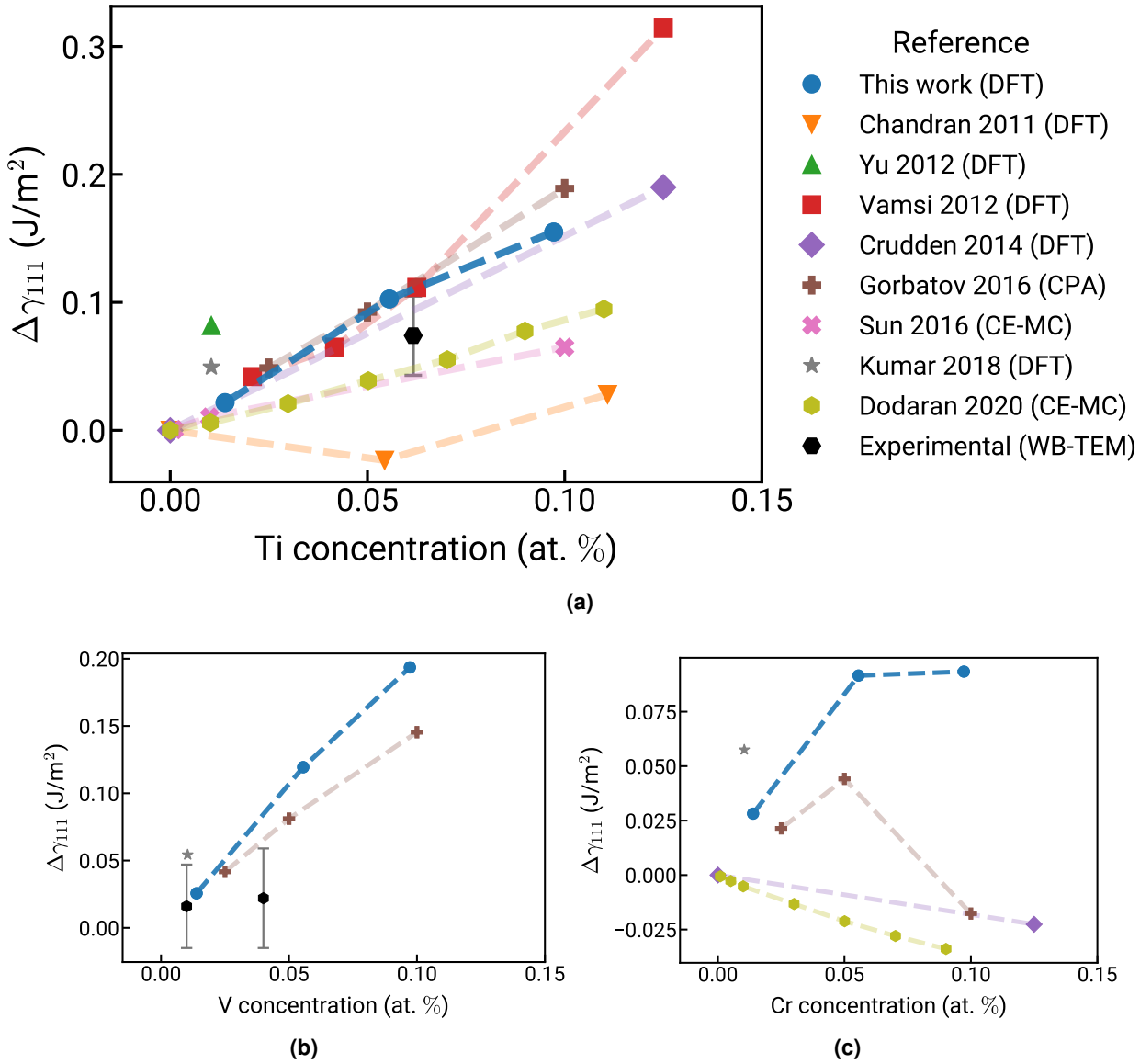
Supplementary Information for Modeling antiphase boundary energies of Ni_3Al-based alloys using automated density functional theory and machine learning

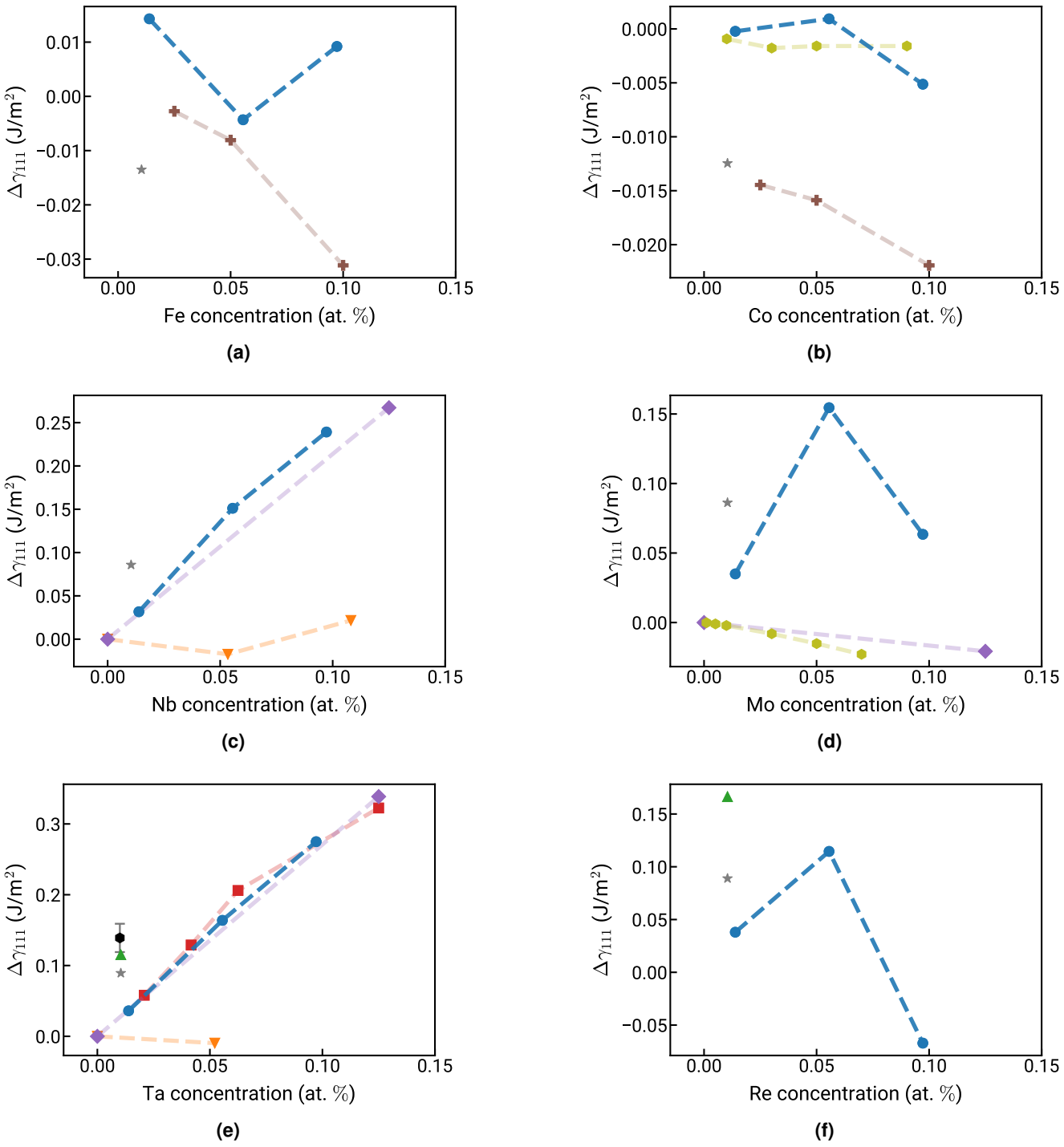
Enze Chen[1,2], Artur Tamm[3], Tao Wang[4], Mario E. Epler[4], Mark Asta[1,2,*], and Timofey Frolov[3,†]

[1] Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA
[2] Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3] Materials Science Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA
[4] Carpenter Technology Corporation, Philadelphia, PA 19103, USA
[*] mdasta@berkeley.edu
[†] frolov2@llnl.gov

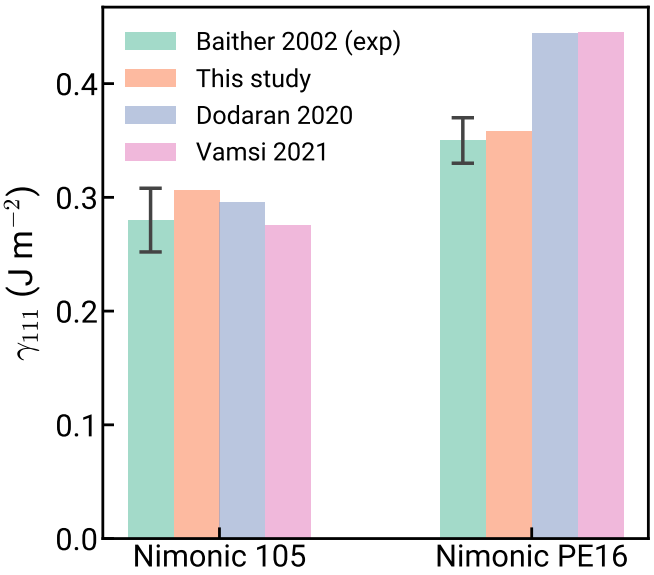
Supplementary Figures



Supplementary Figure 1. Comparing our DFT-calculated APB energies (blue dashed lines) with those in previous studies[1–11]. The change in (111) APB energy is defined as $\Delta\gamma_{111} \overset{\text{def}}{=} \gamma_{111,\text{Ni}_3\text{AlZ}} - \gamma_{111,\text{Ni}_3\text{Al}}$ and plotted as a function of solute concentration. Each panel shows a different solute species and all plots in Supplementary Fig. 1 and 2 share the same legend as panel (a). The method of each previous study is given in parentheses in the legend.



Supplementary Figure 2. Further comparisons between the APB energies obtained in this study and those in the literature. See Supplementary Fig. 1a for the legend.



Supplementary Figure 3. We compare the performance of our ML model against other models in the literature in predicting the APB energy of two commercial superalloys, Nimonic 105 and Nimonic PE16. The green bars are experimental measurements from Baither et al.[12], who report values for γ_{111} between 0.252–0.308 J m^{-2} and 0.330–0.370 J m^{-2} for the two alloys, respectively. The orange, lavender, and pink bars correspond to the predicted APB energies from our ML model, the model by Dodaran et al.[8], and the model by Vamsi and Karthikeyan[13], respectively.

Composition (at.%)	γ_{111} (J m^{-2})
Ni−19.4 Al	0.122(17)
Ni−23.6 Al	0.176(16)
Ni−25.0 Al	0.188
Ni−26.4 Al	0.198(12)
Ni−30.6 Al	0.265(41)

Supplementary Table 1. DFT-calculated values for γ_{111} for on- and off-stoichiometry binary Ni_3Al alloys. The uncertainty in the final two decimal places is the standard deviation of five different DFT calculations of γ_{111} for each composition. γ_{111} increases approximately linearly with increasing Al content near the stoichiometric Ni_3Al composition.

Supplementary Methods

DFT pseudopotentials

We use the following pseudopotentials for our DFT calculations using VASP 5.4[14–17]:

- Semi-core *s* states: Sc, Y, Zr, Ba, Sr
- Semi-core *p* states: Ti, V, Cr, Mn, Nb, Mo, Tc, Hf, Ta, Re, Ca
- Semi-core *d* states: Ga, Ge, Sn, Tl, Pb
- All other elements use the standard pseudopotentials supplied with VASP 5.4.

We initialize magnetic moments (in units of μ_B) as follows:

- 1.0: Fe, Co, Mn, Cr
- 0.5: Ni
- 0.0: Everything else

All other DFT settings may be found in the Methods section III.B in the main text.

The DFT convergence tests presented in Fig. 4a and Fig. 4b use a smaller $1 \times 1 \times 3$ $L1_2$ orthogonal cell (9 layers) than that used in the study. Thus the **k**-point mesh is chosen ($4 \times 7 \times 2$ here; $2 \times 3 \times 2$ in the manuscript) to correspond to the same density in reciprocal space (5000 **k**-points per reciprocal atom).

DFT convergence tests

As a tighter check for convergence, we perform a few higher-accuracy calculations for Ni_3Al and 9.72 at.% Nb with the following stricter settings:

- Energy cutoff of 600 eV.
- Electronic convergence threshold of 1×10^{-7} eV.
- Ionic convergence threshold of 1×10^{-3} eV/atom (norm of the forces).
- For Ni_3Al we also used 15 layers in the supercell.

We then compute the following quantifier of systematic error:

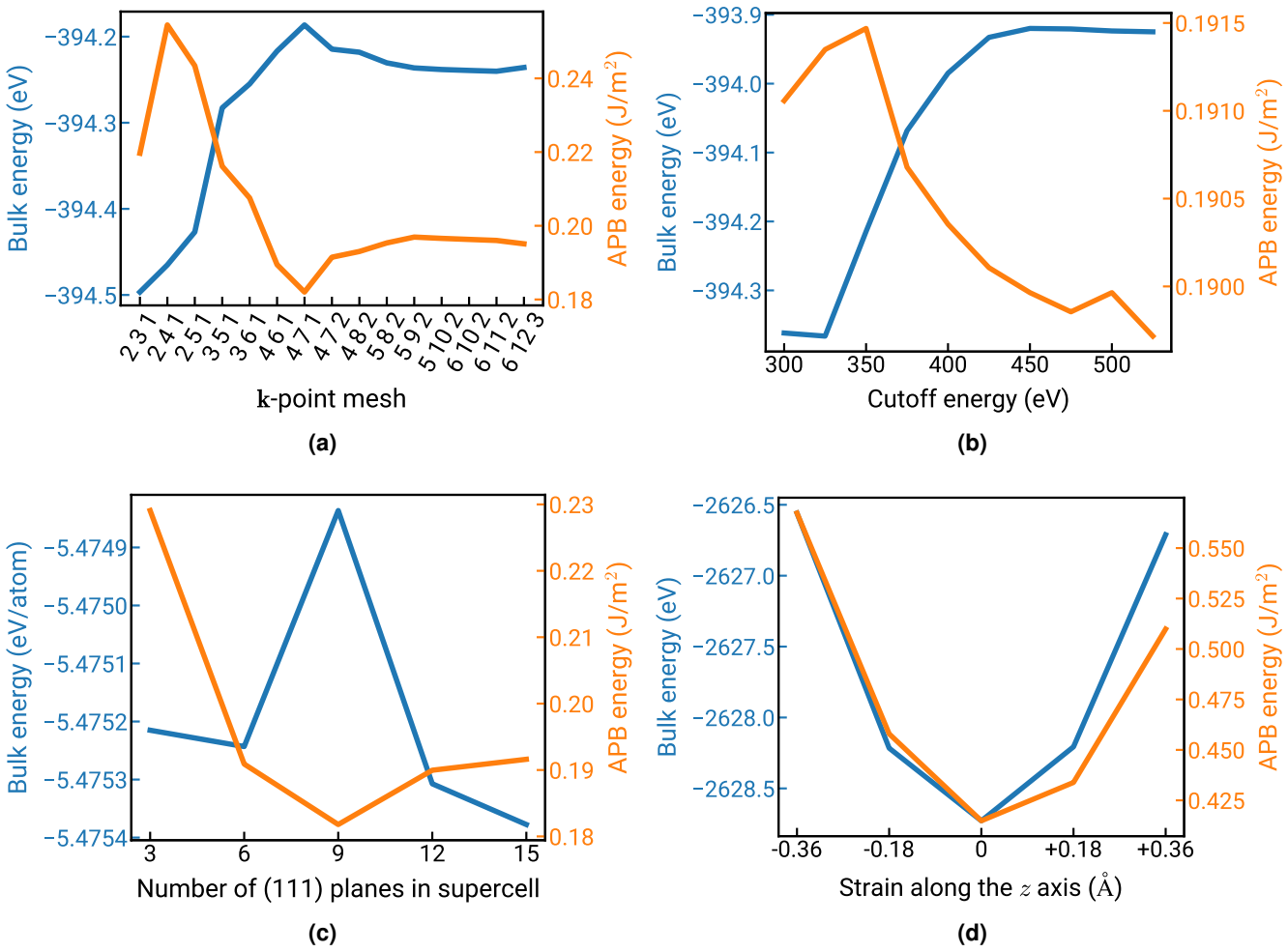
$$\%\ \text{error} = \frac{|\gamma_{111}^{\text{strict}} - \gamma_{111}^{\text{nominal}}|}{\gamma_{111}^{\text{strict}}} \tag{1}$$

For the Ni_3Al system, the higher-accuracy APB energy is $0.185\,\mathrm{J\,m^{-2}}$. This means:

$$\%\ \text{error} = \frac{|0.185 - 0.188|}{0.185} \times 100\,\% = 1.62\,\%$$

For the 9.72 at.% Nb system, the higher-accuracy APB energy is $0.425\,\mathrm{J\,m^{-2}}$. This means:

$$\%\ \text{error} = \frac{|0.425 - 0.428|}{0.425} \times 100\,\% = 0.7\,\%$$



Supplementary Figure 4. Convergence of the total energy in the bulk $L1_2$ structure of Ni_3Al in a $1 \times 1 \times 3$ supercell oriented with (111) planes normal to the z-axis. Convergence of γ_{111} with respect to (a) the **k**-point mesh, (b) the plane wave cutoff energy, and (c) the number of (111) planes separating the APB in Ni_3Al. (d) Convergence of the two quantities with respect to the cell height (measured as strain in Å following volume relaxation) in a 288-atom cell with 9.72 at.% Nb.

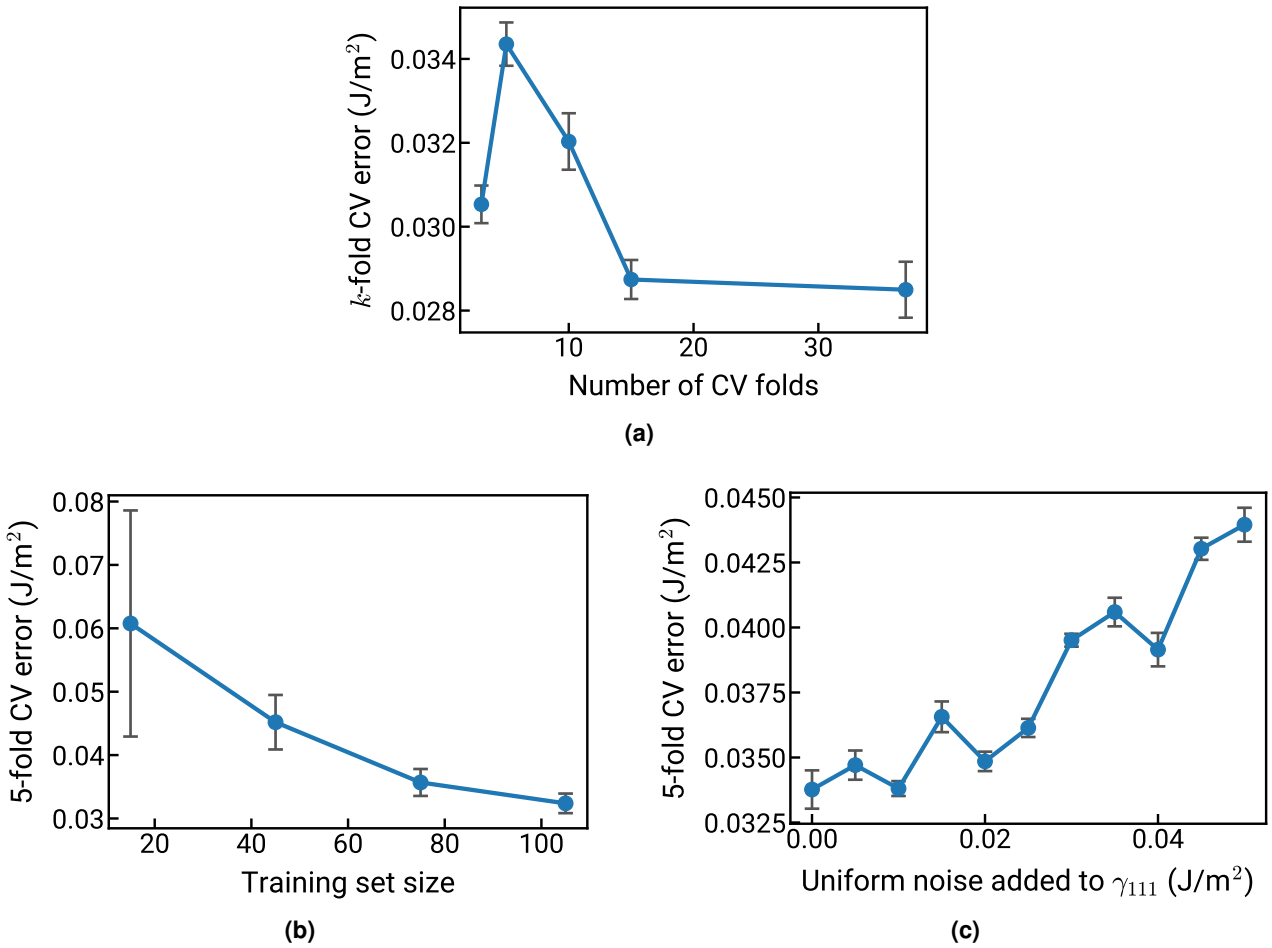
Machine learning methods

Feature	Method
Ordering energy (E_{ord})	DFT[14–17]
Change in lattice parameter (Δa) and density	DFT
Sublattice preference	PyDII[18,19]
Pettifor number[20]	mendeleev package[21]
Magpie elemental properties[22]	matminer package[23]
Pymatgen elemental properties[24]	matminer package
Intermetallic formation enthalpies[25]	matminer package
Mixing thermochemistry and size mismatch[26]	matminer package
Fraction of magnetic transition metals[27]	matminer package

Supplementary Table 2. All of the features that we initially generate for our ML models and their corresponding citations.

While including the ordering energy (E_{ord}) as a feature for modeling the APB energy leverages our domain knowledge[4,5], there may be interest in an ML model without E_{ord}, particularly for multicomponent compositions where this information may not be readily available (see Section I.D in the main text) or vary in estimated values depending on the database used. We use the same procedure outlined in the manuscript to train an ML model without E_{ord} and we obtain a 5-fold grouped CV error of $0.035\,J\,m^{-2}$ (compared to $0.033\,J\,m^{-2}$ with the feature included). Therefore, one might not *need* to calculate E_{ord} in order to predict the APB energies in the model systems used in this study, and we show in Section I.D how a model trained without E_{ord} can extrapolate to multicomponent systems. However, we feel it is an important contribution to include, when available, for several reasons:

- E_{ord} is a physically-motivated correlation that has been proposed in the literature, which we discuss in the Introduction (pg. 4-5). As discussed there, we hope to use ML and data-driven techniques to build on top of this domain knowledge.
- E_{ord} remains after recursive feature elimination, which is a data-driven approach to feature selection as opposed to physical intuition.
- When included, E_{ord} is the fourth-most important feature using the random forest model's feature importance ranking.
- The performance is still marginally better with E_{ord} included. In addition, the good performance achieved with the other, auto-generated features helps support one of our central conclusions that informatics approaches can outperform methods that rely on physical intuition or analytical models alone.



Supplementary Figure 5. (a) We quantify how the k-fold grouped CV error changes with respect to the choice of k. The right-most point at $k = 37$ is equivalent to a leave-one-solute-out approach. Error bars represent one standard deviation above and below the mean of 30 iterations of CV. (b) We quantify how the 5-fold grouped CV error changes with respect to the size of the training set given the small size and sparsity of our data. (c) We quantify how the 5-fold grouped CV error changes with respect to uniform random noise (up to $0.05\,\mathrm{J\,m^{-2}}$) added to the APB energies in the training set to further analyze the effects of data quality.

Supplementary Note 1

One consideration for the results presented here is whether the $L1_2$-ordered structure (γ' phase) is suitable for all of the alloy compositions. While solute elements generally exhibit preferential partitioning between the γ and γ' phases in real superalloys, we do not make that distinction in this work and assume that, in the dilute regime, our results will still capture the dominant chemical contributions to the APB energy from solutes in the γ' phase. For 14 of the ternary solutes (Co, Cr, Cu, Fe, Hf, Mn, Mo, Nb, Pd, Pt, Ta, Ti, V, W), we perform additional CALPHAD calculations at 1000 K using the TCNI9 database to confirm that the γ' phase is stable for those alloys at low (1.39 at.%), medium (5.56 at.%), and high (9.72 at.%) concentrations, and only a few ternary systems (Hf, Mn, Mo, V, W) show competing phases at the highest concentration. The CALPHAD data are also qualitatively consistent with our PyDII calculated sublattice preference for those 14 elements. These CALPHAD results help support the DFT data and validate the use of the latter as training data for ML.

Supplementary References

1. Chandran, M. & Sondhi, S. K. First-principle calculation of APB energy in Ni-based binary and ternary alloys. *Model. Simul. Mater. Sci. Eng.* **19**, 025008 (2011).

2. Yu, X.-X. & Wang, C.-Y. Effect of alloying element on dislocation cross-slip in γ'-Ni_3Al: A first-principles study. *Philos. Mag.* **92**, 4028–4039 (2012).

3. Vamsi, K. V. & Karthikeyan, S. Effect of off-stoichiometry and ternary additions on planar fault energies in Ni_3Al. In *Proceedings of the Twelfth International Symposium on Superalloys*, 521–530 (TMS, Champion, PA, 2012).

4. Crudden, D. J., Mottura, A., Warnken, N., Raeisinia, B. & Reed, R. C. Modelling of the influence of alloy composition on flow stress in high-strength nickel-based superalloys. *Acta Mater.* **75**, 356–370 (2014).

5. Gorbatov, O. I. *et al.* Effect of composition on antiphase boundary energy in Ni_3Al-based alloys: *Ab initio* calculations. *Phys. Rev. B* **93**, 224106 (2016).

6. Sun, R. & van de Walle, A. Automating impurity-enhanced antiphase boundary energy calculations from *ab initio* Monte Carlo. *Calphad* **53**, 20–24 (2016).

7. Kumar, K., Sankarasubramanian, R. & Waghmare, U. V. Tuning planar fault energies of Ni_3Al with substitutional alloying: First-principles description for guiding rational alloy design. *Scripta Mater.* **142**, 74–78 (2018).

8. Dodaran, M. *et al.* Effect of alloying elements on the γ' antiphase boundary energy in Ni-base superalloys. *Intermetallics* **117**, 106670 (2020).

9. Korner, A. Weak-beam study of superlattice dislocations moving on cube planes in Ni_3(Al, Ti) deformed at room temperature. *Philos. Mag. A* **58**, 507–522 (1988).

10. Baluc, N., Schäublin, R. & Hemker, K. J. Methods for determining precise values of antiphase boundary energies in Ni_3Al. *Philos. Mag. Lett.* **64**, 327–334 (1991).

11. Dimiduk, D. M., Thompson, A. W. & Williams, J. C. The compositional dependence of antiphase-boundary energies and the mechanism of anomalous flow in Ni_3Al alloys. *Philos. Mag. A* **67**, 675–698 (1993).

12. Baither, D., Rentenberger, C., Karnthaler, H. P. & Nembach, E. Three alternative experimental methods to determine the antiphase-boundary energies of the γ' precipitates in superalloys. *Philos. Mag. A* **82**, 1795–1805 (2002).

13. Vamsi, K. V. & Karthikeyan, S. Modeling APB energies in multicomponent Ni-base superalloys. *Intermetallics* **132**, 107124 (2021).

14. Kresse, G. & Hafner, J. *Ab initio* molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).

15. Kresse, G. & Hafner, J. *Ab initio* molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251–14269 (1994).

16. Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).

17. Kresse, G. & Furthmüller, J. Efficiency of *ab initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).

18. Ding, H. *et al.* PyDII: A Python framework for computing equilibrium intrinsic point defect concentrations and extrinsic solute site preferences in intermetallic compounds. *Comput. Phys. Commun.* **193**, 118–123 (2015).

19. Python framework for Defects In Intermetallics. *GitHub* https://github.com/pydii/pydii (2015).

20. Pettifor, D. G. A chemical scale for crystal-structure maps. *Solid State Commun.* **51**, 31–34 (1984).

21. Mentel, Ł. mendeleev - A Python resource for properties of chemical elements, ions and isotopes. *GitHub* https://github.com/lmmentel/mendeleev (2014).

22. Ward, L., Agrawal, A., Choudhary, A. & Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Comput. Mater.* **2**, 16028 (2016).

23. Ward, L. *et al.* Matminer: An open source toolkit for materials data mining. *Comput. Mater. Sci.* **152**, 60–69 (2018).

24. Ong, S. P. *et al.* Python Materials Genomics (pymatgen): A robust, open-source Python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).

25. Miedema, A. R. A simple model for alloys. I. rules for the alloying behaviour of transition metals. *Philips Tech. Rev.* **33**, 149–160 (1973).

26. Yang, X. & Zhang, Y. Prediction of high-entropy stabilized solid-solution in multi-component alloys. *Mater. Chem. Phys.* **132**, 233–238 (2012).

27. Deml, A. M., O'Hayre, R., Wolverton, C. & Stevanović, V. Predicting density functional theory total energies and enthalpies of formation of metal-nonmetal compounds by linear regression. *Phys. Rev. B* **93**, 085142 (2016).