

Supporting Information

Single Atom Alloys Aggregation in the Presence of Ligands

Maya Salem and Giannis Mpourmpakis*

Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh,
Pennsylvania 15261, United States

*Corresponding author. E-mail: gmpourmp@pitt.edu

1. (100) structures involved in E_{agg} calculations.

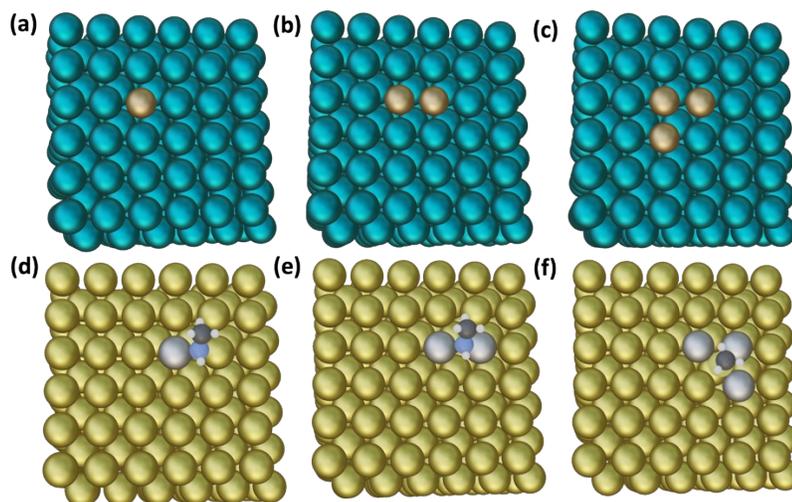


Figure S1. Top view of (a) a single dopant (Cu) on the (100) host metal surface (SAA), (b) Cu dimer on the (100) host metal surface, (c) Cu trimer on the (100) host metal surface, H_3C-NH bridge adsorption on Au (100) surface doped with Ag (d) SAA and (e) dimer, and (f) H_3C-S adsorbed on a hollow position of a trimer-dopant on Au (100) metal host surface.

2. List of descriptors used in the feature importance analysis.

Table S1. DFT calculated CE_{bulk} in this work.

Metals	CE_{bulk} (eV/atom)
Ag	-2.82
Au	-3.34
Cu	-3.74
Ni	-5.85
Pd	-4.17
Pt	-6.83

Table S2. Descriptors used in the feature importance analysis.

Descriptor Name	Symbol
Number of Dopants* Bulk Cohesive Energy/Coordination Number of the Dopant	$\Delta nCE_i/CN$
BCM-Calculated Cohesive Energy	CE_i
Binding Energy of adsorbate on a Single Atom/Coordination Number of the Adsorbate	$\Delta BE_i/CN_{ads}$
Atomic radius ¹	r_i
First Ionization Potential ²	IP_i
Electron affinity ²	EA_i

Features with subscript i (for instance, r_i) indicates that host metal property (X_h), dopant metal property (X_d), and difference between host and dopant metal (i.e. $X_h - X_d$) are taken into account. For example, $CE_{\text{bulk},i}/CN$: $CE_{\text{bulk},h}/CN$, $CE_{\text{bulk},d}/CN$, and $\Delta CE/CN = (CE_{\text{bulk},h} - CE_{\text{bulk},d})/CN$

3. Hyperparameters used in regression models to predict the E_{agg} in non-ligated SAAs.

Table S3. Hyperparameters used in the two-feature regression models ($\Delta nCE_{bulk}/CN$ and Δr) based on the GridSearchCV results.

Model	Hyperparameter
KRR: 2 nd Order Polynomial	Alpha: 0.0001 Gamma: 0.15874
KRR: RBF	Alpha: 0.0001 Gamma: 0.36905
KRR: Laplacian	Alpha: 0.07152 Gamma: 0.05358
SVR: 2 nd Order Polynomial	C: 0.89474 Epsilon: 0.001 Gamma: 1
SVR: 3 rd Order Polynomial	C: 1.842168 Epsilon: 0.001 Gamma: 0.84226
SVR: RBF	C: 4.7368 Epsilon: 0.05358 Gamma: 0.7371
LASSO	Alpha: 0.0527

4. Predicting DFT E_{agg} in non-ligated SAAs using different regression models and criteria.

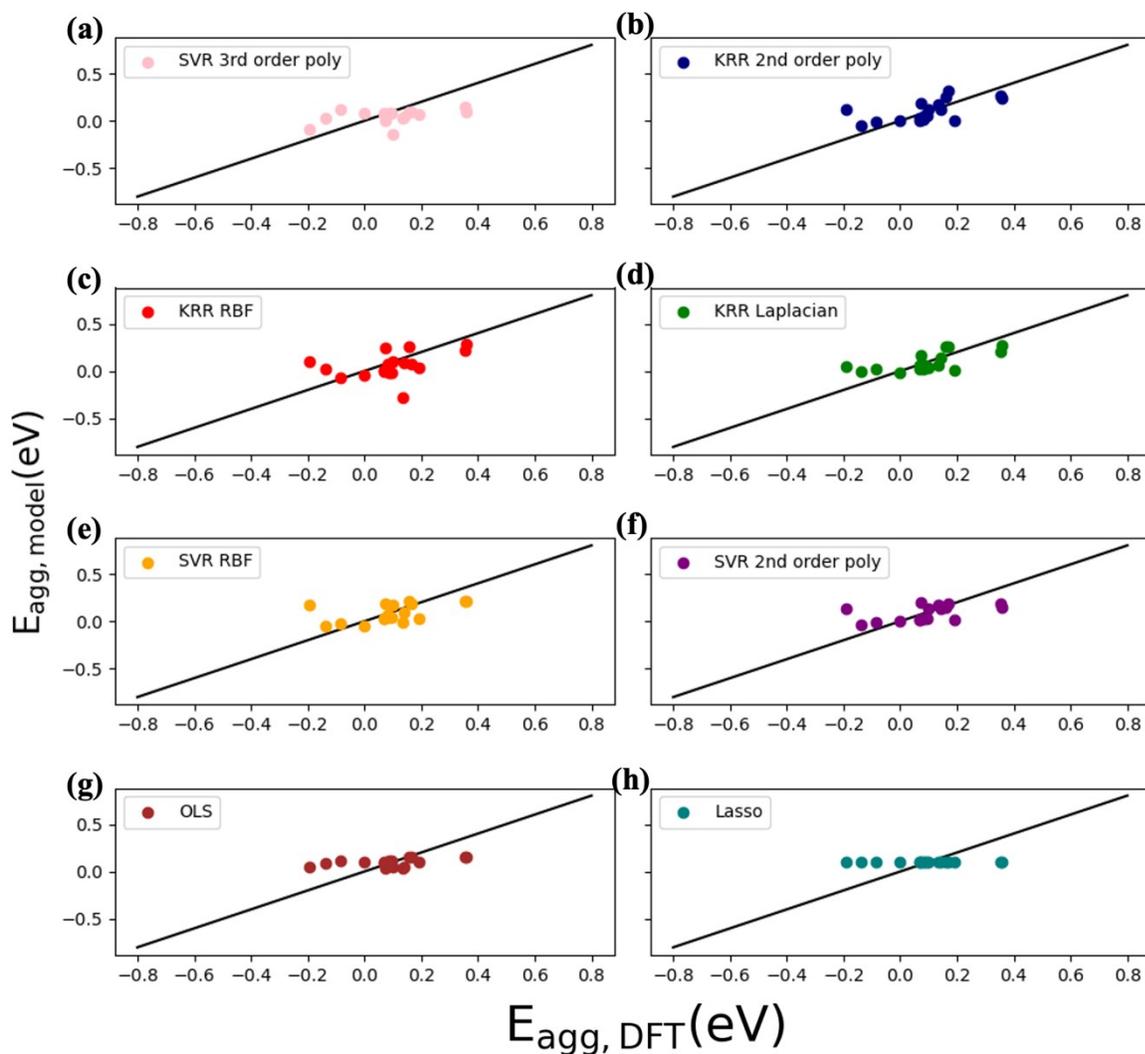


Figure S2. Parity plot between different regression models based on the test set using the two features ($\Delta n\text{CE}_{\text{bulk}}/\text{CN}$ and Δr) and $E_{\text{agg, DFT}}$. (a) Support Vector Regression (SVR): 3rd order polynomial, (b) Kernel Ridge Regression (KRR): 2nd order polynomial, (c) KRR: RBF, (d) KRR: Laplacian, (e) SVR: RBF, (f) SVR: 2nd order polynomial, (g) Linear Regression, and (h) LASSO.

Table S4. Different E_{agg} models of non-ligated systems and their corresponding train, test, and validation MAE.

Model	Test MAE (eV)	Validation MAE (eV)	Train MAE (eV)	Validation-Train Δ MAE (eV)
SVR: 3 rd poly	0.104	0.111	0.097	0.014
KRR: 2nd poly	0.084	0.081	0.075	0.006
KRR: RBF	0.108	0.097	0.053	0.044
KRR: Lap	0.087	0.085	0.066	0.019
SVR: RBF	0.099	0.089	0.063	0.026
SVR: 2 nd order poly	0.087	0.085	0.078	0.007
OLS	0.091	0.109	0.105	0.003
LASSO	0.097	0.110	0.109	0.001

In Figure S3, the segregation energies (E_{seg}) are obtained from our previous work³ and is computed using the following equation⁴:

$$E_{\text{seg}} = E_{\text{pure bulk}} + E_{\text{dopant,1st layer}} - E_{\text{dopant,bulk}} - E_{\text{pure surface}}$$

where $E_{\text{pure bulk}}$ and $E_{\text{pure surface}}$ are the total energies of monometallic bulk and surface, respectively. The $E_{\text{dopant,1st layer}}$ is the total energy of the dopant present in the first layer of the surface, and $E_{\text{dopant,bulk}}$ is the total energy of the dopant present in the bulk. In the presence of ligands, the E_{seg} equation is altered to account for adsorbate effects ($E_{\text{seg}/X}$):

$$E_{\text{seg}/X} = E_{\text{pure bulk}} + E_{\text{dopant,1st layer,X}} - E_{\text{dopant,bulk}} - E_{\text{pure surface,X}}$$

where $E_{\text{pure surface,X}}$ is the total energy of the surface in the presence of an adsorbate. $E_{\text{dopant,1st layer,X}}$ is the total energy of the dopant present in the first layer with the addition of an adsorbate. A negative E_{seg} value indicates that the dopant has the thermodynamic tendency to segregate to the surface, while a positive E_{seg} value denotes that the dopant prefers to stay in bulk.

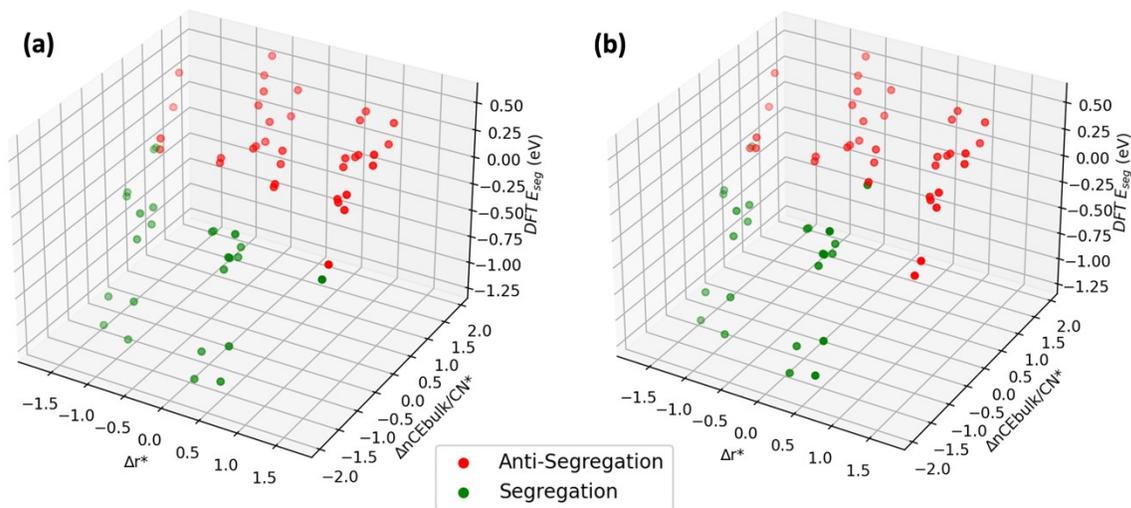


Figure S3. 3D plot of DFT E_{seg} versus the Δr^* and $\frac{\Delta nCE_{\text{bulk}}}{CN}^*$. The data points in (a) are colored based on the actual E_{seg} values (negative E_{seg} = segregation, positive E_{seg} = anti-segregation), while (b) are colored based on the $\Delta r^* < 0.9$ and $\frac{\Delta nCE_{\text{bulk}}}{CN}^* < 0$ criteria. The red and green colors in panel (a) represent the anti-segregation and dopant segregation.

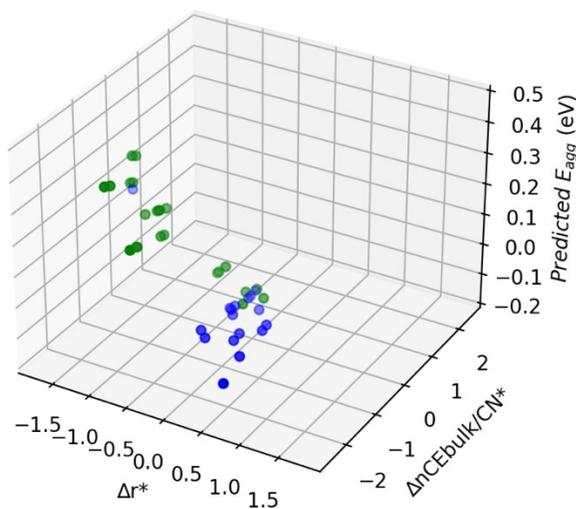


Figure S4. 3D plot of second order polynomial KRR model E_{agg} predictions versus the Δr^* and $\frac{\Delta nCE_{\text{bulk}}}{CN}^*$. The data points are colored based on $\Delta r^* > 0$ and $\frac{\Delta nCE_{\text{bulk}}}{CN}^* < -0.55$ criteria. Systems that satisfy the criteria are represented by the red color, while systems outside this range are SAA, identified using the green color. We excluded the anti-segregation systems (which was first identified and filtered based on Figure S3).

5. DFT Electronic Energy Data for the BE of the Adsorbate to Single Atoms.

Table S5. DFT Electronic Energy of Single Metal Atoms, R-NH, a Single Metal Atom Bonded to R-NH, and the Binding Energy of the Single Atom to the R-NH.

Metal	Single metal atom (Ha)	R-NH (Ha)	M-RNH (Ha)	BE (eV)
Ag	-36.9364	-17.943993	-54.93593	-1.5111831
Au	-33.1443	-17.943993	-51.155727	-1.8339032
Cu	-47.9995	-17.943993	-66.027408	-2.2838738
Ni	-169.106	-17.943993	-187.18503	-3.6679552
Pd	-127.12177	-17.943993	-145.13305	-1.8308724
Pt	-119.9688	-17.943993	-138.04401	-3.5706166

Table S6. DFT Electronic Energy of Single Metal Atoms, R-S, a Single Metal Atom Bonded to R-S, and the Binding Energy of the Single Atom to the R-S.

Metal	Single metal atom (Ha)	R-S (Ha)	M-RS (Ha)	BE (eV)
Ag	-36.9364	-17.61957	-54.639677	-2.2777847
Au	-33.1443	-17.61957	-50.856907	-2.531667
Cu	-47.9995	-17.61957	-65.725453	-2.8948304
Ni	-169.106	-17.61957	-186.86929	-3.9108224
Pd	-127.12177	-17.61957	-144.82736	-2.3407246
Pt	-119.9688	-17.61957	-137.73731	-4.0528659

6. DFT - Geometry Optimization of H₃C-NH and H₃C-S on Ni(111)Ag.

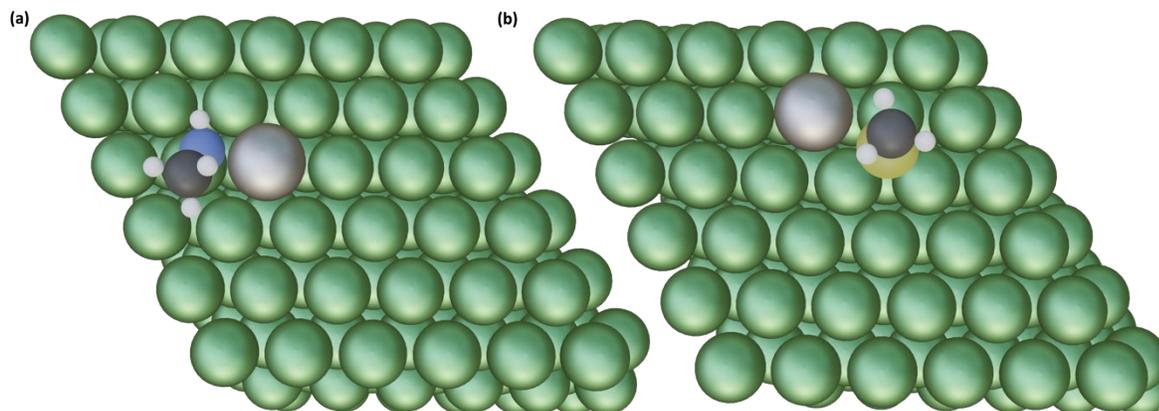


Figure S5. Optimized Ni(111)Ag in the presence of (a) H₃C-NH and (b) H₃C-S. After optimization, the S/N-dopant bond broke and formed a new bond with the metal host.

7. Multi-collinearity check and hyperparameters used in regression models to predict the E_{agg} in non-ligated and ligated SAAs.

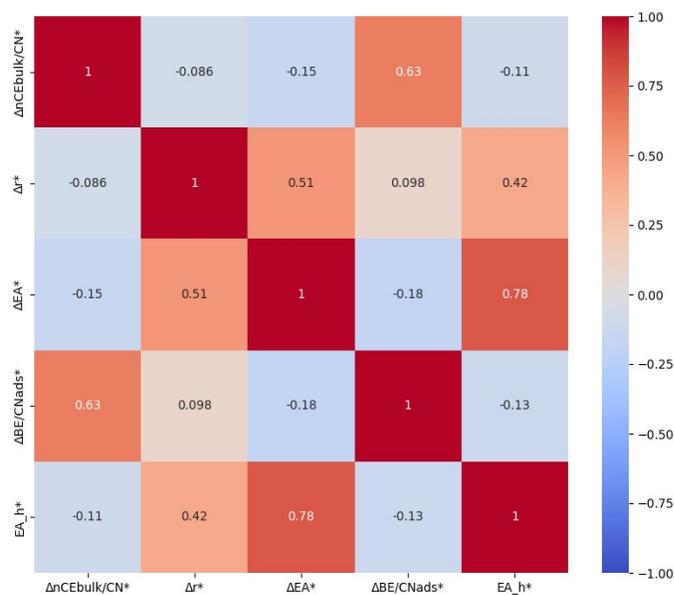


Figure S6. Pearson's correlation table based on the five features (labeled in table) obtained from the Variable Importance plot.

Table S7. Hyperparameters used in the four-feature regression models ($\Delta BE/CN_{ads}$, $\Delta nCE_{bulk}/CN$, ΔEA , and Δr) based on the GridSearchCV results.

Model	Hyperparameter
KRR: 2 nd Order Polynomial	Alpha: 0.6 Gamma: 0.08
KRR: RBF	Alpha: 0.4286 Gamma: 0.7146
KRR: Laplacian	Alpha: 0.4286 Gamma: 0.4291
SVR: 2 nd Order Polynomial	C: 0.21435 Epsilon: 0.001 Gamma: 0.07236
SVR: 3 rd Order Polynomial	C: 0.1429 Epsilon: 0.001 Gamma: 0.2151
SVR: RBF	C: 0.8 Epsilon: 0.09 Gamma: 0.1437
LASSO	Alpha: 0.0621

8. Predicting DFT E_{agg} of non-ligated and ligated systems using different regression models.

Table S8. Different E_{agg} models and their corresponding train, test, and validation MAE.

Model	Test MAE (eV)	Validation MAE (eV)	Train MAE (eV)	Validation-Train Δ MAE (eV)
SVR: 3 rd poly	0.252	0.200	0.182	0.017
KRR: 2 nd poly	0.189	0.189	0.182	0.008
KRR: RBF	0.175	0.167	0.112	0.058
KRR: Lap	0.154	0.161	0.100	0.61
SVR: RBF	0.161	0.160	0.142	0.017
SVR: 2 nd poly	0.325	0.283	0.274	0.009
OLS	0.196	0.209	0.203	0.006
LASSO	0.202	0.209	0.203	0.006

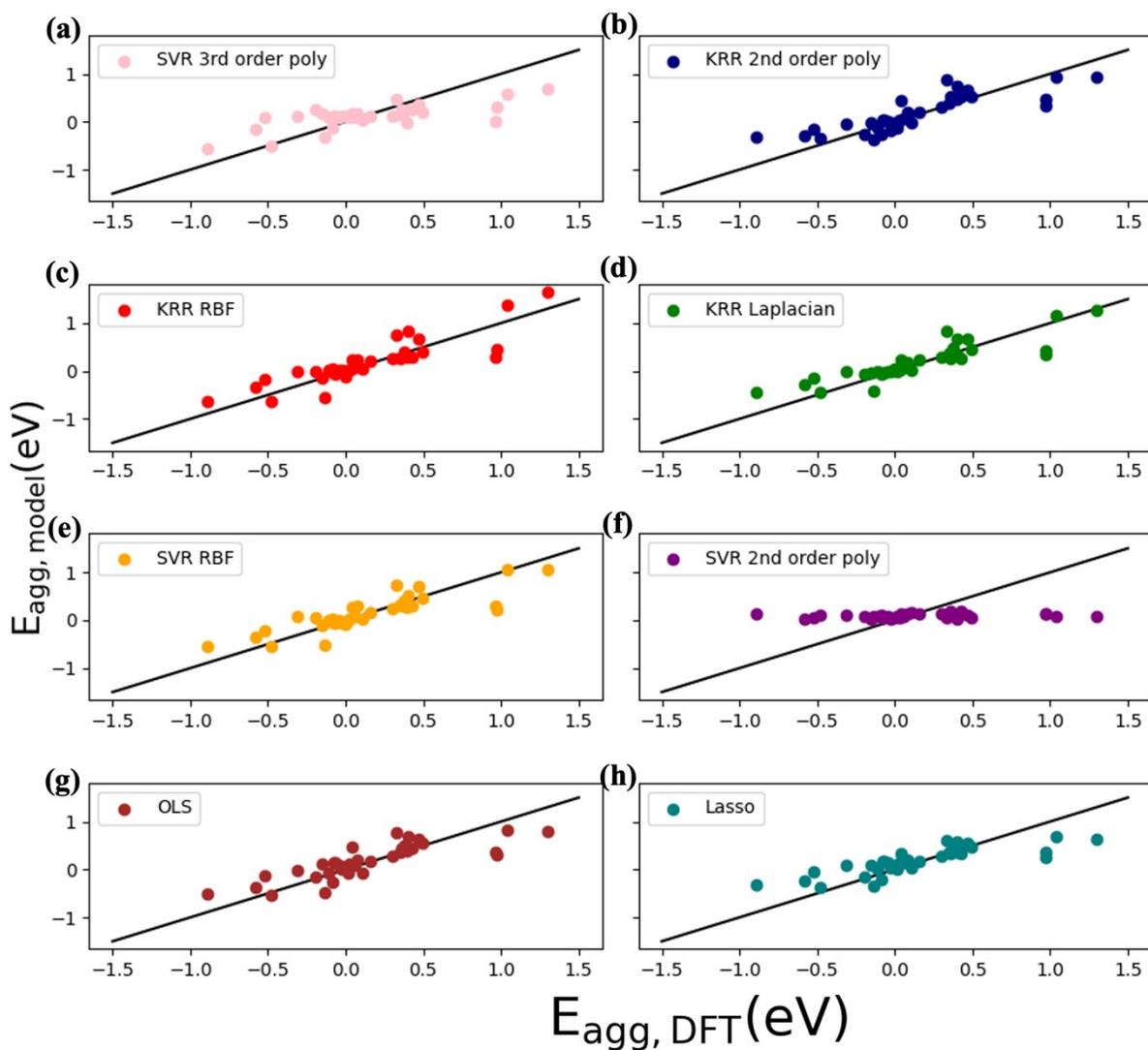


Figure S7. Parity plot between different regression models based on the test set using the two features ($\Delta BE/CN_{ads}$, $\Delta nCE_{bulk}/CN$, ΔEA , and Δr) and $E_{agg,DFT}$. (a) Support Vector Regression (SVR): 3rd order polynomial, (b) Kernel Ridge Regression (KRR): 2nd order polynomial, (c) KRR: RBF, (d) KRR: Laplacian, (e) SVR: RBF, (f) SVR: 2nd order polynomial, (g) Linear Regression, and (h) LASSO.

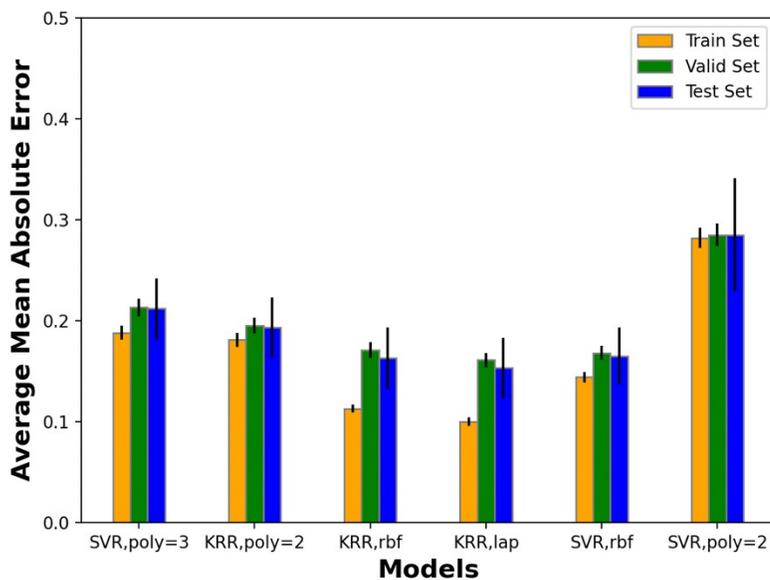


Figure S8. Average of MAE of the different models tested at 100 different train/test splits (using different random seeds). The error bars reflect the standard deviation of the different 100 train/test splits of each model.

Table S9. Comparison between the SVR RBF predictions and experimental observations.

Host	Dopant	Ligand	Experimental Observations	Predictions
Pt	Au	thiol	Well-dispersed ⁵	Au monomers
Au	Pt	thiol	A few agglomerates ⁵	Pt clusters
Cu	Ni	thiol	Ni aggregates ⁶	Ni aggregates
Pd	Au	none	Mixture of Au monomers and dimers ⁷	Au monomers
Cu	Pd	none	Pd uniformly dispersed ⁸	No aggregates
Au	Cu	none	Well-dispersed ⁹	Cu monomers
Au	Pd	none	Pd monomers ⁷	Pd monomers

9. Aggregation Energy in the Ligated SAAs versus Change in the Adsorption Energy of the SAA and Host.

In Figure S9, we plot the aggregation energies of ligated systems against the change in the ligand adsorption energies between the dopant and the host. When the change in adsorption energy is positive, meaning that the ligand is more stable adsorbing on the host metal than on the dopant in the SAA, we find that aggregation is not favored (indicated by the positive sign). This is consistent with the idea that for aggregation to occur, the affinity between the adsorbate and the dopant should be stronger than that between the adsorbate and the host metal. However, we did observe cases where the change in adsorption energy was positive, but the aggregation energy was negative. In these instances, the change in adsorption energy was significantly small (less than 0.06 eV), except for Cu(111)Ni-HNCH₃, Au(100)Ag-SCH₃, Au(100)Pd-SCH₃, and Pd(111)Pt-SCH₃. Furthermore, in these exceptions, segregation energy values indicate that dopant segregation is either unfavorable (i.e., positive segregation energy value) or shows a weak tendency, as in the case of Cu(111)Ni-HNCH₃ (~-0.058 eV). This suggests that when dopant on the surface (segregation energy) is not thermodynamically stable, then aggregation is not favored. Additionally, we note that the larger deviations from the fit, particularly in the thiolate case (pink edge marker), result from a new configuration on the (111) facet.

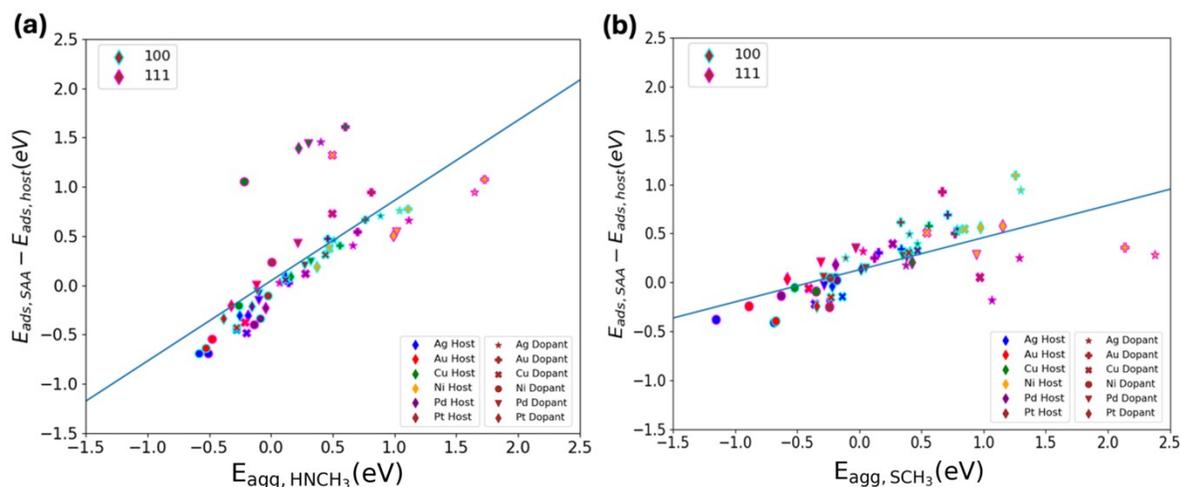


Figure S9. Plot between the aggregation energy of ligated systems (a) amine and (b) thiol groups and the change in the adsorption energies between the SAA and host systems. Color indicates the different metal hosts, marker type indicates the different dopants, and edge color represents the different facets.

References

- 1 E. Clementi, D. L. Raimondi and W. P. Reinhardt, *J Chem Phys*, 1967, **47**, 1300–1307.
- 2 Ł. Mentel, 2014.
- 3 M. Salem, D. J. Loevlie and G. Mpourmpakis, *The Journal of Physical Chemistry C*, 2023, **127**, 22790–22798.
- 4 L. Farsi and N. A. Deskins, *Physical Chemistry Chemical Physics*, 2019, **21**, 23626–23637.
- 5 L. J. Torres-Pacheco, A. De Leon-Rodriguez, L. Álvarez-Contreras, M. Guerra-Balcázar and N. Arjona, *Electrochim Acta*, 2020, **353**, 136593.
- 6 L. Chen, H. Xu, H. Cui, H. Zhou, H. Wan and J. Chen, *Particuology*, 2017, **34**, 89–96.
- 7 N. Takehiro, P. Liu, A. Bergbreiter, J. K. Nørskov and R. J. Behm, *Phys. Chem. Chem. Phys.*, 2014, **16**, 23930–23943.
- 8 F. Xing, J. Jeon, T. Toyao, K. Shimizu and S. Furukawa, *Chem Sci*, 2019, **10**, 8292–8298.
- 9 H. Wang, D. Liu and C. Xu, *Catal Sci Technol*, 2016, **6**, 7137–7150.