METALLURGY

Machine learning–enabled high-entropy alloy discovery

Ziyuan Rao[1], Po-Yen Tung[1,2], Ruiwen Xie[3], Ye Wei[1]*, Hongbin Zhang[3], Alberto Ferrari[4], T.P.C. Klaver[4], Fritz Körmann[1,4], Prithiv Thoudden Sukumar[1], Alisson Kwiatkowski da Silva[1], Yao Chen[1,5], Zhiming Li[1,6], Dirk Ponge[1], Jörg Neugebauer[1], Oliver Gutfleisch[1,3], Stefan Bauer[7], Dierk Raabe[1]*

High-entropy alloys are solid solutions of multiple principal elements that are capable of reaching composition and property regimes inaccessible for dilute materials. Discovering those with valuable properties, however, too often relies on serendipity, because thermodynamic alloy design rules alone often fail in high-dimensional composition spaces. We propose an active learning strategy to accelerate the design of high-entropy Invar alloys in a practically infinite compositional space based on very sparse data. Our approach works as a closed-loop, integrating machine learning with density-functional theory, thermodynamic calculations, and experiments. After processing and characterizing 17 new alloys out of millions of possible compositions, we identified two high-entropy Invar alloys with extremely low thermal expansion coefficients around 2×10^{-6} per degree kelvin at 300 kelvin. We believe this to be a suitable pathway for the fast and automated discovery of high-entropy alloys with optimal thermal, magnetic, and electrical properties.

Alloy design refers to a knowledge-guided approach to the development of high-performance materials. The strategy was established in the Bronze Age and has undergone further developments since that time. Alloy design is the basis for the development of different materials that enable technological progress. Several thousand metallic alloys have been developed so far that serve in engineering applications. The first essential alloy groups developed, such as bronze and steel, are all based on one main element that forms the matrix of the material. Over time, alloys with a higher number of alloying elements in larger fractions, such as austenitic stainless steels, have been developed. Today, with the development of high-entropy alloys (HEAs), we have reached a stage where multiple elements are used in similar fractions (*1*, *2*). Considering only the most used elements of the periodic table, this spans a composition space of at least 10^{50} alloy variants, a space so large that it cannot be managed by conventional alloy design methods (*3*). These conventional methods for designing alloys, which have been applied to small subspaces of the HEA composition realm, include calculation of phase diagrams (CALPHAD) and density-functional theory (DFT) (*4*–*6*). However, CALPHAD provides equilibrium-phase diagrams only, and DFT is computationally costly and cannot be readily applied to higher temperatures and disordered alloys (*5*, *7*). Likewise, combinatorial experiments (*8*) are very labor intensive and only cover the limited composition space of HEAs.

Because of these methodological limitations to finding materials with promising functional and mechanical features, we present a different approach to accelerating the discovery of HEAs. We based our approach on the use of machine learning (ML) techniques, with a focus on probabilistic models and artificial neural networks. Limited by the amount of available composition-property data, conventional ML approaches in alloy design have to predominantly rely on simulation data, often with only limited experimental validation (*9*, *10*). As the experimental microstructure database continues to expand, ML obtains higher accuracy in predicting the phase or microstructure of materials (*11*). However, the direct composition-property prediction is still elusive because of the comparably small databases and the human bias in feature selection. Recently, active learning has emerged as an alternative choice for functional materials discovery (*12*). Active learning is a subfield of ML in which surrogate models iteratively select unseen data points that are most informative to improve the predictive power of the models (*13*). In this approach, the next set of experiments is guided by the previous model trained based upon the results seen so far, yielding data points that will again be used iteratively for updating the model. Active learning has the potential to reduce the computational costs of alloy design and to both incorporate and guide experimental data and routines. However, active learning approaches to guiding the experimental discovery of materials have relied on simple surrogate models and Bayesian optimization methods, which are limited to low-dimensional data, thus showing property improvements only after many iterations (*14*, *15*).

To overcome these obstacles, we propose an active learning framework for the composition discovery of HEAs that is efficient for very sparse experimental datasets. The approach comprises ML-based techniques, DFT, mean-field thermodynamic calculations, and experiments. We focused on the design of high-entropy Invar alloys with a low thermal expansion coefficient (TEC) for several reasons: (i) a high demand exists for different types of Invar alloys to serve emerging markets for the transport of liquid hydrogen, ammonia, and natural gas; (ii) the mechanical properties of the original $Fe_{63.5}Ni_{36.5}$ (wt %) alloy for which Charles Edouard Guillaume received the 1920 physics Nobel Prize leave room for improvement; (iii) alternative Invar alloys (e.g., intermetallic, amorphous, or antiferromagnetic Invar compounds) come at forbiddingly high alloy costs and/or poor ductility (*16*, *17*); (iv) although a few HEAs have the potential to fill this gap (*18*–*20*), the lowest TEC ($\sim 10 \times 10^{-6}$ K^{-1}) of HEAs reported in the literature exceeds the value of the original $Fe_{63.5}Ni_{36.5}$ (wt %) alloy ($\sim 1.6 \times 10^{-6}$ K^{-1}) (*19*); and (v) our active learning framework mainly considers compositional information instead of the alloy manufacturing process, which makes the Invar effect an ideal target because these alloys are mostly determined by composition and less by processing (*6*, *19*) (see fig. S1 and table S1 for more background).

Results and discussion

Generative alloy design

The active learning framework includes three main steps: targeted composition generation, physics-informed screening, and experimental feedback (Fig. 1). Considering the large number of possible composition combinations of HEAs and the small experimental datasets (699 compositions; fig. S2), the challenge is to directly sample new compositions with the desired properties. Therefore, we developed an HEA generative alloy design (HEA-GAD) approach that is based on a generative model (GM) (*21*). First, the HEA-GAD uses GM, mathematical modeling, and sampling to perform a large-scale search of potential Invar alloys. GM learns an efficient and effective representation of the high-dimension data, which not only provides direct data visual representation, but also converts the search in high-dimensional design spaces to those of lower dimensionality (*22*). Different GMs are compared and analyzed on the basis of the evaluation metrics. The results show that the Wasserstein autoencoder (WAE) architecture performs better than other models with similar architectures (*21*) (figs. S3 and S4). The encoder takes

---

[1]Max-Planck-Institut für Eisenforschung GmbH, Düsseldorf, Germany. [2]Department of Earth Sciences, University of Cambridge, Cambridge, UK. [3]Institut für Materialwissenschaft, Technische Universität Darmstadt, Darmstadt, Germany. [4]Materials Science and Engineering, Delft University of Technology, Delft, Netherlands. [5]School of Civil Engineering, Southeast University, Nanjing, China. [6]School of Materials Science and Engineering, Central South University, Changsha, China. [7]KTH Royal Institute of Technology, Stockholm, Sweden.
***Corresponding author. Email: y.wei@mpie.de (Y.W.); d.raabe@mpie.de (D.R.)**

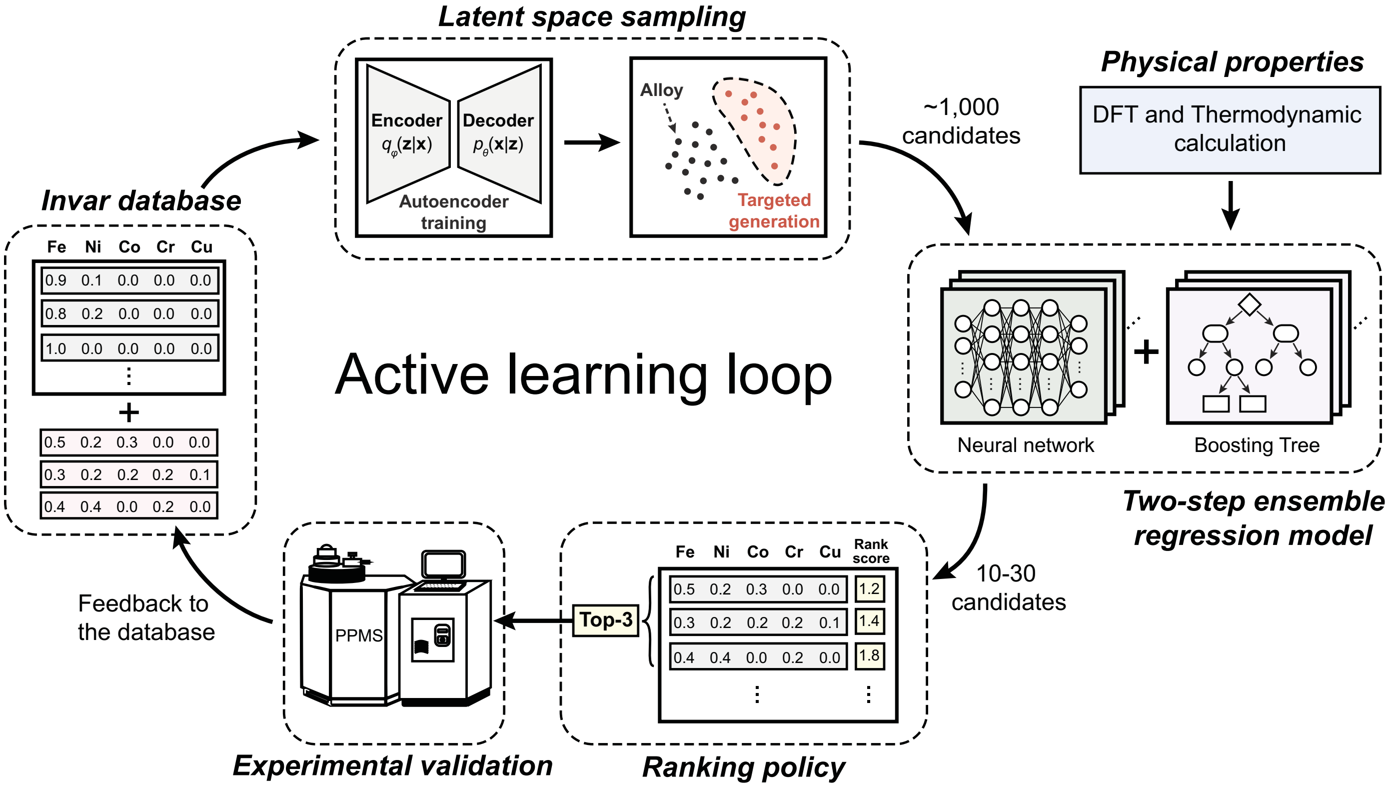

Fig. 1. Approach overview. We developed an active learning framework for the targeted composition design and discovery of HEAs, which combines ML models, DFT calculations, thermodynamic simulations, and experimental feedback. First, the promising candidates are generated under the HEA-GAD framework consisting of two primary steps: (i) an autoencoder for composition generation and (ii) stochastic sampling for composition selection. Second, the selected candidates from the HEA-GAD are further processed by the TERM framework, which includes two ensemble models composed of multilayer perceptrons and gradient-boosting decision trees. In the last step, the most promising compositions are selected by a ranking-based policy. The top three candidates are experimentally measured and fed back to the database. The iteration is repeated until the discovery of Invar alloys.

compositions of alloys as the input and learns to compress them down to low-dimensional representations, and the decoder can then act as a generator for producing alloy compositions given the learned continuous latent z representation. Although WAE is trained with only compositional information of alloys, it may implicitly include information on composition-related properties, which makes the latent space physically meaningful and informative. In our case, Invar alloys show extremely low TEC (hereafter used to refer to the TEC around room temperature unless otherwise specified) values, and the composition-TEC relation obeys specific physical laws. Subsequently, HEA-GAD uses the Gaussian mixture model (GMM) and Markov chain Monte Carlo (MCMC) sampling (*23*, *24*) to perform a large-scale search for the Invar compositions generated from WAE latent representation.

Two-stage ensemble regression

Next, we use the two-stage ensemble regression model (TERM) to further investigate the TEC of the HEA-GAD–generated alloy compositions. The first stage concerns composition-based regression models aiming at fast and large-scale composition inference. Then, the top ~1000 results with potentially low TEC from the HEA-GAD model are screened and enter the second-stage model, where DFT and thermodynamic calculations are included as part of the input, making it a physics-informed model (table S4). In the following section, we demonstrate that incorporating the physical inputs does increase the model accuracy. To increase the robustness of TERM without sacrificing the prediction accuracy, TERM taps into the advantages of the multilayer perceptron (*25*–*27*) and gradient-boosting decision tree approaches (*28*–*30*) by combining both into a single ensemble (*31*). Based on prediction and uncertainty, the exploration and exploitation strategy is used to adaptively guide the discovery of desirable compositions (*31*). Exploration prefers the composition with higher uncertainty (curiosity), whereas exploitation favors the composition with lower predicted TEC (perceived usefulness). Such a baseline strategy is premised on the model's ability to generalize beyond the known data, which is, however, often hampered by the highly nonlinear nature of the composition-property relation and sparsity of the available dataset. To overcome this issue, we designed a rank-order strategy that allows predictions to be rearranged and ranked in a specific order (*32*, *33*). This strategy is particularly advantageous when the underlying distribution of the data is largely unknown. The rank-based strategy ensures that the candidate selection is less affected by model inaccuracy and provides a systematic way to combine model prediction and uncertainty (*31*). Finally, the TEC values of the top three selected candidate materials are experimentally determined by the physical properties measurement system. These experimental results then augment the training database for the next active learning iteration.

Compositional latent space distribution

We produced a large benchmark dataset with 699 data points of Invar alloys mainly from former publications (fig. S2 and table S3) (*34*–*39*). Then, on the basis of the HEA-GAD-TERM framework proposed above, we performed six iterations and cast 18 alloys including 17 new alloys and one $Fe_{63.5}Ni_{36.5}$ (wt %) classic Invar alloy as a reference alloy. Because of the data imbalance (figs. S5 and S6), the discovery of FeNiCoCrCu HEAs is much more difficult than the discovery of FeNiCoCr HEAs. For this reason, we focused on the design of FeNiCoCr HEAs for the first three iterations and on FeNiCoCrCu HEAs for the last three iterations. We show the WAE latent space and GMM-modeled two-dimensional probability density of the first iteration in Fig. 2, A and B. The latent space yields certain islands that indicate the compositional differences. For example, the HEAs tend to stay in the middle, whereas the binary and ternary alloys tend to stay in the edges of the latent space. Also, a smooth transition among the Fe–Ni, Ni–Co binary alloys and the Fe–Ni–Co ternary alloys can be observed. FeNiCoCrCu forms a single island, indicating that features of compositions with nonzero Cu content are indeed captured by HEA-GAD. The new FeCoNiCr HEAs candidates are cross-marked, whereas the best-ranked HEAs are illustrated with white dots in Fig. 2, A and B. We also show the last iteration result of FeCoNiCrCu HEAs discovered by HEA-GAD-TERM in Fig. 2, C and D (in red). The entire latent space is slightly rotated because of the addition of new data into the training dataset from previous iterations. The augmented dataset also leads to a modified GMM-modeled probability density shown in Fig. 2D, in which the left Gaussian ellipse extends more to the left region compared with Fig. 2B. Such phenomena suggest that the HEA-GAD-TERM framework is interpretable and sensitive to the dataset.

Physics-descriptor–informed model

So far, the Masumoto empirical rule (*34*, *35*) has played an important role in the discovery of several Invar alloys. As exemplified in Fig. 3D for the $Fe_{60}Ni_{35}Co_5$ (wt %) Super Invar alloy, according to this rule, the TEC is related to the ratio ω_s/T_c (magnetostriction/Curie temperature): Because of the Invar effect, Invar alloys have lower TEC in the ferromagnetic state (below Curie temperature Q) than in the paramagnetic state (above Q). The TEC in the ferromagnetic state can thus be estimated as

$$\mathrm{TEC} = \frac{QS}{RS} = \frac{QA - SA}{T_c} = \frac{QA}{T_c} - \frac{SA}{T_c} \approx \tan\delta - \frac{\omega_s}{T_c}$$

We demonstrated the correlation between ω_s/T_c and the experimental TEC with DFT and CALPHAD for FeCoNi alloys. The alloys from our experimental dataset were slowly cooled from high-temperature homogenization, so an equilibrium temperature to calculate the phase fractions in our samples cannot be determined unambiguously. We nevertheless calculated ω_s/T_c for the annealing temperatures T_{ann} = 1273 K, 1073 K, and 873 K (Fig. 3, A to C) and observed a good correlation with the experimentally observed TEC values, especially for the values taken at T_{ann} = 873 K. ω_s and T_c are thus useful descriptors that can be exploited to increase the accuracy of TERM. We show the comparison of the model training history with and without the use of the descriptor T_c (Fig. 3E). This history reflects the performance evolution with time (epoch) as more data were fed to the model. The final testing error was notably reduced from 0.19 to 0.14 upon inclusion of DFT and CALPHAD data, a piece of strong evidence that the physics-descriptor–informed model can achieve better accuracy than that based only on compositions.

Learning curve and thermal expansion behavior

We show the measured and predicted TEC values of the 17 alloys experimentally measured

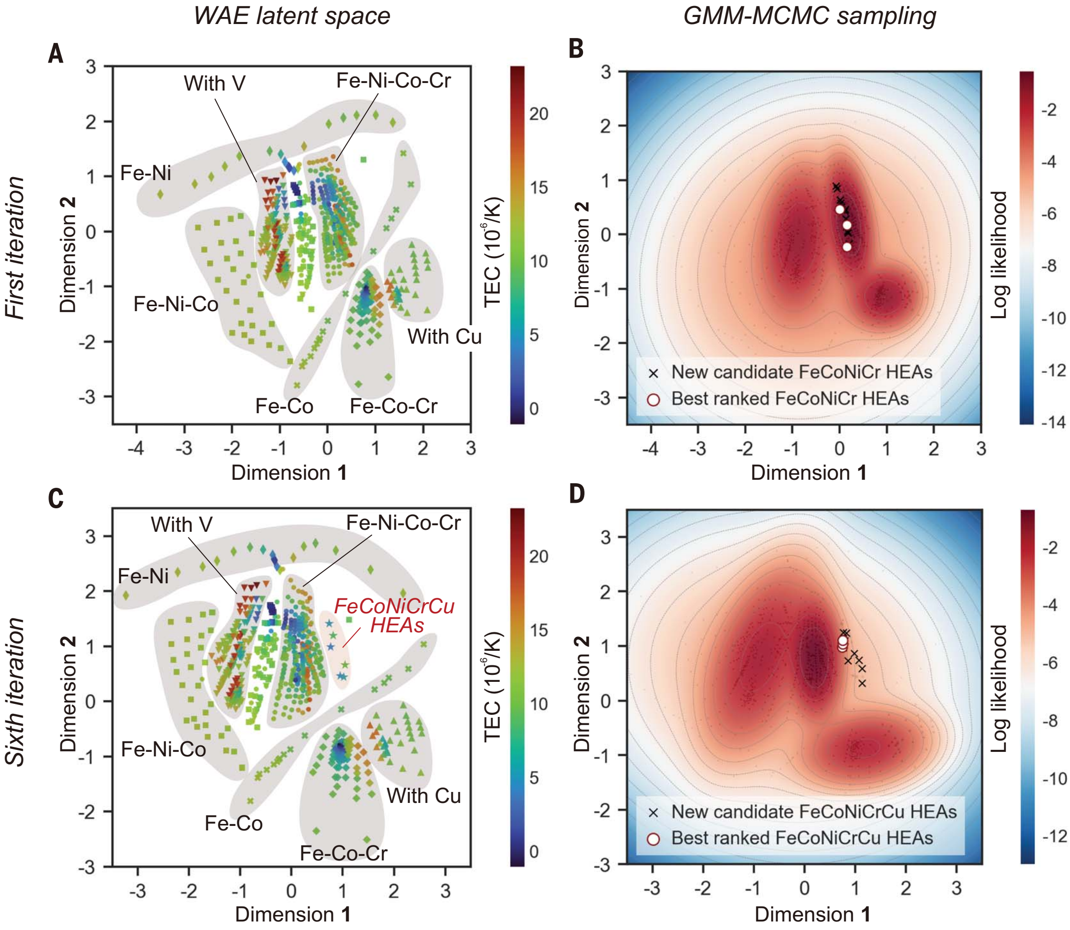

Fig. 2. First and last (sixth) iterations of the HEA-GAD generation. (**A** and **B**) WAE latent space and GMM-modeled density of the first iteration. (**C** and **D**) WAE latent space and GMM-modeled density of the last iteration. The WAE latent space distribution of the different compositions is marked with different symbols. The colors of the data points in the latent space denote their corresponding TEC. The GMM shows the probabilistic density in the latent space. The new candidates proposed by the first stage of the TERM are marked by crosses, and the new compositions proposed by the second stage of the TERM are marked by circles. The learned latent spaces are informative of the TEC of the HEAs.

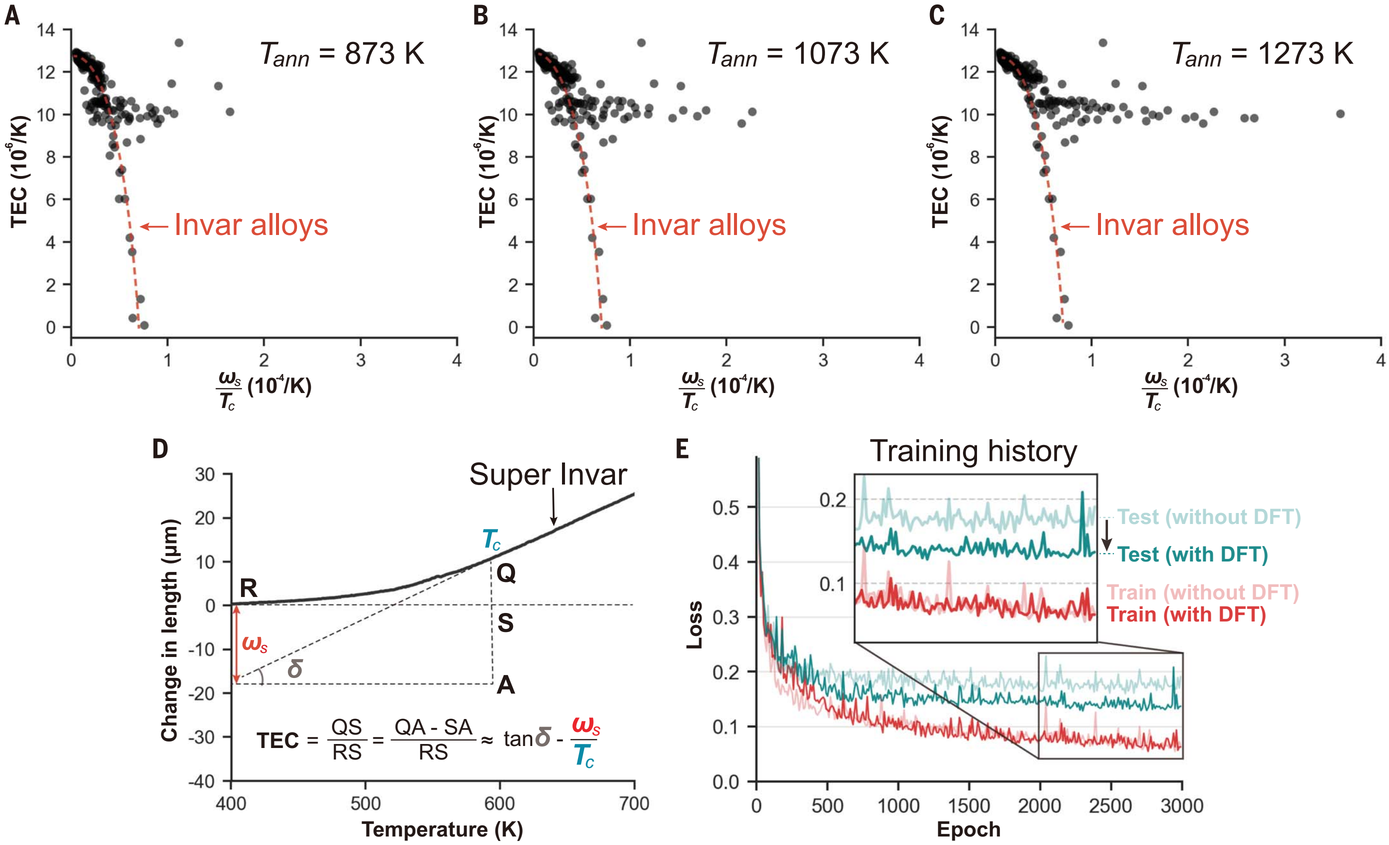

Fig. 3. Importance of the physics-informed descriptors. (**A** to **C**) Correlation between the proposed descriptor ω_s/T_c and the experimental TEC. (**D**) Schematic model of the Masumoto empirical rule for discovering Invar alloys. (**E**) Comparison of training and testing history with and without use of the descriptor ω_s/T_c. Both the final training and testing errors decrease after considering the physics-informed descriptors; for example, the testing error decreases from 19 to 14%.

Table 1. Compositions and TEC of the HEAs designed in this work.*

Alloys	Iteration	Fe (wt %)	Ni (wt %)	Co (wt %)	Cr (wt %)	Cu (wt %)	Predicted TEC ($\times10^{-6}$/K)	Predicted uncertainty ($\times10^{-6}$/K)	Experimental TEC ($\times10^{-6}$/K)
A1	1st	55.2	23.9	16.7	4.2	0	3.41	1.29	7.54
A2	1st	49.2	17.2	27.1	6.5	0	3.13	0.75	10.52
A3	1st	41.8	9.4	40.9	8	0	4.39	0.79	1.41
A4	2nd	52.5	22	20.8	4.7	0	3.91	0.53	7.97
A5	2nd	44	13.8	34.6	7.6	0	4.20	0.96	3.24
A7	3rd	42.4	12.6	37.7	7.3	0	4.58	1.40	4.09
A8	3rd	44.2	15.8	33.2	6.8	0	5.88	2.17	4.83
A9	3rd	54.1	22.8	17.2	5.9	0	5.16	1.43	2.02
B1	4th	40	6.9	39.5	7.9	5.7	7.57	1.45	5.84
B2	4th	48.8	17.8	22.2	6.2	5	5.48	1.01	4.38
B3	4th	57	16.4	14.6	5.1	6.9	4.43	1.33	8.56
B4	5th	40.6	6.9	38.3	9.2	5	8.41	1.70	4.94
B5	5th	57.7	22.9	8.3	5.2	5.9	4.50	1.00	5.31
B6	5th	51.6	6.8	27.5	7.8	6.3	9.32	3.49	9.68
B7	6th	48.3	17.8	20.9	7.9	5.1	5.49	0.92	5.60
B8	6th	50	18.3	18.3	8	5.4	5.65	1.16	5.13
B9	6th	50.7	19.9	15.8	7.9	5.7	5.56	1.05	6.29

*The original $Fe_{63.5}Ni_{36.5}$ Invar (A6) is a reference alloy and is not listed here.

in the six iterations in Table 1. A3 and A9 HEAs with four principal elements show extremely low TECs that are comparable to the classical $Fe_{63.5}Ni_{36.5}$ (wt %) binary Invar alloy. B2 and B4 HEAs with five principal elements show TECs that are comparable to the commercially used $Fe_{54}Co_{17}Ni_{29}$ (wt %) ternary Kovar alloys. In addition, a tabular comparison between HEA-GAD-TERM and trial and error can be found in table S2, where our method shows a fivefold higher discovery rate than that achieved by the trial and error approach alone.

We illustrate the alloy discovery process in two scenarios (Fig. 4, A and B). In the ideal case, the composition-TEC curve is simple and

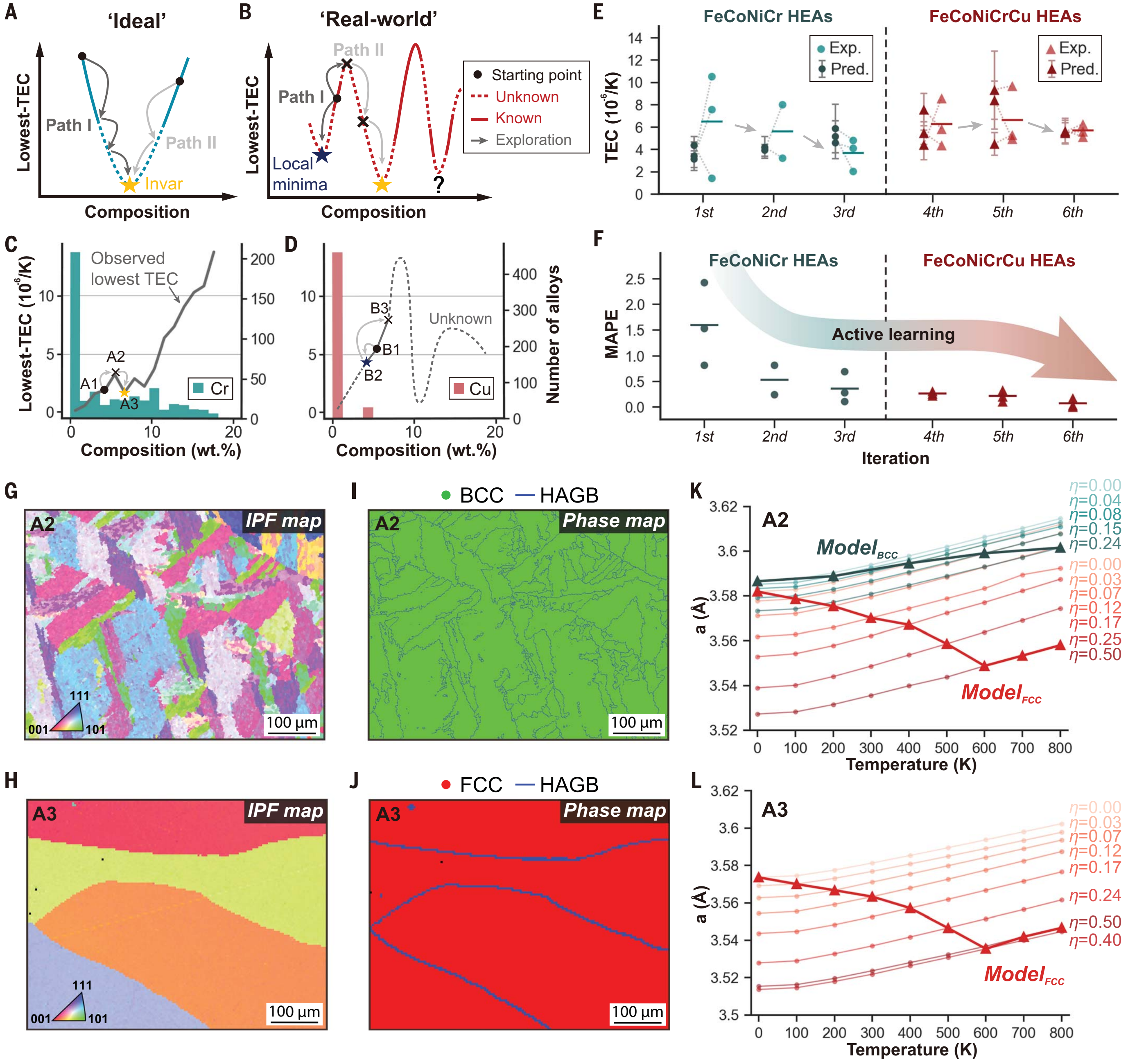

Fig. 4. Analysis of the results after six iterations in the active learning loop. (**A** and **B**) Representation of the alloy discovery process in the ideal scenario and the real world. (**C** and **D**) Cr and Cu distribution histogram. The Cr histogram has various concentrations (from 0 to 20%). By contrast, the vast majority (>95%) of the compositions have zero Cu concentration. The lowest known TEC as a change of composition is plotted as a solid line, and the unknowns are shown as a dashed line. Gray arrows illustrate the discovery paths of HEA-GAD-TERM. (**E**) Experimental and predicted TEC of the FeNiCoCr and FeNiCoCrCu HEAs. (**F**) MAPE of active learning. The dots represent the MAPE between experiment and predictions. Rapid decrease of the MAPE is akin to a natural learning process. (**G** and **I**) Electron backscatter diffraction (EBSD) phase and boundary maps of the A2 alloy. (**H** and **J**) EBSD phase and boundary maps of the A3 alloy. (**K** and **L**) Change of lattice constants with temperature in the A2 $[(Fe^{\uparrow}_{1-\eta}Fe^{\downarrow}_{\eta})_{50.1}(Ni^{\uparrow}_{1-\eta}Ni^{\downarrow}_{\eta})_{16.7}(Co^{\uparrow}_{1-\eta}Co^{\downarrow}_{\eta})_{26.1}(Cr^{\downarrow}_{1-\eta}Cr^{\uparrow}_{\eta})_{7.1}]$ and A3 $[(Fe^{\uparrow}_{1-\eta}Fe^{\downarrow}_{\eta})_{42.7}(Ni^{\uparrow}_{1-\eta}Ni^{\downarrow}_{\eta})_{9.1}(Co^{\uparrow}_{1-\eta}Co^{\downarrow}_{\eta})_{39.5}(Cr^{\downarrow}_{1-\eta}Cr^{\uparrow}_{\eta})_{8.7}]$ alloys for different values of η, where η denotes the pseudo-alloy concentration ($0 \leq \eta \leq 0.50$).

convex, which means that this specific relation is readily learned and "never forgotten." Even with a small dataset present, the global maxima can be easily found regardless of their initial starting points: Both path 1 and path 2 can lead to the Invar point. However, in the reality, the lowest TEC curve is highly nonlinear because of the complex underlying composition-property relations, and the composition landscape remains largely unknown. Both experts with appropriate domain knowledge and algorithms will have to explore the unknown territory and accumulate knowledge about the system by making mistakes. Furthermore, the composition axis is multidimensional and therefore the design space is huge. Therefore, the chosen paths, available data, and starting points will notably influence the final results. Path 1 may lead to local minima, whereas path 2 is rather difficult initially, and multiple high TEC non-Invar HEAs can be discovered before the eventual Invar discovery.

We provide the concentration histogram of Cr and Cu in the current dataset in Fig. 4, C and D. We also plotted the observed lowest TEC curve to illustrate the discovery path in two HEAs. The GAD-TERM framework shows its high efficiency by quickly identifying the Invar points in the first iteration (A3 and B2). However, the algorithm is designed for exploration. The algorithm inevitably discovers some non-Invar alloys along the path (e.g., A4 and A8, denoted by gray arrows in Fig. 4, C and D). As mentioned before, the discovery of FeNiCoCr HEAs and FeNiCoCrCu HEAs are different tasks because of the different data distribution. The distribution of Cu in the alloys is extremely imbalanced (Fig. 4D); that is, by far most of the alloys in the dataset do not contain Cu at all and only a few alloys have 5% Cu. Such distributional difference likely accounts for the substantially different learning behavior observed (Fig. 4, E and F).

We show the measured and predicted TEC values for FeNiCoCr and FeNiCoCrCu HEAs in Fig. 4E and the mean absolute percentage error (MAPE) between experiments and predictions versus experimental iteration in Fig. 4F, with each exploitation and exploration step marked by arrows. For FeNiCoCr HEAs, the average experimental TEC value gradually decreases: 6.49×10^{-6} per degree kelvin (/K) in the first, 5.61×10^{-6}/K in the second, and 3.65×10^{-6}/K in the third iteration (Table 1). Exploration and exploitation take place alternately, akin to a natural learning process, and such a plot represents the "learning curve" of the HEA-GAD-TERM model. The learning curve indicates a progressive trend as the MAPE error decreases notably (from 1.5 to 0.2). Because of the exploration step, the model predictions deviate considerably from their experimental counterparts in the first iteration. Alloy A3 (Table 1) has the highest predicted TEC value ($4.39 \pm 0.79 \times 10^{-6}$/K), but the experimental TEC value shows exactly the opposite, namely, the lowest measured TEC value (1.41×10^{-6}/K). In the second and third iterations, the standard deviation of the experimental TEC values declines substantially (3.34×10^{-6}/K and 1.46×10^{-6}/K, respectively). This demonstrates excellent exploration progress in which HEA-GAD-TERM converges quickly and can predict TEC with high accuracy after only three iterations. Conversely, FeNiCoCrCu shows a different learning behavior. The discovery path shows no significant improvements, from experimentally measured 6.26×10^{-6}/K in the first iteration, to 6.64×10^{-6}/K in the second, and 5.67×10^{-6}/K in the third (for more numerical details, see Table 1). We can attribute this trend to the lack of Cu-containing FeNiCoCrCu data (only three data points are available at the beginning; Fig. 4D). Despite this shortcoming, the experimental mean deviation narrows down, from 33.9% for the first iteration to 10.2% in the last iteration, indicating a gradually improved model accuracy.

To reveal the physical origin behind the properties, we show experimental and DFT analyses of the A2 and A3 alloys (TEC = 10.52×10^{-6}/K and 1.41×10^{-6}/K, respectively, in Fig. 4, G to L). It can be seen in Fig. 4, G to J, that A2 and A3 alloys have a single-phase bcc and fcc structure, respectively. The partial disordered local moment (PDLM) model within the coherent potential approximation simulations (*40*) reveals that the Invar effect is qualitatively related to such volume reduction at finite-temperature PDLM phase compared with the 0 K ferromagnetic ground state (*41*). In contrast to the fcc A3 alloy, the bcc A2 alloy, with a higher T_c around 950 K, exhibits a slight upward trend of the lattice parameter *a*. Using DFT simulations, we also validate that if the A2 alloy can be stabilized in its fcc phase state, then an Invar effect can be realized as well [Fig. 4, K and L, red dash-dot line; for simulation details, see (*31*)]. The TEC value is also affected by the occurrence of phase transformations in some HEAs (*18*, *20*). Our measurements show that the low TEC values of our A3 alloys are not caused by any phase transformation.

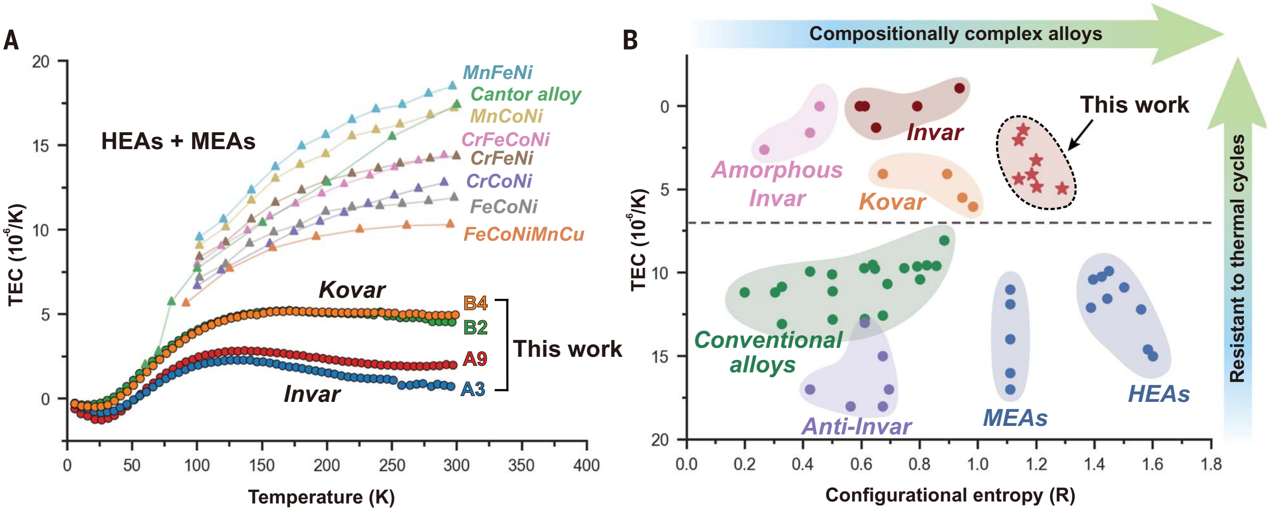

Fig. 5. Summary of the properties of the ML-designed HEAs. (**A**) TEC of the ML-designed HEAs as a function of the change in temperature. As a comparison, we plotted the thermal expansion curve of the HEAs and MEAs. A3 and A9 FeNiCoCr HEAs show extremely low TECs around 2×10^{-6}/K at 300 K, which can be used as Invar alloys. B2 and B4 FeNiCoCrCu HEAs show low TECs around 5×10^{-6}/K at 300 K, which qualifies them as Kovar alloys. (**B**) Configurational entropy plotted against the TEC values for various known alloys and alloys discovered in this work. ML enables this approach to efficiently discover new alloys with excellent properties (high resistance to thermal cycles) in an infinite phase spectrum (compositionally complex alloys).

We show the TEC as a function of temperature for the two Invar alloys (TEC $\approx 2 \times 10^{-6}$/K) and two Kovar alloys (TEC $\approx 5 \times 10^{-6}$/K) that we developed in Fig. 5A, compared with HEAs and medium-entropy alloys (MEAs) (*19*, *42*). The new alloys show abnormally low TEC values compared with the HEAs, MEAs, and conventional alloys previously reported (Fig. 5B) (*43*–*45*). Most Invar alloys show a low TEC but also low configurational entropy. The Invar alloys developed in this work offer a good combination of low TEC and high configurational entropy. This indicates the high potential of the HEA concept for the design of Invar alloys, which, beyond their beneficial thermal expansion response, also offer high strength, ductility, and corrosion resistance.

Conclusions

Understanding the underlying physics behind composition-property relations is the key mission in alloy design, a task particularly challenging in the case of compositionally complex materials. In principle, HEAs with interesting features can hide in practically infinite and vastly unexplored composition space, a scenario that puts targeted alloy design to its hardest test. We have therefore developed a widely applicable active learning framework that combines a generative model, regression ensemble, physics-driven learning, and experiments for the compositional design of HEAs. Our method demonstrates its proficiency in designing high-entropy Invar alloys using very sparse experimental data. The entire workflow required only a few months, in contrast to the conventional alloy design approach, which requires years and many more experiments. We expect that more than one property can be optimized simultaneously using the GAD-TERM framework in the compositional spectrum of HEAs.

REFERENCES AND NOTES

1. J. W. Yeh *et al.*, *Adv. Eng. Mater.* **6**, 299–303 (2004).
2. B. Cantor, I. Chang, P. Knight, A. Vincent, *Mater. Sci. Eng. A* **375–377**, 213–218 (2004).
3. E. P. George, D. Raabe, R. O. Ritchie, *Nat. Rev. Mater.* **4**, 515–534 (2019).
4. H. Mao, H.-L. Chen, Q. Chen, *J. Phase Equilibria Diffus.* **38**, 353–368 (2017).
5. F. Körmann *et al.*, *Appl. Phys. Lett.* **107**, 142404 (2015).
6. Z. Rao *et al.*, *Intermetallics* **111**, 106520 (2019).
7. S. Huang, E. Holmström, O. Eriksson, L. Vitos, *Intermetallics* **95**, 80–84 (2018).
8. Z. Rao, H. Springer, D. Ponge, Z. Li, *Materialia (Oxf.)* **21**, 101326 (2022).
9. B. O. Mukhamedov, K. V. Karavaev, I. A. Abrikosov, *Phys. Rev. Mater.* **5**, 104407 (2021).
10. J. Schmidt, L. Pettersson, C. Verdozzi, S. Botti, M. A. L. Marques, *Sci. Adv.* **7**, eabi7948 (2021).
11. Z. Pei *et al.*, *Adv. Sci. (Weinh.)* **8**, e2101207 (2021).
12. P. V. Balachandran, B. Kowalski, A. Sehirlioglu, T. Lookman, *Nat. Commun.* **9**, 1668 (2018).
13. J. Gubernatis, T. Lookman, *Phys. Rev. Mater.* **2**, 120301 (2018).
14. D. Xue *et al.*, *Nat. Commun.* **7**, 11241 (2016).
15. R. Yuan *et al.*, *Adv. Mater.* **30**, 1702884 (2018).
16. E. Wassermann, M. Acet, in *Magnetism and Structure in Functional Materials*, A. Planes, L. Mañosa, A. Saxena, Eds. (Springer, 2005), pp. 177–197.
17. M. Shiga, *Curr. Opin. Solid State Mater. Sci.* **1**, 340–348 (1996).
18. C. L. Lin *et al.*, *Appl. Phys. Lett.* **119**, 171902 (2021).
19. Z. Rao *et al.*, *Phys. Rev. Mater.* **5**, 044406 (2021).
20. C.-L. Lin *et al.*, *Mater. Chem. Phys.* **271**, 124907 (2021).
21. I. Tolstikhin, O. Bousquet, S. Gelly, B. Schoelkopf, Wasserstein auto-encoders. arXiv:1711.01558 [stat.ML] (2017).
22. R. Gómez-Bombarelli *et al.*, *ACS Cent. Sci.* **4**, 268–276 (2018).
23. A. Christophe, N. de Freitas, A. Doucet, M. I. Jordan, *Mach. Learn.* **50**, 5–43 (2003).
24. W. K. Hastings, *Biometrika* **57**, 97–109 (1970).
25. K. Hornik, M. Stinchcombe, H. White, *Neural Netw.* **2**, 359–366 (1989).
26. R. Kohavi, D. H. Wolpert, *ICML* **96**, 275–283 (1996).
27. U. Von Luxburg, B. Schölkopf, in *Handbook of the History of Logic*, D. M. Gabbay, S. Hartmann, J. Woods, Eds. (Elsevier, 2011), vol. 10, pp. 651–706.
28. G. Ke *et al.*, in *Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, December 2017*, U. von Luxburg, I. Guyon, S. Bengio, H. Wallach, R. Fergus, Eds. (NIPS, 2017), p. 3149–3157.
29. J. H. Friedman, *Comput. Stat. Data Anal.* **38**, 367–378 (2002).
30. A. H. Li, J. Bradic, *J. Am. Stat. Assoc.* **113**, 660–674 (2018).
31. Materials and methods are available as supplementary materials.
32. F. Wilcoxon, in *Breakthroughs in Statistics*, S. Kotz, N. L. Johnson, Eds. (Springer, 1992), pp. 196–202.
33. R. Lowry, "Concepts and applications of inferential statistics" (DOER, 2014); http://doer.col.org/handle//123456789/4853.
34. H. Masumoto, *Science Reports of the Tohoku Imperial University* **20**, 101–123 (1931).
35. H. Masumoto, *Science Reports of the Tohoku Imperial University* **23**, 265–275 (1934).
36. M. Hakaru, S. Hideo, K. Tatsuo, *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **6**, 529–538 (1954).
37. M. Hakaru, S. Hideo, S. Yutaka, *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **7**, 533–540 (1955).
38. M. Hakaru, S. Hideo, G. Kimiyosi, *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **9**, 159–169 (1957).
39. M. Hakaru, S. Hideo, S. Yutaka, *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **9**, 170–175 (1957).
40. I. A. Abrikosov *et al.*, *Phys. Rev. B Condens. Matter Mater. Phys.* **76**, 014434 (2007).
41. A. V. Ruban, *Phys. Rev. B* **95**, 174432 (2017).
42. G. Laplanche *et al.*, *J. Alloys Compd.* **746**, 244–255 (2018).
43. T. Schneider, M. Acet, B. Rellinghaus, E. F. Wassermann, W. Pepperhoff, *Phys. Rev. B Condens. Matter* **51**, 8917–8921 (1995).
44. C. Chanmuang, M. Naksata, T. Chairuangsri, H. Jain, C. E. Lyman, *Mater. Sci. Eng. A* **474**, 218–224 (2008).
45. K. Fukamichi, M. Kikuchi, S. Arakawa, T. Masumoto, *Solid State Commun.* **23**, 955–958 (1977).
46. Code for: Z. Rao *et al.*, Machine learning-enabled high-entropy alloy discovery, GitHub (2022); https://github.com/ziyuanrao11/Machine-learning-enabled-high-entropy-alloy-discovery.
47. Data for: Z. Rao *et al.*, Machine learning-enabled high-entropy alloy discovery, Zenodo (2022); https://doi.org/10.5281/zenodo.7019194.

ACKNOWLEDGMENTS

We thank M. Acet from University of Duisburg-Essen and M. Nellessen, M. Adamek, and F. Schlüter from Max-Planck-Institut für Eisenforschung GmbH. The staff of TU Darmstadt is gratefully acknowledged for providing computational resources with the Lichtenberg high-performance computer for the exact muffin-tin orbital (EMTO) calculations in the present work. **Funding:** Z.R., R.X., O.G., and H.Z. were supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation project ID 405553726-TRR 270). Y.W. was supported by BiGmax, the Max Planck Society's Research Network on Big-Data-Driven Materials Science, and the ERC-CoG-SHINE-771602. P.T. was supported by the Electron and X-Ray Microscopy Community for Structural and Chemical Imaging Techniques for Earth materials (EXCITE grant no. G106564) and the International Max Planck Research School for Interface Controlled Materials for Energy Conservation (IMPRS-SurMat). **Author contributions:** Z.R., Y.W., and D.R. conceived the study. Y.W. and Z.R. designed the active learning framework. Z.R., Y.W., P.T., and S.B. developed the algorithm and analyzed the results. Z.R. performed the experiments. R.X., H.Z., A.F., P.K., and F.K. performed the DFT calculations. P.T.S., A.K.S., and Z.R. performed the thermodynamic calculations. Z.R., Y.W., and P.T. wrote the main parts of the manuscript. P.T. produced the final figures. All authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The training dataset is curated from the previous publications (*34*–*39*) and can be found in (*46*). The necessary data produced by our model in this work can be found in the supplementary materials. The original code used to perform this work is available on GitHub (*46*). We also provide a simplified version of the code integrated into a single Jupyter Notebook on GitHub (*46*), which is easier to perform and understand. In addition, data are provided in the Zenodo repository (*47*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo4940
Materials and Methods
Figs. S1 to S13
Tables S1 to S16
References (*48*–*63*)

Submitted 13 February 2022; resubmitted 30 June 2022
Accepted 30 August 2022
10.1126/science.abo4940