

Screening of Single-Atom Catalysts for CO₂ Electroreduction to CH₄ by DFT Calculations and Machine Learning

Jingzhou Zhou,^a Zhengyu Gu,^a Xiaobing Wang,^a Xiuyun Zhang,^c Yixin Ouyang^{a,b*} and Li Shi^{a,b*}

^aState Key Laboratory of Flexible Electronics (LoFE) & Institute of Advanced Materials (IAM),
Nanjing University of Posts and Telecommunications Nanjing 210023, China

^bKey Laboratory of Quantum Materials and Devices (Southeast University), Ministry of Education,
Nanjing 21189, China

^cCollege of Physics Science and Technology, Yangzhou University, Yangzhou 225002, China.

*Corresponding authors. E-mails: iamlshi@njupt.edu.cn, iamyxouyang@njupt.edu.cn

Computational methods and models

All calculations of energies and structural optimizations were conducted at the spin-polarized DFT level using the Vienna ab initio simulation package (VASP).^{1,2} The projector augmented wave (PAW)³ potentials were employed alongside the generalized gradient approximation parameterized by Perdew, Burke, and Ernzerhof (GGA-PBE)^{4,5} for the exchange-correlation functional. The ICOHP analysis was performed by LOBSTER.⁶ Additionally, To adequately capture the interactions between reaction intermediates and catalysts, the interactions were properly described through the adoption of Grimme's semiempirical DFT-D2 scheme⁷ for dispersion correction. The calculations employed $1 \times 1 \times 1$ k-points Monkhorst-Pack mesh sampling in the surface Brillouin zone. The plane-wave cutoff energy was optimized to a value of 500 eV. The convergence criteria for electronic self-consistent iteration were established as 10^{-5} eV, while the ionic relaxation loop was constrained to forces below -0.02 eV/Å. The Bader charge analysis was employed to evaluate the partial charge densities. The thermal stability of metal clusters was assessed through ab initio molecular dynamics (AIMD) simulations conducted under the NVT ensemble,⁸ utilizing a time step of 1 fs and a total duration of 5 ps. The VASPsol implicit solvent model is used to simulate solvent effects, water was treated as a continuous medium with a dielectric constant of 78.4.^{9,10}

The Gibbs free energy of formation for each elementary step was calculated as $\Delta G = \Delta E + \Delta E_{\text{ZPE}} - T\Delta S$, where ΔE , ΔE_{ZPE} , and ΔS are the differences of DFT total energy, zero point energy, and entropy, respectively. The free energy changes at each electrochemical step involving a proton-electron transfer were computed based on computational hydrogen electrode (CHE) model.^{11,12} The free energy diagram of CO₂ reduction was obtained based on the computational hydrogen electrode method. The ΔG_{max} is the Gibbs free energy of formation for the potential-determine step of CO₂ reduction.

Computational details

1. Formation energy and dissolution potential of SACs

The formation energy refers to the difficulty of synthesizing a catalyst from substrate, namely

thermodynamical stability. The dissolution potential represents electrochemical stability, meaning whether the metal active center will fall off the structure and dissolve into the electrolyte. The formation energy and dissolution energy of SACs are defined as:

$$E_f = E_{TM@C_5N} - E_{C_5N} - E_{TM}$$

$$U_{diss} = U_{diss}^0(metal, bulk) - \frac{E_f}{ne}$$

where $E_{TM@C_5N}$, E_{C_5N} and E_{TM} are the total energies of SACs, substrate and metal atom, $U_{diss}^0(metal, bulk)$ and n are the standard dissolution potential of bulk metal and the number of electrons involved in the dissolution, respectively.

2. Free energy calculation based on CHE

The free energy of elementary reactions of CO₂ reduction is calculated based on the computational hydrogen electrode (CHE), that is, the energy of a H⁺/e⁻ pair is equal to half of the gaseous hydrogen (0.5H₂) at an equilibrium potential. The corrections of zero-point energy (E_{ZPE}), temperature (T) and entropy (S) are introduced into the DFT-calculated total energy (EDFT) to determine the free energy of elementary reactions:

$$\Delta G = \Delta E + \Delta E_{ZPE} - T\Delta S$$

where the ΔE_{ZPE} and $T\Delta S$ for each reaction intermediates can be calculated by the following equations, respectively:

$$\Delta E_{ZPE} = \frac{1}{2} \sum_i h\nu_i$$

$$-TS = k_B T \sum_i \ln \left(1 - e^{-\frac{h\nu_i}{k_B T}} \right) - \sum_i h\nu_i \left(\frac{1}{e^{\frac{h\nu_i}{k_B T}} - 1} \right)$$

In the equations above, h , ν and k_B are Planck constant, vibrational frequencies, and Boltzmann constant, respectively. Besides, only when we calculate the free energy of reaction intermediates, the zero-point energy, entropy, and enthalpy are needed to calculate.

3. Binding energy and limiting potential calculation of SACs

The scaling relations on transition-metal surface are mainly reported by Norskov et al, hence,¹³ in the calculation of scaling relations of binding energies, to compare the scaling relations of single atom catalysts with that of transition-metals, we used the same calculation equations reported by Norskov et al. The binding energies of *CO, *COOH, and *CHO are calculated through

$$E_b(CO) = E(*CO) + \Delta E_{sol}(CO) - E_* - E(CO(g))$$

$$E_b(CHO) = E(*CHO) + \Delta E_{sol}(CHO) - E_* - E(CO(g)) - \frac{1}{2}E(H_2)$$

where $E_b(CO)$ and $E_b(CHO)$ are the binding energies of *CO and *CHO, respectively. The ΔE_{sol} indicates a solvation correction to binding energy. For *CO and *CHO, the solvation correction is -0.1 eV and -0.1 eV, respectively.

The limiting potential (U_L) calculations is defined as:

$$U_L = -\frac{\Delta G_{max}}{e}$$

where ΔG_{max} is the step in the primitive reaction that has the largest free energy barrier.

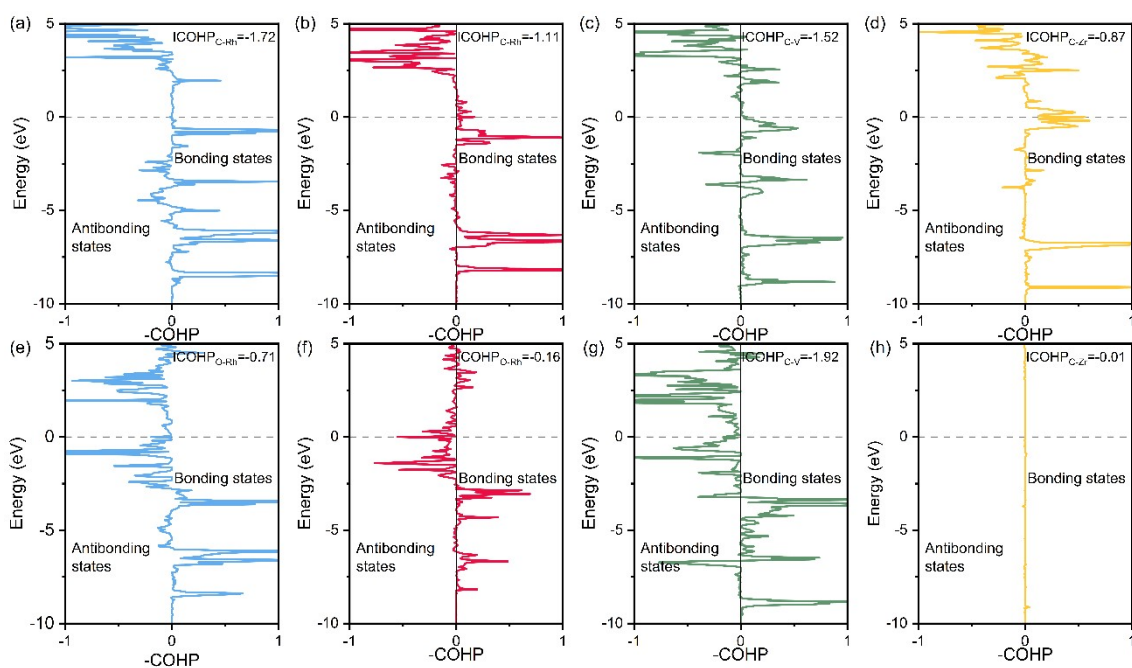


Figure S1. The crystal orbital Hamilton population (COHP) and corresponding ICOHP values of TM-C an TM-O , from left to right including (a) and (e) Rh@C₅N_C1, (b) and (f) Rh@C₅N_2C, (c) and (g) V@C₅N_CN, (d) and (h) Zr@C₅N_CN.

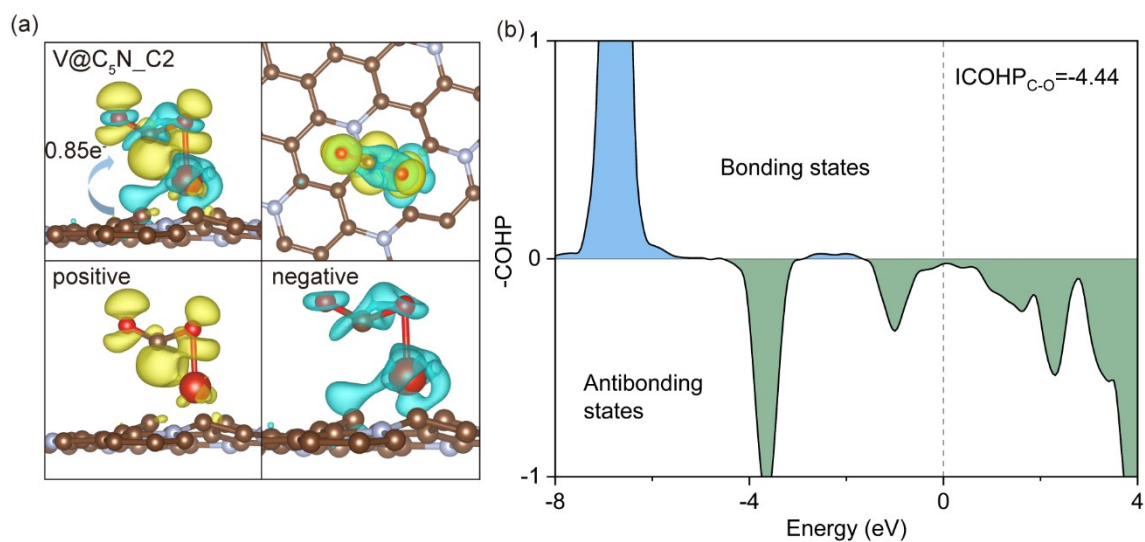


Figure S2. (a) The differential charge density of CO₂ adsorption on V@C₅N₄C₂. The charge accumulation and depletion are depicted by yellow and cyan, respectively. (b) The crystal orbital Hamilton population (COHP) and corresponding ICOHP values of C-O.

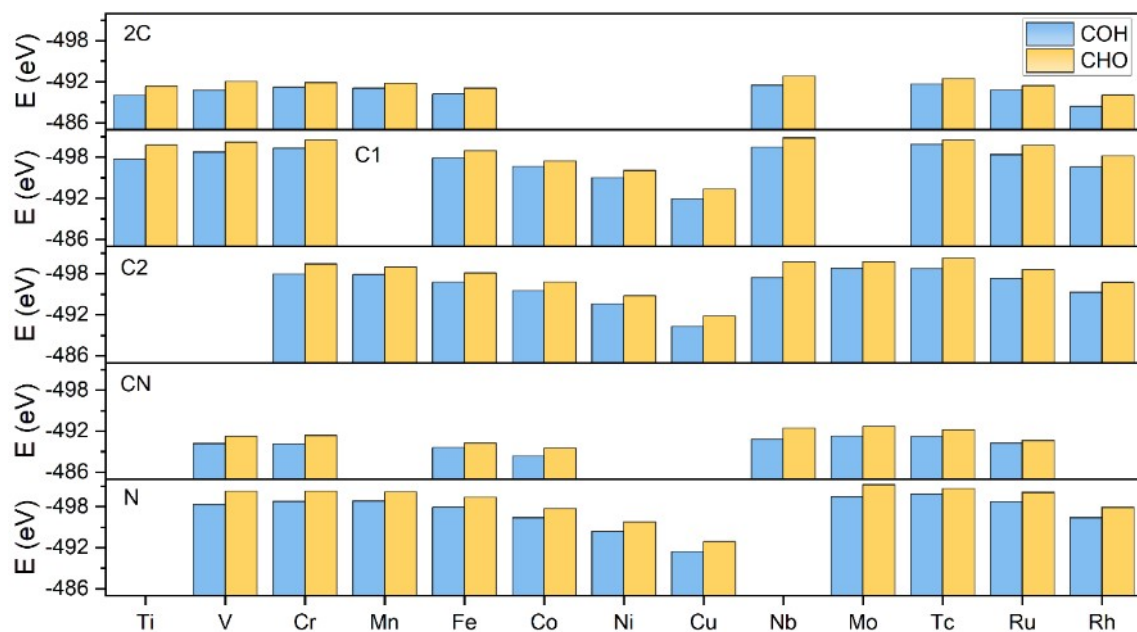


Figure S3. Comparison of *CHO and *COH energies after structural relaxation.

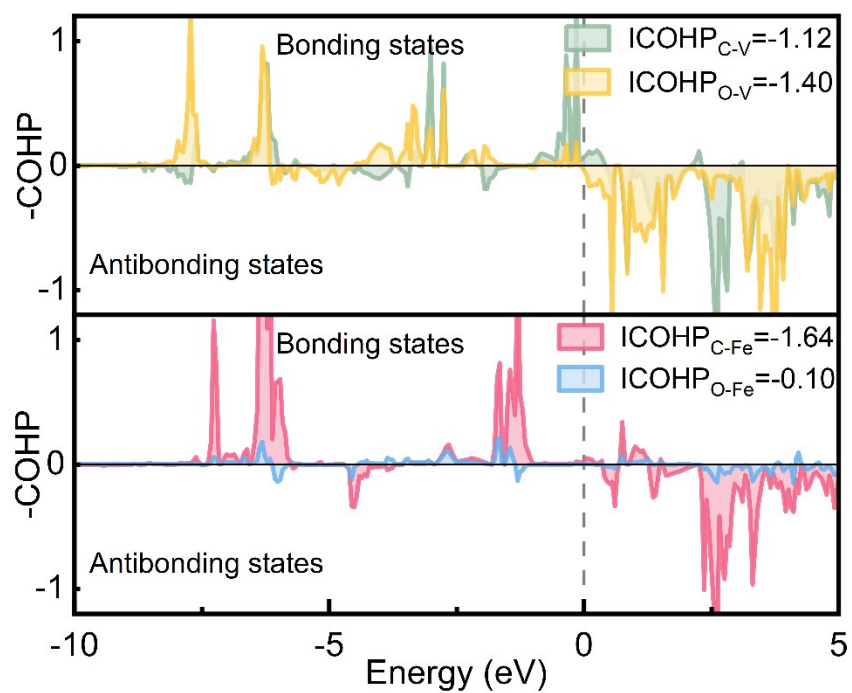


Figure S4. The crystal orbital Hamilton population (COHP) and corresponding ICOHP values of C-V, O-V, C-Fe and O-Fe.

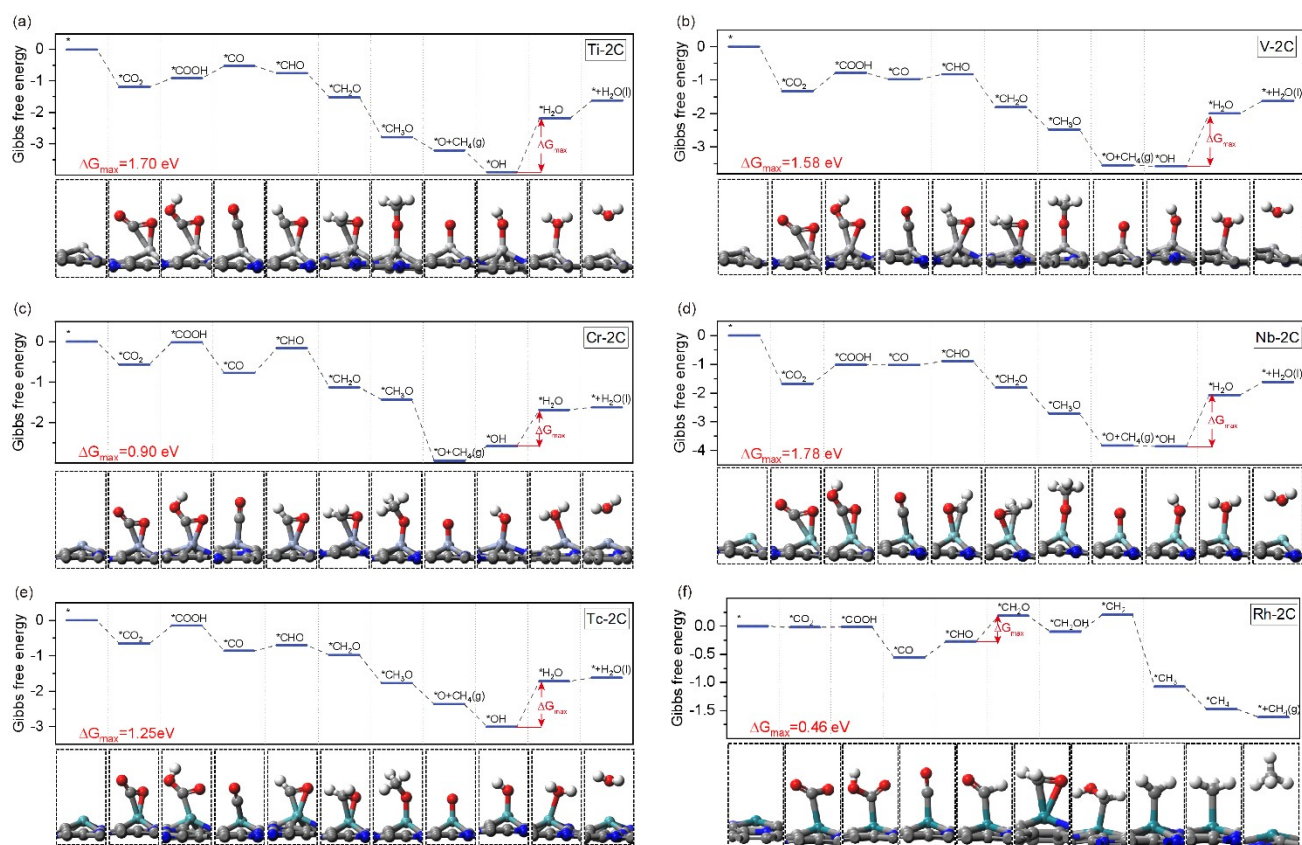


Figure S5. Free energy diagram for the electrochemical reduction of CO_2 on (a) Ti@C₅N₂C, (b) V@C₅N₂C, (c) Cr@C₅N₂C, (d) Nb@C₅N₂C, (e) Tc@C₅N₂C and (f) Rh@C₅N₂C using the computational hydrogen electrode at 0 V vs. RHE. Each step along the horizontal axis represents the transfer of a proton and an electron to the intermediates. The inserts represent all the optimized structures of reaction intermediates.

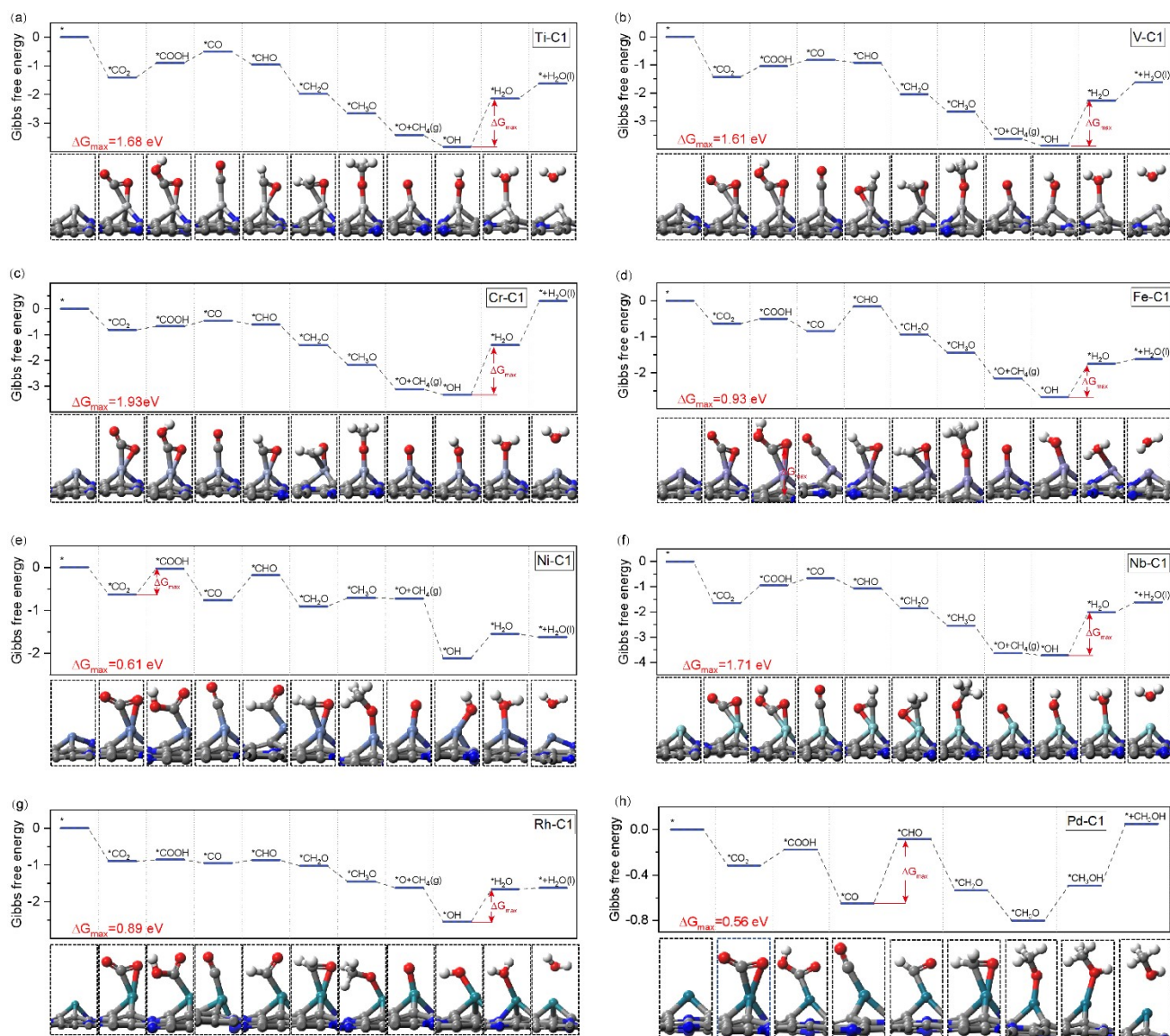


Figure S6. Free energy diagram for the electrochemical reduction of CO₂ on (a) Ti@C₅N₄-C1, (b) V@C₅N₄-C1, (c) Cr@C₅N₄-C1, (d) Fe@C₅N₄-C1, (e) Ni@C₅N₄-C1, (f) Nb@C₅N₄-C1, (g) Rh@C₅N₄-C1 and (h) Pd@C₅N₄-C1 using the computational hydrogen electrode at 0 V vs. RHE. Each step along the horizontal axis represents the transfer of a proton and an electron to the intermediates. The inserts represent all the optimized structures of reaction intermediates.

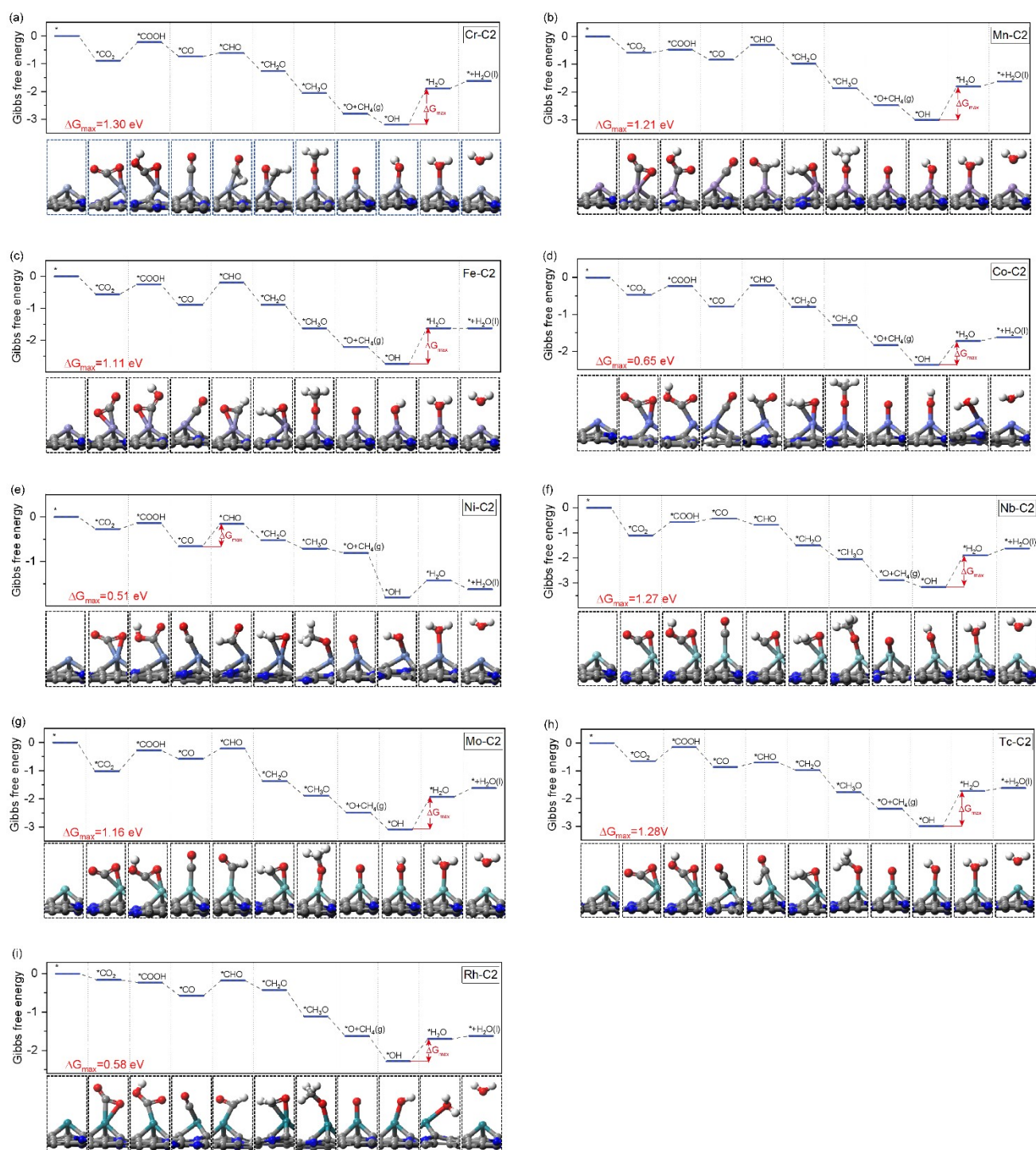


Figure S7. Free energy diagram for the electrochemical reduction of CO₂ on (a) Cr@C₅N₂C₂, (b) Mn@C₅N₂C₂, (c) Fe@C₅N₂C₂, (d) Co@C₅N₂C₂, (e) Ni@C₅N₂C₂, (f) Nb@C₅N₂C₂, (g) Mo@C₅N₂C₂, (h) Tc@C₅N₂C₂ and (i) Rh@C₅N₂C₂ using the computational hydrogen electrode at 0 V vs. RHE. Each step along the horizontal axis represents the transfer of a proton and an electron to the intermediates. The inserts represent all the optimized structures of reaction intermediates.

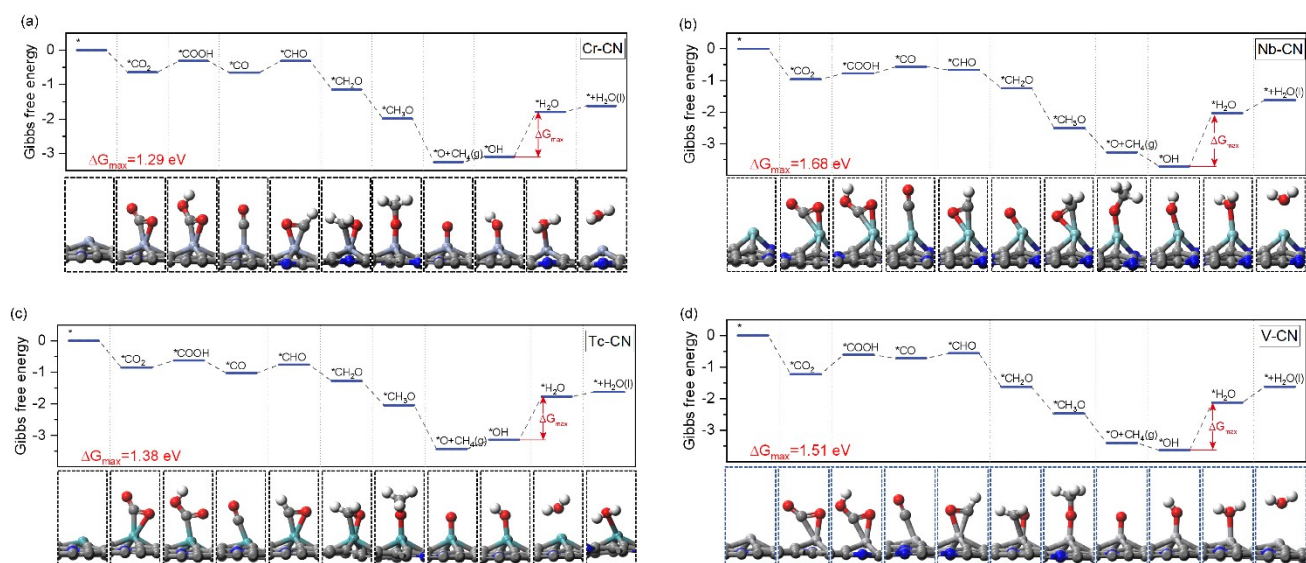


Figure S8. Free energy diagram for the electrochemical reduction of CO_2 on (a) Cr@C₅N₄CN, (b) Nb@C₅N₄CN, (c) Tc@C₅N₄CN and (d) V@C₅N₄CN using the computational hydrogen electrode at 0 V vs. RHE. Each step along the horizontal axis represents the transfer of a proton and an electron to the intermediates. The inserts represent all the optimized structures of reaction intermediates.

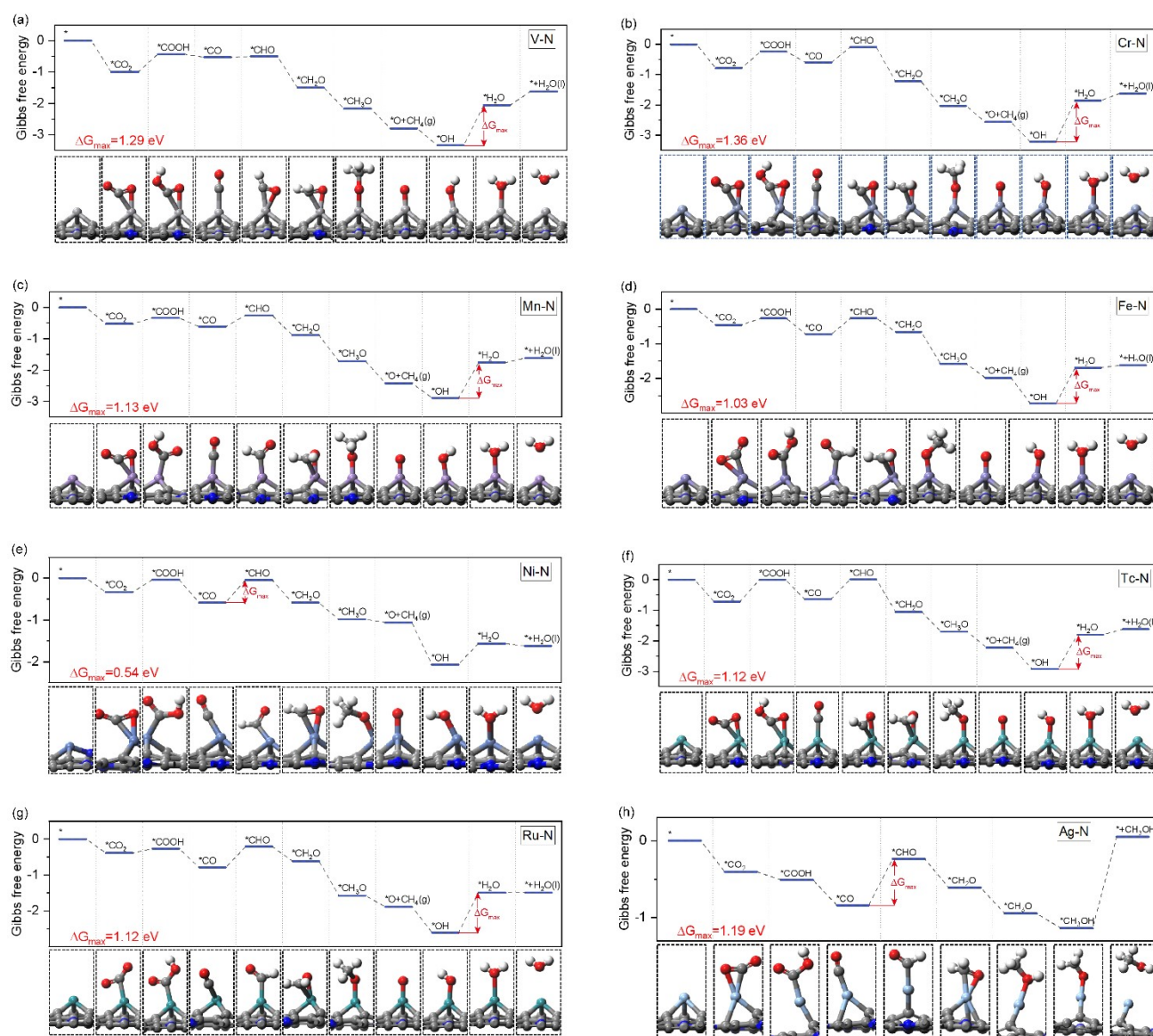


Figure S9. Free energy diagram for the electrochemical reduction of CO_2 on (a) $\text{V@C}_5\text{N}_4$, (b) $\text{Cr@C}_5\text{N}_4$, (c) $\text{Mn@C}_5\text{N}_4$, (d) $\text{Fe@C}_5\text{N}_4$, (e) $\text{Ni@C}_5\text{N}_4$, (f) $\text{Tc@C}_5\text{N}_4$, (g) $\text{Ru@C}_5\text{N}_4$ and (h) $\text{Ag@C}_5\text{N}_4$ using the computational hydrogen electrode at 0 V vs. RHE. Each step along the horizontal axis represents the transfer of a proton and an electron to the intermediates. The inserts represent all the optimized structures of reaction intermediates.

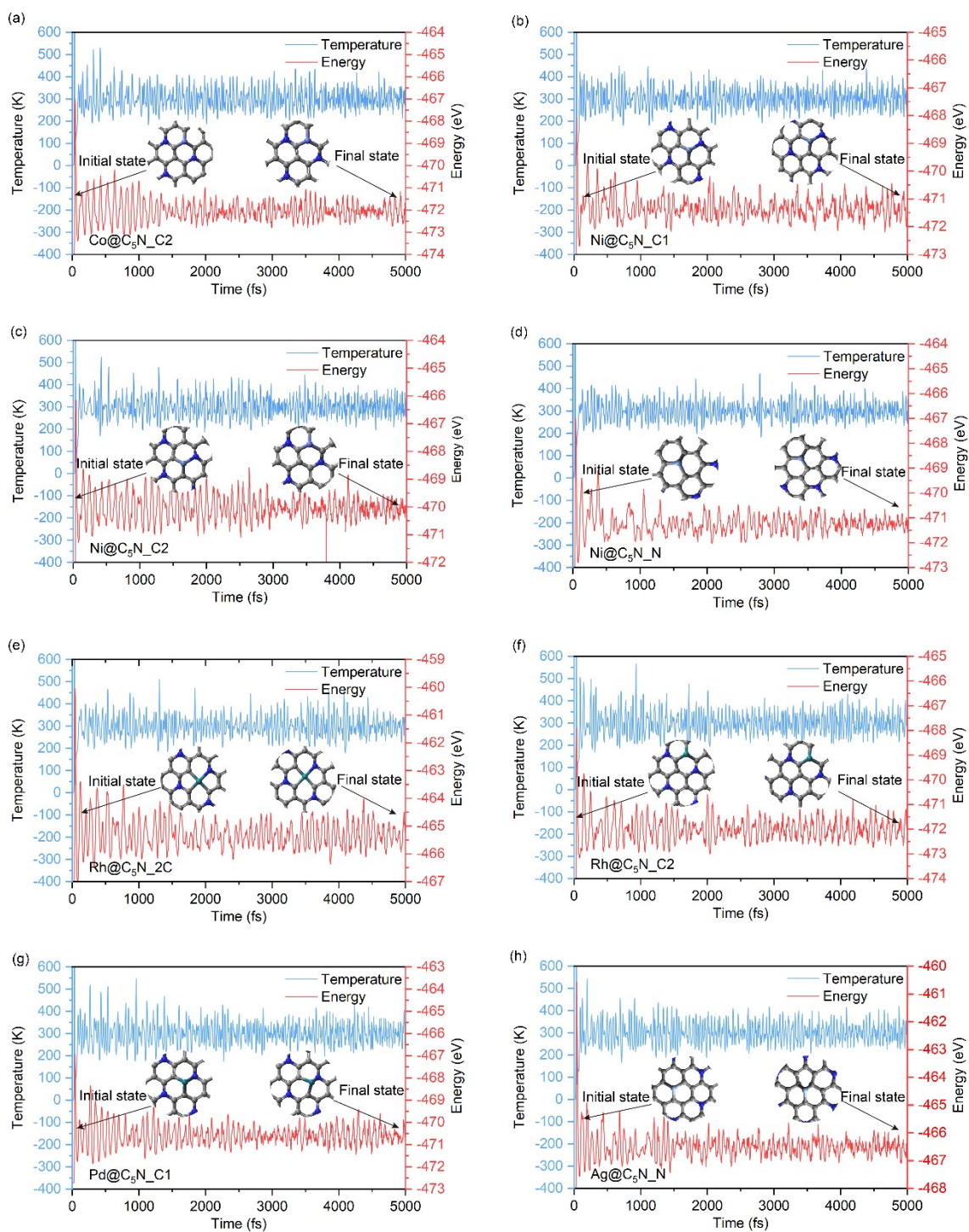


Figure S10. Variations of temperature and energy against the time for AIMD simulations of (a) Co@C₅N_C2, (b) Ni@C₅N_C1, (c) Ni@C₅N_C2, (d) Ni@C₅N_N, (e) Rh@C₅N_2C, (f) Rh@C₅N_C2, (g) Pd@C₅N_C1 and (h) Ag@C₅N_N; the simulation is run at 300 K for 5 ps. Inserts are top views of the snapshot of atomic configuration.

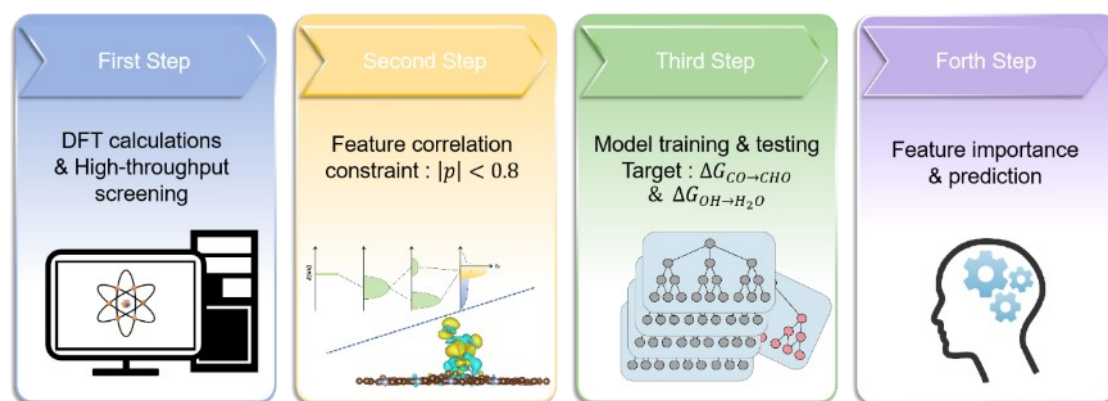


Figure S11. A four-step flowchart for machine learning.

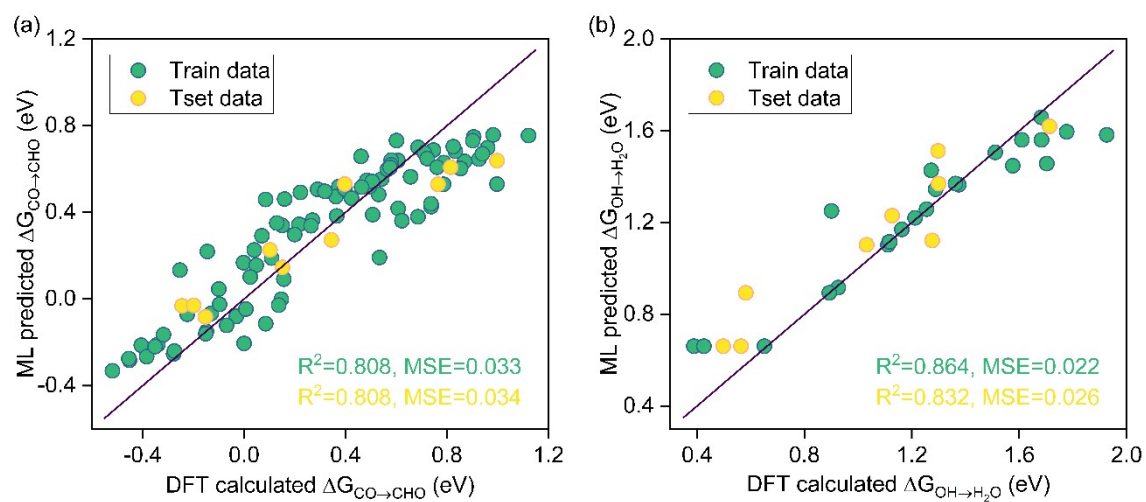


Figure S12. (a) Comparison between DFT and RF-predicted $\Delta G_{CO \rightarrow CHO}$. (b) Comparison between DFT and XGBoost-predicted $\Delta G_{OH \rightarrow H_2O}$.

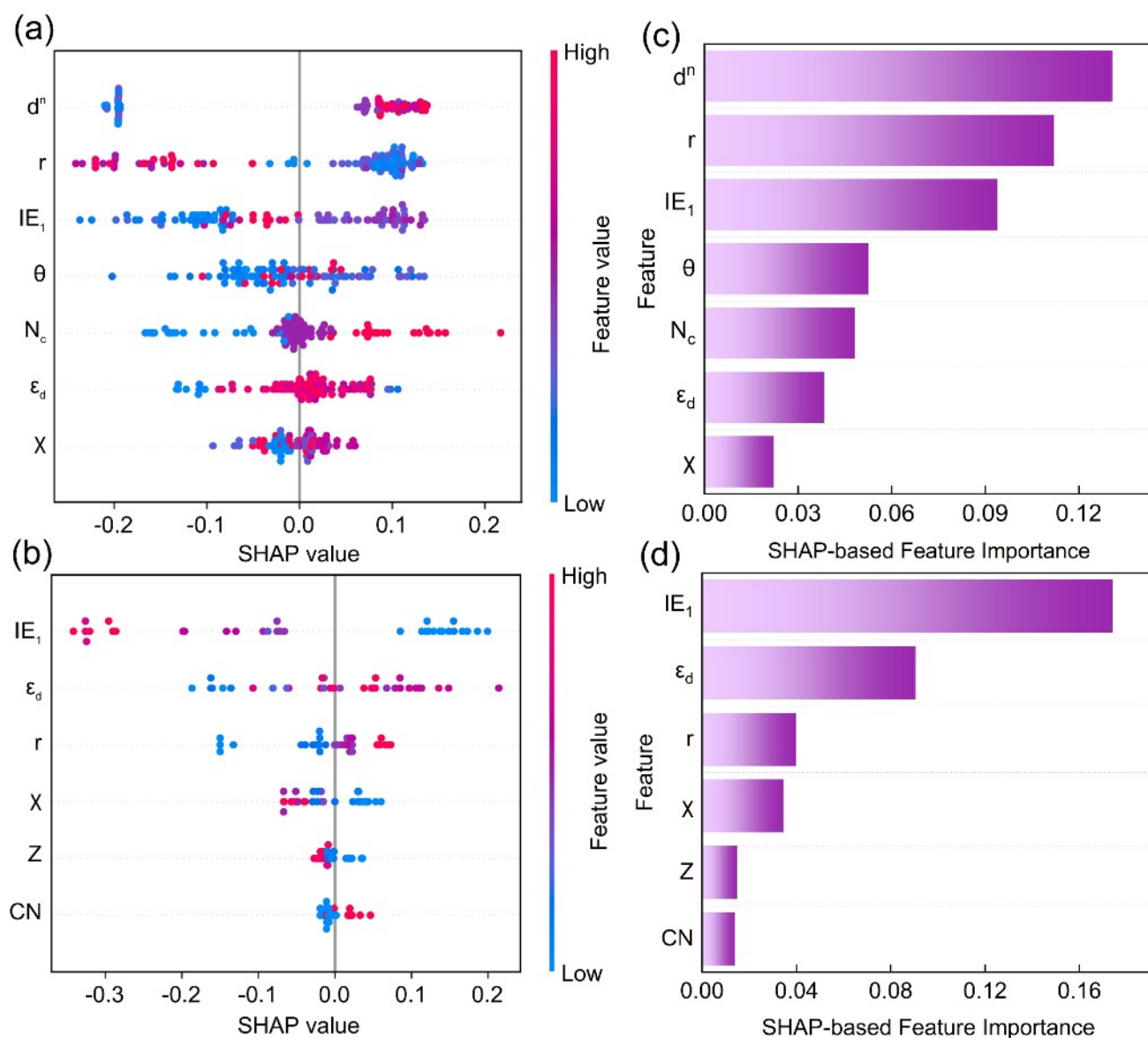
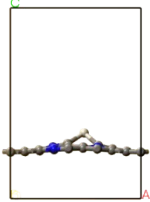
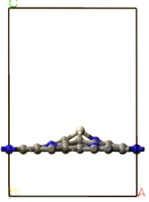
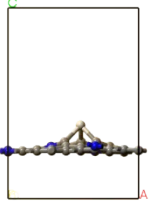
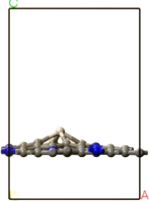
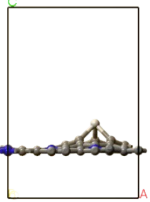
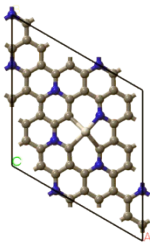
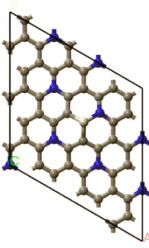
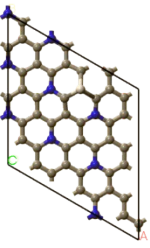
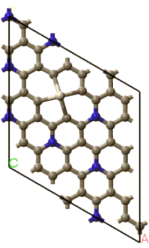
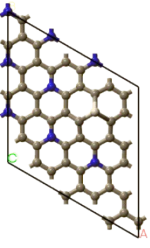
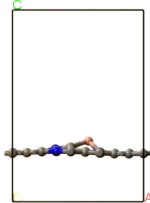
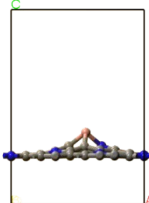
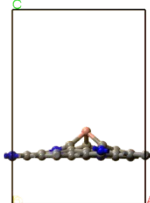
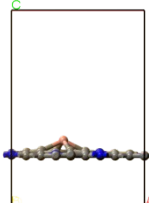
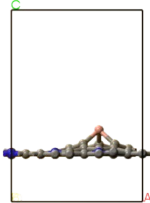
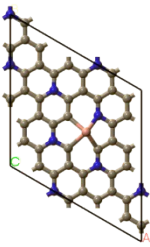
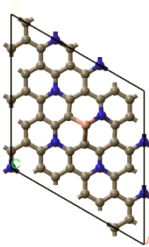
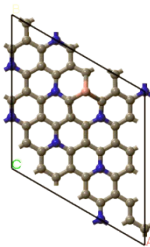
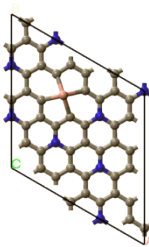
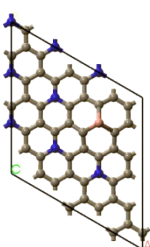
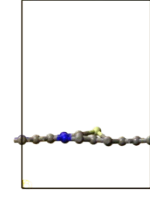
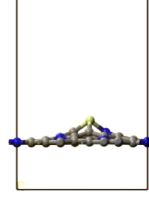
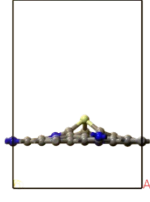
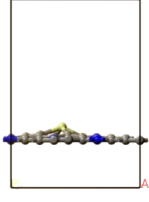
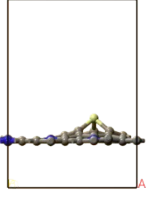
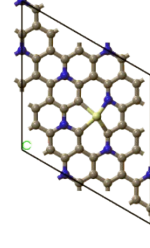
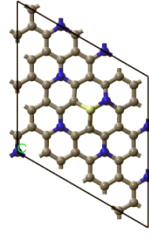
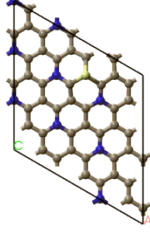
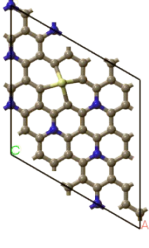
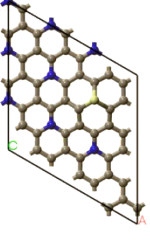
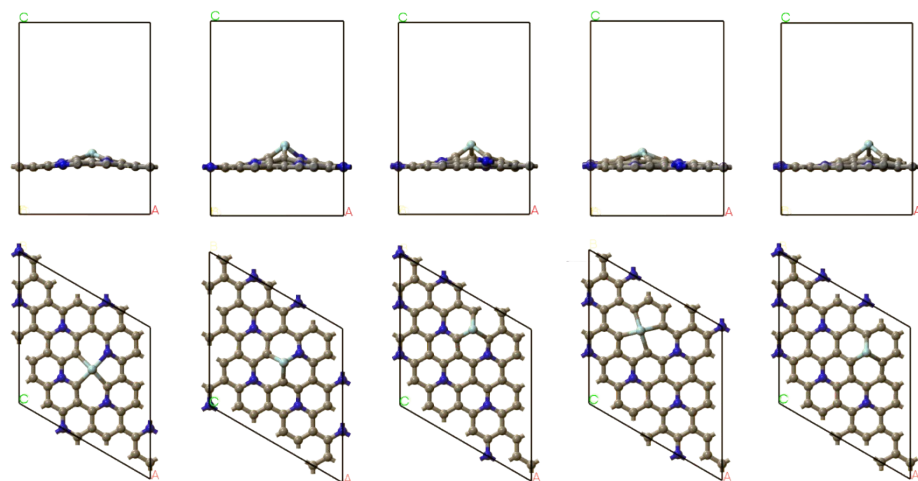


Figure S13 SHAP analysis of the ML model for $\Delta G(*CHO)$ (top) and $\Delta G(H_2O^*)$ (bottom). (a, b) SHAP summary plots showing the impact of each feature on the predicted output. Each dot represents an individual sample (catalyst), and its color indicates the feature value (red = high, blue = low) (c, d) SHAP-based global feature importance rankings calculated from average absolute SHAP values. The most influential descriptors are highlighted, illustrating their relative contributions to model prediction.

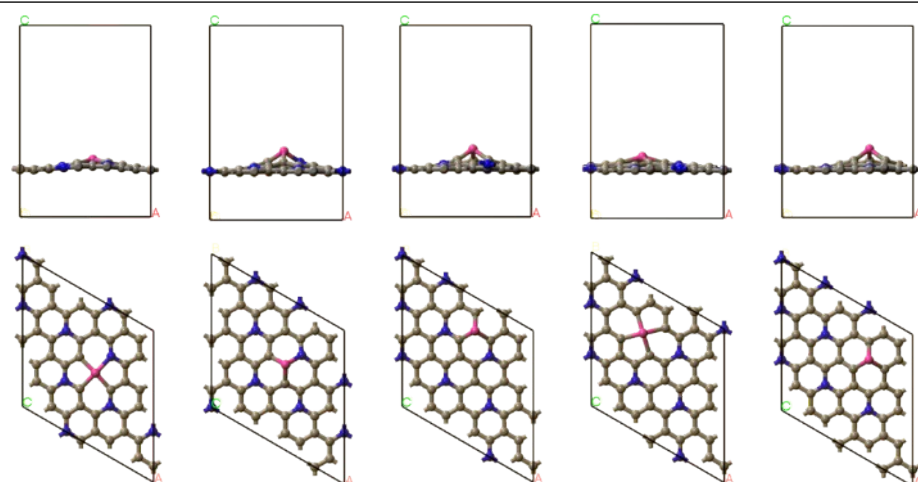
Table S1. 100 distinct TM@C₅N single-atom catalyst configurations

	2C	C1	C2	CN	N
Sc@C ₅ N					
					
Ti@C ₅ N					
					
V@C ₅ N					
					

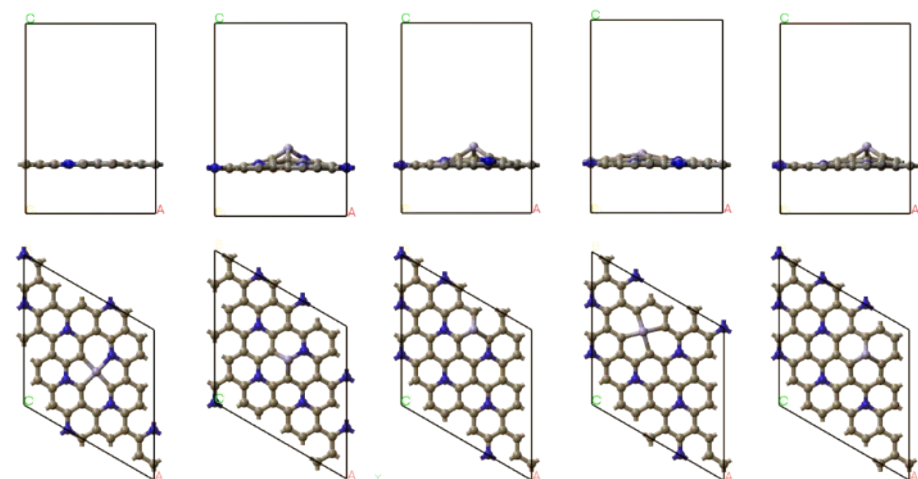
$\text{Cr@C}_5\text{N}$



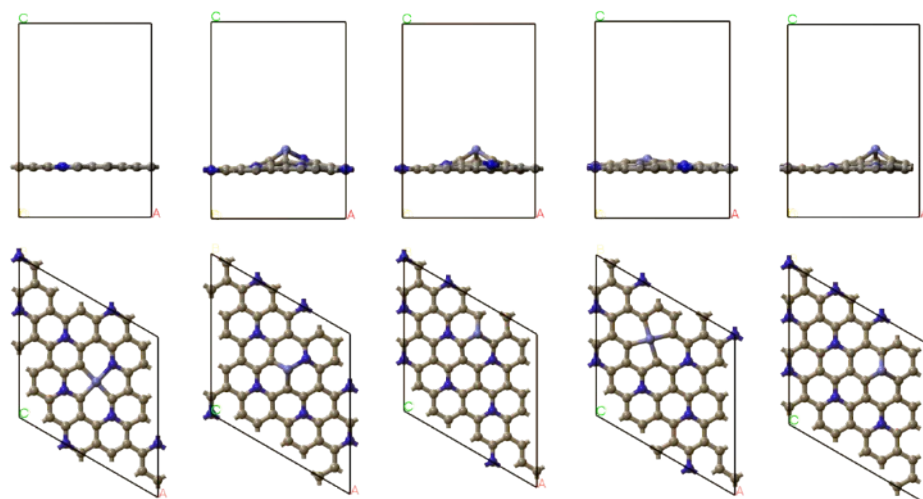
$\text{Mn@C}_5\text{N}$



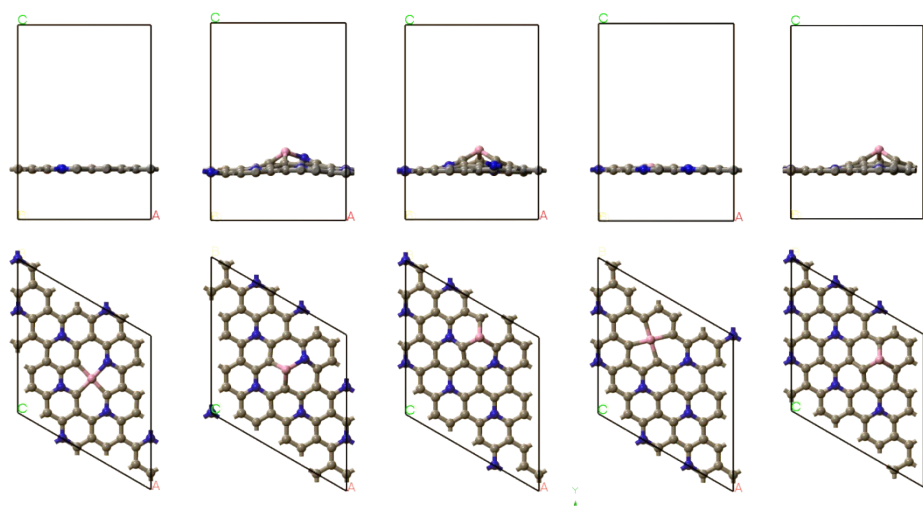
$\text{Fe@C}_5\text{N}$



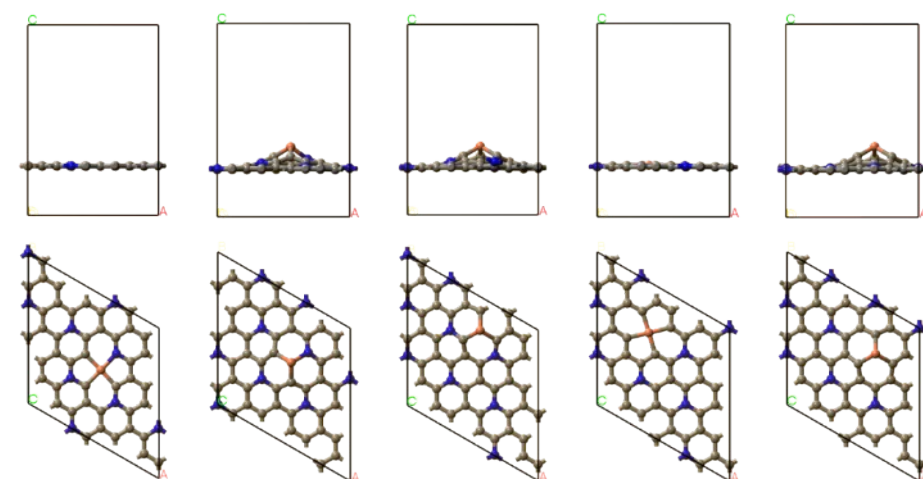
Co@C₅N



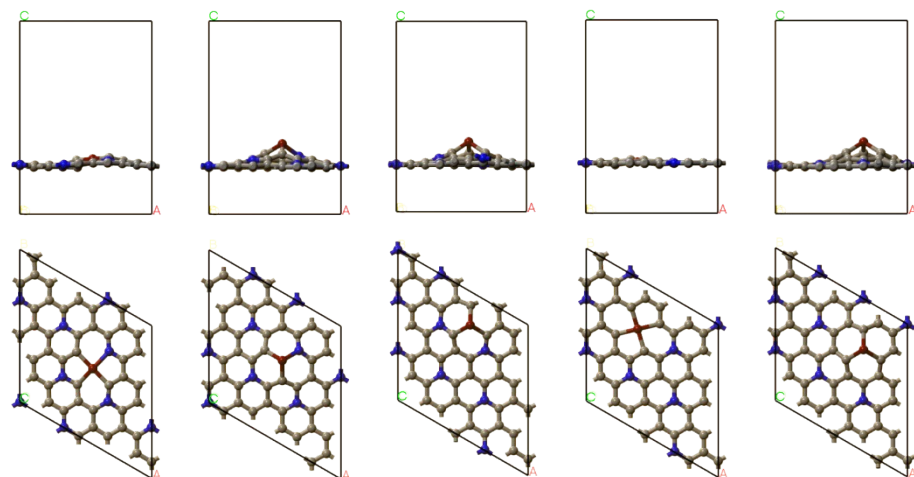
Ni@C₅N



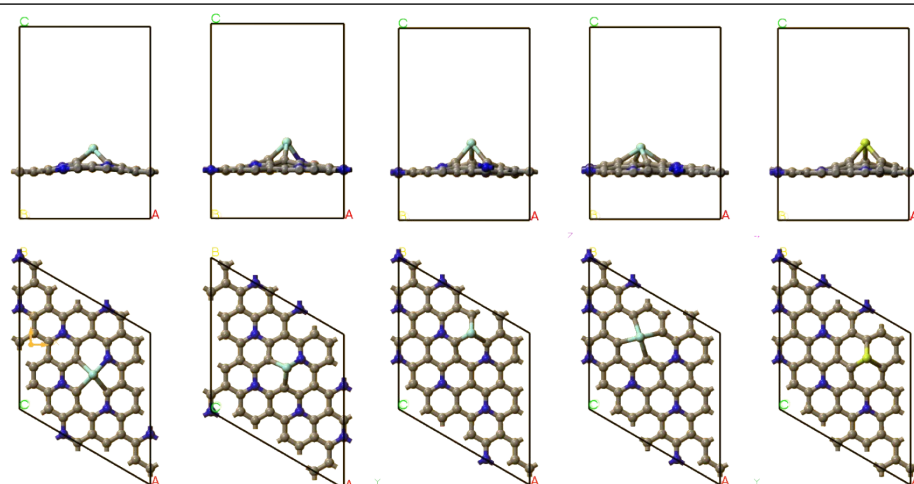
Cu@C₅N



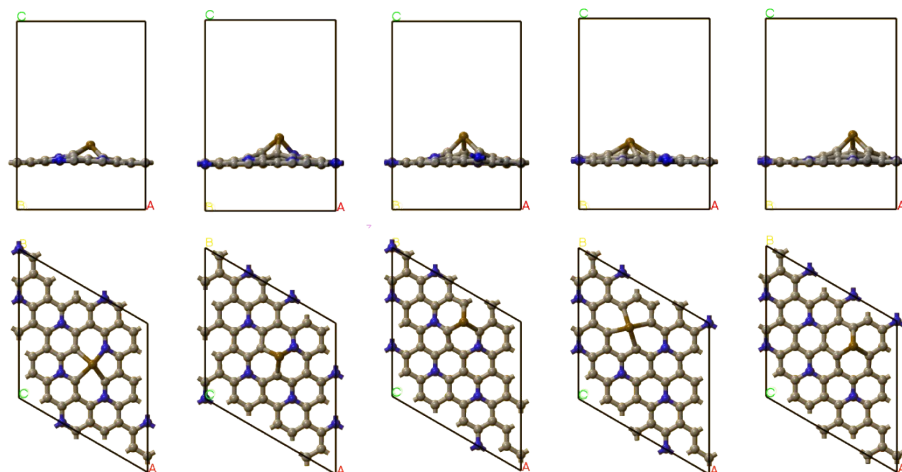
$\text{Zn@C}_5\text{N}$



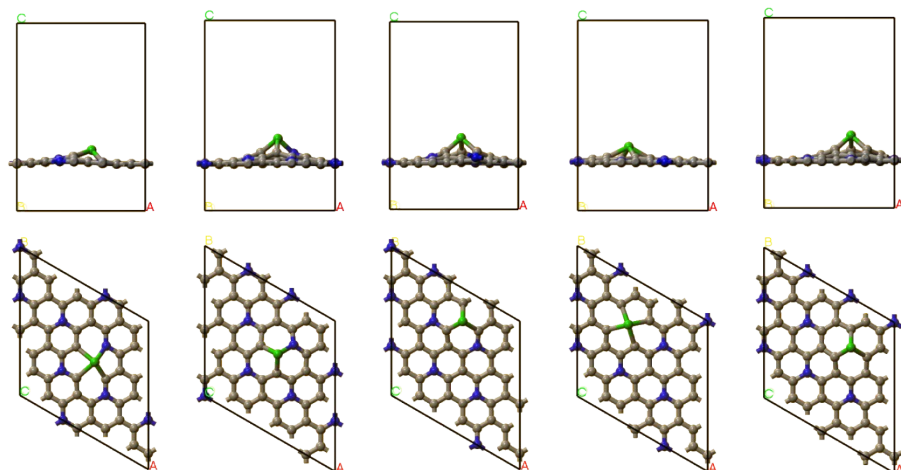
$\text{Y@C}_5\text{N}$



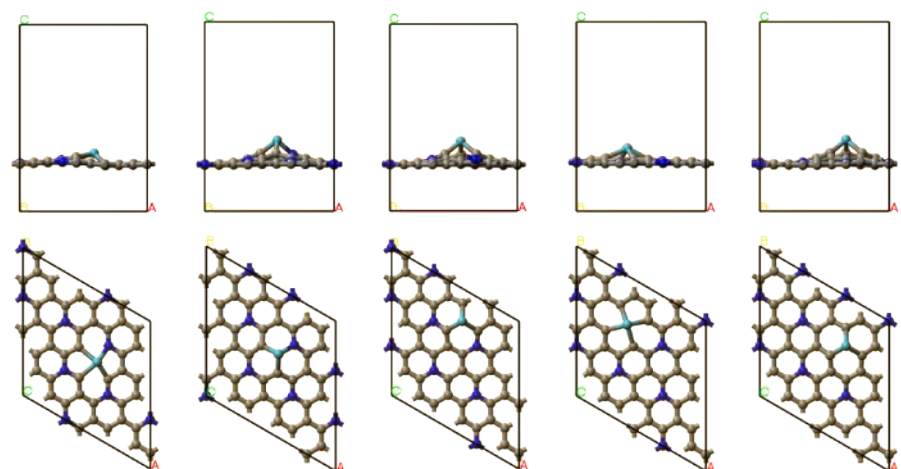
$\text{Zr@C}_5\text{N}$



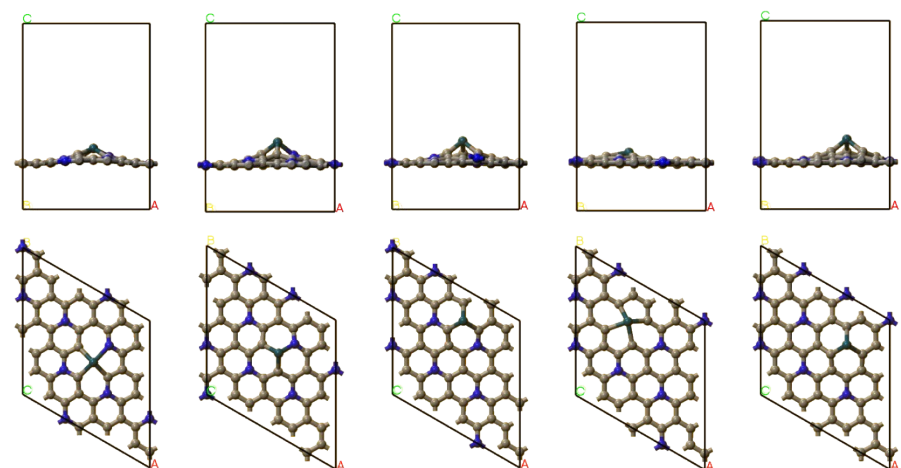
Nb@C₅N



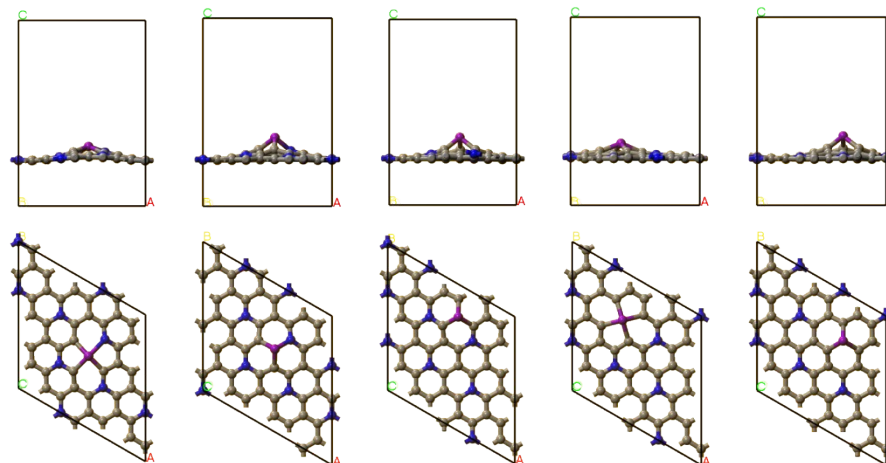
Mo@C₅N



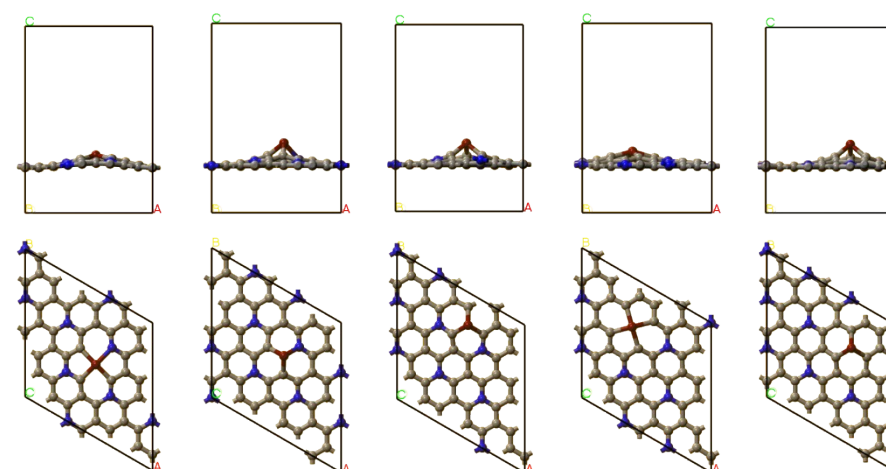
Tc@C₅N



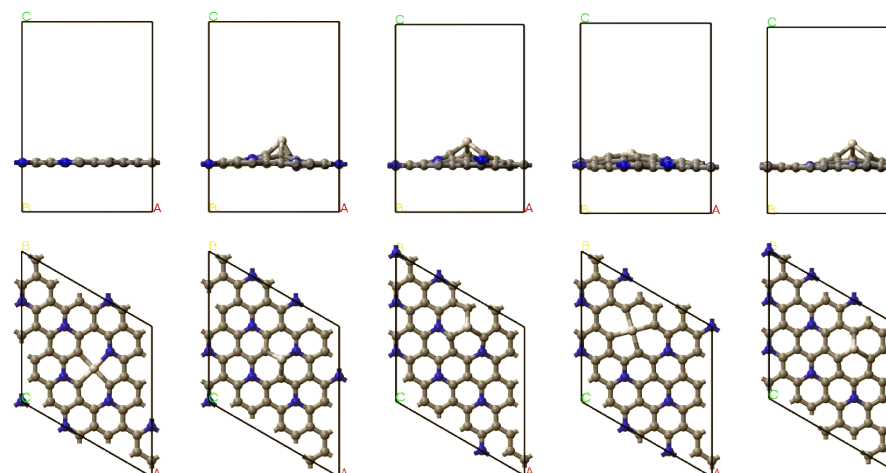
Ru@C₅N



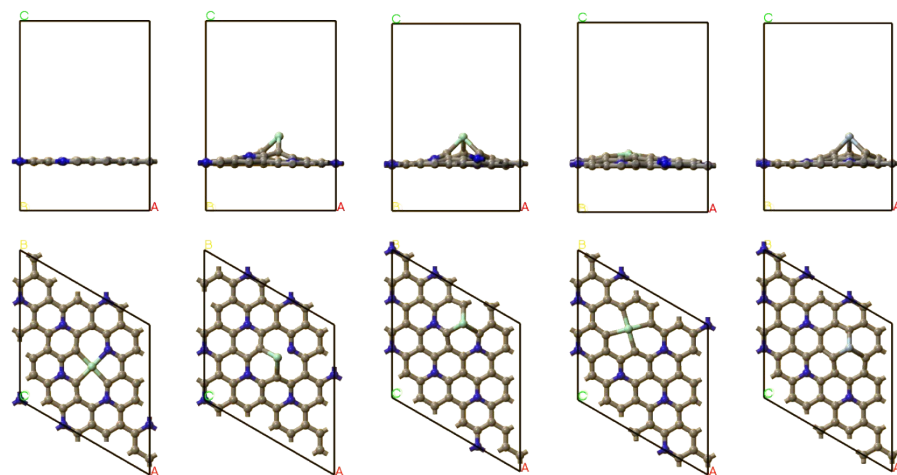
Rh@C₅N



Pd@C₅N



$\text{Ag}@\text{C}_5\text{N}$



$\text{Cd}@\text{C}_5\text{N}$

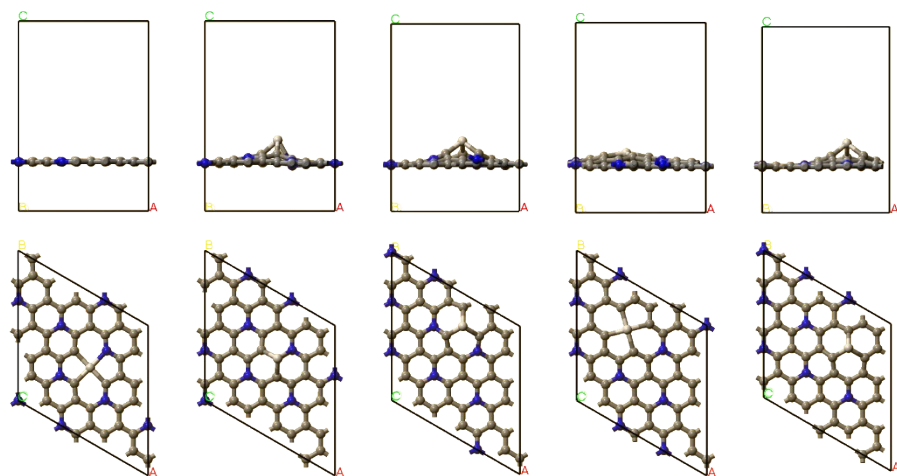


Table S2 The C-O bond lengths of adsorbed CO₂ molecule.

	2C	C1	C2	CN	N
Sc	1.259	1.306	1.302	1.186	1.264
Ti	1.324	1.350	1.338	1.292	1.320
V	1.321	1.331	1.316	1.326	1.315
Cr	1.317	1.296	1.300	1.317	1.303
Mn	1.270	1.323	1.276	1.294	1.291
Fe	1.270	1.323	1.276	1.294	1.291
Co	1.274	1.292	1.263	1.284	1.260
Ni	1.179	1.311	1.308	1.253	1.259
Cu	1.178	1.279	1.279	1.178	1.280
Zn	1.178	1.269	1.244	1.176	1.251
Y	1.269	1.299	1.283	1.183	1.261
Zr	1.336	1.365	1.346	1.301	1.323
Nb	1.352	1.354	1.329	1.338	1.328
Mo	1.342	1.332	1.311	1.304	1.311
Tc	1.349	1.305	1.302	1.312	1.299
Ru	1.287	1.257	1.235	1.307	1.257
Rh	1.238	1.300	1.276	1.258	1.248
Pd	1.177	1.264	1.286	1.177	1.283
Ag	1.177	1.263	1.247	1.178	1.236
Cd	1.179	1.177	1.179	1.179	1.182

Table S3 bending angles of adsorbed CO₂ molecule.

	2C	C1	C2	CN	N
Sc	144.081	133.497	134.394	134.394	142.164
Ti	137.600	130.308	130.988	138.111	132.767
V	135.394	136.891	144.939	134.346	134.961
Cr	135.339	137.217	138.499	134.165	137.707
Mn	141.290	134.601	140.550	137.590	140.837
Fe	140.101	138.926	141.759	139.508	141.714
Co	173.076	136.791	136.069	143.771	143.207
Ni	175.571	142.371	140.348	168.845	140.111
Cu	178.105	142.784	146.319	162.586	146.453
Zn	178.388	179.548	179.638	177.579	179.467
Y	140.546	134.072	124.158	179.631	129.287
Zr	130.715	128.324	129.309	135.388	131.561
Nb	131.334	130.756	132.827	131.784	132.565
Mo	133.050	134.162	136.350	134.762	135.923
Tc	132.315	138.620	138.473	134.345	139.292
Ru	138.286	143.625	145.169	136.009	142.361
Rh	145.040	137.745	140.029	142.416	145.505
Pd	176.277	144.169	138.447	176.849	138.628
Ag	178.208	143.341	145.888	178.437	152.209
Cd	178.865	178.529	176.401	179.399	179.589

Table S4 Free energy of the potential-determining step (PDS) with and without implicit solvent models (ISM).

Free Energy (eV)	Co-C2		Ni-C1		Ni-C2		Ni-N		Rh-2C		Rh-C2	
	OH	H ₂ O	*CO ₂	*COOH	*CO	*CHO	*CO	*CHO	*CHO	*CH ₂ O	*OH	H ₂ O*
G _{ISM}	-488.933	-491.641	-500.639	-503.316	-491.725	-494.636	-492.693	-495.616	-489.970	-492.887	-488.852	-491.720
G	-488.664	-491.372	-500.287	-503.039	-491.557	-494.412	-492.586	-495.404	-489.726	-492.629	-488.645	-491.423
ΔG _{ISM}	0.653		0.683		0.448		0.437		0.443		0.492	
ΔG	0.651		0.608		0.505		0.542		0.458		0.582	

Table S5 16 basic features for ML.

1	Z	proton number
2	n	electron shell number
3	A_r	relative atomic mass
4	ε_d	d-band center
5	r	atomic radius
6	d^n	number of d-electrons
7	χ	Pauling electronegativity
8	IE_1	first ionization energy
9	CN	coordination number
10	N_c	the number of carbon atoms
11	l_{TM-O}	distance between TM and oxygen
12	l_{TM-C}	distance between TM and carbon
13	l_{C-O}	distance between carbon and oxygen
14	θ	the bending angle of the CO ₂ molecule

Table S6 R2 and mean squared error (MSE) values of the XGBoost model for ΔG (*CHO) prediction across 10 Monte Carlo cross-validation (MCCV) trials. Metrics are reported separately for the training and test sets. The final column lists the average values over all trials, reflecting overall prediction accuracy and generalization performance.

		1	2	3	4	5	6	7	8	9	10	Avg
Train	R ²	0.98	0.90	0.97	0.96	0.98	0.96	0.99	0.98	0.96	0.96	0.96
	MSE	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Test	R ²	0.92	0.89	0.85	0.85	0.85	0.88	0.83	0.83	0.86	0.82	0.86
	MSE	0.02	0.02	0.01	0.03	0.02	0.02	0.03	0.06	0.02	0.02	0.02

Table S7 R^2 and mean squared error (MSE) values of the Random Forest model for ΔG (H_2O^*) prediction across 10 Monte Carlo cross-validation (MCCV) trials. Metrics are reported separately for the training and test sets. The final column lists the average values over all trials, reflecting overall prediction accuracy and generalization performance.

		1	2	3	4	5	6	7	8	9	10	Avg
Train	R^2	0.93	0.91	0.89	0.80	0.94	0.86	0.82	0.89	0.94	0.81	0.88
	MSE	0.01	0.01	0.02	0.04	0.01	0.02	0.03	0.01	0.01	0.03	0.02
Test	R^2	0.87	0.86	0.85	0.92	0.81	0.82	0.83	0.84	0.88	0.89	0.86
	MSE	0.02	0.03	0.02	0.01	0.03	0.03	0.02	0.04	0.02	0.02	0.02

Reference

- 1 G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169–11186.
- 2 G. Kresse and D. Joubert, *Phys. Rev. B*, 1999, **59**, 1758–1775.
- 3 P. E. Blöchl, *Phys. Rev. B*, 1994, **50**, 17953–17979.
- 4 J. P. Perdew and Y. Wang, *Phys. Rev. B*, 1992, **45**, 13244–13249.
- 5 J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1992, **46**, 6671–6687.
- 6 R. Nelson, C. Ertural, J. George, V. L. Deringer, G. Hautier and R. Dronskowski, *J. Comput. Chem.*, 2020, **41**, 1931–1940.
- 7 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
- 8 G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.
- 9 K. Mathew, R. Sundararaman, K. Letchworth-Weaver, T. A. Arias and R. G. Hennig, *J. Chem. Phys.*, 2014, **140**, 084106.
- 10 K. Mathew, V. S. C. Kolluru, S. Mula, S. N. Steinmann and R. G. Hennig, *The Journal of Chemical Physics*, 2019, **151**, 234101.
- 11 A. A. Peterson, F. Abild-Pedersen, F. Studt, J. Rossmeisl and J. K. Nørskov, *Energy Environ. Sci.*, 2010, **3**, 1311–1315.
- 12 J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard and H. Jónsson, *J. Phys. Chem. B*, 2004, **108**, 17886–17892.
- 13 X. Liu, J. Xiao, H. Peng, X. Hong, K. Chan and J. K. Nørskov, *Nat. Commun.*, 2017, **8**, 15438.