Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Bayesian Optimization of High-Entropy Alloy Compositions for Electrocatalytic Oxygen Reduction**

*Jack K. Pedersen, Christian M. Clausen, Olga A. Krysiak, Bin Xiao, Thomas A. A. Batchelor, Tobias Löffler, Vladislav A. Mints, Lars Banko, Matthias Arenz, Alan Savan, Wolfgang Schuhmann, Alfred Ludwig, and Jan Rossmeisl**

anie_202108116_sm_miscellaneous_information.pdf

Computational Procedures

All simulated structures, data and scripts necessary for reproducing the simulations herein have made freely accessible at https://nano.ku.dk/english/research/theoretical-electrocatalysis/katladb/bayesian-optimization-of-hea/

Density functional theory simulations

Density functional theory using the revised Perdew-Burke-Ernzerhof (RPBE) exchange-correlation functional[1] as implemented in the GPAW code[2,3] was used to obtain *OH and O* adsorption energies on fcc (111) 2x2-atoms-sized, four-layered surface slabs that were periodically repeated in the direction parallel to the slab. The structures were set up and manipulated in the Atomic Simulation Environment (ASE)[4]. The slabs were constructed with an fcc lattice constant set to the weighted average of the calculated fcc lattice constants of the elements in the surface layer, a vacuum of 7.5 Å was added above and below the slab, and the atoms in the two bottom layers were held fixed during geometry relaxations at which the structures were optimized until the maximum force on any atom was below at least 0.1 eVÅ^{-1}. The wave functions were expanded in plane waves with an energy cutoff set to 400 eV, and the Brillouin zone was sampled with a Monkhorst-Pack grid of 4x4x1 k-points. For training the Ag-Ir-Pd-Pt-Ru quinary alloy adsorption energy regressor, a total of 1304 *OH and 1768 O* adsorption energies were simulated on slabs where the metals in the structure were randomly sampled from an equimolar ratio. Equivalently for the Ir-Pd-Pt-Rh-Ru system where 856 *OH and 997 O* adsorption energies were simulated. The 2x2-atoms-sized, four-layered slabs for the binary alloy systems (Ag-Pd, Ir-Pt and Pd-Ru) were calculated with a similar computational setup with the exception of 10 Å added vacuum. In addition to slabs sampled from equimolar ratios, 25% of the slabs were sampled from a 3:1 ratio e.g. $Ag_{75}Pd_{25}$ and 25% from a 1:3 ratio e.g. $Ag_{25}Pd_{75}$ in order to more sensibly span the compositions of the binary alloy systems.

Adsorption energy prediction

The DFT calculated *OH and O* adsorption energies were used to train a regressor for predicting the *OH and O* adsorption energy at any conceivable on-top and fcc hollow site of an fcc (111) surface. To this end, we applied our previously developed scheme[5] for mapping structures into machine readable features simply by one-hot encoding the identity of the adsorption site ensemble and by counting the number of each element in equidistant positions from the adsorption site. For *OH on-top adsorption the atoms included in our description was the on-top adsorbing atom itself, the surface and subsurface neighboring atoms, and the second-nearest atoms in the third layer as we suggested recently[6]. For O* fcc hollow site adsorption the three-atom site ensemble as well as its surface and subsurface neighboring atoms were included in the description of the site (for an example see Figure S3). These features were used to fit a linear model for each on-top *OH adsorption site, i.e. one for on-top Ag, one for on-top Ir, etc. containing 15 fitted parameters each, as well as a single linear model for fcc hollow adsorbing O*, containing 55 fitted parameters (the fitted linear parameters are shown in Table S2-S5). For the investigated binary alloys a gradient boosted regressor was used on a more elaborate description of the surface site including the neighboring atomic environment up to the third or fourth closest atoms of each layer in order to improve on the prediction accuracy. Since the 2x2-atoms-sized DFT simulated slabs are periodically repeated this will include some zones without any additional information, however all available atoms will be included in the site features. In addition, a gradient boosted model was fitted to each O* adsorption site, i.e. one model for Ag_3, Ag_2Pd, $AgPd_2$ and Pd_3, respectively. Furthermore, when training the gradient boosted regressor the samples of binary alloys were weighted to enhance their representation of the composition span. Thus, each data set of a binary alloy contained two pure metal samples with assigned weights of 1000, approximately 1000 samples drawn equally from the 1:3 and 3:1 molar ratios with assigned weights of 2 and approximately 1000 samples drawn from the equimolar ratio with assigned weights of 1. Linear and gradient boosted regression algorithms were used as implemented in scikit-learn[7] with default hyperparameters.

Current density modeling

The current density predicted by simulations was predicted using equations 1-3. Equation 3 predicts the per site current density based on an Arrhenius-like rate expression and takes as input the difference in the *OH or O* adsorption energy to an optimal value, determined as 0.1 eV[8] and 0.2 eV[9] weaker than on Pt(111). The theoretical framework behind these experimentally validated optimal adsorption energies is a kinetic model that involves the associative mechanism for the ORR. Initially, the electrochemical adsorption-reduction step of O_2 to *OOH must be facilitated, and finally the desorption-reduction steps of *OH or O* to H_2O. We note that other pathways, including the dissociative ORR mechanism, could also contribute to the produced current density. However, we expect these contributions to be minor. In case of the dissociation of O_2 to form adsorbed O* on the surface, the argument is that this reaction becomes increasingly less relevant compared to the associative mechanism as the overpotential increases. Already at about 0.8 V vs. RHE the effect of the dissociative mechanism is minor.[10]

In order to improve on the model's predictive trend, a simple adsorbate interaction between adsorbed *OH and O* was included.[11] This interaction works by ensuring that no two neighboring on-top and hollow sites can adsorb reaction intermediates at the same time. To calculate the current density using equation 1-3, the net coverage and corresponding net adsorption energies were used by accounting for this interaction. In practice these *net* adsorption energies were achieved by predicting the *OH and O* *gross* adsorption energies on all on-top and fcc hollow sites on a randomly constructed surface with a desired composition and measuring 100x100 atoms (the dependence of the predicted current density on the surface size is shown in Figure S1). The surface was then filled with *OH on-top and O* fcc hollow adsorbates starting at the strongest adsorbing sites and filled using the rule that no neighboring on-top and hollow sites can adsorb at the same time, until no more free surface sites remained. The net coverage and net adsorption energies of *OH and O* achieved in this way would then act as the input for equation 1-3 when calculating the predicted current density for the given composition.

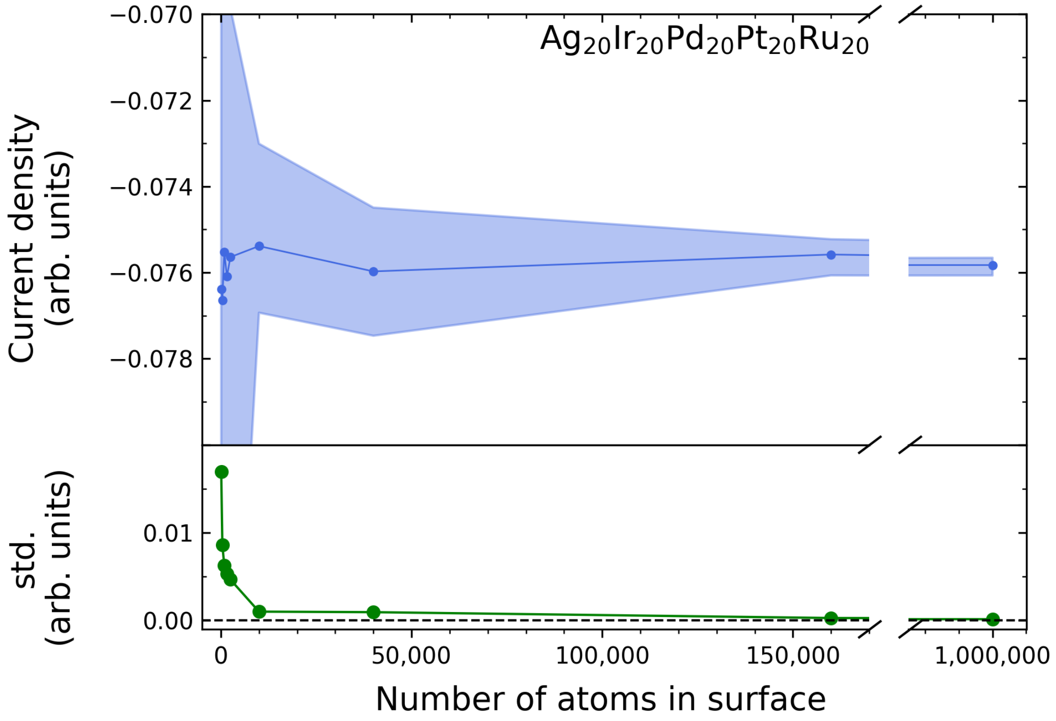

Figure S1. The predicted current density vs. the number of atoms of the simulated surface sampled for at least five random surfaces of equimolar AgIrPdPtRu at each of the sampled sizes. The standard deviation (std.) of the five sampled points are shown in green in the lower plot. At 10,000 atoms in the surface (100x100 atoms) the variation in the simulated current density is appreciably low compared to the variation between compositions (see for instance Figure 3 in the main text).

Gaussian Process

A Gaussian process regressor as implemented in scikit-learn[7] was constructed by setting the prior mean to zero everywhere, and defining the kernel (or covariance function) as shown in Equation 4 in the main text. This choice of kernel and prior mean is often a standard choice in machine learning applications, because of its general applicability and limited prior knowledge about the function being estimated. However, the squared exponential kernel does guarantee continuity and differentiability of the realizations of the Gaussian process, which we would also expect for the current densities of a composition space. It therefore forms a natural starting point for the present implementation.

When training the Gaussian process on the sampled compositions, the hyperparameters C and l were updated as implemented with the default choice of parameters in scikit-learn. The evolution of the hyperparameters as more samples were added are shown for illustration in Figure 2 in the main text.

Bayesian Optimization

In order to find optimal compositions with Bayesian inference, the Gaussian process regressor was initially trained on two randomly selected molar fractions along with their corresponding simulated current densities obtained with the kinetic model. The Gaussian process regressor was then used to predict current densities and surrogate model uncertainties at 1000 randomly selected molar fractions in order to span the quinary composition space in an approximate manner. The selection of the most optimal composition to sample next was performed with the *expected improvement* acquisition function. The principle of this acquisition function is to evaluate the expectation value of the improvement function,[12]

$$\mathrm{E}[I(\boldsymbol{x})]=\mathrm{E}[\max(y_{min}+\xi-Y,0)] \tag{S1}$$

at a molar fraction $\boldsymbol{x}$ for randomly distributed current densities Y and for the highest absolute value of the current density y_{min} sampled by the kinetic model so far. ξ is a tunable parameter that effectively adjusts y_{min}. If ξ is chosen to be greater than zero, the minimum found so far is effectively increased making molar fractions with greater probability of having current densities below the minimum have larger expected improvements and are therefore more likely to be compositions that could further minimize the current density. Assuming that the current densities at $\boldsymbol{x}$ are normally distributed with mean and standard deviation given by the prediction of the surrogate Gaussian process regressor, the expected improvement can be evaluated as

$$\mathrm{E}[I(\boldsymbol{x})] = (y_{min} + \xi - \mu(\boldsymbol{x}))\Phi\left(\frac{y_{min} + \xi - \mu(\boldsymbol{x})}{\sigma(\boldsymbol{x})}\right) + \sigma(\boldsymbol{x})\phi\left(\frac{y_{min} + \xi - \mu(\boldsymbol{x})}{\sigma(\boldsymbol{x})}\right), \qquad \text{(S2)}$$

where μ(x) and σ(x) are the mean and standard deviation supplied by the Gaussian process regressor, respectively, and $\phi(t) = (1/\sqrt{2\pi})\exp(-t^2/2)$ and $\Phi(t) = \int_{-\infty}^{t} \phi(t')dt'$ are the standard normal probability and cumulative distributions, respectively. The expected improvement was evaluated at the same 1000 compositions as the Gaussian process regressor, and the composition with the maximum acquisition value was further optimized by sampling the expected improvement around the composition in molar fraction steps of 0.005 until a maximum was found that was then selected for sampling by the kinetic model. A ξ-value of 0.01 was used throughout. A ξ-value of zero was found to potentially discover the locally optimal compositions very quickly. However, discovery of the global optimum was not guaranteed with 150 samples as was the case for ξ=0.01.

Experimental Procedures

Electrochemical characterization

Binary thin-film composition spreads were analyzed using a high-throughput scanning droplet cell (SDC) coupled with a Jaissle potentiostat/galvanostat. The teflon tip forming the head of the SDC had an opening of 1 mm in diameter, which formed the working electrode in each of the measurement areas (MAs) with a size of $7.35\cdot10^{-3}$ cm^2, allowing local characterization of the samples. Particular MAs on all of the samples were separated from each other by 2.25 mm, which corresponds to composition changes of ca. 1.5 at.% per element. All electrochemical measurements were conducted in 0.1 M $HClO_4$ electrolyte in a three-electrode system with a Ag|AgCl|3M KCl and a Pt wire as a reference and counter electrode, respectively. Linear sweep voltammetry was performed between 1 V and 200 mV vs. the reversible hydrogen electrode (RHE) with a scan rate of 10 mV s^{-1}. All potentials are reported versus the RHE calculated according to the following equation:

$$U_{RHE} = U_{Ag|AgCl|3\,M\,KCl} + 0.210\,V + 0.059\,V\cdot pH\,, \qquad \text{(S3)}$$

where $U_{Ag|AgCl|3M\ KCl}$ is the potential measured vs. the Ag|AgCl|3M KCl reference electrode, 0.210 V is the standard potential of the Ag|AgCl|3M KCl reference electrode at 25 °C. 0.059 V is the result of ln(10)RT/nF, where R is the gas constant, T is the temperature (298 K), F is the Faraday constant and n (=1) is the number of electrons transferred during the reaction.

Composition analysis

The elemental compositions of all MAs in the MLs were determined using automated energy dispersive X-ray spectroscopy (EDX) at 20 kV acceleration voltage in a scanning electron microscope (SEM, JEOL 5800) using a detector (INCA X-act, Oxford Instruments).

Surface roughness analysis by AFM

Topographical images of the Ag-Pd, Pd-Ru, Ir-Pt and Ir-(Hi) Pt thin film libraries were measured by atomic force microscopy (AFM, Bruker Dimension Fastscan) using Fastscan mode. For surfaces, whose roughness is characterized by a single length scale, roughness parameters were calculated by the arithmetic mean roughness Ra.

Phase analysis from XRD

The crystallographic phase analysis was performed using X-ray diffraction (XRD). A Bruker D8 Discover with a Vantec-500 2D-detector in Bragg–Brentano geometry and Cu Kα X-ray source was used. To avoid Si-substrate peaks, measurements were performed in θ–2θ mode with a 2.5° offset on θ. Five frames were taken stepwise at every MA with an increment of θ/2θ 7.5°/15°, starting at 10°/25° and finishing at 40°/85°. In this way an angular 2θ range from approximately 10° to 100° was covered.

Thin-film fabrication Pd-Ru, Ir-Pt, and Ag-Pd

The Pd-Ru and Ir-Pt libraries were fabricated by a combinatorial magnetron sputtering system (DCA Instruments, Finland) equipped with five cathodes. Two of these five cathodes were positioned at 144° form each other to create composition gradients. High purity (Ir: 99.9%, Pt: 99.99%,) 100 mm diameter single-element targets were used. A confocally-placed 100 mm diameter sapphire wafer (c-plane) was used as a substrate for the Ir-Pt system. It was patterned with small numbered crosses by a photolithographic lift-off process to serve as a reference grid and for making local thickness measurements by stylus profilometry. All of the depositions were carried out without intentional heating. Prior to the deposition, the chamber vacuum was on the order of 10^{-5} Pa. During deposition, the pressure was set to 0.667 Pa using Ar (99.9999%) at a flow rate of 60 sccm, and the substrate was kept stationary to obtain continuous compositional gradients. The type of power supply used for each library and sputter powers are listed in Table S1.

The Ag-Pd system was deposited in an alternate vacuum chamber, where cathodes with 38 mm diameter targets (Ag: 99.99%, Pd: 99.95%) are positioned 180° to each other. The substrate used was an approximately 1 cm wide strip cleaved from a 100 mm diameter <100> Si wafer, which was thermally oxidized as a diffusion barrier. The chamber base vacuum was 10^{-4} Pa and deposition was done at an Ar pressure of 0.5 Pa. The 100 mm diameter (100) Si substrate with a 500 nm SiO_2 barrier layer was stationary at the confocal point of the tilted cathodes so that composition gradients were obtained.

Table S1. Sputter parameters for the Pd-Ru, Ir-Pt and Ir-(High)Pt, respectively.

Libraries	Deposition power (W)			
	Pd (RF)[a]	Ru (DC)[b]	Ir (DC)[b]	Pt (RF)[a]
Pd-Ru	182	44	-	-
Ir-Pt	-	-	70	194
Ir-(Hgh)Pt	-	-	40	239

[a] RF: Radio frequency. [b] DC: Direct current.

Results and Discussion

Number of samples in a grid search of an N-component composition space
The number of combinations of alloy compositions in steps of molar fractions of *s* are given by equation S4.

$$N=\frac{\left(\frac{1}{s}+N_{elems}-1\right)!}{\left(\frac{1}{s}\right)!\left(N_{elems}-1\right)!} \tag{S4}$$

where N_{elems} is the number of metals in the alloy system. For example, to uniformly span the composition space of a quinary alloy in 5% intervals 10,626 points are needed. Figure S2 shows the number of combinations needed to span the composition space for various steps of molar fractions. It is observed that as the number of elements increases, the exploration of the composition space becomes increasingly infeasible as the number of samples needed increase combinatorially.

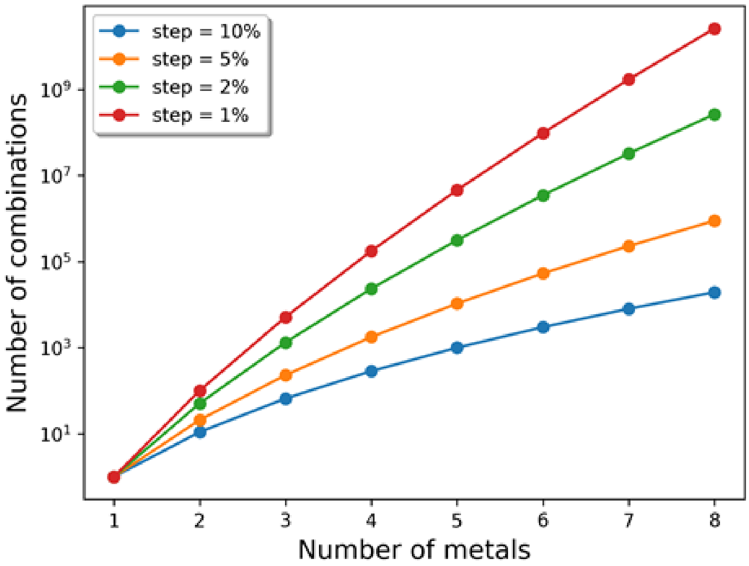

Figure S2. Number of samples needed to span the composition space. Shown for molar fraction step sizes of 1, 2, 5 and 10% as a function of the number of metals in the alloy.

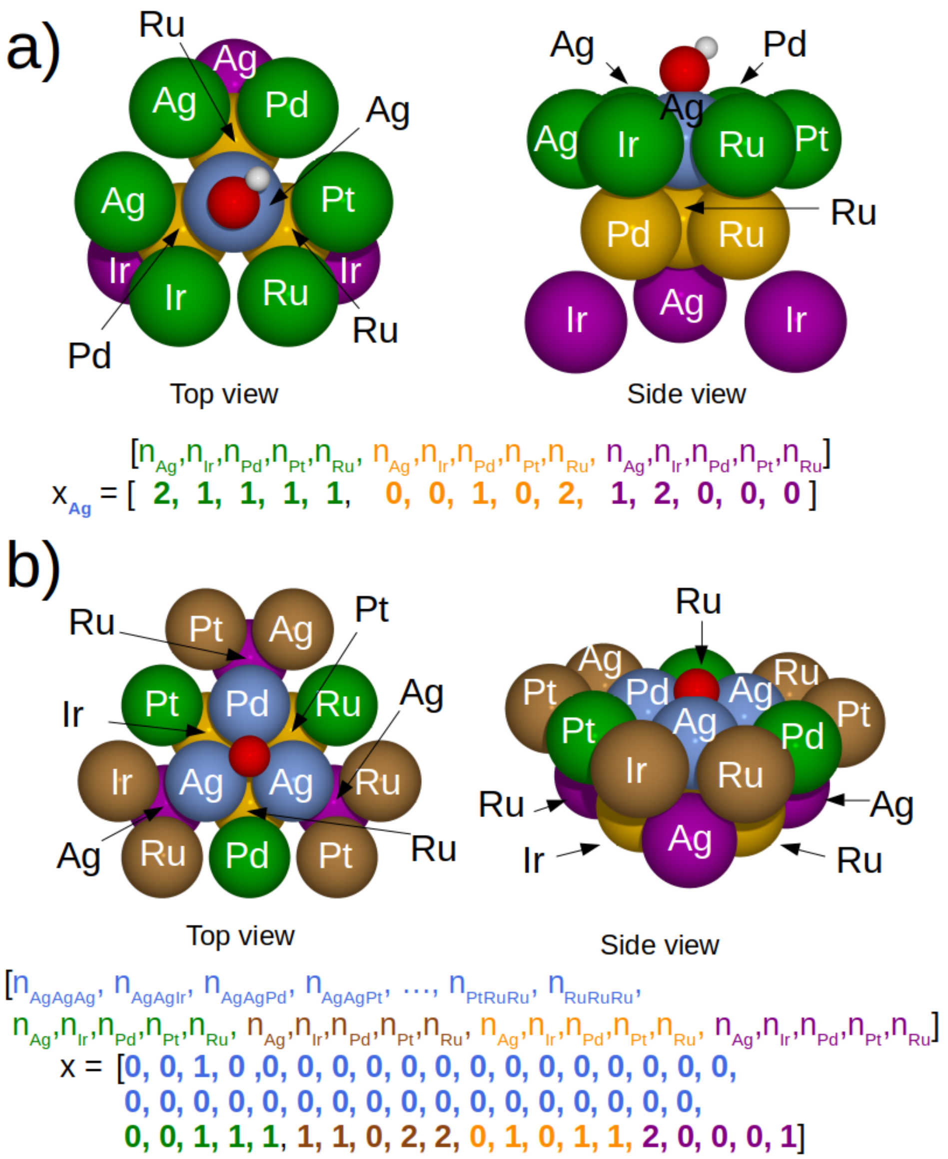

Figure S3. Examples of encoding the features of a structure. Shown in **a)** for an *OH on-top site with a set of features for each possible on-top site element (exemplified for an Ag on-top site) with a total of 15 parameters, and in **b)** for an O* fcc hollow site with one-hot encoding of the adsorption site ensemble (here exemplified for the AgAgPd ensemble) with a total of 55 parameters. The color of the text matches the corresponding colors of the atoms in the structure.

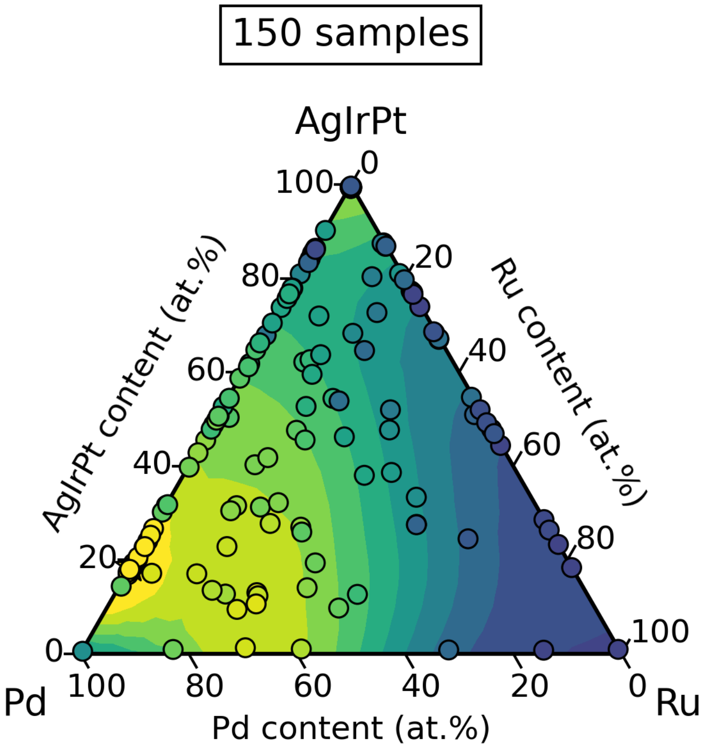

Figure S4. Pseudo-ternary plot of the Ag-Ir-Pd-Pt-Ru modeled current densities after sampling of 150 samples with Bayesian optimization. The Ag, Ir, and Pt concentrations have been grouped into one to highlight the plateau of similar current densities when the binary $Pd_{65}Ru_{35}$ is mixed with other elements in trace amounts. Yellow colors correspond to regions with high absolute values of the current density and blue colors to correspondingly low values. The projection of the current density from the quinary to the pseudo-ternary composition space was accomplished by showing the maximal absolute value of the current density possible for compositions that would otherwise be overlapping.

Table S2. Linear parameters used for on-top *OH adsorption energy prediction on the Ag-Ir-Pd-Pt-Ru system. The parameters will give the electronic energy in eV relative to *OH on Pt(111). The order of the parameters, after the intercept, follows the features given in Figure S3a. In the labels of the parameters the number refers to the layer, and the letter to the proximity to the adsorption site, e.g. “3a Pd” refers to the influence of Pd in the nearest atoms in the 3rd layer below the surface. The intercept has been chosen to yield the prediction for the pure element, since this value is obtained by setting the respective parameters for that element in each zone to zero.

Ag-Ir-Pd-Pt-Ru (eV relative to *OH@Pt(111))

	@Ag	**@Ir**	**@Pd**	**@Pt**	**@Ru**
Intercept	0.262515	-0.363696	0.075474	-0.023367	-0.604279
1b Ag	0	-0.074370	-0.025064	-0.065560	-0.026962
1b Ir	0.124874	0	0.044604	0.014894	0.025734
1b Pd	0.055553	-0.057413	0	-0.042658	-0.021608
1b Pt	0.106736	-0.016379	0.040827	0	0.021586
1b Ru	0.083629	-0.019943	0.007435	-0.016870	0
2a Ag	0	0.020613	0.022701	0.076312	-0.097415
2a Ir	-0.103170	0	-0.050573	-0.022161	-0.022948
2a Pd	-0.025039	0.008149	0	0.046161	-0.070458
2a Pt	-0.057300	-0.012203	-0.032788	0	-0.059442
2a Ru	-0.121723	0.031539	-0.051361	-0.014583	0
3a Ag	0	-0.036480	-0.015738	-0.022794	-0.025380
3a Ir	0.019881	0	0.035545	0.032103	0.013630
3a Pd	-0.001190	-0.014928	0	-0.006961	-0.000195
3a Pt	0.005926	-0.012114	0.000055	0	0.008633
3a Ru	0.023315	-0.002794	0.033945	0.039372	0

Table S3. Linear parameters used for fcc hollow O* adsorption energy prediction on the Ag-Ir-Pd-Pt-Ru system. The parameters will give the electronic energy in eV relative to O* on Pt(111). The order of the parameters follows the features given in Figure S3b. In the labels of the parameters the number refers to the surface layer, and the letter to the proximity to the adsorption site, e.g. “2b Pd” refers to the influence of Pd in the next nearest atoms in the subsurface layer.

Ag-Ir-Pd-Pt-Ru (eV relative to O*@Pt(111))							
AgAgAg	0.823301	**AgAgIr**	-0.470179	**AgAgPd**	0.612573	**AgAgPt**	0.374752
AgAgRu	-0.869420	**AgIrIr**	-0.603051	**AgIrPd**	-0.386478	**AgIrPt**	-0.296955
AgAgRu	-0.971134	**AgPdPd**	0.417371	**AgPdPt**	0.226634	**AgPdRu**	-0.897300
AgPtPt	0.231177	**AgPtRu**	-0.677207	**AgRuRu**	-1.168980	**IrIrIr**	-0.884791
IrIrPd	-0.791251	**IrIrPt**	-0.645038	**IrIrRu**	-1.154468	**IrPdPd**	-0.586752
IrPdPt	-0.435780	**IrPdRu**	-1.101158	**IrPtPt**	-0.331185	**IrPtRu**	-0.912780
IrRuRu	-1.396306	**PdPdPd**	0.142790	**PdPdPt**	0.030062	**PdPdRu**	-0.915004
PdPtPt	0.073705	**PdPtRu**	-0.771205	**PdRuRu**	-1.356419	**PtPtPt**	0.092753
PtPtRu	-0.617444	**PtRuRu**	-1.178394	**RuRuRu**	-1.654559	-	-
1b Ag	-0.081909	**1c Ag**	0.028809	**2a Ag**	-0.040527	**2B Ag**	0.005127
1b Ir	0.062988	**1c Ir**	-0.020630	**2a Ir**	0.030518	**2b Ir**	-0.004639
1b Pd	-0.047180	**1c Pd**	0.014160	**2a Pd**	-0.022949	**2b Pd**	-0.013672
1b Pt	0.025269	**1c Pt**	0.002563	**2a Pt**	-0.011230	**2b Pt**	-0.027344
1b Ru	0.040649	**1c Ru**	-0.025635	**2a Ru**	0.042969	**2b Ru**	0.040405

Table S4. Linear parameters used for on-top *OH adsorption energy prediction on the Ir-Pd-Pt-Rh-Ru system. The parameters will give the electronic energy in eV relative to *OH on Pt(111). The order of the parameters, after the intercept, follows the features given in Figure S3a: In the labels of the parameters the number refers to the layer, and the letter to the proximity to the adsorption site, e.g. “3a Pd” refers to the influence of Pd in the nearest atoms in the 3rd layer below the surface. The intercept has been chosen to yield the prediction for the pure element, since this value is obtained by setting the respective parameters for that element in each zone to zero.

Ir-Pd-Pt-Rh-Ru (eV relative to *OH@Pt(111))

	@Ir	**@Pd**	**@Pt**	**@Rh**	**@Ru**
Intercept	-0.324264	0.044878	-0.008922	-0.323476	-0.639564
1b Ir	0	0.045847	0.011095	0.037634	0.040998
1b Pd	-0.067114	0	-0.041923	-0.012061	-0.024511
1b Pt	-0.022217	0.041869	0	0.028185	0.018334
1b Rh	-0.048245	0.006241	-0.030170	0	-0.002666
1b Ru	-0.026561	0.001899	-0.020977	0.002107	0
2a Ir	0	-0.034497	-0.013930	-0.007092	-0.018194
2a Pd	0.017262	0	0.030018	-0.006571	-0.047992
2a Pt	-0.006588	-0.027407	0	-0.014891	-0.037786
2a Rh	0.009353	-0.020069	0.021633	0	-0.022024
2a Ru	0.021847	-0.030915	0.002223	0.005128	0
3a Ir	0	0.034227	0.029385	0.006703	0.004741
3a Pd	-0.026733	0	-0.016174	-0.021482	-0.015521
3a Pt	-0.016911	0.012355	0	-0.006625	-0.000825
3a Rh	-0.005591	0.015988	0.015289	0	-0.003526
3a Ru	0.006793	0.035460	0.039082	0.007343	0

Table S5. Linear parameters used for fcc hollow O* adsorption energy prediction on the Ir-Pd-Pt-Rh-Ru system. The parameters will give the electronic energy in eV relative to O* on Pt(111). The order of the parameters follows the features given in Figure S3b. In the labels of the parameters the number refers to the surface layer, and the letter to the proximity to the adsorption site, e.g. “2b Pd” refers to the influence of Pd in the next nearest atoms in the subsurface layer.

Ir-Pd-Pt-Rh-Ru (eV relative to O*@Pt(111))							
IrIrIr	-0.802671	**IrIrPd**	-0.698724	**IrIrPt**	-0.549723	**IrIrRh**	-0.867686
IrIrRu	-1.039895	**IrPdPd**	-0.490209	**IrPdPt**	-0.372960	**IrPdRh**	-0.717682
IrIrRu	-1.023565	**IrPtPt**	-0.244375	**IrPtRh**	-0.585337	**IrPtRu**	-0.815201
IrRhRh	-0.894818	**IrRhRu**	-1.128900	**IrRuRu**	-1.293938	**PdPdPd**	0.154385
PdPdPt	0.096998	**PdPdRh**	-0.273295	**PdPdRu**	-0.866591	**PdPtPt**	0.062937
PdPtRh	-0.252892	**PdPtRu**	-0.716471	**PdRhRh**	-0.602280	**PdRhRu**	-1.054314
PdRuRu	-1.262522	**PtPtPt**	0.183877	**PtPtRh**	-0.153542	**PtPtRu**	-0.560802
PtRhRh	-0.506731	**PtRhRu**	-0.874115	**PtRuRu**	-1.077222	**RhRhRh**	-0.856961
RhRhRu	-1.144941	**RhRuRu**	-1.392573	**RuRuRu**	-1.563099	-	-
1b Ir	0.045966	**1c Ir**	-0.020285	**2a Ir**	0.026245	**2B Ir**	0.002915
1b Pd	-0.061963	**1c Pd**	0.029724	**2a Pd**	-0.040319	**2b Pd**	-0.028029
1b Pt	0.006466	**1c Pt**	0.012666	**2a Pt**	-0.022935	**2b Pt**	-0.039050
1b Rh	-0.011702	**1c Rh**	0.000303	**2a Rh**	0.007976	**2b Rh**	0.003210
1b Ru	0.021233	**1c Ru**	-0.022409	**2a Ru**	0.029033	**2b Ru**	0.060954

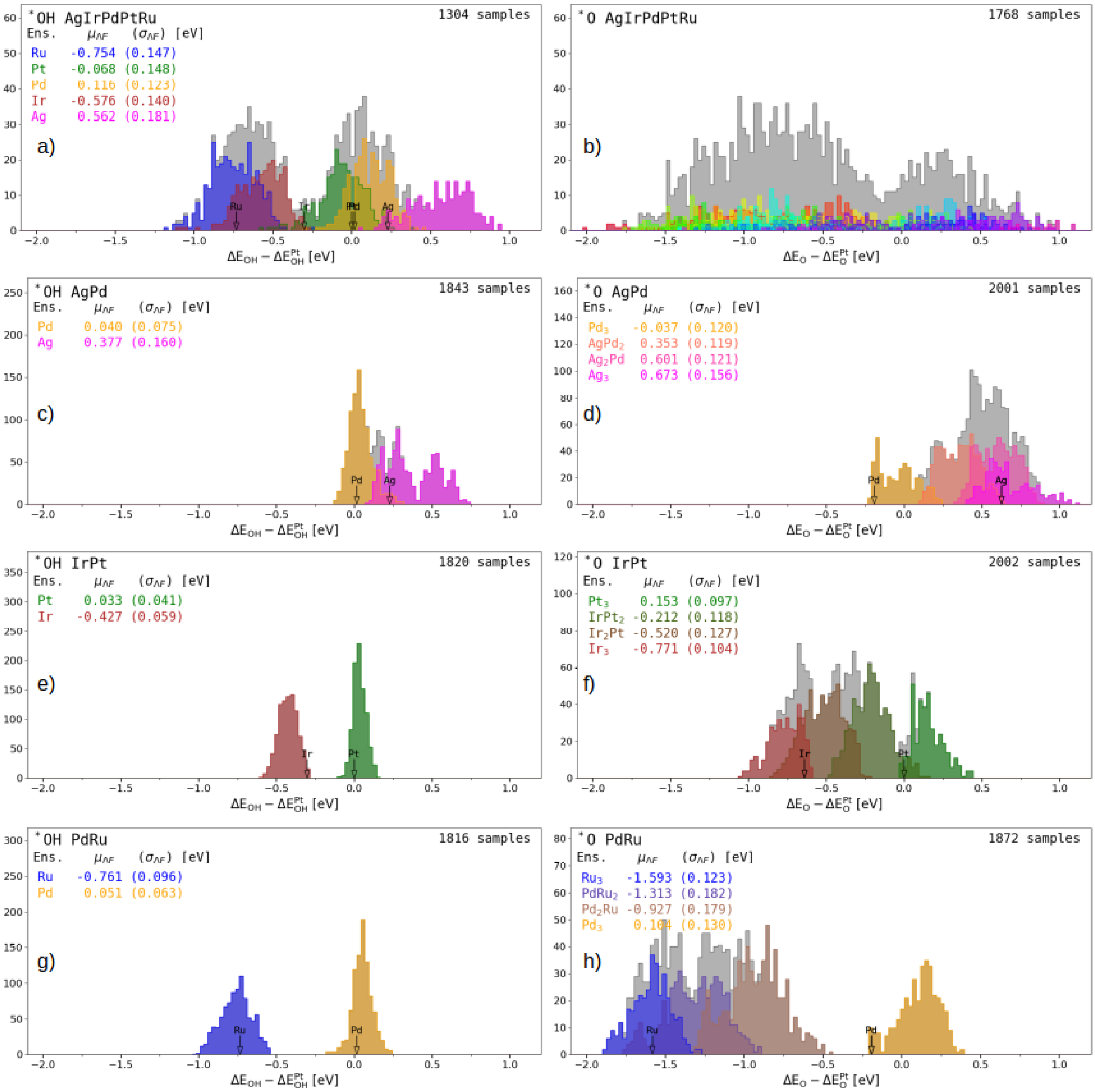

Figure S5. Histograms of DFT calculated adsorption energies of *OH and *O on the quinary alloy and the three binary alloys. Adsorption ensembles are distinguished by different colors with the mean adsorption energy ($\mu_{\Delta E}$) and standard deviation ($\sigma_{\Delta E}$). Adsorption energies of the pure metal fcc(111) surfaces are marked with black arrows.

Table S6. Mean absolute errors (MAEs) in units of eV of several regression algorithms predicting adsorption energies of *OH and O* on the quinary and binary alloys using truncated site features as displayed in Figure S3. 20% of the samples were selected for testing with the remaining samples used to train the regression model. The standard deviations on the last digit(s) of the MAEs are displayed in parentheses.

*OH adsorption energies - Truncated adsorption site features

Regressor type	Ag-Pd (1843 samples)	Ir-Pt (1820 samples)	Pd-Ru (1816 samples)	Ag-Ir-Pd-Pt-Ru (1304 samples)
Dummy (mean)	0.09(7)	0.05(4)	0.06(5)	0.12(9)
Linear regr.	0.04(3)	0.02(2)	0.05(4)	0.06(5)
Ridge regr.	0.04(3)	0.02(2)	0.05(4)	0.06(5)
Gradient Boosting	0.03(2)	0.018(14)	0.03(3)	0.07(6)
Random Forest	0.03(2)	0.018(15)	0.04(3)	0.07(6)

*O adsorption energies - Truncated adsorption site features

Regressor type	Ag-Pd (2001 samples)	Ir-Pt (2002 samples)	Pd-Ru (1872 samples)	Ag-Ir-Pd-Pt-Ru (1768 samples)
Dummy (mean)	0.11(8)	0.10(8)	0.15(10)	0.5(4)
Linear regr.	0.05(4)	0.04(3)	0.08(6)	0.09(7)
Ridge regr.	0.05(4)	0.04(3)	0.08(5)	0.11(8)
Gradient Boosting	0.04(3)	0.04(3)	0.05(4)	0.11(8)
Random Forest	0.03(3)	0.04(3)	0.05(4)	0.10(8)

Table S7. Mean absolute errors (MAEs) in units of eV of several regression algorithms predicting adsorption energies of *OH and O* on the quinary and binary alloys using extended site features (up to fourth-nearest neighboring atoms for all layers). 20% of the samples were selected for testing with the remaining samples used to train the regression model. The standard deviations on the last digit(s) of the MAEs are displayed in parentheses.

*OH adsorption energies - Extended adsorption site features

Regressor type	Ag-Pd (1843 samples)	Ir-Pt (1820 samples)	Pd-Ru (1816 samples)	Ag-Ir-Pd-Pt-Ru (1304 samples)
Dummy (mean)	0.09(7)	0.05(4)	0.06(5)	0.12(9)
Linear regr.	0.04(3)	0.020(17)	0.05(4)	0.08(7)
Ridge regr.	0.03(3)	0.020(19)	0.04(3)	0.06(5)
Gradient Boosting	0.017(16)	0.012(9)	0.03(2)	0.07(5)
Random Forest	0.015(13)	0.011(10)	0.03(2)	0.07(6)

*O adsorption energies - Extended adsorption site features

Regressor type	AgPd (2001 samples)	IrPt (2002 samples)	PdRu (1872 samples)	AgIrPdPtRu (1768 samples)
Dummy (mean)	0.11(8)	0.10(8)	0.15(10)	0.5(4)
Linear regr.	0.04(4)	0.02(2)	0.06(4)	0.09(7)
Ridge regr.	0.05(3)	0.026(19)	0.06(4)	0.10(8)
Gradient Boosting	0.024(19)	0.021(17)	0.04(3)	0.08(6)
Random Forest	0.020(18)	0.024(19)	0.03(3)	0.09(8)

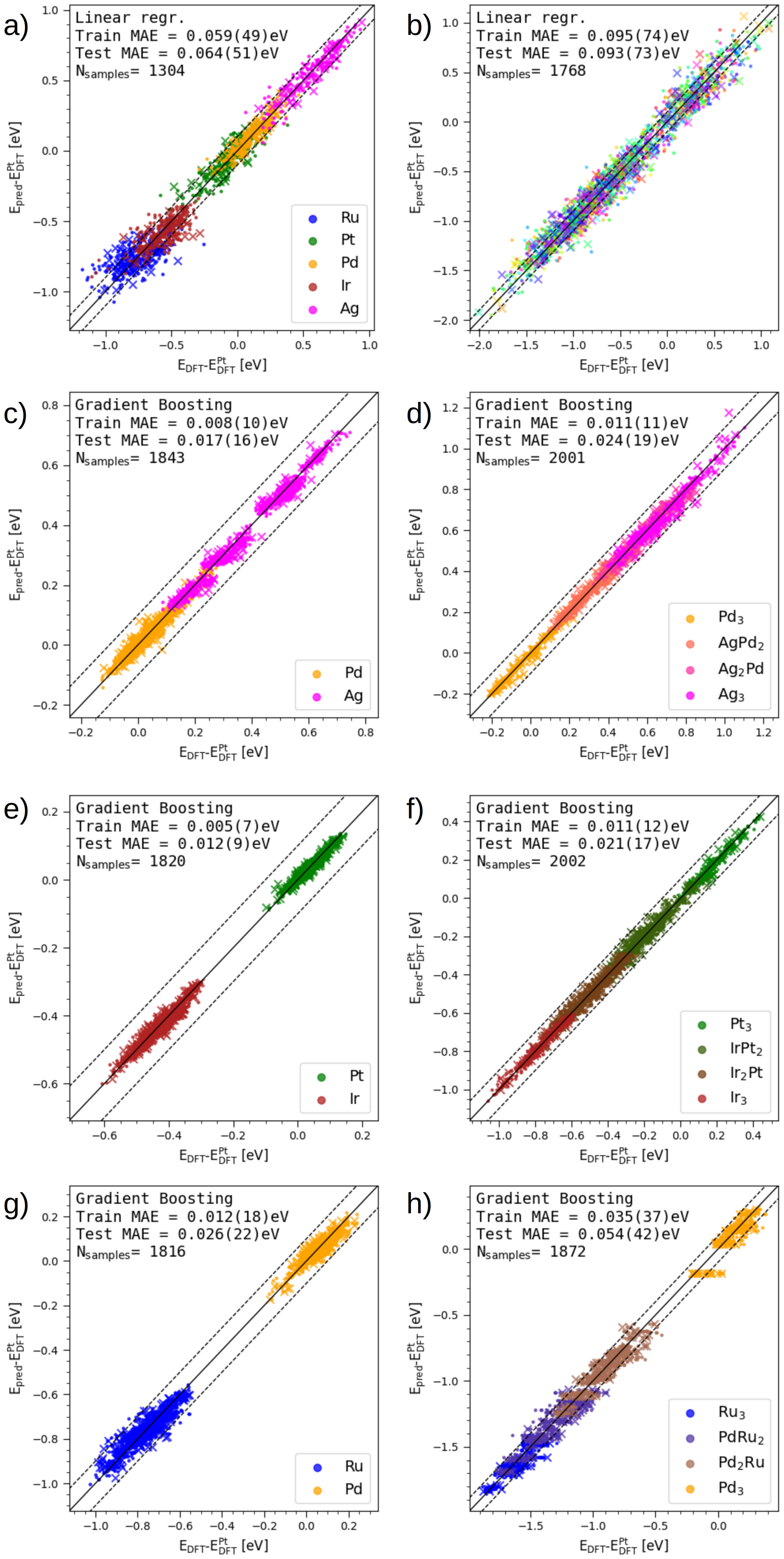

Figure S6. Predicted adsorption energy plotted against DFT calculated adsorption energy for the Ag-Ir-Pd-Pt-Ru linear model (**a**,**b**), Ag-Pd (**c**,**d**), Ir-Pt (**e**,**f**), and Pd-Ru (**g**,**h**) gradient boosted models for on-top *OH (**a**,**c**,**e**,**g**) and fcc hollow O* (**b**,**d**,**f**,**h**) adsorption on fcc(111) surfaces. The colors indicate the identity of the adsorption site as shown in the legend. Mean absolute errors with standard deviations are displayed for both training and test set. 20% of the samples were selected for testing (crosses) with the remaining samples used to train the model (circles). For the quinary alloy the results of the linear regression model trained on the truncated site features are displayed, while for the binary alloys the results of the gradient boosted model trained on the extended site features are displayed.

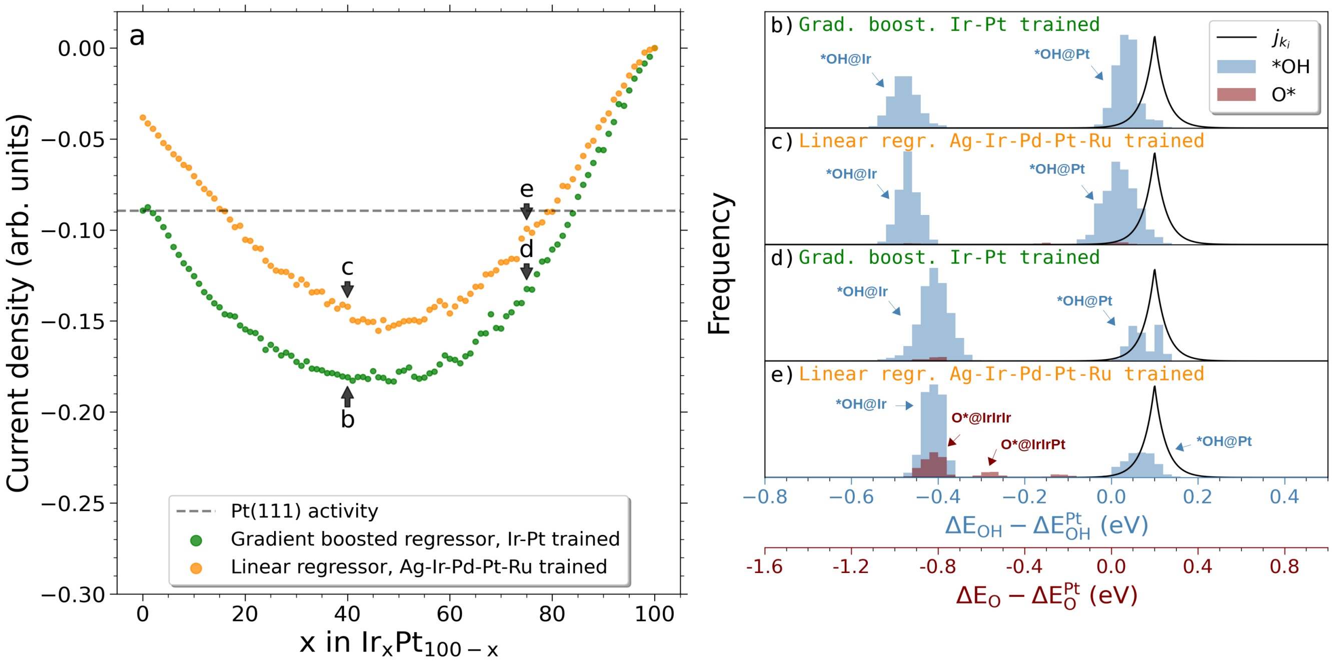

Figure S7. a) Simulated current densities of the Ir-Pt binary system shown as a scan from pure Pt to pure Ir with 1 at.% increments. A linear regression model trained on the DFT calculated samples of the quinary alloy is used alongside a gradient boosted model trained on DFT calculated samples of Ir-Pt to predict the adsorption energies of the simulated surface. These predictions serve as input for Equations 1-3 which yield the resulting current densities. **b-e)** *OH and O* net adsorption energy distributions (after intersite blocking) for selected compositions corresponding to the annotations in **a)**. A scaled visualization of the modeled current density in Equation 3 is shown (black solid line).

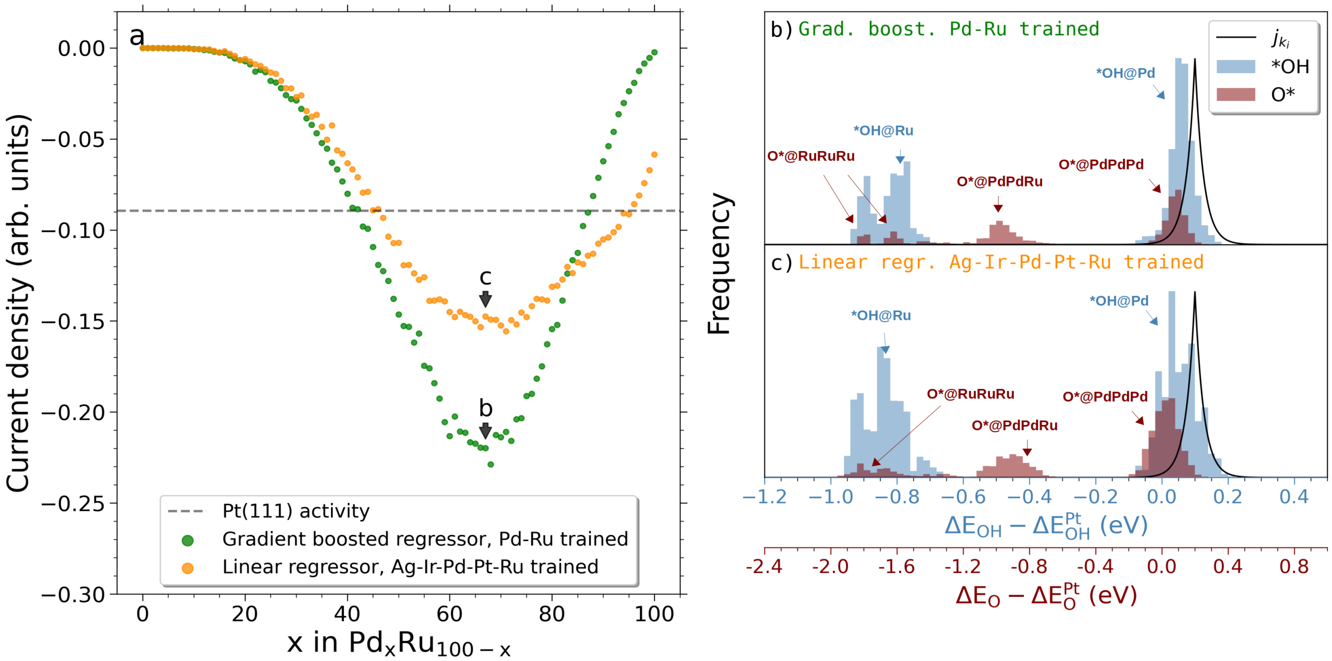

Figure S8. a) Simulated current densities of the Pd-Ru binary system shown as a scan from pure Ru to pure Pd with 1 at.% increments. A linear regression model trained on the DFT calculated samples of the quinary alloy is used alongside a gradient boosted model trained on DFT calculated samples of Pd-Ru to predict the adsorption energies of the simulated surface. These predictions serve as input for Equations 1-3 which yield the resulting current densities. **b,c)** *OH and O* net adsorption energy distributions (after intersite blocking) for selected compositions corresponding to the annotations in **a)**. A scaled visualization of the modeled current density in Equation 3 is shown (black solid line).

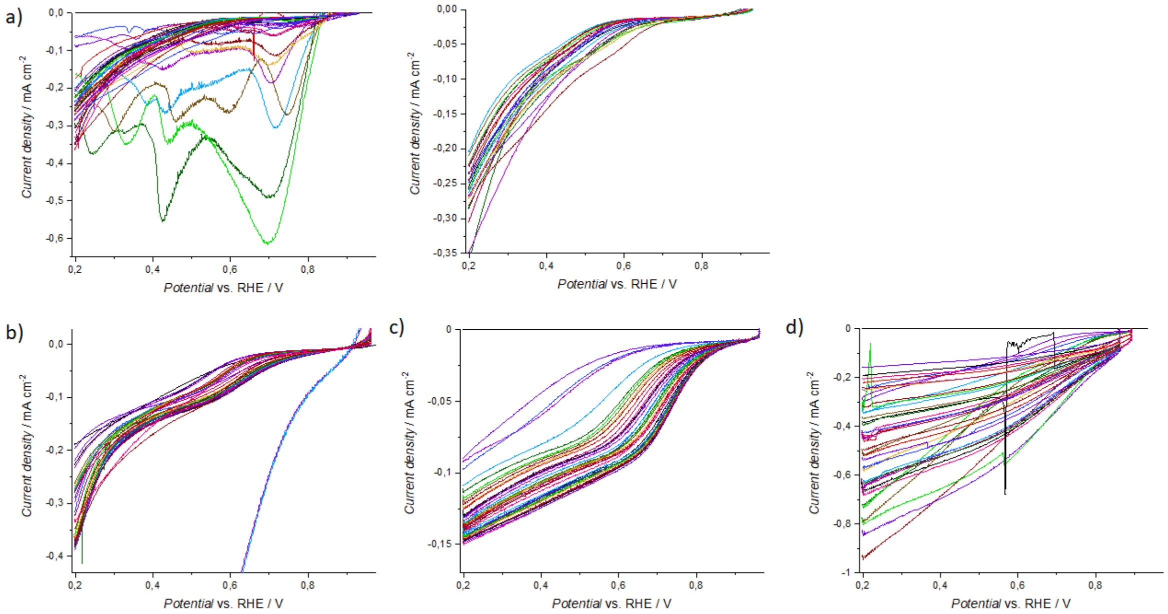

Figure S9. All automatically measured LSV curves. **a)** Ag-Pd (right side: LSV plots from the low-Ag part of the sample, without visible film corrosion), **b)** Pd-Ru, and **c)** and **d)** Ir-Pt binaries.

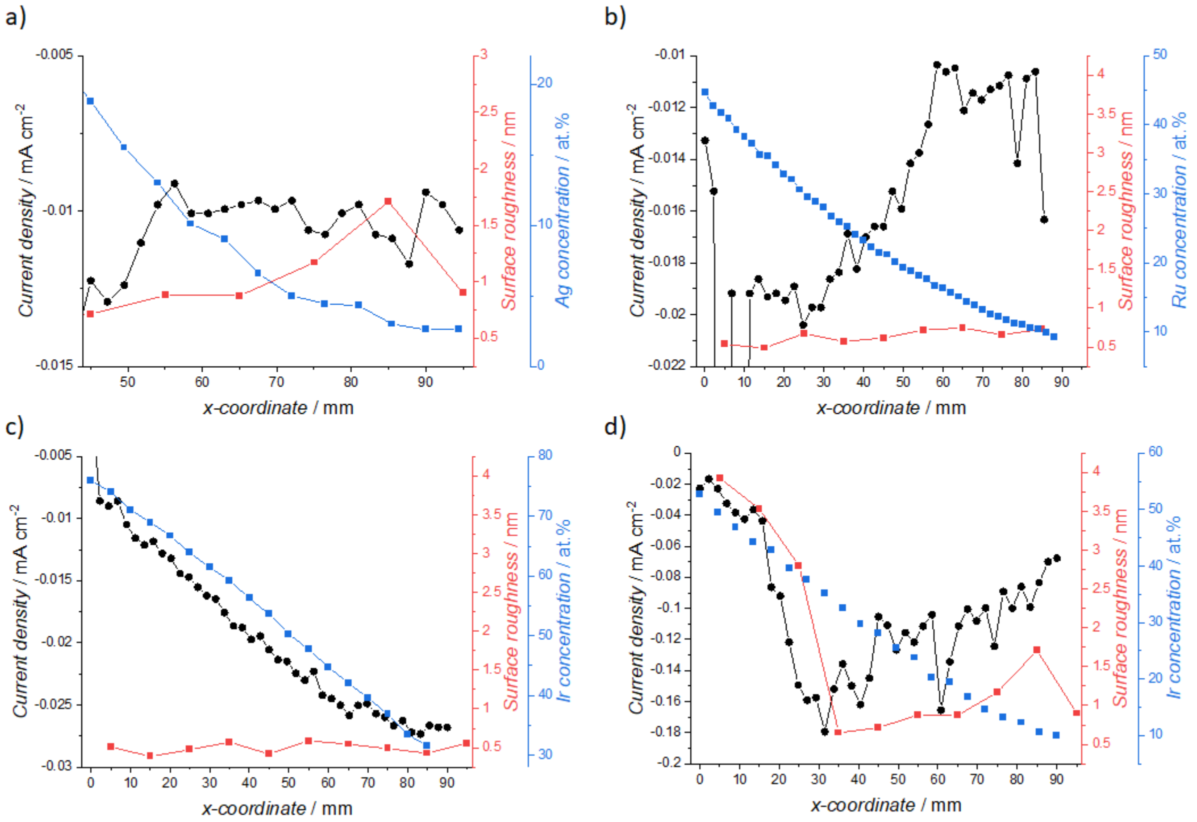

Figure S10. Comparisons of measured ORR current densities (black curve) with sample composition (blue curve) and surface roughness (red curve) for synthesized thin-films of **a)** Ag-Pd, **b)** Pd-Ru, **c)** Ir-Pt, and **d)** Ir-(High)Pt.

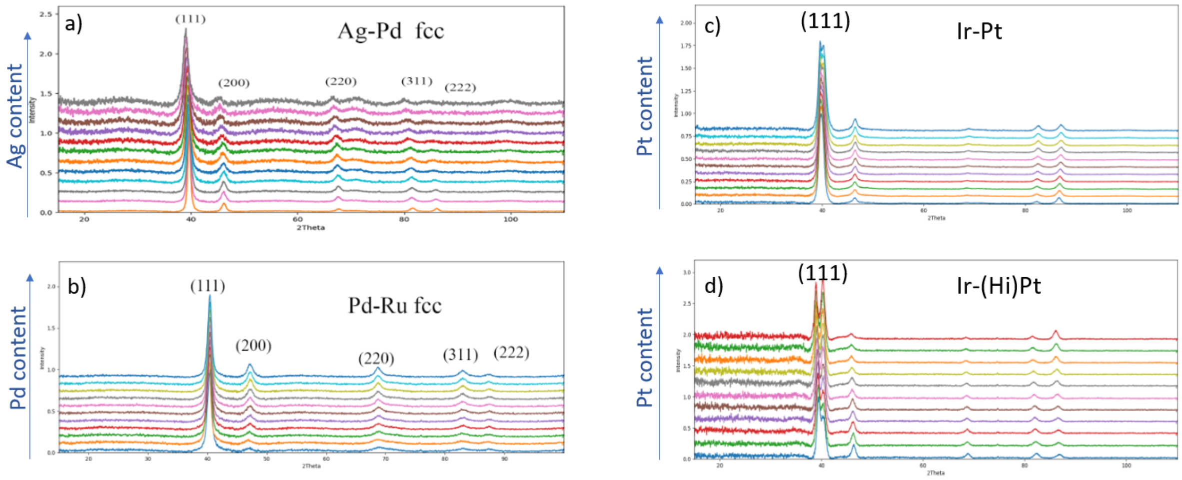

Figure S11. XRD profiles for **a)** Ag-Pd, **b)** Pd-Ru **c)** Ir-Pt, and **d)** Ir-(High)Pt binaries. A single fcc phase is observed for both the Ag-Pd and Pd-Ru systems, regardless of composition, while a dual phase is found for the Ir-Pt system.

References

[1] B. Hammer, L. B. Hansen, J. K. Nørskov, *Phys. Rev. B* **1999**, 59, 7413.

[2] J. J. Mortensen, L. B. Hansen, K. W. Jacobsen, *Phys. Rev. B* **2005**, 71, 035109.

[3] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dułak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, et al., *J. Phys.: Condens. Matter* **2010**, 22, 253202.

[4] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, et al., *J. Phys.: Condens. Matter* **2017**, 29, 273002.

[5] T. A. A. Batchelor, J. K. Pedersen, S. H. Winther, I. E. Castelli, K. W. Jacobsen, J. Rossmeisl, *Joule* **2019**, 3, 834.

[6] C. M. Clausen, T. A. A. Batchelor, J. K. Pedersen, J. Rossmeisl, *Adv. Sci.* **2021**, , 2003357.

[7] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., *J. Mach. Learn. Res.* **2011**, 12, 2825.

[8] I. E. L. Stephens, A. S. Bondarenko, U. Grønbjerg, J. Rossmeisl, I. Chorkendorff, *Energy Environ. Sci.* **2012**, 5, 6744.

[9] J. Greeley, I. E. L. Stephens, A. S. Bondarenko, T. P. Johansson, H. A. Hansen, T. F. Jaramillo, J. Rossmeisl, I. Chorkendorff, J. K. Nørskov, *Nat. Chem.* **2009**, 1, 552.

[10] J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard, H. Jonsson, *J. Phys. Chem. B* **2004**, 108, 17886.

[10] T. A. A. Batchelor, T. Löffler, B. Xiao, O. A. Krysiak, V. Strotkötter, J. K. Pedersen, C. M. Clausen, A. Savan, Y. Li, W. Schuhmann, et al., *Angew. Chem. Int. Ed.* **2021**, 60, 6932.

[11] D. R. Jones, M. Schonlau, W. J. Welch, *J. Global Optim.* **1998**, 13, 455.