

General-purpose machine-learned potential for 16 elemental metals and their alloys

Keke Song,^{1,*} Rui Zhao,^{2,*} Jiahui Liu,^{1,*} Yanzhou Wang,^{3,1} Eric Lindgren,⁴ Yong Wang,⁵ Shunda Chen,^{6,†} Ke Xu,⁷ Ting Liang,⁷ Penghua Ying,⁸ Nan Xu,^{9,10} Zhiqiang Zhao,¹¹ Jiuyang Shi,⁵ Junjie Wang,⁵ Shuang Lyu,¹² Zezhu Zeng,¹² Shirong Liang,¹³ Haikuan Dong,¹⁴ Ligang Sun,¹³ Yue Chen,¹² Zhuhua Zhang,¹¹ Wanlin Guo,¹¹ Ping Qian,¹ Jian Sun,^{5,‡} Paul Erhart,^{4,§} Tapio Ala-Nissila,^{3,15} Yanjing Su,^{1,¶} and Zheyong Fan^{14,**}

¹Beijing Advanced Innovation Center for Materials Genome Engineering,

University of Science and Technology Beijing, Beijing 100083, P. R. China

²School of Materials Science and Engineering, Hunan University, Changsha 410082, China

³MSP group, QTF Centre of Excellence, Department of Applied Physics,

P.O. Box 15600, Aalto University, FI-00076 Aalto, Espoo, Finland

⁴Chalmers University of Technology, Department of Physics, 41926 Gothenburg, Sweden

⁵National Laboratory of Solid State Microstructures,

School of Physics and Collaborative Innovation Center of Advanced Microstructures,

Nanjing University, Nanjing 210093, P. R. China

⁶Department of Civil and Environmental Engineering,

George Washington University, Washington, DC 20052, USA

⁷Department of Electronic Engineering and Materials Science and Technology Research Center,

The Chinese University of Hong Kong, Shatin, N.T., Hong Kong SAR, 999077, P. R. China

⁸Department of Physical Chemistry, School of Chemistry, Tel Aviv University, Tel Aviv, 6997801, Israel

⁹Institute of Zhejiang University-Qzhou, Quzhou 324000, P. R. China

¹⁰College of Chemical and Biological Engineering,

Zhejiang University, Hangzhou 310027, P. R. China

¹¹State Key Laboratory of Mechanics and Control of Mechanical Structures,

Key Laboratory for Intelligent Nano Materials and Devices of Ministry of Education,

and Institute for Frontier Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China

¹²Department of Mechanical Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, P. R. China

¹³School of Science, Harbin Institute of Technology, Shenzhen, 518055, P. R. China

¹⁴College of Physical Science and Technology, Bohai University, Jinzhou 121013, P. R. China

¹⁵Interdisciplinary Centre for Mathematical Modelling, Department of Mathematical Sciences,

Loughborough University, Loughborough, Leicestershire LE11 3TU, UK

(Dated: June 13, 2024)

Machine-learned potentials (MLPs) have exhibited remarkable accuracy, yet the lack of general-purpose MLPs for a broad spectrum of elements and their alloys limits their applicability. Here, we present a feasible approach for constructing a unified general-purpose MLP for numerous elements, demonstrated through a model (UNEP-v1) for 16 elemental metals and their alloys. To achieve a complete representation of the chemical space, we show, via principal component analysis and diverse test datasets, that employing one-component and two-component systems suffices. Our unified UNEP-v1 model exhibits superior performance across various physical properties compared to a widely used embedded-atom method potential, while maintaining remarkable efficiency. We demonstrate our approach's effectiveness through reproducing experimentally observed chemical order and stable phases, and large-scale simulations of plasticity and primary radiation damage in MoTaVW alloys. This work represents a significant leap towards a unified general-purpose MLP encompassing the periodic table, with profound implications for materials science.

Atomistic simulations of elemental metals and their alloys play a crucial role in understanding and engineering materials properties. While quantum-mechanical methods such as density-functional theory (DFT) calculations can be directly used for small simulation cells and short sampling times, their feasibility quickly diminishes with increasing spatial and temporal scales. For large-scale classical atomistic simulations, both molecular dynamics (MD) and Monte Carlo (MC) simulations crucially depend on interatomic potentials. For metallic systems in particular, embedded-atom method (EAM)-type potentials [1, 2] have proven to be useful and been extensively applied over the past decades, especially for elemental metals and their alloys. However, these existing classi-

cal interatomic potentials often lack the required level of accuracy for numerous applications. This deficiency primarily stems from constrained functional forms. Recently, a novel paradigm for developing interatomic potentials has emerged based on machine learning (ML) techniques [3–8]. In a machine-learned potential (MLP), the interatomic potential is modeled using ML methods, allowing for a significantly greater number of fitting parameters and providing versatility as compared to traditional many-body potentials. The functional forms of these MLPs are remarkably flexible, free from the limitations of a small number of analytical functions suggested by physical and chemical intuition or fitting to ground state properties only. The combination of flexible func-

tional forms and a large number of fitting parameters empowers MLPs to achieve a level of accuracy that can be well beyond that of the traditional many-body potentials.

The basic theory behind MLPs is rather mature now. There are two main ingredients of a MLP: the regression model and the descriptors as inputs to the regression model. For the construction of input descriptors, linearly complete basis functions for the atom-environments have been proposed [9, 10]. For the regression model, linear regression [9, 11], artificial neural network (NN) regression [12], and kernel-based regression [13] have all been proven to be feasible approaches. The combination of equivariant (as opposed to invariant) constructions and message passing or graph NNs [14, 15] has also shown great potential in enhancing the regression accuracy of MLPs, albeit at the cost of reduced computational efficiency and challenges in maintaining parallelism.

Despite the higher accuracy offered by MLPs, there are still challenges for applying MLPs in materials modeling, namely the relatively higher computational cost of many MLPs compared to most conventional many-body potentials, and the absence of readily usable databases of MLPs that cover a large number of elements and their compounds. In some cases where an extensive database is available, one can use an available MLP to study a specific problem, but in many cases, one has to train a new one or improve an existing one before being able to study the problem at hand. In particular, there is no simple way to combine MLPs for different elements to build MLPs for their compounds or alloys. This can lead to repeated efforts in the community and the case-by-case approach of developing MLPs is neither optimal nor sustainable in the long run. Regarding the computational cost of MLPs, the neuroevolution potential (NEP) approach [16–18] developed recently has been shown to yield excellent computational efficiency compared to other state-of-the-art methods, thanks to an optimization of the theoretical formalism and an efficient implementation in the GPUMD package [19]. The NEP approach can reach computational speeds unprecedented for MLPs, on par with empirical potentials, paving the way for the application of MLPs to large-scale atomistic simulations.

In this paper, we introduce a sustainable approach for the construction of MLPs. Although our approach can in principle be utilized to construct a comprehensive MLP covering the entire periodic table, we have chosen a more focused task as a proof of concept. Our objective is to develop a general-purpose NEP model encompassing 16 elemental metals and their alloys. Previous attempts to create general-purpose MLPs for numerous elements, or even the entire periodic table, have been initiated by researchers such as Takamoto *et al.* [20, 21] and Chen and Ong [22]. These studies have introduced “universal” MLPs, covering up to 45 elements [21] and 89 elements

[22], respectively. Despite being termed universal, these MLPs have a rather limited application range and are orders of magnitude slower than EAM potentials. General-purpose MLPs have only been conclusively demonstrated for elemental matter such as Si [23], C [24], Fe [25], and Pb [26]. For compounds comprising multiple chemical species, a special class of transition-metal oxides [27], binary Sn alloys with a few metals [28], and Si–O [29] have been successfully modeled using MLPs. However, for metallic alloys, it remains a highly nontrivial task to construct a unified MLP that can be reliably used for arbitrary chemical compositions. Here, our goal is to construct a genuinely general-purpose MLP for a diverse range of elements that matches the speed of EAM and surpasses it in the description of various physical properties.

Apart from achieving high accuracy and efficiency for the unified NEP model, which we term version 1 of unified NEP (UNEP-v1), we also propose an efficient approach for constructing the training dataset. Constructing a training dataset with all the possible chemical compositions is a formidable task. Fortunately, the NEP descriptor parameters depend only on pairs of elements. We will demonstrate that considering unaries and binaries alone for the training dataset is sufficient, yielding a NEP model that is transferable to systems with more components. Using this route, we achieve a transferable UNEP-v1 model for 16 elemental metals (Ag, Al, Au, Cr, Cu, Mg, Mo, Ni, Pb, Pd, Pt, Ta, Ti, V, W, Zr) and their diverse alloys, with only about 100 000 reference structures. This accomplishment is evidenced by accurate predictions of formation energies across various test datasets comprising multi-component alloy systems, reproduction of experimentally observed chemical order and stable phases, and the generality and high efficiency of our UNEP-v1 model in large-scale MD simulations of mechanical deformation and primary radiation damage in MoTaVW refractory high-entropy alloys.

RESULTS

A neural-network architecture for many-component systems. Our starting point is the NEP approach as described in Ref. 18, called NEP3. In this work, we introduce two crucial extensions to NEP3 designed specifically for many-component systems. This extended approach will be called NEP4, which has been implemented in GPUMD during the course of this work and is available starting from version 3.8.

We first briefly introduce NEP3 [18], which is a NN potential that maps a descriptor vector \mathbf{q}^i (with N_{des} components) of a central atom i to its site energy U_i . The total energy of a system of N atoms is expressed as the sum of the site energies $U = \sum_{i=1}^N U^i$. The ML model is a fully connected feedforward NN with a single

hidden layer with N_{neu} neurons,

$$U^i = \sum_{\mu=1}^{N_{\text{neu}}} w_{\mu}^{(1)} \tanh \left(\sum_{\nu=1}^{N_{\text{des}}} w_{\mu\nu}^{(0)} q_{\nu}^i - b_{\mu}^{(0)} \right) - b^{(1)}, \quad (1)$$

where $\tanh(x)$ is the activation function in the hidden layer, $\mathbf{w}^{(0)}$ is the connection weight matrix from the input layer (descriptor vector) to the hidden layer, $\mathbf{w}^{(1)}$ is the connection weight vector from the hidden layer to the output layer, $\mathbf{b}^{(0)}$ is the bias vector in the hidden layer, and $b^{(1)}$ is the bias in the output layer. Denoting the weight and bias parameters in the NN collectively as \mathbf{w} , we can formally express the site energy as

$$U^i = \mathcal{N}(\mathbf{w}; \mathbf{q}^i). \quad (2)$$

The descriptor vector consists of a number of radial and angular components. In this work, we utilize up to five-body angular components. For illustration purposes, we discuss the three-body angular components here. Interested readers are referred to Ref. 18 for the description of higher-order terms up to five-body angular components, which can help improve the model's completeness [30]. A three-body angular descriptor component can be expressed as

$$q_{nl}^i = \sum_{j \neq i} \sum_{k \neq i} g_n(r_{ij}) g_n(r_{ik}) P_l(\theta_{ijk}), \quad (3)$$

where n and l represent the order of the radial and angular expansions, respectively. Here, the summation runs over all neighbors of atom i within a certain cutoff distance, r_{ij} represents the distance between atoms i and j , θ_{ijk} is the angle for the triplet (ijk) with i being the central atom, and $P_l(x)$ is the Legendre polynomial of order l . The functions $g_n(r_{ij})$ depend solely on the distance r_{ij} and are therefore referred to as radial functions. These radial functions are defined as linear combinations of a number of basis functions:

$$g_n(r_{ij}) = \sum_k c_{nk}^{IJ} f_k(r_{ij}). \quad (4)$$

The basis functions f_k are constructed based on Chebyshev polynomials and a cutoff function, ensuring both formal completeness and smoothness. Explicit expressions for these functions can be found in Ref. 18. The expansion coefficients c_{nk}^{IJ} depend on n and k and also on the types (denoted as capitals I and J) of atoms i and j . Due to the summation over neighbors, the descriptor components defined above are invariant with respect to permutation of atoms of the same type. More importantly, these coefficients are treated as trainable parameters [17], which is crucial for efficiently differentiating different atom pairs contributing to the descriptor.

While the descriptor parameters $\{\mathbf{c}^{IJ}\}$ depend on the atom types (species), the NN parameters \mathbf{w} in NEP3 are

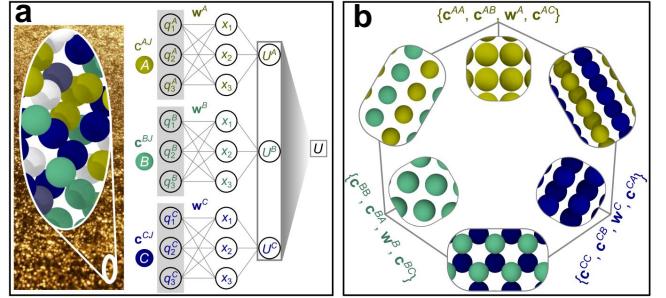


Figure 1. **a** Schematic illustration of the architecture of the NEP4 model with distinct sets of NN parameters for different atom types. For a central atom of type A , the descriptor involves the \mathbf{c}^{AJ} parameters (J can be of any type), while the weight and bias parameters \mathbf{w}^A are specific for type A . Similar rules apply to the central atoms of other types. The total energy U is the sum of the site energies for all the atoms in a given structure. By contrast, in NEP3 all atom types share a common set of NN parameters \mathbf{w} , which restricts the regression capacity. **b** Schematic illustration of the multi-loss evolutionary training algorithm. For a 3-component system, the optimization of the parameters related to atom type A (including \mathbf{w}^A , \mathbf{c}^{AA} , \mathbf{c}^{AB} , and \mathbf{c}^{AC}) is only driven by a loss function defined using the structures with the chemical compositions of A , AB , and AC . In the conventional evolutionary algorithm, which is used in NEP3, a single loss function is used to optimize all parameters, which is less effective for training general-purpose models for many-component systems.

the same for all the atom types. Therefore, as the number of atom types increases, the regression capacity of the NN model for each atom type decreases. To keep a constant regression capacity per atom type, in the present work, we employ different sets of NN parameters \mathbf{w}^I for each atom type I . While this increases the total number of trainable parameters, it *does not* significantly increase the computational cost during MD simulations with the trained model, because it only involves a selection of the correct set of NN parameters for a given atom. With the extension, the site energy can be expressed as

$$U^i = \mathcal{N}(\mathbf{w}^I; \mathbf{q}^i(\{\mathbf{c}^{IJ}\})), \quad (5)$$

which constitutes the NEP4 model introduced in this work (Figure 1a).

A multiple-loss evolutionary training algorithm for many-component systems. While the increase in the number of trainable parameters does not significantly affect the inference speed, it considerably increases the number of iterations required for training, particularly with the approach used for NEP3. It turns out that the training algorithm must be modified to achieve better performance for many-element systems. For training NEP models we use the separable natural evolution strategy (SNES) approach [31], which is a powerful black-box optimization algorithm that is particularly suitable for

problems with many possible solutions [32]. It maintains a mean value and a variance for each trainable parameter that are updated according to the rank of a population of solutions. The rank is determined according to the loss function to be minimized. The loss function L is constructed using predicted and reference data for energies, forces, and virials, and is a function of the trainable parameters, i.e.,

$$L = L(\{\mathbf{w}^I\}; \{\mathbf{c}^{IJ}\}). \quad (6)$$

The rank (or “fitness”) is of crucial importance in evolutionary algorithms, as it determines the relative weight of a solution in the population. However, using a single loss function can lead to ambiguity in rank assignment: Even if the total loss of solution X is smaller than that of solution Y, it does not guarantee that solution X is more accurate for all the subsystems in a many-element system. For example, solution X might offer higher accuracy for Au systems but lower accuracy for Ag systems. To account for this observation we define multiple loss functions for many-element systems. Since we are concerned with alloys, we cannot define a set of loss functions that have no common terms at all, but we can make a definition that minimizes the common parts. Naturally, we define the loss function for element I as the parts in Eq. (6) that are contributed by structures containing element I . For illustration, consider an explicit example with three elements, denoted A , B , and C , respectively. The loss function for element A can be calculated by considering the chemical compositions A , AB , and AC only, excluding B , C , and BC . This loss function is used when training the parameters related to element A , which are \mathbf{w}^A , \mathbf{c}^{AA} , \mathbf{c}^{AB} , and \mathbf{c}^{AC} (Figure 1b). Using this multi-loss evolutionary algorithm, the training converges much faster than using a single-loss function. The efficiency improvement in training becomes more significant with an increasing number of elements, and is crucial for being able to develop models such as UNEP-v1.

Construction of training data for many-component systems based on chemical generalizability. The chemical space for 16 elements consists of $2^{16} - 1 = 65\,535$ chemical combinations, including 16 unaries, 120 binaries, 560 ternaries, etc. It is formidable to construct a training dataset by enumerating all the possible chemical combinations. Fortunately, leveraging the construction of the radial functions in terms of linear combinations of basis functions provides a solution. The descriptor values for a given configuration of n -component ($n > 2$) systems fall within the range spanned by those of the 1-component and 2-component systems derived from the same configuration by element substitution. Given the interpolation capabilities of NNs, a NEP model trained using 1-component and 2-component structures

is expected to predict the behavior of n -component ($n > 2$) systems reasonably well. Therefore, our training dataset focused only on unary and binary systems.

For each unary or binary system, we constructed an initial training dataset with a few hundred structures. These structures included small cells with position and/or cell perturbations, cells with one to a few vacancies, cells with surfaces and various defects (such as grain boundaries) taken from the Materials Project [33] and the Open Quantum Materials Database [34], cells sampled from MD simulations based on an EAM potential [35] at various temperature (up to 5000 K) and pressure conditions including highly deformed structures (see Methods for details). There are initially about 60 000 structures in total for the 16 metals and their binary alloys. In spite of its seemingly modest size, this training dataset is remarkably diverse in configuration space. Reference data (energy, force, and virial) for the structures were generated via DFT calculations using the VASP package (see Methods for details).

The diversity of the initial training dataset ensured a robust initial NEP model that could be used to run MD and hybrid Monte Carlo and molecular dynamics (MCMD) simulations at various thermodynamic conditions. From diverse MD and MCMD trajectories generated by the initial NEP model, structures (still unary and binary only) were sampled and labeled using DFT calculations. Those with relatively large errors (NEP versus DFT) were identified and incorporated into the training set. This iterative process was repeated a few times until no large errors could be detected. This active-learning scheme, while simple, proved to be highly effective. The final training dataset contains 105 464 structures and 6 886 241 atoms in total. The DFT calculations for these structures required about six million CPU hours.

Training and testing results. Using the refined training dataset, we trained a NEP model (see Method and Supplementary Note for details on the hyperparameters) using the NEP4 approach as described above. We refer to this NEP model as UNEP-v1, which represents the first attempt at constructing a unified NEP model for many elements.

The parity plots for energy, force, and stress affirm the high accuracy of this UNEP-v1 model (Figure 2a-c). Despite the large ranges of the three quantities, their root-mean-square errors (RMSEs) are relatively small, at 17.1 meV/atom, 172 meV/Å, and 1.16 GPa, respectively.

To validate the force accuracy of our UNEP-v1 model we consider here three public datasets. Although the public datasets were not computed using exactly the same DFT settings as used for generating the UNEP-v1 training data, the resulting differences in force values are marginal (of the order of a few meV/Å) and are much

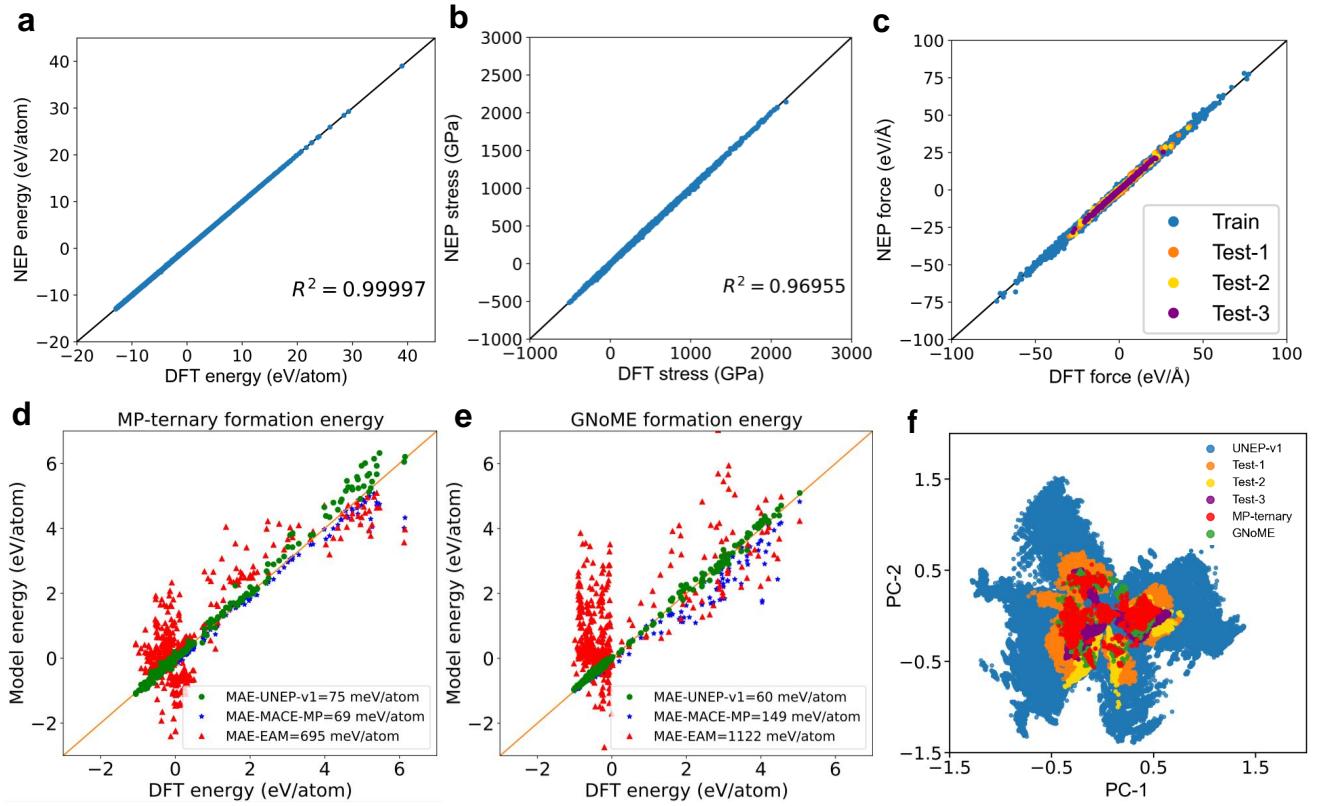


Figure 2. **a–c** Parity plots for energy, stress, and force comparing DFT reference data and UNEP-v1 predictions for the whole training dataset. In **c**, there are three test datasets containing n -component ($n \geq 3$) structures, including one with up to 13 components (Ag, Au, Cr, Cu, Mo, Ni, Pd, Pt, Ta, Ti, V, W, Zr) taken from Lopanitsyna *et al.* [36] (labeled Test-1), one with up to four components (Mo, Ta, V, W) from Byggmästar *et al.* [37] (labeled Test-2), and one with up to three components (Pd, Cu, Ni) from Zhao *et al.* [38] (labeled Test-3). **d–e** Parity plots for formation energies comparing DFT reference data and predictions from UNEP-v1, MACE-MP-0 (medium model) [39], and EAM [35], for structures from the Materials Project (MP-ternary) [33] and the GNoME paper [40]. **f** Distribution of the training dataset (this work, UNEP-v1, comprising 1-component to 2-component systems, blue) and various test datasets, including Test-1 (up to 13-component systems, orange) [36], Test-2 (up to 4-component systems, yellow) [37], Test-3 (up to 3-component systems, purple) [38], MP-ternary alloys (3-component systems, red) [33], and GNoME dataset (2-component to 5-component systems, green) [40], in the 2D principal component (PC) space of the descriptor.

smaller than the force RMSE achieved by UNEP-v1 (Figure 2c). The comparison moreover shows that the UNEP-v1 model trained against 1-component and 2-component structures also performs very well for 3-component [38], 4-component [37], and 13-component [36] structures extracted from the datasets in the previous studies [36–38]. The testing RMSEs of the UNEP-v1 model for these three datasets are respectively 76 meV/Å, 196 meV/Å, and 269 meV/Å, which are comparable to those reported as training RMSEs in the original publications [36–38].

To validate the energy accuracy of our UNEP-v1 model we utilize two public datasets, including all the relevant 3-component structures in the Materials Project database [33] and the structures predicted using the GNoME approach [40] ranging from 2-component to 5-component systems with force components less than 80 eV/Å. We calculate the formation energies using

DFT, an EAM potential [35], a foundation model named MACE-MP-0 (medium version) [39], and our UNEP-v1 model, where the reference energy for each species is based on the most stable allotrope. For the two datasets, the mean absolute errors (MAEs) of our UNEP-v1 model compared to DFT calculations are 75 meV/atom and 60 meV/atom, respectively (Figure 2d and Figure 2e). In contrast, the corresponding values from the EAM potential are 695 meV/atom and 1122 meV/atom, respectively, and thus about one order of magnitude larger. For the Materials Project dataset, which MACE-MP-0 has been trained on while UNEP-v1 has not, MACE-MP-0 is slightly more accurate. However, for the GNoME dataset, on which neither model has been trained, UNEP-v1 demonstrates notably better accuracy.

Besides the Materials Project and GNoME datasets, Figures S1–S3 present parity plots for formation energies

and forces predicted by UNEP-v1, EAM, and MACE-MP-0 compared to DFT for the three test datasets [36–38]. Figures S4–S17 show the formation energies, comparing UNEP-v1, EAM, and DFT for the equation of state curves (for alloys), heating, compressing, and stretching processes with 1 to 5-component materials in various crystalline structures, including face-centered cubic (FCC), body-centered cubic (BCC), hexagonal close packed (HCP), and metallic glasses. The results altogether clearly demonstrate the superior accuracy of UNEP-v1 over EAM and confirm the excellent generalizability of our UNEP-v1 model from the 1- and 2-component structures included in the training dataset to unseen multi-component structures.

As a further test, we trained a NEP model by including relevant n -component ($n \geq 3$) structures from the Open Quantum Materials Database database [34]. The RMSEs for the three public datasets [36–38] obtained using this NEP model are only marginally improved compared to UNEP-v1, which demonstrates that our training dataset with n -component ($n \leq 2$) structures is already sufficient for training a general-purpose NEP model for all the considered elements and their alloys.

As mentioned earlier, our approach to training data generation relies on the chemical generalizability embedded in the radial functions Eq. (4). This feature is illustrated by a principal component analysis of the descriptor space (Figure 2f), which shows that the various n -component ($n \geq 3$) structures fall comfortably within the space spanned by the 1-component and 2-component training structures.

Evaluation of basic physical properties for the 16 metal elements. After having confirmed the high training accuracy of the UNEP-v1 model for 1-component and 2-component systems, and its high testing accuracy for systems with multiple components, we conducted an extensive evaluation of the UNEP-v1 model beyond RMSEs, focusing on various physical properties (see Methods for details on the calculations). Elastic constants C_{ij} , surface formation energies γ , mono-vacancy formation energies E_v , melting points T_m , and phonon dispersion relations were calculated for all 16 elements, using both the UNEP-v1 model and an EAM potential [35]. While there are recent and possibly more accurate EAM potentials [41] for a limited subset of species considered here, we have consistently opted for the widely used EAM potential developed by Zhou *et al.* [35] because it supports all the 16 species and their alloys. Detailed results for phonon dispersion relations are presented in Figs. S18–S20, while other physical properties are listed in Tables S1–S4. Figures 3a–d show the parity plots comparing predictions of various basic properties from EAM and UNEP-v1 against DFT calculations or experimental values. The EAM predictions have some outliers, especially

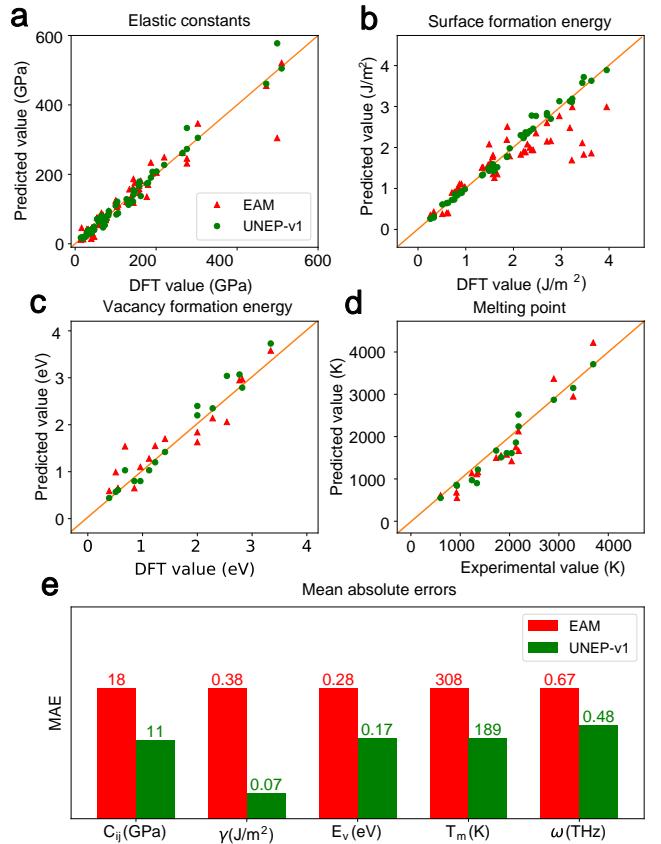


Figure 3. **a–d** Elastic constants C_{ij} , formation energies γ for $\{111\}$, $\{110\}$, and $\{100\}$ surfaces, mono-vacancy formation energies E_v , and melting points T_m as predicted by the EAM potential [35] and the UNEP-v1 model compared to DFT or experimental values for the 16 elements. **e** Mean absolute errors (MAEs) for the above four quantities as well as the phonon frequency ω for EAM and UNEP-v1 models with respect to reference data from DFT calculations and experiment.

in the case of the surface formation energies, while the UNEP-v1 predictions do not show any notable discrepancies.

The MAEs for all the evaluated quantities calculated by averaging the absolute error between predicted (EAM or UNEP-v1) and reference values (DFT or experimental) over all 16 elements are presented in Figure 3e. UNEP-v1 consistently outperforms the EAM potential for all physical properties, and demonstrates a significant advantage in predicting surface formation energies, elastic constants, and vacancy formation energies.

We have additionally trained an ensemble of eight NEP models using different sets of training hyperparameters, and compared the predictions for bulk and shear moduli as well as the equilibrium volume for the ensemble to DFT reference data to estimate the uncertainty in the model predictions (Figure 4a–c). Generally, the devia-

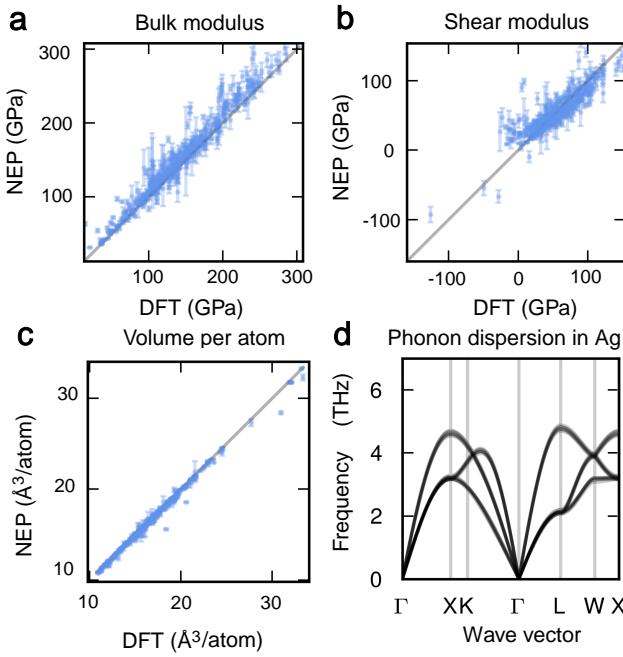


Figure 4. **a–c** Parity plots for NEP model versus DFT data for bulk modulus, shear modulus, and equilibrium volume for roughly 400 different alloys of the 16 elements, with error bars computed as the standard deviation in the predicted properties over an ensemble of eight NEP models. The structures and reference DFT data were taken from the Materials Project [33]. **d** Phonon dispersion relations for FCC Ag calculated by averaging of all models in the ensemble.

tions in the predictions across the ensemble are small, and mostly agree well with the reference data. As a further illustration, we estimated the uncertainty in the phonon dispersion for Ag (Figure 4d), illustrating the very small uncertainty throughout the entire Brillouin zone.

Computational performance. The computational efficiency of a MLP is crucial for its effective applications in large-scale MD simulations. Here, the UNEP-v1 model as implemented in GPUMD exhibits excellent computational performance (Table I). Using a single Nvidia A100 GPU, UNEP-v1 can reach a simulation size of about 14 million atoms and a computational speed of 2.4×10^7 atom step/s, which is only a few times lower than that for the EAM potential (11×10^7 atom step/s) as implemented in LAMMPS [42] using the same hardware. To reach even larger simulation sizes, we implemented a multi-GPU version of NEP that can effectively use the computational power of all the GPUs available on a computational node. With only 8 A100 GPUs, we can reach a simulation size of 100 million atoms, achieving much higher computational efficiency than either the deep potential (DP) (thousands of Nvidia V100 GPUs) [43, 44] or Allegro (128 A100 GPUs) approaches [45].

With 8 A100 GPUs, the overall computational speed of

Table I. Computational performance of UNEP-v1 in comparison with DP [43, 44], Allegro [45], and EAM models (using the GPU package of LAMMPS [42]) for large-scale MD simulations of typical metals. The speed is given in units of 10^3 atom step/(s GPU). The DP results were obtained on V100 GPUs. All other data were generated using A100 GPUs, which offer approximately twice the computational performance of a V100 GPU for this kind of computation.

Model-Element	# atoms	# GPUs	Speed
DP-Cu [43]	127×10^6	27 300	45
DP-Cu [44]	3400×10^6	27 300	330
Allegro-Ag [45]	100×10^6	128	2600
EAM-Cu	23×10^6	1	110 000
EAM-Cu	100×10^6	4	49 300
UNEP-v1-Cu	14×10^6	1	23 500
UNEP-v1-Cu	100×10^6	8	18 800
UNEP-v1-Ag	100×10^6	8	17 200

UNEP-v1 is about 1.5×10^8 atom step/s. The parallel efficiency relative to ideal scaling for UNEP-v1 with 8 A100 GPUs is 80%, while it is only about 50% for EAM with 4 A100 GPUs. The speed per GPU achieved by UNEP-v1 is significantly higher than those for the DP [43, 44] and Allegro approaches [45]. The excellent computational speed of UNEP-v1 allows us to tackle challenging problems in multi-principal-element alloys (MPEAs) as discussed below.

Application to plasticity of multi-principal element alloys. Refractory MPEAs have emerged as materials for high-temperature applications, crystallizing typically in the BCC solid solution phase. These alloys exhibit exceptional properties such as high ductility and mechanical strength at ultra-high temperature [46–49] as well as impressive irradiation resistance [50, 51]. However, their ductility at room temperature is limited [52, 53]. Recent experimental observations in alloys such as HfNbTaTiZr have revealed the presence of numerous straight screw dislocations and a substantial amount of dislocation debris [53, 54], consistent with known behavior in BCC metals [55]. Recent MD simulations have also indicated the possible crucial role of dislocation in the plastic flow of MPEAs [56–58]. Despite these insights, the complex structural and mechanical properties of MPEAs remain incompletely understood. Here, atomistic simulations employing accurate and efficient MLPs can provide further insights into the intricate behavior of these materials. Although there are a few available MLPs limited to specific alloys [56–58], a comprehensive general-purpose potential model capable of encompassing a wide range of elements and their alloys, providing both high efficiency and accuracy and enabling large-scale (up to millions of atoms) MD simulations of BCC MPEAs, is still lacking.

The UNEP-v1 model developed in this work emerges

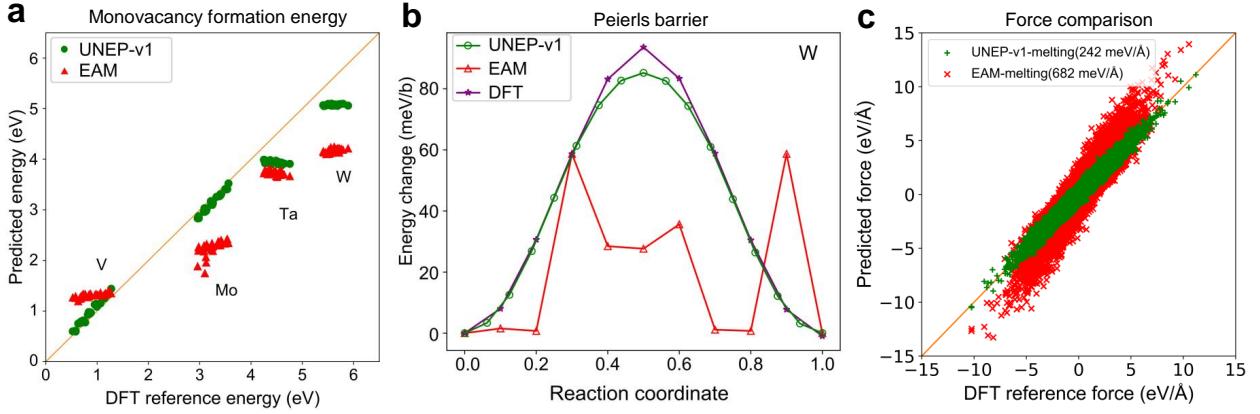


Figure 5. **a** Mono-vacancy formation energies from UNEP-v1 and EAM [35] compared to DFT data for an equimolar MoTaVW alloy with 128 atoms sampled from MCMD simulations. **b** Peierls barrier for $1/2\langle 111 \rangle$ screw dislocation migration in elemental W (see Fig. S21 for the other three species). **c** Comparisons of UNEP-v1, EAM [35], and DFT results for equimolar MoTaVW alloys sampled from MD simulations using 256-atom supercells for a melting process from 10 to 5000 K during 10 ns. UNEP-v1 shows much better predictions than EAM, with a much smaller force RMSE as indicated in the legend (see Fig. S22 for similar comparisons for deformation processes).

as a promising solution, enabling large-scale MD simulations of MPEAs with an accuracy superior to existing models while still achieving very high computational efficiency. To demonstrate its effectiveness in this context, we investigated the mechanism of plastic deformation of a MoTaVW alloy under compression. Our evaluation of the UNEP-v1 model involved comprehensive tests, including checking the vacancy formation energies (Figure 5a) in equimolar MoTaVW alloys, Peierls barriers for the $1/2\langle 111 \rangle$ screw dislocation (Figure 5b) in elemental systems as well as atomic forces in melting (Figure 5c), compression and tensile stretching processes (Fig. S22) of equimolar MoTaVW alloys. The results illustrate the superior performance of UNEP-v1 compared to EAM potentials and its suitability for studying structural and mechanical properties in large-scale MD simulations.

After having confirmed the accuracy, efficiency, and reliability of our NEP model, we modeled an equimolar BCC polycrystalline MoTaVW system containing 100 205 176 atoms and conducted MD simulations to investigate changes in dislocation density under compression. These simulations involved compressive deformation at a strain rate of $4.2 \times 10^8 / \text{s}$ (see Methods for simulation details). The dislocation density decreases during the elastic stage, reaches a minimum at the yield strain $\epsilon = 6\%$, and gradually returns to the original level due to enhanced densification (Figure 6a). The dislocation density plateaus for large strains ($\epsilon \geq 16\%$), consistent with the behavior observed in BCC Ta [59]. It is noteworthy that stress-strain response and dislocation density exhibit contrasting trends under compression.

To gain deeper insight into the plastic deformation mechanisms, we extracted the distribution of the dislocation density in snapshots of the polycrystalline Mo-

TaVW system at selected strains (Figure 6b–e). Notably, all dislocations are confined to grain boundaries of the polycrystalline system under compression, and this pattern remains unchanged throughout the linear response (“elastic”) region of the stress-strain curve (Figure 6b–c). It is worth noting that dislocations transform from other types (labeled 1 and 2 in Figure 6b) to $1/2\langle 111 \rangle$ ones (Figure 6c) in the elastic region ($0 – 2.5\%$), and recover back at the yield strain of 6% (Figure 6d). Subsequently, during the plastic stage (Figure 6d–e), some of the grain boundaries begin to emit, slip, and pin dislocations into the grains along with boundary movement. This finding demonstrates the significant impacts of boundary stability on the hardness of MPEAs, as previously observed in the study of a NiMo alloy [60].

Through 100-million-atom large-scale MD simulations, we have thus illuminated the intricate details of plastic deformation, shedding light on dislocation behavior in grain boundaries. This application of our UNEP-v1 model to the plasticity of MPEAs, exemplified by the MoTaVW alloy, is an important demonstration for the generality and high computational efficiency of our approach.

Application to primary radiation damage in MPEAs. Next, we demonstrate the versatility of the UNEP-v1 model through large-scale MD simulations of primary radiation damage in MPEAs, using again the MoTaVW alloy system for illustration (see Methods for details). Notably, these simulations set a new benchmark with a record-breaking size of 16 million atoms for MLPs in this specific type of simulations. Here, in order to accurately describe interactions at extremely short distances where large forces are at play, we incorporated a

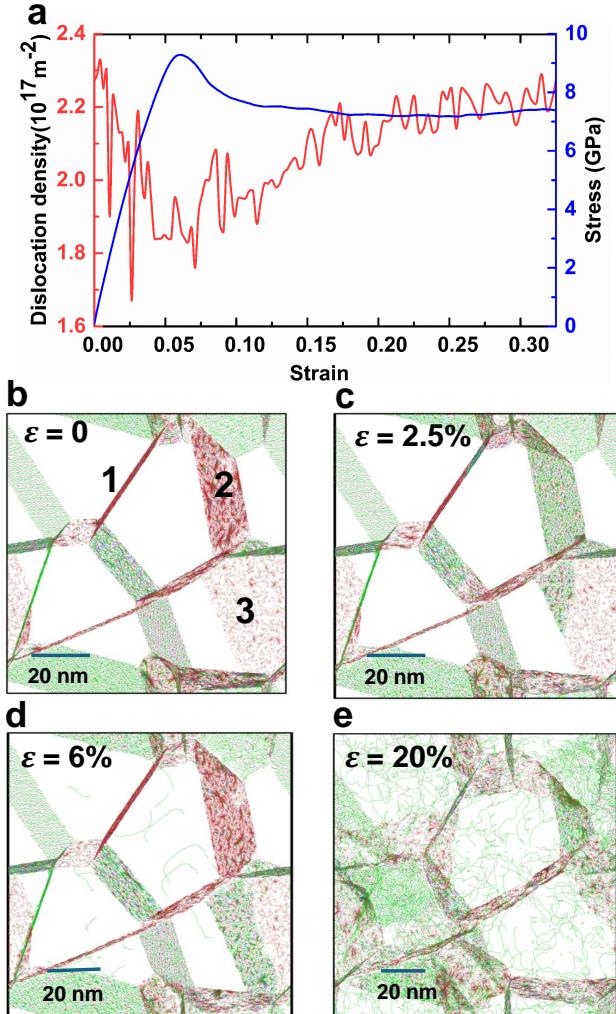


Figure 6. Dislocation density as a function of compressive strain for equimolar polycrystalline MoTaVW alloys containing 12 grains with 100 million atoms at 300 K. **a** Strain-induced dislocation density and stress. **b–e** Distributions of dislocation in 20 nm thick slices at strains of **(b)** $\epsilon = 0\%$, **(c)** $\epsilon = 2.5\%$, **(d)** $\epsilon = 6\%$, and **(e)** $\epsilon = 20\%$, respectively. Grain boundaries are labeled by numbers for reference. The compressed direction is perpendicular to the plane of view. The $1/2\langle 111 \rangle$ dislocations are depicted in green, while other dislocations are shown in red.

two-body Ziegler-Biersack-Littmark (ZBL) potential [61] to train a combined NEP-ZBL model [62]. The UNEP-v1 part and the ZBL part are smoothly connected in the range between 1 and 2 Å. Above 2 Å, only the UNEP-v1 part is active, while below 1 Å, the ZBL part dominates. Illustrative examples demonstrating the seamless connection between UNEP-v1 and ZBL for Al and W dimers are presented in Fig. S23.

Figure 7a shows the defect snapshot of the peak-damage state formed at about 0.6 ps with a primary knock-on atom energy of 100 keV. The defect distribution stabilizes after a few tens of ps. Figure 7b shows

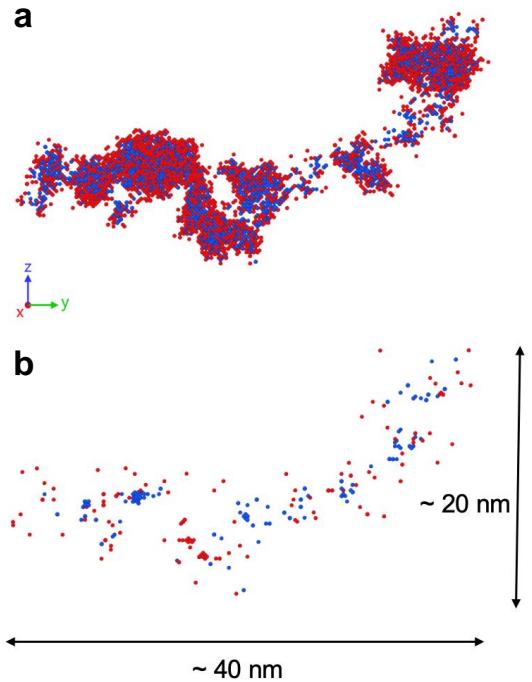


Figure 7. Defect snapshots of a cascade in a MoTaVW alloy at **(a)** the peak damage state (at about 0.6 ps) and **(b)** the final damage state (at 140 ps). The red and blue dots represent interstitial atoms and vacancies, respectively.

the stable defect distribution at 140 ps, revealing 121 residual point defects, including vacancies and interstitial atoms. The maximum cluster sizes for vacancies and interstitials are 15 and 11, respectively. In comparison, a previous study [62] on elemental W at similar simulation conditions reported 183 residual point defects with a maximum defect-cluster size exceeding 200 atoms. The MPEA thus features fewer defects and smaller defect clusters. Our simulation results are consistent with the experimental study of a similar tungsten-based refractory MPEA, which exhibits exceptional radiation resistance, negligible radiation hardening, and no evidence of radiation-induced dislocation loops even at a high dose level [50].

The enhanced radiation resistance of the tungsten-based refractory MPEAs could be attributed to the increased chemical complexity, leading to cascade splitting, as depicted in Figure 7a. Cascade splitting results in the formation of smaller defect clusters and a more dispersed distribution of isolated (non-clustered) point defects. This specific application of our UNEP-v1 model to study primary radiation damage through extensive MD simulations involving 16 million atoms provides further evidence of the generality and high efficiency of our approach. However, more detailed investigations are necessary to comprehensively characterize and understand the role of alloying in influencing radiation resistance.

Comparisons between UNEP-v1 and EAM models in MD and MCMD simulations. Finally, we showcase the reliability of UNEP-v1 in large-scale MD and MCMD (see Methods) simulations across three applications, providing close comparisons with EAM results and experimental data.

In the first application, we use UNEP-v1 to perform MD simulations for the recently synthesized goldene [65], a monolayer form of gold that is not explicitly included in the training dataset. The stable configuration of goldene features a triangular lattice. We first construct a rectangular cell with 1800 atoms and then performed MD simulations in the isothermal-isobaric ensemble with a target temperature of 300 K and a target in-plane pressure of 0 GPa. Figure 8a shows that UNEP-v1 maintains the stability of the goldene sheet at 300 K, exhibiting out-of-plane ripples typical for two-dimensional materials. In contrast, Figure 8b–d shows that the monolayer structure of goldene cannot be maintained by EAM potentials from the literature, which include the one used for most benchmarks [35] and two more recent ones [41, 63]. The trajectories for the MD simulations using the UNEP-v1 model and the three EAM potentials are presented as movies in the SI. The results here demonstrate that the UNEP-v1 model has good generalizability in the configuration space.

In the second application, we use MCMD simulations to study the Mo distribution in a superlattice structure formed by γ -Ni and γ' -Ni₃Al. Starting from a uniform Mo distribution with an overall Mo concentration of 8.1% (Figure 8e), the final ratio of the Mo concentration in γ' -Ni₃Al to that in γ -Ni is $K^{\gamma'}/\gamma = 0.667$ according to our UNEP-v1 model (Figure 8f). This agrees well with experimental observations indicating $K^{\gamma'}/\gamma < 1$ when the initial Mo concentration is above approximately 6% [66, 67]. In contrast, the EAM potential by Zhou *et al.* [35] gives a value of $K^{\gamma'}/\gamma = 4.981$ (Figure 8g), which contradicts the experimental trend.

In the third application, we use MCMD simulations to reproduce the experimentally expected BCC structure in the Al-rich intermetallic Al_{0.31}Cr_{0.06}Cu_{0.22}Ni_{0.32}V_{0.09}, despite the presence of a large fraction of FCC metals, starting from an initial FCC structure (Figure 8h). Our UNEP-v1 model successfully produces both disordered (A2) and ordered (B2) BCC structures (Figure 8i) in full agreement with experiments [64]. In contrast, the EAM potential by Zhou *et al.* [35] keeps the system in the FCC structure (Figure 8j). Similar results for Al_{0.20}Cr_{0.12}Cu_{0.19}Ni_{0.35}V_{0.14} are shown in Fig. S24, further demonstrating the superior reliability of UNEP-v1 over the EAM model. Finally, in Fig. S25, we illustrate that the equimolar TiZrVMo and TiZrVMoTa alloys correctly transform to BCC structures from HCP structures during MCMD simulations with UNEP-v1, in excellent agreement with experimental observations [68]. This indicates that the UNEP-v1 model, trained on 1-

component and 2-component structures, can correctly capture phase transitions occurring in multi-element alloys.

DISCUSSION

In summary, we have developed an advanced NEP approach capable of constructing accurate and efficient general-purpose MLPs for numerous elements and their alloys. Two crucial extensions have been made compared to previous NEP versions. Firstly, we employed distinct NNs for each species, ensuring consistent regression capacity even as the number of species grows. Secondly, we introduced multiple loss functions to optimize different subsets of the parameters, crucially accelerating the training process when using evolutionary algorithms with a large number of trainable parameters. We expect that this concept can more generally boost the application of evolutionary algorithms in solving complex optimization problems.

A pivotal insight driving the success of this approach is the recognition that chemical (species) information can be embedded in the trainable expansion coefficients of radial functions, dependent only on atom pairs and basis functions. As a result, the 1-component and 2-component structures delineate an outer boundary in descriptor space, while n -component structures with $n \geq 3$ represent interpolation points in this space. Leveraging the exceptional interpolation capabilities of NNs, a NEP model trained solely with 1-component and 2-component structures performs very well for n -component structures with $n \geq 3$, provided the configuration space has been sufficiently explored. The effectiveness of this approach has been demonstrated through accurate predictions of formation energies across various multi-component alloys, as well as reproduction of experimentally observed chemical order and stable phases, using our UNEP-v1 model.

While the current study focuses on 16 elements, our approach is scalable and adaptable for constructing NEP models across the entire periodic table. The primary challenge resides in the generation of the reference data, typically via DFT calculations, rather than the regression capabilities of the NEP model. Notably, our approach is also sustainable. Starting from our existing training set for 16 elements, one merely needs to include structures involving 17 chemical compositions (one 1-component and 16 2-component systems) to form a comprehensive training set for 17 elements. This method is far more economical than building an entirely new training set from scratch. Beyond extending the chemical space, one can also broaden the configuration space for existing chemical compositions, through established active-learning approaches, especially with the aid of structure searching methods [69].

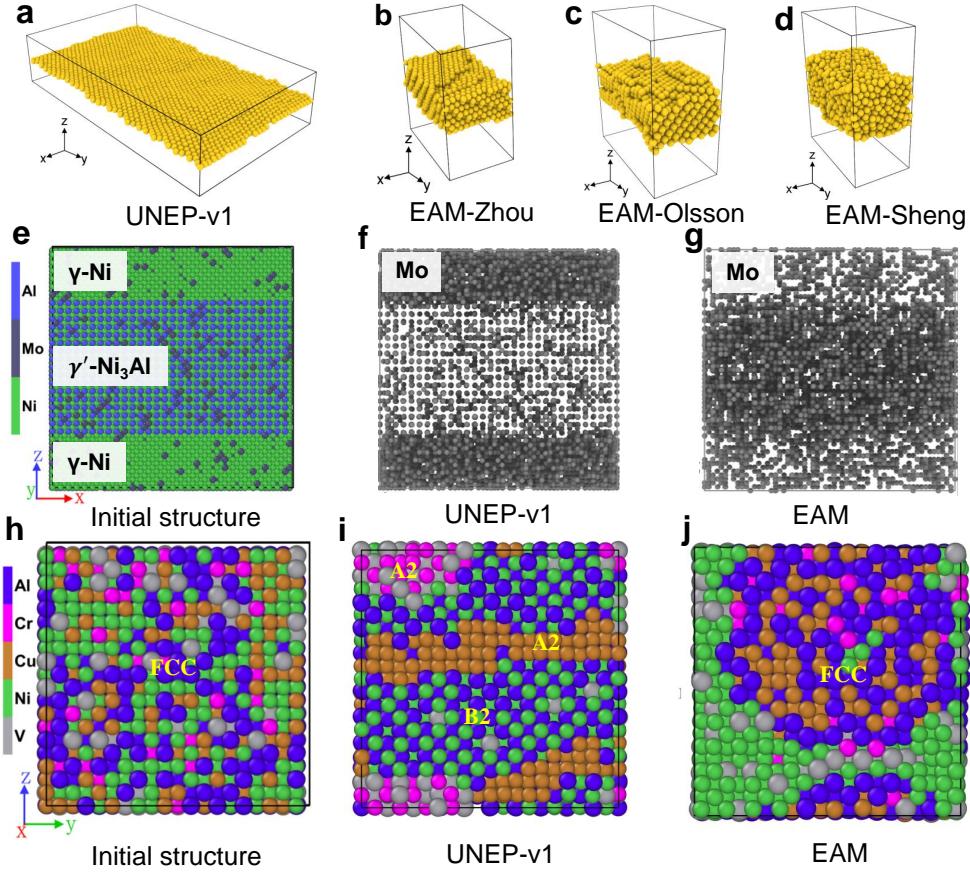


Figure 8. **a-d** Snapshots of structures from MD simulations of 1 ns in the isothermal-isobaric ensemble with a target temperature of 300 K and a target in-plane pressure of 0 GPa, starting with a flat monolayer of goldene, using UNEP-v1 and the EAM potentials by Zhou *et al.* [35], Olsson [63], and Sheng *et al.* [41]. The results demonstrate that UNEP-v1 can maintain goldene's 2D structure at ambient temperature and pressure, while the three EAM potentials cannot. **e** Initial structure of a γ -Ni and γ' -Ni₃Al superlattice with a random Mo distribution. **f-g** Snapshots of the final equilibrium Mo distributions from MCMD simulations using UNEP-v1 and EAM models [35]. UNEP-v1 correctly reproduces the final ratio of the Mo concentration in γ' -Ni₃Al to that in γ -Ni ($K_{\gamma'/\gamma} = 0.667$), in good agreement with experimental observations, while the EAM potential by Zhou *et al.* [35] gives a value ($K_{\gamma'/\gamma} = 4.981$) contradicting experimental trend. **h** Initial FCC structure of Al_{0.31}Cr_{0.06}Cu_{0.22}Ni_{0.32}V_{0.09}. **i-j** Snapshots of the final equilibrium structures from MCMD simulations using UNEP-v1 and EAM models [35]. UNEP-v1 successfully produces both disordered (A2) and ordered (B2) BCC structures in full agreement with experiments [64]. In contrast, EAM potential by Zhou *et al.* [35] keeps the system in the FCC structure, unable to reproduce the experimentally expected BCC structure. (See Fig. S24 for similar results for Al_{0.20}Cr_{0.12}Cu_{0.19}Ni_{0.35}V_{0.14}.)

The successful applications of the UNEP-v1 model in studying plasticity and primary radiation damage in the MoTaVW refractory MPEAs demonstrate the versatility and robustness of the NEP4 approach in general and the UNEP-v1 model in particular, establishing its significant potential for in-depth explorations and insights into the intricate behavior of complex materials such as MPEAs.

In conclusion, our study demonstrates the promise of our approach in constructing a unified general-purpose MLP for the periodic table with remarkable computational efficiency, taking full advantage of the embedded chemical generalizability, the outstanding interpolation capabilities of NNs and an advanced multiple-loss evolutionary training algorithm for many-component systems. By successfully developing a highly accurate and efficient

MLP for a diverse range of elemental metals and alloys, our study showcases the versatility and applicability of our approach across various materials. These advancements mark a significant leap forward in enhancing the practical applications of MLPs in materials modeling, offering new opportunities for more accurate, efficient, and predictive computer simulations in materials research.

METHODS

MD simulations for training structure generation. To create the initial training structures, we used the LAMMPS package (23 Jun 2022) [42] to run MD simulations with cells ranging from 32 to 108 atoms. For each

1-component or 2-component system, we ran MD simulations in the isothermal-isobaric ensemble (zero target pressure) using the EAM potential [35] at 9 temperatures (50, 300, 800, 1300, 1700, 2300, 3000, 4000, and 5000 K), each for 2 ns. For each MD run, we sampled 5 structures. For each structure, we made three copies, one with a subsequent box scaling of 95%, one with 105%, and one with 5% (random) box perturbation. We also ran MD simulations at 300 K with tensile or compressing loading with a strain rate of 2×10^8 /s for 1 to 2 ns and uniformly sampled 35 structures.

DFT calculations for reference data generation. After preparing the initial training structures, we performed quantum-mechanical calculations to obtain reference data, including the energy and virial for each structure and the force on each atom in each structure. DFT calculations as implemented in VASP [70] were performed to generate reference data. The INCAR file for VASP is presented in Supplementary Note S1.

We used the projector augmented wave method [71, 72], the PBE functional [73], an energy cutoff of 600 eV, a Γ -centered k -point mesh with a spacing of 0.2/Å, and a threshold of 10^{-6} eV for the electronic self-consistent loop. We used the blocked Davidson iteration scheme for electronic minimization. The PREC tag in the VASP input file was set to `Accurate` to ensure accurate forces. A Gaussian smearing with a smearing width of 0.02 eV was used. The Gaussian smearing is not the best choice for elemental metals and their alloys but we chose this in view of possible future extension of our approach to the whole periodic table. Our settings can ensure a convergence of the energy to 1 meV/atom for all the materials. In our DFT calculations, we did not consider magnetism, consistent with previous works [36, 38]. We have tested that modeling Ni and Cr as ferromagnetic does not change the energy ordering for the major phases, see Fig. S26. While a proper account of magnetism might help to improve the quality of the DFT results and the resulting potential, it would significantly complicate the training database and require extensive additional computational resources.

The NEP training hyperparameters. We used GPUMD v3.9.3 to train the UNEP-v1 model, which is a NEP4 model as introduced in this paper. The details of the `nep.in` input file we used and the SNES-based multi-loss training algorithm can be found in Supplementary Note S2.

The cutoff radii for radial and angular descriptor parts are 6 Å and 5 Å, respectively. For both the radial and the angular descriptor components, we used 5 radial functions constructed from a linear combination of 9 basis functions. The descriptor vector for one element thus has $5 + 5 \times 6 = 35$ components. There are 80 neu-

rons used in the hidden layer and the NN architecture for each element can be written as 35-80-1, corresponding to 2960 trainable parameters. For each pair of elements, there are $5 \times 9 + 5 \times 9 = 90$ trainable descriptor parameters. The total number of trainable parameters in the UNEP-v1 model for 16 elements is thus $2960 \times 16 + 90 \times 16^2 + 1 = 70\,401$, where a global bias (shifting) parameter is included. The training was performed with a batch size of 10 000 structures for 1 000 000 generations (steps), which took about ten days using four A100 GPUs.

Calculations of basic physical properties. To evaluate the reliability of the UNEP-v1 model in molecular statics and MD simulations, we calculated a set of relevant static and dynamic material properties, with a close comparison with EAM [35], DFT (if available or affordable) [74, 75], and experiments (if available). Energetics, elastic properties, and phonon dispersion relations were calculated with the help of GPUMD-WIZARD, ASE [76], PYNEP [18], CALORINE [77], and PHONOPY [78] packages. Melting points were calculated using the two-phase method as implemented in GPUMD [18] for UNEP-v1 and LAMMPS [42] for EAM, and are compared to experimental values [79]. Vacancy formation energies were evaluated using $4 \times 5 \times 6$ supercells. The formation energies of free surfaces were evaluated with $2 \times 2 \times 10$ supercells (taking a surface perpendicular to z as an example here). The uncertainties in the predictions for the bulk and shear moduli and volume per atom, over different ensemble models, were estimated as the standard deviation using CALORINE [77].

MD simulations for plasticity of MPEAs. We used the UNEP-v1 model to drive MD calculations of the plasticity of MPEAs under compression using the GPUMD package [18, 19]. First, we used the Voronoi algorithm implemented in ATOMSK [80] to build our initial MoTaVW polycrystalline MPEA model by removing overlapping atoms at boundaries. The model is composed of 12 grains with sizes ranging from 96 nm^3 to 195 nm^3 , and contains 100 205 176 atoms which randomly occupy a BCC lattice at equimolar ratios. The initial Mo-TaVW model was further relaxed by MD simulations for 500 ps in the isothermal-isobaric ensemble at 300 K and 0 GPa using the Bernetti-Bussi barostat [81] and Bussi-Donadio-Parrinello thermostat [82]. Finally we simulated uniaxial compressive deformation with a constant engineering strain rate of 4.2×10^8 /s. The time step was kept fixed at 1 fs. The 2D visualization of dislocations perpendicular to the compressive axis was rendered using the OVITO package [83].

MD simulations for primary radiation damage. The MD simulations of the displacement cascade in Mo-

TaVW were performed using the GPUMD package [18, 19] with the UNEP-v1 model and a repulsive two-body ZBL-like potential [61]. A periodic cubic simulation cell with 16 000 000 atoms was constructed by creating a random mixture of the Mo, Ta, V, and W atoms with equimolar ratio in a BCC crystal. We equilibrated this system in the isothermal-isobaric ensemble for 30 ps, with a target temperature of 300 K and a target pressure of 0 GPa. A primary knock-on atom with an energy of 100 keV moving in the high-index direction $\langle 135 \rangle$ (to avoid channeling effects) was then created at the center of the simulation cell. Atoms within a thickness of three lattice constants of the boundaries were maintained at 300 K. The integration time step had an upper limit of 1 fs and was dynamically determined so that the fastest atom could move at most 0.015 Å (less than 0.5% of the lattice constant) within one step. The total number of steps is 200 000, corresponding to 140 ps. Electronic stopping [84] was applied as a frictional force on atoms with a kinetic energy over 10 eV. We used the OVITO package [83] for defect analysis and visualization. The interstitials and vacancies were identified by using the Wigner-Seitz cell method. The defects were grouped into clusters: two vacancies were considered to be in the same cluster if the distance between them was within the second-nearest-neighbor distance, while the third-nearest-neighbor distance was used to identify self-interstitial clusters.

MCMD simulations for multi-component alloys. We utilized MCMD simulations in the canonical MC ensemble (involving swapping atoms of different species) [85] in the study of Mo distribution in a superlattice structure formed by γ -Ni and γ' -Ni₃Al (108 000-atom supercell), the FCC-to-BCC transformation in Al-rich alloys ($\text{Al}_{0.31}\text{Cr}_{0.06}\text{Cu}_{0.22}\text{Ni}_{0.32}\text{V}_{0.09}$ and $\text{Al}_{0.20}\text{Cr}_{0.12}\text{Cu}_{0.19}\text{Ni}_{0.35}\text{V}_{0.14}$, with 4000-atom supercell), and the HCP-to-BCC transformation in the equimolar TiZrVMo and TiZrVMoTa alloys (1600-atom supercell). During these MCMD simulations, MC trials are attempted 2000 times after every 1000 MD steps, totaling about 10^6 MD steps to reach equilibrium, at which the MC acceptance ratio is close to zero.

DATA AVAILABILITY

The training data and trained NEP models are freely available at the Zenodo repository <https://doi.org/10.5281/zenodo.10081676>. The high-throughput-calculation inputs/outputs for the basic physical properties are freely available at <https://github.com/Jonsnow-willow/GPUMD-Wizard>.

CODE AVAILABILITY

The source code and documentation for CALORINE are available at <https://gitlab.com/materials-modeling/calorine> and <https://calorine.materialsmodeling.org/>, respectively. The source code and documentation for GPUMD are available at <https://github.com/brucefan1983/GPUMD> and <https://gpumd.org>, respectively. The source code and documentation for GPUMD-WIZARD are available at <https://github.com/Jonsnow-willow/GPUMD-Wizard>. The source code and documentation for PYNEP are available at <https://github.com/bigd4/PyNEP> and <https://pynep.readthedocs.io/en/latest/>, respectively.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no competing interests.

CONTRIBUTIONS

Keke Song, Rui Zhao, Yanzhou Wang, and Nan Xu prepared the training and test structures. Keke Song, Yanzhou Wang, Zhiqiang Zhao, Ting Liang, Jiuyang Shi, Junjie Wang, Ke Xu, Shuang Lyu, Zezhu Zeng, and Shiron Liang performed the DFT calculations. Keke Song, Shunda Chen, and Yanzhou Wang tested the various hyperparameters and trained the NEP models. Penghua Ying analyzed the descriptor space. Keke Song, Rui Zhao, Jiahui Liu, Ke Xu, Ting Liang, Zhiqiang Zhao, and Haikuan Dong evaluated the basic physical properties. Eric Lindgren trained the first generation of ensemble models, and performed the ensemble model analysis. Yong Wang and Shunda Chen evaluated the computational performance. Jiahui Liu performed the radiation damage simulations and developed the high-throughput calculation tools. Zheyong Fan and Shunda Chen developed the NEP4 model and the multiple-loss evolutionary training algorithm. Ligang Sun, Yue Chen, Zhuhua Zhang, Wanlin Guo, Ping Qian, Jian Sun, Paul Erhart, Tapiro Ala-Nissila, Yanjing Su, and Zheyong Fan supervised the project.

ACKNOWLEDGEMENTS

Keke Song, Jiahui Liu, Yanzhou Wang, Ping Qian, and Yanjing Su acknowledge support from the National Key R & D Program of China (No. 2022YFB3707500) and the National Natural Science Foundation of China (NSFC) (No. 92270001). Yanzhou Wang and Tapiro Ala-Nissila have been supported in part by the Academy of Finland

through its Quantum Technology Finland CoE grant No. 312298 and under the European Union – NextGenerationEU instrument by the Academy of Finland grant 353298. Eric Lindgren and Paul Erhart acknowledge funding from the Swedish Research Council (Nos. 2020-04935 and 2021-05072) and the Swedish Foundation for Strategic Research via the SwedNESS graduate school (GSn15-0008) as well as computational resources provided by the National Academic Infrastructure for Supercomputing in Sweden at NSC and C3SE partially funded by the Swedish Research Council through grant agreement No. 2022-06725. Jian Sun acknowledges support from NSFC (Nos. 12125404, 11974162), the Basic Research Program of Jiangsu, the Fundamental Research Funds for the Central Universities, and computational sources from the High Performance Computing Center of Collaborative Innovation Center of Advanced Microstructures and the high-performance supercomputing center of Nanjing University. Ke Xu and Ting Liang acknowledge support from the National Key R&D Project from Ministry of Science and Technology of China (No. 2022YFA1203100), the Research Grants Council of Hong Kong (No. AoE/P-701/20), and RGC GRF (No. 14220022). Zhiqiang Zhao, Zhuhua Zhang, and Wanlin Guo acknowledge support from the NSFC Projects of International Cooperation and Exchanges (No. 12261160367). Zezhu Zeng, Shuang Lyu, and Yue Chen are grateful for the research computing facilities offered by ITS, HKU.

- [7] Y. Mishin, Machine-learning interatomic potentials for materials science, *Acta Materialia* **214**, 116980 (2021).
- [8] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, Machine Learning Force Fields, *Chemical Reviews* **121**, 10142 (2021).
- [9] A. V. Shapeev, Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials, *Multiscale Modeling & Simulation* **14**, 1153 (2016).
- [10] R. Drautz, Atomic cluster expansion for accurate and transferable interatomic potentials, *Phys. Rev. B* **99**, 014104 (2019).
- [11] A. Thompson, L. Swiler, C. Trott, S. Foiles, and G. Tucker, Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials, *Journal of Computational Physics* **285**, 316 (2015).
- [12] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [13] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, *Phys. Rev. Lett.* **104**, 136403 (2010).
- [14] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials, *Nature communications* **13**, 2453 (2022).
- [15] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csanyi, MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields, in *Advances in Neural Information Processing Systems*, Vol. 35, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Curran Associates, Inc., 2022) pp. 11423–11436.
- [16] Z. Fan, Z. Zeng, C. Zhang, Y. Wang, K. Song, H. Dong, Y. Chen, and T. Ala-Nissila, Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport, *Phys. Rev. B* **104**, 104309 (2021).
- [17] Z. Fan, Improving the accuracy of the neuroevolution machine learning potential for multi-component systems, *Journal of Physics: Condensed Matter* **34**, 125902 (2022).
- [18] Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, A. J. Gabourie, J. Liu, H. Dong, J. Wu, Y. Chen, Z. Zhong, J. Sun, P. Erhart, Y. Su, and T. Ala-Nissila, Gpumd: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations, *The Journal of Chemical Physics* **157**, 114801 (2022).
- [19] Z. Fan, W. Chen, V. Vierimaa, and A. Harju, Efficient molecular dynamics simulations with many-body potentials on graphics processing units, *Computer Physics Communications* **218**, 10 (2017).
- [20] S. Takamoto, S. Izumi, and J. Li, Teanet: Universal neural network interatomic potential inspired by iterative electronic relaxations, *Computational Materials Science* **207**, 111280 (2022).
- [21] S. Takamoto, C. Shinagawa, D. Motoki, K. Nakago, W. Li, I. Kurata, T. Watanabe, Y. Yayama, H. Iriguchi, Y. Asano, *et al.*, Towards universal neural network potential for material discovery applicable to arbitrary com-

- bination of 45 elements, *Nature Communications* **13**, 1 (2022).
- [22] C. Chen and S. P. Ong, A universal graph deep learning interatomic potential for the periodic table, *Nature Computational Science* **2**, 718 (2022).
- [23] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Machine Learning a General-Purpose Interatomic Potential for Silicon, *Phys. Rev. X* **8**, 041048 (2018).
- [24] P. Rowe, V. L. Deringer, P. Gasparotto, G. Csányi, and A. Michaelides, An accurate and transferable machine learning potential for carbon, *The Journal of Chemical Physics* **153**, 034702 (2020).
- [25] R. Jana and M. A. Caro, Searching for iron nanoparticles with a general-purpose Gaussian approximation potential, *Phys. Rev. B* **107**, 245421 (2023).
- [26] J. Kloppenburg, L. B. Pártay, H. Jónsson, and M. A. Caro, A general-purpose machine learning Pt interatomic potential for an accurate description of bulk, surfaces, and nanoparticles, *The Journal of chemical physics* **158**, 134704 (2023).
- [27] N. Artrith, A. Urban, and G. Ceder, Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species, *Phys. Rev. B* **96**, 014112 (2017).
- [28] A. Thorn, D. Gochitashvili, S. Kharabadze, and A. N. Kolmogorov, Machine learning search for stable binary Sn alloys with Na, Ca, Cu, Pd, and Ag, *Phys. Chem. Chem. Phys.* **25**, 22415 (2023).
- [29] L. C. Erhard, J. Rohrer, K. Albe, and V. L. Deringer, Modelling atomic and nanoscale structure in the silicon–oxygen system through active machine learning, *Nature Communications* **15**, 1927 (2024).
- [30] S. N. Pozdnyakov, M. J. Willatt, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, Incompleteness of atomic structure representations, *Phys. Rev. Lett.* **125**, 166001 (2020).
- [31] T. Schaul, T. Glasmachers, and J. Schmidhuber, High Dimensions and Heavy Tails for Natural Evolution Strategies, in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, GECCO '11 (Association for Computing Machinery, New York, NY, USA, 2011) pp. 845–852.
- [32] D. Wierstra, T. Schaul, T. Glasmachers, Y. Sun, J. Peters, and J. Schmidhuber, Natural evolution strategies, *Journal of Machine Learning Research* **15**, 949 (2014).
- [33] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, The Materials Project: A materials genome approach to accelerating materials innovation, *APL Materials* **1**, 011002 (2013).
- [34] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies, *npj Computational Materials* **1**, 15010 (2015).
- [35] X. W. Zhou, R. A. Johnson, and H. N. G. Wadley, Misfit-energy-increasing dislocations in vapor-deposited CoFe/NiFe multilayers, *Phys. Rev. B* **69**, 144113 (2004).
- [36] N. Lopanitsyna, G. Fraux, M. A. Springer, S. De, and M. Ceriotti, Modeling high-entropy transition metal alloys with alchemical compression, *Phys. Rev. Mater.* **7**, 045802 (2023).
- [37] J. Byggmästar, K. Nordlund, and F. Djurabekova, Simple machine-learned interatomic potentials for complex alloys, *Phys. Rev. Mater.* **6**, 083801 (2022).
- [38] R. Zhao, S. Wang, Z. Kong, Y. Xu, K. Fu, P. Peng, and C. Wu, Development of a neuroevolution machine learning potential of Pd-Cu-Ni-P alloys, *Materials & Design* **231**, 112012 (2023).
- [39] I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon, W. J. Baldwin, F. Berger, N. Bernstein, A. Bhowmik, S. M. Blau, V. Čáraře, J. P. Darby, S. De, F. D. Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, F. Falcioni, E. Fako, A. C. Ferrari, A. Genreith-Schriever, J. George, R. E. A. Goodall, C. P. Grey, P. Grigorev, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, J. R. Kermode, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I.-B. Magdäu, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, B. X. Shi, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, T. D. Swinburne, J. Tilly, C. van der Oord, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, F. Zills, and G. Csányi, A foundation model for atomistic materials chemistry (2024), [arXiv:2401.00096 \[physics.chem-ph\]](https://arxiv.org/abs/2401.00096).
- [40] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, Scaling deep learning for materials discovery, *Nature* **624**, 80 (2023).
- [41] H. W. Sheng, M. J. Kramer, A. Cadien, T. Fujita, and M. W. Chen, Highly optimized embedded-atom-method potentials for fourteen fcc metals, *Phys. Rev. B* **83**, 134118 (2011).
- [42] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Computer Physics Communications* **271**, 108171 (2022).
- [43] W. Jia, H. Wang, M. Chen, D. Lu, L. Lin, R. Car, W. E, and L. Zhang, Pushing the Limit of Molecular Dynamics with Ab Initio Accuracy to 100 Million Atoms with Machine Learning, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '20 (IEEE Press, 2020).
- [44] Z. Guo, D. Lu, Y. Yan, S. Hu, R. Liu, G. Tan, N. Sun, W. Jiang, L. Liu, Y. Chen, L. Zhang, M. Chen, H. Wang, and W. Jia, Extending the limit of molecular dynamics with ab initio accuracy to 10 billion atoms, in *Proceedings of the 27th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, PPoPP '22 (Association for Computing Machinery, New York, NY, USA, 2022) p. 205–218.
- [45] A. Musaelian, S. Batzner, A. Johansson, L. Sun, C. J. Owen, M. Kornbluth, and B. Kozinsky, Learning local equivariant representations for large-scale atomistic dynamics, *Nature Communications* **14**, 579 (2023).
- [46] O. Senkov, S. Gorsse, and D. Miracle, High temperature strength of refractory complex concentrated alloys, *Acta Materialia* **175**, 394 (2019).
- [47] E. P. George, D. Raabe, and R. O. Ritchie, High-entropy alloys, *Nature Reviews Materials* **4**, 515 (2019).

- [48] F. G. Coury, M. Kaufman, and A. J. Clarke, Solid-solution strengthening in refractory high entropy alloys, *Acta Materialia* **175**, 66 (2019).
- [49] P. Shi, W. Ren, T. Zheng, Z. Ren, X. Hou, J. Peng, P. Hu, Y. Gao, Y. Zhong, and P. Liaw, Enhanced strength–ductility synergy in ultrafine-grained eutectic high-entropy alloys by inheriting microstructural lamellae, *Nature Communications* **10**, 489 (2019).
- [50] O. El-Atwani, N. Li, M. Li, A. Devaraj, J. K. S. Baldwin, M. M. Schneider, D. Sobieraj, J. S. Wróbel, D. Nguyen-Manh, S. A. Maloy, and E. Martinez, Outstanding radiation resistance of tungsten-based high-entropy alloys, *Science Advances* **5**, eaav2002 (2019).
- [51] O. El Atwani, H. T. Vo, M. A. Tunes, C. Lee, A. Alvarado, N. Krienke, J. D. Poplawsky, A. A. Kohnert, J. Gigax, W.-Y. Chen, M. Li, Y. Q. Wang, J. S. Wróbel, D. Nguyen-Manh, J. K. S. Baldwin, O. U. Tukac, E. Aydogan, S. Fensin, and E. Martinez, A quinary WTaCrVHF nanocrystalline refractory high-entropy alloy withholding extreme irradiation environments, *Nature Communications* **14** (2023).
- [52] O. N. Senkov, D. B. Miracle, K. J. Chaput, and J.-P. Couzinié, Development and exploration of refractory high entropy alloys—a review, *Journal of Materials Research* **33**, 3092–3128 (2018).
- [53] J.-P. Couzinié and G. Dirras, Body-centered cubic high-entropy alloys: From processing to underlying deformation mechanisms, *Materials Characterization* **147**, 533 (2019).
- [54] L. Lilensten, J.-P. Couzinié, L. Perrière, A. Hocini, C. Keller, G. Dirras, and I. Guillot, Study of a bcc multi-principal element alloy: Tensile and simple shear properties and underlying deformation mechanisms, *Acta Materialia* **142**, 131 (2018).
- [55] D. Caillard and J.-L. Martin, *Thermally activated mechanisms in crystal plasticity* (Elsevier, 2003).
- [56] X.-G. Li, C. Chen, H. Zheng, Y. Zuo, and S. P. Ong, Complex strengthening mechanisms in the NbMoTaW multi-principal element alloy, *npj Computational Materials* **6**, 70 (2020).
- [57] S. Yin, Y. Zuo, A. Abu-Odeh, H. Zheng, X.-G. Li, J. Ding, S. P. Ong, M. Asta, and R. O. Ritchie, Atomistic simulations of dislocation mobility in refractory high-entropy alloys and the effect of chemical short-range order, *Nature communications* **12**, 4873 (2021).
- [58] H. Zheng, L. T. Fey, X.-G. Li, Y.-J. Hu, L. Qi, C. Chen, S. Xu, I. J. Beyerlein, and S. P. Ong, Multi-scale investigation of short-range order and dislocation glide in MoNbTi and TaNbTi multi-principal element alloys, *npj Computational Materials* **9**, 89 (2023).
- [59] L. A. Zepeda-Ruiz, A. Stukowski, T. Oppelstrup, and V. V. Bulatov, Probing the limits of metal plasticity with molecular dynamics simulations, *Nature* **550**, 492 (2017).
- [60] J. Hu, Y. Shi, X. Sauvage, G. Sha, and K. Lu, Grain boundary stability governs hardening and softening in extremely fine nanograined metals, *Science* **355**, 1292 (2017).
- [61] J. F. Ziegler and J. P. Biersack, The stopping and range of ions in matter, in *Treatise on Heavy-Ion Science: Volume 6: Astrophysics, Chemistry, and Condensed Matter*, edited by D. A. Bromley (Springer US, Boston, MA, 1985) pp. 93–129.
- [62] J. Liu, J. Byggmästar, Z. Fan, P. Qian, and Y. Su, Large-scale machine-learning molecular dynamics simulation of primary radiation damage in tungsten, *Phys. Rev. B* **108**, 054312 (2023).
- [63] P. A. T. Olsson, Transverse resonant properties of strained gold nanowires, *Journal of Applied Physics* **108**, 034318 (2010).
- [64] J. Yi, S. Tang, M. Xu, L. Yang, L. Wang, and L. Zeng, A novel Al0.5CrCuNiV 3d transition metal high-entropy alloy: Phase analysis, microstructure and compressive properties, *Journal of Alloys and Compounds* **846**, 156466 (2020).
- [65] S. Kashiwaya, Y. Shi, J. Lu, D. G. Sangiovanni, G. Greczynski, M. Magnuson, M. Andersson, J. Rosen, and L. Hultman, Synthesis of goldene comprising single-atom layer gold, *Nature Synthesis* , 1 (2024).
- [66] Y. Tu, Z. Mao, and D. N. Seidman, Phase-partitioning and site-substitution patterns of molybdenum in a model Ni-Al-Mo superalloy: An atom-probe tomographic and first-principles study, *Applied Physics Letters* **101**, 121910 (2012).
- [67] C. Jia, K. Ishida, and T. Nishizawa, Partition of alloying elements between γ (a1), γ' (l12), and β (b2) phases in ni-al base systems, *Metallurgical and Materials Transactions A* **25**, 473 (1994).
- [68] Y. Mu, H. Liu, Y. Liu, X. Zhang, Y. Jiang, and T. Dong, An ab initio and experimental studies of the structure, mechanical parameters and state density on the refractory high-entropy alloy systems, *Journal of Alloys and Compounds* **714**, 668 (2017).
- [69] J. Wang, H. Gao, Y. Han, C. Ding, S. Pan, Y. Wang, Q. Jia, H.-T. Wang, D. Xing, and J. Sun, MAGUS: machine learning and graph theory assisted universal structure searcher, *National Science Review* **10**, nwad128 (2023).
- [70] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [71] P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* **50**, 17953 (1994).
- [72] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* **59**, 1758 (1999).
- [73] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [74] M. De Jong, W. Chen, T. Angsten, A. Jain, R. Notesline, A. Gamst, M. Sluiter, C. Krishna Ande, S. Van Der Zwaag, J. J. Plata, et al., Charting the complete elastic properties of inorganic crystalline compounds, *Scientific data* **2**, 1 (2015).
- [75] R. Tran, Z. Xu, B. Radhakrishnan, D. Winston, W. Sun, K. A. Persson, and S. P. Ong, Surface energies of elemental crystals, *Scientific data* **3**, 1 (2016).
- [76] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbørg, J. Kubal, K. Kaasbørg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Petersen, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, *Journal of Physics: Condensed Matter* **29**, 273002 (2017).

- [77] E. Lindgren, M. Rahm, E. Fransson, F. Eriksson, N. Österbacka, Z. Fan, and P. Erhart, calorine: A Python package for constructing and sampling neuroevolution potential models, *Journal of Open Source Software* **9**, 6264 (2024).
- [78] A. Togo, L. Chaput, T. Tadano, and I. Tanaka, Implementation strategies in phonopy and phono3py, *J. Phys. Condens. Matter* **35**, 353001 (2023).
- [79] W. E. Haynes, *CRC handbook of chemistry and physics (97th ed.)* (CRC Press llc Boca Raton, FL, 2016).
- [80] P. Hirel, Atomsk: A tool for manipulating and converting atomic data files, *Computer Physics Communications* **197**, 212 (2015).
- [81] M. Bernetti and G. Bussi, Pressure control using stochastic cell rescaling, *The Journal of Chemical Physics* **153**, 114107 (2020).
- [82] G. Bussi, D. Donadio, and M. Parrinello, Canonical sampling through velocity rescaling, *The Journal of Chemical Physics* **126**, 014101 (2007).
- [83] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—the Open Visualization Tool, *Modelling and Simulation in Materials Science and Engineering* **18**, 015012 (2009).
- [84] K. Nordlund, Molecular dynamics simulation of ion ranges in the 1-100 keV energy range, *Computational Materials Science* **3**, 448 (1995).
- [85] K. Song, J. Liu, S. Chen, Z. Fan, Y. Su, and P. Qian, Solute segregation in polycrystalline aluminum from hybrid monte carlo and molecular dynamics simulations with a unified neuroevolution potential (2024), arXiv:2404.13694 [cond-mat.mtrl-sci].

SUPPLEMENTARY NOTES

Supplementary Note S1: The INCAR input file for VASP

We have used the following inputs in the INCAR file of the VASP code for training data calculations.

```
GGA      = PE          # Default is POTCAR
ENCUT    = 600         # Default is POTCAR
KSPACING = 0.2        # Default is 0.5
KGAMMA   = .TRUE.     # This is the default
NELM     = 120         # Default is 60
ALGO     = Normal      # This is the default
EDIFF    = 1E-06       # Default is 1E-4
SIGMA    = 0.02        # Default is 0.2
ISMEAR   = 0           # Default is 1
PREC     = Accurate    # Default is Normal
LREAL    = A           # Default is .FALSE.
```

Supplementary Note S2: The nep.in input file for GPUMD and details of the multi-loss SNES training algorithm

We have used the following inputs in the nep.in file of the GPUMD code to train UNEP-v1:

```
type      16 Ag Al Au Cr Cu Mg Mo Ni Pb Pd Pt Ta Ti V W Zr
version   4
cutoff    6 5
n_max    4 4
basis_size 8 8
l_max    4 2 1
neuron   80
lambda_1  0
lambda_e  1
lambda_f  1
lambda_v  0.1
batch    10000
population 60
generation 1000000
zbl      2
```

We explain the hyperparameters one by one:

- The **type** keyword specifies the number and chemical symbols of the species we considered.
- The **version** keyword specifies the version of the NEP model used, which is NEP4 introduced in the present work.
- The **cutoff** keyword specifies that the cutoff radii for the radial (2-body) and angular (from 3-body to 5-body) descriptor components are 6 and 5 Å respectively.
- The **n_max** keyword specifies that the number of radial functions for the radial and angular descriptor components are both $4 + 1 = 5$.
- The **basis_size** keyword specifies that the number of radial basis functions for the radial and angular descriptor components are both $8 + 1 = 9$.
- The **l_max** keyword specifies that the levels of the spherical harmonics for the 3-body, 4-body, and 5-body angular descriptor components are up to 4, 2, and 1 respectively.

- The `neuron` keyword specifies that the number of neurons in the hidden layer is 80.
- The `lambda_1` keyword specifies that no \mathcal{L}_1 regularization is applied here, but the \mathcal{L}_2 regularization is applied by default.
- The `lambda_e` keyword specifies the weight of the energy loss term to be 1.
- The `lambda_f` keyword specifies the weight of the force loss term to be 1.
- The `lambda_v` keyword specifies the weight of the virial loss term to be 0.1.
- The `batch` keyword specifies the batch size of training to be 10 000.
- The `population` keyword specifies the population size to be 60.
- The `generation` keyword specifies the total number of training generations (steps) to be 1 000 000.
- The `zbl` keyword specifies to use the NEP-ZBL model [62] with a cutoff radius of 2 Å for the universal ZBL potential.

The UNEP-v1 models were trained based on the following loss function:

$$L(\mathbf{z}) = L_e(\mathbf{z}) + L_f(\mathbf{z}) + L_v(\mathbf{z}) + L_2(\mathbf{z}), \quad (\text{S7})$$

$$L_e(\mathbf{z}) = \lambda_e \left(\frac{1}{N_{\text{str}}} \sum_{n=1}^{N_{\text{str}}} (U^{\text{NEP}}(n, \mathbf{z}) - U^{\text{ref}}(n))^2 \right)^{1/2}, \quad (\text{S8})$$

$$L_f(\mathbf{z}) = \lambda_f \left(\frac{1}{3N} \sum_{i=1}^N (\mathbf{F}_i^{\text{NEP}}(\mathbf{z}) - \mathbf{F}_i^{\text{ref}})^2 \right)^{1/2}, \quad (\text{S9})$$

$$L_v(\mathbf{z}) = \lambda_v \left(\frac{1}{6N_{\text{str}}} \sum_{n=1}^{N_{\text{str}}} \sum_{\mu\nu} (W_{\mu\nu}^{\text{NEP}}(n, \mathbf{z}) - W_{\mu\nu}^{\text{ref}}(n))^2 \right)^{1/2}, \quad (\text{S10})$$

$$L_2(\mathbf{z}) = \lambda_2 \left(\frac{1}{N_{\text{par}}} \sum_{n=1}^{N_{\text{par}}} z_n^2 \right)^{1/2}. \quad (\text{S11})$$

Here, \mathbf{z} denotes the vector formed by the N_{par} trainable parameters in the model, N_{str} is the number of structures containing the given species in one batch, N is the total number of atoms in these structures, $U^{\text{NEP}}(n, \mathbf{z})$ and $W_{\mu\nu}^{\text{NEP}}(n, \mathbf{z})$ are the energy and virial for the n th structure calculated by the current UNEP-v1 models, $U^{\text{ref}}(n)$ and $W_{\mu\nu}^{\text{ref}}(n)$ are the corresponding reference values, $\mathbf{F}_i^{\text{NEP}}(\mathbf{z})$ is the force on atom i calculated by the current UNEP-v1 models, $\mathbf{F}_i^{\text{ref}}$ is the corresponding reference value, and $L_2(\mathbf{z})$ is the \mathcal{L}_2 regularization term. The SNES training algorithm [31] with the extension of the multi-loss approach we proposed is given below. (1) Create a search distribution in the solution space, including a set of mean values \mathbf{m} and variances \mathbf{s} . The elements in \mathbf{m} are uniformly distributed from -1 to 1 and those in \mathbf{s} take a constant value of 0.1. (2) Starting from the above initial distribution, we perform N_{gen} iterations. First create N_{pop} solutions \mathbf{z}_k ($1 \leq k \leq N_{\text{pop}}$) based on the current \mathbf{m} and \mathbf{s} vectors:

$$\mathbf{z}_k \leftarrow \mathbf{m} + \mathbf{s} \odot \mathbf{r}_k. \quad (\text{S12})$$

Here, N_{pop} is the population size and \mathbf{r}_k is a set of normal-distribution random numbers with zero mean and unit variance. Then evaluate the loss values $L(\mathbf{z}_k)$ for all the solutions \mathbf{z}_k in the population and sort the solutions based on the loss values. Here, *each species has its own set of loss values and sorting scheme*, which is the key point in our multi-loss approach. Last, calculate the natural gradients (where u_k is a set of rank-based utility values [31]),

$$\nabla_{\mathbf{m}} J \leftarrow \sum_{k=1}^{N_{\text{pop}}} u_k \mathbf{r}_k, \quad (\text{S13})$$

$$\nabla_{\mathbf{s}} J \leftarrow \sum_{k=1}^{N_{\text{pop}}} u_k (\mathbf{r}_k \odot \mathbf{r}_k - 1), \quad (\text{S14})$$

and update the means and variances of the search distribution,

$$\mathbf{m} \leftarrow \mathbf{m} + \eta_{\mathbf{m}} (\mathbf{s} \odot \nabla_{\mathbf{m}} J), \quad (\text{S15})$$

$$\mathbf{s} \leftarrow \mathbf{s} \odot \exp \left(\frac{\eta_{\mathbf{s}}}{2} \nabla_{\mathbf{s}} J \right), \quad (\text{S16})$$

where $\eta_{\mathbf{m}} = 1$, and $\eta_{\mathbf{s}} = (3 + \ln N_{\text{par}}) / 5\sqrt{N_{\text{par}}}$.

SUPPLEMENTARY FIGURES

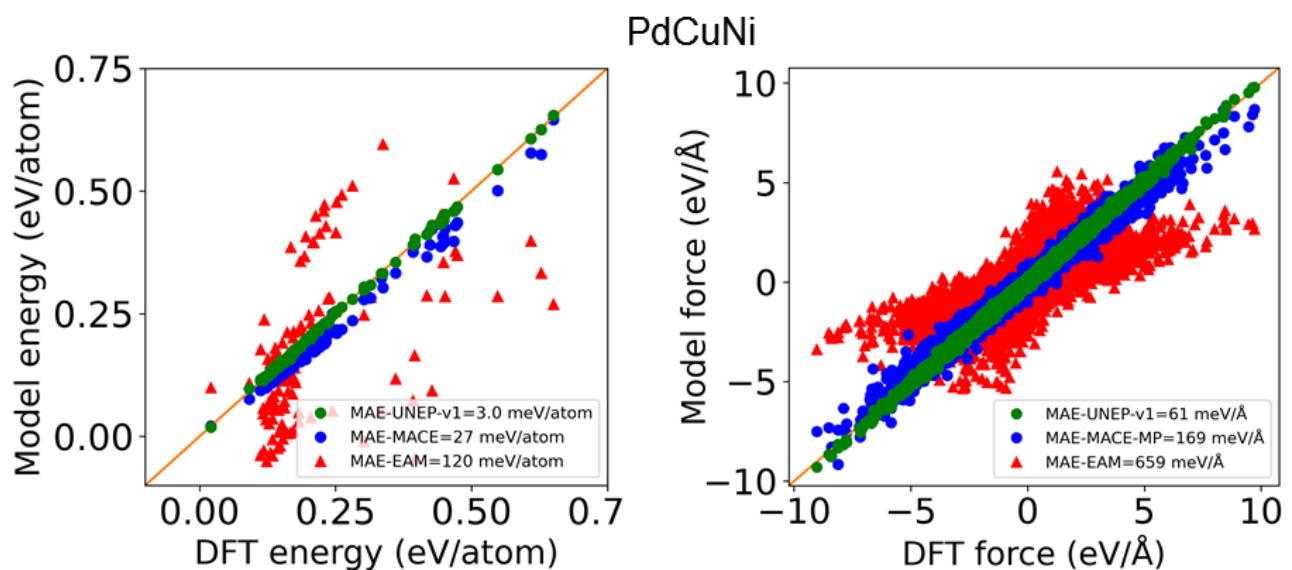


Figure S1. Parity plot for formation energy and force predictions from UNEP-v1, MACE-MP-0 (medium model) and EAM compared to DFT for the test dataset containing up to 3 components (Pd, Cu, Ni) from Zhao *et al.* [38]. UNEP-v1 demonstrates significantly better predictions.

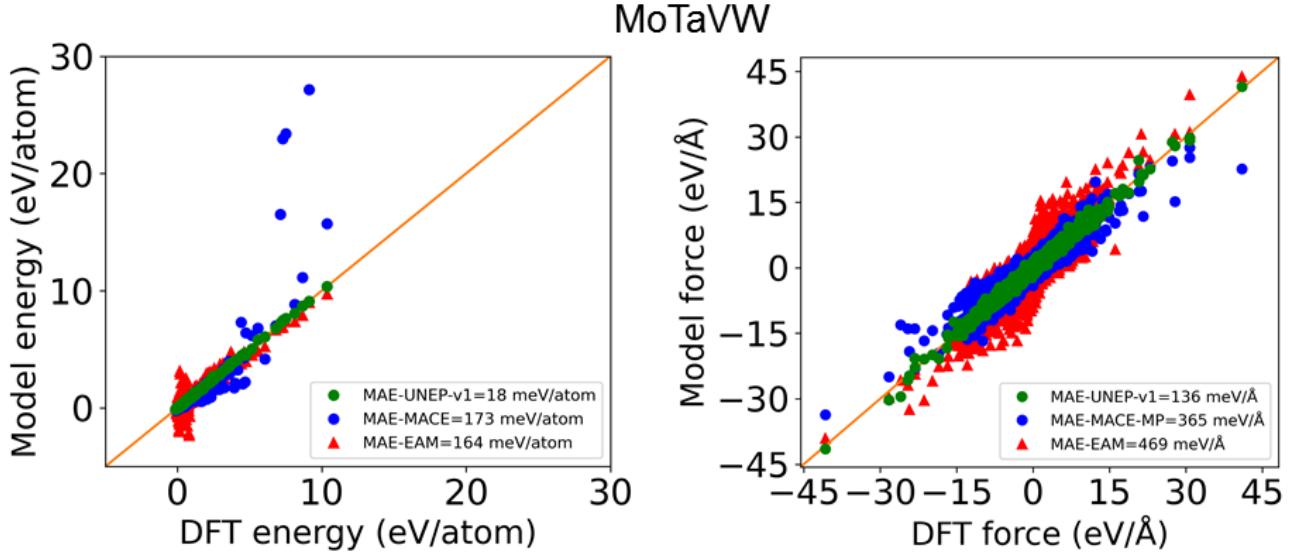


Figure S2. Parity plot for formation energy and force predictions from UNEP-v1, MACE-MP-0 (medium model) and EAM compared to DFT for the test dataset containing up to 4 components (Mo, Ta, V, W) from Byggmästar *et al.* [37]. UNEP-v1 shows much better predictions.

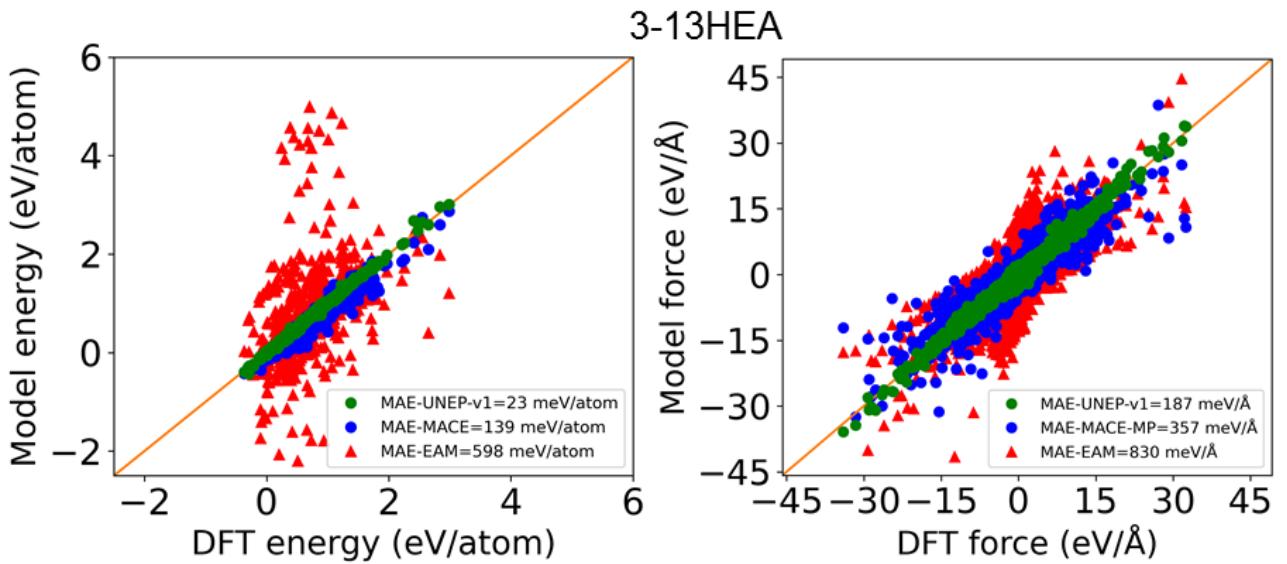


Figure S3. Parity plot for formation energy and force predictions from UNEP-v1, MACE-MP-0 (medium model) and EAM compared to DFT for the test dataset containing up to 13 components (Ag, Au, Cr, Cu, Mo, Ni, Pd, Pt, Ta, Ti, V, W, Zr) from Lopanitsyna *et al.* [36]. UNEP-v1 demonstrates significantly better predictions.

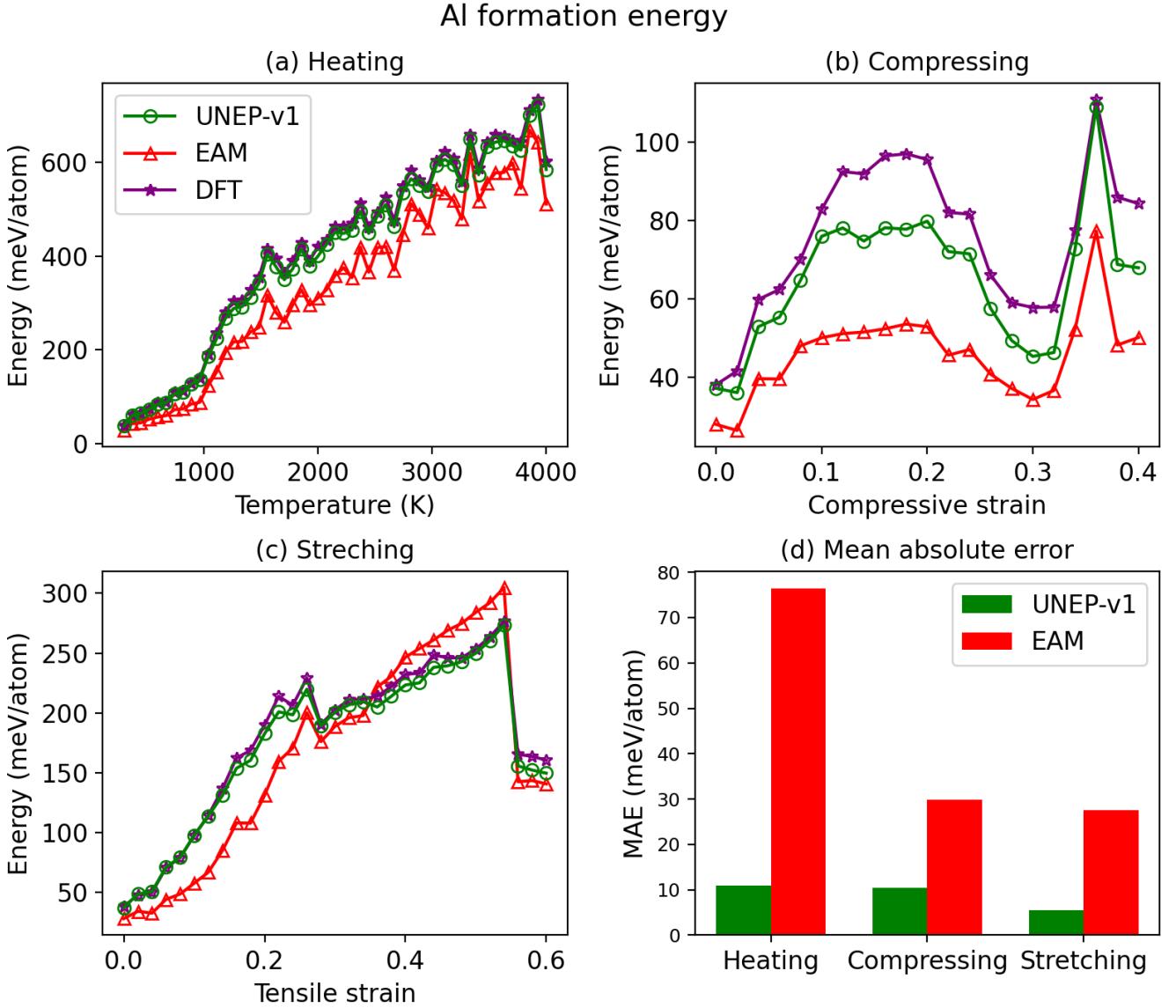


Figure S4. Comparison of the formation energies of FCC Al during (a) a heating process from 300 to 4000 K, (b) a compressing process up to 40% compressive strain, and (c) a stretching process up to 60% tensile strain for UNEP-v1, EAM, and DFT. (d) The mean absolute errors (MAEs) for UNEP-v1 and EAM as compared to DFT. The heating rate in (a) is 7.4 K/ps. The deformation rate in (b) and (c) is 2×10^8 s $^{-1}$. The trajectories were sampled using the UNEP-v1 model. UNEP-v1 shows much better predictions than EAM.

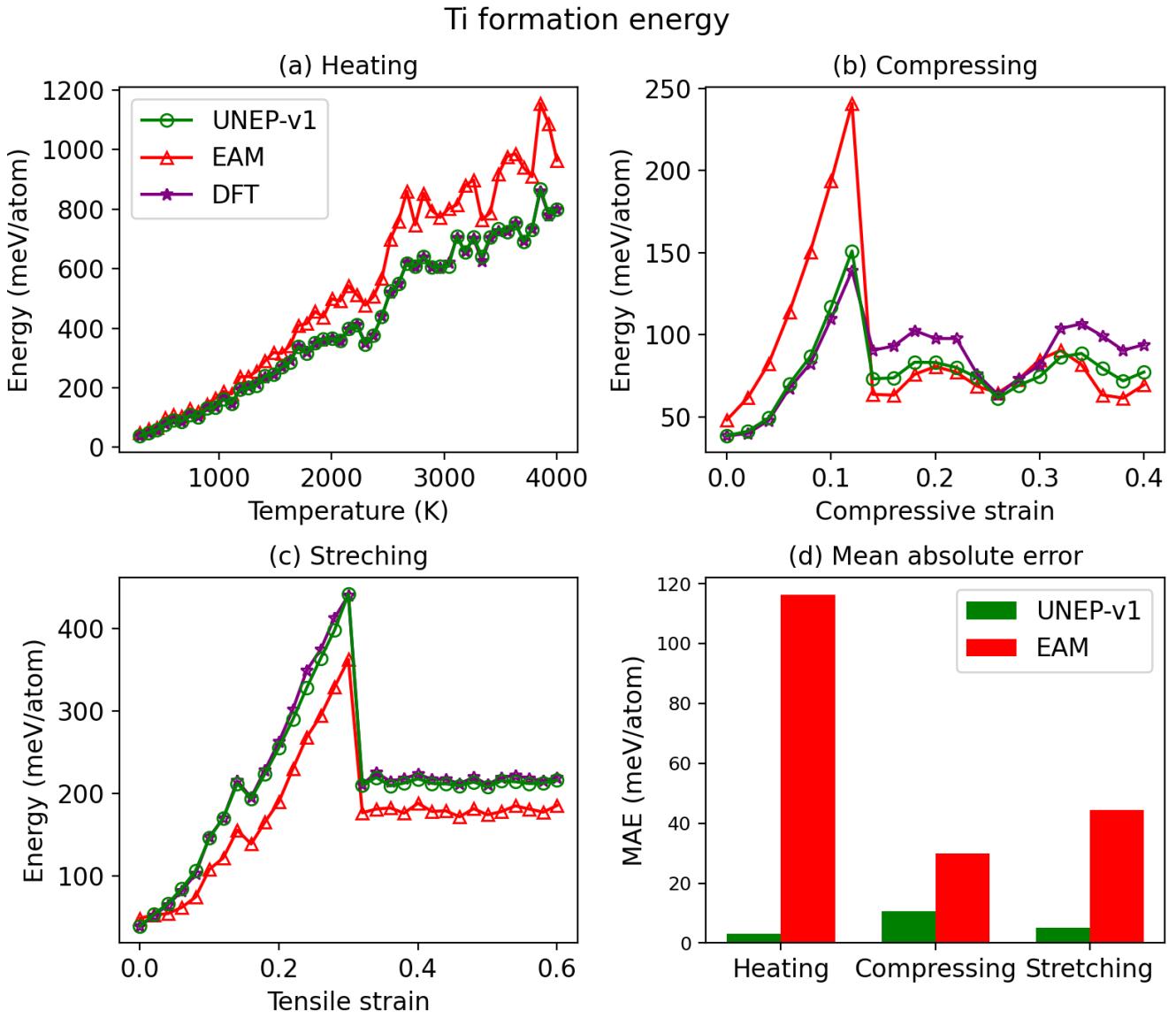


Figure S5. Similar to Fig. S4 but for HCP Ti.

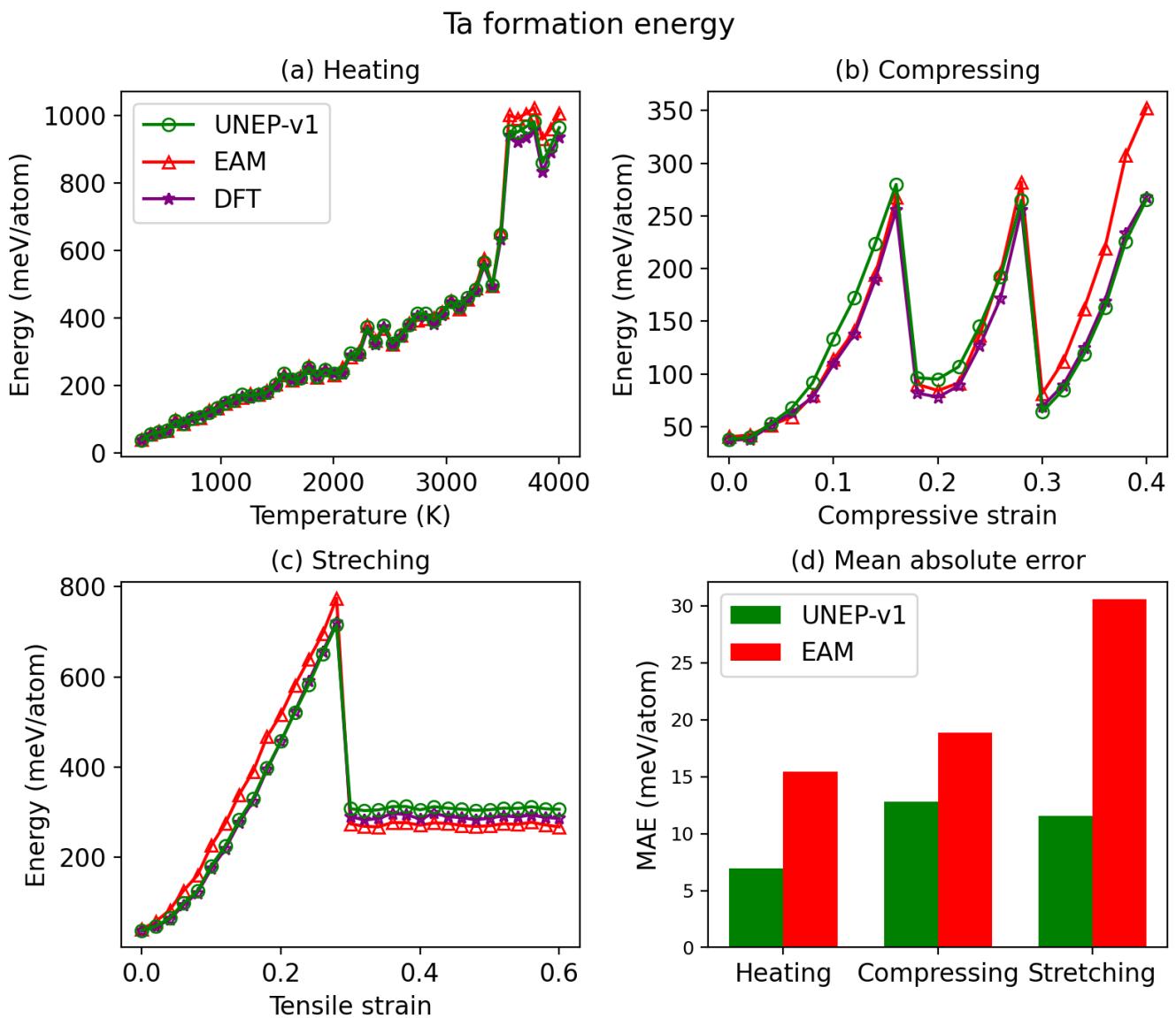


Figure S6. Similar to Fig. S4 but for BCC Ta.

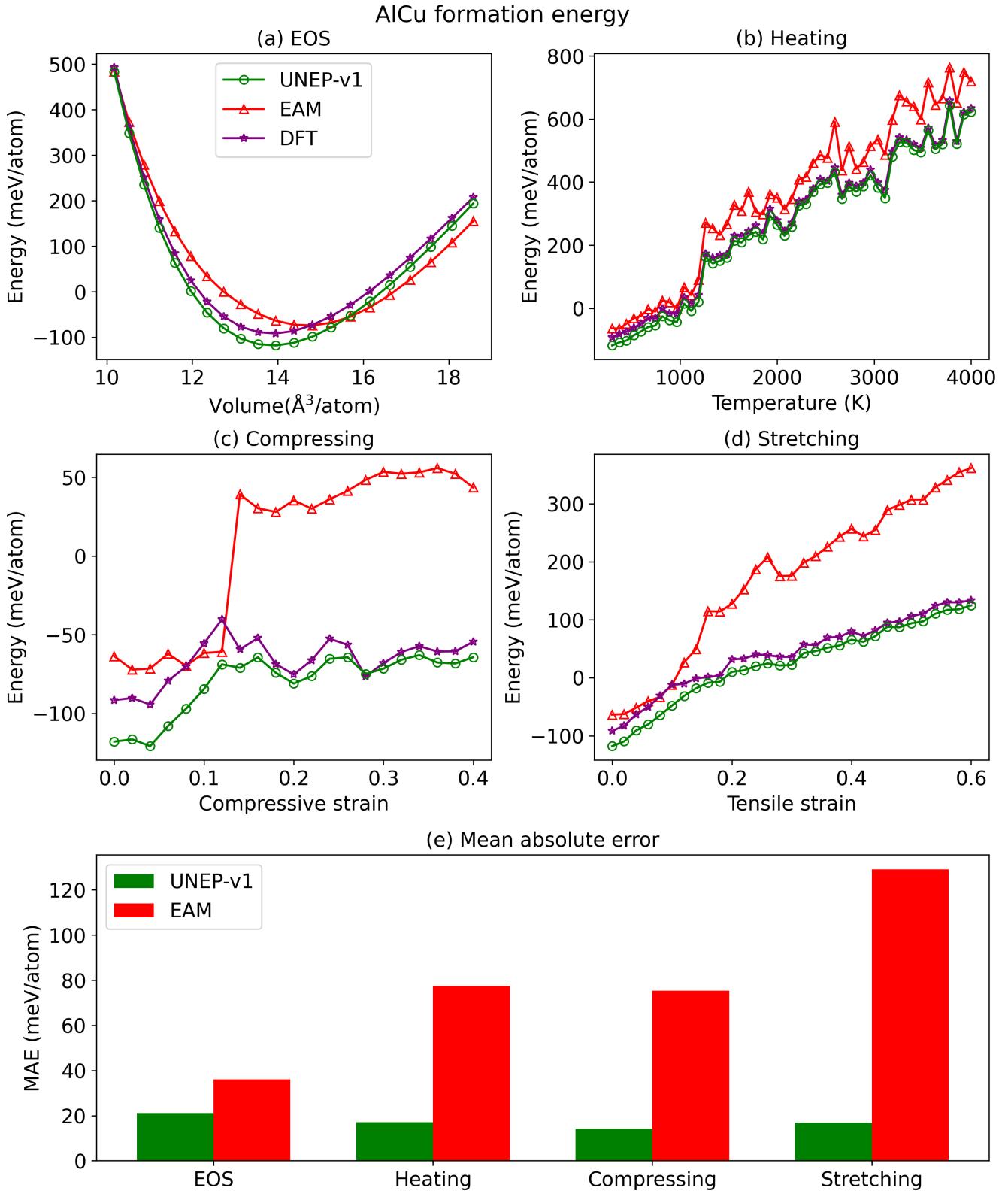


Figure S7. Comparison of the formation energies of FCC AlCu alloy for (a) the equation of state curve, (b) a heating process from 300 to 4000 K, (c) a compressing process up to 40% compressive strain, and (d) a stretching process up to 60% tensile strain as calculated from UNEP-v1, EAM, and DFT. (e) The mean absolute errors (MAEs) for UNEP-v1 and EAM as compared to DFT. The heating rate in (a) is 7.4 K/ps. The deformation rate in (b) and (c) is $2 \times 10^8 \text{ s}^{-1}$. The trajectories were sampled by using the UNEP-v1 model, with the initial structure chemically relaxed through hybrid Monte-Carlo and molecular dynamics (MCMD) simulations.

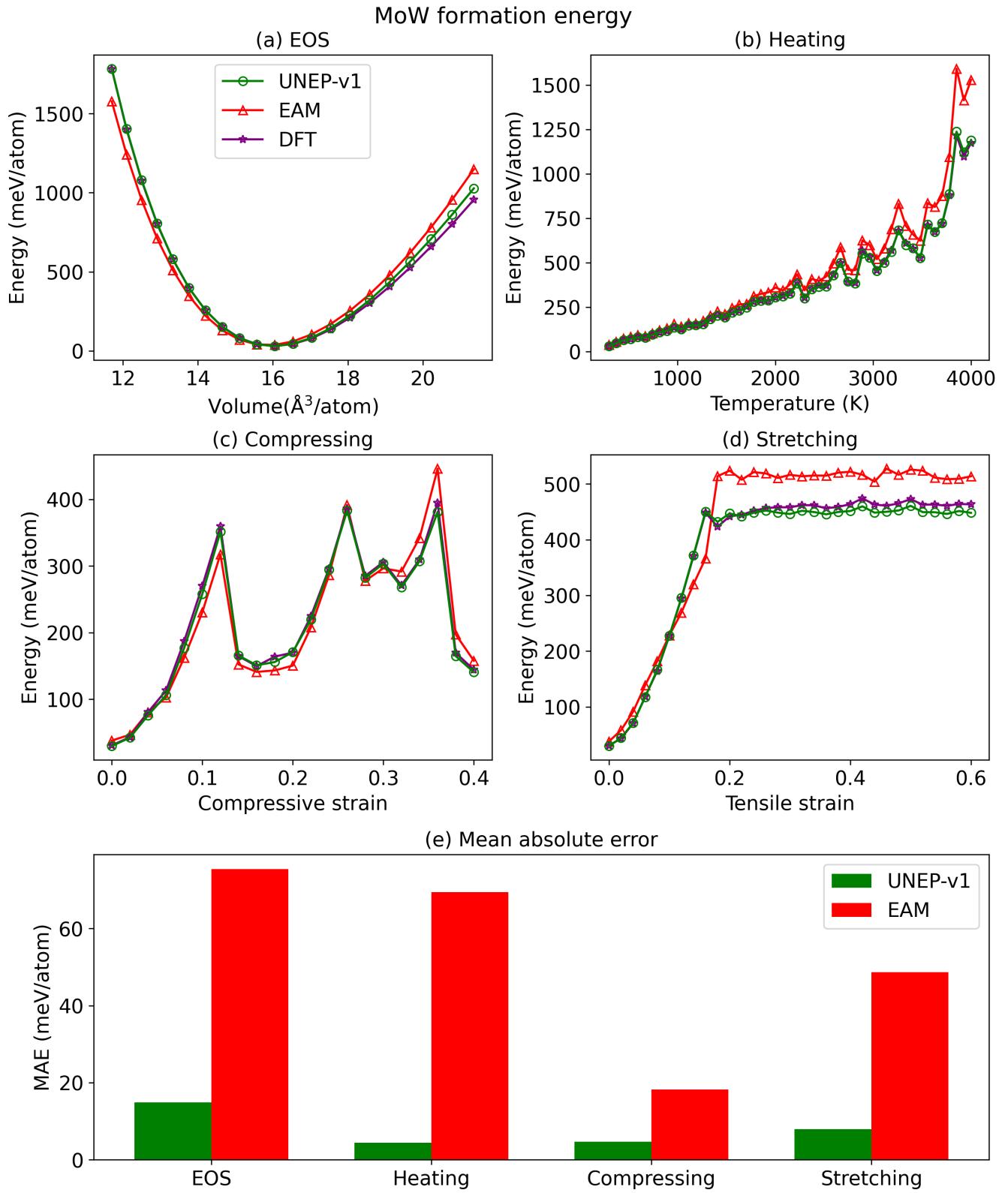


Figure S8. Similar to Fig. S7 but for BCC MoW.

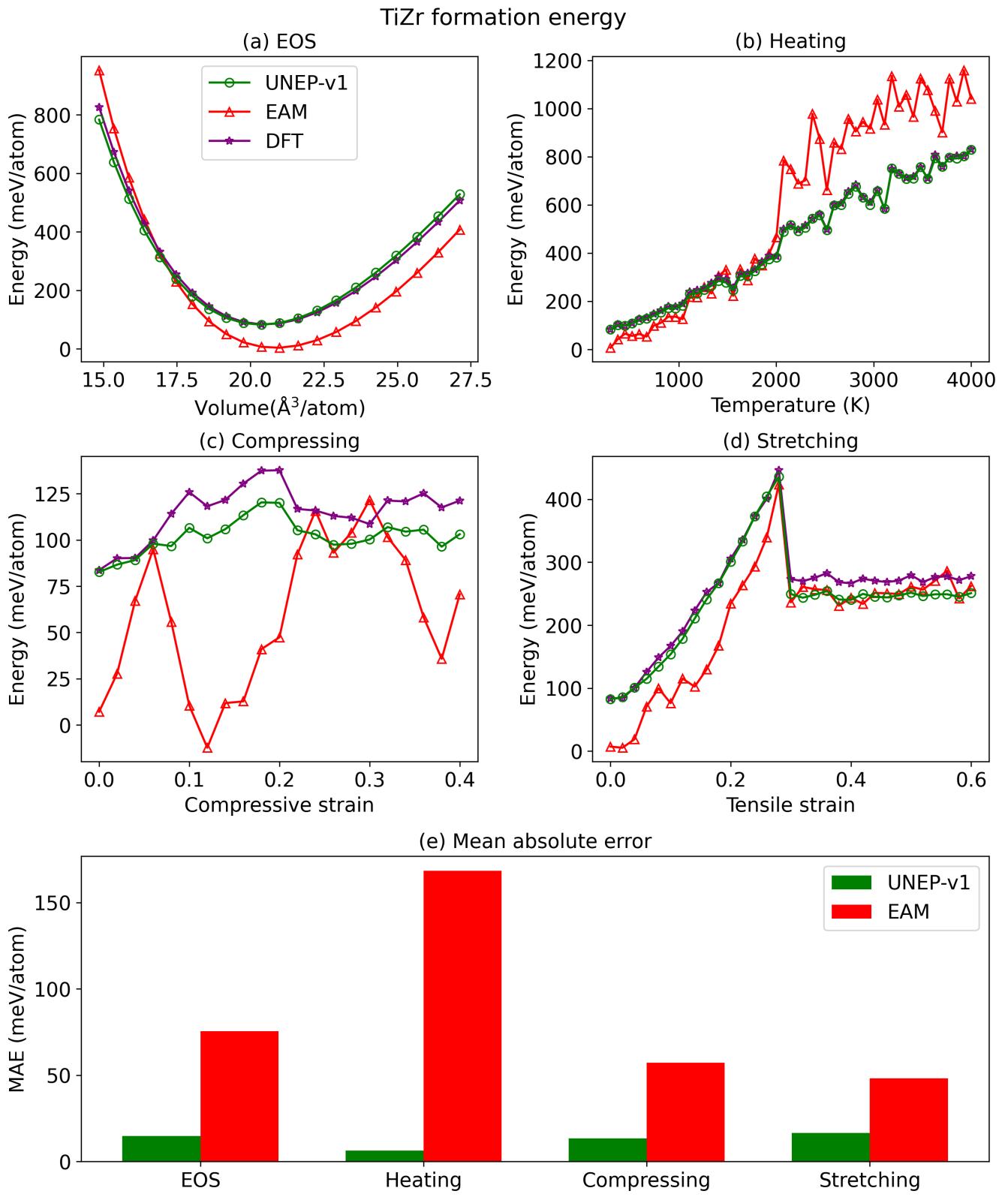


Figure S9. Similar to Fig. S7 but for HCP TiZr.

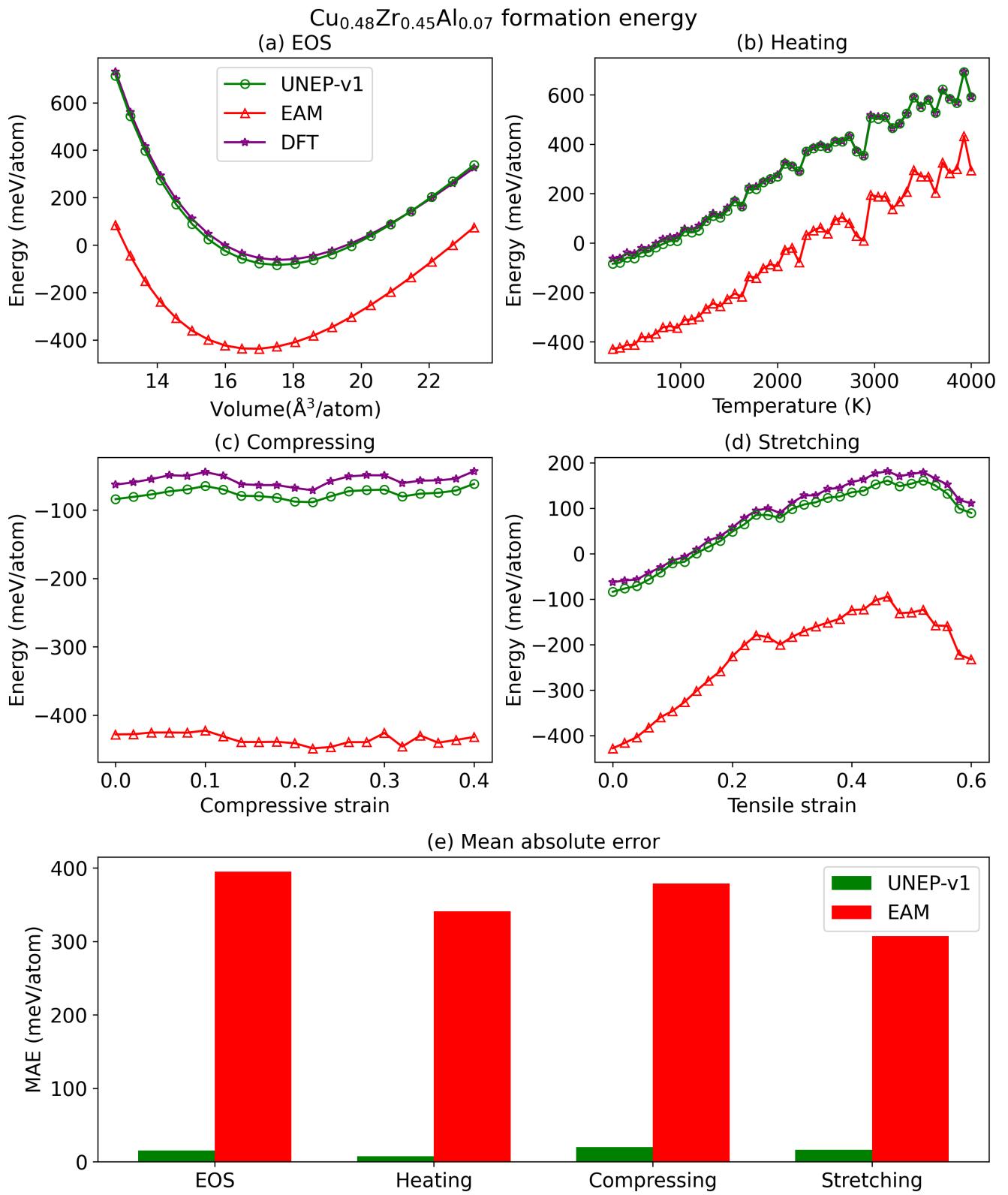


Figure S10. Similar to Fig. S7 but for metallic glassy $\text{Cu}_{0.475}\text{Zr}_{0.451}\text{Al}_{0.074}$.

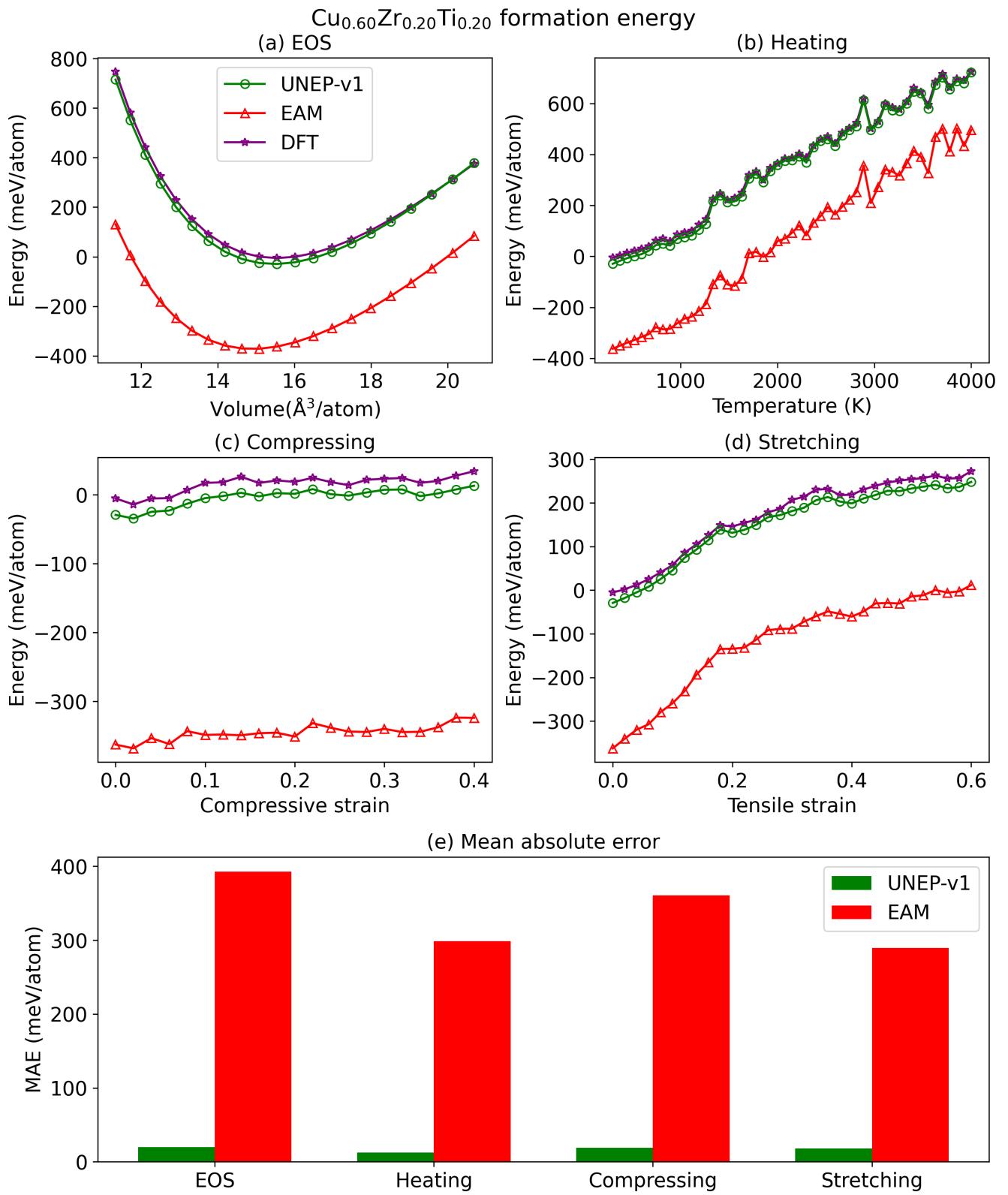


Figure S11. Similar to Fig. S7 but for metallic glassy $\text{Cu}_{0.6}\text{Zr}_{0.2}\text{Ti}_{0.2}$.

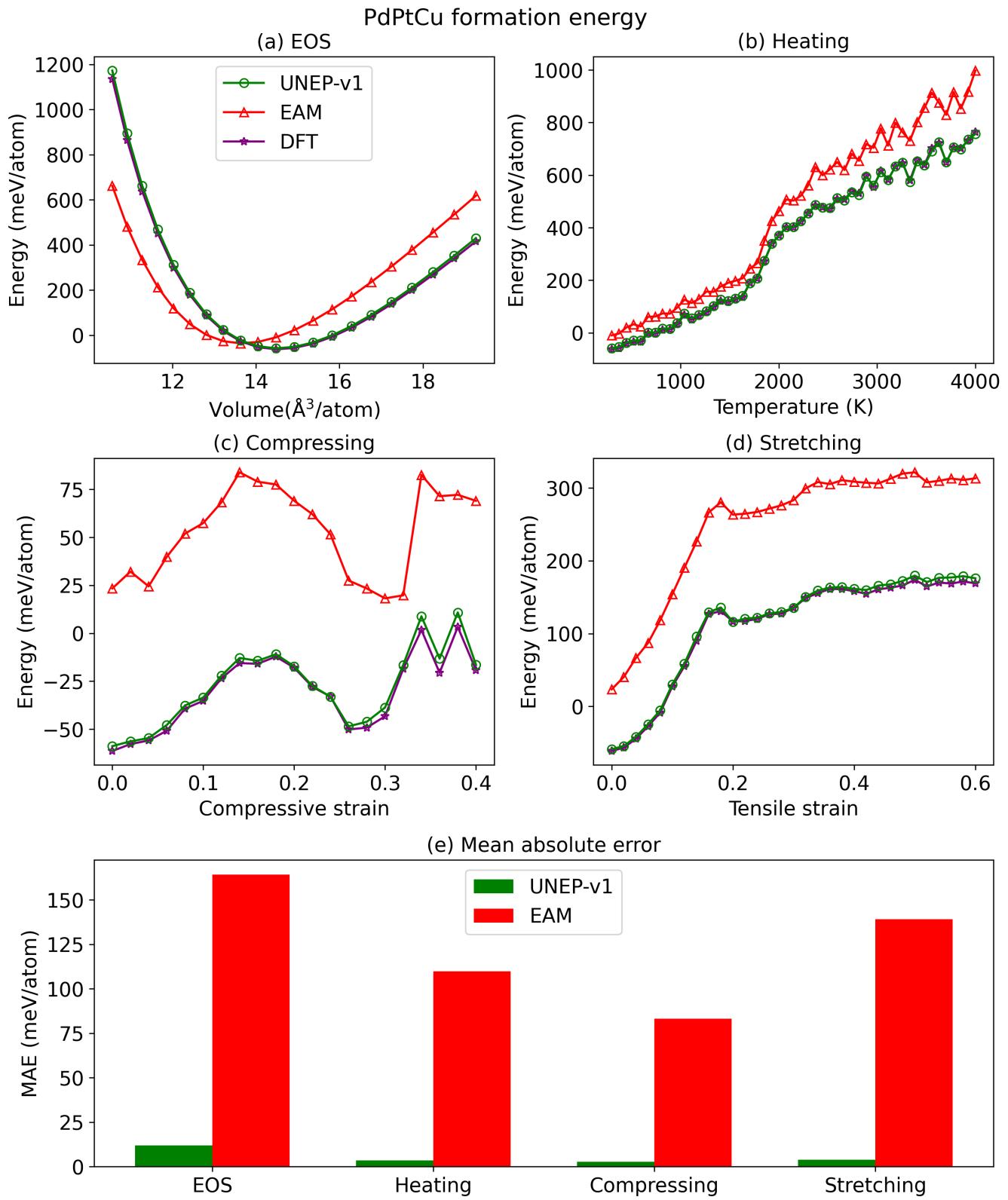


Figure S12. Similar to Fig. S7 but for FCC PdPtCu.

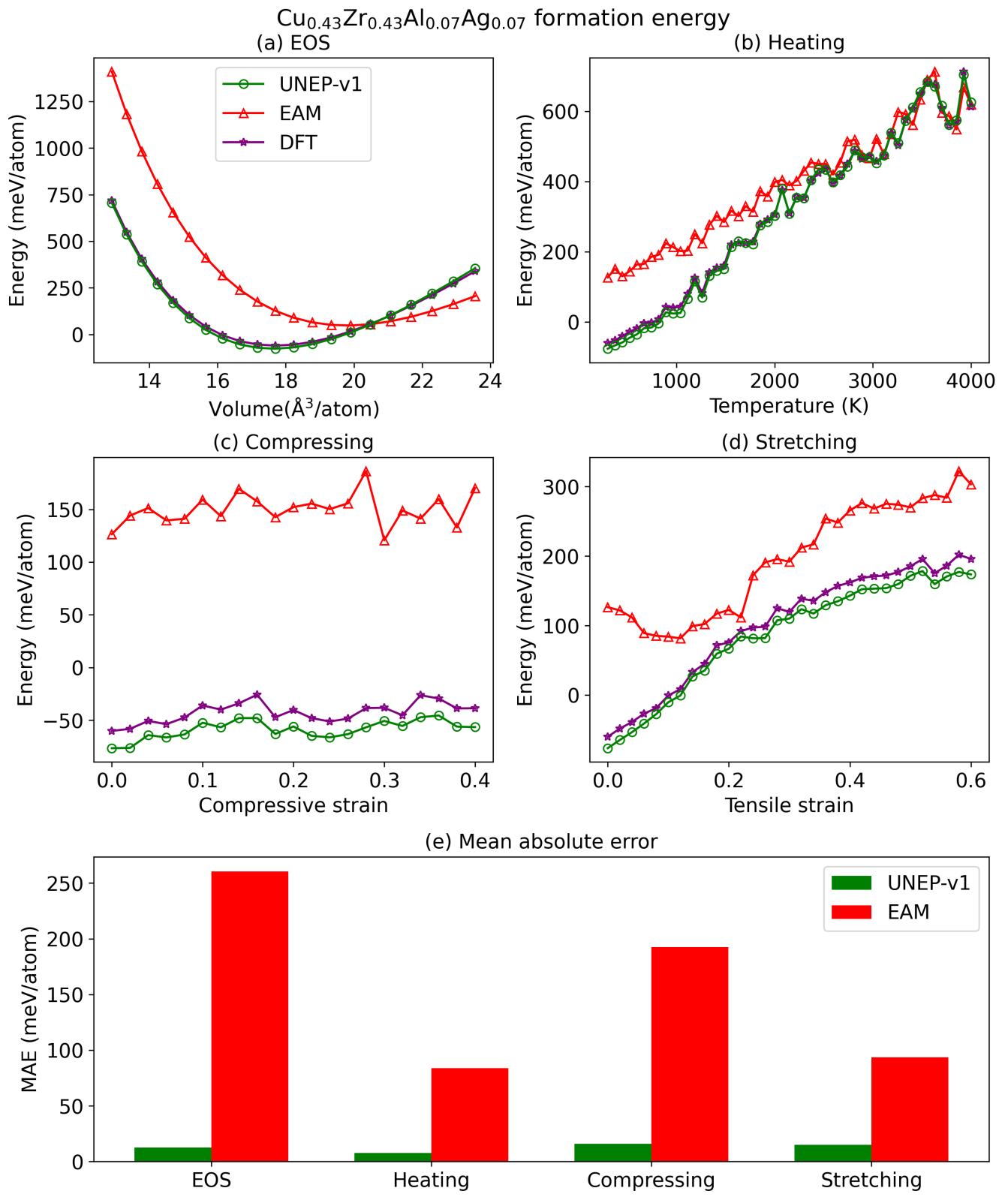


Figure S13. Similar to Fig. S7 but for metallic glassy $\text{Cu}_{0.43}\text{Zr}_{0.43}\text{Al}_{0.07}\text{Ag}_{0.07}$.

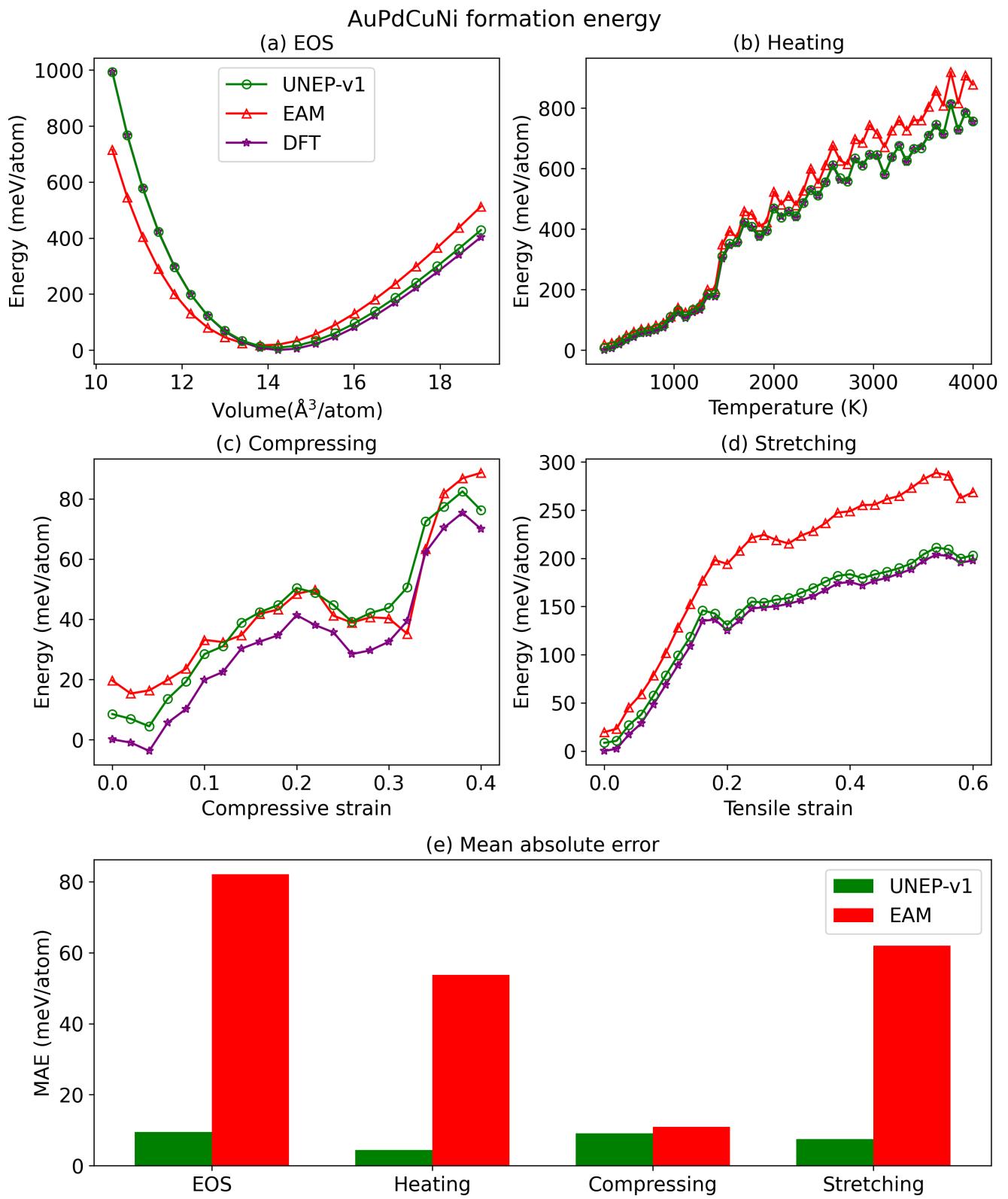


Figure S14. Similar to Fig. S7 but for FCC AuPdCuNi.

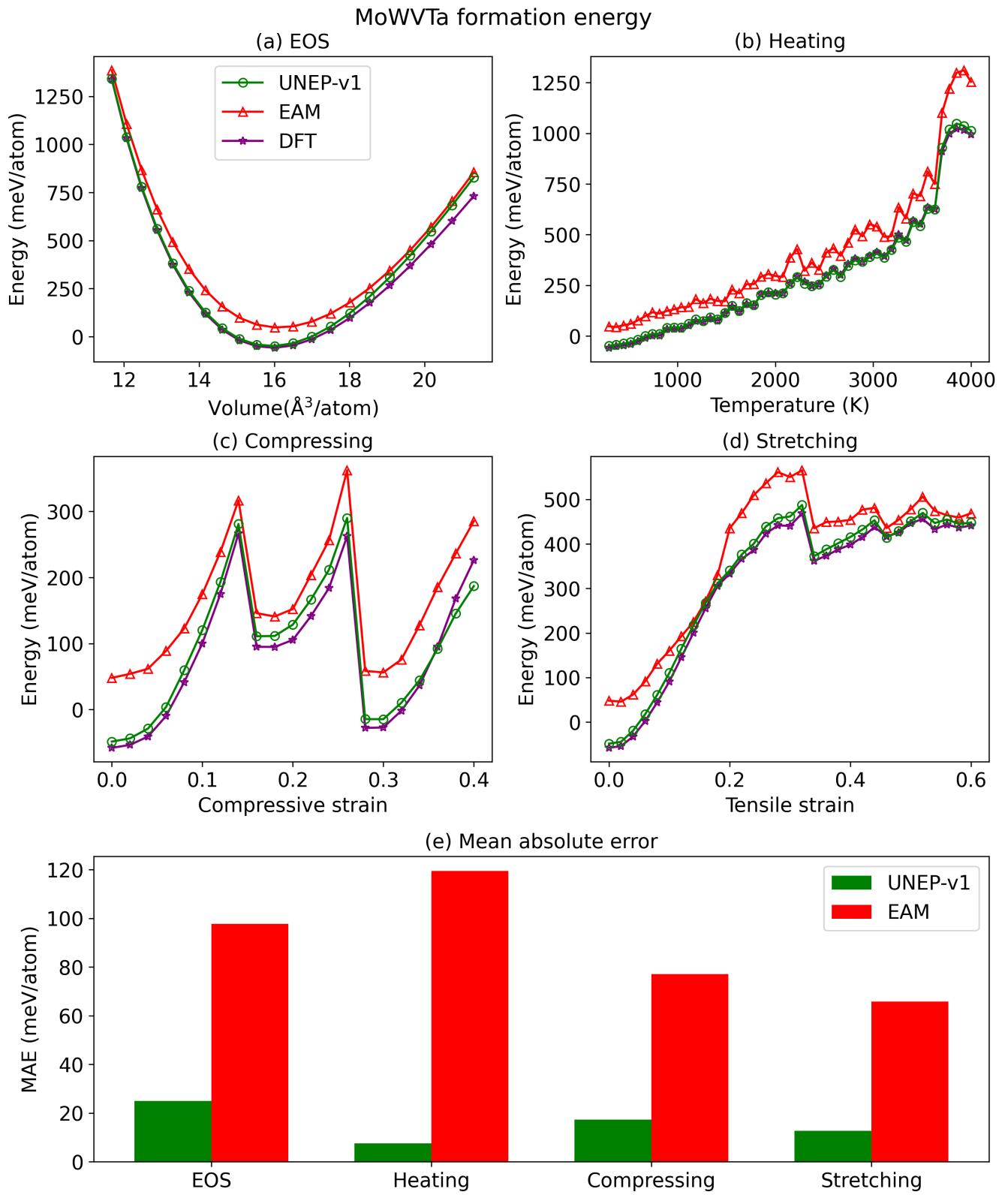


Figure S15. Similar to Fig. S7 but for BCC MoWVTa.

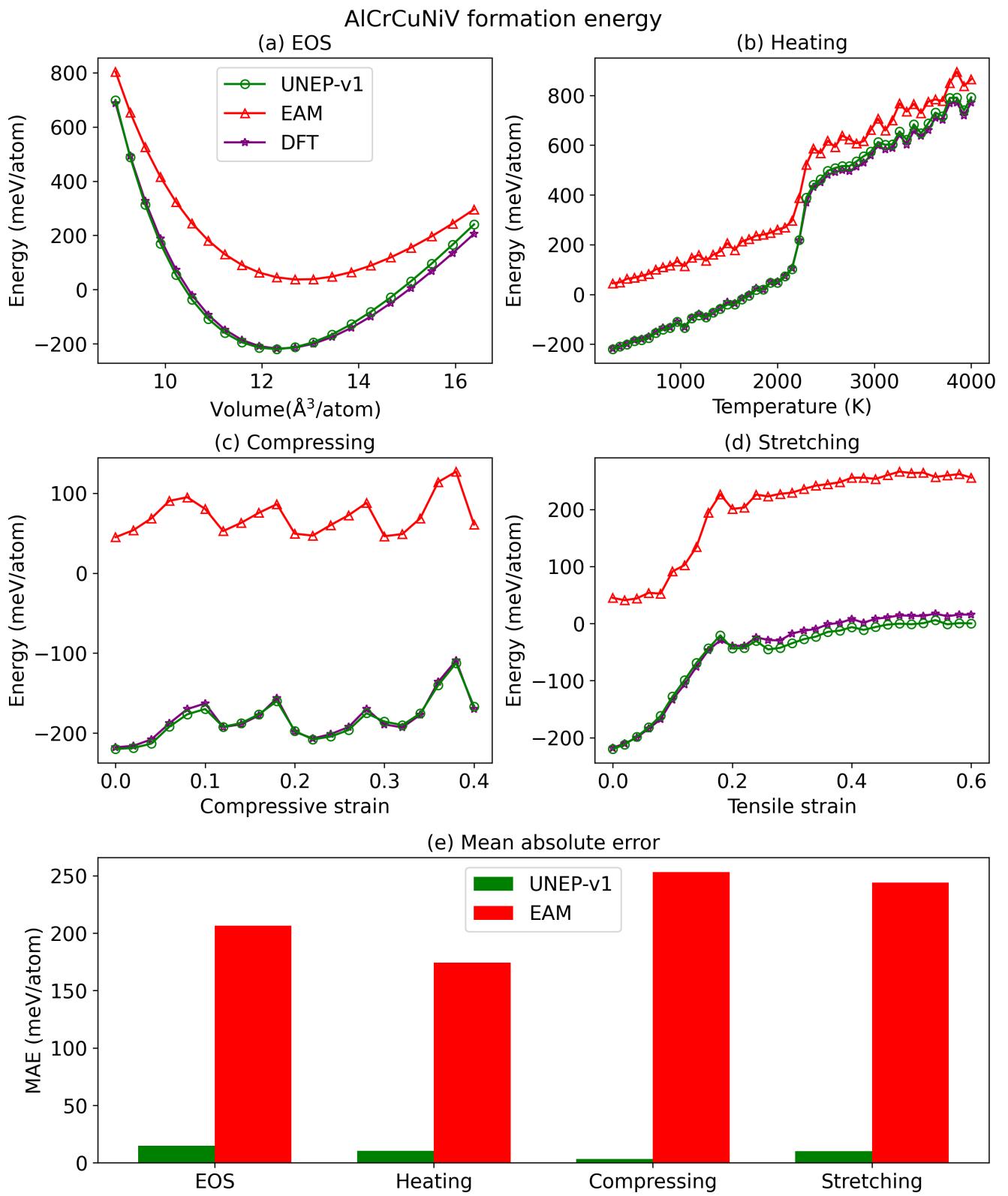


Figure S16. Similar to Fig. S7 but for BCC AlCrCuNiV.

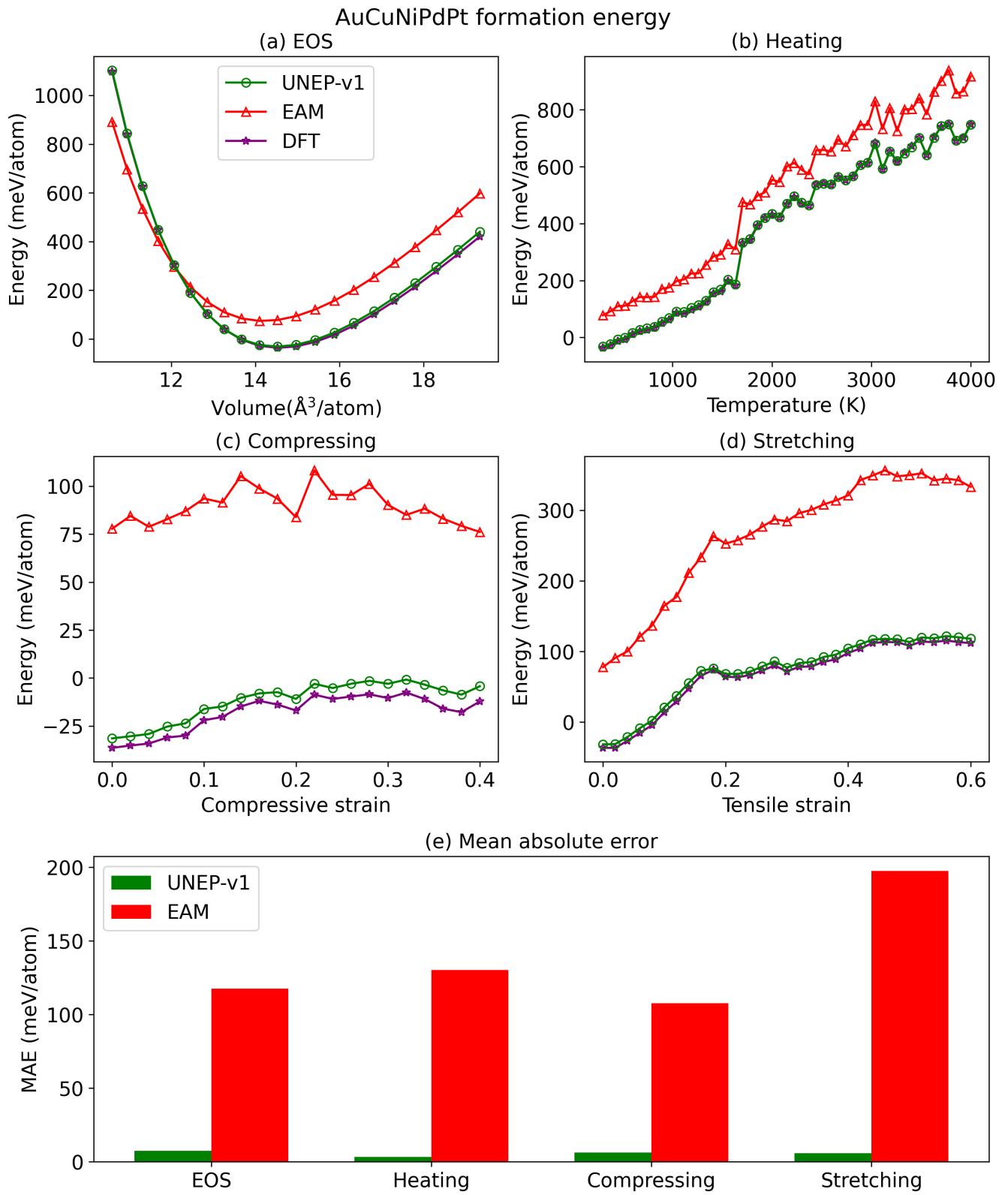


Figure S17. Similar to Fig. S7 but for FCC AuCuNiPdPt.

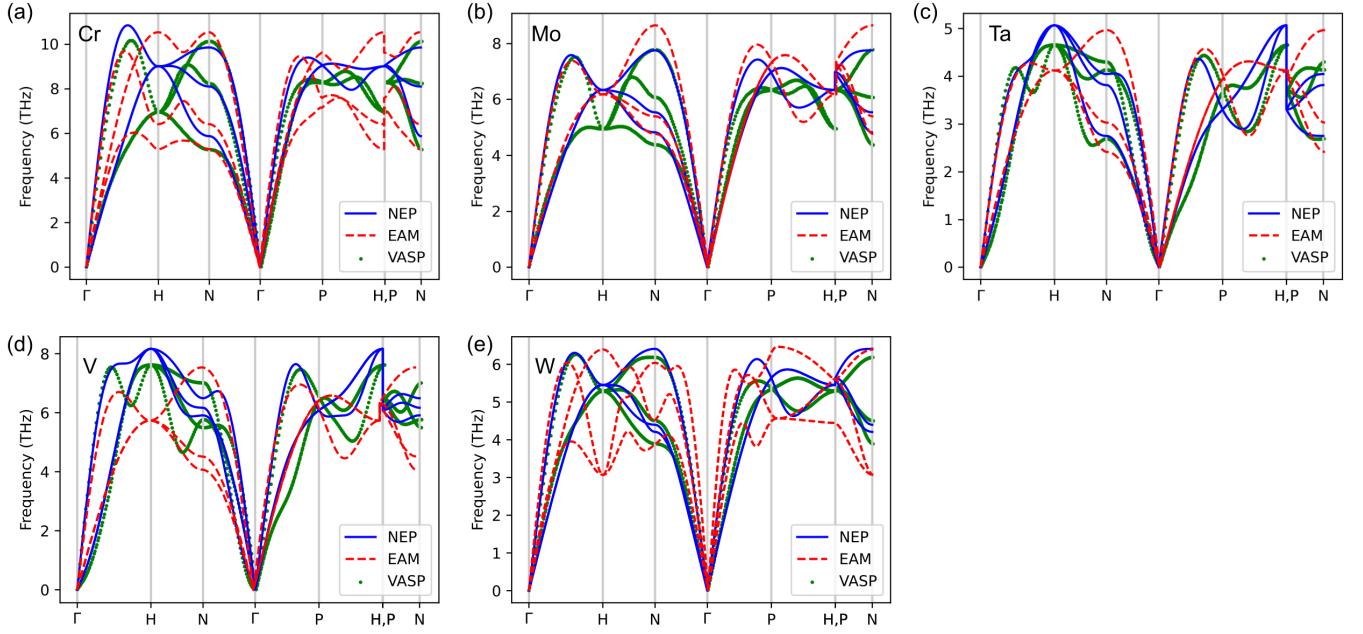


Figure S18. Phonon dispersion relations for the BCC metals from UNEP-v1, EAM and DFT (VASP).

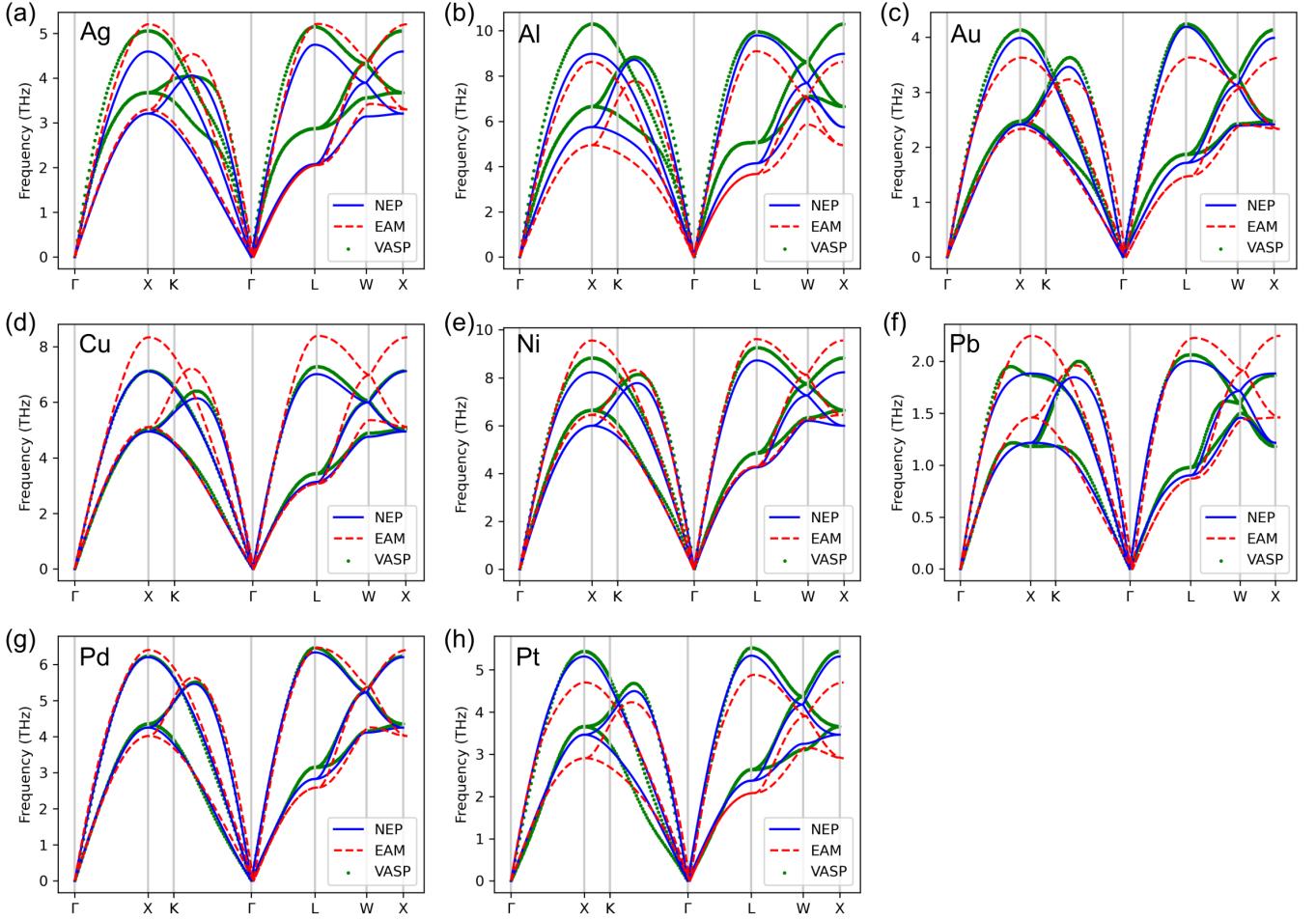


Figure S19. Phonon dispersion relations for the FCC metals from UNEP-v1, EAM, and DFT (VASP).

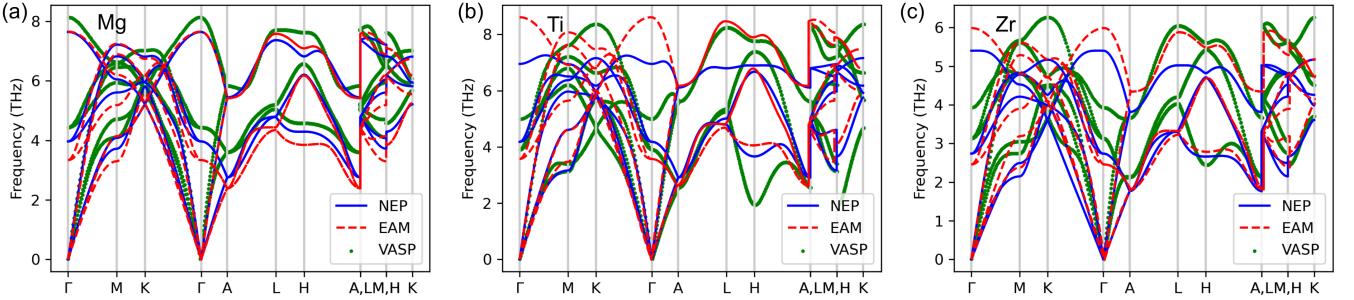


Figure S20. Phonon dispersion relations for the HCP metals from UNEP-v1, EAM, and DFT (VASP).

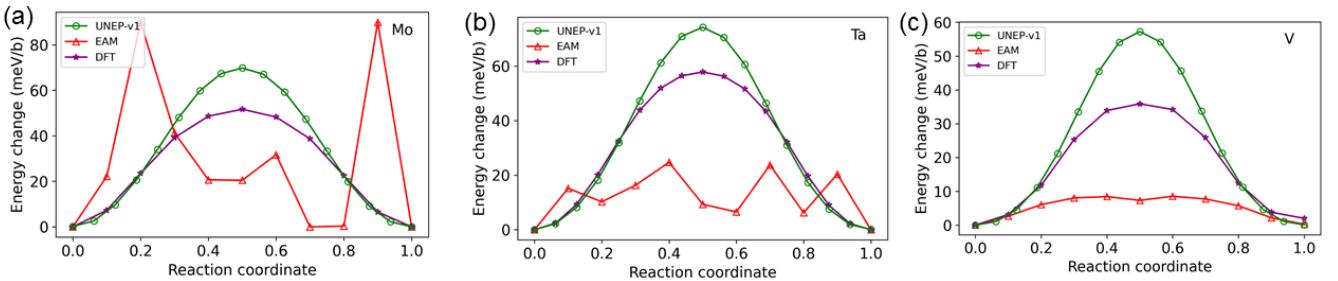


Figure S21. Peierls barrier for the $1/2\langle 111 \rangle$ screw dislocation migration in elemental Mo, Ta, and V, predicted by UNEP-v1, EAM [35], and DFT.

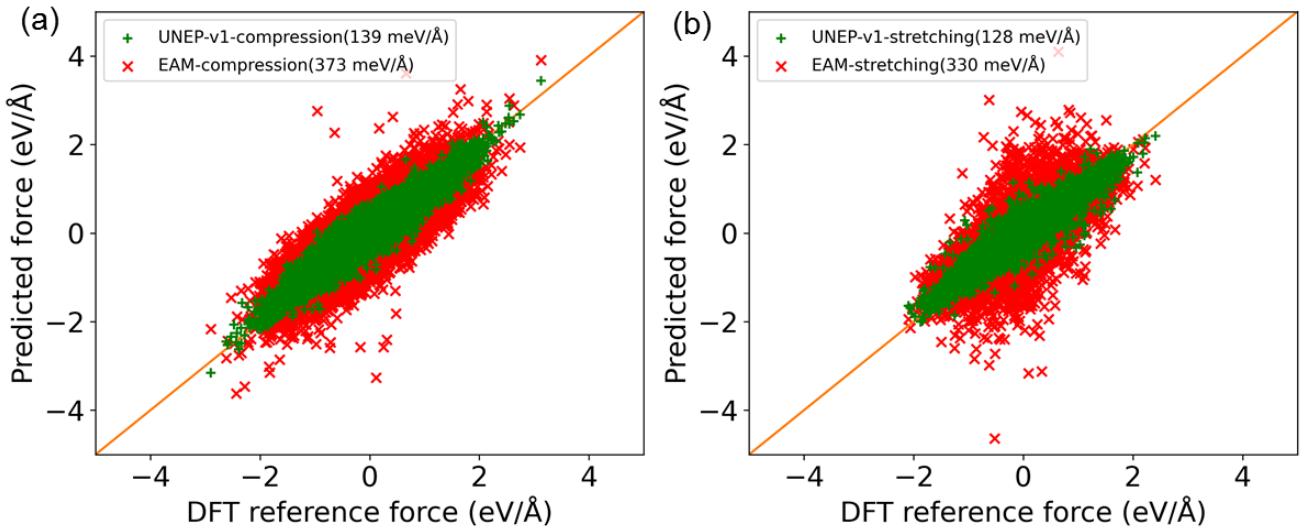


Figure S22. Parity plots for force predictions from UNEP-v1 and EAM [35] compared to DFT for equimolar MoTaVW alloys sampled from various MD simulations using 256-atom supercells, including deformation processes up to 25% compression (a) and stretching (b) at 300K. UNEP-v1 shows much better predictions than EAM [35], with much smaller force RMSEs as indicated in the legends.

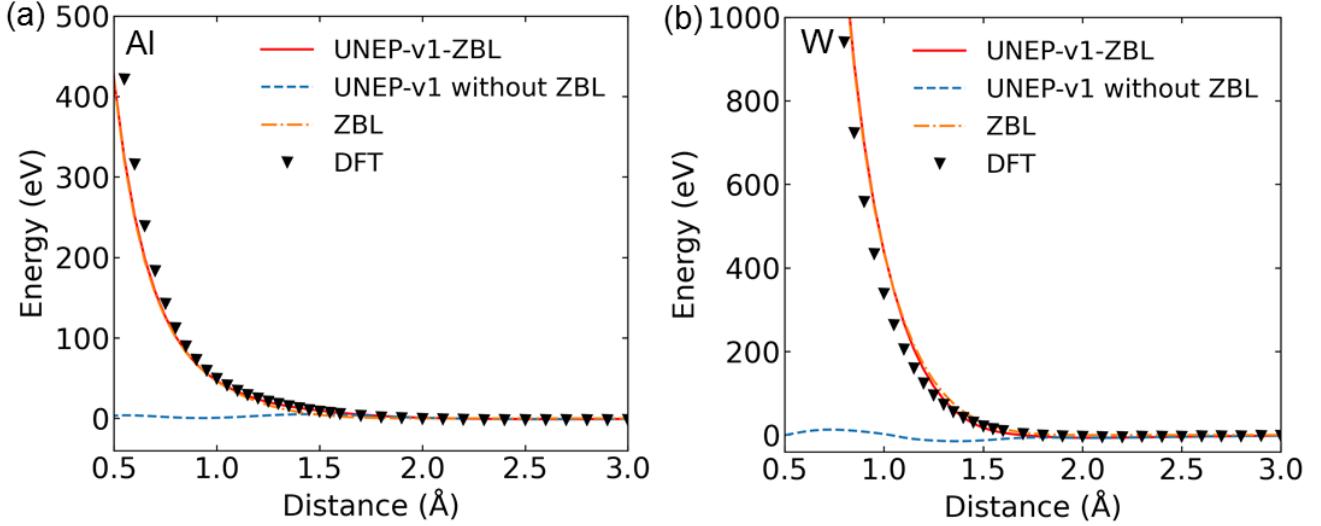


Figure S23. Dimer energies as a function of atom distance for (a) Al and (b) W calculated using combined UNEP-v1 and ZBL (UNEP-v1-ZBL), UNEP-v1 without ZBL, ZBL, and DFT.

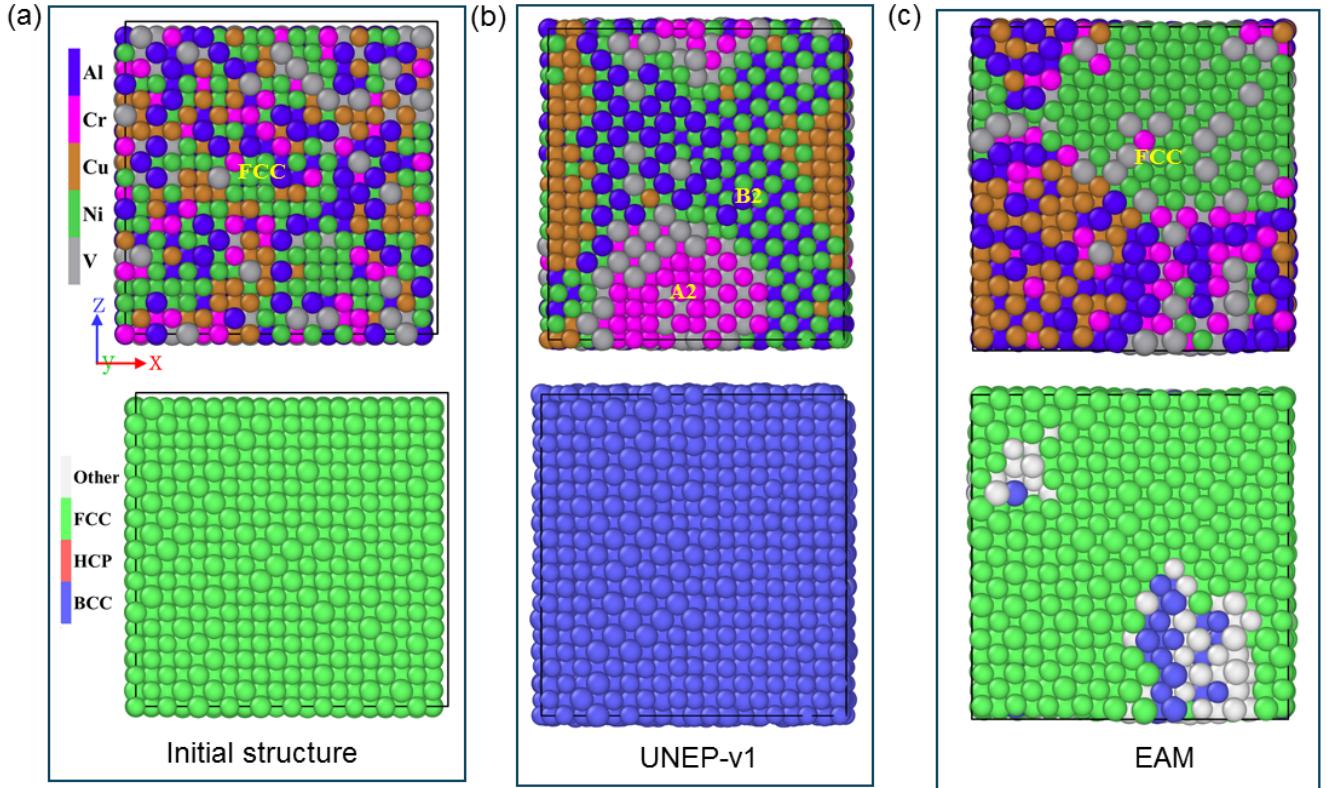


Figure S24. (a) Initial FCC structure of Al_{0.20}Cr_{0.12}Cu_{0.19}Ni_{0.35}V_{0.14}. (b)-(c) Snapshots of the final equilibrium structures from MCMD simulations using UNEP-v1 and EAM [35]. Upper panels show the atomistic structures, while lower panels display the corresponding common neighbor analysis. UNEP-v1 successfully produces both disordered (A2) and ordered (B2) BCC structures in full agreement with experiments [64]. In contrast, EAM potential by Zhou *et al.* [35] keeps the system mostly in FCC structure, unable to reproduce the experimentally expected BCC structure.

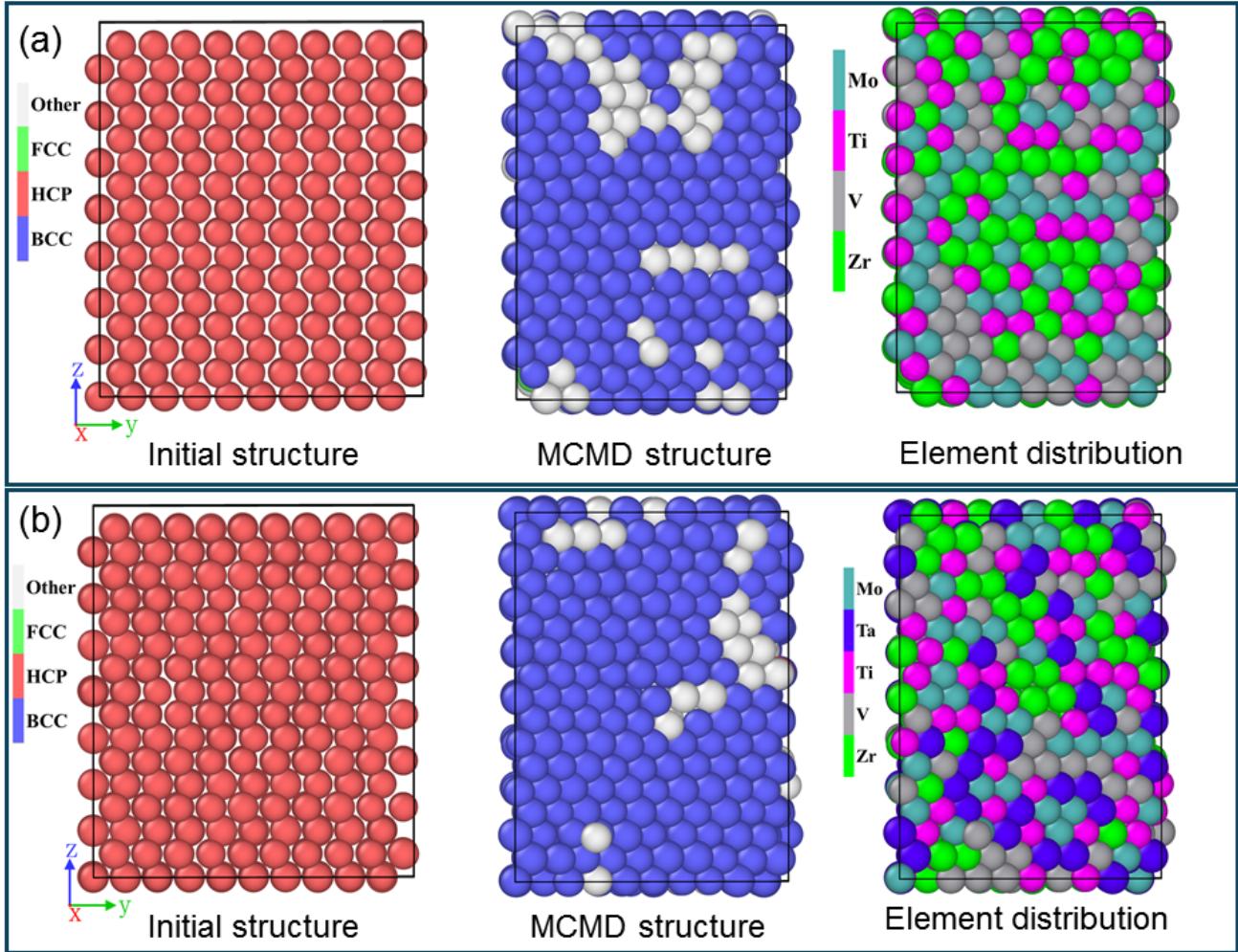


Figure S25. Upper panels demonstrate the transformation of the equimolar TiZrVMo alloy from the initial HCP structure to the BCC structure during MCMD simulations using UNEP-v1 model. Lower panels show a similar transformation for equimolar TiZrVMoTa alloy from the initial HCP structure to the BCC structure during MCMD simulations using UNEP-v1 model. These transformations and the stable phases are in accordance with experimental observations [68], suggesting that sampling binary alloys can correctly capture phase transitions occurring in multi-component alloys.

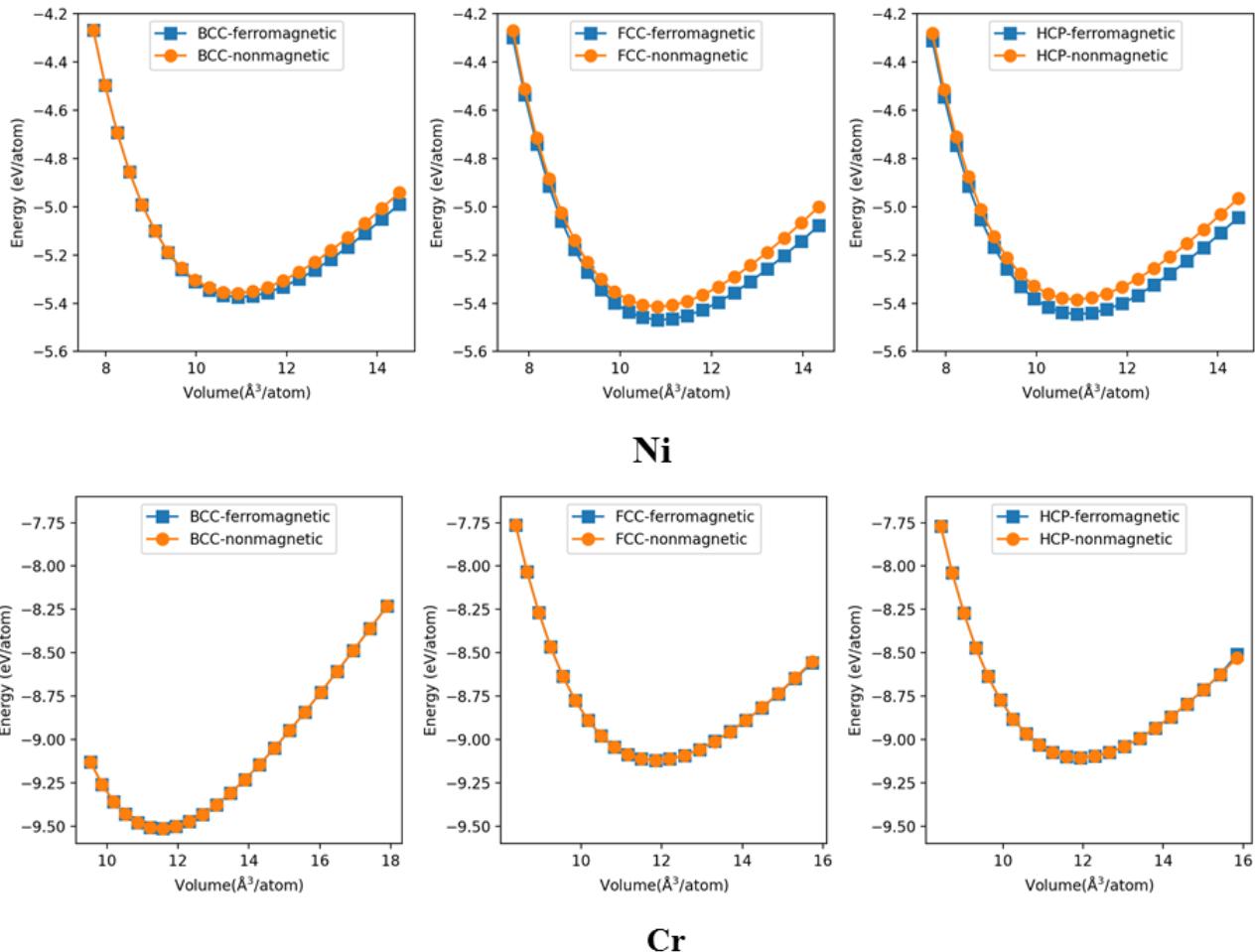


Figure S26. Comparison of the energies for Ni and Cr in BCC, FCC, and HCP structures from DFT calculations with and without considering magnetism.

SUPPLEMENTARY TABLES

Table S1. Elastic constant (in units of GPa).

Species	Component	DFT [74]	EAM	UNEP-v1	Element	Component	DFT [74]	EAM	UNEP-v1
Ag	C11	100	124.7	109.6	Al	C11	104	106.5	121.3
	C12	82	93.4	81.7		C12	73	59.5	54
	C44	41	46.1	41.9		C44	32	27.9	39
Au	C11	144	186.7	154.8	Cu	C11	180	170.0	175.4
	C12	134	156.9	118.5		C12	127	121.0	128.3
	C44	29	42.1	32.7		C44	78	75.5	78.8
Ni	C11	276	245.6	273	Pb	C11	47	49.6	49
	C12	159	147.0	179.2		C12	32	42.1	36.4
	C44	132	124.2	112.3		C44	18	15	15
Pd	C11	187	233.9	190.6	Pt	C11	303	346.4	305.2
	C12	147	175.0	149.7		C12	220	249.1	227.2
	C44	71	71.1	71.8		C44	54	75.8	71.2
Cr	C11	499	304.6	577.4	Mo	C11	472	455.2	460.9
	C12	139	135.8	122.9		C12	158	167.3	174.2
	C44	102	122.8	84.4		C44	106	113.2	88
Ta	C11	265	262.0	261.1	V	C11	276	231.6	333.5
	C12	158	158.2	165.6		C12	131	119.8	123.4
	C44	69	82.4	47		C44	16	45.9	18.4
W	C11	510	520.9	505	Ti	C11	177	135.3	173
	C12	201	205.0	208.2		C12	83	94.5	93.7
	C44	143	160.9	121.5		C13	76	68.5	68.6
Mg	C11	58	54.1	63	Ti	C33	191	203	207.3
	C12	30	30.0	28.3		C44	42	30.3	40.6
	C13	22	20.5	20.3		C66	47	20.4	39.6
Zr	C33	66	67.7	72.5		C11	144	118.3	141.4
	C44	20	13.3	17		C12	65	88.9	81.4
	C66	14	12.0	17.3		C13	67	67.0	62.2
	C11	144	118.3	141.4		C33	162	181.1	137.4
	C12	65	88.9	81.4		C44	26	24.5	21.3
	C13	67	67.0	62.2		C66	40	14.7	30

Table S2. Mono-vacancy formation energy (in units of eV).

Species	DFT	EAM	UNEP-v1
Ag	0.96	1.10	0.80
Al	0.55	0.65	0.61
Au	0.51	0.99	0.57
Cu	1.12	1.28	1.03
Ni	1.41	1.70	1.42
Pb	0.39	0.59	0.44
Pd	1.23	1.55	1.20
Pt	0.68	1.54	1.03
Cr	2.54	2.06	3.04
Mo	2.77	2.95	3.07
Ta	2.82	2.97	2.79
V	2.28	2.14	2.35
W	3.34	3.58	3.73
Mg	0.85	0.65	0.80
Ti	2.00	1.63	2.40
Zr	2.00	1.84	2.20

Table S3. Surface formation energy (in units of J/m²).

Species Surface	DFT [75]			EAM			UNEP-v1		
	100	110	111	100	110	111	100	110	111
Ag	0.82	0.87	0.76	0.99	1.11	0.91	0.83	0.89	0.73
Al	0.91	0.98	0.77	0.93	1.04	0.91	0.93	0.98	0.86
Au	0.86	0.91	0.71	1.01	1.11	0.90	0.88	0.89	0.72
Cu	1.47	1.56	1.34	1.58	1.77	1.51	1.49	1.58	1.32
Ni	2.21	2.29	1.92	1.90	2.08	1.79	2.23	2.34	1.98
Pb	0.33	0.33	0.26	0.39	0.43	0.35	0.29	0.32	0.26
Pd	1.52	1.57	1.36	1.64	1.81	1.52	1.50	1.59	1.34
Pt	1.86	1.87	1.49	2.19	2.51	2.08	1.77	1.78	1.43
Cr	3.63	3.22	3.44	1.86	1.69	2.11	3.63	3.11	3.58
Mo	3.18	2.78	2.96	2.48	2.16	2.77	3.13	2.70	3.13
Ta	2.47	2.34	2.70	2.35	2.00	2.60	2.77	2.40	2.84
V	2.38	2.41	2.70	1.94	1.66	2.15	2.78	2.46	2.79
W	3.95	3.23	3.47	2.99	2.57	1.83	3.89	3.19	3.72
Surface	0001	1010	1011	0001	1010	1011	0001	1010	1011
Mg	0.51	0.60	0.63	0.38	0.40	0.40	0.61	0.64	0.65
Ti	2.15	2.22	2.25	1.83	1.89	1.64	2.30	2.26	2.37
Zr	1.60	1.66	1.57	1.26	1.35	1.34	1.42	1.52	1.50

Table S4. Melting temperature (in units of K).

Species	EAM	UNEP-v1	Exp. [79]
Ag	1135	970	1235
Al	553	840	933
Au	1115	900	1337
Cu	1161	1220	1358
Ni	1500	1670	1728
Pb	611	550	601
Pd	1553	1510	1828
Pt	1422	1610	2041
Cr	2128	2520	2180
Mo	3370	2870	2896
Ta	2950	3150	3290
V	1664	2240	2183
W	4221	3710	3695
Mg	682	860	923
Ti	1578	1610	1941
Zr	1752	1860	2128