Supporting Information

Machine Learning Based Electronic Structure Predictors in Single-Atom Alloys: A Model Study of CO Kink-Site Adsorption Across Transition Metal Substrates

Javad Shirani,[*,†] Hanh D. M. Pham,[‡] Shuaishuai Yuan,[†] Alain B. Tchagang,[¶] Julio J. Valdés,[¶] and Kirk H. Bevan[*,†,§]

†*Division of Materials Engineering, Faculty of Engineering, McGill University, Montréal, Québec, H3A 0C5, Canada*

‡*Department of Chemistry, McGill University, Montréal, Québec, H3A 0B8, Canada*

¶*Digital Technologies Research Centre, National Research Council Canada, Ottawa, ON, Canada K1A 0R6*

§*Centre for the Physics of Materials, Department of Physics, McGill University, Montréal, Québec, H3A 2T8, Canada*

E-mail: mohammadjavad.shirani@mail.mcgill.ca; kirk.bevan@mcgill.ca

Note

All references included in this Supporting Information refer to the list provided at the end of the Supporting Information. However, all references cited in the Supporting Information are also cited in the main manuscript, though with revised ordering.

Supplemental S1: Feature Set Summary

Numerous features were taken into consideration in this study which could be split into 6 electronic structure features extracted from the d-band PDOS (n_d) and 24 chemical (atomic) features adapted from the literature.[1–7] Thus, in total, 30 features were considered.

Electronic Structure Features (extracted from DFT calculations)

In addition to the features detailed in the context of Eqs. (2)-(5) of the main text, we also employed two additional electronic structure features as given below.[8,9]

- γ_1: Skewness of a d-band which is calculated as

$$\gamma_1 = E\Big[\Big(\frac{n_d - \mu}{\sigma}\Big)^3\Big]. \tag{S1}$$

- γ_2: Kurtosis of a d-band which is calculated as

$$\gamma_2 = E\Big[\Big(\frac{n_d - \mu}{\sigma}\Big)^4\Big]. \tag{S2}$$

Here μ is the mean, σ is the standard deviation of the band and E is the expectation operator pertaining to the d-band PDOS distribution.

Chemical Features

- IE_{dopant}: Ionization energy of the substitute doped into the host's surface.

- IE_{host}: Ionization energy of the host.
- IE_{kink}: Ionization energy of the kink site.
- EA_{dopant}: Electron affinity of the substitute doped into the host's surface.
- EA_{host}: Electron affinity of the host.
- EA_{kink}: Electron affinity of the kink site.
- r_{dopant}: Atomic radius of the substitute doped into the host's surface.
- r_{host}: Atomic radius of the host.
- r_{kink}: Atomic radius of the kink site.
- X_{dopant}: Electronegativity of the substitute doped into the host's surface.
- X_{host}: Electronegativity of the host.
- X_{kink}: Electronegativity of the kink site.
- W_{dopant}: Work function of the substitute doped into the host's surface.
- W_{host}: Work function of the host.
- W_{kink}: Work function of the kink site.
- O_{dopant}: Tight-binding d-orbital energy of the substitute doped into the host's surface.
- O_{host}: Tight-binding d-orbital energy of the host.
- O_{kink}: Tight-binding d-orbital energy of the kink site.
- $r_{d,dopant}$: Tight-binding d-orbital radius of the substitute doped into the host's surface.
- $r_{d,host}$: Tight-binding d-orbital radius of the host.
- $r_{d,kink}$: Tight-binding d-orbital radius of the kink site.

- C_{dopant} (V^2_{dopant}): Adsorbate-metal interatomic d coupling matrix element squared for the substitute doped into the host's surface.

- C_{host} (V^2_{host}): Adsorbate-metal interatomic d coupling matrix element squared for the host.

- C_{kink} (V^2_{kink}): Adsorbate-metal interatomic d coupling matrix element squared for the kink site.

Supplemental S2: Machine Learning Hyperparameters

Table S1: The range of hyperparameters tested for each ML model (hyperparameters not mentioned were kept at their default values as per the Scikit-Learn/Keras documentation) – Pertaining to the dataset including all 30 features.

ML Algorithm	Hyperparameters
Ridge regression	Alpha = [0.1, 0.5, 0.8, 1, 10, 100]
Support vector regression	kernel:["linear","rbf"], C: [1,10,10^2,10^3,10^4, 10^5,10^6, 10^7,10^8], gamma: [1,10^{-1},10^{-2}, 10^{-3},10^{-4},10^{-5},10^{-6},10^{-7},10^{-8}, 10^{-9},10^{-10}]
Random forest regression	n_estimators:[50, 100, 200, 300, 400, 500, 600, 700, 800,900], max_depth:[1,2,3,4,5,6,7,8,9,10], max_features:['sqrt'], min_samples_leaf:[3,4,5]
Extra tree regression	n_estimators:[50, 100, 200, 300, 400, 500, 600, 700, 800,900], max_depth:[1,2,3,4,5,6,7,8,9,10], max_features:['sqrt'], min_samples_leaf:[3,4,5], bootstrap: ["True","False"]
Gradient boosting regression	n_estimators:[50, 100, 200, 300, 400, 500, 600, 700, 800,900], learning_rate: [0.000001,0.0001,0.001,0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2], max_depth:[1,2,3,4,5,6,7,8,9,10], criterion: ["friedman_mse"], subsample:[0.5, 0.8, 0.85, 0.9, 0.95, 1.0], max_features:['sqrt'], min_samples_leaf: [20,21,22,23,24,25,26,27,28,29,30]

Table S2: Optimum hyperparameters evaluated for each ML algorithm – Pertaining to the dataset including all 30 features.

ML algorithm	Tuned hyperparameter value
Ridge regression	Alpha = [0.1]
Support vector regression	C: 100000.0, gamma: 0.1, kernel: 'rbf'
Random forest regression	n_estimators:[200], max_depth:[10], max_features:['sqrt'], min_samples_leaf:[3]
Extra tree regression	n_estimators:[600], max_depth:[10], max_features:['sqrt'], min_samples_leaf:[3], bootstrap: ["True"]
Gradient boosting regression	n_estimators:[200], learning_rate: [0.1], max_depth:[8], criterion: ["friedman_mse"], subsample:[0.5], max_features:['sqrt'], min_samples_leaf: [22]

Supplemental S3: Single Feature Trends

Figure S1 shows the binding trends versus some single features obtained via the electronic structure of the systems studied such as: the upper d-band edge (ε_u), d-band width (W_d), and d-band filling (f) – defined per Eqs. (3)-(5) in the main text.

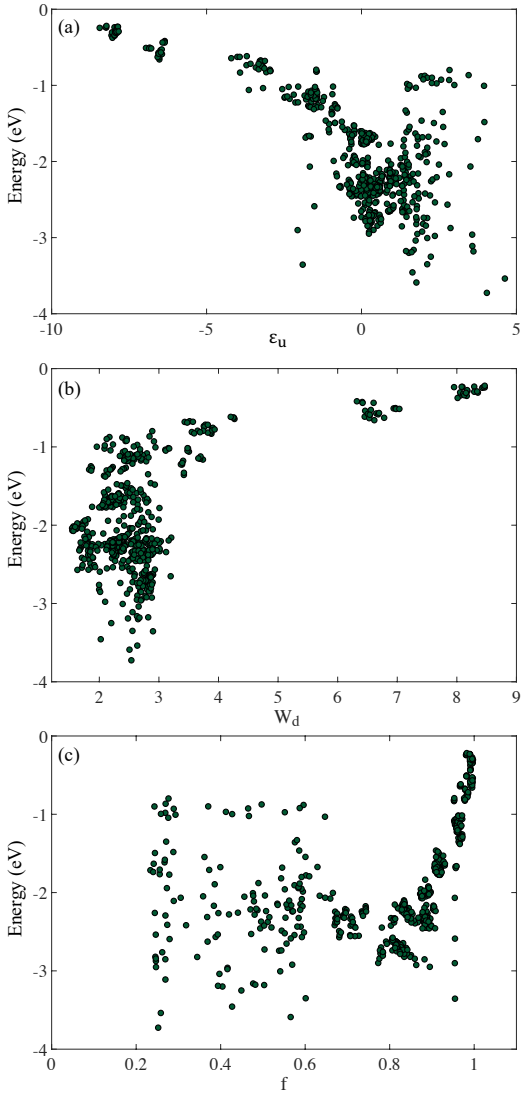

Figure S1: Single-features trends concerning CO adsorption energy at the kink with respect to: (a) the upper d-band edge (ε_u); (b) the d-band width (W_d); and (c) d-band filling (f) at the kink site.

Figure S2 shows the exact same data displayed in Figs. 2c & 2d of the main text re-plotted by element to show the clustering of different elements (dopants). As can be seen in these two cases the adsorption energy is not dominated by the impurity but by the host (substrate) when substitution is not directly at a kink site.

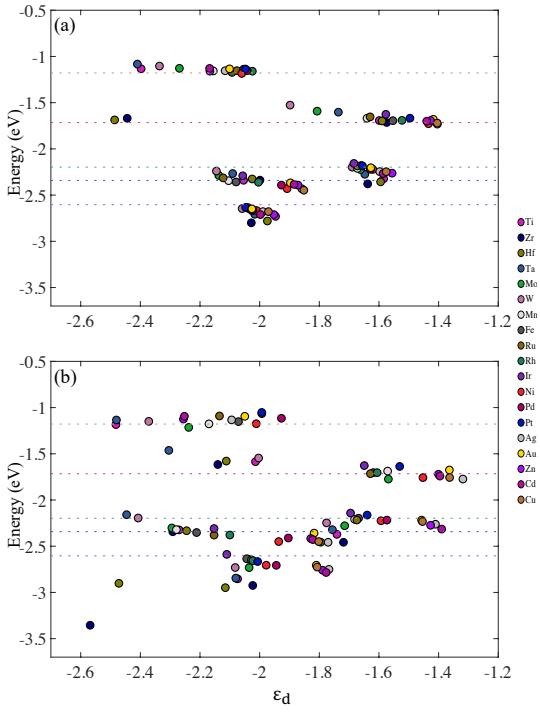

Figure S2: Adsorption energy plots versus ε_d for different subsets of data: (a) when only one impurity is doped at an embedded step edge position distinct from the kink site; and (b) when two impurities are embedded at step edge positions distinct from the kink site. The CO kink-site adsorption energy atop pure slabs is shown for comparison via dotted lines.

Supplemental S4: Comparison to Other Techniques

In addition to the electronic structure approach we adopted in this study, many other methodologies (descriptor spaces) have been proposed in the literature to predict small-molecule binding energies for either bulk metallic alloys, single atom alloys, or transition metal doping in various systems.[8–22] Hence, for the sake of clarity and completeness, it is important to also assess the performance of closely related methodologies[8,23–25] in the context of the chemically diverse single-atom alloys investigated in this work. These comparisons are provided below.

Supplemental S4-A: Comparison with Reference 23

In Ref. 23, a descriptor denoted as φ was proposed formulated as a function of d-band filling, electronegativity and the number of nearest neighbours adjoining the small-molecule binding site. This newly defined descriptor was shown to provide a linear relationship with small-molecule binding energies in graphene-based single-atom catalysts and metal–macrocycle complexes. In this manner Ref. 23 demonstrated that the catalytic activity of single-atom catalysts (in carbon based hosts) is often highly correlated with the local environment of the metal centre, namely its coordination number and electronegativity and the electronegativity of the nearest neighbour atoms. The question naturally arises: can this descriptor be applied with equal success towards single-atom alloys containing a chemically diverse set of impurities? Since our system is slightly different, being non-carbon based and metallic, we therefore marginally modified the descriptor φ in the form given below

$$\varphi = \beta\theta_d \times \frac{E_M + \alpha(n_D \times E_D + n_H \times E_H)}{E_C} \tag{S3}$$

where θ_d is the number of valence electrons in the d-orbital of the adsorption site metal element, E_M is the electronegativity of the adsorption site metal element, n_D is the number

of impurity atom nearest neighbors next to the adsorption site, E_D is the electronegativity of the dopant element, n_H is the number of host atom nearest neighbors to the adsorption site, E_D is the electronegativity of the host atom element, E_C is the electronegativity of the carbon atom (included for the sake of consistency with the original model, though clarification will be applied shortly) and α and β are correction coefficients. In Fig. S3 we first plot the adsorption energy versus φ with $\alpha = 1$ and $\beta = 1$, both for a subset of data with a single dopant at the kink site and for the full dataset. These trends are the first indication that φ does not suffice as a single linear descriptor for the chemically diverse data set investigated in this study.

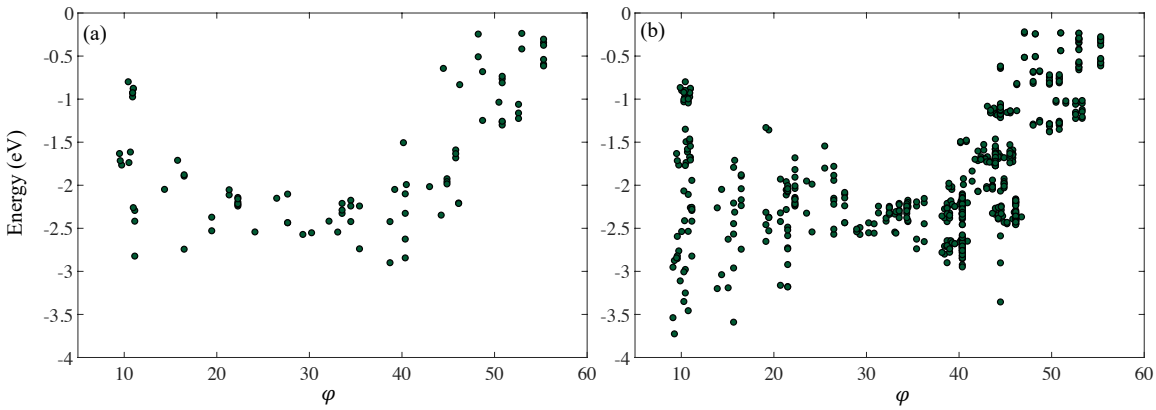

Figure S3: Adsorption energy plots versus φ for different subsets of data: (a) when only one impurity is doped at the kink site; and (b) for the full dataset.

However, the trends in Fig. S3 depend on the parameter E_C (borrowed from Ref. 23) which may not be relevant to the single-atom alloys investigated in this study. To this end we also treated the coefficients α and β in Eq. (S3) as independent adjustable parameters to be fitted via linear regression – this is essentially the same as removing E_C from the fit. The results of this fit can be found in Fig. S4, where it can be seen that the predictive performance of such a single descriptor (φ) remains quite poor for this single-atom alloy dataset when further flexibility is introduced. For these reasons a descriptor of the form given by Eq. (S3) was excluded from consideration in the primary analysis of this study. However, it is important to emphasize that Figs. S3 & S4 do not suggest that this descriptor

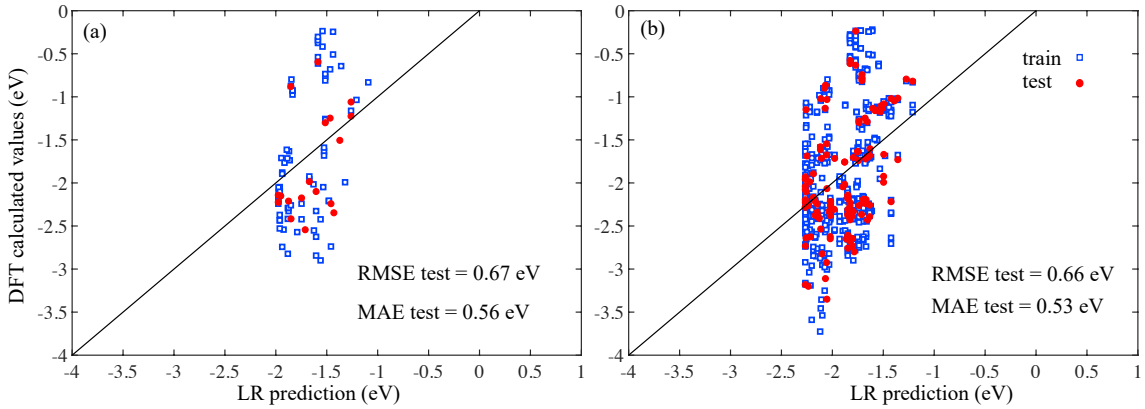

Figure S4: Linear regression prediction performance for φ when α and β are treated as fitting parameters for different subsets of data: (a) when only one impurity is doped at the kink site; and (b) for the full dataset.

is not without utility for the materials examined in Ref. 23.

Note: content continues on the next page.

Supplemental S4-B: Comparison with Reference 24

In Ref. 24 the authors proposed an intriguing model for predicting binding energies of different adsorbates on transition metals and related alloys using easily accessible chemical data.[26] Their model presented binding energies as a function of such features to arrive at linear relations using: (1) the generalized coordination number (as a metric for the geometry of the system); (2) outer-electron number; and (3) electronegativity. These generic descriptors where recast within the modified descriptor ψ which was defined as:

$$\psi = \frac{S_V^2}{\chi^\beta} \tag{S4}$$

where where χ and S_V represent the electronegativity and the valence electron number of the transition metal element, respectively, and β is an index determined by the role of d- and s- orbitals.[24] To further explore this physical chemistry, they also employed the modified descriptor ψ_I in Ref. 24 which can be expressed as:

$$\psi_I = \frac{\left(\Pi_{i=1}^{N} S_{V_i}\right)^{2/N}}{\left(\Pi_{i=1}^{N} \chi_i^\beta\right)^{1/N}}. \tag{S5}$$

Here S_{V_i} and χ_i are the outer-electron number and electronegativity of a given atom, respectively, such that the multiplication series in Eq. (S5) extends over the adsorption site and its nearest neighbours. We have applied the modified electronic descriptors proposed in the above paper, namely ψ and ψ_I, to our data for the subset of data with only one dopant at the kink site and the full dataset. These results are provided in Fig S5, where it can be clearly seen that there is no linear relationship between the adsorption energy and modified descriptors proposed in Ref. 24 per the data set examined in this work. However, it is important to emphasize Fig. S5 does not suggest that this descriptor is not without relevant utility for the subset of data & materials examined in Ref. 24 and supporting follow-up studies.[26]

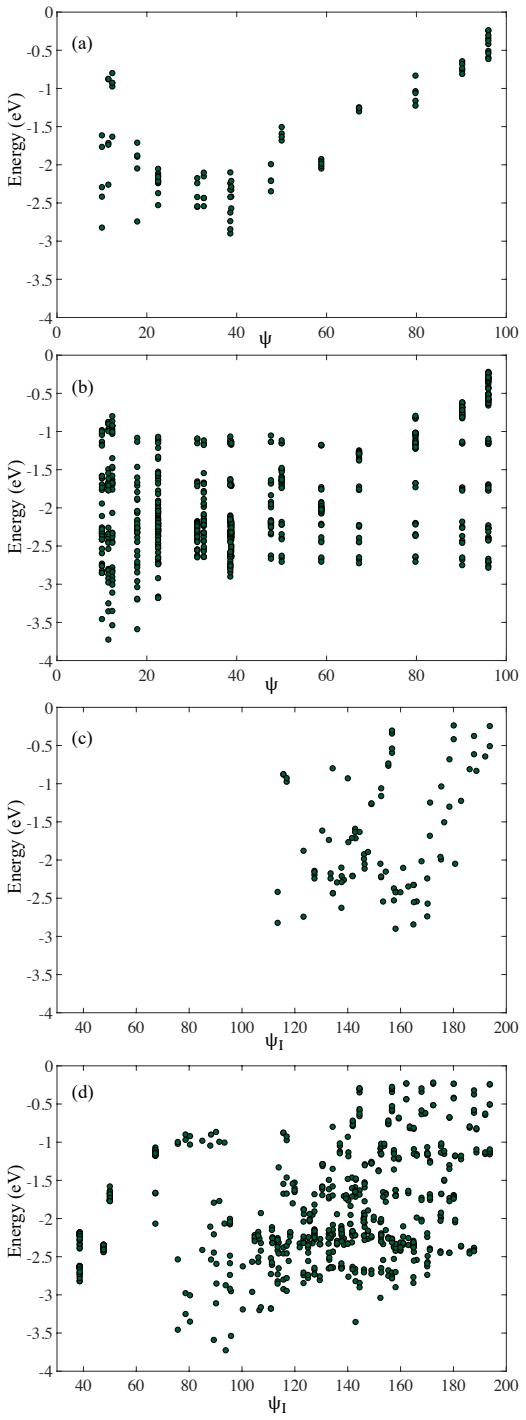

Figure S5: Descriptor trends concerning CO adsorption energies at the Cu(532) kink with respect to modified descriptors proposed in Ref. 24: (a) ψ trends when only one impurity is doped at the kink site; and (b) ψ trends for the full dataset; (c) ψ_I trends when only one impurity is doped at the kink site; and (d) ψ_I trends for the full dataset.

Supplemental S4-C: Comparison with Reference 8

In Ref. 8 an artificial neural network (ANN) model was employed predict the binding energy of CO on bimetallic surfaces. As an earlier example of applying machine learning to catalysis, this work represents an important marker in the field. In a similar manner to our work, they utilized d-band properties as well as atomic properties of the substrate. Importantly, their ANN model provided a low test error corresponding to RMSE ≈ 0.13 eV. However, it was focused on the bimetallic alloys which can be understood to follow linear scaling relations to a greater degree than single atom alloys. To demonstrate this, their binding energy values are plotted with respect to their reported d-band center feature in Fig. S6b – all data was obtained from the Supporting Information to Ref. 8. Next to this result, in Fig. S6a,

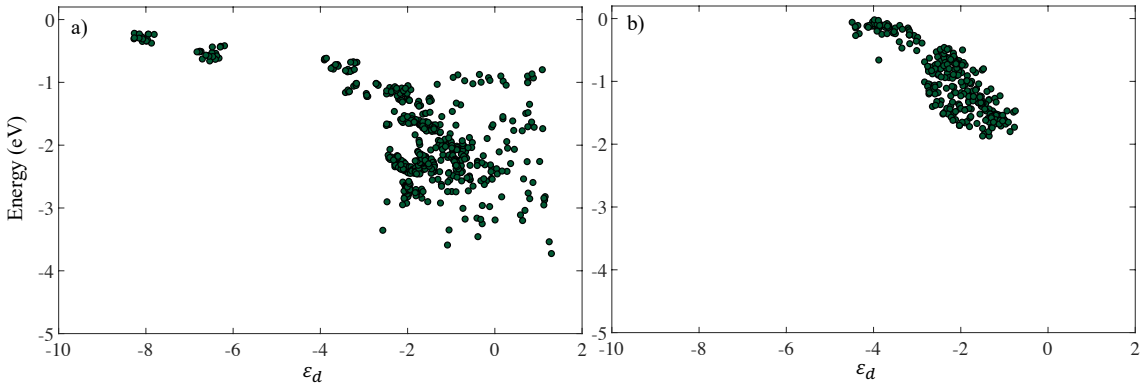

Figure S6: Single-features trends concerning CO adsorption energy with respect to the project density of states (PDOS) d-centre (ε_d) taken from: (a) our manuscript; and (b) Ref. 8.

we have plotted our binding energy versus d-band center trends – just as was provided in Fig. 1 of the main manuscript. As can be seen from the above figure, the subset of alloys they studied follows linearity relations more closely than the non-linear SAAs investigated in our study. Moreover, the data in Ref. 8 is also well clustered (see Fig. S6b). On the other hand, additional diversity diversity is present in our data (see Fig. S6a). This arises from non-linear nature of SAAs, which we have found to require more sophisticated machine modelling methods. Indeed, due to the clustering of their data and its more linear properties,

we were able to attain a fairly accurate linear regression model of the data provided in Ref. 8. This is provided in Fig. S7 and results in an RMSE of 0.15 eV by averaging over 100 time random train/test split of 20%/80%. This is nearly identical to the ANN model RMSE result of ~ 0.13 eV reported in Ref. 8. Hence, it is the diversity of our data set (e.g., stretching out across a wide range of transition metal dopants) and the non-linear nature of single-atom alloys that leads to the slightly higher RMSE quoted in our work.

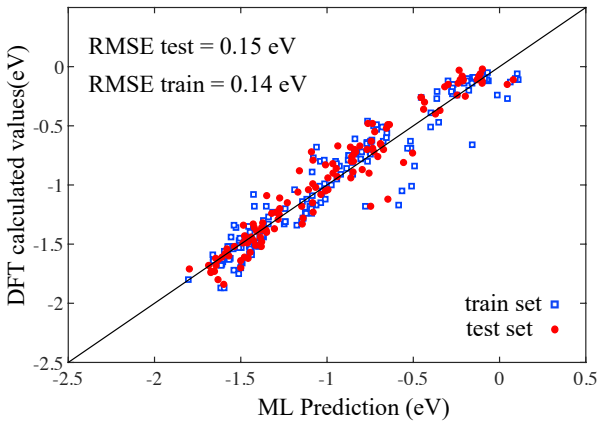

Figure S7: Machine learning prediction performance obtained from a linear regression model applied to the data provided by Ref. 8.

Note: content continues on the next page.

Supplemental S4-D: Comparison with Ref. 25

In Ref. 25 the authors investigated OH and *O binding energies on the surfaces of high entropy alloys (HEAs). To address the complexity of HEAs, they presented a linear regression model which links the local atomic make-up (e.g., number and chemical type of atoms in the binding vicinity) to the binding energy. The binding energy predicted at site i by this model can be encapsulated in the form

$$\Delta E^i_{pred} = \Sigma_k^{metals} C_{1,k} N^i_{1,k} + \Sigma_k^{metals} C_{2,k} N^i_{2,k} + \Sigma_k^{metals} C_{3,k} N^i_{3,k} \tag{S6}$$

where $C_{1,k}$, $C_{2,k}$,and $C_{3,k}$ are the model parameters and $N^i_{1,k}$, $N^i_{2,k}$ and $N^i_{3,k}$ are the numbers of each element (k) found in each zone numbered from 1 to 3. Zone 1 represents the binding site, Zone 2 represents the surface neighbours that are coordinating once to the binding site and Zone 3 represents the subsurface neighbours that are coordinating once to the binding site. Importantly, this linear regression approach provided an impressive root mean squared deviation of 0.063 eV and 0.076 eV for *OH and *O, respectively. Although, the number of fitting parameters was reported to be 55.[25] We applied this same general configurational approach to our data set, the result of which can be found in Fig S8. Notably, the prediction performance of the above configurational approach closely follows the predic-

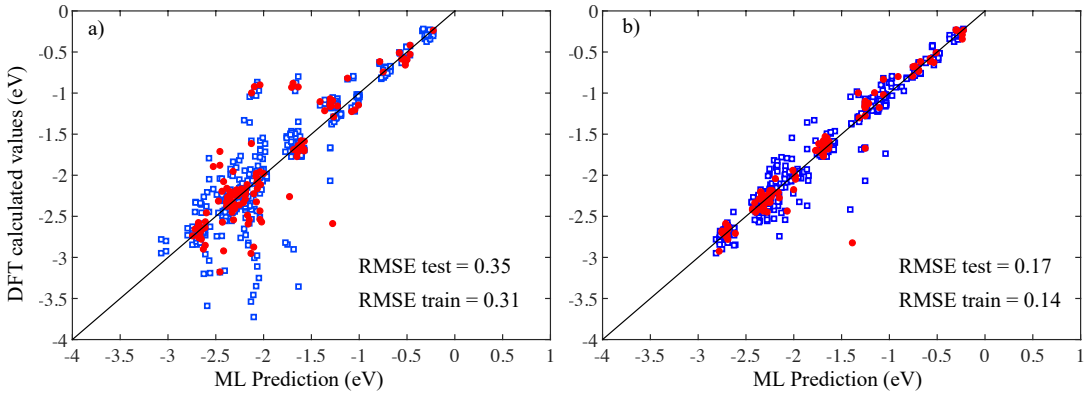

Figure S8: Machine learning prediction performance obtained from linear regression model applied to our data set by employing the methodology in Ref. 8. (a) Includes all the data. (b) Includes the configurationally restricted subset as in the main manuscript.

tion performance in our study (with slightly higher statistical test error) – see Figs. 7 & 8 of the main manuscript. Although the accuracies of the two methods are comparable (ours in Figs. 7 & 8 of the main manuscript and Fig. S8 herein), we would argue that electronic structure features provide a more versatile and intuitive basis for understanding the physics driving catalytic behaviour on distinct metallic surfaces.[27–29]

Features (descriptors) employed for the machine learning driven prediction of adsorption/binding energies are often classified into three main categories: geometric, electronic, and energetic. Geometrical descriptors are acquired from atomic positions and often include corresponding coordination numbers.[30–33] Electronic features are acquired from the electronic structure of a system (e.g., the d-band center).[27–29] Lastly, energetic descriptors include adsorption energies, bulk formation energies, and vacancy formation energies. When such energetic descriptors are derived from DFT calculations, extensive geometry relaxations are required that are often difficult to attain as compared to unrelaxed electronic structure features.[34,35] In this study, we focused on electronic descriptors and their performance as predictors, since it is well known that the electronic structure of a surface is strongly correlated with its surface chemistry.[36,37] Indeed, electronic descriptors (DOS descriptors) are suitable descriptors to consider for the prediction of binding energies, as the mode and degree of a surface's interaction with adsorbates are directly manifest within its electronic DOS through hybridization bonding interactions.[36] Although coordination based geometrical features are sometimes more easily accessible descriptors compared to electronic structure features, there are strong physical concepts that supporting the relation between the electronic structure and binding energies.[27–29,36,37] Lastly, SAAs are known to break linear relationships and yet their relations between electronic structure characteristics and binding properties is not fully understood.[38] These are the physical reasons we have chosen to explore machine learning models in the context of electronic structure features applied to SAAs.

There is, however, also the interpretability of the machine learning model which needs to be taken into account when comparing the two approaches (again, ours in Figs. 7 & 8 of

the main manuscript and Fig. S8 herein). In this context, the notion of of simplicity and complexity in machine learning depends on one's point of view regarding the parameters and the function that links them. On the one hand we have a model that is simple from the function point of view in Fig. S8 (employing linear regression) but complex from the number of parameters, where it is very difficult to understand what is occurring physically in the parameter space to drive the observed behaviour. On the other hand we are choosing a model (in Figs. 7 & 8 of the main manuscript) that is simpler from the parameter perspective, with more physical interpretability with respect to the driving physical variables (e.g., electron filling), but has more complexity in terms of interpreting the model non-linearity. Since these two models are approximately equivalent, producing nearly the same predictive accuracy within this system, it to some degree a matter of interpretability preference regarding which model is applied. However, the linear model (with more parameters) is slightly worse statistically (compare Figs. 7 & 8 of the main manuscript to Fig. S8 herein). Thus, for these additional machine learning feature interpretation reasons we have chosen to employ an electronic structure based model in our analysis.

Supplemental S4: Data Set

The data set employed in this study is attached as a "csv" formatted file with the Supporting Information.

References

(1) Myers, R. T. The periodicity of electron affinity. *J. Chem. Educ.* **1990**, *67*, 307.

(2) Slater, J. C. Atomic radii in crystals. *Chem. Phys.* **1964**, *41*, 3199–3204.

(3) Gordy, W.; Thomas, W. O. Electronegativities of the elements. *Chem. Phys.* **1956**, *24*, 439–444.

(4) Michaelson, H. B. The work function of the elements and its periodicity. *J. Appl. Phys.* **1977**, *48*, 4729–4733.

(5) Williams, M. CRC handbook of chemistry and physics. *Occup Environ Med* **1996**, *53*, 504.

(6) Harrison, W. A. *Electronic structure and the properties of solids: the physics of the chemical bond*; Courier Corporation, 2012.

(7) Kudrnovskỳ, J.; Drchal, V.; Maek, J. Canonical description of electron states in random alloys. *Phys. Rev. B* **1987**, *35*, 2487.

(8) Ma, X.; Li, Z.; Achenie, L. E.; Xin, H. Machine-learning-augmented chemisorption model for CO2 electroreduction catalyst screening. *J. Phys. Chem.* **2015**, *6*, 3528–3533.

(9) Li, Z.; Ma, X.; Xin, H. Feature engineering of machine-learning chemisorption models for catalyst design. *Catal. Today* **2017**, *280*, 232–238.

(10) Xin, H.; Vojvodic, A.; Voss, J.; Nørskov, J. K.; Abild-Pedersen, F. Effects of d-band shape on the surface reactivity of transition-metal alloys. *Phys. Rev. B* **2014**, *89*, 115114.

(11) Fung, V.; Hu, G.; Ganesh, P.; Sumpter, B. G. Machine learned features from density of states for accurate adsorption energy prediction. *Nature Comm.* **2021**, *12*, 88.

(12) Andersen, M.; Levchenko, S. V.; Scheffler, M.; Reuter, K. Beyond scaling relations for the description of catalytic materials. *ACS Catal.* **2019**, *9*, 2752–2759.

(13) Chen, A.; Zhang, X.; Chen, L.; Yao, S.; Zhou, Z. A machine learning model on simple features for CO_2 reduction electrocatalysts. *J. Phys. Chem.* **2020**, *124*, 22471–22478.

(14) Xin, H.; Holewinski, A.; Linic, S. Predictive structure–reactivity models for rapid screening of Pt-based multimetallic electrocatalysts for the oxygen reduction reaction. *ACS Catal.* **2012**, *2*, 12–16.

(15) Artrith, N.; Kolpak, A. M. Understanding the composition and activity of electrocatalytic nanoalloys in aqueous solvents: a combination of DFT and accurate neural network potentials. *Nano Lett.* **2014**, *14*, 2670–2676.

(16) Ulissi, Z. W.; Tang, M. T.; Xiao, J.; Liu, X.; Torelli, D. A.; Karamad, M.; Cummins, K.; Hahn, C.; Lewis, N. S.; Jaramillo, T. F. et al. Machine-learning methods enable exhaustive searches for active bimetallic facets and reveal active site motifs for CO_2 reduction. *ACS Catal.* **2017**, *7*, 6600–6608.

(17) Hoyt, R. A.; Montemore, M. M.; Fampiou, I.; Chen, W.; Tritsaris, G.; Kaxiras, E. Machine learning prediction of H adsorption energies on Ag alloys. *J Chem Inf Model.* **2019**, *59*, 1357–1365.

(18) Singh, A. R.; Rohr, B. A.; Gauthier, J. A.; Nørskov, J. K. Predicting chemical reaction barriers with a machine learning model. *Catal Letters.* **2019**, *149*, 2347–2354.

(19) Gasper, R.; Shi, H.; Ramasubramaniam, A. Adsorption of CO on low-energy, low-symmetry Pt nanoparticles: energy decomposition analysis and prediction via machine-learning models. *J. Phys. Chem. C* **2017**, *121*, 5612–5619.

(20) Han, Z.-K.; Sarker, D.; Ouyang, R.; Mazheika, A.; Gao, Y.; Levchenko, S. V. Single-

atom alloy catalysts designed by first-principles calculations and artificial intelligence. *Nat. Commun.* **2021**, *12*, 1–9.

(21) Dasgupta, A.; Gao, Y.; Broderick, S. R.; Pitman, E. B.; Rajan, K. Machine learning-aided identification of single atom alloy catalysts. *J. Phys. Chem. C.* **2020**, *124*, 14158–14166.

(22) Toyao, T.; Suzuki, K.; Kikuchi, S.; Takakusagi, S.; Shimizu, K.; Takigawa, I. Toward effective utilization of methane: machine learning prediction of adsorption energies on metal alloys. *J. Phys. Chem. C.* **2018**, *122*, 8315–8326.

(23) Xu, H.; Cheng, D.; Cao, D.; Zeng, X. C. A universal principle for a rational design of single-atom electrocatalysts. *Nat. Catal.* **2018**, *1*, 339–348.

(24) Gao, W.; Chen, Y.; Li, B.; Liu, S.-P.; Liu, X.; Jiang, Q. Determining the adsorption energies of small molecules with the intrinsic properties of adsorbates and substrates. *Nat. Commun.* **2020**, *11*, 1196.

(25) Batchelor, T. A.; Pedersen, J. K.; Winther, S. H.; Castelli, I. E.; Jacobsen, K. W.; Rossmeisl, J. High-entropy alloys as a discovery platform for electrocatalysis. *Joule* **2019**, *3*, 834–845.

(26) Li, X.; Li, B.; Yang, Z.; Chen, Z.; Gao, W.; Jiang, Q. A transferable machine-learning scheme from pure metals to alloys for predicting adsorption energies. *J. Mater. Chem. A* **2022**, *10*, 872–880.

(27) Zhao, Z.-J.; Liu, S.; Zha, S.; Cheng, D.; Studt, F.; Henkelman, G.; Gong, J. Theory-guided design of catalytic materials using scaling relationships and reactivity descriptors. *Nat. Rev. Mater.* **2019**, *4*, 792–804.

(28) Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem* **2009**, *1*, 37–46.

(29) Hammer, B.; Norskov, J. K. Why gold is the noblest of all the metals. *Nature* **1995**, *376*, 238–240.

(30) Calle-Vallejo, F.; Loffreda, D.; Koper, M. T.; Sautet, P. Introducing structural sensitivity into adsorption–energy scaling relations by means of coordination numbers. *Nat. Chem* **2015**, *7*, 403–410.

(31) Calle-Vallejo, F.; Tymoczko, J.; Colic, V.; Vu, Q. H.; Pohl, M. D.; Morgenstern, K.; Loffreda, D.; Sautet, P.; Schuhmann, W.; Bandarenka, A. S. Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science* **2015**, *350*, 185–189.

(32) Ma, X.; Xin, H. Orbitalwise coordination number for predicting adsorption properties of metal nanocatalysts. *Phys. Rev. Lett.* **2017**, *118*, 036101.

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

(34) Fernández, E. M.; Moses, P. G.; Toftelund, A.; Hansen, H. A.; Martínez, J. I.; Abild-Pedersen, F.; Kleis, J.; Hinnemann, B.; Rossmeisl, J.; Bligaard, T. et al. Scaling relationships for adsorption energies on transition metal oxide, sulfide, and nitride surfaces. *Angew. Chem. Int. Ed.* **2008**, *47*, 4683–4686.

(35) Abild-Pedersen, F.; Greeley, J.; Studt, F.; Rossmeisl, J.; Munter, T. R.; Moses, P. G.; Skulason, E.; Bligaard, T.; Nørskov, J. K. Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Phys. Rev. Lett.* **2007**, *99*, 016105.

(36) Hoffmann, R. A chemical and theoretical way to look at bonding on surfaces. *Rev. Mod. Phys.* **1988**, *60*, 601.

(37) Burdett, J. K. From bonds to bands and molecules to solids. *Prog. Solid State Chem. 15,* **1984**, *15*, 173–255.

(38) Hannagan, R. T.; Giannakakis, G.; Flytzani-Stephanopoulos, M.; Sykes, E. C. H. Single-Atom Alloy Catalysis. *Chem. Rev.* **2020**, *120*, 12044–12088.