions. Formation probabilities from normal/binomial distributions are integrated into a statistical algorithm used in a "Decision tree classifier" to obtain training/testing datasets. We relate perovskites stability to cuboctahedral formation considering complex crystal chemistry including ionic radii, bond length from ions electronegativity, and bond-valence sums from coordination numbers. Results revealed that cuboctahedral B-sites should be occupied by Ti cations to have a stable structural composition. High formation tendencies are estimated for methylammonium Ti-based chloride or bromide perovskites. Game theory methods determined six formable structures (Cs_2TiX_6 and $RbTiX_3$ for $X = F^-$, Cl^-, Br^-). Probabilities of formation of future datasets with different halides are A_mTiBr_{3m} (0.93) > A_mTiCl_{3m} (0.91) > A_mTiF_{3m} (0.79) > A_mTiI_{3m} (0.52) matching statistical results with ± 8% marginal error and 96.3% classification accuracy; there is optimality in having a (Ti-Br)-like structure for a balanced cuboctahedral. This work shows a potential towards the discovery of novel lead-free perovskites for photoluminescence, tunable bandgaps, tunneling junctions, and photovoltaic applications.

1. Introduction

Organic-inorganic hybrid perovskite crystals are widely used in optoelectronic devices such as solar cells and light-emitting devices [1–3] owing to their strong ability for visible light absorption (0.7×10^7 cm^{-1}) [4] and tunability of their bandgap energy (1.5 ~ 2.43 eV) [2,5,6]. Typical perovskite materials have the chemical formula ABX_3 where A refers to an organic and/or inorganic cation with +1 charge (e.g. Cs^+, MA^+, and Rb^+), B is a metal cation with +2 charge (e.g. Pb^{+2}, Hg^{+2}, Zn^{+2}, Mg^{+2}, and Cd^{+2}) and X is a halide anion with −1 charge (e.g. F^-, Cl^-, Br^-, and I^-) [7–9]. Ideal ABX_3 perovskites consist of a cubic array with three-dimensional A-site corners sharing BX_6 octahedral units in which A-site cation is surrounded by 12 equidistant X anions; the A-site cations fill the cuboctahedral cavities and B-site cations occupy the center of the octahedral sublattice and have 6 nearest X-anion neighbors [10,11]. Previous studies identified that conventional Pb-based perovskites (e.g. $MAPbI_3$) show structural stability problems at high temperatures (> 100 °C) and/or from other environmental factors (moisture, light, and oxygen) due to formation of lead-iodide salts [12–19] that would only occur if internal stresses are enough to dislodge ions form their A and B sites within the cubic and octahedral structures, respectively. Stabilization of inorganic halide perovskites may be achieved through grain encapsulation, surface passivation or most importantly by using stable B-site metals like Ti (e.g. Cs_2TiX_6) for less-toxicity and broader tunable bandgaps [20]. Formability (stability) of halide organic-inorganic Ti-based non-toxic perovskites is critical to study as for the case of other Pb, Sn, Hg, and Ba metal cations based-structures studied by Tanaka et al. and Li et al. [21,22].

In this context, we theoretically investigate on the formability and stability of various Ti-based perovskites through mixing/matching of various cations and anions from random sampling of their reported ionic radii (e.g. Shannon, Pauling, and Stern) for determining Goldschmidt's tolerance factor (t) and octahedral factor (μ). Structural mapping, probabilities of formation, normal and binomial probability distributions, Nash equilibrium, and iterative deletion of strictly dominated strategies (IDSDS) game theory methods are considered for

* Corresponding author.
E-mail address: sbehura1@uic.edu (S.K. Behura).

https://doi.org/10.1016/j.commatsci.2019.109415
Received 13 September 2019; Received in revised form 12 November 2019; Accepted 15 November 2019
Available online 13 December 2019
0927-0256/

Nomenclature

Greeks and Letters

A	Organic and/or inorganic cation
B	Metal cation ($B = Ti^{+2}$)
X	Halide anion
t	Tolerance factor
R	Ionic radius, Å
μ	Octahedral factor
$\bar{X}$	Sample mean
x	Sample number
n	Sample size
σ	Sample standard deviation
y	Normal distribution (normal probability density function)
p	Probability of formation (success)
q	Probability of instability (failure)
P	Binomial probability
k	Desired outcome (desired successful trials)
v	Utility (payoff or preference)
BL	Bond length, Å
χ	Electronegativity
BVS	Bond-valence sums of oppositely charged ions with respect to the ion of interest
υ	Valence or charge state of a bond between two atoms
b	Universal value of 0.4 Å for BVS
CA	Classification accuracy
N	Number of samples that are correctly classified
S	Total number of samples

Subscripts and superscripts

	Probability with respect to tolerance factor (e.g. p_t refers to the percentage of stable tolerance factors; q_t refers to the percentage of unstable tolerance factors)
μ	Probability with respect to octahedral factor (e.g. p_μ refers to the percentage of stable octahedral factors; q_μ refers to the percentage of unstable octahedral factors)
in	Stability factors within the stability range (e.g. t_{in} refers to the number of determined stable tolerance factors)
n	Sample stability factors (e.g. t_n refers to the number of all calculated tolerance factors for a sample with n size)
1	Player 1 as an organic and/or inorganic cation
2	Player 2 as a halide anion
*	Best responses of player 1 or 2 to one another (best replies from cations to anions and vice versa)
ij	Refers to bonding between i and j atoms as in υ_{ij} and BL_{ij}
i	First atom or ion in diatomic bonding for BL and BVS calculations (refers to i atom as in BVS_i)
j	Second atom or ion in diatomic bonding for BL and BVS calculations (refers to j atom as in BVS_i)
0	Refer to a reference tabulated value as in bond length of bond-valence parameter L_0

further statistical analysis. Decision tree classification technique (supervised learning) is used to construct a machine learning model for the various Ti-based perovskites from their training datasets (input/response) to predict future datasets outputs. A statistical algorithm is built to obtain training/testing datasets determined from a random sampling of involved ionic radii. Training/testing datasets address the relation between various ionic radii sizes and structural stability as well as the contribution of the BX_6 octahedral ions/atoms electronegativity, bond length (BL), coordination number (CN), and bond-valence sums (BVS) or formal charge state (oxidation numbers) in optimizing perovskite

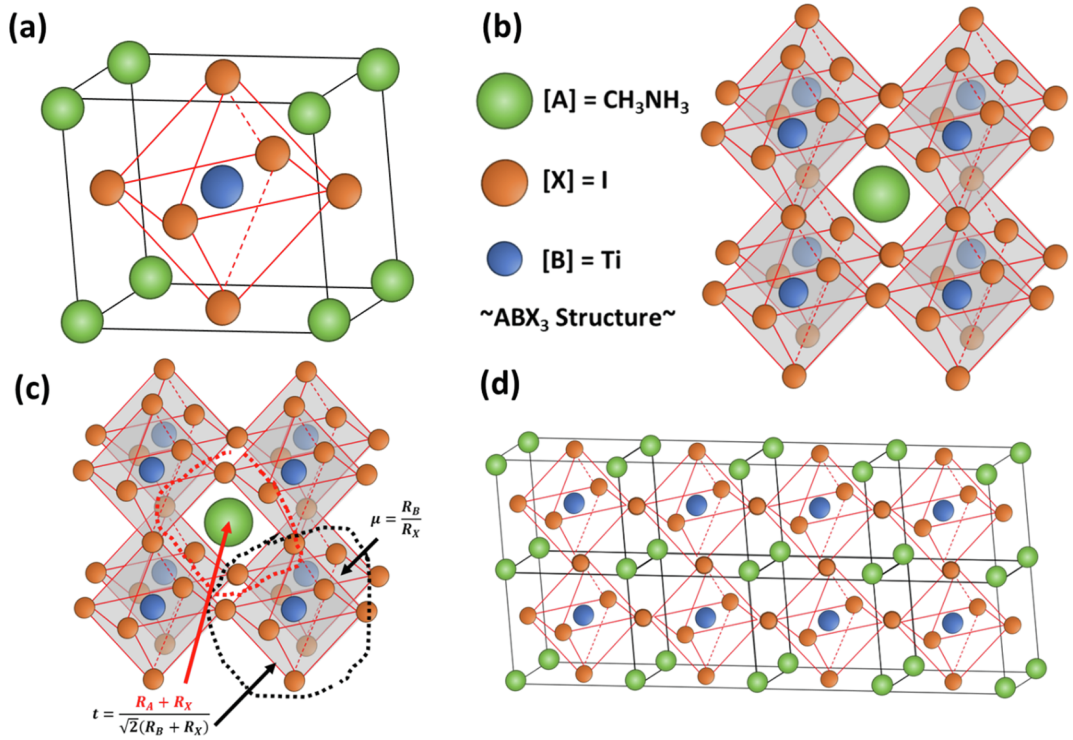

Fig. 1. Ti-based perovskites (ABX_3) typical crystal structure: (a) A cubic structure unit cell; (b) MA^+ cation occupying an A-site and surrounded by eight octahedral structures; (c) Role of stability factors in determining the possibility of MA^+ and Ti^{+2} cations to occupy A-sites and B-sites, respectively, based on the involved elements ionic radii; (d) The ideal cubic structure for $MATiI_3$ when (t = 1) indicating a perfect cation fit.

formability from stability factors. Predicted model results for the new/future datasets are used to test the model validity from comparison with training datasets results for credible formability knowledge.

2. Formability and stability factors

Theoretically, formability and structural stability of A_2TiX_6 and/or $ATiX_3$ perovskite materials (e.g. $MATiI_3$ crystals with MA^+, Ti^{+2}, and I^- ions) are found to be linked to and could be possibly estimated from two main factors: (i) t-factor, and (ii) μ-factor [21]. Goldschmidt's tolerance factor evaluates whether A cations (organic or inorganic) can enter between the formed octahedral structures to fill cubic structure corners in the BX_3 framework. Every cavity is surrounded by eight octahedral BX_6 structures and this might prevent A cations from occupying A-sites if the opening is not large enough to fit the radius size of A cations [10].

The perovskite structure would be empirically formable with feasible structural stability when: (i) t-factor ranges from 0.813 to 1.107, and (ii) μ-factor ranges from 0.442 to 0.895 [21]. For ideal cubic structures: t ~1; otherwise, structural distortions lower crystal symmetry and result in an orthorhombic or tetragonal structure (t = 0.71–0.9) [11]. More details regarding the stability factors roles in defining the crystal chemistry, structural sites occupation, and formability of perovskites can be found in Supplementary (see section 1).

It is worth mentioning that the above stability factors range hold if ionic radii (A, B or X) are assumed to be rigid, spherical shaped and constant regardless of surrounding temperature and ions [10]. This assumption may not be valid for systems with heavier halides (e.g. I^- as compared to O^{-2}, or F^-) due to lower electronegativity and higher atomic number; hence, t-factor might fail in the formability prediction test of many iodide-based perovskites since previous works identified 32 known inorganic ABI_3 perovskites failing the test [10,21]. A better approach to check for the formability of iodide-based perovskites is to lower the t range to $0.8 > t > 0.9$ [10] compensating for the higher atomic number and lower electronegativity that might excessively increase the halide anion radius and the octahedral size leading to narrowing A-site cavities (Fig. 1).

3. Atomic and ionic radii and their impact on crystal chemistry

Variations in the reported ionic radii of involved perovskites ions are widespread where changes are dependent upon electronegativity, BL, CN, and BVS of ions impacting the crystal chemistry. Thus, quantifying ionic radii ranges and determining their corresponding crystal parameters are important to accurately predict structural stabilities and identify the controlling parameters that would tune/change perovskite formability. More details on the relative size of atoms/ions according to the distribution of electrons and the BVS role can be found in the Supplementary (see section 2).

The constituent ions involved in the perovskite structure might be approximated to be hard spheres as reported by Shannon [23,24] for the estimation of the solid-state ionic radii size. However, deviations from standard Shannon radii are possible to occur if there is a strong dependency of cations on anions, and vice versa, which arise from covalency relating to the number of bonds an atom can form within a molecule as in metal-halide (B-X) bonds (much prominent in heavier halides as compared to metal-oxide and fluoride bonds) [10]. Hence, the conventional Shannon radii values might not be sufficient for predicting the formability of organic and/or inorganic perovskite structures (Ti-based perovskites: A_2TiX_6 and/or $ATiX_3$) from their tolerance and octahedral factors. This is attributed to the fact that B-X covalent bonds for all metals show unusual bond lengths different from those reported by Shannon. Variations in the bond lengths would result in deviations from the typical ionic radii (e.g. Shannon); especially when there are heavier halides X = Cl^-, Br^-, and I^- attached to a metal element. Travis et al. [10] found out that d-block metals (e.g. Ti, Cd, Hg) bonded with halogens would show a shorter B-X covalent bond length of 0.075 ~ 0.1 Å less than those of Shannon bonds leading to greater variations in ionic radii. Lower electronegativity between the metal and halogen explains the reason of having greater covalency of atoms and capacity to share more electrons and form covalent bonds (electronegativity of Ti (1.54) is less than Pb (2.33) [25], which might be the reason for more ionic radii variations in Ti^{+2} and less validity of Shannon model); ionic radii of Pb^{+2} is around 1.33 ~ 1.43 Å while for Ti^{+2} is 0.86 ~ 1 Å which both depend on geometry, coordination, and electronic environment [23,26–28]. The ionic radius of methylammonium (MA) which is determined to be between 1.8 ~ 2.7 Å (as approximated from Table 1) would make MA a suitable organic cation candidate for forming the perovskite structure. A more rational option to form Ti-perovskites would be to select Cs and/or Rb monovalent cations that have smaller ionic radii than MA that are in the ranges 1.69 ~ 1.88 Å, and 1.48 ~ 1.72 Å, respectively (as reported in Table 1). Discussions on the introduction of promising Br^- or Cl^- and the use of MA^+, Cs^+, and Rb^+ cations in forming Ti-based perovskites, reported by Tanaka et al. [21], can be found in Supplementary (see section 2).

The obtained results in the earlier study [21] might not represent correct stabilities due to the assumption that ionic radii are considered to be exactly similar to the standard Shannon values (constant values) that do not depend on ions electronegativity, oxidation state, and interactions with neighboring ions. Hence, we have obtained all the possible ionic radii from various studies to recalculate the formation probability of Ti-based perovskites from our determined ionic radii ranges for (A = MA^+, Cs^+, Rb^+), (B = Ti^{+2}) and (X = F^-, Cl^-, Br^-, I^-) elements. In short, we have gathered all the possible ionic radii values for the different atoms/ions involved in the studied Ti-based perovskite structures (Table 1) to carry out our formability test and stability calculations.

Lufaso et al. [29] programmed SPuDS software to predict the octahedral tilt angle with a high degree of accuracy. However, SPuDS does not work well in systems where octahedral distortions result from changes in the B-X bond length (were held constants). The unstudied parameters including changes in the B-X bond length and the possibility of having Ti^{+2}-based perovskites have been incorporated in this study using bond valences regarded to ionic radii and composition to explain octahedral stability in the presence of (B-X)-based distortions. More details about the pioneering work by Lufaso et al. [29] on predicting $CaTiO_3$ using SPuDS can be found in Supplementary (see section 2).

Table 1
Possible ionic radii values of cations/anions reported in previous studies and literature with their expected ranges used in the theoretical and probabilistic calculations.

Ions	MA^+	Cs^+	Rb^+	Ti^{+2}	F^-	Cl^-	Br^-	I^-	Ref.
Ionic Radius (Å)	2.17	1.88	1.72	0.86	1.35	1.81	1.96	2.20	[21]
	1.80	1.75	1.48	0.92	1.36	1.81	1.92	2.19	[4,23,26]
	–	1.69	–	0.96	1.33	1.81	1.95	2.16	[26,30]
	2.23	–	1.52	1.00	1.33	1.84	1.96	2.16	[31–33]
	2.70	1.81	1.66	1.00	1.33	1.81	–	2.03	[28,33,34]
	1.8 ~ 2.7	1.69 ~ 1.88	1.48 ~ 1.72	0.86 ~ 1	1.33 ~ 1.36	1.81 ~ 1.84	1.92 ~ 1.96	2.03 ~ 2.2	Radii ranges; This study

4. Ti-based perovskites *via* machine learning

Application of machine learning in materials science for the prediction of material property and in the discovery of new materials and/or perovskites have been already proposed in earlier works [35–39]. High-dimensional data require an automated analytical model building technique such as machine learning allowing computers to learn from input/response data using various algorithms; hence, classification of discrete information or regression of continuous data would be applicable *via* iterative construction of the model. In other words, models obtained from machine learning methods and input/output samples can produce reliable and credible decisions and results regarding the stability or formability of materials of interest. The basic steps involved in machine learning are sample construction, model building, and model evaluation. Built models are usually evaluated from the model accuracy in predicting results similar to values from experiments and/or from reliable theoretical models [35].

Li et al. (2018) and Li et al. (2019) [36,37] combined machine learning with density functional theory (DFT) calculations to study the formability and stability of 354 kinds of halide perovskites from their decomposition energies according to the elemental ionic radius. The machine learning model, including stability factors, was able to discover stabilities of $\sim 10^3$ rare-earth-metal perovskites from the total tested halide double perovskites of $\sim 10^4$. Pilania et al. [38] classified

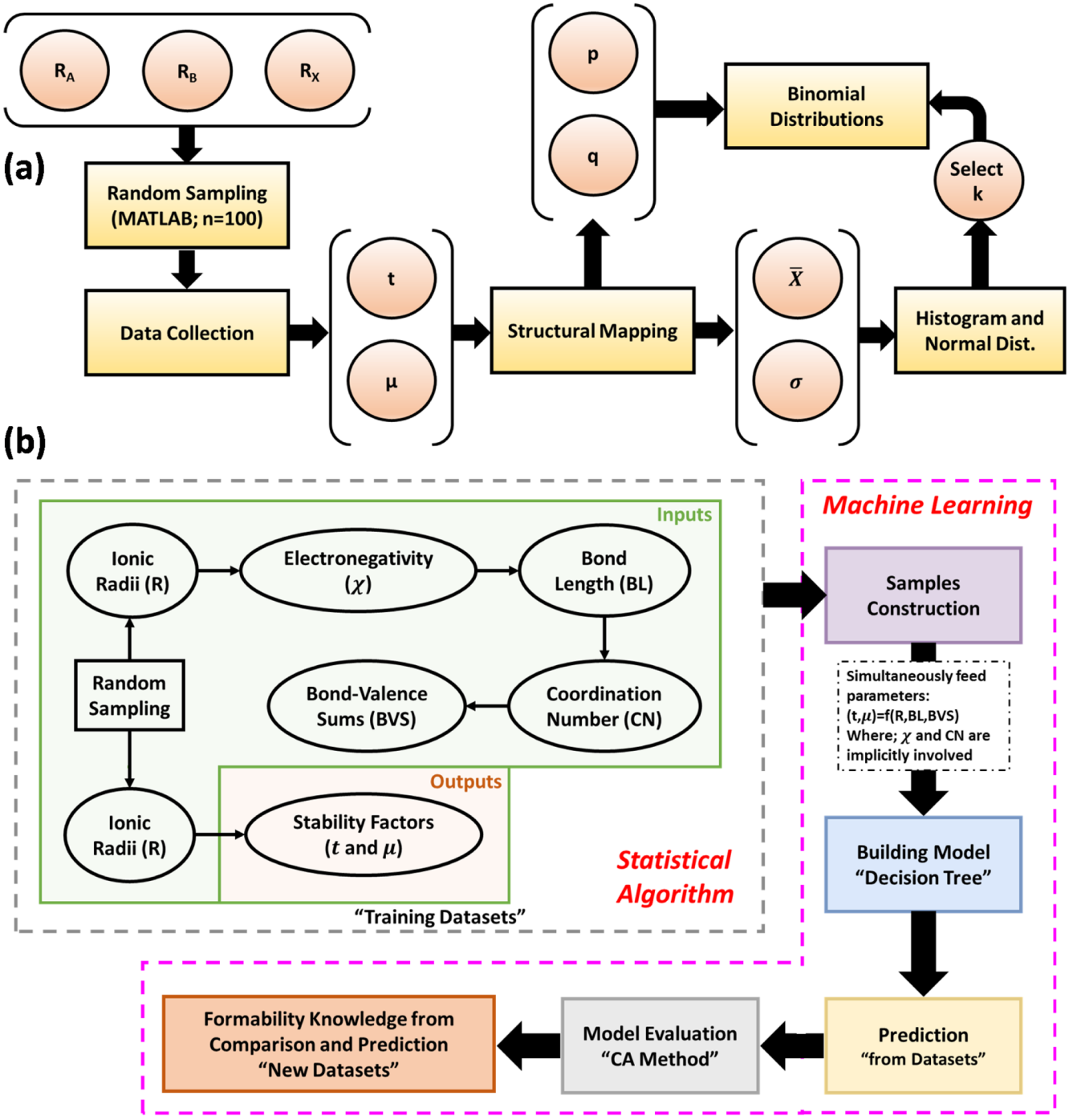

Fig. 2. (a) Theoretical framework used for random sampling of various ionic radii and their probabilistic calculations for formability of Ti-based perovskites; (b) Statistical algorithm proposed for calculating training datasets (inputs/outputs) used for machine learning and "Decision tree" model building analysis.

various ABO_3 *via* "the gradient tree boosting classifier" depending upon A and B ionic radii relative to the radius of O (Shannon radii values [23]) to identify possible structural chemistry for forming perovskite solid crystals; Wagner et al. [40] applied a "Decision tree" algorithm to determine perovskite formability based on gini (GII) impurity (see Supplementary; section 3).

5. Theoretical framework

5.1. Statistical analysis

The theoretical study was initiated by following the suggested framework as shown in Fig. 2(a). Circles in the framework figure refer to input and/or output data where squares indicate the process and/or method considered for data collection and analysis. In brief, radii values of various elements (components) of the Ti-based perovskite structure (A_2BX_6 and/or ABX_3; where A = Cs_2^+, Rb^+, or MA^+; B = Ti^{+2}; and X = F^-, Cl^-, Br^- or I^-) have been randomly selected, mixed, and matched with one another by a random sampling technique using a pre-coded MATLAB program (Random_Generator). Three matrices of [100x1] are randomly generated for each of the 12 studied perovskites from the three available radii ranges of R_A, R_B, and R_X. This is also equivalent to saying that we have generated a one [100x3] matrix for each of the 12 studied Ti-based perovskites to plot the structural maps (total of 12 matrices of [100x3]). Collected data from [4,21,23,26,28,30–34] were then utilized to determine possible tolerance and octahedral factors of various studied Ti-based perovskites (Table 1). Obtained stability factors were used to plot structural maps of different perovskites; hence, giving us a clear idea to distinguish between stable (formable) and non-stable perovskite structures which determines the probability of formation and/or success. The rationale behind the random selection is attributed to the many reasons that can impact the ionic radii of involved ions. Variations in the ionic radii are expected due to inevitable changes in ions oxidation state, electronegativity, covalency, bond length, and coordination from neighboring atoms interactions. Gathered ionic radii values from the three models (e.g. Shannon, Pauling, and Stern) can be randomly mixed with unevenly distributed percentages. This would be translated to a wider range of radii possibilities for checking ABX_3 and A_2BX_6 formability. Since the three models are correct to an extent, using ionic radii mixed from these models will give us much more accurate results mimicking all the possible combinations that could occur. Normal distribution curves (histogram) of both t-factor and μ-factor were plotted. Normal distribution curves must show up for any drawn sample size of > 30 (n > 30). The sample mean and sample standard deviation of the calculated tolerance and octahedral factors (n = 100; for each) were determined to get the population density of the formation. Probability of success as p (formation) and probability of failure as q (instability) obtained from Eq. (6) and Eq. (7), respectively, were used in the binomial distribution analysis at different desired outcomes (k = 20, 40, 60, 80, 100). The desired outcome (k) refers to the number of successfully formed perovskites (success) in the sample size n = 100 (e.g. if k = 40 refers to getting 40 successful trials (formation) out of 100 calculated from p and q). Normal and binomial probabilities of the formation along with Nash equilibrium and IDSDS strategies were used later to select the most stable combinations.

5.2. Machine learning

A proposed statistical algorithm was used to calculate the impact of various randomly sampled ionic radii (see Table S1 and Figure S1) on the crystal chemistry parameters, Fig. 2(b). Then, obtained radii values were utilized to get perovskite crystal chemistry data (inputs) and stability factors (outputs) for constructing training datasets. Briefly, random sampling of ionic radii is our starting point to identify the impact of ionic radii size on structural stability attributed to changes in input parameters as ions/atoms electronegativity, bond length, coordination number, and bond-valence sums (oxidation numbers). Regarding the machine learning framework, determined training datasets were simultaneously fed into a MATLAB classification learner "Decision trees". Only ionic radii, ions/atoms bond length, and BVS are considered as explicit inputs since atoms electronegativity and CN are implicitly involved in Eq. (11) and Eq. (13). A fine tree classifier learner was initiated with our training datasets to build four separate "Decision tree" models for A_mTiI_{3m}, A_mTiF_{3m}, A_mTiCl_{3m}, and A_mTiBr_{3m}, where m = 1 for A = MA^+ or Rb^+ and m = 2 for A = Cs^+. We had four generated [300x10] matrices as future/trained datasets for the studied A_mTiX_{3m} structure. The first, second, and third 100 row-values in the matrices were obtained for MA^+, Cs^+, and Rb^+, respectively, from their known ionic radii ranges (see Table S1 and Figure S1). Predictions of structural stabilities of future datasets (obtained from another trial of random sampling) were determined from the built "Decision trees". Model accuracy and evaluation were carried out by five cross-validation (CV) split as well as classification accuracy (CA) of the trained model, reported in Eq. (14).

6. Equations and methods

t-factor and μ-factor were calculated for the different studied Ti-based perovskites from their ionic radii from Eq. (1) and Eq. (2), respectively, as reported in earlier works [10,21,22]. Structural maps of the Ti-perovskites were plotted and then utilized in carrying out normal and binomial probability distributions (discrete). Obtained probabilities of formation from stability factors would allow us to focus more on studying the samples that showed stability/instability behavior (p ≠ 1). In the results section, we will find that six samples were showing stability/instability behavior based on their probabilities of formation: Cs_2TiI_6, $MATiF_3$, $MATiCl_3$, $MATiBr_3$, $MATiI_3$, and $RbTiI_3$ (or in short, Cs_2TiI_6, $MATiX_3$, and $RbTiI_3$) which have been studied further to determine their normal and binomial probability distributions.

$$t = \frac{R_A + R_X}{\sqrt{2}\,(R_B + R_X)} \tag{1}$$

$$\mu = \frac{R_B}{R_X} \tag{2}$$

The sample mean and sample standard deviations were obtained for the Ti-perovskites with p < 1 from Eq. (3) and Eq. (4), respectively, to carry out normal distribution and histogram analysis. Thus, normal distribution and histogram curves can be plotted for the six stable/unstable Ti-based perovskites (p < 1) from Eq. (5). Probability normal distribution allows us to determine the population percentage that fits within the desired ranges of t and μ [41,42].

$$\bar{X} = \frac{\sum x}{n} \tag{3}$$

$$\sigma = \sqrt{\frac{\sum (x - \bar{X})^2}{n - 1}} \tag{4}$$

$$y = \frac{e^{-(x-\bar{X})^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}} \tag{5}$$

The probability of formation and/or instability of Ti-based perovskites have been defined in Eq. (6) and Eq. (7), respectively, as the product and/or summation of the probabilities of having t-factor and μ-factor within our stability/instability ranges. The general binomial probability formula used in the binomial distribution analysis is shown in Eq. (8) which gives us an estimation about the mass distribution of successful trials (formation) at various k desired successful outcomes (0 to 100) from the total sample size (n = 100) [41,42]. Several assumptions have been made to perform our probabilistic analysis, including: (i) discrete distributions, (ii) population is infinite (N→∞), (iii)

sample sizes of various ionic radii are 100 numbers (n = 100) chosen by the random sampling method using a pre-coded MATLAB random generation program; hence, according to the central limit theorem, (iv) our probability distributions of all samples are expected to approach a normal distribution behavior (this is only true if n > 30) [41,43,44].

$$p = p_t\, p_\mu = \left(\frac{t_{in}}{t_n}\right)\left(\frac{\mu_{in}}{\mu_n}\right) \tag{6}$$

$$q = q_t + q_\mu \approx 1 - p = 1 - \left[\left(\frac{t_{in}}{t_n}\right)\left(\frac{\mu_{in}}{\mu_n}\right)\right] \tag{7}$$

$$P(k) = \frac{n!}{k!(n-k)!}p^k(1-p)^{(n-k)} \tag{8}$$

Selection strategies and combinations of various ions have been done by applying the mix/match strategy (treating cations and anions as players based on game theory concepts [45]). Studied Ti-based perovskites are as shown in Table 2 which involve mixing/matching between four halide anions and three organic/inorganic cations with Ti as a metal cation. Nash equilibrium from best replies (p*, P(k)*) and IDSDS methods, as in Eq. (9) and Eq. (10) [45–47], were applied to our studied perovskites to choose the best strategies that maximize our formation probabilities (probability and binomial probability); hence, we would be able to identify the best selections yield in the formation of Ti-based perovskites.

Nash equilibrium from best replies for player 1 (A ions):

$$v_1(p^*, P(k)^*) \geqslant v_1(p, P(k)^*) \tag{9}$$

Nash equilibrium from best replies for player 2 (X ions):

$$v_2(p^*, P(k)^*) \geqslant v_2(p^*, P(k)) \tag{10}$$

For the statistical algorithm used in the machine learning analysis, the bond length of involved covalent and hydrogen bonds in the BX_6 octahedral was calculated from Schomaker-Stevenson Eq. (11) [48,49]; this equation accounts for the impact of high electronegativity differences in the hetero-diatomic configuration. The bond-valence model shown in Eq. (12) and Eq. (13), as reported by Brown [50], determines bond lengths with consideration of coordination numbers – which would identify deviations of the theoretical BVS from our predicted values. The number of bonds exists in the studied atoms yield in the determination of oxidation numbers. Briefly, a valence assigned to an ion is the sum of the bond valences associated with covalent bonds with neighboring ions of the opposite charge. The parameter b has a nearly universal value of 0.4 Å, where BL_0 is taken from tabulated anion-cation pairs bond lengths defining chemical bonding and crystal stability [38].

$$BL_{ij} = R_i + R_j - 0.09(\chi_i - \chi_j) \tag{11}$$

$$BVS = \sum v_{ij} \tag{12}$$

$$v_{ij} = \exp\left[(BL_0 - BL_{ij})/b\right] \tag{13}$$

Classification accuracy (CA) of the observed machine learning models was evaluated from using Eq. (14) as reported elsewhere [35]; where S and N denote the number of samples that are correctly classified and the total number of samples, respectively.

$$CA = S/N \tag{14}$$

7. Results and discussion

7.1. Stability factors and probability of formation

Structural mapping for Cs_2TiX_6, $MATiX_3$, and $RbTiX_3$, as shown in Fig. 3, estimated that Cs_2TiX_6 and $RbTiX_3$ are formable Ti-based perovskites at all times (for any radii combinations) except when halide anion is chosen to be I^- (~0.5 formation probability). Ionic radii of F^-, Cl^-, and Br^- are smaller than those radii values of I^- halide anions (Table 1) which would allow Ti^{+2} ions to enter the octahedral structure occupying B-sites. This explains the formability of Cs_2TiX_6 and $RbTiX_3$ for X = F^-, Cl^-, Br^- (μ-factor = 0.45 ~ 0.75 within the formability range). The t-factor condition was met with no stability problems for inorganic Cs^+ and/or organic Rb^+ cations in occupying the A-sites between the eight octahedral structures. Cs^+ and Rb^+ cations with reasonably small radii sizes compared to MA cations allow them to enter cubic corner cavities in the BX_3 framework (Fig. 3 and Table 3). However, if the μ-factor condition is not met, the formed structure would not be considered as a perovskite material. In terms of $MATiX_3$, both t-factor and μ-factor played a key role in determining $MATiX_3$ structural probabilities owing to the larger radii range of MA cations (affecting t-factor) and relatively large iodide ions (affecting μ-factor). Thus, more controversial results were found for $MATiX_3$ with formability probabilities in the range 0.465 to 0.87 dependent upon the selected halide anion. Very small ionic radii in F^- and Cl^- would make it easy for MA^+ cations to escape from A-sites due to the much smaller octahedral structures; whereas very large ions as in I^- would overlap in the octahedral framework blocking the A-site region from being occupied by MA. Therefore, the optimal choice in $MATiX_3$ determined to be with the use of Br^- anions giving us an ideal t-factor range with a maximum formation probability of 0.87 (Fig. 3 and Table 3).

Probability of formation, probability of instability and the corresponding t-factor and μ-factor mean and sample standard deviation have been theoretically calculated for all the studied Ti-based perovskites as shown in Table 3. Probabilities of formation were obtained from the product of both p_t and p_μ (being in the range from Eq. (6)) since those were two independent events occurring simultaneously when forming perovskites.

7.2. Statistical normal distribution and histogram

Normal distributions of t-factor and μ-factor for metastable Ti-perovskites (Fig. 4 and Table 3) perfectly fitted the histogram bell curves indicating the reliability of randomization of generated samples. From mean and standard deviation in normal distribution curves (Fig. 4, Table 3, Table S2, and Table S3), frequency of occurrence (population density) met the t-factor condition in all combinations with t = 0.856 ~ 1.042 (except for $MATiF_3$). However, mean values of μ = 0.438 ~ 0.690 found to be sometimes out of the stable range (μ < 0.438) lowering formation probabilities in specific combinations (e.g. Cs_2TiI_6, $MATiI_3$, and $RbTiI_3$ with p ≤ 0.5). In metastable Ti-perovskites, the highest formation probabilities were observed for $MATiBr_3$ (0.87) and $MATiCl_3$ (0.81) as shown in Fig. 5(a), respectively. This is attributed to the optimum ionic radii of mixed cations and anions. In general, stable Pb-based perovskites have larger Pb^{+2} radius size and higher electronegativity than Ti^{+2} leading to μ-factors within the stability range. From normal distribution curves and (68%–95%–99.7%) statistical rule, mean and sample standard deviations of μ and t were used to determine the population percentage meeting the stability constraints; Fig. 5(b). Combinations with the highest population density of formation (stability percentage for both μ and t > 67%) recorded the highest formation in agreement with the predetermined probabilities for $MATiBr_3$ and $MATiCl_3$ (see Supplementary for explanations

Table 2
Studied Ti-based perovskites from selected organic/inorganic cations and halide anions.

Cation \ Halide Anion	F^-	Cl^-	Br^-	I^-
Cs^+	Cs_2TiF_6	Cs_2TiCl_6	Cs_2TiBr_6	Cs_2TiI_6
MA^+	$MATiF_3$	$MATiCl_3$	$MATiBr_3$	$MATiI_3$
Rb^+	$RbTiF_3$	$RbTiCl_3$	$RbTiBr_3$	$RbTiI_3$

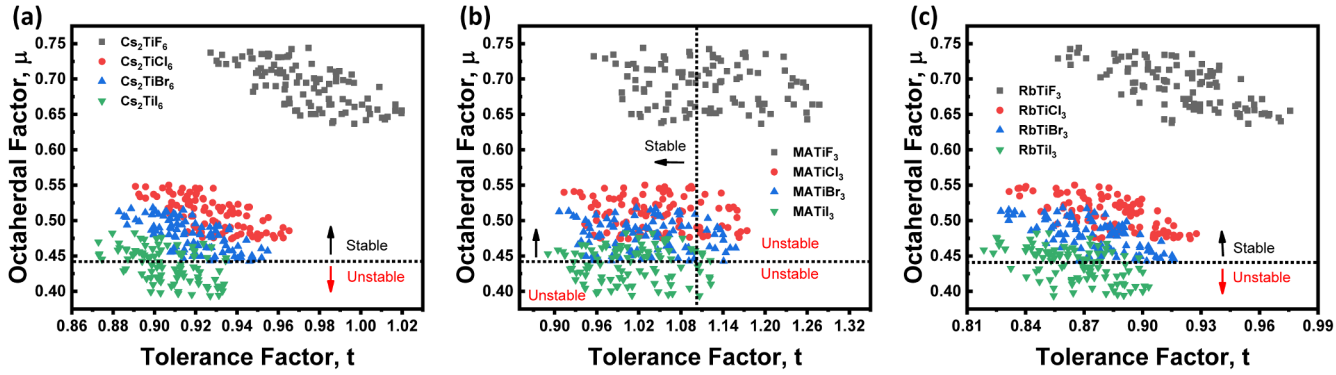

Fig. 3. Structural maps of the various studied Ti-based perovskites: (a) Cs_2TiX_6; (b) $RbTiX_3$; (c) $MATiX_3$.

Table 3
Formation and instability probabilities of Ti-based perovskites with mean and standard deviation of identified t-factor and μ-factor from random sampling.

Ti-Perovskite	Probability of formation (p)*	Probability of instability (q)*	Sample Mean ($\bar{X}$)*	Sample Std. Deviation (σ)*
Cs_2TiF_6	1, 1 = 1	0, 0 = 0	0.973, 0.690	0.022, 0.030
Cs_2TiCl_6	1, 1 = 1	0, 0 = 0	0.927, 0.509	0.017, 0.023
Cs_2TiBr_6	1, 1 = 1	0, 0 = 0	0.918, 0.478	0.016, 0.022
Cs_2TiI_6	1, 0.5 = **0.5**	0, 0.5 = **0.5**	0.905, 0.438	0.015, 0.023
$MATiF_3$	0.49, 1 = **0.49**	0.51, 0 = **0.51**	1.113, 0.690	0.079, 0.030
$MATiCl_3$	0.81, 1 = **0.81**	0.19, 0 = **0.19**	1.042, 0.509	0.065, 0.023
$MATiBr_3$	0.87, 1 = **0.87**	0.13, 0 = **0.13**	1.028, 0.478	0.062, 0.022
$MATiI_3$	0.93, 0.5 = **0.465**	0.07, 0.5 = **0.57**	1.009, 0.438	0.058, 0.023
$RbTiF_3$	1, 1 = 1	0, 0 = 0	0.919, 0.690	0.027, 0.030
$RbTiCl_3$	1, 1 = 1	0, 0 = 0	0.882, 0.509	0.022, 0.023
$RbTiBr_3$	1, 1 = 1	0, 0 = 0	0.875, 0.478	0.021, 0.022
$RbTiI_3$	1, 0.5 = **0.5**	0, 0.5 = **0.5**	0.865, 0.438	0.020, 0.023

*(#, #) = (t, μ); probability of formation and/or instability is totally dependent on t and μ; $p = p_t * p_\mu$ (independent events for formation); $q = q_t + q_m$ (dependent and cumulative events for instability).

regarding the other metastable structures; section 6).

7.3. Statistical binomial distribution

According to the binomial distribution analysis in Fig. 5(c), it was observed that the highest probabilities were determined for $MATiCl_3$ (0.096 @ k = 80) > $MATiI_3$ (0.035 @ k = 40) > $MATiF_3$ (0.016 @ k = 40) > $MATiBr_3$ (0.015 @ k = 80) > Rb-based and Cs-based perovskites (0.01 @ k = 40, 60); (see Table S4 and Figure S2). Hence, one may conclude that perovskite formability would be best for $MATiCl_3$ > $MATiBr_3$ > (Cs_2TiI_6, $MATiF_3$, $RbTiI_3$) > $MATiI_3$ regarding the metastable Ti-based combinations according to the binomial distribution analysis (see detailed explanations in Supplementary; section 7).

7.4. Game theory and Nash equilibrium decisions

Calculations associated with game theory concepts were carried out using Eq. (9) and Eq. (10); yielded in giving the "Nash equilibrium options" that are the best selections (decisions) based on the best responses of both players (A^+Ti^{+2} and X^-) among the quasi-stable Ti-perovskites. The strategic form of the game is shown in Table 4 where payoff values (utilities) for every player were assigned from the corresponding probabilities. The best reply (BR) method determined the highest utility values (1, 1) attributed to the best decisions. Best selections would give us the highest probabilities of formation where the utility of 1 indicates 100% formability for the corresponding mixed/matched ions. Nash equilibrium or dominant strategies in Table 5 were found for the stable Cs_2TiX_6 and $RbTiX_3$ for X = F^-, Cl^-, Br^- in agreement with probabilistic analysis (Table 3).

In other words, Nash equilibrium gave us the six formable structures (Cs_2TiF_6, Cs_2TiCl_6, Cs_2TiBr_6, $RbTiF_3$, $RbTiCl_3$, $RbTiBr_3$) since ionic radii sizes in those compounds will yield in a 100% probability of formation. IDSDS method showed the same six strategies because IDSDS does not lose dominant Nash strategies [45–47]. Deletions of dominated strategies (IDSDS) occur sequentially as (i) $Cs_2^+Ti^{+2}$ and Rb^+Ti^{+2} strictly dominates MA^+Ti^{+2} for player 1 (cation-Ti), and (ii) F^-_{3m}, Cl^-_{3m}, and Br^-_{3m} strictly dominate I^-_{3m} for player 2 (halide anion) that occurs after deletion of MA^+Ti^{+2}. The six non-Nash dominated strategies in Table 4 show stable/unstable structures with highest to lowest formation probability as $MATiBr_3$ > $MATiCl_3$ > (Cs_2TiI_6 and $RbTiI_3$) > $MATiF_3$ > $MATiI_3$. Ti-perovskites with iodides had the lowest formation tendency because of the large iodide anions (> 2Å) that might overlap in the octahedral. The undesired overlapping prevents Ti cations from occupying B-sites as observed in the earlier analysis.

7.5. Machine learning and supervised decision tree analysis

Input/output pairs of various Ti-based perovskites (Fig. 6 and Figure S1) were used for fine tree training (Decision trees) *via* the classification learner toolbox in MATLAB relying on patterns and inference. Tree training algorithms were carried out *via* 7 predictors (R_A, R_B, R_X, BL for Ti-X, BL for X···H, BVS for Ti and BVS for X) and 2 responses for structural formability (Yes/No); trained data were used to build mathematical decision tree models for the formability of future input datasets (Figure S1) with high predictions accuracy (96.3% ~ 99%). We believe that BL of Ti-X and X···H in the BX_6 octahedral would explain the stability of the octahedral surrounded by the cubical structure, respectively (Figures S3-S6). Minimum and maximum BLs of Ti-X and X···H were estimated from ionic radii and their corresponding electronegativities (Table S5). From electronegativities: F^- (3.98) > Cl^- (3.16) > Br^- (2.96) > I^- (2.66) > H^+ (2.2) > Ti^{+2} (1.54) [25,51,52], they showed prominent impact on BL as shown in Table 6. With the consideration of maximum ionic radii (in Å): I^- (2.2) > Br^-

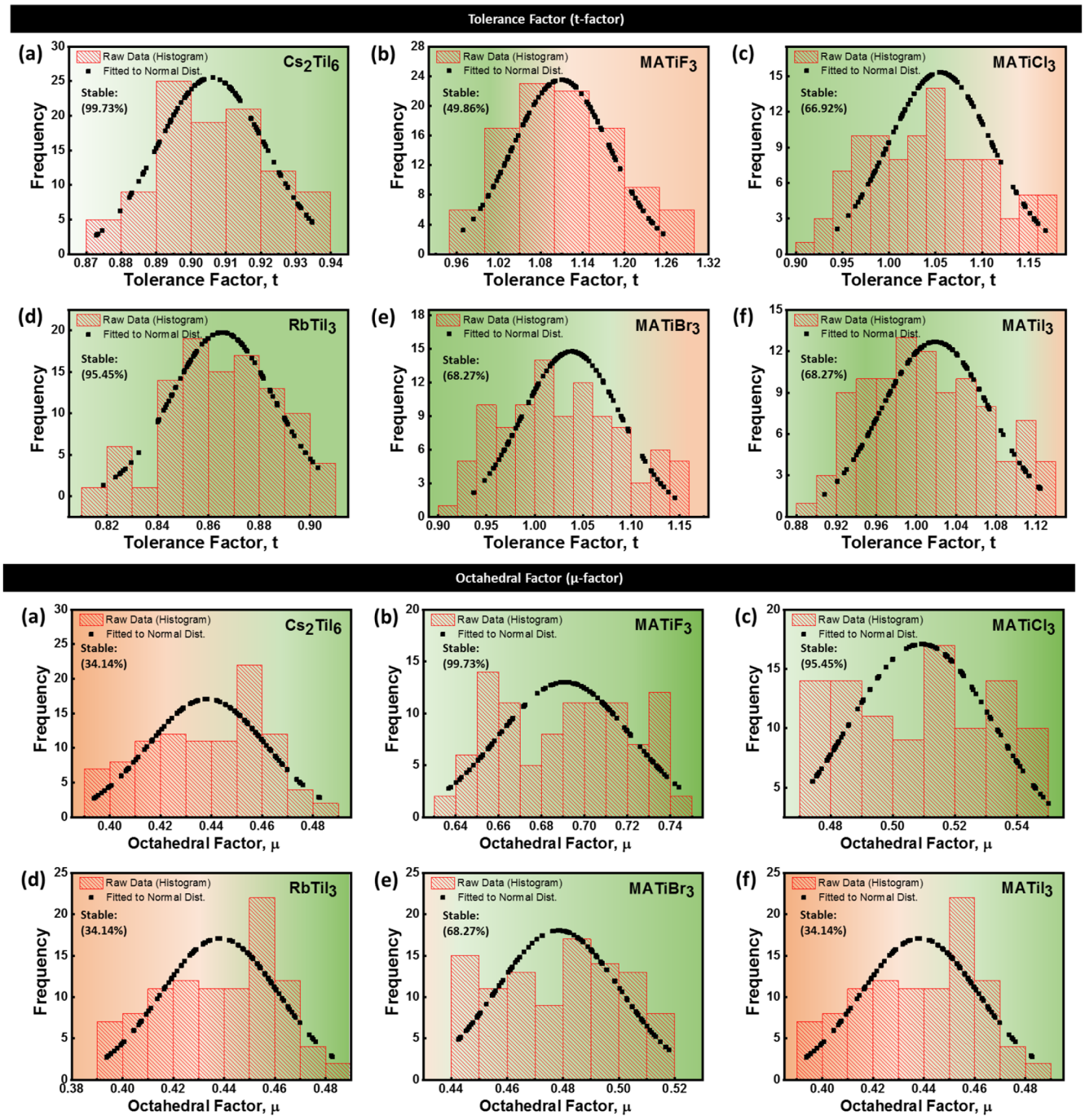

Fig. 4. t-factor and μ-factor histogram curves and normal distribution fittings of unstable Ti-perovskites from random mixing/matching of selected ionic radii (random samplings): (a) Cs_2TiI_6; (b) $MATiF_3$; (c) $MATiCl_3$; (d) $RbTiI_3$; (e) $MATiBr_3$; (f) $MATiI_3$. Background color green for stable and red for unstable populations.

(1.96) > Cl^- (1.84) > F^- (1.36) > Ti^{+2} (1) > H^+ (0.012), we conclude that iodide-based perovskites have the longest BLs with a maximum of 3.3 Å for Ti-I and 2.17 Å for I···H among the other halides. Minimum and maximum H^+ ionic radii are approximated as 8.4×10^{-6} Å [53] and 0.012 Å [23,26].

The large iodide BLs and low electronegativity could be the reason behind the addressed instabilities in the BX_6 iodide-octahedrons. Precisely, Seshadri [54] defined that BVS should be equal to the formal oxidation state of the cation or anion to have a formable/stable structure. If BVS for X = −1 and Ti = +2, (that is for each Ti atom, there will be two attached iodides) there will be a complete stable structure. However, our BVS analysis for iodide-based structures confirms that both X and Ti BVS have a wide range of possibilities (Table 6 and Table S6). The wider the BVS range, the lower the stability we might have due to the higher chances for Ti to be over-bonded or under-bonded with iodide.

Pb-free $ATiX_3$ perovskites with A = MA^+ or Rb^+ are identified with formability of stable/unstable structures (from highest to lowest) as $ATiI_3$ > $ATiBr_3$ > $ATiCl_3$ > $ATiF_3$ which were in good agreement with our statistical and game theory analysis. In $ATiI_3$, the observed high deviated BVS agrees with the high structural instabilities (low stabilities in I-based perovskites could be from the wide BVS ranges of Ti-I and I···H found as 1.31 and 0.66, respectively).

According to the absolute deviations of BVS from formal oxidations in Cs-based perovskites, stability from high to low is found to be as $CsTiCl_3$ > $CsTiBr_3$ > $CsTiF_3$ > $CsTiI_3$ (results revealed good

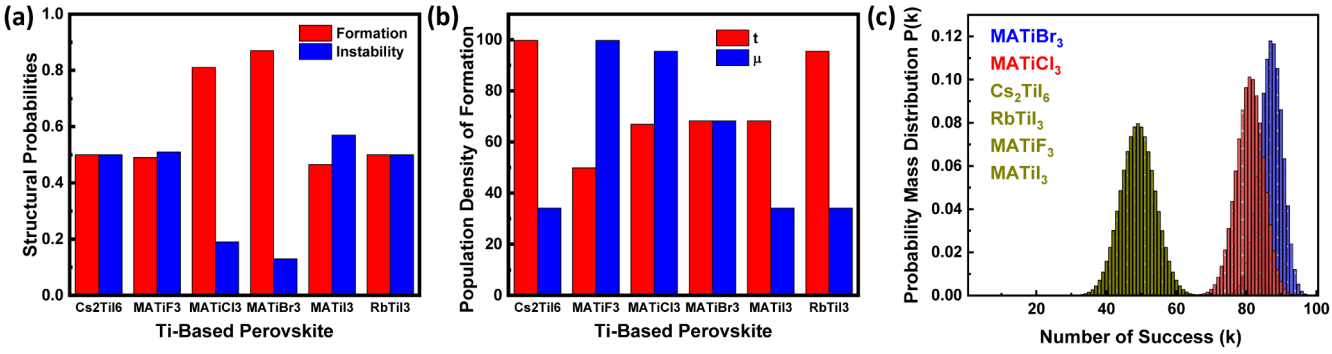

Fig. 5. Probabilistic and distribution analysis for stable/unusable (probably formable) Ti-based perovskites: (a) Structural probabilities of formation; (b) Population density meeting our formability constraints and stability factors; (c) Binomial distribution curves and probability of formation mass distribution, P(k), at (k = 0 – 100) and (n = 100).

Table 4
Strategic form of the game (Ti-based perovskite formation) showing the possible decisions (mix/match) with their assigned utilities calculated from probability and binomial probability of formation*

Cation-Ti / Halide Anion	F^{-}_{3m}	Cl^{-}_{3m}	Br^{-}_{3m}	I^{-}_{3m}
$Cs_2^{+}Ti^{+2}$	1, 1	1, 1	1, 1	0.5, 0.08
$MA^{+}Ti^{+2}$	0.49, 0.08	0.81, 0.1	0.87, 0.12	0.465, 0.08
$Rb^{+}Ti^{+2}$	1, 1	1, 1	1, 1	0.5, 0.08

*m = 2 for $Cs_2^{+}Ti^{-2}$; otherwise m = 1; utilities refer to the payoff and/or satisfaction and preference of both players (cations/anions) defined as: (#, #) = "$p, P(k)$" where $P(k)$ was determined from binomial probabilities at the mean = np.

Table 5
Nash equilibrium analysis showing dominant strategies and the highest formation probabilities among the studied Ti-based perovskites.

Cation-Ti / Halide Anion	F^{-}_{3m}	Cl^{-}_{3m}	Br^{-}_{3m}
$Cs_2^{+}Ti^{+2}$	1, 1	1, 1	1, 1
$Rb^{+}Ti^{+2}$	1, 1	1, 1	1, 1

agreement of models with previous analysis). Machine learning predicted that structural formability of Cs_2TiX_6 > $RbTiX_3$ > $MATiX_3$ described by the higher CN and lower BVS deviations in Cs-based perovskites. Among all the studied combinations, Cs-based perovskites show the longest BLs and lowest BVS deviations owing to the greater CN (8 ~ 12) of Ti cuboctahedral atoms. Fig. 6(i,j) shows an illustration of the taken CN for Ti and X in the two studied perovskite structures. In a unit cell of A_mTiX_{3m} cuboctahedral structure, the optimal CN of X should be (3 for m = 1, and 6 for m = 2); where CN selections for Ti would be ideal with (6 for m = 1 and 12 for m = 2) knowing that m = 2 for A = Cs^{+}; elsewhere, m = 1 (for A = MA^{+} and Rb^{+}) as explained in a previous work [56].

Probabilities of formation of training datasets concerning used halides in A_mTiX_{3m} were identified as 0.49, 0.83, 0.94, and 0.96 for I^{-}, F^{-}, Cl^{-}, and Br^{-}, respectively, matching our earlier results. To check for the validity of determined machine learning "Decision trees" models, new randomly generated datasets showed estimated formation probabilities as 0.52, 0.79, 0.91, and 0.93 for I^{-}, F^{-}, Cl^{-}, and Br^{-}, respectively (Fig. 7). These results were comparable to the actual calculated probabilities 0.44, 0.80, 0.90, and 0.93 for I^{-}, F^{-}, Cl^{-}, and Br^{-}, respectively. Moreover, similarities in probabilities of formation between training and future datasets were identified with a marginal error of ± 8%. Hence, formability probability (from high to low) were found to be A_mTiBr_{3m} > A_mTiCl_{3m} > A_mTiF_{3m} > A_mTiI_{3m} (Fig. 7).

We might generalize that it is optimal to have ions with electronegativity and ionic radius close to the Br^{-} as 2.96 and 1.96 Å, respectively (i.e. electronegativity-to-ionic-radius ratio ~ 1.51 $Å^{-1}$). In contrast, low $Å^{-1}$ ratio in I^{-} (1.2) or high $Å^{-1}$ ratio in Cl^{-} (1.71) and F^{-} (2.96) are not favorable for having a balanced cuboctahedral due to not yet known halide interactions or chemical bonding with neighboring ions/atoms.

The supervised machine learning model "Decision tree" classified our discrete inputs into two categories (Yes/No) for formability possibility. Model results were illustrated in branch/node plots representing outcomes (class labels) in the branches according to chosen predictors or controlling stability parameters in the nodes. Decision trees analysis of the training datasets (Fig. 7 and Table S7) revealed the controlling parameters responsible for formability of the different halide perovskites, which found to be R_A in (A_mTiBr_{3m}, A_mTiCl_{3m}, A_mTiF_{3m}) and R_B in A_mTiI_{3m} describing the importance of ionic radii role in defining perovskite structural stabilities (e.g. Ti atoms properties determine iodide-based perovskite stabilities). Classification accuracy of predicted results showed high model reliability of 99%, 98.3%, 96.3%, and 97.7% for A_mTiBr_{3m}, A_mTiCl_{3m}, A_mTiF_{3m}, and A_mTiI_{3m}, respectively (obtained from plotting R_B vs. R_X, see Supplementary; section 10). As we go down the trees (Fig. 7), defined predictors for input data classification become less important to impact perovskites stability. Analyzed predictors used in machine learning analysis include R_A, R_B, R_X, BL for Ti-X, BL for X···H, BVS for Ti in Ti-X and BVS for X in X···H bonds (where BLs and BVS determine BX_6 stability). Using a predictor importance function in MATLAB, we estimated the relative importance of the seven predictors in controlling perovskites formability. Predictor importance results revealed that R_A (0.1244 compared to 0 for other predictors) is the only important parameter for studying the formability of A_mTiBr_{3m}. However, A_mTiCl_{3m} can be stabilized *via* tuning R_A (0.0343), R_B (0.0065), or R_X (0.0015) where numbers refer to predictor importance relative to 0 values in the remaining parameters. Interestingly, both R_A (0.1606) and BVS of F (0.0033) are the controlling formability indicators in A_mTiF_{3m}. As expected, a more complicated case was observed in A_mTiI_{3m} where all ionic radii R_A (0.0014), R_B (0.0602), R_X (0.0093), and BL for Ti-I (0.0076) are linked to the cuboctahedral stability. In summary, formability of A_mTiBr_{3m} is much easier to control by only changing R_A to stabilize the unstable cubic structures surrounding stable octahedrons in the crystal lattice; A_mTiCl_{3m} requires 3 parameters to control instability issues by changing/matching possible A ions with Ti with more emphasis on forming the cubical structures; A_mTiF_{3m} could be formed by tuning fluoride BLs mixed with different oxidation states of Ti to achieve formable octahedrons; and lastly, A_mTiI_{3m} stability is more controversial and might be controlled by changing the 3 ionic radii sizes or tuning Ti-I BL for enhanced octahedral stability allowing cubical structure to form afterward.

Comparisons of our findings with the reported experimental results will show the validity of the constructed theoretical algorithm. Chen

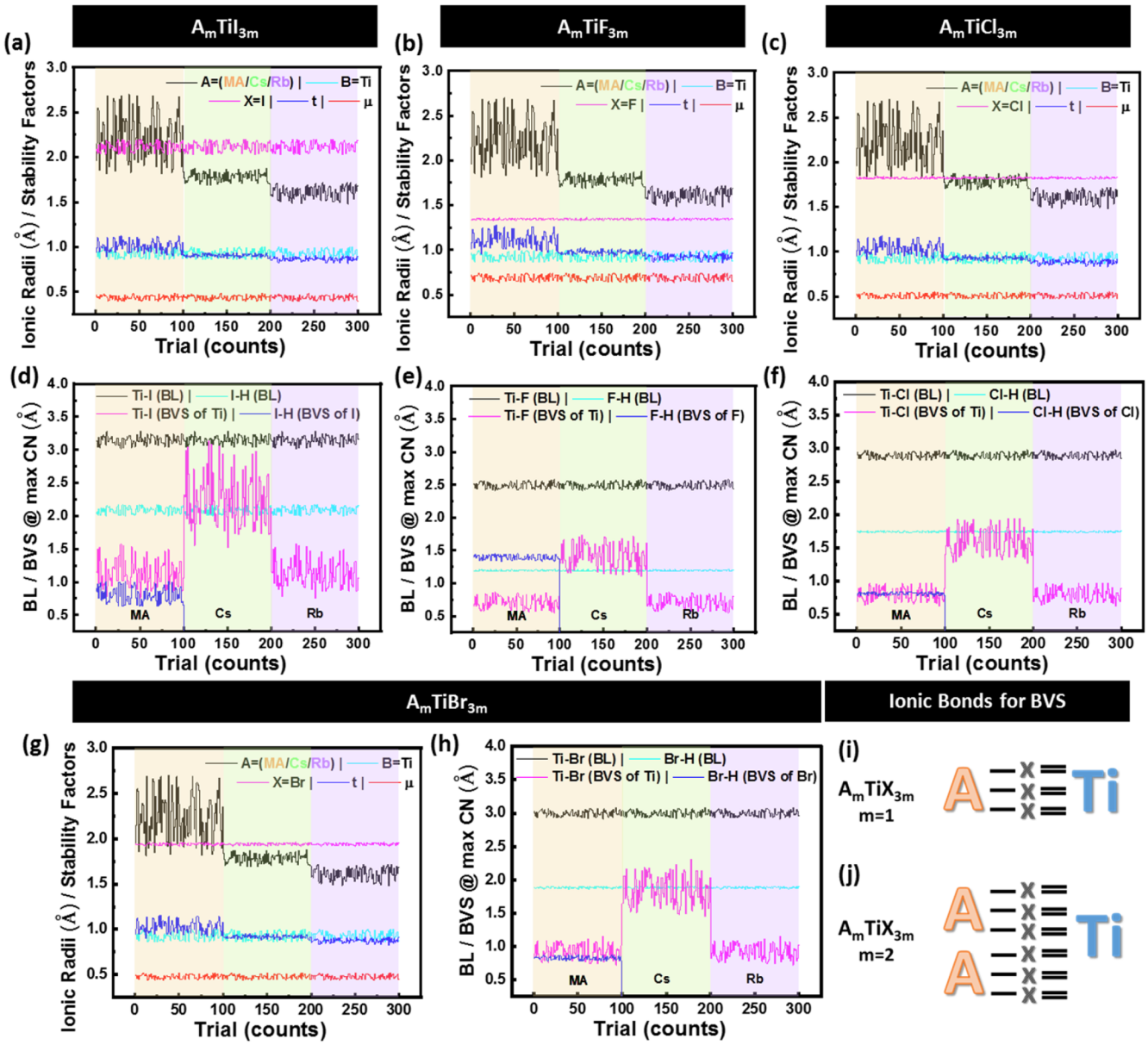

Fig. 6. Training datasets used in the machine learning analysis obtained from inputs (R, BL = f(R, χ), and BVS = f(R,CN)) and response (stability factors t and μ as outputs) from random sampling of ionic radii in Ti-based perovskites: (a) and (d): A_mTiI_{3m}; (b) and (e): A_mTiF_{3m}; (c) and (f): A_mTiCl_{3m}; (g) and (h): A_mTiBr_{3m}. Note that m = 2 for A = Cs^+; elsewhere, m = 1 (for A = MA^+ and Rb^+); background colors orange for MA, green for Cs, and purple for Rb. For BVS calculations, ionic bonds and used coordination numbers (CN) are estimated from (i) for MA-Ti-based and Rb-Ti-based perovskites and (j) for Cs-Ti-based perovskites.

et al. (2018) [57] fabricated thin-film-double perovskites Cs_2TiBr_6 with 1.8 eV bandgap showing a highly stable structure (with solar conversion efficiency ~ 3.3%) matching our estimated structural stability (balanced cuboctahedral) for (Ti-Br)-based perovskites. The association between highly stable structure and estimated structural stability can be thought of as the right occupation and/or strong attachment of A, B, and X ions to their corresponding ionic sites (i.e. the structure is thermodynamically in equilibrium when phases are stable with lowest energy state due to ion migration to their intended sites form ions potential gradients creating a cuboctahedron). High stability refers to minimal loss/gain of thermal energy preventing salt formation degradation reactions arising from internal stresses due to UV or high-temperature environment. Other studies estimated the possibility of creating thermodynamically stable Ti-based perovskites (with high-pressure phase equilibria) from using Ca inorganic cations in $CaTiO_3$ [58–60] and $CaTiO_3$-$CaSiO_3$ [61].

The thermodynamic phase stability is a key parameter governing the perovskite synthesizability. Future works on the evaluation of Ti-based perovskite thermodynamic stability (lowest energy state) are important to validate our results. A common computational approach involves studying the stability from phase diagrams (e.g. using Pymatgen) determined from DFT-[convex hull] (DFT-CH). Phase stability is determined by constructing the energy convex hull of a given region of composition space. The energy above the convex hull (E_{hull}) provides a direct measure of the stability [62]. The convex hull consists of phases with energy lower than any other phase or a linear combination of phases at the respective compositions [63]. E_{hull} (meV/atom) is a measure of the decomposition energy of the compound into a linear combination of the stable phases. Thermodynamically stable compounds exhibit an E_{hull} of zero (i.e., they are on the convex hull and are stable with equilibrium phases at T = 0 K lower in energy than any other phase or combination of phases) [63]. More positive values of E_{hull} indicate decreasing stability [64]. The stability or convex hull distance can be defined as $[\Delta H_{stab}^{ABX_3} = \Delta H_f^{ABX_3} - \Delta H_f]$, where $\Delta H_f^{ABX_3}$ is

Table 6
Bond length, oxidation state, and coordination numbers for BVS investigation in the involved structural bonding in the BX_6 octahedral of the studied Ti-based perovskites.

Perovskite Structure	Bond	BL (Å)*	Ox.*	CN (No. of Bonds)* Ti	X	L_0*	BVS for Ti or X*	Dev. of BVS (%)*
$MATiX_3$ / $RbTiX_3$	Ti-F	2.41 ~ 2.58	2	3,4,5,6	1,2,3	1.72	0.44 ~ 0.87	−77.9 ~ −56.6
	F···H*	1.17 ~ 1.21	1	N/A	1,2,3	0.92	0.50 ~ 1.45	−50~+45
	Ti-Cl	2.82 ~ 2.99	2	3,4,5,6	1,2,3	2.17	0.50 ~ 0.97	−75.1 ~ −51.5
	Cl···H*	1.72 ~ 1.77	1	N/A	1,2,3	1.28	0.29 ~ 0.85	−70.8 ~ −15
	Ti-Br	2.91 ~ 3.09	2	3,4,5,6	1,2,3	2.32	0.59 ~ 1.15	−70.7 ~ −42.5
	Br···H*	1.85 ~ 1.90	1	N/A	1,2,3	1.42	0.30 ~ 0.87	−69.8 ~ −13
	Ti-I	2.99 ~ 3.30	2	3,4,5,6	1,2,3	2.54	0.86 ~ 2.17	−57.1~+8.5
	I···H*	1.99 ~ 2.17	1	N/A	1,2,3	1.61	0.35 ~ 1.01	−65.1~+1.0
Cs_2TiX_6	Ti-F	2.41 ~ 2.58	2	8,9,10,11,12	1,2,3	1.72	1.11 ~ 1.73	−44.5 ~ −13.5
	Ti-Cl	2.82 ~ 2.99	2	8,9,10,11,12	1,2,3	2.17	1.27 ~ 1.95	−36.5 ~ −2.5
	Ti-Br	2.91 ~ 3.09	2	8,9,10,11,12	1,2,3	2.32	1.45 ~ 2.31	−27.5~+15.5
	Ti-I	2.99 ~ 3.30	2	8,9,10,11,12	1,2,3	2.54	1.52 ~ 3.14	−24~+57

*H-bonds only exist in the $MATiX_3$ structure; BL: Bond length; Ox.: Oxidation number and is either for Ti or X [for Ti in Ti-X bonds and halide anions or X in X···H bonds]; CN: Coordination number referring to number of bonds attached to the atom of interest; L_0: Bond-valence parameter of specified bonds taken from tabulated values in [55]; BVS: Bond-valence sums for either Ti in Ti-X or X in X···H bonds; Dev.: Deviation of BVS from formal oxidation states of anions or cations.

the formation energy of ABX_3 calculated from the total energy of the perovskite, E(ABX_3), from $[\Delta H_f^{ABX_3} = E(ABX_3) - \mu_A - \mu_B - 3\mu_X]$ and μ_A, μ_B, and μ_X are the chemical potentials of A, B, and halide, respectively (for most elements, chemical potentials are equal to the DFT total energies of their ground states); ΔH_f is the convex hull energy at the ABX_3 composition. A positive convex hull distance indicates unstable phases whereas negative distances confirm compound stability [63]. Metastable phases can often be synthesized whereby a structure predicted to lie near, but somewhat above the CH [65]. Previous studies [64–66] focused only on studying ABO_3 using DFT-CH (analysis of stability attributed to decomposition into other phases) using the Open Quantum Materials Database (OQMD) [63] for total energy calculations to evaluate constructed machine learning models. Thus, checking the studied Ti-based perovskites (A_2TiX_6 and $ATiX_3$) thermodynamic stability *via* DFT-CH is recommended in future works to validate our results from formation/decomposition energy of compounds.

8. Conclusion

The formability of Ti-based perovskites has been theoretically investigated through mixing/matching of various cations and anions relying on a random sampling of reported ionic radii (e.g. Shannon, Pauling, and Stern) and common stability factors for structural mapping. Our theoretical analysis predicts that Cs_2TiX_6 and $RbTiX_3$ are structurally stable with all halogens except for I^- with only ~ 0.5 formation probability. In metastable Ti-perovskites, the highest formation probabilities are observed for $MATiBr_3$ (0.87) and $MATiCl_3$ (0.81) owing to the optimum ionic radii allowing MA to occupy A-sites and Ti metals to fill B-sites for structural formation. However, the other unstable structures (Cs_2TiI_6, $RbTiI_3$, $MATiF_3$, $MATiI_3$) showed a 50/50 formation chance which might be associated with changes in ionic radii resultant from actual oxidation state and covalency. Also, instability of Ti-iodide-based structures might be associated with large anion radius (I^- ~2 Å → overlapping) preventing Ti atoms from occupying B-sites in the octahedral. Hence, one may conclude that perovskite formability would be best for $MATiCl_3$ > $MATiBr_3$ > (Cs_2TiI_6, $MATiF_3$, $RbTiI_3$) > $MATiI_3$ obtained from binomial distribution analysis of the metastable combinations. Nash equilibrium determined the six stable structures as (Cs_2TiX_6 and $RbTiX_3$ for X = F^-, Cl^-, Br^-) in agreement with the probabilistic analysis. Moreover, the IDSDS method yielded in the same six stable combinations (strategies) identified by Nash equilibrium from deletions of dominated strategies.

Seven predictors are utilized in the machine learning analysis including ionic radii (R_A, R_B, R_X), bond length (BL for Ti-X, BL for X···H), and bond-valence sums (BVS for Ti in Ti-X and BVS for X in X···H bonds). BLs and BVS implicitly involve both ions electronegativity and atoms coordination number (CN), respectively, for the determination of BX_6 octahedral stability. Studied predictors give us the important crystal chemistry parameters responsible for controlling cuboctahedral stability. Critical predictors are determined as R_A for A_mTiBr_{3m}; R_A, R_B, and R_X for A_mTiCl_{3m}; R_A and BVS of F for A_mTiF_{3m} with a more complicated case observed in A_mTiI_{3m} where all ionic radii (R_A, R_B, R_X) and BL for Ti-I are critical for the determination of the A_mTiI_{3m} cuboctahedral stability. Predicted model results for the new/future datasets (and their maximum and minimum BL and BVS from Table S8) are used to test the model validity. We concluded that Cs-based perovskites have high stabilities (Cs_2TiX_6 > $RbTiX_3$ > $MATiX_3$) because of their higher CN and lower BVS deviations. Moreover, formability probabilities of future datasets are found as 0.52, 0.79, 0.91, and 0.93 for I^-, F^-, Cl^-, and Br^-, respectively (A_mTiBr_{3m} > A_mTiCl_{3m} > A_mTiF_{3m} > A_mTiI_{3m}) with a marginal error ± 8% and a classification accuracy > 96.3%. The wider the BVS range, the lower the stability we might get due to the higher chances for Ti to be over-bonded or under-bonded with halides. Structural optimality is found in (Ti-Br)-like structures with electronegativity-to-ionic-radius ratio ~ 1.51 $Å^{-1}$ for a balanced cuboctahedral. Machine learning results are found in good agreement with our statistical and game theory analysis, with the addressed contradictions in $ATiI_3$ where high BVS deviations showed high structural instabilities. Ti might become a promising metal cation candidate in forming stable and formable perovskites (A_2TiX_6 and/or $ATiX_3$) substituting the conventional toxic Pb-based degradable perovskites and to be futuristically used in photoluminescence, tunable bandgaps, tunneling junctions, and high thermal and/or moisture photovoltaic applications.

CRediT authorship contribution statement

Hisham A. Maddah: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing - original draft. **Vikas Berry:** Resources, Supervision. **Sanjay K. Behura:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

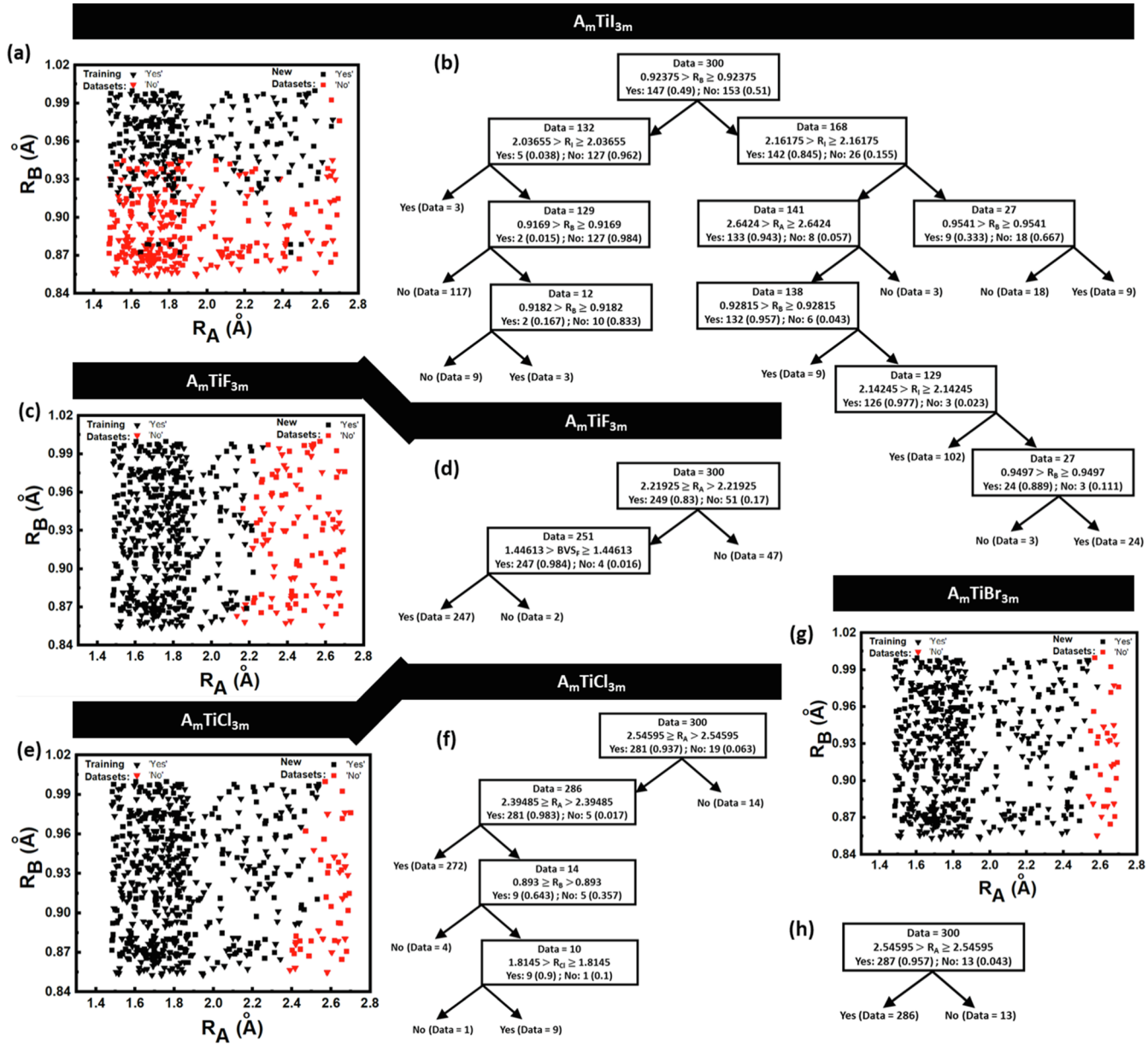

Fig. 7. Predicted perovskites stabilities of both the training datasets and their corresponding new datasets (for testing) obtained from the trained model (input/response of the training datasets): (a) A_mTiI_{3m}; (c) A_mTiF_{3m}; (e) A_mTiCl_{3m}; (g) A_mTiBr_{3m}. Classification tree graphs from fine tree training of the training datasets: (b) A_mTiI_{3m}; (d) A_mTiF_{3m}; (f) A_mTiCl_{3m}; (h) A_mTiBr_{3m}; R refers to radii values, Data or Yes/No number corresponds to datasets and structural formability (Yes/No) with number in the parenthesis for the associated probabilities. Note that m = 2 for A = Cs^+; elsewhere, m = 1 (for A = MA^+ and Rb^+); See Supplementary; section 7 (pages S9-S28).

Acknowledgements

HAM would like to acknowledge the Saudi Arabian Cultural Mission (SACM) and King Abdulaziz University (KAU) for their support and funding to pursue graduate studies. SKB and VB thank Dimerond Technologies, LLC for the support to conduct renewable energy research at the University of Illinois at Chicago. All the authors thank the University of Illinois at Chicago for the support. VB thanks funding support from the National Science Foundation (grant: 1054877) and Office of Naval Research (grants: N000141110767 and N000141812583).

Data Availability

All the source codes and training/testing datasets generated for machine learning are available in the Supplementary information (for more details regarding the generated classification trees, training models .m files or input/output data .mat files, please contact the authors).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.commatsci.2019.109415.

References

[1] S. Tombe, G. Adam, H. Heilbrunner, H. Apaydin, E. Iwuoha, M.C. Scharber, Optical and electronic properties of mixed halide (2017) 1714–1723, https://doi.org/10.1039/c6tc04830g.
[2] H. Hu, T. Salim, B. Chen, Y.M. Lam, Molecularly Engineered Organic-Inorganic Hybrid Perovskite with Multiple Quantum Well Structure for Multicolored Light-Emitting Diodes, Sci. Rep. (2016), https://doi.org/10.1038/srep33546.
[3] H. Hu, T. Salim, B. Chen, Y.M. Lam, Molecularly Engineered Organic- Inorganic Hybrid Perovskite with Multiple Quantum Well Structure for Multicolored Light-Emitting Diodes, Nat. Publ. Gr. (2016) 1–8, https://doi.org/10.1038/srep33546.
[4] N.G. Park, Perovskite solar cells: An emerging photovoltaic technology, Mater. Today. (2015), https://doi.org/10.1016/j.mattod.2014.07.007.
[5] L. Tzounis, T. Stergiopoulos, A. Zachariadis, C. Gravalidis, ScienceDirect Perovskite

solar cells from small scale spin coating process towards roll-to-roll printing : Optical and Morphological studies.
[6] L. Tzounis, T. Stergiopoulos, A. Zachariadis, C. Gravalidis, A. Laskarakis, S. Logothetidis, Perovskite solar cells from small scale spin coating process towards roll-to-roll printing: Optical and Morphological studies, Mater. Today Proc. (2017), https://doi.org/10.1016/j.matpr.2017.04.117.
[7] C.C. Stoumpos, D.H. Cao, D.J. Clark, J. Young, J.M. Rondinelli, J.I. Jang, J.T. Hupp, M.G. Kanatzidis, Ruddlesden − Popper Hybrid Lead Iodide Perovskite 2D Homologous, Semiconductors (2016), https://doi.org/10.1021/acs.chemmater.6b00847.
[8] D.B. Mitzi, C.A. Feild, Z. Schlesinger, R.B. Laibowitz, Transport, optical, and magnetic properties of the conducting halide perovskite $CH_3NH_3SNi_3$, J. Solid State Chem. (1995), https://doi.org/10.1006/jssc.1995.1023.
[9] H. Bloemen, I. Interstellar Processes, D.J. Hollen-bach, H.A. Thronson Jr, D.B. Mitzi, S. Wang, C.A. Feild, C.A. Chess, A.M. Guloy, I.T.J. Watson, Conducting Layered Organic-inorganic Halides Containing (110)-Oriented Perovskite Sheets, Science (1995).
[10] W. Travis, E.N.K. Glover, H. Bronstein, D.O. Scanlon, R.G. Palgrave, On the application of the tolerance factor to inorganic and hybrid halide perovskites: A revised system, Chem. Sci. (2016), https://doi.org/10.1039/c5sc04845a.
[11] J. Shi, S. Yun, First-Principles DFT Calculations for Perovskite Solar Cells, in, Count. Electrodes Dye. Perovskite Sol. Cells (2018), https://doi.org/10.1002/9783527813636.ch19.
[12] D. Wang, M. Wright, N.K. Elumalai, A. Uddin, Stability of perovskite solar cells, Sol. Energy Mater. Sol. Cells. (2016), https://doi.org/10.1016/j.solmat.2015.12.025.
[13] G. Niu, W. Li, F. Meng, L. Wang, H. Dong, Y. Qiu, Study on the stability of $CH_3NH_3PbI_3$ films and the effect of post-modification by aluminum oxide in all-solid-state hybrid solar cells, J. Mater. Chem. A. (2014), https://doi.org/10.1039/c3ta13606j.
[14] M.I. Asghar, J. Zhang, H. Wang, P.D. Lund, Device stability of perovskite solar cells – A review, Renew. Sustain. Energy Rev. (2017), https://doi.org/10.1016/j.rser.2017.04.003.
[15] G. Niu, X. Guo, L. Wang, Review of recent progress in chemical stability of perovskite solar cells, J. Mater. Chem. A. (2015), https://doi.org/10.1039/c4ta04994b.
[16] H.S. Kim, J.Y. Seo, N.G. Park, Material and Device Stability in Perovskite Solar Cells, ChemSusChem. (2016), https://doi.org/10.1002/cssc.201600915.
[17] I. Hwang, I. Jeong, J. Lee, M.J. Ko, K. Yong, Enhancing Stability of Perovskite Solar Cells to Moisture by the Facile Hydrophobic Passivation, ACS Appl. Mater. Interfaces. (2015), https://doi.org/10.1021/acsami.5b04490.
[18] J. Chen, X. Cai, D. Yang, D. Song, J. Wang, J. Jiang, A. Ma, S. Lv, M.Z. Hu, C. Ni, Recent progress in stabilizing hybrid perovskites for solar cell applications, J. Power Sources. (2017), https://doi.org/10.1016/j.jpowsour.2017.04.025.
[19] Q. Fu, X. Tang, B. Huang, T. Hu, L. Tan, L. Chen, Y. Chen, Recent Progress on the Long-Term Stability of Perovskite Solar Cells, Adv. Sci. (2018), https://doi.org/10.1002/advs.201700387.
[20] Y. Zhou, Y. Zhao, Chemical stability and instability of inorganic halide perovskites, Energy Environ. Sci. (2019).
[21] H. Tanaka, T. Oku, N. Ueoka, Structural stabilities of organic–inorganic perovskite crystals, in, Jpn. J. Appl. Phys. (2018), https://doi.org/10.7567/JJAP.57.08RE12.
[22] S. Arrhenius, C. Li, X. Lu, W. Ding, L. Feng, Y. Gao, Z. Guo, Formability of ABX_3 (X = F, Cl, Br, I) halide perovskites, Acta Crystallogr. Sect. B Struct. Sci. (2008), https://doi.org/10.1107/S0108768108032734.
[23] R.D. Shannon, Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, Acta Crystallogr. Sect. A. (1976), https://doi.org/10.1107/S0567739476001551.
[24] R.D. Shannon, No Title, Acta Crystallogr., Sect. B Struct. Crystallogr. Cryst. Chem. 25 (1969) 925.
[25] W.M. Haynes, CRC Handbook Chemistry and Physics, CRC Press. (2016), https://doi.org/10.1016/0022-2860(92)85083-S.
[26] L. Pauling, The sizes of ions and the structure of ionic crystals, J. Am. Chem. Soc. (1927), https://doi.org/10.1021/ja01402a019.
[27] M. Rahm, R. Hoffmann, N.W. Ashcroft, Atomic and Ionic Radii of Elements 1–96, Chem. - A Eur. J. (2016), https://doi.org/10.1002/chem.201602949.
[28] C. Yoder, Metallic Covalent and Ionic, Radii (r) (2019).
[29] M.W. Lufaso, P.M. Woodward, Prediction of the crystal structures of perovskites using the software program SPuDS, Acta Crystallogr. Sect. B Struct. Sci. (2001), https://doi.org/10.1107/S0108768101015282.
[30] E.S. Stern, K.H. Amis, Ionic size, Chem. Rev. 59 (1959) 1–64.
[31] R.A. Jishi, Modifed Becke-Johnson exchange potential: improved modelling of lead halides for solar cell, applications (2016).
[32] C. Baka, Closedmouth, Discospinster, D. Zvonimir, Dreadstar, EagleFan, G. Nygaard, G. Gonzonoir, Ionic radius, https://resources.saylor.org/wwwresources/archived/site/wp-content/uploads/2011/06/Ionic-Radius.pdf.
[33] J.E. Huheey, E. a. Keiter, R.L. Keiter, Inorganic Chemistry: Principles of Structure and Reactivity (4th Edition), 1993. doi:10.1016/j.jcorp.
[34] A. Amat, E. Mosconi, E. Ronca, C. Quarti, P. Umari, M.K. Nazeeruddin, M. Grätzel, F. De Angelis, Cation-induced band-gap tuning in organohalide perovskites: Interplay of spin-orbit coupling and octahedra tilting, Nano Lett. (2014), https://doi.org/10.1021/nl5012992.
[35] Y. Liu, T. Zhao, W. Ju, S. Shi, S. Shi, S. Shi, Materials discovery and design using machine learning, J. Mater. (2017), https://doi.org/10.1016/j.jmat.2017.08.002.
[36] Z. Li, Q. Xu, Q. Sun, Z. Hou, W.-J. Yin, Stability Engineering of Halide Perovskite via Machine, Learning, Arxiv, 2018.
[37] Z. Li, Q. Xu, Q. Sun, Z. Hou, W.J. Yin, Thermodynamic Stability Landscape of Halide Double Perovskites via High-Throughput Computing and Machine Learning, Adv. Funct. Mater. (2019), https://doi.org/10.1002/adfm.201807280.
[38] G. Pilania, P.V. Balachandran, J.E. Gubernatis, T. Lookman, Classification of ABO_3 perovskite solids: A machine learning study, Acta Crystallogr, Sect. B Struct. Sci. Cryst. Eng. Mater. (2015), https://doi.org/10.1107/S2052520615013979.
[39] H. Im, J. Lee, S. Ko, T.W. Kim, H.W. Hyon, Y. Chang, Identifying Pb-free perovskites for solar cells by machine learning, Npj Comput, Mater. 5 (2019) 37.
[40] N. Wagner, J.M. Rondinelli, Theory-Guided Machine Learning in Materials Science, Front. Mater. (2016), https://doi.org/10.3389/fmats.2016.00028.
[41] A.G. Bluman, Elementary Statistics A Step by Step Approach Eighth Edition (2012), https://doi.org/10.1017/CBO9781107415324.004.
[42] R. Bott, Bluman - Elementary Statistics Step by Step Approach, Igarss (2014), https://doi.org/10.1007/s13398-014-0173-7.2.
[43] S. Sinharay, Discrete probability distributions, in: Int. Encycl. Educ., 2010. doi:10.1016/B978-0-08-044894-7.01721-8.
[44] A.C. Aitken, H. Cramer, Random Variables and Probability Distributions, Math. Gaz. (2007), https://doi.org/10.2307/3607491.
[45] J.E.J. Harrington, Games Strategies and Decision Making, Macmillan (2009).
[46] A.K. Jagannatham, V. Kumar, Introduction to game theory, in: Decis. Sci. Theory Pract., 2016. doi:10.1201/9781315183176.
[47] M.J. Osborne An introduction to game theory 2004.
[48] V. Schomaker, D.P. Stevenson, Some Revisions of the Covalent Radii and the Additivity Rule for the Lengths of Partially Ionic Single Covalent Bonds, J. Am. Chem. Soc. (1941), https://doi.org/10.1021/ja01846a007.
[49] R.J. Gillespie, E.A. Robinson, Bond Lengths in Covalent Fluorides. A New Value for the Covalent Radius of Fluorine, Inorg. Chem. (1992), https://doi.org/10.1021/ic00036a045.
[50] I.D. Brown, Recent developments in the bond valence model of inorganic bonding, Phys. Chem. Miner. (1987), https://doi.org/10.1007/BF00307605.
[51] L. Pauling, Nature of chemical Bond, in: Nat. Chem. Bond (3rd Ed.), 1960. doi:10.1021/jp020500+.
[52] L.C. Allen, Electronegativity Is the Average One-Electron Energy of the Valence-Shell Electrons in Ground-State Free Atoms, J. Am. Chem. Soc. (1989), https://doi.org/10.1021/ja00207a003.
[53] D. Castelvecch, Just How Small Is the Proton? - Scientific American, Springer Nat, 2010 https://www.scientificamerican.com/article/just-how-small-is-the-proton/.
[54] R. Seshadri, Bond valence nets and sums, Univ, St. Barbar. Mater. Dep, California, 2011 https://www.mrl.ucsb.edu/~seshadri/2011_218/2011_218_BondValenceSums.pdf.
[55] N.E. Brese, M. O'Keeffe, Bond-valence parameters for solids, Acta Crystallogr. Sect. B. (1991), https://doi.org/10.1107/S0108768190011041.
[56] A. Varadwaj, P.R. Varadwaj, H.M. Marques, K. Yamashita, Halogen in materials design: Revealing the nature of hydrogen bonding and other non-covalent interactions in the polymorphic transformations of methylammonium lead tribromide perovskite, Mater. Today Chem. (2018), https://doi.org/10.1016/j.mtchem.2018.04.003.
[57] M. Chen, M.G. Ju, A.D. Carl, Y. Zong, R.L. Grimm, J. Gu, X.C. Zeng, Y. Zhou, N.P. Padture, Cesium Titanium(IV) Bromide Thin Films Based Stable Lead-free Perovskite Solar Cells, Joule. (2018), https://doi.org/10.1016/j.joule.2018.01.009.
[58] A.M. Teweldeberhan, J.L. Dubois, S.A. Bonev, Stability of the high-pressure phases of $CaTiO_3$ perovskite at finite temperatures, Phys. Rev. B - Condens. Matter Mater. Phys. (2012), https://doi.org/10.1103/PhysRevB.86.064104.
[59] H.F. Kay, P.C. Bailey, Structure and properties of $CaTiO_3$, Acta Crystallogr. (1957), https://doi.org/10.1107/s0365110x57000675.
[60] C.J. Ball, B.D. Begg, D.J. Cookson, G.J. Thorogood, E.R. Vance, Structures in the System CaTiO3/SrTiO3, J. Solid State Chem. (1998), https://doi.org/10.1006/jssc.1998.7836.
[61] A. Kubo, T. Suzuki, M. Akaogi, High pressure phase equilibria in the system $CaTiO_3$-$CaSiO_3$: Stability of perovskite solid solutions, Phys. Chem. Miner. (1997), https://doi.org/10.1007/s002690050063.
[62] X.G. Zhao, D. Yang, Y. Sun, T. Li, L. Zhang, L. Yu, A. Zunger, Cu-In Halide Perovskite Solar Absorbers, J. Am. Chem. Soc. (2017), https://doi.org/10.1021/jacs.7b02120.
[63] S. Kirklin, J.E. Saal, B. Meredig, A. Thompson, J.W. Doak, M. Aykol, S. Rühl, C. Wolverton, The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies, Npj Comput. Mater. (2015), https://doi.org/10.1038/npjcompumats.2015.10.
[64] W. Li, R. Jacobs, D. Morgan, Predicting the thermodynamic stability of perovskite oxides using machine learning models, Comput. Mater. Sci. (2018), https://doi.org/10.1016/j.commatsci.2018.04.033.
[65] A.A. Emery, C. Wolverton, High-Throughput DFT calculations of formation energy, stability and oxygen vacancy formation energy of ABO_3 perovskites, Sci. Data. (2017), https://doi.org/10.1038/sdata.2017.153.
[66] P.V. Balachandran, A.A. Emery, J.E. Gubernatis, T. Lookman, C. Wolverton, A. Zunger, Predictions of new ABO_3 perovskite compounds by combining machine learning and density functional theory, Phys. Rev. Mater. (2018), https://doi.org/10.1103/PhysRevMaterials. 2.043802.