High-throughput screening and DFT characterization of bimetallic alloy catalysts for the nitrogen reduction reaction

Mehdi Shamekhi,[a,b] Arash Toghraei [c] , Daniel Guay[*b] and Gilles H. Peslherbe [*a,b,d,e]

[a]*Department of Physics, Concordia University, Montreal, Quebec, Canada*

[b]*Centre for Research in Molecular Modeling, Concordia University, Montreal, Quebec, Canada*

[c]*Centre of Energy, Materials and Telecommunications, INRS, Varennes, Quebec, Canada*

[d]*Department of Chemistry and Biochemistry, Concordia University, Montreal, Quebec, Canada*

[e]*Department of Chemical and Materials Engineering, Montreal, Quebec, Canada*

Table of Contents

Computational hydrogen electrode……2
d-band model……3
Hyperparameter tuning……3
Model performance……6
Formation energy……10

Figures

Figure S1: MLA feature distributions ……5
Figure S2: Training and validation loss curves of the ANN. ……6
Figure S3: Relative feature importance of the train set used for MLA training. ……7
Figure S4: Features of the distal vs enzymatic pathway for the NRR over Au@Au_3Re and Au@Au_3Mo. ……8
Figure S5: Crystal orbital Hamilton population for the interactions of Re in Au_3Re and adsorbed N_2. ……9
Figure S6: Adsorption sites considered in the search of minimum energy configurations. ……10
Figure S7: Net charge on active sites (Re in Au_3Re and Mo in Au_3Mo)……11
Figure S8: PDOS of the Mo active site in Au-Mo surface alloys ……12
Figure S9: Free energy profile of the NRR distal pathway ……13
Figure S10: Minimum energy configuration of NRR intermediates ……14
Figure S11: MLA-predicted vs DFT-calculated NRR limiting potentials……16
Figure S12: PDOS of N_2 before (blue) and after adsorption (red) on Au_3Re. ……17
Figure S13: PDOS of (a) Re and (b) Mo active sites in Au_3Re and Au_3Mo ……18
Figure S14: Hydrogenation sequence of NRR intermediates……19
Figure S15: Minimum energy configuration of NRR intermediates ……20
Figure S16: Free energy profile of the HER Volmer-Heyrovsky pathway. ……21

Tables

Table S1: Experimental and calculated gas-phase thermochemistry data 4
Table S2: Hyperparameters of machine learning algorithms 4
Table S3: Performance of machine learning algorithms.....5
Table S4 Absolute variance of training features used in the MLA model. 7
Table S5 Formation energy of surface alloys in eV. 10
Table S6: Adsorption free energy of NRR intermediates over selected solid supports (eV) 11
Table S7: Adsorption free energies of N_2 and H along with selectivity factor (eV).....15

1 Computational hydrogen electrode

Electrochemical reactions involve proton-coupled electron transfer steps which are affected by the electrode potential. This potential dependence is accounted for in the computational hydrogen electrode model originally developed by Norskov.[1] The reversible hydrogen electrode (RHE) is used as reference in this approach.

$$\frac{1}{2}H_2 \leftrightarrow H^+ + e^- \tag{S1}$$

At an applied potential *U* vs. RHE, the chemical potential of the proton-electron pair is shifted by -e*U*:

$$\mu(H^+ + e^-) = 1/2\ \mu(H_2) - eU \tag{S2}$$

Using Equation (2), the reaction free energy of an elementary step $*A + (H^+ + e^-) \rightarrow *AH$ with an applied potential *U* can be calculated as:

$$\Delta G = \Delta G_0 + eU \tag{S3}$$

where ΔG_0 is the reaction free energy at 0 V vs. RHE. The limiting potential is defined as the lowest applied potential (vs RHE) that makes the reaction (elementary steps) downhill in free energy.

$$U_{limiting} = -\frac{\Delta G_{max}}{e} \quad \text{(S4)}$$

2 d-band model

The interaction of the adsorbate and the catalyst surface involves interaction of *sp*-states of the adsorbate with both *sp*-states and *d*-states of the catalyst,[2,3] with an adsorption energy that can be decomposed as:

$$\Delta E = \Delta E_{sp} + \Delta E_{d} \quad \text{(S5)}$$

where ΔE_{sp} is the interaction of the adsorbate with *sp*-states of the surface and ΔE_{d} is the interaction of the adsorbate states with the *d*-states of the surface. Since all transition metals (TMs) have filled and broad *sp* orbitals, the adsorbate interaction with *sp* states (ΔE_{sp}) of the catalyst is the same for all TMs. Thus, ΔE depends primarily on the interaction of the *sp*-states of the adsorbate and the *d*-states of the catalyst. The projected density of states (PDOS) onto the metal *d*-orbitals that interact with the adsorbate *sp*-orbitals is characterized by its moments: the first moment is the *d*-band center (ϵ_d) and the second moment is the d-band width (W_d). As seen from the Pearson correlation map in the main text (Figure 3), the combination of ϵ_d and W_d is a good descriptor to estimate trends in adsorption energy for TMs with different adsorbates, as reported before in different contexts.[4-6]

3 Hyperparameter tuning

The hyperparameters that could be tuned for the ANN, SVR and RF algorithms are listed in Table S1. The optimal hyperparameter set for ANN is as follows: number of hidden layers (2 layers), learning rate (0.01), nodes (2 layers with 13 → 11 nodes/layer), batch size (16), activation function (relU) and total number of epochs (300) with early stopping. The loss function is set as the mean squared error (MSE). During the training phase, the loss function is minimized using the Adam optimization algorithm.

Table S1: Experimental and calculated gas-phase thermochemistry data

Species	E_{DFT}	ZPE (Exp.)	C_p (Exp.)	TS (Exp.)	ZPE (DFT)	TS (DFT)
N_2(g)	-35.13	0.15	0.09	0.59	0.14	0.59
H_2(g)	-8.01	0.19	0.09	0.40	0.25	0.40
NH_3(g)	-30.18	0.88	0.10	0.60	0.91	0.59

E_{DFT}: electronic energy; ZPE: zero-point energy; C_p: heat capacity; TS: vibrational entropy (S) multiplied by temperature T=298 K. All values are in eV. Experimental data from NIST (https://webbook.nist.gov/chemistry/form-ser/) and calculated data evaluated with BEEF-vdW/PAW (see section 3.1 of main text).

Table S2: Hyperparameters of machine learning algorithms

MLA	Hyperparameters	Range
ANN	Activation function	relU, gelU, selU
	Number of layers	1-2
	Neurons	4-18
	Epochs	150, 200, 250, 300
	Optimizer	Adam, rmprop
	Learning rate	0.01, 0.005, 0.001
	Batch size	16, 32
SVR	kernel	linear, rbf
	C	1, 10, 100, 1000
	Gamma	1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}
RF	n estimators	50, 100, 200, 300, 400
	max depth	1,2,3,4,5
	max features	sqrt
	min samples leaf	3,4,5

MLA: machine learning algorithm; ANN: artificial neural network: SVR: support vector regressor; RF: random forest.

Table S3: Performance of machine learning algorithms

MLA	MAE	MSE	RMSE
Linear regression	0.28 **(0.32)**	0.16 **(0.21)**	0.40 **(0.46)**
Ridge regression	0.31 **(0.34)**	0.17 **(0.22)**	0.42 **(0.47)**
Lasso	0.35 **(0.38)**	0.21 **(0.26)**	0.46 **(0.51)**
Elastic net	0.33 **(0.36)**	0.20 **(0.25)**	0.44 **(0.50)**
SVR	0.27 **(0.32)**	0.19 **(0.28)**	0.43 **(0.53)**
Random forest	0.17 **(0.25)**	0.06 **(0.15)**	0.26 **(0.39)**
Neural network	0.15 **(0.23)**	0.06 **(0.13)**	0.25 **(0.34)**

MAE: mean absolute error, MSE: mean squared error; RMSE: root mean squared error. All values are in eV.

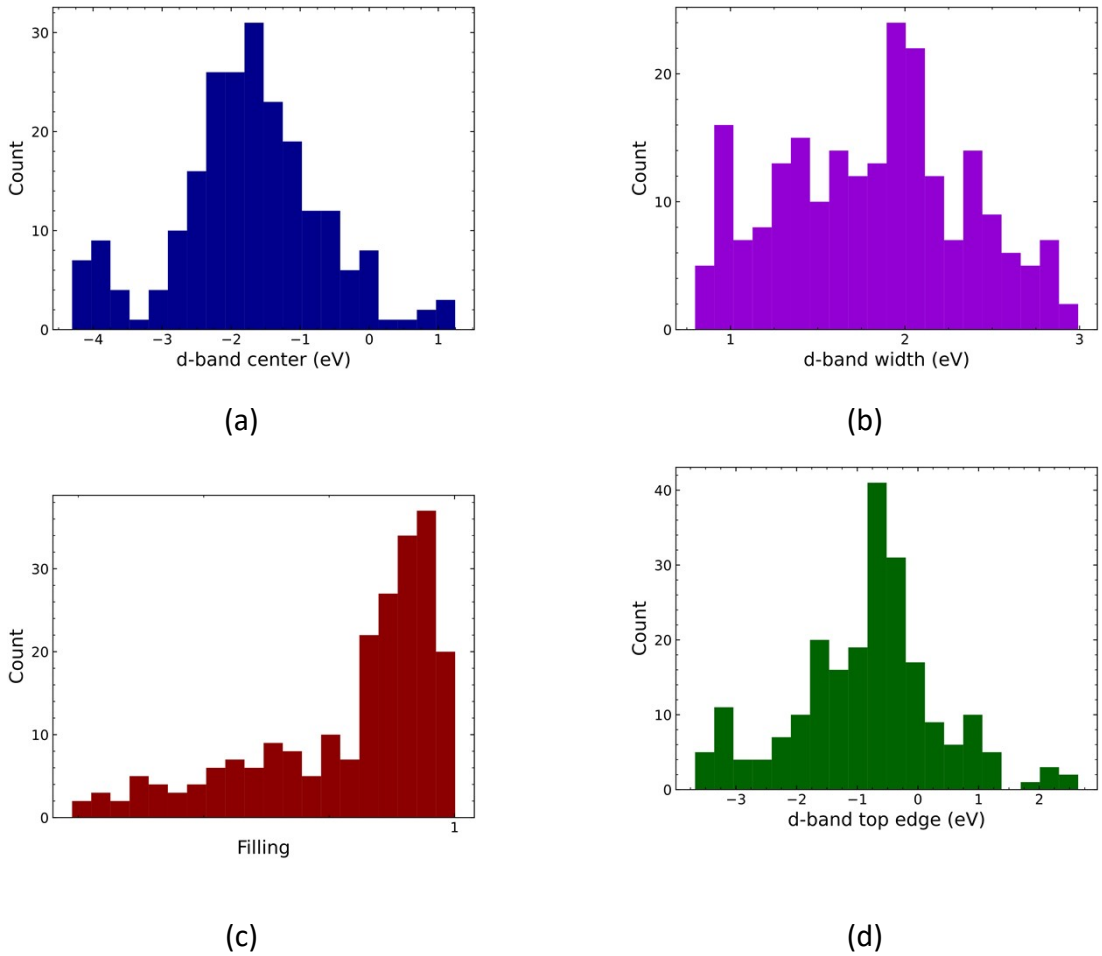
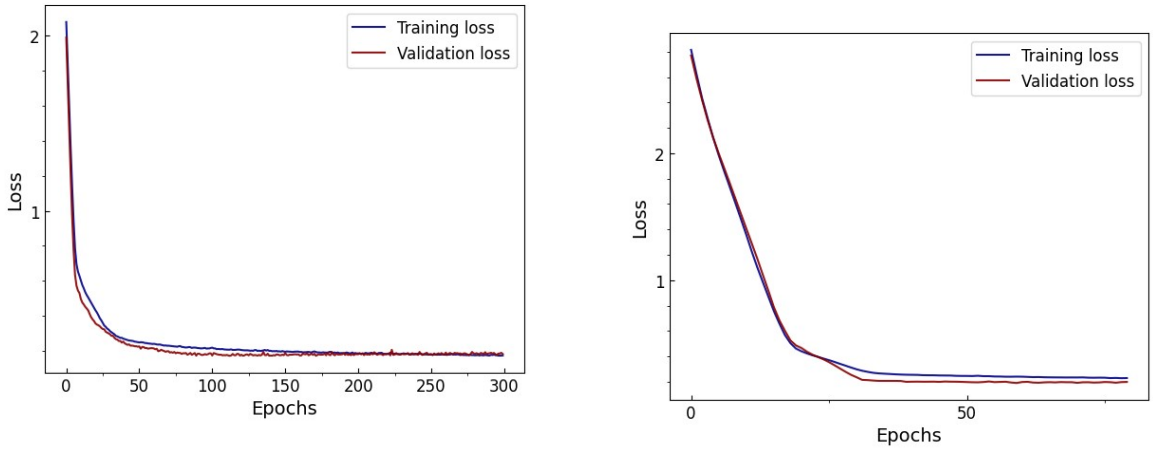


Figure S1: MLA feature distributions, (a) d-band center, (b) d-band width, (c) d-band filling and (d) top-edge of the d-band.



(a) (b)

Figure S2: Training and validation loss curves of the ANN: (a) without early stopping, showing signs of overfitting, and (b) with early stopping, where training is stopped at epoch 80 to prevent overfitting.

4 Model performance

All models were initially trained using all available features discussed in the text but following feature importance analysis (Fig. S1), only the important features, namely the d-band center, filling and top edge (dtop), and the electron affinity, were used to reduce the model complexity and evaluate the performance of the SVR, RF and ANN models. Using only four important features, the ANN MAE for train and test sets increased to 0.21 and 0.28 eV, respectively, which remains relatively small. However, the RMSE increased to 0.33 and 0.44 eV, respectively. We also evaluated the ANN performance using only elemental descriptors and the RMSE for train and test sets increased to 0.42 and 0.54 eV, which clearly shows the effect of electronic descriptors on the model accuracy. Another common approach for feature elimination/filtering, based on feature variance, was found to be questionable for the problem at hand; d-band filling has the lowest variance and should be eliminated based on the feature variance approach, while feature importance analysis clearly identifies d-band filling as one of the important features.

Table S4. Absolute variance of training features used in the MLA model.

Feature	Variance	Unit [a]
Center	1.16	eV
Filling	0.03	-
Width	0.28	eV
d-band top edge	1.47	eV
Ionization potential	0.45	eV
Electron affinity	0.27	eV
Electronegativity	0.03	-
Workfunction	0.34	eV
Atomic radius	0.01	Å

[a] Units refer to the original features; variance is expressed in squared units (e.g., eV^2, $Å^2$)

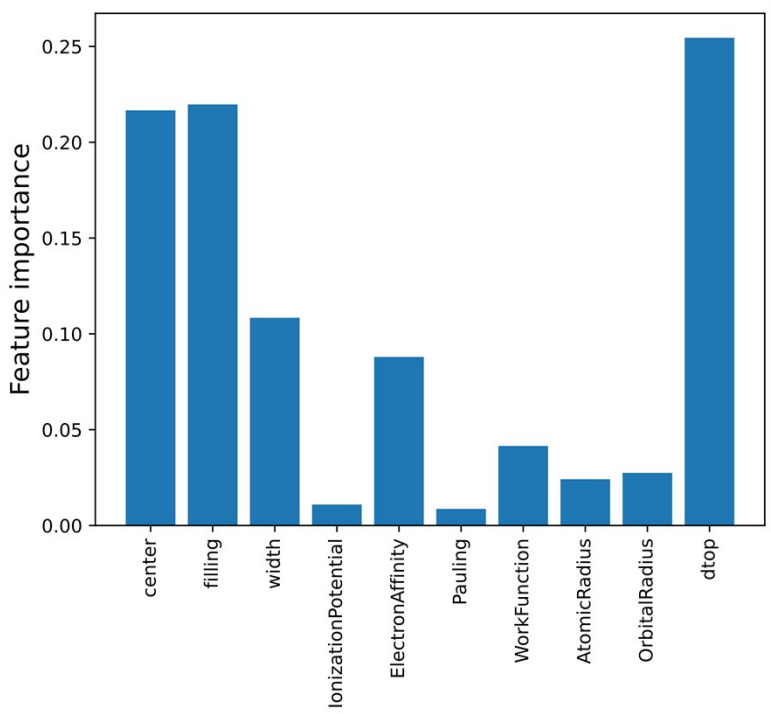


Figure S3: Relative feature importance of the train set used for MLA training. Values reflect the influence of each input feature on prediction performance.

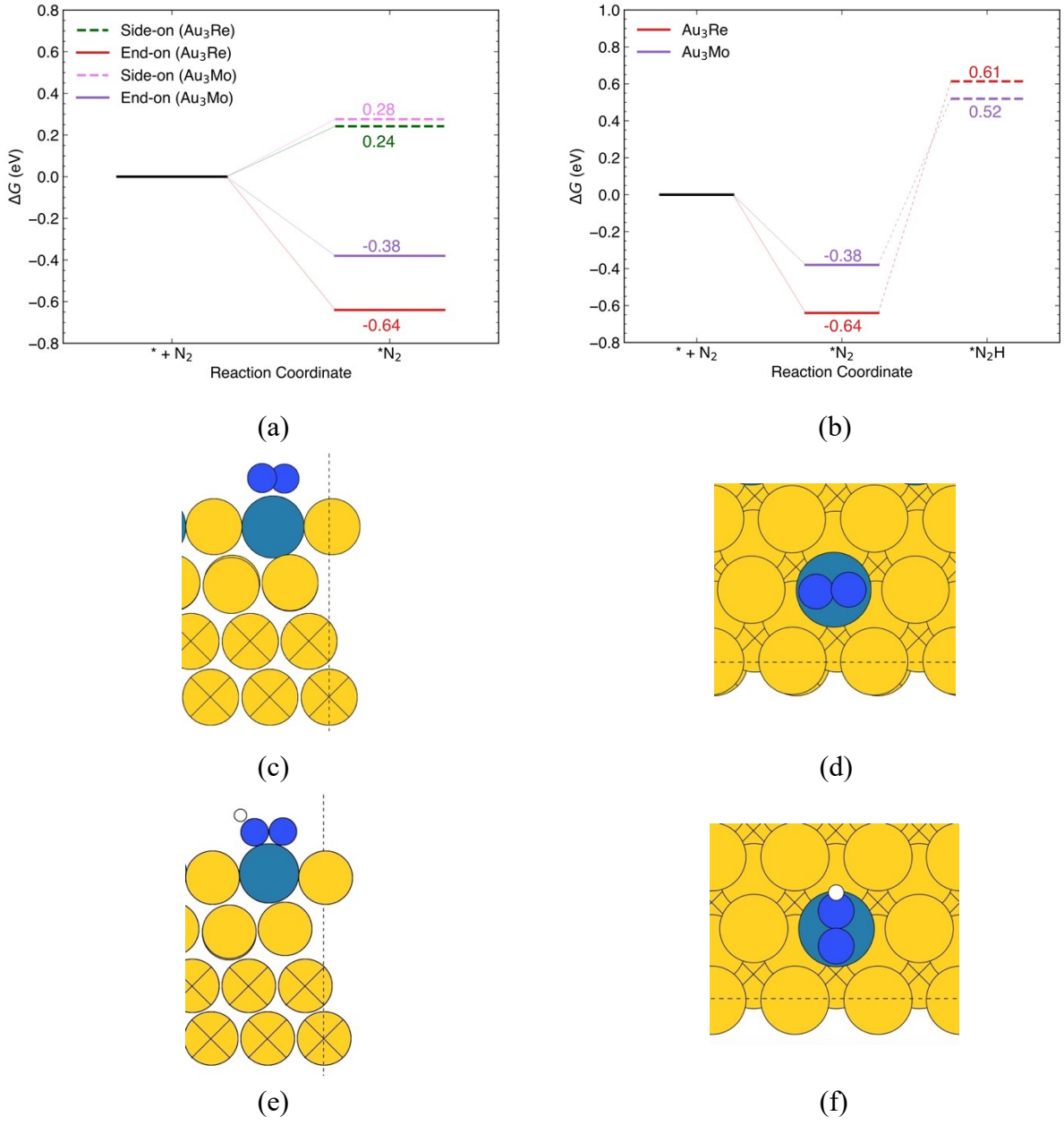


Figure S4: Features of the distal *vs* enzymatic pathway for the NRR over Au@Au_3Re and Au@Au_3Mo. (a) adsorption free energy of N_2 with end-on and side-on configurations, (b) first NRR step of the enzymatic pathway with $*N_2$ and $*N_2H$ side-on configuration (to be compared with Fig. 6 of the main text for the distal pathway with end-on configurations); (c, e) side and (d, f) top view of intermediates.

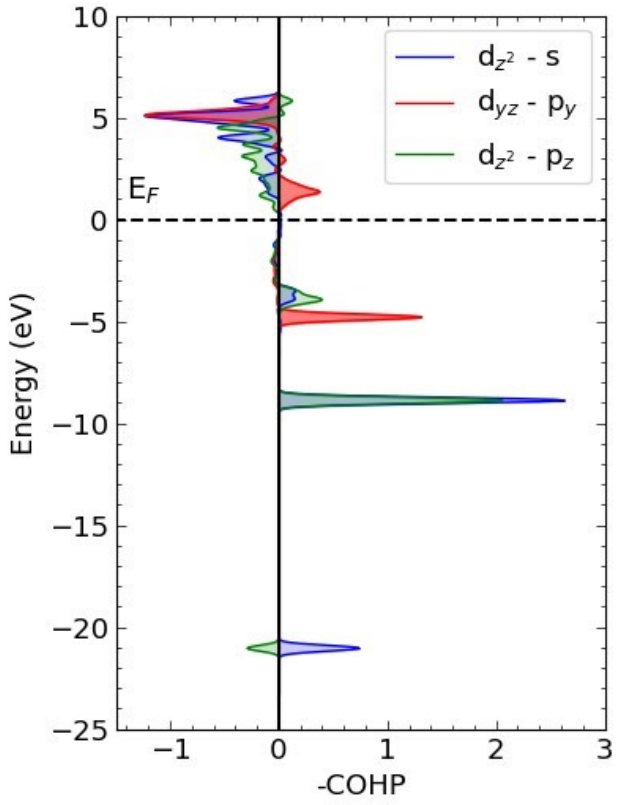


Figure S5: Crystal orbital Hamilton population for the interactions of Re in Au_3Re and adsorbed N_2. The Fermi level is indicated by the dashed line.

5 Formation energy

The formation energy, which provides a measure of the stability of a compound or mixture with respect to its pure elemental components, is calculated for the alloys as:

$$E_{Formation} = E_{A_{(4-x)}B_x} + x(E_A - E_B) - E_{A\,(slab)} \quad \text{(S6)}$$

where $E_{A_{(4-x)}B_x}$, E_A and E_B are the electronic energies of the four-layer 2 × 2 alloy slab, isolated host and guest metal atoms, respectively, x is the number of guest metals in the alloy, and $E_{A\,(slab)}$ is the electronic energy of the four-layer 2 × 2 pure host slab (16 host atoms). The second term in the right-hand side of the equation accounts for guest (B) substitution of a host (A) atom in the slab. For example, the so-defined formation energy of Au@Au_3Re is:

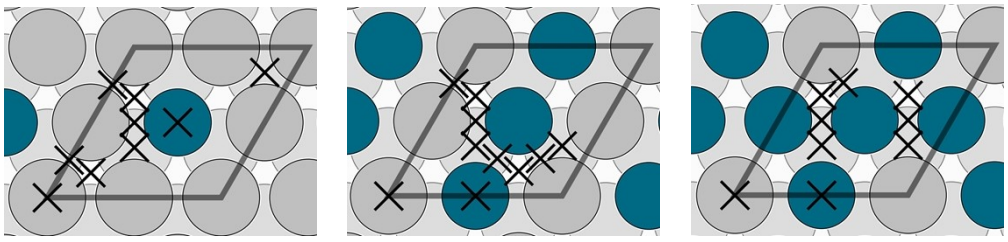
$$E_{Formation} = E_{Au@Au_3Re} + E_{Au} - E_{Re} - E_{Au(slab)} \quad \text{(S7)}$$

where $E_{Au@Au3Re}$, E_{Au}, E_{Re} and $E_{Au(slab)}$ are the energies of the alloy slab, isolated Au and Re atoms, and pure Au slab, respectively."

Table S5 Formation energy of surface alloys in eV.[a]

Structure	$E_{Formation}$
Au_3Re	-5.37
Au_3Mo	-3.20

[a]The formation energies are calculated using equation S6.



(a) (b) (c)

Figure S6: Adsorption sites considered in the search of minimum energy configurations of each intermediate overA_3B, A_2B_2 and AB_3 surface alloys.

Table S6: Adsorption free energy of NRR intermediates over selected solid supports (eV)

Structure	ΔG_{*N2}	ΔG_{*N2H}	ΔG_{*N2H2}	ΔG_{*N}	ΔG_{*NH}	ΔG_{*NH2}	ΔG_{*NH3}	ΔG_{*H}
Au_3Re	-0.64	-0.15	0.09	-1.13	-1.04	-0.69	-0.57	-0.26
Au_3Mo	-0.37	0.29	0.22	-0.74	-0.91	-0.59	-0.46	0.08
Mo	-0.07	0.10	0.08	-1.39	-1.37	-0.69	-0.13	-0.40
Re	-0.16	0.43	0.44	-1.28	-1.20	-0.25	-0.09	-0.46

The adsorption free energies (ΔG) are calculated using Equation 2 of the main text.

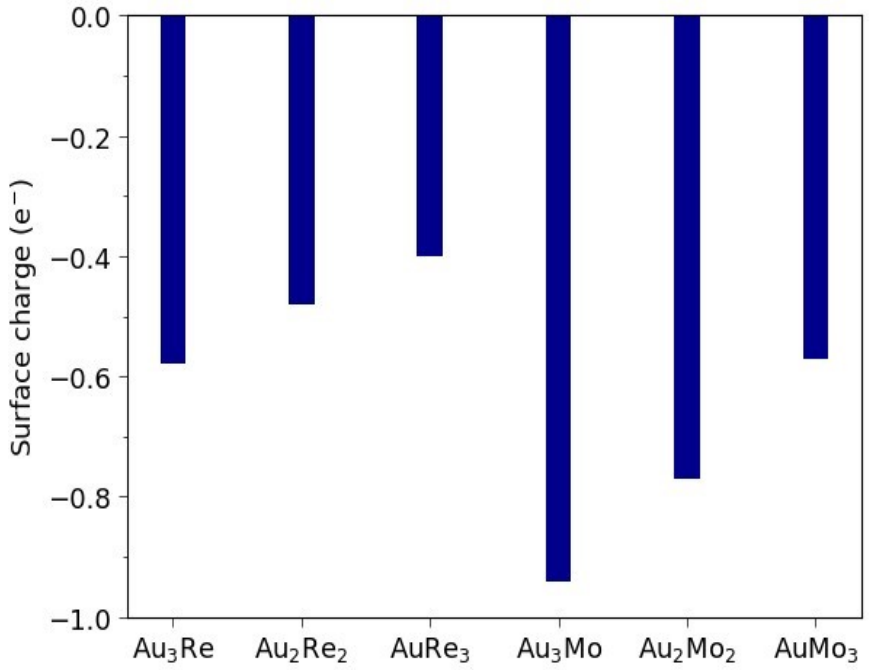
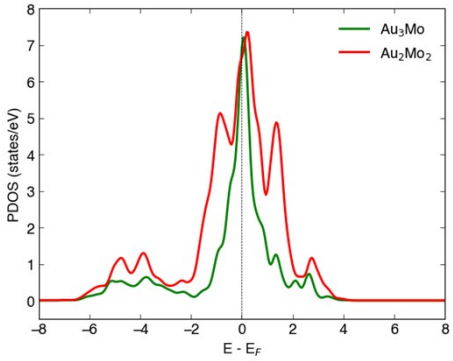
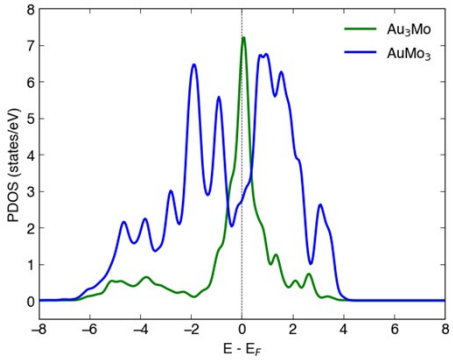


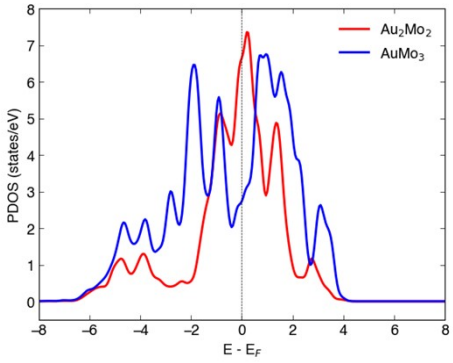
Figure S7: Net charge on active sites (Re in Au_3Re and Mo in Au_3Mo) relative to that of in pure Re (001) and Mo (100)



(a)



(b)



(c)

Figure S8: PDOS of the Mo active site in Au-Mo surface alloys; Au_3Mo (green), Au_2Mo2 (red) and $AuMo_3$ (blue); (a) Au_3Mo and Au_2Mo_2, (b) Au_3Mo and $AuMo_3$ and Au_2Mo_2 and $AuMo_3$

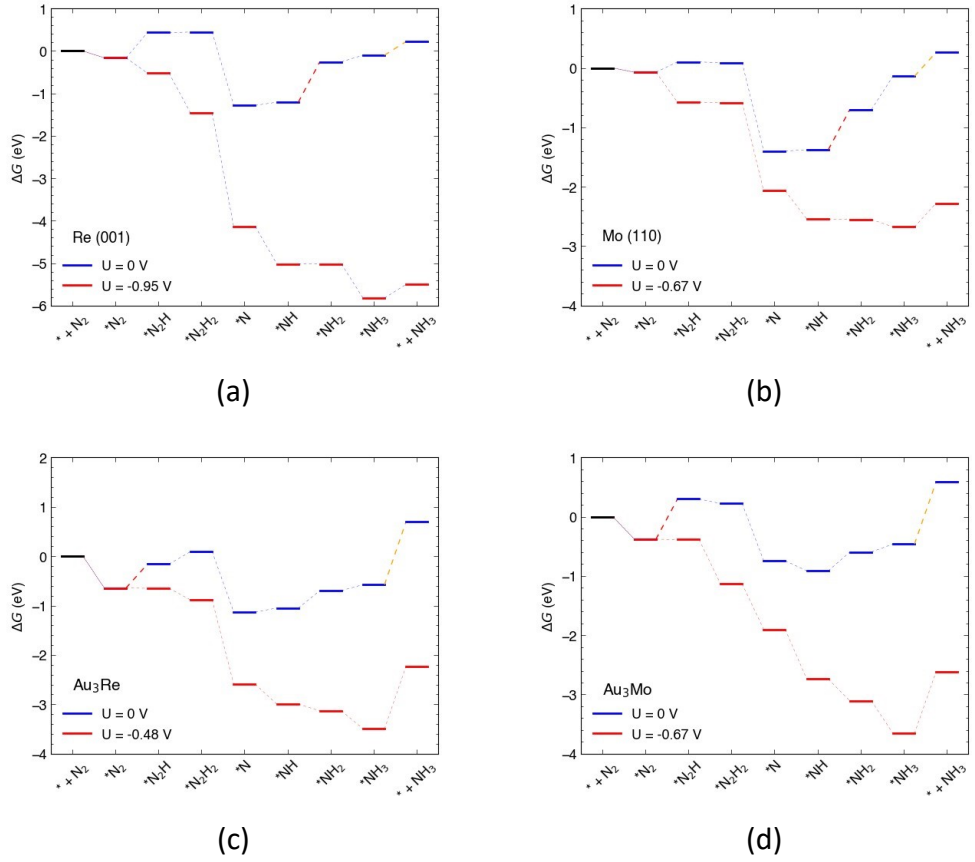


Figure S9: Free energy profile of the NRR distal pathway on (a) Re (001), (b) Mo (110), (c) Au_3Re (111) and (d) Au_3Mo (111) surfaces.

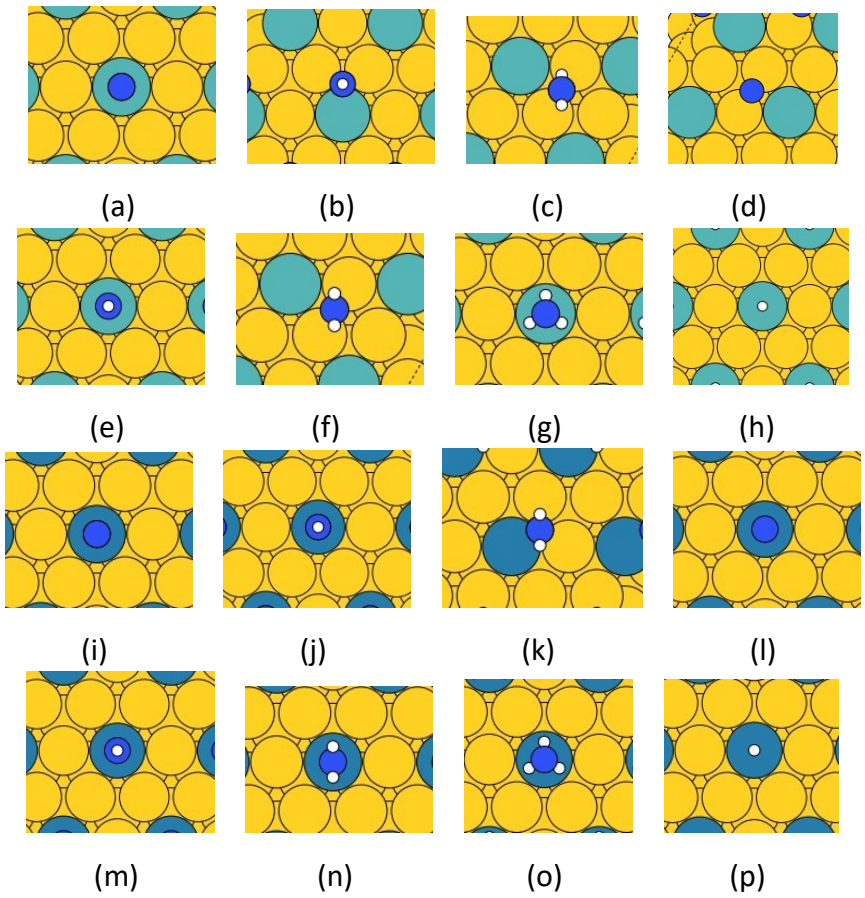


Figure S10: Minimum energy configuration of NRR intermediates over (a-h) Au_3Mo and (i-p) Au_3Re (top view).

Table S7: Adsorption free energies of N_2 and H along with selectivity factor (eV)

structure	ΔG_{*N2}	ΔG_{*H}	PLS	$U_{limiting}$	ΔG_S
Au_3Re	-0.64	-0.26	b	-0.48	-0.38
Au_3Mo	-0.37	0.08	b	-0.67	-0.46
Mo	-0.07	-0.40	f	-0.67	0.33
Re	-0.16	-0.46	f	-0.95	0.30

PLS: potential limiting step. Step b ($*N_2 \rightarrow *N_2H$) and step f ($*NH \rightarrow *NH_2$) are illustrated in the NRR distal mechanism of Equation 1 of the main text.

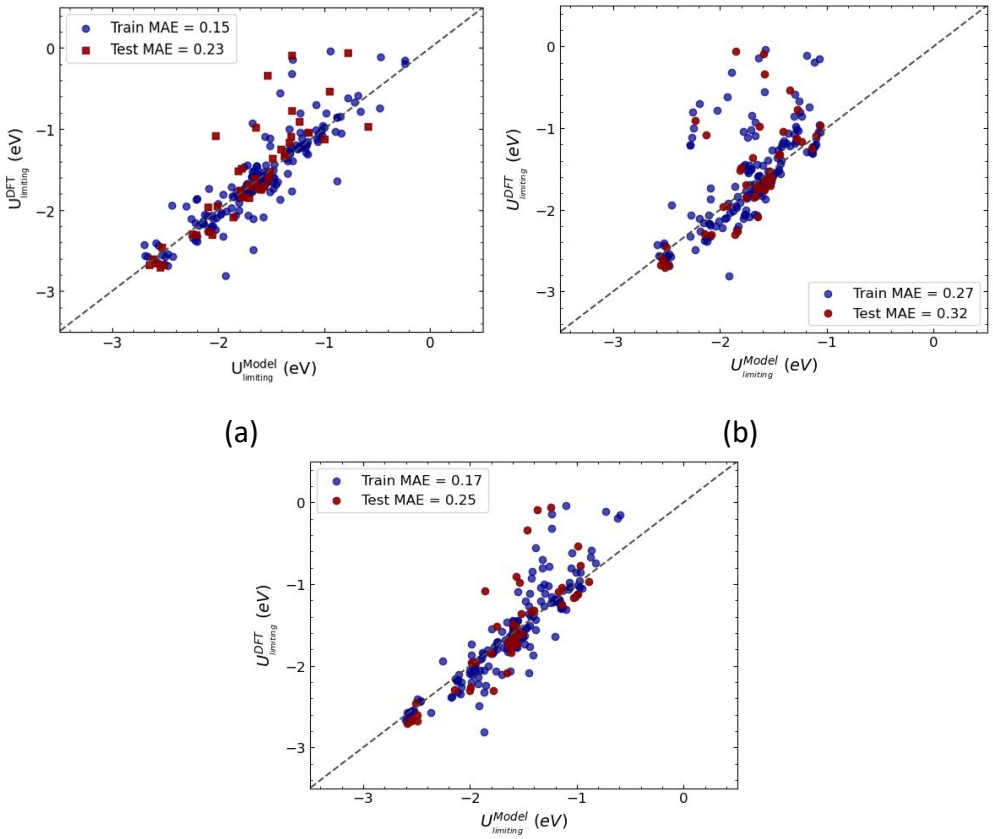


Figure S11: MLA-predicted vs DFT-calculated NRR limiting potentials for (a) ANN, (b) SVR and (c) RF for train (blue) and test (red) datasets

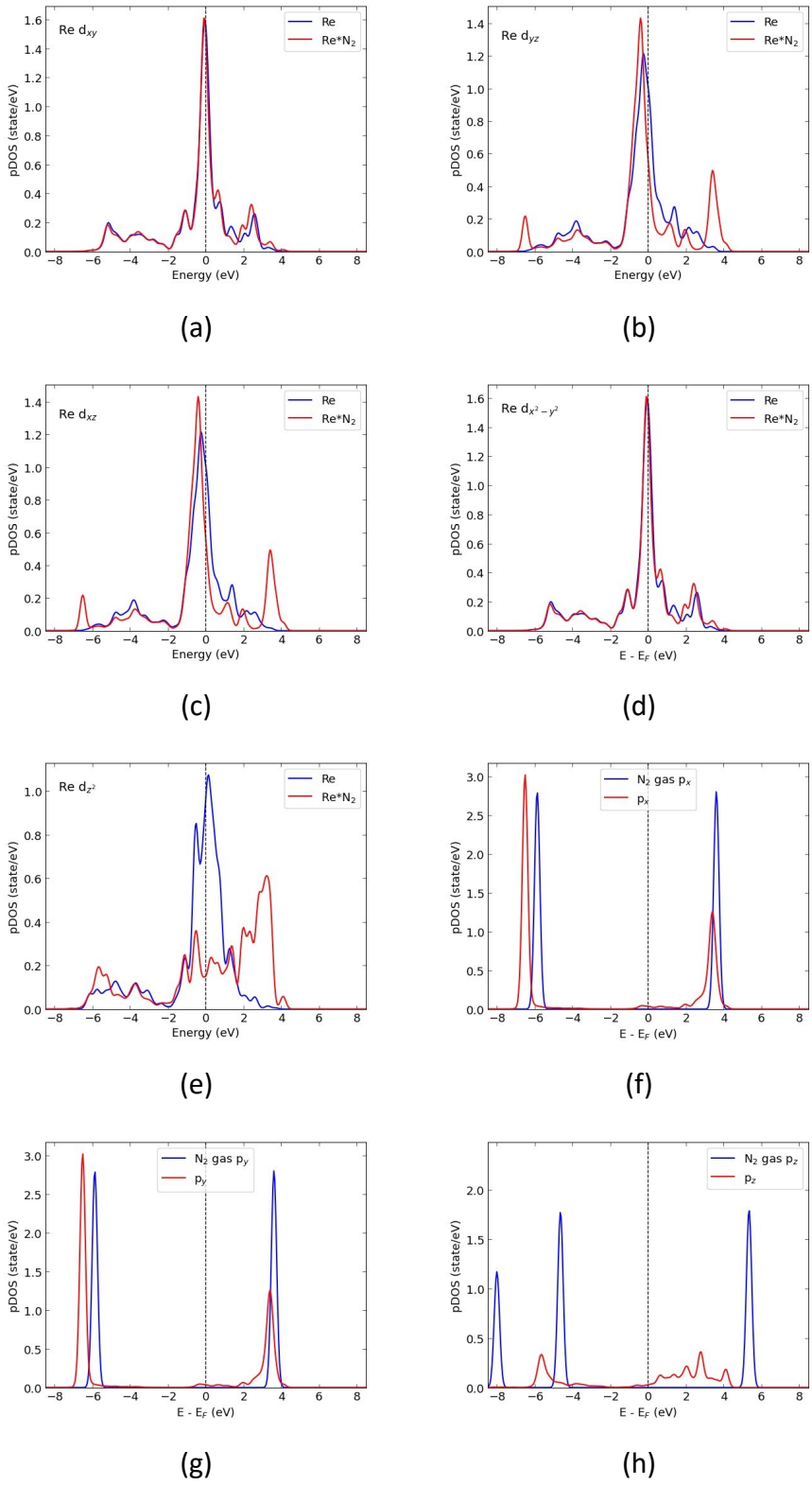
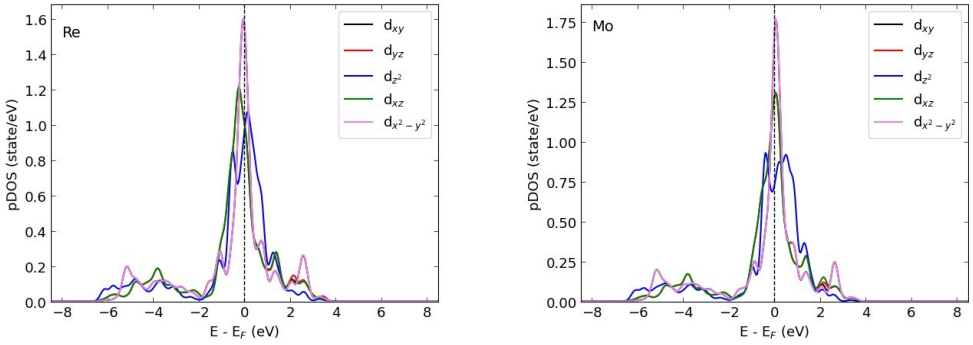
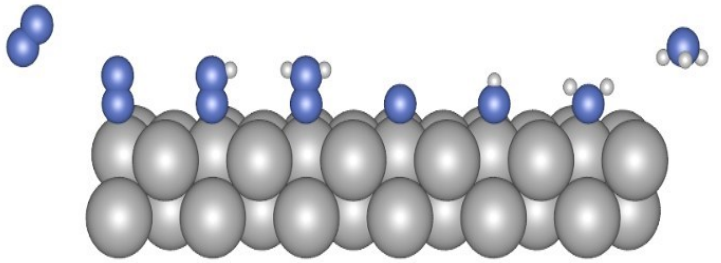


Figure S12: PDOS of N_2 before (blue) and after adsorption (red) on Au_3Re.

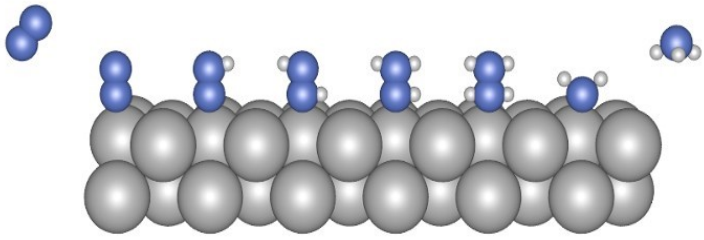


(a) (b)

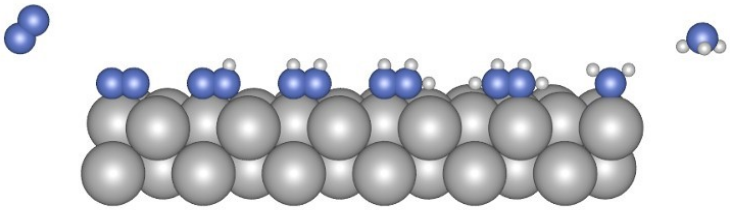
Figure S13: PDOS of (a) Re and (b) Mo active sites in Au_3Re and Au_3Mo, respectively. The dashed line represents the Fermi level.



(a)



(b)



(c)

Figure S14: Hydrogenation sequence of NRR intermediates for the (a) distal, (b) alternating and (c) enzymatic mechanism. Nitrogen, hydrogen and the catalyst surface are represented as blue, white and gray balls, respectively.

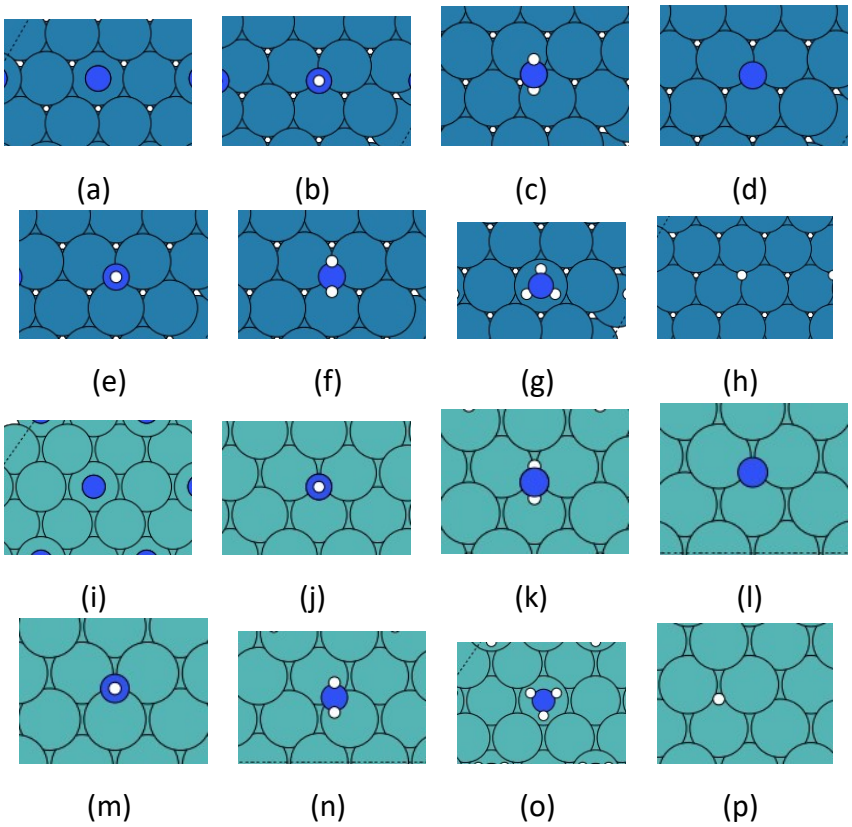


Figure S15: Minimum energy configuration of NRR intermediates over (a-h) Re (001) and (i-p) Mo (110) surfaces (top view).

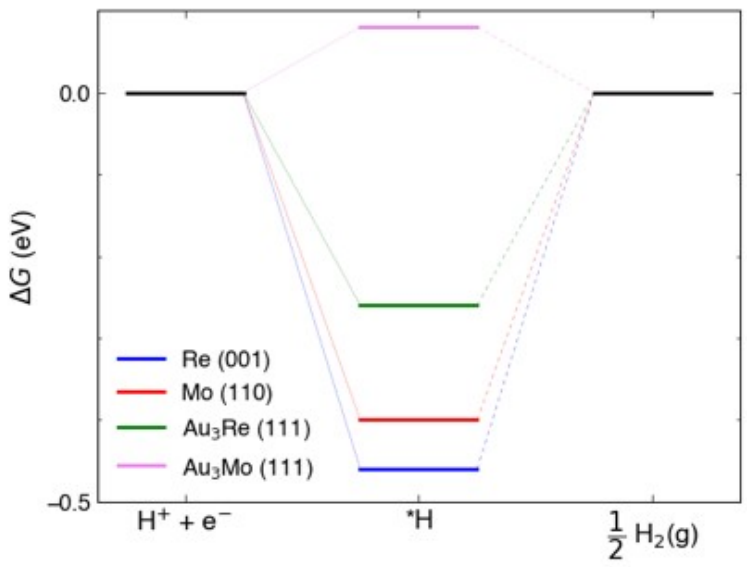


Figure S16: Free energy profile of the HER Volmer-Heyrovsky pathway over Au3Mo (111) (violet), Au3Re (111) (green), Mo (110) (red) and Re (001) (blue) surfaces.

Notes and references

(1) Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jonsson, H. “Origin of the overpotential for oxygen reduction at a fuel-cell cathode.” *The Journal of Physical Chemistry B* **2004**, *108*, 17886–17892.

(2) Hammer, B.; Nørskov, J. K. “Electronic factors determining the reactivity of metal surfaces.” *Surface science* **1995**, *343*, 211–220.

(3) Hammer, B.; Morikawa, Y.; Nørskov, J. K. “CO chemisorption at metal surfaces and overlayers.” *Physical review letters* **1996**, *76*, 2141.

(4) Zheng L. et al. "High-throughput screening of bimetallic catalysts enabled by machine learning." *Journal of Materials Chemistry A* **2017,** 5, 24131-24138

(5) Tran, K.; Ulissi Z. K. "Active learning across intermetallics to guide discovery of electrocatalysts for CO2 reduction and H2 evolution." *Nature Catalysis* **2018**, 1, 696-703.

(6) Guolin C. et al. "Electronic descriptors for designing high-entropy alloy electrocatalysts by leveraging local chemical environments." *Nature Communications* **2025**, 16, 1251.