Article:
Sheriff, K., Xiao, D.Z., Cao, Y. et al. (2026) Machine learning potentials for modeling alloys across compositions. Science Advances, 12 (25). eaea9951. ISSN: 2375-2548

https://doi.org/10.1126/sciadv.aea9951

MATERIALS SCIENCE

Machine learning potentials for modeling alloys across compositions

Killian Sheriff[1], Daniel Z. Xiao[1], Yifan Cao[1], Lewis R. Owen[2], Rodrigo Freitas[1]*

Materials properties depend strongly on chemical composition, i.e., the relative amounts of each chemical element. Changes in composition lead to entirely different chemical arrangements, which vary in complexity from perfectly ordered (i.e., stoichiometric compounds) to completely disordered (i.e., solid solutions). Accurately capturing this range of chemical arrangements remains a major challenge, limiting the predictive accuracy of machine learning potentials (MLPs) in materials modeling. Here, we combine information theory and machine learning to optimize the sampling of chemical motifs and design MLPs that effectively capture the behavior of metallic alloys across their entire compositional and structural landscape. The effectiveness of this approach is demonstrated by predicting the compositional dependence of various materials properties—including stacking-fault energies, short-range order, heat capacities, and phase diagrams—for the AuPt and CuAu binary alloys, the ternary CrCoNi, and the TiTaVW high-entropy alloy. Extensive comparison against experimental data demonstrates the robustness of this approach in enabling materials modeling with high physical fidelity.

INTRODUCTION

The properties of materials are heavily influenced by their chemical composition, i.e., the relative amounts of each constituent chemical element. Changes in composition lead to notable variations in the chemical arrangements of atoms that drive macroscopic behavior. For example, in stoichiometric compounds (Fig. 1A), the elements combine in fixed and precise ratios and are organized in a periodic and highly ordered phase with little chemical complexity. In contrast, solid solutions (Fig. 1B) are phases in which chemical elements are distributed in a nearly disordered manner, allowing greater flexibility in composition. These differences in atomic arrangements often translate into distinct macroscopic properties. For example, in the CuAu alloy, the ordered intermetallic phase Cu_3Au is brittle, whereas the disordered CuAu solid solution is ductile (*1*).

Although machine learning potentials (MLPs) have demonstrated remarkable predictive accuracy for stoichiometric compounds (*2*–*7*), extending this success to solid solutions has proven challenging (*8*–*11*). Addressing this limitation is important because solid solutions are prevalent across the materials design space. For example, they are the dominant phase over wide compositional ranges in the phase diagrams of metallic alloys (*12*). Notably, much of the widespread interest in high-entropy materials stems from the chemical complexity of solid solutions and the flexibility they enable in designing mechanical and functional properties (*13*–*15*).

The main challenge impeding the development of high-accuracy MLPs for solid solutions is the large variability in the chemical arrangement of atoms in this phase (*10*, *16*), which is immensely amplified when changes in chemical composition are taken into account. Generating an MLP training set that faithfully represents this chemical complexity is notoriously difficult (*9*). The current state-of-the-art strategy (*8*, *17*) uses atomistic configurations obtained from density functional theory–based Monte Carlo (DFT-MC) trajectories. Although physically grounded, this approach is prohibitively expensive—often requiring more than 100,000 CPU hours to create a training set for a single composition (*8*). Furthermore, this approach is limited to equilibrium states, which is a critical drawback because MLPs trained on equilibrium datasets generalize poorly to out-of-distribution non-equilibrium configurations commonly encountered during materials synthesis and processing (*2*, *18*–*20*). Emerging approaches such as active learning (*21*–*24*), although efficient for optimizing first-principles datasets, often lack objective functions capable of capturing chemical complexity, relying instead on system representations that have been shown to obscure the relevant variations in local chemical environments (*10*).

Here, we demonstrate the design of MLPs that accurately model metallic alloys across their compositional and structural landscape. This is achieved by combining information theory with geometric descriptors to construct training datasets optimized to capture chemical complexity through motif-based sampling (MBS). Extensive comparisons against experimental data show that this approach produces high-fidelity models of complex material systems—ranging from simple binary systems to high-entropy alloys.

RESULTS

Performance of universal ML potentials on solid solutions

Universal machine learning potentials (uMLPs) are a class of models trained on broad datasets encompassing a wide range of chemistries, compositions, and crystal structures. These models aim to achieve robust transferability across diverse material systems. However, this promise of universality often breaks down in chemically disordered phases. For example, consider Fig. 2A, which shows the accuracy of the MatterSim (*3*) uMLP across the composition space of the CrCoNi solid solution—a prototypical high-entropy alloy (see the Materials and Methods section "Training and testing datasets" for dataset details). The mean absolute error (MAE; $\delta_{\mathrm{MAE}} = |E - E_{\mathrm{DFT}}|$) between predicted and reference DFT energies reaches up to 287 meV/atom, with as much as 603% variation across compositions (i.e., two orders of magnitude). This compositional sensitivity implies that the reliability of predicted physical properties can vary markedly with composition.

This behavior is not unique to MatterSim. Other state-of-the-art uMLPs—namely, Orb (*2*) and MACE (*25*)—exhibit similar mean MAE

[1]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]School of Chemical, Materials and Biological Engineering, University of Sheffield, Sheffield, UK.
*Corresponding author. Email: rodrigof@mit.edu

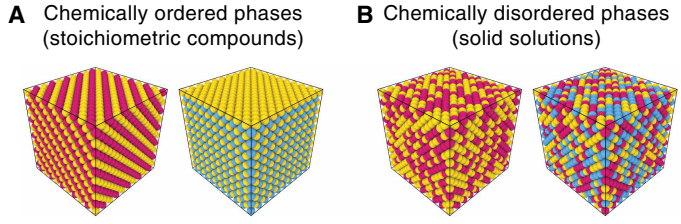

Fig. 1. Chemical complexity of phases in metallic alloys. (**A**) Stoichiometric compounds exhibit ordered atomic arrangements with fixed elemental ratios. (**B**) Solid solutions are chemically disordered phases with flexible composition and greater chemical complexity.

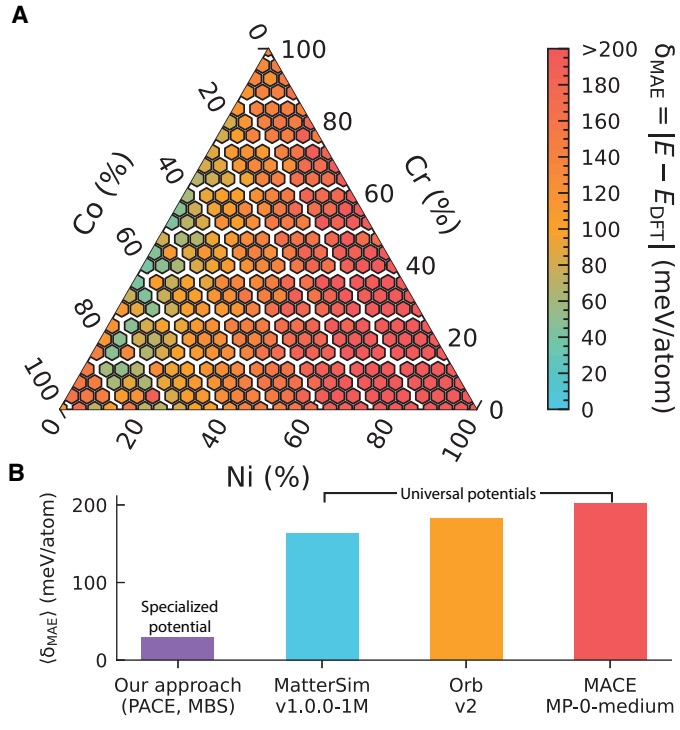

Fig. 2. Error of uMLPs on solid solutions. (**A**) Energy MAE of MatterSim-v1.0.0- 1 M (*3*) across the composition space of the CrCoNi random fcc solid solution. (**B**) Comparison of error averaged across the compositional space for our MBS approach (a specialized potential) and various uMLPs (*2*, *3*, *25*) for the same alloy.

across all compositions $\langle\delta_{MAE}\rangle$, as shown in Fig. 2B (see Supplementary Section S1 for more details). These large errors reveal that none of the tested uMLPs can accurately resolve the fine energy differences associated with chemical disorder, highlighting a key shortcoming of compositionally agnostic MLP training. This motivates the development of strategies for designing training dataset that systematically and selectively sample the full range of chemical complexity present in solid solutions.

MBS for solid solutions

The accuracy of MLPs in modeling solid solutions hinges on how well their training datasets capture chemically relevant configurations—a complicated task due to the overwhelming chemical complexity of solid solutions. Recently (*10*, *16*, *18*), we demonstrated that this complexity can be systematically characterized by decomposing the structure into local coordination polyhedra—called local chemical motifs—and analyzing their frequency and spatial correlations. Building on this framework, we introduce an MBS method to construct MLP training datasets that more uniformly represent local chemical environments, as illustrated in Fig. 3A (see the Materials and Methods section "Sampling of chemical motifs").

The impact of MBS on model performance is evaluated by comparing it against a baseline random sampling (RS) approach, in which chemically random configurations are drawn uniformly across the ternary phase space (see the Materials and Methods section "Training and testing datasets"). Both datasets contain identical numbers of structures, chemical compositions, and thermal noise levels. Whereas the RS dataset reflects unaltered chemical randomness, the MBS approach refines these configurations via intracell atomic swaps that promote a more uniform distribution of chemical motifs—reducing bias in motif frequencies and capturing a broader diversity of atomic environments.

Although both datasets maintain comparable degrees of chemical short-range order (SRO; see Supplementary Section S2), MBS achieves higher motif packing density—defined as the percentage of unique motifs sampled in the dataset—rising by 27% in MBS relative to RS. Equivalently, the Jensen-Shannon divergence between the sampled and target uniform motif distribution is reduced by 0.0442 bits in MBS (this divergence quantifies the similarity between two probability distributions, with lower values indicating a closer match to the target). This improvement scales with dataset size. For example, in a larger dataset with 702 configurations (compared to 66 configurations used to obtain Fig. 3B), the divergence drops from 0.4814 to 0.3661 bits, and the motif packing density increases by 38%, such that 72.3% of all possible motifs are sampled in the MBS dataset.

As a complementary approach, we also evaluated the widely used special quasirandom structure [SQS (*26*, *27*)] method, which is commonly used to generate chemically disordered configurations for training machine-learned potentials. SQS constructs atomic arrangements that replicate the pair correlations of an ideal random alloy using optimization techniques. To assess its effectiveness, we applied SQS to each configuration in the 702-structure RS dataset. However, this did not lead to improved results: The Jensen-Shannon divergence slightly increased from 0.4814 to 0.4857 bits, and the fraction of unique chemical motifs remained approximately constant at 52.2%.

To quantify how the improved motif diversity of MBS training sets affects model accuracy, we trained an ensemble of 10 MLPs per dataset (see the Materials and Methods section "MLP model training") and evaluated them on five test sets constructed with controlled degrees of SRO. These test sets were generated using reverse Monte Carlo (rMC) (see the Materials and Methods section "Training and testing datasets"), with the total amount of SRO quantified by the metric in Eq. 1. Because all training datasets share identical thermal noise and lattice parameters, any differences in model performance can be attributed specifically to motif representation. As shown in Fig. 3B, MBS-trained models consistently outperform their RS counterparts, with the performance gap widening as the amount of SRO increases. This difference ranges from sub-milli–electron volts per atom in near-random regimes to as much as 5 meV/atom in high-SRO configurations (Fig. 3B; see Supplementary Section S3 for more details).

This improvement is physically meaningful: Typical energy differences associated with SRO in alloys are on the order of 10 meV/

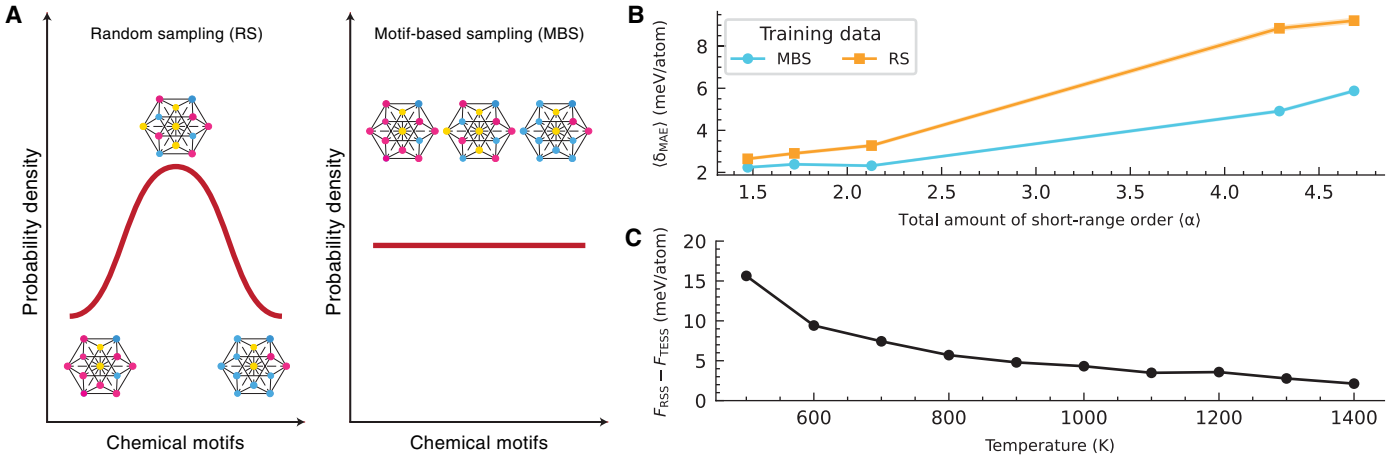

Fig. 3. Impact of MBS on the predictive performance of MLPs in solid solutions. (**A**) Illustration of the chemical motif population density in a random solid solution compared to that in an MBS dataset. (**B**) Mean absolute energy error as a function of the total amount of SRO (Eq. 1) in the test set (averaged across the compositional space), highlighting that MBS-trained models consistently outperform RS. SEM across 10 independently initialized models are shown as shaded regions but are too small to be visible. (**C**) Free-energy difference between a CrCoNi random solid solution and its thermally equilibrated counterpart. The energy scale demonstrates that the accuracy gains provided by MBS are physically significant: They are necessary to resolve energetically competing configurations in alloy thermodynamics.

atom; thus, this accuracy is necessary to resolve energetically competing configurations in alloy thermodynamics (Fig. 3D; see the Materials and Methods section "Free energy" and Supplementary Section S4). By increasing motif diversity, MBS enables finer energetic resolution across chemically distinct configurations, supporting more accurate prediction of order-disorder energetics and related thermodynamic quantities such as free energies.

Completing the training set: Structural and thermal components

To achieve robust predictive performance, the training set must incorporate not only chemical diversity but also structural and thermal variations. Although MBS accounts for chemical complexity, two additional components—namely, phase sampling and thermal perturbations—are needed to ensure coverage of distinct lattice types and vibrational states. These components are briefly summarized below (see the Materials and Methods sections "Sampling of thermal effects" and "Training and testing datasets" for a full description).

Phase sampling is achieved by including metastable phases that differ in structure and chemical order. In practice, this is performed by applying MBS to solid solutions of different lattices [e.g., face-centered cubic (fcc), body-centered cubic (bcc), and hexagonal close-packed (hcp)] and by including a variety of stoichiometric compounds (i.e., ordered intermetallics). These additions expose the model to variations in local coordination and long-range order that are absent from chemically disordered phases. (See Supplementary Section S16 for more information on chemically ordered phases.)

Thermal perturbations are introduced via synthetic thermal noise that mimics vibrational effects and by thermal expansion. In practice, configurations are isotropically expanded to match target thermal expansion factors and further perturbed by random atomic displacements and strain fields. These augmentations are applied uniformly across all phases to ensure smooth interpolation over the vibrational configuration space.

We evaluate this comprehensive sampling strategy—combining chemical, structural, and thermal components—on the CrCoNi system. The resulting MLP achieves a substantial reduction in energy prediction error across the compositional space, as shown in Fig. 4. Compared to uMLPs (Fig. 2A), this approach produces significantly lower and more uniform errors across the full ternary composition triangle. The average MAE, shown in Fig. 2B, also confirms this improvement.

We now shift from developing and benchmarking the MBS framework to validating its predictive capabilities. To this end, we apply MBS to construct MLP training sets for additional systems, including the AuPt and AuCu binary alloys, as well as the high-entropy TiTaVW alloy. Validation is carried out across a range of experimentally accessible properties, including phase diagrams, chemical SRO, thermoelastic and calorimetric behavior, and microstructural properties. Details of dataset construction, including DFT parameters and MBS parameters, are provided in the Materials and Methods sections "Training and testing datasets" and "Density functional theory." For each system, we trained an MLP using the atomic cluster expansion [ACE (*28*)] model (see the Materials and Methods section "MLP model training").

Phase diagrams

To assess the thermodynamic fidelity of our MLPs, we evaluated their ability to reproduce experimental phase behavior in both binary alloys and compositionally complex high-entropy systems. This provides a stringent test of the model's accuracy beyond static energy prediction, requiring them to capture equilibrium thermodynamic transitions across a range of compositions, structures, and temperature conditions.

We first test our approach on the CrCoNi system. Using a single MLP trained on the full CrCoNi composition space (i.e., the model in Fig. 4), we computed phase diagrams for two of its binary subsystems: CrNi and CrCo.

For all cases, solid-state phase boundaries were evaluated using the multicell Monte Carlo approach (*29*), modified to account for vibrational degrees of freedom (see the Materials and Methods section "Phase diagrams"). The predicted phase diagrams are shown in Fig. 5 (A and B), where it can be seen that the MLP predictions

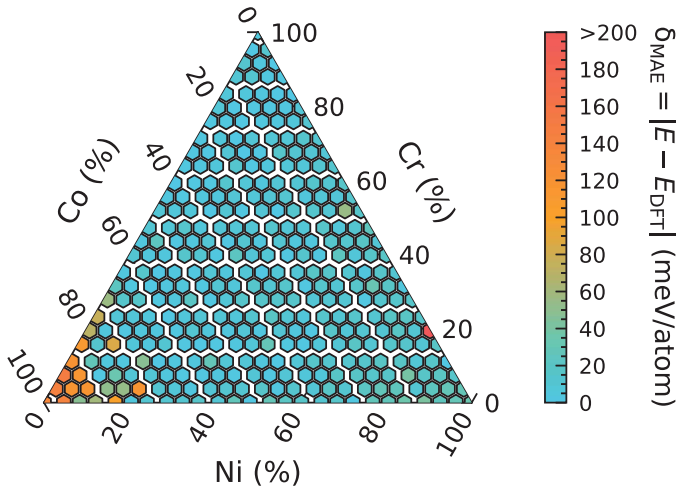

Fig. 4. Energy prediction error across the CrCoNi composition space. The MAE of MLP predictions using combined motif-based, phase, and thermal sampling. Compared to the universal MLP shown in Fig. 2A, this approach yields significantly lower and more uniform errors across the composition triangle.

closely match available experimental data and CALPHAD models, accurately capturing key phase boundaries. For example, in the CrNi binary the fcc-bcc boundaries align well with the experimental measurements, demonstrating the MLP's ability to resolve competing crystal structures. To further illustrate the versatility of our approach, we also examined the AuPt system. The results are shown in Fig. 5C, where it can be seen that the model accurately captures the main feature of this phase diagram, namely, the miscibility gap and its evolution with temperature. Notably, the miscibility gap marks a solid-solid transition where phase decomposition occurs within the same crystal lattice. This contrasts sharply with the fcc-bcc transition observed in CrNi (Fig. 5A) and the order-disorder transitions (fcc-sigma-bcc) in CrCo (Fig. 5A and Supplementary Sections S11 and S16), which arise from fundamentally different thermodynamic driving forces and exhibits distinct kinetic behavior. Across all systems, Fig. 4 shows that the discrepancy between MLP predictions and experiment is comparable to the spread in experimental datasets themselves, indicating that the remaining differences are likely within the limits of experimental uncertainty.

Having validated the model against well-characterized solid-solid transitions in binary alloys, we next evaluate melting points in chemically complex high-entropy alloys, for which phase diagram data are typically sparse or unavailable. To evaluate model performance in this regime, we used the solid-liquid phase coexistence method (*30*) to estimate melting points. This technique isolates a phase interface and tracks the steady-state growth rate of the solid phase as a function of temperature (see the Materials and Methods section "Melting temperature"). For the equiatomic fcc CrCoNi alloy, the melting point was identified as the temperature at which the growth rate vanishes (*31*) (Fig. 6A). The resulting estimate, $T_m = 1641$ K, agrees within 3% of experimental reports (*32*) (note that the solidus-liquidus temperature gap in CrCoNi is small and has not yet been experimentally characterized). Because the liquid phase itself is chemically disordered, accurately predicting melting temperatures further demonstrates the MBS robustness in capturing chemical disorder beyond solid-state configurations.

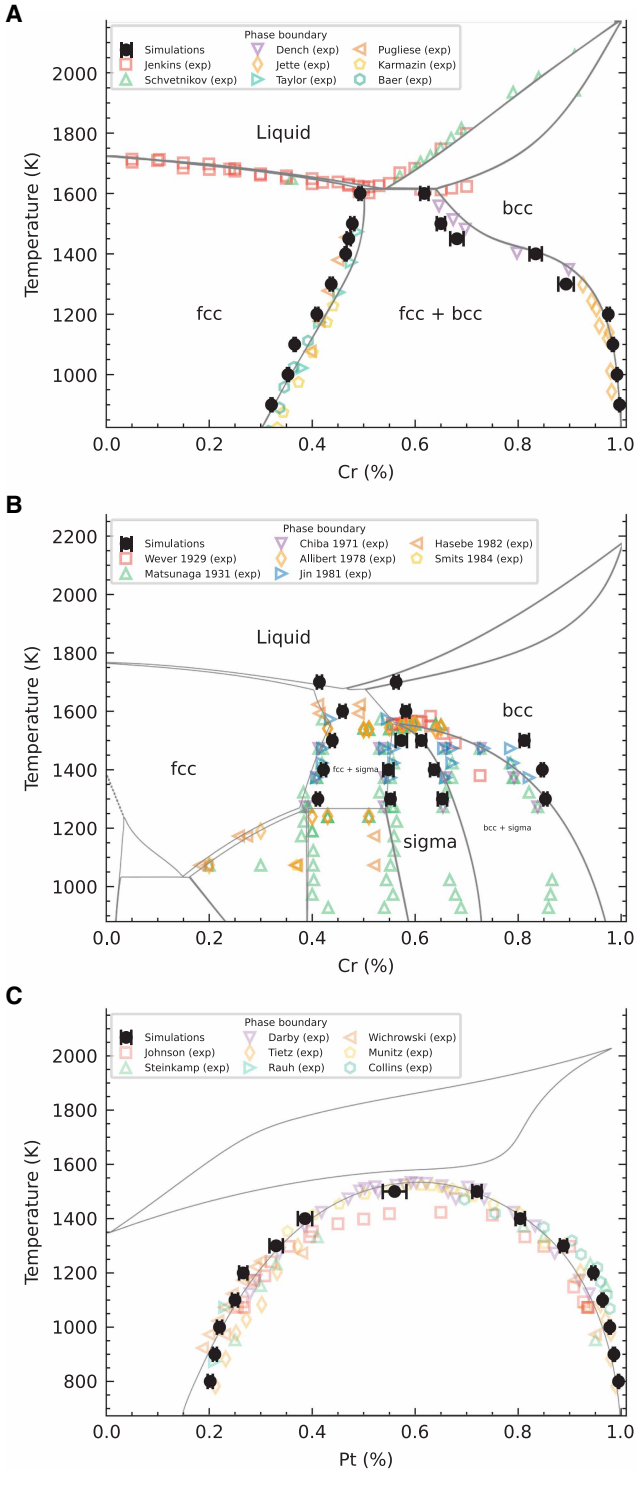

Fig. 5. Phase diagram predictions for binary alloy systems. Predicted solid-state phase boundaries along with corresponding experimental data (*78–101*) and CALPHAD models (*12*, *102–104*) for the binary alloys: (**A**) CrNi, (**B**) CrCo, and (**C**) AuPt. The CrNi and CrCo predictions were obtained using a single MLP trained on the ternary CrCoNi alloy. In contrast, the AuPt predictions were made using an independently trained binary MLP. See Supplementary Section S12 for a comparison of phase diagrams obtained with other potentials.

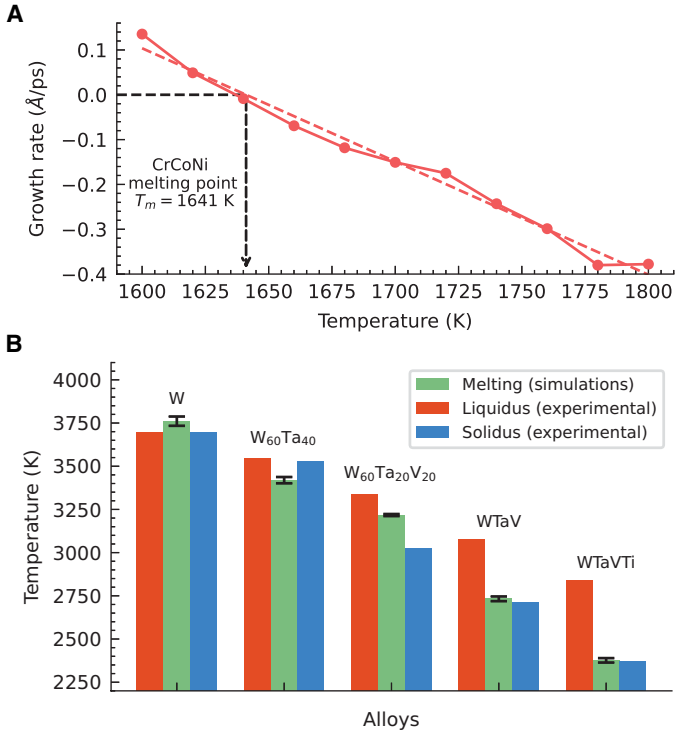

Fig. 6. Predicted melting temperatures for high-entropy alloys. (**A**) Solid-liquid coexistence simulation results for the equiatomic fcc CrCoNi alloy, with melting temperature identified at zero growth rate. The predicted melting temperature (T_m = 1641 K) is within 3% of the experimental value (*32*). (**B**) Predicted melting temperatures for the bcc refractory high-entropy alloy TaTiVW and several derivative compositions, compared against experimentally reported solidus (T_s) and liquidus (T_ℓ) temperatures (*33*). Predictions for the alloys typically fall within the expected coexistence window between T_s and T_ℓ (note that, for pure elements, $T_s = T_\ell$).

We then applied the same procedure to the bcc TaTiVW refractory high-entropy alloy and several of its derivative alloys (Fig. 6B). For each composition, the predicted melting temperature was compared with experimentally reported solidus (T_s) and liquidus (T_ℓ) values (*33*). Our predictions are within 2% for pure W, and they typically fall within the experimental $T_s - T_\ell$ interval, as expected for the alloys (*31*): The melting point as measured in our simulations corresponds to the coexistence of solid and liquid phases and should lie between the onset and completion of melting.

Together, the results in Figs. 5 and 6 highlight the versatility and transferability of the MBS-trained MLPs. Without retraining or fine-tuning, the models generalize from chemically diverse training sets to accurately reproduce equilibrium phase behavior of binary subsystems and complex high-entropy alloys (see Supplementary Section S12 for a comparison of phase diagrams obtained with other potentials). This makes them a powerful tool for exploring phase stability and guiding design of alloys across broad composition spaces.

Chemical SRO

Chemical SRO is the tendency of solid solutions to deviate from perfect randomness, i.e., a measure of local chemical correlations. Although solid solutions are chemically disordered phases lacking periodic chemical arrangements, they nonetheless exhibit subtle energetic biases toward certain local chemical configurations. These biases, captured quantitatively through Warren-Cowley parameters (*34*), fundamentally characterize the statistical fluctuations in local chemical environments and directly reflect underlying thermodynamic equilibria. However, typical training datasets generated from purely random configurations or SQS inherently miss these energetically significant correlations, limiting the accuracy of conventional MLPs (*8*, *10*).

To address this limitation, our MBS explicitly targets motif diversity, thus providing richer training datasets capable of capturing the delicate energetic gradients associated with SRO. To quantitatively assess the accuracy of our MLPs in predicting SRO, we computed the first-nearest-neighbor Warren-Cowley parameters, α_{ij} (Eq. 7), from atomistic simulations of thermally equilibrated solid solutions (see the Materials and Methods section "Warren-Cowley parameters").

Figure 7A compares the temperature dependence of Warren-Cowley parameters predicted for various $Cr_xNi_{(1-x)}$ alloys against experimental values obtained from diffuse scattering measurements (*35*–*37*) (with x ranging from 19.9 to 33.3%). The close agreement demonstrates that our MBS training captures the nuanced energetic biases responsible for SRO across the compositional space. Similarly, Fig. 7B compares the predictions for the Cu_3Au alloy with experimental data from ref. (*38*), where Warren-Cowley parameters are derived from in situ x-ray total scattering and rMC refinement. Although predicted magnitudes of α_{ij} differed slightly from experiments (see Supplementary Section S5)—likely due to known systematic binding energy biases of the DFT functional affecting lattice parameters and thus local ordering—the experimental trends can be accurately captured by standardizing the Warren-Cowley parameters (i.e., transforming the data to have a mean of zero and an SD of 1) to correct for the strength of interactions, removing some of the DFT biases (*39*–*42*).

Overall, our results confirm that MBS effectively captures the subtle energetic biases that drive local chemical ordering and that MLPs trained on such data can resolve SRO contributions.

Thermoelastic and calorimetric behavior

To further evaluate the transferability of our MLPs, we investigated their capacity to capture temperature-dependent thermophysical properties—specifically, thermal expansion and constant-pressure specific heat c_p across the CrCoNi and TaTiVW alloy systems.

Figure 8A shows the thermal expansion of CrCoNi and selected binary derivatives (see the Materials and Methods section "Thermal expansion"). The predicted temperature dependence of lattice parameters align closely with experimental observations. For example, the $Cr_{20}Ni_{80}$ and NiCo systems exhibit a maximum deviations from experimental results of only 0.1 and 0.3%, respectively, across the entire temperature range. The coefficients of thermal expansion (Supplementary Section S6) deviate at most 15% from their experimental values. This close agreement demonstrates the model's accuracy in capturing anharmonic vibrational effects and associated thermal expansion behavior in complex metallic alloy systems.

Figure 8B shows constant-pressure heat capacities for TaTiVW-based alloys (see the Materials and Methods section "Specific heat"). The simulations reproduce the expected monotonic increase in c_p with temperature, reflecting vibrational anharmonicity and mode excitation effects. Comparison with measurements from the NIST (*43*) and recent measurements on refractory high-entropy alloys (*33*) further validates the quantitative reliability of our predictions. The largest observed discrepancy from experimental data was 13%

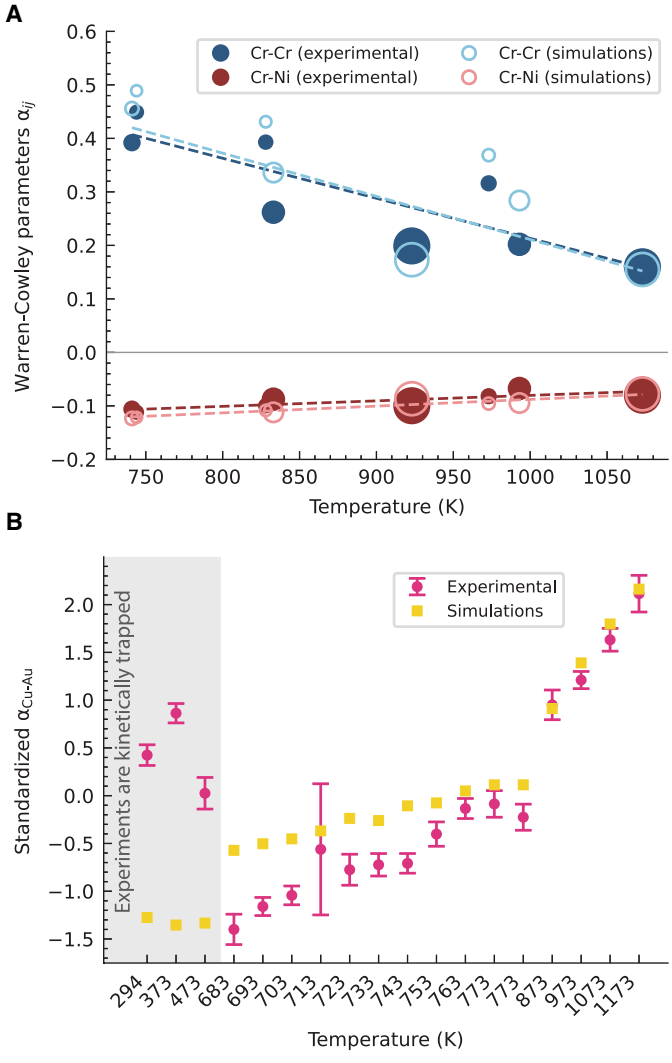

Fig. 7. SRO predictions compared with experimental data. (**A**) Temperature dependence of Warren-Cowley parameters for $Cr_xNi_{(1-x)}$ alloys compared to diffuse scattering experimental measurements (*35*–*37*) across the compositional space. The sizes of the data points are proportional to the Cr concentration, with *x* ranging from 19.9 to 33.3%. Dashed lines are a linear fit to the data with corresponding color. See Supplementary Section S15 for a comparison against other potentials. (**B**) Standardized Warren-Cowley parameters for the Cu_3Au alloy, comparing predicted values against experimental data derived from x-ray total scattering and rMC modeling (*38*). Standardization highlights consistent qualitative trends despite quantitative discrepancies attributed to known systematic energy biases in the DFT functional (*39*–*41*). SEM values of all simulation results are smaller than the data points.

for $Ti_{20}Ta_{20}V_{20}W_{40}$, whereas TiTaVW and $Ta_{40}W_{60}$ deviate by a maximum of 4 and 6%, respectively, and the other three compositions have deviations of less than 2%.

Computationally aided microstructure design

The preceding sections have highlighted the accuracy and fidelity of MBS-trained MLPs in capturing diverse thermophysical and thermodynamic properties across alloy compositions. In this section, we illustrate how such computational accuracy could potentially support materials design and development, particularly in scenarios where experimental exploration is prohibitively challenging or expensive.

A notable example is the stacking-fault energy (SFE; γ_{SFE}), which significantly influences the deformation mechanisms and thus the damage tolerance of fcc alloys. In particular, the CrCoNi alloy has attracted considerable attention due to its exceptional mechanical properties at low temperatures, attributed largely to its low SFE. This low SFE facilitates the activation of alternative deformation pathways such as mechanical twinning and transformation-induced plasticity (*44*), particularly at low temperatures when traditional dislocation slip mechanisms are insufficient (*45*). However, despite its central role in deformation behavior, the SFE has not traditionally served as a practical alloy design parameter. Experimental determination of the SFE is notoriously challenging, requiring careful interpretation of fault structures via electron microscopy or indirect inference methods. Furthermore, recent investigations have underscored the complexity and subtlety involved in reliably extracting accurate SFE values from experiments, highlighting potential pitfalls and uncertainties in current methodologies (*46*).

To address this gap, we used our MLP framework to systematically predict the SFE across the entire compositional space of the CrCoNi alloy at 500 K, explicitly accounting for thermal effects such as SRO at the appropriate length scales (*8*) (Fig. 9; see the Materials and Methods section "Stacking-fault energy"). This computational approach captures subtle energetic contributions that are essential for accurate SFE prediction (*8*, *47*). As shown in Fig. 8, the SFE exhibits a pronounced dependence on composition. Despite the complexity of this variation, clear trends emerge that can inform materials design. For example, reducing the Cr concentration along the $Cr_xCo_{0.5-x/2}Ni_{0.5-x/2}$ path (with $x \leq 80\%$) leads to a marked decrease in the SFE, as indicated by the arrow in Fig. 9. The influence of SRO on the SFE across the compositional space is detailed in Supplementary Section S7, along with additional results at 800 K (showing similar compositional trends).

These computational predictions combined with the predicted phase stability of Fig. 5 illustrate how systematically exploring compositional dependencies of the SFE can provide valuable insights for alloy design. By leveraging this predictive approach, one is able to efficiently identify promising alloy compositions, complementing and extending experimental capabilities in materials development.

DISCUSSION

This work demonstrates that chemical motif diversity in the training dataset is a critical determinant of the predictive accuracy of MLPs for chemically disordered phases. The MBS strategy introduced here systematically improves model performance across compositions and structures by promoting a more uniform and physically meaningful representation of local chemical motifs. The success of this strategy hinges on recognizing that the fidelity of an MLP is limited not only by its architectural complexity or dataset size but also by the representational diversity of the atomic environments on which it is trained. To contextualize our focus on chemically disordered materials, we note that chemically ordered phases (e.g., $L1_2$, $L1_0$, B_2, and σ) are already well described by modern MLP, as demonstrated in prior studies (*2*, *3*, *6*, *7*, *25*). These phases are explicitly included in our training sets and phase diagram validations (see Supplementary Section S16), but the methodological innovations in this work primarily target the historically challenging regime of chemical disorder and temperature-dependent orderings.

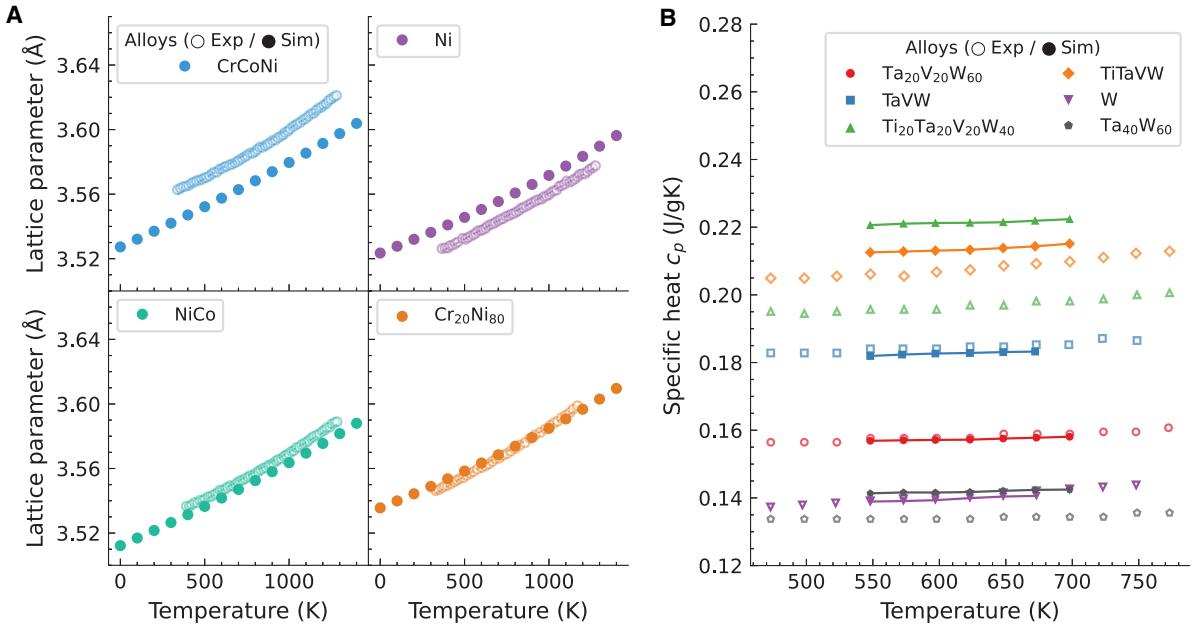

Fig. 8. Thermophysical properties of CrCoNi and TaTiVW. (**A**) Temperature dependence of lattice parameters for CrCoNi and selected binary derivatives compared to experiments from ref. (*109*). (**B**) Specific heat of TaTiVW-based alloys compared with experimental data from ref. (*33*), except for pure W, which is compared to ref. (*43*).

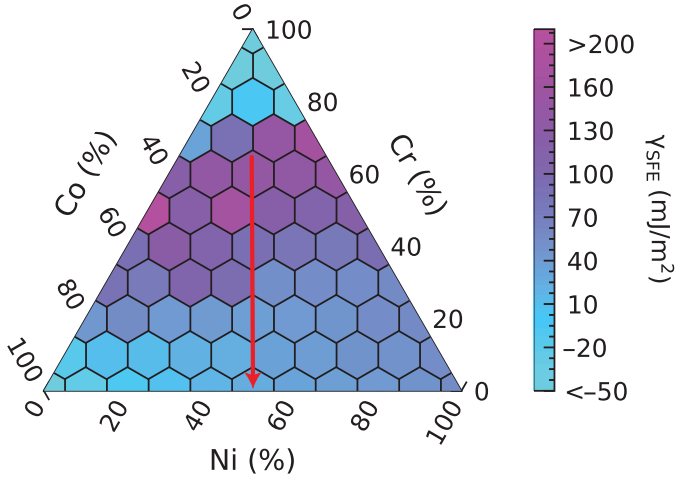

Fig. 9. Computationally aided microstructure design. Predicted SFE (γ_{SFE}) across the CrCoNi composition space at 500 K. The arrow indicates the compositional path $Cr_xCo_{0.5-x/2}Ni_{0.5-x/2}$ (with $x \leq 80\%$), along which the SFE decreases significantly. Trends such as this one can inform materials design.

The generality of the MBS approach is evidenced by its consistent performance across a chemically diverse set of metallic alloys. These include systems with different magnetic, electronic, and structural complexity, all trained from datasets generated with the same consistent first-principles workflow. MBS can be incorporated with negligible computational cost into existing training pipelines to enhance the performance of both MLPs and uMLPs without requiring model architectural changes. Although RS or SQS has historically been the default in many MLP training protocols that account for disordered phases (*48*–*50*), our results demonstrate that MBS offers a more robust and physically informed alternative and should be preferred in applications where transferability and error control are critical.

Recent uMLP architectures—particularly those trained on chemically diverse datasets—show promise for accurately capturing the properties of materials across the compositional space (*51*–*54*). However, the sheer breadth of their training data often obscures the origins of their performance, underscoring the need for systematic validation. In this context, the MBS training and test datasets introduced here provide a valuable benchmark for evaluating the applicability of uMLPs across the compositional space. In addition, we note that our showcase of uMLPs in this work (Fig. 1) serves primarily to reveal their systematic limitations in chemically disordered regimes, not to claim architectural superiority. These models, although promising for small-cell property prediction and materials screening, remain too computationally expensive for the large-scale atomistic simulations required to study processing, synthesis, and microstructure evolution. Specialized MBS-trained potentials, by contrast, retain high accuracy across compositional and thermal conditions while remaining tractable for simulations involving millions of atoms and long timescales. This capability is essential for computing engineering-relevant quantities such as phase diagrams and deformation energetics, which demand extensive thermodynamic sampling that current uMLP architectures cannot feasibly support.

The ability to reliably compute key materials properties across compositions—from phase stability and thermophysical behavior to microstructure energetics—hinges on the availability of accurate and efficient interatomic models. Within this context, MBS improves the practicality of MLPs for alloy design by enhancing accuracy at fixed training set size, thereby reducing the data burden typically associated with high-throughput property predictions. These predictions are central to materials development; for example, accurate free-energy predictions allow the construction of phase diagrams that inform heat treatment protocols, whereas the SFE guides assessments of deformation mechanisms. When combined with uncertainty quantification, such predictions provide actionable guidance for materials development. A promising future direction is

the development of automated frameworks for high-throughput phase diagram generation. Although our current approach achieves accurate predictions of phase boundaries, it still relies on prior knowledge of the relevant competing phases. Generative models (*55*) offer a potential solution by autonomously proposing both stable and metastable structures across compositional and structural landscapes.

It is essential to acknowledge, however, that the ultimate accuracy of the predicted properties remains bounded by the quality of the reference data used to train the MLPs. The exchange-correlation functionals used in DFT calculations in this work have well-known limitations in their predictive accuracy (*40*, *41*). Nevertheless, functional development remains an active area of research, with recent advances [e.g., r2SCAN (*56*) and ML density functionals (*57*–*59*)] offering improvements in physical accuracy. Higher-fidelity quantum chemistry methods may provide an even more reliable foundation for MLP training sets (*60*).

Crucially, even if a perfect electronic structure method was available—capable of producing exact quantum mechanical results—the methodological framework introduced here would remain indispensable. High-accuracy reference data alone do not guarantee generalization across chemical composition space, nor do they obviate the need for thoughtful sampling of chemical disorder during training. The MBS approach offers a systematic path toward generalizable and practically useful MLPs. Moreover, our findings suggest that the current generation of functionals may already be sufficient for property predictions of practical relevance, so long as the compositional and configurational diversity of the training data is carefully managed.

In conclusion, this work advances the development of MLPs by addressing two persistent limitations: the absence of systematic strategies for chemical sampling of disordered phases and the difficulty of maintaining accuracy across broad compositional ranges. By introducing a chemical motif–based framework for training set construction and demonstrating its effectiveness across multiple alloy systems, we provide a foundation for building MLPs that are both accurate and robust—delivering reliable predictions across structures and compositions. This predictive performance, further supported by close agreement with experimental data across a range of thermodynamic and thermophysical properties, underscores the potential of MLPs to transition from niche research tools to a core component of multiscale simulation pipelines in materials design.

MATERIALS AND METHODS

Sampling of chemical motifs

The MBS datasets are constructed using the rMC (*61*) algorithm to generate atomic configurations that approximate a uniform probability distribution over chemical motifs, $P^{\text{target}}(\mathcal{M})$.

We begin with N chemically random supercells of size $n_x \times n_y \times n_z$, constructed using a lattice parameter equal to the composition-weighted average of the DFT-relaxed values of the pure elements at 0 K. The chemical composition C_i of the supercells are selected to evenly span the concentration space. For example, in a binary system A_xB_{1-x} with $N = 5$, the initial supercells are chemically random solid solutions with compositions $x \in \{0, 0.25, 0.5, 0.75, 1\}$.

The rMC simulation is run for N_{rMC} intracell swap attempts, with the acceptance probability for an atomic swap at step k given by $\min\left[1, \exp\left(-\beta_{k+1}\Delta\chi^2\right)\right]$, where $\beta_{k+1} = (1+k)\beta_k$ and $\Delta\chi^2 = \chi^2_{k+1} - \chi^2_k$ with

$$\chi^2_k = D_{\text{JS}}\left[P^{(k)}(\mathcal{M}) \,\|\, P^{\text{target}}(\mathcal{M})\right]$$

where D_{JS} is the Jensen-Shannon divergence used to quantify the difference between the current motif distribution $P(\mathcal{M})$ and the target uniform distribution $P^{\text{target}}(\mathcal{M}) = \frac{1}{N_{\text{ms}}}$, where N_{ms} is the number of unique motifs in an n-component system. At each step k, the motif distribution $P^{(k)}(\mathcal{M})$ is evaluated using our chemical motif identification framework (*10*, *16*), which uses an E(3)-equivariant graph neural network to classify local chemical environments. This framework has been shown to capture the full spectrum of chemical motifs of alloys with up to five elements across the fcc, bcc, and hcp crystal structures.

Sampling of thermal effects

To incorporate thermal expansion and vibrations, each supercell is modified after the MBS using a strategy inspired by our prior works (*8*, *10*, *62*).

To account for thermal expansion, supercells are isotropically expanded using a linear expansion coefficient, corresponding to thermal expansion at an effective temperature T_e. Half of the supercells are expanded to reflect thermal expansion at 300 K, and the other half at $0.9T_e$. After this thermal expansion, a random strain ϵ uniformly drawn from $\left[-\epsilon_{\min}, \epsilon_{\max}\right]$ is also applied to each dimension isotropically.

To account for thermal vibrations, each atom is displaced by a random vector drawn from a uniform distribution within a sphere. To capture both low-temperature and high-temperature vibrational amplitudes, the maximum displacement is set proportional to $\alpha \times d_{\text{nn}} / 2$, where α is uniformly sampled in the interval $\left[\alpha_{\min}, \alpha_{\max}\right]$, and where $d_{\text{nn}} / 2$ is half of the nearest-neighbor distance.

The choice of the parameters (i.e., α, T_e, $\alpha_{\min}$, and $\alpha_{\max}$) is system specific and based on their chemistry, crystal structure, and dataset intended use (e.g., training or test set).

Training and testing datasets

This section details the datasets used to train and evaluate the MLPs presented in this work. We start by describing the construction of the training and test sets for CrCoNi used to generate the results in Figs. 2 and 3.

The CrCoNi test set used in Fig. 2 (and Fig. 4) consists of 351 configurations that sample the CrCoNi compositional space evenly across the ternary triangle; the chemical configuration of each snapshot is random, i.e., these are random solid solutions. Each snapshot is a $3 \times 3 \times 3$ fcc supercell (108 atoms). Artificial expansions, atomic displacements, and strain were applied using the following parameters: $T_e = 1800\,\text{K}$, $\epsilon_{\min} = -3\,\%$, $\epsilon_{\max} = 3\,\%$, $\alpha_{\min} = 0.01$, and $\alpha_{\max} = 0.5$.

The CrCoNi MBS and RS training sets used to train the MLPs in Fig. 3 were designed to highlight the predictive differences between potentials trained using these two approaches. Each dataset consists of 66 configurations spanning the CrCoNi ternary composition space evenly. Each snapshot is a $3 \times 3 \times 3$ fcc supercell containing 108 atoms. The MBS dataset was generated by applying the MBS procedure to the RS structures to enhance motif diversity, using $N_{\text{rMC}} = 1{,}069{,}200$ and $\beta_0 = 100$. To retain force information during training while isolating chemical effects, a small thermal noise perturbation was applied uniformly to all systems ($\alpha = 0.05$). No strain augmentations were included. All structures were isotropically expanded to a thermal expansion of 300 K.

The CrCoNi test sets used in Fig. 3B were designed to represent a range of artificial SRO configurations in CrCoNi to avoid biases associated with using MLP-specific SRO configurations. Five datasets were created, each consisting of 66 3 × 3 × 3 fcc supercell containing 108 atoms, evenly spanning the ternary compositional space. No strain augmentations or thermal noise were applied to isolate the effect of chemistry on the test errors. All structures were isotropically expanded to a thermal expansion of 300 K. To introduce artificial SRO into the systems, rMC simulations were performed to enforced targeted amounts of SRO, quantified by a scalar measure of SRO called the total amount of order, defined as

$$\alpha = \sum_{ij} |\alpha_{ij}| \quad (1)$$

where α_{ij} denotes the first nearest-neighbor Warren-Cowley parameter for pair *ij* (Eq. 7). Each of the five datasets corresponds to a distinct target value of α, representing varying degrees of chemical ordering.

Next, we describe the complete datasets used to train the MLPs for the AuCu, AuPt, CrCoNi, and TaTiVW systems. These MLPs were used to generate the results in Figs. 4 to 9. These datasets were designed to capture a broad range of thermodynamic and structural conditions across all four material systems.

The AuCu training dataset is composed of 202 3 × 3 × 3 fcc supercells, each containing 108 atoms, evenly spaced across the binary compositional space, and with two cells per composition. The systems were constructed using MBS with $N_{\mathrm{rMC}} = 1{,}000{,}000$ and $\beta_0 = 100$. Artificial expansion, thermal noise, and strain were applied to each system using the parameters: $T_e = 1650\,\mathrm{K}$, $\epsilon_{\min} = -3\,\%$, $\epsilon_{\max} = 3\,\%$, $\alpha_{\min} = 0.01$, and $\alpha_{\max} = 0.5$. No stoichiometric compounds (i.e., ordered intermetallics) were included because none were expected (from experimental phase diagrams) in the temperature and composition range used in our simulations. Additional configurations of the pure elements were included, with additional 76 2 × 2 × 2 fcc supercells (32 atoms each) generated for each element. Artificial expansion, thermal noise, and strain were applied using the following parameters: $T_e^{\mathrm{Au}} = 1337\,\mathrm{K}$, $T_e^{\mathrm{Cu}} = 2041\,\mathrm{K}$, $\epsilon_{\min} = -5\,\%$, $\epsilon_{\max} = 5\,\%$, $\alpha_{\min} = 0.01$, and $\alpha_{\max} = 0.6$.

The AuPt training dataset follows the same construction strategy as the AuCu dataset but with different parameters during: $T_e = 2100\,\mathrm{K}$, $\epsilon_{\min} = -3\,\%$, $\epsilon_{\max} = 3\,\%$, $\alpha_{\min} = 0.01$, and $\alpha_{\max} = 0.5$.

The CrCoNi training dataset is composed of 702 3 × 3 × 3 fcc supercells (each containing 108 atoms), 702 4 × 3 × 3 hcp supercells (each containing 144 atoms), and 351 4 × 4 × 4 bcc supercells (each containing 128 atoms). The supercells are evenly spaced across the ternary compositional space, with two supercells per composition for the fcc and hcp structures. Each of the fcc, hcp, and bcc datasets was optimized using MBS, with $\beta_0 = 100$ and $N_{\mathrm{rMC}} = 1{,}895{,}400$, $N_{\mathrm{rMC}} = 2{,}021{,}760$, and $N_{\mathrm{rMC}} = 898{,}560$, respectively. Artificial expansion, thermal noise, and strain were applied to each using the following parameters: $T_e = 1800\,\mathrm{K}$, $\epsilon_{\min} = -3\,\%$, $\epsilon_{\max} = 3\,\%$, $\alpha_{\min} = 0.01$, and $\alpha_{\max} = 0.5$. The dataset also includes pure elements, as well as the following stoichiometric compounds (i.e., ordered intermetallics): $L1_2$, $L1_0$, $L1_1$, and B_2 for all pairs of elements, and $CrNi_2$ (space group $Fd\bar{3}m$) and Cr_8Co_7 (space group $P4_2nm$). Twelve configurations of each (pure elements and intermetallics) were generated as 3 × 3 × 3 supercells. Thermal noise was applied using the same parameters as in the MBS datasets.

The TaTiVW training dataset is composed of 572 2 × 2 × 2 fcc supercells (each containing 32 atoms), 572 4 × 3 × 3 hcp supercells (containing 144 atoms), and 572 3 × 3 × 3 bcc supercells (each containing 54 atoms), evenly spaced across the quaternary compositional space, with two supercells per composition for the fcc, bcc, and hcp structures. Each of the fcc, hcp, and bcc datasets was optimized using MBS, with $\beta_0 = 100$ and $N_{\mathrm{rMC}} = 1{,}000{,}000$, $N_{\mathrm{rMC}} = 660{,}000$, and $N_{\mathrm{rMC}} = 1{,}000{,}000$, respectively. Artificial expansion, thermal noise, and strain were applied to each with parameters: $T_e = 3700\,\mathrm{K}$, $\epsilon_{\min} = -3\,\%$, $\epsilon_{\max} = 3\,\%$, $\alpha_{\min} = 0.01$, and $\alpha_{\max} = 0.5$. The dataset also includes pure elements, as well as the B_2 intermetallics for all element pairs. For the pure elements, a total of 76 2 × 2 × 2 supercells were generated in fcc, bcc, and hcp structures for each element. Artificial expansion, thermal noise, and strain were applied to each system with parameters: $T_e^{\mathrm{Ta}} = 3293\,\mathrm{K}$, $T_e^{\mathrm{W}} = 3695\,\mathrm{K}$, $T_e^{\mathrm{Ti}} = 1941\,\mathrm{K}$, $T_e^{\mathrm{V}} = 2183\,\mathrm{K}$, $\epsilon_{\min} = -5\,\%$, $\epsilon_{\max} = 5\,\%$, $\alpha_{\min} = 0.01$, and $\alpha_{\max} = 0.6$. Twelve systems of each intermetallic were generated as 3 × 3 × 3 supercells. Thermal noise was applied using the same parameters as in the MBS datasets.

Density functional theory

The Vienna ab initio simulation package [VASP (*63*–*67*), version 6.2.1] was used to perform collinear spin-polarized DFT calculations. The Perdew-Burke-Ernzerhof [PBE (*68*)] exchange-correlation functional and projector-augmented wave (*69*) pseudopotentials (version potpaw.PBE.54) were used for all systems, except for AuPt, where local density approximation [LDA (*70*, *71*)] was used instead because it yields a lattice parameter in better agreement with experiment than PBE.

For the test set shown in Fig. 2, we used the Materials Project static calculation parameters (MPStaticSet) as implemented in pymatgen to ensure consistency with the uMLPs, which were trained on first-principles data with similar settings.

For the rest of the datasets, however, the valence configurations used in pseudopotentials were as follows Cr: $3p^6 3d^5 4s^1$, Co: $3d^8 4s^1$, Ni: $3d^8 4s^2$, Au: $5d^{10} 6s^1$, Pt: $5d^9 6s^1$, Cu: $3d^{10} 4s^1$, Ta: $5p^6 5d^4 6s^1$, Ti: $3s^2 3p^6 3d^3 4s^1$, V: $3s^2 3p^6 3d^4 4s^1$, and W: $5s^2 5p^6 5d^5 6s^1$. For the AuPt, CuAu, CrCoNi, and TaTiVW systems, electronic states were sampled using Monkhorst-Pack meshes of 9 × 9 × 9, 8 × 8 × 8, 9 × 9 × 9, and 14 × 14 × 14, respectively, for 4-atom fcc unit cells. In addition, in the CrCoNi, and TaTiVW systems, electronic states were sampled using Monkhorst-Pack meshes of 12 × 9 × 9 and 14 × 14 × 14, respectively, for 4-atom hcp unit cells. For bcc unit cells with 2 atoms, meshes of 12 × 12 × 12 and 15 × 15 × 15 were used. For larger supercells, the *k*-point density was scaled inversely with the real-space supercell size (e.g., a 3 × 3 × 3 fcc CrCoNi supercell was sampled using a 3 × 3 × 3 mesh). A second-order Methfessel-Paxton smearing scheme was used, with widths of 0.26, 0.12, 0.10, and 0.15 eV, and plane-wave energy cutoffs of 350, 450, 430, and 500 eV, respectively. Spin polarization was included only for calculations involving the CrCoNi system. For all systems, the energy threshold for self-consistency and the force threshold for structure relaxation were 10^{-5} eV and 0.02 eV / Å, respectively.

MLP model training

The MLPs for all alloy (AuCu, AuPt, CrCoNi, and TaTiVW) were trained using the performant implementation of the ACE (*28*) framework via the python-ace package. Models construction followed the general body-ordered ACE formalism with increasing correlation order and included both one-body and many-body terms. The global

neighbor list cutoff radius was set to 5.3/5.3/5.0/5.5Å for each system, respectively. We used a Chebyshev-based radial basis (ChebExpCos) and a smooth cosine cutoff function, with radial supports ranging from 1.75/1.75 /1.75/1.5Å to 5.3/5.3 /4.75/5.5Å, and a spline bin width of 0.001 Å. Angular and radial resolutions were controlled using body-order–specific parameters. For unary contributions, we used $\ell_{\max} = [0,3,2,1,1,0]$ ([0,4,3,2,2,1] for TaTiVW) and $n_{\text{rad}} = [15,6,3,2,2,1]$ ([20,8,4,3,3,2] for TaTiVW) across body-orders 1 to 6. For two-body to five-body interactions, we used $\ell_{\max} = [0,2,2,1,1]$ ([0,3,3,2,2] for TaTiVW) and $n_{\text{rad}} = [15,3,3,2,1]$ ([20,4,4,3,2] for TaTiVW). Each element's basis was limited to a maximum of 700 functions (1000 for TaTiVW). For the density embedding, we used a shifted and scaled Finnis-Sinclair–type functional (FinnisSinclairShiftedScaled), parameterized using a two-channel density and internal parameters [1,1,1,0.5] per element.

The training process minimized a combined force and energy loss function with regularization. Specifically, we applied both $\mathcal{L}_1$ and $\mathcal{L}_2$ regularization penalties of $10^{-8}/10^{-8}/10^{-8}/10^{-9}$ to the model coefficients and used a damping factor κ = auto/auto/0.1/0.05 to weight the energy contributions. Optimization was performed using a Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimizer with a maximum of 1400/1400/1400/2000 iterations and a hierarchical body-order fitting schedule (ladder_type: body_order; ladder_step: 25/25/25/15). We used a batch-wise evaluation scheme with an initial batch size of 400, which was adaptively reduced during training to improve memory efficiency if needed.

Free energy

The free energy in Fig. 3C was computed by accounting for the vibrational and configurational contributions.

For the vibrational contributions we used the nonequilibrium Frenkel-Ladd (FL) method (*72*) implemented in the large-scale atomic/molecular massively parallel simulator [LAMMPS (*73*)]. This method estimates the free energy of a solid by performing a dynamic interpolation between the system Hamiltonian H_0 and a reference Einstein crystal H_E, according to

$$H(\lambda) = (1-\lambda)H_0 + \lambda H_E, \quad \lambda \in [0,1] \tag{2}$$

where λ is a time-dependent switching parameter. The free energy is obtained by combining the results from forward and backward switching paths

$$F_0 = F_E + \frac{1}{2}\left(W_{\text{irr}}^{\text{fwd}} - W_{\text{irr}}^{\text{bwd}}\right) + \delta F_{\text{CM}} \tag{3}$$

where F_E is the known analytic free energy of the Einstein crystal, δF_{CM} accounts for the fixed center-of-mass constraint, and $W_{\text{irr}}^{\text{fwd}}$ and $W_{\text{irr}}^{\text{bwd}}$ are the irreversible work contributions from the forward and backward paths, respectively.

All molecular dynamics simulations used a timestep of 1 fs and were performed using the velocity-Verlet integration algorithm with a Langevin thermostat with a damping constant of 0.1 ps. Each trajectory included a 10-ps equilibration period followed by 100 ps of switching in both the forward (λ: 0 → 1) and backward (λ: 1 → 0) directions. The switching was carried out using the fix ti/spring command in LAMMPS, with the function 2 polynomial form to reduce dissipation. The Einstein crystal's spring constant k were set to 2.5 eV/Å^2. Each free energy value was averaged over five forward and backward switching trajectories. Final values were reported as the average over all runs, and statistical uncertainties were estimated as the SEM.

The vibrational contribution to the free energy of random solid solutions (F_{RSS}) was obtain by applying this procedure to chemically random solid solutions. Meanwhile, for the vibrational contribution of thermally equilibrated solid solutions (F_{RSS}) the chemical configurations used were obtained from the final step of Monte Carlo simulations with the acceptance probability for an atomic swap at any given step $\min\left[1, \exp(-\beta\Delta E)\right]$, where $\beta^{-1} = k_B T$, k_B is the Boltzmann constant, and $\Delta E = E_f - E_i$ with E_f the energy after the atomic swap and E_i the energy before it. These Monte Carlo simulations were initialized from chemically random 14 × 14 × 14 supercells (10,976 atoms) using lattice parameters adjusted for thermal expansion. Each simulation ran for 25 Monte Carlo steps per atom.

The configurational contribution to the free energy is $-TS_{\text{conf}}$, where the configurational entropy was computed using the pair approximation within the cluster variation method (CVM) (*74*)

$$S_{\text{conf}}^{\text{pair}} = k_B N\left[(2w-1)\sum_i L(x_i) - w\sum_{ij} L(y_{ij}) + (w-1)\right] \tag{4}$$

where N is the number of lattice sites. The variable x_i is the probability of finding species i at a site, and y_{ij} is the probability of finding species i and j on a nearest-neighbor pair. The factor w is half the coordination number (equal to 6 for fcc lattices). The function $L(x) = x\ln x - x$ is used to express the entropy terms compactly.

Phase diagrams

The phase diagrams in Fig. 5 were computed using a multicell Monte Carlo method within the semigrand isobaric ensemble, obtained by adapting the MC^2 algorithm (*29*, *75*–*77*) to operate within the semigrand ensemble. This approach models coexisting phases in distinct simulation cells under a shared chemical potential difference Δμ, which enables equilibrium between coexisting phases to emerge naturally, without interfacial energy penalties. A key advantage of this approach is the natural equilibrium stopping condition provided by the semigrand isobaric ensemble: Phase equilibrium is achieved when all coexisting phases have equal semigrand potentials at the imposed Δμ. This thermodynamic criterion ensures that the system has reached true equilibrium once each cell minimizes its semigrand free energy, which provides a robust and unambiguous determination of phase boundaries.

For each temperature point, two or more simulation cells were initialized with different crystal structures (fcc and bcc for CrNi; fcc and sigma or sigma and bcc for CrCo; fcc only for AuPt) to represent potentially coexisting phases. Each cell contained 8 × 8 × 8 unit cells for fcc and bcc structures (2048 atoms for fcc and 1024 atoms for bcc) or 4 × 4 × 4 unit cells for the sigma phase (1920 atoms) and was subject to periodic boundary conditions in all directions. The cells equilibrated under identical temperature T, pressure $P = 0$, and chemical potential difference $\Delta\mu = \mu_B - \mu_A$ conditions.

Within each cell, semigrand Monte Carlo moves were performed by randomly selecting atoms and attempting to change their chemical identity (e.g., Cr → Ni or Ni → Cr). The acceptance probability for these transmutation moves followed the Metropolis criterion

$$P_{\text{acc}} = \min\left\{1, \exp\left[-\beta\left(\Delta E - \Delta\mu\cdot\Delta N_B\right)\right]\right\} \tag{5}$$

where ΔE is the change in potential energy, $\beta = 1/k_B T$, and ΔN_B is the change in the number of B-type atoms.

To capture vibrational contributions, short molecular dynamics trajectories were performed between Monte Carlo cycles using a Nosé-Hoover thermostat and barostat (damping parameters of 100 timesteps) with a timestep of 0.001 ps. Each Monte Carlo cycle consisted of N attempted transmutations (where N is the number of atoms) followed by 10 molecular dynamics steps.

Phase fraction moves were implemented to allow virtual mass transfer between cells without physical particle exchange. Small perturbations to the phase fractions $\Delta\varphi$ were proposed and accepted according to

$$P_{acc} = \min\left\{1,\ \exp\left[-\beta N_{tot} \times \sum_i \left(\Omega_i - \Omega_{min}\right) \cdot \Delta\varphi_i\right]\right\} \quad (6)$$

where Ω_i is the semigrand potential per atom of phase i, Ω_{min} is the minimum among all phases, and N_{tot} is the total number of atoms in the system.

For each temperature, the appropriate $\Delta\mu$ value was determined using a parallel grid search to achieve equal semigrand potentials between coexisting phases. Simulations were equilibrated for 100,000 Monte Carlo cycles, followed by 50,000 production cycles (1,000,000 and 500,000 molecular dynamics steps, respectively). Phase compositions were extracted by averaging over the production period, with error bars calculated as the SEM.

For validation purposes, the predictions were compared to those of experimental data from refs. (*78*–*85*) as compiled by ref. (*86*) for CrNi, from refs. (*87*–*92*) as compiled by ref. (*93*) for CrCo, and from refs. (*94*–*100*) as compiled by ref. (*101*) for AuPt. Predictions are also compared to CALPHAD phase diagrams of CrNi (*102*), CrCo (*103*), and AuPt (*104*) as compiled by the ASM Alloy Phase Diagram Database (*12*).

Melting temperature

The melting temperatures in Fig. 6 were computed using the phase-coexistence method (*30*).

A total of 196,608 (fcc) and 200,000 (bcc) atoms were used to construct the initial crystal-liquid interface structure. At the center of the simulation box, 25% of the atoms formed an fcc slab ($16 \times 16 \times 16$ unit cells with its orthogonal axes oriented along the $[\bar{1}12]$, $[1\bar{1}1]$, and [110] directions). For the bcc structure, 67% of the atoms formed a bcc slab ($20 \times 28 \times 60$ unit cells with its orthogonal axes oriented along the $[\bar{1}10]$, [001], and [110] directions). The remaining 75%/33% of atoms, representing the liquid phase, were randomly distributed in the regions above and below the slab, occupying the same x-y cross-sectional area and distributed along the [110] normal. The simulation cell dimensions reflected the equilibrium lattice constant of the fcc or bcc phase and the average density of the liquid at the corresponding temperature. Periodic boundary conditions were applied in all three spatial dimensions.

The crystalliquid system was equilibrated through a multistage procedure. First, the liquid atoms were relaxed for 100 steps (using the conjugate gradient algorithm) whereas the atoms in the solid slab remained fixed. Subsequently, the liquid region was equilibrated for 2 ps at the target temperature. This was followed by a 2-ps equilibration of the solid region at 0.95× the target temperature. In the final stage, the liquid atoms were further equilibrated for 2 ps at the relevant temperature and zero pressure along the direction normal to the crystal-liquid interface. The Nosé-Hoover thermostat and/or barostat was used for all equilibration procedures described above.

The phase-coexistence simulations were conducted for 300 ps using a timestep of 2 fs. Temperature and pressure were controlled using a Nosé-Hoover thermostat and barostat with zero pressure maintained along the direction normal to the crystal-liquid interface, whereas the box dimensions parallel to the interface were held fixed. To prevent center-of-mass drift, the total linear momentum of the system was reset to zero every 2 ps, and temperature calculations excluded the center-of-mass velocity.

Simulations were performed at multiple temperatures, spaced by increments of 20 K around the estimated melting temperature. At each temperature, the interface velocity was determined by measuring the rate at which liquid atoms converted to solid atoms, using the polyhedral template matching (PTM) method (*105*, *106*) with a root mean square deviation cutoff of 0.15 to identify liquid and solid atoms. Following an initial 20-ps equilibration period, the number of liquid phase atoms was sampled every 2 ps. The melting temperature was identified as the temperature at which the net crystal-liquid interface velocity was zero.

For validation purposes, the predictions were compared against experimental data from refs. (*32*, *33*).

Warren-Cowley parameters

SRO was quantified using the Warren-Cowley parameters (*34*), defined as

$$\alpha_{AB} = 1 - \frac{p(A|B)}{c_A} \quad (7)$$

where $p(A|B)$ is the conditional probability of finding an A-type atom in the first nearest-neighbor shell of a B-type atom, and c_A is the concentration of A-type atoms in the system.

The Warren-Cowley parameters in Fig. 7 were obtained by thermally equilibrating solid solutions using Monte Carlo simulations (described in the Materials and Methods section "Free energy") as follows. Chemically random $10 \times 10 \times 10$ supercell structures of 4000 atoms were generated with lattice parameters that account for thermal expansion. Each simulation ran for 50 Monte Carlo steps per atom and 20 independent simulations were performed for each temperature and concentration. A total of 1000 configurations were saved evenly along the entire Monte Carlo trajectory, and the last 50 saved configurations of each simulation were used to compute the average Warren-Cowley parameters with the WarrenCowleyParameters OVITO modifier (*107*). The error bars in Fig. 7 represent the SEM, evaluated using 20 independent simulations.

For validation purposes, the predictions were compared against experimental data from refs. (*35*–*38*, *108*).

Thermal expansion

Lattice parameters in Fig. 8A were computed using molecular dynamics simulations with LAMMPS (*73*). For each composition, we generated random solid solutions with an fcc crystal structure in supercells of $20 \times 20 \times 20$ unit cells. The initial lattice parameter was set to 3.526 Å. The system was evolved for 100 ps in the NPT ensemble using a Nosé-Hoover thermostat and barostat with damping parameters of 0.1 and 1 ps, respectively. A 2.5-fs timestep was used, and periodic boundary conditions were applied in all dimensions. Simulations were performed at temperatures ranging from 0 to 1500 K

in increments of 100 K. The lattice parameters were computed from the time-averaged simulation cell dimensions, excluding the first half of each simulation to ensure equilibration.

For validation purposes, the predictions were compared against experimental data digitized from ref. (*109*).

Specific heat

The specific heat (c_p) in Fig. 8B was computed using molecular dynamics simulations performed with LAMMPS (*73*). Simulations were performed in the isothermal-isobaric (NPT) ensemble with a Nosé-Hoover thermostat and barostat, using a timestep of 0.001 ps. The system was maintained at zero pressure, with damping parameters of 1 ps for both the thermostat and barostat. At each temperature, the system was equilibrated for 30,000 timesteps (30 ps), and the enthalpy was recorded every 50 steps. The specific heat c_p at each temperature was determined via linear regression on the enthalpy versus temperature data, using a window of ±3 temperature points. Only the second half of each simulation was used in the analysis to ensure equilibration.

For validation purposes, the computational results were compared against experimental data digitized from ref. (*33*). In the case of pure tungsten, our results were benchmarked against the standard NIST data (*43*), which uses the Shomate equation

$$c_p = A + Bt + Ct^2 + Dt^3 + \frac{E}{t^2} \tag{8}$$

where $t = T/1000$ K, and the coefficients (A, B, C, D, and E) are temperature range dependent, specifically optimized for the intervals of 298 to 1900 K and 1900 to 3680 K.

Stacking-fault energy

The (111) SFE (γ_{sf} in Fig. 9) was computed as follows. First, thermally equilibrated solid solutions at 500 K were generated at each composition via Monte Carlo simulations (as described in the Materials and Methods section "Free energy"). Chemically random $14 \times 8 \times 6$ supercells of 4032 atoms were generated with coordinate axes aligned with the $[110]$, $[1\bar{1}2]$, and $[1\bar{1}\bar{1}]$ directions. After adjusting the lattice parameters for thermal expansion, each simulation ran for 50 Monte Carlo steps per atom. The periodic boundary conditions along the $[1\bar{1}\bar{1}]$ direction were then removed, and the system was relaxed by minimizing the potential energy at a fixed volume while holding the atoms in the top and bottom close-packed planes at their initial positions. An intrinsic stacking fault was introduced between two $(1\bar{1}\bar{1})$ planes, followed by a second relaxation under the same conditions. The SFE was determined by computing the energy difference between the relaxed structures before and after introducing the fault and dividing it by the cross-sectional area of the supercell in the $[1\bar{1}\bar{1}]$ direction. The energy values were calculated for all 18 $(1\bar{1}\bar{1})$ planes in the final configurations of each Monte Carlo simulation. γ_{sf} in Fig. 9 corresponds to the mean of these 18 values.

Supplementary Materials

This PDF file includes:

Supplementary Sections S1 to S16
Figs. S1 to S16
Table S1
References

REFERENCES

1. P. M. Anderson, J. P. Hirth, J. Lothe, *Theory of Dislocations* (Cambridge Univ. Press, 2017).
2. M. Neumann, J. Gin, B. Rhodes, S. Bennett, Z. Li, H. Choubisa, A. Hussey, J. Godwin, Orb: A fast, scalable neural network potential. arXiv:2410.22570 [cond-mat.mtrl-sci] (2024).
3. H. Yang, C. Hu, Y. Zhou, X. Liu, Y. Shi, J. Li, G. Li, Z. Chen, S. Chen, C. Zeni, M. Horton, R. Pinsler, A. Fowler, D. Zügner, T. Xie, J. Smith, L. Sun, Q. Wang, L. Kong, C. Liu, H. Hao, Z. Lu, Matter sim: A deep learning atomistic model across elements, temperatures and pressures. arXiv:2405.04967 [cond-mat.mtrl-sci] (2024).
4. H. Tang, Y. Zhang, Q.-J. Li, H. Xu, Y. Wang, Y. Wang, J. Li, High accuracy neural network interatomic potential for NiTi shape memory alloy. *Acta Mater.* **238**, 118217 (2022).
5. A. Thorn, D. Gochitashvili, S. Kharabadze, A. N. Kolmogorov, Machine learning search for stable binary Sn alloys with Na, Ca, Cu, Pd, and Ag. *Phys. Chem. Chem. Phys.* **25**, 22415–22436 (2023).
6. E. Fransson, J. Wiktor, P. Erhart, Phase transitions in inorganic halide perovskites from machine-learned potentials. *J. Phys. Chem. C* **127**, 13773–13781 (2023).
7. A. Seko, Machine learning potentials for multicomponent systems: The Ti-Al binary system. *Phys. Rev. B* **102**, 174104 (2020).
8. Y. Cao, K. Sheriff, R. Freitas, Capturing short-range order in high-entropy alloys with machine learning potentials. *npj Comput. Mater.* **11**, 268 (2025).
9. Y.-W. Zhang, V. Sorkin, Z. H. Aitken, A. Politano, J. Behler, A. Thompson, T. W. Ko, S. P. Ong, O. Chalykh, D. Korogod, E. Podryabinkin, A. V. Shapeev, J. Li, Y. Mishin, Z. Pei, X. Liu, J. Kim, Y. Park, S. Hwang, S. Han, K. Sheriff, Y. Cao, R. Freitas, Roadmap for the development of machine learning-based interatomic potentials. *Modelling Simul. Mater. Sci. Eng.* **33**, 063001 (2025).
10. K. Sheriff, Y. Cao, T. Smidt, R. Freitas, Quantifying chemical short-range order in metallic alloys. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2322962121 (2024).
11. K. Song, R. Zhao, J. Liu, Y. Wang, E. Lindgren, Y. Wang, S. Chen, K. Xu, T. Liang, P. Ying, N. Xu, Z. Zhao, J. Shi, J. Wang, S. Lyu, Z. Zeng, S. Liang, H. Dong, L. Sun, Y. Chen, Z. Zhang, W. Guo, P. Qian, J. Sun, P. Erhart, T. Ala-Nissila, Y. Su, Z. Fan, General purpose machine-learned potential for 16 elemental metals and their alloys. *Nat. Commun.* **15**, 4017 (2024).
12. H. Okamoto, M. E. Schlesinger, E. M. Mueller, Eds., *ASM Handbook, Volume 3: Alloy Phase Diagrams* (ASM International, 2016).
13. L. Han, S. Zhu, Z. Rao, C. Scheu, D. Ponge, A. Ludwig, H. Zhang, O. Gutfleisch, H. Hahn, Z. Li, D. Raabe, Multifunctional high-entropy materials. *Nat. Rev. Mater.* **9**, 846–865 (2024).
14. E. P. George, D. Raabe, R. O. Ritchie, High-entropy alloys. *Nat. Rev. Mater.* **4**, 515–534 (2019).
15. C. Oses, C. Toher, S. Curtarolo, High-entropy ceramics. *Nat. Rev. Mater.* **5**, 295–309 (2020).
16. K. Sheriff, Y. Cao, R. Freitas, Chemical-motif characterization of short-range order with E(3)-equivariant graph neural networks. *npj Comput. Mater.* **10**, 160 (2024).
17. M. R. K. Musa, Y. Qian, J. Peng, D. Cereceda, Accelerating the discovery of low-energy structure configurations: A computational approach that integrates first-principles calculations, Monte Carlo sampling, and machine learning. *Scr. Mater.* **259**, 116499 (2025).
18. M. Islam, K. Sheriff, Y. Cao, R. Freitas, Nonequilibrium chemical short-range order in metallic alloys. *Nat. Commun.* **16**, 10 (2025).
19. K. Sheriff, R. Freitas, A. Trewartha, S. Torrisi, "Simultaneous discovery of reaction coordinates and committor functions using equivariant graph neural networks," in *NeurIPS 2024 AI for Accelerated Materials Design* (NeurIPS, 2024).
20. A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, E. D. Cubuk, Scaling deep learning for materials discovery. *Nature* **624**, 80–85 (2023).
21. E. Podryabinkin, K. Garifullin, A. Shapeev, I. Novikov, MLIP-3: Active learning on atomic environments with Moment Tensor Potentials. arXiv:2311.10427 [cond-mat.mtrl-sci] (2023).
22. A. R. Tan, S. Urata, S. Goldman, J. C. B. Dietschreit, R. Gómez-Bombarelli, Single model uncertainty quantification in neural network potentials does not consistently outperform model ensembles. *npj Comput. Mater.* **9**, 169 (2023).
23. S. Pathrudkar, S. Taylor, A. Keripale, A. S. Gangan, P. Thiagarajan, S. Agarwal, J. Marian, S. Ghosh, A. S. Banerjee, Electronic structure prediction of medium and high entropy alloys across composition space. *npj Comput. Mater.* **11**, 378 (2025).
24. J. Vandermause, Y. Xie, J. S. Lim, C. J. Owen, B. Kozinsky, Active learning of reactive Bayesian force fields applied to heterogeneous catalysis dynamics of H/Pt. *Nat. Commun.* **13**, 3365 (2022).
25. I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovacs, J. Riebesell, X. R. Advincula, M. Asta, W. J. Baldwin, N. Bernstein, A. Bhowmik, S. M. Blau, V. Carare, J. P. Darby, S. De, F. D. Pia, V. L. Deringer, R. Elijosius, Z. E. Machachi, E. Fako, A. C. Ferrari, A. G. Schriever, J. George, R. E. A. Goodall, C. P. Grey, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I. B. Magdau, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, C. Oord, E. V. Umbrich, T. Vegge, M. Vondrak, Y. Wang, W. C. Witt, F. Zills, G. Csanyi, A foundation model for atomistic materials chemistry. arXiv:2312.15199 [physics.chem-ph] (2023).

26. A. Zunger, S. H. Wei, L. G. Ferreira, J. E. Bernard, Special quasirandom structures. *Phys. Rev. Lett.* **65**, 353–356 (1990).
27. A. van de Walle, P. Tiwary, M. de Jong, J. Olmsted, M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen, Z.-K. Liu, Efficient stochastic generation of special quasirandom structures. *Calphad* **42**, 13–18 (2013).
28. Y. Lysogorskiy, C. van der Oord, A. Bochkarev, S. Menon, M. Rinaldi, T. Hammerschmidt, M. Mrovec, A. Thompson, G. Csányi, R. Drautz, Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. *npj Comput. Mater.* **7**, 97 (2021).
29. C. Niu, Y. Rao, W. Windl, M. Ghazisaeidi, Multi-cell Monte Carlo method for phase prediction. *npj Comput. Mater.* **5**, 120 (2019).
30. J. R. Morris, C. Wang, K. Ho, C. T. Chan, Melting line of aluminum from simulations of coexisting phases. *Phys. Rev. B* **49**, 3109–3115 (1994).
31. B. Echebarria, R. Folch, A. Karma, M. Plapp, Quantitative phase-field model of alloy solidification. *Phys. Rev. E* **70**, 061604 (2004).
32. Z. Wu, H. Bei, G. M. Pharr, E. P. George, Temperature dependence of the mechanical properties of equiatomic solid solution alloys with face-centered cubic crystal structures. *Acta Mater.* **81**, 428–441 (2014).
33. I. H. Kim, H. S. Oh, K. S. Lee, E. S. Park, Optimization of conflicting properties via engineering compositional complexity in refractory high entropy alloys. *Scr. Mater.* **199**, 113839 (2021).
34. J. M. Cowley, An approximate theory of order in alloys. *Phys. Rev.* **77**, 669–675 (1950).
35. B. Schönfeld, L. Reinhard, G. Kostorz, W. Bührer, Short-range order and atomic displacements in Ni–20 at% Cr single crystals. *Phys. Status Solidi B* **148**, 457–471 (1988).
36. B. Schönfeld, G. E. Ice, C. J. Sparks, H.-G. Haubold, W. Schweika, L. B. Shaffer, X-ray study of diffuse scattering in Ni–20 at% Cr. *Phys. Status Solidi B* **183**, 79–95 (1994).
37. R. Caudron, M. Sarfati, A. Finel, F. Solal, In situ diffuse scattering of neutrons in alloys and application to phase diagram determination. *J. Phys. I France* **2**, 1145–1171 (1992).
38. L. Owen, H. Playford, H. Stone, M. Tucker, Analysis of short-range order in Cu_3Au using X-ray pair distribution functions. *Acta Mater.* **125**, 15–26 (2017).
39. Y. Zhang, G. Kresse, C. Wolverton, Nonlocal first-principles calculations in Cu-Au and other intermetallic alloys. *Phys. Rev. Lett.* **112**, 075502 (2014).
40. M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, K. A. Lyssenko, Density functional theory is straying from the path toward the exact functional. *Science* **355**, 49–52 (2017).
41. G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Constantin, J. Paier, J. Ángyán, Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B* **79**, 155107 (2009).
42. C. Zeni, K. Rossi, T. Pavloudis, J. Kioseoglou, S. de Gironcoli, R. E. Palmer, F. Baletto, Data-driven simulation and characterisation of gold nanoparticle melting. *Nat. Commun.* **12**, 6046 (2021).
43. M. Chase, *NIST-JANAF Thermochemical Tables* (American Institute of Physics, ed. 4, 1998).
44. H. S. Oh, S. J. Kim, J. S. K.-B. Park, H. J. Chang, S. I. Baik, E. S. Park, Composition-dependent transformation-induced plasticity in Co-based complex concentrated alloys. *Acta Mater.* **262**, 119349 (2024).
45. D. Liu, Y. Yu, S. Yin, F. Shuang, P. Zhang, T. Lei, T. Jiang, R. O. Ritchie, R. Yu, Exceptional fracture toughness of CrCoNi-based medium- and high-entropy alloys at 20 kelvin. *Science* **378**, 978–983 (2022).
46. M. Shih, J. Miao, M. Mills, M. Ghazisaeidi, Stacking fault energy in concentrated alloys. *Nat. Commun.* **12**, 3574 (2021).
47. J. Ding, Q. Yu, M. Asta, R. O. Ritchie, Tunable stacking fault energies by tailoring local chemical order in CrCoNi medium-entropy alloys. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8919–8924 (2018).
48. X.-G. Li, C. Chen, H. Zheng, Y. Zuo, S. P. Ong, Complex strengthening mechanisms in the NbMoTaW multi-principal element alloy. *npj Comput. Mater.* **6**, 70 (2020).
49. H. Zheng, T. Nguyen, C. Wang, Y. Zuo, S. P. Ong, Multi-scale investigation of short-range order and dislocation glide in MoNbTi and TaNbTi multi-principal element alloys. *npj Comput. Mater.* **9**, 89 (2023).
50. S. Yin, Y. Zuo, A. Abu-Odeh, H. Zheng, X.-G. Li, J. Ding, S. P. Ong, M. Asta, R. O. Ritchie, Atomistic simulations of dislocation mobility in refractory high-entropy alloys and the effect of chemical short-range order. *Nat. Commun.* **12**, 4873 (2021).
51. L. Barroso-Luque, M. Shuaibi, X. Fu, B. M. Wood, M. Dzamba, M. Gao, A. Rizvi, C. L. Zitnick, Z. W. Ulissi, Open Materials 2024 (OMat24) inorganic materials dataset and models. arXiv:2406.11306 [cond-mat.mtrl-sci] (2024).
52. F. Shuang, Z. Wei, K. Liu, W. Gao, P. Dey, Universal machine learning interatomic potentials poised to supplant dft in modeling general defects in metals and random alloys. *Acta Mater.* **284**, 120530 (2025).
53. S. Zhu, D. Sarıtürk, R. Arróyave, Accelerating CALPHAD-based phase diagram predictions in complex alloys using universal machine learning potentials: Opportunities and challenges. *Acta Mater.* **286**, 120747 (2025).
54. D. S. Levine, M. Shuaibi, E. W. C. Spotte-Smith, M. G. Taylor, M. R. Hasyim, K. Michel, I. Batatia, G. Csányi, M. Dzamba, P. Eastman, N. C. Frey, X. Fu, V. Gharakhanyan, A. S. Krishnapriyan, J. A. Rackers, S. Raja, A. Rizvi, A. S. Rosen, Z. Ulissi, S. Vargas, C. L. Zitnick, S. M. Blau, B. M. Wood, The Open Molecules 2025 (OMol25) dataset, evaluations, and models. arXiv:2501.07720 [physics.chem-ph] (2025).
55. C. Zeni, R. Pinsler, D. Zügner, A. Fowler, M. Horton, T. Xie, J. Smith, H. Yang, H. Hao, L. Sun, Q. Wang, L. Kong, C. Liu, J. Schrier, B. Wood, T. S. Miller, K. Huang, Y. Zhou, C. Hu, X. Liu, Y. Shi, G. Li, S. Chen, J. Li, Z. Chen, Z. Lu, A generative model for inorganic materials design. *Nature* **639**, 624–632 (2025).
56. J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, J. Sun, Accurate and numerically efficient r^2SCAN meta-generalized gradient approximation. *J. Phys. Chem. Lett.* **11**, 8208–8215 (2020).
57. K. Bystrom, B. Kozinsky, CIDER: An expressive, nonlocal feature set for machine learning density functionals with exact constraints. *J. Chem. Theory Comput.* **18**, 2180–2192 (2022).
58. K. Bystrom, S. Falletta, B. Kozinsky, Training machine-learned density functionals on band gaps. *J. Chem. Theory Comput.* **20**, 7516–7532 (2024).
59. R. Akashi, M. Sogal, K. Burke, Can machines learn density functionals? Past, present, and future of ML in DFT. arXiv:2502.03541 [physics.chem-ph] (2025).
60. M. S. Chen, J. R. Gissinger, R. K. Lindsey, L. E. Smeeton, N. Goldman, Data-efficient machine learning potentials from transfer learning of periodic correlated electronic structure methods: Liquid water at AFQMC, CCSD, and CCSD(T) accuracy. *J. Chem. Theory Comput.* **19**, 4510–4519 (2023).
61. R. L. McGreevy, Reverse Monte Carlo modelling. *J. Phys. Condens. Matter* **13**, R877–R913 (2001).
62. H. W. Chung, R. Freitas, G. Cheon, E. J. Reed, Data-centric framework for crystal structure identification in atomistic simulations using machine learning. *Phys. Rev. Mater.* **6**, 043801 (2022).
63. G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
64. G. Kresse, J. Furthmüller, Efficient iterative schemes forab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
65. G. Kresse, J. Furthmüller, Efficiency of ab initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
66. G. Kresse, J. Hafner, Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251–14269 (1994).
67. G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
68. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
69. P. E. Blöchl, Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
70. D. M. Ceperley, B. J. Alder, Ground State of the electron gas by a stochastic method. *Phys. Rev. Lett.* **45**, 566–569 (1980).
71. J. P. Perdew, A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048–5079 (1981).
72. R. Freitas, M. Asta, M. de Koning, Nonequilibrium free-energy calculation of solids using LAMMPS. *Comput. Mater. Sci.* **112**, 333–341 (2016).
73. S. Plimpton, Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
74. R. Kikuchi, CVM entropy algebra. *Prog. Theor. Phys. Suppl.* **115**, 1–26 (1994).
75. M. Ghazisaeidi, Alloy thermodynamics via the Multi-cell Monte Carlo $(MC)^2$ method. *Comput. Mater. Sci.* **193**, 110322 (2021).
76. E. Antillon, M. Ghazisaeidi, Efficient determination of solid-state phase equilibrium with the multicell Monte Carlo method. *Phys. Rev. E* **101**, 063306 (2020).
77. P. Ungerer, A. Wender, G. Demoulin, B. Bourasseau, P. Mougin, Application of Gibbs ensemble and NPT Monte Carlo simulation to the development of improved processes for H_2S-rich gases. *Mol. Simul.* **30**, 631–648 (2004).
78. C. Jenkins, E. Bucknall, C. Austin, G. Mellor, Some alloys for use at high temperatures. Part IV. *J. Iron Steel Inst.* **136**, 187–222 (1937).
79. V. N. Svechnikov, V. M. Pan, Characteristics of the equilibrium diagram and processes of solution and precipitation in the Cr-Ni system. *Sb. Nauchn. Rab. Inst. Metallofiz. Akad. Nauk Ukr. SSR* **15**, 164–178 (1962). [In Russian]
80. W. Dench, Adiabatic high-temperature calorimeter for the measurement of heats of alloying. *Trans. Faraday Soc.* **59**, 1279–1290 (1963).
81. E. Jette, V. Nordstrom, B. Queneau, F. Foote, X-ray studies on the nickel-chromium system. *Trans. Am. Inst. Min. Metall. Eng.* **111**, 361–373 (1934).
82. A. Taylor, R. Floyd, The constitution of nickel rich alloys of the Ni-Cr-Ti system. *J. Inst. Metals* **80**, 577–587 (1951).
83. L. Pugliese, G. Fitterer, Activities and phase boundaries in the Cr-Ni system using a solid electrolyte technique. *Metall. Trans.* **1**, 1997–2002 (1970).
84. L. Karmazin, Remarks on the ordering in the Ni-Cr alloys in the vicinity of Ni_2Cr. *Czech. J. Phys. B* **28**, 1065–1068 (1978).
85. H. G. Baer, Superstructure and K-state in the Cr-Ni system. *Z. Metallkd.* **49**, 614–622 (1958). [In German]

86. P. Nash, The Cr-Ni (chromium-nickel) system. *Bull. Alloy Phase Diagrams* **7**, 466–476 (1986).
87. F. Wever, U. Hashimoto, On the properties of the Co-Cr binary system. *Mitt. Kaiser-Wilhelm-Inst. Eisenforsch. Düsseldorf* **11**, 293–330 (1929). [In German]
88. Y. Matsunaga, On the equilibrium diagram of the cobalt-chromium system. *Kinzoku-no-Kenkyu* **8**, 549–564 (1931). [In Japanese]
89. A. Chiba, "Mobility of interphase boundary in metals," thesis, Tohoku University, Sendai, Japan (1971).
90. C. Allibert, C. Bernard, N. Valignat, M. Dombre, Co-Cr binary system: Experimental re-determination of the phase diagram and comparison with the diagram calculated from the thermodynamic data. *J. Less-Common Metals* **59**, 211–228 (1978).
91. Z. Jin, A study of the range of stability of σ phase in some ternary systems. *Scand. J. Metall.* **10**, 279–287 (1981).
92. M. Hasebe, K. Oikawa, T. Nishizawa, Computer calculation of phase diagrams of Co-Cr and Co-Mn systems. *J. Japan Inst. Metals* **46**, 577–583 (1982). [In Japanese]
93. K. Ishida, T. Nishizawa, The Co-Cr (cobalt-chromium) system. *Bull. Alloy Phase Diagrams* **11**, 357–370 (1990).
94. C. H. Johansson, J. O. Linde, Kristallstruktur, elektrischer Widerstand, Thermokräfte, Wärmeleitfähigkeit, magnetische Suszeptibilität, Härte und Vergütungserscheinungen des Systems Au Pt in Verbindung mit dem Zustandsdiagramm. *Ann. Phys.* **397**, 762–792 (1930). [In German]
95. W. Stenzel, J. Weerts, "X-ray examination of alloys of the gold-platinum system" in *Festschrift zum 50-jährigen Bestehen der Platinschmelze* (G. Siebert G.m.b.H., 1931) [In German].
96. A. Darling, R. Mintern, J. Chaston, The gold-platinum system. *J. Inst. Metals* **81**, 125–132 (1952).
97. T. Tiedema, J. Bouman, W. Burgers, Precipitation in gold-platinum alloys. *Acta Metall.* **5**, 310–321 (1957).
98. E. Raub, G. Falkenburg, The system gold-platinum-rhodium and the binary systems of its components. *Z. Metallkd.* **55**, 392–397 (1964). [In German]
99. C. G. Wictorin, Two-phase boundary of the Au-Pt system. *Ark. Mat. Astron. Fys. B* **36**, 509–515 (1949).
100. A. Münster, K. Sagel, Entmischungskurve und kritischer Punkt des Systems gold—platin. *Z. Phys. Chem.* **23**, 415–425 (1960). [In German]
101. H. Okamoto, T. B. Massalski, The Au-Pt (gold-platinum) system. *Bull. Alloy Phase Diagrams* **6**, 46–56 (1985).
102. P. Turchi, L. Kaufman, Z. Liu, Modeling of Ni-Cr-Mo based alloys: Part I—Phase stability. *Calphad* **30**, 70–87 (2006).
103. H. Okamoto, Co-Cr (cobalt-chromium). *J. Phase Equilib.* **24**, 377–378 (2003).
104. H. Okamoto, T. B. Massalski, "Au-Pt (gold-platinum)" in *Binary Alloy Phase Diagrams*, T. B. Massalski, Ed. (ASM International, ed. 2, 1990).
105. P. M. Larsen, S. Schmidt, J. Schiøtz, Robust structural identification via polyhedral template matching. *Modelling Simul. Mater. Sci. Eng.* **24**, 055007 (2016).
106. A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—The Open Visualization Tool. *Modelling Simul. Mater. Sci. Eng.* **18**, 015012 (2010).
107. K. Sheriff, R. Freitas, WarrenCowleyParameters, an OVITO python modifier to compute the Warren-Cowley parameters (2024). https://github.com/killiansheriff/WarrenCowleyParameters.
108. F. Walsh, A. Abu-Odeh, M. Asta, Reconsidering short-range order in complex concentrated alloys. *MRS Bull.* **48**, 753–761 (2023).
109. K. Jin, S. Mu, K. An, W. D. Porter, G. Samolyuk, G. M. Stocks, B. C. Sales, Thermophysical properties of Ni-containing single-phase concentrated solid solution alloys. *Mater. Des.* **117**, 185–192 (2017).
110. K. Sheriff, killiansheriff/ChemicalMotifIdentifier, version v_0.0.10, Zenodo (2025); https://doi.org/10.5281/zenodo.16129603.
111. K. Sheriff, Ridul, killiansheriff/Machine-learning-potentials-for-modeling-alloys-across-compositions, v1.0.0, Zenodo (2026); https://doi.org/10.5281/zenodo.20010513.
112. K. Sheriff, LovelyPlots, a collection of Matplotlib style sheets to nicely format figures for scientific papers, thesis and presentations, v_1.0.2, Zenodo (2022); https://doi.org/10.5281/zenodo.10784160.
113. Q. J. Li, H. Sheng, E. Ma, Strengthening in multi-principal element alloys with local-chemical order roughened dislocation pathways. *Nat. Commun.* **10**, 3563 (2019).
114. C. J. O'Brien, C. M. Barr, P. M. Price, K. Hattar, S. M. Foiles, Grain boundary phase transformations in PtAu and relevance to thermal stabilization of bulk nanocrystalline metals. *J. Mater. Sci.* **53**, 2911–2927 (2018).

Acknowledgments

Funding: This work was supported by the MathWorks Ignition Fund, MathWorks Engineering Fellowship Fund, and the Portuguese Foundation for International Cooperation in Science, Technology and Higher Education in the MIT–Portugal Program. This material is based on work supported by the Air Force Office of Scientific Research (AFOSR) under award number FA9550-25-1-0199, through the Young Investigator Program. L.R.O. would like to thank the Royal Academy of Engineering (RAEng) for ongoing support during his research fellowship. The Cu3Au data were obtained through provision of beam time at Diamond Light Source Ltd. (EE11665). K.S. acknowledge the support of the Exponent Fellowship. D.Z.X. was supported by the National Science Foundation Graduate Research Fellowship Program. This material is based on work supported by the National Science Foundation Graduate Research Fellowship Program under grant no. DGE-2141064. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work used the Expanse supercomputer at the San Diego Supercomputer Center through allocation MAT210005 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by the National Science Foundation grant nos. 2138259, 2138286, 2138307, 2137603, and 2138296. **Author contributions:** K.S. and R.F. conceived the project. K.S. developed all methodologies and carried out all calculations, with the exception of the multicell Monte Carlo simulations, which were performed by D.Z.X. Y.C. helped K.S. with the melting point and SFE calculations. K.S., D.Z.X., Y.C., L.R.O., and R.F. contributed to the interpretation of the results. K.S. and R.F. prepared the manuscript, which was reviewed and edited by K.S., D.Z.X., Y.C., L.R.O., and R.F. Project administration, supervision, and funding acquisition was performed by R.F. **Competing interests:** The authors declare that they have no competing interests. **Data, code, and materials availability:** All data and code needed to evaluate and reproduce the results in the paper are present in the paper and/or the Supplementary Materials. The software for SRO quantification can be found in our ChemicalMotifIdentifier Python package (*110*) (https://github.com/killiansheriff/ChemicalMotifIdentifier; https://doi.org/10.5281/zenodo.16129603). The MLPs and datasets can be found in our GitHub repository (*111*) (https://github.com/killiansheriff/Machine-learning-potentials-for-modeling-alloys-across-compositions; https://doi.org/10.5281/zenodo.20010513). Our figure style is implemented in LovelyPlots (*112*) (https://github.com/killiansheriff/LovelyPlots; https://doi.org/10.5281/zenodo.10784160) under the paper style. This study did not generate new materials.

Submitted 28 July 2025
Accepted 5 May 2026
Published 19 June 2026
10.1126/sciadv.aea9951

ScienceAdvances

Machine learning potentials for modeling alloys across compositions

Killian Sheriff, Daniel Z. Xiao, Yifan Cao, Lewis R. Owen, and Rodrigo Freitas

Sci. Adv. **12** (25), eaea9951. DOI: 10.1126/sciadv.aea9951

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.