Data-Driven Prediction of Grain Boundary Segregation and Disordering in High-Entropy Alloys in a 5D Space

Journal:	*Materials Horizons*
Manuscript ID	MH-COM-07-2021-001204.R1
Article Type:	Communication
Date Submitted by the Author:	02-Dec-2021
Complete List of Authors:	Hu, Chongze; UCSD Luo, Jian; UCSD

New Concepts:

The data-driven prediction of grain boundary (GB) properties in high-entropy alloys (HEAs) as functions of four independent compositional degrees of freedom and temperature in a 5D space opens a new paradigm. The interaction of multiple elements and interfacial disordering can induce a new region of segregation not predicted by the classical theory. A significant data-driven discovery is the existence of a critical (isoequilbrium) temperature with the simultaneous vanishing of GB segregation due to an enthalpy-entropy compensation effect. A physics-informed data-driven model can predict GB properties based on parameters with clear physical meanings. This study expands our fundamental knowledge of GB segregation and paves the way for tailoring properties of HEAs via controlling GBs.

Data-Driven Prediction of Grain Boundary Segregation and Disordering in High-Entropy Alloys in a 5D Space

Chongze Hu[1, 2, #] and Jian Luo[1, 2, *]

¹Department of Nanoengineering; ²Program of Materials Science and Engineering, University of California San Diego, La Jolla, California 92093, USA

*Corresponding author. E-mail: jluo@alum.mit.edu (J.L.)

#Current address: Center for Integrated Nanotechnologies, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

Abstract

Grain boundaries (GBs) can critically influence the microstructural evolution and various materials properties. However, a fundamental understanding of GBs in high-entropy alloys (HEAs) is lacking because of the complex couplings of the segregations of multiple elements and interfacial disordering, which can generate new phenomena and challenge the classical theories. Here, by combining large-scale atomistic simulations and machine learning models, we demonstrate the feasibility of predicting the GB properties as functions of four independent compositional degrees of freedoms and temperature in a 5D space, thereby enabling the construction of GB diagrams for quinary HEAs. Artificial neural network (ANN), support vector machine (SVM), regression tree, and rational quadratic Gaussian models are trained and tested, and the ANN model yields the best machine learning based predictions. A data-driven discovery further reveals new coupled segregation and disordering effects in HEAs. For instance, interfacial disordering can enhance the co-segregation of Cr and Mn at CrMnFeCoNi GBs. A physics-informed data-driven model is constructed to provide more physical insights and better transferability. Density functional theory (DFT) calculations are used to validate the prediction generality and reveal underlying segregation mechanisms. This study not only provides a new paradigm enabling the prediction of GB properties in a 5D space, but also uncovers new GB segregation phenomena in HEAs beyond the classical GB segregation models.

Introduction

Since the Bronze and Steel Ages, the development of every major class of metallic alloys, such as the Cu, Fe, Al, Ti, and Ni-based alloys, have revolutionized technologies and changed our daily lives. High-entropy alloys (HEAs), also known as multi-principal element alloys (MPEAs) or complex concentrated alloys (CCA), represent the newest class of alloys that attract significant research interest.[1-4] The vast composition space of HEAs offers immense opportunities for designing materials for various applications.

In every class of polycrystalline alloys, grain boundaries (GBs) exist ubiquitously.[5, 6] The elemental segregation (*a.k.a.* adsorption in the interfacial thermodynamics) at GBs is a critical phenomenon that can change microstructural evolution[7-9] and govern a broad range of materials properties[10-12]. Even though the GB segregation have been extensively researched for decades, most prior studies and models are based on alloys with one primary (principal) element.[13-20] Moreover, the effects of interfacial disordering on segregation are typically not considered in the classical site-occupying models.[13-15] The GB segregation in the emerging HEAs containing five or more principal elements are hitherto only investigated by few experimental[21, 22] and theoretical[23-25] studies and only for a few compositions. The underlying mechanisms of GB segregation in HEAs are elusive and a predictive model does not exist, which motivate this study.

In a broader perspective, GBs can be considered as two-dimensional (2D) interfacial phases [6], which are also named as "complexions" [26, 27] to differentiate them from thin precipitated layers of 3D bulk phases. Notably, GB complexion (phase) diagrams, which represent GB thermodynamic states or properties as functions of thermodynamic variables such as temperature and bulk composition (representing chemical potentials), have been developed as the GB counterparts to bulk phase diagrams. To date, various GB diagrams have been constructed for binary and ternary systems,[19, 20, 26, 28-30] but they are rarely developed for multicomponent systems,[31] certainly not for HEAs, owing to the increasing complexity of a large, multi-dimensional compositional space. Furthermore, the more general GBs (asymmetric GBs with mixed twist and tilt features), which are ubiquitous in polycrystalline materials and often the weak links chemically and mechanically,[28, 32] are still scarcely studied.

Herein, by combining the large-scale hybrid Monte Carlo and molecular dynamics (MC/MD) simulations and an artificial neutral network (ANN) model, we demonstrate the feasibility of predicting the GB properties as functions of four independent compositional DOFs and temperature in a 5D space for a representative general GB in $Cr_xMn_yFe_zCo_lNi_m$ HEAs. In summary, 1032 individual hybrid MC/MD simulations were performed (for 258 compositions × 4 different temperatures). Each was formed on an asymmetric Σ81 GB with a large simulation cell of 11,664 atoms for ~10^5 hybrid MC/MD simulation steps to achieve convergence. This produced by far the largest dataset to enable further data-driven analysis.

Our MC/MD simulations further reveal the unrecognized importance of interfacial disorder in influencing GB segregation that can produce new phenomena in HEAs. Notably, an analysis of

the large dataset uncovers the almost (albeit incomplete) vanishing of GB segregation/depletion of all elements simultaneously at a critical temperature.

Furthermore, a physics-informed data-driven model (PIDDM) is created to represent GB segregation and disordering in HEAs analytically. While the machine learning model can make predictions via a black-box approach, this PIDDM describes the GB properties in a 5D space via analytical formulae where all parameters have clear physical meanings, and it offers better transferability.

The Workflow

Figure 1 displays the workflow of predicting GB properties and GB diagrams, investigating the underlying mechanisms, uncovering new interfacial phenomena, and developing a PIDDM via combined large-scale hybrid MC/MD simulations, machine learning models, various data analysis approaches, and density functional theory (DFT) calculations.

First, we selected an asymmetric Σ81 GB of mixed tilt and twist characters to represent general GBs. We randomly generated 258 compositions out of 1371 possible choices by varying the amount of each element from 5% to 35% with a step of 5%. Second, principal components analysis (PCA) was used to ensure the 258 selected compositions are sufficiently randomly (Supplementary Discussion 1). Third, the large-scale isothermal-isobaric (constant-NPT) ensemble hybrid MC/MD simulations were carried out to calculate the adsorption amounts (*i.e.*, GB excesses of solutes: Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, Γ_{Co}, Γ_{Ni}), GB excess of disorder (Γ_{Dis}), and GB free volume (V_{Free}), as well as bulk composition. Subsequently, we mapped out a relation between bulk composition and GB properties, from 1000 K to 1300 K (where CrMnFeCoNi alloys can maintain a single face-centered cubic or FCC phase[33]). Here, the GB excess of disorder Γ_{Dis} was computed based on: $\Gamma_{Dis} = \int_{-\infty}^{+\infty}[\eta_{Dis}(z) - \eta_{Dis}(\pm\infty)]dz$, where the η_{Dis} (= 1 for a liquid and = 0 for a perfect crystal) was calculated following a procedure proposed by Chua *et al.*.[34] Other GB excess qualities are defined based on the standard Gibbs definitions. See Method section for details.

Fourth, the MC/MD-simulated dataset was used to develop an ANN model, along with support vector machine (SVM), regression tree, and rational quadratic Gaussian models, to predict six important GB properties (Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, Γ_{Co}, Γ_{Ni}, and Γ_{Dis}), each in a 5D space. Among the four different machine learning models trained and tested, the ANN model yielded the best predictions (Fig. 2J) so it was selected for the subsequent machine learning predictions. Fifth, GB diagrams of thermodynamic properties were constructed for the first time for quinary HEAs; as an illustrative example, Fig. 1E shows isothermal sections of Γ_{Cr} for $Cr_xMn_yFe_zCo_{0.2}Ni_{0.2}$ subsystem, where $x + y + z = 0.6$.

Sixth, a large dataset generated by hybrid MC/MD simulations was also used to analyze and investigate the new coupled interfacial disordering and segregation phenomena in HEAs beyond the classical models via a data-driven discovery approach. Seventh, additional hybrid MC/MD simulations were conducted for other GBs to show the generality of our findings. Eighth, DFT

calculations were conducted, based on the GB configurations obtained by hybrid MC/MD simulations, to investigate the unique segregation mechanisms in HEAs. Ninth, a physics-informed data-driven model or PIDDM was developed to predict the GB segregation and disordering to provide more physics insights with better transferability than the "black-box" ANN model, which can also achieve good accuracies.

It is worth noting that before we conducted large-scale hybrid MC/MD simulations, we first performed careful benchmark simulations to validate our *NPT*-based hybrid MC/MD method by comparing it with prior *NVT*-based MC simulations[24] as well as prior experiments[21, 22, 35]; see Supplementary Discussion 2.

ANN prediction of GB diagrams of thermodynamic properties

The dataset generated by 1032 MC/MD simulations have been used to train, evaluate, and test one-layer single-task ANN models to predict six GB properties (Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, Γ_{Co}, Γ_{Ni}, and Γ_{Dis}). The histogram of root-mean-square errors (RMSEs) was used to assess the ANN performance. Notably, the ANN models are fairly accurate to predict six GB properties with small RMSEs (Suppl. Fig. S4). The accuracies are further supported by the parity plots between ANN predictions and MC/MD simulations, where the promising linear relations are achieved for Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, and Γ_{Co} (see Fig. 2A and Suppl. Fig. S3(B)). Relatively large deviations are found for Γ_{Ni} and Γ_{Dis}, which can be ascribed to the weak segregation of Ni and large uncertainty in quantifying Γ_{Dis}. Overall, the ANN models are robust to predict GB properties, especially for moderate and strong segregation (*e.g.*, Cr and Mn) and depletion elements (*e.g.*, Fe and Co) at HEA GBs.

In addition to the ANN model, we also trained and tested SVM, regression tree, and rational quadratic Gaussian models as alternative machine learning models to compare their performances with the ANN model. We found that the ANN model yielded the lowest RMSE values for predicting all six GB properties (Fig. 2J). Thus, we adopted the ANN model as our machine learning engine for the data-driven discovery of GB properties.

To further validate our ANN model, we adopt a structural similarity index (SSIM; 1 = same and 0 = different) to compare the similarity of ANN-predicted binary GB diagrams *vs.* MC/MD simulations; representative GB diagrams are shown in Fig. 2B-C. The SSIM histogram (Suppl. Fig. S5) shows the high values (~0.88-0.89) for most Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, and Γ_{Co} diagrams, but relatively low values (~0.63-0.66) for Γ_{Ni} and Γ_{Dis} diagrams. This is consistent with the prior analysis based on RMSEs. See elaboration in Supplementary Discussion 3.

Interestingly, ANN predictions can outperform MC/MD simulations in two aspects. First, the ANN model can suppress the MC/MD errors caused by the large thermal noises at high temperatures (Suppl. Fig. S6) by a smoothing effect. Second, the ANN model becomes more convenient than MC/MD simulations to predict GB diagrams with multiple variables. For example, Figs. 2D-I show the ANN-predicted ternary GB segregation and disorder diagrams in $Cr_xMn_{0.2}Fe_yCo_{0.2}Ni_z$ (where $x + y + z = 0.6$) at 1000 K. More ANN-predicted GB diagrams can

be found in Suppl. Figs. S23-S29.

Notably, the efficient ANN model makes it possible to map out GB properties as functions of four independent compositional DOFs and temperature in a 5D space for HEAs.

New GB segregation phenomena in HEAs and interpreting an experimental observation

Beyond the ANN model prediction, we further conducted a series of in-depth data analyses, along with additional focused simulations, to elucidate new interfacial phenomena that are unique to HEAs. As the first example to illustrate new GB segregation phenomena in HEAs, we have conducted and analyzed MC/MD simulations of nine representative equimolar ternary (medium-entropy) to quinary (high-entropy) alloys, including: FeCoNi, CrMnNi, CrMnFe, CrFeNi, CrCoFe, CrFeCoNi, CrMnFeNi, CrMnFeCo, and CrMnFeCoNi. New observations of GB disorder enhanced co-segregation are documented in Supplementary Discussion 4 and discussed below.

On the one hand, we find that the competition of segregation sites in relatively ordered GBs can suppress the (*e.g.*, Cr) segregation in ternary alloys (*e.g.*, CrMnNi and CrMnFe). On the other hand, the segregation of multiple elements and GB disordering can enhance each other to lead to stronger co-segregation at more disordered GBs in quaternary and quinary alloys (*e.g.*, co-segregation of Cr and Mn with high levels of GB disorder in CrMnFeCo and CrMnFeCoNi).

Two representative cases, CrMnNi *vs.* CrMnFeCoNi, are schematically illustrated in Fig. 3A *vs.* 3B, and additional examples are shown in Suppl. Fig. S7. Specifically, the MC/MD-simulated GB structure of the CrMnNi ternary alloy at 1000 K (Fig. 3C) shows that the relatively ordered GB (Γ_{Dis} of ~39 nm^{-2}) has weak segregation of Cr and virtually no segregation of Mn (Γ_{Cr} = 5.3 nm^{-2} and Γ_{Mn} = 0.8 nm^{-2}). However, the more disordered GB (Γ_{Dis} of ~43 nm^{-2}) in CrMnFeCoNi exhibits strong co-segregation of Cr and Mn (Γ_{Cr} = 18.6 nm^{-2} and Γ_{Mn} = 7.0 nm^{-2}). The compositional profiles shown in Fig. 3C-D confirm the strong GB segregation of both Cr and Mn in CrMnFeCoNi, but weak GB segregation in CrMnNi. Moreover, the computed profile of the disorder parameter also verifies a more disordered GB core of ~0.88 nm thick in CrMnFeCoNi *vs.* a less disordered GB core of ~0.75 nm thick in CrMnNi.

It is interesting to further note that for the systems without Mn element (*e.g.*, CrCoFe and CrFeCoNi), Cr atoms are not favorable to segregate at relatively ordered GBs (Suppl. Figs. S7-8). The structural analysis based on polyhedral template matching (PTM) approach[36] shows that Mn segregation can induce GB disordering and hexagonal structures locally (Suppl. Fig. S8). This implies that the Mn segregation can change the interaction between Cr and GBs to further induce the Cr segregation. A similar phenomenon has been observed and explained in terms of chemical affinity.[37] As an alternative view, Mn segregation can induce GB disordering and subsequently alter the interaction between Cr and GBs to enhance the Cr segregation. In contrast, the weak Cr segregation in absent of Mn can be attributed to relatively ordered GB structures.

Notably, the abovementioned phenomena are not only consistent with the finding in nanocrystalline Al alloys reported by Babicheva *et al.*,[38] but also applicable to other ternary,

quaternary, and quinary HEA systems studied in this work (Suppl. Figs. S7-8), thereby suggesting the generality of the coupling effects. See Supplementary Discussion 4 for more results.

Using atom probe tomography, Li *et al.* observed GB spinodal decomposition that formed Cr-rich and Mn-rich regions.[35] The new insight obtained by modeling here offers an explanation of this observation. Since GB disorder (more significant at high temperatures) promotes the co-segregation of Cr and Mn, it could separate into Cr-rich and Mn-rich GB regions upon cooling with less GB disorder.

Analysis of the large dataset of GB thermodynamic properties

To achieve more in-depth understanding of couplings among GB properties, we further calculated Pearson correlation coefficients (PCCs) among five GB adsorption (*i.e.*, Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, Γ_{Co}, and Γ_{Ni}) and two structural properties (Γ_{Dis} and V_{Free}) based on the MC/MD-simulated dataset. Notably, the heat map of PCC shows that GB disorder is correlated with GB adsorption properties (Figs. 4A-B). Specifically, the segregation of Mn (Γ_{Mn}) has the strongest correlation with GB disorder (Γ_{Dis}) among all elements, which agree with the MC/MD simulations showing that Mn segregation can induce large GB disordering. In addition, by calculating the PCCs at different temperatures, we found that the correlations between GB disorder and adsorption properties decrease with increasing temperature, while the correlations between other GB properties remain almost unchanged. This analysis (Supplementary Discussion 5) further verifies the importance of interfacial disordering on GB segregation in HEAs. Note that no abrupt (first-order) GB transition with temperature or composition was observed in the current case.

It is generally expected that GB disorder may increase free volume in unary and binary alloys. However, we found that GB free volume (V_{Free}) is small and has almost no correlation with GB disorder or other GB properties in HEAs (Fig. 4B). This may be ascribed to the packing of multiple elements of different radii at HEA GBs (as a rather unique character of HEAs).

The correlations of GB segregations of different elements in HEAs also show interesting trends and suggest new interfacial phenomena. On one hand, GB disordering can promote the segregations of both Cr and Mn (*i.e.*, both Γ_{Cr} and Γ_{Mn} are positively correlated with Γ_{Dis}; see Fig. 4A). Consequently, Γ_{Cr} and Γ_{Mn} are positively correlated (Fig. 4A), despite that a positive Cr-Mn mixing enthalpy (Suppl. Table S3) suggests repulsion between them. On the other hands, a significant negative Co-Cr mixing enthalpy (Suppl. Table S3) suggests that they should attract one another in the bulk phase, but their GB adsorptions are negatively correlated (Fig. 4A) because Γ_{Cr} is positively, but Γ_{Co} is negatively, correlated with Γ_{Dis} (Fig. 4A). These findings again suggest the critical role of interfacial disordering in influencing GB segregation in HEAs.

Next, we examine the correlation of Γ_i *vs.* Γ_{Dis} at different temperatures. Linear correlation analysis (Fig. 4C and Suppl. Fig. S10) shows the following statistical correlation:

$$\Gamma_i(T,X) - \Gamma_i^0 = \bar{\alpha}_{Dis}^i(T) \cdot \left[\Gamma_{Dis}(T,X) - \Gamma_{Dis}^0\right], \tag{1}$$

where T is temperature, $X = \{X_i\}$ is the bulk composition, $(\Gamma_i^0, \Gamma_{Dis}^0)$ is the intersection point of all linear regression lines that is virtually independent of temperature, and $\bar{\alpha}_{\text{Dis}}^i(T)$ is the linear correlation coefficient. Here, we use an overbar on $\bar{\alpha}_{\text{Dis}}^i$ to denote that it is a statistical average value. We do not suggest a simple linear relation between Γ_i and Γ_{Dis} (and we will subsequently show that $\Gamma_i^0 = \Gamma_i^0(X)$ is not a constant, but a function of the composition $X = \{X_i\}$ for any individual alloy). Scattered data points in Fig. 4C can be ascribed to 258 different alloy compositions. However, a hidden statistical correlation exists after averaging over the 258 random compositions. Notably, excellent linear correlations exist for $\bar{\alpha}_{\text{Dis}}^i$ vs. T for all elements (Fig. 4D):

$$\bar{\alpha}_{\text{Dis}}^i(T) = \beta_i \cdot (T - T_C), \tag{2}$$

where β_i is the slope. Notably, the linear regression lines of $\bar{\alpha}_{\text{Dis}}^i$ vs. T for all five elements cross over at nearly the same point on the T axis (Fig. 4D), which is denoted as T_C (~ 1388 ± 51 K).

Taking Cr as one example, the MC/MD-simulated Γ_{Cr} has linear relation with Γ_{Dis} statistically (Fig. 4C). The positive slopes ($\bar{\alpha}_{\text{Dis}}^{\text{Cr}} > 0$) of the Γ_{Cr} vs. Γ_{Dis} regression lines are related to the positive Cr segregation at the GB. The fitted $\bar{\alpha}_{\text{Dis}}^{\text{Cr}}$ value linearly decays by increasing the temperature with a negative slope of β_{Cr}, and intersects with the T axis at $T_C = 1347$ K (Fig. 4D). Similar behavior can also be found for Mn (with moderate positive GB segregation), where $\bar{\alpha}_{\text{Dis}}^{\text{Mn}} > 0$ (Suppl. Fig. S10(b)), $\beta_{\text{Mn}} < 0$, and $T_C = 1464$ K (albeit a high uncertainty in T_C due to the small slope). In contrast, the slopes of $\Gamma_{\text{Fe(Co)}}$ vs. Γ_{Dis} regression lines are negative ($\bar{\alpha}_{\text{Dis}}^{\text{Fe(Co)}} < 0$) due to the depletion of Fe or Co (Suppl. Fig. S10(c-d)); consequently, $\bar{\alpha}_{\text{Dis}}^{\text{Fe(Co)}}$ linearly increases with increasing temperature ($\beta_{\text{Fe(Co)}} > 0$; $T_C = 1370$ K for Fe and 1371 K for Co, respectively, in Fig. 4D). Finally, there is only a weak correlation in Γ_{Ni} vs. Γ_{Dis} with a small negative slope due the small Γ_{Ni} values, which also results in large relative noises in Suppl. Fig. S10(e) and possibly a large error on the projected T_C (due to small slope and large uncertainty).

A physics-informed data-driven model (PIDDM)

Next, we propose a physics-informed data-driven model or PIDDM based on above analysis of the large MC/MD-simulated dataset. The detailed derivation and data-fitting of this PIDDM can be found in Methods section and Supplementary Discussion 6. Here, Γ_i and Γ_{Dis} as functions of T and $X (= \{X_i\})$ of an HEA can be expressed as:

$$\Gamma_i(T,X) = \beta_i \cdot (T - T_C) \cdot \left[\Gamma_{Dis}(T,X) - \Gamma_{\text{Dis}}^0\right] + \Sigma_j\left(\kappa_{i,j}^{\text{Seg}} \cdot X_j\right), \tag{3}$$

and

$$\Gamma_{Dis}(T,X) = \Sigma_i\left(\kappa_i^{\text{Dis}} \cdot X_i\right) \cdot \exp\left(-\frac{E_{\text{A}}^{\text{Dis}}}{k_{\text{B}}T}\right). \tag{4}$$

Here, $\kappa_{i,j}^{\text{Seg}}$ is the coupling coefficient for the GB segregation between component i and j, $E_{\text{A}}^{\text{Dis}}$ is an activation energy (in an Arrhenius relation), k_{B} is the Boltzmann constant, and κ_i^{Dis} is the coupling coefficient for GB disorder and component i. By using the best fitted parameters listed in Suppl.

Tables S1 and S2, the predicted GB properties from this PIDDM agree with MC/MD simulations with a small root-mean square error (RMSE) of ~2.3 nm^{-2} (Suppl. Fig. S11).

Here, a distinct merit of this PIDDM (in contrast to the ANN model) is that all the model parameters have clear physical meanings. The fitted κ_i^{Dis} and $\kappa_{i,j}^{Seg}$ values (Suppl. Tables S1 and S2) represent the couplings between segregation and disorder, as well as segregation of different elements, which are fully consistent with the trends observed in our MC/MD simulations, as discussed in Supplementary Discussion 6. Notably, we can predict the $\Gamma_i(T,X)$ with relatively small RMSEs for each element using this simple analytical Eq. (3), as shown in Suppl. Table S2. Furthermore, the parity plots show that the PIDDM predictions agree well with the hybrid MC/MD simulations for all elements (Suppl. Fig. S11).

Interestingly, this PIDDM provides a new physics insight via decoupling the effect of GB disorder on segregation. The first term, $\beta_i \cdot (T - T_C) \cdot [\Gamma_{Dis}(T,X) - \Gamma_{Dis}^0]$, represents the "disorder contribution" (albeit it is in fact a coupled disorder and segregation effect). The second (typically smaller) term $\Gamma_i^0(X) = \sum_i(\kappa_i^{Dis} \cdot X_i)$ in Eq. (3) is the composition contribution to the GB adsorption at the minimum disorder. Thus, we can further quantify the fractions of this disorder contribution (the first term) to the 1032 model-predicted Γ_i values and plot them in histograms for all five elements in Fig. 5A. The large factions of 0.70 for Γ_{Cr}, 0.71 for Γ_{Fe}, and 0.66 for Γ_{Co}, respectively, suggest the significant roles of GB disorder in influencing the GB segregation of Cr, Fe, and Co. However, the fractions of disorder contributions are moderate (~0.46) for Γ_{Mn} and almost zero for Γ_{Ni} (Fig. 4A). Interestingly, the fraction of disorder contribution is proportional to the absolute value of coupling coefficient $|\beta_i|$. These findings demonstrate the importance of GB disorder on influencing GB segregation in HEAs, particularly for Cr, Fe, and Co in this case. Again, Eq. (3) shows that $\overline{\alpha}_{Dis}^i$ obtained from fitting Eq. (1) only represents a statistical correlation between Γ_i and Γ_{Dis} after averaging the last term $\sum_j(\kappa_{i,j}^{Seg} \cdot X_j) \equiv \Gamma_i^0(X)$ over 258 different $X = \{X_i\}$ values.

The physical meaning and origin of T_C are briefly discussed here. Eq. (3) implies that at $T = T_C$, $\Gamma_i(T_C,X) = \sum_i(\kappa_i^{Dis} \cdot X_i)$ and Fig. 4C shows that $\overline{\sum_i(\kappa_i^{Dis} \cdot X_i)} = \Gamma_i^0 \sim 0$ is a small number. Thus, T_C represents an isoequilibrium temperature due to an (incomplete) enthalpy-entropy compensation effect[39-41] to produce $\Gamma_i(T_C,X) \sim 0$. Note that "isoequilibrium" is the preferred term adopted here (vs. "compensation" [39]) based on the terminology discussion by Liu and Guo.[41] Here, we can use a general adsorption isotherm to discuss this isoequilibrium effect:

$$\frac{X_i^{GB}}{X_1^{GB}} = \frac{X_i^{Bulk}}{X_1^{Bulk}}\exp\left(-\frac{\Delta G_{i \to 1}^{Seg(eff)}}{k_B T}\right). \tag{5}$$

where X_i^{GB} and X_i^{Bulk} are the fractions of solute element i at the GB and bulk, respectively, $i = 1$ is a "reference" element (e.g., Ni), and $\Delta G_{i \to 1}^{Seg}$ is the free energy of segregation by swapping of a solute atom i inside the bulk (grain) with an atom 1 at the GB. In general, an isoequilibrium effect[39-41] can occur at T_C if the effective GB segregation entropy ($\Delta S_{i \to 1}^{Seg\,(eff)}$) and enthalpy ($\Delta H_{i \to 1}^{Seg(eff)}$) are

linearly related:

$$\Delta H_{i\rightarrow 1}^{Seg(eff)} = \Delta S_{i\rightarrow 1}^{Seg(eff)} \cdot T_C + C, \tag{6}$$

so that $\Delta G_{i\rightarrow 1}^{Seg(eff)} = C$ (Fig. 5B). In the current case, the constant C must vanish for a perfect isoequilibrium effect due to the network constrain ($\sum_i X_i^{Bulk} = \sum_i X_i^{GB} = 1$), albeit the enthalpy-entropy compensation is not perfect (leading to small residual $C_i = \Gamma_i^0(X) = \sum_i (\kappa_i^{Dis} \cdot X_i) \neq 0$).

Comparing Eq. (6) with Eq. (3), we conclude that this entropic effect must be related to the increased GB disorder $\Delta \Gamma_{Dis}$. Thus, we can now envision the following physical meaning and origin of T_C. The increased GB disorder $\Delta \Gamma_{Dis}$ can reduce the effective GB free energy of segregation ($\Delta G_{i\rightarrow 1}^{Seg(eff)} = \Delta H_{i\rightarrow 1}^{Seg(eff)} - T \cdot \Delta S_{i\rightarrow 1}^{Seg(eff)}$) through the entropy of GB segregation, where the reduction is proportional to $\Delta H_{i\rightarrow 1}^{Seg(eff)}$ and more significant for strong segregating or depleting element. Thus, with increasing GB disorder $\Delta \Gamma_{Dis}$, GB segregation (or depletion) for different elements is reduced and equalized due to this enthalpy-entropic compensation effect. The effective $\Delta G_{i\rightarrow 1}^{Seg(eff)}$ virtually vanishes (or is minimized) at T_C, as shown in Fig. 5B.

It should be noted that this compensation effect is likely only an approximate relation because $\Gamma_i(X)$ is small but not exactly zero. Our data (Fig. 4D and Suppl. Table S2) also show variations in the best fitted T_C values for different elements ($\sim 1388 \pm 51$ K). We should also note that this predicted T_C is from an extrapolation. As the temperature approaches the bulk solidus curve, premelting-like interfacial phases[6, 31, 42, 43] can develop at GBs to change the projection. See Supplementary Discussion 12 for further elaboration about the origin and physical meaning of T_C.

It is interesting to further compare the fitted compositional coefficients ($\kappa_{i,j}^{Seg}$) with the corresponding segregation enthalpies in binary alloys. In our sign convention, a positive $\left(-\Delta H_{i,j}^{Seg}\right)$ (*i.e.*, a negative $\Delta H_{i,j}^{Seg}$) indicates a preference for segregation (in the opposite to that used in a prior model [44] where we obtained the data). Taking Cr as one example, Fig. 5C shows the parity plot of Cr segregation enthalpies $\left(-\Delta H_{Cr,j}^{Seg}\right)$ (j = Mn, Fe, Co, Ni)[44] *vs.* corresponding compositional coefficients $\kappa_{Cr,j}^{Seg}$ fitted from our PIDDM. The excellent linear relation with $R^2 = 0.95$ indicates a strong positive correlation. In addition, signs of the $\left(-\Delta H_{Cr,j}^{Seg}\right)$ and $\kappa_{i,j}^{Seg}$ are always consistent. For instance, both a positive $\left(-\Delta H_{Cr,Fe}^{Seg}\right)$ or $\left(-\Delta H_{Cr,Co}^{Seg}\right)$ in the classical segregation model and a positive $\kappa_{Cr,Fe}^{Seg}$ or $\kappa_{Cr,Co}^{Seg}$ in our PIDDM indicates preferred segregation of Cr at the GB of Fe or Co. Likewise, negative $\left(-\Delta H_{Cr,Mn}^{Seg}\right)$ or $\left(-\Delta H_{Cr,Ni}^{Seg}\right)$ and $\kappa_{Cr,Mn}^{Seg}$ (or $\kappa_{Cr,Ni}^{Seg}$) suggest preferred depletion of Cr at the GB of Mn or Ni. Thus, the compositional coefficients ($\kappa_{i,j}^{Seg}$) are well correlated with binary segregation enthalpies.

Both the PIDDM and ANN models can be used to map out the GB thermodynamic properties for HEAs in a 5D space as functions of four independent compositional DOFs and temperature. In comparison with the PIDDM, the ANN model is more accurate for predicting GB properties with smaller RMSEs (Suppl. Tables S1-S2). However, the ANN model predicts GB properties in

a "black-box" approach without offering any physics insights. Moreover, lacking the physical interpretation inhibits the model transferability. In contrast, the simple analytical formulae of the PIDDM, where all model parameters have clear physical meanings, can provide understandings of the underlying physical interactions (including their signs and strengths in a quantitative way) in HEAs. Thus, this PIDDM represents a general and transferrable model for HEAs.

Comparisons with classical and other existing segregation models

First, we compare the GB segregation predicted by classical segregation models for binary alloys *vs.* the hybrid MC/MD-simulated GB segregation in HEAs. Here, we adopt the Wynblatt-Ku model[47] (considering both chemical and elastic contribution to GB segregation; see Supplementary Discussion 7) to compute GB fractions (X_{GB}) of Cr as functions of bulk fractions of Cr ($x = X_{Cr}$) for four Cr_xM_{1-x} (M = Mn, Fe, Co, Ni) binary alloys at 1000 K (Fig. 5D). Then, we select four HEAs, including $Cr_xMn_{0.4-x}Fe_{0.2}Co_{0.2}Ni_{0.2}$ ($0.05 \leq x \leq 0.35$), as well as three variants where we swap Mn with Fe, Co, or Ni; we further plot MC/MD-simulated Γ_{Cr} curves as functions of x in Fig. 5E. We notice several major differences in the trends of segregation in binary alloys *vs.* HEAs. The segregation strengths of Cr in binary alloys are ranked as Fe > Co > Ni > Mn (Fig. 5D), while they are ranked as Mn $\approx$ Ni > Co $\approx$ Fe in HEAs (Fig. 5E). More complex and intriguing compositional dependences, *e.g.*, saturation of Cr segregation with $x > 0.2$ in $Cr_xMn_{0.2}Fe_{0.4-x}Co_{0.2}Ni_{0.2}$ and $Cr_xMn_{0.2}Fe_{0.2}Co_{0.4-x}Ni_{0.2}$ *vs.* acceleration of Cr segregation after $x > 0.2$ in $Cr_xMn_{0.2}Fe_{0.2}Co_{0.2}Ni_{0.4-x}$, are also observed in HEAs (Fig. 6B).

Second, Xing *et al.* developed a lattice-type model for predicting GB segregation for ternary alloys[48], and we further extended this model to quinary alloys in Supplementary Discussion 9. Although this model can predict some general trends, *e.g.*, the positive segregation enthalpies for Cr, Mn, and Ni (segregation) *vs.* negative segregation enthalpies for Fe and Co (depletion), we cannot make quantitative predictions of GB segregation for non-equimolar HEAs.

Third, Li *et al.* proposed a density-based thermodynamic model for GB segregation.[23] This phenomenological model assumed that GB energy can be written as a function of GB density, which suggested the importance of GB free volume (V_{Free}). In contrast, the PCC heat map (Fig. 4A) shows that V_{Free} almost has no correlations with GB adsorption properties (Γ_i). Instead, Γ_{Dis} exhibits strong correlation with Γ_i (Fig. 4A). Thus, we suggest that GB disorder (instead of density or free volume) should be treated as a key parameter for developing future phenomenological models. See Supplementary Discussion 8 for further discussion.

Key characters of various GB segregation models are compared in Fig. 5F.

It is important to note that most of the above simplified models and analyses (discussed in this and last sections) do not consider segregation entropies, as well as multilayer adsorption and different segregation free energies, which are important and can affect GB segregation substantially. The entropic effects of GB segregation have been considered in prior studies of binary and ternary alloys.[45, 46] Here, we have discovered a T_C due to an (incomplete) enthalpy-

entropy compensation effect (Eq. (5) and (6)), which represents an entropic effect of the segregation in this five-component system.

Generality of the predictions and DFT validation

In this study, most MC/MD simulations are based on an asymmetric Σ81 (mixed tilt and twist) GB to represent the behaviors of general GBs. To test generality of our predictions, we have also performed MC/MD simulations for three other GBs, including an asymmetric Σ15 (mixed tilt and twist) GB, a Σ41 symmetric tilt GB, and a Σ13 symmetric twist GB. For each of them, four non-equimolar HEAs selected based in the simulations of the asymmetric Σ81 GB diagrams, where the first three (HEA1-3) exhibit strong Cr segregation while last one (HEA4) has weak Cr segregation, were examined. Notably, MC/MD simulations show similar and consistent trends for all four GBs: HEA1-3 always have large Γ_{Cr}, but HEA4 always has small Γ_{Cr} (Suppl. Table S4 and Fig. S15). Furthermore, DFT calculations also confirm that E_{Seg}^{Cr} (around -0.026 eV/atom) of HEA1-3 is significantly lower than that for HEA4 (~ 0.0001 eV/atom), as shown in Supplementary Table S4. In conclusion, the trends predicted based on the asymmetric Σ81 (mixed tilt and twist) GB are likely representative. See Supplementary Discussion 10 for elaboration.

Probing segregation mechanisms by first-principles calculations of electronic structures

We have also calculated sum of bond ordering (SBO) values for the four non-equimolar HEAs discussed above to further understand how the bonding environment affects the Cr segregation (see Methods). Since SBO represents the total number of electrons that form bonds, similar SBO values indicate similar bonding environments. Interestingly, Fe, Cr, and Co atoms always have similar SBO values, which are ~4.04, ~3.95, and ~3.78, respectively. In contrast, Mn and Ni exhibit two distinct SBO values of ~4.20 and ~3.49, respectively (Suppl. Fig. S16). Therefore, the preferred Cr segregation at the Fe- or Co-rich GBs can be understood because Fe or Co can provide more favorable segregation sites with similar bonding environments. On the other hand, the different bonding environments at Mn- or Ni-rich GBs can inhibit Cr segregation.

A recent study suggested that SBO can be used as a descriptor to predict and subsequently tailor GB segregation.[28] For example, if we want to promote segregation of a certain element (*e.g.*, Cr) in HEAs, we can increase the composition of the elements with similar SBO values (*e.g.*, Fe and Co) and/or reduce the composition of those with different SBO values (*e.g.*, Mn and Ni). It is interesting to note that large chemical-affinity disparity of different elements can foster segregation in HEAs.[37] Here, it is shown that the atomic pairs of Cr-Fe and Cr-Co have relatively larger chemical-affinity disparity than Cr-Mn and Cr-Ni, thereby suggesting that Fe and Co will promote Cr segregation to CrMnFeCoNi GBs.[37] This is consistent with the SBO calculations.

Conclusions

In this study, we used large-scale hybrid MC/MD simulations to generate a large dataset of GB properties for $Cr_x Mn_y Fe_z Co_l Ni_m$ HEAs. The machine learning technique was firstly used to predict the GB properties of HEA as a function of four compositional DOFs and temperature in a 5D

space. In addition, we found that interfacial disorder, as well as the interactions among segregation/depletion of five elements and GB disorder, can induce new and complex interfacial phenomena in HEAs, beyond the prediction of classical GB segregation theory. Notably, we discovered a GB critical compensation temperature in HEAs. Based on a careful analysis of the large dataset, we further created a physics-informed data-driven model or PIDDM to represent GB segregation and disordering in the 5D space, where all parameters have clear physical meanings (*vs.* the black-box machine learning model). This work has enriched the classical GB segregation theory and developed a predictive and transferrable model for a range of HEA systems.

Methods

Composition Selection and Principal Component Analysis (PCA). In this work, the composition of each element was fixed in a range from 5 at% to 35 at% with a step of 5 at% for $Cr_xMn_yFe_zCo_lNi_m$. Since $x + y + z + l + m = 1$, there are 1371 possible compositions in total. Among them, we randomly selected 258 subsystems to perform high-throughput MC/MD simulations to generate a dataset. Principal component analysis (PCA) was used to analyze the composition distribution for these 258 subsystems to ensure the randomness of our selection, where the singular value decomposition (SVD) algorithm was chosen. The PCA were performed by Matlab2019a.

Interatomic Potential: Validity and Limitation. We adopted the second nearest-neighbor modified embedded atom method (2NN MEAM) potential.[49] This potential was initially developed for the equiatomic alloys but is extended to model non-equiatomic compositions from 5 to 35 at% in this study. It was based on validated MAEM potentials for unary, binary, and ternary alloys, which suggests its extendibility to predict short-range interactions in non-equiatomic compositions. This is perhaps the best available interatomic potential for the CrMnFeCoNi system. Wynblatt and Chatain have also applied this MEAM potential to study the GB thermodynamic properties of CrMnFeCoNi for Σ13 twist GB and produced simulation results were supported by experiments (albeit most of GBs in the experimental studies have random orientations).[24] In addition to Σ13 twist GB, our MC/MD simulations were performed for Σ15 asymmetric tilt GB, Σ41 symmetric tilt, and Σ81 mixed GB. Our simulations showed that Cr and Mn elements always have the strongest segregation for all four GBs (Suppl. Fig. S15), thereby suggesting the generality of our predictions. Moreover, we believe this study can offer insights on new physics of segregation phenomena in HEAs (even if the prediction is limited by the potential). Notably, this study can help us to validate the new methodology of modeling GB segregation in HEAs (*e.g.*, computing GB diagrams and establish new analytical models), which can be used for other systems and/or when improved potentials are developed in future.

Hybrid MC/MD simulation for GB diagrams. The GBstudio[50] website was used to construct a mixed twist-tilt Σ81 GB with boundary planes $(1\bar{1}0)//(7\bar{8}7)$ to represent general GBs of a large simulation cell containing 11,664 atoms. The energy minimization for each GB was first performed at 0 K by conjugate gradient (CG) algorithm. Subsequently, the hybrid Monte Carlo and molecular dynamics (hybrid MC/MD) simulations in constant NPT ensembles were carried out to swap atoms and find energetically favorable GB structure. Five MC trial moves were conducted between each MD step with a 0.1 fs MD time step and 10^5 hybrid MC/MD steps performed for each simulation to achieve convergence. The NPT-based MC/MD simulations can consider the coupling of GB disordering and segregation. All MD and MC/MD simulations were performed using the LAMMPS code.[51] 1032 individual hybrid MC/MD simulations were performed (for 258 compositions × 4 different temperatures) to generate a large 1032×6 (5 bulk compositions plus temperature) dataset to train ANN models. Note that all MC/MD simulations were performed from 1000 K to 1300 K because experiments show that CrMnFeCoNi HEA has single FCC phase at this temperature range.[33]

The methods used to calculate GB excess of solute (*i.e.*, GB adsorption amount Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, Γ_{Co}, Γ_{Ni}) and disorder (Γ_{Dis}) diagrams were same as our prior studies. The GB excess of disorder Γ_{Dis} was computed based on: $\Gamma_{Dis} = \int_{-\infty}^{+\infty}[\eta_{Dis}(z) - \eta_{Dis}(\pm\infty)]dz$, where the disorder parameter η_{Dis} ($\eta_{Dis} = 1$ means liquid and $\eta_{Dis} = 0$ represents perfect crystal) was calculated of each atom following a procedure proposed by Chua *et al.*.[34] More detailed procedures are described in our prior works.[20, 28, 52] The exact value of Γ_{Dis} is difficult to be measured directly by experiments. However, the relative levels of structural disorder at the GB can be examined by STEM and compared with simulated Γ_{Dis} (see, *e.g.*, our recent study[52]). To calculate GB free volume (V_{Free}), we used the relation of $V_{Free} = V_{Total} \cdot \Sigma\Gamma_i$, where V_{Total} is the total volume of GB structure and i = Cr, Mn, Fe, Co, or Ni. To minimize the thermal noise effect, we calculated each GB property based on the average of five random structures during the last five MC/MD steps.

It should be noted that we set an overall global composition in a hybrid MC/MD simulation. The bulk composition is recalculated based the grain composition (away from the GB region) after achieving the chemical equilibrium, which is subsequently used for both training the ANN model and further analysis and developing a data-based analytical model. It is important to note that we always plot the GB segregation and other properties as a function of actual bulk composition measured at the center of the grain (instead of the global composition). Thus, the grain size effect is not included.

Machine Learning models. To identify the best machine learning model for the data-driven prediction and discovery of GB properties, we trained and tested an artificial neural network (ANN) model, along with three other linear regression-based machine learning models: support vector machine (SVM), regression tree (Tree), and rational quadratic Gaussian models. To avoid overfitting, 5 folds was chosen as a cross-validation value for training the SVM, regression tree and Gaussian models. All data processing and development of machine learning models were performed by Deep Learning Toolbox and Regression Learner Toolbox in Matlab2019a.

By comparing the RMSE values of predicting six GB properties of training ANN model and other three linear regression models (Fig. 2(J)), we found that the RMSEs of the ANN model always have the lowest values among all four machine learning models. Thus, we adopted the ANN model as the machine learning engine for this data-driven prediction of GB properties.

For training and validating the ANN models, we split the data set into training, validation, and test subsets in a ratio of 0.7:0.15:0.15. The Levenberg-Marquardt backpropagation function was adopted to train ANN models. We found the optimized network architectures for the ANN (n^i-$n^{[i]}$-1, where n^i is the number of input parameter, $n^{[i]}$ (the number of neurons in the single layer) is set to be 6-20-1. The random dropout values were set to default setting 0.5. To prevent over-fitting issues for training ANN models, we always train ANN models until the prediction of training, validation, and test data have similar MSE errors.

Derivation of a physics-informed data-driven model (PIDDM). Based on the linear regression analyses shown in Fig. 4C and Suppl. Fig. S10, the adsorption amount, $\Gamma_i(T,X)$, is statistically correlated with the GB excess of disorder, $\Gamma_{Dis}(T,X)$, linearly with the slope $\bar{\alpha}_{\text{Dis}}^i(T)$ at a given temperature T, where $X = \{X_i\}$ is a concise form to note the bulk composition of the HEA. Thus, we statistically have the following linear correlation:

$$\Gamma_i(T,X) - \Gamma_i^0 = \bar{\alpha}_{\text{Dis}}^i(T) \cdot [\Gamma_{Dis}(T,X) - \Gamma_{Dis}^0], \tag{7}$$

where $(\Gamma_i^0, \Gamma_{Dis}^0)$ is the intersection point of all linear regression lines for different temperatures in each panel of Fig. 4C, and they are virtually independent of temperature. Here, we use an overbar on $\bar{\alpha}_{\text{Dis}}^i$ to denote that it is a statistical average value. In other words, we do not suggest a simple linear relation between Γ_i and Γ_{Dis}; on the contrary, we know it is not (because of $\Gamma_i^0 = \Gamma_i^0(X)$), where overbar means a statistically averaged value, as the individual $\Gamma_i^0(X)$ is not a constant based Eq. (9) below). Scattered data points in Fig. 4C can be ascribed to 258 different compositions $X = \{X_i\}$. However, a hidden statistical correlation exists after averaging over the 258 random compositions X. We can observe in Fig. 4C that Γ_i^0 is a relatively small number: $\Gamma_i^0 = \overline{\Gamma_i^0(X)} \ll \overline{\Gamma_i(T,X)}$. Furthermore, the linear regression analyses shown in Fig. 4D suggest:

$$\bar{\alpha}_{\text{Dis}}^i(T) = \beta_i \cdot (T - T_C), \tag{8}$$

where β_i is slope of the linear regression line in Fig. 4D. Here, $T_C \approx 1388 \pm 51$ K is a critical temperature shown in Fig. 4D. At $T = T_C$, $\Gamma_i(T_C,X) = \Gamma_i^0 \sim 0$ (see Suppl. Fig. S10), so this critical temperature is the so-called "isoequilibrium" (or "compensation") temperature of GB segregation.[39, 40] Eqs. (7)-(8) are same as Eqs. (1)-(2) in the main text. We further propose:

$$\Gamma_i(T,X) = \beta_i \cdot (T - T_C) \cdot [\Gamma_{Dis}(T,X) - \Gamma_{Dis}^0] + \Gamma_i^0(X). \tag{9}$$

14

Here, we can assume $\Gamma_{Dis}^0 = \Gamma_{\mathrm{Dis}}^{\mathrm{min}}$ (or the minimum among all possible HEAs compositions), which are approximately held based on Fig. 4C and Suppl. Fig. S10, except for the case of Ni, where there are too high noises due to the small values of Γ_{Ni}. In Eq. (7) and Fig. 4 and S10, Γ_i^0 is a fitted constant independent of X. In Eq. (9), we further generalize Eq. (7) to allow this constant Γ_i^0 to be a function of X to enable more accurate fitting, where we have $\Gamma_i^0 \equiv \Gamma_i^0(X)$. Here, we may adopt a linear expression as a first-order approximation:

$$\Gamma_i^0(X) = \sum_j \left(\kappa_{i,j}^{\mathrm{Seg}} \cdot X_j \right) \tag{10}$$

where $\kappa_{i,j}^{\mathrm{Seg}}$ is a coupling coefficient for the GB segregation. Thus, we have:

$$\Gamma_i(T,X) = \beta_i \cdot (T - T_{\mathrm{C}}) \cdot \left[\Gamma_{Dis}(T,X) - \Gamma_{\mathrm{Dis}}^0 \right] + \sum_j \left(\kappa_{i,j}^{\mathrm{Seg}} \cdot X_j \right). \tag{11}$$

Since GB disorder should increase with temperature, we propose the following relation:

$$\Gamma_{Dis}(T,X) = \Gamma_{Dis,\,0}(X) \cdot \exp\left(-\frac{E_A^{Dis}}{k_{\mathrm{B}}T} \right), \tag{12}$$

where E_A^{Dis} is the activation energy of disordering, and k_{B} is the Boltzmann constant. We again adopt a linear expression as a first-order approximation for the temperature-independent pre-factor:

$$\Gamma_{Dis,\,0}(X) = \sum_i \left(\kappa_i^{\mathrm{Dis}} \cdot X_i \right) \tag{13}$$

Next, we can use all hybrid MC/MD-simulated data points to fit Eqs. (12) and (13). Finally, by combining Eqs. (11)-(13), we can obtain:

$$\Gamma_i(T,X) = \beta_i \cdot (T - T_{\mathrm{C}}) \cdot \left[\sum_i \left(\kappa_i^{\mathrm{Dis}} \cdot X_i \right) \exp\left(-\frac{E_A^{Dis}}{k_{\mathrm{B}}T} \right) - \Gamma_{\mathrm{Dis}}^0 \right] + \sum_j \left(\kappa_{i,j}^{\mathrm{Seg}} \cdot X_j \right). \tag{14}$$

Further discussions about the physics-informed data-driven model (PIDDM) and the physical meaning and origin of T_{C} can be found in Supplementary Discussions 5 and 6.

Density functional theory (DFT) calculations. The first-principles DFT calculations were performed by using the Vienna *ab initio* Simulations Package (VASP).[53, 54] The Kohn-Sham equation was used to solve the projected-augmented wave (PAW) method[55, 56] along with standard PAW potentials. All GB structures were fully relaxed until the Hellmann-Feynman forces were smaller than 0.02 eV/Å. The Brillouin-zone integrations were sampled on a Γ-centered $2\times2\times1$ k-point grids. The kinetic energy cutoff for plane waves was set to 368 eV. The convergence criterion for the electronic self-consistency was set to 10^{-4} eV. The "high" precision setting was adopted to avoid wrap around errors. The spin-polarization was not considered due to weak effect on atomic arrangement.[57] The SBO was calculated by using the state-of-the-art DDEC06 method[58] following the all-electron static calculations.

Author contributions: J. L. conceived the idea and supervised the work. C. H. performed simulations and calculations. Both authors wrote, reviewed, and revised the manuscript.

Competing interests: The authors declare no conflict of interests.

Data availability: The MC/MD-simulated dataset for the PIDDM and ANN models are available on GitHub website https://github.com/huhuhhhh/HEAGB/tree/main/raw_data. The supporting data for

generating binary and ternary GB diagrams are also available in this repository.

Code availability: The MATLAB scripts for fitting the PIDDM and training ANN models are available on GitHub website https://github.com/huhuhhhh/HEAGB/tree/main.

Electronic Supplementary Information:
Supplementary Discussions 1-13
Supplementary Tables S1-S4
Supplementary Figures S1–S29
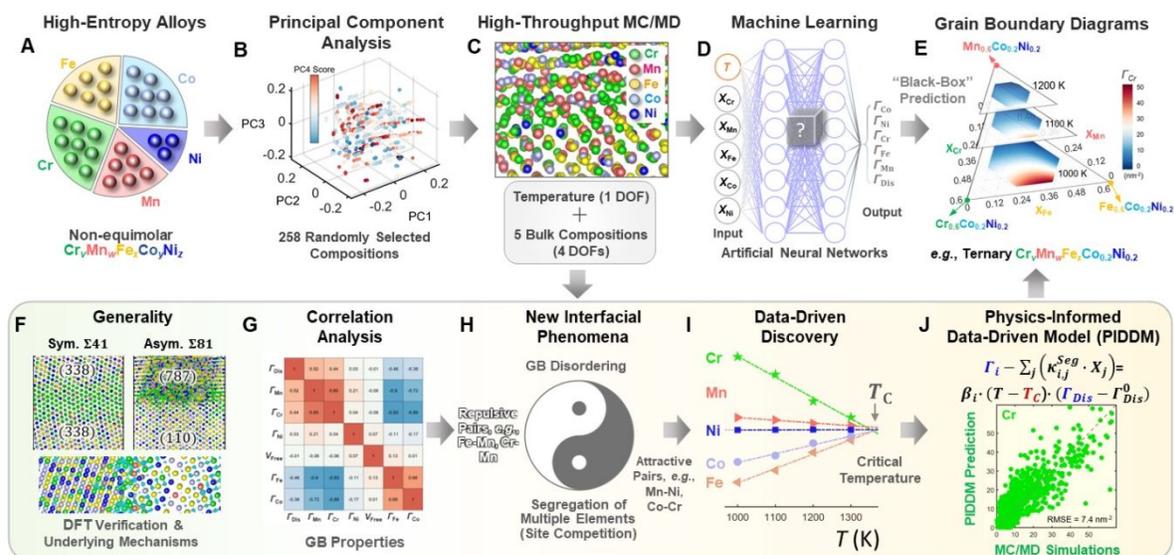
Fig. 1. Workflow of the machine learning prediction of grain boundary (GB) diagrams, data-based discovery of new interfacial phenomena, and development of a physics-informed data-driven model (PIDDM). (A) Schematic diagram of non-equimolar five-element $Cr_xMn_yFe_zCo_lNi_m$ alloys. (B) Principal component analysis (PCA) verifying the randomness in the selection of 258 HEAs. (C) The equilibrium structure of an asymmetric Σ81 GB in $Co_{0.2}Ni_{0.2}Cr_{0.2}Fe_{0.35}Mn_{0.05}$ at 1000 K obtained by hybrid Monte Carlo and molecular dynamics (hybrid MC/MD) simulations. In total, 1032 such individual hybrid MC/MD simulations were performed for 258 HEAs at four different temperatures to calculate GB excesses of solutes (*i.e.*, Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, Γ_{Co}, Γ_{Ni}) and disorder (Γ_{Dis}), and free volume (V_{Free}). (D) Schematic diagram of an artificial neural network (ANN) model for predicting six GB properties (excluding V_{Free}, because of its small values and weak correlations with other variables); three other machine learning models have also been trained and tested, while the ANN model is selected because of its best performance. (E) An example of GB diagrams predicted by the ANN model for a ternary $Cr_xMn_yFe_zCo_{0.2}Ni_{0.2}$ ($x + y + z = 0.6$) subsystem, showing three isothermal sections of the Cr adsorption (Γ_{Cr}) diagrams. (F) Screenshot of strong Cr segregation in different GBs, which is also verified by DFT calculations. (G) Correlation analysis of GB properties. (H) Schematic of new interfacial phenomena in HEAs. (I) A data-driven discovery of a GB critical temperature T_C. (J) A PIDDM that can be used to predict GB properties with parameters of clear physical meanings.
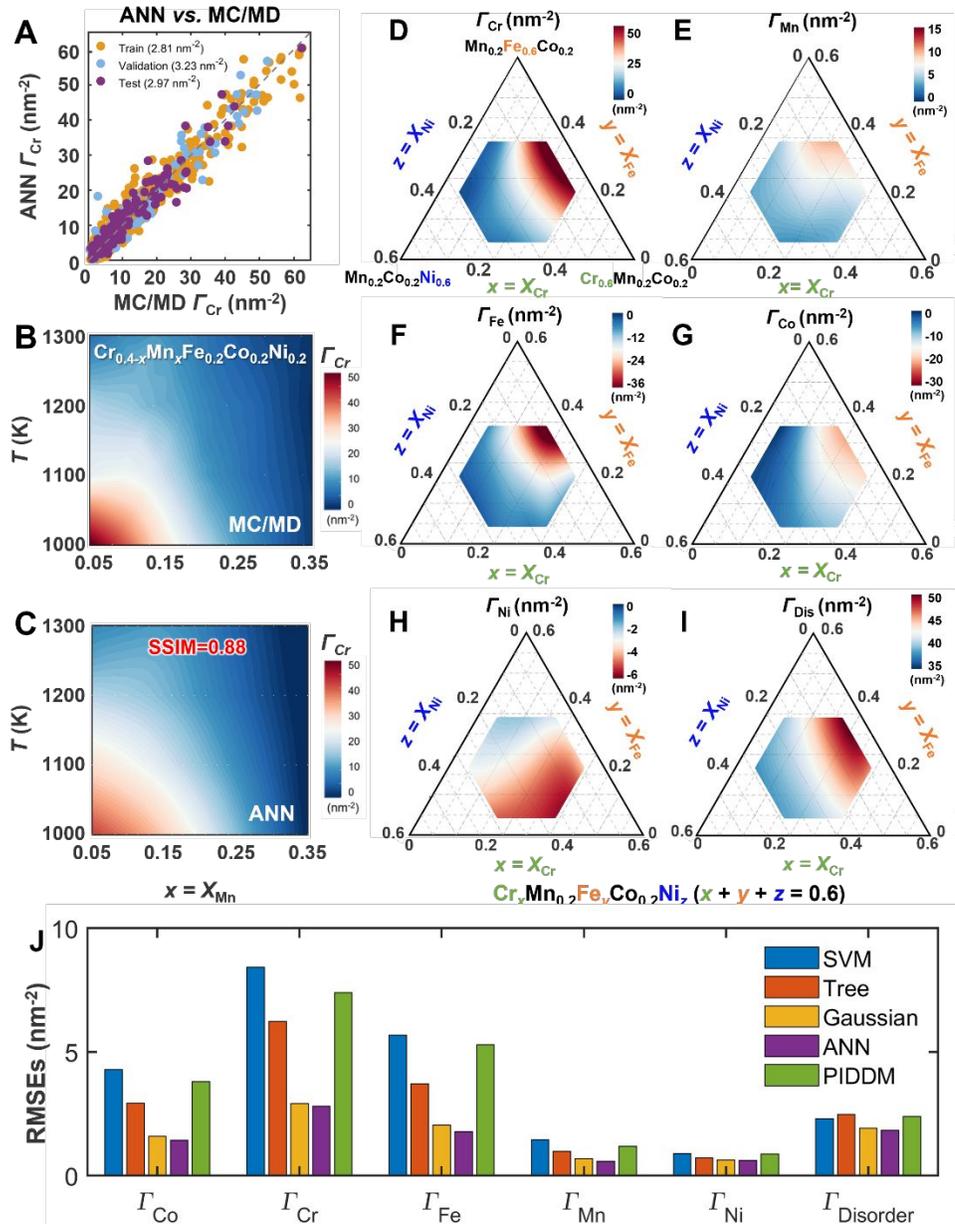
Fig. 2. ANN performance for predicting GB properties. (A) Parity plot of ANN predictions *vs*. MC/MD simulations for the GB excess of Cr adsorption (Γ_{Cr}). (B-C) MC/MD-simulated *vs*. ANN-predicted isopleths of Γ_{Cr} diagrams as functions of temperature and Mn bulk composition ($x = X_{Mn}$) for the $Cr_{0.4-x}Mn_xFe_{0.2}Co_{0.2}Ni_{0.2}$ system. (D-I) Representative ternary isothermal sections of ANN-predicted GB diagrams of Γ_{Cr}, Γ_{Mn}, Γ_{Fe}, Γ_{Co}, Γ_{Ni}, and Γ_{Dis} for $Cr_xMn_{0.2}Fe_yCo_{0.2}Ni_z$ ($x + y + z = 0.6$; $x = X_{Cr}$, $y = X_{Fe}$, $z = X_{Ni}$) at 1000 K. Note that different color bars are used for different diagrams better represent different ranges. Specifically, the color bars of Γ_{Cr} and Γ_{Mn} diagrams represent from zero (blue for weak segregation) to positive values (red for strong segregation) in Panel (D, E, I), while those in the Γ_{Fe}, Γ_{Co}, and Γ_{Ni} diagrams represent from zero (blue for weak depletion) to negative values (red for strong depletion) in Panel (F, G, H). (J) Comparison of RMSEs for six GB properties predicted by the ANN model, three other machine learning models, *i.e.*, support vector machine (SVM), regression tree (Tree), and rational quadratic Gaussian models, and the PIDDM. The ANN model is selected to predict GB diagrams because it yielded the lowest RMSEs among the four machine learning models.
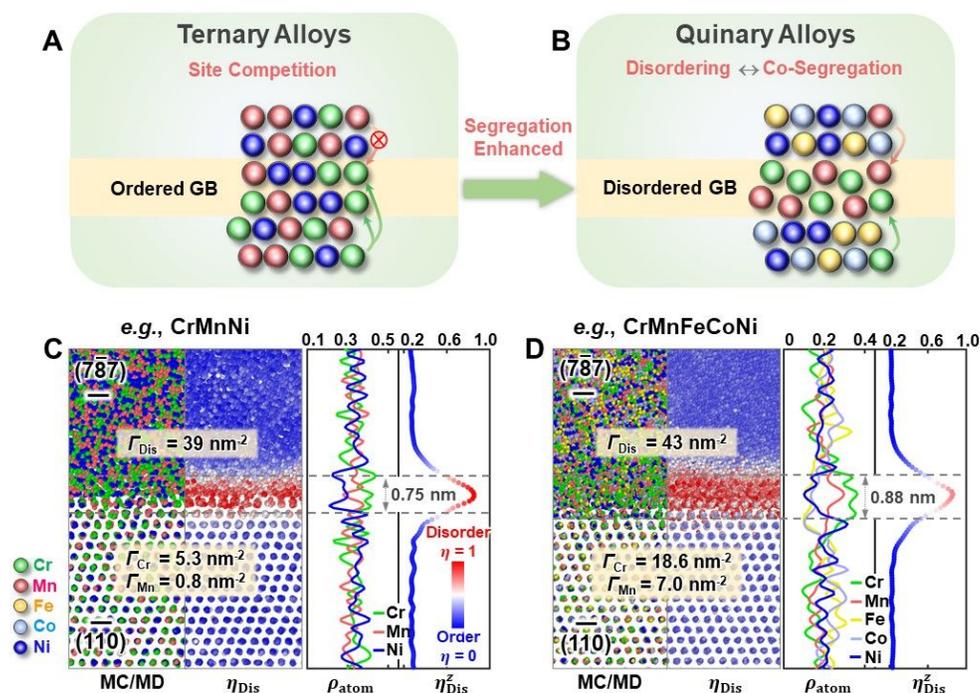
Fig. 3. **Unique coupled interfacial disordering and GB co-segregation behaviors in HEAs, exemplified by comparing the same asymmetric Σ81 GB in equimolar CrMnNi *vs.* CrMnFeCoNi alloys at 1000 K.** (A) Schematic of the site competition in relatively ordered GBs. (B) Schematic of the coupling of interfacial disordering and co-segregation of Cr and Mn in quinary alloys, which can enhance each other. (C) MC/MD-simulated GB structure of the equimolar CrMnNi alloy and the corresponding disorder parameter (η_{Dis}) and atomic density profiles. This GB exhibits an GB excess disorder Γ_{Dis} of ~39 nm^{-2}, moderate segregation of Cr ($\Gamma_{\text{Cr}} = $ ~5.3 nm^{-2}), and weak segregation of Mn ($\Gamma_{\text{Mn}} = $ ~0.8 nm^{-2}). (D) MC/MD-simulated GB structure of the equimolar CrMnFeCoNi and the corresponding disorder parameter (η_{Dis}) and atomic density profiles. In comparison with the same GB in the ternary CrMnCr alloy, this GB in the quinary Cantor alloy is more disordered with a larger Γ_{Dis} of ~43 nm^{-2} and strong co-segregation of Cr and Mn ($\Gamma_{\text{Cr}} = $ ~18.3 nm^{-2} and $\Gamma_{\text{Cr}} = $ ~7.0 nm^{-2}, which represent ~3.5× and ~9× increases, respectively, from those in the ternary alloy). More examples and further discussion can be found in Supplementary Discussion 4.
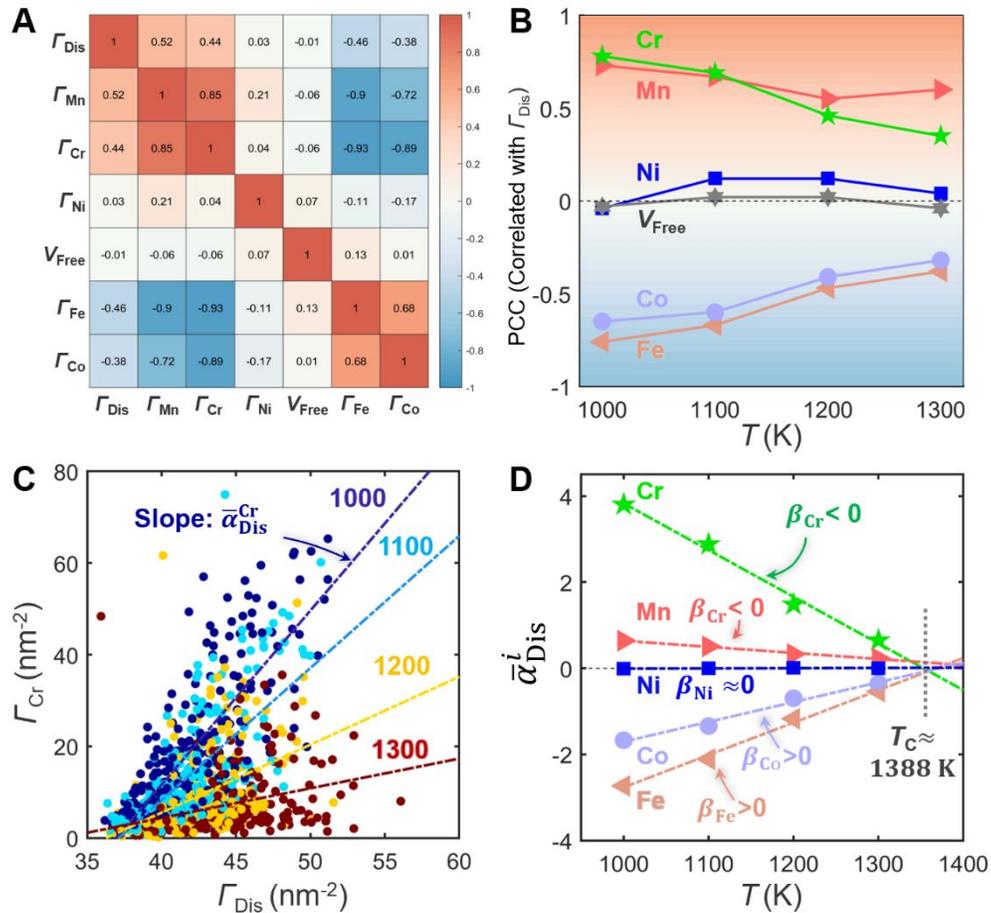
Fig. 4. **Correlation analyses of GB thermodynamic properties.** (A) Heat map of Pearson correlation coefficients (PCCs) between all pairs of the seven GB properties. (B) Calculated correlation coefficients between GB excess of disorder (Γ_{Dis}) and six other GB properties (GB excesses of Cr, Fe, Co, Ni and Mn, as well as GB free volume) at different temperatures. (C) Plots of GB excess of Cr (Γ_{Cr}) *vs*. GB excess of disorder (Γ_{Dis}) at 1000 K, 1100 K, 1200 K, and 1300 K for 258 different HEAs. The dashed lines are regression lines of Γ_{Cr} *vs*. Γ_{Dis} at four different temperatures, where the statistical linear correlation coefficient is denoted as $\bar{\alpha}_{Dis}^{Cr}$. The plot for other four GB adsorption properties *vs*. Γ_{Dis} can be found in Suppl. Fig. S10. (D) The fitted $\bar{\alpha}_{Dis}^{i}$ (i = Cr, Mn, Fe, Co, Ni) as functions of temperature (T) for five elements. The slopes of $\bar{\alpha}_{Dis}^{i}$ *vs*. T regression lines are labelled as β_i. Notably, all five fitted linear lines cross over at nearly one point on the horizontal T-axis ($\bar{\alpha}_{Dis}^{i} = 0$) at $T_C \approx 1388$ K.
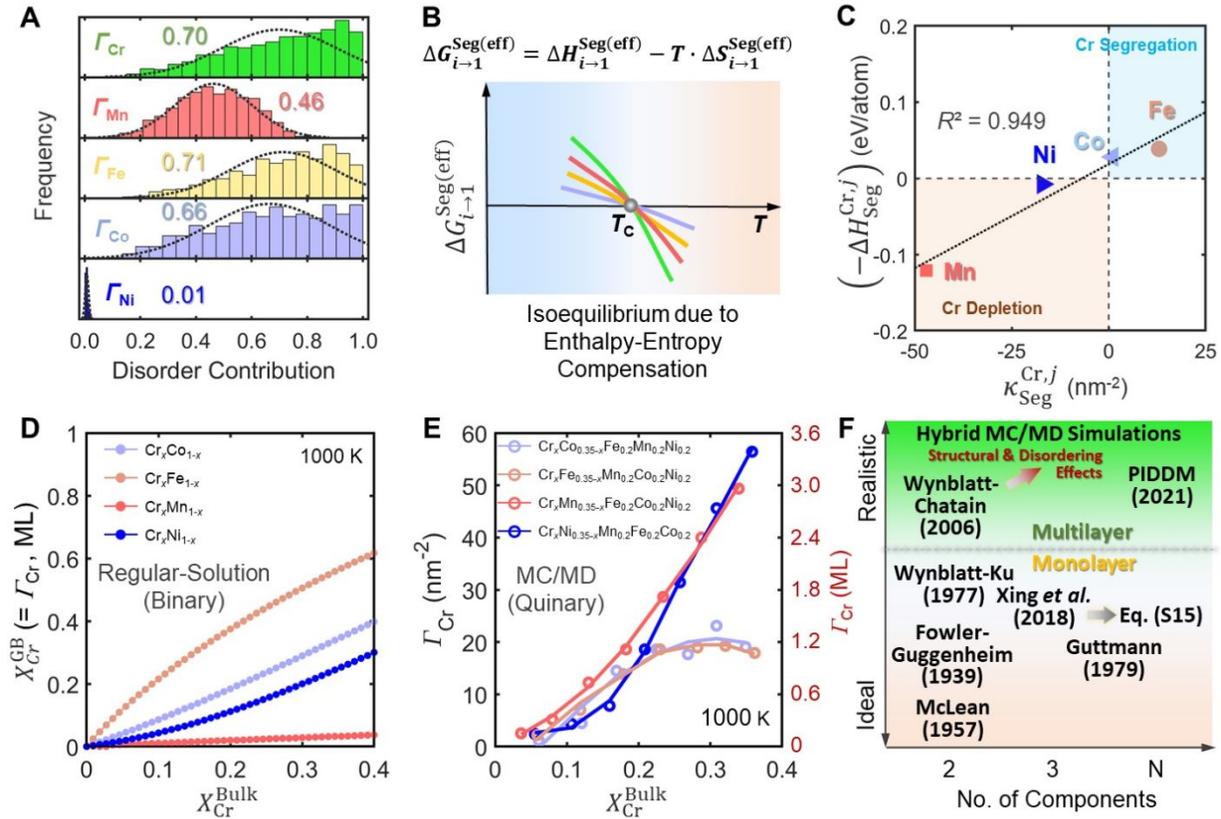
Fig. 5. **Analyses and comparisons of the classical segregation models, MC/MD simulations, and PIDDM.** (A) Histograms of the disorder contribution to GB adsorption of each element based on the DBAM. (B) Schematic illustration of an isoequilibrium effect[40] due to a perfect enthalpy-entropy compensation (albeit this compensation is likely incomplete in the current case). (C) Parity plot of $\Delta H_{Cr,j}^{Seg}$ (the segregation enthalpy of Cr in the binary Cr-j alloy, where j = Mn, Fe, Co, and Ni, calculated used a lattice-type model[44]) *vs*. $\kappa_{Cr,j}^{Seg}$ (the compositional coupling coefficients in the PIDDM). The positive (or negative) values of $\Delta H_{Cr,j}^{Seg}$ or $\kappa_{Cr,j}^{Seg}$ indicate Cr is favorable (or unfavorable) to segregate. (D) Calculated GB adsorption ($X_{Cr}^{GB} = \Gamma_{Cr}$) *vs*. the bulk Cr fraction (X_{Cr}^{bulk}) for four Cr$_x$-j_{1-x} (j = Mn, Fe, Co, and Ni) binary alloys at 1000 K using the Wynblatt-Ku model[47]. (E) MC/MD-simulated Γ_{Cr} *vs*. X_{Cr}^{bulk} for four HEAs at 1000 K. The compositions are noted in the legend, where in each case the increase in the Cr fraction is compensated by one selected element while keeping the fractions of the three other elements at the constant level of 0.2. (F) Comparison of various GB segregation models discussed in the text and supplementary discussions.

27. E. W. Hart, *Scripta Metallurgica*, 1968, **2**, 179-&.

28. C. Hu, Y. Zuo, C. Chen, S. Ping Ong and J. Luo, *Materials Today*, 2020, **38**, 49-57.

29. C. Hu, Y. Li, Z. Yu and J. Luo, *npj Computational Materials*, 2021, **7**, 159.

30. N. Zhou, C. Hu and J. Luo, *Acta Materialia*, 2021, **221**, 117375.

31. N. Zhou, T. Hu and J. Luo, *Current Opinion in Solid State and Materials Science*, 2016, **20**, 268-277.

32. A. Morawiec and K. Glowinski, *Acta Materialia*, 2013, **61**, 5756-5767.

33. F. Otto, A. Dlouhý, K. G. Pradeep, M. Kuběnová, D. Raabe, G. Eggeler and E. P. George, *Acta Materialia*, 2016, **112**, 40-52.

34. A. L. S. Chua, N. A. Benedek, L. Chen, M. W. Finnis and A. P. Sutton, *Nature Materials*, 2010, **9**, 418-422.

35. L. Li, Z. Li, A. Kwiatkowski da Silva, Z. Peng, H. Zhao, B. Gault and D. Raabe, *Acta Materialia*, 2019, **178**, 1-9.

36. P. M. Larsen, S. Schmidt and J. Schiøtz, *Modelling and Simulation in Materials Science and Engineering*, 2016, **24**, 055007.

37. S. Chen, Z. H. Aitken, S. Pattamatta, Z. Wu, Z. G. Yu, R. Banerjee, D. J. Srolovitz, P. K. Liaw and Y.-W. Zhang, *Acta Materialia*, 2021, **206**, 116638.

38. R. I. Babicheva, S. V. Dmitriev, Y. Zhang, S. W. Kok, N. Srikanth, B. Liu and K. Zhou, *Computational Materials Science*, 2015, **98**, 410-416.

39. P. Wynblatt and D. Chatain, *Metallurgical and Materials Transactions a-Physical Metallurgy and Materials Science*, 2006, **37A**, 2595-2620.

40. L. Liu and Q.-X. Guo, *Chemical Reviews*, 2001, **101**, 673-696.

41. P. Lejček, S. Hofmann, M. Všianská and M. Šob, *Acta Materialia*, 2021, **206**, 116597.

42. J. Luo and X. M. Shi, *Applied Physics Letters*, 2008 **92**, 101901

43. J. Luo, *Journal of the American Ceramic Society*, 2012, **95**, 2358-2371.

44. H. A. Murdoch and C. A. Schuh, *Journal of Materials Research*, 2013, **28**, 2154-2163.

45. J. Du Plessis and G. Van Wyk, *Journal of Physics and Chemistry of Solids*, 1988, **49**, 1441-1450.

46. P. Lejček and S. Hofmann, *Journal of Materials Science*, 2021, **56**, 7464-7473.

47. P. Wynblatt and R. C. Ku, *Surface Science*, 1977, **65**, 511-531.

48. W. Xing, A. R. Kalidindi, D. Amram and C. A. Schuh, *Acta Materialia*, 2018, **161**, 285-294.

49. W.-M. Choi, Y. H. Jo, S. S. Sohn, S. Lee and B.-J. Lee, *npj Computational Materials*, 2018, **4**, 1.

50. H. Ogawa, *MATERIALS TRANSACTIONS*, 2006, **47**, 2706-2710.

51. S. Plimpton, *Journal of Computational Physics*, 1995, **117**, 1-19.

52. J. Nie, C. Hu, Q. Yan and J. Luo, *Nature Communications*, 2021, **12**, 2374.

53. G. Kresse and J. Hafner, *Phys. Rev. B*, 1993, **47**.

54. G. Kresse and G. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169-11186.

55. P. E. Blöchl, *Phys. Rev. B* 1994, **50**, 17953-17979.

56. G. Kresse and D. Joubert, *Phys. Rev. B* 1999, **59**.

57. Z. Leong, J. S. Wróbel, S. L. Dudarev, R. Goodall, I. Todd and D. Nguyen-Manh, *Scientific Reports*, 2017, **7**, 39803.

58. T. A. Manz, *RSC Advances*, 2017, **7**, 45552-45581.