npj computational materials

www.nature.com/npjcompumats

ARTICLE OPEN

Check for updates

Machine learning guided discovery of ternary compounds involving La and immiscible Co and Pb elements

Renhai Wang[1,2], Weiyi Xia[2,3], Tyler J. Slade[2,3], Xinyu Fan[4], Huafeng Dong[1], Kai-Ming Ho[2,3], Paul C. Canfield[2,3] and Cai-Zhuang Wang[2,3] ✉

Ternary compounds with an immiscible pair of elements are relatively unexplored but promising for novel quantum materials discovery. Exploring what third element and its ratio that can be added to make stable ternary compounds out of an immiscible pair of elements remains a great challenge. In this work, we combine a machine learning (ML) method with ab initio calculations to efficiently search for the energetically favorable ternary La-Co-Pb compounds containing immiscible elements Co and Pb. Three previously reported structures are correctly captured by our approach. Moreover, we predict a ground state La_3CoPb compound and 57 low-energy La-Co-Pb ternary compounds. Attempts to synthesize La_3CoPb via multiple techniques produce mixed or multi-phases samples with, at best, ambiguous signals of the predicted lowest-energy La_3CoPb and the second lowest-energy $La_{18}Co_{28}Pb_3$ phases. The calculated results of Gibbs free energy are consistent with experiments, and will provide very useful guidance for further experimental synthesis.

npj Computational Materials (2022)8:258 ; https://doi.org/10.1038/s41524-022-00950-0

INTRODUCTION

Regions of ternary phase space involving two immiscible elements (such as Co-Pb, Ba-Fe or La-Cr) with a third element added are relatively unexplored but promising for novel quantum materials discovery[1]. Ternary compounds containing such immiscible pairs are relatively rare, but when they do form, the immiscible elements are normally segregated from each other with the third element encapsulating or separating them. This leads to well defined reduced dimensionality and clear one-dimensional (1-D) and two-dimensional (2-D) structures[2,3] (i.e., bulk materials with packing chains or stacking layers). When a 3*d*-transition metal (TM) is forced to adopt such reduced dimensionality, complex electronic and magnetic states such as superconductivity (**BaFe**$_2$As$_2$)[4], fragile magnetism (**LaCr**Ge$_3$)[5,6] or even ferromagnetism (La$_6$**Co**$_{13}$**Pb**)[7] can emerge (with the immiscible pair elements shown in **bold**). The heart of the "antagonistic-pairs" search is the question, "*Which third element and at what ratio can be added to make stable ternary compounds out of the immiscible pair of elements*?" In order to address this outstanding question, knowledge about the relationship among the chemical compositions, crystal structures and their relative thermodynamic stability are crucial.

However, for a system even with three specified chemical elements, the number of possible combination of chemical compositions and the crystal structures they may take is enormous. It will be very time consuming for experiment to try all these possibilities or for direct ab initio high-throughput calculation to examine these huge number of combinations to discover the stable and metastable ternary compounds. As a necessary first step toward addressing the outstanding question posted above, the purpose of this paper is to integrate available deep ML methods with database and the state-of-the-art computational tools to efficiently explore such a complex relationship through accelerated discovery of compounds.

We primarily focus on the ternaries containing 3*d* TM or/and lanthanide elements, as such systems are likely to possess rich emergent electronic and magnetic properties relevant to the quantum materials community. We use La-Co-Pb system involving the involving immiscible pairs of Co and Pb elements as a prototype to demonstrate the "proof-of-principle" of our approach, while the ML-harnessed approach developed from this study is robust for general novel complex materials discovery and can be well applied to other systems of interest.

The Co-Pb binary phase diagram[8] clearly shows the two elements to be highly immiscible, with no stable binary compounds and effectively zero solubility up to at least 1900 K, as shown in the Supplementary Fig. 1 in the Supporting Information. However, several ternary compounds containing this immiscible pair have been reported, with most containing a rare Earth element (i.e., La$_6$Co$_{13}$Pb, La$_5$CoPb$_3$, La$_{12}$Co$_6$Pb)[1]. We show that the ML approach can correctly capture the three known ternary compounds even they are excluded from the database used in our ML training. Moreover, our ML approach predicts a ground state La_3CoPb compound with its formation energy below the currently known convex hull for La-Co-Pb ternary system. 57 metastable La-Co-Pb ternary phases with formation energy <100 meV/atom above the convex hull are also predicted, among them 12 structures have the formation energy within 50 meV/atom above the convex hull. Experimental attempts to synthesize the two low-energy ternary compounds (i.e., La_3CoPb and $La_{18}Co_{28}Pb_3$) are also made. We also studied the entropy effects on the relative thermodynamic stability of the five competing ternary phases (the three known and the two from our prediction) as the function of temperature to understand the experimental synthesis results.

[1]School of Physics and Optoelectronic Engineering, Guangdong University of Technology, 510006 Guangzhou, China. [2]Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA. [3]Ames Laboratory-USDOE, Iowa State University, Ames, IA 50011, USA. [4]Department of Physics, University of Science and Technology of China, 230026 Hefei, China. ✉email: wangcz@ameslab.gov

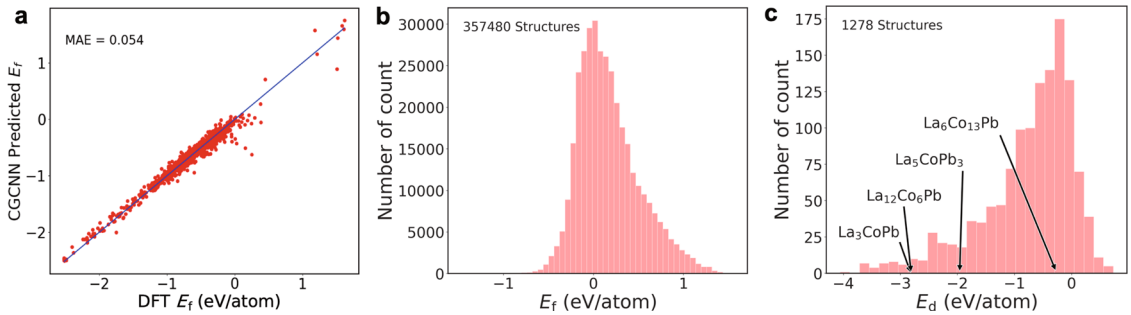

Fig. 1 Machine learning screening of La-Co-Pb ternaries. a The relationship between the formation energy predicted by CGCNN and the values calculated by DFT. MAE stands for mean absolute error. **b** Distribution of formation energy E_f for 357480 hypothetical La-Co-Pb ternary compounds (see text) predicted by the trained CGCNN mode. **c** Distribution of E_d (formation energy relative to the convex hull) calculated by our CGCNN-ML model for the 1278 La-Co-Pb ternary compounds that have negative E_f predicted by our ML model. Note that the three structures with the same chemical composition as the known compound ($La_6Co_{13}Pb$, La_5CoPb_3, $La_{12}Co_6Pb$) as indicated are in this short list.

RESULTS AND DISCUSSIONS

ML-guided DFT prediction of low-energy ternary compounds

We adopt a crystal graph convolutional neural network (CGCNN) model[9] for our deep ML approach. In CGCNN, a crystal structure is represented by a crystal graph that encodes both atomic information and bonding interactions between atoms. A convolutional neural network is then built on top of the graph to construct the proper descriptors that are optimal for predicting target properties. The CGCNN model is used here to provide fast predictions for the relationship among the crystal structures, chemical compositions, and formation energies. For more details about the CGCNN model, we refer the readers to Ref. [9]. We will use the CGCNN approach to perform rapid screening for the energetically favorable candidate La-Co-Pb ternary compounds. A small set of promising structures from the CGCNN ML model prediction are then selected for further optimization and validation by ab initio calculations. There are several newer versions and models of CGCNN[10,11] since Ref. [9] was published. These newer versions mainly modify the graph representations from the original CGCNN to better fit the systems of their interest. Although these modified CGCNN models may perform better than the original one for the systems of their interest, there are not enough data to demonstrate the newer versions are substantially better than the original one in general cases. While these modified models can be further explored, we found the original CGCNN is simpler and can provide faster high-throughput screening to help us quickly select promising candidate for density functional theory (DFT) calculations.

We train a CGCNN model for predicting the formation energy E_f (see the method section for definition of E_f) of compounds according to Ref. [9], using the structures and the ab initio calculated formation energies of 46744 binary and ternary compounds (excluding the three known La-Co-Pb ternaries) from the Materials Project (MP) database[12]. We split the dataset randomly into 80%, 10% and 10% as training set, validation set and test set, respectively. Parameters are all optimized under the guidance of Ref. [9]. The best-performing model that provides the lowest mean absolute error (MAE) for the validation set is selected from all the 200 epochs during the computation. More details of the CGCNN set up and training in this study are also given in the method section. To validate the accuracy of the CGCNN model for ternary compounds, we select all ternary structures with experimental ID (11916 crystal structures in total) from the MP database, and use the trained CGCNN model to predict their formation energies. The comparison between the CGCNN predicted values and the DFT calculated values are shown in Fig. 1a. It can be seen that the overall prediction performs well, with the MAE only 0.054 eV/atom. This result suggests that the CGCNN model describes reasonably well the formation energies of these ternary compounds.

In order to efficiently search for possible stable La-Co-Pb ternary compounds over a wide range of compositions, we replace the three elements on the lattices of the 11916 ternary compounds from the MP database by La, Co, and Pb and rapidly screen the formation energies E_f of these hypothetical La-Co-Pb compounds by our trained CGCNN model. There are 6 ways to shuffle the order of the 3 elements on the lattice of a given ternary compound. We also allow the volume of the crystals to vary by a scaling factor from 0.96 to 1.04 in an interval of 0.02. Therefore, the CGCNN model predicts the formation energies for 357480 La-Co-Pb ternary structures. The formation energy distribution from the CGCNN prediction is shown in Fig. 1b. It can be seen that the ML model predicts a noticeable fraction of La-Co-Pb ternaries to have negative formation energies E_f.

Using the trained CGCNN energy model, we then evaluate the formation energies E_d relative to the currently known convex hull for the hypothetical La-Co-Pb ternary compounds with negative formation energies E_f from the CGCNN model (see the method section for definition of E_d). The distribution of E_d from our ML model calculations is shown in Fig. 1c. E_d shown in Fig. 1c is calculated using the total energies of the ternary structures predicted from the CGCNN model while the energies of the reference stable ground state phases are from the MP database which are obtained by DFT calculations. Therefore, the E_d values have a systematic down-shift because the total energy from CGCNN is systematically lower than the corresponding values from DFT calculations (see Supplementary Fig. 11 in the Supporting Information). We only use the relative values of E_d from this plot to help select promising candidate structures for further accurate evaluation by DFT calculations. Note that for structures with the same chemical stoichiometry, only the structure with lowest formation energy is considered in the plot. There are 1278 structures in Fig. 1c, which is less than 0.36% of the initial structures used in the CGCNN screening. Therefore, the ML helps to select only small fraction of the hypothetical La-Co-Pb ternary structures (1278 structures in the present case) for further studies by ab initio calculations, thus greatly accelerating the pace of materials discovery.

We first validate the predictive capability of our ML-guided approach by examining whether the three known ternary La-Co-Pb compounds (i.e., $La_6Co_{13}Pb$, La_5CoPb_3, $La_{12}Co_6Pb$) are captured even though these known phases were excluded from our ML training database and from the hypothetical structure pool. We can see that three structures with the compositions the same to the known compounds are in the 1278 candidate structures selected by our CGCNN prediction, as shown in Fig. 1c. After further relaxation by ab initio calculations, these three hypothetical structures convert to the experimentally known $La_6Co_{13}Pb$, La_5CoPb_3, $La_{12}Co_6Pb$ compounds, as shown in Supplementary Fig. 2 in the supporting information. We therefore verify that the

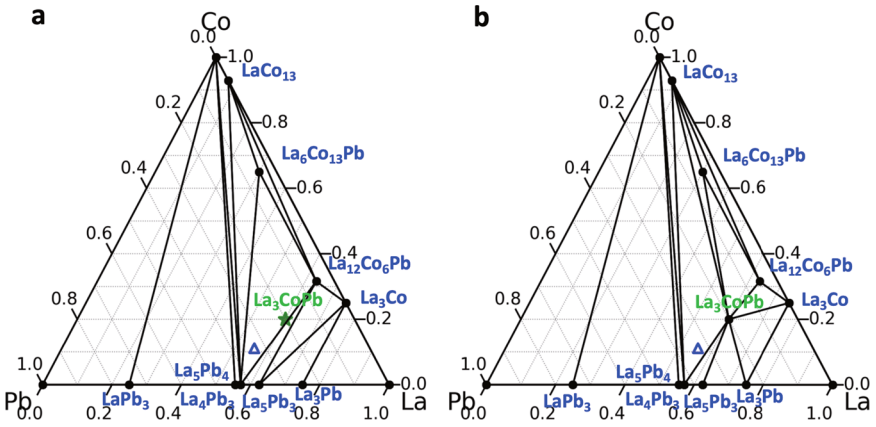

Fig. 2 Convex hull of La-Co-Pb system. Calculated convex hull of the formation energies in the La-Co-Pb system: **a** based on previously reported phases and **b** after including the predicted La_3CoPb phase. All calculations were carried out at 0 K. Black solid circles in the ternary phase diagram represent stable compounds and the green solid star represents a predicted stable phase, while blue hollow triangles indicate the metastable ternary phases. The black lines separate the composition space into Gibbs triangles.

three known La-Co-Pb ternary compounds can be captured by our CGCNN-ML prediction in combination with ab initio calculations.

In addition to capturing the three known ternary compounds, our ab initio calculations for the 1278 candidate structures selected from our CGCNN prediction also lead to the discovery a set of low-energy ternary La-Co-Pb compounds. In order to determine if any of the ternary compounds is at a global energy minimum after the DFT relaxation, we construct the convex hull for the La-Co-Pb system as shown in Fig. 2. The La-Co-Pb ternary convex hull shown in Fig. 2a is constructed based on ab initio calculations using the known stable binary and ternary structures formed by La, Co, and Pb from the commonly used databases, including MP[12], Springer Materials (*materials.springer.com*) and Inorganic Crystal Structure Database (*icsd.fiz-karlsruhe.de*). There are some reported binary and ternary phases (see Supplementary Table 1 in Supporting Information) for this system. Among them, seven binary phases La_3Co, $LaCo_{13}$, $LaPb_3$, La_5Pb_4, La_4Pb_3, La_5Pb_3 and La_3Pb and two ternary phases, $La_{12}Co_6Pb$, and $La_6Co_{13}Pb$ are energetically stable according to our ab initio calculations. Interestingly, our ab initio calculations show that a La_3CoPb compound discovered by our ML-guided approach discussed above has formation energy below the known La-Co-Pb convex hull, indicating that the La_3CoPb is at a global energy minimum. The thermodynamic stability of La_3CoPb is determined by the Gibbs triangle formed by $La_6Co_{13}Pb$, $La_{12}Co_6Pb$ and La_4Pb_3 as shown in Fig. 2a: $La_3CoPb \rightarrow (2La_6Co_{13}Pb + 51La_{12}Co_6Pb + 93La_4Pb_3)/332$. Our ab initio calculations show that the formation energy of La_3CoPb is about 32.5 meV/atom below the known convex hull. An updated convex hull for the La-Co-Pb system including this La_3CoPb phase is shown in Fig. 2b. It is interesting to note that although La_5CoPb_3 compound has been synthesized by experiment, its formation energy is above the updated convex hull by 35 meV/atom according to the present ab initio calculations and shown by the blue triangle in Fig. 2a, b. Therefore, La_5CoPb_3 is a metastable phase at zero temperature.

The crystal structure of the discovered La_3CoPb compound is plotted in Fig. 3a. It is a typical perovskite structure, with a *Pm-3m* space group symmetry and a lattice constant of 5.331 Å at the GGA-PBE level of calculation. The crystallographic details of the La_3CoPb structure are given in Supplementary Table 2 in the Supporting Information. There is one Wyckoff site for each of La (3*c*), Co (1b) and Pb (1a). To further investigate the dynamical stability, we calculated the phonon dispersion for La_3CoPb using the density functional perturbation theory via Phonopy code[13]. The phonon dispersion curve and the phonon density of states (DOS) of the La_3CoPb phase are plotted in Fig. 3b, showing that the La_3CoPb structure is also dynamically stable without any soft (negative) phonon modes. The electronic band structure and DOS in Fig. 3c, d shows a metallic state with somewhat localized La states near the Fermi level.

Besides predicting the stable La_3CoPb compound, our ML-guided ab initio approach also reveals 57 ternary La-Co-Pb compounds over a wide range of compositions which have formation energy within 100 meV/atom above the convex hull. Among them, 12 compounds (not including the known La_5CoPb_3 compound which has E_d of 35 meV/atom) have formation energy within 50 meV/atom above the convex hull. In Fig. 4, the structures of 8 of the low-energy metastable structures with $E_d < 40$ meV/atom above the convex hull are plotted. More detailed information about the 57 ternary La-Co-Pb compounds which have formation energies within 100 meV/atom above the convex hull are given in the Supporting Information. Some of these low-energy metastable structures could be synthesizable by experiment, especially under non-equilibrium synthesis conditions.

To validate the ML model after obtaining the DFT energies, we made a scatter plot between the CGCNN predicted total energy and our DFT calculated total energy for the 1278 candidate structures we selected from CGCNN screening, as shown in Supplementary Fig. 11. Through the validation, we get a MAE of 0.320 eV/atom by comparing the DFT energies and the values predicted with CGCNN model, indicating that the CGCNN model is reasonably good for total energy prediction. From the plot in Supplementary Fig. 11, we can also see that the energies from CGCNN are systematically lower than those from DFT but the trend of relative energies obtained by the CGCNN is in a reasonable good accuracy.

To further verify the novelty of our predicted structures, we compared the energetic stability of these structures with those listed in the popular AFLOWlib[14] database. We found 96 predicted ternary La-Co-Pb structures listed in AFLOWlib. These structures cover 10 different La-Co-Pb compositions. We emphasize that the thermodynamic stability of ternary compounds is more closely determined by their formation energy relative to the convex hull E_d, rather than the formation energy E_f which is calculated with respect to the energies of the three elementary phases. We examined the lowest-energy structure from each of the 10 compositions from AFLOWlib and found all of them were included in our 1278 candidate structures selected by our ML model, although we did not include the AFLOWlib database in our ML. We calculated the E_d by DFT for the lowest-energy structures from each of the 10 compositions, and found that all of them have E_d higher than 150 meV/atom, as listed in Supplementary Table 13.

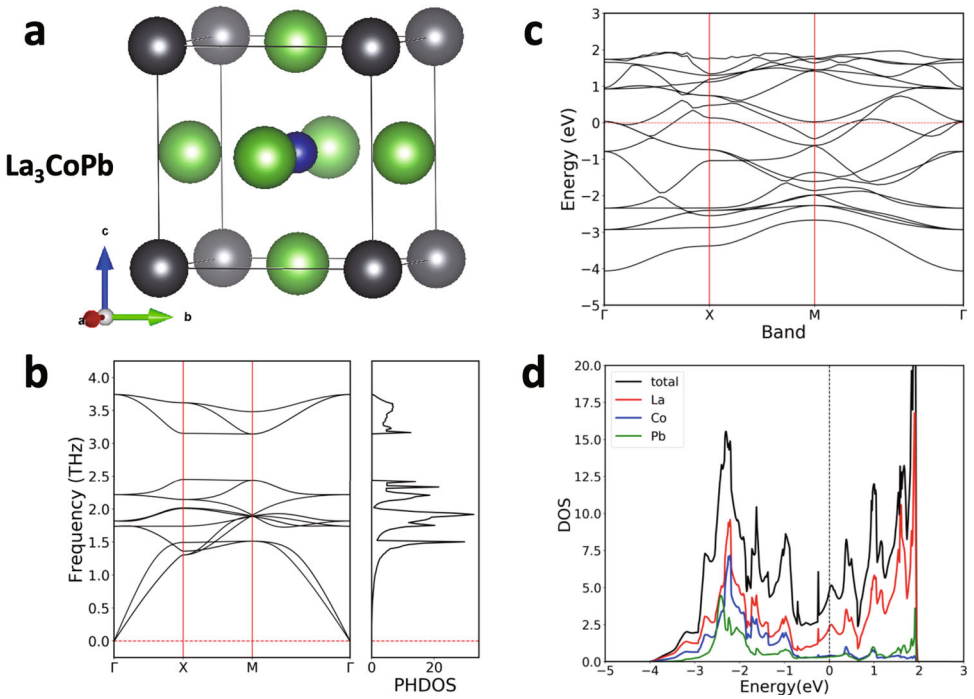

Fig. 3 Calculated electronic and thermal dynamic properties of La_3CoPb. **a** Low-energy La_3CoPb ternary compound from our prediction. The green, blue, and gray balls represent the La, Co, and Pb atoms respectively. **b** The phonon dispersion and phonon density-of-states of the predicted La_3CoPb compound. **c** and **d** Electronic band structure and density-of-states of La_3CoPb.

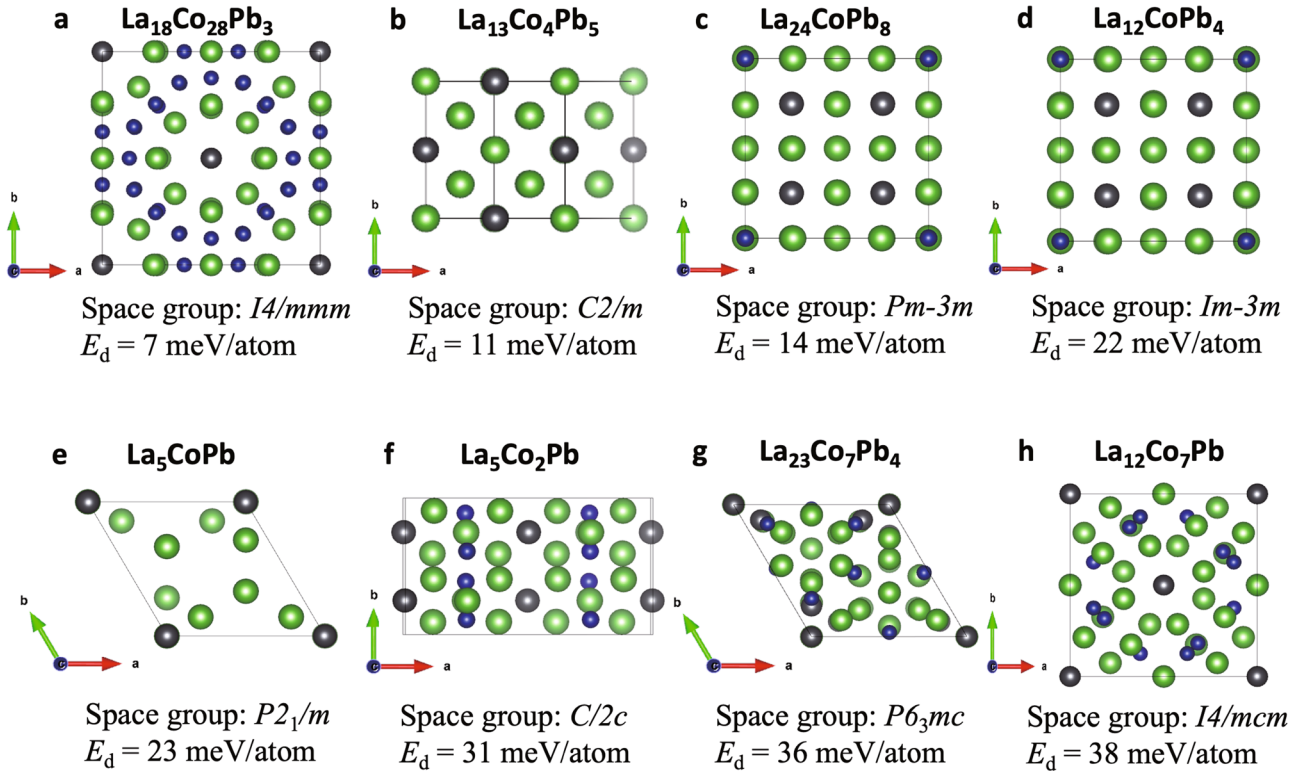

Fig. 4 Crystal structures of predicted low-energy La-Co-Pb compounds. The crystal structures of the 8 low-energy metastable structures of La-Co-Pb ternary compounds obtained by our ML + DFT prediction: **a** $La_{18}Co_{28}Pb_3$, **b** $La_{13}Co_4Pb_5$, **c** $La_{24}CoPb_8$, **d** $La_{12}CoPb_4$, **e** La_5CoPb, **f** La_5Co_2Pb, **g** $La_{23}Co_7Pb_4$, and **h** $La_{12}Co_7Pb$. The space group symmetries and the formation energy relative to the convex hull (E_d) of these structures are also shown.

These results indicate that all the ternary La-Co-Pb compound predicted by AFLOWlib have the formation energy at least 150 meV/atom above the convex hull. In comparison, our present study predicted 58 structures with their formation energies <100 meV/atom above the convex hull. Since our study aims at predicting novel ternary compounds with E_d <100 meV/atom, the structures listed in AFLOWlib will not be further discussed in this paper.

Experimental verification attempts

To determine if the predicted La_3CoPb and $La_{18}Co_{28}Pb_3$ compounds can be synthesized, we carried out multiple synthetic attempts using a variety of techniques. We first carried out several exploratory reactions as follows. Elemental La (Ames Laboratory), Co (Alfa-Aesar, 99.99%), and Pb (Alfa-Aesar, 99.99%) were mixed in nominally 3:1:1 and 18:28:3 molar ratios respectively and arc melted together for each mixture. Each button was flipped and re-melted five times to ensure homogeneity. After melting, the buttons were placed in a Ta crucible and flame sealed under vacuum inside a fused silica ampule. The two samples obtained from such arc melt procedure were then annealed at 650 °C for 1 week, and at 700 °C for 2 weeks respectively. After the annealing step, the sample was quenched by immersing the ampule in water. Upon opening the ampule, the original La-Co-Pb buttons clearly had partially melted during annealing. We refer to the samples prepared with different molar ratios of element mixtures as 3:1:1 and 18:28:3 samples respectively. In addition to the above trials, we furthermore attempted to synthesize the La_3CoPb phase by annealing under pressure using a high temperature, high pressure Rockwell furnace. Here, we prepared and arc-melted La, Co, and Pb in a nominally 3:1:1 molar ratio as described above. The button was then ground to a fine powder using a mortar and pestle inside an N_2 filled glovebox. The powder was contained in an MgO crucible, which was held inside a small cubic graphite heater and placed inside a pyrophyllite cube, the latter of which was used as the pressure medium. The sample was brought to ~3.2 GPa at room temperature, then heated up to 700 °C, where the sample was annealed for 72 h under the applied pressure. After cooling to ambient temperature, the pressure was released, and the sample was mechanically separated from the MgO crucible inside an N_2 filled glovebox.

To determine what compounds formed after arc melting and annealing, we analyzed pieces of the button with powder X-ray diffraction (PXRD). Figure 5 shows the powder patterns obtained on the 3:1:1 sample (left column) and the 18:28:3 sample (right column) respectively after annealing at 650 °C for 1 week. The PXRD patterns for the unannealed samples and 700 °C/2-weeks annealing samples are similar to those shown in Fig. 5, as can be seen from the Supporting Information. The experimental PXRD patterns are clearly complex with many partially overlapping peaks. We find that each of our experimental samples consists of mixture phases and it is difficult to unambiguously determine these phases through fitting analysis, but the vast majority of the peaks can be associated with known compounds.

In order to see if the experimental samples contain any trace of the predicted La_3CoPb and $La_{18}Co_{28}Pb_3$ ternary phases, we calculated the PXRD patterns for these two phases using the structures from our prediction and compare the calculation results with the experimental PXRD patterns as shown in Fig. Fig. 5a--d respectively. The predicted stable La_3CoPb compound has high symmetry, so the calculated XRD exhibits only 6 peaks in the range of 10–60 degrees. These 6 peaks can find partial alignment with the peaks in the experimental PXRD pattern. The agreement between experimental and predicted patterns is better in the 3:1:1 sample than the 18:28:3 sample. We note that if we allow the lattice constant of the La_3CoPb phase to reduce by 2.0%, the alignment between the calculated PXRD peaks and the experimental ones are better, as shown in Supplementary Fig. 3 in the Supporting Information. Based on the accuracy of the GGA-PBE calculations, discrepancy of 2.0% in the lattice constant between calculation and experiment for La-Co-Pb compounds is possible. For the predicted metastable $La_{18}Co_{28}Pb_3$ compound (which has formation energy 7 meV/atom above the convex hull) shown in Fig. 4a, some of the peaks in the calculated PXRD pattern overlap with experimental PXRD pattern, but the agreement with experiment is marginal. We can also see in Fig. 5d that the calculated pattern for the $La_{18}Co_{28}Pb_3$ phase agrees better with experimental data from the sample with a nominally 18:28:3 (La:Co:Pb) molar ratio. From these comparisons, we can see that the predicted La_3CoPb and $La_{18}Co_{28}Pb_3$ ternary phases may be included in the mixture of phases in the experimental samples, but they are not the dominant ones, and the results are ultimately ambiguous given the complexity of the patterns.

We also calculate the PXRD patterns for the three known ternary compounds i.e., $La_6Co_{13}Pb$, La_5CoPb_3, $La_{12}Co_6Pb$, to see to what extent the experimental samples contain these three known compounds. We find that among these three phases, the calculated PXRD pattern for the $La_6Co_{13}Pb$ agrees better with our experimental data. In Fig. 5e, f, the calculated PXRD pattern from the $La_6Co_{13}Pb$ phase is plotted in comparison of our experimental pattern from the 3:1:1 and 18:28:3 samples after annealing for one week at 650 °C. Comparisons with the pattern from the two experimental samples with no annealing and after annealing at 700 °C for 2 weeks are similar to that shown in Fig. 5e, f, as one can see from Supplementary Fig. 4 and Supplementary Fig. 5 in the Supporting Information. Comparisons of the PXRDs of the other two known phases with the PXRD data from our experimental samples are also shown in Supplementary Fig. 6 and Supplementary Fig. 7 in the Supporting Information. We can see that some of the peaks of these three known ternary phases overlap with experimental PXRD spectra, but the agreement with experiment is also far from perfect. These comparisons suggest the three known ternary phases are also among the mixture of phases in the experimental samples, but they are also not the dominant phases.

Furthermore, we note that our experimental samples also contain binary phases. For example, since the La_3CoPb phase is not the dominant phase in the arc melted sample prepared with in a nominally 3:1:1 (La:Co:Pb) molar ratio, other La-rich phases are expected. In Supplementary Fig. 8 and Supplementary Fig. 9 in the Supporting Information, we show the calculated PXRDs for two La-rich binary phases La_3Co and La_5Pb_3. One can see that many of the peaks of these two binary phases also overlap with experimental PXRD spectra, strongly suggesting these two phases are present the final mixtures.

Finally, our attempts to synthesize La_3CoPb under pressure also produced mixed phase samples. The powder data is shown in Supplementary Fig. 10 in the Supporting Information. Prior to applying pressure, the powder pattern is effectively the same as that shown in Fig. 5. Significant changes are observed after annealing under pressure; however, again, the majority of the peaks can be at least partially matched with those expected from known binary or ternary phases in the La-Co-Pb phase space. In particular, La_5Pb_3, $La_2Co_{1.7}$, $La_{12}Co_6Pb$, and $La_6Co_{13}Pb$ are all likely present in the final mixture. Like in the ambient pressure experiments, the experimental data does show partial agreement with some of the peaks expected from La_3CoPb, but again the results are ambiguous, and it would be at best present as a secondary phase.

In order to gain insights into the thermodynamic stability of the competing compounds at finite temperature, we evaluate the Gibbs free energy as a function of temperature for relevant low-energy binary and ternary phases in the La-Co-Pb system by including the vibrational entropy and electronic entropy contributions, as shown in more detail in Method section. With the values

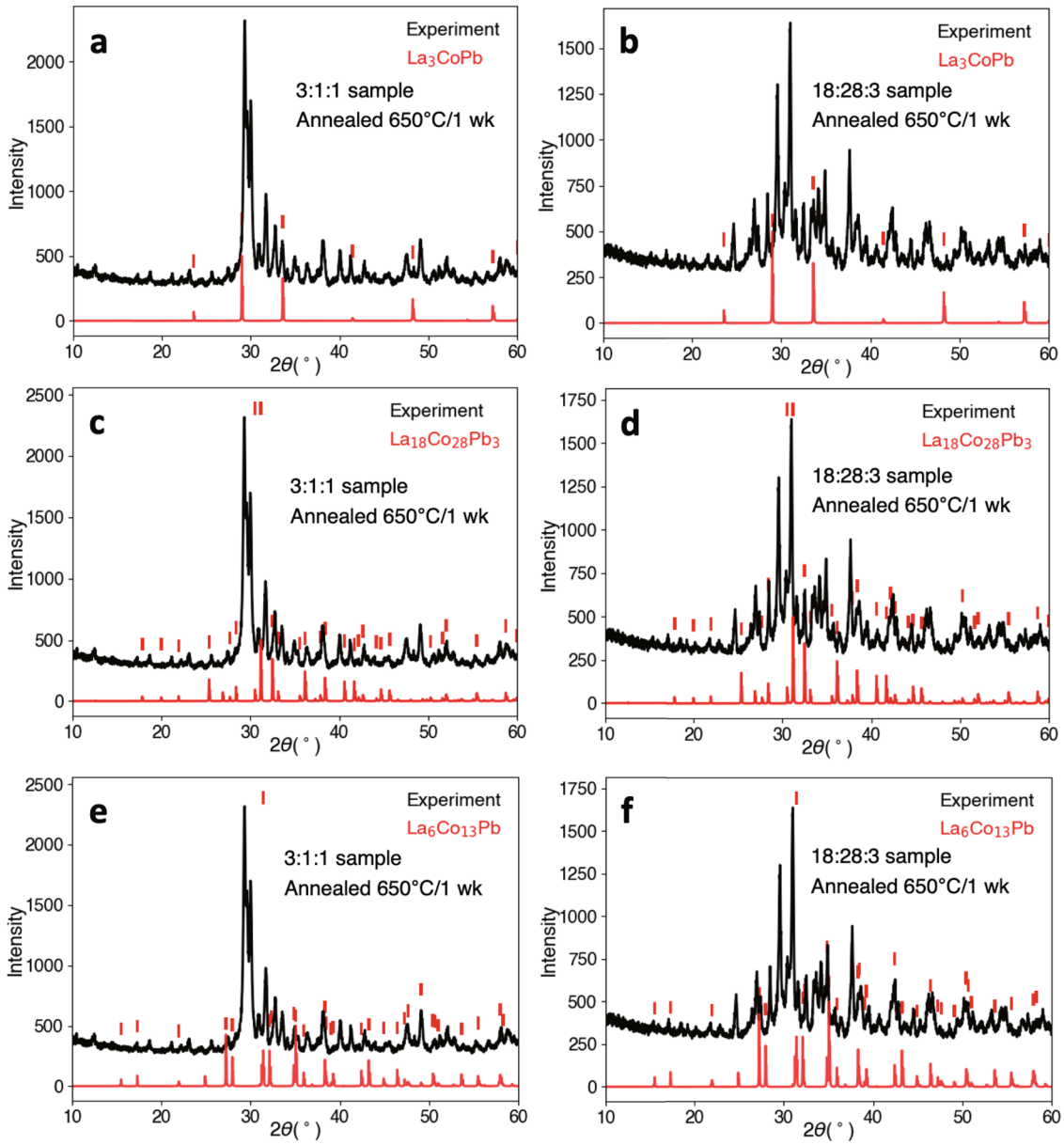

Fig. 5 Powder diffraction patterns of experimental samples. Powder diffraction patterns obtained from pieces of the arc melted La-Co-Pb pellet (left column (**a**, **c** and **e**): the 3:1:1 sample; right column (**b**, **d** and **f**): the 18:28:3 sample) after annealing 650 °C for 1 week, compared to calculated PXRD of the predicted La_3CoPb (**a** and **b**) and $La_{18}Co_{28}Pb_3$ (**c** and **d**) ternary phases, and the known $La_6Co_{13}Pb$ phase (**e** and **f**), respectively.

of Gibbs free energies as the function of temperature for these phases, the formation Gibbs energy (G_d, also see the Methods section for more details) of each ternary phase with respect to the nearby competing binary and ternary phases in the ternary convex hull as the function of temperature can be calculated. In order to obtained G_d for the five competing phases (i.e., La_3CoPb, La_5CoPb_3, $La_{12}Co_6Pb$, $La_6Co_{13}Pb$ and $La_{18}Co_{28}Pb_3$), Gibbs free energies of three nearby binary phases ($LaCo_{13}$, La_3Co and La_4Pb_3) are also calculated. The G_d of the five low-energy ternary La-Co-Pb phases as the function of temperature is plotted in Fig. 6. We can see that the known $La_{12}Co_6Pb$ phase is the most stable one at low temperature (below 120 K) but becomes the least stable among the five ternary phases at high temperature. The predicted La_3CoPb and $La_{18}Co_{28}Pb_3$ phases are the most stable ones between 120 K and 200 K. Above 200 K, the known $La_6Co_{13}Pb$ phase is much more stable than other phases. These calculation results are consistent with the results from our experimental synthesis attempts that although some traces of the predicted La_3CoPb and $La_{18}Co_{28}Pb_3$ phases can be detected in the experimental samples, large fraction of such phases are difficult to be produced by direct high temperature synthesis methods. We can also see that the calculation of the Gibbs free energy of the structures at high temperature is crucial to understand the thermodynamic stability of the compounds.

In summary, we searched for possible thermodynamically stable ternary La-Co-Pb crystal structures containing the immiscible Co-Pb pair of elements with different compositions using a CGCNN ML method and ab initio calculations. Our approach successfully captured the three known ternary compounds ($La_6Co_{13}Pb$, La_5CoPb_3, $La_{12}Co_6Pb$) in the La-Co-Pb system. We also predicted a perovskite structure of La_3CoPb to be energetically and dynamically stable based on ab initio GGA-PBE calculations, and further predicted a number of metastable ternary structures which may be synthesizable experimentally. The convex hull for ternary

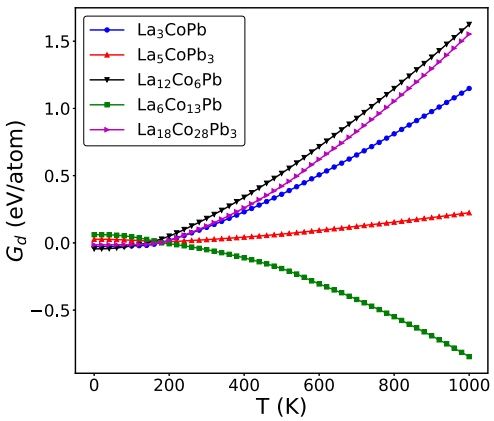

Fig. 6 Calculated Gibbs free energy for La-Co-Pb compounds. Calculated G_d as a function of temperature for the five competitive ternary La-Co-Pb compounds.

La-Co-Pb system has also been updated based on the compounds discovered from our ML-guided approach.

We demonstrated that the combination of ML with ab initio calculation is a powerful approach to accelerate the exploration of low-energy crystal structures for complex compounds. In the studies presented in this paper, DFT calculations are only needed for 1278 structures (which is less than 0.36% of total structures screened by ML) selected by ML to recover the three known ternary phases and predict a ground state and a number of potentially metastable ternary phases for La-Co-Pb system. Such ab initio calculations are easily manageable. According to the timing from our calculations, we estimate that such calculations can be done in 1 day on a cluster computer of 500 nodes (24–32 cores per node). Therefore, such an approach can provide timely feedback between computation and experiment to greatly accelerate the materials exploration and discovery. Moreover, such an approach will continually provide relevant structures and energy/property data to improve accuracy of the prediction of the ML model, which in turn will further reduce the workload of ab initio calculations while retaining the reliability of the predictions.

Our basic attempts to synthesize the predicted La_3CoPb via multiple synthetic techniques, produced mixtures of phases with, at best, highly ambiguous signals of the predicted low-energy La_3CoPb and $La_{18}Co_{28}Pb_3$ phases. These results suggest that either (i) the synthesis of these ternary La-Co-Pb compounds is very challenging, or (ii) these predicted compounds are not indeed stable. While ab initio calculations assisted by ML can greatly accelerate the exploration of crystal structures and their relative energetic stabilities at 0 K, successful synthesis is certainly not guaranteed. Gibbs free energies of the structures at high temperature is crucial to understand the thermodynamic stability of the compounds and provide more useful guidance for experimental synthesis. Our calculation results on Gibbs free energies for the five competing ternary phases show that synthesis of the predicted La_3CoPb and $La_{18}Co_{28}Pb_3$ phases would be very challenging since they are thermodynamically stable only in a narrow temperature range below 200 K. Kinetics of phase nucleation and growth, in addition to the relative thermodynamic stability of the competing phases, also need to be further addressed in order to provide more useful guidance to successful synthesis the novel materials.

Our study demonstrates the importance of the vibration contribution to future attempts to use ML models trained on zero-temperature DFT energies to predict the synthesizability of materials. Ideally, such models would be trained at finite temperature Gibbs free energy, but large computed databases of such energies are not available due to high computational cost of phonon calculations. Several attempts have been made to train ML models to predict vibrational energies[15,16], and our work demonstrate the importance of these efforts.

METHODS

The CGCNN model is built with convolutional neural networks on top of crystal graph consist of convolutional layers and pooling layers. Crystal structures are converted to crystal graphs with nodes representing atoms in the unit cell and edges representing atom connections. Nodes are characterized by feature vectors using nine atomic properties including group number, period number, electronegativity, covalent radius, valence electrons, first ionization energy, electron affinity, block (s, p, d, f), and atomic volume. Edges are characterized by neighboring bonds for each atom/node. Convolutional layers and hidden layers are built on top of these nodes, resulting in a graph with each node representing the local environment of each atom. After pooling, a vector representing the entire crystal is connected to hidden layers, followed by the output layer to provide the prediction. Dataset is selected from formation energies of 46,744 binary and ternary compounds from the MP database[12], and then randomly divided into training set (80%), validation set (10%) and test set (10%). In this study, we set the hyperparameters mostly to default values in the code provided in Ref. [9]. Batch size is set to 256 and total number of epochs to run is set to 200. Stochastic Gradient Descent is used as optimization algorithm. After 200 epochs of run, the best model is selected with the minimum MAE of the validation set, which is 0.04 eV/atom for our CGCNN model. This model is then used to predict the formation energies of our generated 357480 La-Co-Pb hypothetical structures, as shown in Fig. 1.

Our ab initio calculations were carried out using the projector augmented wave method[17] within density functional theory (DFT) as implemented in VASP code[18,19]. The exchange and correlation energy is treated by the generalized gradient approximation (GGA) and parameterized by the Perdew-Burke-Ernzerhof formula (PBE)[20]. A plane-wave basis set with a kinetic energy cutoff of 520 eV was used to expand the electronic wave functions, and the convergence criterion for the total energy was set to 10^{-5} eV. Monkhorst−Pack's sampling scheme[21] was adopted for Brillouin zone sampling with a k-point grid of $2\pi \times 0.033$ Å^{-1}, and the unit cell lattice vectors (both the unit cell shape and size) are fully relaxed together with the atomic coordinates until the force on each atom is <0.01 eV × Å^{-1}.

To characterize the energetic stability of the structures at different compositions, the formation energy with respect to the elemental ground-state bulk phases of the constituent elements (denoted as E_f) and the formation energy with respect to the known convex hull of the ternary system (denoted as E_d) are calculated.

For any given structure $La_mCo_nPb_p$, E_f is defined as:

$$E_f = \frac{E(La_mCo_nPb_p) - mE(La) - nE(Co) - pE(Pb)}{m+n+p} \quad (1)$$

where $E(La_mCo_nPb_p)$ is the total energy of the $La_mCo_nPb_p$ compound. E(La), E(Co) and E(Pb) are the per-atom energy of the ground state of La, Co and Pb crystals, respectively.

E_d of any given phase on the ternary convex hull can be calculated by comparing its formation energy with respect to the nearby three known phases on the convex hull. These known phases can be ternary, binary, or elementary crystalline phases, and the chemical compositions of the three phases are located at the vertexes of a triangle (called a Gibbs triangle) that encloses the composition of the phase for which E_d is calculated. The surface formed by all the Gibbs triangles is the convex hull. The formation energies of all known stable structures are therefore located on the surface of this convex cladding. Any structure having formation energy below this convex hull surface will be another stable structure and the convex hull

surface will then be re-constructed by including the formation energy of this structure. Therefore, E_d determines the thermodynamic stability of the given phase against decomposition into the nearby three known phases at $T = 0$ K.

We also evaluate the Gibbs free energy as a function of temperature for different low energy binary and ternary phases in the La-Co-Pb system by including the vibrational entropy and electronic entropy contributions. The vibrational entropy is calculated based on the vibrational DOS using the quasi-harmonic approximation (QHA) and the electronic entropy is calculated based on the electronic density of state and Fermi-Dirac distribution function from the DFT calculations.

Under zero pressure and at given temperature T and volume V, the Gibbs free energy G of a crystalline compound is given by[22]

$$G = E_0 + \frac{1}{2}\sum_i \hbar\omega_i - TS_{vib} - TS_{el} \quad (2)$$

where E_0 is the DFT total energy at $T = 0$K, $\frac{1}{2}\sum_i \hbar\omega_i$ is the zero point energy from phonon calculations, S_{vib} and S_{el} are the vibrational and electronic entropy respectively. The vibrational entropy S_{vib} can be calculated from phonon frequencies within the QHA[13,23], and the electronic entropy S_{el} is calculated based on the electronic DOS and Fermi-Dirac distribution function from the DFT calculations, i.e.,

$$S = -k_B \int n(E)[f \ln f + (1-f)\ln(1-f)]dE \quad (3)$$

where k_B is Boltzmann constant, $n(E)$ is the electronic DOS. The probability of occupation of each eigenstate is given by the Fermi function f:

$$f = \frac{1}{e^{(E-E_f)/k_BT} + 1} \quad (4)$$

To characterize the thermodynamic stability of the ternary structures at different compositions and different temperature, the formation Gibbs energies G_d of each ternary phase with respect to the nearby competing binary and ternary phases in the ternary convex hull at each T can be calculated in the same way as E_d calculation, but replacing the total energies of the competing phases by the corresponding Gibbs free energies G of the phases. Therefore, G_d determines the thermodynamic stability of the given phase against decomposition into the nearby three known phases at different T.

DATA AVAILABILITY

The data that support the findings of this paper, including both experimental and computational results, are available from the corresponding author, C-Z.W. (email: wangcz@ameslab.gov), upon reasonable request.

CODE AVAILABILITY

All codes used in this work are accessible through their websites. We use VASP 5.4.1 version in this work.

Received: 21 April 2022; Accepted: 10 December 2022;
Published online: 24 December 2022

REFERENCES

1. Canfield, P. C. New materials physics. *Rep. Prog. Phys.* **83**, 016501 (2019).
2. Gvozdetskyi, V. et al. Computationally Driven Discovery of a Family of Layered LiNiB Polymorphs. *Angew. Chem. Int. Ed.* **58**, 15855–15862 (2019).
3. Wang, R. et al. Theoretical search for possible Li–Ni–B crystal structures using an adaptive genetic algorithm. *J. Appl. Phys.* **127**, 094902 (2020).
4. Canfield, P. C. & Bud'Ko, S. L. FeAs-based superconductivity: a case study of the effects of transition metal doping on $BaFe_2As_2$. *Annu. Rev. Condens. Matter Phys.* **1**, 27–50 (2010).
5. Kaluarachchi, U. S., Bud'ko, S. L., Canfield, P. C. & Taufour, V. Tricritical wings and modulated magnetic phases in $LaCrGe_3$ under pressure. *Nat. Commun.* **8**, 1–6 (2017).
6. Taufour, V. et al. Ferromagnetic quantum critical point avoided by the appearance of another magnetic phase in $LaCrGe_3$ under pressure. *Phys. Rev. Lett.* **117**, 037207 (2016).
7. Pöttgen, R. & Rodewald, U. C. Rare earth-transition metal-plumbides. *Handb. Phys. Chem. rare earths* **38**, 55–103 (2008).
8. Massalski, T. B., Okamoto, H., Subramanian, P., Kacprzak, L. & Scott, W. W. *Binary alloy phase diagrams*. Vol. 1 (American society for metals Metals Park, OH, 1986).
9. Xie, T. & Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
10. Park, C. W. & Wolverton, C. Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. *Phys. Rev. Mater.* **4**, 063801 (2020).
11. Noh, J., Gu, G. H., Kim, S. & Jung, Y. Uncertainty-Quantified Hybrid Machine Learning/Density Functional Theory High Throughput Screening Method for Crystals. *J. Chem. Inf. Model.* **60**, 1996–2003 (2020).
12. Jain, A. et al. Commentary: The Materials Project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
13. Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).
14. Curtarolo, S. et al. AFLOW: An automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.* **58**, 218–226 (2012).
15. Legrain, F., Carrete, J., van Roekeghem, A., Curtarolo, S. & Mingo, N. How Chemical Composition Alone Can Predict Vibrational Free Energies and Entropies of Solids. *Chem. Mater.* **29**, 6220–6227 (2017).
16. Legrain, F. et al. Vibrational Properties of Metastable Polymorph Structures by Machine Learning. *J. Chem. Inf. Model.* **58**, 2460–2466 (2018).
17. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
18. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
19. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
20. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
21. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
22. Wang, C. Z., Ho, K. M., Shirk, M. D. & Molian, P. A. Laser-Induced Graphitization on a Diamond (111) Surface. *Phys. Rev. Lett.* **85**, 4092–4095 (2000).
23. Togo, A., Chaput, L., Tanaka, I. & Hug, G. First-principles phonon calculations of thermal expansion in Ti_3SiC_2, Ti_3AlC_2, and Ti_3GeC_2. *Phys. Rev. B* **81**, 174301 (2010).

ACKNOWLEDGEMENTS

Work at Ames Laboratory was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Science and Engineering Division including a grant of computer time at the National Energy Research Scientific Computing Centre (NERSC) in Berkeley, CA. Ames Laboratory is operated for the U.S. DOE by Iowa State University under Contract No. DE-AC02-07CH11358. Work at Guangdong University of Technology was supported by the Guangdong Basic and Applied Basic Research Foundation (Grant No. 2021A1515110328 & 2022A1515012174) and the Guangdong Natural Science Foundation of China (Grant No. 2019B1515120078). R. H. Wang and H. F. Dong also thank Center of Campus Network & Modern Educational Technology, Guangdong University of Technology, Guangdong, China for providing computational resources and technical support for this work. T.J.S was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, through Ames Laboratory under its Contract with Iowa State University (No. DE-AC02-07CH11358) and through the Center for Advancement of Topological Semimetals (CATS). T.J.S. was also supported by the Gordon and Betty Moore Foundation (Grant No. GBMF4411).

AUTHOR CONTRIBUTIONS

R.W. and W.X. contributed equally to this paper. R.W. and W.X. performed machine learning predictions, first-principles calculations as well as XRD simulations. X.F. contributed to the CGCNN ML studies. T.S. and P.C. synthesized the samples and analyzed the experimental PXRD patterns. C.-Z.W. P.C. and K.-M.H. conceptualized and planned the research and C.-Z.W. coordinated the research. All authors contributed to the discussions, analyses of the data, and writing of the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-022-00950-0.

Correspondence and requests for materials should be addressed to Cai-Zhuang Wang.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.