




ARTICLE



<https://doi.org/10.1038/s41467-021-22048-9>

OPEN

Single-atom alloy catalysts designed by first-principles calculations and artificial intelligence

Zhong-Kang Han^{1,5}, Debalaya Sarker^{1,5}, Runhai Ouyang^{2,5}, Aliaksei Mazheika³ , Yi Gao⁴  & Sergey V. Levchenko¹ 

Single-atom-alloy catalysts (SAACs) have recently become a frontier in catalysis research. Simultaneous optimization of reactants' facile dissociation and a balanced strength of intermediates' binding make them highly efficient catalysts for several industrially important reactions. However, discovery of new SAACs is hindered by lack of fast yet reliable prediction of catalytic properties of the large number of candidates. We address this problem by applying a compressed-sensing data-analytics approach parameterized with density-functional inputs. Besides consistently predicting efficiency of the experimentally studied SAACs, we identify more than 200 yet unreported promising candidates. Some of these candidates are more stable and efficient than the reported ones. We have also introduced a novel approach to a qualitative analysis of complex symbolic regression models based on the data-mining method subgroup discovery. Our study demonstrates the importance of data analytics for avoiding bias in catalysis design, and provides a recipe for finding best SAACs for various applications.

¹Center for Energy Science and Technology, Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Moscow, Russia. ²Materials Genome Institute, Shanghai University, Shanghai, P.R. China. ³Technische Universität Berlin, BasCat—UniCat BASF JointLab, Berlin, Germany. ⁴Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, P.R. China. ⁵These authors contributed equally: Zhong-Kang Han, Debalaya Sarker, Runhai Ouyang. ✉email: alex.mazheika@gmail.com; gaoyi@zjlab.org.cn; S.Levchenko@skoltech.ru

Recently, single-atom dispersion has been shown to dramatically reduce the usage of rare and expensive metals in heterogeneous catalysis, at the same time providing unique possibilities for tuning catalytic properties^{1,2}. The pioneering work by Sykes and co-workers² has demonstrated that highly dilute bimetallic alloys, where single atoms of Pt-group are dispersed on the surface of an inert metal host, are highly efficient and selective in numerous catalytic reactions. These alloy catalysts are now extensively used in the hydrogenation-related reactions such as hydrogenation of CO₂, water–gas shift reaction, hydrogen separation, and many others^{3–5}. The outstanding performance of SAACs is attributed to a balance between efficiency of H₂ dissociation and binding of H at the surface of metallic alloys^{2,6,7}.

Using desorption measurements in combination with high-resolution scanning tunneling microscopy, Kyriakou et al. have shown that isolated Pd atoms on a Cu surface can substantially reduce the energy barrier for both hydrogen uptake and subsequent desorption from the Cu metal surface². Lucci and co-workers have observed that isolated Pt atoms on the Cu(111) surface exhibit stable activity and 100% selectivity for the hydrogenation of butadiene to butenes⁸. Liu et al. have investigated the fundamentals of CO adsorption on Pt/Cu SAAC using a variety of surface science and catalysis techniques. They have found that CO binds more weakly to single Pt atoms in Cu(111), compared to larger Pt ensembles or monometallic Pt. Their results demonstrate that SAACs offer a new approach to design CO-tolerant materials for industrial applications⁹. To date, Pd/Cu^{10–12}, Pt/Cu^{7–9,13–15}, Pd/Ag^{12,16}, Pd/Au¹², Pt/Au¹⁷, Pt/Ni¹⁸, Au/Ru¹⁹, and Ni/Zn²⁰ SAACs have been synthesized and found to be active and selective towards different hydrogenation reactions. However, the family of experimentally synthesized SAACs for hydrogenation remains small and comparisons of their catalytic properties are scarce.

Conventional approaches to designing single-atom heterogeneous catalysts for different industrially relevant hydrogenation reactions mainly rely on trial-and-error methods. However, challenges in synthesis and in situ experimental characterization of SAACs impose limitations on these approaches. With advances in first-principles methods and computational resources, theoretical modeling opens new opportunities for rational catalyst design^{6,21–48}. A general simple yet powerful approach is the creation of a large database with first-principles based inputs, followed by intelligent interrogation of the database in search of materials with the desired properties^{35,48}. Significant efforts have been made in developing reliable descriptor-based models following the above general approach^{6,21–35,48}. In catalysis, a descriptor is a parameter (a feature) of the catalytic material that is easy to evaluate and is correlated with a complex target property (e.g., activation energy or turnover frequency of a catalytic reaction). A notable amount of research has been devoted to searching for and using descriptors with a simple (near-linear) relation to target properties^{22–30}. For example, the linear relationship between the reaction energies and the activation energies is known as the Brønsted–Evans–Polanyi relationship (BEP) in heterogeneous catalysis^{29,30,45–47}. Also, the linear correlation between *d*-band center of a clean transition-metal surface and adsorption energies of molecules on that surface have been studied in great detail and widely applied^{22–24,36,44}. In catalysis, near-linear correlations between adsorption energies of different adsorbates are referred to as scaling relations^{26,28,37}. The advantages of such correlations are their simplicity and usually clear physical foundations. However, they are not exact, and there is an increasing number of studies focused on overcoming limitations imposed by the corresponding approximations^{6,31–34,38–41,48}. The nonlinear and intricate relationship between the catalysts' properties and surface reactions at realistic conditions^{42,43} has held

back the reliable description of catalytic properties. Note that, although the stability of SAACs is of no less significance in designing a potential catalyst than their catalytic performance, it hasn't received the same level of attention.

In this work, combining first-principles calculations and compressed-sensing data-analytics methodology, we address the issues that inhibit the wider use of SAAC in different industrially important reactions. By identifying descriptors based only on properties of the host surfaces and guest single atoms, we predict the binding energies of H (BE_H), the dissociation energy barriers of H₂ molecule (*E*_b), the segregation energies (SE) of the single guest atom at different transition metal surfaces, and the segregation energies in the presence of adsorbed hydrogen (SE_H). The state-of-the-art compressed-sensing based approach employed here for identifying the key descriptive parameters is the recently developed SISSO (sure independence screening and sparsifying operator)⁴⁹. SISSO enables us to identify the best low-dimensional descriptor in an immensity of offered candidates. The computational time required for our models to evaluate the catalytic properties of a SAAC is reduced by at least a factor of one thousand compared to first-principles calculations, which enables high-throughput screening of a huge number of SAAC systems.

Results and discussion

The BE_H for more than three hundred SAACs are calculated within the framework of DFT with RPBE exchange-correlation functional. This large dataset consists of BE_H values at different low-index surface facets including fcc(111), fcc(110), fcc(100), hcp(0001), and bcc(110) and three stepped surface facets including fcc(211), fcc(310), and bcc(210) of SAACs with twelve transition-metal hosts (Cu, Zn, Cr, Pd, Pt, Rh, Ru, Cd, Ag, Ti, Nb, and Ta). On each TM host surface, one of the surface atoms is substituted by a guest atom to construct the SAACs. BE_H for pristine surfaces (where the guest atom is the same with the host metal) are also included. H atom is placed at different non-equivalent high-symmetry sites close to the guest atom (Supplementary Fig. 1), and the BE_H for the most favorable site is included in the data set. Complete information on adsorption sites and the corresponding BE_H is given in Supplementary Data 1. The BE_H are further validated by a comparison with previous calculations^{6,21}.

To better understand the variation in BE_H for different guest atoms, we first investigate correlation between BE_H and the *d*-band center of the *d* orbitals that are projected to the single guest atom for the alloyed systems. We find that this way of calculating *d*-band center provides better correlation with other properties than *d*-band centers for the *d* orbitals projected on (i) the single guest atom plus its 1st nearest neighbor shell or (ii) the whole slab⁵⁰. The correlation is shown in Fig. 1a (Supplementary Fig. 2) for different SAACs on Ag(110) host surface [Pt(111) host surface]. According to the *d*-band center theory^{21,23,36,44}, the closer the *d*-band center is to the Fermi level, the stronger the BE_H should be. However, it is evident from Fig. 1a (Supplementary Fig. 2) that the expected linear correlation, as predicted by the *d*-band model, is broken for SAACs for H adsorption. This is due to the small size of the atomic H orbitals, leading to a relatively weak coupling between H *s* and the TM *d*-orbitals²¹. Furthermore, we check the validity of the BEP relations between the *E*_b and the H₂ dissociation reaction energy for SAACs (Fig. 1b), which is commonly used to extract kinetic data for a reaction on the basis of the adsorption energies of the reactants and products^{29,45–47}. As shown in Fig. 1b, the highlighted SAACs inside the blue dotted circle significantly reduce *E*_b while reducing reaction energy only moderately. As a result, SAACs provide small reaction energy and

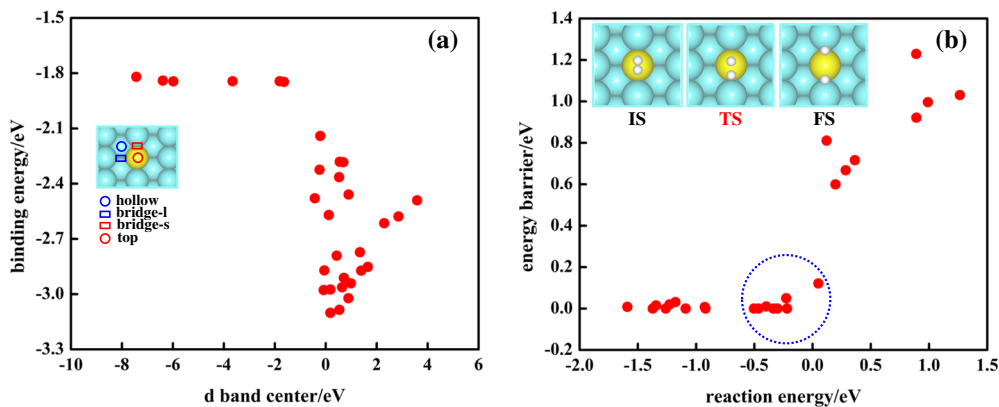


Fig. 1 Correlation between simple descriptors and target properties. Correlation between **a** H-atom binding energy BE_H and the d -band center and **b** the H_2 dissociation energy barrier E_b and the H_2 dissociation reaction energy for Ag(110) based SAACs. Only most stable adsorption sites are included (the hollow site for all systems on this plot). The SAACs inside the blue dotted circle in **b** significantly reduce E_b while reducing reaction energy only moderately.

Table 1 Primary features used for the descriptor construction.

System	Class	Name	Abbreviation
Host	Atomic	Energy of the highest-occupied Kohn-Sham level	H^*
		Energy of the lowest-unoccupied Kohn-Sham level	L^*
		Electron affinity (Atomic radius)	$EA^*(R^*)^a$
		Ionization potential	IP^*
		Binding energy of H with single host metal atom	$EH^*(EB^*)^a$
	Bulk	(Binding energy of host metal dimers)	
		Binding distance of H with single host metal atom	$dH^*(dd^*)^a$
		(Binding distance of host metal dimer)	
		Cohesive energy	EC^*
		d -band center	DC^*
Guest atom	Surface ^b	d -band center of the top surface layer	DT^*
		d -band center of the subsurface layer	DS^*
		Slab Fermi level	F^*
	Atomic	Energy of the highest-occupied Kohn-Sham level	H
		Energy of the lowest-unoccupied Kohn-Sham level	L
		Electron affinity (Atomic radius)	$EA(R)^a$
		Ionization potential	IP
		Binding energy of H with single guest metal atom	$EH(EB)^a$
	Bulk	(Binding energy of guest metal dimers)	
		Binding distance of H with single guest metal atom	$dH(dd)^a$
		(Binding distance of guest metal dimers)	
		Cohesive energy	EC
		d -band center	DC

^aThe feature in parentheses is used for the model of segregation energy (SE), while the feature outside parentheses is used for the models of H binding energy (BE_H) and H_2 dissociation energy barrier (E_b).

^bThe host metal-based features are marked by *. The surface-based primary features were calculated using the slab unit cell consisting of one atom per atomic layer.

low activation energy barrier, which leads to breaking BEP relations and thus optimized catalytic performance. The BEP relations are also found to be broken for other reactions catalyzed by SAACs⁶.

Thus, the standard simple correlations (from d -band center theory and the BEP relations) fail for H adsorption on SAACs. Moreover, the calculation of the d -band center for each SAAC is highly computationally demanding, considering the very large number of candidates. These facts emphasize the necessity to find new accurate, but low-cost descriptors for computational screening of SAACs. In the SISSO method, a huge pool of more than 10 billion candidate features is first constructed iteratively by combining 19 low-cost primary features listed in Table 1 using a set of mathematical operators. A compressed-sensing based procedure is used to select one or more most relevant candidate features and construct a linear model of the target property (see Supplementary Methods for details on the SISSO procedure).

Note that the three primary surface features are properties of the pure host surfaces (elemental metal systems). This is undoubtedly much more efficient than obtaining the properties of SAACs (alloyed metal systems). In the latter case, due to the interaction between the single guest atom and its images, a large supercell of the whole periodic system containing guest atom and host surface needs to be computed. On the contrary, only smallest unit cell is needed to compute the pristine surface features.

To test the predictive power of obtained models, we employ 10-fold cross validation (CV10). The dataset is first split into ten subsets, and the descriptor identification along with the model training is performed using nine subsets. Then the error in predicting properties of the systems in the remaining subset is evaluated with the obtained model^{51–53}. The CV10 error is defined as the average value of the test errors obtained for each of the ten subsets. In SISSO over-fitting may occur with increasing dimensionality of the descriptor (i.e., the number of complex

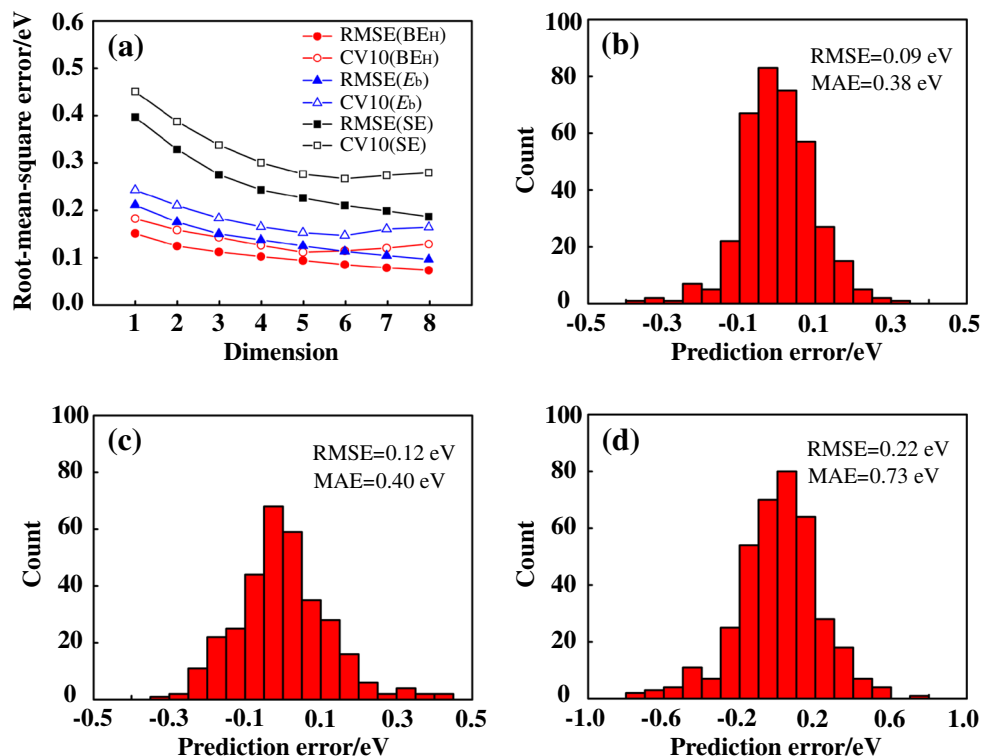


Fig. 2 SISSO errors and their distribution for different target properties. **a** RMSE and the averaged RMSE of the 10 fold cross-validation. **b–d** Distribution of errors for the best models versus RPBE results for BE_H (**b**), E_b (**c**), and SE (**d**).

Table 2 The identified descriptors and the coefficients and correlations in corresponding SISSO models for BE_H , E_b , and SE.

Property	d^m	Descriptor	Coefficient	Correlation
BE_H	d_1^5	$(EA^* + 2 F^* - EC) \cdot DT^* \cdot EH^* / (EC^* + F^*)$	0.12653E+00	0.8964
	d_2^5	$\sqrt[3]{DC} \cdot H^* \cdot DT^* \cdot (EA^* - EH^* - EC - EC^*)$	-0.20440E-02	0.5891
	d_3^5	$ EH^* - L^* - EH - F^* / (DC^2 + EC \cdot EC^*)$	-0.50891E+00	0.4850
	d_4^5	$ EH - F^* - EH^* - EC^* - EC - DT^* - F^* $	0.34705E-01	0.3849
E_b	d_5^5	$L \cdot EC \cdot (EA^* + DS^* - H - EH / L^* - EH^*)$	-0.48772E-04	0.3862
	d_6^6	$((IP^* - L) - EC^* - DT^*) / EC/DC - L^*/IP^* $	-0.87339E-01	0.7643
	d_7^6	$(EA^* + DC^* + DC - DT^*) / (EA^* + EH^* + L^* - F^*)$	-0.19577E-01	0.5726
	d_8^6	$(DC + EH^*) \cdot (EC^* - F^*) \cdot (L - EC - EC - EH)$	-0.13173E-01	0.4568
	d_9^6	$(DT^* - EH) \cdot DC \cdot (H/EC + EA^*/L^*)/EC^*$	-0.19172E-01	0.4414
	d_{10}^6	$e^{EC} \cdot EH \cdot DS^* / ((L^* - DS^*) + H^* - EC^*)$	0.33549E-01	0.3768
SE	d_{11}^6	$DC^2 \cdot (EC^* - F^*) / (DT^* - F^* - EA + EC)$	-0.14362E-02	0.3643
	d_{12}^6	$(EC + IP + F^* - DT^*) / (IP^*/R + H^*/dd^*)$	-0.82665E+00	0.8969
	d_{13}^6	$ DC - EB^* \cdot (L - DC - EC) / EB^2$	0.30742E+00	0.5346
	d_{14}^6	$ EC^* - L^* + DC - DS^* - DC - F^* - EC - F^* $	0.11317E+00	0.5386
	d_{15}^6	$ H - IP - L + IP^* / ((DC/EC) + (EC/H))$	0.17455E+00	0.3913
	d_{16}^6	$(F^* - EC) \cdot (L^* - DT^* - IP) / (F^* - EB^*)$	-0.51761E-02	0.3982
	d_{17}^6	$EC^* \cdot DC \cdot (EB^* - L) \cdot (L + L^* - EC - DS^*)$	-0.80032E-03	0.3379

features that are used in construction of the linear model)⁴⁹. The descriptor dimension at which the CV10 error starts increasing identifies the optimal dimensionality of the descriptor (details of the validation approach can be found in Supplementary Methods). For the optimal dimensionality, the same set of primary features is found during CV10 in 9, 8, and 8 cases for the SISSO models of BE_H , E_b , and SE, respectively. The root-mean-square errors (RMSE), together with the CV10 errors of the SISSO models for BE_H , E_b , and SE are displayed in Fig. 2a. The obtained optimal descriptor dimensionalities for BE_H , E_b , and SE of the SAACs are 5, 6, and 6, respectively. Distribution of errors for the best models versus RPBE results is displayed in Fig. 2b–d. The RMSE and maximum absolute error (MAE) of the models are

also shown. The error distributions for all the lower-dimensional models relative to the best ones are displayed in Supplementary Figs. 4–6.

From the Table 2 one can see that the d -band center features DC , DC^* , DT , DT^* , DS , and DS^* appear in every dimension of the descriptors for BE_H and E_b , consistent with the well-established importance of d -band center for adsorption at transition-metal surfaces^{21,23,36,44}. The cohesive energies of guest (EC) and host (EC*) bulk metals are selected in each dimension of the descriptor for SE. This is due to the fact that the segregation is driven by the imbalance of binding energy between host and guest–host atoms. Interestingly, most of the descriptor components include only simple mathematical operators (+, −, /, ||),

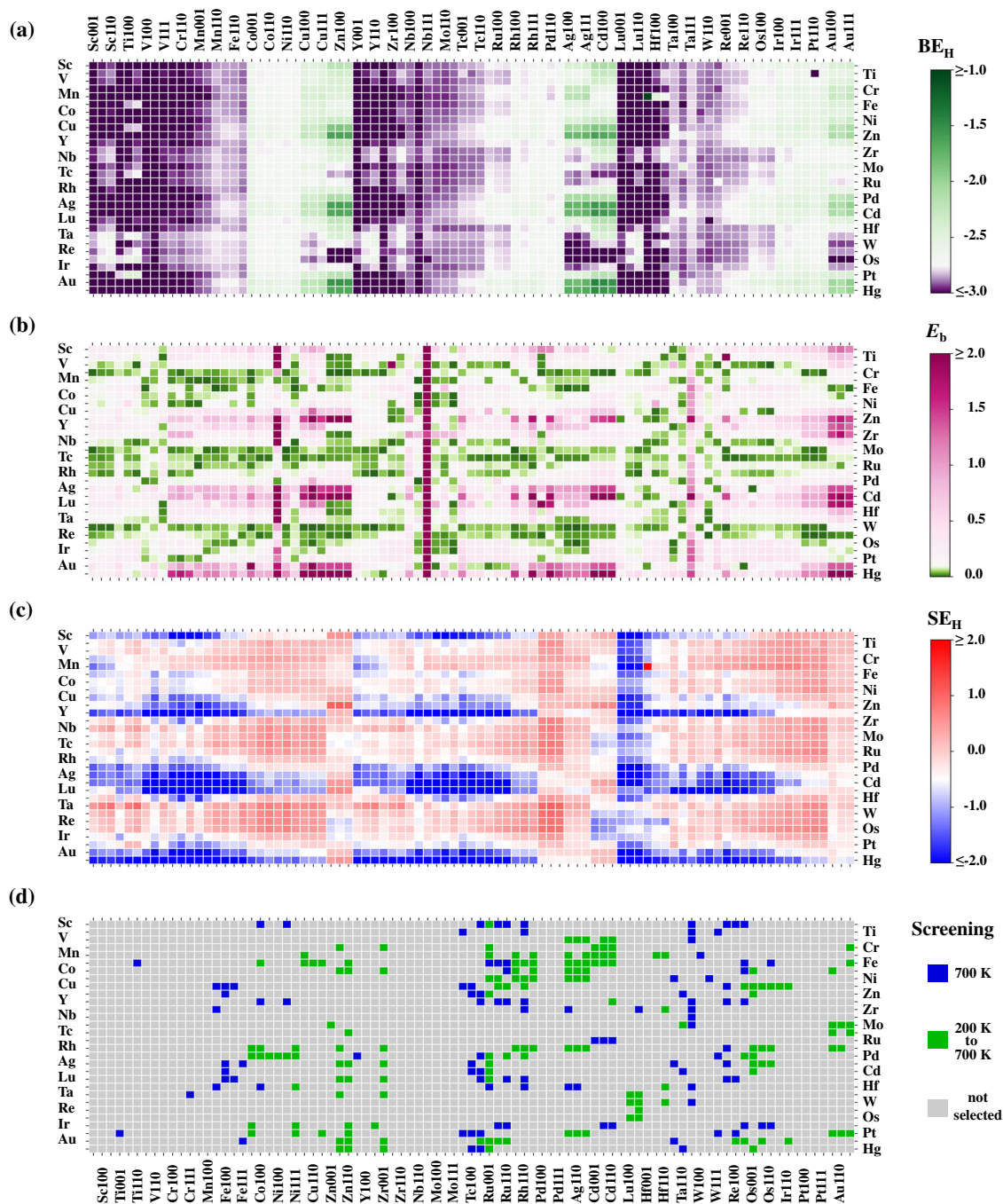


Fig. 3 Results of high-throughput screening of SAACs with SISSO. Results for **a** BE_H , **b** E_b , and **c** SE_H are shown. The promising candidates at different temperatures T are highlighted in **d**. Vertical axis displays the guest atom type, and the horizontal axis displays the host metal surfaces with different surface cuts.

indicating that the primary features already capture most of the complexity of the target properties.

We employ the identified computationally cheap SISSO models to perform high-throughput screening of SAACs to find the best candidates for the hydrogenation reactions. The results for BE_H , E_b , and SE_H (the segregation energy when surface H adatom is present, where the H adatom induced segregation energy change is included, see the “Methods” part for details) of the flat surfaces are displayed in Fig. 3a–c (see Supplementary Fig. 7 for the results for the stepped surfaces, the values of BE_H , E_b , and SE_H for all the SAACs are given in Supplementary Data 1).

The choice of the screening criteria for the three properties BE_H , E_b , and SE_H , which are related to the activity and stability of

SAACs, plays the central role in the screening processes and determines the candidates to be chosen. Previous work demonstrates that for the high performance in hydrogenation reactions, SAACs should exhibit weaker binding of H and lower H_2 dissociation energy barrier simultaneously². However, different criteria are applicable for different reaction conditions. For example, at low temperatures SAACs can maintain their stability for a longer time. At higher temperatures H atoms will desorb from the surfaces and larger energy barriers can be overcome, resulting in a requirement for stronger binding and higher upper limit of the dissociation barrier E_b . Keeping this variability in mind, we consider temperature-dependent and pressure-dependent selection criteria (see “Methods” section below for details on the

selection criteria). We have screened more than five thousand SAAC candidates (including about the same number of flat and stepped surfaces; the values of the primary features for all the candidates can be found in the Supplementary Data 2) at both low temperature (200 K) and high temperature (700 K) at partial H_2 pressure $p = 1$ atm. We find 160 flat-surface SAACs (Fig. 3d, in green) and 134 stepped-surface SAACs (Supplementary Fig. 7d, in green) that are both active and stable at a low temperature (200 K). At a higher temperature (700 K), 102 flat-surface SAACs (Fig. 3d, in blue and green) and 136 stepped-surface SAACs (Supplementary Fig. 7d, in blue and green) are classified as promising SAACs for hydrogenation reactions. Moreover, we have identified the SAACs that are promising in a wide range of temperatures (green squares in Fig. 3d for flat surfaces and Supplementary Fig. 7d for stepped surfaces).

Note that, without the stability selection criterion based on SE_H , all experimentally established SAACs (Pd/Cu, Pt/Cu, Pd/Ag, Pd/Au, Pt/Au, Pt/Ni, Au/Ru, and Ni/Zn) are predicted to be good catalysts in the temperature range of $200\text{ K} < T < 700\text{ K}$, which is further confirmed by DFT calculations. However, some of these systems (Pd/Ag and Pd/Au) are experimentally shown to have low stability^{12,16}. Thus, inclusion of the stability-related property SE_H is of immense importance for a reliable prediction of catalytic performance, as is confirmed by our results. We note that a machine-learning study on stability of single-atom metal alloys has recently been reported⁵⁴. However, our analysis takes into account effects of adsorbates on the segregation energy, which has not been considered previously. For example, the SE for Pd/Ag (110) and Pt/Ag(110) systems are 0.33 eV and 0.46 eV, respectively, implying that the Pd and Pt impurities tend to segregate into the bulk of the Ag(110) systems. However, SE_H for Pd/Ag (110) and Pt/Ag(110) systems are -0.10 eV and -0.21 eV , respectively, suggesting Pd and Pt impurities will segregate to the surface in the presence of H adatom. These results are also consistent with the experimental observations that the efficiency of Pd/Ag single-atom catalysts towards the selective hydrogenation of acetylene to ethylene was highly improved with the pre-treatment of the samples under H_2 conditions¹⁶.

We define an activity (or efficiency) indicator involving both the free energy of H adsorption (ΔG) and the energy barrier (E_b) as $\sqrt{\Delta G^2 + E_b^2}$ to construct an activity-stability map. As shown in Fig. 4, some of the new discovered candidates (bottom-left corner of activity-stability map) are predicted to have both higher stability and efficiency than the reported ones, making them optimized for practical applications (see Supplementary Fig. 8 for the results for the stepped surfaces). As expected, stability and activity are inversely related, which can be seen from the negative slope of the general trend in Supplementary Fig. 8 (showing selected materials) and Supplementary Fig. 9 (showing all explored materials), as well as a cut-off in population of the lower left-hand corner of these plots. Nevertheless, we have found several materials that are predicted to be better SAACs than the so-far reported ones. Considering stability, activity, abundance, and health/safety, two discovered best candidates Mn/Ag(111) and Pt/Zn(0001) are highlighted in Fig. 4. The aggregation energies for Mn/Ag(111), Pt/Zn(0001), and the experimentally established SAACs are also tested and displayed in Supplementary Table 9.

Although the SISSO models are analytic formulas, the corresponding descriptors are complex, reflecting the complexity of the relationship between the primary features and the target properties. While potentially interpretable, the models do not provide a straightforward way of evaluating relative importance of different features in actuating desirable changes in target properties. To facilitate physical understanding of the actuating mechanisms, we apply the subgroup discovery (SGD) approach^{55–60}. SGD finds local

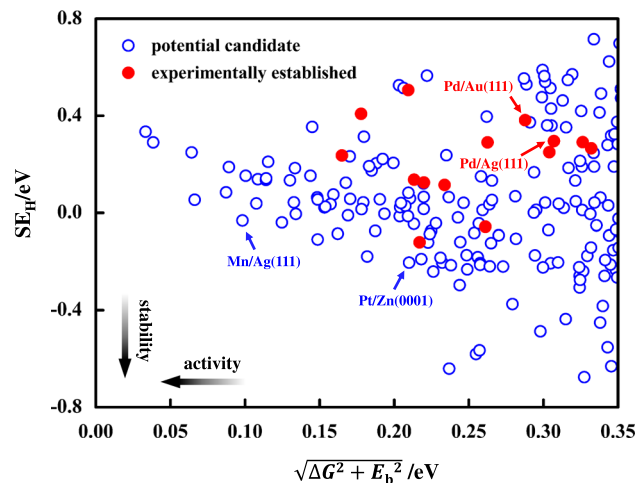


Fig. 4 Stability vs. activity map for flat SAACs surfaces at $T = 298\text{ K}$ and $p = 1\text{ atm}$. The SE_H on y-axis represents stability and activity parameter $\sqrt{\Delta G^2 + E_b^2}$ is shown on x-axis. Experimentally established SAACs are denoted with red solid spheres and the blue open circles represent new predicted candidates.

patterns in the data that maximize a quality function. The patterns are described as an intersection (a selector) of simple inequalities involving provided features, e.g., (feature1 < a1) AND (feature2 > a2) AND... The quality function is typically chosen such that it is maximized by subgroups balancing the number of data points in the subgroup, deviation of the median of the target property for the subgroup from the median for the whole data set, and the width of the target property distribution within the subgroup⁶⁰.

Here, we apply SGD in a novel context, namely as an analysis tool for symbolic regression models, including SISSO. The primary features that enter the complex SISSO descriptors of a given target property are used as features for SGD (see Table 2). The data set includes all 5200 materials and surfaces used in the high-throughput screening. The target properties are evaluated with the obtained SISSO models. Five target properties are considered: $\sqrt{\Delta G^2 + E_b^2}$, SE, SE_H , E_b , $|\Delta G|$, and BE_H . Since we are interested mainly in catalysts that are active at normal conditions, ΔG is calculated at $T = 300\text{ K}$. Our goal is to find selectors that minimize these properties within the subgroup. Such selectors describe actuating mechanisms for minimization of a given target property. For SE, the following best selector is found: ($EC^* \leq -3.85\text{ eV}$) AND ($-3.36\text{ eV} < EC \leq -0.01\text{ eV}$) AND ($IP \geq 7.45\text{ eV}$). The corresponding subgroup contains 738 samples (14% of the whole population), and the distribution of SE within the subgroup is shown in Supplementary Fig. 10. Qualitatively, the first two conditions imply that the cohesive energy of the host material is larger in absolute value than the cohesive energy of the guest material. Physically this means that bonding between host atoms is preferred over bonding between guest atoms and therefore over intermediate host–guest binding. This leads to the tendency of maximizing the number of host–host bonds by pushing guest atom to the surface. We note that this stabilization mechanism has been already discussed in literature⁶¹, and here we confirm it by data analysis. In addition, we find that stability of SAACs requires that the ionization potential of the guest atom is high. This can be explained by the fact that lower IP results in a more pronounced delocalization of the s valence electrons of the guest atom, and partial charge transfer to the surrounding host atoms. The charge transfer favors larger number of neighbors due to increased Madelung potential, and therefore destabilizes surface position of the guest atom.

We calculate SE_H using SISSO models for SE and BE_H [see Eq. (3) in the “Methods” section]. Therefore, SGD for SE_H is performed using primary features present in the descriptors of both SE and BE_H . The top subgroup contains features related to binding of H to the host and guest metal atoms, e.g., ($EB^* < -5.75$ eV) AND ($EH^* \leq -2.10$ eV) AND ($EH \geq -2.88$ eV) AND ($IP^* \leq 7.94$ eV) AND ($IP > 8.52$ eV) AND ($R \geq 1.29$ Å). However, the distribution of SE for this subgroup is very similar to the distribution of SE_H , which means that the stability of guest atoms at the surface is weakly affected by H adsorption when guest atoms are already very stable at the surface. The important effect of H adsorption is revealed when we find subgroups minimizing directly SE_H —SE (in this case only primary features that appear in the SISSO descriptor of BE_H are considered for SGD analysis). The top subgroup we found contains 72 samples (1.4% of the whole population) and is described by several degenerate selectors, in particular (-2.35 eV $\leq EH^* \leq -2.32$ eV) AND ($EC^* > -2.73$ eV) AND ($EC < -5.98$ eV) AND ($H \geq -5.12$ eV). This is a very interesting and intuitive result. Distributions of SE_H and SE for this subgroup are shown in Supplementary Fig. 11. The SE for all materials in the subgroup is above 0 eV. However, SE_H is much closer to 0 eV, and is below 0 eV for a significant number of materials in this subgroup. The conditions on the cohesive energy of guest and host metals (very stable bulk guest metal and less stable bulk host metal) are reversed with respect to SE, i.e., adsorption of hydrogen affects strongly the systems where guest atom is unstable at the surface. This increases the reactivity of the guest atom towards an H atom. The condition ($EH^* \geq -2.35$ eV) selects materials where interaction of H with a host atom is not too strong, so that H can bind with the guest atom and stabilize it at the surface. The condition ($EH^* \leq -2.32$ eV) makes the subgroup narrower, which further decreases median difference SE_H —SE but has no additional physical meaning. The condition ($H \geq -5.12$ eV) has a minor effect on the subgroup.

One of the top selectors (among several describing very similar data subsets) for minimizing $\sqrt{\Delta G^2 + E_b^2}$ (calculated at $T = 300$ K) is: (-2.85 eV $\leq DC \leq 1.95$ eV) AND ($DT^* \leq -0.17$ eV). The corresponding subgroup contains 1974 samples (38% of the whole population). The distribution of E_b within the subgroup is shown in Supplementary Fig. 10. The selector implies that systems providing low barrier for H_2 dissociation, and at the same time balanced binding of H atoms to the surface are characterized by (i) d -band center of the bulk guest metal around the Fermi level and (ii) d -band center of the host surface top layer below the Fermi level. This can be understood as follows. Condition (i) implies that there is a significant d -electron density that can be donated to the adsorbed H_2 molecule, facilitating its dissociation. A very similar (apart from slightly different numerical values) condition appears in the selector for the best subgroup for E_b target property alone [(-2.05 eV $\leq DC \leq 1.46$ eV) AND ($EC^* \geq -6.33$ eV)]. Condition (ii) implies that the surface d -band is more than half-filled, so that additional electrons are available for transferring to the H_2 molecule for its activation without causing excessive binding and therefore minimizing $|\Delta G|$ in accordance with Sabatier principle. Indeed, several subgroups of surfaces binding H atoms strongly (minimizing BE_H) are described by selectors including condition $DT^* > -0.17$, which is exactly opposite to condition (ii). Analysis of BE_H and $|\Delta G|$ also shows that the strong and intermediate binding of H atoms to the surface is fully controlled by the features of host material.

We note that SGD is capable of finding several alternative subgroups, corresponding to different mechanisms of actuating interesting changes in target properties. These subgroups have a lower quality according to the chosen quality function, but they still contain useful information about a particular mechanism. In

fact, they can be rigorously defined as top subgroups under additional constraint of zero overlap (in terms of data points) with previously found top subgroups. Analysis of such subgroups can be a subject of future work. We also note that quality function used in SGD is a parameter and can affect the found subgroups. It should be chosen based on the physical context of the problem. Exploring the role of different factors in the quality function and taking into account proposition degeneracy (no or minor effect of different conditions in the selectors due to correlation between the features) can significantly improve interpretability of the selectors. The interpretability also depends crucially on our physical understanding of the features and relations between them. Nevertheless, in combination with human knowledge SGD analysis allows for development of understanding, that would not be possible without the help of artificial intelligence.

In summary, by combining first-principles calculations and the data-analytics approach SISSO, we have identified accurate and reliable models for the description of the hydrogen binding energy, dissociation energy, and guest-atom segregation energy for SAACs, which allow us to make fast yet reliable prediction of the catalytic performance of thousands SAACs in hydrogenation reactions. The model correctly evaluates performance of experimentally tested SAACs. By scanning more than five thousand SAACs with our model, we have identified over two hundred new SAACs with both improved stability and performance compared to the existing ones. We have also introduced a novel approach to a qualitative analysis of complex SISSO descriptors using data-mining method subgroup discovery. It allows us to identify actuating mechanisms for desirable changes in the target properties, e.g., reaction barrier reduction or an increase in catalyst's stability, in terms of basic features of the material. Our methodology can be easily adapted to designing new functional materials for various applications.

Methods

All first-principles calculations are performed with the revised Perdew-Burke-Ernzerhof (RPBE) functional⁶² as implemented in the all-electron full-potential electronic-structure code FHI-aims⁶³. The choice of functional is validated based on a comparison of calculated H_2 adsorption energies to the available experimental results⁶⁴ (see Supplementary Table 1). Nevertheless, it is expected that, because of the large set of systems inspected and the small variations introduced by the functional choice, the main trends will hold even when using another functional (see Supporting Information for more details on the computational setup). The climbing-image nudged elastic band (CI-NEB) algorithm is employed to identify the transition state structures⁶⁵.

BE_H are calculated using Eq. (1), where $E_{H/support}$ is the energy of the total H/support system, $E_{support}$ is the energy of the metal alloy support, and E_H is the energy of an isolated H atom.

$$BE_H = E_{H/support} - E_{support} - E_H \quad (1)$$

The surface segregation energy in the dilute limit, SE, is defined as the energy difference of moving the single impurity from the bulk to the surface. In this work, it is calculated using Eq. (2), where $E_{top-layer}$ and $E_{nth-layer}$ correspond to the total RPBE energies of the slab with the impurity in the top and n th surface layer, respectively. The value of n is chosen so that the energy difference between $E_{nth-layer}$ and $E_{(n-1)th-layer}$ is less than 0.05 eV.

$$SE = E_{surface} - E_{nth-layer} \quad (2)$$

The surface segregation energy when surface H adatom is present (the H is put at the most stable adsorption site for each system), SE_H , is calculated using Eq. (3).

$$SE_H = SE + \Delta E_H, \quad (3)$$

where $\Delta E_H = BE_{H-top-layer} - BE_{H-pure}$ is the H adatom-induced segregation energy change.

Here $BE_{H-top-layer}$ and BE_{H-pure} are the hydrogen adatom binding energies with the impurity in the top layer and the BE_H of the pure system without impurity. Thus, the SE_H can be derived from the models of SE and BE_H .

Using first-principles inputs as training data, we have employed SISSO to single out a physically interpretable descriptor from a huge number of potential candidates. In practice, a huge pool of more than 10 billion candidate descriptors is first constructed iteratively by combining user-defined primary features with a set of

mathematical operators. The number of times the operators are applied determines the complexity of the resulting descriptors. We consider up to three levels of complexity (feature spaces) Φ_1 , Φ_2 , and Φ_3 . Note that a given feature space Φ_n also contains all of the lower rung (i.e., $n - 1$) feature spaces. Subsequently, the desired low-dimensional representation is obtained from this pool⁴⁹. The details of the feature space (Φ_n) construction and the descriptor identification processes can be found in the Supplementary Methods. The proper selection of primary features is crucial for the performance of SISSO-identified descriptors. Inspired by previous studies^{31,38}, we consider three classes of primary features (see Table 1) related to the metal atom, bulk, and surface. The more detailed description and values of all the primary features are given in the Supplementary Table 2, Supplementary Table 3, Supplementary Data 1, and Supplementary Data 2.

The selection of the promising candidates at various temperatures and hydrogen partial pressures is performed based on ab initio atomistic thermodynamics⁶⁶. H adsorption/desorption on SAAC surfaces as a function of temperature and H_2 partial pressure (T , p) is characterized by the free energy of adsorption ΔG :

$$\Delta G = E_{H/\text{support}} - E_{\text{support}} - \mu_H(T, p) \quad (4)$$

with the chemical potential of hydrogen $\mu_H = \frac{1}{2}\mu_{H_2}$ obtained from:

$$\mu_H = \frac{1}{2} \left(E_{H_2} + \Delta\mu_{H_2}(T, p) \right), \quad (5)$$

where $\Delta\mu_{H_2}(T, p) = \mu_{H_2}(T, p^0) - \mu_{H_2}(T^0, p^0) + k_B T \ln\left(\frac{p}{p^0}\right)$.

Here $T^0 = 298$ K and $p^0 = 1$ atm. The first two terms are taken from JANAF thermochemical tables⁶⁷. In the following, we set $p = 1$ atm.

According to Sabatier principle the optimum heterogeneous catalyst should bind the reactants strong enough to allow for adsorption, but also weak enough to allow for the consecutive desorption²⁵. In this work, a BE_H range is defined by the conditions:

$$|BE_H - \frac{1}{2}(E_{H_2} - 2E_H) - \frac{1}{2}\Delta\mu_{H_2}(T)| < 0.3 \text{ eV}, \quad (6)$$

where $E_{H_2} - 2E_H$ is the hydrogen binding energy of the hydrogen molecule. The experimental value of -4.52 eV⁶⁸ was used in this work.

The above conditions correspond to the free-energy bounds:

$$|\Delta G| < 0.3 \text{ eV}, \quad (7)$$

Conditions on energy barrier (E_b) are defined by considering Arrhenius-type behaviour of the reaction rate on E_b and T . Assuming that acceptable barriers are below 0.3 eV for $T^0 = 298$ K, we estimate acceptable barrier at any temperature as:

$$E_b < \frac{0.3T}{T^0} \text{ eV}. \quad (8)$$

Similarly the bounds for SE_H are determined by imposing a minimum 10% ratio for top-layer to subsurface-layers dopant concentration by assuming an Arrhenius-type relation with SE_H interpreted as activation energy:

$$SE_H < k_B T \ln(10). \quad (9)$$

The subgroup discovery was performed using RealkD package (<https://bitbucket.org/realKD/realKD/>). Each feature was split to 15 subsets using 15-means clustering algorithm. The borders between adjacent data clusters (a_1, a_2, \dots) are applied further for construction of inequalities (feature1 < a_1), (feature2 $\geq a_2$), etc. While final result might depend on the number of considered clusters, in our previous study we found that relatively high numbers of considered clusters provide essentially the same result⁶⁰. The candidate subgroups are built as conjunctions of obtained simple inequalities. The main idea of SGD is that the subgroups are unique if the distribution of the data in them is as different as possible from the data distribution in the whole sampling. Here the data distribution is the distribution of a target property ($\sqrt{\Delta G^2 + E_b^2}$, SE , E_b , $|\Delta G|$, and BE_H). The uniqueness is evaluated with a quality function. In this study we used the following function:

$$Q(S) = \frac{s(S)}{s(P)} \left(\frac{\text{med}(P) - \text{med}(S)}{\text{med}(P) - \text{min}(P)} \right) \left(1 - \frac{\text{amd}(S)}{\text{amd}(P)} \right) \quad (10)$$

with S —subgroup, P —whole sampling, s —size, med and min —median and minimal values of a target property, amd —absolute average deviation of the data around the median of target property. With this function the algorithm is searching for subgroups with lower values of target properties. The search was done with an adapted for such purposes Monte-Carlo algorithm⁵⁹, in which first a certain number of trial conjunctions (seeds) is generated. Afterwards, for each seed (accompanied with pruning of inequalities) the quality function is calculated. We have tested here several numbers of initial seeds: 10,000, 30,000, 50,000, and 100,000. The subgroups with the overall high quality function value were selected.

Data availability

All relevant data are available from the corresponding authors upon reasonable request.

Code availability

FHI-aims: <https://aimsclub.fhi-berlin.mpg.de>.

SISSO: <https://github.com/rouyang2017/SISSO>.

SGD: <https://bitbucket.org/realKD/realKD/>.

Received: 24 May 2020; Accepted: 8 February 2021;

Published online: 23 March 2021

References

- Qiao, B. et al. Single-atom catalysis of CO oxidation using Pt 1/FeO x. *Nat. Chem.* **3**, 634 (2011).
- Kyriakou, G. et al. Isolated metal atom geometries as a strategy for selective heterogeneous hydrogenations. *Science* **335**, 1209–1212 (2012).
- Choi, K. I. & Vannice, M. A. CO oxidation over Pd and Cu catalysts V. Al₂O₃-supported bimetallic Pd• Cu particles. *J. Catal.* **131**, 36–50 (1991).
- Greeley, J., Nørskov, J. K., Kibler, L. A., El-Aziz, A. M. & Kolb, D. M. Hydrogen evolution over bimetallic systems: Understanding the trends. *ChemPhysChem* **7**, 1032–1035 (2006).
- Kamakoti, P. et al. Prediction of hydrogen flux through sulfur-tolerant binary alloy membranes. *Science* **307**, 569–573 (2005).
- Darby, M. T., Réocreux, R., Sykes, E. C. H., Michaelides, A. & Stamatakis, M. Elucidating the stability and reactivity of surface intermediates on single-atom alloy catalysts. *ACS Catal.* **8**, 5038–5050 (2018).
- Sun, G. et al. Breaking the scaling relationship via thermally stable Pt/Cu single atom alloys for catalytic dehydrogenation. *Nat. Commun.* **9**, 1–9 (2018).
- Lucci, F. R. et al. Selective hydrogenation of 1, 3-butadiene on platinum–copper alloys at the single-atom limit. *Nat. Commun.* **6**, 1–8 (2015).
- Liu, J. et al. Tackling CO poisoning with single-atom alloy catalysts. *J. Am. Chem. Soc.* **138**, 6396–6399 (2016).
- Tierne, H. L., Baber, A. E. & Sykes, E. C. H. Atomic-scale imaging and electronic structure determination of catalytic sites on Pd/Cu near surface alloys. *J. Phys. Chem. C* **113**, 7246–7250 (2009).
- Boucher, M. B. et al. Single atom alloy surface analogs in Pd 0.18 Cu 15 nanoparticles for selective hydrogenation reactions. *Phys. Chem. Chem. Phys.* **15**, 12187–12196 (2013).
- Pei, G. X. et al. Performance of Cu-alloyed Pd single-atom catalyst for semihydrogenation of acetylene under simulated front-end conditions. *ACS Catal.* **7**, 1491–1500 (2017).
- Marcinkowski, M. D. et al. Selective formic acid dehydrogenation on Pt-Cu single-atom alloys. *ACS Catal.* **7**, 413–420 (2017).
- Simonovis, J. P., Hunt, A., Palomino, R. M., Senanayake, S. D. & Waluyo, I. Enhanced stability of Pt-Cu single-atom alloy catalysts: in situ characterization of the Pt/Cu (111) surface in an ambient pressure of CO. *J. Phys. Chem. C* **122**, 4488–4495 (2018).
- Marcinkowski, M. D. et al. Pt/Cu single-atom alloys as coke-resistant catalysts for efficient C–H activation. *Nat. Chem.* **10**, 325 (2018).
- Pei, G. X. et al. Ag alloyed Pd single-atom catalysts for efficient selective hydrogenation of acetylene to ethylene in excess ethylene. *ACS Catal.* **5**, 3717–3725 (2015).
- Duchesne, P. N. et al. Golden single-atomic-site platinum electrocatalysts. *Nat. Mater.* **17**, 1033–1039 (2018).
- Li, Z. et al. Atomically dispersed Pt on the surface of Ni particles: synthesis and catalytic function in hydrogen generation from aqueous ammonia–borane. *ACS Catal.* **7**, 6762–6769 (2017).
- Chen, C. H. et al. Ruthenium-based single-atom alloy with high electrocatalytic activity for hydrogen evolution. *Adv. Energy Mater.* **9**, 1803913 (2019).
- Studt, F. et al. Identification of non-precious metal alloy catalysts for selective hydrogenation of acetylene. *Science* **320**, 1320–1322 (2008).
- Greeley, J. & Mavrikakis, M. Alloy catalysts designed from first principles. *Nat. Mater.* **3**, 810–815 (2004).
- Pallassana, V., Neurock, M., Hansen, L. B., Hammer, B. & Nørskov, J. K. Theoretical analysis of hydrogen chemisorption on Pd (111), Re (0001) and Pd ML/R e (0001), Re ML/P d (111) pseudomorphic overlayers. *Phys. Rev. B* **60**, 6146 (1999).
- Mavrikakis, M., Hammer, B. & Nørskov, J. K. Effect of strain on the reactivity of metal surfaces. *Phys. Rev. Lett.* **81**, 2819 (1998).
- 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* **89**, 115114 (2014).
- Nørskov, J. K., Bligaard, T., Rossmeisl, J. & Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem.* **1**, 37 (2009).
- Montemore, M. M. & Medlin, J. W. Scaling relations between adsorption energies for computational screening and design of catalysts. *Catal. Sci. Technol.* **4**, 3748–3761 (2014).
- Greeley, J. Theoretical heterogeneous catalysis: scaling relationships and computational catalyst design. *Annu. Rev. Chem. Biomol. Eng.* **7**, 605–635 (2016).
- Abild-Pedersen, F. et al. Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Phys. Rev. Lett.* **99**, 016105 (2007).

29. Michaelides, A. et al. Identification of general linear relationships between activation energies and enthalpy changes for dissociation reactions at surfaces. *J. Am. Chem. Soc.* **125**, 3704–3705 (2003).
30. Logadottir, A. et al. The Brønsted–Evans–Polanyi relation and the volcano plot for ammonia synthesis over transition metal catalysts. *J. Catal.* **197**, 229–231 (2001).
31. Andersen, M., Levchenko, S. V., Scheffler, M. & Reuter, K. Beyond scaling relations for the description of catalytic materials. *ACS Catal.* **9**, 2752–2759 (2019).
32. 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.* **7**, 403 (2015).
33. O'Connor, N. J., Jonayat, A., Janik, M. J. & Senftle, T. P. Interaction trends between single metal atoms and oxide supports identified with density functional theory and statistical learning. *Nat. Catal.* **1**, 531–539 (2018).
34. Xu, H., Cheng, D., Cao, D. & Zeng, X. C. A universal principle for a rational design of single-atom electrocatalysts. *Nat. Catal.* **1**, 339–348 (2018).
35. Curtarolo, S. et al. The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).
36. Hammer, B. & Nørskov, J. K. Why gold is the noblest of all the metals. *Nature* **376**, 238–240 (1995).
37. Roling, L. T. & Abild-Pedersen, F. Structure-sensitive scaling relations: adsorption energies from surface site stability. *ChemCatChem* **10**, 1643–1650 (2018).
38. Li, Z., Wang, S., Chin, W. S., Achenie, L. E. & Xin, H. High-throughput screening of bimetallic catalysts enabled by machine learning. *J. Mater. Chem. A* **5**, 24131–24138 (2017).
39. Tran, K. & Ulissi, Z. W. Active learning across intermetallics to guide discovery of electrocatalysts for CO₂ reduction and H₂ evolution. *Nat. Catal.* **1**, 696–703 (2018).
40. Kitchin, J. R. Machine learning in catalysis. *Nat. Catal.* **1**, 230–232 (2018).
41. Li, Z., Wang, S. & Xin, H. Toward artificial intelligence in catalysis. *Nat. Catal.* **1**, 641–642 (2018).
42. Reuter, K. Ab initio thermodynamics and first-principles microkinetics for surface catalysis. *Catal. Lett.* **146**, 541–563 (2016).
43. Reuter, K., Frenkel, D. & Scheffler, M. The steady state of heterogeneous catalysis, studied by first-principles statistical mechanics. *Phys. Rev. Lett.* **93**, 116105 (2004).
44. Nørskov, J. Catalysis—calculations and concepts. *Adv. Catal.* **45**, 71 (2001).
45. Van Santen, R. A., Neurock, M. & Shetty, S. G. Reactivity theory of transition-metal surfaces: a Brønsted–Evans–Polanyi linear activation energy–free-energy analysis. *Chem. Rev.* **110**, 2005–2048 (2009).
46. Fajin, J. L., Cordeiro, M. N. D., Illas, F. & Gomes, J. R. Generalized Brønsted–Evans–Polanyi relationships and descriptors for O–H bond cleavage of organic molecules on transition metal surfaces. *J. Catal.* **313**, 24–33 (2014).
47. Viñes, F., Vojvodic, A., Abild-Pedersen, F. & Illas, F. Brønsted–Evans–Polanyi relationship for transition metal carbide and transition metal oxide surfaces. *J. Phys. Chem. C* **117**, 4168–4171 (2013).
48. Zhong, M. et al. Accelerated discovery of CO₂ electrocatalysts using active machine learning. *Nature* **581**, 178–183 (2020).
49. Ouyang, R., Curtarolo, S., Ahmetcik, E., Scheffler, M. & Ghiringhelli, L. M. SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. *Phys. Rev. Mater.* **2**, 083802 (2018).
50. Thirumalai, H. & Kitchin, J. R. Investigating the reactivity of single atom alloys using density functional theory. *Top. Catal.* **61**, 462–474 (2018).
51. Bengio, Y. & Grandvalet, Y. No unbiased estimator of the variance of k-fold cross-validation. *J. Mach. Learn. Res.* **5**, 1089–1105 (2004).
52. Rodriguez, J. D., Perez, A. & Lozano, J. A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 569–575 (2009).
53. Wong, T.-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognit.* **48**, 2839–2846 (2015).
54. Rao, K. K., Do, Q. K., Pham, K., Maiti, D. & Grabow, L. C. Extendable machine learning model for the stability of single atom alloys. *Top. Catal.* **63**, 728–741 (2020).
55. Wrobel, S. *European Symposium on Principles of Data Mining and Knowledge Discovery* 78–87 (Springer, 1997).
56. Friedman, J. H. & Fisher, N. I. Bump hunting in high-dimensional data. *Stat. Comput.* **9**, 123–143 (1999).
57. Atzmueller, M. Subgroup discovery. *Wiley Interdiscip. Rev.* **5**, 35–49 (2015).
58. Boley, M., Goldsmith, B. R., Ghiringhelli, L. M. & Vreeken, J. Identifying consistent statements about numerical data with dispersion-corrected subgroup discovery. *Data Min. Knowl. Discov.* **31**, 1391–1418 (2017).
59. Goldsmith, B. R., Boley, M., Vreeken, J., Scheffler, M. & Ghiringhelli, L. M. Uncovering structure-property relationships of materials by subgroup discovery. *N. J. Phys.* **19**, 013031 (2017).
60. Mazheika, A. et al. Ab initio data-analytics study of carbon-dioxide activation on semiconductor oxide surfaces. <https://arxiv.org/1912.06515> (2019).
61. Chelikowsky, J. R. Predictions for surface segregation in intermetallic alloys. *Surf. Sci.* **139**, L197–L203 (1984).
62. Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew–Burke–Ernzerhof functionals. *Phys. Rev. B* **59**, 7413 (1999).
63. Blum, V. et al. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **180**, 2175–2196 (2009).
64. Silbaugh, T. L. & Campbell, C. T. Energies of formation reactions measured for adsorbates on late transition metal surfaces. *J. Phys. Chem. C* **120**, 25161–25172 (2016).
65. Henkelman, G., Uberuaga, B. P. & Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901–9904 (2000).
66. Reuter, K., Stampf, C. & Scheffler, M. in *Handbook of Materials Modeling* 149–194 (Springer, 2005).
67. Tables, J. T. *JANAF Thermochemical Tables*, (eds. Stull, D. R. & Prophet, H.) (National Bureau of Standards Publication, 1971).
68. Lide, D. R. *CRC Handbook of Chemistry and Physics: A Ready-Reference Book of Chemical and Physical Data* (CRC Press, 1995).

Acknowledgements

S.V.L. is supported by Skolkovo Foundation Grant. The machine-learning methodology development was funded by RFBR and INSF, project number 20-53-56065. Y.G. is supported by the National Natural Science Foundation of China (11604357, 11574340). R.O. is supported by the National Key Research and Development Program of China (2018YFB0704400) and the Program of Shanghai Youth Oriental Scholars.

Author contributions

S.V.L. created the idea and conceived the work. S.V.L. and Y.G. designed and supervised the project. S.V.L. and A.M. supervised the SGD analysis. R.O. supervised the SISSO analysis. Z.-K.H. performed all the calculations. Z.-K.H. and D.S. wrote the manuscript with inputs from all the authors. All authors contributed to the analysis and interpretation of the results. All the authors commented on the manuscript and have given approval to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-22048-9>.

Correspondence and requests for materials should be addressed to A.M., Y.G. or S.V.L.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

Supplementary Information to the manuscript: Single-Atom Alloy Catalysts Designed by First-Principles Calculations and Artificial Intelligence

Zhong-Kang Han,^{1#} Debalaya Sarker,^{1#} Runhai Ouyang,^{2#} Aliaksei Mazheika,^{3*} Yi Gao,^{4*} Sergey V. Levchenko^{1*}

¹ Center for Energy Science and Technology, Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Moscow, 143026, Russia

² Materials Genome Institute, Shanghai University, 333 Nanchen Road, Shanghai, 200444, P.R. China

³ Technische Universität Berlin, BasCat – UniCat BASF JointLab, Hardenbergstraße 36, 10623 Berlin, Germany

⁴ Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, 201210, P. R. China

[#] These authors contributed equally.

Correspondence Email: S.Levchenko@skoltech.ru; gaoyi@zjlab.org.cn; alex.mazheika@gmail.com

Supplementary Methods

All *ab initio* calculations were performed with the revised Perdew-Burke-Ernzerhof (RPBE) functional¹ as implemented in the all-electron full-potential electronic-structure code FHI-aims² using density-functional theory (DFT) and numerical atom-centered basis functions. *Light* numerical settings are used which are tested to be converged compared to *tight* numerical settings, resulting in accuracy in total energy differences within 0.05 eV per supercell. The choice of functional is validated based on a comparison of calculated H₂ adsorption energies to the available experimental results³ (Supplementary Table 1). Spin-polarization effects are tested and included where appropriate for the system with Fe, Co, and Ni atoms. Slabs of at least nine metal layers were considered with the two to four bottom layers fixed, based on the convergence of BE_H and SE (within 0.05 eV) with respect to the thickness and supercell size of the slab. The lattice vector along the direction parallel to the vacuum gap was 50 Å. All atoms in the systems except for the fixed bottom have been allowed to relax until the maximum remaining force fell below 10⁻² eV/Å. The climbing-image nudged elastic band (CI-NEB) algorithm is employed to identify the transition state structures.⁴ H atom is placed at different non-equivalent high-symmetry sites close to the guest atom (Supplementary Figure 1), and the BE_H for the most favorable site is included in the data set. The host metal surfaces considered in this work to construct the training data set are Cu(100), Cu(310), Zn(0001), Cr(110), Pd(111), Pd(211), Pt(111), Rh(111), Ru(0001), Cd(0001), Ag(100), Ag(110), Ti(0001), Nb(210), and Ta(210) with more than three hundred points for each considered properties. All the DFT calculated BE_H, SE, and *E_b* can be found in the file “Supplementary Data 1”.

For constructing the Φ_1 , Φ_2 , and Φ_3 feature spaces we made use of the set of algebraic/functional operators given in eq. 1.

$$\hat{H}^{(m)} \equiv \{+, -, \cdot, /, \log, \exp, \exp^{-1}, ^2, ^3, \sqrt{\cdot}, ^3\sqrt{\cdot}, |-\| \}, \quad (1)$$

The superscript m indicates that when applying $\hat{H}^{(m)}$ to primary features φ_1 and φ_2 a dimensional analysis is performed, which ensures that only physically meaningful combinations are retained (e.g. only primary features with the same unit are added or subtracted). All primary features included in this study were obtained either from the literature (see Supplementary Table 2), or from DFT calculations (see Supplementary Table 3). The values of the primary features for the training data sets can be found in the file “Supplementary Data 1” and the values of the primary features for all the high-throughput screening SAAC candidates can be found in the file “Supplementary Data 2”.

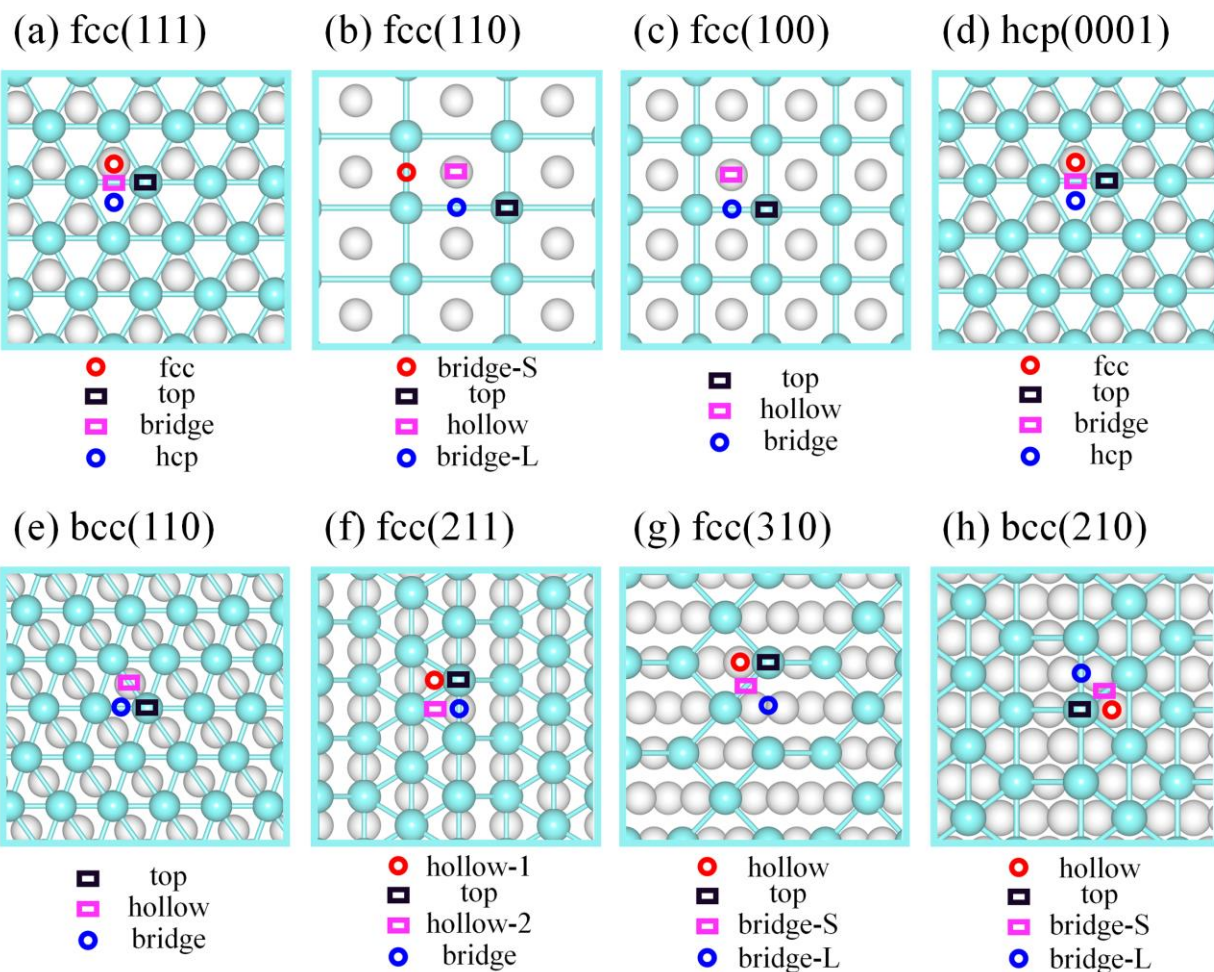
The sparsifying ℓ_0 constraint is applied to a smaller feature subspace selected by a screening procedure (sure independence screening (SIS)), where the size of the subspace is equal to a user-defined SIS value times the dimension of the descriptor. The SIS value is not an ordinary hyperparameter and its optimization through a validation data set is not straightforward. Ideally, one would want to search the entire feature space for the optimal descriptor. However, this is not computationally tractable since the computational cost of the sparsifying ℓ_0 constraint grows exponentially with the size of the searched feature space. Instead, the SIS value should be chosen as large as computationally possible. The reasonable SIS values were chosen based on the convergence of the training error.

To confirm the reliability of the SISO model optimization approaches, the data were initially divided into training and test sets. The Pd(211) and Pt(111) based systems are used as test set while all the other data were contained in the training set. As mentioned in the main text, 10-fold cross validation (CV10) method was used to determine the dimensionality of the descriptor. First, the best descriptors were selected by SISO based on only the training data. The RMSR and CV10 errors on the training set for the descriptors of BE_H , E_b , and SE are displayed in Supplementary Figure 3a. Second, the predictive power of the SISO selected descriptors was tested using the test set. We display in Supplementary Figure 3b the distribution of errors on the training set and test set for the descriptors of BE_H (green box), E_b (blue box), and SE (cyan box). To check the predictive power of SISO selected descriptors on different types of surfaces, we divide the test set into two groups: one with only flat surface and the other one with only stepped surface. In the first group, we considered a new transition metal of Pt which are not contained in the training set (Cu, Zn, Cr, Pd, Rh, Ru, Cd, Ag, Ti, Nb, and Ta), while in the second group, we considered a new surface cut of fcc(211) which is also not included in the training set. The root-mean-square errors, RMSEs, (maximum absolute errors, MAEs) of the SISO selected descriptors for the test flat surface set are found to be BE_H : 0.10 eV (0.43 eV), E_b : 0.12 eV (0.62 eV), and SE: 0.22 eV (0.78 eV), while for the test stepped surface set are found to be BE_H : 0.11 eV (0.54 eV), E_b : 0.14 eV (0.71 eV), and SE: 0.24 eV (0.87 eV). The moderate errors of the RMSEs for both test flat surface set and test stepped surface set showed that the transferability of the descriptors is good. The little larger RMSEs and MAEs for test stepped surface set compared to that for test flat surface set can be rationalized by the fact that in the training set, more flat surfaces are considered compared to stepped surfaces. We would expect the errors for the test set to further decrease when more data are included in the training set. Thus, after confirming the reliability of the of the SISO model optimization approaches, finally we included all the data into the training set for the SISO selection of the best descriptors. As mentioned in the main text, the obtained optimal descriptor dimensionalities for BE_H , E_b , and SE of the SAACs are 5, 6, and 6, respectively. The same optimal descriptor dimensionalities for BE_H , E_b , and SE were found when the Pt(111) and Pd(211) based

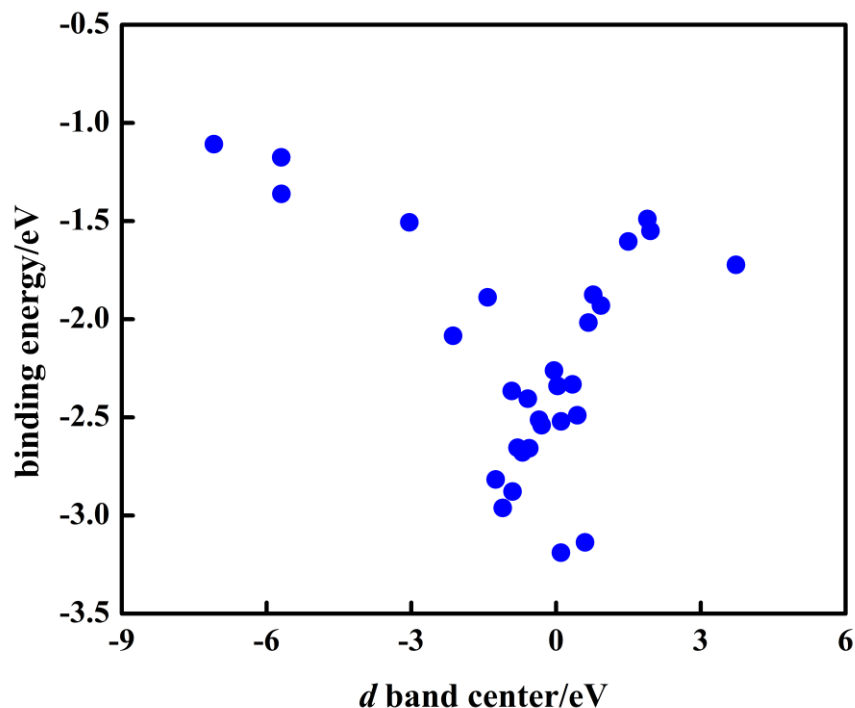
systems are not included, further confirming the reliability the used SISSO model optimization approaches. The error distributions for all the lower dimensional models are displayed in Supplementary Figures 4-6. The RMSE and MAE of the models are also shown. The identified lower-dimensional descriptors and the coefficients and correlations in corresponding SISSO models for BE_H , E_b , and SE are displayed in Supplementary Tables 4-6. The top five largest deviations between calculated and predicted BE_H , E_b , and SE are collected in Supplementary Table 7. As can be seen, the deviation for SE is larger than that of BE_H and E_b . However, we found even for SE_H our model's precision is higher than 95% (Supplementary Table 8).

Supplementary Table 1. The experimental and theoretical adsorption energies (in eV) of H_2 at different transition metal surfaces.

method	Pt(111)	Ru(0001)	Pd(111)
RPBE	-0.56	-1.06	-1.01
PBE	-0.32	-0.78	-0.70
Experiments ³	-0.75	-1.20	-0.91



Supplementary Figure 1. The considered hydrogen adsorption sites on the fcc(111) (a), fcc(110) (b), fcc(100) (c), hcp(0001) (d), bcc(110) (e), fcc(211) (f), fcc(310) (g), bcc(210) (h) of pure transition metal surfaces. The atom below the top site for each surface facet is either the host atom (pure transition metal surfaces) or the single guest atom (single atom alloy metal surfaces).



Supplementary Figure 2. Correlation between H-atom binding energy BE_H and the d -band center for Pt(111) based SAACs.

Supplementary Table 2. Primary features obtained from the literature.⁵ Electron affinity (EA in eV), ionization potential (IP in eV), and covalent radius (R) of the metal atom.

system	class	name	abbreviation
host	atomic	Electron affinity	EA*
		Ionization potential	IP*
		Atomic radius	R*
guest atom	atomic	Electron affinity	EA
		Ionization potential	IP
		Atomic radius	R

Supplementary Table 3. Primary features obtained from DFT-RPBE calculations (spin-polarization effects are tested for and included where appropriate). Energy of the highest-occupied Kohn-Sham level (H in eV), energy of the lowest-unoccupied Kohn-Sham level (L in eV), binding energy of H with isolated metal atom (E_H in eV as calculated by equation (2)), binding energy of

metal dimers (EB in eV as calculated by equation (3)), binding distance of H with isolated metal atom (dH in Å), and binding distance of metal dimers of the metal atom; cohesive energy (EC in eV as calculated by equation (4)) and *d*-band center (DC in eV) of the bulk metal; *d*-band center of the top surface layer (DT in eV), *d*-band center of the subsurface layer (DS in eV), and the slab Fermi level (F in eV) of the metal surface.

system	class	name	abbreviation
host	atomic	Energy of the highest-occupied Kohn-Sham level	H*
		Energy of the lowest-unoccupied Kohn-Sham level	L*
		Binding energy of H with single host metal atom	EH*
		Binding energy of host metal dimers	EB*
		Binding distance of H with single host metal atom	dH*
		Binding distance of host metal dimer	dd*
	bulk	Cohesive energy	EC*
		<i>d</i> -band center	DC*
	surface [#]	<i>d</i> -band center of the top surface layer	DT*
		<i>d</i> -band center of the subsurface layer	DS*
		Slab Fermi level	F*
guest atom	atomic	Energy of the highest-occupied Kohn-Sham level	H
		Energy of the lowest-unoccupied Kohn-Sham level	L
		Binding energy of H with single guest metal atom	EH
		Binding energy of guest metal dimers	EB
		Binding distance of H with single guest metal atom	dH
		Binding distance of guest metal dimers	dd
	bulk	Cohesive energy	EC
		<i>d</i> -band center	DC

[#]the surface based primary features were calculated by using the unit cell consisting of one atom per atomic layer.

$$EH = E_{H-metal} - E_{metal} - E_H \quad (2)$$

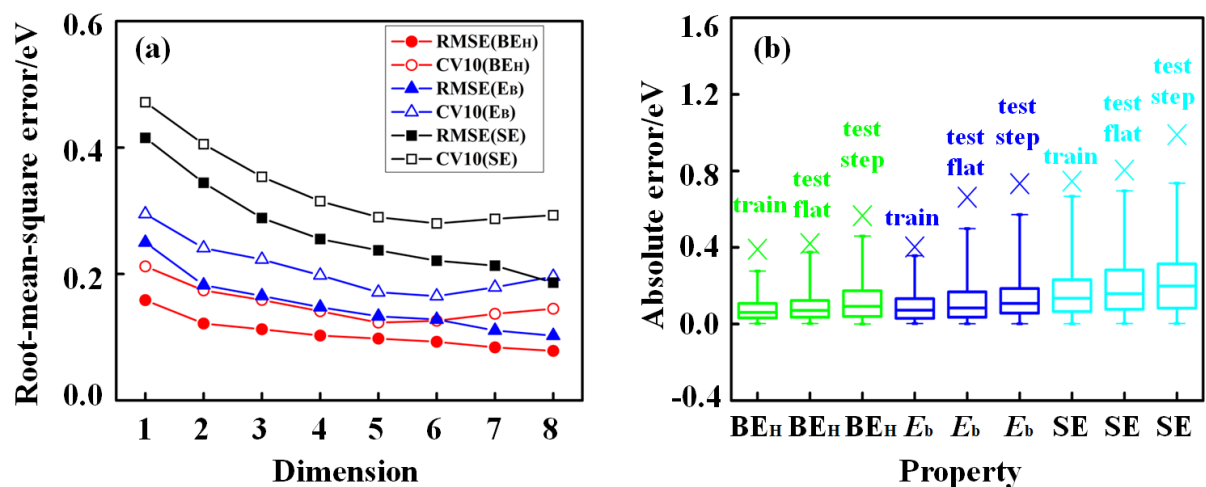
where $E_{H-metal}$ is the energy of the total H-metal system, E_{metal} is the energy of the isolated metal atom, and E_H is the energy of the isolated H atom.

$$EB = E_{dimer} - 2E_{single} \quad (3)$$

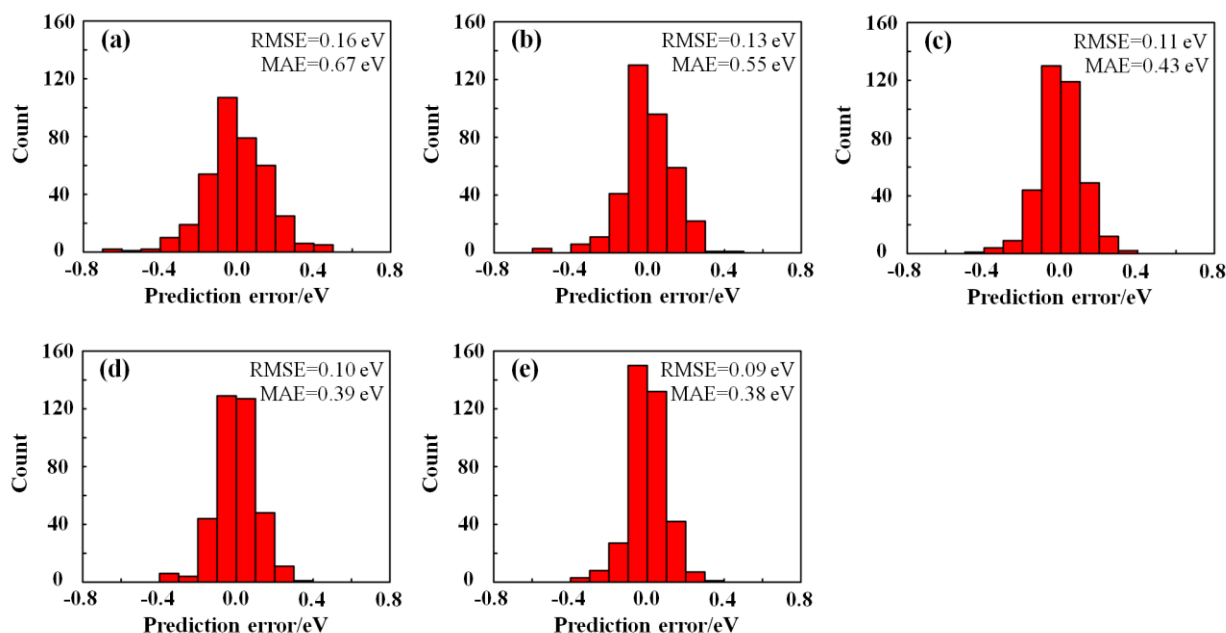
where E_{dimer} is the energy of the total metal-metal dimer system, E_{single} is the energy of the isolated metal atom.

$$EC = E_{bulk/n} - E_{single} \quad (4)$$

where E_{bulk} is the energy of the bulk metal system, E_{single} is the energy of the isolated metal atom, and n is the number of metal atoms in the bulk unit cell.



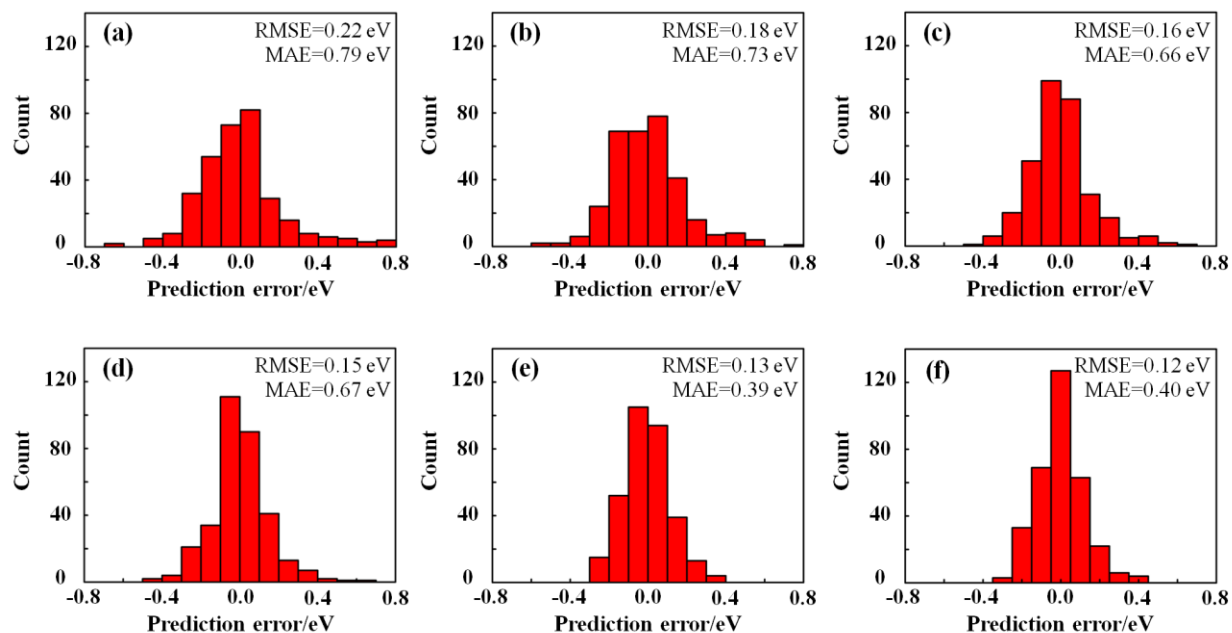
Supplementary Figure 3. (a) RMSE and the averaged RMSE of the 10 fold cross-validation. (b) Box plots of the absolute errors for the training set and test set for the SISSO selected best models of BE_H (green), E_b (blue), and SE (cyan). The test set is divided into two parts: one part contains only flat surfaces and the other one contains only stepped surfaces. The upper and lower limits of the rectangles represent the 75th and 25th percentiles of the distribution, the internal horizontal lines mark the median (50th percentile), and the upper and lower limits of the error bars indicate the 99th and 1st percentiles. The crosses depict the maximum absolute errors.



Supplementary Figure 4. The error distributions for all the lower-dimensional models of BE_H: (a) 1D, (b) 2D, (c) 3D, and (d) 4D, and the SISSO selected best model (e) 5D.

Supplementary Table 4. The identified descriptors and the coefficients and correlations for all the lower-dimensional models of BE_H.

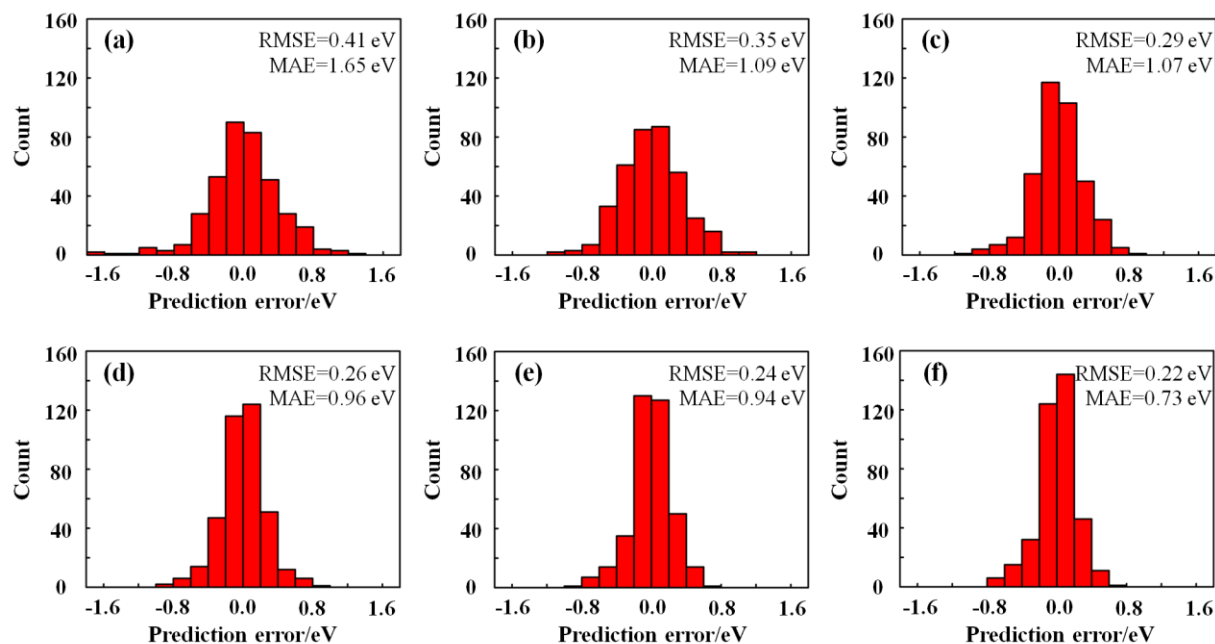
d^m		descriptor	coefficient	correlation
d^1	d_1^1	$(EA^*+2F^*-EC)\cdot DT^*\cdot EH^*/(EC^*+F^*)$	0.12286E+00	0.8964
d^2	d_1^2	$(EA^*+2F^*-EC)\cdot DT^*\cdot EH^*/(EC^*+F^*)$	0.12093E+00	0.8964
	d_2^2	$\sqrt[3]{DC}\cdot H^*\cdot DT^*\cdot (EA^*-EH^* - EC-EC^*)$	-0.20756E-02	0.5891
d^3	d_1^3	$(EA^*+2F^*-EC)\cdot DT^*\cdot EH^*/(EC^*+F^*)$	0.12393E+00	0.8964
	d_2^3	$\sqrt[3]{DC}\cdot H^*\cdot DT^*\cdot (EA^*-EH^* - EC-EC^*)$	-0.19673E-02	0.5891
	d_3^3	$ EH^*-L^*- EH-F^* /(DC^2+EC\cdot EC^*)$	-0.56460E+00	0.4850
d^4	d_1^4	$(EA^*+2F^*-EC)\cdot DT^*\cdot EH^*/(EC^*+F^*)$	0.11932E+00	0.8964
	d_2^4	$\sqrt[3]{DC}\cdot H^*\cdot DT^*\cdot (EA^*-EH^* - EC-EC^*)$	-0.18876E-02	0.5891
	d_3^4	$ EH^*-L^*- EH-F^* /(DC^2+EC\cdot EC^*)$	-0.60955E+00	0.4850
	d_4^4	$ IP^*\cdot F^*/(EH+F^*)- EC^*-DS^* - DC^*-F^* $	0.31619E-01	0.3864
d^5	d_1^5	$(EA^*+2F^*-EC)\cdot DT^*\cdot EH^*/(EC^*+F^*)$	0.12653E+00	0.8964
	d_2^5	$\sqrt[3]{DC}\cdot H^*\cdot DT^*\cdot (EA^*-EH^* - EC-EC^*)$	-0.20440E-02	0.5891
	d_3^5	$ EH^*-L^*- EH-F^* /(DC^2+EC\cdot EC^*)$	-0.50891E+00	0.4850
	d_4^5	$ EH-F^*-EH^* - EC^*-EC - DT^*-F^* $	0.34705E-01	0.3849
	d_5^5	$L\cdot EC\cdot (EA^*+DS^*- H-EH / L^*-EH^* $	-0.48772E-04	0.3862



Supplementary Figure 5. The error distributions for all the lower dimensional models of E_b : (a) 1D, (b) 2D, (c) 3D, (d) 4D, (e) 5D and the SISSO selected best model (f) 6D.

Supplementary Table 5. The identified descriptors and the coefficients and correlations for all the lower-dimensional models of E_b .

d^m		descriptor	coefficient	correlation
d^1	d_1^1	$((IP^*-L)- EC^*-DT^*)/ EC/DC-L^*/IP^* $	0.92945E-01	0.7643
d^2	d_1^2	$((IP^*-L)- EC^*-DT^*)/ EC/DC-L^*/IP^* $	0.98762E-01	0.7643
	d_2^2	$(EA^*+DC^*+ DC-DT^*)/(EA^*+EH^*+ L^*-F^*)$	-0.23925E-01	0.5726
d^3	d_1^3	$((IP^*-L)- EC^*-DT^*)/ EC/DC-L^*/IP^* $	0.91192E-01	0.7643
	d_2^3	$(EA^*+DC^*+ DC-DT^*)/(EA^*+EH^*+ L^*-F^*)$	-0.22522E-01	0.5726
	d_3^3	$(DC+EH^*)\cdot(EC^*-F^*)\cdot(L-EC - EC-EH)$	-0.14550E-01	0.4568
d^4	d_1^4	$((IP^*-L)- EC^*-DT^*)/ EC/DC-L^*/IP^* $	0.90281E-01	0.7643
	d_2^4	$(EA^*+DC^*+ DC-DT^*)/(EA^*+EH^*+ L^*-F^*)$	-0.21259E-01	0.5726
	d_3^4	$(DC+EH^*)\cdot(EC^*-F^*)\cdot(L-EC - EC-EH)$	-0.14089E-01	0.4568
	d_4^4	$(DT^*-EH)\cdot DC\cdot(H/EC+EA^*/L^*)/EC^*$	-0.18463E-01	0.4414
d^5	d_1^5	$((IP^*-L)- EC^*-DT^*)/ EC/DC-L^*/IP^* $	0.84777E-01	0.7643
	d_2^5	$(EA^*+DC^*+ DC-DT^*)/(EA^*+EH^*+ L^*-F^*)$	-0.18289E-01	0.5726
	d_3^5	$(DC+EH^*)\cdot(EC^*-F^*)\cdot(L-EC - EC-EH)$	-0.13888E-01	0.4568
	d_4^5	$(DT^*-EH)\cdot DC\cdot(H/EC+EA^*/L^*)/EC^*$	-0.19258E-01	0.4414
	d_5^5	$DC^2\cdot DT^*/IP\cdot(L^*-DS^*+ H^*-EC^*)$	-0.57167E-02	0.3975
d^6	d_1^6	$((IP^*-L)- EC^*-DT^*)/ EC/DC-L^*/IP^* $	-0.87339E-01	0.7643
	d_2^6	$(EA^*+DC^*+ DC-DT^*)/(EA^*+EH^*+ L^*-F^*)$	-0.19577E-01	0.5726
	d_3^6	$(DC+EH^*)\cdot(EC^*-F^*)\cdot(L-EC - EC-EH)$	-0.13173E-01	0.4568
	d_4^6	$(DT^*-EH)\cdot DC\cdot(H/EC+EA^*/L^*)/EC^*$	-0.19172E-01	0.4414
	d_5^6	$e^{EC}\cdot EH\cdot DS^*/((L^*-DS^*)+ H^*-EC^*)$	0.33549E-01	0.3768
	d_6^6	$DC^2\cdot(EC^*-F^*)/(DT^*-F^*-EA+EC)$	-0.14362E-02	0.3643



Supplementary Figure 6. The error distributions for all the lower dimensional models of SE: (a) 1D, (b) 2D, (c) 3D, (d) 4D, (e) 5D and the SISSO selected best model (f) 6D.

Supplementary Table 6. The identified descriptors and the coefficients and correlations for all the lower-dimensional models of SE.

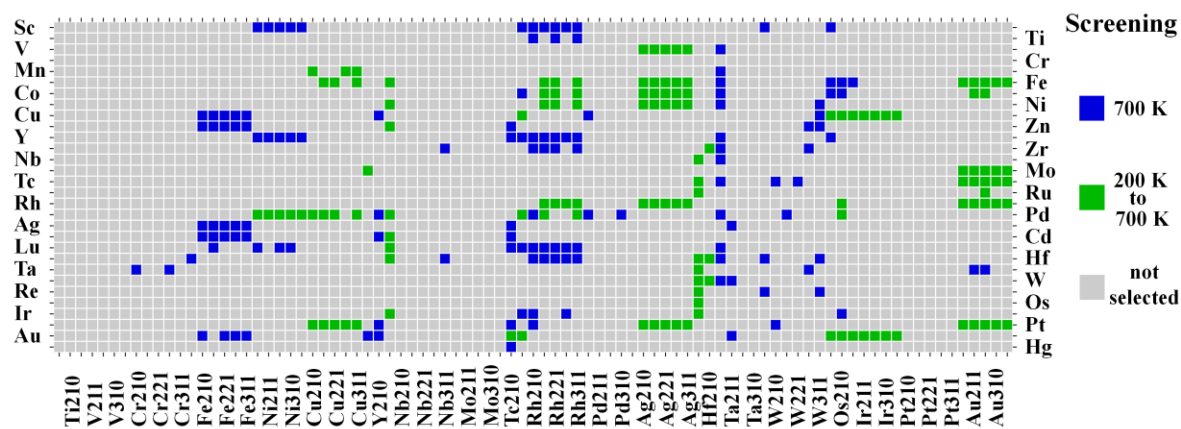
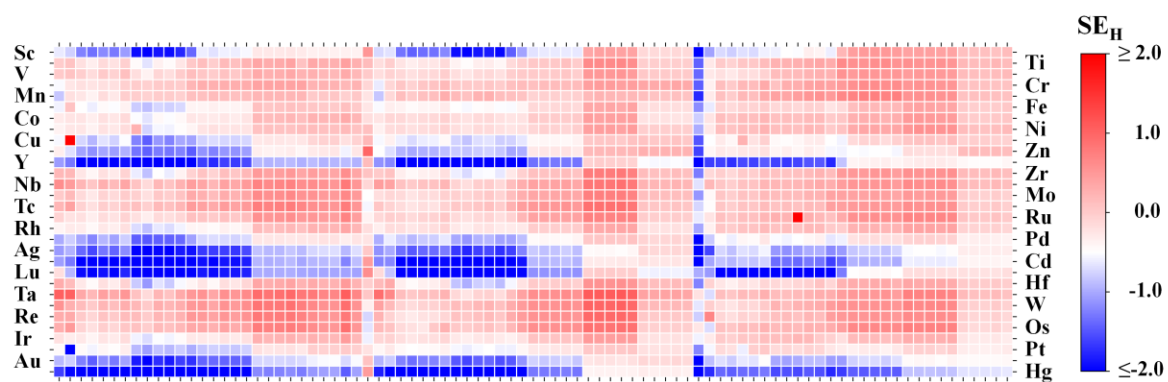
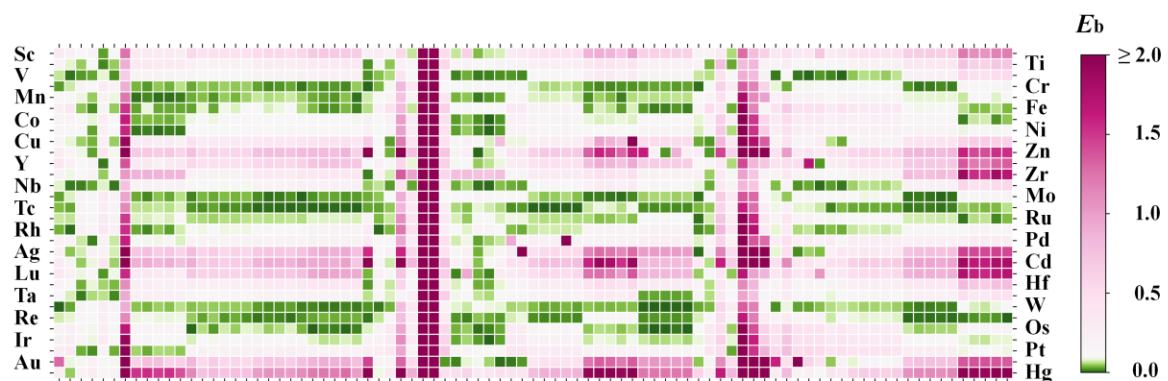
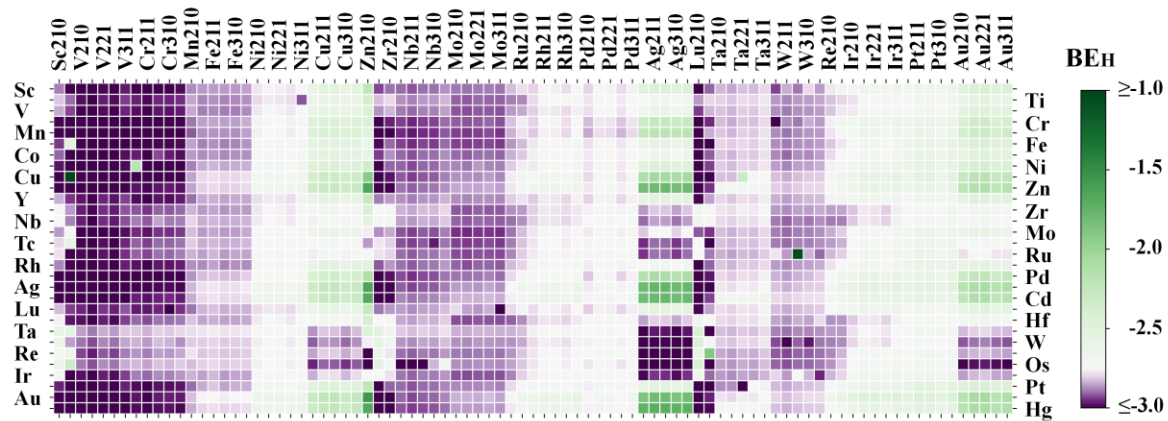
d^m		descriptor	coefficient	correlation
d^1	d_1^1	$(EC+IP+ F^*-DT^*)/(IP^*/R+H^*/dd^*)$	-0.75479E+00	0.8969
d^2	d_1^2	$(EC+IP+ F^*-DT^*)/(IP^*/R+H^*/dd^*)$	-0.75372E+00	0.8969
	d_2^2	$ DC-EB^* \cdot (L-DC-EC)/EB^2$	0.21861E+00	0.5346
d^3	d_1^3	$(EC+IP+ F^*-DT^*)/(IP^*/R+H^*/dd^*)$	-0.81257E+00	0.8969
	d_2^3	$ DC-EB^* \cdot (L-DC-EC)/EB^2$	0.24095E+00	0.5346
	d_3^3	$\ EC^*-L^* + DC-DS^* - DC-F^* - EC-F^*\ $	0.12052E+00	0.5386
d^4	d_1^4	$(EC+IP+ F^*-DT^*)/(IP^*/R+H^*/dd^*)$	-0.82171E+00	0.8969
	d_2^4	$ DC-EB^* \cdot (L-DC-EC)/EB^2$	0.30402E+00	0.5346
	d_3^4	$\ EC^*-L^* + DC-DS^* - DC-F^* - EC-F^*\ $	0.11671E+00	0.5386
	d_4^4	$ H-IP-L+IP^* /((DC/EC)+(EC/H))$	0.16261E+00	0.3913
d^5	d_1^5	$(EC+IP+ F^*-DT^*)/(IP^*/R+H^*/dd^*)$	-0.82332E+00	0.8969
	d_2^5	$ DC-EB^* \cdot (L-DC-EC)/EB^2$	0.30382E+00	0.5346
	d_3^5	$\ EC^*-L^* + DC-DS^* - DC-F^* - EC-F^*\ $	0.10926E+00	0.5386
	d_4^5	$ H-IP-L+IP^* /((DC/EC)+(EC/H))$	0.16324E+00	0.3913
	d_5^5	$(F^*-EC) \cdot (L^*-DT^*-IP)/(F^*-EB^*)$	-0.55996E-02	0.3982
d^6	d_1^6	$(EC+IP+ F^*-DT^*)/(IP^*/R+H^*/dd^*)$	-0.82665E+00	0.8969
	d_2^6	$ DC-EB^* \cdot (L-DC-EC)/EB^2$	0.30742E+00	0.5346
	d_3^6	$\ EC^*-L^* + DC-DS^* - DC-F^* - EC-F^*\ $	0.11317E+00	0.5386
	d_4^6	$ H-IP-L+IP^* /((DC/EC)+(EC/H))$	0.17455E+00	0.3913
	d_5^6	$(F^*-EC) \cdot (L^*-DT^*-IP)/(F^*-EB^*)$	-0.51761E-02	0.3982
	d_6^6	$EC^* \cdot DC \cdot (EB^*-L) \cdot (L+L^*-EC-DS^*)$	-0.80032E-03	0.3379

Supplementary Table 7. The top five largest deviations between calculated and predicted BE_H (in eV), E_b (in eV), and SE (in eV).

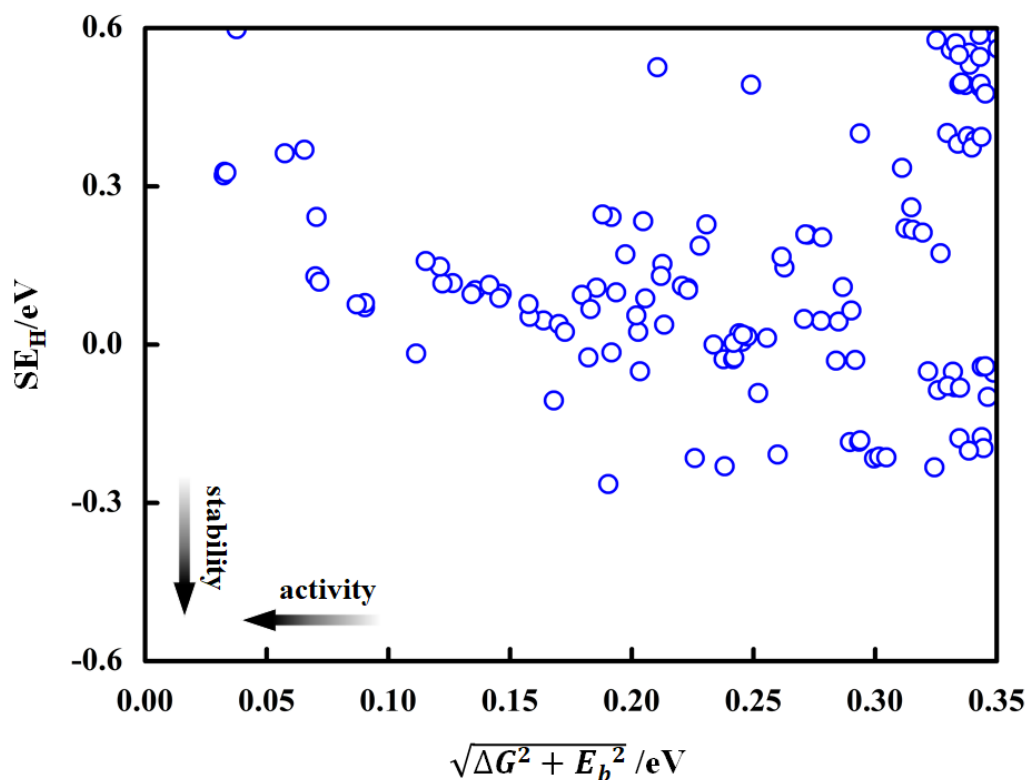
property	system	calculated	predicted	deviation
BE_H	Co/Ag(100)	-2.98	-2.60	0.38
	Co/Ag(110)	-2.98	-2.65	0.33
	Ni/Ti(0001)	-2.95	-3.27	0.33
	Ta/Ti(0001)	-3.19	-2.89	0.30
	Sc/Zn(0001)	-1.91	-2.21	0.30
E_b	Hf/Pt(111)	0.78	0.38	0.40
	Cu/Ag(110)	0.92	0.52	0.40
	Ir/Zr(0001)	0.68	0.28	0.40
	Pd/Ag(110)	0.60	0.24	0.36
	Zr/Cu(100)	0.71	0.38	0.33
SE	V/Pd(111)	-0.27	0.46	0.73
	Hg/Zn(0001)	-0.24	0.48	0.71
	Os/Zn(0001)	0.96	0.25	0.71
	Cd/Zn(0001)	-0.13	0.50	0.62
	Hg/Cd(0001)	-0.19	0.43	0.62

Supplementary Table 8. Number of systems with the predicted and calculated SE_H that meet the same condition of $SE_H < kT \ln(10)$ (N_{meet}), the total number of calculated systems (N_{total}), and the SE model's precision ($P = N_{\text{meet}} / N_{\text{total}}$).

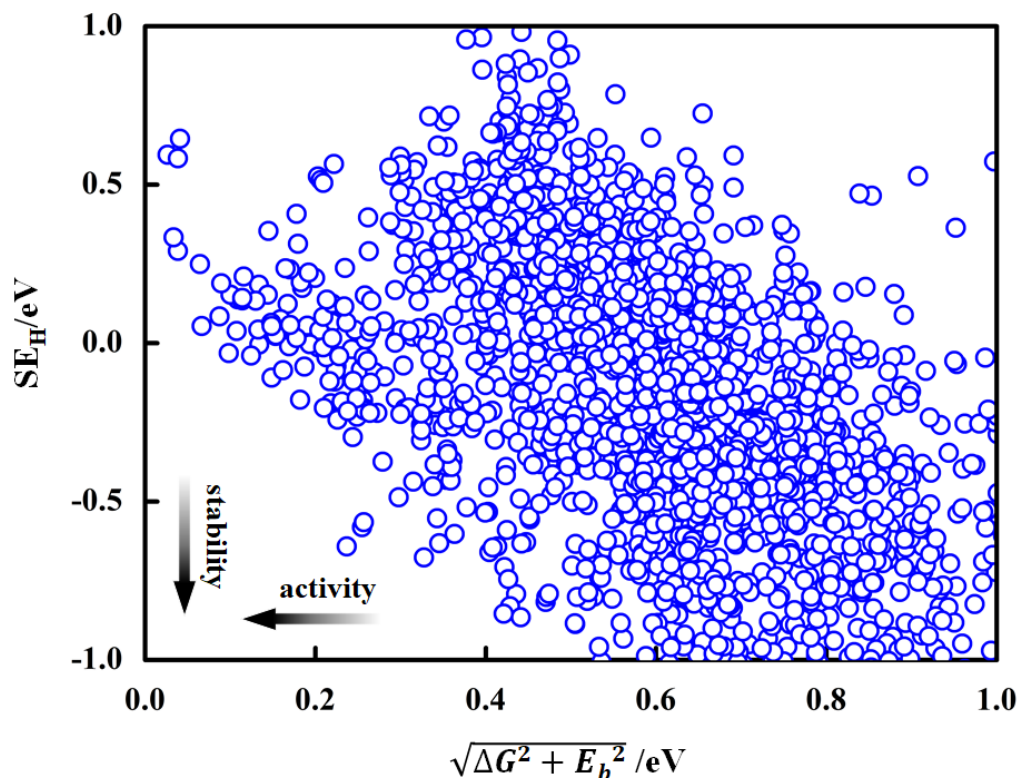
temperature	N_{meet}	N_{total}	P
200 K	345	360	95.83%
700 K	346	360	96.11%



Supplementary Figure 7. High-throughput screening of SAACs for (a) BE_H , (b) E_b , and (c) SE_H . The screened candidates are highlighted in (d). Vertical axis displays the guest atom type, and the vertical horizontal axis displays the host metal surfaces with different stepped surface cuts.



Supplementary Figure 8. Stability vs. activity map for stepped SAACs surfaces at $T = 298\text{ K}$ and $p = 1\text{ atm}$. The SE_H on y-axis represents stability and activity parameter $\sqrt{\Delta G^2 + E_b^2}$ is shown on x-axis.

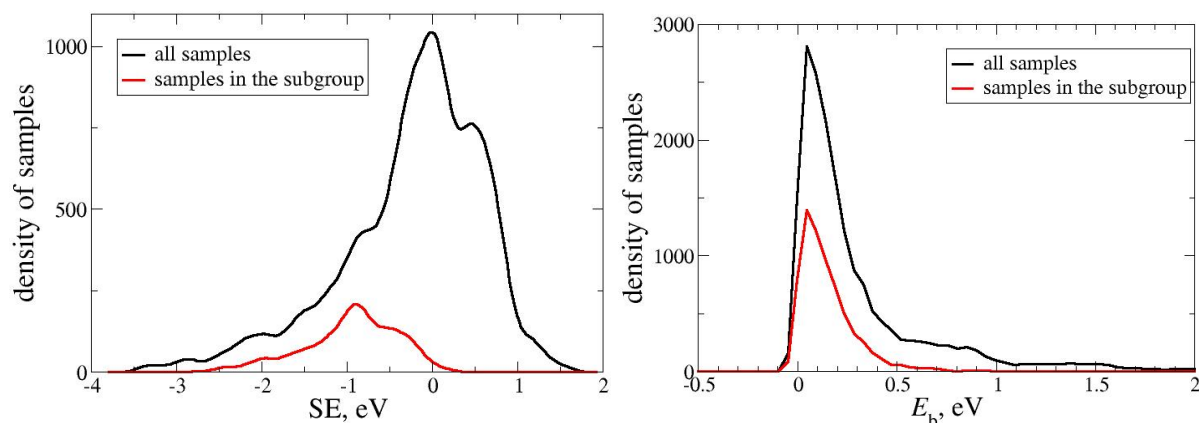


Supplementary Figure 9. Stability vs. activity map for SAACs surfaces at $T = 298\text{ K}$ and $p = 1$ atm. The SE_H on y-axis represents stability and activity parameter $\sqrt{\Delta G^2 + E_b^2}$ is shown on x-axis.

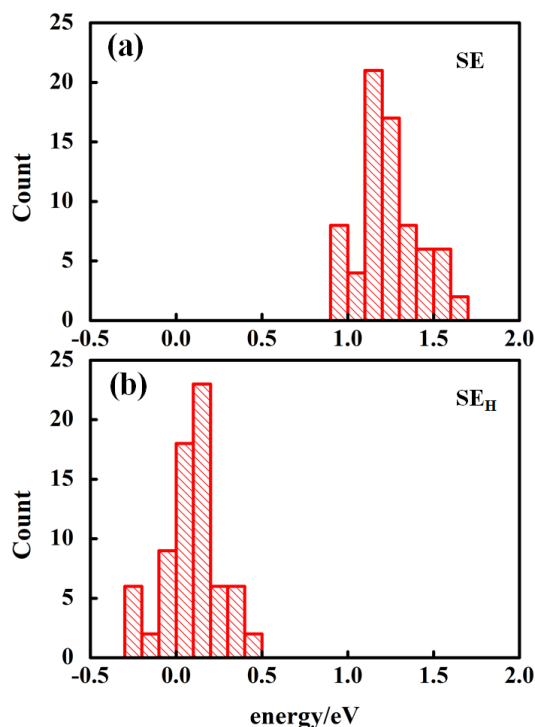
Supplementary Table 9. The aggregation energies (E_A , in eV) for Mn/Ag(111), Pt/Zn(0001), and the experimentally established SAACs.

System	Mn/Ag(111)	Pt/Zn(0001)	Pd/Cu(111)	Pt/Cu(111)	Pd/Au(111)	Pt/Au(111)	Au/Ru(111)	Ni/(Zn(0001)
E_A	2.43	0.29	0.12	0.17	0.08	0.03	-0.16	0.16

The aggregation energies were calculated as the formation energy of guest atom dimer from isolated single guest atom. Positive values mean repulsive interactions between the two guest atoms.



Supplementary Figure 10. The distribution of data samples for segregation energies (left) and H₂ dissociation energy barriers (right) in the whole sampling and obtained subgroups from SGD minimization of SE and $\sqrt{\Delta G^2 + E_b^2}$.



Supplementary Figure 11. The distribution of data samples for SE (a) and SE_H (b) for the subgroup $(-2.35 \text{ eV} \leq E_H^* \leq -2.32 \text{ eV})$ AND $(E_C^* > -2.73 \text{ eV})$ AND $(E_C < -5.98 \text{ eV})$ AND $(H \geq -5.12 \text{ eV})$ obtained from SGD minimization of $SE_H - SE$.

Supplementary References

1. Hammer, B., Hansen, L.B. & Nørskov, J.K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Physical Review B* **59**, 7413 (1999).

2. Blum, V. et al. Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications* **180**, 2175-2196 (2009).
3. Silbaugh, T.L. & Campbell, C.T. Energies of formation reactions measured for adsorbates on late transition metal surfaces. *The Journal of Physical Chemistry C* **120**, 25161-25172 (2016).
4. Henkelman, G., Uberuaga, B.P. & Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics* **113**, 9901-9904 (2000).
5. Dean, J.A. Lange's handbook of chemistry. (New york; London: McGraw-Hill, Inc., 1999).

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

This interesting work leverages the recently developed SISO (sure independence screening and sparsifying operator) algorithm to develop descriptors of stability and activity for screening single atom alloy catalysts (SAACs). The main impact of the work seems to be in identifying a number of new SAACs for potential experimental study. Two SAACs (Mn/Ag(111) and Pt/Zn(0001)) are identified as particularly promising. The paper is application-based in nature and doesn't appear to contribute significant conceptual advances for SAACs or machine learning applications. While the work seems to be well done and the paper is well-written, and the SAAC and ML topics are of great interest these days, the paper scope is not particularly ambitious. This work may be suitable for Nat. Comm. if its scope is slightly broadened and fundamental insight is improved. My comments and suggestions are provided below.

- Page 3: Figure 1a shows the H-atom binding energy vs. the d-band center for the Ag(110) host surface. The d-band center is typically calculated from the projected DOS, and it is unclear which surface atom the d-states are being taken from based on the text. Please clarify. Furthermore, using the Ag(110) surface to show that the d-band model is broken because of SAACs is not very convincing. Ag(110) should not follow the d-band model due to the fact that its d-band is completely full; therefore binding trends on Ag(110) should depend more on changes in sp electron density and Pauli repulsion (see DOI: https://link.springer.com/chapter/10.1007/978-94-015-8911-6_11 for details on why Ag(110) should not follow the d-band model). Do you have other examples of SAAC breaking scaling relations besides a host metal with a full d-band?

- Page 6: The manuscript may benefit from some discussion on the robustness of the SISO models identified in Table 2. Does adding or removing a training data point lead to a different descriptor? If so, does the new descriptor still yield similar model behavior?

- Page 6: Can you comment at all on the relative importance of the primary features included in the model relative to their impact on the overall prediction? This may lead to interesting general conclusions, like elucidating whether the primary features of the guest or host metal play a larger role for a given target property. Discussing the relative importance of guest vs. host metal features would be quite informative.

- Page 7: It seems the authors identify Tc alloys as promising SAACs. It's worth noting that there may be other health/safety considerations when using Tc in catalytic applications due to the fact that all Tc isotopes are radioactive.

- Page 8: I believe that the manuscript would benefit from an expanded discussion of Figure 3 that explains the general trends that emerge from the high-throughput screening results (e.g., in general, what types of guest atoms yield SAACs with low hydrogen dissociation barriers? What guest/host combinations lead to small segregation energies and why in terms of atomic radii size or other features?)

Minor Comments Main Text

- Figure 1: If you change solid red circles to be different symbols for hollow bridge, bridge, top that would be more information-rich and potentially informative (just a suggestion).
- Table S1 caption. "the surface-based primary features were calculated using the slab unit cell consisting of one atom per atomic layer." Should be "The".
- Page 6: The text indicates that the primary features DC, DC*, DT, DT*, DS, and DS* appear in every dimension of the descriptors for hydrogen binding energy and dissociation barrier. However, based on Table 1, it is unclear what the DT and DS primary features are as opposed to the DT* and DS* primary features. From reading the SI, it seems * denotes host metal from guest atom feature. I think this * notation can be clarified in Table 1.
- Page 9: "higher stability and efficiency than the reported ones, making them perfectly optimized for practical applications." Perfectly optimized seems to be a strong choice of words here. Perhaps remove the word "perfectly".

Minor Comments on Supporting Information

- Page 1: "Spin-polarization effects are tested for and included where appropriate." Is it noted somewhere for which spin polarization effects are included? This is a vague statement and could perhaps be made more explicit
- Figure S1 caption. "bcc(110) e," should be bcc(110) (e)
- Table S3: "Binding energy of host metal dimers", So this is a dimer energy for $A(g) + A(g) \rightarrow A_2(g)$? Could perhaps be clarified.
- Font size for the captions in Figures S3-S5 are smaller than the other Figure S captions (i.e., font size 10 vs. 12).
- Table S5: "Number of system with the predicted and calculated segregation energy meet the same condition of $SE < kT \ln(10) (N_{meet})$..." Perhaps it should read as "Number of systems with the predicted and calculated segregation energies that meet the same condition..."

Reviewer #2 (Remarks to the Author):

The manuscript presents machine learning models of single atom catalysts and screening procedure for design of hydrogenation catalysts based on this new type of alloys emerged in recent years. The features designed are easily available properties that are tabulated including electronic structure, bulk properties, etc. The target properties include the binding energy, activation barrier and the segregation. Those properties are crucial for screening high performance hydrogenation catalysts. While the work is thoroughly done in those aspects, this does not reach the standard of Nat Comm.

1. The novelty of the approach is lacking. Compressed sensing is used recently in M. Andersen, S. V. Levchenko, M. Scheffler, K. Reuter, Beyond Scaling Relations for the Description of Catalytic Materials. ACS Catal. 9, 2752–2759 (2019).
2. While the SISSO with cross validation is reasonably accurate for training a small dataset, its generalization to new systems is still the biggest problems for all current learning framework. Active learning approach was used to tackle this problem (K. Tran, Z. W. Ulissi, Active learning across intermetallics to guide discovery of electrocatalysts for CO₂ reduction and H₂ evolution. Nature Catalysis. 1, 696–703 (2018).), while a large amount of calculations are required. The current study used only ~300 datapoints for training and extend the model to ~5000 space without validation of model prediction.
3. The criteria for screening catalysts used in this study is arbitrary. Without detailed kinetics, the approach can only provide a rough screening of candidate materials.
4. For segregation, a recent study by Grabow et al. (K. K. Rao, Q. K. Do, K. Pham, D. Maiti, L. C. Grabow, Extendable Machine Learning Model for the Stability of Single Atom Alloys. Top. Catal. (2020), doi:10.1007/s11244-020-01267-2.). Even the *H binds weakly on the metals, its effect on the segregation is not considered in any of those studies.
5. The most fundamental problem of this study and the approach in general is their lacking of understanding the uniqueness of single atom alloys. Although the SISSO method comes up formula in reduced feature space, the physics is missing. The message to the community by the study is rather incremental while does not provide a way forward to tackle all those issues.

Reviewer #3 (Remarks to the Author):

The authors report the use of modern data analytics towards the reliable prediction of activity and stability of dilute alloy "single atom catalysts" for hydrogenation. The topic of particular interest as single atom catalysts have made massive strides for oxidation reactions but have had limited success for reductions particularly due to lack of activity and/or abysmal stability. The strength of the authors approach is that it addresses catalyst screening beyond the simple approximation BEP, d-band center etc. etc. etc. These concepts are embedded in the psyche of computational catalysis so deep that we forget they are simple models and, in many instances, too simple for quantitative predictions-but excellent for rationalizations on small data sets.

The authors show that by assembling a large number of atomic, bulk and alloy descriptors (table 1) they are able to perform a high dimensional correlation with the ab initio data to yield property predictions FAR more accurate than the existing simple concepts. On the one hand this is a great step forward for screening studies on the other hand if I have a more complex fitting function, I do expect a better fit. The one worry I have is this then becomes a brute force approach without the intellectual understanding that can be provided by a simple model. In this respect it might have been more intellectually pleasing for the authors to consider if there was a smaller subset of parameters (2-3) that might do a reasonable job (better than linear fits but not the full-blown set) which might hint at a simpler model. As is, the approach is fine I do worry about both overfitting/underfitting of data but do believe the authors have covered this ground adequately.

Finally, then the result of this study is that using their model they can rapidly predict the results of DFT calculations and use that data to make predictions about activity and stability based on simple energetic parameters such as presented in Figure 4. In my opinion this is the most important plot in the whole paper and the authors did not really deal with its ramifications very well. The wisdom in single atom catalysts (particularly for hydrogenation) is that the more active the species the less stable it will be—hence the scarcity of single atoms (dilute alloys) that are reported. If the authors are correct there is a large abundance of materials far in the lower right-hand corner (active and stable) that should break this trend whereas those that do exist are mostly in the upper right-hand corner (active but less stable). This is the most significant discovery/prediction in the paper as far as I am concerned, and the authors barely comment on it. Sadly, a follow-on experimental study making targets and validating the prediction would be a breakthrough and this is also not done.

Ultimately my problem is that screening for screening sake, without understanding new things, and without verifying that my parameters to define the screening criteria are valid is a reasonable technical accomplishment and not consistent with an advance I would expect in a Nature Journal. I think this paper would be highly appropriate for a journal such as ACS Catal or a chem-informatics journal but other than a more advanced fitting procedure for predicting DFT data I see no real advance here.

Response to Referees

Reviewer 1: This interesting work leverages the recently developed SISSO (sure independence screening and sparsifying operator) algorithm to develop descriptors of stability and activity for screening single atom alloy catalysts (SAACs). The main impact of the work seems to be in identifying a number of new SAACs for potential experimental study. Two SAACs (Mn/Ag(111) and Pt/Zn(0001)) are identified as particularly promising. The paper is application-based in nature and doesn't appear to contribute significant conceptual advances for SAACs or machine learning applications. While the work seems to be well done and the paper is well-written, and the SAAC and ML topics are of great interest these days, the paper scope is not particularly ambitious. This work may be suitable for Nat. Comm. if its scope is slightly broadened and fundamental insight is improved. My comments and suggestions are provided below.

Response: We thank the reviewer for the positive comments on our work. Indeed, by identifying the model with both the activity and stability parameters of the SAACs we could confirm the experimentally studied high performance SAACs. Moreover, we predict two new particularly promising systems. Keeping the reviewer's suggestions in mind, in the revised manuscript we now analyze the correlations of each component of the selected best descriptor with the target properties and discuss their physical significance. We also highlight the importance of using the combination of features rather than focusing on individual feature's role in the description of the target properties. Thus, we have stepped beyond the well-established *d*-band center theory, scaling relationships, and the Brønsted-Evans-Polanyi relationship, and have focused on the importance of data analytics in finding new SAACs.

1) Page 3: Figure 1a shows the H-atom binding energy vs. the *d*-band center for the Ag(110) host surface. The *d*-band center is typically calculated from the projected DOS, and it is unclear which surface atom the *d*-states are being taken from based on the text. Please clarify. Furthermore, using the Ag(110) surface to show that the *d*-band model is broken because of SAACs is not very convincing. Ag(110) should not follow the *d*-band model due to the fact that its *d*-band is completely full; therefore binding trends on Ag(110) should depend more on changes in *sp* electron density and Pauli repulsion (see DOI: https://link.springer.com/chapter/10.1007/978-94-015-8911-6_11 for details on why Ag(110) should not follow the *d*-band model). Do you have other examples of SAAC breaking scaling relations besides a host metal with a full *d*-band?

Response: We thank the reviewer for highlighting this important aspect. In the present study, the *d*-band centers are calculated from the *d* orbitals projected on the single guest atom only. Note that, to validate the choice of our *d*-band centers, we have calculated *d*-band centers for the *d* orbitals projected on (i) the single guest atom and its 1st nearest neighbor shell and (ii) the whole slab. However, the correlation between the binding energy and these later two *d*-band centers are found to be worse compared to the *d*-band center of the single guest atom. This is now clarified in the revised manuscript [page 3].

In the revised manuscript we have included the correlations between the binding energy and the *d*-band center for another system as well [Pt(111) surface], whose *d*-band is not completely full.

Changes made:

1. We have replaced the sentence “we first investigate correlation between BE_H and the d -band center for the alloyed systems” by “we first investigate correlation between BE_H and the d -band center for the alloyed systems. Note that, d -band centers are calculated from the d orbitals projected on the single guest atom only. We find that this way of calculating d -band center provides better correlation with other properties than d -band centers for the d orbitals projected on (i) the single guest atom plus it's 1st nearest neighbor shell or (ii) the whole slab [Topics in Catalysis 61, 462-474 (2018)].” on page 3 of the revised manuscript.
2. We have added Figure S2 by including also the system of Pt(111) surfaces which is reproduced as Figure R1 below.

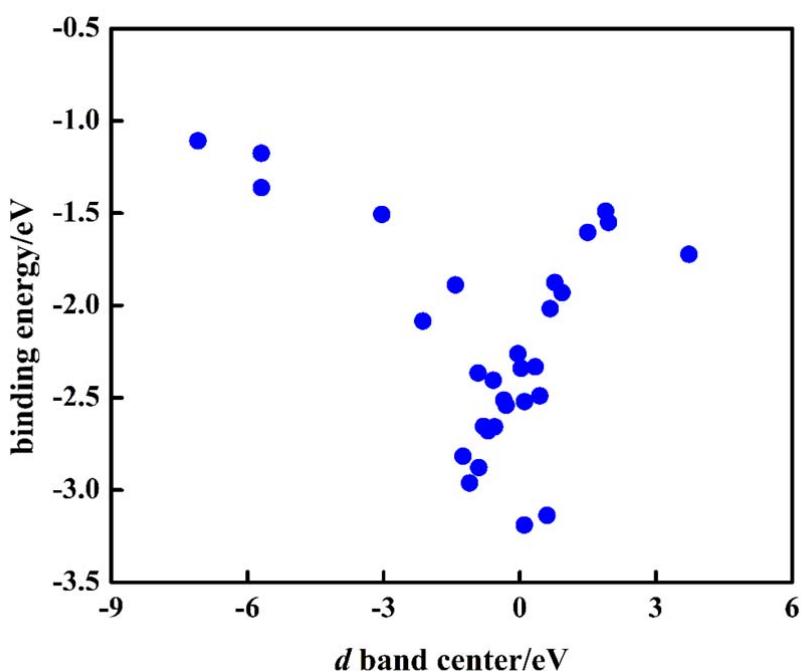


Figure R1. Correlation between (a) H-atom binding energy BE_H and the d-band center and (b) the H_2 dissociation energy barrier E_b and the H_2 dissociation reaction energy for Pt(111) based SAACs.

2) Page 6: The manuscript may benefit from some discussion on the robustness of the SISSO models identified in Table 2. Does adding or removing a training data point lead to a different descriptor? If so, does the new descriptor still yield similar model behavior?

Response: The descriptor is robust and remains unchanged upon randomly removing one training data point. We have randomly kept out one data point for each model and have repeated the process 5 times to check the robustness of the descriptor. Moreover, for the optimal dimensionality, the same set of primary features is found during CV10 in 9, 8, and 8 cases for the SISSO models of BE_H , E_b , and SE, respectively. Also, new systems that were not included in the training set were used as test set to further confirm the high transferability of our model. Finally, some of the high-throughput screening selected high performance SAACs, including all the experimentally evidenced systems and

our suggested top two best systems, are validated confirmed by density-functional theory calculations.

Changes made:

We have added the sentence “For the optimal dimensionality, the same set of primary features is found during CV10 in 9, 8, and 8 cases for the SISSO models of BE_H , E_b , and SE, respectively” on page 6 of the revised manuscript.

3) Page 6: Can you comment at all on the relative importance of the primary features included in the model relative to their impact on the overall prediction? This may lead to interesting general conclusions, like elucidating whether the primary features of the guest or host metal play a larger role for a given target property. Discussing the relative importance of guest vs. host metal features would be quite informative.

Response: We thank the referee for this suggestion. In this work we highlight the importance of the combinations of the primary features rather than using each feature individually to describe the target properties. However keeping the referee’s advice in mind, we have now introduced a novel general approach to the analysis of complex symbolic-regression models, based on the data mining approach called subgroup discovery. This have allowed us to uncover physical role of particular features, as well as relative role of guest versus host features.

Changes made:

We have added the following paragraphs on page 10 and 11 of the revised manuscript.

“Although the SISSO models are analytic formulas, the corresponding descriptors are complex, reflecting the complexity of the relationship between the primary features and the target properties. While potentially interpretable, the models do not provide a straightforward way of evaluating relative role of different features in actuating desirable changes in target properties. To facilitate physical understanding of the actuating mechanisms, we apply the subgroup discovery (SGD) approach.⁵⁵⁻⁶⁰ SGD finds local patterns in data that maximize a quality function. The patterns are described as an intersection (a selector) of simple inequalities involving provided features, e.g., (feature1<a1) AND (feature2>a2) AND... . The quality function is typically chosen such that it is maximized by subgroups balancing the number of data points in the subgroup, deviation of the median of the target property for the subgroup from the median for the whole data set, and the width of the target property distribution within the subgroup.⁶⁰”

“Here, we apply SGD in a novel context, namely as an analysis tool for symbolic regression models, including SISSO. The primary features that enter the complex SISSO descriptors of a given target property are used as features for SGD (see Table 2). The data set includes all 5200 materials and surfaces used in the high-throughput screening. The target properties are calculated using the obtained SISSO models. Five target properties are considered: $\sqrt{\Delta G^2 + E_b^2}$, SE, SE_H , E_b , $|\Delta G|$, and BE_H . Since we are interested mainly in catalysts that are active at normal conditions, ΔG is calculated at $T = 300$ K. Our goal is to find selectors that *minimize* these properties within the subgroup. Such selectors describe actuating mechanisms for minimization of a given target property. For SE, the following best selector is found: $(EC^* \leq -3.85 \text{ eV}) \text{ AND } (-3.36 \text{ eV} < EC \leq -0.01 \text{ eV})$

AND ($IP \geq 7.45$ eV). The corresponding subgroup contains 738 samples (14% of the whole population), and the distribution of SE within the subgroup is shown in Figure S10. Qualitatively, the first two conditions imply that the cohesive energy of the host material is larger in absolute value than the cohesive energy of the guest material. Physically this means that bonding between host atoms is preferred over bonding between guest atoms and therefore over intermediate host-guest binding. This leads to the tendency of maximizing number of host-host bonds by pushing guest atom to the surface. This stabilization mechanism has been discussed in literature,⁶¹ and here we confirm it by *data analysis*. In addition, we find that stability of SAACs requires that ionization potential of the guest atom is high. This can be explained by the fact that lower IP results in more pronounced delocalization of the *s* valence electrons of the guest atom and partial charge transfer to the surrounding host atoms. The charge transfer favors larger number of neighbors due to increased Madelung potential, and therefore destabilizes surface position of the guest atom.

We calculate SE_H using SISSO models for SE and BE_H [see equation (3) in the Methods section]. Therefore, SGD for SE_H is performed using primary features appearing in the descriptors of both SE and BE_H . The top found subgroup contains features related to binding of H to the host and guest metal atoms, e.g. ($EB^* < -5.75$ eV) AND ($EH^* \leq -2.10$ eV) AND ($EH \geq -2.88$ eV) AND ($IP^* \leq 7.94$ eV) AND ($IP > 8.52$ eV) AND ($R \geq 1.29$ Å). However, the distribution of SE for this subgroup is very similar to the distribution of SE_H , which means that the stability of guest atoms at the surface is weakly affected by H adsorption when the surface guest atoms are already very stable. The important effect of H adsorption is revealed when we find subgroups minimizing directly $SE_H - SE$ (in this case only primary features that appear in the SISSO descriptor of BE_H are considered for SGD analysis). The top subgroup we found contains 72 samples (1.4% of the whole population) and is described by several degenerate selectors, in particular (-2.35 eV $\leq EH^* \leq -2.32$ eV) AND ($EC^* > -2.73$ eV) AND ($EC < -5.98$ eV) AND ($H \geq -5.12$ eV). This is a very interesting and intuitive result. Distributions of SE_H and SE for this subgroup are shown in Figure S11. The SE for all materials in the subgroup is above 0 eV. However, SE_H is much closer to 0 eV, and is below 0 eV for a significant number of materials in this subgroup. The conditions on the cohesive energy of guest and host metals (very stable bulk guest metal and less stable bulk host metal) are reversed with respect to SE, i.e., adsorption of hydrogen affects strongly the systems where guest atom is unstable at the surface. This increases the reactivity of the guest atom towards an H atom. The condition ($EH^* \geq -2.35$ eV) selects materials for which interaction of H with a host atom is not too strong, so that H can bond with the guest atom and stabilize it at the surface. The condition ($EH^* \leq -2.32$ eV) makes the subgroup narrower, which further decreases median difference $SE_H - SE$ but has no additional physical meaning. The condition ($H \geq -5.12$ eV) has a minor effect on the subgroup.

One of the top selectors (among several describing very similar data subsets) for minimizing $\sqrt{\Delta G^2 + E_b^2}$ (calculated at $T = 300$ K) is: (-2.85 eV $\leq DC \leq 1.95$ eV) AND ($DT^* \leq -0.17$ eV).

The corresponding subgroup contains 1974 samples (38% of the whole population), and the distribution of E_b within the subgroup is shown in Figure S10. The selector implies that systems providing low barrier for H_2 dissociation and at the same time balanced binding of H atoms to the surface are characterized by (i) *d*-band center of the bulk guest metal around the Fermi level and (ii) *d*-band center of the host surface top layer below the Fermi level. This can be understood as follows.

Condition (i) implies that there is a significant *d*-electron density that can be donated to the adsorbed H₂ molecule, facilitating its dissociation. A very similar (apart from slightly different numerical values) condition appears in the selector for the best subgroup for *E_b* target property alone [$-2.05 \text{ eV} \leq \text{DC} \leq 1.46 \text{ eV}$) AND ($\text{EC}^* \geq -6.33 \text{ eV}$)]. Condition (ii) implies that the surface *d*-band center is more than half filled, which provides additional electrons for transferring to the H₂ molecule, but without excessive binding, to minimize $|\Delta G|$ in accordance with Sabatier principle. Indeed, several subgroups of strongly bound H atoms (minimizing BE_H) are described by selectors including condition $\text{DT}^* > -0.17$, which is exactly opposite to condition (ii). Analysis of BE_H and $|\Delta G|$ also shows that the strong and intermediate binding of H atoms to the surface is fully controlled by the features of host material.

We note that SGD is capable of finding several alternative subgroups, corresponding to different mechanisms of actuating interesting changes in target properties. These subgroups have a lower quality according to the chosen quality function, but they still contain useful information about a particular mechanism. In fact, they can be rigorously defined as top subgroups under additional constraint of zero overlap (in terms of data points) with previously found top subgroups. Analysis of such subgroups can be a subject of future work. We also note that quality function used in SGD is a parameter and can affect the found subgroups. It should be chosen based on the physical context of the problem. Exploring the role of different factors in the quality function and taking into account proposition degeneracy (no or minor effect of different conditions in the selectors due to correlation between the features) allows us to develop an understanding that may not be possible without the SGD analysis.”

4) Page 7: It seems the authors identify Tc alloys as promising SAACs. It’s worth noting that there may be other health/safety considerations when using Tc in catalytic applications due to the fact that all Tc isotopes are radioactive?

Response: We agree with the referee that health/safety considerations are very important for catalytic applications. This point is now duly mentioned on page 11 of the revised manuscript.

Changes made:

We have changed the sentence “Considering stability, activity, and abundance, two discovered best candidates Mn/Ag(111) and Pt/Zn(0001) are highlighted in Figure 4” to “Considering stability, activity, abundance, and health/safety, two discovered best candidates Mn/Ag(111) and Pt/Zn(0001) are highlighted in Figure 4” on page 11 of the revised manuscript.

5) Page 8: I believe that the manuscript would benefit from an expanded discussion of Figure 3 that explains the general trends that emerge from the high-throughput screening results (e.g., in general, what types of guest atoms yield SAACs with low hydrogen dissociation barriers? What guest/host combinations lead to small segregation energies and why in terms of atomic radii size or other features?)

Response: We thank the referee for these suggestions. In the revised manuscript, we apply the subgroup discovery (SGD) approach to evaluate relative role of different features in actuating

desirable changes in target properties and to facilitate physical understanding of the actuating mechanisms. Please referee to comment 3) for detailed discussion.

6) Minor Comments Main Text

- Figure 1: If you change solid red circles to be different symbols for hollow bridge, bridge, top that would be more information-rich and potentially informative (just a suggestion).
- Table S1 caption. “the surface-based primary features were calculated using the slab unit cell consisting of one atom per atomic layer.” Should be “The”.
- Page 6: The text indicates that the primary features DC, DC*, DT, DT*, DS, and DS* appear in every dimension of the descriptors for hydrogen binding energy and dissociation barrier. However, based on Table 1, it is unclear what the DT and DS primary features are as opposed to the DT* and DS* primary features. From reading the SI, it seems * denotes host metal from guest atom feature. I think this * notation can be clarified in Table 1.
- Page 9: “Higher stability and efficiency than the reported ones, making them perfectly optimized for practical applications.” Perfectly optimized seems to be a strong choice of words here. Perhaps remove the word “perfectly”.

Minor Comments on Supporting Information

- Page 1: “Spin-polarization effects are tested for and included where appropriate.” Is it noted somewhere for which spin polarization effects are included? This is a vague statement and could perhaps be made more explicit
- Figure S1 caption. “bcc(110) e,” should be bcc(110) (e)
- Table S3: “Binding energy of host metal dimers”, So this is a dimer energy for $A(g) + A(g) \rightarrow A_2(g)$? Could perhaps be clarified.
- Font size for the captions in Figures S3-S5 are smaller than the other Figure S captions (i.e., font size 10 vs. 12).
- Table S5: “Number of system with the predicted and calculated segregation energy meet the same condition of $SE < kT \ln(10)$ (Nmeet)...” Perhaps it should read as “Number of systems with the predicted and calculated segregation energies that meet the same condition...”

Response: We thank the referee for pointing these issues/errors. We have modified all these issues/errors accordingly in the revised manuscript and supporting information.

Reviewer 2: The manuscript presents machine learning models of single atom catalysts and screening procedure for design of hydrogenation catalysts based on this new type of alloys emerged in recent years. The features designed are easily available properties that are tabulated including electronic structure, bulk properties, etc. The target properties include the binding energy, activation barrier and the segregation. Those properties are crucial for screening high performance hydrogenation catalysts. While the work is thoroughly done in those aspects, this does not reach the standard of Nat Comm.

Response: We thank the referee for the critical comments. In the revised manuscript we have applied the subgroup discovery (SGD) approach to evaluate relative role of different features in actuating desirable changes in target properties and to facilitate physical understanding of the actuating mechanisms. The combined SISSO and SGD data analytic approaches are novel which provide us

not only predictive models but also new understanding. This allows us to go beyond the well-established *d*-band center theory, scaling relationships, and the Brønsted-Evans-Polanyi relationship.

1) The novelty of the approach is lacking. Compressed sensing is used recently in M. Andersen, S. V. Levchenko, M. Scheffler, K. Reuter, Beyond Scaling Relations for the Description of Catalytic Materials. ACS Catal. 9, 2752–2759 (2019).

Response: We would like to emphasize the advancement and novelty of our work as follows:

- (i) The aim of our work is to predict potential SAACs for hydrogenation reactions, which are not only active but also stable and thus suitable for several practical applications. It is noteworthy that SAACs has attracted significant research interests lately due to their immense potential in cost-effective large-scale industrial usages. Thus, while the methodological workhorse of this study, i.e., SISO with DFT inputs, has already been discussed before, the knowledge and understanding presented here are novel and suitable for Nat. Comm..
- (ii) An important difference of our work from previous similar efforts [Nature Catalysis 1, 531-539 (2018); ACS Catalysis 9, 2752-2759 (2019)] is that we have used only features that are very easy to evaluate. This makes high-throughput screening of a huge number of SAAC systems practically accessible, and we have exploited this in our study. Using our models, we have identified more than two hundred new SAACs with quantitative hierarchy for experimental validation, and have highlighted two new SAACs (Mn/Ag(111) and Pt/Zn(0001)) as particularly promising candidates. Moreover, in the updated manuscript we have also developed a novel strategy of analyzing complex models obtained by symbolic regression, based on the data-mining approach subgroup discovery (SGD).
- (iii) Besides the thermodynamic properties (i.e., binding energy, adsorption energy, and adsorption free energy) used in previous work [Nature Catalysis 1, 339-348 (2018); Nature Catalysis 1, 696-703 (2018); Nature 581, 178-183 (2020)] to describe the performance of catalysts, we have also included the kinetic property (energy barrier) and a stability indicator. As a result, our models both explain well the experimental results and enable design of high-performance catalysts with not only higher activity and but also stability.

2) While the SISO with cross validation is reasonably accurate for training a small dataset, its generalization to new systems is still the biggest problems for all current learning framework. Active learning approach was used to tackle this problem (K. Tran, Z. W. Ulissi, Active learning across intermetallics to guide discovery of electrocatalysts for CO₂ reduction and H₂ evolution. Nature Catalysis. 1, 696–703 (2018).), while a large amount of calculations are required. The current study used only ~300 datapoints for training and extend the model to ~5000 space without validation of model prediction.

Response: Indeed, the referee is correct that active learning can ensure reliability of the model. However, combination of SISO with active learning is a non-trivial task, because typical SISO model construction is computationally expensive. To address referee's concern, we analyze in more detail the cross-validation results, in particular the stability of descriptor selection during the

cross-validation. In addition, we validate our predictions by performing DFT calculations for some of the identified high-performance SAACs, including all the experimentally studied systems and our suggested top two best systems. We would also like to mention that the number of data points/systems in our training set is almost three times larger than that in the study of oxide-supported single-atom catalyst systems studied by Nolan and co-workers using the compressed-sensing LASSO approach [Nature Catalysis. 1, 531–539 (2018)].

Changes made:

We have added the sentence “For the optimal dimensionality, the same set of primary features are selected is found during CV10 in 9, 8, and 8 cases for the SISSO models of BE_H , E_b , and SE, respectively” on page 6 of the revised manuscript.

3) The criteria for screening catalysts used in this study is arbitrary. Without detailed kinetics, the approach can only provide a rough screening of candidate materials.

Response: We agree with the referee that detailed kinetics would improve reliability of the predictions. However, this is currently not feasible. Nevertheless, we do include a kinetic property (dissociation barrier), while previously only thermodynamic properties (binding energy, adsorption energy, adsorption free energy) were considered [Nature Catalysis 1, 339-348 (2018); Nature Catalysis 1, 696-703 (2018); Nature 581, 178-183 (2020)]. Thus, our study is the first to include a kinetic property in data-driven catalyst design. Alternatively, one can validate results of the predictions against experimental measurements. Our model consistently predicts high efficiency of the experimentally studied Pd/Cu, Pt/Cu, Pd/Ag, Pt/Au, Pd/Au, Pt/Ni, Au/Ru, and Ni/Zn SAACs (the first metal is the dispersed component), which further confirms its validity.

4) For segregation, a recent study by Grabow et al. (K. K. Rao, Q. K. Do, K. Pham, D. Maiti, L. C. Grabow, and Extendable Machine Learning Model for the Stability of Single Atom Alloys. Top. Catal. (2020), doi:10.1007/s11244-020-01267-2.). Even the *H binds weakly on the metals, its effect on the segregation is not considered in any of those studies.

Response: Actually, adsorption energies of H on metal surfaces are not small for some systems. For example, at room temperature and partial pressure of $H_2 = 1$ atm, free energy of adsorption for the experimentally established Pt/Ag(111) system is -0.23 eV and the H adatom induced segregation energy change is as high as 0.49 eV.

Changes made:

The following was added to the main text:

“We note that a machine-learning study of stability of single-atom metal alloys has been recently reported [Topics in Catalysis (2020) 63:728–741]. However, our analysis takes into account effects of adsorbates on the segregation energy, which has not been done previously.”

5) The most fundamental problem of this study and the approach in general is their lacking of understanding the uniqueness of single atom alloys. Although the SISSO method comes up formula in reduced feature space, the physics is missing. The message to the community by the study is rather

incremental while does not provide a way forward to tackle all those issues.

Response: We are grateful to referee for this comment, as it shows that important implications of our study were unclear. Our results show that it is exactly the uniqueness of SAACs that requires advanced data analysis techniques to predict their properties. As we demonstrate, the easy-to-understand correlations that work well for simple metal surfaces are not applicable to SAACs. We use a methodology (compressed sensing) that not only provides a model based on easily accessible features, but also identifies the level of complexity of the problem in terms of those features.

Nevertheless, we admit that additional data analysis that identifies common features of good SAACs would be useful. Therefore, we applied the subgroup discovery (SGD) approach to evaluate relative role of different features in actuating desirable changes in target properties and to facilitate physical understanding of the actuating mechanisms.

Changes made:

We have added the following paragraphs on page 10 and 11 of the revised manuscript.

“Although the SISSO models are analytic formulas, the corresponding descriptors are complex, reflecting the complexity of the relationship between the primary features and the target properties. While potentially interpretable, the models do not provide a straightforward way of evaluating relative role of different features in actuating desirable changes in target properties. To facilitate physical understanding of the actuating mechanisms, we apply the subgroup discovery (SGD) approach.⁵⁵⁻⁶⁰ SGD finds local patterns in data that maximize a quality function. The patterns are described as an intersection (a selector) of simple inequalities involving provided features, e.g., (feature1<a1) AND (feature2>a2) AND... . The quality function is typically chosen such that it is maximized by subgroups balancing the number of data points in the subgroup, deviation of the median of the target property for the subgroup from the median for the whole data set, and the width of the target property distribution within the subgroup.⁶⁰”

“Here, we apply SGD in a novel context, namely as an analysis tool for symbolic regression models, including SISSO. The primary features that enter the complex SISSO descriptors of a given target property are used as features for SGD (see Table 2). The data set includes all 5200 materials and surfaces used in the high-throughput screening. The target properties are calculated using the obtained SISSO models. Five target properties are considered: $\sqrt{\Delta G^2 + E_b^2}$, SE, SE_H, E_b, |ΔG|, and BE_H. Since we are interested mainly in catalysts that are active at normal conditions, ΔG is calculated at $T = 300$ K. Our goal is to find selectors that *minimize* these properties within the subgroup. Such selectors describe actuating mechanisms for minimization of a given target property. For SE, the following best selector is found: $(EC^* \leq -3.85 \text{ eV}) \text{ AND } (-3.36 \text{ eV} < EC \leq -0.01 \text{ eV}) \text{ AND } (IP \geq 7.45 \text{ eV})$. The corresponding subgroup contains 738 samples (14% of the whole population), and the distribution of SE within the subgroup is shown in Figure S10. Qualitatively, the first two conditions imply that the cohesive energy of the host material is larger in absolute value than the cohesive energy of the guest material. Physically this means that bonding between host atoms is preferred over bonding between guest atoms and therefore over intermediate host-guest binding. This leads to the tendency of maximizing number of host-host bonds by pushing guest atom to the surface. This stabilization mechanism has been discussed in literature,⁶¹ and here we confirm it by *data analysis*. In addition, we find that stability of SAACs requires that ionization potential of the

guest atom is high. This can be explained by the fact that lower IP results in more pronounced delocalization of the s valence electrons of the guest atom and partial charge transfer to the surrounding host atoms. The charge transfer favors larger number of neighbors due to increased Madelung potential, and therefore destabilizes surface position of the guest atom.

We calculate SE_H using SISSO models for SE and BE_H [see equation (3) in the Methods section]. Therefore, SGD for SE_H is performed using primary features appearing in the descriptors of both SE and BE_H . The top found subgroup contains features related to binding of H to the host and guest metal atoms, e.g. ($EB^* < -5.75$ eV) AND ($EH^* \leq -2.10$ eV) AND ($EH \geq -2.88$ eV) AND ($IP^* \leq 7.94$ eV) AND ($IP > 8.52$ eV) AND ($R \geq 1.29$ Å). However, the distribution of SE for this subgroup is very similar to the distribution of SE_H , which means that the stability of guest atoms at the surface is weakly affected by H adsorption when the surface guest atoms are already very stable. The important effect of H adsorption is revealed when we find subgroups minimizing directly $SE_H - SE$ (in this case only primary features that appear in the SISSO descriptor of BE_H are considered for SGD analysis). The top subgroup we found contains 72 samples (1.4% of the whole population) and is described by several degenerate selectors, in particular (-2.35 eV $\leq EH^* \leq -2.32$ eV) AND ($EC^* > -2.73$ eV) AND ($EC < -5.98$ eV) AND ($H \geq -5.12$ eV). This is a very interesting and intuitive result. Distributions of SE_H and SE for this subgroup are shown in Figure S11. The SE for all materials in the subgroup is above 0 eV. However, SE_H is much closer to 0 eV, and is below 0 eV for a significant number of materials in this subgroup. The conditions on the cohesive energy of guest and host metals (very stable bulk guest metal and less stable bulk host metal) are reversed with respect to SE, i.e., adsorption of hydrogen affects strongly the systems where guest atom is unstable at the surface. This increases the reactivity of the guest atom towards an H atom. The condition ($EH^* \geq -2.35$ eV) selects materials for which interaction of H with a host atom is not too strong, so that H can bond with the guest atom and stabilize it at the surface. The condition ($EH^* \leq -2.32$ eV) makes the subgroup narrower, which further decreases median difference $SE_H - SE$ but has no additional physical meaning. The condition ($H \geq -5.12$ eV) has a minor effect on the subgroup.

One of the top selectors (among several describing very similar data subsets) for minimizing $\sqrt{\Delta G^2 + E_b^2}$ (calculated at $T = 300$ K) is: (-2.85 eV $\leq DC \leq 1.95$ eV) AND ($DT^* \leq -0.17$ eV).

The corresponding subgroup contains 1974 samples (38% of the whole population), and the distribution of E_b within the subgroup is shown in Figure S10. The selector implies that systems providing low barrier for H_2 dissociation and at the same time balanced binding of H atoms to the surface are characterized by (i) d -band center of the bulk guest metal around the Fermi level and (ii) d -band center of the host surface top layer below the Fermi level. This can be understood as follows. Condition (i) implies that there is a significant d -electron density that can be donated to the adsorbed H_2 molecule, facilitating its dissociation. A very similar (apart from slightly different numerical values) condition appears in the selector for the best subgroup for E_b target property alone [$(-2.05$ eV $\leq DC \leq 1.46$ eV) AND ($EC^* \geq -6.33$ eV)]. Condition (ii) implies that the surface d -band center is more than half filled, which provides additional electrons for transferring to the H_2 molecule, but without excessive binding, to minimize $|\Delta G|$ in accordance with Sabatier principle. Indeed, several subgroups of strongly bound H atoms (minimizing BE_H) are described by selectors including condition $DT^* > -0.17$, which is exactly opposite to condition (ii). Analysis of BE_H and $|\Delta G|$ also

shows that the strong and intermediate binding of H atoms to the surface is fully controlled by the features of host material.

We note that SGD is capable of finding several alternative subgroups, corresponding to different mechanisms of actuating interesting changes in target properties. These subgroups have a lower quality according to the chosen quality function, but they still contain useful information about a particular mechanism. In fact, they can be rigorously defined as top subgroups under additional constraint of zero overlap (in terms of data points) with previously found top subgroups. Analysis of such subgroups can be a subject of future work. We also note that quality function used in SGD is a parameter and can affect the found subgroups. It should be chosen based on the physical context of the problem. Exploring the role of different factors in the quality function and taking into account proposition degeneracy (no or minor effect of different conditions in the selectors due to correlation between the features) allows us to develop an understanding that may not be possible without the SGD analysis.”

Reviewer 3: The authors report the use of modern data analytics towards the reliable prediction of activity and stability of dilute alloy “single atom catalysts” for hydrogenation. The topic of particular interest as single atom catalysts have made massive strides for oxidation reactions but have had limited success for reductions particularly due to lack of activity and/or abysmal stability.

1) The strength of the authors approach is that it addresses catalyst screening beyond the simple approximation BEP, d-band center etc. etc. etc. These concepts are embedded in the psyche of computational catalysis so deep that we forget they are simple models and, in many instances, to too simple for quantitative predictions-but excellent for rationalizations on small data sets.

Response: We thank the reviewer for this comment. It correctly outlines the important aspect of our work.

2) The authors show that by assembling a large number of atomic, bulk and alloy descriptors (table1) they are able to perform a high dimensional correlation with the ab initio data to yield property predictions FAR more accurate than the existing simple concepts. On the one hand this is a great step forward for screening studies on the other hand if I have a more complex fitting function, I do expect a better fit. The one worry I have is this then become a brute force approach without the intellectual understanding that can be provided by a simple model. In this respect it might have been more intellectually pleasing for the authors to consider if there was a smaller subset of parameters (2-3) that might do a reasonable job (better than linear fits but not the full-blown set) which might hint at a simpler model. As is, the approach is fine I do worry about both overfitting/underfitting of data but do believe the authors have covered this ground adequately.

Response: This is a very important comment that overlaps with similar concerns of the other referees. Indeed, we perform a careful cross-validation of our models and validate them on a test set never used for training, to ensure models’ predictive power. However, the training and test sets are unavoidably limited, and there is never a guarantee that we capture all important physical variations present in the larger data set. This makes our mind crave for additional consistency check that we

call “physical understanding”. It justifies extrapolation of the models, possibly even to a different class of systems. Such extrapolation can be very useful, but also very misleading, as our study demonstrates. Nevertheless, we admit that additional data analysis that identifies common features of good SAACs would be useful. Therefore, we applied the subgroup discovery (SGD) approach to evaluate relative role of different features in actuating desirable changes in target properties and to facilitate physical understanding of the actuating mechanisms.

Changes made:

We have added the following paragraphs on page 10 and 11 of the revised manuscript.

“Although the SISSO models are analytic formulas, the corresponding descriptors are complex, reflecting the complexity of the relationship between the primary features and the target properties. While potentially interpretable, the models do not provide a straightforward way of evaluating relative role of different features in actuating desirable changes in target properties. To facilitate physical understanding of the actuating mechanisms, we apply the subgroup discovery (SGD) approach.⁵⁵⁻⁶⁰ SGD finds local patterns in data that maximize a quality function. The patterns are described as an intersection (a selector) of simple inequalities involving provided features, e.g., (feature1<a1) AND (feature2>a2) AND... . The quality function is typically chosen such that it is maximized by subgroups balancing the number of data points in the subgroup, deviation of the median of the target property for the subgroup from the median for the whole data set, and the width of the target property distribution within the subgroup.⁶⁰”

“Here, we apply SGD in a novel context, namely as an analysis tool for symbolic regression models, including SISSO. The primary features that enter the complex SISSO descriptors of a given target property are used as features for SGD (see Table 2). The data set includes all 5200 materials and surfaces used in the high-throughput screening. The target properties are calculated using the obtained SISSO models. Five target properties are considered: $\sqrt{\Delta G^2 + E_b^2}$, SE, SE_H, E_b, |ΔG|, and BE_H. Since we are interested mainly in catalysts that are active at normal conditions, ΔG is calculated at $T = 300$ K. Our goal is to find selectors that *minimize* these properties within the subgroup. Such selectors describe actuating mechanisms for minimization of a given target property. For SE, the following best selector is found: (EC* ≤ -3.85 eV) AND (-3.36 eV < EC ≤ -0.01 eV) AND (IP ≥ 7.45 eV). The corresponding subgroup contains 738 samples (14% of the whole population), and the distribution of SE within the subgroup is shown in Figure S10. Qualitatively, the first two conditions imply that the cohesive energy of the host material is larger in absolute value than the cohesive energy of the guest material. Physically this means that bonding between host atoms is preferred over bonding between guest atoms and therefore over intermediate host-guest binding. This leads to the tendency of maximizing number of host-host bonds by pushing guest atom to the surface. This stabilization mechanism has been discussed in literature,⁶¹ and here we confirm it by *data analysis*. In addition, we find that stability of SAACs requires that ionization potential of the guest atom is high. This can be explained by the fact that lower IP results in more pronounced delocalization of the *s* valence electrons of the guest atom and partial charge transfer to the surrounding host atoms. The charge transfer favors larger number of neighbors due to increased Madelung potential, and therefore destabilizes surface position of the guest atom.

We calculate SE_H using SISSO models for SE and BE_H [see equation (3) in the Methods

section]. Therefore, SGD for SE_H is performed using primary features appearing in the descriptors of both SE and BE_H . The top found subgroup contains features related to binding of H to the host and guest metal atoms, e.g. ($EB^* < -5.75$ eV) AND ($EH^* \leq -2.10$ eV) AND ($EH \geq -2.88$ eV) AND ($IP^* \leq 7.94$ eV) AND ($IP > 8.52$ eV) AND ($R \geq 1.29$ Å). However, the distribution of SE for this subgroup is very similar to the distribution of SE_H , which means that the stability of guest atoms at the surface is weakly affected by H adsorption when the surface guest atoms are already very stable. The important effect of H adsorption is revealed when we find subgroups minimizing directly $SE_H - SE$ (in this case only primary features that appear in the SISSO descriptor of BE_H are considered for SGD analysis). The top subgroup we found contains 72 samples (1.4% of the whole population) and is described by several degenerate selectors, in particular (-2.35 eV $\leq EH^* \leq -2.32$ eV) AND ($EC^* > -2.73$ eV) AND ($EC < -5.98$ eV) AND ($H \geq -5.12$ eV). This is a very interesting and intuitive result. Distributions of SE_H and SE for this subgroup are shown in Figure S11. The SE for all materials in the subgroup is above 0 eV. However, SE_H is much closer to 0 eV, and is below 0 eV for a significant number of materials in this subgroup. The conditions on the cohesive energy of guest and host metals (very stable bulk guest metal and less stable bulk host metal) are reversed with respect to SE, i.e., adsorption of hydrogen affects strongly the systems where guest atom is unstable at the surface. This increases the reactivity of the guest atom towards an H atom. The condition ($EH^* \geq -2.35$ eV) selects materials for which interaction of H with a host atom is not too strong, so that H can bond with the guest atom and stabilize it at the surface. The condition ($EH^* \leq -2.32$ eV) makes the subgroup narrower, which further decreases median difference $SE_H - SE$ but has no additional physical meaning. The condition ($H \geq -5.12$ eV) has a minor effect on the subgroup.

One of the top selectors (among several describing very similar data subsets) for minimizing $\sqrt{\Delta G^2 + E_b^2}$ (calculated at $T = 300$ K) is: (-2.85 eV $\leq DC \leq 1.95$ eV) AND ($DT^* \leq -0.17$ eV). The corresponding subgroup contains 1974 samples (38% of the whole population), and the distribution of E_b within the subgroup is shown in Figure S10. The selector implies that systems providing low barrier for H_2 dissociation and at the same time balanced binding of H atoms to the surface are characterized by (i) d -band center of the bulk guest metal around the Fermi level and (ii) d -band center of the host surface top layer below the Fermi level. This can be understood as follows. Condition (i) implies that there is a significant d -electron density that can be donated to the adsorbed H_2 molecule, facilitating its dissociation. A very similar (apart from slightly different numerical values) condition appears in the selector for the best subgroup for E_b target property alone [$(-2.05$ eV $\leq DC \leq 1.46$ eV) AND ($EC^* \geq -6.33$ eV)]. Condition (ii) implies that the surface d -band center is more than half filled, which provides additional electrons for transferring to the H_2 molecule, but without excessive binding, to minimize $|\Delta G|$ in accordance with Sabatier principle. Indeed, several subgroups of strongly bound H atoms (minimizing BE_H) are described by selectors including condition $DT^* > -0.17$, which is exactly opposite to condition (ii). Analysis of BE_H and $|\Delta G|$ also shows that the strong and intermediate binding of H atoms to the surface is fully controlled by the features of host material.

We note that SGD is capable of finding several alternative subgroups, corresponding to different mechanisms of actuating interesting changes in target properties. These subgroups have a lower quality according to the chosen quality function, but they still contain useful information about a

particular mechanism. In fact, they can be rigorously defined as top subgroups under additional constraint of zero overlap (in terms of data points) with previously found top subgroups. Analysis of such subgroups can be a subject of future work. We also note that quality function used in SGD is a parameter and can affect the found subgroups. It should be chosen based on the physical context of the problem. Exploring the role of different factors in the quality function and taking into account proposition degeneracy (no or minor effect of different conditions in the selectors due to correlation between the features) allows us to develop an understanding that may not be possible without the SGD analysis.”

3) Finally, then the result of this study is that using their model they can rapidly predict the results of DFT calculations and use that data to make predictions about activity and stability based on simple energetic parameters such as presented in Figure 4. In my opinion this is the most important plot in the whole paper and the authors did not really deal with its ramifications very well. The wisdom in single atom catalysts (particularly for hydrogenation) is that the more active the species the less stable it will be—hence the scarcity of single atoms (dilute alloys) that are reported. If the authors are correct there is a large abundance of materials far in the lower right-hand corner (active and stable) that should break this trend whereas those that do exist are mostly in the upper right-hand corner (active but less stable). This is the most significant discovery/prediction in the paper as far as I am concerned, and the authors barely comment on it. Sadly, a follow-on experimental study making targets and validating the prediction would be a breakthrough and this is also not done.

Response: There seems to be a misunderstanding regarding Fig. 4. The most active and stable materials are in the lower LEFT-hand corner. Just as the referee points out, this corner is scarcely populated compared to the whole area covered by all calculated materials. However, this does not mean there are no materials that can be better than the experimentally tested ones. To clarify this aspect, we have now added a discussion to the main text and new Figure S9 in the revised supporting information, which is reproduced as Figure R1 below.

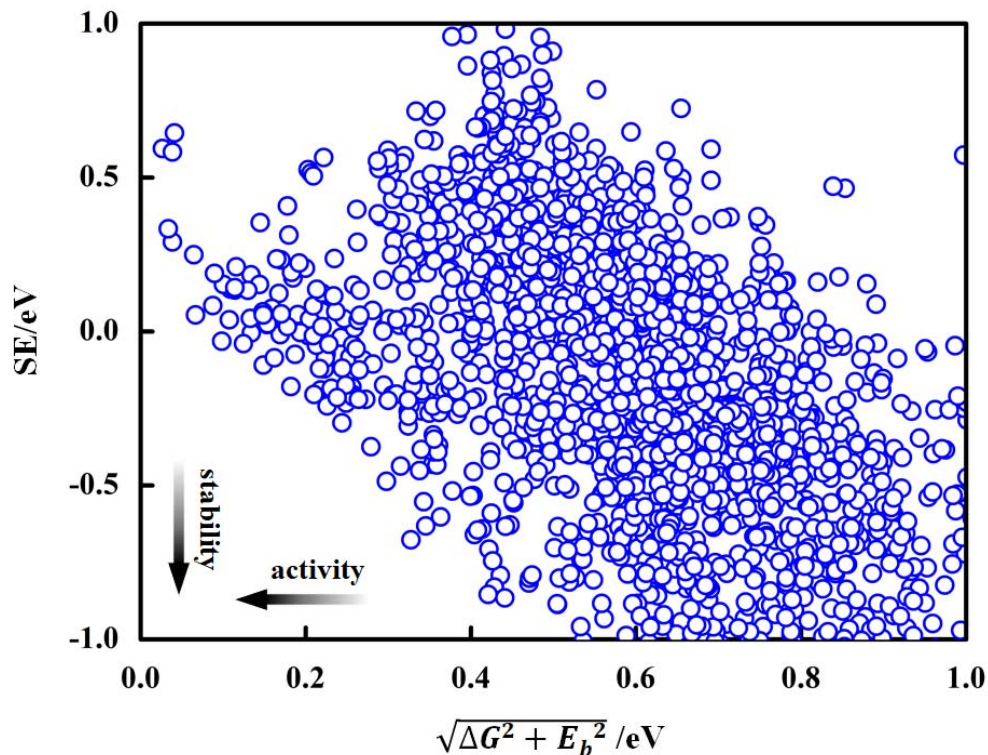


Figure R1. Stability vs. activity map for flat SAACs surfaces at T=298 K and p=1 atm. The SE on y-axis represents stability and activity parameter $\sqrt{\Delta G^2 + E_b^2}$ is shown on x-axis.

Changes made:

1) We have added the sentences:

“As expected, stability and activity are inversely related, which can be seen from the negative slope of the general trend in Figure 8 (showing selected materials) and Figure S9 (showing all explored materials), as well as a cut-off in population of the lower left-hand corner of these plots. Nevertheless, there are several materials that are predicted to be better SAACs than the so-far reported ones.” on page 11 of the revised manuscript.

2) We added Figure S9 in the revised supporting materials.

4) Sadly, a follow-on experimental study making targets and validating the prediction would be a breakthrough and this is also not done.

Response: This work was conceived as a theoretical one. We are happy to share methodology and predictions with the community as soon as possible. We very much hope that our findings will encourage experimental groups to validate our predictions.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have greatly expanded their work based on the reviewer comments. Importantly, they now utilize a data mining algorithm called Subgroup Discovery to analyze their SAAC dataset in combination with their SISO model. This added analysis enables the authors to give much more satisfying and general insights regarding the stability and activity of the SAACs, which should prove useful for the catalysis community. Additionally, Subgroup Discovery has not been used yet in the catalysis/surface science fields (and SISO algorithm has only been used once before in catalysis field to my knowledge), thus this work also introduces cutting-edge data science tools to the broader scientific community. Therefore, this paper should be of broad interest to multiple communities. I believe the work is suitable for publication.

Reviewer #2 (Remarks to the Author):

Authors addressed most of the comments. However, the physical insights by subgroup discovery is rather limited. I stick to my opinion that this work is not a significant step toward ML method itself or SAAC discovery. It might be appropriate to a more specialized catalysis journal.

1. The SISO machine learning method employed in this study is not new. With the same set of features, a regular neural network can be more easily trained and coupled with active learning. With existing alloy database published in community, a convolutional neural net can also be used since the local environment of single atom alloys is analogous to the traditional fcc-type alloys, e.g. A3B, in the first coordination shell. In term of physical interpretation, they are all black-box models. SISO can give a formula instead, although its direct understanding by a catalysis expert is still not there. The formula can be considered as symbolic regression rather than physical models. Interpreting black-box models are not necessarily providing physical insights that can be translated to design.

2. Subgroup discovery is a half-way approach to extract conditions of features optimizing a defined quality function. It is monte carlo based algorithm. The identified boundary values will depend on runs and hyperparameters. The approach has been used in materials science and catalysis. It is overstated in terms of novelty in the context. The rule identified by the method is convoluted rather than being insightful.

3. The design space of SAAs is relatively small compared to complex alloys. The indication in abstract for hundreds of thousands is misleading.

4. It says the energy BEH and the d-band center and (b) the H₂ dissociation energy barrier E_b and the H₂ dissociation reaction energy for Pt(111) based SAACs. But the (b) panel is missing.

5. It claimed a step away from the d-band theory, BEP, and scaling relations. While machine learning models can be considered as a further step away from the d-band center type of theory level, it is not fair to say that for the original d-band theory since machine learning models are regression based only. It is not close to go beyond BEP and scaling relations in this work since it simply does not consider full reaction pathways. The claim is irrelevant.

6. The d-band center of the bonding guest atom is obvious choice for atop adsorption, but not quite for hollow, bridge. The averaged d-band center of a collection of atoms in the revision is not the right since the coupling strength decays rapidly with distance.

Reviewer #3 (Remarks to the Author):

After carefully considering the previous reviewers' comments and the revised manuscript I can say most of the technical concerns I have about this work are resolved and I may have even softened (but not changed) my stance about not really bringing new understanding. I still do not like these screening/data analytics papers for the sake of data analytics but in this case the decision point for me is that the authors predict many new catalysts so, in principle, the way to test and validate this model is on the table.

IF the authors are right then this is a breakthrough, if they are wrong ... I think this may well be worth publishing in Nature Comm and I look forward to seeing this work validated (or not).

The text does require significant proof reading and improving on the English, particularly the new parts and should be proof read careful before it is published.

Reviewer 1:

The authors have greatly expanded their work based on the reviewer comments. Importantly, they now utilize a data mining algorithm called Subgroup Discovery to analyze their SAAC dataset in combination with their SISSO model. This added analysis enables the authors to give much more satisfying and general insights regarding the stability and activity of the SAACs, which should prove useful for the catalysis community. Additionally, Subgroup Discovery has not been used yet in the catalysis/surface science fields (and SISSO algorithm has only been used once before in catalysis field to my knowledge), thus this work also introduces cutting-edge data science tools to the broader scientific community. Therefore, this paper should be of broad interest to multiple communities. I believe the work is suitable for publication.

Reply: We thank the reviewer for the positive assessment of our work, and most of all for the reviewer's critical comments that helped us to improve our manuscript.

Reviewer 2:

Authors addressed most of the comments. However, the physical insights by subgroup discovery is rather limited. I stick to my opinion that this work is not a significant step toward ML method itself or SAAC discovery. It might be appropriate to a more specialized catalysis journal.

Reply: We thank the reviewer for sharing his/her opinion, and address the reviewer's concerns below.

1. The SISSO machine learning method employed in this study is not new. With the same set of features, a regular neural network can be more easily trained and coupled with active learning. With existing alloy database published in community, a convolutional neural net can also be used since the local environment of single atom alloys is analogous to the traditional fcc-type alloys, e.g. A3B, in the first coordination shell.

In term of physical interpretation, they are all black-box models. SISSO can give a formula instead, although its direct understanding by a catalysis expert is still not there. The formula can be considered as symbolic regression rather than physical models. Interpreting black-box models are not necessarily providing physical insights that can be translated to design.

Reply: Whether it would be easier to train a convolutional neural network coupled with active learning remains to be seen, in particular if one takes into account typically much larger number of training data needed for a reliable NN model. Also, it remains to be seen whether the list of primary features plus the existing data on fcc-type alloys are sufficient for a better prediction model. These questions are an interesting topic for future work, as there is no evidence in the literature regarding these issues. However, they are beyond the scope of this work. Compressed-sensing based descriptor identification is not a new idea in general, but it is much newer than NN, and we are aware of only one published application of SISSO in catalysis (with a contribution by one of the authors), which however did not include a high-throughput study. Using SISSO to specifically predict properties of SAACs is in fact a new idea. As the referee points out, SISSO provides analytic formulas as models. While it is indeed sometimes not easy to interpret these models, in our manuscript we developed a

novel approach that allows us to move forward also in this direction. We were inspired by comments of this and the other referees to do this. Our approach allows us to extract physical understanding based on thorough data analysis rather than intuition based on a limited set of data. We believe this is an important step that is of general interest to the community, in addition to our actual high-throughput predictions.

2. Subgroup discovery is a half-way approach to extract conditions of features optimizing a defined quality function. It is monte carlo based algorithm. The identified boundary values will depend on runs and hyperparameters. The approach has been used in materials science and catalysis. It is overstated in terms of novelty in the context. The rule identified by the method is convoluted rather than being insightful.

Reply: We are not sure what referee means by “half-way approach”. We have not seen such a characteristics in existing literature. It is for sure a novel approach in catalysis. Although there is a publication that employed SGD in the context of catalysis, there are no examples of using SGD for catalyst design. However, we do not even claim novelty of SGD, we claim novelty of how we use it: to interpret complex symbolic regression models.

Yes, SGD results depend on the run if the Monte Carlo approach is used (as in our case). However, this dependence comes mainly from feature correlations, provided an extensive sampling was performed. The realkd implementation takes special care in distributing sampling points optimally to improve sampling efficiency, and we performed an extensive sampling. In fact, we always see consistency among several top subgroups. The problem is not the convolution but the idea that one can always just take an arbitrary set of features, look at the top subgroup for a given property and hope to get an insight just looking at it. As we describe in the text, the insight should come from understanding the physics behind the primary features and analysis of several top selectors and a joint analysis of subgroups for different target properties. To emphasize this, we have updated the following paragraph:

“Exploring the role of different factors in the quality function and taking into account proposition degeneracy (no or minor effect of different conditions in the selectors due to correlation between the features) can significantly improve interpretability of the selectors. The interpretability also depends crucially on our physical understanding of the features and relations between them. Nevertheless, in combination with human knowledge SGD analysis allows for development of understanding that would not be possible without the help of artificial intelligence.”

3. The design space of SAAs is relatively small compared to complex alloys. The indication in abstract for hundreds of thousands is misleading.

Reply: For binary alloys we agree with the referee, but if we consider multi-component alloys there are easily hundreds of thousands. However, since we consider only binary SAACs, we follow referee’s suggestion and remove that statement.

4. It says the energy BEH and the d-band center and (b) the H₂ dissociation energy barrier E_b and the H₂ dissociation reaction energy for Pt(111) based SAACs. But the (b) panel is missing.

Reply: Thank you for noticing this. Perhaps something went wrong with formatting. We have checked to make the changes correspondingly.

5. It claimed a step away from the d-band theory, BEP, and scaling relations. While machine learning models can be considered as a further step away from the d-band center type of theory level, it is not fair to say that for the original d-band theory since machine learning models are regression based only. It is not close to go beyond BEP and scaling relations in this work since it simply does not consider full reaction pathways. The claim is irrelevant.

Reply: Indeed, for a neural network model based on hand-picked descriptors this would be a fair criticism. But we use SISSO to IDENTIFY descriptors, and we can directly compare their performance with the d-band center descriptor. Whether we can provide a physical interpretation of the identified descriptor is a different matter, and we show how to do it with SGD.

6. The d-band center of the bonding guest atom is obvious choice for atop adsorption, but not quite for hollow, bridge. The averaged d-band center of a collection of atoms in the revision is not the right since the coupling strength decays rapidly with distance.

Reply: For all Ag(110) and Pt(111) based SAACs the most stable adsorption sites are hollow sites. We found the correlation between BE_H and the d-band center of the d orbitals that are projected to the single guest atom for the alloyed systems provides better correlation with other properties than d-band centers for the d orbitals projected on (i) the single guest atom plus its 1st nearest neighbor shell or (ii) the whole slab. This is consistent with a previous study [Topics in Catalysis 61, 462-474 (2018)].

Reviewer 3:

After carefully considering the previous reviewers' comments and the revised manuscript I can say most of the technical concerns I have about this work are resolved and I may have even softened (but not changed) my stance about not really bringing new understanding. I still do not like these screening/data analytics papers for the sake of data analytics but in this case the decision point for me is that the authors predict many new catalysts so, in principle, the way to test and validate this model is on the table.

IF the authors are right then this is a breakthrough, if they are wrong ... I think this may well be worth publishing in Nature Comm and I look forward to seeing this work validated (or not).

The text does require significant proof reading and improving on the English, particularly the new parts and should be proof read careful before it is published.

Reply: We thank the referee for the clear opinion and critical comments that helped to significantly improve our manuscript. Indeed, to the best of our knowledge this work is the first one that uses SISSO for high-throughput PREDICTIONS of catalytic properties rather than just data analysis. From this point of view, our work is both a guide for a rational design of SAACs and an important step towards testing and further development of data-analytics methodology for catalysis.

Following the advice by the referee, we have carefully proofread the text.