

Supplementary information: Machine learning interatomic potentials for solid-state precipitation

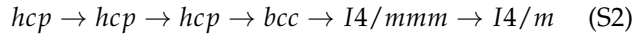
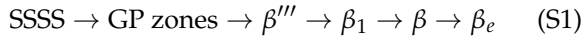
Lorenzo Piersante¹ and Anirudh Raju Natarajan^{1,2,*}

¹Laboratory of materials design and simulation (MADES),
Institute of Materials, École Polytechnique Fédérale de Lausanne

²National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne

SI. CRYSTALLOGRAPHY OF THE MG-ND BINARY SYSTEM

Figure S1 illustrates the crystal structures of all known metastable and stable phases of the Mg-Nd system [1–3]. Precipitation in a dilute Mg-Nd alloy ($x_{\text{Nd}} \leq 0.5$ at.%) proceeds as follows [3]:



In the early stages of aging, Guinier-Preston (GP) zones form first. Next, the β''' precipitate emerges via an order-disorder transformation. The β''' phase comprises an infinite family of structures generated by stacking zigzag and hexagonal chains of Nd atoms along the $[10\bar{1}0]$ direction of hcp [3]. Figure S1 shows one such arrangement, labeled β'''_{01} . As heat treatment continues, the β_1 phase forms, followed by the metastable β and equilibrium β_e phases at longer aging times. Precipitation of these later phases reduces alloy strength from its peak value. At higher Nd concentrations, two additional equilibrium phases occur: the high-temperature C15 Laves phase (Mg_2Nd) and an equiatomic B2 (MgNd) intermetallic compound.

SII. MEAM BENCHMARKS

Figure S2(a) shows the metastable hcp convex hull obtained with the MEAM potential of Kim et al. [4]. All symmetry-distinct orderings identified as hcp from their relaxed DFT structures were fully relaxed with the MEAM potential, and their formation energies are plotted in Fig. S2(a). A large discrepancy between DFT and MEAM predictions is evident. Two spurious ground states appear at compositions $x_{\text{Nd}} > 1/4$ with formation energies so negative that the experimentally observed β''' phases no longer lie on the hcp convex hull. As a result, the MEAM potential is not suitable for studying precipitation in dilute Mg-Nd alloys.

SIII. ATOMIC CLUSTER EXPANSION POTENTIAL

The ACE MLIP fitted in this work uses spherical Bessel functions as radial basis functions, with a radial parameter of 5.25 and a cosine cutoff [5]. At short interatomic distances, the potential interpolates to a ZBL repulsion. An inner cutoff function mediates this transition by smoothly switching off the radial basis functions within the region $[r_{\text{in}} - \Delta_{\text{in}}, r_{\text{in}}]$. At short distances, when $r < r_{\text{in}} - \Delta_{\text{in}}$, the radial basis functions vanish and only the ZBL repulsion contributes [5]. Table SI summarizes the values of r_{in} and Δ_{in} used in this work.

Block	r_{in} (Å)	Δ_{in} (Å)
Mg	1.2	0.2
Nd	2.2	0.4
Mg-Nd	2.0	0.2

TABLE SI. Inner cutoff parameters of the ACE potential.

The Mg basis set was taken from [6]. The Nd basis set was optimized to capture the energy ordering of dhcp, fcc, and hcp polymorphs with a small number of basis functions. The binary basis set was adjusted to balance computational efficiency and accuracy. Figure S2(b) shows the training RMSE as a function of basis set size. All potentials were trained for 200 epochs on Mg-Nd symmetry-distinct orderings enumerated with the CASM code (see Table SIV). The figure also reports the computational speed for CPU evaluation with LAMMPS (8 Intel i7-12700 processors). From the largest to the smallest basis set, computational efficiency improves by approximately 300%, while the training error increases by around 30%. We selected the smallest basis set without quintuplets as the optimal compromise between computational efficiency and accuracy.

* anirudh.natarajan@epfl.ch

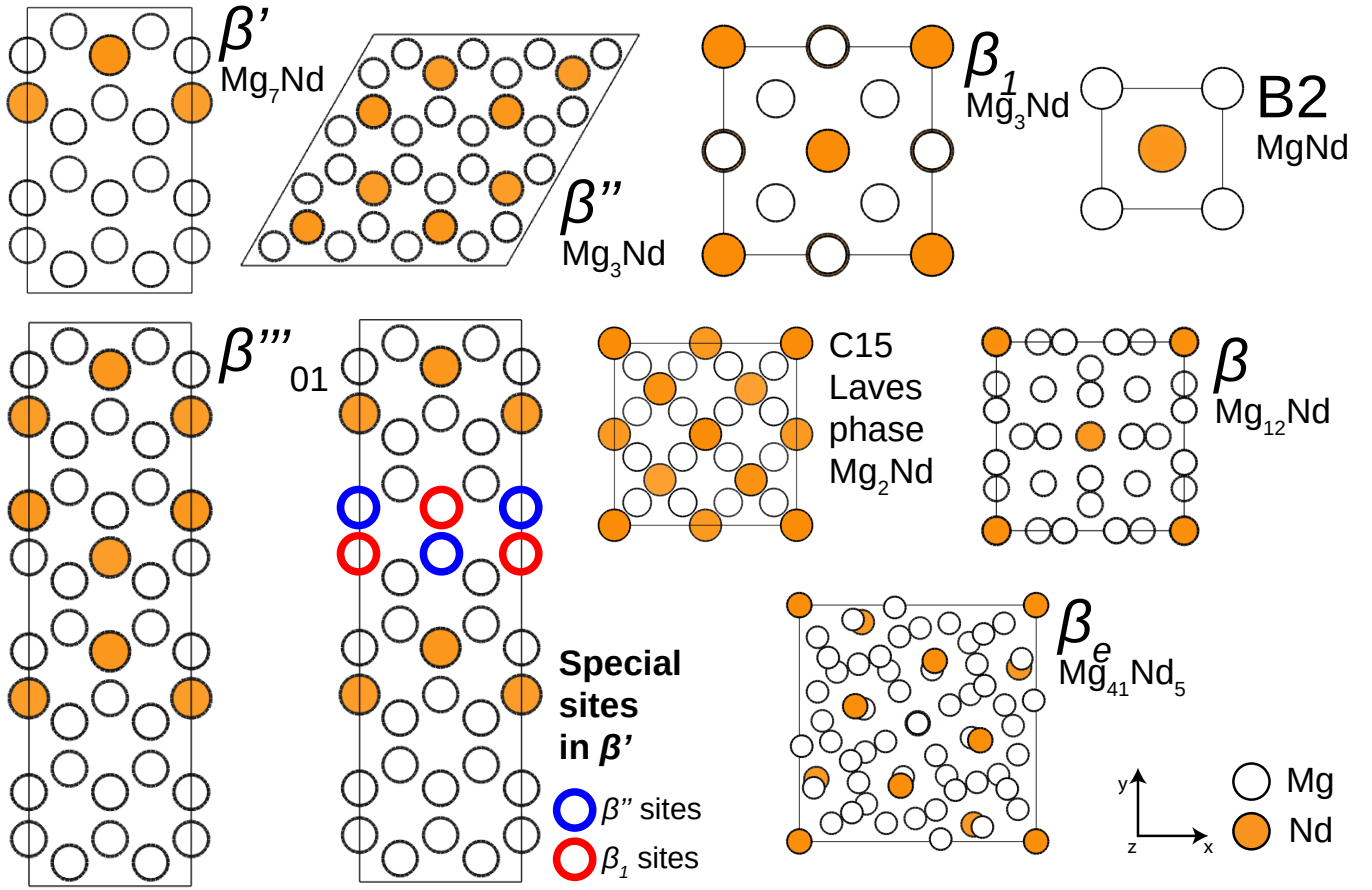


FIG. S1. Schematic crystal structures of the stable and metastable phases in the Mg-Nd system. All crystal structures are projected along the c -axis.

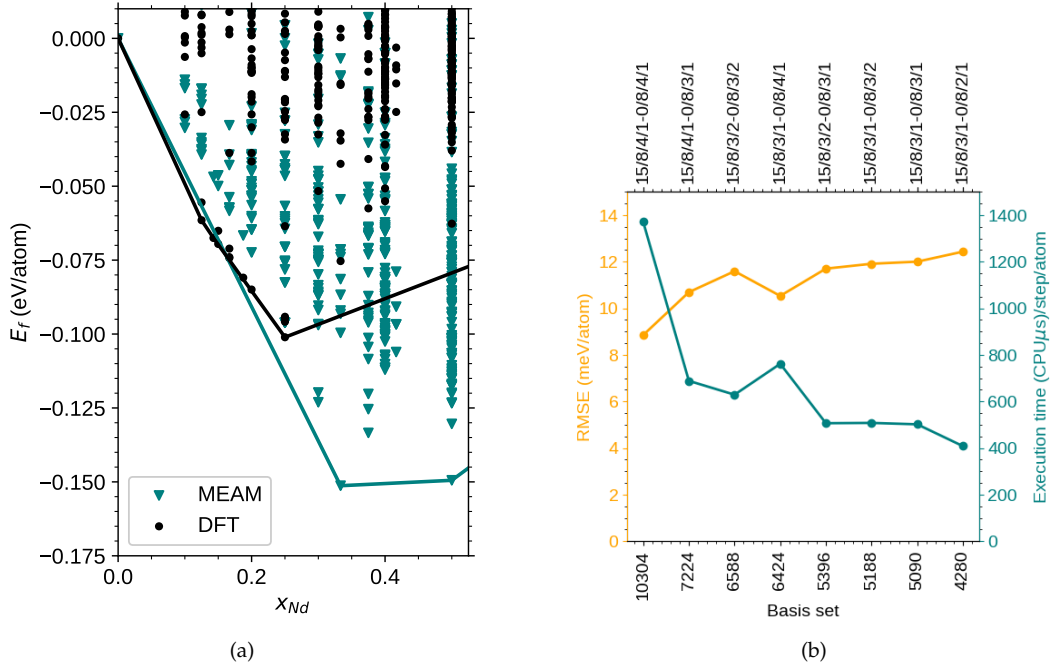


FIG. S2. (a) Comparison of the hcp convex hull obtained with the MEAM potential and DFT calculations. (b) Variation of training RMSE and computational time as a function of binary basis set.

SIV. REFERENCE DATABASE

Tables SII to SIV list all structures contained in our reference database of DFT calculations. Calculations along each relaxation trajectory were retained, resulting in more DFT data points than structures.

The database for pure Mg includes structures along the well-known Bain ($bcc \leftrightarrow fcc$), Burgers ($bcc \leftrightarrow hcp$), and trigonal ($bcc \rightarrow SC \rightarrow fcc$) transformation pathways. We also included structures along strain distortions connecting different orientational variants of bcc relative to hcp . Ab initio molecular dynamics (AIMD) simulations were performed with VASP [7–10] in the NVT ensemble at various temperatures and densities for pure Mg and Nd. Symmetrically-distinct chemical decorations were enumerated in hcp , bcc , fcc , and C15 Laves parent crystals with the CASM code [11]. Liquid structures in binary Mg-Nd were sampled at compositions $x_{Nd} = 1/8, 1/4, 1/3$, and $1/2$ with Monte Carlo/molecular dynamics in the NPT ensemble as implemented in LAMMPS [12] using the MEAM potential of Kim et al. [4]. Crystal structures with atomic displacements were created either by applying random displacements to atomic coordinates or by using the enumeration routines in phonopy [13]. In pure Mg, atomic displacements were enumerated in hcp , bcc , fcc , $dhcp$, and ω structures. In Mg-Nd, we created rattled structures for all stable and metastable convex hull states up to $x_{Nd} = 1/2$. Lastly, strain-shuffle maps denote structures that systematically sample the Burgers transformation across strain and displacement.

Type	Structures	Data points
Intermetallics [14]	379	1644
Dimers	39	39
hcp strains [15]	228	515
hcp cohesive energy	36	36
hcp GSFs	81	81
hcp SSFs	3	3
hcp bulk/dot/rod geometries	14	14
VD of 39 prototypes	907	2471
Vacancy structures	1510	31328
hcp/fcc surfaces [16]	27	83
hcp surface decohesion	184	184
hcp/fcc/bcc Σ GBs [16]	27	75
Phonon displacements [13]	8	8
AIMD snapshots	1820	1820
Transformation pathways	1167	1167
bcc strain-shuffle maps	966	966
TOTAL	7396	40434

TABLE SII. Number of structures and data points in the reference database for pure Mg. Abbreviations: generalized stacking fault (GSF), stable stacking fault (SSF), volumetric distortion (VD), grain boundary (GB), ab-initio molecular dynamics (AIMD).

Type	Structures	Data points
Intermetallics [14]	349	1169
Dimers	39	39
VD of 34 prototypes	714	3256
AIMD snapshots	2104	2104
TOTAL	3206	6568

TABLE SIII. Number of structures and data points in the reference database for pure Nd. Abbreviations: volumetric distortion (VD), ab-initio molecular dynamics (AIMD).

Type	Structures	Data points
Experimental phases [1, 2]	14	33
Symmetry-distinct orderings [11, 17]	1851	7444
Data generation scheme orderings	297	1261
Dimers	39	39
Burgers pathways	248	435
h - β_1 strain-shuffle maps	4605	90247
Nd dissolution in hcp	2	101
Nd-Nd binding in hcp	12	215
Nd in hcp Σ GB	12	38
Nd in hcp SSFs	40	674
VD of convex hull states	273	1476
$\beta'/\beta''/\beta_1/\beta_e$ strains [15]	455	4067
Nd-vacancy interaction in hcp	16	184
$\beta'/\beta''/\beta'''/\beta_1$ /NdMg ₂ vacancies	34	435
$\beta'/\beta''/\beta_1$ GSFE profiles	72	72
β'/β'' APBs	5	27
Random phonon displacements	1300	1300
Liquid structures	1594	1594
TOTAL	10869	109642

TABLE SIV. Number of structures and data points in the reference database for binary Mg-Nd. Abbreviations: grain boundary (GB), stable stacking fault (SSF), volumetric distortion (VD), generalized stacking fault energy (GSFE), anti-phase boundary (APB).

SV. DATA SAMPLING WITH MAHALANOBIS DISTANCES

The mean and covariance matrix used to compute the Mahalanobis distance in Fig. 5 are updated with the following two equations:

$$\mu_{N+1} = \frac{N}{N+1}\mu_N + \frac{1}{N+1}B_i^{new} \quad (S3)$$

$$\Sigma_{N+1} = \frac{N-1}{N}\Sigma_N + \frac{(\mu_N - B_i^{new})(\mu_N - B_i^{new})^T}{N+1} \quad (S4)$$

where N is the number of atomic sites in the initial dataset, μ_N and Σ_N are the mean and covariance matrix of the initial dataset, and B_i^{new} is the descriptor of the

newly selected site. In the workflow, we select the structure containing the atom with the largest Mahalanobis distance. We then apply Eq. (S3) and Eq. (S4) iteratively to update the mean and covariance matrix over all sites in the newly selected structure.

SVI. DESCRIPTOR-BASED CLASSIFICATION

The active learning loop described in Fig. 4 requires assigning nearly 0.5 million structures to four different parent crystal structures. Direct application of the structure-matching algorithm of Thomas et al. [17] is infeasible at this scale, so we instead use a method based on geometric descriptors. A simple Euclidean distance between the average descriptor of a relaxed structure and that of the reference parent structure is also difficult to apply because assignment thresholds are sensitive to small lattice and atomic distortions.

To address this, we compute the Mahalanobis distance of each relaxed structure relative to a small set of structures labeled using the accurate structure-matching method. Because crystal structures on a given convex hull (e.g., fcc) deviate slightly from the ideal parent structure, the Mahalanobis distance relative to such a dataset naturally accounts for the spread in descriptor values. We classified an initial set of symmetry-distinct orderings (Table SIV) using the structure-matching algorithm in CASM [11, 17], obtaining four separate convex hulls for hcp, bcc, fcc, and C15 Laves parent lattices. These labeled datasets then served as references for computing the Mahalanobis distance of structures generated during active learning. Algorithm SI outlines the classification procedure. We used an ACE basis set of 15/5-0/4 throughout the classification routine.

Algorithm SI: Descriptor-based classification

- 1: mask the atomic species of the new ordering;
 - 2: calculate the average descriptor;
 - 3: evaluate the Mahalanobis distance, d_M , of the average descriptor relative to each reference dataset (hcp, bcc, fcc, C15);
 - 4: **if** $d_M^{hcp}, d_M^{bcc}, d_M^{fcc}, d_M^{C15} > \text{tolerance}$ **then**
 - 5: assign the ordering to no lattice;
 - end**
 - else**
 - 6: assign the ordering to the nearest lattice;
 - end**
-

SVII. TRAINING THE MLIP

Figure S3 shows the distribution of data points across composition for the reference database and the training datasets used in fitting the Mg-Nd MLIP. Construction of the training datasets combined domain knowledge with

the Mahalanobis distance sampling workflow shown in Fig. 5. Three training datasets were developed to parameterize the Mg, Nd, and Mg-Nd MLIPs. Each training dataset began with a set of initial structures chosen based on domain knowledge. Additional data points were then selected using the procedure in Fig. 5 to enhance structural and chemical diversity. For the Mg-Nd dataset, we also included structures containing atoms closer than 2.5 Å to capture the core repulsion. The abbreviations employed in the following paragraphs are listed in the captions of Tables SII to SIV.

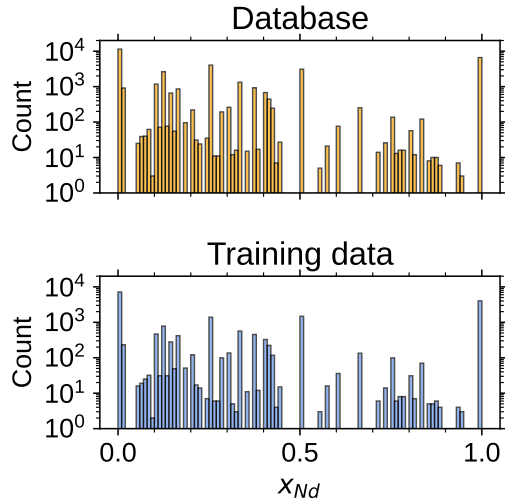


FIG. S3. Histograms of crystal structures in the reference database (top) and training datasets (bottom) as a function of composition. Data points from Mg vacancy relaxation trajectories and Mg-Nd h - β_1 strain-shuffle maps are excluded for clarity.

The pure Mg training set began with 4405 data points comprising prototypical intermetallic structures, dimers, deformations of the bulk hcp structure (strains, GSFs, SSFs, bulk/rod/corner geometries, and Σ GBs), hcp slab geometries, transformation pathways, volumetric distortions of hcp and bcc, strain-shuffle distortions of bcc related to the Burgers transformation, and vacancies in several prototypical structures (hcp, bcc, fcc, and C15). Using this initial dataset as a reference, we sampled 2650 additional data points from the remaining pool of structures. The final training set for the pure Mg potential contained 7055 data points.

The pure Nd training set was initialized with 2097 data points comprising prototypical intermetallic structures, dimers, and volumetric distortions of all structural prototypes with energies below that of hcp-Nd. Using this initial dataset as a reference, we sampled an additional 708 data points from the volumetric distortions of all other structural prototypes. The updated dataset then served as the reference for sampling 1200 AIMD snapshots. The final pure Nd training set contained 4005 data points.

The initial Mg-Nd training set contained 5563 data points comprising experimental phases, dimers, Nd dissolution configurations, Nd-Nd pairs at varying distances embedded within a hcp-Mg matrix, symmetry-distinct orderings enumerated with CASM, symmetry-distinct orderings identified with the iterative data generation scheme in Fig. 4, volumetric distortions of phases on the convex hull, Nd-SSF interactions, vacancies in $\beta'/\beta''/\beta_1/\text{Mg}_2\text{Nd}$, Nd-vacancy interactions, symmetry-distinct Burgers pathways starting from hcp orderings in supercells containing up to 4 atoms, and symmetry-distinct Burgers pathways of β' with occupation of some $h\text{-}\beta_1$ sites. Using this initial data as a reference set, we sampled 75 liquid structures at four compositions ($x_{\text{Nd}} = 1/8, 1/4, 1/3$, and $1/2$), 30 rattled structures for each of 13 structural prototypes corresponding to all global and metastable ground states up to $x_{\text{Nd}} = 1/2$, 500 strained configurations of $\beta'/\beta''/\beta_1/\beta_e$, 400 structures from the Burgers transformation of $h\text{-}\beta_1$, 30 Nd- Σ GB structures, and 625 data points from the remaining volumetric distortions, iterative data generation scheme, Nd-SSF, Nd-vacancy, Nd dissolution, and Nd-Nd binding data. The final binary Mg-Nd training set contained 7985 data points.

SVIII. INFLUENCE OF THE WEIGHTING SCHEME

This section investigates the influence of the weighting scheme on the τ_w plots. Figures S4(a) to S4(d) reproduce the analyses of Fig. 11(a) and Figs. 13(a) to 13(c) using different weighting schemes. We considered two possibilities. The first is the additive or multiplicative combination of weights defined by Vigna [18]. The second is the additive or multiplicative combination of weights following the definition used in the pacemaker code [5]. pacemaker uses the weighting function:

$$w(i) = \frac{n_e}{(e_i^{\text{DFT}} - e_{\text{min}}^{\text{DFT}} + \epsilon)^2} \quad (\text{S5})$$

where e_i^{DFT} is the normalized DFT energy of structure i , ϵ is a softening factor, and n_e is a normalizing constant. We set $\epsilon = 0.01$ eV to assign greater weight to structures close to the ground state. In the binary case, the energy is replaced by the semi-grand canonical potential evaluated at the desired chemical potential.

As Figs. S4(a) to S4(d) illustrate, τ_w follows similar trends whether the weights are additive or multiplicative. The additive curves generally lie above the multiplicative counterparts, and the drops in τ_w occur at the same locations. The multiplicative scheme more readily identifies low-energy phase stability errors because it responds more strongly to phase stability swaps. For a given weight combination method, the Vigna and pacemaker weighting schemes produce qualitatively similar curves.

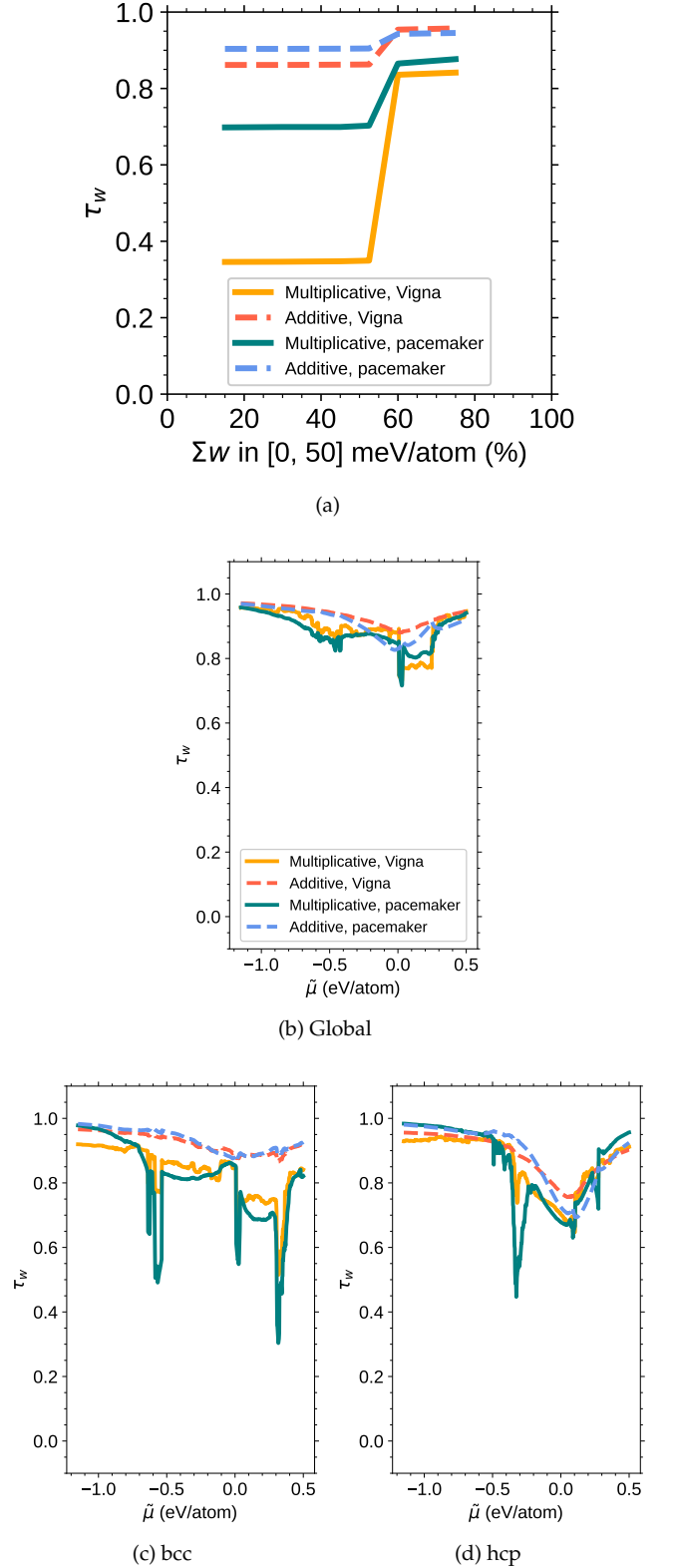


FIG. S4. (a) τ_w as a function of the total weight assigned to data points in the 0–50 meV/atom range for Mg and Nd. $\tilde{\mu}-\tau_w$ traces for (b) the global system, (c) the bcc phase, and (d) the hcp phase.

SIX. MLIP BENCHMARKS

Property	Experiment	DFT	MLIP
Lattice parameters			
a (Å)	3.21 (298K) [19]	3.21	3.20
c/a	1.624 (298K) [19]	1.614	1.617
Mechanical properties (GPa)			
C_{11}	63.48 [20]	62.5	62.4
C_{12}	25.94 [20]	22.3	23.4
C_{13}	21.7 [20]	21	21
C_{33}	66.45 [20]	63.8	64.8
C_{44}	18.42 [20]	17.6	17.4
C_{66}	18.75 [20]	20.1	19.5
B	36.7 [21]	36.2	36.4
Surface/interface properties (mJ/m ²)			
γ_{SSF} Basal I1	33 [22]	34 [23]	38
γ_{SSF} Prism	-	225	187
γ_{SSF} Pyramid I	-	160	149
γ_{SSF} Pyramid II	-	155	165
E_{surf} Basal (0001)	-	549	573
E_{surf} Prism (10 $\bar{1}$ 0)	-	607	615
E_{surf} Pyramid I (10 $\bar{1}$ 1)	-	639	665
E_{surf} Pyramid II (11 $\bar{2}$ 1)	-	771	767
E_{surf} Σ GB (0001)	-	168	177
Vacancy properties			
E_{vac} (eV)	0.79 [24]	0.813	0.818
E_{act} basal hop (meV)	-	390	401
E_{act} pyramidal hop (meV)	-	403	420

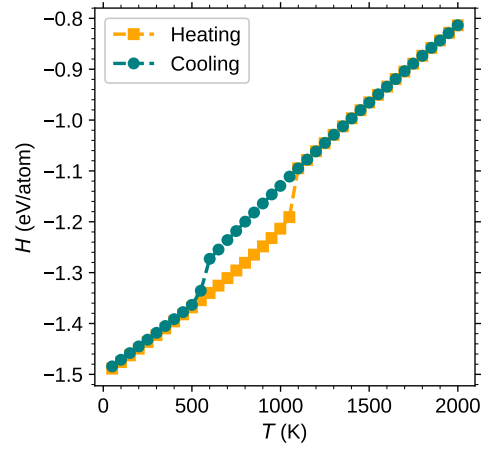
TABLE SV. Comparison of selected hcp-Mg properties from experiment, DFT, and MLIP.

Property	Exp.	DFT	MLIP
a (Å)	3.65 [25]	3.71	3.70
c/a	3.226 [25]	3.216	3.220
B (GPa)	31.8 [26]	33.4	33.8

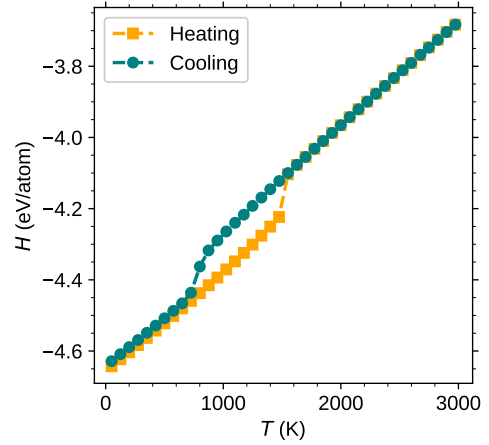
TABLE SVI. Comparison of lattice parameters and mechanical properties of dhcp-Nd from experiment, DFT, and MLIP.

Type	Data points
Common intermetallics	1644
Dimers	39
Sampled VD data	1600
Sampled AIMD snapshots	1200

TABLE SVII. Training dataset to fit the Mg potentials as a function of the low-energy weights. Abbreviations: volumetric distortion (VD), ab-initio molecular dynamics (AIMD).



(a) hcp-Mg



(b) dhcp-Nd

FIG. S5. Ensemble-averaged enthalpy as a function of temperature computed via NPT molecular dynamics. Trajectories were initiated at low temperature for (a) hcp-Mg and (b) dhcp-Nd.

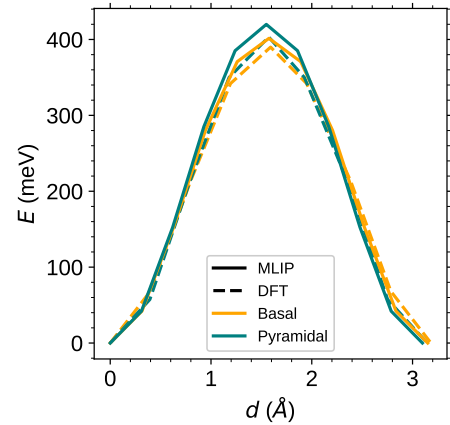


FIG. S6. Comparison of vacancy hops in hcp-Mg computed via the NEB method using DFT and MLIP.

A. Unaries: Mg, Nd

The first set of benchmarks concerns the properties of the pure elements. Tables SV and SVI summarize the materials properties of hcp-Mg and dhcp-Nd. Figure S5(a) shows the change in enthalpy during a heating-cooling cycle carried out with molecular dynamics in the NPT ensemble at zero pressure, starting from hcp-Mg. The structure after cooling contained some defects and had a slightly higher enthalpy. Figure S5(b) shows the corresponding heating-cooling cycle for dhcp-Nd. The structure after cooling was predominantly fcc-Nd and had a higher enthalpy than the dhcp-Nd starting point. Lastly, Fig. S6 compares DFT and MLIP energies of vacancy hops in hcp-Mg along the pyramidal and basal planes computed with the nudged elastic band (NEB) method [27]. The simulations were performed in a $4 \times 4 \times 3$ supercell.

Table SVII summarizes the training dataset used to fit the Mg potentials used in the investigation of the influence of the weighting scheme in Fig. 11(a). The initial dataset contained the prototypes of intermetallics and the dimers. This dataset was the starting point for sampling volumetric distortions of several prototypes and AIMD snapshots. The final dataset had 4483 data points and closely resembled the dataset used to train the Nd potentials.

B. Binary: Mg-Nd

1. General benchmarks

This section reports several results about the training and validation of the binary MLIP. Figure S7 compares the convex hulls across the hcp, bcc, C15 Laves, and fcc convex hulls as predicted by the preliminary MLIP used in the data generation loop and the finalized MLIP. This potential was fitted after the sampled orderings were added to the training dataset. Table SVIII reports the lattice parameters and the bulk moduli of the Mg_2Nd C15 Laves phase and the MgNd B2 phase. Figures S8(a) to S8(c) report the Nd-Nd binding energy as a function of distance between solutes, the binding energy of a Nd atom to different stable stacking fault geometries as a function of distance from the faulty plane, and the binding energy of a Nd atom to the Σ grain boundary again as a function of distance. Figure S9 summarizes the vacancy formation energies of Mg and Nd vacancies in the β' , β'' , and β_1 phases. Figure S9(a) is a diagram enumerating the locations of the symmetry-distinct vacancy sites. Figures S9(b) and S9(c) collect the formation energy for creating vacancies at each of these sites. Lastly, Fig. S10 shows the binding energy between a Nd atom and a vacancy as a function of their separation.

Property	Mg_2Nd		MgNd	
	DFT	MLIP	DFT	MLIP
a (Å)	3.88	3.89	6.14	6.15
B (GPa)	38.6	41.3	40.0	41.4

TABLE SVIII. Comparison of structural and mechanical properties for phases with $x_{\text{Nd}} \geq 1/4$ from DFT, and MLIP.

2. Interface properties

Figures S11(b) to S11(d), and Table SIX compare generalized stacking fault energy (GSFE) profiles and interface energies between DFT and the MLIP. Figure S11(a) shows the basis vectors (red arrows) used to map the GSFE profiles. We compare the MLIP with DFT only along these directions. The green dots in the β' schematic mark the locations where we computed anti-phase boundary (APB) energies. We also compare interface energies predicted by DFT and the MLIP at different levels of lattice strain. We evaluated interface energies when two phases were forced to coexist at the lattice parameters of each phase and at their average. Cells were relaxed in the direction perpendicular to the interface. We computed interface energies with the MLIP, followed by static DFT calculations at the MLIP-predicted lattice parameters for validation.

The energy profiles in Figs. S11(b) to S11(d) demonstrate close agreement between the MLIP and DFT. The APB produced by a shift along the $[2\bar{1}\bar{1}0]_{\text{hcp}}$ vector in β' has an MLIP-predicted energy of 173 mJ/m^2 against a DFT value of 180 mJ/m^2 . This result also agrees with the value reported in [28]. The other two APBs have MLIP-computed energies of 129 and 55 mJ/m^2 , matching the DFT values of 117 and 66 mJ/m^2 .

Phases		Planes		Latt. 1		Average		Latt. 2		Lit.
1	2	1	2	DFT	IP	DFT	IP	DFT	IP	
β'	hcp	(100)	($2\bar{1}\bar{1}0$)	26	44	17	42	18	38	46 [29]
		(010)	(01 $\bar{1}0$)	69	55	75	61	5	47	51 [29]
		(001)	(0001)	49	45	31	39	23	23	28 [29]
β'	β_1	(100)	(001)	36	22	36	21	37	18	-
		(010)	($\bar{1}10$)	24	23	21	22	18	21	-
		(001)	(110)	6	5	8	5	11	5	-
β_1	hcp	($\bar{1}12$)	(01 $\bar{1}0$)	93	83	88	81	85	79	98 [30]
β_1	β'	($\bar{1}12$)	(010)	142	128	212	176	148	107	-

TABLE SIX. Interfacial energies for various precipitate/matrix combinations. Orientation relationships are detailed in the text. Latt. and Lit. abbreviate lattice parameters and literature values, respectively.

Table SIX compares several interface energies. We analyzed various interfaces between hcp-Mg, β' , and β_1 . We considered two experimentally observed interfaces [31–33]: the β' -hcp interface with orientation relationship

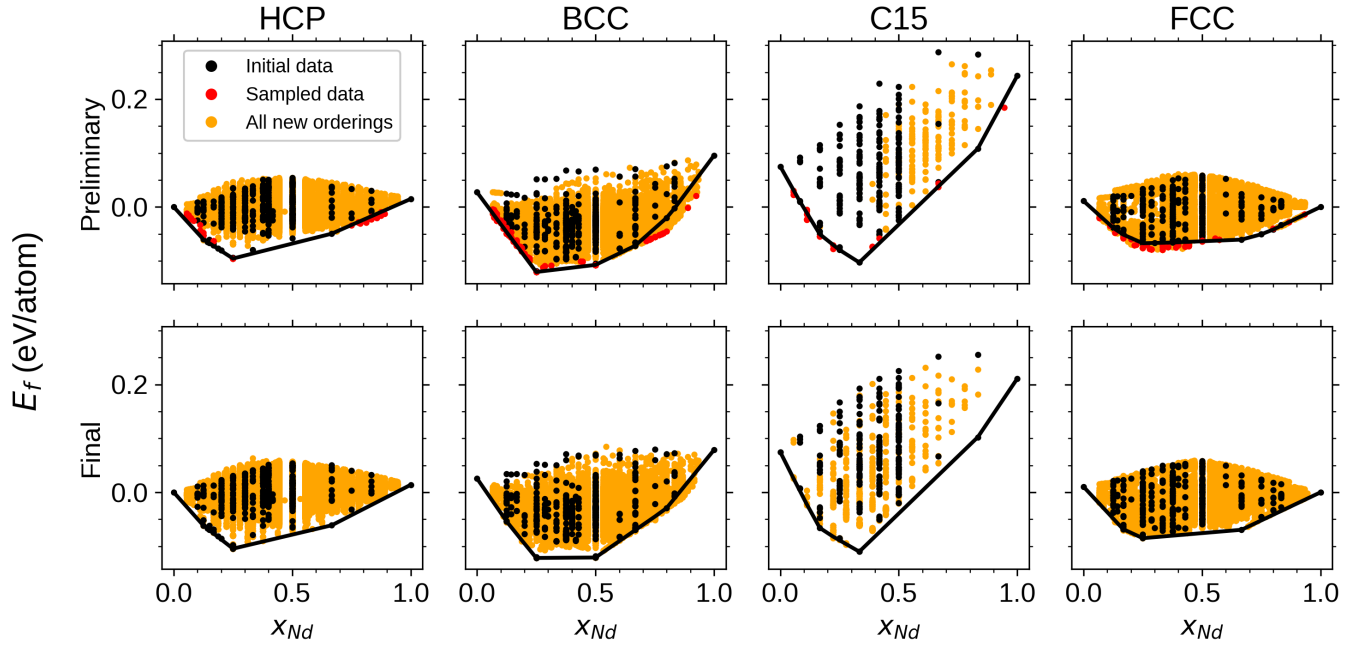
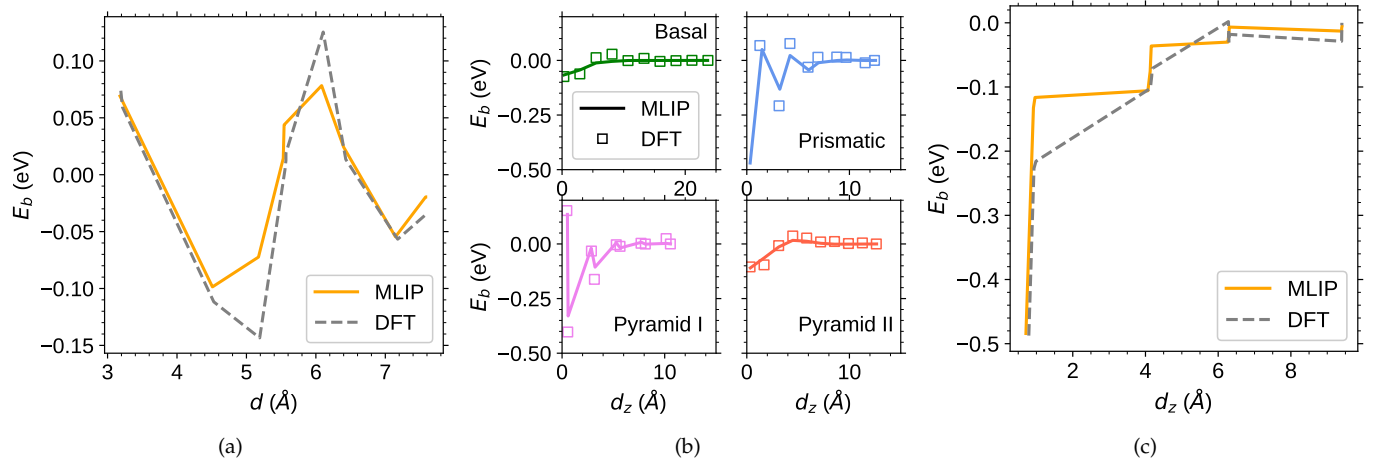
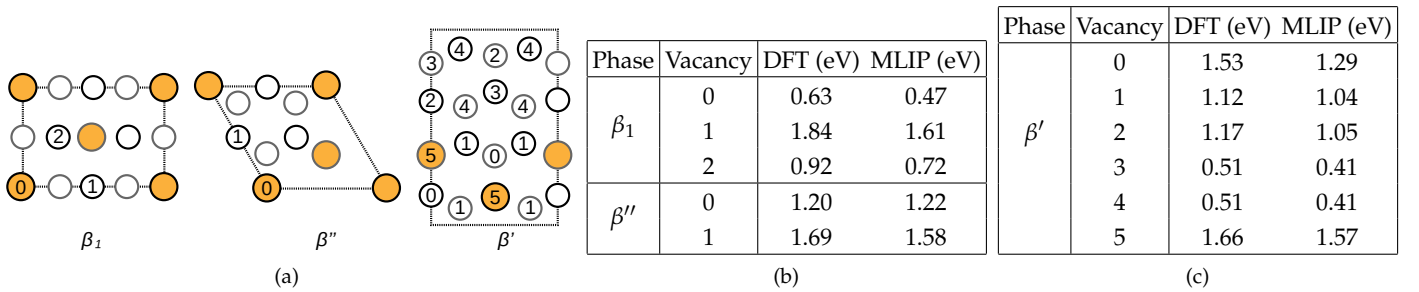


FIG. S7. Convex hulls after one round of iterative data generation.

FIG. S8. Comparison of DFT and MLIP energies for (a) Nd-Nd binding, (b) Nd-SSF binding, and (c) Nd- Σ grain boundary interaction.FIG. S9. (a) Vacancy sites in the β_1 , β'' , and β' structures. (b)-(c) Comparison of vacancy formation energies computed via DFT and MLIP.

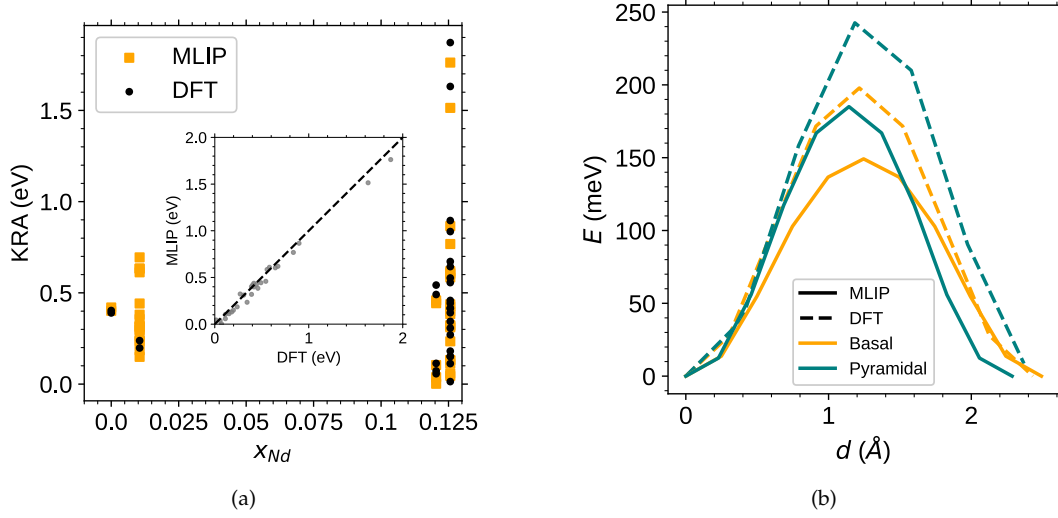


FIG. S12. (a) KRA energy barriers for Nd compositions up to $x_{\text{Nd}} = 1/8$. Scatter points clustered around $x_{\text{Nd}} = 0$ and $1/8$ represent hops in hcp-Mg and β' , respectively. Inset: Parity plot comparing KRA barriers computed with DFT and MLIP. (b) KRA energy profiles for Nd-vacancy exchange in hcp-Mg computed with the NEB method using DFT and MLIP.

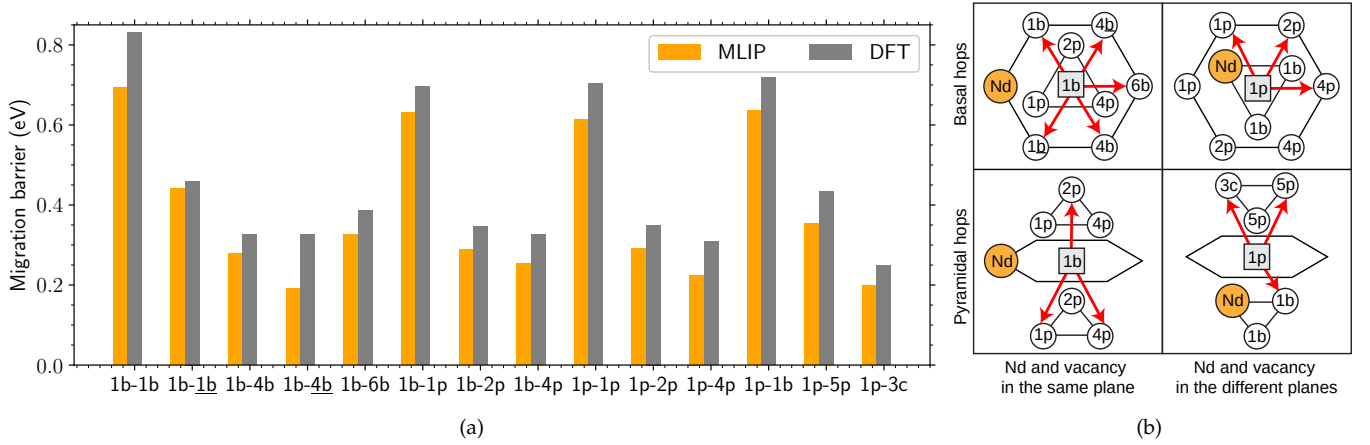


FIG. S13. (a) Comparison of migration energy barriers for vacancy hops away from a Nd atom in hcp-Mg computed with MLIP and DFT (DFT data from [34]). (b) Schematic of the hop geometry. The gray square and white circles indicate the vacancy and Mg atoms, respectively. Labels specify the neighbor shell (1st NN, 2nd NN, etc.) and geometric location (b : basal, p : pyramidal, c : column), following the convention in [34].

starting vacancy configuration. Figure S13(b) illustrates the vacancy jump geometry and details the labeling convention. We compare the MLIP results to DFT data available in [34]. Although the MLIP slightly underestimates the migration barriers, it correctly reproduces the overall trends.

Figure S14 provides an overview of the 26 symmetry-distinct vacancy hops in β' . The crystal structure at the top of the diagram marks the different jumps with arrows, distinguishing basal (solid) and pyramidal (dashed) hops for Mg (red) and Nd (blue). The KRA barriers evaluated with the MLIP are plotted at Nd compositions around $x_{\text{Nd}} = 1/8$ in Fig. S12(a). The corre-

sponding DFT values come from static calculations on structures obtained from NEB calculations with the MLIP. The plots in Fig. S14 indicate good agreement between the MLIP and DFT. Interestingly, several vacancy hops in β' have very small barriers or are unstable. This is the case for hops number 2, 6, 7, 10, 11, 15, 19, 20, 22, and 23.

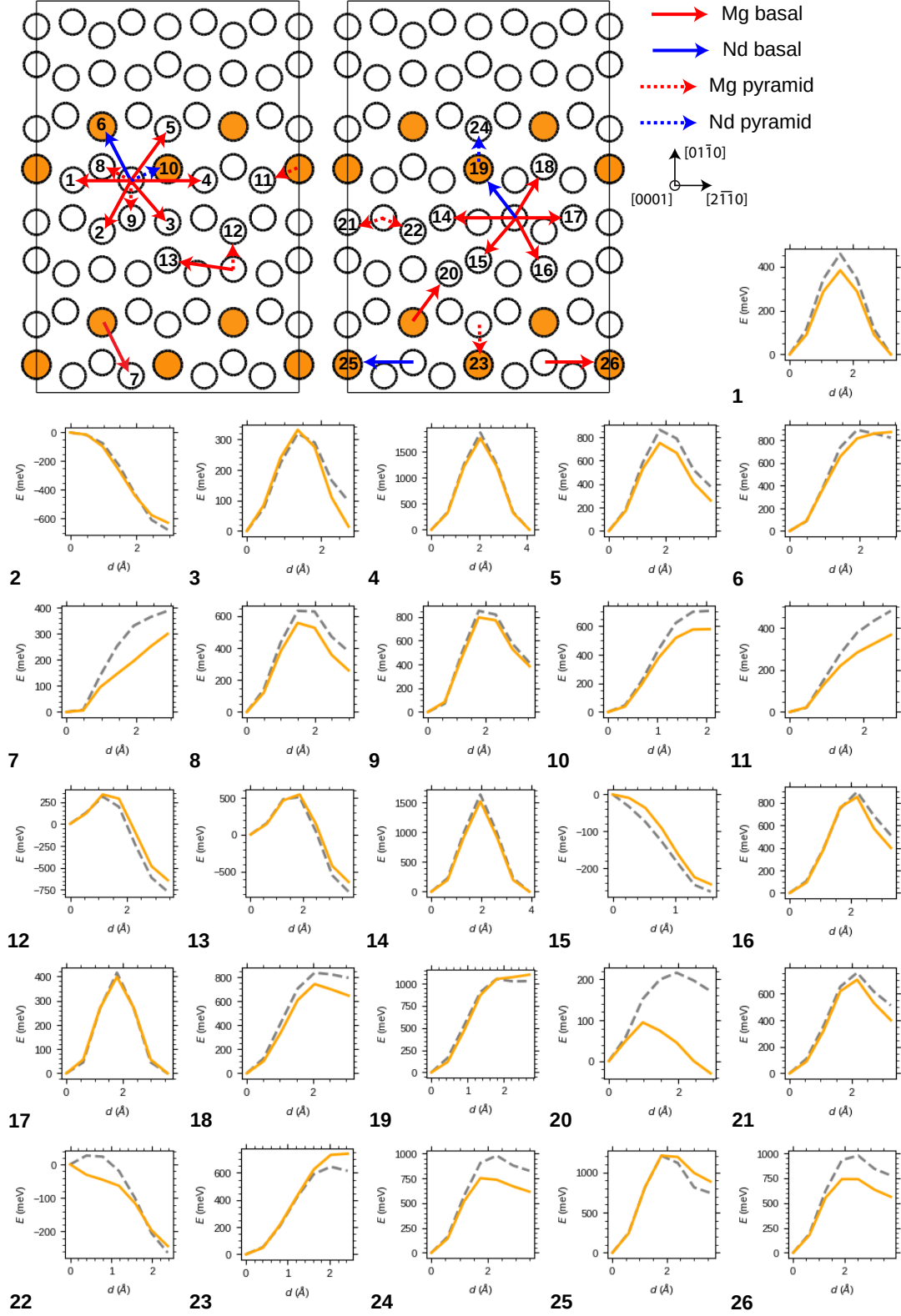


FIG. S14. Energy profiles for 26 symmetry-distinct vacancy hops in β' . In the schematic crystal structure, arrows indicate vacancy jumps from the source (tail) to the numbered target atom (head). Numbers in each energy profile correspond to these target atoms. Energies are referenced to the initial state. The orange solid line is for the MLIP, the gray dashed line is for DFT.

-
- [1] Scott Kirklin, James E Saal, Bryce Meredig, Alex Thompson, Jeff W Doak, Muratahan Aykol, Stephan Rühl, and Chris Wolverton. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *npj Computational Materials*, 1(1):15010, December 2015.
- [2] D. Zagorac, H. Müller, S. Ruehl, J. Zagorac, and S. Rehme. Recent developments in the Inorganic Crystal Structure Database: theoretical crystal structure data and related features. *Journal of Applied Crystallography*, 52(5):918–925, October 2019.
- [3] Anirudh Raju Natarajan, Ellen L.S. Solomon, Brian Puchala, Emmanuelle A. Marquis, and Anton Van Der Ven. On the early stages of precipitation in dilute Mg–Nd alloys. *Acta Materialia*, 108:367–379, April 2016.
- [4] Ki-Hyun Kim and Byeong-Joo Lee. Modified embedded-atom method interatomic potentials for Mg–Nd and Mg–Pb binary systems. *Calphad*, 57:55–61, June 2017.
- [5] Anton Bochkarev, Yury Lysogorskiy, Sarath Menon, Minaam Qamar, Matous Mrovec, and Ralf Drautz. Efficient parametrization of the atomic cluster expansion. *Physical Review Materials*, 6(1):013804, January 2022.
- [6] Eslam Ibrahim, Yury Lysogorskiy, Matous Mrovec, and Ralf Drautz. Atomic Cluster Expansion for a General-Purpose Interatomic Potential of Magnesium, May 2023. arXiv:2305.03577 [cond-mat].
- [7] G. Kresse and J. Hafner. *Ab initio* molecular dynamics for liquid metals. *Physical Review B*, 47(1):558–561, January 1993.
- [8] G. Kresse and J. Hafner. *Ab initio* molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Physical Review B*, 49(20):14251–14269, May 1994.
- [9] G. Kresse and J. Furthmüller. Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, July 1996.
- [10] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B*, 59(3):1758–1775, January 1999.
- [11] Brian Puchala, John C. Thomas, Anirudh Raju Natarajan, Jon Gabriel Goiri, Sesha Sai Behara, Jonas L. Kaufman, and Anton Van Der Ven. CASM — A software package for first-principles based study of multicomponent crystalline solids. *Computational Materials Science*, 217:111897, January 2023.
- [12] Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. In 'T Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, Ray Shan, Mark J. Stevens, Julien Tranchida, Christian Trott, and Steven J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, 271:108171, February 2022.
- [13] Atsushi Togo, Laurent Chaput, Terumasa Tadano, and Isao Tanaka. Implementation strategies in phonopy and phono3py. *Journal of Physics: Condensed Matter*, 35(35):353001, September 2023.
- [14] Sanjeev Krishna Kolli, Anirudh Raju Natarajan, John C. Thomas, Tresa M. Pollock, and Anton Van Der Ven. Discovering hierarchies among intermetallic crystal structures. *Physical Review Materials*, 4(11):113604, November 2020.
- [15] John C. Thomas and Anton Van Der Ven. The exploration of nonlinear elasticity and its efficient parameterization for crystalline materials. *Journal of the Mechanics and Physics of Solids*, 107:76–95, October 2017.
- [16] Richard Tran, Xiang-Guo Li, Joseph H. Montoya, Donald Winston, Kristin Aslaug Persson, and Shyue Ping Ong. Anisotropic work function of elemental crystals. *Surface Science*, 687:48–55, September 2019.
- [17] John C. Thomas, Anirudh Raju Natarajan, and Anton Van Der Ven. Comparing crystal structures with symmetry and geometry. *npj Computational Materials*, 7(1):164, October 2021.
- [18] Sebastiano Vigna. A Weighted Correlation Index for Rankings with Ties. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1166–1176, Florence Italy, May 2015. International World Wide Web Conferences Steering Committee.
- [19] G.B Walker and M Marezio. Lattice parameters and zone overlap in solid solutions of lead in magnesium. *Acta Metallurgica*, 7(12):769–773, December 1959.
- [20] L. J. Slutsky and C. W. Garland. Elastic Constants of Magnesium from 4.2°K to 300°K. *Physical Review*, 107(4):972–976, August 1957.
- [21] M Nishimura, K Kinoshita, Y Akahama, and H Kawamura. Volume compression of Mg and Al to multimegabar pressure. In *Proceedings of the Joint 20th AIRAPT - 43rd EH-PRG Conference on Science and Technology of High Pressure*, Postfach 3640, 76021, Karlsruhe, 2005. Forschungszentrum Karlsruhe GmbH.
- [22] Rasool Ahmad, Binglun Yin, Zhaoxuan Wu, and W.A. Curtin. Designing high ductility in magnesium alloys. *Acta Materialia*, 172:161–184, June 2019.
- [23] Markus Stricker, Binglun Yin, Eleanor Mak, and W. A. Curtin. Machine learning for metallurgy II. A neural-network potential for magnesium. *Physical Review Materials*, 4(10):103602, October 2020.
- [24] P. Tzanetakis, J. Hillairet, and G. Revel. The Formation Energy of Vacancies in Aluminium and Magnesium. *Physica Status Solidi (b)*, 75(2):433–439, 1976.
- [25] Nd Crystal Structure: Datasheet from “PAULING FILE Multinaries Edition – 2022” in SpringerMaterials.
- [26] Jr. K. A. Gschneidner and L. Eyrin. Handbook on the physics and chemistry of rare earths, 1978. ISSN: 0168-1273.
- [27] Hannes Jónsson, Greg Mills, and Karsten W. Jacobsen. Nudged elastic band method for finding minimum energy paths of transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations*, pages 385–404. World Scientific, Singapore, June 1998.
- [28] Yanlin Guo, Bin Liu, Wei Xie, Qun Luo, and Qian Li. Antiphase boundary energy of β series precipitates in Mg–Y–Nd system. *Scripta Materialia*, 193:127–131, March 2021.
- [29] Stephen DeWitt, Ellen L.S. Solomon, Anirudh Raju Natarajan, Vicente Araullo-Peters, Shiva Rudraraju, Larry K. Aagesen, Brian Puchala, Emmanuelle A. Marquis, Anton Van Der Ven, Katsuyo Thornton, and John E. Allison. Misfit-driven β''' precipitate composition and morphology in Mg–Nd alloys. *Acta Materialia*, 136:378–389, September 2017.

- [30] D. Choudhuri, R. Banerjee, and S. G. Srinivasan. Interfacial structures and energetics of the strengthening precipitate phase in creep-resistant Mg-Nd-based alloys. *Scientific Reports*, 7(1):40540, January 2017.
- [31] Jian-Feng Nie. Precipitation and Hardening in Magnesium Alloys. *Metallurgical and Materials Transactions A*, 43(11):3891–3939, November 2012.
- [32] H. Liu, Y. Gao, J.Z. Liu, Y.M. Zhu, Y. Wang, and J.F. Nie. A simulation study of the shape of β' precipitates in Mg–Y and Mg–Gd alloys. *Acta Materialia*, 61(2):453–466, January 2013.
- [33] H. Liu, Y.M. Zhu, N.C. Wilson, and J.F. Nie. On the structure and role of β' in β_1 precipitation in Mg–Nd alloys. *Acta Materialia*, 133:408–426, July 2017.
- [34] Ravi Agarwal and Dallas R. Trinkle. Exact Model of Vacancy-Mediated Solute Transport in Magnesium. *Physical Review Letters*, 118(10):105901, March 2017.