



# Crystal Structure Prediction of Binary Alloys via Deep Potential

Haidi Wang<sup>1</sup>, Yuzhi Zhang<sup>2,3</sup>, Linfeng Zhang<sup>4\*</sup> and Han Wang<sup>5\*</sup>

<sup>1</sup> School of Electronic Science and Applied Physics, Hefei University of Technology, Hefei, China, <sup>2</sup> Yuanpei College of Peking University, Beijing, China, <sup>3</sup> Beijing Institute of Big Data Research, Peking University, Beijing, China, <sup>4</sup> Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ, United States, <sup>5</sup> Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Beijing, China

## OPEN ACCESS

### Edited by:

Wei Hu,  
Lawrence Berkeley National  
Laboratory, United States

### Reviewed by:

Honghui Shang,  
Chinese Academy of Sciences (CAS),  
China  
Zp Hu,  
Nankai University, China

### \*Correspondence:

Linfeng Zhang  
linfengz@princeton.edu  
Han Wang  
wang\_han@iapcm.ac.cn

### Specialty section:

This article was submitted to  
Theoretical and Computational  
Chemistry,  
a section of the journal  
Frontiers in Chemistry

Received: 31 July 2020

Accepted: 28 August 2020

Published: 26 November 2020

### Citation:

Wang H, Zhang Y, Zhang L and  
Wang H (2020) Crystal Structure  
Prediction of Binary Alloys via Deep  
Potential. *Front. Chem.* 8:589795.  
doi: 10.3389/fchem.2020.589795

Predicting crystal structure has been a challenging problem in physics and materials science for a long time. A reliable energy calculation engine combined with an efficient global search algorithm, such as particle swarm optimization algorithm or genetic algorithm, is needed to conduct crystal structure prediction. In recent years, machine learning-based interatomic potential energy surface models have been proposed, potentially allowing us to perform crystal structure prediction for systems with the accuracy of density functional theory (DFT) and the speed of empirical force fields. In this paper, we employ a previously developed Deep Potential model to predict the intermetallic compound of the aluminum–magnesium system, and find six meta-stable phases with negative or nearly zero formation energy. In particular, Mg<sub>12</sub>Al<sub>8</sub> shows excellent ductility and Mg<sub>5</sub>Al<sub>27</sub> has a high Young's modulus. Based on our benchmark results, we propose a relatively robust structure screening criterion that selects potentially stable structures from the Deep Potential-based convex hull and performs DFT refinement. By using this criterion, the computational cost needed to construct the convex hull with *ab initio* accuracy can be dramatically reduced.

**Keywords:** many-body potential energy, deep learning, crystal structure prediction, Al-Mg, alloy

## INTRODUCTION

In recent years, crystal structure prediction has played an increasingly important role, not only for understanding the ground-state structure of matter, but also for designing materials and drug molecules with desired functionality (Oganov, 2018; He et al., 2019; Zhao et al., 2019; Xie et al., 2020). Generally speaking, a ground-state crystal structure prediction method involves three components: a model that generates the interatomic potential energy surface (PES) and forces, a sampling technique for exploring different conformations in the configuration space, and a relaxation procedure to find the local minima on the PES (Podryabinkin et al., 2019). While the relaxation procedure is relatively standard, different sampling techniques have been championed by different software packages. For instance, the genetic evolutionary algorithm, the particle swarm algorithm, and the firefly algorithm have been widely used in the USPEX software (Glass et al., 2006), the CALYPSO package (Wang et al., 2010), and the PyChemia library (Avenidaño-Franco and Romero, 2016), respectively. Due to the high accuracy required by both sampling and relaxation, density functional theory (DFT) (Kohn and Sham, 1965) is typically used for generating the PES. Despite its widespread success, DFT has a high computational cost that typically scales cubically with the system size, which, to some extent, hinders routine applications to large and complex systems.

Many empirical PES models for popular solid-state systems have been proposed (Mendeleev et al., 2009; Jelinek et al., 2012; Dickel et al., 2018) to address the efficiency issue of DFT. Due to the relatively simple and analytical expressions adopted by these empirical models, an acceleration of many orders of magnitude in terms of computational cost can be gained, but presumably at the price of accuracy and transferability. As such, trial-and-error processes are typically required for developing such models, yet challenges have remained for systems involving multiple elements, complex and exotic phases, or bond breaking and formation events.

In recent years, a few machine learning (ML) techniques have been proposed for representing the PES (Behler and Parrinello, 2007; Bartók et al., 2010; Artrith and Urban, 2016; Khorshidi and Peterson, 2016; Shapeev, 2016; Han et al., 2018; Zhang et al., 2018a,b). Unlike typically empirical PES models, representations coming from ML tasks, such as kernel functions and neural networks (NNs), have shown great promise to fit high-dimensional functions. When trained on a suitably generated dataset of atomic configurations and corresponding potential energies and forces, a good ML-based PES model can be used with an accuracy of the reference DFT model, and an efficiency comparable to that of empirical PES models. Not surprisingly, ML-based PES models have been employed in recent work for structure search tasks. For instance, boron has been studied by several groups: Podryabinkin et al. (2019) adopted the moment tensor potential (Shapeev, 2016) and the USPEX evolutionary algorithm; Huang et al. (2018) used the Behler–Parrinello potential (Behler and Parrinello, 2007) and the stochastic surface walking global optimization method; Tong et al. (2018) used the Gaussian Approximation Potential (Bartók et al., 2010) and the CALYPSO approach.

In this work, we target at using ML-based PES models for crystal structure prediction of alloys. We adopt the smooth version of the Deep Potential (DP) model (Zhang et al., 2018b), which employs NN architectures to parameterize two networks, the embedding network that defines a list of symmetry-preserving descriptors, and the fitting network that maps these descriptors to local energy contributions. The versatile architecture of DP makes it particularly suitable for multicomponent systems and those involving bond breaking and formation, for which most methodologies are hard to handle. The aluminum–magnesium (Al–Mg) binary alloy system is selected as an example based on the following reasons: First, Al–Mg binary alloys are important in real-life applications. They are widely used in automotive, aerospace, and electronic device industries (Gupta and Ling, 2011) due to their lightweight nature and excellent mechanical properties. However, only a limited number of intermetallic compounds of the Al–Mg binary system have been documented in well-known databases, such as the American Society for Metals (ASM) Alloy Phase Diagram Database<sup>1</sup>, the Inorganic Crystal Structure Database (ICSD)<sup>2</sup>, the Open Quantum Materials Database (OQMD)<sup>3</sup>,

and the Material Project database (MP)<sup>4</sup>. Second, our previous study has established an Al–Mg DP model (Zhang et al., 2019), which describes well the basic physicochemical properties and has been carefully tested. As such, this DP model can be readily used for crystal structure prediction and can be download online<sup>5</sup>.

Combining the particle swarm optimization (PSO) method and the DP model, potential intermetallic compounds of the Al–Mg system are systematically explored. Compared with a previous study (Zhuang et al., 2017), which only explored the Mg-rich phases, our simulation covers a much wider concentration range. Six new Al–Mg intermetallic compounds (Mg<sub>12</sub>Al<sub>8</sub>, Mg<sub>7</sub>Al<sub>9</sub>, Mg<sub>14</sub>Al<sub>18</sub>, Mg<sub>6</sub>Al<sub>10</sub>, Mg<sub>8</sub>Al<sub>16</sub>, and Mg<sub>5</sub>Al<sub>27</sub>) are found to be meta-stable. The mechanical properties of these new compounds are further investigated. To facilitate future investigations of more complicated tasks, special attention is given to the whole simulation protocol and the selection criterion for further DFT validations. Direct comparisons with popular empirical PES models and DFT show the advantage of DP in terms of both accuracy and efficiency.

## COMPUTATIONAL METHODS

We adopt the PSO method, as implemented in the CALYPSO package (Wang et al., 2010), to search potentially stable and meta-stable Mg–Al intermetallic structures. PSO is inspired by the choreography of a bird flock and can be seen as a distributed behavior algorithm that performs multidimensional search. In the CALYPSO package, there are three steps for a global structure prediction task. First, a group of structures called population is generated randomly with symmetric constraints to allow a diverse sampling of the PES. The number of structures employed here is defined by a parameter called population size (`PopSize`). Second, a local relaxation of the population is performed based on a PES engine, which is typically a DFT model, and here we replace it with a DP model. A procedure that eliminates similar structures by using the so-called bond characterization matrix is followed up to enhance the search efficiency. Third, a certain number of new structures (the best 60% of the population size) are generated by PSO. Within the PSO scheme, a velocity vector associated with each structure is updated using the information of the previously proposed and optimized structure, as well as the globally best structure, that is, the structure with the lowest enthalpy, at the current generation. The new structures are generated based on the current structures and the velocity vectors. The last two steps continue iteratively until the predefined largest number of generations (`GenNumb`) is reached. The parameter `GenNumb` is typically selected to be large enough so that the structure with the lowest energy can sustain for several generations. Generally speaking, the more atoms (`Natom`) in a structure, the larger `PopSize` and `GenNumb` are required. We refer to Yanchao Wang and Ma (2012) for more details of the CALYPSO

<sup>1</sup>ASM <https://www.asminternational.org/phase-diagrams>

<sup>2</sup>ICSD <https://icsd.fiz-karlsruhe.de/search/basic.xhtml>

<sup>3</sup>OQMD <http://www.oqmd.org>

<sup>4</sup>MP <https://www.materialsproject.org>

<sup>5</sup>deepmd <http://www.deepmd.org>

code. We set the CALYPSO parameters according to the following criteria:

$$\begin{cases} \text{PopSize} = 30, & \text{GenNumb} = 60; & \text{if } N_{\text{atom}} \leq 10, \\ \text{PopSize} = 40, & \text{GenNumb} = 80; & \text{if } 10 < N_{\text{atom}} \leq 20, \\ \text{PopSize} = 50, & \text{GenNumb} = 100; & \text{if } 20 < N_{\text{atom}} \leq 32. \end{cases} \quad (1)$$

The DP model (Zhang et al., 2018b) used here employs NN functions to represent the PES. In short, the total energy of a system is described as the sum of atomic energies,

$$E = \sum_{i=1}^N \epsilon_i, \quad (2)$$

where  $\epsilon_i$  is the  $i$ th atomic energy. The atomic energy is represented as

$$\epsilon_i = N_{\omega(i)}(\{R_{ij}\}_{j \in \mathcal{N}(i)}), \quad (3)$$

where  $N_{\omega(i)}$  is called a sub-network that computes the atomic contribution to the total energy, and  $\omega(i)$ , which depends on the chemical species of atom  $i$ , denotes the weights used to parameterize the sub-network. The neighbors of atom  $i$  within the cut-off radius  $R_c$  are denoted by  $\mathcal{N}(i)$ .  $R_{ij}$  is the position of atom  $j$  relative to  $i$  used to describe the local environment of atom  $i$ . To generate uniformly accurate DP models in a way that minimizes both human intervention and the computational cost for data generation and model training, a concurrent learning strategy called the Deep Potential GENERator (DP-GEN) (Zhang et al., 2019) has been proposed. In this strategy, an initial dataset (random Al-Mg alloy structures) labeled by DFT calculations is used to train an ensemble of DP models, and molecular dynamics is driven by one of the DP models to sample the configuration space. An error indicator serves to select a small fraction out of the new samples as candidates, which are labeled with *ab initio* energies and forces and added to the database. Such iterations are repeated until the configuration space has been explored sufficiently, and a decent DP model has been obtained with high accuracy and transferability. The training is performed using the DeePMD-kit package (Wang et al., 2018) and the concurrent learning strategy is realized by the DP-GEN software package (Zhang et al., 2020). In details, the DeepPot-SE model is used with a cutoff radius of 8.0 Å. The size of the embedding and fitting NNs are  $25 \times 50 \times 100$  and  $240 \times 240 \times 240$ , respectively. During the training, the learning rate decreases exponentially with respect to the starting value of 0.0005. The decay rate and decay step are set to 0.95 and 128,000, respectively. In addition, the prefactors of loss functions are set to  $p_e^{\text{start}} = 0.02$ ,  $p_e^{\text{limit}} = 2$ ,  $p_f^{\text{start}} = 1000$ ,  $p_f^{\text{limit}} = 1$ ,  $p_v^{\text{start}} = 0.0$ ,  $p_v^{\text{limit}} = 0.0$ . Both DeePMD-kit and DP-GEN are publicly available online<sup>6</sup>. For more details, we refer the reader to Wang et al. (2018) and Zhang et al. (2019, 2020).

All DFT calculations are carried out with the Vienna Ab-Initio Simulation Package (VASP, version 5.4.4) (Kresse and

Furthmüller, 1996). The generalized gradient approximation within the Perdew–Burke–Ernzerhof (Perdew et al., 1996) (PBE) functional is used to model the exchange–correlation energy. The plane wave basis sets with kinetic energy cutoff of 600 eV are used to expand the valence electron wave functions. For all structural relaxations, the convergence criterion for the energy in electronic SCF iterations and the Hellmann–Feynman forces in ionic step iterations are set to  $1.0 \times 10^{-6}$  eV and  $1.0 \times 10^{-2}$  eV/Å, respectively. The Brillouin zone is represented by Monkhorst–Pack (Pack and Monkhorst, 1977) special k-point mesh with a grid spacing of  $0.08 \text{ \AA}^{-1}$ . The phonon spectra are obtained based on finite-difference method as implemented in the Atomic Simulation Environment (ASE) (Bahn and Jacobsen, 2002; Larsen et al., 2017) software, where the forces are calculated by the python interface of DeePMD-kit. To calculate the phonon density of states, the q-point mesh is set to  $20 \times 20 \times 20$ . The local structure relaxation is carried out by the LAMMPS package (Plimpton, 1995), and the DP model used here has been reported and extensively tested in Zhang et al. (2019).

All structure data and convex hulls are analyzed by pymatgen software (Ong et al., 2013).

## RESULTS AND DISCUSSIONS

To demonstrate the validity of CALYPSO+DP scheme, we perform some preliminary tests for several different stoichiometric proportions. Here, we take  $\text{Mg}_{12}\text{Al}_8$  as an example to show the evolution of the energies of all the structures (Figure 1A), as well as the lowest energy (Figure 1C), during the CALYPSO structure search process. According to the energy histogram in Figure 1B, it can be found that there are about 458 structures, out of a total number of 3,200 structures, within an energy range  $< 20$  meV/atom (compared with the ground state structure). The evolution of the lowest energy for all generations shows that the global optimization converged quickly. It is remarkable to find that one potentially stable structures can be readily obtained in the first few generations (labeled by red star in Figure 1A). When refined by DFT, it shows that this structure is the ground state structure of the corresponding combination. In the tests of some other stoichiometric proportions, that is,  $\text{Mg}_3\text{Al}_3$ ,  $\text{Mg}_2\text{Al}_2$ ,  $\text{Mg}_1\text{Al}_1$ , and  $\text{Mg}_3\text{Al}$ , the corresponding known structures in the materials project database, that is, mp-1038779, mp-1094987, mp-1038934, and the  $L1_2$  phase (Mendeleev et al., 2009), are found by the CALYPSO+DP scheme, which further confirms the validity of our approach.

Next, we use the DP+CALYPSO scheme to construct the convex hull of the Al-Mg system systematically. We restrict the number of atom in the supercell to  $< 32$  atoms. In other words, we consider the systems  $\text{Mg}_x\text{Al}_y\{x + y \leq 32, x \geq 1, y \geq 1, x, y \in \mathbb{Z}\}$ , which represent 496 combinations, or 323 concentrations, in total. According to these parameter settings, the total number of local relaxations is up to  $2 \times 10^7$ . In the following, we consider three prerequisites that should be satisfied for a stable crystal structure: (i) thermodynamic stability, which is estimated by the formation energy and convex hull; (ii) dynamic stability, which

<sup>6</sup>See <https://github.com/deepmodeling>

can be assessed by phonon dispersions; and (iii) mechanical stability, which is evaluated via elastic constants (Xu et al., 2019).

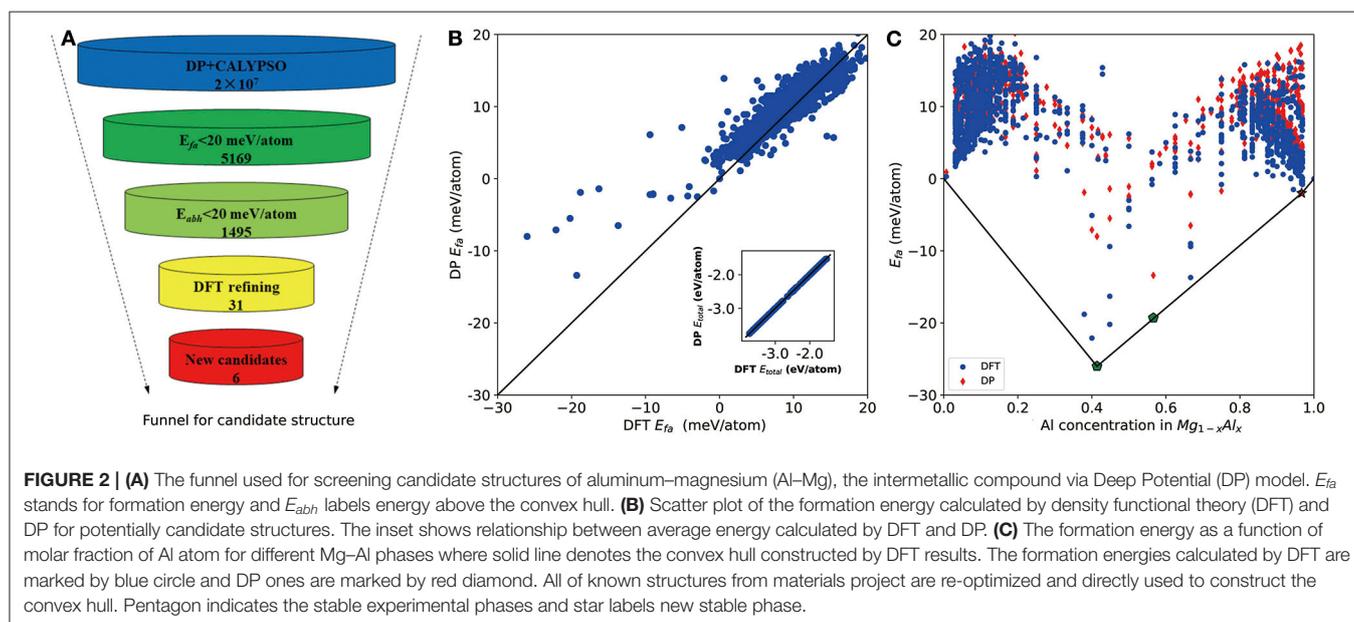
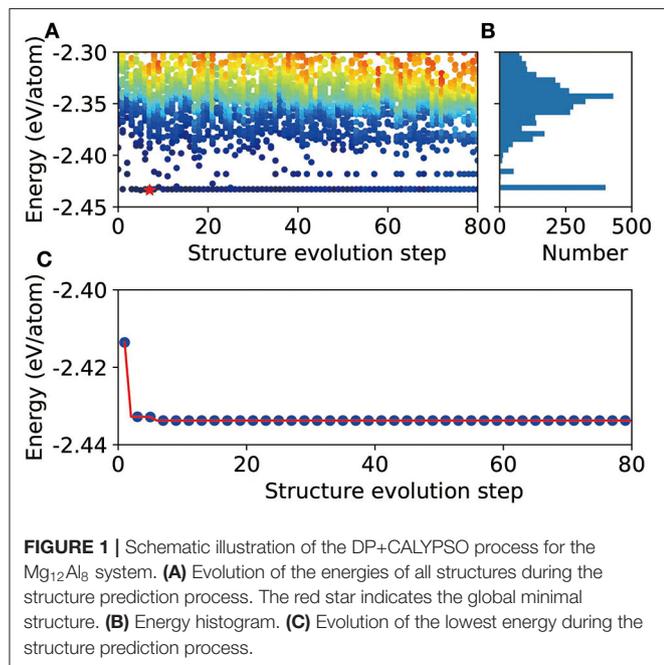
According to the preliminary tests, we notice that although the DP model can generate energies and forces that are very close to the DFT reference model, small intrinsic error still exists. Therefore, if our goal is to calculate properties with the accuracy of the DFT-based PES landscape, an additional refinement step has to be adopted based on structures selected from a DP+CALYPSO process. Two concepts we pay particular attention to are the formation energy ( $E_{fa}$ ) and the energy

above convex hull ( $E_{abh}$ ). The formation energy (Haastруп et al., 2018) of an alloy system is the energy required to produce the system from the most stable crystal structures of the individual components, which is defined as

$$E_{fa} = \frac{E(\text{Mg}_x\text{Al}_y) - xE(\text{Mg}) - yE(\text{Al})}{x + y} \quad (4)$$

where  $E(\text{Mg}_x\text{Al}_y)$  is the total energy of the material  $\text{Mg}_x\text{Al}_y$ , and  $E(\text{Mg})$  and  $E(\text{Al})$  are the average energies of the elements Mg and Al in their stable crystal at 0 K.  $E_{abh}$  measures the energy for a material to decompose into the set of most stable materials with the same chemical composition. A positive  $E_{abh}$  indicates that this material is unstable with respect to such decomposition. A zero  $E_{abh}$  indicates that this is the most stable material at its composition. To accurately determine these properties, we use two criteria for an additional DFT refinement:  $E_{fa}$  being less than 20 meV/atom, and  $E_{abh}$  being <20 meV/atom. We use these two criteria at the same time based on the following reasons. Our goal is to find potential stable or meta-stable structures. Due to the error of the DP model, some structures with positive DP-predicted  $E_{fa}$  may turn negative if we refine it by DFT and vice versa. At the same time, the thermodynamic stability is also controlled by  $E_{abh}$ . If  $E_{abh}$  is too high, this structure will decompose into other phases even if this structure has a negative formation energy. In the tested example, we will show that since in general DP exhibits a  $\sim 2$  meV/atom average error. Compared with DFT, the criteria used here are fairly robust. In contrast, for a previously established empirical model, due to its large intrinsic error, the procedure introduced above is no longer practical, since the number of DFT refinements is so large that little efficiency can be gained.

As shown in Figure 2A, we first use the formation energy  $E_{fa}$  to screen the candidate structures, from which the number of structures is significantly reduced from  $2 \times 10^7$  to 5,169. Based



on these 5,169 DP-optimized structures, a crude convex hull is constructed. Then the criterion for  $E_{abh}$  is introduced to remove the thermodynamically unstable structures, and, finally, 1,495 structures are selected for further DFT refinements. As shown in the inset of **Figure 2B**, for all DFT refined structures, the energies per atom calculated by DP and DFT are in excellent agreement. The largest and root mean square error (RMSE) of the total energy per atom are about 17 and 2 meV/atom, respectively. As for the formation energy in **Figure 2B**, the largest error and RMSE are about 15 and 6 meV/atom, respectively. Among them, the errors of the formation energies of experimentally stable phases (Zhuang et al., 2017)  $Mg_{17}Al_{12}$  and  $Mg_{23}Al_{30}$  (labeled by pentagon in **Figure 2C**) are 15 and 6 meV/atom, which confirms the validity of our 20 meV/atom selection criteria.

The convex hull based on DFT results is then constructed and presented in **Figure 2C**, including two experimentally stable structures  $Mg_{17}Al_{12}$  and  $Mg_{23}Al_{30}$  labeled by green pentagon. One new stable structure with a formula of  $MgAl_{29}$  is discovered and denoted by red star. Based on the DFT-refined convex hull, we look for new structures that are potentially synthesizable by experiments. We use the following criteria:  $E_{abh} < 20$  meV/atom

and  $E_{fa} < 1$  meV/atom, where  $E_{abh}$  and  $E_{fa}$  are DFT-calculated values, and obtain 31 potentially candidates, including 1 stable structure and 30 meta-stable structures. However, we may note that most of those newly proposed meta-stable structures locate at the boundary region of the concentration range. That is to say, most of these structures have very low concentration of Al or Mg. Compared with these phases, the phases in the middle region are of more interesting, from which we propose six new meta-stable structures, namely  $Mg_{12}Al_8$ ,  $Mg_7Al_9$ ,  $Mg_{14}Al_{18}$ ,  $Mg_6Al_{10}$ ,  $Mg_8Al_{16}$ , and  $Mg_5Al_{27}$ . The corresponding side view of the crystal structures are shown in **Figures 3A–F** and the geometric structure parameters are listed in **Table 1**.

As listed in **Table 1**, the meta-stable structures can be divided into two groups according to their lattice types.  $Mg_7Al_9$ ,  $Mg_{14}Al_{18}$ , and  $Mg_6Al_{10}$  have a tetragonal lattice, whereas  $Mg_{12}Al_8$ ,  $Mg_8Al_{16}$ , and  $Mg_5Al_{27}$  have a cubic lattice. Moreover, all structures have nearly zero or negative formation and small energy above convex hull, which indicates that these structures may be synthesizable by experiments in future.

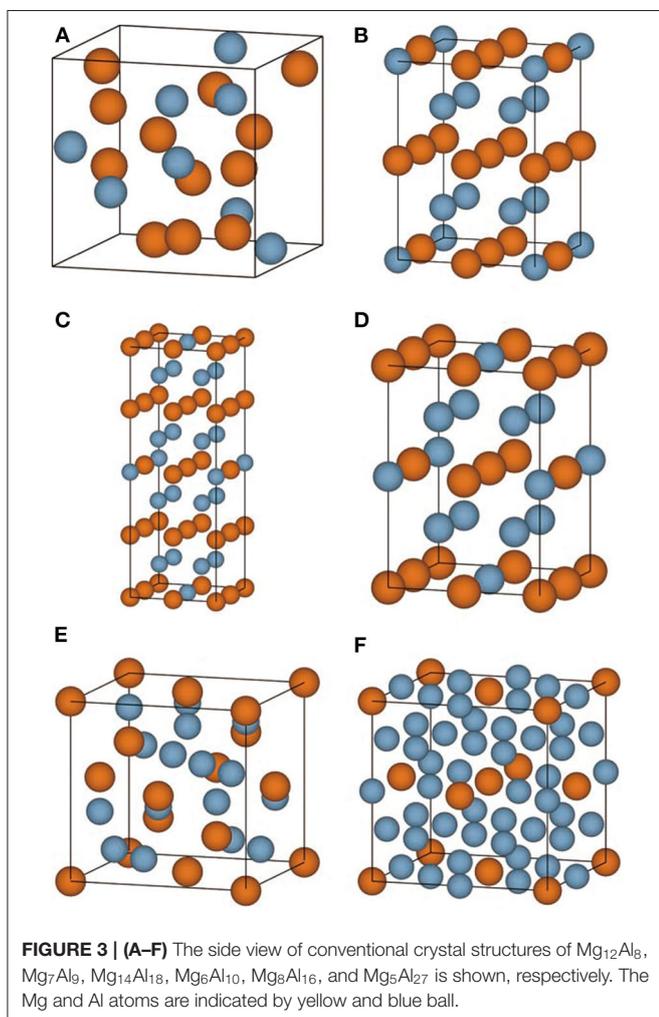
Given the encouraging stability metrics above, we proceed to study the dynamic and mechanical stability of these newly proposed intermetallic compounds via DP model. As shown in **Figure 4**, the phonon structures show no imaginary frequency, which indicates that all of those intermetallic compounds are dynamically stable. As for the mechanical aspect, the elastic stability conditions (Mouhat and Coudert, 2014) for cubic crystals are given as:

$$C_{11} - C_{12} > 0, C_{11} + 2C_{12} > 0, C_{44} > 0, \quad (5)$$

and those for tetragonal crystals are given as:

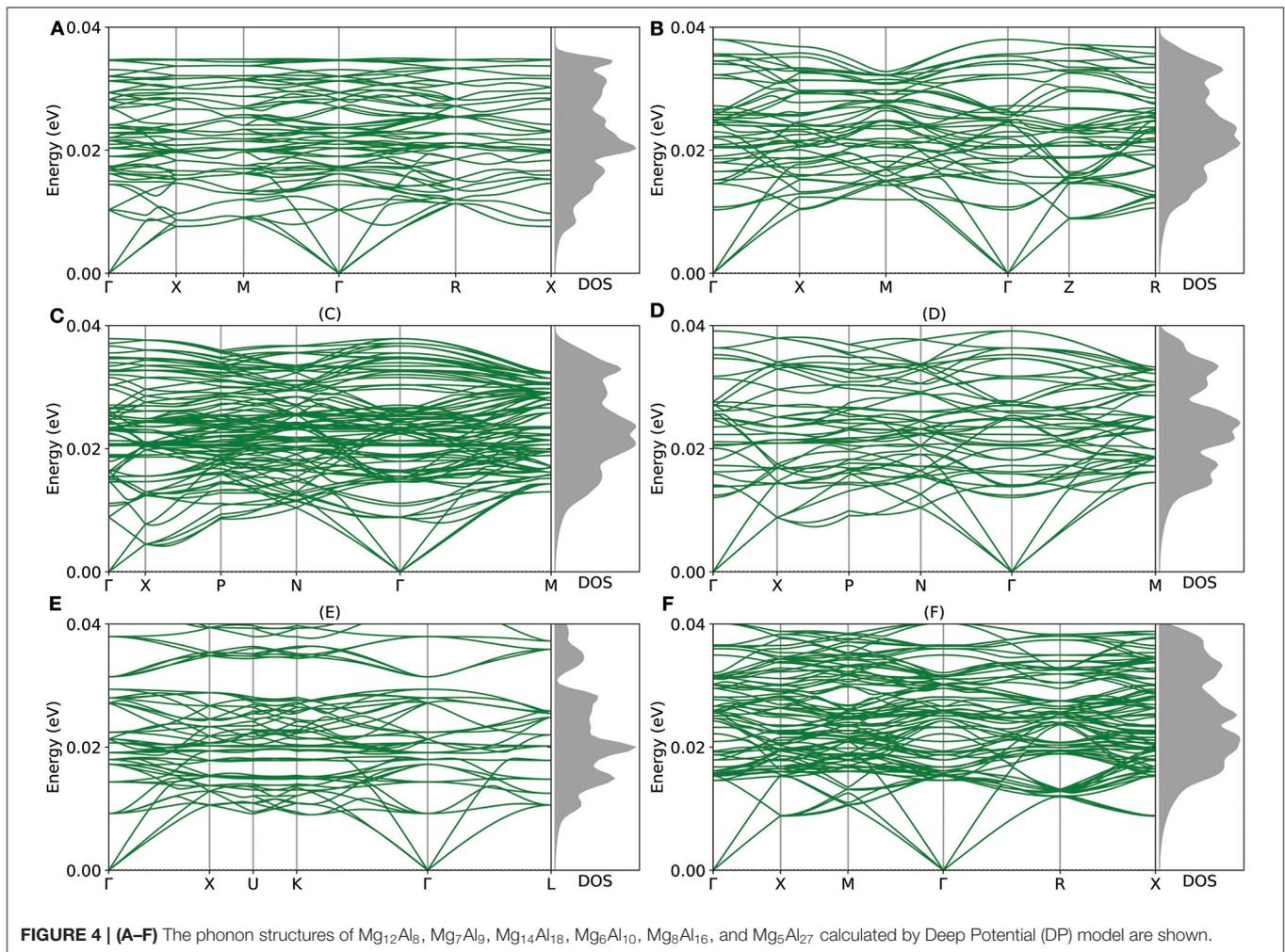
$$C_{11} > |C_{12}|, 2C_{13}^2 < C_{33}(C_{11} + 2C_{12}), C_{44} > 0, C_{66} > 0. \quad (6)$$

According to **Table 2**, both groups of structures meet the elastic stability conditions, which indicates that these 6 intermetallic compounds are mechanically stable. Further, we use the Pugh's ratio  $B_v/G_v$  to assess the expected average ductility (Pugh, 1954). According to Pugh, a larger  $B_v/G_v$  value implies a better ductility property. As shown in **Table 2**, the Pugh's ratio of both  $Mg_8Al_{16}$  and  $Mg_{12}Al_8$  are larger than that of hcp Mg (2.08). In particular, the Pugh's ratio of  $Mg_{12}Al_8$  (2.42) is comparable to that of Al (2.47), so it may have excellent ductility. Considering its



**TABLE 1 |** Lattice parameters  $a(\text{\AA})$ ,  $b(\text{\AA})$ ,  $c(\text{\AA})$ , density  $\rho$  ( $\text{g/cm}^3$ ), space group symbol  $S_m$ , lattice type, formation energy  $E_{fa}$  (meV/atom), and energy above the convex hull  $E_{abh}$  (meV/atom) of  $Mg_{12}Al_8$ ,  $Mg_7Al_9$ ,  $Mg_{14}Al_{18}$ ,  $Mg_6Al_{10}$ ,  $Mg_8Al_{16}$ , and  $Mg_5Al_{27}$  calculated by DFT.

Formula	$N_{atom}$	$a$	$b$	$c$	$\rho$	$S_m$	Lattice type	$E_{fa}$	$E_{ah}$
$Mg_{12}Al_8$	20	7.38	7.38	7.38	2.10	$P4_{332}$	Cubic	-22.05	3.09
$Mg_7Al_9$	16	5.98	5.98	8.44	2.27	$P4/mmm$	Tetragonal	-0.82	18.74
$Mg_{14}Al_{18}$	32	5.98	5.98	16.92	2.27	$I4/mmm$	Tetragonal	-0.10	19.45
$Mg_6Al_{10}$	16	5.93	5.93	8.43	2.33	$I4/mmm$	Tetragonal	0.87	17.71
$Mg_8Al_{16}$	24	7.67	7.67	7.67	2.30	$Fd-3m$	Cubic	-13.73	1.30
$Mg_5Al_{27}$	32	8.21	8.21	8.21	2.55	$Pm-3m$	Cubic	0.60	7.95



**TABLE 2 |** Elastic constants, bulk modulus  $B_V$  (GPa), shear modulus  $G_V$  (GPa), Young's modulus  $E_V$  (GPa), Pugh's ratio ( $B_V/G_V$ ), and Poisson's ratio  $\nu$  of  $Mg_{12}Al_8$ ,  $Mg_7Al_9$ ,  $Mg_{14}Al_{18}$ ,  $Mg_6Al_{10}$ ,  $Mg_8Al_{16}$  and  $Mg_5Al_{27}$  calculated by the DP model.

Formula	$C_{11}$	$C_{12}$	$C_{13}$	$C_{33}$	$C_{44}$	$C_{66}$	$B_V$	$G_V$	$B_V/G_V$	$E_V$	$\nu$
$Mg_9^{a,1}$	59.57	26.54	21.05	72.94	15.08	16.12	36.53	17.59	2.08	45.48	0.29
$Mg_{12}Al_8$	71.26	35.75	36.50	73.15	21.19	21.38	48.14	19.89	2.42	52.45	0.32
$Mg_7Al_9$	94.75	30.55	39.20	84.63	32.91	22.00	53.81	28.08	1.92	71.77	0.28
$Mg_{14}Al_{18}$	89.19	31.60	40.27	85.57	32.16	23.48	54.57	27.49	1.99	70.60	0.28
$Mg_6Al_{10}$	101.77	36.06	42.59	92.14	31.37	23.23	59.65	29.05	2.05	74.97	0.29
$Mg_8Al_{16}$	97.47	49.80	48.72	95.07	30.68	31.73	65.06	28.35	2.29	74.27	0.31
$Mg_5Al_{27}$	95.40	39.17	39.87	104.71	37.48	40.21	58.91	35.09	1.68	87.84	0.25
$Al_4^{a,2}$	117.64	63.34	58.20	108.46	32.55	40.44	78.15	31.65	2.47	83.66	0.32

The subscript  $v$  denotes the Voigt expressions. The same properties of Mg and Al element are also calculated for comparison (here, the minor inconsistency of elastic constants comes from inherent error of DP model and computation error).

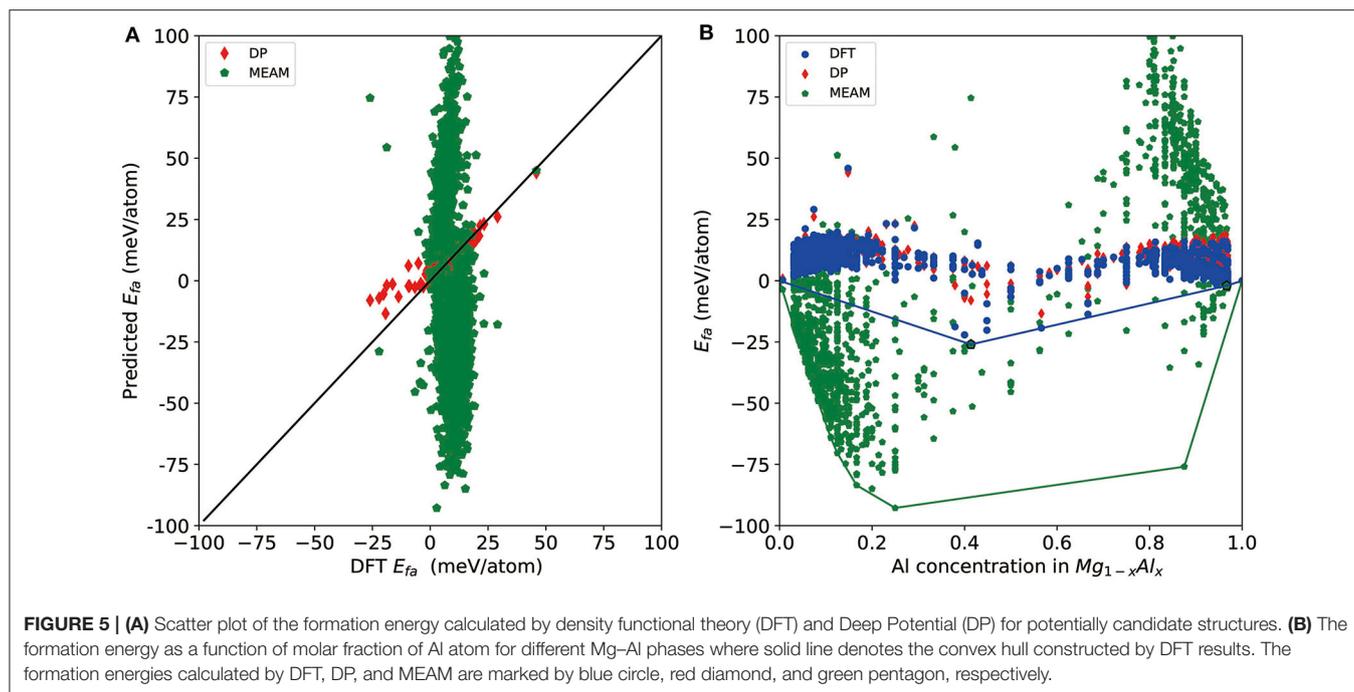
<sup>a</sup>All values are calculated by authors based on DP model.

<sup>1</sup>Zhuang et al. (2017) give DFT values of  $C_{11} = 66$  GPa,  $C_{12} = 25$  GPa,  $C_{13} = 19$  GPa,  $C_{33} = 70$  GPa,  $C_{44} = 20$  GPa,  $B_V = 37$  GPa,  $G_V = 21$  GPa,  $B_V/G_V = 1.76$  GPa, and  $E_V = 54$  GPa.

<sup>2</sup>Zhang et al. (2019) give DFT values of  $C_{11} = 111.2$  GPa,  $C_{12} = 61.4$  GPa,  $C_{44} = 36.8$  GPa,  $B_V = 78.0$  GPa, and  $G_V = 32.1$  GPa.

lower density ( $2.10 \text{ g/cm}^3$ ) compared with Al ( $2.72 \text{ g/cm}^3$ ), this intermetallic compound may have potential applications in automotive, aerospace, electronic, and device industries if it can be synthesized. In addition, the  $Mg_5Al_{27}$  has a higher Young's modulus among all these structures, which indicates that this material may be applied to manufacture high strength devices.

Finally, we test the accuracy of a recent version of the MEAM potential (Jelinek et al., 2012) and see whether it can be used for a similar task or not. For a direct comparison, we test it on all DFT-refined structures and report the results in Figure 5. As shown by the red diamonds and green pentagons in Figure 5A, MEAM exhibits much larger errors compared with DP for most of the structures, and there are MEAM predictions outside the range of the plot ( $\pm 100 \text{ meV/atom}$ ) due to large errors. MEAM results show a large spread on the convex hull plot constructed by DFT results (Figure 5B). For the tested structures, the largest error of  $E_{fa}$  is  $204 \text{ meV/atom}$ , and the RMSE is  $44 \text{ meV/atom}$ . Moreover, the largest error of per-atom total energy is  $465 \text{ meV/atom}$ , and the mean error is  $236 \text{ meV/atom}$ . As such, if we use MEAM+CALYPSO to do a screening of the structures, the selection criteria for further



DFT refinement would be  $E_{fa}$  and  $E_{abh}$  larger than at least 200 meV/atom. As a rough estimation, according to these criteria,  $\sim 1 \times 10^6$  structures will need to be refined by DFT, which is not computationally feasible at all. Above all, we conclude that although MEAM is very efficient, it cannot be used to improve the efficiency of constructing a convex hull at the level of DFT accuracy.

To demonstrate the efficiency of our DP-based procedure, we use two groups of structures to compare the time performance. One group is composed of several  $Mg_{31}Al$  structures and the other group consists of  $MgAl_{31}$  structures. The test results shows that, compared with DFT relaxation, DP has an average speed-up ratio about 3,700 and 650 for  $Mg_{31}Al$  and  $MgAl_{31}$ , respectively, which indicates DP has better time scaling and can be used for larger systems. All tests are performed on Intel(R) Xeon(R) Gold 6248 CPU @ 2.50 GHz.

## CONCLUSIONS

In this paper, we demonstrate that the DP+CALYPSO scheme is reliable for crystal structure prediction for binary alloy system in a wide concentration range. As a concrete example, we use this scheme to predict potentially stable intermetallic compounds of the Al–Mg binary system. Six new meta-stable Al–Mg intermetallic compounds are successfully predicted, including  $Mg_{12}Al_8$ ,  $Mg_7Al_9$ ,  $Mg_{14}Al_{18}$ ,  $Mg_6Al_{10}$ ,  $Mg_8Al_{16}$ , and  $Mg_5Al_{27}$ . All the meta-stable structures are predicted to have thermodynamic stability, dynamic stability, and mechanical stability. In particular,  $Mg_{12}Al_8$  shows excellent ductility and  $Mg_5Al_{27}$  has high Young's modulus. We remark that the exploration strategy proposed in this work can be combined

with the DP-GEN protocol (Zhang et al., 2019, 2020) to generate more training data and improve the DP potential. Moreover, to serve a larger community, DeePMD-kit can be interfaced with other popular general-purpose crystal structure prediction software such as CALYPSO, USPEX, and Pychemia. However, some disadvantages also exists for current scheme, such as expensive cost for training a model, low interface efficiency with CALYPSO, and so on, which limits its application to search complex multicomponent systems with larger number of atoms. We will leave these problems in our future work.

## DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

HaiW, HanW, and LZ conceived the idea. HaiW performed the calculations and wrote the script. All authors revised the manuscript.

## FUNDING

The work of HaiW was supported by the Fundamental Research Funds for the Central Universities. The work of LZ was supported in part by the Center Chemistry in Solution and at Interfaces (CSI) funded by the DOE Award DE-SC0019394. The work of HanW was supported by the National Science Foundation of China under Grant No. 11871110, the National Key Research and Development

Program of China under Grants Nos. 2016YFB0201200 and 2016YFB0201203, and Beijing Academy of Artificial Intelligence (BAAI).

## ACKNOWLEDGMENTS

We are grateful for computing resources provided by the High-performance Computing Platform of Peking University, and the

Beijing Institute of Big Data Research, University of Science and Technology of China.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2020.589795/full#supplementary-material>

## REFERENCES

- Artrith, N., and Urban, A. (2016). An implementation of artificial neural-network potentials for atomistic materials simulations: performance for  $TiO_2$ . *Comput. Mater. Sci.* 114, 135–150. doi: 10.1016/j.commatsci.2015.11.047
- Avenda no-Franco, G., and Romero, A. H. (2016). Firefly algorithm for structural search. *J. Chem. Theory Comput.* 12, 3416–3428. doi: 10.1021/acs.jctc.5b01157
- Bahn, S. R., and Jacobsen, K. W. (2002). An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.* 4, 56–66. doi: 10.1109/5992.998641
- Bartók, A. P., Payne, M. C., Kondor, R., and Csányi, G. (2010). Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* 104:136403. doi: 10.1103/PhysRevLett.104.136403
- Behler, J., and Parrinello, M. (2007). Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* 98:146401. doi: 10.1103/PhysRevLett.98.146401
- Dickel, D. E., Baskes, M. I., Aslam, I., and Barrett, C. D. (2018). New interatomic potential for Mg-Al-Zn alloys with specific application to dilute Mg-based alloys. *Model. Simul. Mater. Sci. Eng.* 26:045010. doi: 10.1088/1361-651X/aaabaad
- Glass, C. W., Oganov, A. R., and Hansen, N. (2006). USPEX-Evolutionary crystal structure prediction. *Comput. Phys. Commun.* 175, 713–720. doi: 10.1016/j.cpc.2006.07.020
- Gupta, M., and Ling, S. N. M. (2011). *Magnesium, Magnesium Alloys, and Magnesium Composites*. Singapore: John Wiley & Sons. doi: 10.1002/9780470905098
- Haastруп, S., Strange, M., Pandey, M., Deilmann, T., Schmidt, P. S., Hinsche, N. F., et al. (2018). The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals. *2D Mater.* 5:042002. doi: 10.1088/2053-1583/aaacf1
- Han, J., Zhang, L., Car, R., and E, W. (2018). Deep potential: a general representation of a many-body potential energy surface. *Commun. Comput. Phys.* 23, 629–639. doi: 10.4208/cicp.OA-2017-0213
- He, J., Xia, Y., Naghavi, S. S., Ozolins, V., and Wolverton, C. (2019). Designing chemical analogs to PbTe with intrinsic high band degeneracy and low lattice thermal conductivity. *Nat. Commun.* 10, 1–8. doi: 10.1038/s41467-019-08542-1
- Huang, S. D., Shang, C., Kang, P. L., and Liu, Z. P. (2018). Atomic structure of boron resolved using machine learning and global sampling. *Chem. Sci.* 9, 8644–8655. doi: 10.1039/C8SC03427C
- Jelinek, B., Groh, S., Horstemeyer, M. F., Houze, J., Kim, S. G., Wagner, G. J., et al. (2012). Modified embedded atom method potential for Al, Si, Mg, Cu, and Fe alloys. *Phys. Rev. B* 85:245102. doi: 10.1103/PhysRevB.85.245102
- Khorshidi, A., and Peterson, A. A. (2016). AMP: a modular approach to machine learning in atomistic simulations. *Comput. Phys. Commun.* 207, 310–324. doi: 10.1016/j.cpc.2016.05.010
- Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140:A1133. doi: 10.1103/PhysRev.140.A1133
- Kresse, G., and Furthmüller, J. (1996). Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Phys. Rev. B* 6, 15–50. doi: 10.1016/0927-0256(96)00008-0
- Larsen, A. H., Mortensen, J. J., Blomqvist, J., Castelli, I. E., Christensen, R., Dulak, M., et al. (2017). The atomic simulation environment—a python library for working with atoms. *J. Phys.* 29:273002. doi: 10.1088/1361-648X/aa680e
- Mendeleev, M. I., Asta, M., Rahman, M. J., and Hoyt, J. J. (2009). Development of interatomic potentials appropriate for simulation of solid-liquid interface properties in Al-Mg alloys. *Philos. Mag.* 89, 3269–3285. doi: 10.1080/14786430903260727
- Mouhat, F., and Coudert, F. X. (2014). Necessary and sufficient elastic stability conditions in various crystal systems. *Phys. Rev. B* 90:224104. doi: 10.1103/PhysRevB.90.224104
- Oganov, A. R. (2018). Crystal structure prediction: reflections on present status and challenges. *Faraday Discuss.* 211, 643–660. doi: 10.1039/C8FD90033G
- Ong, S. P., Richards, W. D., Jain, A., Hautier, G., Kocher, M., Cholia, S., et al. (2013). Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci.* 68, 314–319. doi: 10.1016/j.commatsci.2012.10.028
- Pack, J. D., and Monkhorst, H. J. (1977). Special points for Brillouin-zone integrations. *Phys. Rev. B* 16, 1748–1749. doi: 10.1103/PhysRevB.16.1748
- Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77, 3865–3868. doi: 10.1103/PhysRevLett.77.3865
- Plimpton, S. (1995). Fast parallel algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* 117, 1–19. doi: 10.1006/jcph.1995.1039
- Podryabinkin, E. V., Tikhonov, E. V., Shapeev, A. V., and Oganov, A. R. (2019). Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Phys. Rev. B* 99:064114. doi: 10.1103/PhysRevB.99.064114
- Pugh, S. (1954). XCII. Relations between the elastic moduli and the plastic properties of polycrystalline pure metals. *Lond. Edinburgh Dubl. Philos. Mag. J. Sci.* 45, 823–843. doi: 10.1080/14786440808520496
- Shapeev, A. V. (2016). Moment tensor potentials: a class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* 14, 1153–1173. doi: 10.1137/15M1054183
- Tong, Q., Xue, L., Lv, J., Wang, Y., and Ma, Y. (2018). Accelerating CALYPSO structure prediction by data-driven learning of a potential energy surface. *Faraday Discuss.* 211, 31–43. doi: 10.1039/C8FD00055G
- Wang, H., Zhang, L., Han, J., and E, W. (2018). DeepMD-kit: a deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* 228, 178–184. doi: 10.1016/j.cpc.2018.03.016
- Wang, Y., Lv, J., Zhu, L., and Ma, Y. (2010). Crystal structure prediction via particle-swarm optimization. *Phys. Rev. B* 82:094116. doi: 10.1103/PhysRevB.82.094116
- Wang, Y., Lv, J., Zhu, L., and Ma, Y. (2012). CALYPSO: A method for crystal structure prediction. *Comput. Phys. Commun.* 183, 2063–2070. doi: 10.1016/j.cpc.2012.05.008
- Xie, H., Zhang, W., Duan, D., Huang, X., Huang, Y., Song, H., et al. (2020). Superconducting zirconium polyhydrides at moderate pressures. *J. Phys. Chem. Lett.* 11, 646–651. doi: 10.1021/acs.jpclett.9b03632
- Xu, W. W., Yu, Q., Xia, F., Yin, G., Huang, Y., and Chen, L. (2019). Exploration of crystal structure and the origin of unexpected intrinsic ductility of  $\chi$   $Co_7Ta_2$ . *J. Alloys Compounds* 797, 1198–1204. doi: 10.1016/j.jallcom.2019.05.223
- Zhang, L., Han, J., Wang, H., Car, R., and E, W. (2018a). Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* 120:143001. doi: 10.1103/PhysRevLett.120.143001
- Zhang, L., Han, J., Wang, H., Saidi, W., Car, R., and E, W. (2018b). End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. *Adv. Neural Inform. Process. Syst.* 31, 4441–4451.
- Zhang, L., Lin, D.-Y., Wang, H., Car, R., and E, W. (2019). Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys. Rev. Mater.* 3:023804. doi: 10.1103/PhysRevMaterials.3.023804

- Zhang, Y., Wang, H., Chen, W., Zeng, J., Zhang, L., Wang, H., et al. (2020). DP-GEN: a concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.* 253:107206. doi: 10.1016/j.cpc.2020.107206
- Zhao, Z., Zhang, S., Yu, T., Xu, H., Bergara, A., and Yang, G. (2019). Predicted pressure-induced superconducting transition in electride Li<sub>6</sub>P. *Phys. Rev. Lett.* 122:97002. doi: 10.1103/PhysRevLett.122.097002
- Zhuang, H. L., Chen, M., and Carter, E. A. (2017). Prediction and characterization of an Mg-Al intermetallic compound with potentially improved ductility via orbital-free and Kohn-Sham density functional theory. *Model. Simul. Mater. Sci. Eng.* 25:075002. doi: 10.1088/1361-651X/aa7e0c

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Wang, Zhang, Zhang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.