Catalysis
Science &
Technology

The effects of near-surface atomic order on the catalytic properties of Cu_3Au and $CuAu_3$ intermetallics for the CO_2 reduction reaction

Journal:	*Catalysis Science & Technology*
Manuscript ID	CY-ART-05-2023-000658.R1
Article Type:	Paper
Date Submitted by the Author:	21-Sep-2023
Complete List of Authors:	Garcia Verga, Lucas; University of Sao Paulo, Chemistry; Johns Hopkins University, Department of Materials Science and Engineering Wang, Yunzhe; Johns Hopkins University, Department of Materials Science and Engineering Chakraborty, Tanmoy; University of Warwick, Department of Chemistry; Johns Hopkins University, Department of Materials Science and Engineering Da Silva, Juarez; University of Sao Paulo, Mueller, Tim; Johns Hopkins University,

The effects of near-surface atomic order on the catalytic properties of Cu_3Au and $CuAu_3$ intermetallics for the CO_2 reduction reaction

Lucas G. Verga[1,2], Yunzhe Wang[2], Tanmoy Chakraborty[2,3], Juarez L. F. Da Silva[1], and Tim Mueller[2*]

[1] São Carlos Institute of Chemistry, University of São Paulo, PO Box 780, 13560-970, São Carlos, SP, Brazil.

[2] Department of Materials Science and Engineering, Johns Hopkins University, Baltimore, Maryland, 21218, United States.

[3] Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, United Kingdom.

E-mail*: tmueller@jhu.edu

Abstract

Density Functional Theory (DFT) calculations combined with Cluster Expansions (CE) were employed to explore the catalytic activity of stoichiometric and non-stoichiometric $Cu_3Au(100)$, $Cu_3Au(111)$, $CuAu_3(100)$, and $CuAu_3(111)$ surfaces for CO_2 reduction to CO. We identified several adsorption sites which otherwise would not be considered using slab models generated by cleaving ordered bulk structures. For $Cu_3Au(111)$ and $CuAu_3(111)$, the adsorption sites with the highest and lowest estimated onset potential values found through our approach would be missed if modeled from bulk-derived slabs, showing the ability of our strategy to explore the richness of alloy catalysts that arises from the formation of near-surface atomic ordering effects. Finally, we performed simulated annealing calculations based on CE fit to predict CO and COOH adsorption energies. This analysis led us to find adsorption site ensembles that can deviate from the linear scaling relationship that usually hinders the discovery of more active and selective catalysts for CO production.

1. Introduction

The risks associated with CO_2 accumulation in the atmosphere have motivated the scientific community to develop economically viable technologies for carbon capture, storage, and utilization.(*1*) The electroreduction of CO_2, known as the CO_2 reduction reaction (CO_2RR), can contribute to such a challenge by enabling the recycling of CO_2 into value-added products. Moreover, if the CO_2RR is performed with clean energy sources, it can also help with problems associated with the intermittency of wind or solar energy.(*2*) The CO_2RR can yield various products, such as CO, CH_4, C_2H_5OH, and C_2H_4, depending on the catalyst used for the chemical process. However, CO_2RR generally requires high overpotentials using known catalysts. Therefore, there is a need to develop improved CO_2RR catalysts, which can be accomplished by better understanding the relationship between the atomic structure of a catalyst and its activity and selectivity.

After the pioneering works of Hori et al.,(*3-5*) copper (Cu) and Cu-based materials have been widely studied as possible catalysts for the CO_2RR. One of the reasons for the interest in Cu catalysts is their unique ability to reduce CO_2 to a variety of molecules such as CO, CH_4, C_2H_4, and C_2H_5OH, totaling up to 16 different products.(*6*) Studies of the CO_2RR on Cu have investigated possible reaction mechanisms,(*7-10*) structural characteristics such as specific facet exposure or nanoparticle size shape and size that might impact the mechanisms and activity,(*11-16*) and other effects.(*17-23*) Meanwhile, gold (Au) and Au-based catalysts are also interesting options for the CO_2RR due to their ability to produce CO with high faradaic efficiency.(*5*) The generated CO, together with H_2, constitutes a mixture called syngas, which can be further reformed to other products such as methane or methanol through thermally activated reactions.(*24*) Several studies involving the CO_2RR on Au catalysts have explored the reasons for the high CO selectivity or ways to increase the catalyst activity and/or selectivity.(*25-29*)

A strategy commonly employed to tune the activity and selectivity of Cu- and Au-based catalysts is the addition of one or more elements to the catalyst composition to generate an alloy. Alloying creates new variables that can be manipulated to advance beyond single-metal catalysts, such as the choice of the metals used for the alloys, the ratio between the elements, and the exploration of different phases, such as ordered or disordered alloys. The use of alloys can change the catalytic behavior as compared to single-metal catalysts due to the presence of strain effects,(*30-34*) ensemble effects,(*33-36*) ligand effects,(*33, 34, 37*) or by generating bifunctional catalysts.(*38-40*) The increased number of variables provides new routes for optimization but increases the complexity of the exploration, making in-depth atomistic knowledge about each effect important information for catalyst design.

For the above reasons, Cu-Au bimetallic systems become natural catalyst candidates. The results observed in the literature for Cu_xAu_y catalysts vary depending on the catalyst morphology and composition. Kim et al.(*41*) observed that Cu_3Au, CuAu, and $CuAu_3$ nanoparticles with average diameter of approximately 11 nm generate CO as the main product. While the selectivity does not vary significantly for these nanoparticles, the activity follows a volcano shape with $CuAu_3$ at the peak. The $CuAu_3$ and CuAu turnover rates for CO were 93.1 and 40.4 times higher than Cu, respectively. Similar results indicating high CO selectivity were also obtained for other Cu_xAu_y compositions, such as $Cu_{90}Au_{10}$, $Cu_{80}Au_{20}$, and CuAu, with the latter being the most active among the tested compositions.(*42*)

CO production from the CO_2RR was also observed for Cu_xAu_y foils and 6 and 2 nm nanoparticles, with the small alloy nanoparticles having a yield per surface area up to 175 times higher than bulk Au, which was attributed to both the increase of low-coordinated sites and synergistic effects from alloys.(*43*) In addition to nanoparticle size and composition, the atomic ordering of CuAu nanoparticles was also demonstrated to impact the activity for such alloys, with ordered CuAu nanoparticles being up to 3.2 times more active than disordered nanoparticles with the same composition.(*32*) While the selectivity towards CO seems to be predominant on certain Cu_xAu_y surfaces and nanoparticles, other strategies in the literature can be employed to favor other products. For instance, Cu_3Au nanowire arrays have been shown to generate C_{2+} products,(*44*) while a Au-Cu catalysts with 7% atomic percentage of Au in the surface shows controlled availability of adsorbed CO due to regulated concentration of CO_2 and are selective for CH_4 production.(*45*)

Because of the several effects that impact the activity and selectivity of bimetallic catalysts, the generation of realistic models to treat these systems computationally remains a challenge. One commonly employed approach is to generate models for bimetallic catalysts from experimental information. For example, upon experimental evidence of Au shell formation, the CO_2RR on ordered CuAu nanoparticles was modeled using strained Au slabs.(*32*) However, experimental information about alloy catalysts will not always be readily available. When this is the case, a possible solution is the generation of slabs cleaved along a particular direction from ordered bulk alloys. However, slabs derived from ordered bulk alloys can neglect the important atomic ensembles that could form in the surface due to the break of symmetry, which can induce different atomic order than that found in the bulk structure. For example, for the oxygen reduction reaction, calculations indicate that the surface and subsurface elemental distribution for Pt_3Ni intermetallic alloys play a significant role in the catalyst activity.(*46*)

Here, we use grand canonical Monte Carlo (GCMC) and DFT calculations to explore different adsorption sites based on first-nearest neighbors on Cu_3Au and $CuAu_3$ (100) and (111) slabs and how the diversity of sites can change the CO_2RR activity towards CO. We discuss the elemental distribution on surface and subsurface layers as a function of chemical potential used in the GCMC calculations. Subsequently, we present a method that can be used to distinguish sites using information on the atom closest to the adsorbate and its nearest neighbors. This method is used within GCMC calculations, allowing us to find, without prior experimental information, several types of adsorption sites that would not be present on slabs generated by cleaving the bulk structure. For all distinct and frequent types of adsorption sites, we compute the adsorption of CO and COOH using DFT calculations, which are used to describe the CO_2RR activity towards CO on each site. Interestingly, we find sites responsible for the lowest and highest CO_2RR onset potentials that would not be present on bulk-derived slabs. We finish our exploration of atomic ordering effects by finding atomic ensembles that deviate from linear-scaling relationships between CO and COOH adsorption energies by up to −0.33 eV when both molecules are assumed to adsorb on the same top site.

2. Theoretical Approach and Computational Details

2.1. Energy Calculations

All density functional theory calculations were performed using the Vienna *ab initio* simulation package (VASP)(*47-49*) together with the projected augmented wave (PAW) method.(*49, 50*) We employ the Perdew-Burke-Ernzerhof (PBE) exchange-correlation energy functional(*51*) and the semi-empirical D3 correction proposed by Grimme.(*52*) The Cu_3Au and $CuAu_3$ slabs were constructed using their optimized bulk lattice constants, 3.717 Å and 3.993 Å, respectively, obtained with the same DFT-PBE+D3 approach. For the Brillouin zone integration, we employed efficient grids generated by the *k*-point grid server (*53, 54*) to generate a grid corresponding to a real-space supercell with a minimum distance of at least 45.0 Å between lattice points. This was found to converge the formation energy within 1 meV/atom, which is sufficient for our approach, (see ESI†). The plane wave cutoff energy was set to 489 eV, which is 12.5% higher than the maximum cutoff energy suggested for all the PAW potentials used in our calculations. For the slab calculations that are used to train the cluster expansions, the convergence criteria for the electronic self-consistent iteration was set to be 10^{-4} eV/simulation cell and the ionic relaxation loop stopped when the change in energy between two ionic steps was smaller than 10^{-3} eV/simulation cell. For the adsorption of molecules, we used as convergence criteria 10^{-5} eV/simulation cell for the electronic iterations and 0.05 eV/ Å for the atomic forces on all atoms. A vacuum thickness of at least 12 Å was added perpendicularly to the surface to ensure negligible interaction among periodic images. Further technical details related to the DFT calculations and a summary of the employed convergence thresholds can be found in the supplementary information.

2.2. Cluster Expansion and Monte Carlo Calculations

We calculated energies as a function of atomic order using the cluster expansion method,(*55, 56*) which is a generalized Ising model including many-body interactions. The cluster expansion method is widely used to study atomic order in materials including bulk systems,(*57-59*) surfaces,(*46, 60, 61*) and nanoparticles.(*62-64*) A brief introduction to cluster expansion is provided within this section, while further details on the theory can be found in review and software implementation papers.(*65-67*) In the cluster expansion framework, the structure of the material is represented as a set of lattice sites that may be occupied by different species, and the occupancy of the i^{th} site is given by the site variable s_i. For the case of a binary alloy, a site variable value of +1 indicates one species is present at the site, and -1 indicates the other species is present, as illustrated in Figure 1a). Each configuration, i.e., distribution of chemical species at the different sites, is therefore represented by the set of all site variables, $\mathbf{s}$. Using this representation one can express a property of the material, such as the total energy, as a linear combination of basis functions known as cluster functions, where each cluster function is the product of all site variables in a cluster of sites:

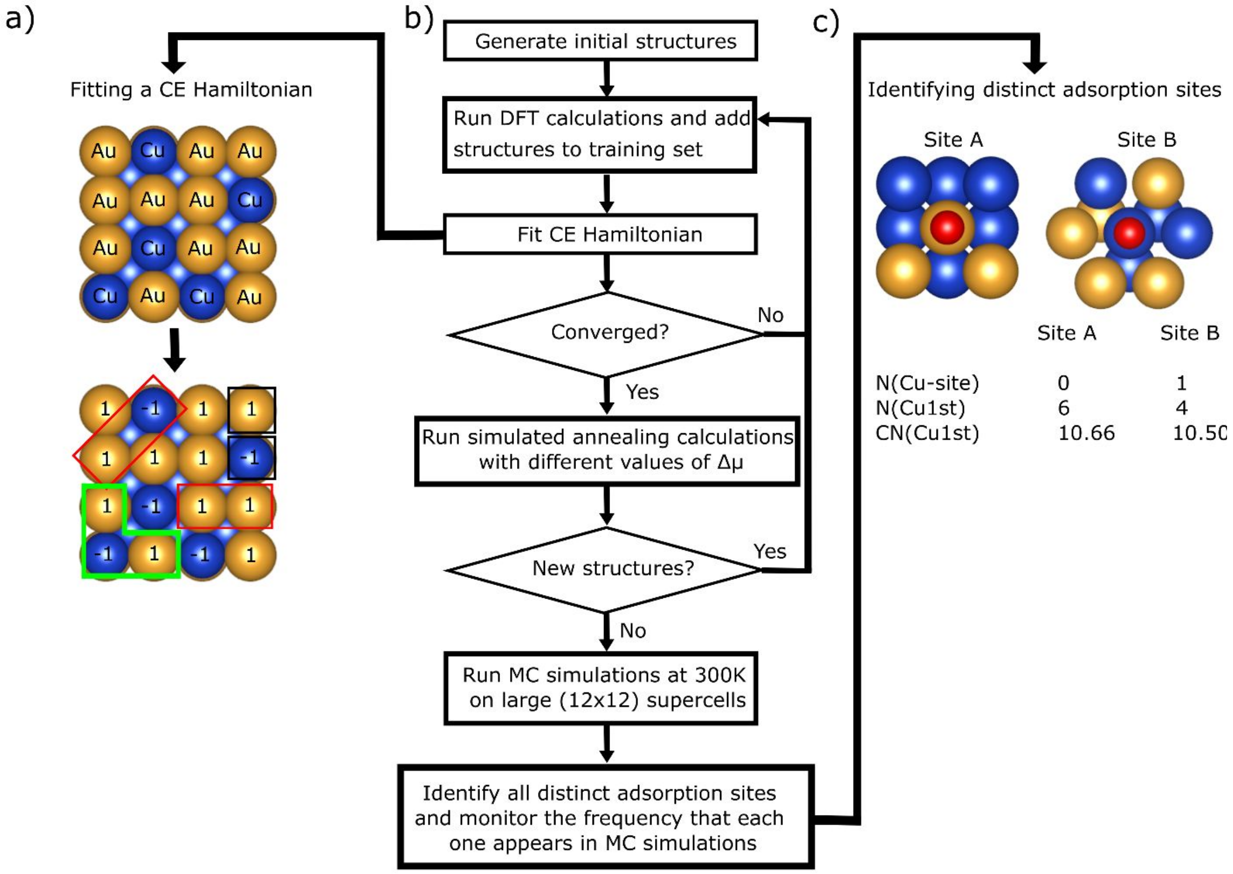
$$E(\mathbf{s}) = V_0 + \sum_{clusters} V_{cluster} \prod_{i \in cluster} s_i, \tag{2}$$

where, V_0 is the empty cluster and $V_{cluster}$ are the coefficients of this expansion, known as effective cluster interactions (ECIs). We fit the ECIs to DFT-calculated formation energies using a Bayesian multivariable linear regression.(*68*) The formations are defined as:

$$E_{Form} = E_{tot}^{Cu_xAu_y} - xE_{tot}^{Cu} - yE_{tot}^{Au}, \tag{3}$$

where $E_{tot}^{Cu_xAu_y}$ is the energy of a slab with x Cu atoms and y Au atoms, and E_{tot}^{Cu} and E_{tot}^{Au} represent the energies per atom of bulk Cu and Au, respectively.

Figure 1a) shows examples of single-site, nearest-neighbor and second-nearest-neighbor two-site, and three-site clusters for a given configuration. Any subset of sites could form a cluster, and if all the possible clusters are included in the expansion, it is exact. In practice, the number of clusters is truncated, and only symmetrically distinct small and compact clusters are included.(*65*) Our cluster expansion construction includes the empty cluster, the one-body cluster, all 2-body clusters with no more than 10 Å between sites, and all 3- and 4-body clusters with no more than 4 Å between sites. As we have used the maximum distance between sites as a parameter to truncate the clusters that were included in the cluster expansion construction, the systems with larger lattice parameters end up with fewer clusters in the expansion than systems with small lattice parameters. In total, we had 188, 196, 167, and 174 symmetrically distinct cluster functions for the CE of $Cu_3Au(100)$, $Cu_3Au(111)$, $CuAu_3(100)$, and $CuAu_3(111)$, respectively. The effective cluster interactions were fit to DFT data using the Bayesian approach(*68, 69*) with a similar prior distribution as used in a previous study,(*46*) as detailed in Section 5-ESI†.



	Site A	Site B
N(Cu-site)	0	1
N(Cu1st)	6	4
CN(Cu1st)	10.66	10.50

Figure 1: Schematics of employed methodology a). Illustration of how clusters and site variables are defined within a CE. b) Flowchart of the methodology used to generate slab models and identify and monitor distinct adsorption sites. c) Illustration of the criteria used to identify distinct adsorption sites based on the atom that is nearest to the adsorbate and its first-nearest neighbors.

Figure 1b) shows the general methodology used to train the CE and run the Monte Carlo (MC) simulations to search for slabs and adsorption sites that represent stoichiometric and off-stoichiometric alloy surfaces. For each composition and surface orientation, we generated 50 nine-layer slabs as initial training data using the computed lattice constant for the underlying bulk composition. Here we assume that the compositions and lattice constants of the underlying bulk intermetallic phases are relatively constant across a range of chemical potentials, consistent with previous analysis.(*70*) The structures for the initial training data were selected using the methodology described in Mueller et al.(*69*) in which the variance in the predicted property values is minimized. The use of nine layers provides sufficient sub-surface depth to explore the formation of near-surface alloys.(*46*) After the initial training of the CE, we searched for new ground state structures using grand canonical Monte Carlo (GCMC) simulated annealing calculations on (3×3) and (4×4) slabs. The GCMC simulations are parameterized by the difference in chemical potentials between Cu and Au:

$$\Delta\mu = \mu(Au) - \mu(Cu) \tag{4}$$

where $\mu(Cu)$ is the chemical potential of Cu and $\mu(Au)$ is the chemical potential of Au. The simulated annealing calculations were performed at temperatures ranging from 1500K to 20K, where the temperature changed in each iteration as $T(i+1) = 0.9T(i)$. For each combination of T and $\Delta\mu$ values, we performed $10000 * N_{sites}$ equilibration and recording steps, where N_{sites} is the number of sites in the system.

To estimate the windows in which Cu_3Au and $CuAu_3$ should be in equilibrium, we constructed a 0 K convex hull of DFT-calculated formation energies (Figure S3b)). The slopes of the tie lines on the convex hull on either side of a given phase provide the upper and lower limits of $\Delta\mu$ at which that phase is calculated to be stable at 0 K. To search for new ground state structures, we ran MC simulations at chemical potentials within this range, sampling at increments of 0.01 eV. Using the formation energies, we determined the structures in the convex hull for each facet and composition (see Section 5-ESI†). We included the distinct ground-state structures found in this process in our training set and repeated the process until no new structures were found. In total, we calculated 112, 252, 113, and 157 structures for the Cu_3Au(100), Cu_3Au(111), $CuAu_3$(100), and $CuAu_3$(111) surfaces, respectively. Additional details are provided in Section 5-ESI †.

To determine the room-temperature distribution of adsorption sites on each of the facets of each of the phases, we ran GCMC simulations on (12×12) slabs at 300K with different values of chemical potential within the values at which each phase was calculated to be stable sampling at increments of 0.01 eV. For each GCMC calculation, we took 100 snapshots at intervals of $20000 \times N_{sites}$ MC steps after an initial $200000 \times N_{sites}$ MC steps used for equilibration. Using the snapshots, we tracked the frequency at which different types of adsorption sites appeared in the outermost layer. The procedure for defining these adsorption sites is discussed in Section 3.2.

2.3. Reaction free energies

To assess the onset potentials and estimate the catalytic activity of the different adsorption sites for the CO_2 reduction toward CO, we computed the adsorption energies, ΔE, for reaction intermediates using the following equation:

$$\Delta E = E_{tot}(*C_nO_mH_p) - E_{tot}(slab) - nE_{tot}(C) - mE_{tot}(O) - pE_{tot}(H), \tag{5}$$

where $E_{tot}(*C_nO_mH_p)$ is the energy of a molecule with composition $C_nO_mH_p$ adsorbed on the slab, $E_{tot}(slab)$ is the energy of the clean slab, $E_{tot}(O)$ is the energy difference between H_2O and H_2, $E_{tot}(C)$ is calculated as the difference between the energy of CO_2 and 2 times $E_{tot}(O)$, and $E_{tot}(H)$ is half of the energy of the H_2 molecule. Using this convention, more negative values of the adsorption energy, ΔE, indicate stronger bonds.

To study the reduction of CO_2 to CO, we considered the reaction pathway CO_2(g) → *COOH → *CO → $CO^{‡}$ → CO(g), which has been identified in previous studies.(*12, 29, 36*) We considered that the reduction happens through a sequence of proton coupled electron steps (PCETs) and modeled them under a low-coverage assumption. We also assumed that contributions obtained from vibrational frequencies and the solvation energies would not present strong dependency on the different adsorption sites. Following this methodology, the adsorption energies of *CO and *COOH on each adsorption site becomes sufficient to generate free energy reaction diagrams for CO_2RR towards CO using the computational hydrogen electrode (CHE)(*7, 71*) model. Additional details of the approximations, reaction mechanism, CHE model, and the equations showing how to obtain free energy diagrams based on adsorption energies are described in the electronic supporting information, Section 2-ESI†.

In addition to using the CE to investigate the adsorption site distribution on different surface facets and composition, we also trained two cluster expansions to predict *CO and *COOH adsorption energies on top sites. These CEs were used to explore the construction of new descriptors for adsorption energies and explore other adsorption site ensembles that could be interesting for CO_2RR. The formation energies, as described in equation 2, were used to fit the cluster expansions for adsorbate-free slabs. For the systems with adsorbed molecules, we redefined the formation energies as:

$$E_{Form} = E_{tot}^{Cu_xAu_y}(*C_nO_mH_p) - xE_{tot}^{Cu} - yE_{tot}^{Au} - nE_{tot}(C) - mE_{tot}(O) - pE_{tot}(H), \tag{6}$$

where $E_{tot}^{Cu_xAu_y}(*C_nO_mH_p)$ is the energy of a $C_nO_mH_p$ molecule adsorbed on a Cu_xAu_y slab, E_{tot}^{Cu} and E_{tot}^{Au} are the energies per atom of bulk Cu and Au, respectively, while $E_{tot}(O)$, $E_{tot}(C)$, and $E_{tot}(H)$ are the energies used as reference for oxygen, carbon and hydrogen atoms as defined for equation 4. In this way, subtracting the formation energy with (equation 5) and without an adsorbed molecule (equation 2) yields the adsorption energies, ΔE, as defined in equation 4.

For each cluster expansion used to obtain the adsorption energies, we included all the previous structures and additional 23 systems with an adsorbed molecule (*CO or *COOH). The energies used in fitting the cluster expansions were the lowest energy configuration of the molecule on each top site. In Section 3.4, these cluster expansions are also used within simulated annealing calculations to search for adsorption site ensembles that would favor CO_2RR by deviating from linear-scaling relationships (LSR) between adsorption energies of *CO and *COOH. These simulated annealing calculations were performed at temperatures ranging from 2000K to 0.5K, where the temperature changed in each iteration as $T(i+1) = 0.9T(i)$. For each value of T, we performed $10000 \times N_{sites}$ equilibration and recording steps, where N_{sites} is the number of sites in the system.

3. Results and Discussion

To explore the impact of atomic ordering on CO_2RR using Cu_3Au and $CuAu_3$ (100) and (111) slabs as models, we separate our results and discussion into four sections. We first show how our CEs compare to DFT-calculated formation energies and discuss the elemental distribution as a function of the chemical potential for different facets and underlying bulk compositions. Then, we discuss our method for distinguishing adsorption sites and explore their prevalence as a function of slab Au composition. The sites distinguished in this process are tested for CO_2RR by computing *CO and *COOH adsorption energies and analyzing them with the CHE method. Finally, we employ new CEs and simulated annealing calculations to find adsorption site ensembles that can deviate from linear-scaling relationships between *CO and *COOH binding, further exploring and discussing the impact of atomic ordering on CO_2RR.

3.1. Cluster Expansion and Surface Atomic Ordering

The cluster expansions for $Cu_3Au(100)$, $Cu_3Au(111)$, $CuAu_3(100)$, and $CuAu_3(111)$ had root-mean-squared error leave-one-out cross-validation errors (LOOCV) of 2.18, 3.18, 2.00, 1.61, meV/atom, respectively, with very good agreement between the CE-calculated and DFT-calculated energies (Figure 2). The LOOCV values compare well to other machine-learned energy models for alloys that have recently shown accuracies from 1 to 13 meV/atom. (*72-75*)

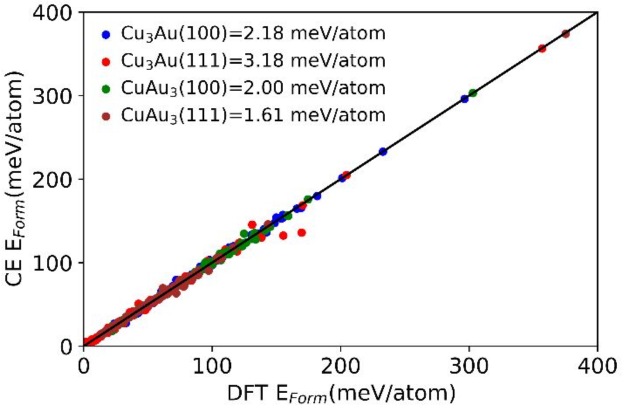


Figure 2: Parity plot between the formation energies obtained via the CE Hamiltonian as predicted using leave-one-out cross-validation (LOOCV) and DFT calculations, together with the LOOCV errors for each CE.

We used the cluster expansions in Monte Carlo simulations to predict near-surface atomic structure. Here we investigate the properties of clean surfaces to study the relationship between local atomic order and adsorption, but we note that potential adsorbate-induced structural changes would be an interesting topic for future study. Even when the bulk of material is in an ordered, intermetallic phase, there can be significant variation in the near-surface structure and composition as a function of Au and Cu chemical potentials. In general, we find that the near-surface Au composition increases as a function of $\Delta\mu$, as expected (Figure 3). Moreover, we observe a Au enrichment of the slab first layer as compared to the overall slab composition for all values of $\Delta\mu$, facets, and underlying bulk composition, which is expected due to the tendency of atoms with lower surface energy to segregate to the surface.(*76, 77*) For Cu_3Au at 300 K, we observe that across the entire chemical potential range the surface layer is the most Au-rich, with this effect being more significant for (100) than for (111) slabs (Figure 3a). For the subsurface, we observe the opposite effect, with $Cu_3Au(100)$ slabs showing a second layer composed solely of Cu atoms. On $Cu_3Au(100)$ at low values of $\Delta\mu$, the first and second layers have 50% and 0% Au, respectively, which is also the pattern found in the bulk $L1_2$ structure. As $\Delta\mu$ increases, nearly all the additional Au resides in the outermost layer, increasing the composition of this layer to about 72% Au. For (111) facets, there is less dramatic segregation between the first and second layers. We observe a first layer that varies from 33 to 51% Au composition, while the subsurface layer is more Cu-rich, with Au composition between 0% and 25%. For the $Cu_3Au(111)$ slab with nearly the ideal 25% Au composition, we observed the first layer having around 42% Au.

The Au enrichment of the topmost surface layer is consistent with results from the literature that present Au compositions ranging from 39 to 51%.(*78, 79*) The results from our analysis are also comparable with the insights previously obtained with different computational approaches.(*73, 80, 81*) These show the tendency to form a Au-rich shell on Cu_xAu_y nanoparticles. Moreover, some of the different behaviour between (111) and (100) facets that we observed is also present, e.g nanoparticles with Cu atoms in the surface are observed more commonly on the (111) facet than on the (100) facet.(*80*)

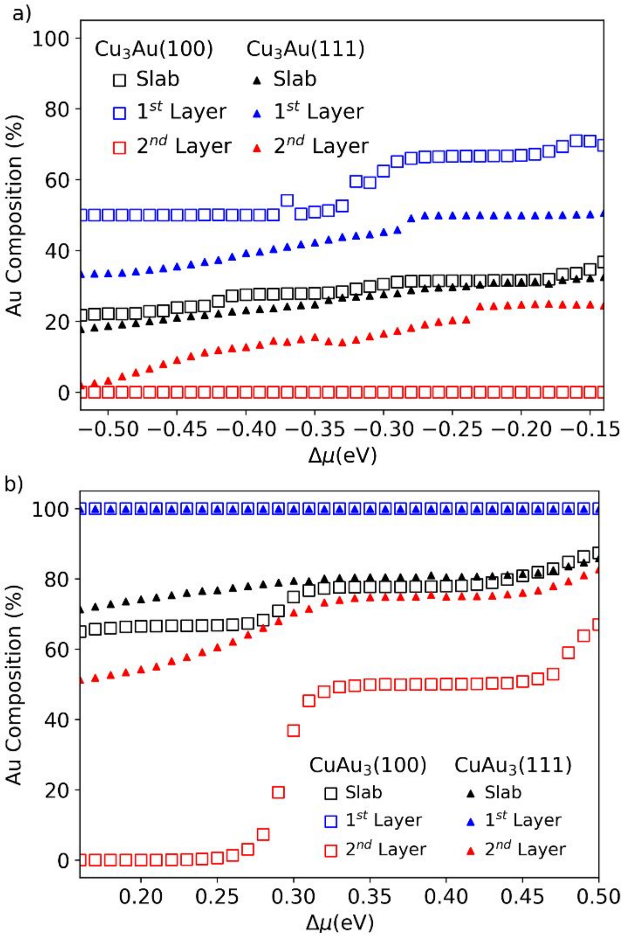


Figure 3: Near-surface composition profiles of the slabs obtained through GCMC calculations for the systems at 300 K for a) Cu_3Au(100) and (111) surfaces and b) $CuAu_3$(100) and (111) surfaces. Blue (red) curves show the Au composition as a function of $\Delta\mu$ for the first (second) layer of the slab, while black curves show the Au composition for the 9-layer slab. Triangles and squares represent the data obtained for (111) and (100) facets, respectively.

For $CuAu_3$(100) and (111) surfaces, we observe that the first layer contains only Au atoms across the entire chemical potential range at which the bulk structure is estimated to be stable. For the subsurface, we observe a similar trend as the one observed for Cu_3Au(100) and (111) slabs, i.e., this layer is more Cu-rich than the overall slab. For (100) slabs, we observe two distinct surface phases: at low values of $\Delta\mu$, we observe a nearly 100% Cu subsurface layer, followed by a phase with nearly 50% Au composition at higher $\Delta\mu$ values. Near the highest values of $\Delta\mu$ for which $CuAu_3$ is stable, the second-layer Au composition rises above 50%. For (111) slabs, the increase in Au composition happens more continuously, with the Au content from the subsurface layer going from 51% to up to 83% across the range of $\Delta\mu$ values. There is a plateau at about 75% Au that corresponds to the formation of an ordered second layer similar to what one obtains for the (111) surface cleaved from the ordered $CuAu_3$ bulk.

3.2. Selection of Adsorption Sites

To better understand how the variations in surface structure may affect the catalytic activity, we identify the most common types of adsorption sites that frequently occurred in Monte Carlo simulations, classified by their local environments. We base our classification on top sites following findings from a previous study (*14*) that indicated that COOH would preferentially adsorb on such sites. For *COOH, we consider different molecule orientations with the C atom adsorbed on the top site. For *CO, we calculate the adsorption on the classified top site and on all distinct bridge sites nearby. We selected top and bridge adsorption sites for *CO after preliminary calculations showed these as the preferred ones for *CO adsorption on Cu_xAu_y alloys. For both molecules, the lowest energy configuration was used to create free energy diagrams.

Thus, each top site was represented with a single vector containing three values: 1) the number of Cu atoms in the surface site directly below the adsorbed molecule (1 or 0); 2) the number of Cu nearest neighbors for the adsorption site; 3) the average of the coordination number of Cu nearest-neighbors. We employ a cutoff radius larger than the average bond length for each underlying bulk to determine the first-nearest-neighbors of each atom. As we include the number of Cu-nearest neighbors as a descriptor, both surface and subsurface sites are accounted for in the representation and the proportion of Cu atoms as surface or subsurface neighbors is distinguished by using the average coordination number as an additional descriptor. As this analysis is applied on snapshots of MC calculations, there are no concerns about bond length increases due to relaxation processes, making the determination of first-nearest-neighbors straightforward. A few examples of how this vector is constructed are provided in Figure 1c). This approach allows us to identify whether a Cu or Au atom is present at the adsorption site, the composition of the nearest neighbors, and how the nearest neighbors are distributed in the surface or subsurface. This approach is similar to the one employed in a recent study focused on Ni_xGa_y catalysts.(*82*) The main difference is that here we apply the descriptor always considering a single surface orientation and bulk stoichiometry, which allow us to reduce the number of fingerprints from seven to three.

To identify the types of sites to be studied regarding their catalytic activity for the CO_2RR, we select all types of sites with a prevalence higher than 0.1% during the GCMC calculations at 300 K on (12×12) surface models for each facet and bulk composition. Then, representations of the selected adsorption sites were searched within the (3×3) and (4×4) models that were used to generate the CE. This step was necessary to ensure that we had systems that we could use for DFT calculations that contained the representative adsorption sites. The catalytic activity for CO_2 reduction toward CO was estimated by generating free energy diagrams using DFT-calculated *CO and *COOH adsorption energies.

For stoichiometric $Cu_3Au(100)$, we identify five different types of common adsorption sites. One of these is a Cu site surrounded by an equal number of Cu and Au atoms, and the other four are Au sites with different numbers of Cu and Au first-nearest-neighbors. For $Cu_3Au(100)$, the adsorption sites marked as Site 1 and Site 2 have the same characteristics of first-nearest-neighbors as the adsorption sites one should obtain by cleaving an ordered Cu_3Au bulk to expose a (100) facet. Figure 4 a) shows that the surface of the slabs obtained through GCMC at 300K is dominated by Site 1 and Site 2 for all slabs between 21% and 28% Au composition. For slabs with Au composition greater than 28%, the Site 2 population drops dramatically, being replaced by Au adsorption sites with a higher number of Au first-nearest-neighbors due to the excess Au in the slab's first layer.

For $Cu_3Au(111)$, the number of distinct adsorption sites observed during GCMC calculations was much higher, with 11 sites appearing more than 1% of the time compared to the total number of sites. From the 11 distinct adsorption sites, 4 represent Cu adsorption sites and 7 Au sites with different numbers of Cu and Au first-nearest-neighbors. For $Cu_3Au(111)$, only the adsorption site marked as Site 1 has the same characteristics of adsorption sites that can be obtained from (111) facets derived by the as-cleaved Cu_3Au bulk. Differently from what was observed for $Cu_3Au(100)$, the distribution of sites for the (111) facet is smoother, and several atomic ensembles could appear and contribute to the reactivity of the surface, depending on the slab composition.

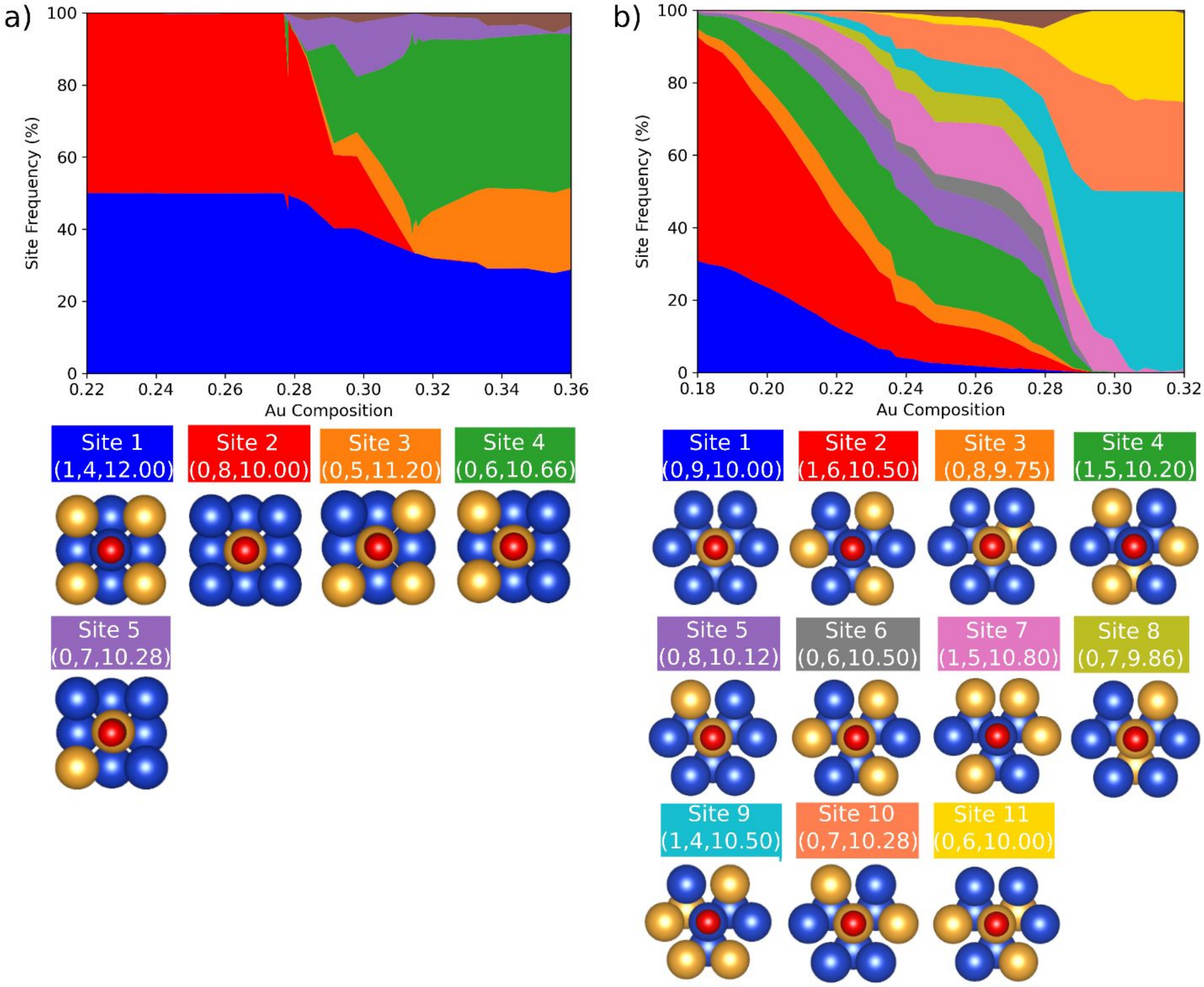


Figure 4: Frequency that each adsorption site is observed during the GCMC calculations versus the Au slab composition with the systems at 300 K, followed by representative structures for each adsorption site together with its representation vector for a) $Cu_3Au(100)$ and b) $Cu_3Au(111)$ surfaces.

Figure 5 a) and Figure 5 b) show types of identified adsorption sites for $CuAu_3(100)$ and (111), respectively. For all cases, the adsorption sites are Au sites, with all first-nearest-neighbors in the first layer being Au atoms and with different number of Cu and Au neighbors in the subsurface. For $CuAu_3(100)$, Site 3 has the same characteristics of the adsorption site present on $Cu_3Au(100)$ facets cleaved from the ordered bulk. Meanwhile, Site 1, which only appears for Cu-rich compositions and rapidly disappears for slab compositions with more than 68% Au, is the adsorption site one can obtain from the cleavage of an ordered CuAu(100) facet. For $CuAu_3(111)$, no site obtained through our approach would be observed by cleaving the bulk because such a process cannot yield a Au surface. For both $CuAu_3(100)$ and (111), the change in adsorption site frequency as a function of slab composition happens because of the increase in Au content in the subsurface. Interestingly, the number of distinct adsorption sites for this composition is smaller than for Cu_3Au due to the full Au surface layer that characterized all the obtained slabs.

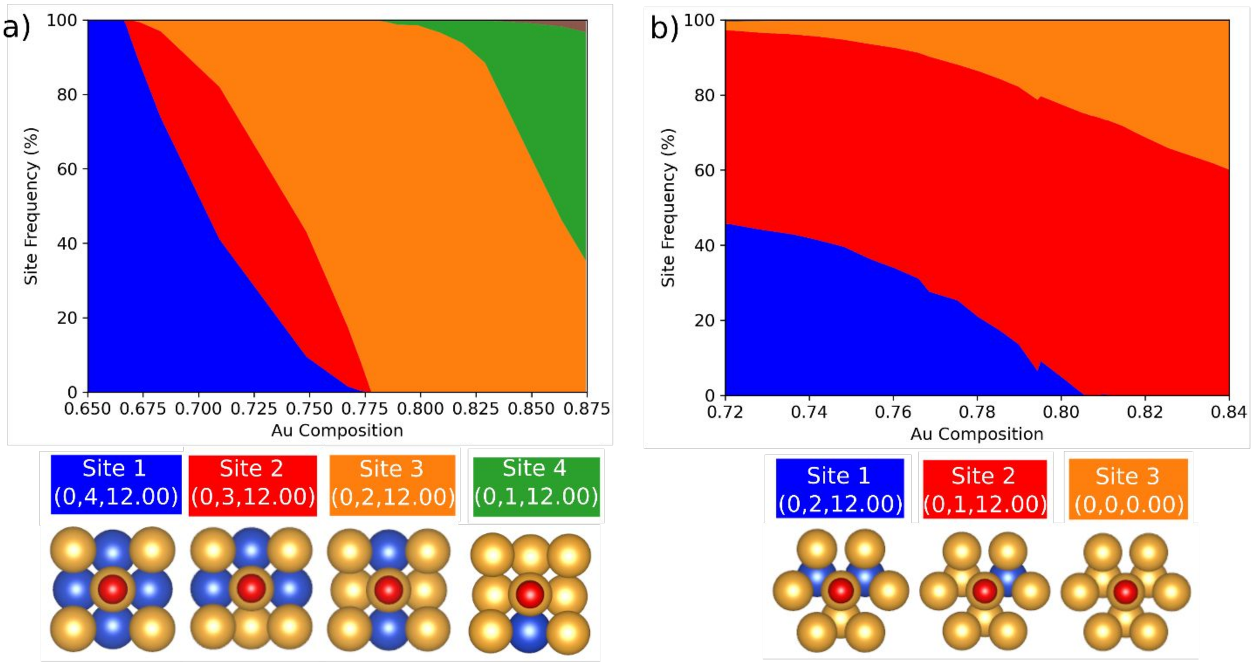


Figure 5: Frequency that each adsorption site is observed during the GCMC calculations versus the Au slab composition with the systems at 300 K, followed by the representation of each adsorption site together with its representation vector for a) $CuAu_3(100)$ and b) $CuAu_3(111)$ surfaces.

To assess the catalytic activities of different site types, we calculated *CO and *COOH adsorption energies on a variety of sites of different types. The complete set of sites and the calculated adsorption energies are provided in Section 6 - ESI†. We observe that most of the sites with the same fingerprints show similar adsorption energies for both *CO and *COOH. However, for some adsorption sites that have the same fingerprints, we find variation in *CO adsorption energies as large as −0.18 eV, implying that sites of the same type can have different catalytic properties. We find that this problem could be partially addressed by extending our fingerprint vector to include information about the second-nearest neighbors and their coordination numbers, distinguishing sites that otherwise look similar based on their first-nearest neighbors (Section 6 - ESI†). For the site types (based on first-nearest neighbors) that had the best onset potentials we performed additional sampling to find a variety of sites of each type that were distinguished by their second-nearest neighbors. We also calculated *CO and *COOH adsorption energies on these sites, to better quantify the spread in adsorption energies on each site type and increase the likelihood of finding a highly active ensemble. These calculations reveal that in nearly all cases the sites that have the same first-nearest-neighbor feature vector had adsorption energies within 0.1 eV of each other, with the exception of site 2 for $Cu_3Au(111)$ surfaces, for which the spreads in *CO and *COOH adsorption energies were 0.18 and 0.23 eV, respectively.

3.3. CO_2RR to CO Catalytic Behavior for Different Adsorption Sites

To investigate the impact of the atomic order in the catalytic behavior of the near-surface Cu-Au alloys, for each facet and bulk composition we calculated the *CO and *COOH adsorption energies on representative adsorption sites identified based on the three-value fingerprints defined as described above. Adsorption sites distinguished by the second-nearest neighbors were not included in this analysis. Figure 6 a) and b) show free energy reaction diagrams for the CO_2 reduction to CO on (100) and (111) facets for Cu, Au, Cu_3Au, and $CuAu_3$ slabs, while Table 1 summarizes the main findings of the free energy reaction diagrams. For Cu and Au, we show the diagram using the adsorption configuration with the strongest ΔE for each adsorbate. For the alloys, we show the reaction diagram using the best and worst adsorption sites from all the calculated sites in terms of onset potential, U_{onset}.

Figure 6 a) shows that the most endergonic reaction step for the alloys and Au(100) surface is the formation of *COOH from CO_2, while for Cu(100) is the *CO desorption to $CO^{\ddagger}$. However, the *CO desorption step is not directly affected by the applied potential within our model, making *COOH formation the potential determining step, PDS, for all surfaces. Figure 6 b) shows the same effect for (111) facets. The study of the best and worst adsorption sites for each alloy surface also helps to illustrate the importance of modeling the alloy through the sampling of different sites compared to using a slab derived from an ordered bulk for each case. For both surfaces, we observe that the best site for both alloys show U_{onset} values higher than Cu and smaller, or at least comparable, to Au, showing a trend where the higher the Au content, the higher the U_{onset}.

More specifically, the U_{onset} for Cu_3Au and $CuAu_3$ for the best sites on (100) facets are, respectively, −0.59 and −0.63 eV, showing a slight improvement as compared to the −0.67 eV computed for Au(100) and a more endergonic reaction step as compared to the –0.32 eV from Cu(100). Moreover, the *CO desorption from the best sites of both alloy surfaces demands less energy as compared to Cu(100) and Au(100) surfaces. Meanwhile, for the worst sites from both surfaces, the *CO desorption is still facile compared to monometallic surfaces; however, the U_{onset} for Cu_3Au and $CuAu_3$ increases to −0.94 eV and −0.88 eV, respectively.

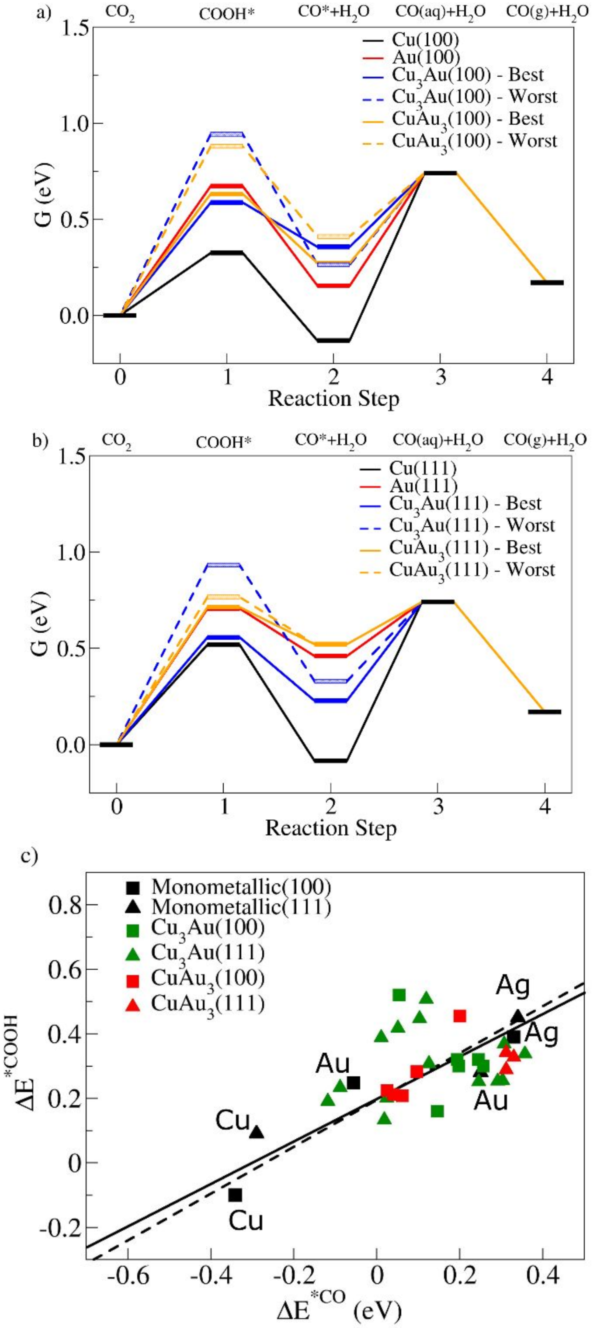


Figure 6: Reaction free energy diagrams for CO_2 reduction to CO on Cu, Au, Cu_3Au, and $CuAu_3$ slabs for a) (100) and b) (111) facets. For Cu_3Au and $CuAu_3$ slabs, we show the reaction diagram using the adsorption sites that yielded the lowest (best) and highest (worst) U_{onset}. c) ΔE values for *CO and *COOH obtained for all the adsorption sites used to represent $Cu_3Au(100)$, $Cu_3Au(111)$, $CuAu_3(100)$, and $CuAu_3(111)$ surfaces, as well as the values calculated for Cu(100), Cu(111), Au(100), Au(111), Ag(100), Ag(111). The linear scaling relationships are drawn from the monometallic calculations for (111) and (100) surfaces together with the values from Pt(100), and Pt(111) and are displayed through solid and dashed lines for the (111) and (100) systems, respectively. Squares and triangles represent ΔE values obtained for (100) and (111) facets, respectively.

Table 1: Results from reaction free energy diagrams, including the site related to the best and worst U_{onset} calculated for CO_2RR to CO, the U_{onset} value, and a note indicating if the site would be observed on bulk-derived slabs or not. *Site 12 for $Cu_3Au(111)$ is not represented in Figure 4 because its prevalence on GCMC was smaller than 1%.

System	Site	U_{onset}(eV)	Observable from bulk-derived slab
Cu(100)	-	−0.32	-
Au(100)	-	−0.67	-
Cu_3Au(100) - Best	Site 2 - (0,8,10.00)	−0.59	Yes
Cu_3Au(100) - Worst	Site 1 - (1,4,12.00)	−0.94	Yes
$CuAu_3$(100) - Best	Site 3 - (0,2,12.00)	−0.63	Yes
$CuAu_3$(100) - Worst	Site 1 - (0,4,12.00)	−0.88	No
Cu(111)	-	−0.52	-
Au(111)	-	−0.71	-
Cu_3Au(111) - Best	Site 3 - (0,8,9.75)	−0.56	No
Cu_3Au(111) - Worst	Site 12* - (1,4,11.25)	−0.93	No
$CuAu_3$(111) - Best	Site 2 - (0,1,12.00)	−0.71	No
$CuAu_3$(111) - Worst	Site 3 - (0,0,0.00)	−0.76	No

For Cu_3Au(100), the best and worst sites are the two sites that one would observe by cleaving an ordered Cu_3Au bulk material to expose the (100) facet. For $CuAu_3$, the best site represents the site one would obtain from bulk-derived $CuAu_3$(100), while the worst site is the one that represents the (100) facet cleaved from an ordered CuAu system, which only appears in Cu-rich stoichiometries in our GCMC calculations. Thus, for (100) surfaces, a naïve approach of using as models the surfaces obtained directly from ordered bulk structures can be sufficient to represent the best and worst scenarios one would obtain from our strategy.

For (111) facets, the best sites of Cu_3Au and $CuAu_3$ show U_{onset} values of −0.56 eV and −0.71 eV that are, respectively, lower and comparable to the −0.71 eV computed for Au(111) and, again, more endergonic as compared to the –0.52 eV from Cu(111). The energy change of the *CO desorption step for the best site of Cu_3Au(111) is also smaller than what is observed for Cu(111), while the best site for $CuAu_3$(111) shows smaller free energy changes for *CO desorption than both Cu(111) and Au(111). For the worst sites, we also observe a significant increase in U_{onset} for Cu_3Au(111) to −0.93 eV, while the difference between the best and worst sites for $CuAu_3$ was only 0.05 eV.

The best and worst sites from Cu_3Au(111) are different from the sites that are present in bulk-derived Cu_3Au(111) surfaces. The best results are obtained with Site 3 from Figure 3b), which shows a higher prevalence for slabs with Au composition lower than 30%. For $CuAu_3$(111), both sites would not be present in bulk-derived surfaces; however, the energy spread between the best and worst calculations is small, and both results are also similar to the ones from the Au(111) surface. Differently to the (100) surfaces, the best and worst adsorption sites from our strategy would not be present in slabs obtained directly from ordered bulk structures, showing that the usage of surface models obtained by cleaving ordered bulk structures could hamper the description of the activity of these surfaces.

The reaction diagrams also help to illustrate one of the limiting factors that one faces when searching for improved catalysts, known as the linear scaling relationship, LSR. The material must strengthen the *COOH adsorption to facilitate the first reaction step without strengthening the *CO adsorption, which would hinder product desorption. Figure 6 c) shows the LSR between *CO and *COOH adsorption energies drawn from calculations on monometallic catalysts. Adsorption sites that deviate from the LSR, staying below the lines in Figure 6 c), would strengthen the *COOH formation as expected from their *CO adsorption energies. All the values obtained from the distinct alloy adsorption sites are near or between Cu and Au values, with no deviation from the LSR larger than −0.14 eV. Figure 6 c) also shows that the approach adopted here can provide a detailed view of the distinct adsorption sites that might affect the activity of alloy catalysts. The similarity in adsorption energies on different ensembles and the difficulty in deviating from the LSR for the different adsorption sites may result from the comparable binding properties of Cu and Au, implying that the same approach, when applied to other systems, could lead to a larger spread in adsorption energies as a function of atomic ordering.

3.4. Exploring the Impact of Near-surface Atomic Ordering on Adsorption Energies Through Cluster Expansions

As previously discussed, our approach to classifying adsorption sites based on the atom near the adsorbate and its first neighbors is adequate for a first-order classification but may have large uncertainties in predicted adsorption energies. To yield insights on how to construct new descriptors and predict adsorption energies using a more complete set of interactions, we trained two cluster expansions to predict *CO and *COOH adsorption on top sites and explored which atoms, in addition to those already included in our approach, could contribute the most to the calculated adsorption energies. We performed this analysis only for $Cu_3Au(111)$ surfaces, as this facet and underlying bulk composition has the largest number of distinct sites. Moreover, $Cu_3Au(111)$ surfaces presented the largest deviation for *CO and *COOH adsorption energies for sites with the same fingerprints based on first-nearest neighbors and distinct second-nearest neighbors.

Figure *7* shows the parity plot of the leave-one-out cross-validation ΔE values from the CE and the same ΔE values calculated through DFT. The plots for Figure *7* a) *CO and b) *COOH, together with the values displayed in Table 2, show that the inclusion of trimer (three-body) clusters is essential to improving the prediction of adsorption energies. When pair and trimers are included with r_{cut} of 4.8 Å, which ensures that only the atom nearest to the adsorbate and its first neighbors are included in our CE, we observe predictions with MAE of 0.06 and 0.05 eV for *CO and *COOH, respectively. By increasing the cutoff of pair clusters to 7 Å, we slightly improve the description for *COOH, lowering by another 0.03 eV the maximum error in the prediction. The MAE values found in our approach are comparable with values from models used in the literature to explore high-entropy alloys for both the CO_2 reduction and oxygen reduction reactions.(*83, 84*)

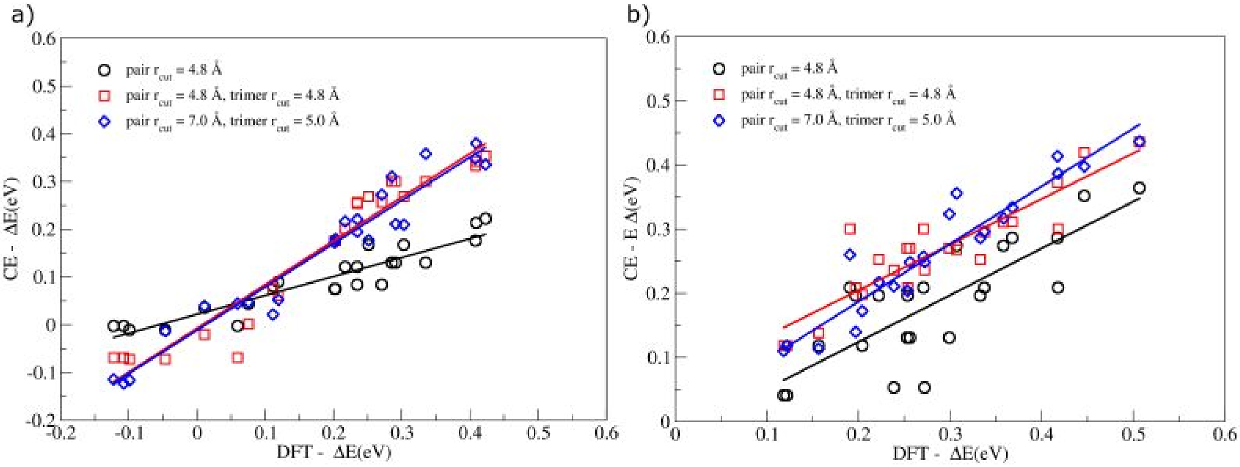


Figure 7: Parity plot of adsorption energies obtained by CE (LOOCV values) and DFT for a) *CO and b) *COOH adsorption on $Cu_3Au(111)$. The results from CE were obtained considering only pair or pair and trimer clusters with different values of cutoff radius, r_{cut}, for the clusters included in the expansion.

Table 2: Parameters obtained for the parity plots between adsorption energies obtained by the CE (LOOCV values) and DFT for *CO and *COOH adsorption on $Cu_3Au(111)$ surfaces. We show the equations slope, offset, MAE, RMSE, R^2 and maximum absolute error obtained with only pair or pair and trimer clusters with different values of the cutoff radius, r_{cut}, used to determine the clusters included in the expansion.

	Slope	Offset	MAE (eV)	RMSE (eV)	R^2	Max Error (eV)
*CO Adsorption						
Pair r_{cut}=4.8 Å, no trimers	0.33	0.03	0.13	0.14	0.87	0.24
Pair r_{cut}=4.8 Å, Trimer r_{cut}=4.8 Å	0.84	0.00	0.06	0.07	0.84	0.13
Pair r_{cut}=7.0 Å, Trimer r_{cut}=5.0 Å	0.79	0.00	0.07	0.08	0.79	0.16
*COOH Adsorption						
Pair r_{cut}=4.8 Å, no trimers	0.70	−0.02	0.11	0.12	0.58	0.24
Pair r_{cut}=4.8 Å, Trimer r_{cut}=4.8 Å	0.58	0.09	0.05	0.07	0.53	0.16
Pair r_{cut}=7.0 Å, Trimer r_{cut}=5.0 Å	0.85	0.02	0.05	0.06	0.74	0.13

The results corroborate our analysis that a vector including the site near the adsorbate and its first nearest neighbors is a good tool to distinguish sites, but the importance of three-body clusters suggests that it would be best to also consider interactions among the first nearest neighbors to distinguish types of sites based on adsorption energies. If further accuracy is necessary for the prediction of adsorption energies, one would need to include more information, as evidenced by the CE prediction of COOH with higher r_{cut} for pair clusters.

The importance of considering a wider variety of interactions when predicting adsorption energies is also supported by recent findings in the literature. For instance, comparison between coordination numbers and generalized coordination numbers demonstrate the importance of considering second-nearest-neighbors when predicting adsorbate adsorption energies.(*85*) A similar conclusion for monometallic and alloy systems has been reached based on other physics-based models derived using machine learning.(*86*) Including more information to distinguish the adsorption sites, however, leads to a higher number of representative sites, which would rapidly increase the computational cost of a thorough investigation of all sites via DFT. In these situations, machine-learned surrogate models such as the CE could facilitate the search for active ensembles.

In the CE approach, the predicted adsorption energies are determined only by the clusters that contain the adsorbate binding site. We use this characteristic of CE models to gain insight into which simple descriptors are most important for predicting adsorption energies. We analyze the CE models trained with r_{cut} of 7 Å and 4.8 Å for pair and trimer clusters, respectively. The cluster that makes the largest contribution to the adsorption energies in these CEs, apart from the clusters that contain the atom near the adsorbate and its first nearest neighbors, includes atoms in the surface that are the next in line to the first neighbors of the atom near the adsorbate, see Figure 8. Thus, the number of Au or Cu atoms at this location relative to the adsorbate could be used as an additional fingerprint in future studies, in addition to including interactions between first-nearest neighbors.

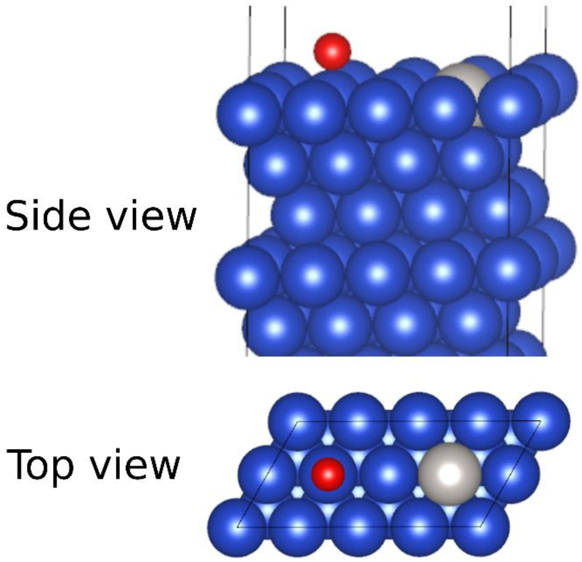


Figure 8: Representation of pair cluster with the largest contribution to the *CO and *COOH adsorption energies as predicted by the cluster expansion, apart from the clusters that contain the atom near the adsorbate and its first nearest neighbors. The red sphere represents the adsorbate binding site, while gray (blue) spheres represent atoms that are (are not) present in the cluster.

Our exploration of different adsorption sites formed in Cu_xAu_y near-surface alloy catalysts, was not able to find adsorption site ensembles that significantly deviate from the linear scaling relationship (LSR) between the adsorption energies of *COOH and *CO that one finds when studying monometallic catalysts. So, we employed our CEs fit to predict the *CO and *COOH adsorption energies in simulated annealing calculations designed to minimize the following quantity:

$$E_{target} = \Delta E^{*COOH} - \Delta E^{*COOH}_{LSR} = \Delta E^{*COOH} - (0.657\Delta E^{*CO} + 0.196) \tag{7}$$

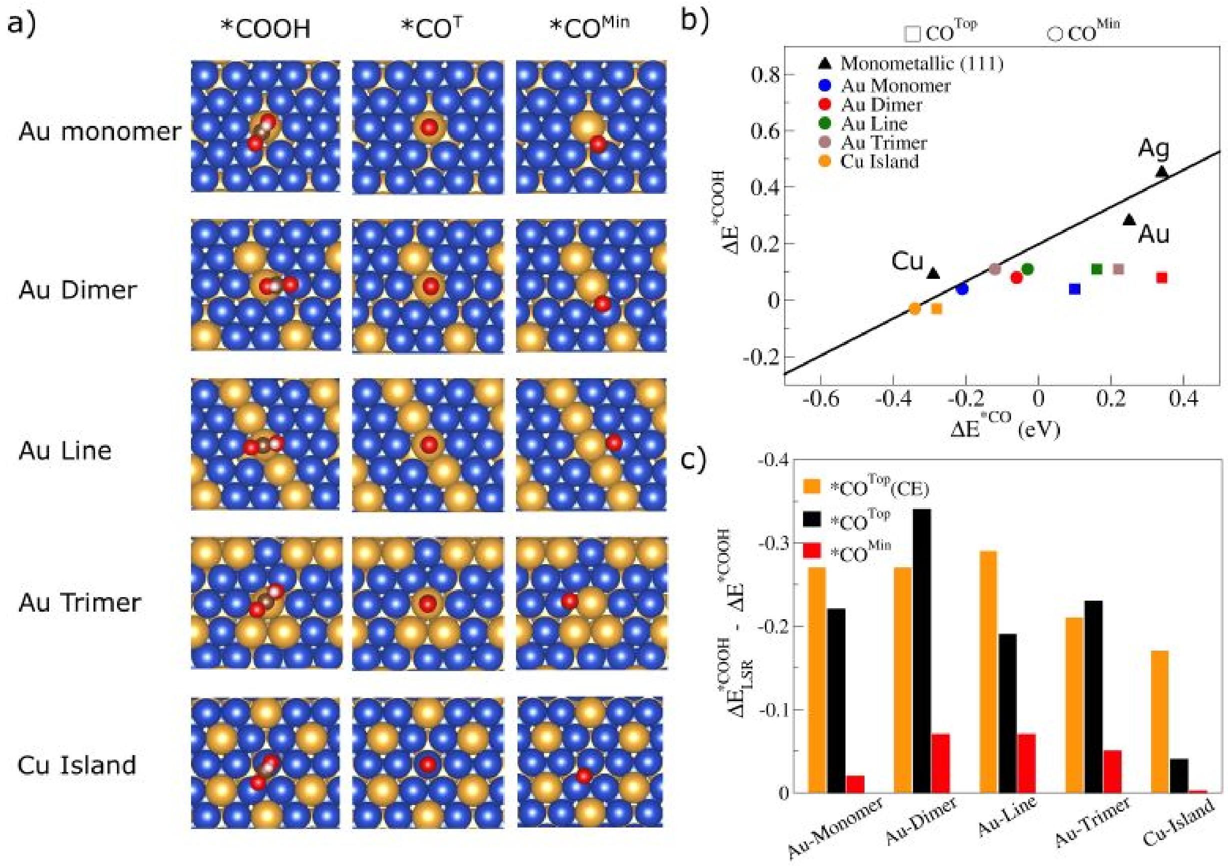


Figure 9: DFT calculations for *CO and *COOH on different adsorption site ensembles found through a simulated annealing approach designed to maximize the deviation of the linear scaling relationship, LSR. a) Representations of the lowest energy configuration for *COOH, the configuration with *CO on the top site suggested via the simulated annealing calculations, CO^{Top}, and the configuration with the lowest energy for *CO considering the same top site and all its neighboring bridge sites, CO^{Min}. b) ΔE values for *CO and *COOH obtained for all the ensembles, where squares and circles represent the ΔE values for CO^{Top} and CO^{Min}, respectively. We also include the values calculated for Cu(111), Au(111), and Ag(111) and the linear scaling relationships drawn as a solid line based on these monometallic surfaces together with the values from Pt(111). c) Deviations from the LSR, ΔE^{*COOH}-ΔE^{*COOH}_{LSR}, for each ensemble. Black (red) bars represent the data when CO^{Top} (CO^{Min}) were used as input for the calculation of ΔE^{*COOH}_{LSR}, while yellow bars represent the values calculated via CE.

During the simulated annealing calculations, we kept the atom nearest to the adsorbate fixed as either Au or Cu to predict different atom arrangements for both cases. For calculations with Au as the nearest atom to the adsorbate, we observe four different ensembles that are predicted by the cluster expansion to deviate from the LSR by more than –0.2 eV. When the adsorption site atop was a Cu atom, the simulated annealing calculations predict that only one ensemble would significantly deviate from LSR. The predicted ensembles are illustrated in Figure 9 a). We also searched for ensembles that would deviate from the LSR in the other direction, which would hinder the reaction. For this case, we observe Cu single atoms on surfaces mainly composed of Au as possible ensembles, but as these cases are not predicted to improve the catalytic activity for this reaction, we decided not to further investigate the predictions with DFT.

The sites predicted to break the scaling relationship in a favorable direction were tested using *CO and *COOH adsorption calculations via DFT. The cluster expansions used for this strategy were fitted against DFT data for *CO and *COOH adsorbed on top sites. However, we followed the same approach as discussed for the whole work and tested with DFT the *CO adsorption on the top site and on all neighboring bridge sites.

Figure 9 b) shows that when considering the results when both *CO and *COOH adsorbed on Au top sites, all the isolated Au ensembles show interesting features. For these ensembles, we observe stronger ΔE for *COOH than Au(111) without leading to *CO ΔE values that hinder *CO desorption as much as Cu(111). This effect is most significant for the Au dimer. For all the ensembles found in the simulated annealing with Au as the nearest atom to the adsorbate, a significant deviation of the LSR only happens when considering the results from $^*CO^{Top}$, Figure 9 b) and c)). These results indicate that if the *CO molecule can move to a neighboring bridge site during the reaction, the benefit of these ensembles, as compared to a Cu(111) catalyst, would not be attainable. However, if other effects, such as a high *CO coverage, hinder the *CO migration to a neighboring bridge site and both reaction steps happen on the same site, one could expect that such isolated Au clusters on Cu(111) would favor active and selective catalysts for CO production. The CE-predicted deviations from the LSR are generally in good agreement with the DFT-calculated values on the top sites, with a notable exception on the pure Cu site (Figure 9 c)). For the Cu island, despite the prediction using the CEs, the results with DFT do not show any significant deviation from the LSR.

4. Conclusions

Generating models to treat bimetallic catalysts without prior experimental information can be challenging since surfaces cleaved from ordered bulk can miss atomic ensembles important for the overall catalytic activity of the system. Here, we explored this topic using Cu_3Au and $CuAu_3(111)$ and (100) surfaces applied to the CO_2RR to CO. We employed grand canonical Monte Carlo (GCMC) calculations together with a cluster expansion approach trained on DFT calculations to monitor the presence and frequency of distinct adsorption sites on stoichiometric and non-stoichiometric slabs.

The GCMC calculations showed the presence of several types of adsorption sites with a prevalence higher than 1% as compared to the total number of sites, as characterized using a simple three-value descriptor based on first-nearest neighbors. Comparing the adsorption sites for each system in terms of U_{onset} values, we observed differences of 0.35, 0.25, 0.37, 0.05 V between the highest and lowest U_{onset} from the calculated sites for $Cu_3Au(100)$, $CuAu_3(100)$, $Cu_3Au(111)$, and $CuAu_3(111)$, respectively. The best and worst sites determined based on U_{onset} values for $Cu_3Au(100)$ and the best site for $CuAu_3(100)$ are present on surfaces that could have been obtained by cleaving ordered bulk alloys, but this does not necessarily mean that such surfaces should generally be expected to be stable and / or highly active.

For $CuAu_3(111)$, the calculated sites are not present for bulk-derived surfaces, but the difference between the calculated U_{onset} values for different sites is small and similar to the ones from Au(111), showing that this system could potentially be modeled as a Au overlayer following similar strategies used in the literature.(*32*) The most significant differences were observed for $Cu_3Au(111)$, in which bulk-derived slabs would not contain the sites responsible for the highest and lowest U_{onset} values observed for the system, showing that more simple approaches for the generation of models could miss important adsorption sites. None of the adsorption sites selected through this approach significantly deviate from LSR between adsorption energies of *CO and *COOH, which hinders the search for more active catalysts for this reaction.

Cluster expansions designed to predict *CO and *COOH adsorption energies show the limits of the simple strategy to distinguish adsorption sites, providing information on many-body and longer-range interactions that could be included in future attempts to classify sites based on adsorption energies. The cluster expansions trained on top-site adsorption energies were used to search for adsorption site ensembles that deviate from the LSR between *CO and *COOH. Their predictions were confirmed, to an extent, with DFT calculations showing ensembles that deviate by up to -0.33 eV from the LSR when both *CO and *COOH are assumed to adsorb on the same top site. However, in these cases it was found that the energy of *CO could be significantly lowered, offsetting the deviation from the scaling relationship, if it were able to migrate to a nearby vacant bridge site.

5. Conflicts of interest

There are no conflicts of interest to declare.

6. Acknowledgements

The authors gratefully acknowledge support from the São Paulo Research Foundation (FAPESP, grant numbers 2017/11631-2, 2018/21401-7, 2019/05561-7, and 2021/07129-5), Shell and the strategic importance of the support given by ANP (Brazil's National Oil, Natural Gas and Biofuels Agency) through the R&D levy regulation. The authors also acknowledge the Laboratory of Advanced Scientific Computing (University of São Paulo), and the Department of Information Technology - Campus São Carlos for hosting our computer cluster. This work is supported by the National Science Foundation (CBET-1930013). The computational resources are provided by the Maryland Advanced Research Computing Center (MARCC) and the Extreme Science and Engineering Discovery Environment (XSEDE) through NSF award DMR-140068.

7. References

1. T. P. Senftle, E. A. Carter, Holy Grail: Chemistry Enabling an Economically Viable CO_2 Capture, Utilization, and Storage Strategy. *Accounts of Chemical Research* **50**, 472-475 (2017).

2. Q. Lu, F. Jiao, Electrochemical CO_2 reduction: Electrocatalyst, reaction mechanism, and process engineering. *Nano Energy* **29**, 439-456 (2016).
3. Y. Hori, K. Kikuchi, S. Suzuki, Production of CO and CH_4 in Electrochemical Reduction of CO_2 at Metal Electrodes in Aqueous Hydrogencarbonate Solution. *Chemistry Letters* **14**, 1695-1698 (1985).
4. Y. Hori, A. Murata, R. Takahashi, Formation of hydrocarbons in the electrochemical reduction of carbon dioxide at a copper electrode in aqueous solution. *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* **85**, 2309-2326 (1989).
5. Y. Hori, H. Wakebe, T. Tsukamoto, O. Koga, Electrocatalytic process of CO selectivity in electrochemical reduction of CO2 at metal electrodes in aqueous media. *Electrochimica Acta* **39**, 1833-1839 (1994).
6. K. P. Kuhl, E. R. Cave, D. N. Abram, T. F. Jaramillo, New Insights into the Electrochemical Reduction of Carbon Dioxide on Metallic Copper Surfaces. *Energy & Environmental Science* **5**, 7050-7059 (2012).
7. A. A. Peterson, F. Abild-Pedersen, F. Studt, J. Rossmeisl, J. K. Nørskov, How Copper Catalyzes the Electroreduction of Carbon Dioxide into Hydrocarbon Fuels. *Energy & Environmental Science* **3**, 1311-1315 (2010).
8. M. T. Tang, H. Peng, P. S. Lamoureux, M. Bajdich, F. Abild-Pedersen, From Electricity to Fuels: Descriptors for C1 Selectivity in Electrochemical CO_2 Reduction. *Applied Catalysis B: Environmental* **279**, 119384 (2020).
9. W. Luo, X. Nie, M. J. Janik, A. Asthagiri, Facet Dependence of CO_2 Reduction Paths on Cu Electrodes. *ACS Catalysis* **6**, 219-229 (2016).
10. F. Calle-Vallejo, M. T. M. Koper, Theoretical Considerations on the Electroreduction of CO to C2 Species on Cu(100) Electrodes. *Angewandte Chemie International Edition* **52**, 7282-7285 (2013).
11. D. Raciti *et al.*, Low-Overpotential Electroreduction of Carbon Monoxide Using Copper Nanowires. *ACS Catal.* **7**, 4467-4472 (2017).
12. L. Cao *et al.*, Mechanistic Insights for Low-Overpotential Electroreduction of CO2 to CO on Copper Nanowires. *ACS Catalysis* **7**, 8578-8587 (2017).
13. K. J. P. Schouten, E. Pérez Gallent, M. T. M. Koper, Structure Sensitivity of the Electrochemical Reduction of Carbon Monoxide on Copper Single Crystals. *ACS Catalysis* **3**, 1292-1295 (2013).
14. L. G. Verga, P. C. D. Mendes, V. K. Ocampo-Restrepo, J. L. F. Da Silva, The Role of Site Coordination on the CO_2 Electroreduction Pathway on Stepped and Defective Copper Surfaces. *Catalysis Science & Technology* **11**, 2770-2781 (2021).
15. L. G. Verga, P. C. D. Mendes, V. K. Ocampo-Restrepo, J. L. F. Da Silva, Exploring the Adsorption Site Coordination as a Strategy to Tune Copper Catalysts for CO_2 Electro-reduction. *Catalysis Science & Technology* **12**, 869-879 (2022).
16. A. Bagger, W. Ju, A. S. Varela, P. Strasser, J. Rossmeisl, Electrochemical CO_2 Reduction: Classifying Cu Facets. *ACS Catalysis* **9**, 7894-7899 (2019).
17. R. Reske, H. Mistry, F. Behafarid, B. Roldan Cuenya, P. Strasser, Particle Size Effects in the Catalytic Electroreduction of CO_2 on Cu Nanoparticles. *Journal of the American Chemical Society* **136**, 6978-6986 (2014).
18. E. Pérez-Gallent, G. Marcandalli, M. C. Figueiredo, F. Calle-Vallejo, M. T. M. Koper, Structure- and Potential-Dependent Cation Effects on CO Reduction at Copper Single-Crystal Electrodes. *Journal of the American Chemical Society* **139**, 16412-16419 (2017).
19. Y. Huang, A. D. Handoko, P. Hirunsit, B. S. Yeo, Electrochemical Reduction of CO_2 Using Copper Single-Crystal Surfaces: Effects of CO* Coverage on the Selective Formation of Ethylene. *ACS Catalysis* **7**, 1749-1756 (2017).

20. R. B. Sandberg, J. H. Montoya, K. Chan, J. K. Nørskov, CO-CO coupling on Cu facets: Coverage, strain and field effects. *Surface Science* **654**, 56-62 (2016).
21. L. D. Chen, M. Urushihara, K. Chan, J. K. Nørskov, Electric Field Effects in Electrochemical CO_2 Reduction. *ACS Catalysis* **6**, 7133-7139 (2016).
22. G. Iijima, T. Inomata, H. Yamaguchi, M. Ito, H. Masuda, Role of a Hydroxide Layer on Cu Electrodes in Electrochemical CO_2 Reduction. *ACS Catalysis* **9**, 6305-6319 (2019).
23. M. Fields, X. Hong, J. K. Nørskov, K. Chan, Role of Subsurface Oxygen on Cu Surfaces for CO_2 Electrochemical Reduction. *The Journal of Physical Chemistry C* **122**, 16209-16215 (2018).
24. S. G. Jadhav, P. D. Vaidya, B. M. Bhanage, J. B. Joshi, Catalytic Carbon Dioxide Hydrogenation to Methanol: A Review of Recent Studies. *Chemical Engineering Research and Design* **92**, 2557-2567 (2014).
25. A. Bagger, W. Ju, A. S. Varela, P. Strasser, J. Rossmeisl, Electrochemical CO_2 Reduction: A Classification Problem. *ChemPhysChem* **18**, 3266-3273 (2017).
26. W. Zhu *et al.*, Monodisperse Au Nanoparticles for Selective Electrocatalytic Reduction of CO_2 to CO. *Journal of the American Chemical Society* **135**, 16833-16836 (2013).
27. W. Zhu *et al.*, Active and Selective Conversion of CO_2 to CO on Ultrathin Au Nanowires. *Journal of the American Chemical Society* **136**, 16132-16135 (2014).
28. S. Back, M. S. Yeom, Y. Jung, Understanding the Effects of Au Morphology on CO_2 Electrocatalysis. *The Journal of Physical Chemistry C* **122**, 4274-4280 (2018).
29. Y. Wang *et al.*, Undercoordinated Active Sites on 4H Gold Nanostructures for CO2 Reduction. *Nano Letters* **20**, 8074-8080 (2020).
30. P. Strasser *et al.*, Lattice-strain Control of the Activity in Dealloyed Core–shell Fuel Cell Catalysts. *Nature Chemistry* **2**, 454-460 (2010).
31. R. Reske *et al.*, Controlling Catalytic Selectivities during CO_2 Electroreduction on Thin Cu Metal Overlayers. *The Journal of Physical Chemistry Letters* **4**, 2410-2413 (2013).
32. D. Kim *et al.*, Electrochemical Activation of CO_2 through Atomic Ordering Transformations of AuCu Nanoparticles. *Journal of the American Chemical Society* **139**, 8329-8336 (2017).
33. J. R. Kitchin, J. K. Nørskov, M. A. Barteau, J. G. Chen, Role of Strain and Ligand Effects in the Modification of the Electronic and Chemical Properties of Bimetallic Surfaces. *Physical Review Letters* **93**, 156801 (2004).
34. H. Li, K. Shin, G. Henkelman, Effects of Ensembles, Ligand, and Strain on Adsorbate Binding to Alloy Surfaces. *The Journal of Chemical Physics* **149**, 174705 (2018).
35. M. Karamad, V. Tripkovic, J. Rossmeisl, Intermetallic Alloys as CO Electroreduction Catalysts—Role of Isolated Active Sites. *ACS Catalysis* **4**, 2268-2273 (2014).
36. Y. Wang *et al.*, Ensemble Effect in Bimetallic Electrocatalysts for CO2 Reduction. *Journal of the American Chemical Society* **141**, 16635-16642 (2019).
37. Z. Yin *et al.*, Highly Selective Palladium-Copper Bimetallic Electrocatalysts for the Electrochemical Reduction of CO_2 to CO. *Nano Energy* **27**, 35-43 (2016).
38. J. Wang *et al.*, Silver/Copper Interface for Relay Electroreduction of Carbon Dioxide to Ethylene. *ACS Applied Materials & Interfaces* **11**, 2763-2767 (2019).
39. H. A. Hansen, C. Shi, A. C. Lausche, A. A. Peterson, J. K. Nørskov, Bifunctional Alloys for the Electroreduction of CO_2 and CO. *Physical Chemistry Chemical Physics* **18**, 9194-9201 (2016).
40. Z. P. Jovanov *et al.*, Opportunities and Challenges in the Electrocatalysis of CO_2 and CO Reduction Using Bifunctional Surfaces: A Theoretical and Experimental Study of Au–Cd Alloys. *Journal of Catalysis* **343**, 215-231 (2016).
41. D. Kim, J. Resasco, Y. Yu, A. M. Asiri, P. Yang, Synergistic Geometric and Electronic Effects for Electrochemical Reduction of Carbon Dioxide Using Gold–Copper Bimetallic Nanoparticles. *Nature Communications* **5**, 4948 (2014).

42. J. Christophe, T. Doneux, C. Buess-Herman, Electroreduction of Carbon Dioxide on Copper-Based Electrodes: Activity of Copper Single Crystals and Copper–Gold Alloys. *Electrocatalysis* **3**, 139-146 (2012).
43. E. Andrews, Y. Fang, J. Flake, Electrochemical Reduction of CO_2 at CuAu Nanoparticles: Size and Alloy Effects. *Journal of Applied Electrochemistry* **48**, 435-441 (2018).
44. W. Zhu *et al.*, Low-overpotential Selective Reduction of CO_2 to Ethanol on Electrodeposited CuxAuy Nanowire Arrays. *Journal of Energy Chemistry* **37**, 176-182 (2019).
45. X. Wang *et al.*, Gold-in-copper at Low *CO Coverage Enables Efficient Electromethanation of CO_2. *Nature Communications* **12**, 3387 (2021).
46. L. Cao, T. Mueller, Rational Design of Pt_3Ni Surface Structures for the Oxygen Reduction Reaction. *The Journal of Physical Chemistry C* **119**, 17735-17747 (2015).
47. G. Kresse, J. Hafner, Ab initio Molecular Dynamics for Open-shell Transition Metals. *Physical Review B* **48**, 13115-13118 (1993).
48. G. Kresse, J. Furthmüller, Efficient Iterative Schemes for Ab initio Total-Energy Calculations Using a Plane-wave Basis Set. *Physical Review B* **54**, 11169-11186 (1996).
49. G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B* **59**, 1758-1775 (1999).
50. P. E. Blochl, Projector Augmented-Wave Method. *Physical Review B* **50**, 17953-17979 (1994).
51. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple. *Physical Review Letters* **77**, 3865-3868 (1996).
52. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *The Journal of Chemical Physics* **132**, 154104 (2010).
53. P. Wisesa, K. A. McGill, T. Mueller, Efficient generation of generalized Monkhorst-Pack grids through the use of informatics. *Physical Review B* **93**, 155109 (2016).
54. Y. Wang, P. Wisesa, A. Balasubramanian, S. Dwaraknath, T. Mueller, Rapid Generation of Optimal Generalized Monkhorst-Pack Grids. *Computational Materials Science* **187**, 110100 (2021).
55. J. M. Sanchez, F. Ducastelle, D. Gratias, Generalized Cluster Description of Multicomponent Systems. *Physica A: Statistical Mechanics and its Applications* **128**, 334-350 (1984).
56. J. M. Sanchez, Cluster expansion and the configurational theory of alloys. *Physical Review B* **81**, 224202 (2010).
57. V. Blum, G. L. W. Hart, M. J. Walorski, A. Zunger, Using genetic algorithms to map first-principles results to model Hamiltonians: Application to the generalized Ising model for alloys. *Physical Review B* **72**, 165113 (2005).
58. M. Sanati, G. L. W. Hart, A. Zunger, Ordering tendencies in octahedral MgO-ZnO alloys. *Physical Review B* **68**, 155210 (2003).
59. F. Zhou, T. Maxisch, G. Ceder, Configurational Electronic Entropy and the Phase Diagram of Mixed-Valence Oxides: The Case of ${\mathrm{Li}}_{x}{\mathrm{FePO}}_{4}$. *Physical Review Letters* **97**, 155704 (2006).
60. W. Chen, D. Schmidt, W. F. Schneider, C. Wolverton, Ordering and Oxygen Adsorption in Au–Pt/Pt(111) Surface Alloys. *The Journal of Physical Chemistry C* **115**, 17915-17924 (2011).
61. W. Chen, P. Dalach, W. F. Schneider, C. Wolverton, Interplay between Subsurface Ordering, Surface Segregation, and Adsorption on Pt–Ti(111) Near-Surface Alloys. *Langmuir* **28**, 4683-4693 (2012).
62. L. Cao, T. Mueller, Theoretical Insights into the Effects of Oxidation and Mo-Doping on the Structure and Stability of Pt–Ni Nanoparticles. *Nano Letters* **16**, 7748–7754 (2016).

63. T. Mueller, G. Ceder, Effect of Particle Size on Hydrogen Release from Sodium Alanate Nanoparticles. *ACS Nano* **4**, 5647-5656 (2010).
64. L.-L. Wang, T. L. Tan, D. D. Johnson, Configurational Thermodynamics of Alloyed Nanoparticles with Adsorbates. *Nano Letters* **14**, 7077-7084 (2014).
65. L. Cao, C. Li, T. Mueller, The Use of Cluster Expansions To Predict the Structures and Properties of Surfaces and Nanostructured Materials. *Journal of Chemical Information and Modeling* **58**, 2401-2413 (2018).
66. A. van de Walle, M. Asta, G. Ceder, The alloy theoretic automated toolkit: A user guide. *CALPHAD* **26**, 539-553 (2002).
67. M. Ångqvist *et al.*, ICET – A Python Library for Constructing and Sampling Alloy Cluster Expansions. *Advanced Theory and Simulations* **2**, 1900015 (2019).
68. T. Mueller, G. Ceder, Bayesian Approach to Cluster Expansions. *Physical Review B* **80**, 024103 (2009).
69. T. Mueller, G. Ceder, Exact Expressions for Structure Selection in Cluster Expansions. *Physical Review B* **82**, 184107 (2010).
70. L. Cao, L. Niu, T. Mueller, Computationally generated maps of surface structures and catalytic activities for alloy phase diagrams. *Proceedings of the National Academy of Sciences* **116**, 22044-22051 (2019).
71. J. K. Nørskov *et al.*, Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *The Journal of Physical Chemistry B* **108**, 17886-17892 (2004).
72. J. Weinreich, A. Römer, M. L. Paleico, J. Behler, Properties of α-Brass Nanoparticles. 1. Neural Network Potential Energy Surface. *The Journal of Physical Chemistry C* **124**, 12682-12695 (2020).
73. N. Artrith, A. M. Kolpak, Understanding the Composition and Activity of Electrocatalytic Nanoalloys in Aqueous Solvents: A Combination of DFT and Accurate Neural Network Potentials. *Nano Letters* **14**, 2670-2676 (2014).
74. A. Shapeev, Accurate representation of formation energies of crystalline alloys with many components. *Computational Materials Science* **139**, 26-30 (2017).
75. G. L. W. Hart, T. Mueller, C. Toher, S. Curtarolo, Machine learning for alloys. *Nature Reviews Materials* **6**, 730-755 (2021).
76. R. Tran *et al.*, Surface energies of elemental crystals. *Scientific Data* **3**, 160080 (2016).
77. P. C. D. Mendes *et al.*, Ab Initio Insights into the Formation Mechanisms of 55-Atom Pt-Based Core–Shell Nanoalloys. *The Journal of Physical Chemistry C* **124**, 1158-1164 (2020).
78. O. Bauer *et al.*, Au Enrichment and Vertical Relaxation of the $Cu_3Au(111)$ Surface Studied by Normal-incidence X-ray Standing Waves. *Physical Review B* **93**, 235429 (2016).
79. M. Okada *et al.*, Experimental and Theoretical Studies on Oxidation of Cu-Au Alloy Surfaces: Effect of Bulk Au Concentration. *Scientific Reports* **6**, 31101 (2016).
80. J.-Q. Goh, J. Akola, R. Ferrando, Geometric Structure and Chemical Ordering of Large AuCu Clusters: A Computational Study. *The Journal of Physical Chemistry C* **121**, 10809-10816 (2017).
81. J. Dean, M. J. Cowan, J. Estes, M. Ramadan, G. Mpourmpakis, Rapid Prediction of Bimetallic Mixing Behavior at the Nanoscale. *ACS Nano* **14**, 8171-8180 (2020).
82. Z. W. Ulissi *et al.*, Machine-Learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for CO2 Reduction. *ACS Catalysis* **7**, 6600-6608 (2017).
83. T. A. A. Batchelor *et al.*, High-Entropy Alloys as a Discovery Platform for Electrocatalysis. *Joule* **3**, 834-845 (2019).
84. J. K. Pedersen, T. A. A. Batchelor, A. Bagger, J. Rossmeisl, High-Entropy Alloys as Catalysts for the CO2 and CO Reduction Reactions. *ACS Catalysis* **10**, 2169-2176 (2020).

85. F. Calle-Vallejo, J. I. Martínez, J. M. García-Lastra, P. Sautet, D. Loffreda, Fast Prediction of Adsorption Properties for Platinum Nanocatalysts with Generalized Coordination Numbers. *Angewandte Chemie International Edition* **53**, 8316-8319 (2014).

86. J. Dean, M. G. Taylor, G. Mpourmpakis, Unfolding adsorption on metal nanoparticles: Connecting stability with catalysis. *Science Advances* **5**, eaax5101.