

SUPPLEMENTAL INFORMATION

A Deep Generative Modeling Architecture for Designing Lattice-Constrained Perovskite Materials

Ericsson Tetteh Chenebuah^{1,2,3,*}, Michel Nganbe¹, Alain Beaudelaire Tchagang^{1,2}

¹ Department of Mechanical Engineering, University of Ottawa, 161 Louis-Pasteur, Ottawa, ON, K1N 6N5, Canada.

² Digital Technologies Research Center, National Research Council of Canada, 1200 Montréal Road, Ottawa, ON, K1A 0R6, Canada.

³ Lead contact

* Correspondence: echen013@uottawa.ca

1. Invertible mesh-grid descriptor

The mesh-grid descriptor concept for representing perovskite materials consists of three feature arrays, namely: (1) label mesh array; (2) property mesh array; and (3) XRD mesh array. In *one-hot* encoded form, the label mesh stacks the atomic numbers (Z) of distinctive atoms in the perovskite's unit cell along the column axis. The present research considers only the first 100 elements in the periodic table of elements (i.e. Hydrogen to Fermium) due to their prevalence in the training dataset. The label mesh also one-hot encodes the stoichiometrical type of perovskite samples in the training set (i.e. either ABX_3 or $A_2BB'X_6$ corresponding to [1 0] or [0 1]), and the number of all atoms in the conventional unit cell (n_{atoms}). In order to ensure each distinctive atom in the label mesh is representative of a crystallographic site, the atomic listings are reorganized to perfectly align to their corresponding site occupancies. For instance, considering the ternary ABX_3 perovskite formula, which can be written in several forms, atom A belongs to Site 1, atom B belongs to Site 2, and atom X belongs to Site 3. For compatibility, the label array matrix is augmented (i.e. *zero-padded*) to fit the overall size arrays with respect to the other meshes. As illustrated using the Fig. S1, the label mesh is organized into a 128×8 two-dimensional array and is reshaped to produce a 32×32 square matrix. Moreover, the property mesh

consists of one-hot encoded features that characterize the thermochemistry behavior of each atom that belongs to a distinctive site occupancy in the stoichiometry of the unit cell. Hence, for ABX_3 and $A_2BB'X_6$ stoichiometries, there are three and four encoded thermochemistry columns, respectively, which are stacked together. As illustrated using the Fig. S1, the thermochemistry properties are twelve in total and are concatenated along their row axis. Table S1 lists the numerical range of each property, in addition to the bin size used in the one-hot encoding process. To fit the overall matrix size with respect to the other meshes, the property mesh is zero-padded to produce a 132×8 two-dimensional array and is reshaped to form a 32×32 square matrix. Finally, the X-ray diffraction (XRD) mesh fingerprints the scaled XRD computed peaks/intensities onto a 32×32 matrix array. The XRD pattern is simulated using a $\text{CuK}\alpha$ wavelength ($\sim 1.5 \text{ \AA}$) prober for 2θ diffraction angle ranging from 0° to 180° . Assembling all meshes together produces a $32 \times 32 \times 3$ invertible RGB image. Figure S2 displays the mean squared error (MSE) distribution in pixel attributes between decoded and actual/calculated XRD patterns for over 70,000 newly generated crystals by the LCMGM, as it relates to a crystal system type. The source codes for calculating and projecting XRD peaks were enabled using the XRD Calculator in the `pymatgen.analysis.diffraction.xrd` submodule [1, 2]. This research makes available the codes and other relevant information for designing the invertible mesh-grid descriptor at www.github.com/chenebuah/LCMGM.

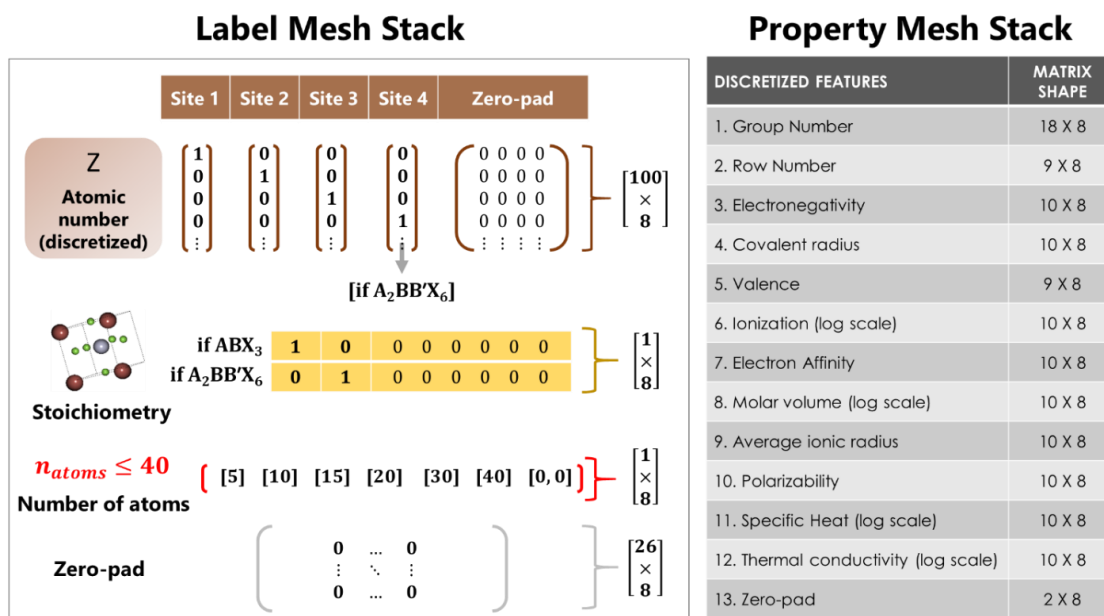


Figure S1. Stacking arrangement for organizing discrete features with respect to the label mesh and property mesh arrays.

Table S1. Thermochemistry properties used in developing the property mesh array of the invertible mesh-grid descriptor. Each property is discretized using one-hot encoding features based on the bin size and numerical range of real values.

Ref	Thermochemistry property	Unit	Range	Number of bins
-	Group number	-	1,2,...,18	18
-	Row number	-	1,2,...,9	9
[3]	Pauling electronegativity	Pauling	0.7 - 3.98	10
[4]	Covalent radius	Angstrom	0.28 - 2.6	10
-	Valence	-	1,2,...,9	9
[3]	First ionization energy (log scale)	eV	3.89 - 24.59	10
[3]	Electron affinity	eV	-2.33 - 5.94	10
[1]	Molar volume (log scale)	cm ³	4.39 - 70.94	10
[1]	Average ionic radius	Angstrom	0 - 1.94	10
[3]	Static average electric dipole Polarizability	10 ⁻²⁴ cm ³	0.21 - 59.42	10
[3]	Specific heat (log scale)	KJ/kg.K	0.06 - 14.3	10
[5]	Thermal conductivity (log scale)	W/m.K	0.0036 - 430	10

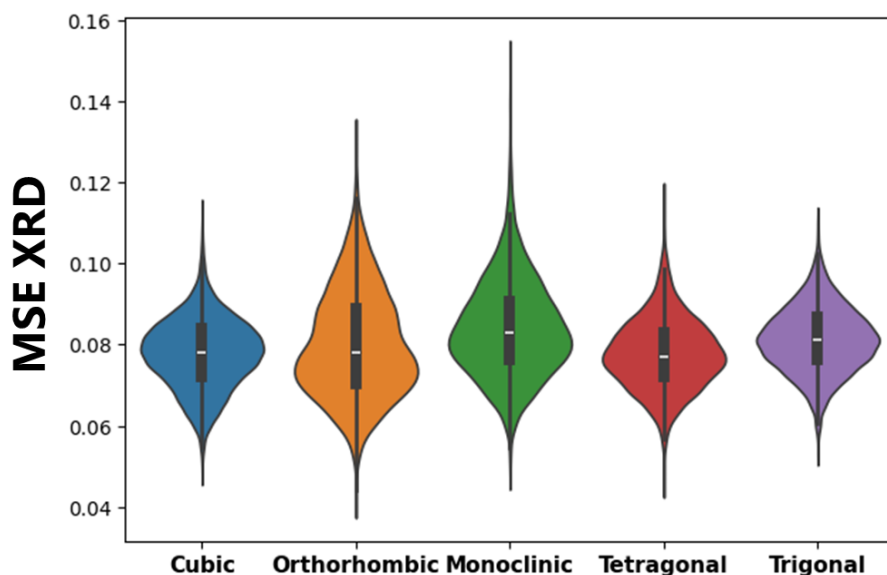


Figure S2. Mean squared error (MSE) distribution in pixel attributes between decoded and actual/calculated XRD patterns for over 70,000 newly generated crystals by the LCMGM.

2. LCMGM neural network architecture and hyperparameters

For addressing the inverse design scheme, the Lattice Constrained Materials Generative Model (LCMGM) comprise of three modeling phases, namely: (1) semi-supervised variational autoencoder (SS-VAE); (2) auxiliary generative adversarial network (A-GAN); and (3) Geometrical optimization in Bayesian optimization (BO) and density functional theory (DFT). Both the SS-VAE and A-GAN are target-learning generative models and are architected using deep neural networks. The Tables S2 and S3 details the modeling architectures and hyperparameters for designing the SS-VAE and A-GAN, respectively. The SS-VAE comprise of four modeling networks (i.e. encoder, regressor, classifier, and decoder) that are graphically tied together in backpropagation. The A-GAN on the other hand consists of two typical neural networks (i.e. generator and discriminator) that competes against each other. Each neural network contains sub-networks that reflects the auxiliary generative approach for learning the geometrical lattice constraints. Given a perovskite unit cell, the geometrical constraints are the atomic coordinates ($\vec{R}_{(x,y,z)} \in \mathbb{R}^{(40 \times 3)}$) and lattice parameters ($L_{p[a,b,c;\alpha,\beta,\gamma]} \in \mathbb{R}^{(2 \times 3)}$). The modeling weights are continuously updated, as the predictor sub-model of the A-GAN’s discriminator regressively predicts the unknown geometrical constraints and evaluates the error in prediction. Batch normalization are typically used between interconnecting layers for standardizing and stabilizing the networks. All neural network-modeling designs were scripted using Python, and were enabled on a Keras API [6] of the TensorFlow backend [7]. Further information on coding the LCMGM are made available at www.github.com/chenebuah/LCMGM.

Table S2. Modeling architecture and hyperparameters for designing the SS-VAE model of the LCMGM

Network	Layer API class	Filter/Units	Kernel	Stride	Padding	Activation
Encoder (VAE)	Input shape	(32 × 32 × 3)				
	Conv2D	32	3 × 3	2	same	LeakyReLU (α=0.2)
	Conv2D	64	3 × 3	2	same	LeakyReLU (α=0.2)
	Conv2D	128	3 × 3	1	same	LeakyReLU (α=0.2)
	Flatten	1 × 8192				
	Dense	1024				Sigmoid
	\mathbf{z}_{mean} Dense	256				Linear
	\mathbf{z}_{var} Dense	256				Linear

	z (encoded sampling vector)	$z_{\text{mean}} + \exp(0.5 \times z_{\text{var}}) \cdot \epsilon$ where $\epsilon \sim \mathcal{N}(0, I)$				
Regressor (MLP)	Input z					ReLU
	Dense	256				ReLU
	Dense	128				ReLU
	Dense	64				ReLU
	Dense	32				ReLU
	Output Dense	1				Linear
Classifier (MLP)	Input z					ReLU
	Dense	128				ReLU
	Dense	64				ReLU
	Dense	32				ReLU
	Class output	5				Softmax
Decoder	Dense	1024				Sigmoid
	Dense	8192				LeakyReLU ($\alpha=0.2$)
	Reshape	$(8 \times 8 \times 128)$				
	Conv2DTranspose	64	3×3	1	same	LeakyReLU ($\alpha=0.2$)
	Conv2DTranspose	32	3×3	2	same	LeakyReLU ($\alpha=0.2$)
	Reconstructed Conv2DTranspose	3	3×3	2	same	Sigmoid
	Output shape	$(32 \times 32 \times 3)$				
Hyperparameters						
Learning rate	1.00e-04					
Decay	1e-4/200					
Batch size	32					
Epoch	1500					
Optimizer	Legacy Adam					
Loss metrics	VAE \Rightarrow KL divergence + Reconstruction (Mean Squared Error) Regressor MLP \Rightarrow Mean Squared Error Classifier MLP \Rightarrow Categorical cross-entropy					

Table S3. Modeling architecture and hyperparameters for designing the A-GAN model of the LCMGM

Network	Sub-network	Layer API class	Filter/Units	Kernel	Stride	Padding	Activation
Generator	Constraints [atom coordinates + lattice parameters]	Input shape	(42×3)				
		Conv1D	24	3	1	same	LeakyReLU ($\alpha=0.2$)
		Flatten	(1×1008)				
		Dense	256				LeakyReLU ($\alpha=0.2$)
		Reshape output	(1×256)				

	Noise	Input shape	(1 × 100)				
		Dense	256				LeakyReLU ($\alpha=0.2$)
		Reshape output	(1 × 256)				
	Merge [Constraints + Noise]	Concatenate	(2 × 256)				
		Main	Reshape	(16 × 16 × 2)			
	Conv2D		4	3 × 3	1	same	LeakyReLU ($\alpha=0.2$)
	Conv2D		16	3 × 3	1	same	LeakyReLU ($\alpha=0.2$)
Flatten	(1 × 4096)						
Dense	512					LeakyReLU ($\alpha=0.2$)	
Dense output	256				tanh		
Discriminator	Input z	(256 × 1)					
		Dense	512			LeakyReLU ($\alpha=0.2$)	
	Critic	Reshape	(16 × 16 × 2)				
		Conv2D	4	3 × 3	1	same	LeakyReLU ($\alpha=0.2$)
		Conv2D	16	3 × 3	1	same	LeakyReLU ($\alpha=0.2$)
		Flatten	(1 × 4096)				
		Dense	64				LeakyReLU ($\alpha=0.2$)
		Dense	16				LeakyReLU ($\alpha=0.2$)
		Dense output	1				Sigmoid
	Predictor	Dense	1134				LeakyReLU ($\alpha=0.2$)
		Reshape	(42 × 27)				
		Conv1D	9	3	1	same	LeakyReLU ($\alpha=0.2$)
		Conv1D	3	3	1	same	Sigmoid
		Output shape	(42 × 3)				
Hyperparameters							
Learning rate	1.00e-04						
Decay	1e-4/200						
Batch size	64						
Epoch	100						
Optimizer	Legacy Adam						
Loss metrics	Critic ⇒ Binary crossentropy Predictor ⇒ Mean Squared Error						

3. Bayesian Optimization (BO) pseudocode

The BO algorithm is used to find the optimized lattice configuration L_p (i.e. edge vectors and inter-axial angles) that best minimizes the total DFT energy (E_{tot}). The BO algorithm as implemented in study was

developed using Scikit-Optimize (*skopt*) [8], which is a sequential model-based library for multi-objective optimization. As demonstrated in the main article, the pseudocode can be described as follows:

1. For a newly generated perovskite \hat{x} , initialize the surrogate model with a dataset D of initial points $L_p \subseteq \hat{x} \in \mathbb{R}^{(32 \times 32 \times 3)}$ and their corresponding true solutions E_{tot} .
2. For $t = 1, 2, \dots, N$:
 - a. Fit the surrogate model to the current perovskite dataset D .
 - b. Choose the next L_p point by optimizing the acquisition function over the surrogate model.
 - c. Evaluate the objective function E_{tot} which is a function of L_p .
 - d. Augment the dataset D with the new point $\{L_{p|t+1}, E_{tot|t+1}(L_{p|t+1})\}$.
3. Return the best point found in D .

In the present research, the applied surrogate model is based on Gaussian Processes and the acquisition function is *expected improvement*. The surrogate model is initialized using $D = 40$ points and the number of calls (t) are observed for $N = 200$ iterations. It should be noted moreover that the mesh-grid descriptor concept is modified to include the L_p features in order to effect the dynamic changes within the search space for finding the optimized configuration. For each considered crystal system, the search space dimensionality for guiding the acquisition function is outlined in the Table 7 of the main article. To further demonstrate the good effect of the BO algorithm in performing pre-DFT relaxation, the Fig. S3 (A) visualizes the variation in optimized versus non-optimized/generated candidates with respect to their final DFT-relaxed forms, as investigated singularly on the lattice a (Å) edge vectors. The Fig. S3 (B) displays the average absolute error difference in matching lattice parameters for all newly generated candidates by the BO. Except for the b -edge vector, performing pre-relaxation with the BO algorithm yields better lattice configurations that are closer to their final DFT-relaxed states. Further information on implementing the BO algorithm are made available at www.github.com/chenebuah/LCMGM.

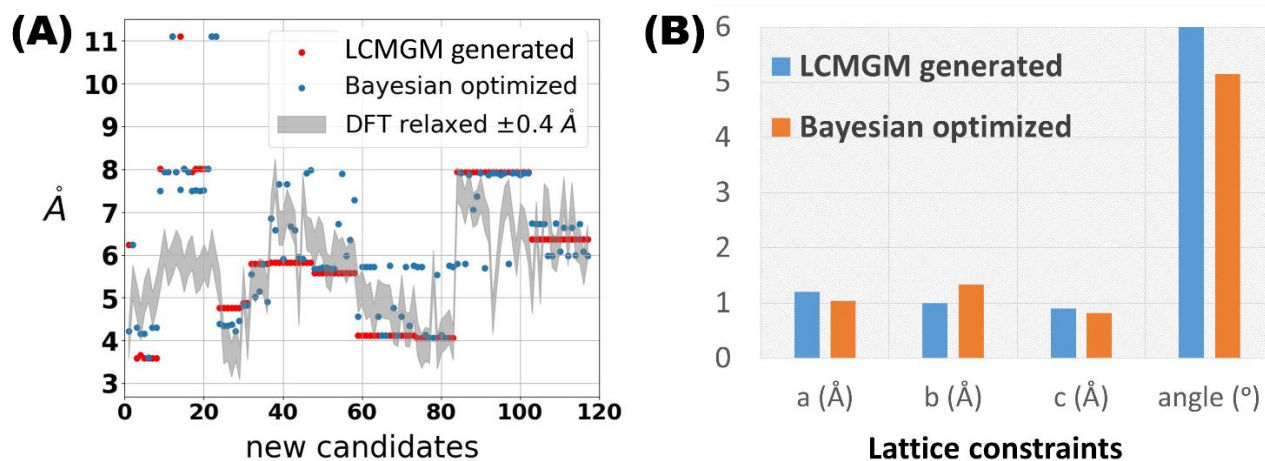


Figure S3. Effect of BO algorithm for finding the optimized three-dimensional lattice configuration. (A) optimized versus non-optimized (i.e. generated) candidates with respect to their final DFT-relaxed forms for lattice a -edge vectors; (B) average absolute difference in DFT-relaxed parameters for optimized and non-optimized.

4. Prototypical geometries for querying the A-GAN model

For generating new candidates with higher geometrical conformity, the fully trained A-GAN model is queried using proven perovskite prototypes as reference geometries. In selecting prototypical structures, the present research prioritizes Inorganic Crystal Structure Database (ICSD) [9] perovskites due to their higher validation from physical experiments. However, for special instances where no suitable ICSD prototype (e.g. cubic ABX_3 stoichiometry), hypothetically discovered compounds that have been extensively investigated via *ab initio* computations are used. The Table S4 outlines the prototypes used in this study, in addition to their data entries ID from the Materials Project (MP) database [10].

Table S4. Examples of prototypical perovskite structures from the Materials Project (MP) database used in querying the A-GAN model of the fully trained LCMGM. n_{atoms} is the number of atoms in the conventional unit cell.

Crystal system	MP ID	Prototype	Space group	Lattice parameters	n_{atoms}	ICSD ID
Cubic	mp-1004524	HPbI ₃	$Pm\bar{3}m$	$a = 6.244 \text{ \AA}$	5	FALSE
	mp-976185	NaBeO ₃	$Pm\bar{3}m$	$a = 3.589 \text{ \AA}$	5	FALSE
	mp-1183176	AlNiO ₃	$Pm\bar{3}m$	$a = 3.658 \text{ \AA}$	5	FALSE
	mp-1078783	La ₂ VReO ₆	$Fm\bar{3}m$	$a = 8.014 \text{ \AA}$	40	TRUE
	mp-1079027	Sr ₂ ZrCoO ₆	$Fm\bar{3}m$	$a = 7.945 \text{ \AA}$	40	TRUE
	mp-1079458	Cs ₂ NaCeCl ₆	$Fm\bar{3}m$	$a = 11.112 \text{ \AA}$	40	TRUE
Monoclinic	mp-6989	BaSeO ₃	$P2_1/m$	$a = 4.761 \text{ \AA}, b = 5.771 \text{ \AA}, c = 7.062 \text{ \AA}, \beta = 107.92^\circ$	10	TRUE
	mp-998711	TlGeF ₃	$P2_1/m$	$a = 4.872 \text{ \AA}, b = 6.058 \text{ \AA}, c = 7.537 \text{ \AA}, \beta = 107.38^\circ$	10	FALSE
	mp-554410	Ba ₂ LaRuO ₆	$C2/m$	$a = 10.556 \text{ \AA}, b = 6.082 \text{ \AA}, c = 6.152 \text{ \AA}, \beta = 125.22^\circ$	20	TRUE
	mp-1181527	Eu ₂ HoTaO ₆	$P2_1/c$	$a = 5.793 \text{ \AA}, b = 5.902 \text{ \AA}, c = 10.070 \text{ \AA}, \beta = 125.00^\circ$	20	TRUE
Orthorhombic	mp-5246	KNbO ₃	$Amm2$	$a = 5.824 \text{ \AA}, b = 5.860 \text{ \AA}, c = 4.018 \text{ \AA}$	10	TRUE
	mp-623098	Ca ₂ FeWO ₆	$Pmm2$	$a = 5.584 \text{ \AA}, b = 5.674 \text{ \AA}, c = 7.885 \text{ \AA}$	20	TRUE
Tetragonal	mp-19990	BaTiO ₃	$P4/mmm$	$a = 4.112 \text{ \AA}, c = 5.036 \text{ \AA}$	5	TRUE
	mp-1205663	Sr ₂ ZrTiO ₆	$P4mm$	$a = 4.058 \text{ \AA}, c = 8.283 \text{ \AA}$	10	FALSE
Trigonal	mp-1079481	AlSBr ₃	$P\bar{3}m1$	$a = 7.935 \text{ \AA}, c = 6.968 \text{ \AA}$	10	TRUE
	mp-6654	Cs ₂ LiGaF ₆	$P\bar{3}m1$	$a = 6.361 \text{ \AA}, c = 5.180 \text{ \AA}$	10	TRUE

5. Newly designed LCMGM perovskites

Altogether, 124 perovskite materials were newly designed in this research. Among them, 72 are suggested to be new chemistries (i.e. unique and novel), as they cannot be found in established materials databases such as Materials Project (MP) [10], Open Quantum Materials Database (OQMD) [11], and Novel Materials discovery (NOMAD) [12]. Tables S5 provides a full breakdown of the new materials with respect to their determined lattice features. In addition, Table S6 reports the calculated thermodynamic and bandgap properties of the new materials. Except otherwise stated, the materials are reported based on DFT evaluation using GGA-PBE functionals. The new materials dataset can be accessed from the Mendeley data repository at <https://doi.org/10.17632/m262xxpgn2.1>. The raw

Crystallographic Information Files (CIF) and Quantum Espresso output files can also be downloaded from <https://github.com/chenebuah/LCMGM>.

Table S5. Lattice features of new materials designed by the LCMGM. For the listed materials, DFT-calculations are strictly reported using GGA-PBE functionals.

CIF ID	Composition	Space group	DFT-relaxed lattice parameters	DFT-relaxed density (g/cm ³)	Volume (Å ³)	Number of atoms in unit cell	Prevalence
CUBIC							
ABX₃ compounds							
001	CsMnO ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 3.987 Å	6.181	63.378	5	OQMD
002	CsZnBr ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 5.364 Å	4.714	154.336	5	OQMD
003	SrCuI ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 4.943 Å	7.313	120.774	5	novel and unique
004	MnPdS ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 4.412 Å	4.979	85.883	5	OQMD
005	SrMnS ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 5.053 Å	3.073	149.214	5	OQMD
006	SrPdI ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 5.304 Å	6.395	149.214	5	novel and unique
007	SrPdS ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 4.729 Å	4.558	105.757	5	novel and unique
008	BeAuI ₃	<i>Pm</i> $\bar{3}$ <i>m</i>	a = 5.113 Å	7.287	133.668	5	novel and unique
A₂BB'X₆ compounds							
009	Co ₂ VPbO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 7.944 Å	6.254	501.323	40	OQMD
010	Cs ₂ LiScF ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.786 Å	4.229	678.225	40	MP & OQMD
011	Dy ₂ HfIrO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.02 Å	10.193	515.850	40	novel and unique
012	Eu ₂ LaVO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.343 Å	6.746	580.720	40	OQMD
013	K ₂ GePtF ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.751 Å	4.557	670.152	40	novel and unique
014	Lu ₂ HoTlO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.469 Å	8.913	607.430	40	novel and unique
015	Nb ₂ CuWO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 7.827 Å	7.331	479.497	40	novel and unique
016	Rb ₂ LiMnF ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.348 Å	3.960	581.765	40	MP & OQMD
017	Rb ₂ AlPdF ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.575 Å	4.406	630.525	40	novel and unique
018	Sr ₂ VCuO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 7.771 Å	5.459	469.279	40	novel and unique
019	Sr ₂ CWO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 7.912 Å	6.264	495.289	40	OQMD
020	Sr ₂ VPbO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.275 Å	6.204	566.636	40	OQMD

021	U ₂ CuWO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 7.952 Å	10.824	502.839	40	novel and unique
022	Yb ₂ LaHoO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.766 Å	7.354	673.604	40	novel and unique
023	Yb ₂ TlVO ₆	<i>Fm</i> $\bar{3}$ <i>m</i>	a = 8.054 Å	8.865	522.438	40	novel and unique
MONOCLINIC							
ABX₃ compounds							
024	AlAsO ₃	<i>P2</i> ₁ / <i>m</i>	a = 5.212 Å, b = 5.274 Å, c = 5.215 Å, β = 92.21°	3.475	143.350	10	OQMD
025	NbAlO ₃	<i>P2</i> ₁ / <i>m</i>	a = 3.857 Å, b = 6.247 Å, c = 5.305 Å, β = 104.05°	4.497	127.822	10	OQMD
026	TaAlO ₃	<i>P2</i> ₁ / <i>m</i>	a = 3.929 Å, b = 6.456 Å, c = 6.701 Å, β = 91.45°	5.002	169.975	10	OQMD
027	TaAuO ₃	<i>P2</i> ₁ / <i>m</i>	a = 3.548 Å, b = 5.890 Å, c = 8.132 Å, β = 91.77°	8.327	169.940	10	OQMD
028	LaNbO ₃	<i>P2</i> ₁ / <i>m</i>	a = 3.937 Å, b = 5.983 Å, c = 6.977 Å, β = 106.32°	5.892	164.344	10	OQMD
029	TaPtO ₃	<i>P2</i> ₁ / <i>m</i>	a = 3.493 Å, b = 6.486 Å, c = 6.260 Å, β = 105.97°	10.328	141.824	10	OQMD
030	RbCrBr ₃	<i>P2</i> ₁ / <i>m</i>	a = 5.328 Å, b = 7.905 Å, c = 8.035 Å, β = 108.78°	3.910	338.417	10	OQMD
031	VSbO ₃	<i>P2</i> ₁ / <i>m</i>	a = 3.953 Å, b = 4.971 Å, c = 7.285 Å, β = 103.31°	5.262	143.153	10	OQMD
032	AgRhO ₃	<i>Pm</i>	a = 3.832 Å, b = 3.746 Å, c = 5.196 Å, β = 103.94°	5.937	74.587	5	OQMD
033	TiSN ₃	<i>Pm</i>	a = 4.806 Å, b = 3.049 Å, c = 4.809 Å, β = 113.07°	3.124	70.469	5	novel and unique
034	VPN ₃	<i>Pm</i>	a = 3.660 Å, b = 2.959 Å, c = 5.027 Å, β = 111.01°	4.050	54.442	5	novel and unique
035	VSN ₃	<i>Pm</i>	a = 4.922 Å, b = 2.797 Å, c = 4.980 Å, β = 115.64°	3.359	68.559	5	novel and unique
A₂BB'X₆ compounds							
036	Nd ₂ SiWO ₆	<i>Cm</i>	a = 9.182 Å, b = 5.277 Å, c = 6.096 Å, β = 96.93°	6.754	295.372	20	OQMD
037	Sb ₂ TiCuO ₆	<i>P2</i> ₁ / <i>c</i>	a = 5.452 Å, b = 6.046 Å, c = 9.152 Å, β = 125.65°	6.109	301.675	20	novel and unique
038	Y ₂ TiCrO ₆	<i>P2</i> ₁ / <i>c</i>	a = 5.289 Å, b = 5.751 Å, c = 9.147 Å, β = 124.78°	5.431	278.225	20	novel and unique
039	Y ₂ SrTiO ₆	<i>P2</i> ₁ / <i>c</i>	a = 5.519 Å, b = 5.801 Å, c = 9.968 Å, β = 121.54°	4.997	319.133	20	novel and unique
040	Yb ₂ NdTiO ₆	<i>P2</i> ₁ / <i>c</i>	a = 5.473 Å, b = 5.708 Å, c = 9.828 Å, β = 120.99°	8.003	307.026	20	novel and unique
041	Zr ₂ TiCuO ₆	<i>P2</i> ₁ / <i>c</i>	a = 5.038 Å, b = 5.29 Å, c = 10.115 Å, β = 118.39°	5.459	269.575	20	novel and unique
ORTHORHOMBIC							
ABX₃ compounds							
042	KPdCl ₃	<i>Amm</i> 2	a = 7.000 Å, b = 7.022 Å, c = 4.952 Å	3.437	243.411	10	OQMD
043	KSrCl ₃	<i>Amm</i> 2	a = 7.844 Å, b = 7.975 Å, c = 5.542 Å	2.233	346.685	10	OQMD
044	KSrF ₃	<i>Amm</i> 2	a = 6.635 Å, b = 6.926 Å, c = 4.746 Å	2.798	218.098	10	OQMD

045	RbGeH ₃	<i>Amm2</i>	a = 6.631 Å, b = 6.337 Å, c = 4.005 Å	3.179	168.293	10	novel and unique
046	RbMgCl ₃	<i>Amm2</i>	a = 7.144 Å, b = 7.143 Å, c = 5.036 Å	2.793	256.985	10	MP & OQMD
047	RbZnH ₃	<i>Amm2</i>	a = 6.628 Å, b = 6.780 Å, c = 3.336 Å	3.409	149.913	10	OQMD
048	LiTiH ₃	<i>Amm2</i>	a = 4.759 Å, b = 5.094 Å, c = 3.167 Å	2.502	76.776	10	novel and unique
049	RbTiF ₃	<i>Amm2</i>	a = 7.291 Å, b = 7.496 Å, c = 3.471 Å	3.332	189.702	10	OQMD
050	GdTlO ₃	<i>Amm2</i>	a = 6.503 Å, b = 9.747 Å, c = 3.526 Å	6.087	223.495	10	OQMD
051	GdYO ₃	<i>Amm2</i>	a = 6.455 Å, b = 9.861 Å, c = 3.508 Å	4.375	223.294	10	OQMD
A₂BB'X₆ compounds							
052	Al ₂ GdTlO ₆	<i>Pmm2</i>	a = 6.138 Å, b = 7.153 Å, c = 7.997 Å	4.617	351.109	20	novel and unique
053	Al ₂ PdCO ₆	<i>Imm2</i>	a = 5.598 Å, b = 6.064 Å, c = 7.457 Å	3.521	253.137	20	novel and unique
054	Al ₂ ErPdO ₆	<i>Pmm2</i>	a = 5.872 Å, b = 6.504 Å, c = 6.596 Å	5.585	251.911	20	novel and unique
055	Al ₂ TaPdO ₆	<i>Imm2</i>	a = 5.712 Å, b = 6.312 Å, c = 7.261 Å	5.548	261.789	20	novel and unique
056	Ge ₂ TaSiO ₆	<i>Pmm2</i>	a = 5.660 Å, b = 5.734 Å, c = 7.333 Å	6.283	237.988	20	novel and unique
057	La ₂ TaMoO ₆	<i>Pmm2</i>	a = 5.543 Å, b = 5.976 Å, c = 8.119 Å	8.034	268.942	20	novel and unique
058	Mg ₂ GdTlO ₆	<i>Imm2</i>	a = 6.194 Å, b = 6.857 Å, c = 7.513 Å	5.025	319.094	20	novel and unique
059	Sr ₂ GdTlO ₆	<i>Pmm2</i>	a = 5.942 Å, b = 6.077 Å, c = 8.427 Å	6.652	304.295	20	MP & OQMD
060	Sr ₂ PdCO ₆	<i>Imm2</i>	a = 5.696 Å, b = 6.327 Å, c = 6.545 Å	5.486	235.873	20	novel and unique
061	Tl ₂ GdTlO ₆	<i>Imm2</i>	a = 6.014 Å, b = 6.139 Å, c = 8.407 Å	9.021	310.386	20	novel and unique
062	Tl ₂ TaMgO ₆	<i>Pmm2</i>	a = 5.751 Å, b = 6.081 Å, c = 7.866 Å	8.573	275.088	20	OQMD
063	Tl ₂ YTlO ₆	<i>Pmm2</i>	a = 5.979 Å, b = 6.092 Å, c = 8.363 Å	8.446	304.614	20	novel and unique
064	V ₂ YTlO ₆	<i>Imm2</i>	a = 5.918 Å, b = 6.246 Å, c = 8.185 Å	5.135	302.549	20	OQMD
TETRAGONAL							
ABX₃ compounds							
065	ErAgS ₃	<i>P4/mmm</i>	a = 5.067 Å, c = 5.039 Å	4.767	129.374	5	novel and unique
066	EuErS ₃	<i>P4/mmm</i>	a = 5.243 Å, c = 5.106 Å	4.914	140.359	5	novel and unique
067	ErTaS ₃	<i>P4/mmm</i>	a = 4.889 Å, c = 5.099 Å	6.055	121.878	5	novel and unique
068	ErTiS ₃	<i>P4/mmm</i>	a = 4.973 Å, c = 5.144 Å	4.064	127.215	5	novel and unique
069	HoAgO ₃	<i>P4/mmm</i>	a = 4.453 Å, c = 4.043 Å	6.645	80.169	5	OQMD

070	HoAgS ₃	<i>P4/mmm</i>	a = 5.034 Å, c = 5.068 Å	4.770	128.429	5	novel and unique
071	HoTiS ₃	<i>P4/mmm</i>	a = 4.944 Å, c = 5.188 Å	4.047	126.811	5	novel and unique
072	KTaS ₃	<i>P4/mmm</i>	a = 4.910 Å, c = 5.897 Å	3.693	142.165	5	OQMD
073	TaPS ₃	<i>P4/mmm</i>	a = 4.897 Å, c = 4.823 Å	4.423	115.658	5	OQMD
074	TiSiO ₃	<i>P4/mmm</i>	a = 4.319 Å, c = 3.295 Å	3.348	61.464	5	OQMD
075	SmTaO ₃	<i>P4/mmm</i>	a = 4.266 Å, c = 4.158 Å	8.322	75.670	5	OQMD
076	TbAgS ₃	<i>P4/mmm</i>	a = 5.132 Å, c = 5.084 Å	4.500	133.899	5	novel and unique
077	TbNbO ₃	<i>P4/mmm</i>	a = 4.265 Å, c = 4.113 Å	6.655	74.816	5	OQMD
078	TbTaO ₃	<i>P4/mmm</i>	a = 4.254 Å, c = 4.109 Å	8.660	74.359	5	OQMD
079	TbTaS ₃	<i>P4/mmm</i>	a = 4.888 Å, c = 5.118 Å	5.923	122.282	5	novel and unique
080	RbVCl ₃	<i>P4/mmm</i>	a = 4.878 Å, c = 4.818 Å	3.517	114.644	5	MP & OQMD
A₂BB'X₆ compounds							
081	Ho ₂ EuAgO ₆	<i>P4mm</i>	a = 4.032 Å, c = 9.372 Å	7.472	152.361	10	novel and unique
082	Ho ₂ EuScO ₆	<i>P4mm</i>	a = 3.759 Å, c = 10.690 Å	6.847	151.051	10	novel and unique
083	Ho ₂ RbTaO ₆	<i>P4mm</i>	a = 3.920 Å, c = 10.546 Å	7.093	162.054	10	novel and unique
084	Ho ₂ YTiO ₆	<i>P4mm</i>	a = 3.875 Å, c = 10.470 Å	7.596	157.214	10	novel and unique
085	K ₂ RbAgI ₆	<i>P4mm</i>	a = 5.715 Å, c = 12.109 Å	4.338	395.495	10	novel and unique
086	K ₂ TaVO ₆	<i>P4mm</i>	a = 3.747 Å, c = 10.437 Å	4.602	146.536	10	novel and unique
087	La ₂ RbTaO ₆	<i>P4mm</i>	a = 4.101 Å, c = 9.463 Å	6.678	159.151	10	novel and unique
088	Sr ₂ GdSbO ₆	<i>P4mm</i>	a = 4.196 Å, c = 8.922 Å	5.817	157.084	10	MP & OQMD
089	Sr ₂ LuUO ₆	<i>P4mm</i>	a = 4.330 Å, c = 8.605 Å	7.042	161.334	10	MP
090	V ₂ EuTaO ₆	<i>P4mm</i>	a = 3.943 Å, c = 10.168 Å	5.575	158.084	10	novel and unique
TRIGONAL							
ABX₃ compounds							
091	RbHgCl ₃	<i>P$\bar{3}$m1</i>	a = 7.648 Å, c = 8.574 Å	3.001	501.510	10	OQMD
092	RbPtCl ₃	<i>P$\bar{3}$m1</i>	a = 7.495 Å, c = 7.168 Å	3.685	402.663	10	OQMD
093	LuBiSe ₃	<i>P$\bar{3}$m1</i>	a = 7.124 Å, c = 6.662 Å	7.041	338.106	10	novel and unique
094	RbBiSe ₃	<i>P$\bar{3}$m1</i>	a = 7.578 Å, c = 7.924 Å	4.477	455.044	10	novel and unique
095	VBiSe ₃	<i>P$\bar{3}$m1</i>	a = 7.046 Å, c = 6.337 Å	6.056	314.607	10	novel and unique
096	LuCrSe ₃	<i>P6₃/mmc</i>	a = 7.121 Å, c = 5.536 Å	6.336	280.723	10	novel and unique
097	RbCrSe ₃	<i>P$\bar{3}$m1</i>	a = 7.436 Å, c = 7.413 Å	3.502	409.895	10	OQMD
098	LaLuSe ₃	<i>P$\bar{3}$m1</i>	a = 6.946 Å, c = 7.555 Å	5.794	364.505	10	OQMD

099	LaVSe ₃	$P\bar{3}m1$	a = 6.676 Å, c = 7.111 Å	5.163	316.930	10	novel and unique
100	PtPbO ₃	$P\bar{3}m1$	a = 4.563 Å, c = 8.906 Å	9.311	185.432	10	OQMD
101	PtPbSe ₃	$P\bar{3}m1$	a = 5.619 Å, c = 9.014 Å	8.613	284.600	10	novel and unique
102	RhPbSe ₃	$P\bar{3}m1$	a = 6.573 Å, c = 6.662 Å	7.288	287.827	10	novel and unique
103	SiPbSe ₃	$P\bar{3}m1$	a = 7.202 Å, c = 6.553 Å	5.326	339.896	10	novel and unique
104	VPbSe ₃	$P\bar{3}m1$	a = 7.376 Å, c = 5.708 Å	6.114	310.546	10	novel and unique
105	ZnPbSe ₃	$P\bar{3}m1$	a = 6.636 Å, c = 7.200 Å	6.162	317.063	10	novel and unique
106	ScAuSe ₃	$P\bar{3}m1$	a = 5.911 Å, c = 7.153 Å	7.348	249.925	10	novel and unique
107	LuScSe ₃	$P\bar{3}m1$	a = 6.817 Å, c = 6.913 Å	5.453	321.257	10	novel and unique
108	LuUSE ₃	$P\bar{3}m1$	a = 6.681 Å, c = 7.214 Å	7.740	322.002	10	novel and unique
109	URhSe ₃	$P\bar{3}m1$	a = 6.195 Å, c = 6.469 Å	8.925	248.267	10	novel and unique
A₂BB'X₆ compounds							
110	Ag ₂ AsClF ₆	$P\bar{3}m1$	a = 5.026 Å, c = 7.020 Å	4.758	177.330	10	novel and unique
111	Ag ₂ RbHoF ₆	$P\bar{3}m1$	a = 6.854 Å, c = 7.250 Å	3.266	340.586	10	novel and unique
112	Ag ₂ TcClF ₆	$P\bar{3}m1$	a = 5.253 Å, c = 7.035 Å	4.576	194.124	10	novel and unique
113	Cd ₂ LuClF ₆	$P\bar{3}m1$	a = 7.270 Å, c = 5.555 Å	3.583	293.598	10	novel and unique
114	Cd ₂ RbAuF ₆	$P\bar{3}m1$	a = 7.137 Å, c = 6.217 Å	3.761	316.674	10	novel and unique
115	K ₂ HoAuF ₆	$P\bar{3}m1$	a = 6.121 Å, c = 6.203 Å	4.571	232.406	10	novel and unique
116	K ₂ LiHoF ₆	$P\bar{3}m1$	a = 6.003 Å, c = 5.534 Å	3.501	199.423	10	OQMD
117	K ₂ RbHoF ₆	$P\bar{3}m1$	a = 6.511 Å, c = 6.629 Å	3.020	281.024	10	OQMD
118	Rb ₂ LiAuBr ₆	$P\bar{3}m1$	a = 7.421 Å, c = 6.511 Å	4.568	358.569	10	OQMD
119	Rb ₂ LiDyF ₆	$P\bar{3}m1$	a = 6.150 Å, c = 5.592 Å	4.119	211.503	10	OQMD
120	Rb ₂ LiNiBr ₆	$P\bar{3}m1$	a = 7.300 Å, c = 6.366 Å	4.047	339.244	10	OQMD
121	Rb ₂ AuLuF ₆	$P\bar{3}m1$	a = 6.281 Å, c = 6.223 Å	5.129	245.503	10	novel and unique
122	Rb ₂ DyNiF ₆	$P\bar{3}m1$	a = 6.206 Å, c = 5.755 Å	4.378	221.651	10	novel and unique
123	Rb ₂ TcClF ₆	$P\bar{3}m1$	a = 5.705 Å, c = 7.136 Å	3.454	232.256	10	novel and unique
124	Rb ₂ CaTiF ₆	$P\bar{3}m1$	a = 6.295 Å, c = 5.893 Å	3.062	233.522	10	OQMD

Table S6. Thermodynamic stability and bandgap properties of new materials designed by the LCMGM. For the listed materials, DFT-calculations are strictly reported using GGA-PBE functionals.

CIF ID	Composition	DFT-relaxed total energy per atom (Ry/atom)	DFT-relaxed total energy (Ry)	Model-predicted total energy (eV)	Energy above convex hull (eV/atom)	Formation energy (eV/atom)	DFT - bandgap (eV)
CUBIC							
<i>ABX₃</i> compounds							
001	CsMnO ₃	-165.444	-827.218	-32.799	0.186	-1.580	0
002	CsZnBr ₃	-402.512	-2012.560	-15.806	0.000	-1.748	0.0859
003	SrCuI ₃	-480.071	-2400.356	-15.578	0.044	-1.043	0
004	MnPdS ₃	-199.452	-997.261	-28.994	0.423	-0.449	0
005	SrMnS ₃	-149.084	-745.419	-28.024	0.635	-0.953	0
006	SrPdI ₃	-485.199	-2425.997	-17.542	0.000	-1.221	0
007	SrPdS ₃	-194.133	-970.666	-21.460	0.981	-0.437	0
008	BeAuI ₃	-692.147	-3460.736	-13.387	0.016	-1.269 *	0
<i>A₂BB'X₆</i> compounds							
009	Co ₂ VPbO ₆	-52.188	-2087.527	-68.347	0.643	-1.165	0
010	Cs ₂ LiScF ₆	-34.030	-1361.204	-54.248	0.017	-3.569	6.7233
011	Dy ₂ HfIrO ₆	-164.404	-6576.141	-79.647	1.214	-2.195	0
012	Eu ₂ LaVO ₆	-50.194	-2007.745	-85.684	1.115	-2.139	0
013	K ₂ GePtF ₆	-64.325	-2573.017	-40.223	0.873	-1.879	0
014	Lu ₂ HoTlO ₆	-75.623	-3024.913	-79.941	0.000	-3.427	0
015	Nb ₂ CuWO ₆	-61.029	-2441.147	-75.774	1.543	-0.883	0
016	Rb ₂ LiMnF ₆	-28.467	-1138.661	-52.492	0.000	-2.894	0
017	Rb ₂ AlPdF ₆	-34.714	-1388.549	-45.750	0.411	-2.634	0
018	Sr ₂ VCuO ₆	-29.958	-1198.316	-68.379	0.298	-2.213	0
019	Sr ₂ CWO ₆	-52.477	-2099.069	-71.026	1.101	-1.577	2.0118
020	Sr ₂ VPbO ₆	-46.394	-1855.777	-66.692	0.572	-2.083	0
021	U ₂ CuWO ₆	-102.606	-4104.243	-72.444	2.610	-0.312	0
022	Yb ₂ LaHoO ₆	-65.341	-2613.651	-75.541	0.111	-3.326	0
023	Yb ₂ TlVO ₆	-65.075	-2602.987	-64.440	0.586	-2.023	0.1409
MONOCLINIC							
<i>ABX₃</i> compounds							
024	AlAsO ₃	-104.316	-1043.162	-67.583	0.428	-2.109	3.6212
025	NbAlO ₃	-99.824	-998.242	-80.558	0.734	-2.318	0
026	TaAlO ₃	-278.819	-2788.188	-83.671	0.866	-2.278	1.4996
027	TaAuO ₃	-626.344	-6263.444	-78.559	0.563	-1.861	0.3946
028	LaNbO ₃	-183.609	-1836.089	-87.842	0.547	-2.808	0
029	TaPtO ₃	-605.823	-6058.231	-72.194	1.828	-0.665	0
030	RbCrBr ₃	-368.234	-3682.336	-40.936	0.156	-0.985	0

031	VSbO ₃	-162.983	-1629.831	-67.787	0.977	-1.168	0
032	AgRhO ₃	-230.445	-1152.225	-25.581	0.811	-0.113	0
033	TiSN ₃	-66.931	-334.655	-35.996	0.877	-0.021	0
034	VPN ₃	-70.868	-354.342	-42.206	0.058	-0.771	0
035	VSN ₃	-72.978	-364.888	-35.891	0.913	0.237	0
A₂BB'X₆ compounds							
036	Nd ₂ SiWO ₆	-130.300	-2605.992	-72.259	1.825	-1.465	0
037	Sb ₂ TiCuO ₆	-159.870	-3197.390	-141.698	0.071	-2.091	0
038	Y ₂ TiCrO ₆	-124.841	-2496.813	-170.561	0.981	-2.511	0
039	Y ₂ SrTiO ₆	-126.064	-2521.275	-160.096	0.987	-2.784	3.2363
040	Yb ₂ NdTiO ₆	-223.631	-4472.616	-151.034	0.538	-3.009	0
041	Zr ₂ TiCuO ₆	-153.609	-3072.182	-157.729	1.339	-2.009	0
ORTHORHOMBIC							
ABX₃ compounds							
042	KPdCl ₃	-172.686	-1726.859	-17.820	0.000	-1.554	0
043	KSrCl ₃	-122.414	-1224.137	-19.168	0.108	-2.521	4.6614
044	KSrF ₃	-110.424	-1104.244	-23.240	0.495	-3.236	5.4440
045	RbGeH ₃	-111.795	-1117.949	-17.102	0.000	-0.357	1.0458
046	RbMgCl ₃	-124.004	-1240.038	-17.871	0.000	-2.306	4.1539
047	RbZnH ₃	-158.675	-1586.746	-14.677	0.000	-0.544	1.656
048	LiTiH ₃	-40.454	-404.538	-24.877	0.000	-1.071	0
049	RbTiF ₃	-120.435	-1204.349	-27.462	0.313	-2.865	0
050	GdTlO ₃	-431.943	-4319.428	-39.368	0.892	-1.617	0
051	GdYO ₃	-321.215	-3212.153	-47.808	1.464	-2.484	0
A₂BB'X₆ compounds							
052	Al ₂ GdTlO ₆	-275.688	-5513.768	-188.418	0.346	-3.110	0
053	Al ₂ PdCO ₆	-85.965	-1719.301	-67.402	0.830	-1.582	0
054	Al ₂ ErPdO ₆	-240.312	-4806.246	-145.773	0.362	-2.596	0
055	Al ₂ TaPdO ₆	-207.095	-4141.893	-75.720	0.836	-2.151	0
056	Ge ₂ TaSiO ₆	-215.748	-4314.963	-159.215	0.286	-2.340	0.2051
057	La ₂ TaMoO ₆	-331.488	-6629.754	-160.818	1.573	-1.815	0
058	Mg ₂ GdTlO ₆	-296.084	-5921.674	-91.504	0.179	-3.269	0
059	Sr ₂ GdTlO ₆	-319.882	-6397.642	-169.879	0.959	-2.594	0
060	Sr ₂ PdCO ₆	-130.118	-2602.354	-65.313	0.449	-1.784	0.0011
061	Tl ₂ GdTlO ₆	-435.106	-8702.126	-93.901	0.000	-3.356	0
062	Tl ₂ TaMgO ₆	-329.450	-6588.991	-147.939	0.000	-2.610	0
063	Tl ₂ YTlO ₆	-343.597	-6871.939	-162.669	0.000	-2.860	0
064	V ₂ YTlO ₆	-219.065	-4381.293	-85.706	1.126	-1.953	0
TETRAGONAL							
ABX₃ compounds							
065	ErAgS ₃	-463.556	-2317.782	-25.788	0.169	-1.196	0
066	EuErS ₃	-575.862	-2879.308	-32.905	1.160	-1.127	0

067	ErTaS ₃	-597.913	-2989.564	-33.687	0.974	-0.970	0
068	ErTiS ₃	-388.828	-1944.139	-32.054	0.697	-1.436	0
069	HoAgO ₃	-388.603	-1943.017	-29.822	0.702	-1.513	0
070	HoAgS ₃	-443.858	-2219.289	-24.004	0.503	-0.836	0
071	HoTiS ₃	-369.124	-1845.620	-32.807	0.551	-1.584	0
072	KTaS ₃	-308.218	-1541.088	-31.028	0.371	-2.486 *	0
073	TaPS ₃	-296.670	-1483.348	-25.165	1.995	0.903	0
074	TiSiO ₃	-71.039	-355.195	-38.376	1.107	-2.043	0
075	SmTaO ₃	-479.691	-2398.457	-45.258	0.687	-2.768	0
076	TbAgS ₃	-407.505	-2037.527	-22.947	0.770	-0.614	0
077	TbNbO ₃	-348.093	-1740.466	-42.225	0.845	-2.529	0
078	TbTaO ₃	-527.082	-2635.411	-45.243	0.716	-2.781	0
079	TbTaS ₃	-541.854	-2709.272	-35.035	0.723	-1.227	0
080	RbVCl ₃	-276.953	-1384.764	-23.907	0.000	-2.017	0
A₂BB'X₆ compounds							
081	Ho ₂ EuAgO ₆	-485.244	-4852.441	-77.667	0.344	-2.569	0
082	Ho ₂ EuScO ₆	-445.271	-4452.705	-94.589	0.000	-3.911	0
083	Ho ₂ RbTaO ₆	-464.585	-4645.848	-79.910	0.725	-2.822	0
084	Ho ₂ YTiO ₆	-429.607	-4296.068	-80.472	0.016	-3.279	0
085	K ₂ RbAgI ₆	-433.181	-4331.811	-25.887	0.000	-1.071	0
086	K ₂ TaVO ₆	-192.019	-1920.189	-78.647	0.172	-2.580	0
087	La ₂ RbTaO ₆	-319.073	-3190.733	-79.573	0.766	-2.718	0
088	Sr ₂ GdSbO ₆	-244.495	-2444.951	-71.593	0.980	-2.032	0
089	Sr ₂ LuUO ₆	-368.999	-3689.987	-80.529	0.552	-3.165	0
090	V ₂ EuTaO ₆	-302.663	-3026.625	-89.658	1.003	-1.966	0
TRIGONAL							
ABX₃ compounds							
091	RbHgCl ₃	-297.921	-2979.212	-24.552	0.054	-1.446	2.5859
092	RbPtCl ₃	-430.590	-4305.899	-32.057	0.417	-1.043	0
093	LuBiSe ₃	-555.890	-5558.899	-49.192	0.000	-1.140	0.0804
094	RbBiSe ₃	-470.351	-4703.510	-37.882	0.000	-0.717	0
095	VBiSe ₃	-465.169	-4651.687	-56.483	0.000	-0.957	0
096	LuCrSe ₃	-424.236	-4242.363	-55.385	0.582	-0.606	0
097	RbCrSe ₃	-338.737	-3387.368	-46.942	0.214	-0.470	0
098	LaLuSe ₃	-466.304	-4663.037	-55.438	0.394	-1.555	0.1010
099	LaVSe ₃	-375.596	-3755.964	-59.404	0.336	-1.039	0
100	PtPbO ₃	-533.946	-5339.457	-55.898	0.521	-0.664	0
101	PtPbSe ₃	-750.135	-7501.348	-40.603	0.451	-0.006	0
102	RhPbSe ₃	-509.222	-5092.220	-42.127	0.581	0.095	0
103	SiPbSe ₃	-424.548	-4245.479	-46.444	0.000	-0.720	0
104	VPbSe ₃	-458.016	-4580.155	-53.025	0.000	-0.646	0
105	ZnPbSe ₃	-525.183	-5251.833	-42.930	0.000	-1.201	0

106	ScAuSe ₃	-627.979	-6279.789	-50.205	0.000	-1.002	0
107	LuScSe ₃	-406.029	-4060.285	-54.567	0.589	-1.188	0
108	LuUSe ₃	-580.340	-5803.396	-65.350	0.208	-1.275	0
109	URhSe ₃	-540.873	-5408.732	-56.525	1.044	0.171	0
A₂BB'X₆ compounds							
110	Ag ₂ AsClF ₆	-190.717	-1907.172	-41.889	0.000	-2.179	0.7603
111	Ag ₂ RbHoF ₆	-317.470	-3174.699	-50.041	0.000	-3.029	0
112	Ag ₂ TcClF ₆	-194.672	-1946.716	-43.352	0.041	-1.755	0
113	Cd ₂ LuClF ₆	-353.812	-3538.123	-44.738	0.258	-2.863	1.7495
114	Cd ₂ RbAuF ₆	-358.508	-3585.077	-40.590	0.000	-2.600	0.5285
115	K ₂ HoAuF ₆	-382.348	-3823.483	-48.402	0.273	-2.980	0
116	K ₂ LiHoF ₆	-206.161	-2061.606	-54.417	0.000	-3.718	0
117	K ₂ RbHoF ₆	-228.643	-2286.429	-52.103	0.034	-3.580	0
118	Rb ₂ LiAuBr ₆	-497.836	-4978.356	-27.283	0.092	-1.032	0
119	Rb ₂ LiDyF ₆	-222.202	-2222.019	-51.676	0.107	-3.468	0.0008
120	Rb ₂ LiNiBr ₆	-363.021	-3630.207	-32.499	0.000	-1.303	0.1248
121	Rb ₂ AuLuF ₆	-450.354	-4503.540	-49.319	0.112	-3.104	1.6765
122	Rb ₂ DyNiF ₆	-263.580	-2635.796	-46.060	0.770	-2.519	0
123	Rb ₂ TcClF ₆	-131.235	-1312.350	-45.770	0.064	-2.367	0
124	Rb ₂ CaTiF ₆	-115.453	-1154.526	-50.753	0.425	-3.037	0

Asterisked (*) E_f values are corrected using DFT-computed E_{rel} evaluations

References

- Ong, S.P. et al. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
<https://doi.org/10.1016/j.commatsci.2012.10.028>.
- De Graef, M. & McHenry, M.E. Structure of materials: An introduction to crystallography, diffraction and symmetry. Cambridge University Press (2012).
- Lide, D. CRC Handbook of Chemistry and Physics, 85th Edition. Taylor & Francis (2004).
- Cordero, B. et al. Covalent radii revisited. *Dalton Trans.* **21**, 2832–2838 (2008).
<https://doi.org/10.1039/B801115J>
- Ho, C. Y., Powell, R. W. & Liley, P. E. Thermal conductivity of the elements. *J. Phys. Chem. Ref. Data* **1**, 279-421 (1972). <https://doi.org/10.1063/1.3253100>
- Chollet, F. & others. Keras. GitHub. (2015). <https://github.com/keras-team/keras>
- TensorFlow. GitHub. (2015). <https://github.com/tensorflow/tensorflow>

8. Scikit-Optimize. GitHub. (2019). <https://github.com/scikit-optimize/scikit-optimize>
9. Belsky, A., Hellenbrandt, M., Karen, V.L., & Luksch, P. New developments in the inorganic crystal structure database (ICSD): accessibility in support of materials research and design. *Acta Cryst.* **B58**, 364–369 (2002). <https://doi.org/10.1107/S0108768102006948>
10. Jain, A. et al. Commentary: the materials project: A materials genome approach to accelerating materials innovation. *Apl. Mater.* **1**, 011002 (2013). <https://doi.org/10.1063/1.4812323>
11. Kirklin, S. et al. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *npj Comput. Mater.* **1**, 15010 (2015). <https://doi.org/10.1038/npjcompumats.2015.10>
12. Draxl, C. & Scheffler, M. Nomad: the fair concept for big data-driven materials science. *MRS Bull.* **43**, 676–682 (2018). <https://doi.org/10.1557/mrs.2018.208>