

MATERIALS SCIENCE

Unfolding adsorption on metal nanoparticles: Connecting stability with catalysis

James Dean, Michael G. Taylor, Giannis Mpourmpakis*

Metal nanoparticles have received substantial attention in the past decades for their applications in numerous areas, including medicine, catalysis, energy, and the environment. Despite these applications, the fundamentals of adsorption on nanoparticle surfaces as a function of nanoparticle size, shape, metal composition, and type of adsorbate are yet to be found. Herein, we introduce the first universal adsorption model that accounts for detailed nanoparticle structural characteristics, metal composition, and different adsorbates by combining first principles calculations with machine learning. Our model fits a large number of data and accurately predicts adsorption trends on nanoparticles (both monometallic and alloy) that have not been previously seen. In addition to its application power, the model is simple and uses tabulated and rapidly calculated data for metals and adsorbates. We connect adsorption with stability behavior of nanoparticles, thus advancing the design of optimal nanoparticles for applications of interest, such as catalysis.

INTRODUCTION

Metal nanoparticles (NPs) find tremendous applications in catalysis, ranging from producing fuels and chemicals (1) to transforming solar to chemical energy (2). However, the stability of the NPs and their catalytic functionality commonly show opposite trends, with very active catalysts typically dying after some cycles of operation and needing energy input to regenerate. An example of nanocatalyst death is on structure-sensitive reactions, such as the carbon monoxide (CO) oxidation on Au (3–6), where low-coordinated sites on the surface of the NPs are required, giving rise to activity at very small NP sizes (6). In these reactions, a sintering of the small NPs (7, 8) on the surface of the support during catalytic operation (e.g., elevated temperatures) results to larger average NP size, decreasing the number of low-coordinated sites on the surface of the NPs, thus altering catalytic behavior. As a result, the NPs grow in size, become more stable in terms of cohesive energy (CE) (9) (increase the number of metal atoms and bonds per particle), but lose catalytic activity.

A key feature determining to a large extent the catalytic functionality of metals is the adsorption strength of the various species on the surface of the catalyst. The Sabatier principle (10) postulated more than 100 years ago states that an active catalyst should bind the adsorbates with a binding strength that is not very strong neither too weak. In the first case, the strongly adsorbed species poison the catalyst surface, whereas in the second case, the weakly bound reactants prefer to desorb rather than reside on the catalyst surface. As a result, in an intermediate case, the reactants meet each other and react on the catalyst surface. Nowadays, with the great advances in computational power and in theoretical chemistry methods, we can simulate catalytic behavior on metal catalysts with great accuracy and guide experiments (11–13), avoiding trial and error experimentation in the laboratory. This is because the adsorption energy of the species that participate (reactants or products) often correlates with the activation energy of the reaction. These well-known Bronsted-Evans-Polanyi relationships, developed more than half a century ago, inherently connect thermodynamics with kinetics, highlighting once again the importance of adsorption in catalysis (14, 15). In addition, the adsorption behavior of a functional group (e.g., CH₃)

on a specific site of the catalyst scales with the adsorption of a similar chemical group (e.g., CH₂) on the same site of the catalyst of different metals (16). Together, one can create reactivity plots, so-called volcano plots, relating catalytic activity with a single descriptor being the binding energy (BE) of key intermediates (17, 18). As a result, the computational efforts have vastly focused on screening different metal catalysts for the discovery of the “magic” BE of chemical species on the surface of the catalyst that will result to the design of very active catalysts. To this end, the d-band model developed in the 1990s (19) revealed that the adsorption behavior on the surface of a metal catalyst depends on the average energy of the d-electron band of the catalyst. This electronic descriptor of the catalyst rationalizes the energy shifts in the adsorption strength of chemical species on the catalyst surface that, in turn, can regulate catalytic behavior (17).

Despite all these advances in computational catalysis, still the *in silico* design of catalytically active materials is far from being realized. One of the reasons is that the catalyst design efforts often neglect the stability of the catalyst, leading to predictions of materials that are either difficult to be synthesized or unstable under catalytic operation (e.g., catalyst reconstruction under a chemically reactive environment). Another reason is that NPs are catalysts with a high degree of site heterogeneity on their surface exhibiting different planes (e.g., 111, 100 etc.), corners, and edges, which lead to a heterogeneity on their surface properties, such as adsorption and catalysis. In an effort to capture this site-specific adsorption response, adsorption models have been developed in literature relating the BE of the adsorbates with surface characteristics of the NPs, such as the coordination numbers (CNs) (6, 20–22). One common feature of these models is that just the CN of the site cannot adequately describe the BE variation, and secondary descriptors are needed, such as the curvature (20), the electronic properties [orbital information (22)], and the neighbors of the adsorption site [generalized CNs (GCNs) (21)].

In this work, we apply density functional theory (DFT) and machine learning techniques to derive a simple physics-based model that can accurately capture the variation of the adsorption energy as a function of the local adsorption site environment on the NP surface and the type of metal NP. The power of this model is that, although it captures catalyst properties, it is based solely on structural information of the catalyst and tabulated values of metals, avoiding any calculations of its electronic structure. Therefore, it is general and applicable on any metal

Copyright © 2019
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

Department of Chemical Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA.

*Corresponding author. Email: gmpourmp@pitt.edu

nanostructure. This model enables a direct relation between the adsorption behavior on the NP catalyst and the stability of the catalyst. As a result, we demonstrate the efficient screening of NP morphologies of different metals based on their adsorption response and thermodynamic stability. The screening of materials for their adsorption response is important not only to the design of catalysts but also to the numerous fields and applications, which depend on adsorption phenomena, such as sensors, separations and purifications, mineral extraction, corrosion prevention, printing, cosmetics, and medicine (23). Hence, although we focus on species that are most relevant to the catalyst design, the development of a physics-based, generalizable model to predict adsorption has the potential for wide-ranging impact.

RESULTS

We expect several factors to contribute to the adsorption interaction between a site on a monometallic NP and an adsorbate based on prior literature (6, 20). We hypothesize that the most important factors are (i) the stability of the adsorbate in the gas phase (24), (ii) the tendency of the metal and adsorbate to interact and form chemical bonds, (iii) the stability of the bare NP (6), and (iv) the stability of the binding site (20). We use stability here to express the thermodynamic strength of the bonds that are on NPs of different sizes, shapes, and composition or of adsorbates at different metals and sites on the NPs. As a result, we refer to energy differences (thermodynamics) arising from different bond formations (e.g., metal-metal and metal-adsorbate). The rationale behind our hypothesis can be found in the energetic terms used in the calculation of a BE in DFT, which is the difference between the electronic energy of the adsorbate-metal complex ($E_{\text{Ads-M}}$) and that of the adsorbate (E_{Ads}) and the metal NP or surface (E_{M}). For instance, E_{Ads} is related to the stability of the adsorbate, E_{NP} to the stability of the NP, and $E_{\text{Ads-M}}$ to the strength of adsorbate-metal interactions and the local energy landscape of the adsorption site, i.e., the stability of the site. Using these physical descriptors, we generally represent the BE as

$$E_{\text{binding,true}} = \hat{f}(\text{Stab}_{\text{Ads}}, \text{Stab}_{\text{NP}}, \text{Int}_{\text{Ads-M}}, \text{Stab}_{\text{Site}}) + \epsilon \quad (1)$$

where $E_{\text{binding,true}}$ refers to the true BE of the adsorbate to a specific site on an NP. In this work, we take the true binding interaction to be the DFT BE (Eq. 5). Stab_{Ads} , Stab_{NP} , and $\text{Stab}_{\text{Site}}$ refer to the stability descriptors for the adsorbate, NP, and adsorption site, respectively. Last, $\text{Int}_{\text{Ads-M}}$ refers to a descriptor for the strength of adsorbate-metal interactions. Because it is unreasonable to expect a simple linear model to capture all the physics of adsorption, we acknowledge this limitation with an irreducible error term, ϵ . In this work, we consider our chosen physical descriptors (Stab_{Ads} , Stab_{NP} , $\text{Int}_{\text{Ads-M}}$, and $\text{Stab}_{\text{Site}}$) to all be in units of energy, thus forming a first-order polynomial relation with the BE ($E_{\text{binding,true}}$). As a consequence, we do not investigate cross-terms between these descriptors.

Defining a local CE

Intimately related to the stability of the metal and the strength of metal-metal bonds is the CE of the bulk metal, which is defined as the amount of energy required for the atoms in the system to achieve infinite separation. CE is captured in NPs by the bond-centric (BC) model of Yan *et al.* (9), which asserts CE as the sum of every metal-metal bond energy in the NP. In this work, we apply the same concept to describe

the stability of binding sites, with the justification that chemically unsaturated sites (i.e., less stable because of having fewer metal-metal bonds) tend to bind adsorbates stronger. To this end, we introduce the local CE descriptor, CE_{local} , defined by applying the BC model only to the metal atom participating directly in the binding interaction. In other words, we take the summation of BC model-approximated bond energies connecting with the metal atom in the binding site, which is defined as

$$\begin{aligned} \text{Stab}_{\text{site}} &\approx \text{CE}_{\text{local}} = \sum_{i=1}^{\text{CN}} E_{\text{bond A-B}_i} \\ &= \sum_{i=1}^{\text{CN}} \left(\frac{\gamma_{\text{A}} \text{CE}_{\text{bulk,A}}}{\text{CN}_{\text{A}}} \sqrt{\frac{\text{CN}_{\text{A}}}{\text{CN}_{\text{bulk,A}}}} + \frac{\gamma_{\text{B}_i} \text{CE}_{\text{bulk,B}_i}}{\text{CN}_{\text{B}_i}} \sqrt{\frac{\text{CN}_{\text{B}_i}}{\text{CN}_{\text{bulk,B}_i}}} \right) \end{aligned} \quad (2)$$

In Eq. 2, A-B_i refers to the atoms in the neighborhood of the binding site. Atom A refers to the metal atom directly participating in the top-site adsorption. Atom B_i is one of the metal atoms directly bound to atom A. The summation index i refers to the bonds between atom A and atoms B_i and ranges between 1 and the CN of atom A. As an example, an adsorbate bonding to a site with CN 6 would have six bonds, meaning the summation would have six terms (bonds A-B₁, A-B₂, ..., A-B₆). Each bond energy E_{bond} can be approximated with the BC model, which we directly substitute inside the summation parentheses in Eq. 2. The other terms in the equation can be described as follows: $\text{CE}_{\text{bulk,X}}$ is the bulk CE of atom X, CN_{X} is the CN of atom X, $\text{CN}_{\text{bulk,X}}$ is the bulk CN of atom X, and γ_{X} is the BC model γ coefficient of atom X (9). This definition of $\text{Stab}_{\text{site}}$ (Eq. 2) allows us to capture not only the CN and metal identity of the atom directly bound to the adsorbate but also the electronic effects that come from the geometry of the local site, such as stronger bonds resulting from undercoordinated atoms adjacent to the binding atom. This formulation of CE_{local} additionally differentiates between atoms of the same CN and different metal types.

Using this description of the local geometry, we can focus on its ability to describe the binding of a single adsorbate-metal pair. In Fig. 1A, we plot the DFT-calculated top-site BE of CO to a 172-atom Au cube and a 147-atom Au cuboctahedron/icosahedron. We note a strongly inverse relationship between CE_{local} and BE. In addition, we visualize both the BE (Fig. 1, A to D) and CE_{local} (Fig. 1, E to G) that the different surface sites on the NPs exhibit. Overall, Fig. 1 demonstrates that the strongest adsorption sites are the ones exhibiting the weakest local cohesion.

Additional adsorption descriptors

Other descriptors chosen for the NP binding were (i) the CE of the entire NP, (ii) the metal atom binding to the adsorbate, and (iii) a term involving the ionization potential (IP) and electron affinity (EA) of the adsorbate. The total NP CE was chosen because we expect the stability (or lack thereof) of a given NP to play cumulatively a role in its ability to bind a molecule (25). The gas-phase BE between a single metal atom and the adsorbate was chosen because we expect to provide a good (and fast to calculate) descriptor of the tendency of the adsorbate to bind a specific metal. Last, the negative average of the IP and EA, which we call "IPEA," was specifically chosen because it has been shown to be a first-order finite difference approximation of the adsorbate's chemical potential within hard-soft acid base theory (26). In addition to this, it is also the negative of Mulliken's definition of electronegativity (26). The values of each of these descriptors can be found in tables S1 to S3, along with more detailed justifications. Notably, the adsorbate-metal atom BE

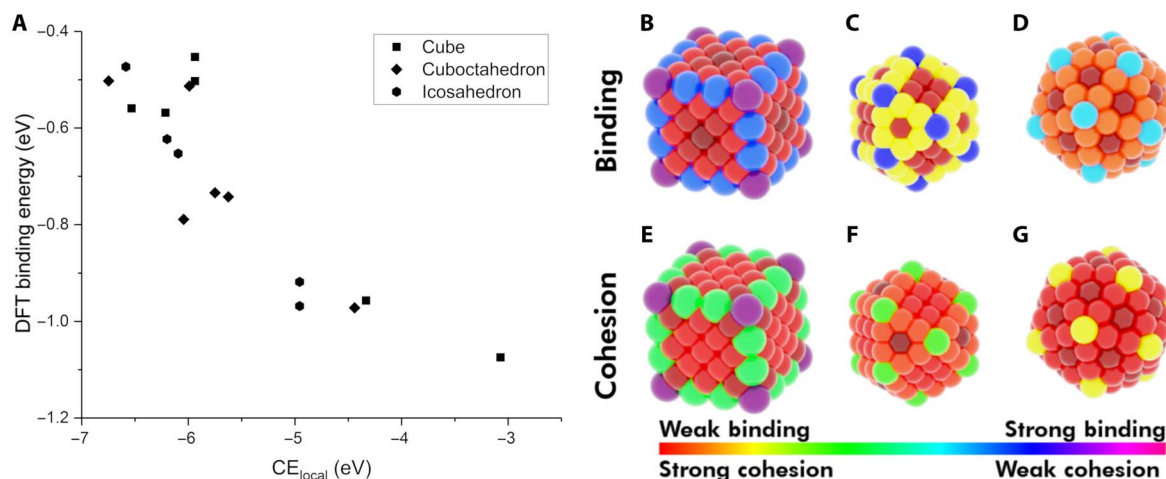


Fig. 1. Demonstration of CE_{local} as a descriptor for adsorption energy. (A) The BE of CO on various sites of Au NPs as a function of CE_{local} : 172-atom cube (rectangles), 147-atom icosahedron (hexagons), and 147-atom cuboctahedron (rhombus). Heat map of different sites on the NPs with respect to their BE of CO (B to D) and to their CE_{local} (E to G). The color scheme follows the range of strongest CO binding to weakest CE_{local} (violet) and of weakest binding to strongest CE_{local} (red).

is the only DFT calculation required to parameterize the model (once it is trained), which is computationally inexpensive to perform. Furthermore, all other quantities (CE_{local} , CE_{NP} , and IPEA) can be rapidly determined through a combination of simple algebra and tabulated physical properties.

Adsorption on monometallic NPs and slabs

Following our choice of chemical descriptors, we conducted linear regression using the Caret package (27) as implemented in R (28), performing ordinary least squares (OLS) regression in conjunction with 10-fold cross-validation for Eq. 3 (which is the first-order polynomial equation we define in our discussion of Eq. 1)

$$E_{\text{bind,model}} = a + b \times CE_{\text{local}} + c \times CE_{\text{NP}} + d \times \text{IPEA} + e \times \text{MADs} \quad (3)$$

where $E_{\text{bind,model}}$ is the model's predicted BE and a , b , c , d , and e are constants. CE_{local} is the local site's cohesion. CE_{NP} is the CE of the whole NP. IPEA is the negative average of the IP and EA. MADs is the Metal-Adsorbate BE.

Using this model, we perform an OLS regression using as a training set the top-site adsorption of three different adsorbates [i.e. methyl radical (CH_3), CO, and hydroxyl radical (OH)] onto three different metals (Cu, Ag, and Au) with five different NP morphologies (172-atom cube, 55- and 147-atom icosahedron, and 55- and 147-atom cuboctahedron), as shown in Fig. 2A. Although we only focus on adsorption on top sites (fig. S4, F to I), we note that previous work (16) has shown that the top-site BE of various adsorbates (including CH_3 and OH) correlates with other surface site BEs (e.g., bridge and hollow sites) over different metals (including Cu/Ag/Au). Regression statistics can be found on Table 1 [case (i)].

In the final functional form of the model, the coefficients regressed make excellent physical sense (Table 1). A negative sign on the CE_{local} coefficient suggests that as the cohesion of the local site increases (becomes more negative), its binding affinity to the adsorbates decreases. Similarly, a positive sign on IPEA, which is related to the adsorbate's chemical potential, denotes that a less stable adsorbate

(higher chemical potential) will have a higher tendency to bind to a metal. Last, the direct correlation with MADs, the BE between the adsorbate and a single metal atom in the gas phase, is also very intuitive as it describes the intrinsic tendency of the metal to bind the adsorbate, with a more negative MADs indicating a stronger binding interaction.

For this model [Table 1, case (i)], we observed a cross-validated R^2 of 0.936 and a cross-validated root mean square error (RMSE) of 0.179 eV. Notably, here is the high P value and error associated with the CE of the NP. As a result, we conducted a second regression where we did not include CE_{NP} , but instead we only included the remaining three regressors and an intercept. This model is plotted in Fig. 2B, and its regression information can be found in the middle of Table 1. We see that the model [Table 1, case (ii)] fits nearly as well, with an overall model R^2 of 0.933 and a cross-validated RMSE of 0.179 eV. If we additionally plot the BE as a function of NP cohesion (fig. S1), then we can see that, at least in the case of our chosen adsorbate-NP combinations, it does not provide an adequate descriptor for the binding interaction. Because the inclusion of the NP CE offers no improvement over a less complex model, we decide to ignore this term and use the form outlined in Eq. 4 for the remainder of our investigation

$$E_{\text{bind,model}} = a + b \times CE_{\text{local}} + c \times \text{IPEA} + d \times \text{MADs} \quad (4)$$

The P value, cross-validated RMSE, and cross-validated R^2 all indicate that our model is generalizable to other metal-adsorbate systems. In addition, the remaining degrees of freedom (DOF), calculated as the dimensionality of the dataset minus the dimensionality of the regression equation, for every model in Table 1, are large, indicating that overfitting is unlikely. Overall, this shows that our model's predictions correlate well with the DFT-calculated values. This is important because we ultimately wish to provide a computationally efficient framework for predicting DFT adsorption energies. The error metrics, RMSE, mean absolute error (MAE), and R^2 offer a picture of how likely the model is to fail or succeed compared to DFT. The cross-validated RMSE and MAE have values of 0.179 and 0.144 eV, respectively, which indicates that, on average, the model can be expected to be within 0.144 to 0.179 eV of the true DFT-calculated value. Last, the good R^2 values of

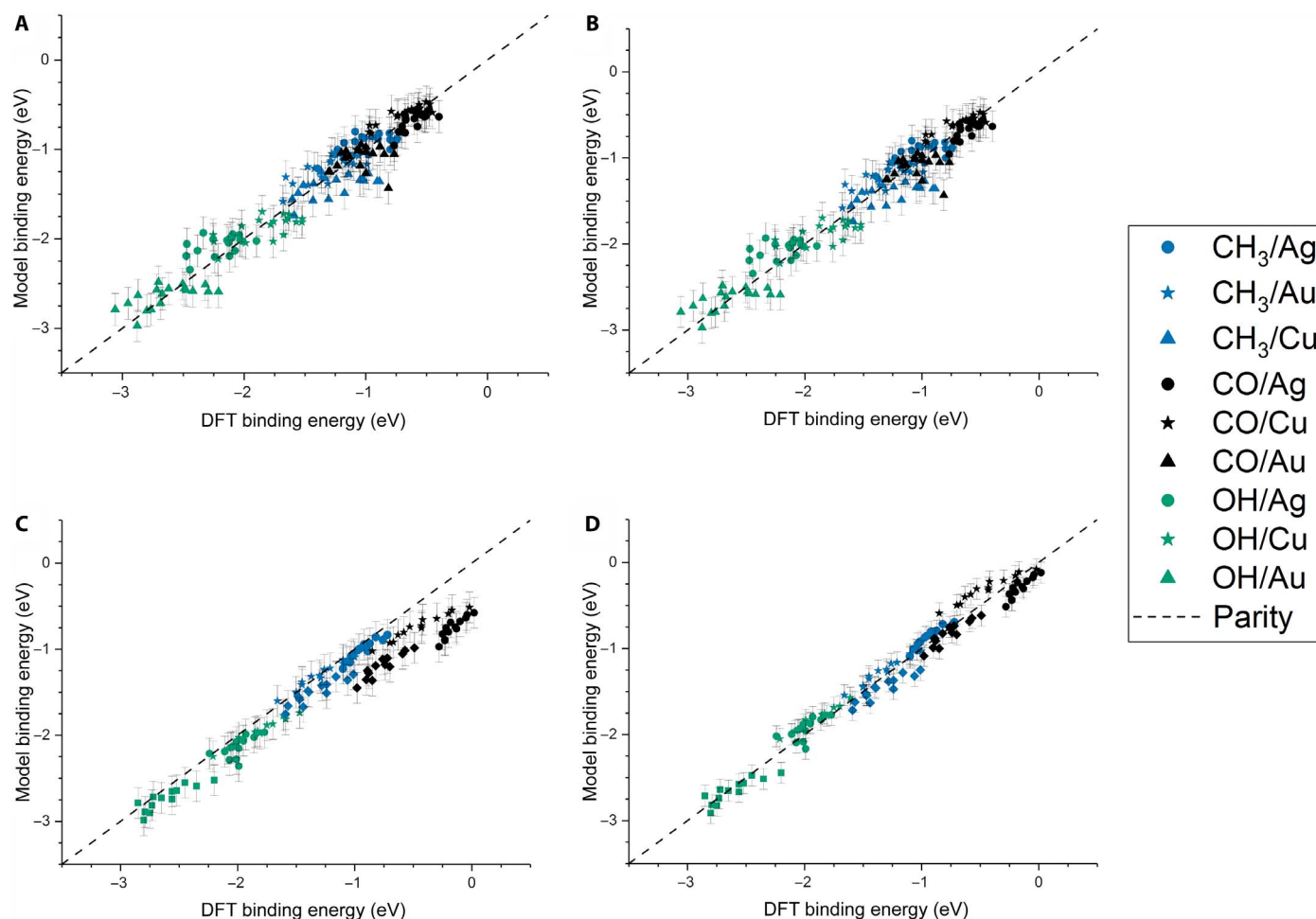


Fig. 2. Parity plot of the model-predicted BE of adsorbates (OH, CO, and CH_3) on various metal systems versus the DFT BE (eV). (A) The model both trained and tested on PBE DFT data for NPs (Au/Ag/Cu, 55 to 172 atoms), which includes the CE_{NP} term. This model corresponds to case (i) of Table 1. (B) The model both trained and tested on PBE DFT data for NPs (Au/Ag/Cu, 55 to 172 atoms), which does not include the CE_{NP} term. This model corresponds to case (ii) of Table 1. (C) The model trained on PBE DFT data for NPs (Au/Ag/Cu, 55 to 172 atoms) and tested against RPBE DFT data for top-site adsorptions on metal surfaces (Au/Ag/Cu) from the literature slab dataset (29). This model corresponds to case (ii) of Table 1. (D) The model both trained and tested on RPBE DFT data for top-site adsorptions on metal surfaces (Au/Ag/Cu) from the slab dataset. This model corresponds to case (iii) of Table 1. In all cases, error bars are determined from the 10-fold cross-validated RMSE on the training set. Our DFT-calculated BEs of the different adsorbates on the various sites of the metal NPs are shown in table S6.

0.933 and 0.936 indicate that our model is strongly correlated with the actual DFT predictions.

We tested our model's generalizability by splitting our data into training and test sets based on adsorbate, metal identity, and morphology. In these series of "leave-one-in" tests, we restrict the training set to only a single morphology, adsorbate, or metal and then test the developed model on the data that were left out. The results of these tests can be found in table S4 and fig. S2. Overall, they provide strong evidence that our model is capturing the underlying physics of the binding interaction: When we train on a single metal or morphology, we still capture the other metals or morphologies with good accuracy. In the case of training on a single adsorbate, we see that, although the model fits are worse, they still capture the trends of the binding interactions (e.g., the adsorbates that were not tested are off parity by a constant amount). This is because the model has only a single possible value in the case of the IPEA term. Therefore, it reduces to a model that has only three terms, which are the intercept, CE_{local} , and MADs. However, this indicates that our model is highly robust. Even when the ad-

sorbate descriptor is missing, the model captures trends in the binding interactions. Therefore, our CE_{local} and MADs descriptors may be applicable to a wide array of adsorbates and NPs.

In addition to nonperiodic NP systems, we also investigated periodic (slab) systems. Recently, Roling and Abild-Pedersen (29) developed several scaling relations for adsorption on metal slabs and, alongside these relations, reported a large body of DFT calculations describing the adsorption of CH_3 , CO, and OH on the surface of several metal slabs of Ag, Au, Cu, Ir, Ni, Pd, Pt, and Rh. Using the scaling relations, they derive a model that describes metal-adsorbate BEs to a high degree of accuracy but is parameterized by several DFT calculations. In their model, the metal-adsorbate BE is asserted as a function of the DFT-calculated BE of the metal atom to the surface and of the gas-phase BE of an adsorbate to a metal atom (which we call MADs in this work). Although the gas-phase metal-adsorbate BE needs only to be parameterized once per metal-adsorbate system, calculating the metal's BE to its surface needs to be done once for every potential binding site, which requires a DFT optimization for the slab both with and without the

Table 1. OLS regression information for (i) four-descriptor model that includes CE_{local} , IPEA, MADs, and CE_{NP} ; (ii) three-descriptor model that excludes CE_{NP} ; and (iii) equivalent three-descriptor model using the slab dataset found in the literature. All cases are trained using datasets where CH_3 , CO, or OH adsorb to Cu, Ag, or Au (29). The maximum error corresponds to the largest deviation of a single data point. MAE, mean absolute error; RMSE, root mean square error; DOF, degrees of freedom.

| (i) Trained on NPs (four-descriptor model) (RMSE: 0.179 eV, MAE: 0.145 eV, R^2 : 0.936, maximum error: 0.619 eV, and remaining DOF: 157) | | | |
|--|----------------------|---------|------------------------|
| | Coefficient estimate | SE | P value |
| Intercept | 1.51477 | 0.15876 | $<2 \times 10^{-16}$ |
| CE_{local} | -0.1450 | 0.01663 | 3.85×10^{-15} |
| IPEA | 0.33171 | 0.01280 | $<2 \times 10^{-16}$ |
| MADs | 0.67858 | 0.01522 | $<2 \times 10^{-16}$ |
| CE_{NP} | -0.0002 | 0.05388 | 0.998 |
| (ii) Trained on NPs (three-descriptor model) (RMSE: 0.179 eV, MAE: 0.144 eV, R^2 : 0.933, maximum error: 0.619 eV, and remaining DOF: 158) | | | |
| | Coefficient estimate | SE | P value |
| Intercept | 1.51509 | 0.12148 | $<2 \times 10^{-16}$ |
| CE_{local} | -0.14502 | 0.01410 | $<2 \times 10^{-16}$ |
| IPEA | 0.33171 | 0.01274 | $<2 \times 10^{-16}$ |
| MADs | 0.67857 | 0.01501 | $<2 \times 10^{-16}$ |
| (iii) Trained on slab dataset (three-descriptor model) (RMSE: 0.122 eV, MAE: 0.102 eV, R^2 : 0.979, maximum error: 0.259 eV, and remaining DOF: 113) | | | |
| | Coefficient estimate | SE | P value |
| Intercept | 1.67677 | 0.09220 | $<2 \times 10^{-16}$ |
| CE_{local} | -0.14590 | 0.01079 | $<2 \times 10^{-16}$ |
| IPEA | 0.28743 | 0.01005 | $<2 \times 10^{-16}$ |
| MADs | 0.79516 | 0.01187 | $<2 \times 10^{-16}$ |

metal atom. In scenarios where adsorption energies across many different sites are to be investigated, the metal-surface BE term becomes a limiting step and may prove impossible for larger NPs (or slabs) that are infeasible to be calculated with DFT. Hence, our model, which does not need to be reparameterized for every binding site, allows for significantly higher throughput.

Here, we use the revised Perdew-Burke-Ernzerhof (RPBE)-calculated top-site adsorption energies on slabs reported by Roling and Abild-Pedersen (29) and focus on the Cu, Ag, and Au slabs adsorbing CH_3 , CO, and OH. We choose this particular dataset because it is a recent study encompassing a variety of face-centered cubic metal slabs from the d-block (including those we investigate from our own NP calculations) using the same adsorbates present in our calculations. It is also a large dataset, which allows us to have a more accurate assessment of our model's generalizability to these systems. We take our model, which has been fit to only our PBE DFT NP data, and use their data as a test set. In Fig. 2C, we plot the DFT-calculated BEs of the slabs

and compare them with the predictions of our model. We see that, although there is some deviation from parity, we still capture the overall trends of the binding interaction. This deviation is most likely because our training set used the PBE functional with a Gaussian planewave basis set and our model was trained entirely on NPs (no slabs were included in the training set). The slab dataset was produced from calculations with the RPBE functional and a planewave basis set. This results to the important observation that, although the adsorption data are with different functional, basis set, and adsorption environment (a slab not an NP), we are still able to predict the DFT BE to a high degree of accuracy based on our developed model. This is further evidence that our model is capturing the underlying physical trends of the systems. We note, however, that RPBE tends to produce more accurate CO adsorption energies than PBE, which overbinds CO relative to experiments (30). We observe this overbinding behavior in Fig. 2C, where the PBE-trained model predicts CO adsorption energies to be more negative relative to the RPBE DFT data. Last, to prove that our introduced descriptors and developed model are universal and applicable to periodic systems, we plot in Fig. 2D the model's behavior when it is both trained and tested on the slab dataset. In Fig. 2D, we immediately note that the discrepancy with CO moves back to parity. As expected, this indicates that, when the model is parameterized using the same functional used to calculate benchmark BEs, the results are more accurate. As it can be seen, all the data impressively fall on the parity line. Overall, this result combined with our leave-one-in tests (table S4 and fig. S2) shows that our model only needs to be fed a small subset of the possible morphologies that an NP or slab may take on and can then predict binding to other morphologies. The implication is that computationally inexpensive systems can be used to parameterize the model via DFT and then it can extend to much larger systems, which would otherwise be computationally intractable to investigate.

Extension to bimetallics

So far, our analysis has been entirely based on monometallic systems. Although monometallic systems are important to understand, a much larger and challenging chemical space can be found in the realm of bimetallics. For example, although there are 40 d-block metals (including La and Ac), there are $\frac{40^2-40}{2} = 780$ possible unique bimetallic alloys, without considering the additional materials space dimensions of NP size, shape, chemical ordering, etc. If we additionally consider chemical ordering in a bimetallic 55-atom NP (a computationally tractable metal NP size with DFT) of a single arbitrary morphology, there are a maximum of 2^{55} different ways by which both elements could be arranged in the NP. This number is already very large and further grows when different morphologies are considered. Therefore, fast and accurate descriptions of bimetallic alloys are a necessity if we are to explore even a fraction of these systems.

All evidence so far indicates that our model is physically sound, and therefore, we hypothesize that it should be extendable to bimetallic systems. Moreover, because CE_{local} is a direct extension of the BC model as we describe in Eq. 2 and takes into account the atomic identity of both the metal atom involved directly in the adsorption and every metal atom that has formed bonds with (first neighbors), this descriptor extends naturally to bimetallic systems. In addition, IPEA has no dependence on binding site, and MADs is only dependent on the atom directly bound to the adsorbate and does not require modification to capture top-site binding in bimetallic systems. Therefore, we continue using the identical physics-based descriptors for bimetallic NPs and slabs (as were used with monometallic systems), investigating top-site adsorption

energetics. In Fig. 3, we plot the BE of the CH_3 , CO, and the OH as predicted by our model when trained only on monometallic NPs against the DFT-calculated binding on several sites of the icosahedral 55-atom NPs ($\text{Cu}_{31}\text{Ag}_{24}$ and $\text{Cu}_{22}\text{Ag}_{33}$).

We can immediately see that the model very accurately captures trends in adsorption on the bimetallic Cu/Ag NPs. This further points toward the physical basis of our model because the model is only trained on monometallic systems and can still accurately capture the binding of the adsorbates on bimetallic NPs, which the model has never seen before. Furthermore, this demonstrates that our model is generalizable and applicable to both monometallic and bimetallic NPs. In addition, this is also indicative of the good performance of the CE_{local} descriptor. Other descriptors such as the CN and GCN (31) either do not have obvious extensions to systems of different metals or require extensive modification to capture multimetallic environments (32). The BC model (9) from which we derive CE_{local} , however, is deliberately formulated to describe bond strength between different metals and has been shown to perform very well on bimetallic alloys. We note, however, that Cu and Ag have similar electronegativities (33). Therefore, the tested $\text{Cu}_{55-x}\text{Ag}_x$ NPs do not show strong charge transfer because our calculations indicate point charges of less than $\pm 0.1 |e|$. Bimetallic systems that potentially develop significant charge transfer could affect the BE of adsorbates [and affect, for example, electrocatalytic behavior (18)] and may require an additional descriptor, such as binding site electronegativity, to capture these effects (34).

Extension to d7 and d8 metal slabs

Having shown that we are able to train on just one d9 metal and capture the adsorption trends of the other d9 metals (fig. S2), we believe that we should be able to extend our model to other systems. We again used the slab dataset (29), in which several d9 (Cu, Ag, and Au), d8 (Ni, Pd, and Pt), and d7 (Rh and Ir) metals were investigated. In Fig. 4A, we only trained the model on the Roling and Abild-Pedersen dataset of CH_3 , CO, and OH adsorbed to Cu, Ag, and Au NPs and showed that it is

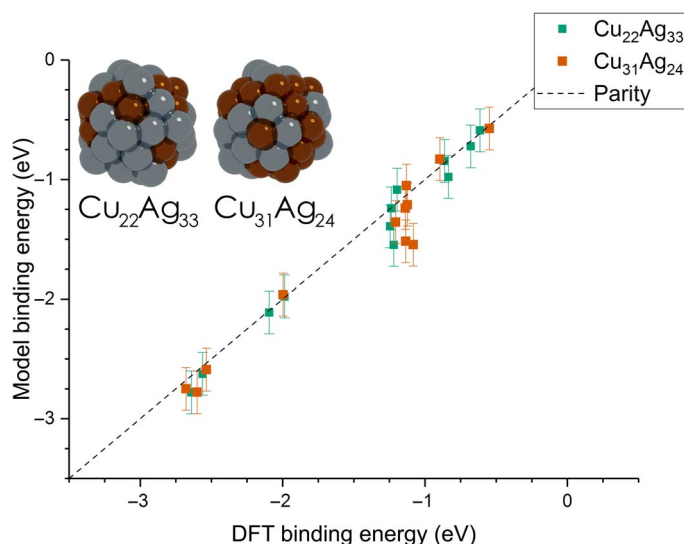


Fig. 3. Parity plot between our developed model and DFT calculations on icosahedral bimetallic ($\text{Cu}_{55-x}\text{Ag}_x$, $x = 24, 33$) NPs. The model is trained on CH_3 , CO, and OH adsorbing on monometallic Ag, Cu, and Au NPs and is able to capture adsorption on bimetallic NPs. Images of the two NPs are shown as inset, with copper and silver atoms colored in brown and gray, respectively.

able to capture the general adsorption trends for the other columns of the periodic tables, although deviations occur on a per adsorbate and per metal basis.

We then expanded our training set to include every periodic system in the slab dataset (CH_3 , CO, and OH adsorbed to Rh, Ir, Ni, Pd, Pt, Cu, Ag, or Au slabs) and attempted to fit our three-descriptor model to it. This resulted in a poor fit: Although it captures to some degree of accuracy the different metal-adsorbate pairs, it is unable to capture differences in the binding site of each slab (fig. S3A). This suggests that our model is missing some physical descriptor that can accurately differentiate between different columns of the periodic table. Searching for such a descriptor, we then investigated a variety of additional physical descriptors: The CN of the local site, the same first-order approximation of the chemical potential (of the metal), the hardness of the metal, the d-count of the metal, the covalent radius of the metal, the resistivity of the metal (related to the electronic structure), and the melting point of the metal (related to the strength of bonds). Even allowing the model to overfit by including all 10 of the potential physical descriptors we investigate, we do not see a substantial improvement over our original model. We present the summary of our search for an effective descriptor via OLS in table S5 and fig. S3.

We attempted to leverage more complex machine learning techniques to provide additional avenues to improve the adsorption model. Using LASSO (least absolute shrinkage and selection operator; see the “LASSO” subsection in section S3), we find the best descriptors to be the CN, the chemical potential of the metal (calculated in the same way as IPEA), the covalent radius of the metal, CE_{local} , IPEA, and MADs. Of these, both the covalent radius and the chemical potential of the metal exhibit low coefficients. Higher importance is given to the other descriptors, which, in the absence of CN (which can be justified by how CE_{local} is a demonstrably better descriptor), reduces to the model we have developed with slightly different coefficients (which is to be expected, considering that LASSO is not the same as OLS).

We additionally use symbolic regression (see the “Symbolic regression with Eureka” subsection in section S3), which is a highly flexible (but also interpretable) machine learning technique. Despite this enhanced flexibility, we again converge to the same three descriptors: CE_{local} , IPEA, and MADs. This is a further evidence that our model has a physically sound basis: The symbolic regression was given free rein to combine any of the investigated set of descriptors using addition, subtraction, multiplication, division, or any combination thereof yet still quickly predicted the form of our model, albeit with different coefficients (which is again expected, as the coefficients were generated via genetic algorithm and not formal OLS regression).

Although this provides excellent support for our physics-based chosen descriptors, it also indicates that none of the descriptors we have investigated so far is able to capture the difference in elements from column to column. In the absence of a good tabulated physical descriptor for this, if one even exists, we instead perform fits on a per column basis (Fig. 4B). We observe a general agreement with parity on the d8 and d7 metals and excellent agreement for the d9 metals.

Extension to Rh NPs and NH_3

We also tested our model to an entirely different adsorbate-metal pair: NH_3 and Rh. We see that, when we train on only Ag/Cu/Au with the three adsorbates (OH, CH_3 , and CO), the Rh and NH_3 systems are off by a consistent amount, as shown in Fig. 5A. Thus, although there is a deviation from parity, this deviation appears to be consistent, and the binding trend is still captured.

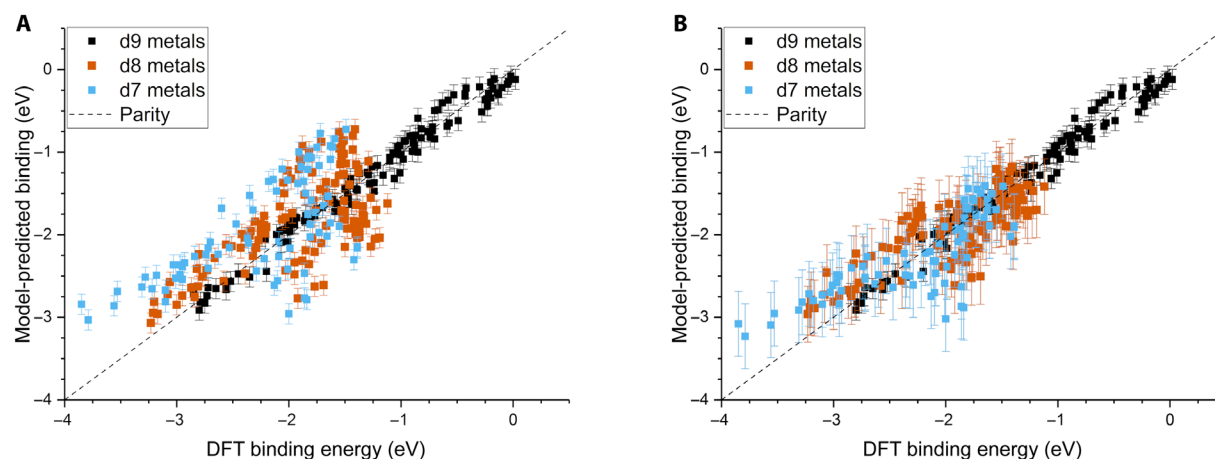


Fig. 4. Our three-descriptor model extended to slab dataset. (A) The model trained on the slab dataset (29) on Cu, Ag, and Au surfaces and tested against the Rh, Ir, Ni, Pd, Pt, Cu, Ag, and Au surfaces from the slab dataset. (B) The equivalent model when trained separately for each column of the d-block, still using the slab dataset. Error bars in every case are the 10-fold cross-validated RMSE of the training set.

Applying a constant correction to Rh and NH_3 to minimize the RMSE (and applying both corrections simultaneously to the Rh- NH_3 systems) results to bringing the data into parity (Fig. 5B). We again see that, although we cannot find a physical descriptor able to capture the differences between the columns of the periodic table (we trained on d9 metals with $\text{CH}_3/\text{CO}/\text{OH}$ and tested on the completely different system Rh with NH_3), we still capture the overall binding trend, albeit with an offset from parity.

Last, we parameterized the model on Rh using all three adsorbates (Fig. 5C) and showed that, although the model has difficulties finding differences in the BEs of each site, it is still able to capture well the trend between different adsorbates. We observe general trends in the different systems when we train on Ag, Au, and Cu, but offsets appear by some constant amount from the graph by metal and adsorbate. Furthermore, when attempting to simultaneously fit a model to all data (fig. S3), we see a stratification by adsorbate-metal pair, where roughly similar energies are being predicted for the same pairs with little difference in the CN and thus the CE_{local} .

DISCUSSION

In this work, we develop a simple yet powerful physics-based model for capturing trends in the strength of binding interactions between different adsorbates and metal NPs using machine learning techniques. The model introduces three simple descriptors, namely, the CE_{local} , IPEA, and MADs, that are able to capture adsorption on any site of metal NPs including monometallic and bimetallic systems.

Regarding the generalizability of the model, we are able to train the model on DFT-calculated adsorption results on NPs and capture the adsorption behavior on periodic surfaces. Furthermore, we use both LASSO and symbolic regression to search a larger space of potential descriptors that may differentiate between the columns in the d-block and to validate our model formulation. Both techniques demonstrated the importance of these three descriptors in adsorption. These simple descriptors can effectively model a wide range of binding interactions, including variations on the type of metals and composition, adsorption sites, and adsorbates.

Beyond the functional form of this model, we present a new descriptor for the chemistry of the local site, CE_{local} . This descriptor, which repre-

sents the stability of the adsorption site, is a localization of the BC model for NP CE (9), where only the bonds between the binding site metal atom and its neighbors are taken into account. The CE_{local} descriptor allows describing the stability of the local binding site by taking into account not only the CN of its atoms but also the coordination of its neighbors in addition to the chemical identity of the metal and its neighbors.

Our CE_{local} descriptor addresses shortcomings of other similar descriptors for local site reactivity. CN and GCN models (31) only contain geometric information, neglecting elemental composition of the site. In addition, these descriptors do not have an obvious extension to multimetallic systems, whereas CE_{local} requires no modification to describe adsorption on multimetallic binding sites. It is able to describe accurately the adsorption on bimetallic NPs even when the model is parameterized exclusively on monometallic systems. Other descriptors, such as the orbitalwise CN (22), are difficult to calculate and require knowledge or approximations of the electronic structure of the material of interest.

For high-throughput screening, where many different metals need to be investigated, computationally inexpensive and readily accessible descriptors are essential. We highlighted that, using data tabulated for most experimentally relevant adsorbates and metals and parameterizing our model with only a single DFT calculation per metal atom-adsorbate pair, we allow for rapid high-throughput screening of the adsorbate-NP and adsorbate-surface search space. Impressively, the model captures adsorption trends on bimetallic systems. Although we have not tested ternary systems here, we have no reason to believe that these physical properties will not remain relevant to accurately model multimetallic systems. Furthermore, we note that CE_{local} is constructed to account for practically any combination of metals on the local site. Future work will address whether this model is directly extendable to ternary systems.

Beyond the work of Roling and Abild-Pedersen (29), highlighted earlier in extending our model to periodic systems, recent work has also focused on developing adsorption models using statistical techniques. The automated screening and modeling approach of Tran and Ulissi (32) describes binding sites in terms of their atomic number, Pauling electronegativity, CN, and median of adsorption energies between the adsorbate and the pure metal. This approach yields accurate adsorption

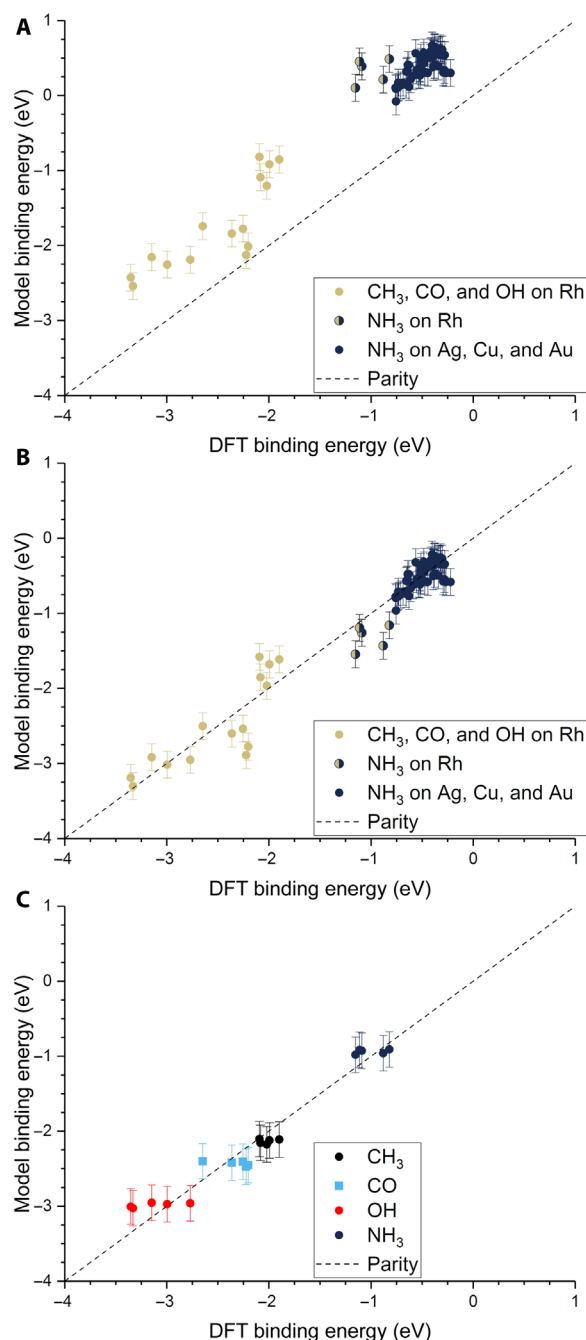


Fig. 5. Extension of our model to Rh and NH_3 . (A) The model parameterized on our Ag, Cu, and Au NPs adsorbing CH_3 , CO, and OH and tested against Rh and NH_3 . (B) The equivalent model with empirical (constant) corrections for Rh and NH_3 . In the case of NH_3 bound to Rh, both corrections are simultaneously applied and indicated by two-colored dots. (C) The model trained on CH_3 , CO, OH, and NH_3 adsorbing on icosahedral/cuboctahedral Rh_{55} .

energies showing good agreement with DFT but uses statistical models with enhanced complexity such as k -nearest neighbors and variations on the random forest. Enhanced model complexity typically requires larger datasets to prevent overfitting; this highlights the advantage of using a simpler linear model (see discussion on the “leave-one-in” tests in fig. S2).

The work of Andersen *et al.* (34) similarly uses a variety of DFT (e.g., d-band and densities of states) and experimentally derived descriptors (e.g., IP, EA, and CN), showing that DFT-calculated descriptors tend to provide better adsorption predictions than tabulated experimental descriptors such as the IP and EA. Using SISSO (Sure Independence Screening and Sparisfying Operator), they present several models, including models that take the form of first-order polynomials (form similar to Eq. 3). Comparing with the first-order polynomial models, they show training set RMSE ranging from approximately 0.38 to 0.16 eV as the number of terms increases, which is similar to what we observe in our multi-d-column linear regressions. Although the DFT-based approach can provide better predictions (due to comparing DFT adsorption energies with DFT-derived properties), we note that, in this work, we use tabulated data for more rapid screening, parameterizing our model with minimal DFT calculations.

Overall, our model generally exhibits cross-validated RMSE of about 0.12 to 0.18 eV. This error is comparable to other models in literature (29, 32, 34) and relatively small compared to the simplicity and applicability of our model. However, we suggest that our model should be primarily used for screening potential catalytic or adsorption materials to retrieve qualitative materials performance trends. For a quantitative property estimation, especially for properties that are very sensitive to error propagation as turnover frequencies in catalysis, these should be calculated by higher-fidelity methods once candidate materials are identified by our model (screening).

Our choice of descriptors has a strong physical basis. The IPEA can be viewed as the tendency of an adsorbate to react with other species, the CE_{local} is the equivalent tendency of the adsorption site to form bonds with adsorbates, and the MADs is a tuning of interaction between a metal and an adsorbate. Our model can accurately capture adsorption on both NPs and periodic surfaces, despite not having a descriptor for the stability of the whole surface or NP (i.e., CE_{NP}). This indicates that long-range interactions on a metal surface/NP play a minor role in determining the binding strength of the adsorbates. It is the stability of the local site on a surface or NP (CE_{local}) that contributes the most to adsorption on this site. We should acknowledge, however, that the adsorbates we investigate in this work are all relatively small and that most of their atoms either are directly participating in the metal-adsorbate bond or are nearest neighbors to the bonding interaction. We hypothesize that, for larger adsorbates, this IPEA descriptor may need to be replaced by a new descriptor characterizing the electronics of only the portion of the adsorbate directly involved in the bonding interaction. This can be justified via well-known effects such as induction and hyperconjugation, which (in the case of σ bonds) tend to be limited to the range of only one or two bonds. Fortunately, there is a large body of metal-catalyzed reactions such as the Haber-Bosch (35, 36), Fischer-Tropsch (37), water-gas shift reaction (38), etc., where small molecules similar to the ones we have investigated comprise most of the reactants and intermediates.

In summary, we introduce an adsorption model that is able to accurately describe the binding strength of molecules on any site of metal NPs, including alloys. Our model is simple in its form and uses data that can be readily accessible (or calculated on the fly). With surface adsorption being a critical step in catalysis, we anticipate it to be highly applicable as a screening tool for the high-throughput search of potential catalysts. With the rise of large databases and recent advancements in machine learning, these high-throughput searches tend to require cheap but physically relevant descriptors for reactivity of metals. In addition, our model can advance the discovery of nanosensors because it allows

for screening of adsorbates and metals at the same time with regard to their interaction strength. To the best of our knowledge, this is the first demonstration connecting adsorption properties of metal NPs (including bimetallics of random composition and chemical ordering) with the stability of the adsorption site. As a result, our model can also substantially affect materials optimization by designing nanostructures that exhibit the desired adsorption response within a stability cutoff.

MATERIALS AND METHODS

NP adsorption calculations

All adsorbate-NP BEs were calculated with DFT using QUICKSTEP (39), as implemented in CP2K (40). Exchange-correlation was accounted using the PBE functional. The DZVP (double-zeta valence polarized) basis set (41) was used with the Goedecker, Teter, and Hutter (GTH) pseudopotentials (42) at a 500-rydberg cutoff. Self-consistent field cycles were performed with a convergence criterion of 10^{-7} Ha. Geometry relaxations were converged to forces below 0.02 eV/Å. Visual depictions of the NPs with labeled CNs can be found in fig. S4 (A to E), and depictions of the top adsorption configurations studied can be found in fig. S4 (F to I). DFT-calculated BEs for each NP-adsorbate pair we investigated can be found in table S6.

Single-metal adsorption calculations

To calculate the MADs, we used the molecular structures of the complexes between a single metal and the adsorbate, as illustrated in Fig. 2. The energy for these complexes was evaluated using the PBE functional with the def2-SV(P) (split valence polarized) basis set (43) and the RI/MARJ [Resolution of Identity/Multipole Accelerated RI-J, where RI-J is the RI approximation for coulomb (J) integrals] approximations, as implemented in the Turbomole package (44, 45). Each structure was relaxed using a quasi-Newton-Raphson method for all multiplicities at or lower than septet (to account for spin state of each complex), and the lowest-energy multiplicity was used to calculate the MADs in Eq. 5 as

$$\text{MADs} = E_{\text{bind}, \text{M-Ads}} = E_{\text{M-Ads, complex}} - E_{\text{M}} - E_{\text{Ads}} \quad (5)$$

where E_X represents the electronic energy of species X, M is a metal, and Ads is the adsorbate. Visual depictions of the gas-phase metal-adsorbate binding description can be found in fig. S4 (J to M). Calculated values for MADs can be found in table S4.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/5/9/eaax5101/DC1>

Section S1. Thermodynamic data

Section S2. Leave-one-in tests

Section S3. Investigations on slabs

Section S4. Adsorption configurations

Table S1. Ionization energies and electron affinities for CH_3 , CO, OH, and NH_3 .

Table S2. BC model-calculated CEs of the relevant NPs we investigated.

Table S3. Calculated local CEs for relevant adsorption sites.

Table S4. Regression statistics for the various leave-one-in tests we performed.

Table S5. Regression statistics for intentionally overfit model plot of fig. S3B, with coefficients generated via OLS regression.

Table S6. DFT-calculated BEs for all studied adsorbate-NP pairs, ordered first by adsorbate, then by morphology, then by element, and lastly by CN.

Fig. S1. Adsorbate BEs versus BC model-calculated NP CEs.

Fig. S2. Parity plots for the various leave-one-in tests we performed.

Fig. S3. Characterization of all metal-adsorbate pairs in the slab dataset (29) simultaneously (e.g., there is one training set, which includes all adsorption interactions from the dataset).

Fig. S4. Illustration of initial configurations for several DFT calculations performed.

References (46–58)

REFERENCES AND NOTES

- L. Liu, A. Corma, Metal catalysts for heterogeneous catalysis: From single atoms to nanoclusters and nanoparticles. *Chem. Rev.* **118**, 4981–5079 (2018).
- S. Linic, P. Christopher, D. B. Ingram, Plasmonic-metal nanostructures for efficient conversion of solar to chemical energy. *Nat. Mater.* **10**, 911–921 (2011).
- M. Haruta, Novel catalysis of gold deposited on metal oxides. *Catal. Surv. Asia* **1**, 61–73 (1997).
- M. Valden, X. Lai, D. W. Goodman, Onset of catalytic activity of gold clusters on titania with the appearance of nonmetallic properties. *Science* **281**, 1647–1650 (1998).
- A. T. Bell, The impact of nanoscience on heterogeneous catalysis. *Science* **299**, 1688–1691 (2003).
- M. G. Taylor, N. Austin, C. E. Gounaris, G. Mpourmpakis, Catalyst design based on morphology- and environment-dependent adsorption on metal nanoparticles. *ACS Catal.* **5**, 6296–6301 (2015).
- D. R. Rolison, Catalytic nanoarchitectures—The importance of nothing and the unimportance of periodicity. *Science* **299**, 1698–1701 (2003).
- F. Yang, M. S. Chen, D. W. Goodman, Sintering of Au particles supported on $\text{TiO}_2(110)$ during CO oxidation. *J. Phys. Chem. C* **113**, 254–260 (2009).
- Z. Yan, M. G. Taylor, A. Mascareno, G. Mpourmpakis, Size-, shape-, and composition-dependent model for metal nanoparticle stability prediction. *Nano Lett.* **18**, 2696–2704 (2018).
- P. Sabatier, Hydrogénations et déshydrogénations par catalyse. *Ber. Dtsch. Chem. Ges.* **44**, 1984–2001 (1911).
- J. Greeley, M. Mavrikakis, Alloy catalysts designed from first principles. *Nat. Mater.* **3**, 810–815 (2004).
- D. A. Hansgen, D. G. Vlachos, J. G. Chen, Using first principles to predict bimetallic catalysts for the ammonia decomposition reaction. *Nat. Chem.* **2**, 484–489 (2010).
- J. K. Nørskov, T. Bligaard, J. Rossmeisl, C. H. Christensen, Towards the computational design of solid catalysts. *Nat. Chem.* **1**, 37–46 (2009).
- T. Bligaard, J. K. Nørskov, S. Dahl, J. Matthiesen, C. H. Christensen, J. Sehested, The Brønsted–Evans–Polanyi relation and the volcano curve in heterogeneous catalysis. *J. Catal.* **224**, 206–217 (2004).
- F. Viñes, A. Vojvodic, F. Abild-Pedersen, F. Illas, Brønsted–Evans–Polanyi relationship for transition metal carbide and transition metal oxide surfaces. *J. Phys. Chem. C* **117**, 4168–4171 (2013).
- F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skúlason, T. Bligaard, J. K. Nørskov, Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Phys. Rev. Lett.* **99**, 016105 (2007).
- J. K. Nørskov, F. Abild-Pedersen, F. Studt, T. Bligaard, Density functional theory in surface chemistry and catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 937–943 (2011).
- Z. W. Seh, J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Nørskov, T. F. Jaramillo, Combining theory and experiment in electrocatalysis: Insights into materials design. *Science* **355**, eaad4998 (2017).
- B. Hammer, J. K. Nørskov, Electronic factors determining the reactivity of metal surfaces. *Surf. Sci.* **343**, 211–220 (1995).
- G. Mpourmpakis, A. N. Andriotis, D. G. Vlachos, Identification of descriptors for the CO interaction with metal nanoparticles. *Nano Lett.* **10**, 1041–1045 (2010).
- F. Calle-Vallejo, J. I. Martínez, J. M. García-Lastra, P. Sautet, D. Loffreda, Fast prediction of adsorption properties for platinum nanocatalysts with generalized coordination numbers. *Angew. Chem. Int. Ed.* **53**, 8316–8319 (2014).
- X. Ma, H. Xin, Orbitalwise coordination number for predicting adsorption properties of metal nanocatalysts. *Phys. Rev. Lett.* **118**, 036101 (2017).
- A. Dąbrowski, Adsorption—From theory to practice. *Adv. Colloid Interface Sci.* **93**, 135–224 (2001).
- J. Wassenaar, E. Jansen, W.-J. van Zeist, F. M. Bickelhaupt, M. A. Sieglar, A. L. Spek, J. N. H. Reek, Catalyst selection based on intermediate stability measured by mass spectrometry. *Nat. Chem.* **2**, 417–421 (2010).
- N. Austin, J. K. Johnson, G. Mpourmpakis, Au_{13} : CO adsorbs, nanoparticle responds. *J. Phys. Chem. C* **119**, 18196–18202 (2015).
- Y. I. Delchev, A. I. Kuleff, J. Maruani, T. Mineva, F. Zahariev, in *Recent Advances in the Theory of Chemical and Physical Systems* (Springer Netherlands, 2006), pp. 159–176.
- M. Kuhn, Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26 (2008).
- R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).

29. L. T. Roling, F. Abild-Pedersen, Structure-sensitive scaling relations: Adsorption energies from surface site stability. *ChemCatChem* **10**, 1643–1650 (2018).
30. B. Hammer, L. B. Hansen, J. K. Nørskov, Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **59**, 7413–7421 (1999).
31. Z. Zhao, Z. Chen, X. Zhang, G. Lu, Generalized surface coordination number as an activity descriptor for CO₂ reduction on Cu surfaces. *J. Phys. Chem. C* **120**, 28125–28130 (2016).
32. K. Tran, Z. W. Ulissi, Active learning across intermetallics to guide discovery of electrocatalysts for CO₂ reduction and H₂ evolution. *Nat. Catal.* **1**, 696–703 (2018).
33. W. M. Haynes, *CRC Handbook of Chemistry and Physics* (CRC Press, ed. 92, 2011).
34. M. Andersen, S. V. Levchenko, M. Scheffler, K. Reuter, Beyond scaling relations for the description of catalytic materials. *ACS Catal.* **9**, 2752–2759 (2019).
35. P. Mehta, P. Barboun, F. A. Herrera, J. Kim, P. Rumbach, D. B. Go, J. C. Hicks, W. F. Schneider, Overcoming ammonia synthesis scaling relations with plasma-enabled catalysis. *Nat. Catal.* **1**, 269–275 (2018).
36. A. Vojvodic, A. J. Medford, F. Studt, F. Abild-Pedersen, T. S. Khan, T. Bligaard, J. K. Nørskov, Exploring the limits: A low-pressure, low-temperature Haber–Bosch process. *Chem. Phys. Lett.* **598**, 108–112 (2014).
37. M. D. Porosoff, B. Yan, J. G. Chen, Catalytic reduction of CO₂ by H₂ for synthesis of CO, methanol and hydrocarbons: Challenges and opportunities. *Energ. Environ. Sci.* **9**, 62–73 (2016).
38. M. Flytzani-Stephanopoulos, Gold atoms stabilized on various supports catalyze the water–gas shift reaction. *Acc. Chem. Res.* **47**, 783–792 (2014).
39. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, QUICKSTEP: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103–128 (2005).
40. J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, cp2k: Atomistic simulations of condensed matter systems. *Wires Comput. Mol. Sci.* **4**, 15–25 (2014).
41. J. VandeVondele, J. Hutter, Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **127**, 114105 (2007).
42. S. Goedecker, M. Teter, J. Hutter, Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **54**, 1703–1710 (1996).
43. F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **294**, 143–152 (1998).
44. F. Weigend, M. Häser, RI-MP2: First derivatives and global consistency. *Theor. Chem. Acc.* **97**, 331–340 (1997).
45. R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, Electronic-structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* **162**, 165–169 (1989).
46. C.-G. Zhan, J. A. Nichols, D. A. Dixon, Ionization potential, electron affinity, electronegativity, hardness, and electron excitation energy: Molecular properties from density functional theory orbital energies. *J. Phys. Chem. A* **107**, 4184–4195 (2003).
47. J. Berkowitz, G. B. Ellison, D. Gutman, Three methods to measure RH bond energies. *J. Phys. Chem.* **98**, 2744–2765 (1994).
48. G. B. Ellison, P. C. Engelking, W. C. Lineberger, An experimental determination of the geometry and electron affinity of methyl radical. *J. Am. Chem. Soc.* **100**, 2556–2558 (1978).
49. P. Erman, A. Karawajczyk, E. Rachlew-Källne, C. Strömholm, J. Larsson, A. Persson, R. Zerne, Direct determination of the ionization potential of CO by resonantly enhanced multiphoton ionization mass spectroscopy. *Chem. Phys. Lett.* **215**, 173–178 (1993).
50. K. M. A. Refaey, J. L. Franklin, Endoergic ion–molecule–collision processes of negative ions. III. Collisions of I[−] on O₂, CO, and CO₂. *Int. J. Mass Spectrom. Ion Phys.* **20**, 19–32 (1976).
51. P. A. Schulz, R. D. Mead, P. L. Jones, W. C. Lineberger, OH[−] and OD[−] threshold photodetachment. *J. Chem. Phys.* **77**, 1153–1165 (1982).
52. R. T. Wiedmann, R. G. Tonkyn, M. G. White, K. Wang, V. McKoy, Rotationally resolved threshold photoelectron spectra of OH and OD. *J. Chem. Phys.* **97**, 768–772 (1992).
53. F. Qi, L. S. Sheng, Y. W. Zhang, S. Q. Yu, W. K. Li, Experimental and theoretical study of the dissociation-energies D₀(H₂N-H) and D₀(H₂N⁺-H) and other related quantities. *Chem. Phys. Lett.* **234**, 450–454 (1995).
54. *NIST Computational Chemistry Comparison and Benchmark Database* (NIST Standard Reference Database Number 101, 2018).
55. R. L. Wasserstein, N. A. Lazar, The ASA statement on p-values: Context, process, and purpose. *Am. Stat.* **70**, 129–133 (2016).
56. J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).
57. R. Tibshirani, Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Series B Stat. Methodol.* **58**, 267–288 (1996).
58. M. Schmidt, H. Lipson, Distilling free-form natural laws from experimental data. *Science* **324**, 81–85 (2009).

Acknowledgments: We would like to acknowledge the computational support from the Center for Research Computing at the University of Pittsburgh and the Extreme Science and Engineering Discovery Environment, which is supported by the NSF (ACI-1548562). **Funding:** This material is based on work supported by the National Science Foundation under grant no. 1634880 (CMMI). **Author contributions:** J.D. and M.G.T. performed all the calculations. G.M. conceived the project and carried out the advising. All authors contributed to the preparation of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 28 March 2019
 Accepted 12 August 2019
 Published 13 September 2019
 10.1126/sciadv.aax5101

Citation: J. Dean, M. G. Taylor, G. Mpourmpakis, Unfolding adsorption on metal nanoparticles: Connecting stability with catalysis. *Sci. Adv.* **5**, eaax5101 (2019).

Unfolding adsorption on metal nanoparticles: Connecting stability with catalysis

James Dean, Michael G. Taylor and Giannis Mpourmpakis

Sci Adv **5** (9), eaax5101.
DOI: 10.1126/sciadv.aax5101

ARTICLE TOOLS

<http://advances.sciencemag.org/content/5/9/eaax5101>

SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2019/09/09/5.9.eaax5101.DC1>

REFERENCES

This article cites 54 articles, 6 of which you can access for free
<http://advances.sciencemag.org/content/5/9/eaax5101#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science Advances* is a registered trademark of AAAS.