



Full Length Article



Exploration of V–Cr–Fe–Co–Ni high-entropy alloys with high yield strength: A combination of machine learning and molecular dynamics simulation

Lu Chen ^{a,b}, Asker Jarlöv ^{b,c}, Hang Li Seet ^c, Mui Ling Sharon Nai ^c, Yefei Li ^{a,*}, Kun Zhou ^{b,*}^a State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an 710049, China^b School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore^c Singapore Institute of Manufacturing Technology, Additive Manufacturing Division, 5 Cleantech Loop, #01-01 Cleantech Two Block B, Singapore 636732, Singapore

ARTICLE INFO

ABSTRACT

Keywords:

Machine learning
Molecular dynamics
High-entropy alloy
Alloy design
Bayesian optimization

Improving the strength of Cr–Fe–Co–Ni high-entropy alloys is a key issue in expanding their applicability. Herein, a framework combining machine learning and molecular dynamics is employed to improve the yield strength of Cr–Fe–Co–Ni high-entropy alloys through vanadium addition. The results indicate that by specifying the valence electron concentration, vacancy formation energy, and mismatch in cohesive energy as input features, a support vector regression model with a radial bias function kernel displays the highest performance among numerous combinations of machine learning models and material features. According to the Shapley additive explanation, the vacancy formation energy and valence electron concentration show a negative correlation with the yield strength above 1.66 eV and 7.60, respectively, and a positive correlation otherwise. The mismatch in the cohesive energy always shows a negative correlation. By utilizing a Bayesian adaptive alloy design, $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$ has been identified to have the highest yield strength. The simulated tensile deformation of polycrystalline high-entropy alloys confirms the predicted trend in yield strength, and the events observed during plastic deformation are consistent with previous experimental observations. The proposed framework provides a promising prospect for accelerating the design of high-entropy alloys with reduced dependence on costly trial-and-error experiments.

1. Introduction

With the rapid development of modern industry, traditional alloys can no longer meet the ever-increasing application requirements. High-entropy alloys (HEAs) are the result of accelerated material discovery where multiple elements are mixed in near equiatomic ratios (5–35 at%) [1,2]. As an extensively studied HEA, the CrFeCoNi-based HEA exhibits impressive mechanical and thermodynamic properties, including excellent ductility and fracture toughness at cryogenic temperatures [3], superior strength and stability at high temperatures [4], and high wear-, corrosion-, and fatigue-resistance [5]. However, the low yield strength at room temperature hinders their application as practical structural materials [6–8]. The addition of new alloying elements, such as Al [9], V [10–12], Cu [13], Mo [14], and Ta [15], is a heavily explored strengthening method, where V has been identified as a suitable strength-enhancing candidate, owing to its larger atomic radius and misfit volume, thus providing solid solution hardening [16].

However, the multi-element nature of HEAs introduces huge

complexity in optimizing the chemical composition, where the formation of secondary phases may severely restrict their application. In V–Cr–Fe–Co–Ni HEAs, both V and Cr can destabilize the face-centered cubic (FCC) structure in favor of either a body-centered cubic (BCC) or tetragonal structure (sigma phase), thereby increasing the strength but severely reducing the ductility [12]. The chemical composition of the tetragonal structure can be described as $A_{50}B_{50}$, where A refers to elements from groups VB–VIIIB and B to elements from groups VIIIB–VIIIB in the periodic table [17]. For V–Cr–Fe–Co–Ni HEAs, V and Cr represent the A elements while Fe, Co, and Ni represent the B elements. To combat the issue of undesired secondary phase formation, novel alloy design strategies must be adopted to predict the phase composition before predicting the optimal chemical composition.

Compared to expensive and time-consuming trial-and-error experiments, machine learning (ML), which models the relationship between material features and properties based on preexisting datasets, have recently been used to optimize the composition of different alloys. The effectiveness of ML was demonstrated by accelerating the discovery of

* Corresponding authors.

E-mail addresses: liyfei@xjtu.edu.cn (Y. Li), kzhou@ntu.edu.sg (K. Zhou).

novel aluminum alloys with enhanced strength, and the designed alloy had higher tensile strength than commercial alloys [18]. ML was also utilized to design copper alloys with both high strength and electrical conductivity, and the results were in good agreement with experiments [19].

High-precision datasets are critical to the robustness and performance of ML. However, the limited experimental studies on V-containing HEAs are insufficient to construct accurate datasets. Fortunately, high-throughput molecular dynamics (MD) simulations are well suited to produce datasets for ML because of their larger temporal and spatial scales as compared to other atomistic simulation techniques [20]. Additionally, MD has been able to capture key phenomena in HEAs including the effect of chemical composition on the stacking fault energy [21], the deformation behavior under cyclic loading [22], and the origin of nanoscale amorphization at cryogenic temperatures [23]. The synergy between ML and MD was recently capitalized to optimize the composition of Cr-Mn-Fe-Co-Ni and Cr-Co-Ni HEAs [24,25]. However, only the relationship between the chemical composition and the strength was considered, and no deeper understanding was provided regarding the underlying features that govern the strengthening. Cr-Mn-Fe-Co-Ni and Cr-Co-Ni HEAs have been extensively studied [26], and there is a need to go beyond such alloys.

In our current study, a yield strength-oriented framework is proposed, combining ML and MD, to reveal the mechanical and phase features that significantly affect the yield strength. Utilizing the proposed design strategy, the composition of FCC-structured V-Cr-Fe-Co-Ni HEAs is optimized to enhance the yield strength. Simulated tensile deformation of polycrystalline HEAs is also performed using MD and compared against experimental results to validate the framework.

2. Methodology

2.1. Yield strength-oriented machine learning framework

The yield strength-oriented ML framework consists of five main processes including data collection, feature pool construction, model and feature combination screening, adaptive alloy design, and validation using MD simulations, as shown in Fig. 1.

High-quality data is a prerequisite for high-performance ML models. As an excess of Cr and V might trigger the transformation from an FCC structure to either a BCC or tetragonal structure [12,17], the phase stability is calculated using MD simulations to ensure that the FCC structure is the most stable structure (see Section 2.2 for further details). The database is then constructed by performing 223 simulated tensile deformations of FCC-structured V-Cr-Fe-Co-Ni HEAs.

The concentration C_i and the related characteristic f_i of element i are

used to calculate the average F for the mechanical and phase features of each HEA composition as

$$F = \sum_{i=1}^n C_i f_i, \quad (1)$$

where n is the number of elements in the alloy (five in the current work).

The difference δ_F and mismatch values D_F of the corresponding features are obtained by

$$\delta_F = \sqrt{\sum_{i=0}^n C_i \left(1 - \frac{f_i}{F}\right)^2}, \quad (2)$$

$$D_F = \sum_{i=1}^n \sum_{j=1, i \neq j}^n C_i C_j |f_i - f_j|. \quad (3)$$

Some features, such as the configurational entropy ΔS , geometrical descriptors γ and Λ , the phase stability descriptor Ω related to the enthalpy, only have F values, and their respective formulas are listed in Table S1, similar to what was done in [27–29]. A total of 47 mechanical and phase features are included in the feature pool.

An optimal combination of model and features is crucial for the performance of ML. Nine classic regression models, including linear regression (LIN), k-nearest neighbor (KNN), support vector regression model with radial bias function kernel (SVR.r), linear kernel (SVR.l), and polynomial kernel (SVR.p), decision tree (DT), random forest regressor (RFR), gradient descent regressor (GBR), and extreme gradient boosting regressor (XGBR) are utilized. A three-step screening procedure is used to identify suitable features in each model that achieve a good balance between model performance and complexity, including correlation screening, recursive elimination, and exhaustive elimination. The first step is to calculate the Pearson correlation coefficient r between any two features. Once $|r|$ is larger than 0.95, the features are considered linearly related [30], and only the feature with a higher influence on the yield strength is retained. In the subsequent recursive elimination, one material feature is omitted, and the remaining features are used to build the model. The set of features that results in the lowest model error is retained, and the process is repeated until the error between two subsequent iterations increases. Finally, exhaustive elimination enumerates all possible subsets of remaining features to identify the optimal combination. The dataset is split into 80% training and 20% testing data, and the model error is calculated by fivefold cross validation. Scikit-learn is employed to perform all ML calculations [32].

Bayesian optimization enables the efficient design of alloy compositions by balancing the exploitation (i.e., searching for HEAs with the highest predicted yield strength) and exploration (i.e., exploring uncertain areas) of the search space. In this work, a utility function is employed to choose three HEA compositions for the next MD

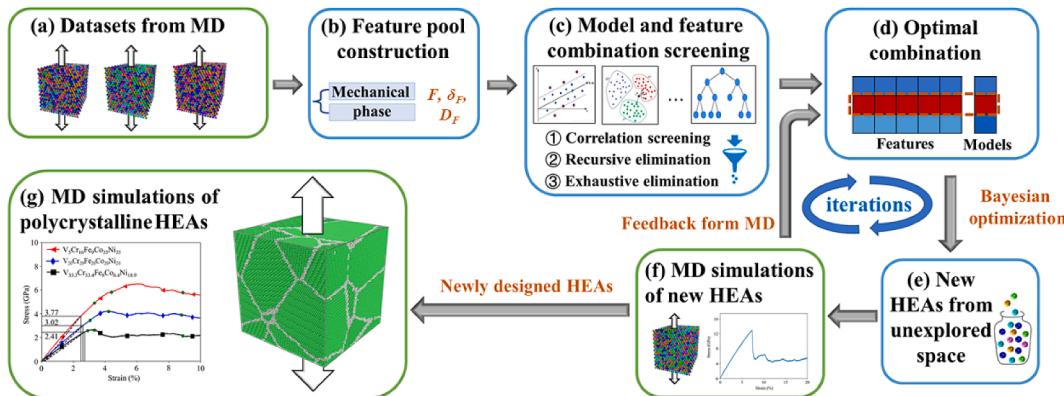


Fig. 1. Schematic diagram of the yield strength-oriented framework based on ML (blue boxes) and MD (green boxes). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

simulations and then iteratively improve the ML model by incorporating the data of the selected HEA compositions, material features, and yield strength into the dataset. The Expected Improvement (EI) function is the most widely used utility function [28] and is implemented as.

$$\text{EI}(z) = \sigma[\varphi(z) + \phi(z)], \quad (4)$$

$$z = (\mu - \mu^*)/\sigma, \quad (5)$$

where $\varphi(z)$ is the distribution function, $\phi(z)$ is the standard normal density, μ and σ are the predicted mean yield strength and standard deviation, respectively, and μ^* is the current maximum yield strength in the dataset. The μ and σ values of each HEA composition are obtained by the Gaussian process [29], which can find the global maximum within a small number of iterations. From each iteration, three sets of HEA composition with the highest $\text{EI}(z)$ value are then selected for subsequent MD simulations, and the newly obtained yield strength are added into the dataset for the next iteration.

To ensure that the predicted trend in the yield strength is consistent for more complex simulation cells that more closely resemble actual engineering materials, simulated uniaxial tensile deformation are performed on polycrystalline HEAs using MD simulations. The V₅Cr₁₆Fe₉Co₃₅Ni₃₅ (predicted to have the highest yield strength by the framework), V₂₀Cr₂₀Fe₂₀Co₂₀Ni₂₀, and V_{18.9}Cr_{33.3}Fe_{33.4}Co₈Ni_{6.4} (predicted to have among the lowest yield strength in the initial dataset) compositions are selected as polycrystalline tensile samples for MD (see additional details in Section 2.2).

2.2. Molecular dynamics simulations

MD simulations are employed to obtain the range of V and Cr content in V–Cr–Fe–Co–Ni HEAs across which the FCC structure is the most stable structure, as previous experimental works reveal that BCC and tetragonal structures are detrimental to the ductility [12,17]. Additionally, MD simulations of tensile deformation are performed, both for single-crystalline and polycrystalline simulation cells, to provide the dataset for the ML algorithm and to validate the results, respectively.

The range of V and Cr content for which the FCC structure is the most stable structure is determined by calculating the phase stability of V–Cr–Fe–Co–Ni HEAs based on the cohesive energy E_c . This value is calculated for different V–Cr–Fe–Co–Ni compositions by varying the V plus Cr content between 0 and 75 at% with increments of 5 at% while keeping the remaining three elements at an equiatomic ratio, as previous

experimental studies highlight the importance of these elements in destabilizing the FCC structure [17]. Three structures are considered: FCC, BCC, and tetragonal structures, with dimensions and orientations as shown in Fig. 2a–c. For each composition, five simulations with different atomic configurations are performed and averaged. The tetragonal structure is similar to what has been reported in experimental studies of V-containing HEAs [12,17]. Periodic boundary conditions are applied in all three directions. The FCC and BCC structures are solely prepared using the Atomsk software [32] whereas the tetragonal structure is prepared using the crystallographic information file obtained from ref. [33].

MD simulations are performed in the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software [34] using the modified embedded atom method (MEAM) potential developed by Choi et al. [35]. The interatomic potential is validated as described in the Supplementary Information. The energy of the simulation cell is minimized by the conjugate gradient (CG) algorithm, followed by thermal relaxation at 300 K in the isothermal-isobaric (NPT) ensemble for 60 ps before E_c is obtained.

The single-crystalline simulation cells used for uniaxial tensile deformation have dimensions and orientations as shown in Fig. 2d. After energy minimization by the CG algorithm and thermalization in the NPT ensemble at 300 K for 60 ps, tensile deformation at the same temperature is performed with a strain rate of $5 \times 10^8 \text{ s}^{-1}$ along the y-axis up to a strain of 20%. The high strain rate is a requirement in MD simulations due to the limited temporal scale [21,36]. A high strain rate will increase the yield strength in both FCC-[37,38] and BCC-structured [39,40] HEAs, meaning that it is not suitable to directly compare quantitative values obtained from MD simulations against those obtained from experiments. However, the qualitative trends observed during MD simulations still agree well with those observed in experiments, and the fundamental deformation mechanisms have been successfully investigated using MD simulations [23,41–43].

The polycrystalline simulation cells are prepared using the same structure to prevent differences in grain orientations from affecting the results. The structure is obtained by the Voronoi tessellation and consists of six grains with an average grain size of 170 Å. The dimension and orientation of the structure are provided in Fig. 2e. The conditions imposed on the polycrystalline simulation cells during tensile deformation are the same as those for the single-crystalline simulation cells described above. The Common Neighbor Analysis [44] and the dislocation analysis method [45] in the Open Visualization Tool (OVITO)

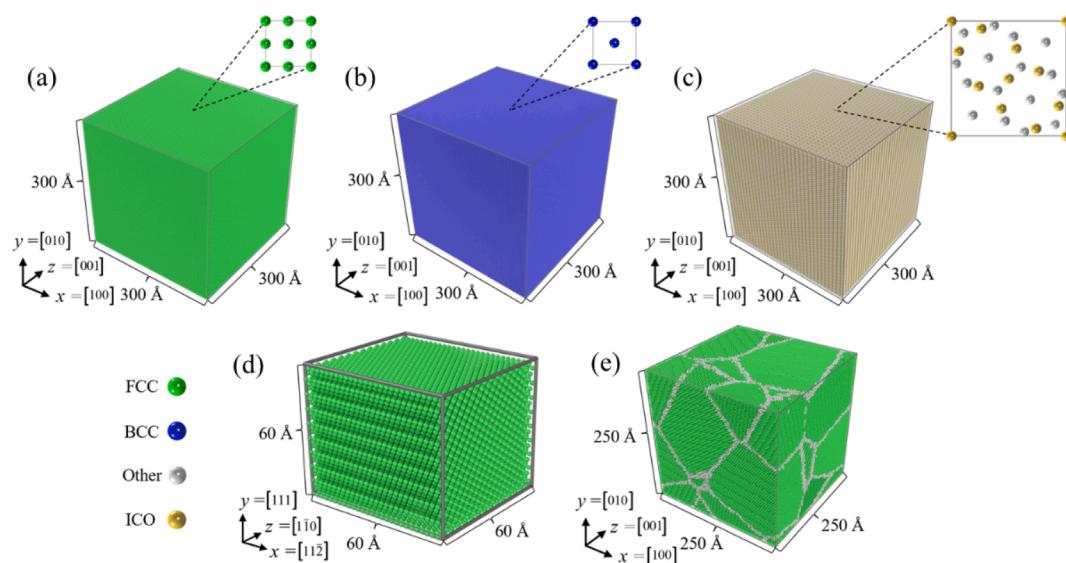


Fig. 2. Simulation cells used in the MD simulations to calculate E_c for the (a) FCC, (b) BCC, and (c) tetragonal structures, (d) for tensile deformation of single-crystalline HEAs, and (e) for tensile deformation of polycrystalline HEAs.

software [46] are used to analyze the deformation process.

3. Results

3.1. Phase stability of V–Cr–Fe–Co–Ni high-entropy alloys

Fig. 3 shows the variations in E_c for the FCC, BCC, and tetragonal structures of $(\text{VCr})_x(\text{FeCoNi})_{100-x}$ HEAs (x ranging from 0 to 75 at%) at 300 K. The interatomic potential used in the present work has been validated to reproduce the elastic constants and lattice parameters of the BCC and tetragonal structures for the binary elemental combinations [35]. To ensure that the interatomic potential also produces accurate results for multi-element systems, the lattice parameters of the BCC VCrFeCoNi and tetragonal VCrFeNi HEAs are calculated and compared against experimental values. The interatomic potential predicts an a value of 2.865 Å for the BCC VCrFeCoNi HEA and a and c values of 4.618 and 8.942 Å, respectively, for the tetragonal VCrFeNi HEA (both structures are relaxed at 0 K). This is in good agreement with experimental studies, where a is equal to 2.865–2.866 Å for the BCC VCrFeCoNi [11], while a and c are equal to 4.570 and 9.153 Å, respectively, for the tetragonal VCrFeNi [47].

Fig. 3 shows that the FCC structure has lower E_c when the V and Cr content is less than 55 at%, indicating that it is the most stable phase. Above 55 at%, both the BCC and tetragonal structures instead become the phases with the lowest E_c . Thus, to prevent the formation of the BCC and tetragonal structures, the V plus Cr content is restricted to be less than 55 %. To further limit the search space, the original definition of HEAs is used, *i.e.*, each element must have an atomic concentration between 5 and 35 at% [26].

3.2. Model and feature combination screening

The Pearson correlation map between any two features is shown in **Fig. S2**. In total, 17 features are eliminated because they are linearly related to other features but have a low influence on the yield strength. The SVR.r model is used as an example for the remaining screening process. **Fig. 4a** shows the trend of the model error during the recursive elimination process. Initially, a steep decrease is observed, indicating that the eliminated features have a low influence on the yield strength. When the number of features is reduced to eight, subsequent elimination increases the model error, indicating the importance of the remaining features. As shown in **Fig. 4b**, decreasing the number of features from eight to three during the exhaustive screening process only slightly

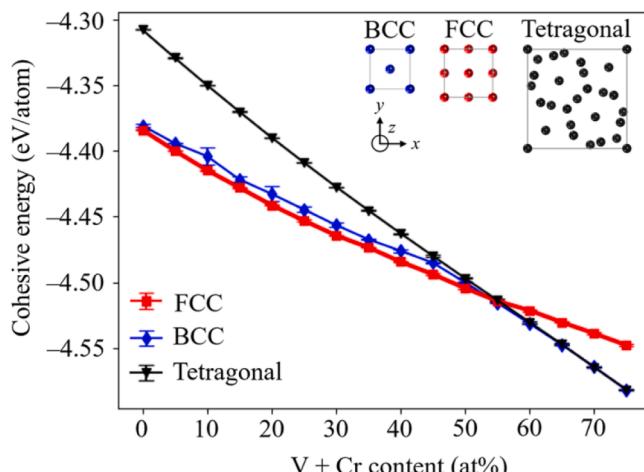


Fig. 3. Values of E_c for the FCC, BCC, and tetragonal structures of $(\text{VCr})_x(\text{FeCoNi})_{100-x}$ HEAs at 300 K. The insets show one unit cell of the FCC, BCC, and tetragonal structures. The x , y , and z axis of the unit cells are oriented along the [100], [010], and [001] directions, respectively.

increases the model error while it significantly reduces the model complexity. Hence, the number of input features is limited to three in the following analysis. The three input features used for the SVR.r model are the valence electron concentration VEC, vacancy formation energy E_v , and mismatch in cohesive energy D_{Ec} .

Table 1 lists the most influential features for different ML models, along with the mean squared error (MSE). The feature E_v is used for nearly all ML models, indicating that it has a significant influence on the yield strength. The SVR.r, SVR.l, and RFR models show the lowest MSE and variation on the training data and thus the best performance (**Fig. 4c**). The SVR.r model also shows the lowest MSE, lowest mean absolute percentage error (MAPE), and highest coefficient of determination R^2 for the testing data (**Fig. 4d**). Thus, the SVR.r model using E_v , VEC, and D_{Ec} as the input parameters is selected as the model used in the adaptive alloy design iteration to screen for HEAs with high yield strength in the virtual space.

3.3. Adaptive alloy design

According to the compositional upper and lower boundaries of HEAs whose yield strength exceeds 15 GPa in the dataset, a limited compositional candidate space is designed to improve the search efficiency. The composition of the five elements in the $\text{V}_a\text{Cr}_b\text{Fe}_c\text{Co}_d\text{Ni}_e$ HEAs is limited to 5–11 at% for a , 10–35 at% for b , 9–28 at% for c , 20–35 at% for d , and 20–35 at% for e . **Table 2** lists the elemental composition and the yield strength of the HEAs obtained from the adaptive iterative loop. Note that the listed yield strength is higher than that observed in experiments because of the high strain rate employed in MD simulations [37,38]. All newly designed HEAs show better performance than the initial dataset. The $\text{V}_5\text{Cr}_{16}\text{Fe}_9\text{Co}_{35}\text{Ni}_{35}$ HEA, suggested during the second and fourth iterations, has a yield strength of 18.28 GPa, which is 11.5 % higher than the highest value in the initial dataset (16.4 GPa).

Fig. 5 shows a scatterplot of the measured and predicted yield strength of the new HEAs obtained after each iteration, represented by filled and empty circles, respectively. As the adaptive alloy design progresses, the gap between the measured and predicted values decreases, indicating that the robustness of the adaptive model is gradually improving.

During the fourth iteration, the composition of the newly designed HEAs stabilizes with the Co and Ni content reaching its upper limits and the Fe content reaching its lower limit. This agrees with experimental results, where an increase in the Co content at the expense of the Fe content increased the yield strength [48]. However, the interatomic potential might not be able to reproduce the effect of adding a great amount of V on the yield strength because of an overestimation of the strengthening effect of the other constituent elements.

As the iteration progresses further, the yield strength starts to decline, and one of the suggested HEA compositions in the fifth iteration loop exhibits a measured yield strength that is almost as low as the yield strength in the training dataset. These observations indicate that the adaptive iterative search has located the global maximum, and thus the iteration process is stopped.

3.4. MD simulation of polycrystalline samples

Fig. 6 shows the MD validation of the developed framework using polycrystalline tensile samples. Because of the high strain rate, the observed stress and dislocation densities are much higher than those observed during experiments [21,38–40]. Additionally, apart from grain boundaries, the simulation cells are almost defect-free, which will further contribute to the higher stress [36]. Because of these discrepancies between MD simulations and experiments, the analysis of the MD results will focus on the qualitative trend instead of the quantitative values.

The deformation of all HEAs starts with an elastic stage, where the stress is linearly proportional to the strain. The Young's modulus

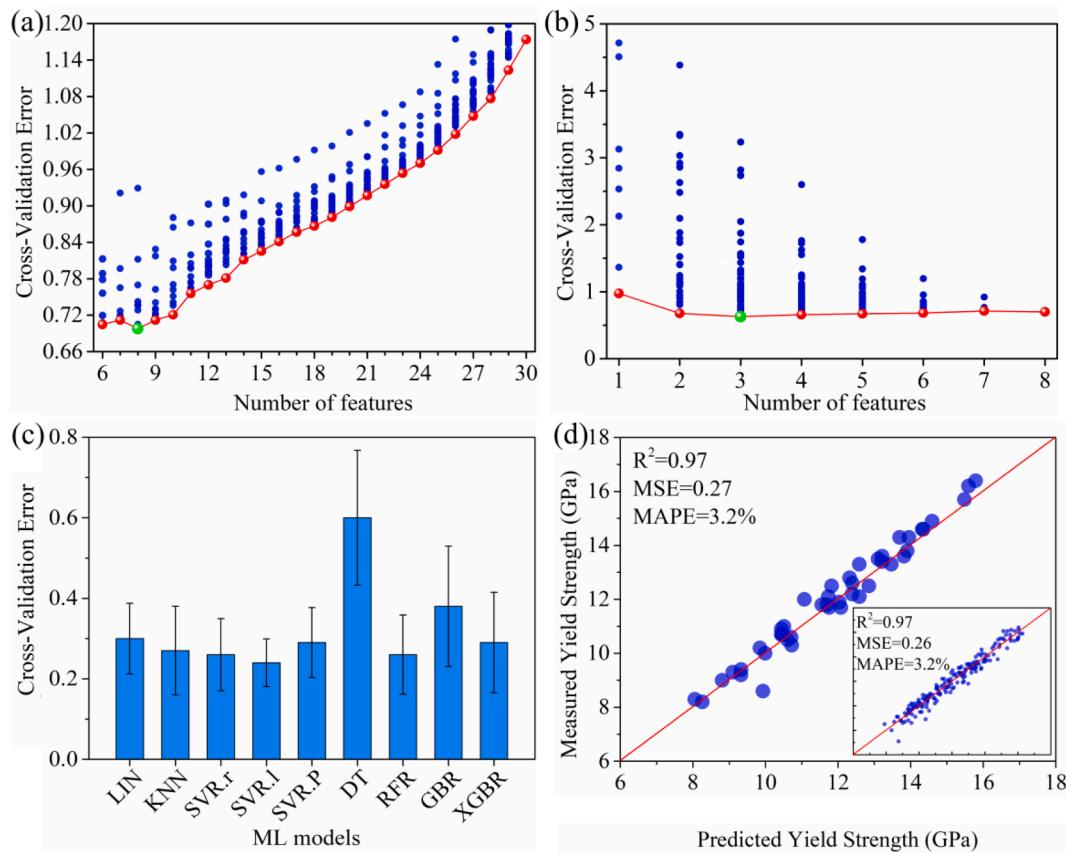


Fig. 4. ML model and material feature combination: (a) recursive elimination process, (b) exhaustive screening process, (c) model error of all the tested ML models, and (d) scatterplot of the measured and predicted yield strength for the SVR.r model based on the testing data. The inset in (d) shows the performance of the training data.

Table 1

Most influential material features on the yield strength for different ML models as identified by the feature screening. See supplementary information for a closer description of each material feature.

Machine learning model	Feature	Mean squared error (training/testing data)
LIN	$E_v, E_c, \eta B$	0.30/0.36
KNN	E_v, w, δ_{EI-2nd}	0.27/0.38
SVR.r	E_v, VEC, D_Ec	0.26/0.27
SVR.I	$E_v, D_Ec, \delta_{EI-2nd}$	0.24/0.33
SVR.p	D_B, δ_{VEG}, w	0.29/0.40
DT	$E_v, D_{EI-1st}, \delta_{EV}$	0.60/0.59
RFR	$E_v, \delta_{EI-2nd}, D_{e/a}$	0.26/0.30
GBR	$E_v, D_{e/a}, \delta_{EV}$	0.38/0.48
XGBR	$E_v, \delta_{EI-2nd}, D_{EI-1st}$	0.29/0.31

decreases in the order of $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$, $V_{20}Cr_{20}Fe_{20}Co_{20}Ni_{20}$, and $V_{18.9}Cr_{33.3}Fe_{33.4}Co_8Ni_{6.4}$. The yield strength, calculated by the 0.2 % offset rule, follows the same trend as for the Young's modulus, in good agreement with the results from the ML framework. However, the yield strain decreases in the order of $V_{18.9}Cr_{33.3}Fe_{33.4}Co_8Ni_{6.4}$, $V_{20}Cr_{20}Fe_{20}Co_{20}Ni_{20}$, and $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$, revealing a strength–ductility trade-off commonly observed for metallic materials [49]. After yielding, the stress of all samples reaches a peak value, after which it decreases to a stable value known as the flow stress. The peak stress and strain are in the order from lowest to highest: $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$, $V_{20}Cr_{20}Fe_{20}Co_{20}Ni_{20}$, and $V_{18.9}Cr_{33.3}Fe_{33.4}Co_8Ni_{6.4}$.

Although the ML framework focuses on obtaining HEAs with high yield strength, an analysis of the plastic deformation and dislocation evolution is necessary for evaluating the suitability of the HEA as an

Table 2

Elemental composition and yield strength of the newly designed alloys after five iterations. For each iteration, three HEA compositions are suggested for the sub (a, b, and c).

Iterations	Alloy Composition (at%)					Yield strength (GPa)
	V	Cr	Fe	Co	Ni	
1a	6	31	9	34	20	17.84
1b	5	32	9	34	20	17.97
1c	5	30	9	35	21	18.20
2a	5	15	10	35	35	18.05
2b	5	16	9	35	35	18.28
2c	5	18	9	35	33	17.99
3a	6	16	9	35	34	17.33
3b	6	14	10	35	35	17.56
3c	6	15	9	35	35	17.59
4a	5	25	9	35	26	17.85
4b	5	17	9	35	34	17.93
4c	5	16	9	35	35	18.28
5a	5	21	9	35	30	17.74
5b	6	10	14	35	35	16.46
5c	5	10	15	35	35	17.09

engineering alloy and for further validating that the current theoretical approach agrees with experimental studies. Fig. 6b shows the evolution of stair-rod dislocations in the three HEAs. Stair-rod dislocations are known to contribute to higher flow stress by providing strain hardening in HEAs and are formed when Shockley partial dislocations on intersecting {111} planes interact [50,51]. The density of stair-rod dislocations in the three HEAs decreases in the order of $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$, $V_{20}Cr_{20}Fe_{20}Co_{20}Ni_{20}$, and $V_{18.9}Cr_{33.3}Fe_{33.4}Co_8Ni_{6.4}$, which agrees well

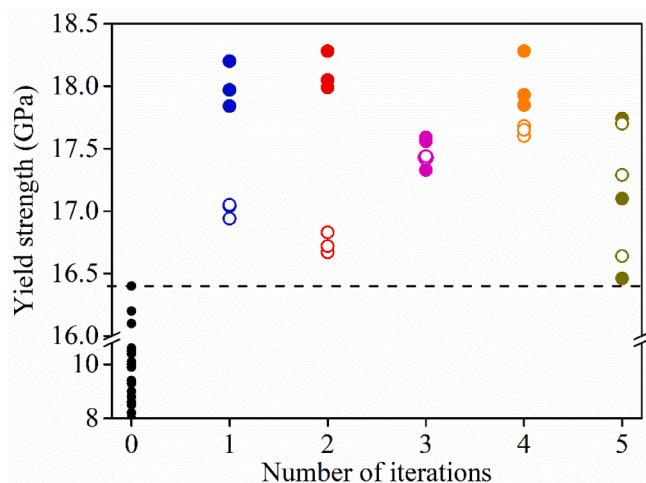


Fig. 5. Scatterplot of the measured and predicted yield strength as a function of the iteration number. The filled and empty circles represent the measured and predicted yield strength, respectively. The black points represent points in the training dataset, and the dashed line indicates the highest yield strength in the training dataset.

with the trend observed for the flow stress, indicating that the main strengthening mechanism in these alloys is the formation of stair-rod dislocations. Other dislocations, such as Hirth dislocations, which have been reported to contribute to strain hardening [36], are also observed, but their density is much lower.

Fig. 6c and d show snapshots taken during the deformation process. The selected grain has the same geometry for all three HEAs. Around the yield strain, the first Shockley partial dislocation nucleates on the FCC slip planes, *i.e.*, the {111} planes, and the dislocation density rapidly increases as the deformation progresses. Fig. 6c(i)–e(i) show that for all the composition, the first Shockley partial dislocation nucleates from the grain boundary on the same slip plane, indicating the grain boundaries' role as a dislocation source [25,36]. However, as the strain increases, the dislocation evolution for the three compositions starts to differ. While multiple Shockley partial dislocations nucleate from the grain boundaries and travel across the grains, leaving behind stacking faults of HCP atoms in the V_{18.9}Cr_{33.3}Fe_{33.4}Co₈Ni_{6.4} and V₂₀Cr₂₀Fe₂₀Co₂₀Ni₂₀ HEAs, the V₅Cr₁₆Fe₉Co₃₅Ni₃₅ HEA shows the greatest resistance to nucleation of Shockley partial dislocations for the investigated grain (Fig. 6c(ii)–(iv)). This resistance to dislocation nucleation provides further explanation as to why the yield strength is higher in the V₅Cr₁₆Fe₉Co₃₅Ni₃₅ HEA, as the nucleation of the first Shockley partial dislocation is closely associated with the yield strength in MD simulations [21,52].

Fig. 6c(iv) and e(iv) show that the transformation induced plasticity (TRIP) mechanism is active in V₅Cr₁₆Fe₉Co₃₅Ni₃₅ and V_{18.9}Cr_{33.3}Fe_{33.4}Co₈Ni_{6.4}, where the FCC structure transforms into a martensite phase, *i.e.*, four or more consecutive layers of HCP atoms [53]. The martensite phase is known to provide extensive work hardening and has been observed in experimental studies of V–Cr–Fe–Co–Ni HEAs during the early stages of tensile deformation [10,11]. Although the TRIP effect in V–Cr–Fe–Co–Ni HEAs is normally associated with a transformation from FCC to HCP, followed by HCP to BCC, this effect is normally suppressed at elevated temperatures [54] and high Ni content [10].

To summarize, the uniaxial tensile deformation of polycrystalline HEAs predicted using MD simulations shows that the V₅Cr₁₆Fe₉Co₃₅Ni₃₅ HEA possesses the highest yield strength, in good agreement with the ML framework, and a high density of stair-rod dislocations during the plastic deformation, thus providing high flow stress. Additionally, the dislocation evolution during the plastic deformation agrees well with previous experimental studies, lending further credibility to the current

theoretical approach. Future work will experimentally validate the proposed framework.

4. Discussion

To further clarify the relationship between the three input features in the SVR.r model (*i.e.*, the E_v , VEC, and D_{Ec}) and the yield strength, Shapley additive explanations (SHAP), an algorithm for interpreting the importance of features in complex ML models [55], is used to explain the model predictions. The distribution of the SHAP values for samples in the training data is plotted in Fig. 7a, where each point denotes a set of HEA composition. A positive or negative SHAP value for a feature indicates that an increase in the feature increases or decreases the yield strength, respectively. The mean SHAP value of a feature can therefore be regarded as the importance of that feature. As shown in Fig. 7b, this importance is in the order from lowest to highest: E_v , D_{Ec} , and VEC.

Fig. 7c shows the scatter plot of the SHAP values as a function of E_v . When E_v is larger than a critical value (1.66 eV), a negative correlation is found between E_v and the yield strength, which can be explained by lattice distortion: the smaller the E_v , the easier it is to form vacancies. The presence of vacancies will induce lattice distortion, thereby increasing the yield strength [26]. However, when E_v is sufficiently small, vacancy clusters are induced, which in turn can lead to a violent motion of dislocations, and the alloy enters the plastic stage earlier, thereby reducing the yield strength [56]. The E_v value and yield strength are positively correlated in this regime.

The second most important feature is D_{Ec} . The more elements there are with similar E_c , the smaller the D_{Ec} value for the HEA, as seen in Eq. (3). Based on this inference, the increase in the Co and Ni content will improve the yield strength of V–Cr–Fe–Co–Ni HEAs owing to their small difference in the elemental E_c value (4.39 and 4.44 eV, respectively). This conclusion is consistent with the above alloy design and previous references [21,25].

Finally, the negative correlation between the yield strength and VEC is supported by previous studies [57,58]. However, Fig. 7d shows that VEC has a positive correlation with the yield strength below 7.60, which matches the critical VEC value used to predict whether the HEA consists of an FCC or BCC structure. The HEA consists of a single FCC structure when the VEC is larger than 8, a dual FCC and BCC structure when the VEC is between 6.87 and 8, and a single BCC structure when the VEC is less than 6.87 [59]. Although our E_c calculation in Section 3.1 ensures a stable FCC structure, a lower VEC value translates to a higher propensity for the BCC structure to exist that might increase the yield strength. The higher strength of the BCC structure is attributed to the BCC structure having more severe lattice distortions than the FCC structure which results in significant solid solution strengthening [57]. Increasing the content of elements with lower VEC, such as Cr, will result in increased yield strength, which may also explain the above alloy search results. Alloys with higher Cr content obtained in the second and third iterations (2a, 2b, 3b, and 3c) have larger yield strength compared to alloys with lower Cr content in the fifth iteration (5b and 5c).

5. Conclusion

A yield strength-oriented ML framework is implemented by combining ML and MD to optimize the chemical composition of V–Cr–Fe–Co–Ni HEAs for high yield strength. The results are validated by performing simulated tensile deformation of polycrystalline samples using MD to compare against previous experimental studies. The following conclusions are made:

- Among numerous combinations of nine ML models and 47 material features, the SVR.r model, with E_v , D_{Ec} , and VEC as input features, performs best, with an accuracy of 97%.
- During the adaptive alloy design, the difference between the measured and predicted yield strength decreased, indicating

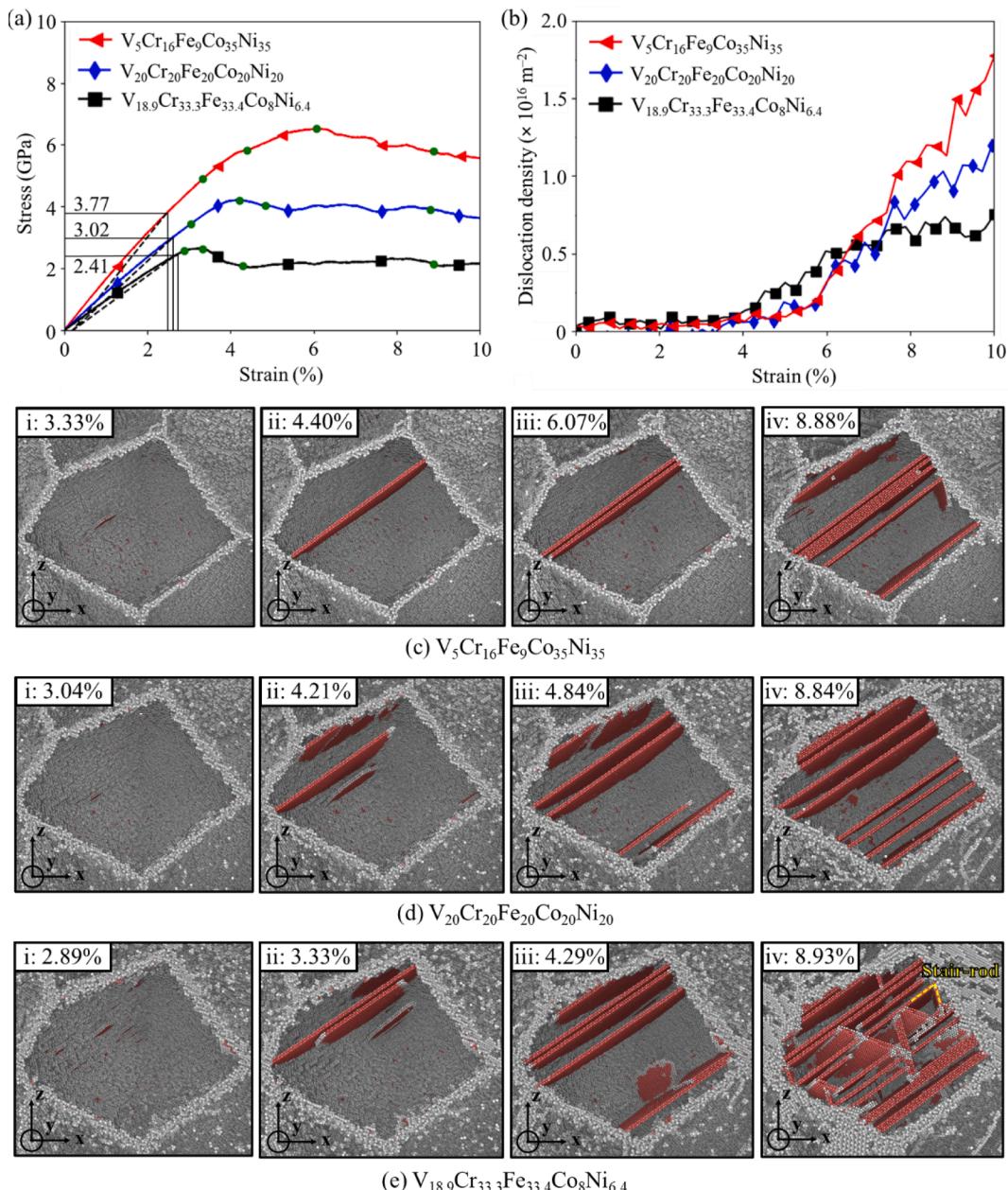


Fig. 6. Uniaxial tensile deformation of polycrystalline $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$, $V_{20}Cr_{20}Fe_{20}Co_{20}Ni_{20}$, and $V_{18.9}Cr_{33.3}Fe_{33.4}Co_8Ni_{6.4}$ HEAs: (a) stress–strain curves, (b) evolution of the stair-rod dislocations, and (c–e) snapshots of the dislocation process. The black dashed lines and the numbers in (a) represent the 0.2% strain offset rule and the yield strength, respectively. The green points along each curve represent the strain at which the snapshots in (c–e) are taken. To better observe the dislocation evolution in (c–e), atoms with an FCC structure are not shown, while atoms with HCP and amorphous structures are colored red and gray, respectively. The size of the atoms belonging to the grain boundaries is set to be 50% larger. The x , y , and z axis in (c–e) are oriented along the $[101]$, $[010]$, and $[\bar{1}01]$ directions, respectively. The yellow dashed lines in e(iv) represent a stair-rod dislocation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

increased robustness of the ML model as the number of iterations increased. The $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$ HEA is identified to have the highest yield strength, with an 11.5% increase as compared to the HEA with the highest yield strength in the initial dataset.

- The polycrystalline tensile deformation shows a similar trend as the ML framework, with the $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$ HEA showing the highest yield strength due to greater resistance to dislocation nucleation. Additionally, the $V_5Cr_{16}Fe_9Co_{35}Ni_{35}$ HEA shows high flow stress due to the formation of a high density of stair-rod dislocations and HCP martensite, in agreement with previous experimental studies of V–Cr–Fe–Co–Ni HEAs.

- The SHAP algorithm reveals that the influence of each material feature on the yield strength is in the order from lowest to highest: E_v , D_{Ec} , and VEC. The E_v and VEC values show a negative correlation with the yield strength above 1.66 eV and 7.60, respectively, and a positive correlation otherwise. The D_{Ec} value always shows a negative correlation with the yield strength, and thus, high concentration of elements with similar E_c , such as Co and Ni, will increase the yield strength.

With the continuous development of accurate interatomic potentials for new HEA composition, it is expected that the ML and MD framework can be employed to optimize the composition in a wider range of HEAs.

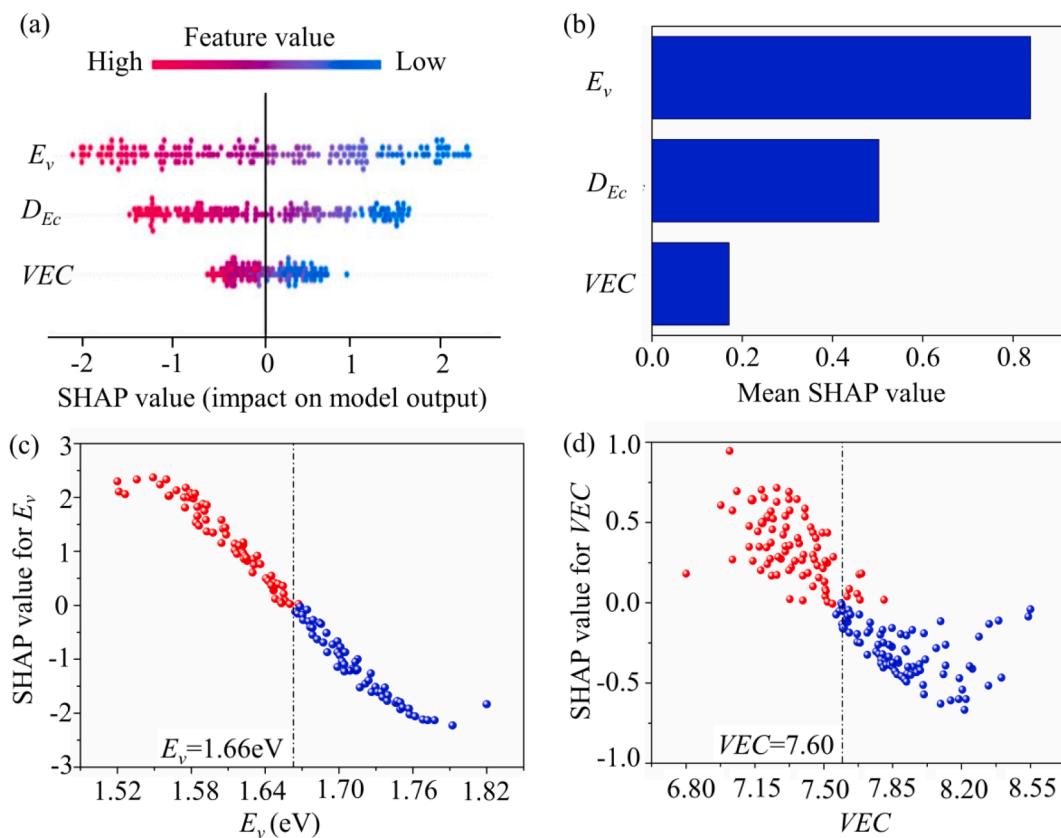


Fig. 7. Influences of material features on the yield strength via the SHAP analysis: (a) distribution of SHAP values for different samples, (b) importance ranking of E_v , D_{Ec} , and VEC , and (c, d) scatterplot of the SHAP values with varying (c) E_v , and (d) VEC . Each point represents a specific HEA composition in the training data. Red and blue colors represent positive and negative correlations, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Additionally, the flexibility in focusing on other mechanical properties, such as hardness, wear resistance, and flow stress, by adapting the MD simulations enables the framework to optimize alloys for numerous applications. Future work should use the suggested framework to guide experiments and accelerate the design of novel HEAs.

CRediT authorship contribution statement

Lu Chen: Conceptualization, Methodology, Software, Data curation, Investigation, Visualization, Formal analysis, Writing – original draft. **Asker Jarlöv:** Conceptualization, Methodology, Software, Data curation, Investigation, Visualization, Formal analysis, Writing – original draft. **Hang Li Seet:** Conceptualization, Methodology, Resources, Funding acquisition, Writing – review & editing. **Mui Ling Sharon Nai:** Conceptualization, Methodology, Resources, Validation, Funding acquisition, Project administration, Supervision, Writing – review & editing. **Yefei Li:** Conceptualization, Methodology, Resources, Validation, Funding acquisition, Project administration, Supervision, Writing – review & editing. **Kun Zhou:** Conceptualization, Methodology, Software, Resources, Validation, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

Lu Chen and Asker Jarlöv contributed equally to this manuscript. The authors are grateful for the high-performance computing resource offered by the National Supercomputing Center, Singapore (<https://www.nscc.sg>). Asker Jarlöv is grateful for the ongoing research scholarship from the A*STAR Graduate Academy.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.commatsci.2022.111888>.

References

- [1] J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, S.-Y. Chang, Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes, *Adv. Eng. Mater.* 6 (5) (2004) 299–303, <https://doi.org/10.1002/adem.200300567>.
- [2] B. Cantor, I.T.H. Chang, P. Knight, A.J.B. Vincent, Microstructural development in equiatomic multicomponent alloys, *Mater. Sci. Eng., A* 375–377 (2004) 213–218, <https://doi.org/10.1016/j.msea.2003.10.257>.
- [3] F. Otto, A. Dlouhý, C.h. Somsen, H. Bei, G. Eggeler, E.P. George, The influences of temperature and microstructure on the tensile properties of a CoCrFeMnNi high-entropy alloy, *Acta Mater.* 61 (2013) 5743–5755, <https://doi.org/10.1016/j.actamat.2013.06.018>.
- [4] Y. Zhang, T.T. Zuo, Z. Tang, M.C. Gao, K.A. Dahmen, P.K. Liaw, Z.P. Lu, Microstructures and properties of high-entropy alloys, *Prog. Mater. Sci.* 61 (2014) 1–93, <https://doi.org/10.1016/j.pmatsci.2013.10.001>.

- [5] Y.F. Kao, T.D. Lee, S.K. Chen, Y.S. Chang, Electrochemical passive properties of $\text{Al}_x\text{CoCrFeNi}$ ($x = 0, 0.25, 0.50, 1.00$) alloys in sulfuric acids, *Corros. Sci.* 52 (2010) 1026–1034, <https://doi.org/10.1016/j.corsci.2009.11.028>.
- [6] A. Gali, E.P. George, Tensile properties of high- and medium-entropy alloys, *Intermetallics* 39 (2013) 74–78, <https://doi.org/10.1016/j.intermet.2013.03.018>.
- [7] 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 (2014) 428–441, <https://doi.org/10.1016/j.actamat.2014.08.026>.
- [8] B. Gludovatz, A. Hohenwarter, D. Catoor, E.H. Chang, E.P. George, R.O. Ritchie, A fracture-resistant high-entropy alloy for cryogenic applications, *Science* 345 (6201) (2014) 1153–1158, <https://doi.org/10.1126/science.1254581>.
- [9] T.F. Yang, Q.X. Song, L. Shi, C. Wang, S. Liu, Y. Zhang, et al., Effects of Al addition on microstructure and mechanical properties of $\text{Al}_{1-x}\text{CoCrFeNi}$ High-entropy alloy, *Mater. Sci. Eng., A* 648 (2015) 15–22, <https://doi.org/10.1016/j.msea.2015.09.034>.
- [10] Y.H. Jo, W.M. Choi, D.G. Kim, A. Zargaran, S.S. Sohn, H.S. Kim, B.J. Lee, N.J. Kim, S. Lee, FCC to BCC transformation-induced plasticity based on thermodynamic phase formation in novel $\text{V}_{10}\text{Cr}_{10}\text{Fe}_{45}\text{Co}_x\text{Ni}_{35-x}$ medium-entropy alloys, *Sci. Rep.* 9 (1) (2019), <https://doi.org/10.1038/s41598-019-39570-y>.
- [11] Y.H. Jo, D.G. Kim, M.C. Jo, K.-Y. Doh, S.S. Sohn, D. Lee, H.S. Kim, B.J. Lee, S. Lee, Effects of deformation-induced BCC martensitic transformation and twinning on impact toughness and dynamic tensile response in metastable VCrFeCoNi high-entropy alloy, *J. Alloy. Compd.* 785 (2019) 1056–1067, <https://doi.org/10.1016/j.jallcom.2019.01.293>.
- [12] N.D. Stepanov, D.G. Shaysultanov, G.A. Salishchev, M.A. Tikhonovsky, E. Oleynik, A.S. Tortika, O.N. Senkov, Effect of V content on microstructure and mechanical properties of the CoCrFeMnNi_x high entropy alloys, *J. Alloy. Compd.* 628 (2015) 170–185, <https://doi.org/10.1016/j.jallcom.2014.12.157>.
- [13] Z.J. Xu, Z.T. Li, Y. Tong, W.D. Zhang, Z.G. Wu, Microstructural and mechanical behavior of a CoCrFeNiCu_4 non-equiautomic high entropy alloy, *J. Mater. Sci. Technol.* 60 (2021) 35–43, <https://doi.org/10.1016/j.jmst.2020.03.078>.
- [14] W.P. Li, X.G. Wang, B. Liu, Q.H. Fang, C. Jiang, Fracture mechanisms of a Mo alloyed CoCrFeNi high entropy alloy: in-situ SEM investigation, *Mater. Sci. Eng., A* 723 (2018) 79–88, <https://doi.org/10.1016/j.msea.2018.03.032>.
- [15] H. Jiang, K.M. Han, D.X. Qiao, Y.P. Lu, Z.Q. Cao, T.J. Li, Effects of Ta addition on the microstructures and mechanical properties of CoCrFeNi high entropy alloy, *Mater. Chem. Phys.* 210 (2018) 43–48, <https://doi.org/10.1016/j.matchemphys.2017.05.056>.
- [16] B.L. Yin, F. Maresca, W.A. Curtin, Vanadium is an optimal element for strengthening in both fcc and bcc high-entropy alloys, *Acta Mater.* 188 (2020) 486–491, <https://doi.org/10.1016/j.actamat.2020.01.062>.
- [17] M.-H. Tsai, K.-Y. Tsai, C.-W. Tsai, C. Lee, C.-C. Juan, J.-W. Yeh, Criterion for sigma phase formation in Cr- and V-containing high-entropy alloys, *Mater. Res. Lett.* 1 (4) (2013) 207–212, <https://doi.org/10.1080/21663831.2013.831382>.
- [18] J. Li, Y. Zhang, X. Cao, Q.i. Zeng, Y.e. Zhuang, X. Qian, H. Chen, Accelerated discovery of high-strength aluminum alloys by machine learning, *Commun. Mater.* 1 (1) (2020), <https://doi.org/10.1038/s43246-020-00074-2>.
- [19] C.S. Wang, H.D. Fu, L. Jiang, D.Z. Xue, J.X. Xie, A property-oriented design strategy for high performance copper alloys via machine learning, *NPJ Comput. Mater.* (2019) 87, <https://doi.org/10.1038/s41524-019-0227-7>.
- [20] J. Li, Q.H. Fang, P.K. Liaw, Microstructures and properties of high-entropy materials: modeling, simulation, and experiments, *Adv. Eng. Mater.* 23 (2020), <https://doi.org/10.1002/adem.202001044>, 2001044–1–2001044–30.
- [21] A. Jarlöv, W. Ji, Z. Zhu, Y. Tian, R. Babicheva, R. An, H.L. Seet, M.L.S. Nai, K. Zhou, Molecular dynamics study on the strengthening mechanisms of Cr-Fe-Co-Ni high-entropy alloys based on the generalized stacking fault energy, *J. Alloy. Compd.* 905 (2022) 164137, <https://doi.org/10.1016/j.jallcom.2022.164137>.
- [22] J. Peng, F. Li, B. Liu, Y. Liu, Q. Fang, J. Li, P.K. Liaw, Mechanical properties and deformation behavior of a refractory multiprincipal element alloy under cycle loading, *J. Micromech. Mole. Phys.* 05 (04) (2020) 2050014, <https://doi.org/10.1142/S2424913020500149>.
- [23] W. Ji, M.S. Wu, Nanoscale origin of the crystalline-to-amorphous phase transformation and damage tolerance of Cantor alloys at cryogenic temperatures, *Acta Mater.* 226 (2022), <https://doi.org/10.1016/j.actamat.2022.117639>, pp. 117639–1–11.
- [24] J. Li, B.B. Xie, Q.H. Fang, B. Liu, Y. Liu, P.K. Liaw, High-throughput simulation combined machine learning search for optimum elemental composition in medium entropy alloy, *J. Mater. Sci. Technol.* 68 (2021) 70–75, <https://doi.org/10.1016/j.jmst.2020.08.008>.
- [25] L. Zhang, K. Qian, J. Huang, M. Liu, Y. Shibuta, Molecular dynamics simulation and machine learning of mechanical response in non-equiautomic FeCrNiCoMn high-entropy alloy, *J. Mater. Res. Technol.* 13 (2021) 2043–2054, <https://doi.org/10.1016/j.jmrt.2021.06.021>.
- [26] D.B. Miracle, O.N. Senkov, A critical review of high entropy alloys and related concepts, *Acta Mater.* 122 (2017) 448–511, <https://doi.org/10.1016/j.actamat.2016.08.081>.
- [27] C. Wen, Y. Zhang, C. Wang, D. Xue, Y. Bai, S. Antonov, L. Dai, T. Lookman, Y. Su, Machine learning assisted design of high entropy alloys with desired property, *Acta Mater.* 170 (2019) 109–117, <https://doi.org/10.1016/j.actamat.2019.03.010>.
- [28] H. Zhang, H. Fu, S. Zhu, W. Yong, J. Xie, Machine learning assisted composition effective design for precipitation strengthened copper alloys, *Acta Mater.* 215 (2021), <https://doi.org/10.1016/j.actamat.2021.117118>, pp. 117118–1–10.
- [29] X. Huang, C. Jin, C. Zhang, H. Zhang, H. Fu, Machine learning assisted modelling and design of solid solution hardened high entropy alloys, *Mater. Des.* 211 (2021), <https://doi.org/10.1016/j.matdes.2021.110177>, pp. 110177–1–13.
- [30] R. Yuan, Z. Liu, P.V. Balachandran, D. Xue, Y. Zhou, X. Ding, J. Sun, D. Xue, T. Lookman, Accelerated discovery of large electrostrains in BaTiO_3 -based piezoelectrics using active learning, *Adv. Mater.* 30 (7) (2018) 1702884, <https://doi.org/10.1002/adma.201702884>.
- [31] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, Scikit-learn: machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830, <https://dl.acm.org/doi/abs/10.5555/1953048.2078195>.
- [32] P. Hirrel, Atomsk: a tool for manipulating and converting atomic data files, *Comput. Phys. Commun.* 197 (2015) 212–219, <https://doi.org/10.1016/j.cpc.2015.07.012>.
- [33] G. Bergman, D.P. Shoemaker, The determination of the crystal structure of the σ phase in the iron-chromium and iron-molybdenum systems, *Acta Crystallogr. A* 7 (1954) 857–865, <https://doi.org/10.1107/S0365110X54002605>.
- [34] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, *J. Comput. Phys.* 117 (1995) 1–19, <https://doi.org/10.1006/jcph.1995.1039>.
- [35] W.-M. Choi, J.-S. Kim, W.-S. Ko, D.G. Kim, Y.H. Jo, S.S. Sohn, S. Lee, B.-J. Lee, Computational design of $\text{V}-\text{CoCrFeMnNi}$ high-entropy alloys: an atomistic simulation study, *Calphad* 74 (2021) 102317, <https://doi.org/10.1016/j.calphad.2021.102317>.
- [36] W.R. Jian, Z.C. Xie, S.Z. Xu, Y.Q. Su, X.H. Yao, I.J. Beyerlein, Effects of lattice distortion and chemical short-range order on the mechanisms of deformation in medium entropy alloy CoCrNi , *Acta Mater.* 199 (2020) 352–369, <https://doi.org/10.1016/j.actamat.2020.08.044>.
- [37] F. Yuan, W. Cheng, S. Zhang, X. Liu, X. Wu, Atomistic simulations of tensile deformation in a CrCoNi medium-entropy alloy with heterogeneous grain structures, *Materialia* 9 (2020), <https://doi.org/10.1016/j.mtla.2019.100565>, pp. 100565–1–9.
- [38] C. Varvenne, A. Luque, W.A. Curtin, Theory of strengthening in fcc high entropy alloys, *Acta Mater.* 118 (2016) 164–176, <https://doi.org/10.1016/j.actamat.2016.07.040>.
- [39] F. Maresca, W.A. Curtin, Theory of screw dislocation strengthening in random BCC alloys from dilute to “High-Entropy” alloys, *Acta Mater.* 182 (2020) 144–162, <https://doi.org/10.1016/j.actamat.2019.10.007>.
- [40] F. Maresca, W.A. Curtin, Mechanistic origin of high strength in refractory BCC high entropy alloys up to 1900K, *Acta Mater.* 182 (2020) 235–249, <https://doi.org/10.1016/j.actamat.2019.10.015>.
- [41] W.M. Choi, Y.H. Jo, S.S. Sohn, S. Lee, B.J. Lee, Understanding the physical metallurgy of the CoCrFeMnNi high-entropy alloy: an atomistic simulation study, *NPJ Comput. Mater.* 4 (1) (2018) 1–9, <https://doi.org/10.1038/s41524-017-0060-9>.
- [42] C. Han, R. Babicheva, J.D.Q. Chua, U. Ramamurti, S.B. Tor, C.-N. Sun, K. Zhou, Microstructure and mechanical properties of $(\text{TiB}_2 + \text{TiC})/\text{Ti}$ composites fabricated in situ via selective laser melting of Ti and B4C powders, *Addit. Manuf.* 36 (2020) 101466, <https://doi.org/10.1016/j.addma.2020.101466>.
- [43] P. Wang, Y. Bu, J. Liu, Q. Li, H. Wang, W. Yang, et al., Atomic deformation mechanism and interface toughening in metastable high entropy alloy, *Mater. Today* 37 (2020) 64–73, <https://doi.org/10.1016/j.mattod.2020.02.017>.
- [44] A. Stukowski, Structure identification methods for atomistic simulations of crystalline materials, *Model. Simul. Mater. Sci. Eng.* 20 (4) (2012), <https://doi.org/10.1016/j.mattod.2020.02.017>, pp. 045021–1–16.
- [45] A. Stukowski, V.V. Bulatov, A. Arsenlis, Automated identification and indexing of dislocations in crystal interfaces, *Model. Simul. Mater. Sci. Eng.* 20 (8) (2012) 085007, <https://doi.org/10.1088/0965-0393/20/8/085007>.
- [46] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool, *Model. Simul. Mater. Sci. Eng.* 18 (1) (2010) 015012, <https://doi.org/10.1088/0965-0393/18/1/015012>.
- [47] Y.H. Jo, W.M. Choi, D.G. Kim, A. Zargaran, K. Lee, H. Sung, S.S. Sohn, H.S. Kim, B. J. Lee, S. Lee, Utilization of brittle σ phase for strengthening and strain hardening in ductile VCrFeNi high-entropy alloy, *Mater. Sci. Eng., A* 743 (2019) 665–674, <https://doi.org/10.1016/j.msea.2018.11.136>.
- [48] D. Wei, X. Li, W. Heng, Y. Koizumi, F. He, W.-M. Choi, B.-J. Lee, H.S. Kim, H. Kato, A. Chiba, Novel Co-rich high entropy alloys with superior tensile properties, *Mater. Res. Lett.* 7 (2) (2019) 82–88, <https://doi.org/10.1080/21663831.2018.1553803>.
- [49] S.Z. Han, E.A. Choi, S.H. Lim, S. Kim, J. Lee, Alloy design strategies to increase strength and its trade-offs together, *Prog. Mater. Sci.* 117 (2021), <https://doi.org/10.1016/j.pmatsci.2020.100720>, pp. 100720–1–51.
- [50] S.L. Wei, C.C. Tasan, Deformation faulting in a metastable CoCrNiW complex concentrated alloy: A case of negative intrinsic stacking fault energy? *Acta Mater.* 200 (2020) 992–1007, <https://doi.org/10.1016/j.actamat.2020.09.056>.
- [51] Q. Xiao, L. Wang, Y.J. Liang, Y.F. Xue, Annealing hardening in cryo-rolled high-entropy alloys by belated deformation twinning, *Mater. Sci. Eng. A* 801 (2021), <https://doi.org/10.1016/j.msea.2020.140403>, pp. 140403–1–6.
- [52] V. Borovikov, M.I. Mendelev, A.H. King, Effects of stable and unstable stacking fault energy on dislocation nucleation in nano-crystalline metals, *Model. Simul. Mater. Sci. Eng.* 24 (8) (2016) 085017, <https://doi.org/10.1088/0965-0393/24/8/085017>.
- [53] J.W. Xiao, N. Wu, O. Ojo, C. Deng, Stacking fault and transformation-induced plasticity in nanocrystalline high-entropy alloys, *J. Mater. Res.* 36 (2021) 2705–2714, <https://doi.org/10.1557/s43578-021-00140-6>.
- [54] Y.H. Jo, K.-Y. Doh, D.G. Kim, K. Lee, D.W. Kim, H. Sung, S.S. Sohn, D. Lee, H. S. Kim, B.-J. Lee, S. Lee, Cryogenic-temperature fracture toughness analysis of non-equi-atomic $\text{V}_{10}\text{Cr}_{10}\text{Fe}_{45}\text{Co}_{20}\text{Ni}_{15}$ high-entropy alloy, *J. Alloy. Compd.* 809 (2019) 151864, <https://doi.org/10.1016/j.jallcom.2019.151864>.
- [55] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, *Advances in Neural Information Processing Systems* 30 (2017) 4766–4775. <http://arxiv.org/abs/1705.07874>.

- [56] A. Durand, L. Peng, G. Laplanche, J.R. Morris, E.P. George, G. Eggeler, Interdiffusion in Cr-Fe-Co-Ni medium-entropy alloys, *Intermetallics* 122 (2020), <https://doi.org/10.1016/j.intermet.2020.106789>, pp. 106789-1–15.
- [57] C. Yang, C. Ren, Y.F. Jia, G. Wang, M.J. Li, W.C. Lu, A machine learning-based alloy design system to facilitate the rational design of high entropy alloys with enhanced hardness, 117431–1 117431-10, *Acta Mater.* 222 (2022), <https://doi.org/10.1016/j.actamat.2021.117431>.
- [58] R. Chen, G. Qin, H. Zheng, L. Wang, Y. Su, YuLung Chiu, H. Ding, J. Guo, H. Fu, Composition design of high entropy alloys using the valence electron concentration to balance strength and ductility, *Acta Mater.* 144 (2018) 129–137, <https://doi.org/10.1016/j.actamat.2017.10.058>.
- [59] S. Guo, C. Ng, J. Lu, C.T. Liu, Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys, *J. Appl. Phys.* 109 (2011) 103505-1–103505-6, <https://doi.org/10.1063/1.3587228>.

1 **Supplementary information for**

2 **Exploration of V–Cr–Fe–Co–Ni high-entropy alloys with high yield strength: a**
3 **combination of machine learning and molecular dynamics simulation**

4 Chen Lu^{1,2}, Asker Jarlöv^{2,3}, Hang Li Seet³, Mui Ling Sharon Nai³,
5 Yefei Li^{1,*}, Kun Zhou^{2,*}

6 ¹ *State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an,*
7 *710049, China*

8 ² *School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang*
9 *Avenue, 639798, Singapore*

10 ³ *Singapore Institute of Manufacturing Technology, 73 Nanyang Drive, 637662, Singapore*

11 **Table of contents**

- 12 1. Feature pool
13 2. Validation of the interatomic potential
14 3. Correlation screening

15 Supplementary Figures S1–S3

16 Supplementary Tables S1–S4

17 1. Feature pool

18 Table S1 shows the feature, description, and formula of all the 47 mechanical and phase
 19 features used.

20 **Table S1.** Detailed description and formula used to calculate each feature.

Mechanical Feature	Description	Formula
G	Shear modulus	
B	Bulk modulus	
ν	Poisson's ratio	$\sum_{i=1}^n C_i f_i$
R	Radius	
η_G	Shear module mismatch in strengthening model	$\frac{C_i}{f_i + F} \frac{2(f_i - F)}{f_i + F}$
η_B	Bulk module mismatch in strengthening model	$\sum_{i=1}^n \frac{2(f_i - F)}{1 + 0.5 \left \frac{2(f_i - F)}{f_i + F} \right }$
δ_G	Difference in shear modulus	
δ_B	Difference in bulk modulus	
δ_ν	Difference in Poisson's ratio	$\sqrt{\sum_{i=1}^n C_i (1 - \frac{f_i}{F})^2}$
δ_R	Difference of radius	
D_G	Local shear modulus mismatch	
D_B	Local bulk modulus mismatch	
D_ν	Local poisson's ratio mismatch	$\sum_{i=1}^n \sum_{j=1, i \neq j}^n C_i C_j f_i - f_j $
D_R	Local radius mismatch	
A	Energy term in the strengthening model	$G * \delta R * (1 + \nu) / (1 - \nu)$
μ	Lattice distortion energy	$\frac{1}{2} E * \delta R$
w	Work function to the power of six	$(\sum_{i=1}^n C_i W_i)^6$
Phase Feature	Description	Formula
X	Pauling electronegativity	
T_m	Melting point	
VEC	Valence electron concentration	
e/a	Free electron concentration	
E_c	Cohesive energy	$\sum_{i=1}^n C_i f_i$
E_v	Vacancy formation energy	
$E_{1\text{-}1st}$	First ionization energy	
$E_{1\text{-}2nd}$	Second ionization energy	
ΔS	Configurational entropy	$-R \sum_{i=1}^n C_i \ln(C_i)$
ΔH	Mixing enthalpy	$\sum_{i=1, j > i}^n 4C_i C_j \Delta H_{ij}^{\text{mix}}$

Table S1. Continued

δ_X	Difference in electronegativity	
δ_{Tm}	Difference in melting temperature	
δ_{VEC}	Difference in valence electron concentration	
$\delta_{e/a}$	Difference in free electron concentration	$\sqrt{\sum_{i=1}^n C_i (1 - \frac{f_i}{F})^2}$
δ_{Ec}	Difference in cohesive energy	
δ_{Ev}	Difference in vacancy formation energy	
δ_{EI-1st}	Difference in first ionization energy	
δ_{EI-2nd}	Difference in second ionization energy	
δ_{AH}	Difference in mixing enthalpy	$\sqrt{\sum_{i=1, i \neq j}^n C_i C_j (1 - \frac{\Delta H_{ij}^{mix}}{\Delta H})}$
D_X	Local electronegativity mismatch	
D_{Tm}	Local melting temperature mismatch	
D_{VEC}	Local valence electron concentration mismatch	
$D_{e/a}$	Local free electron concentration mismatch	$\sum_{i=1}^n \sum_{j=1, i \neq j}^n C_i C_j f_i - f_j $
D_{Ec}	Local cohesive energy mismatch	
D_{Ev}	Local vacancy formation energy mismatch	
D_{EI-1st}	Local first ionization energy mismatch	
D_{EI-2nd}	Local second ionization energy mismatch	
Ω	Phase stability descriptor	$T_m * \frac{\Delta S}{ \Delta H }$
Γ	Packing misfit parameter	$\frac{1 - \sqrt{\frac{(R + R_{\min})^2 - R^2}{(R + R_{\min})^2}}}{1 - \sqrt{\frac{(R + R_{\max})^2 - R^2}{(R + R_{\max})^2}}}$
Λ	Λ parameter	$\frac{\Delta S}{\Delta R^2}$

21 2. Validation of the interatomic potential

22 Table S2 shows the elastic constants C_{11} , C_{12} , and C_{44} for six selected V–Cr–Fe–Co–Ni
23 high-entropy alloys (HEAs) calculated by molecular dynamics (MD) simulations at 300 K
24 using standard straining methods and the modified embedded atom method (MEAM) potential
25 developed in ref. [1].

26 **Table S2.** Calculated elastic constants for six selected HEAs based on MD simulations. The
27 number (#) refers to the numbers used in Table S3 and Fig. S1.

Alloy	#	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)
Cr _{19.4} Fe _{12.3} Co _{34.3} Ni _{24.1} V _{9.9}	1	236.4	151.9	80.0
Cr ₁₈ Fe _{20.5} Co _{13.4} Ni _{19.1} V ₂₉	2	242.3	193.6	75.2
Cr _{19.3} Fe _{20.6} Co _{17.2} Ni _{27.4} V _{15.5}	3	170.0	170.0	79.6
Cr _{31.3} Fe ₅ Co _{34.3} Ni ₁₆ V _{22.6}	4	245.1	177.2	74.9
Cr _{20.7} Fe _{10.1} Co _{25.7} Ni _{13.5} V _{30.1}	5	245.0	182.5	70.7
Cr _{27.1} Fe _{21.8} Co _{25.4} Ni _{20.6} V _{5.1}	6	231.7	158.4	80.8

28 The elastic constants are then used to calculate the shear modulus G and Poisson's ratio ν
29 according to

$$G = (G_v + G_{\bar{v}})/2, \quad (1)$$

$$G_v = (C_{11} - C_{12} + 3C_{44})/5, \quad (2)$$

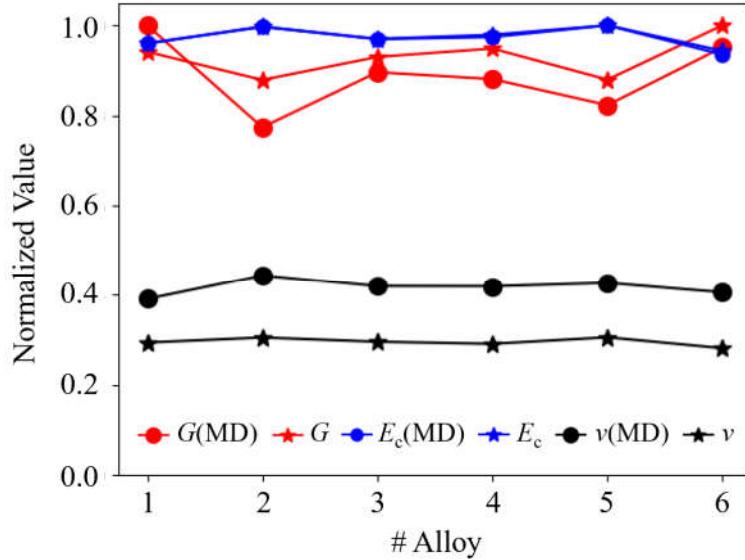
$$G_r = \frac{5C_{44}(C_{11} - C_{12})}{4C_{44} + 3(C_{11} - C_{12})}, \quad (3)$$

$$v = \frac{1}{1 + C_{11}/C_{12}}. \quad (4)$$

Table S3 shows ν , G , and the cohesive energy E_c (defined here as the energy required to separate the crystal into its building blocks) for the six HEAs in Table S2 calculated by MD simulations and the equations shown in Table S1. Figure S1 shows how the normalized values of ν , G , and E_c (calculated by dividing them by their largest value) change for the selected HEAs. The trend predicted by MD simulations agrees well with that from the equations in Table S1, so the interatomic potential is deemed reliable.

Table S3. Estimated v , G , and E_c obtained by both MD simulations and equations. The number (#) refers to the numbers used in Table S2 and Fig. S1.

#	G (MD)	G (Eq.)	v (MD)	v (Eq.)	E_c (MD)	E_c (Eq.)
1	61.92	81.09	0.3912	0.2941	4.4022	4.4233
2	47.91	75.71	0.4442	0.3053	4.5743	4.5916
3	55.48	80.10	0.4200	0.2959	4.4471	4.4677
4	54.54	81.70	0.4196	0.2913	4.4686	4.5097
5	50.95	75.77	0.4270	0.3057	4.5879	4.6069
6	58.90	86.14	0.4060	0.2816	4.2901	4.3447



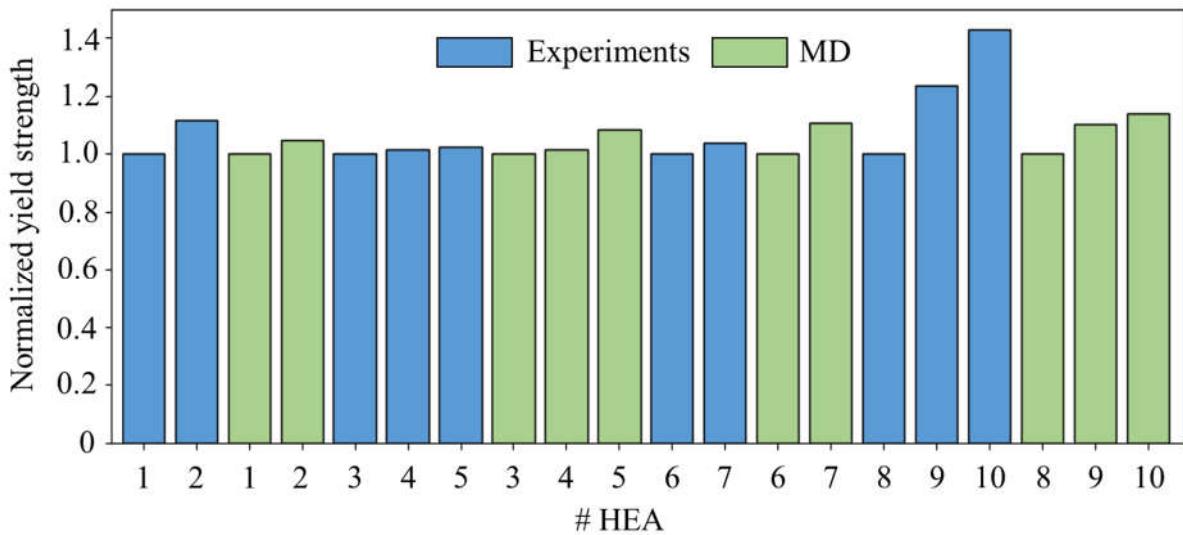
42 **Fig. S1.** Estimated v , G , and E_c obtained by both MD simulations and equations shown in
 43 Table S1. '# Alloy' refers to the number given to the HEAs in Table S1 and S2.

44 The interatomic potential is further validated by reproducing the qualitative trends of how
 45 the composition affects the yield strength observed in several experimental studies. Table 1
 46 shows the predicted yield strength by MD simulations for 10 V–Cr–Mn–Fe–Co–Ni HEAs as
 47 well as the experimentally measured YS. Note that only HEAs that consist of a single FCC
 48 structure prior to deformation, which is the focus of our work, are considered. Comparison
 49 should only be made between alloys from the same study due to the differences in the testing
 50 condition and processing route.

51 **Table S4.** Predicted and actual yield strength (YS) obtained from MD simulations and
 52 experiments, respectively. The '# HEA' refers to the labels used in Fig. S2.

# HEA	Composition	YS (MD) (GPa)	YS (Experiment) (MPa)	Ref.
1	V ₁₂ Cr ₂₀ Fe ₁₆ Co ₂₅ Ni ₂₇	13.906	543	[1]
2	V ₁₅ Cr ₂₀ Fe ₈ Co ₂₄ Ni ₃₃	14.596	607	[1]
3	V ₁₀ Cr ₁₀ Fe ₄₅ Co ₁₀ Ni ₁₅	12.833	339	[2]
4	V ₁₀ Cr ₁₀ Fe ₄₅ Co ₂₀ Ni ₁₅	13.030	345	[2]
5	V ₁₀ Cr ₁₀ Fe ₄₅ Co ₃₀ Ni ₅	13.934	348	[2]
6	V ₁₀ Cr ₁₀ Fe ₄₅ Co ₃₀ Ni ₅	13.934	359	[3]
7	V ₁₀ Cr ₁₀ Fe ₄₅ Co ₃₀ Ni _{2.5} Mn _{2.5}	12.597	345	[3]
8	Cr ₂₅ Fe ₂₅ Co ₂₅ Ni ₂₅	16.510	210	[4]
9	Cr ₂₅ Fe ₂₀ Co ₃₅ Ni ₂₀	18.174	260	[4]
10	Cr ₂₅ Fe ₁₅ Co ₄₅ Ni ₁₅	18.831	300	[4]

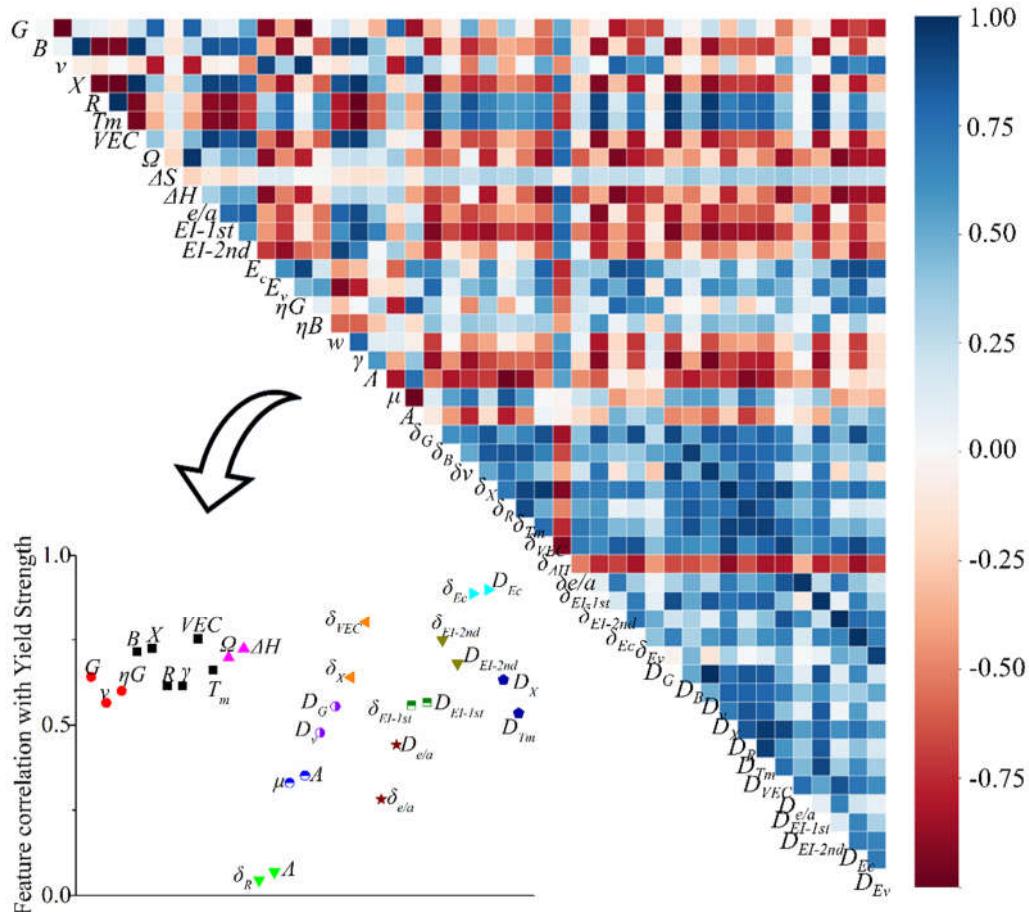
53 Due to the limited spatial and temporal scale of atomistic simulations, including MD
 54 simulations, a very high strain rate must be used (10^7 – 10^9 s $^{-1}$) as compared to what is actually
 55 used in experiments (10^{-3} – 10^0 s $^{-1}$). The high strain rate in combination with the small and
 56 defect-free simulation cells results in higher stress being recorded in MD simulations than in
 57 experiments, as well as an overestimation of the yield strength due to the stress overshoot [5–7].
 58 It is thus not possible to directly compare quantitative results obtained from MD directly with
 59 those obtained from experiments. However, by calculating the normalized yield strength for
 60 the MD simulations and experiments, respectively, the qualitative trend can be analyzed, which
 61 tends to show good agreement with experiments [8, 9]. Fig. S2 shows the normalized YS from
 62 MD simulations and experiments. The qualitative trend observed from experiments can be
 63 accurately captured by MD simulations.



64 **Fig. S2.** Normalised yield strength obtained by experiments and MD simulations.

65 3. Correlation screening

66 Figure S3 shows the Pearson correlation map between any two features, where blue
 67 and red represent positive and negative correlation, respectively. In the bottom left plot,
 68 features that are highly correlated with each other are indicated by the same color, and for
 69 each color, the feature with the highest correlation with the yield strength is retained.



70

Fig. S3. Pearson correlation map between any two features.

71 **References**

- 72
- 73 [1] Choi, W.M., Kim, J.S., Ko, W.S., Kim, D.G., Jo, Y.H., Sohn, S.S., *et al.*, Computational
74 design of V-CoCrFeMnNi high-entropy alloys: An atomistic simulation study.
75 *Calphad*, 2021. **74**: pp. 102317-1–9. <https://doi.org/10.1016/j.calphad.2021.102317>.
- 76
- 77 [2] Jo, Y.H., Choi, W.M., Kim, D.G., Zargaran A., Sohn S.S., Kim, H.S., *et al.*, FCC to
78 BCC transformation-induced plasticity based on thermodynamic phase stability in
79 novel V₁₀Cr₁₀Fe₄₅Co_x Ni_{35-x} medium-entropy alloys. *Scientific reports*, 2019. **9**: pp. 1–
80 14. <https://doi.org/10.1038/s41598-019-39570-y>.
- 81
- 82 [3] Yang, J., Jo, Y.H., Kim, D.W., Choi, W.M., Kim, H.S., Lee, B.J., *et al.*, Effects of
83 transformation-induced plasticity (TRIP) on tensile property improvement of
84 Fe₄₅Co₃₀Cr₁₀V₁₀Ni_{5-x}Mnx high-entropy alloys. *Materials Science and*
85 *Engineering: A*, 2020. **772**: 138809-1–9. <https://doi.org/10.1016/j.msea.2019.138809>.
- 86
- 87 [4] Wei, D., Li, X., Jiang, J., Heng, W., Koizumi, Y., Choi, W.M., *et al.* Novel Co-rich
88 high performance twinning-induced plasticity (TWIP) and transformation-induced
89 plasticity (TRIP) high-entropy alloys. *Scripta Materialia*, 2019. **165**: pp. 39–43.
90 <https://doi.org/10.1016/j.scriptamat.2019.02.018>.
- 91
- 92 [5] Fang, Q., Chen, Y., Li, J., Jiang, C., Liu, B., Liu, Y., *et al.* Probing the phase
93 transformation and dislocation evolution in dual-phase HEA. *International Journal of*
94 *Plasticity*, 2019. **114**: pp. 161–173. <https://doi.org/10.1016/j.ijplas.2018.10.014>.
- 95
- 96 [6] Jian, W.R., Xie, Z., Xu, S., Su, Y., Yao, X., and Beyerlein, I.J. Effects of lattice
97 distortion and chemical short-range order on the mechanisms of deformation in medium
98 entropy alloy CoCrNi. *Acta Materialia*, 2020. **199**: pp. 352–369.
99 <https://doi.org/10.1016/j.actamat.2020.08.044>.
- 100
- 101 [7] Yuan, F., Cheng, W., Zhang, S., Liu, X., and Wu, X. Atomistic simulations of tensile
102 deformation in a CrCoNi medium-entropy alloy with heterogeneous grain structures.
103 *Materialia*, 2020. **9**: pp. 100565-1–9. <https://doi.org/10.1016/j.mtla.2019.100565>
- 104
- 105 [8] Ji, W., and Wu, M.S., Nanoscale origin of the crystalline-to-amorphous phase
106 transformation and damage tolerance of Cantor alloys at cryogenic temperatures. *Acta*
107 *Materialia*, 2022: **226**: pp. 117639-1–11.
108 <https://doi.org/10.1016/j.actamat.2022.117639>.
- 109
- 110 [9] Choi, W.M., Jo, Y.H., Sohn, S.S., Lee, S., and Lee, B.J. Understanding the physical
111 metallurgy of the CoCrFeMnNi high-entropy alloy: an atomistic simulation study. *npj*
112 *Computational Materials*, 2018. **4**(1): pp. 1–9. <https://doi.org/10.1038/s41524-017-0060-9>.
- 113