

Supporting Information

Machine-learned Charge State Prediction Enables Nitrogen Reduction Reaction on Transition Metals Doped C₃B

Chengwei Yang^{a,b,†}, Chao Yang^{c,†}, Yunxia Liang^a, Hongxia Yan^a, Aodi Zhang^d, Guixian Ge^{a,*}, Wentao Wang^{b,c,*}, Pengfei Ou^{c,*}

^a Key Laboratory of Ecophysics and Department of Physics, College of Science, Shihezi University, Shihezi 832003, PR China

^b Guizhou Provincial Key Laboratory of Computational Nano-Material Science, Guizhou Education University, Guiyang 550018, PR China.

^c Department of Chemistry, National University of Singapore, Singapore, 117543, Singapore

^d School of Physics and Electronics, Henan University, Kaifeng 475004, China

† These authors contributed equally.

* Corresponding authors: pengf.ou@nus.edu.sg (P. Ou); wtwang@gznc.edu.cn (W. Wang); geguixian@126.com (G. Ge)

Note S1. The formation energy of defect

The defect formation energy $E_f(D^q)$ in 2D C₃B for a defect D in a charge state q is evaluated as ¹:

$$E_f(D^q) = E_{\text{tot}}(D^q) - E_{\text{tot}}(\text{bulk}) + \sum_i n_i \mu_i + q(E_F + E_{VBM}) \quad (\text{S1})$$

- $E_{\text{tot}}(D^q)$ and $E_{\text{tot}}(\text{bulk})$ are the calculated total energies of the supercells with defect X in charge state q and without the defect, respectively.
- n_i is the number of atoms of type i ($n_i < 0$, if atoms are added; and $n_i > 0$, if atoms are removed).
- μ_i is the corresponding atomic chemical potential. The chemical potentials of C and B are thermodynamic constraints that can be used to represent experimental synthesis conditions. Under the thermal equilibrium growth conditions, the stable generation of the host material C₃B should satisfy the equation:

$$3\mu_C + \mu_B = \mu_{C_3B} \quad (\text{S2})$$

Where μ_C and μ_B represent the chemical potentials of C and B in C₃B and μ_{C_3B} is the chemical potential of C₃B. The optimized static energy E_{C_3B} for per formula unit in C₃B can be expressed as

$$E_{C_3B} = 3E_C + E_B + \Delta H_f(C_3B) \quad (\text{S3})$$

Where E_C is the energy of a C atom in the graphene, E_B is the energy of a C atom in the borophene, and $\Delta H_f(C_3B)$ is the formation enthalpy per formula C_3B . Under the thermal equilibrium growth conditions, $E_{C3B} = \mu_{C3B}$; then, from **Eqs. (S2)** and **(S3)**, we have the equation

$$\Delta H_f(C_3B) = 3(\mu_C - E_C) + (\mu_B - E_B) = 3\Delta\mu_C + \Delta\mu_B \quad (S4)$$

where, $\Delta\mu_C = \mu_C - E_C$ and $\Delta\mu_B = \mu_B - E_B$ are defined as the chemical potential changes referenced to the elemental bulk/gaseous state, depending on the growth environment. With the **Eqs. (3)** and **(4)**, the calculated PBE formation enthalpies per formula unit C_3B is -0.68 eV. Then, under C-poor (B-rich) conditions $\Delta\mu_B=0$, $\mu_B=E_B$, $\Delta\mu_C=\Delta H_f(C_3B)/3$, thus, we can get μ_C . Similarly, we can get the chemical potential of C/B atom under B-rich (C-poor) conditions. The acquired chemical potentials for μ_B and μ_C under the two extreme growth conditions are shown in **Table S1**.

Table S1. The chemical potentials calculated using the PBE functional at the representative limits imposed by the formation of competing impurity-related phases for transition metal single atoms are presented. (unit:eV)

	B rich	C rich
$\mu(B)$	6.70	5.78
$\mu(C)$	8.91	9.22
$\mu(Ti)$	11.02	12.87
$\mu(V)$	11.21	13.06
$\mu(Cr)$	10.57	12.39
$\mu(Mn)$	10.00	10.92
$\mu(Fe)$	8.99	9.92
$\mu(Co)$	7.84	8.77
$\mu(Ni)$	5.95	6.88
$\mu(Cu)$	3.71	3.71
$\mu(Zr)$	11.58	13.42
$\mu(Nb)$	14.99	14.68
$\mu(Mo)$	12.31	13.23
$\mu(Ru)$	10.10	11.95
$\mu(Rh)$	8.05	8.98
$\mu(Pd)$	5.57	6.12
$\mu(Ag)$	1.16	3.01
$\mu(Hf)$	12.98	14.83

$\mu(\text{Ta})$	13.75	15.60
$\mu(\text{W})$	13.44	14.37
$\mu(\text{Re})$	13.68	15.53
$\mu(\text{Os})$	11.86	13.71
$\mu(\text{Ir})$	9.27	10.20
$\mu(\text{Pt})$	6.47	7.39
$\mu(\text{Au})$	1.22	3.07

When the extrinsic element doping is studied theoretically, the initial step also involves determining the chemical potential of the doping element^{2, 3}. According to the general rules²⁻⁷, to prevent the precipitation of source elements during doping using TM elements, it is required that

$$\Delta\mu_X < 0 \quad (\text{S5})$$

Where, $\Delta\mu_X = \mu_X - E_X$, is defined as the change in the chemical potential of the reference element in the bulk/gas state, depending on the growth environment. E_X is the energy of one X atom in an X crystal. We take Ti doping as an example, the upper bound is $\Delta\mu_{\text{Ti}} < 0$ as. On the other hand, to prevent the formation of secondary phases. Such as TiC, TiB, and TiB_2 , because of the extrinsic doping in C_3B , it is required that

$$\Delta\mu_{\text{Ti}} + \Delta\mu_{\text{C}} < \Delta H_f(\text{TiC}) \quad (\text{S6})$$

$$\Delta\mu_{\text{Ti}} + \Delta\mu_{\text{B}} < \Delta H_f(\text{TiB}) \quad (\text{S7})$$

$$\Delta\mu_{\text{Ti}} + 2\Delta\mu_{\text{B}} < \Delta H_f(\text{TiB}_2) \quad (\text{S8})$$

In which, $\Delta H_f(\text{TiB})$, $\Delta H_f(\text{TiB}_2)$ and $\Delta H_f(\text{TiC})$ are the formation enthalpies of per molecular unit of TiB, TiB_2 and TiC, respectively, and it can be calculated by^{5, 8}

$$\Delta H_f(\text{TiC}) = E(\text{TiC}) - E_{\text{Ti}} - E_{\text{C}} \quad (\text{S9})$$

$$\Delta H_f(\text{TiB}) = E(\text{TiB}) - E_{\text{Ti}} - E_{\text{B}} \quad (\text{S10})$$

$$\Delta H_f(\text{TiB}_2) = E(\text{TiB}_2) - E_{\text{Ti}} - 2E_{\text{B}} \quad (\text{S11})$$

Where, $E(\text{TiC})$, $E(\text{TiB})$ and $E(\text{TiB}_2)$ are the total energy per formula TiC, TiB and TiB_2 . The control of chemical potentials of TM anchored on the C_3B monolayer is crucial to avoid undesired phase formation^{4, 8, 9}. We conducted PBE functional calculations using experimental lattice constants of impurity-related phases for TMs such as Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zr, Nb, Mo, Ru, Rh, Pd, Ag, Hf, Ta, W, Re, Os, Ir, Pt, and Au anchored on the C_3B monolayer, as summarized in **Table S2**. The PBE calculated formation enthalpy (given per formula unit) for the

impurity-related phases are shown in **Table S3**, together with the Oqmd^{10, 11} (the Open Quantum Materials Database will be freely available for public use at <https://oqmd.org/>) values. It is found that the calculated values are in good agreement with the Oqmd ones.

Table S2. Experimental lattice constants of impurity-related phases (Ti, V, Cr, Mn, Fe, Co, Ni, Cu,

Ru, Rh, Pd, and Ag) are used for the PBE functional calculations. The values a, b, and c are in Å.

	Space group	Experiment
TiB ₂	P6/mmm	a=b=3.029, c=3.221; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ¹²
TiB	Pnma	a= 6.111, b=3.045, c=4.557; $\alpha=\beta=\gamma=90^\circ$ ¹³
TiC	Fm $\bar{3}$ m	a=b=c=4.331; $\alpha=\beta=\gamma=90^\circ$ ¹⁴
VB ₂	P6/mmm	a=b=2.994, c=3.018; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ¹⁵
VB	Cmcm	a=b=4.286, c=2.963; $\alpha=\beta=90^\circ$, $\gamma=138.454^\circ$ ¹⁶
VC	Fm $\bar{3}$ m	a=b=c=2.935; $\alpha=\beta=\gamma=60^\circ$ ¹⁷
CrB ₂	P6/mmm	a=b=2.956, c=3.023; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ¹⁸
CrB	I4 ₁ /amd	a=b=c=8.100; $\alpha=\beta=159.256^\circ$, $\gamma=29.502^\circ$ ¹⁹
Cr ₃ C ₂	Pnma	a=5.477, b=2.786, c=11.453; $\alpha=\beta=\gamma=90^\circ$ ²⁰
Mn ₇ C ₃	Pnma	a=4.439, b=6.794, c=11.721; $\alpha=\beta=\gamma=90^\circ$ ²¹
MnB	Pnma	a=5.445, b=2.983, c=4.114; $\alpha=\beta=\gamma=90^\circ$ ²²
Fe ₂ B	I4/mcm	a= b=5.052, c= 4.231; $\alpha=\beta=\gamma=90^\circ$ ²³
FeB	I4 ₁ /amd	a=b=c=7.758; $\alpha=\beta=158.258^\circ$, $\gamma=30.938^\circ$ ²⁴
Fe ₃ C	Pnma	a=5.30, b=6.711, c=4.475; $\alpha=\beta=\gamma=90^\circ$ ²⁵
FeC	Fm $\bar{3}$ m	a=b=c=3.985; $\alpha=\beta=\gamma=90^\circ$ ²⁶
CoB	Pnma	a=5.152, b=3.058, c=3.917; $\alpha=\beta=\gamma=90^\circ$ ²⁷
Co ₂ C	Pnma	a=4.359, b=4.363, c=2.874; $\alpha=\beta=\gamma=90^\circ$ ²⁸
Ni ₃ B	Pnma	a= 5.202, b= 6.641, c= 4.389; $\alpha=\beta=\gamma=90^\circ$ ²⁹
Ni ₂ B	I4/mcm	a=b=c=4.119; $\alpha=\beta=105.592^\circ$, $\gamma=117.544^\circ$ ²³
NiB	Cmcm	a=b=3.973, c=2.991; $\alpha=\beta=90^\circ$, $\gamma=136.575^\circ$ ³⁰
Ni ₃ C	R $\bar{3}$ c	a=b=c=5.079; $\alpha=\beta=\gamma=53.819^\circ$ ³¹
ZrB ₂	P6/mmm	a=b=3.167, c=3.529; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ¹²
ZrC	Fm $\bar{3}$ m	a=b=c=4.692; $\alpha=\beta=\gamma=90^\circ$ ³²
NbB ₂	P6/mmm	a=b=3.102, c=3.321; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ³³
NbB	Cmcm	a=3.309, b=8.792, c=3.171; $\alpha=\beta=\gamma=90^\circ$ ¹⁶
NbC	Fm $\bar{3}$ m	a=b=c=4.473; $\alpha=\beta=\gamma=90^\circ$ ³⁴
MoB ₂	R $\bar{3}$ m	a=b=3.02, c=20.02; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ³⁵
MoB	I4 ₁ /amd	a=b=3.126, c=3.221; $\alpha=\beta=\gamma=90^\circ$ ³⁶

Mo ₂ C	P $\bar{3}$ m2	a=b=2.884, c=5.462; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ³⁷
MoC	P $\bar{6}$ m2	a=b=2.903, c=2.828; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ³⁸
RuB ₂	Pmmn	a=4.645, b=2.865, c=4.045; $\alpha=\beta=\gamma=90^\circ$ ³⁹
Ru ₂ B ₃	P6 ₃ /mmc	a=b=2.917, c=12.856; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁴⁰
RuC	P $\bar{6}$ m2	a=b=2.908, c=2.822; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁴¹
RhB	Cmcm	a=3.382, b=5.86, c=4.157; $\alpha=\beta=\gamma=90^\circ$ ⁴²
Rh ₂ B	Pnma	a=5.36, b=3.836, c=6.813; $\alpha=\beta=\gamma=90^\circ$ ⁴³
Pd ₅ B ₂	C12/c1	a=12.776, b=4.95, c=5.47; $\alpha=90^\circ$, $\beta=90.05^\circ$, $\gamma=90^\circ$ ⁴⁴
Pd ₃ B	Pnma	a=5.464, b=7.563, c=4.85; $\alpha=\beta=\gamma=90^\circ$ ⁴⁵
PdC	Fm $\bar{3}$ m	a=b=c=4.432; $\alpha=\beta=\gamma=90^\circ$ ⁴⁶
AgB ₂	P6/mmm	a=b=3.016, c=4.079; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁴⁷
HfB ₂	P6/mmm	a=b=3.141, c=3.484; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁴⁸
HfC	Fm $\bar{3}$ m	a=b=c=3.284; $\alpha=\beta=\gamma=60^\circ$ ⁴⁹
TaB ₂	P6/mmm	a=b=3.10, c=3.318; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ¹²
TaC	Fm $\bar{3}$ m	a=b=c=3.166; $\alpha=\beta=\gamma=60^\circ$ ⁵⁰
TaB	Cmcm	a=b=4.662, c=3.164; $\alpha=\beta=90^\circ$, $\gamma=138.638^\circ$ ⁵¹
W ₂ B	I4/mcm	a=b=c=4.621; $\alpha=\beta=\gamma=117.737^\circ$ ⁵¹
WB	Cmcm	a=b=4.662, c=2.853; $\alpha=\beta=90^\circ$, $\gamma=138.638^\circ$ ⁵¹
WC	P $\bar{6}$ m2	a=b=2.923, c=3.221; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁵²
ReB	P $\bar{6}$ m2	a=b=2.897, c=2.941; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁵³
OsB ₂	Pmmn	a=2.884, b=4.693, c=4.082; $\alpha=\beta=\gamma=90^\circ$ ³⁹
OsB	P $\bar{6}$ m2	a=b=2.895, c=2.876; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁵⁴
OsC	P $\bar{6}$ m2	a=b=2.994, c=2.730; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁵⁵
IrB	Cmcm	a=b=3.553, c=3.939; $\alpha=\beta=90^\circ$, $\gamma=111.102^\circ$ ⁵⁶
IrC	P $\bar{6}$ m2	a=b=3.045, c=10.431; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁵⁷
Pt ₂ B	P6 ₃ /mmc	a=b=2.936, c=3.221; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁵⁸
PtB	I4/mmm	a=b=c=4.651; $\alpha=\beta=131.311^\circ$, $\gamma=71.319^\circ$ ⁵⁹
PtC	Fm $\bar{3}$ m	a=b=c=3.163; $\alpha=\beta=\gamma=60^\circ$ ⁶⁰
AuB ₂	P6/mmm	a=b=3.005, c=4.233; $\alpha=\beta=90^\circ$, $\gamma=120^\circ$ ⁴⁷

Table S3. The formation enthalpy per formula unit for impurity-related phases in TM@C₃B (B and C sites) was calculated using the PBE functional. For comparison, we referenced 67 values from the Open Quantum Materials Database (OQMD), which are publicly accessible at OQMD^{10, 11}. (unit: eV)

	PBE	Oqmd
--	-----	------

$\Delta\text{Hf}(\text{TiB}_2)$	3.18	3.20
$\Delta\text{Hf}(\text{TiB})$	1.67	1.70
$\Delta\text{Hf}(\text{TiC})$	1.64	1.61
$\Delta\text{Hf}(\text{VB}_2)$	2.23	2.20
$\Delta\text{Hf}(\text{VB})$	1.71	1.67
$\Delta\text{Hf}(\text{VC})$	0.87	0.78
$\Delta\text{Hf}(\text{CrB}_2)$	1.04	1.01
$\Delta\text{Hf}(\text{CrB})$	1.07	1.05
$\Delta\text{Hf}(\text{Cr}_3\text{C}_2)$	0.66	0.52
$\Delta\text{Hf}(\text{Mn}_7\text{C}_3)$	0.83	0.85
$\Delta\text{Hf}(\text{MnB})$	1.00	1.00
$\Delta\text{Hf}(\text{Fe}_2\text{B})$	0.92	0.94
$\Delta\text{Hf}(\text{FeB})$	0.75	0.76
$\Delta\text{Hf}(\text{Fe}_3\text{C})$	0.21	0.19
$\Delta\text{Hf}(\text{FeC})$	1.20	1.20
$\Delta\text{Hf}(\text{CoB})$	0.82	0.82
$\Delta\text{Hf}(\text{Co}_2\text{C})$	0.27	0.28
$\Delta\text{Hf}(\text{Ni}_3\text{B})$	1.07	1.07
$\Delta\text{Hf}(\text{Ni}_2\text{B})$	0.89	0.88
$\Delta\text{Hf}(\text{NiB})$	0.49	0.49
$\Delta\text{Hf}(\text{Ni}_3\text{C})$	0.22	0.24
$\Delta\text{Hf}(\text{ZrB}_2)$	3.05	2.98
$\Delta\text{Hf}(\text{ZrC})$	1.72	1.62
$\Delta\text{Hf}(\text{NbB}_2)$	2.21	2.09
$\Delta\text{Hf}(\text{NbB})$	1.61	1.55
$\Delta\text{Hf}(\text{NbC})$	4.47	4.06
$\Delta\text{Hf}(\text{MoB}_2)$	2.72	2.61
$\Delta\text{Hf}(\text{MoB})$	1.02	1.00
$\Delta\text{Hf}(\text{Mo}_2\text{C})$	1.64	1.61
$\Delta\text{Hf}(\text{MoC})$	0.19	0.18
$\Delta\text{Hf}(\text{RuB}_2)$	1.72	1.66
$\Delta\text{Hf}(\text{Ru}_2\text{B}_3)$	1.64	1.61

$\Delta Hf(RuC)$	1.35	1.30
$\Delta Hf(RhB)$	0.78	0.78
$\Delta Hf(Rh_2B)$	0.62	0.63
$\Delta Hf(Pd_5B_2)$	1.83	1.32
$\Delta Hf(Pd_3B)$	1.05	1.04
$\Delta Hf(PdC)$	2.91	2.97
$\Delta Hf(AgB_2)$	1.54	1.56
$\Delta Hf(HfB_2)$	3.05	3.07
$\Delta Hf(HfC)$	1.89	1.90
$\Delta Hf(TaB_2)$	1.94	1.95
$\Delta Hf(TaC)$	1.16	1.17
$\Delta Hf(TaB)$	1.63	1.63
$\Delta Hf(W_2B)$	0.80	0.78
$\Delta Hf(WB)$	0.50	0.71
$\Delta Hf(WC)$	0.29	0.25
$\Delta Hf(ReB)$	0.56	0.56
$\Delta Hf(OsB_2)$	0.62	0.64
$\Delta Hf(OsB)$	0.46	0.47
$\Delta Hf(OsC)$	1.82	1.82
$\Delta Hf(IrB)$	0.42	4.28
$\Delta Hf(IrC)$	2.46	2.45
$\Delta Hf(Pt_2B)$	0.58	0.36
$\Delta Hf(PtB)$	0.38	0.36
$\Delta Hf(PtC)$	2.92	2.95
$\Delta Hf(AuB_2)$	1.98	2.00

- q is the defect charge state ($q > 0$, when electrons are removed and $q < 0$, when electrons are added).
- E_F is the Fermi energy, which varies in the band gap range of the defect-free supercell. E_{VBM} is referenced to the defect-free supercell's valence-band maximum (VBM). The obtained PBE band gap is used as the changing range of Fermi energy.

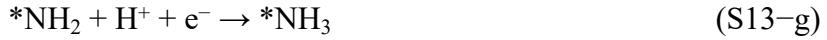
From the calculated defect formation energy, the transition level ($\varepsilon(q_1/q_2)$) is defined as the Fermi energy, for which the formation energies of the two charge states q_1 and q_2 of a defect are equal, and given as⁶¹

$$\varepsilon(q_1/q_2) = \frac{E_f(D^{q_1}) - E_f(D^{q_2})}{q_2 - q_1} \quad (\text{S12})$$

where $E_f(D^q)$ is the above-defined defect formation energy for defect D in a charge state q . The transition level of a defect is the Fermi level position, where the defect can accept/donate electrons. A deep-level means that removing a hole or electron from the trap to the valence or conduction band requires relatively high energy^{61, 62}.

Note S2 Computational hydrogen electrode

The pathways of the alternating, distal and enzymatic have been summarized in detail. In an acid electrolyte, the distal mechanism can be decomposed into the following elementary steps:



The alternating or enzymatic mechanisms can be decomposed into the following elementary steps:



where * represents the active site on the C₃B catalysts, and g represent the gas phase, respectively.

To compute the Gibbs free energy change (ΔG) of each elementary step of

electrochemical NH₃ synthesis, we adopted the computational hydrogen electrode (CHE) model developed by Nørskov et al.⁶³, according to which the ΔG of an electrochemical reaction is defined as:

$$\Delta G = \Delta E - T\Delta S + \Delta ZPE + \Delta G_U + \Delta G_{pH} \quad (S15)$$

Where ΔE is the reaction energy directly obtained from DFT calculations, and T is the temperature (T=298.15 K). ΔZPE and ΔS are the differences between the adsorbed species and the gas phase molecules in zero-point energy and entropy, respectively. ΔG_U is the free energy contribution related to applied potential U, which can be determined as ΔG_U = -eU. The U_L is defined by the equation U_L = -ΔG_{max}/e, where maximum Gibbs free energy change (ΔG_{max}) represents the highest positive Gibbs free energy change during the reduction process. In each NRR pathway, the step with the ΔG_{max} is referred to as the potential-determining step (PDS). ΔG_{pH} is the correction of the free energy of H⁺ ions by the concentration dependence of the entropy: ΔG_{pH} = 2.303 × k_B × pH, where k_B is the Boltzmann constant, and the value of pH is set to be 0.

Thus, the Gibbs free energy change for each step of the distal mechanism can be described as follows:

$$\Delta G_1 = \Delta G_{*NN} - \Delta G_{N2(g)} - G^* \quad (S16-a)$$

$$\Delta G_2 = \Delta G_{*NNH} - \Delta G_{*N2} - 1/2G^*H_2 \quad (S16-b)$$

$$\Delta G_3 = \Delta G_{*NNH2} - \Delta G_{*NNH} - 1/2G^*H_2 \quad (S16-c)$$

$$\Delta G_4 = \Delta G_{NH3} + \Delta G_{*N} - \Delta G_{*NNH2} - 1/2G^*H_2 \quad (S16-d)$$

$$\Delta G_5 = \Delta G_{*NH} - \Delta G_{*N} - 1/2G^*H_2 \quad (S16-e)$$

$$\Delta G_6 = \Delta G_{*NH2} - \Delta G_{*NH} - 1/2G^*H_2 \quad (S16-f)$$

$$\Delta G_7 = \Delta G_{*NH3} - \Delta G_{*NH2} - 1/2G^*H_2 \quad (S16-g)$$

The Gibbs free energy change for each step of the alternating or enzymatic mechanism can be described as follows:

$$\Delta G_1 = \Delta G_{*N2} - \Delta G_{N2(g)} - G^* \quad (S17-a)$$

$$\Delta G_2 = \Delta G_{*NNH} - \Delta G_{*N2} - 1/2G^*H_2 \quad (S17-b)$$

$$\Delta G_3 = \Delta G_{*NHNH} - \Delta G_{*NNH} - 1/2G^*H_2 \quad (S17-c)$$

$$\Delta G_4 = \Delta G_{*NHNH2} + \Delta G_{*N} - \Delta G_{*NHNH} - 1/2G^*H_2 \quad (S17-d)$$

$$\Delta G_5 = \Delta G_{*NH2NH2} - \Delta G_{*NHNH2} - 1/2G^*H_2 \quad (S17-e)$$

$$\Delta G_6 = \Delta G_{*NH2} - \Delta G_{*NH2NH2} - 1/2G^*H_2 \quad (S17-f)$$

$$\Delta G_7 = \Delta G_{*NH3} - \Delta G_{*NH2} - 1/2G^*H_2 \quad (S17-g)$$

The zero-point energies and entropies of the reaction species were calculated from

the vibrational frequencies. During these frequency calculations, all atoms of substrate were rigidly constrained so that no additional degrees of freedom, due to the catalyst, are introduced in to the reacting system. For gas phase molecule, the entropy term can be expressed as the sum of the translational, rotational and vibrational contributions, whereas for adsorbates the translational and rotational entropy were not taken into account due to negligible contributions. The thermodynamic quantities for the N₂, H₂, NH₃ species in the gas phase are shown in **Table S4**.

Table S4. Computed total energies (E) and thermodynamic quantities, in eV, for the gas phase N₂, H₂ and NH₃ species. (T = 298.15 K, P = 1 bar), at the 298.15 K, 0.035 bar. In comparison, the experimental entropies (TS_{exp}) of the gas phase N₂, H₂ and NH₃ are also shown, which are from NIST standard reference database (<https://doi.org/10.18434/T4D303>).

Species	E	TS	ZPE	G	TS _{exp}
N ₂	-16.63	0.59	0.15	-17.06	0.59
H ₂	-6.77	0.40	0.27	-6.90	0.41
NH ₃	-19.53	0.60	0.91	-19.22	0.60

Note S3 Machine learning (ML)

All the ML algorithms were carried out utilizing the open-source code Scikit-learn in Python3 environment⁶⁴. The ML process was conducted using the Gradient Boosted Regression (GBR), Random Forest Regression (RFR), Support Vector Regression, and the K-Neighbour Regression (KNR) models.

a. The computational method of the GBR regression algorithm

The best performing Gradient Boosted Regression (GBR) model is an integrated ML algorithm that is generated by the integration of weak regression trees^{65, 66}. Given the training samples D = (x₁, y₁), (x₂, y₂), ... (x_n, y_n) the number of leaf nodes in every regression tree is J. We divided the input data into J disjoint areas and defined each regression tree as t_m(x). The training goal of GBR is to minimize the loss function L, and the parameters of decision tree θ_m are determined through empirical risk

minimization:

$$\theta_m = \operatorname{argmin}_{\theta} \sum_{i=1}^n L(y_i, f_{m-1}(x_i) + t_m(x_i)) \quad (S18)$$

The process of GBR training is as follows:

- (a) Initialize a regression tree function $f_0(x)$.
- (b) Train GBR in the gradient decline direction, and compute the negative gradient value of the loss function as the estimated value of the residual. For the m_{th} iteration, GBR generates a regression tree according to the residual and updates the current function $f_m(x)$.
- (c) The final regression model is the weighted sum of several weak regression trees, which is defined as:

$$f_M(x) = \sum_{m=1}^M t(x; \theta_m) \quad (S19)$$

b. The computational method of the RFR regression algorithm

The Random Forest Regression (RFR) model is an ensemble learning method for regression by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees⁶⁷. Random decision forests correct for decision trees' habit of overfitting to their training set. The training algorithm for random forests applies the general technique of bootstrap aggregating, or bagging, to tree learners. Given a training set $X = x_1 \dots x_n$ with responses $Y = y_1 \dots y_n$, bagging repeatedly (B times) selects a random sample with replacement of the training set and fits trees to these samples:

- (a) For $b = 1 \dots B$: Sample, with replacement, n training examples from X, Y; call these X_b, Y_b ; Train a regression tree f_b on X_b, Y_b .
- (b) After training, predictions for unseen samples x' can be made by averaging the predictions from all the individual regression trees on x'

$$\tilde{f} = \frac{1}{B} \sum_{b=1}^B f_b(x) \quad (S20)$$

c. The computational method of the SVR regression algorithm

The Support Vector Regression (SVR) model is a supervised learning model with associated learning algorithms that analyze data used for regression analysis⁶⁸. Given a

set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on the side of the gap on which they fall. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. Training the original SVR to:

$$(a) \text{Minimize}_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|^2 \quad (S21)$$

$$(b) \text{Subject to } |y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b| \leq \varepsilon \quad (S22)$$

where \mathbf{x}_i is a training sample with a target value y_i . The inner product plus intercept $\langle \mathbf{w}, \mathbf{x}_i \rangle - b$ is the prediction for that sample, and ε is a free parameter that serves as a threshold: all predictions have to be within a ε range of the true predictions. Slack variables are usually added into the above to allow for errors and to allow approximation in the case the above problem is infeasible.

d. The computational method of the KNR regression algorithm

The K-Neighbour Regression (KNR) model is a non-parametric method proposed by Thomas Cover used for classification and regression⁶⁹. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k -NN is used for classification or regression. In k -NN regression, the output is the property value for the object. This value is the average of the values of k nearest neighbors.

In the classification phase, k is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the k training samples nearest to that query point. A commonly used distance metric for continuous variables is Euclidean distance. For discrete variables, such as for text classification, another metric can be used, such as the overlap metric (or Hamming distance). Often, the classification accuracy of k -NN can be improved significantly if the distance metric is learned with specialized algorithms such as Large Margin Nearest

Neighbor or Neighborhood components analysis. This algorithm works as follows:

- (a) Compute the Euclidean or Mahalanobis distance from the query example to the labeled examples.
- (b) Order the labeled examples by increasing distance.
- (c) Find a heuristically optimal number k of nearest neighbors, based on RMSE. This is done using cross-validation
- (d) Calculate an inverse distance weighted average with the k -nearest multivariate neighbors.

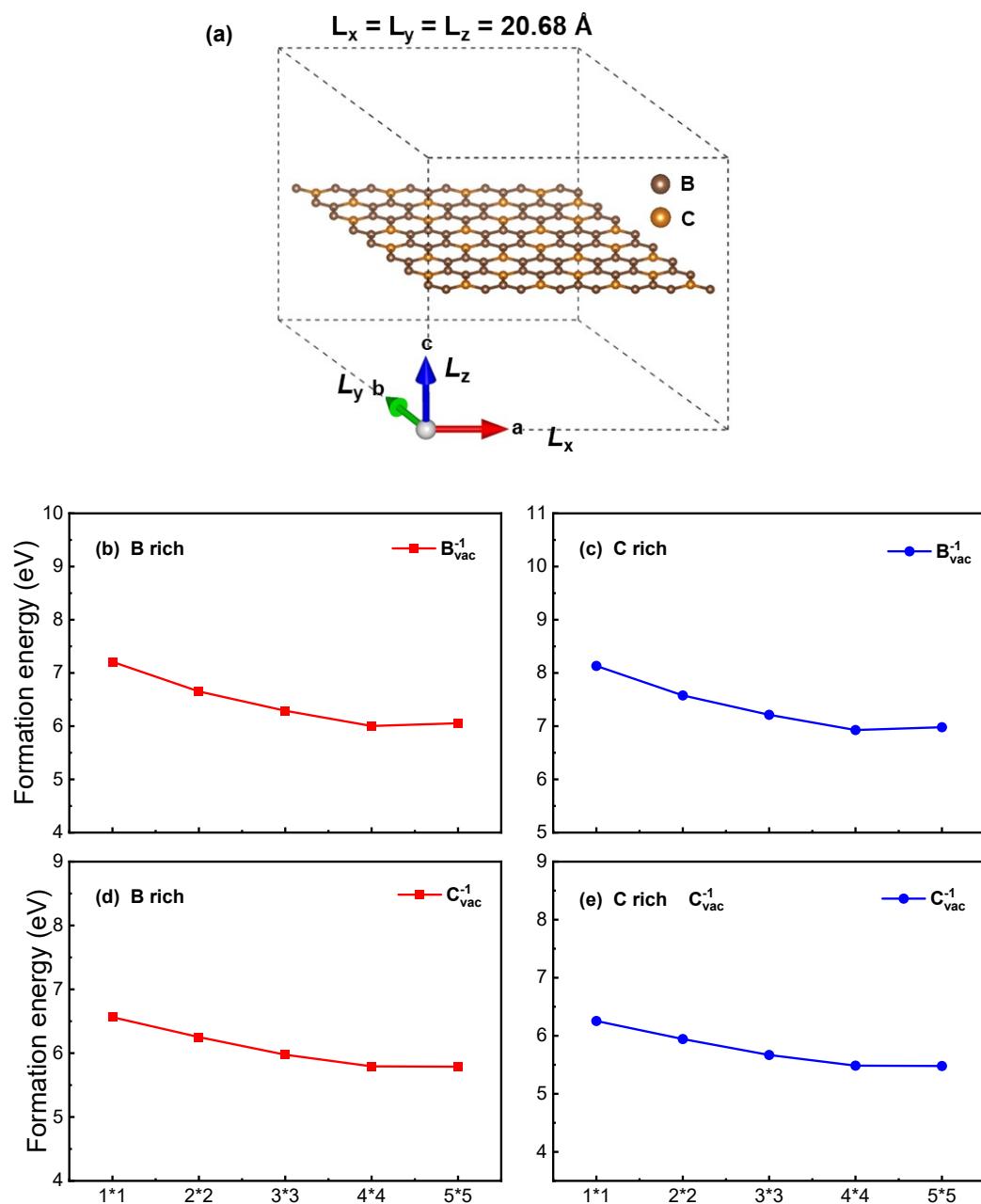


Fig. S1. (a) Uniformly scaled supercell of C₃B monolayer with the lateral lengths equal. (b)(c) (d)(e) Convergence of defect (B and C vacancies in C₃B in -1 charge state) formation energies as a function of uniformly scaled supercell size in all the directions (lateral length: L_x = L_y = L_z = 20.68 Å). The bottom x-axis represents the supercell size, which * denotes the cubic supercell size. For example, 4*4 illustrates 4×4 × 1 supercell size.

Our calculated lattice parameters of the C₃B monolayer are $a = b = 5.17$ Å, with C-C and C-B bond lengths of 1.42 Å and 1.56 Å, respectively, which align well with previously reported data^{70, 71}. In this paper, different charge states of the system are modeled by introducing a uniform background charge (positive or negative) in the supercell. However, this method may induce spurious Coulomb interactions, leading to errors in calculating defect formation energy in supercells. In order to correct the spurious Coulomb interaction, our calculations for the defect formation energies in the C₃B monolayer are performed uniformly in a supercell with L_x = L_y = L_z^{72, 73}. This method has been used successfully to investigate the defect formation energies of several 2D materials in the field of electrocatalysis.⁷⁴⁻⁷⁷ To assess the impact of supercell sizes on defect formation energy, we calculate the formation energies of C and B vacancy defects in the -1 charge state using different supercell models: 1*1 (8 atoms), 2*2 (32 atoms), 3*3 (72 atoms), 4*4 (128 atoms), and 5*5 (200 atoms) with equal lateral lengths (L_x = L_y = L_z). Here, * denotes the cubic supercell size. For example, 4*4 illustrates 4 × 4 × 1 supercell size. As shown in **Fig. S1**, our results show that the defect formation energy converges quite well at the supercell size of 4 * 4 (L_x = L_y = L_z = 20.68 Å). Thus, we use a supercell size of 4 × 4 × 1 (128 atoms) for all calculations. Correspondingly, the vacuum layer in this supercell is 20.68 Å, which can effectively avoid the periodic interactions along the vacuum direction.

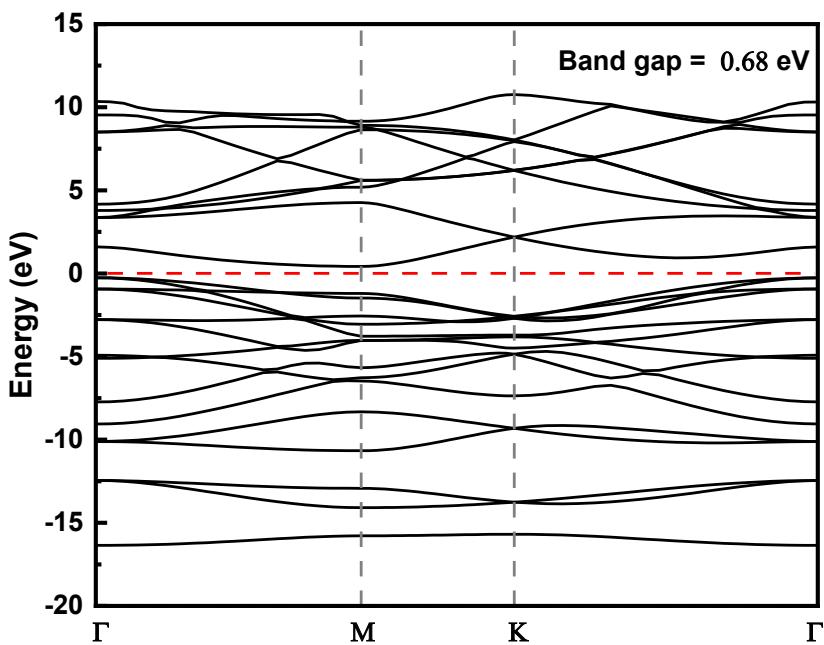


Fig. S2. Band structure of pristine unit cell of C₃B monolayer calculated with PBE functional. The band gap is 0.68 eV, and the Fermi level is set to 0 eV.

Note S4 GGA+U

Considering the strong correlation effect of transition metals, we used the GGA+U method to further refine the overpotential of V_C¹⁺@C₃B. In this paper, U is introduced as an empirical parameter, with U_{eff} = 3.1 eV applied to the 3d orbital of V, as derived from earlier computational and experimental studies⁷⁸. The Gibbs free energy diagrams for V_C¹⁺@C₃B, calculated using the GGA + U method, are presented in **Fig. S3** and compared with results obtained from the PBE method. In the first hydrogenation step, the energy difference between PBE and PBE+U is 0.39 eV and 0.63 eV, respectively, while in the subsequent steps, the difference between them is small. The overpotential of V_C¹⁺@C₃B for the PBE approach is 0.39 V, whereas it is 0.63 V for the PBE + U approach. These findings suggest that the PBE method is adequate for describing the electrocatalytic activity, in line with previous research^{79, 80}. However, small variations in the U-value of transition metals can significantly affect the NRR overpotential. As a result, the U value has a substantial impact on the NRR overpotential. Currently, the specification of U is often weak, with many authors arbitrarily selecting U values for transition metals, leading to confusion in the conclusions of NRR overpotential.

calculations. To avoid this ambiguity, in this paper we employed the PBE method for our calculations and focused on screening systems with promising NRR activity.

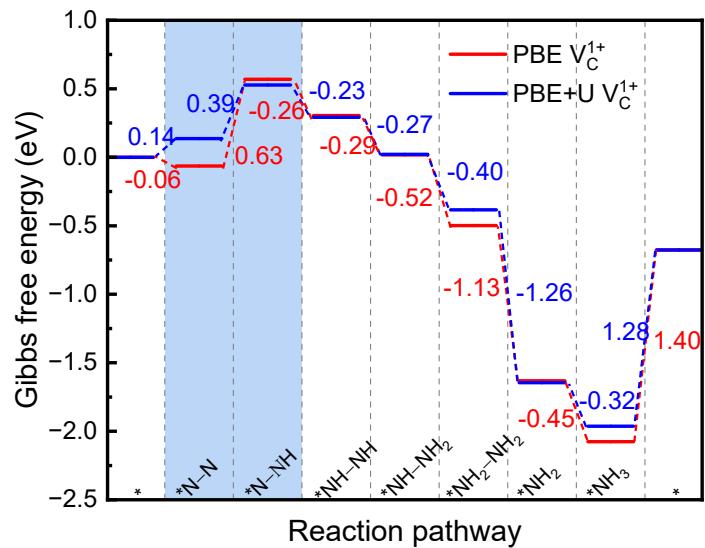


Fig. S3. Gibbs free energy diagrams of $V_c^{1+}@\text{C}_3\text{B}$ using PBE calculation and PBE +U calculation.

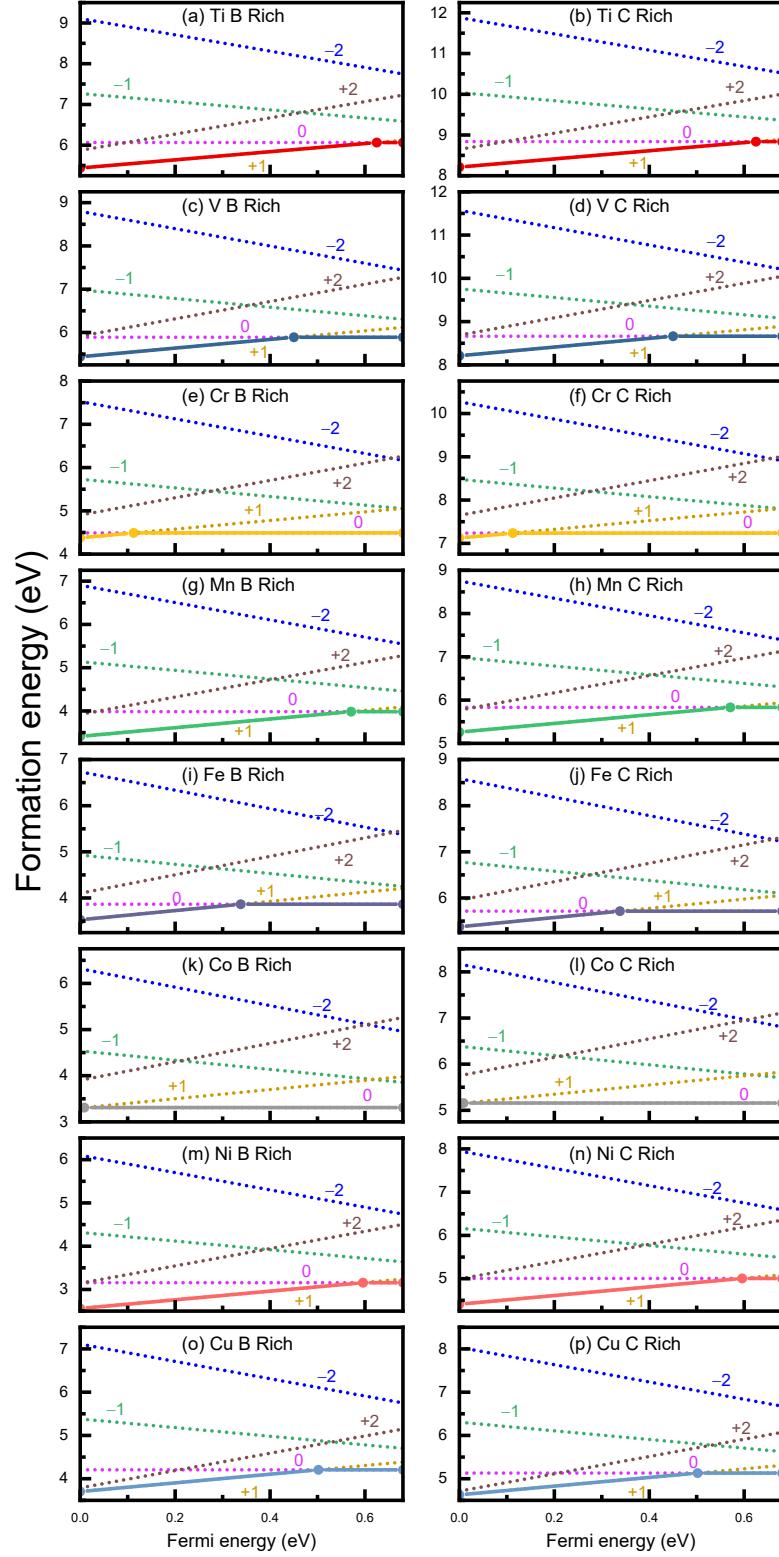


Fig. S4. The formation energies of 3d–TM_B@C₃B under B-rich and C-rich conditions are plotted as a function of the Fermi level with respect to the VBM. The Fermi level (E_F) at the VBM and the conduction band minimum (CBM) is set to 0.00 and 0.68 eV, respectively. The solid line represents the charge state with the lowest formation energy of all possible charge states.

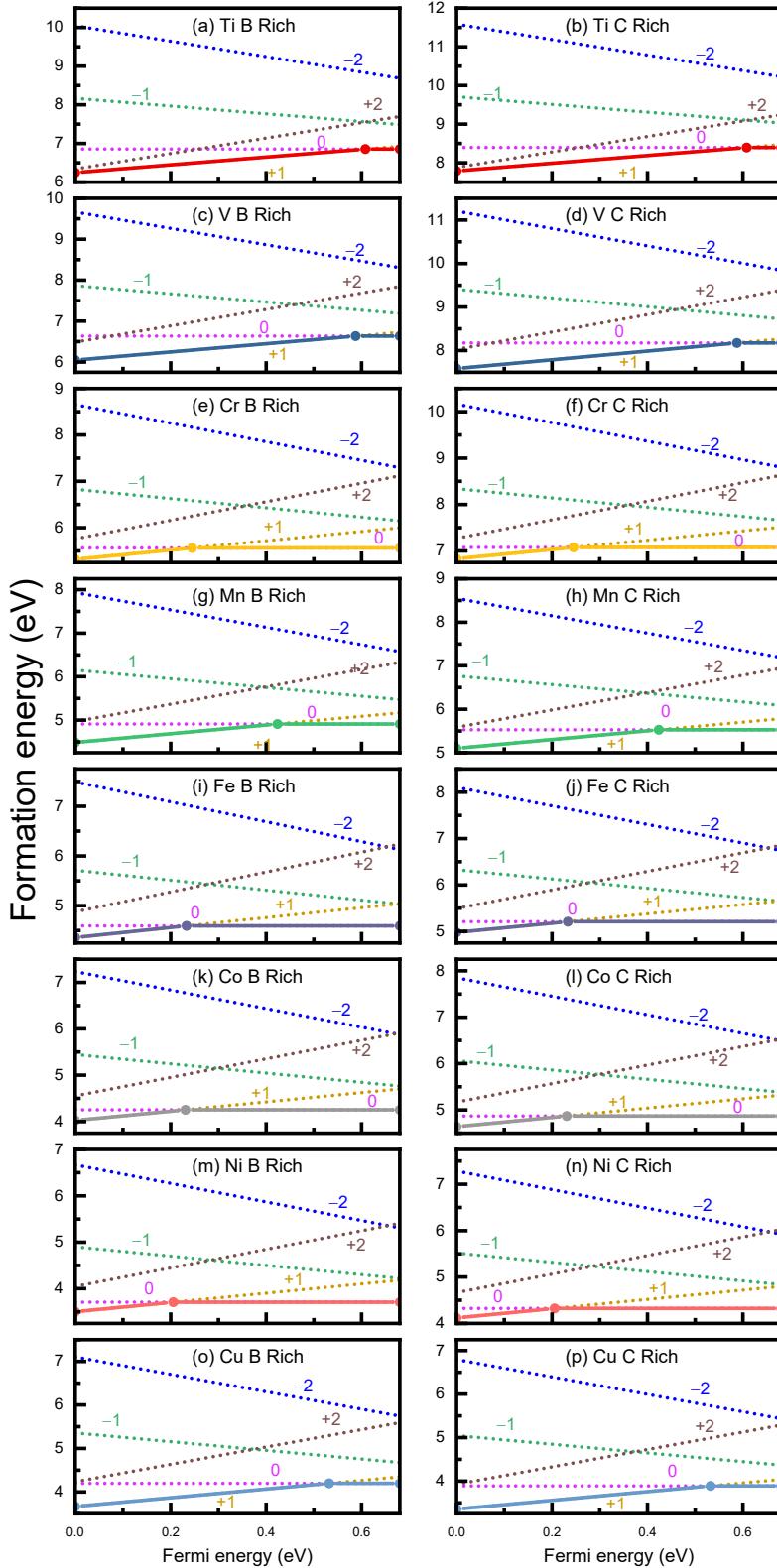


Fig. S5. The formation energies of 3d-TM_C@C₃B under B-rich and C-rich conditions are plotted as a function of the Fermi level with respect to the VBM. The Fermi level (E_F) at the VBM and the conduction band minimum (CBM) is set to 0.00 and 0.68 eV, respectively. The solid line represents the charge state with the lowest formation energy of all possible charge states.

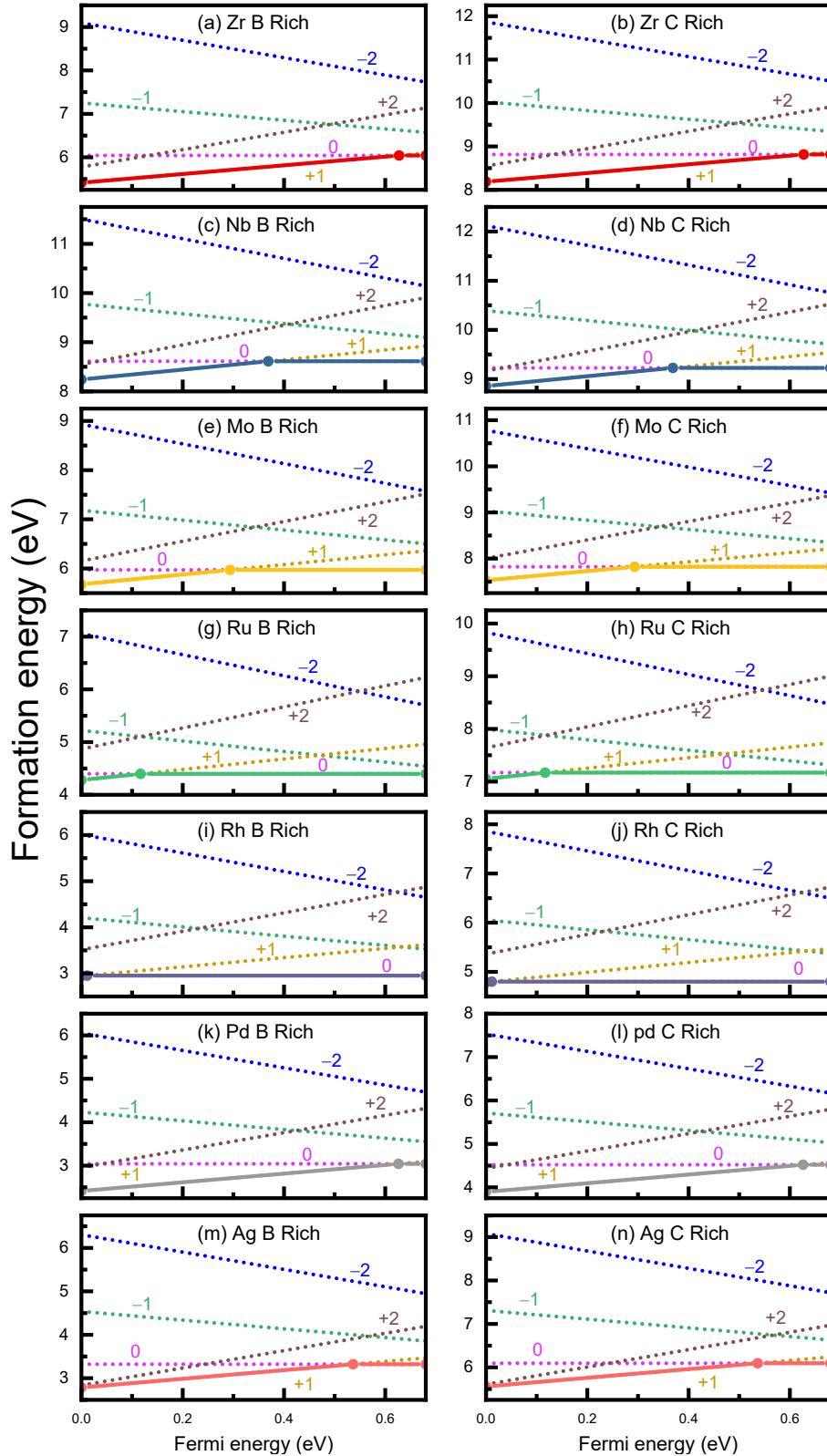


Fig. S6. The formation energies of 4d-TM_B@C₃B under B-rich and C-rich conditions are plotted as a function of the Fermi level with respect to the VBM. The Fermi level (E_F) at the VBM and the conduction band minimum (CBM) is set to 0.00 and 0.68 eV, respectively. The solid line represents the charge state with the lowest formation energy of all possible charge states.

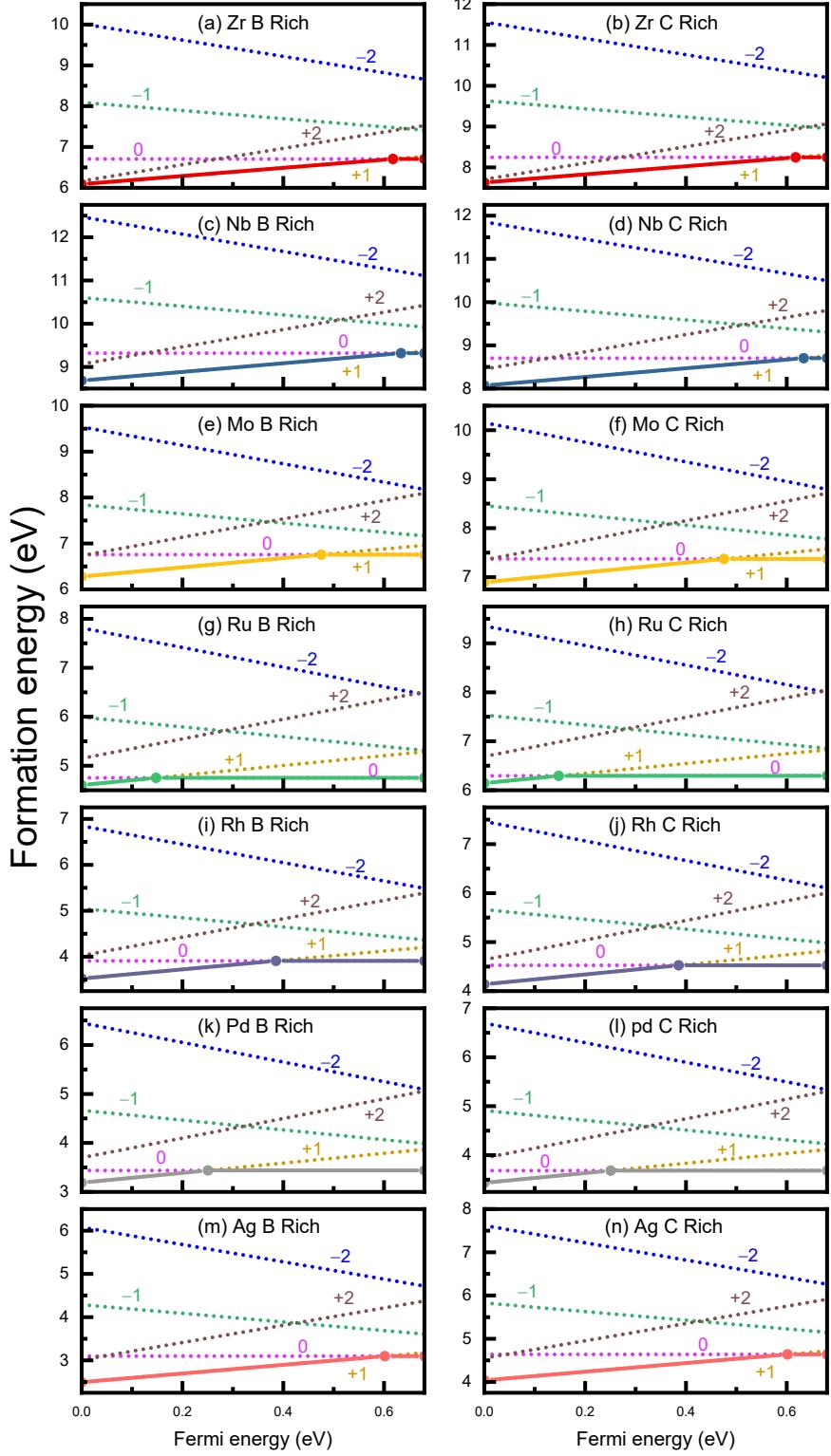


Fig. S7. The formation energies of 4d-TM_C@C₃B under B-rich and C-rich conditions are plotted as a function of the Fermi level with respect to the VBM. The Fermi level (E_F) at the VBM and the conduction band minimum (CBM) is set to 0.00 and 0.68 eV, respectively. The solid line represents the charge state with the lowest formation energy of all possible charge states.

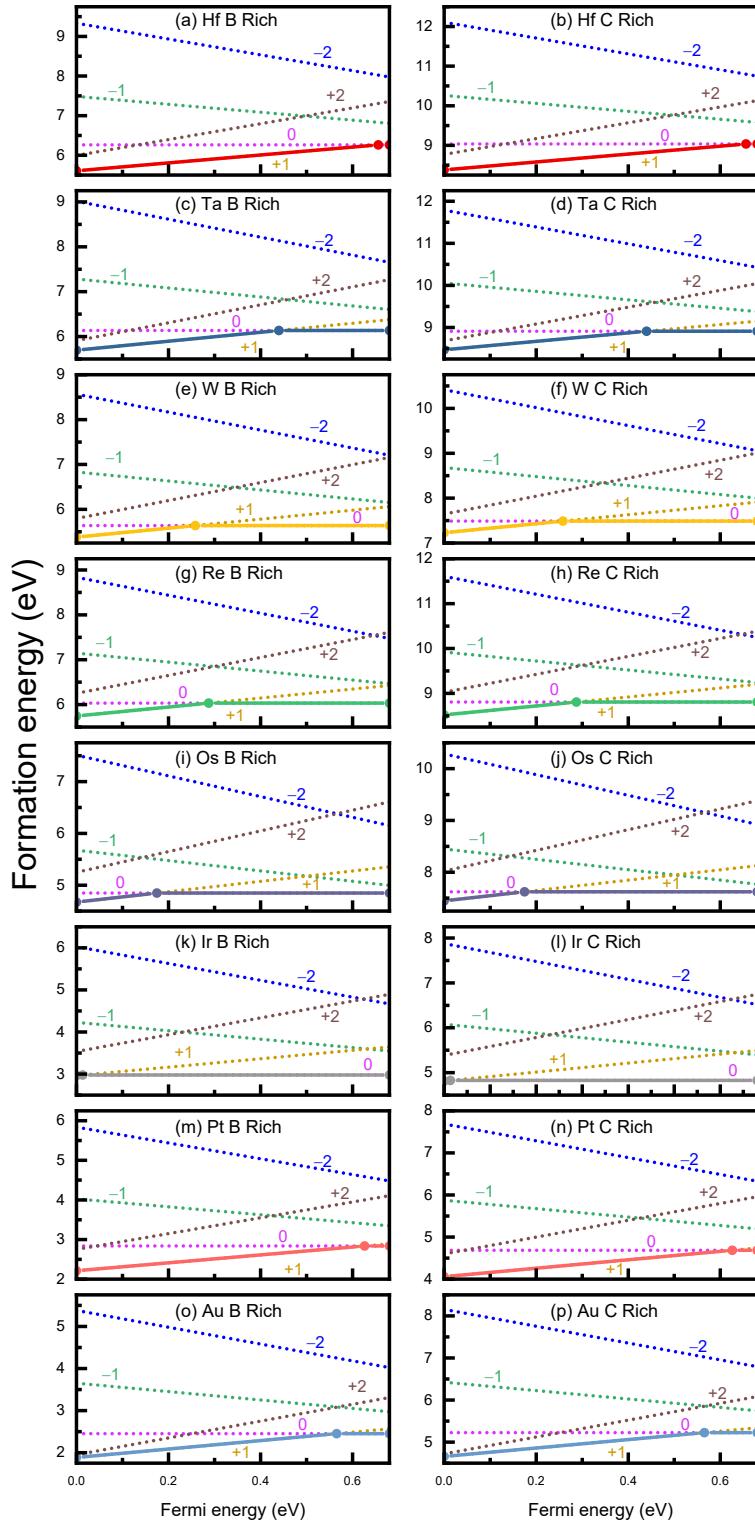


Fig. S8. The formation energies of 5d–TM_B@C₃B under B–rich and C–rich conditions are plotted as a function of the Fermi level with respect to the VBM. The Fermi level (E_F) at the VBM and the conduction band minimum (CBM) is set to 0.00 and 0.68 eV, respectively. The solid line represents the charge state with the lowest formation energy of all possible charge states.

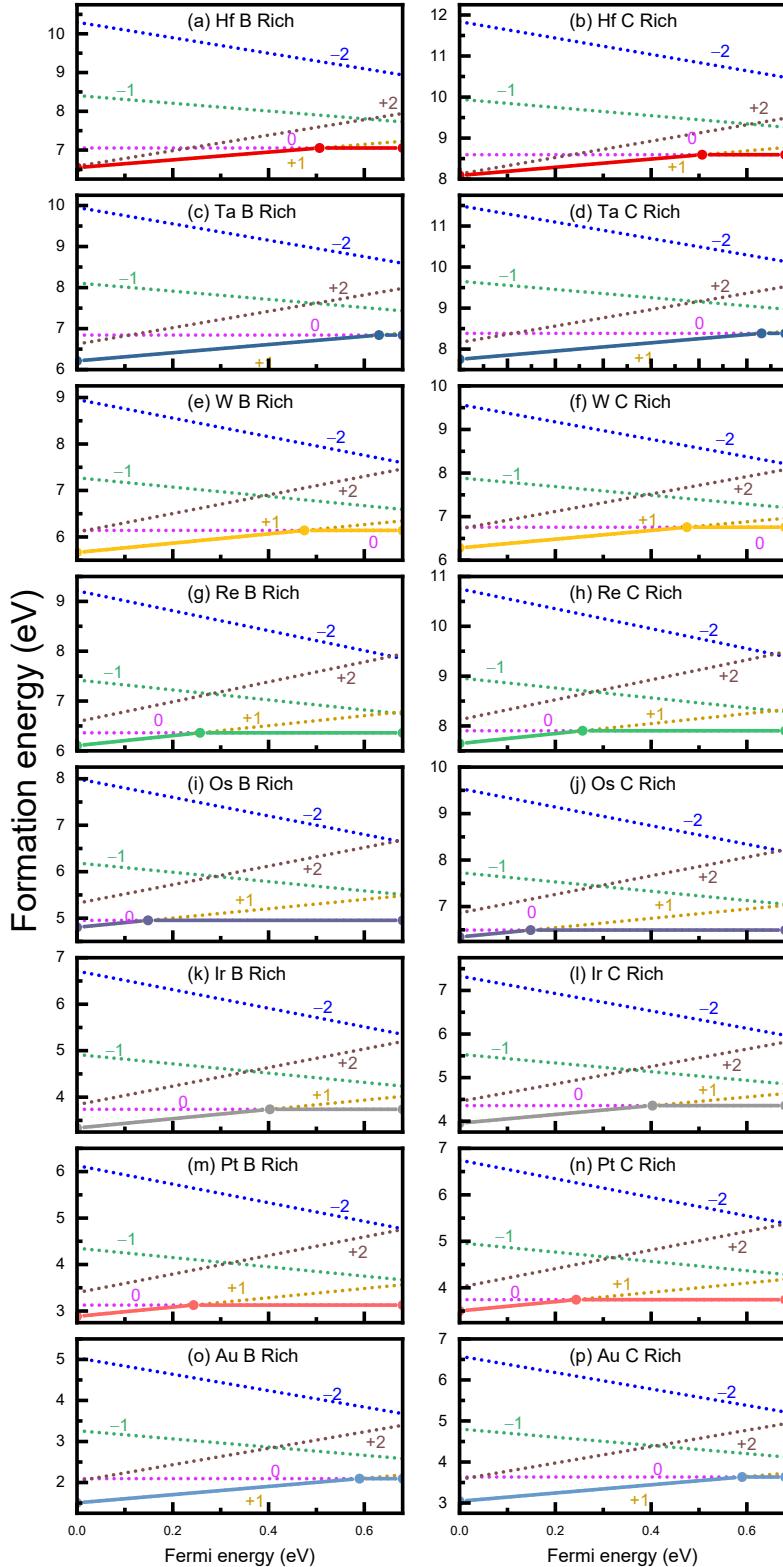


Fig. S9. The formation energies of 5d–TM_C@C₃B under B–rich and C–rich conditions are plotted as a function of the Fermi level with respect to the VBM. The Fermi level (E_F) at the VBM and the conduction band minimum (CBM) is set to 0.00 and 0.68 eV, respectively. The solid line segments represent the charge state with the lowest formation energy of all possible charge states.

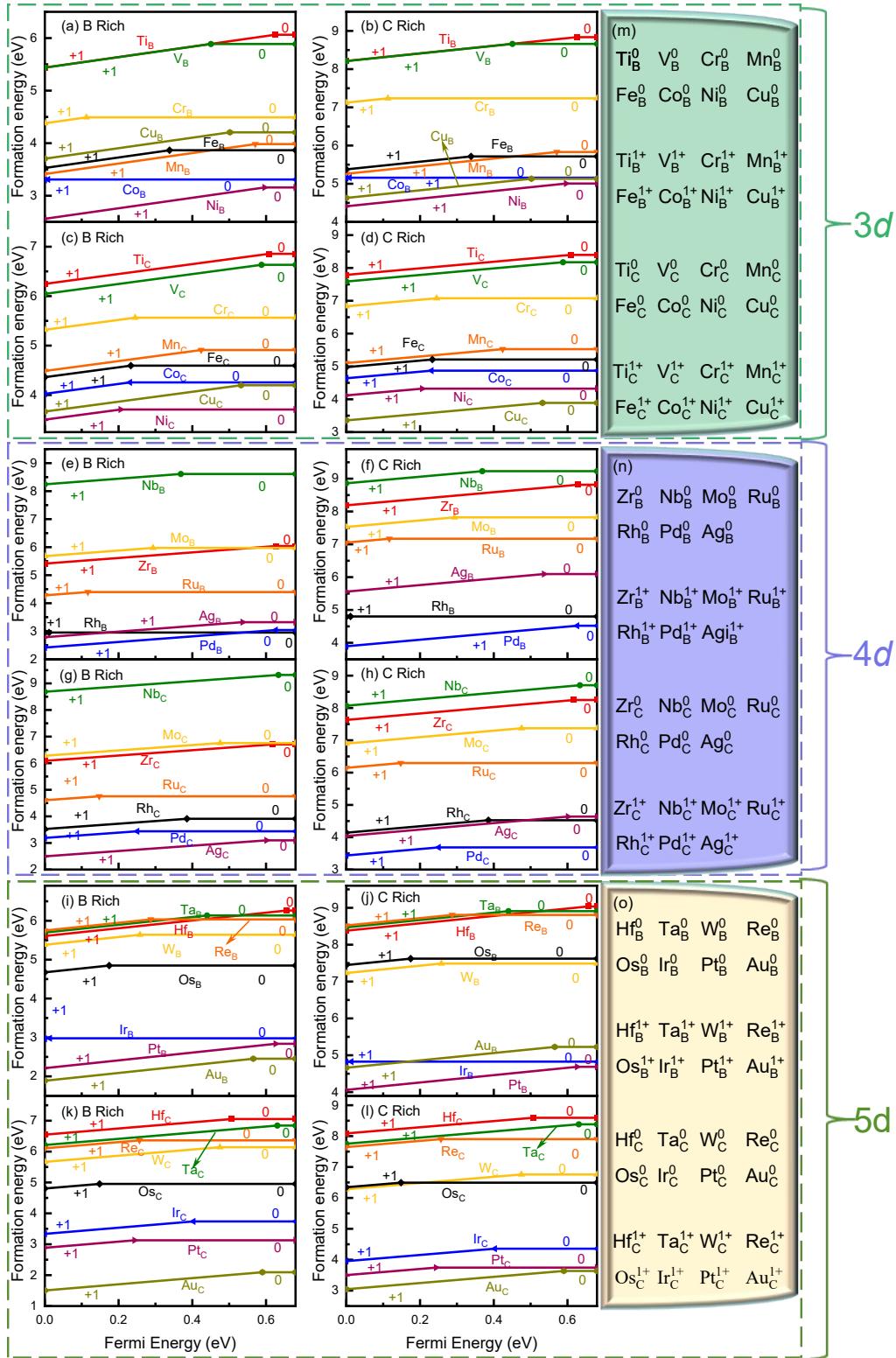


Fig. S10. The formation energies of 3d (a-d), 4d (e-f), and 5d (i-l) TM doped B and C sites C_3B monolayer under B-rich and C-rich conditions as a function of the E_F . C_3B monolayer stabilized charge states at the B and C sites of 3d (m), 4d (n), and 5d (o) doped TMs. The E_F varies between the valence band maximum (VBM, 0.00 eV) and the conduction band minimum (CBM, 0.68 eV). The TM atoms and their stable charge states considered in this paper are shown.

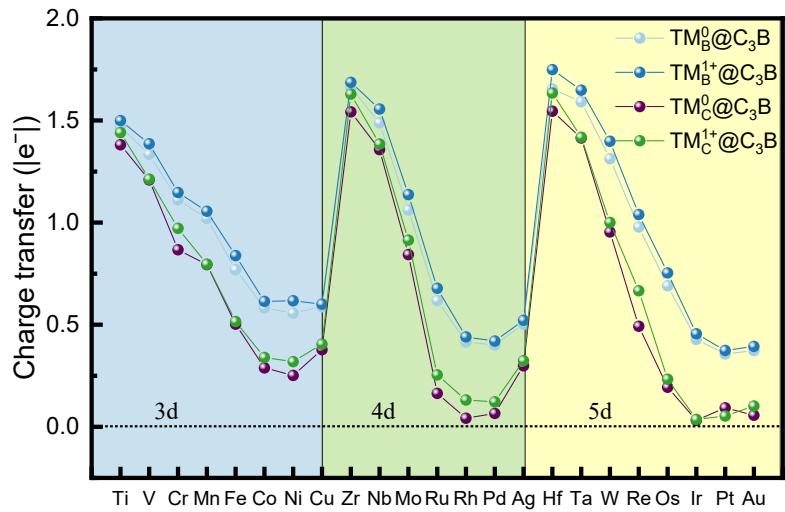


Fig. S11. The charge transfer of $3d$, $4d$, and $5d$ TM@ C_3B (B and C sites) in the 0 and +1 charge states.

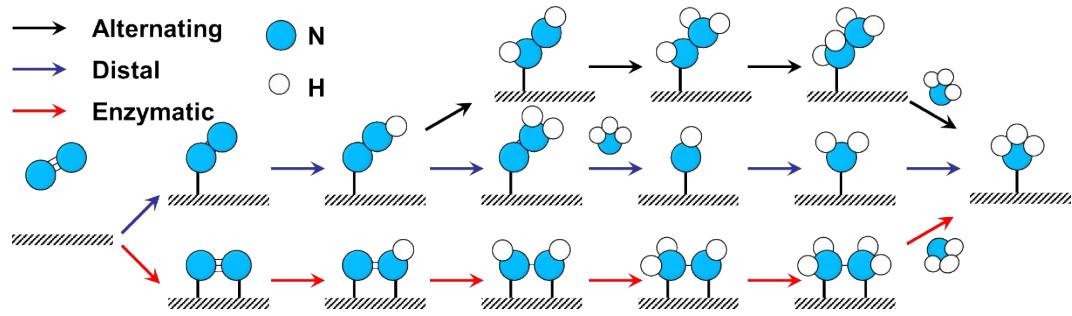


Fig. S12. Alternating, distal, and enzymatic reaction pathways for NRR.

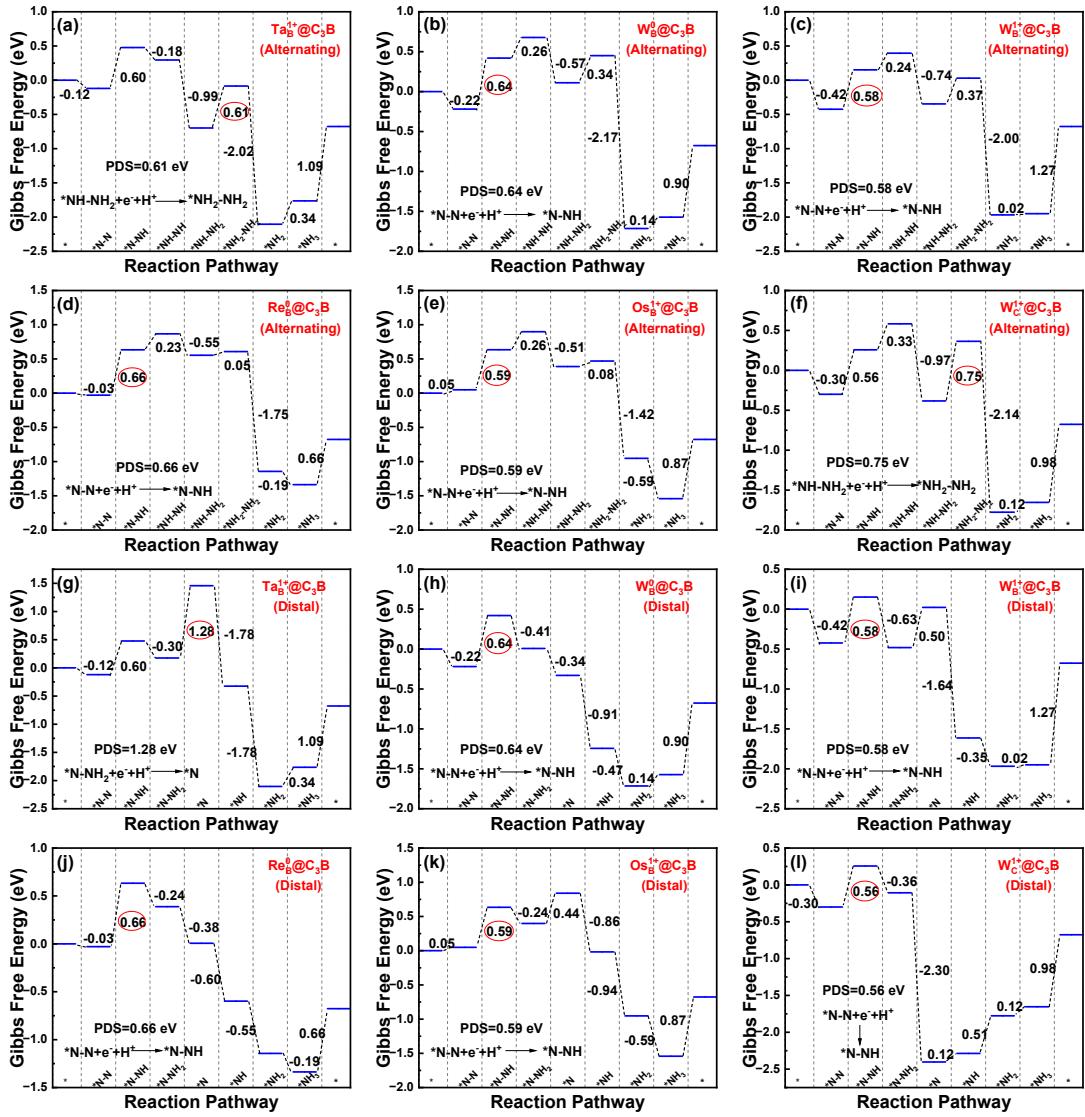
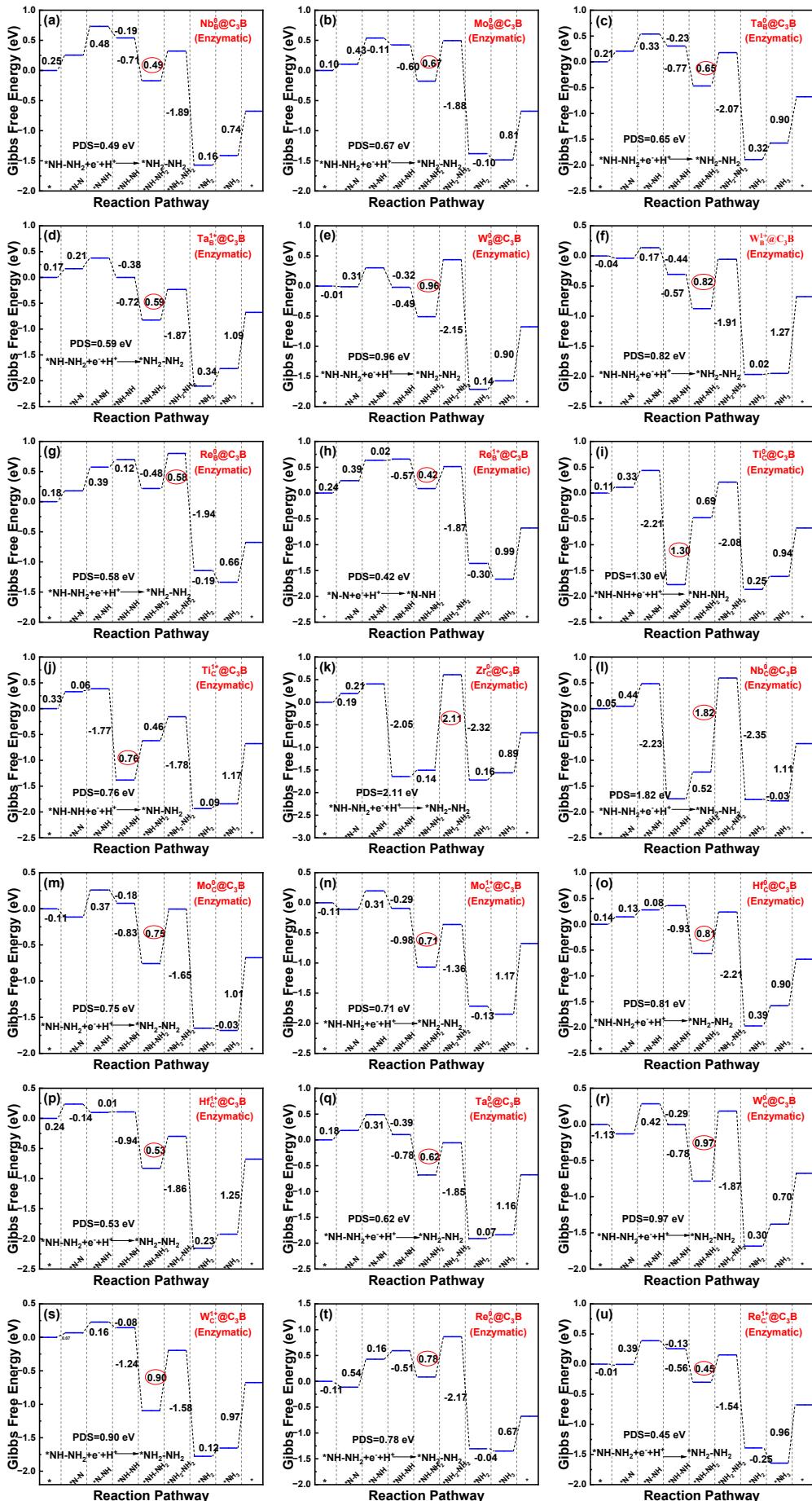


Fig. S13. Calculated free energy diagrams for NRR through alternating and distal mechanism catalyzed by TM@C₃B (TM = Ta_B¹⁺, W_B⁰, W_B¹⁺, Re_B⁰, Os_B¹⁺, and W_C¹⁺).



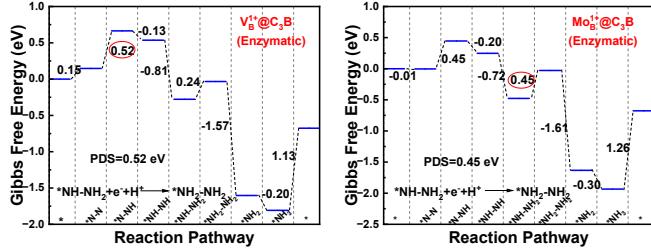


Fig. S14. Calculated free energy diagrams for NRR through enzymatic mechanism catalyzed by TM@C₃B (TM = V_B¹⁺, Nb_B⁰, Mo_B⁰, Mo_B¹⁺, Ta_B⁰, Ta_B¹⁺, W_B⁰, W_B¹⁺, Re_B⁰, Re_B¹⁺, Ti_C⁰, Ti_C¹⁺, Zr_C⁰, Nb_C⁰, Mo_C⁰, Mo_C¹⁺, Hf_C⁰, Hf_C¹⁺, Ta_C⁰, W_C⁰, W_C¹⁺, Re_C⁰, and Re_C¹⁺).

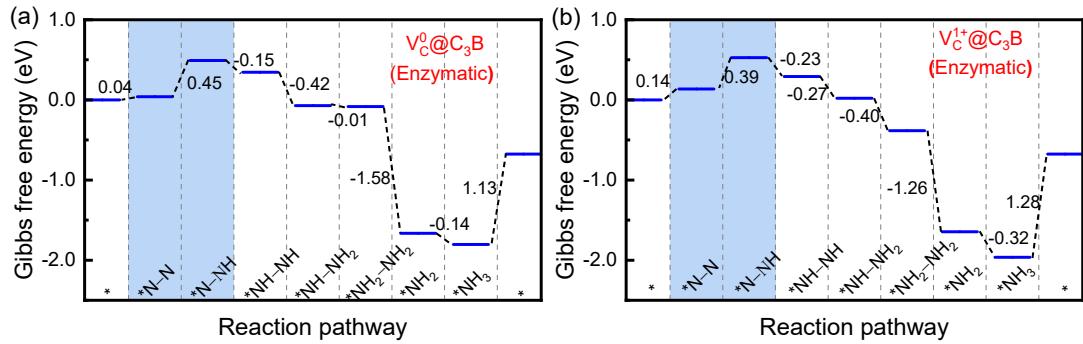


Fig. S15. Calculated free energy diagrams for NRR through enzymatic mechanism catalyzed by (a) $V_C^0@C_3B$, and (b) $V_C^{1+}@C_3B$.

We explore the details of the enzymatic mechanism of NRR on the promising $V_C^0@C_3B$ and $V_C^{1+}@C_3B$, as shown in **Fig. S15**. In the reduction process, adsorbed *N₂ can interact with H⁺/e⁻ pairs to form *NNH. The PDS occurs during the first hydrogenation step, forming the *NNH species with ΔG values of 0.45 and 0.39 eV, respectively. This small barrier can easily be overcome at room temperature or decreased with more negative voltage, which is superior to electrocatalysts reported in the literature, such as V@MoP (-0.65 V)⁸¹, RuB@GY (-0.53 V)⁸² and Fe@Nb₂C (-0.47 V)⁸³.

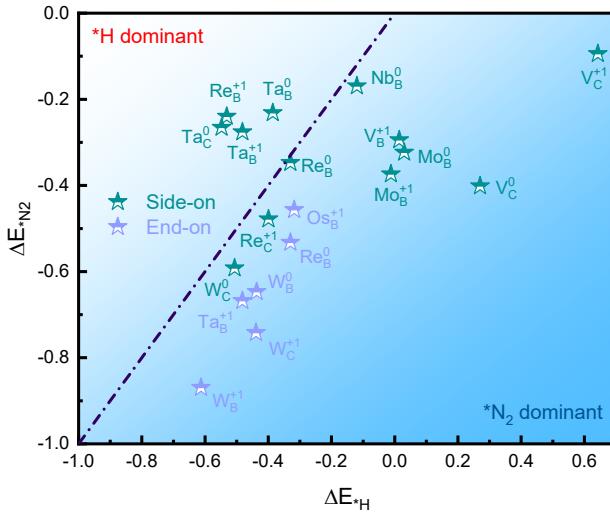


Fig. S16. Adsorption energies of N_2 and H atom for the HER and NRR on various catalysts.

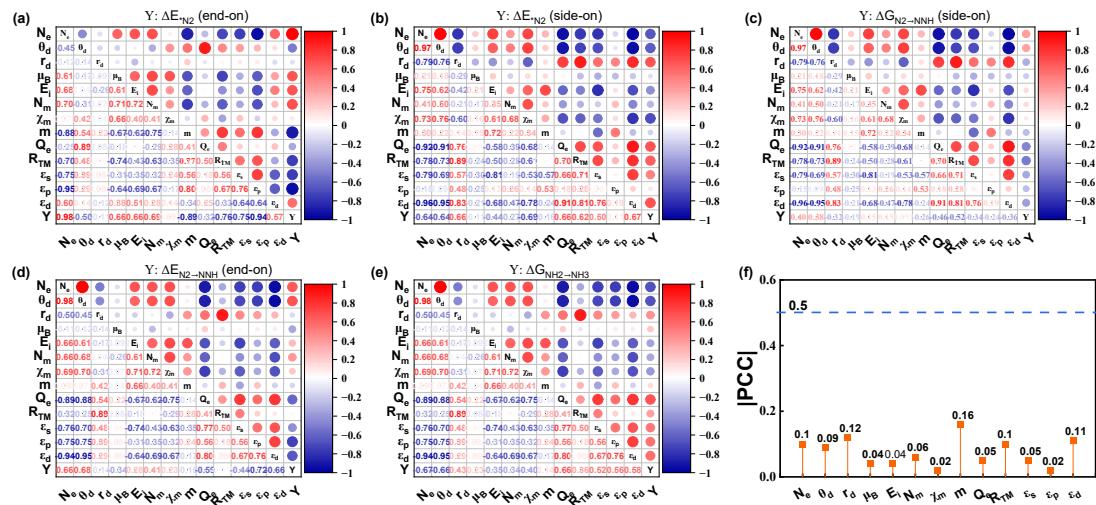


Fig. S17. Pearson correlation coefficients (PCC) plot of 13 input features with (a) $\Delta E_{\text{N}2}$ (end-on), (b) $\Delta E_{\text{N}2}$ (side-on), (c) $\Delta G_{\text{N}2 \rightarrow \text{N}2\text{H}}$ (side-on), (d) $\Delta G_{\text{N}2 \rightarrow \text{N}2\text{H}}$ (end-on) and (e) $\Delta G_{\text{NH}2 \rightarrow \text{NH}3}$. Red, blue, and white colors represent direct, indirect, and no correlations, respectively. (f) $|\text{PCC}|$ of the 13 input features.

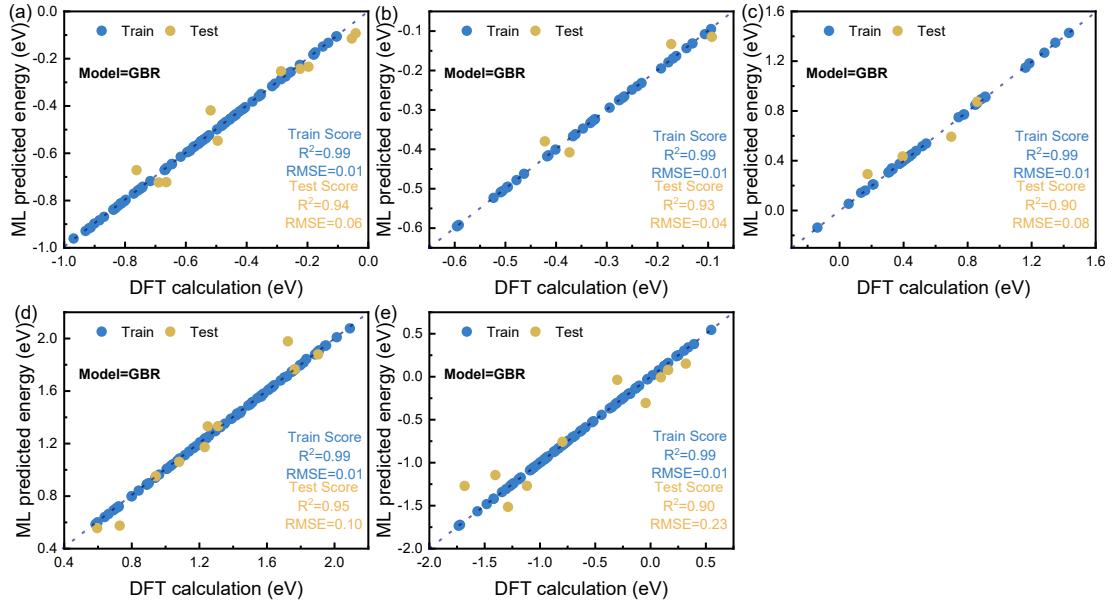


Fig. S18. (a)–(e) ΔE_{*N_2} (end-on), ΔE_{*N_2} (side-on), $\Delta G_{N_2 \rightarrow NNH}$ (enzymatic), $\Delta G_{N_2 \rightarrow NNH}$ (distal), $\Delta G_{NH_2 \rightarrow NH_3}$ of DFT calculation vs ML predicted energy in the neutral and charged state.

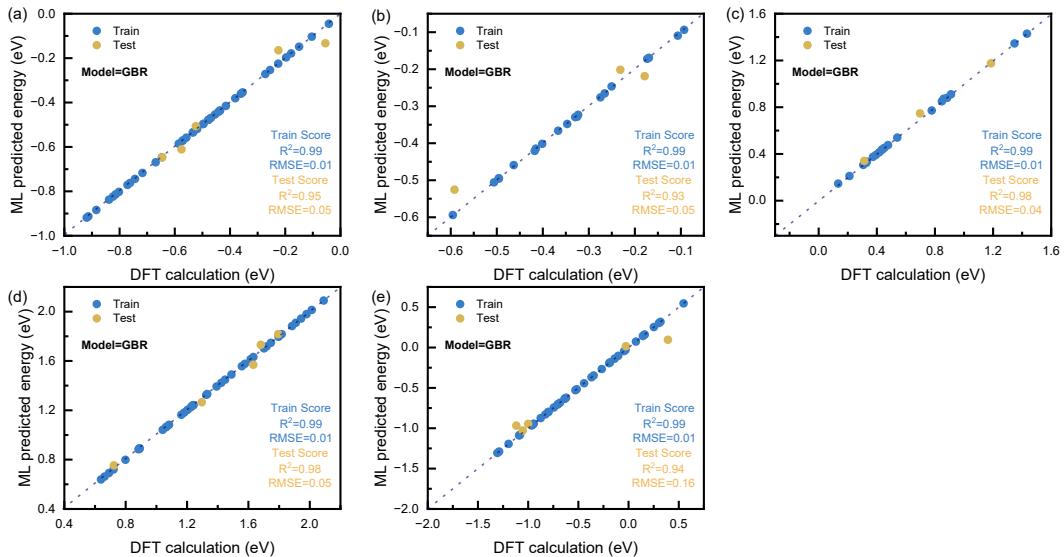


Fig. S19. (a)–(e) ΔE_{*N_2} (end-on), ΔE_{*N_2} (side-on), $\Delta G_{N_2 \rightarrow NNH}$ (enzymatic), $\Delta G_{N_2 \rightarrow NNH}$ (distal), $\Delta G_{NH_2 \rightarrow NH_3}$ of DFT calculation vs ML predicted energy in the neutral state.

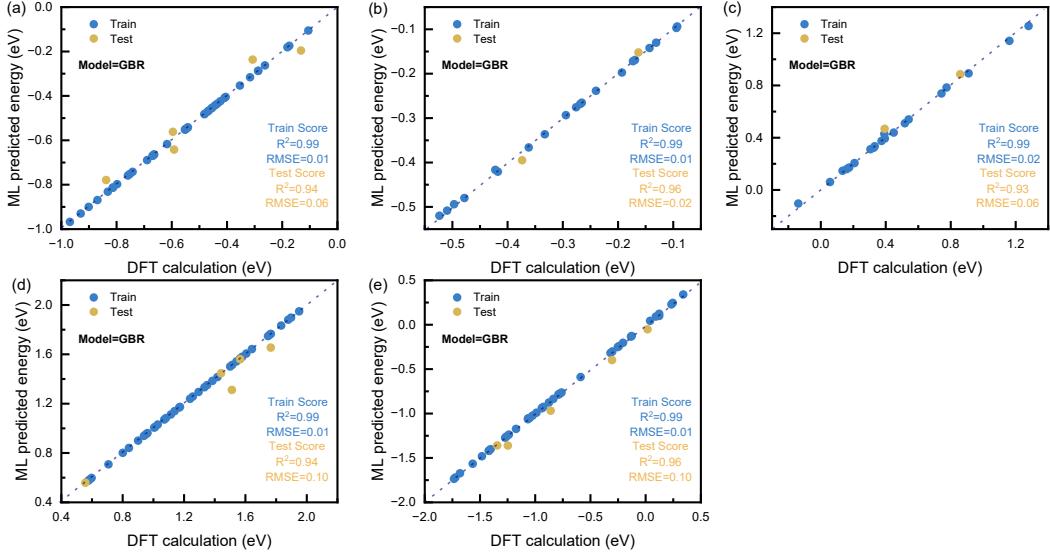


Fig. S20. (a)-(e) ΔE_{*N_2} (end-on), ΔE_{*N_2} (side-on), $\Delta G_{N_2 \rightarrow NNH}$ (enzymatic), $\Delta G_{N_2 \rightarrow NNH}$ (distal), $\Delta G_{NH_2 \rightarrow NH_3}$ of DFT calculation vs ML predicted energy in the charged state.

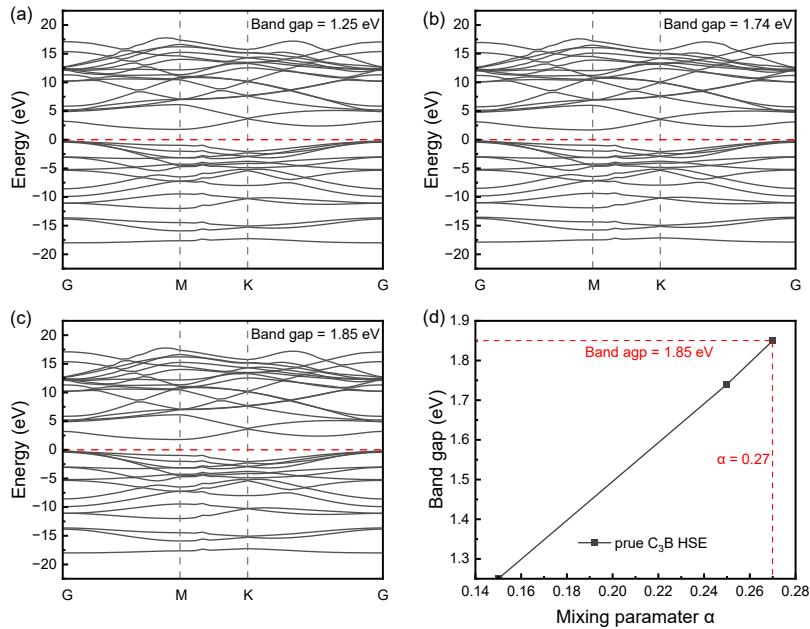


Fig. S21. The band structure of C_3B (unit cell) using the HSE06 method with different Hartree–Fock mixing parameters α , (a) $\alpha=0.15$, (b) $\alpha=0.25$, and (c) $\alpha=0.27$. The valence band maximum (VBM) position is placed at zero. (d) The bandgap as a function of Hartree–Fock mixing parameters α .

The HSE hybrid functional is adopted to accurately describe the PDOS of $V_C^0@C_3B$ and $V_C^{1+}@C_3B$ with the screening parameter (ω) set to 0.2 \AA^{-1} . We test the bandgap of pure C_3B by adjusting the Hartree–Fock mixing parameter (α) from 0.15 to 0.25. As shown in **Fig. S21**, we find that when $\alpha = 0.27$, the calculated bandgap is 1.85 eV for pure C_3B , close to the bandgap (1.84 eV) obtained by Zhang et al.⁸⁴.

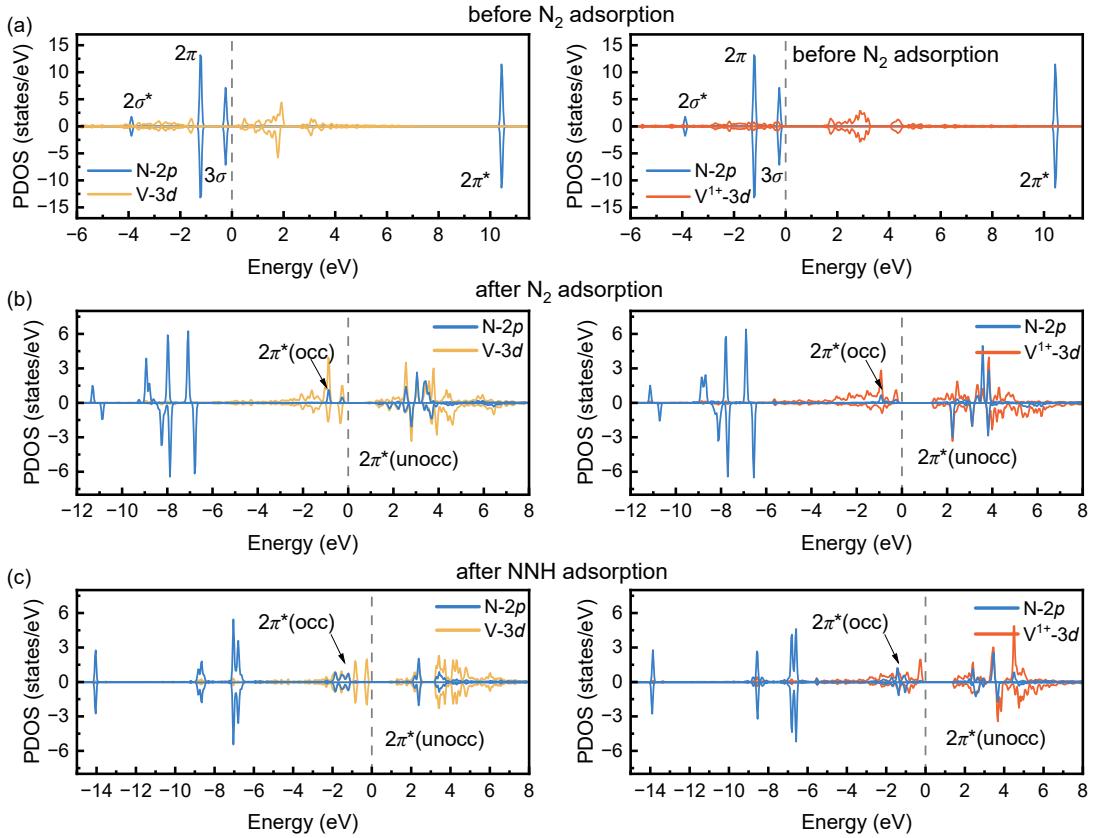


Fig. S22. (a) The partial density of states (PDOS) of V-3d in $V_{C^0}@C_3B$ and $V_{C^{1+}}@C_3B$ before N_2 adsorption; the PDOS of N_2 is also given for comparison. (b) The PDOS of V-3d and N-2p in $V_{C^0}@C_3B$ and $V_{C^{1+}}@C_3B$ after N_2 adsorption. (c) The PDOS of V-3d and N-2p in $V_{C^0}@C_3B$ and $V_{C^{1+}}@C_3B$ after NNH adsorption. PDOS uses the Heyd–Scuseria–Ernzerhof (HSE) hybrid functional approach, where the E_F is set to zero.

In **Fig. S22a**, the frontier orbitals of the N_2 can be observed, labeled as $2\sigma^*$, 2π , 3σ , and $2\pi^*$, respectively. For the free N_2 , the $2\sigma^*$, 2π , and 3σ orbitals lie below the E_F , while the antibonding $2\pi^*$ orbital is located above the E_F , accounting for the high stability of the N_2 . However, after N_2 adsorption (**Fig. S22b**), the $2\pi^*$ orbital of N_2 obviously decreases, and there is a significant hybridization between the N-2p orbitals and the V-3d orbitals, especially below the E_F . This hybridization primarily involves the transfer of electrons from N_2 to the vacant V-3d orbitals. Simultaneously, electrons are transferred from the V-3d orbitals to the vacant $2\pi^*$ orbitals of N_2 , causing the $2\pi^*$ orbitals to split and hybridize with the 3d orbitals of V, which forms occupied (unoccupied) $2\pi^*$ orbitals located above (below) the E_F . This "donation and back-donation" mechanism⁸⁵ strengthens the TM-N bond while weakening the N≡N bond. After NNH adsorption, in the $V_{C^0}@C_3B$ and $V_{C^{1+}}@C_3B$, there is a significant hybridization between the N-2p and V-3d orbitals, as shown in **Fig. S22c**.

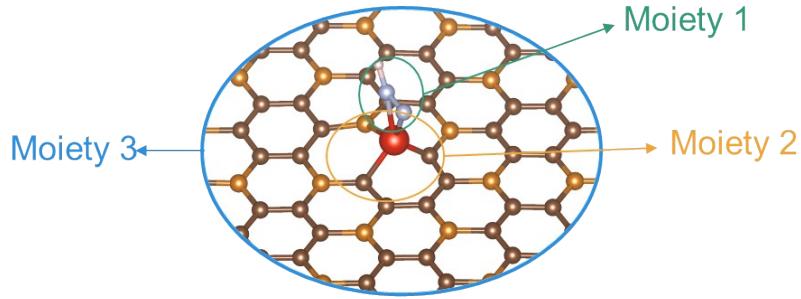


Fig. S23. schematic diagram of three moieties for $\text{V}_\text{C}^0@\text{C}_3\text{B}$ and $\text{V}_\text{C}^{1+}@\text{C}_3\text{B}$.

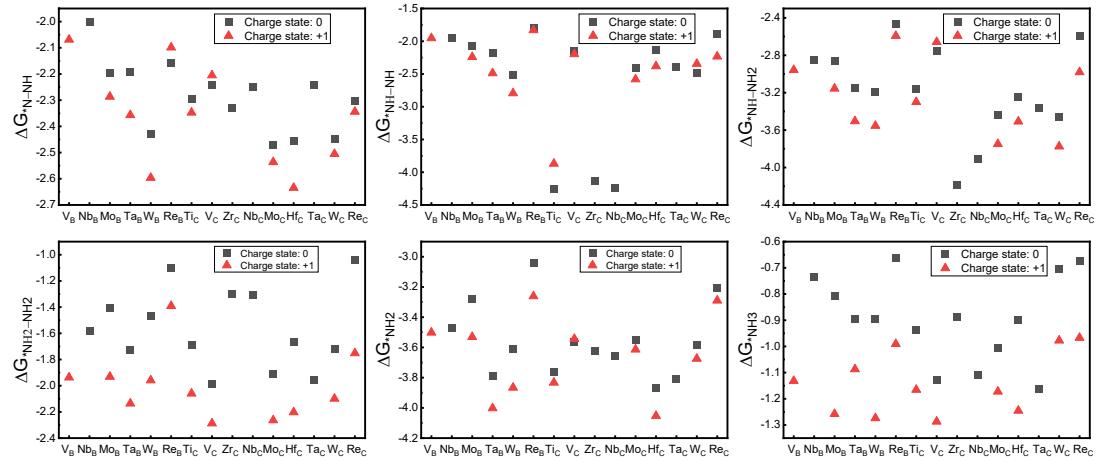


Fig. S24. Adsorption free energies of intermediates (*NNH, *NHNH, *NHNH₂, *NH₂NH₂, *NH₃) for TM@C₃B (V_B^{1+} , Nb_B^0 , Mo_B^0 , Mo_B^{1+} , Ta_B^0 , Ta_B^{1+} , W_B^0 , W_B^{1+} , Re_B^0 , Re_B^{1+} , Ti_C^0 , Ti_C^{1+} , V_C^0 , V_C^{1+} , Zr_C^0 , Nb_C^0 , Mo_C^0 , Mo_C^{1+} , Hf_C^0 , Hf_C^{1+} , Ta_C^0 , W_C^0 , W_C^{1+} , Re_C^0 , Re_C^{1+}).

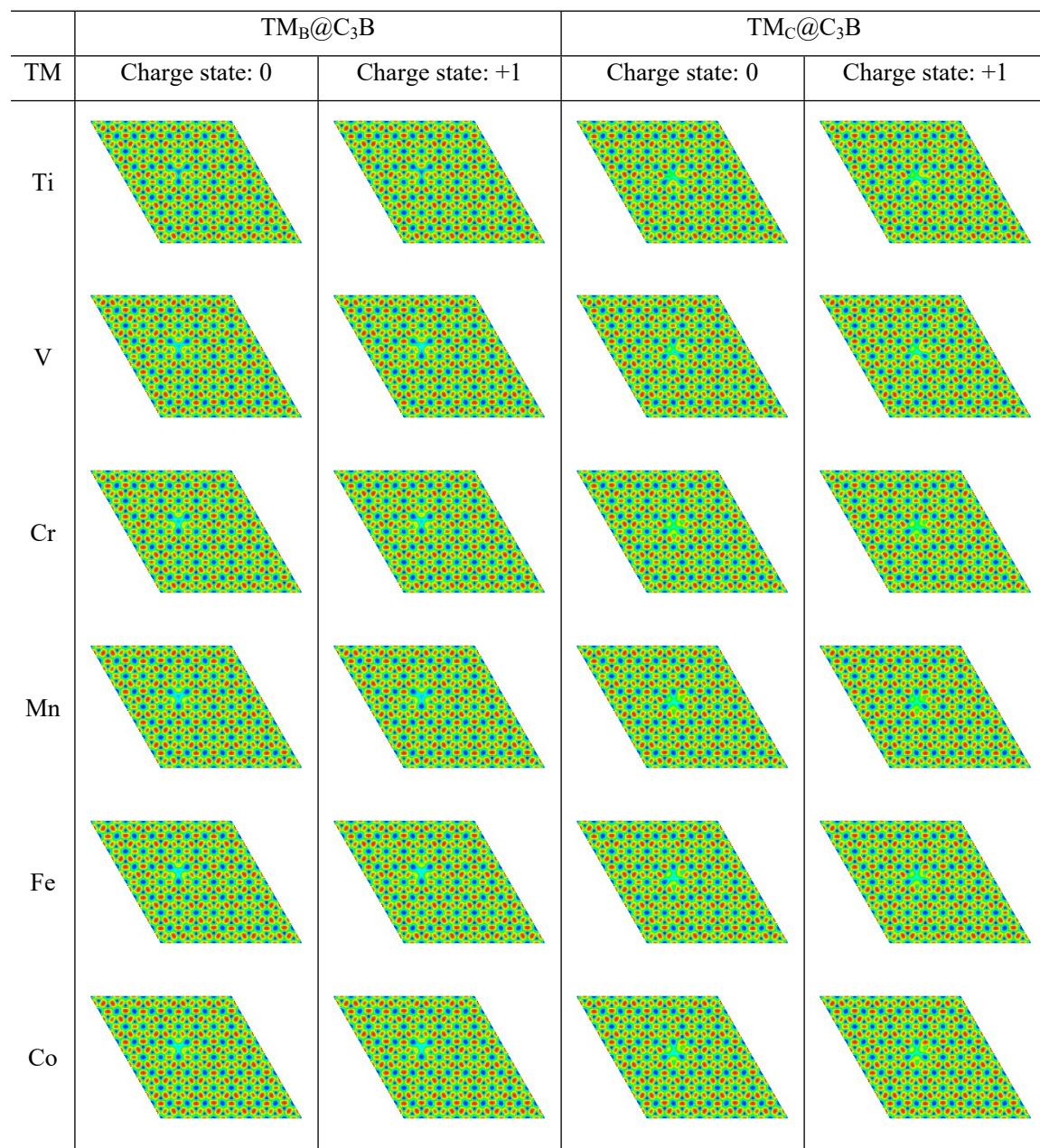
Note S5 Solvation effect

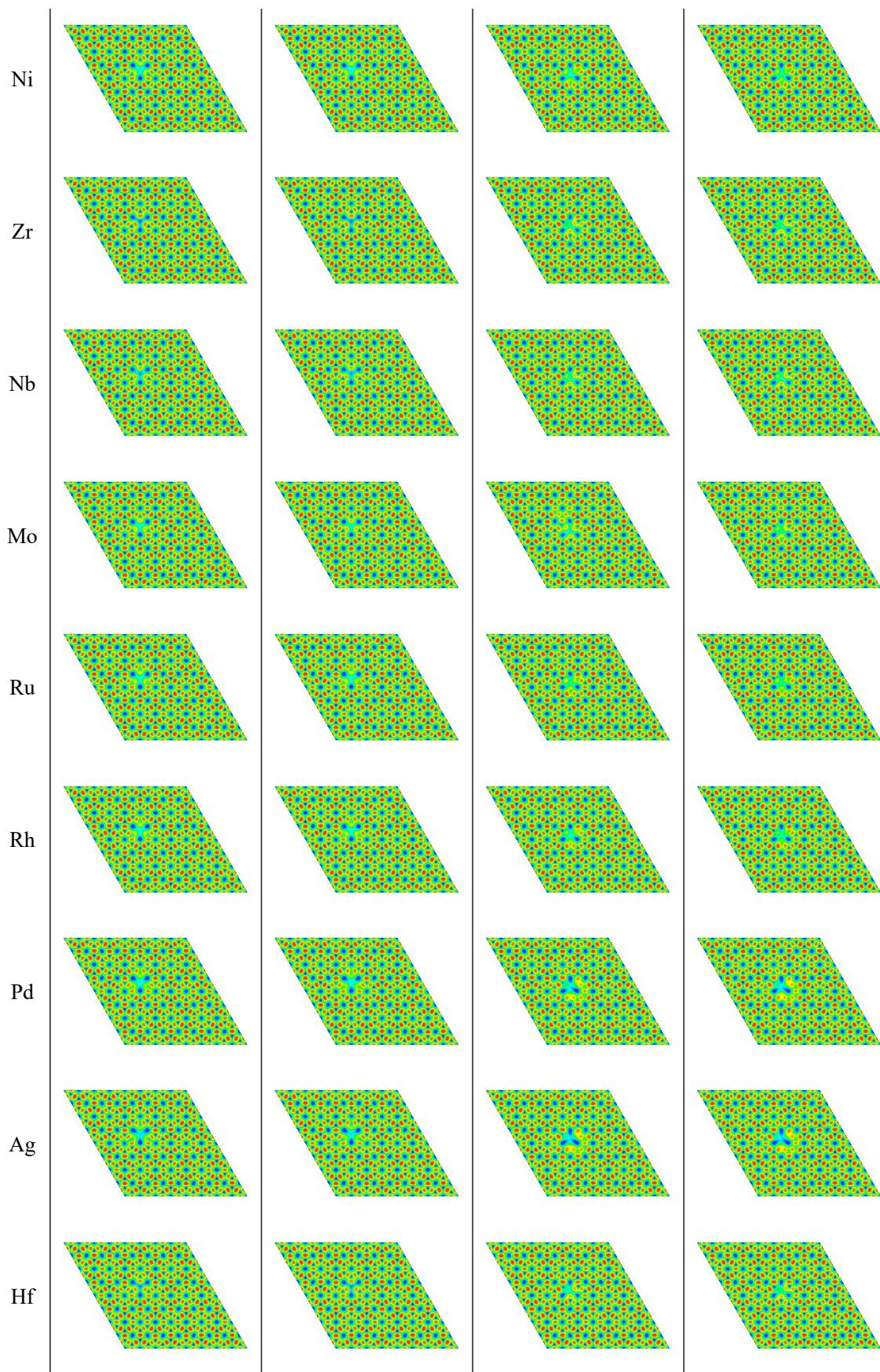
Given that the NRR process typically occurs in aqueous electrolyte solutions, this study takes the solvation effect into account. Using $\text{V}_\text{C}^0@\text{C}_3\text{B}$ and $\text{V}_\text{C}^{1+}@C_3\text{B}$ as examples, the enzymatic reaction pathways in the solution are calculated. An implicit solvation model, as implemented in VASPsol, is employed in this paper. As indicated in **Table S5**, the PDS during the hydrogenation process remains the first protonation step when considering the solvation effect. However, the limiting potentials are generally reduced, with the maximum reduction being 0.19 V for $\text{V}_\text{C}^0@\text{C}_3\text{B}$. This indicates that while the catalytic activity of TM@C₃B (V_C^0 and V_C^{1+}) in solution slightly improves.

Table S5. The $\Delta G_{N2 \rightarrow NNH}$, $\Delta G_{NNH \rightarrow NHNH}$, $\Delta G_{NHNH \rightarrow NHNH_2}$, $\Delta G_{NHNH_2 \rightarrow NH}$, $\Delta G_{NH \rightarrow NH_2}$, $\Delta G_{NH_2 \rightarrow NH_3}$ are calculated using the PBE (without solvation) and PBE+ SOL (with solvation) methods in $V_C^0 @ C_3B$ and $V_C^{1+} @ C_3B$ systems (unit: eV).

TM	$\Delta G_{N2 \rightarrow}$	$\Delta G_{NNH \rightarrow NHN}$	$\Delta G_{NHNH \rightarrow NHN}$	$\Delta G_{NHNH_2 \rightarrow N}$	$\Delta G_{NH \rightarrow N}$	$\Delta G_{NH_2 \rightarrow N}$
	NNH	H	H2	H	H2	H3
V_C^0 (PBE)	0.45	-0.15	-0.42	-0.01	-1.58	-0.14
V_C^0 (PBE+ SOL)	0.34	-0.34	-0.83	-0.56	-0.55	-0.78
V_C^{1+} (PBE)	0.39	-0.23	-0.27	-0.40	-1.26	-0.32
V_C^{1+} (PBE+ SOL)	0.12	-0.27	-0.64	-0.42	-0.73	-0.95

Table S6. The ELF map of 3d, 4d, and 5d TM@C₃B (B and C sites) in the 0 and +1 charge states.





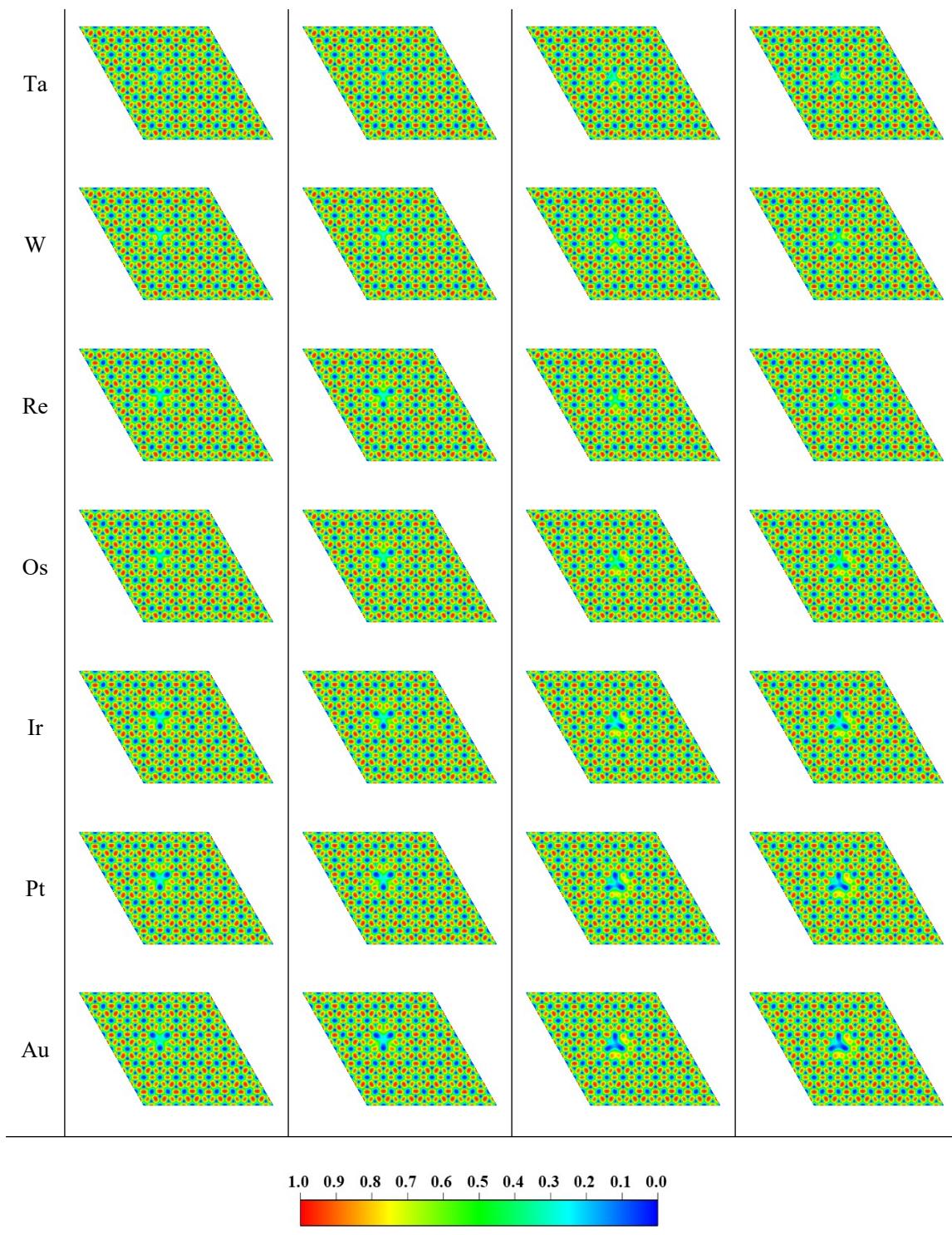


Table S7. The bond lengths between TM and the adjacent atoms (C and B) at the 3d, 4d, and 5d-TM@C₃B (B and C sites) in the 0 and +1 charge states.

TM _B @C ₃ B		Charge state: 0			Charge state: +1		
TM		d _{TM-C1}	d _{TM-C2}	d _{TM-C3}	d _{TM-C1}	d _{TM-C2}	d _{TM-C3}
Ti		1.97	1.97	1.97	1.97	1.97	1.97
V		1.90	1.90	1.90	1.89	1.89	1.89
Cr		1.89	1.89	1.89	1.89	1.89	1.89
Mn		1.88	1.88	1.88	1.88	1.88	1.88
Fe		1.81	1.81	1.84	1.83	1.83	1.83
Co		1.80	1.80	1.80	1.80	1.80	1.80
Ni		1.84	1.84	1.84	1.84	1.84	1.84
Cu		1.91	1.91	1.92	1.91	1.91	1.92
Zr		2.07	2.07	2.09	2.07	2.08	2.09
Nb		2.01	2.01	2.01	2.00	2.00	2.00
Mo		1.99	1.99	1.99	2.00	2.00	2.00
Ru		1.90	1.90	1.90	1.90	1.90	1.90
Rh		1.91	1.91	1.91	1.91	1.91	1.91
Pd		1.96	1.96	1.96	1.97	1.97	1.97
Ag		2.09	2.09	2.09	2.10	2.10	2.10
Hf		2.06	2.06	2.06	2.06	2.06	2.06
Ta		1.99	1.99	1.99	1.99	1.99	1.99
W		1.96	1.96	1.96	1.94	1.94	1.94
Re		1.92	1.92	1.92	1.94	1.94	1.94
Os		1.90	1.90	1.90	1.90	1.90	1.90
Ir		1.92	1.92	1.92	1.92	1.92	1.92
Pt		1.95	1.95	1.95	1.95	1.95	1.95
Au		2.04	2.04	2.04	2.04	2.04	2.04

TM _C @C ₃ B		Charge state: 0			Charge state: +1		
TM		d _{TM-B}	d _{TM-C1}	d _{TM-C2}	d _{TM-B}	d _{TM-C1}	d _{TM-C2}
Ti		2.34	1.92	1.92	2.60	1.90	1.90
V		2.23	1.83	1.83	2.22	1.83	1.83
Cr		2.02	1.85	1.85	2.14	1.87	1.87
Mn		2.03	1.85	1.85	2.04	1.83	1.83
Fe		1.89	1.82	1.82	1.89	1.80	1.81
Co		1.86	1.78	1.78	1.88	1.78	1.78
Ni		1.89	1.81	1.81	1.92	1.82	1.82
Cu		2.00	1.90	1.90	2.01	1.90	1.90

Zr	2.47	2.06	2.05	2.67	2.04	2.04
Nb	2.32	1.94	1.94	2.32	1.94	1.94
Mo	1.95	1.91	2.02	2.12	1.92	1.93
Ru	1.94	1.88	1.88	1.94	1.89	1.89
Rh	1.96	1.91	1.91	1.98	1.90	1.90
Pd	2.00	1.97	1.97	2.02	1.98	1.98
Ag	2.16	2.12	2.12	2.17	2.13	2.13
Hf	2.40	2.04	2.04	2.61	2.01	2.01
Ta	2.28	1.93	1.93	2.28	1.94	1.94
W	2.05	1.90	1.90	2.07	1.91	1.91
Re	1.96	1.89	1.89	2.04	1.89	1.89
Os	1.96	1.88	1.88	1.96	1.88	1.88
Ir	1.97	1.90	1.90	1.98	1.90	1.90
Pt	1.99	1.95	1.95	2.01	1.96	1.96
Au	2.09	2.06	2.06	2.10	2.07	2.07

Table S8. Physical properties for the transition metal atoms.

TM	χ_m	E_i	r_d	N_m	m	N_e	θ_d
Ti	1.54	6.83	1.46	0.08	47.88	4	2
V	1.63	6.74	1.35	0.53	50.94	5	3
Cr	1.66	6.76	1.28	0.67	52	6	5
Mn	1.55	7.43	1.26	-0.5	54.94	7	5
Fe	1.83	7.9	1.27	0.15	55.85	8	6
Co	1.88	7.86	1.25	0.67	58.93	9	7
Ni	1.91	7.63	1.25	1.16	58.69	10	8
Cu	1.9	7.72	1.28	1.24	63.55	11	10
Zr	1.33	6.95	1.6	0.43	91.22	4	2
Nb	1.6	6.77	1.47	0.89	92.91	5	4
Mo	2.16	7.18	1.4	0.75	95.96	6	5
Ru	2.2	7.36	1.34	1.06	101.1	8	7
Rh	2.28	7.46	1.35	1.14	102.9	9	8
Pd	2.2	8.33	1.38	0.56	106.4	10	10
Ag	1.93	7.57	1.45	1.31	107.9	11	10
Hf	1.3	7	1.58	0.01	178.5	4	2
Ta	1.5	7.88	1.47	0.32	180.9	5	3
W	2.36	7.98	1.41	0.82	183.9	6	4
Re	1.9	7.87	1.38	0.15	186.2	7	5
Os	2.2	8.7	1.35	1.1	190.2	8	6

Ir	2.2	9	1.36	1.57	192.2	9	7
Pt	2.28	8.96	1.39	2.14	195.1	10	9
Au	2.54	9.22	1.44	2.32	197	11	10

Table S9. Characteristics of TM@C₃B (B and C sites) in the 0 and +1 charge states. magnetic moment (μ_{TM} , μ_{B}), s-band center (ε_s , eV), p-band center (ε_p , eV), d-band center (ε_d , eV),

TM _B @C ₃ B				Charge state: 0		
TM	μ_{TM}	ε_s	ε_p	ε_d	R_{TM}	Q_e
Ti	0.87	-0.82	1.57	1.05	1.97	1.47
V	2.00	-1.53	1.05	0.52	1.90	1.33
Cr	3.00	-1.49	1.41	0.51	1.89	1.11
Mn	2.00	-2.15	0.85	-0.83	1.88	1.02
Fe	1.00	-2.38	0.72	-0.79	1.82	0.77
Co	0.00	-2.36	0.36	-1.24	1.80	0.58
Ni	0.86	-2.89	-0.19	-2.58	1.84	0.56
Cu	1.96	-2.44	-0.03	-4.06	1.91	0.59
Zr	0.90	-0.98	1.78	1.16	2.08	1.66
Nb	2.00	-0.58	2.07	0.50	2.01	1.49
Mo	3.00	-1.35	1.54	-0.09	1.99	1.06
Ru	1.00	-1.48	1.92	-0.93	1.90	0.62
Rh	0.00	-4.09	-0.40	-2.30	1.91	0.42
Pd	0.87	-3.30	-0.71	-3.84	1.96	0.40
Ag	1.96	-2.40	-0.20	-5.68	2.09	0.50
Hf	0.86	-1.80	2.11	1.52	2.06	1.65
Ta	1.99	-1.62	1.87	0.36	1.99	1.59
W	3.00	-2.11	1.74	-0.08	1.96	1.31
Re	2.00	-2.89	1.43	-0.83	1.92	0.98
Os	1.00	-2.89	1.71	-1.01	1.90	0.69
Ir	0.00	-3.04	1.21	-1.95	1.92	0.43
Pt	0.87	-3.18	0.31	-3.67	1.95	0.36
Au	1.96	-2.38	0.40	-5.30	2.04	0.37
TM _B @C ₃ B				Charge state: +1		
TM	μ_{TM}	ε_s	ε_p	ε_d	R_{TM}	Q_e
Ti	0.00	-0.70	1.69	0.98	1.97	1.50
V	1.00	-1.14	1.52	0.33	1.89	1.39
Cr	2.52	-1.06	1.73	0.72	1.89	1.15
Mn	3.00	-1.53	1.09	-0.77	1.88	1.06

Fe	2.00	-2.08	1.01	-0.92	1.83	0.84
Co	-0.36	-2.07	0.82	-0.61	1.80	0.61
Ni	0.00	-2.77	-0.01	-2.68	1.84	0.62
Cu	0.98	-2.60	-0.04	-4.23	1.91	0.60
Zr	0.00	-0.14	2.04	1.39	2.08	1.69
Nb	1.00	-0.27	2.04	0.46	2.00	1.56
Mo	2.03	-1.00	1.80	-0.22	2.00	1.14
Ru	1.75	-1.56	1.90	-1.02	1.90	0.68
Rh	0.00	-1.96	0.77	-1.49	1.91	0.44
Pd	0.00	-3.15	-0.23	-3.55	1.97	0.42
Ag	0.98	-2.21	0.14	-5.60	2.10	0.52
Hf	0.00	-1.24	2.50	1.74	2.06	1.75
Ta	1.00	-2.23	2.03	0.82	1.99	1.65
W	2.00	-0.69	2.21	0.03	1.94	1.40
Re	2.99	-2.38	1.48	-0.90	1.94	1.04
Os	1.85	-2.64	1.78	-1.13	1.90	0.75
Ir	0.00	-2.39	1.64	-1.50	1.92	0.46
Pt	0.00	-2.74	0.75	-3.40	1.95	0.37
Au	0.98	-1.63	0.73	-5.19	2.04	0.39

TM	TM _C @C ₃ B			Charge state: 0		
	μ_{TM}	ϵ_s	ϵ_p	ϵ_d	R_{TM}	Q_e
Ti	0.00	0.09	1.64	1.35	2.06	1.38
V	0.93	-0.58	1.27	0.17	1.96	1.21
Cr	-2.00	-1.41	1.51	0.27	1.91	0.87
Mn	3.00	-1.85	1.28	-0.36	1.91	0.79
Fe	2.00	-2.61	0.87	-1.01	1.84	0.50
Co	1.00	-2.40	0.90	-1.21	1.81	0.29
Ni	0.00	-2.68	0.90	-1.63	1.84	0.25
Cu	0.91	-2.35	0.35	-3.66	1.93	0.38
Zr	0.00	0.31	1.60	1.46	2.19	1.54
Nb	0.92	-0.32	1.44	0.33	2.07	1.36
Mo	0.00	-1.17	1.80	-0.26	1.96	0.84
Ru	0.00	-2.08	1.55	-1.52	1.90	0.16
Rh	1.00	-2.74	0.58	-2.02	1.93	0.04
Pd	0.00	-2.61	0.58	-3.03	1.98	0.07
Ag	0.84	-2.01	0.50	-5.10	2.13	0.30
Hf	0.00	-0.69	2.04	1.68	2.16	1.55
Ta	0.90	-1.50	1.54	0.39	2.05	1.41

W	0.00	-2.12	1.93	-0.16	1.95	0.95
Re	1.00	-2.77	1.85	-0.80	1.92	0.49
Os	0.00	-3.03	1.54	-1.44	1.90	0.19
Ir	0.99	-3.28	1.26	-2.18	1.92	-0.03
Pt	0.00	-3.22	1.04	-3.11	1.97	-0.09
Au	0.81	-3.04	0.52	-4.85	2.07	0.06
TM_C@C₃B			Charge state: +1			
TM	μ_{TM}	ε_s	ε_p	ε_d	R_{TM}	Q_e
Ti	1.00	0.36	1.36	0.98	2.14	1.44
V	0.00	-0.71	1.36	0.03	1.96	1.21
Cr	3.00	-0.66	1.89	0.54	1.96	0.97
Mn	2.00	-1.62	1.38	-0.54	1.90	0.80
Fe	1.04	-2.00	1.28	-0.90	1.83	0.52
Co	0.00	-2.27	0.97	-1.19	1.81	0.34
Ni	0.95	-2.04	1.00	-1.80	1.85	0.32
Cu	0.00	-2.08	0.61	-3.61	1.94	0.41
Zr	1.00	0.26	1.37	1.35	2.25	1.63
Nb	0.00	0.00	1.62	0.42	2.07	1.38
Mo	1.00	-1.17	1.89	-0.55	1.99	0.91
Ru	-0.74	-1.45	2.16	-1.12	1.90	0.25
Rh	0.00	-2.55	0.86	-2.15	1.93	0.13
Pd	0.97	-2.35	0.75	-2.91	1.99	0.12
Ag	0.00	-1.71	0.81	-4.99	2.14	0.32
Hf	1.00	-0.41	1.82	1.66	2.21	1.63
Ta	0.00	-1.44	1.66	0.46	2.05	1.42
W	1.00	-1.89	2.03	-0.31	1.97	1.00
Re	1.99	-2.59	1.84	-0.89	1.94	0.67
Os	0.78	-2.64	1.85	-1.27	1.91	0.23
Ir	0.00	-2.93	1.47	-2.19	1.93	0.04
Pt	0.97	-2.83	1.24	-2.96	1.98	-0.05
Au	0.00	-2.70	0.80	-4.84	2.08	0.10

Table S10. Summary of prediction results of R^2 scores and RMSE for charge and neutral state of ΔE_{*N2} (end on), ΔE_{*N2} (side on), $\Delta G_{N2 \rightarrow NNH}$ (distal), $\Delta G_{N2 \rightarrow NNH}$ (enzymatic) and $\Delta G_{NH2 \rightarrow NH3}$ based on machine learning models.

Regressors	ΔE_{*N2} (end-on)				ΔE_{*N2} (side-on)			
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
	(train)	(train)	(test)	(test)	(train)	(train)	(test)	(test)
GBR	0.999	0.002	0.943	0.057	0.999	0.001	0.942	0.033
KNR	0.819	0.096	0.770	0.117	0.813	0.061	0.771	0.062
RFR	0.924	0.056	0.884	0.066	0.902	0.038	0.759	0.061
SVR	0.819	0.096	0.770	0.117	0.690	0.079	0.688	0.068
Regressors	$\Delta G_{N2 \rightarrow NNH}$ (distal)				$\Delta G_{N2 \rightarrow NNH}$ (enzymatic)			
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
	(train)	(train)	(test)	(test)	(train)	(train)	(test)	(test)
GBR	0.999	0.001	0.960	0.096	0.999	0.001	0.929	0.071
KNR	0.901	0.122	0.861	0.178	0.712	0.203	0.702	0.135
RFR	0.952	0.006	0.088	0.913	0.843	0.113	0.819	0.109
SVR	0.949	0.090	0.891	0.141	0.750	0.183	0.649	0.224
Regressors	$\Delta G_{NH2 \rightarrow NH3}$							
	R^2	RMSE	R^2	RMSE				
	(train)	(train)	(test)	(test)				
GBR	0.999	0.001	0.902	0.220				
KNR	0.801	0.254	0.791	0.249				
RFR	0.913	0.146	0.866	0.161				
SVR	0.721	0.305	0.715	0.249				

References

1. C. G. Van de Walle and J. Neugebauer, First-principles calculations for defects and impurities: Applications to III-nitrides, *J. Appl. Phys.*, 2004, 95, 3851-3879.
2. G. Wang, H. Chen, Y. Li, A. Kuang, H. Yuan and G. Wu, A hybrid density functional study on the visible light photocatalytic activity of (Mo,Cr)-N codoped KNbO₃, *Phys. Chem. Chem. Phys.*, 2015, 17, 28743-28753.
3. G. Wang, L. Zhao, S.-D. Guo, J. Chang, B. Wang, W. Zhao, B. Yuan, X. Long, W. Zhang and P. Su, Bandgap engineering of KTaO₃ for water-splitting by different doping strategies, *Int. J. Hydrogen Energy*, 2021, 46, 38663-38677.
4. W. J. Yin, S. H. Wei, M. M. Al-Jassim, J. Turner and Y. F. Yan, Doping properties of monoclinic BiVO₄ studied by first-principles density-functional theory, *Phys. Rev. B*, 2011, 83, 155102.
5. M. Lan, R. Wang, Z.-H. Yang, X. Wang, S. Sun and S.-H. Wei, Enhancing magnetic dipole emission in Eu-doped SrMO₃ (M=Ti,Zr,Hf): First-principles calculations, *Phys. Rev. B*, 2021, 103, 245201.

6. V. Wang, W. Xiao, L. J. Kang, R. J. Liu, H. Mizuseki and Y. Kawazoe, Sources of n-type conductivity in GaInO₃, *J. Phys. D: Appl. Phys.*, 2015, 48, 015101.
7. Y. Wei, X. Li, J. Yang, C. Liu, J. Zhao, Y. Liu and S. Dong, Interaction between hydrogen and gallium vacancies in β-Ga₂O₃, *Sci. Rep.*, 2018, 8, 10142.
8. J. Zhang, X. Chen, M. Deng, H. Shen, H. Li and J. Ding, Effects of native defects and cerium impurity on the monoclinic BiVO₄ photocatalyst obtained via PBE+U calculations, *Phys. Chem. Chem. Phys.*, 2020, 22, 25297-25305.
9. J. Zhang, P. Deng, M. Deng, H. Shen, Z. Feng and H. Li, Hybrid Density Functional Theory Study of Native Defects and Nonmetal (C, N, S, and P) Doping in a Bi₂WO₆ Photocatalyst, *ACS Omega*, 2020, 5, 29081-29091.
10. S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl and C. Wolverton, The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies, *npj Comput. Mater.*, 2015, 1, 15010.
11. J. E. Saal, S. Kirklin, M. Aykol, B. Meredig and C. Wolverton, Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD), *JOM*, 2013, 65, 1501-1509.
12. J. T. Norton, H. Blumenthal and S. J. Sindeband, Structure of diborides of titanium, zirconium, columbium, tantalum and vanadium, *JOM*, 2017, 1, 749-751.
13. A. Wittmann, H. Nowotny and H. Boller, Ein Beitrag zum Dreistoff Titan-Molybdän-Bor, *Monatshefte für Chemie* 1960, 91, 608-615.
14. M. Arbuzov, B. Khaenko, E. Kachkovskaya and S. Y. Golub, X-ray study of titanium carbide in homogeneity region, *Ukrain. Fiz. Zh.*, 1984, 19, 497-501.
15. J. Li, H. Zhang, X. Cheng and D. Li, First-principles study of the structural and electronic properties of VB₂ under high pressure, *Physica B*, 2010, 405, 2768-2771.
16. O. Schob and E. Parthé, AB compounds with Sc, Y and rare earth metals. I. Scandium and yttrium compounds with CrB and CsCl structure, *Acta Crystallogr.*, 1965, 19, 214-224.
17. H. Nowotny and R. Kieffer, Röntgenographische Untersuchung von Karbidsystemen, *Int J Mater Res*, 1947, 38, 257-265.
18. B. Lönnberg, Thermal expansion studies on the group IV–VII transition metal diborides, *J. Less-common Met.*, 1988, 141, 145-156.
19. G. Papesch, H. Nowotny and F. Benesovsky, Untersuchungen in den Systemen: Chrom-Bor-Kohlenstoff, Mangan-Bor-Kohlenstoff und Mangan-Germanium-Kohlenstoff, *Monatshefte für Chemie/Chemical Monthly*, 1973, 104, 933-942.
20. K. Hellbom and A. Westgren, Den rombiska kromkarbidens kristallbyggnad, *Svensk Kemisk Tidskrift*, 1933, 45, 141-150.
21. B. Hájek, P. Karen and V. Brožek, Heptamanganese tricarbide Mn₇C₃, *Collect. Czech. Chem. Commun.*, 1983, 48, 2740-2750.
22. R. Kiessling, The Borides of Manganese, *Acta Chem. Scand.*, 1950, 4, 146-159.
23. E. E. Havinga, H. Damsma and P. Hokkeling, Compounds and pseudo-binary alloys with the CuAl₂(C16)-type structure I. Preparation and X-ray results, *J. Less-common Met.*, 1972, 27, 169-186.
24. A. N. Kolmogorov, S. Shah, E. R. Margine, A. F. Bialon, T. Hammerschmidt and R. Drautz, New superconducting and semiconducting Fe-B compounds predicted with an ab initio evolutionary search, *Phys. Rev. Lett.*, 2010, 105, 217003.

25. F. H. Herbstein and J. Smuts, Comparison of X-ray and neutron-diffraction refinements of the structure of cementite Fe₃C, *Acta Crystallogr.*, 1964, 17, 1331-1332.
26. X. Chong, Y. Jiang and J. Feng, Exploring the intrinsic ductile metastable Fe-C compounds: Complex chemical bonds, anisotropic elasticity and variable thermal expansion, *J. Alloys Compd.*, 2018, 745, 196-211.
27. T. Bjurström, *Röntgenanalyse der systeme eisen-bor, kobalt-bor und nickel-bor*, Almqvist & Wiksell, 1933.
28. J. Clarke and K. Jack, The preparation and the crystal structures of Cobalt Nitride, Co₂N, of cobalt carbonitrides, Co₂ (C, N), and of Cobalt Carbide, Co₂C, *Chem. Ind.*, 1951, 46, 1004-1005.
29. S. Rundqvist, S. Pramatus, B. Lamm, A. Haug, H. Theorell, R. Blinc, S. Paušak, L. Ehrenberg and J. Dumanović, Crystal structure refinements of Ni₃B, o-Ni₄B₃, and m-Ni₄B₃, *Acta Chem. Scand.*, 1967, 21, 21-0191.
30. P. Blum, La structure du borure de nickel Ni b, *Journal de Physique et le Radium*, 1952, 13, 430-431.
31. S. Nagakura, Study of metallic carbides by electron diffraction part II. Crystal structure analysis of nickel carbide, *J. Phys. Soc. Jpn.*, 1958, 13, 1005-1014.
32. V. F. Funke, I. V. Pshenichnyi, P. V. Zubarev, L. A. Pliner, D. M. Lyakhov and V. M. Golomazov, Some physicomechanical properties of ZrC-NbC-MoC_{1-x} alloys, *Soviet Powder Metallurgy and Metal Ceramics*, 1977, 16, 968-973.
33. I. Higashi, Y. Takahashi and T. Atoda, Crystal growth of borides and carbides of transition metals from molten aluminum solutions, *J. Cryst. Growth*, 1976, 33, 207-211.
34. O. Matsumoto, Y. Yaguchi, Y. Shiota and Y. Kanzaki, Formation of cubic solid solutions in the Mo-Nb-C and Mo-Ta-C systems by the carbonization of oxides in the plasma arc, *High Temp. Sci.*, 1983, 16, 243-250.
35. B. Kodess, L. Butman, S. Sambueva and S. KIRSCH, Refinement of Mo₂B₅ structural type, *Soviet physics. Crystallography*, 1992, 37, 30-32.
36. R. Steinitz, I. Binder and D. Moskowitz, System Molybdenum-Boron and Some Properties of The Molybdenum-Borides, *JOM*, 2017, 4, 983-987.
37. J. E. Senkin, *Some Physical, Mechanical, and Thermodynamic Properties of Transition Metal Refractory Carbides: A Bibliography*, University of California Lawrence Radiation Laboratory, 1963.
38. J. Schuster, E. Rudy and H. Nowotny, Die "MoC"-Phase mit WC-Struktur, *Monatshefte für Chemie* 1976, 107, 1167-1176.
39. B. Aronsson, E. Stenberg and J. ÅSelius, Borides of Ruthenium, Osmium and Iridium, *Nature*, 1962, 195, 377-378.
40. L. T, Structure of Ru₂B₃ and WB_{2.0} as Determined by Single-Crystal Diffractometry and Some Notes on WB System, *Arkiv for Kemi*, 1969, 30, 115.
41. C. P. Kempfer and M. R. Nadler, Preparation and Crystal Structures of RuC and OsC, *J. Chem. Phys.*, 1960, 33, 1580-1581.
42. J. H. Buddery and A. J. Welch, Borides and silicides of the platinum metals, *Nature*, 1951, 167, 362.
43. R. W. Mooney and A. J. E. Welch, The crystal structure of Rh₂B, *Acta Crystallogr.*, 1954, 7, 49-53.

44. M. Beck, M. Ellner and E. J. Mittemeijer, Powder diffraction data for borides PdB and PdB and the formation of an amorphous boride PdB, *Powder Diffr.*, 2001, 16, 98-101.
45. R. Gumeniuk, H. Borrmann and A. Leithe-Jasper, Refinement of the crystal structures of trinickel boron, NiB, and tripalladium boron, Pd₃B, *Z Krist-New Cryst St*, 2006, 221, 425-426.
46. H. R. Soni, S. K. Gupta and P. K. Jha, Ab initio total energy calculation of the dynamical stability of noble metal carbides, *Physica B*, 2011, 406, 3556-3561.
47. W. Obrowski, Die Struktur der Diboride von Gold und Silber, *Die Naturwissenschaften*, 1961, 48, 428-428.
48. J.-D. Zhang and X.-L. Cheng, The predicted structural and electronic properties of HfB₂ under high pressure: First-principles calculations, *Physica B*, 2010, 405, 3532-3536.
49. P. G. Cotter and J. A. Kohn, Industrial Diamond Substitutes: I, Physical and X-Ray Study of Hafnium Carbide, *J. Am. Ceram. Soc.*, 2006, 37, 415-420.
50. K. Becker and F. Ebert, Die Kristallstruktur einiger binärer Carbide und Nitride, *Zeitschrift für Physik*, 1925, 31, 268-272.
51. E. Rudy, F. Benesovsky and L. Toth, Untersuchung der Dreistoffsysteme der Va- und VIa-Metalle mit Bor und Kohlenstoff, *Int J Mater Res*, 1963, 54, 345-353.
52. J. Leciejewicz, A note on the structure of tungsten carbide, *Acta Crystallogr.*, 1961, 14, 200-200.
53. S. Zeng, G. Li, Y. Zhao, R. Wang and J. Ni, Machine Learning-Aided Design of Materials with Target Elastic Properties, *J. Phys. Chem. C*, 2019, 123, 5042-5047.
54. H.-H. Chen, Z. Li, Y. Cheng, Y. Bi and L.-C. Cai, Thermodynamic properties of OsB under high temperature and high pressure, *Physica B*, 2011, 406, 3338-3341.
55. F. Peng, D. Chen, H. Fu and T. Gao, The phase transition and elastic property of osmium carbide under pressure, *physica Status Solidi B: Basic Solid State Physics*, 2010, 248, 1222-1226.
56. J. Schmidt, N. Hoffmann, H.-C. Wang, P. Borlido, P. J. Carriço, T. F. Cerqueira, S. Botti and M. A. Marques, Large-scale machine-learning-assisted exploration of the whole materials space, *arXiv preprint arXiv*, 2022, DOI: 10.48550/arXiv.2210.00579, 00579.
57. J. Yang and F. M. Gao, Hardness calculations of 5d transition metal monocarbides with tungsten carbide structure, *physica Status Solidi B: Basic Solid State Physics*, 2010, 247, 2161-2167.
58. E. Hassler, T. Lundström and L.-E. Tergenius, The crystal chemistry of platinum metal borides, *J. Less-common Met.*, 1979, 67, 567-572.
59. H.-C. Wang, S. Botti and M. A. L. Marques, Predicting stable crystalline compounds using chemical similarity, *npj Comput. Mater.*, 2021, 7, 12.
60. V. Mankad, N. Rathod, S. D. Gupta, S. K. Gupta and P. K. Jha, Stable structure of platinum carbides: A first principles investigation on the structure, elastic, electronic and phonon properties, *Mater. Chem. Phys.*, 2011, 129, 816-822.
61. S. B. Zhang, S. H. Wei and A. Zunger, Intrinsinc-type versusp-type doping asymmetry and the defect physics of ZnO, *Phys. Rev. B*, 2001, 63, 075205.
62. B. Kong, X. Y. An, T. X. Zeng and J. Zhang, Hybrid density functional studies of native defects and H impurities in wurtzite CdSe, *Phys. Chem. Chem. Phys.*, 2020, 22, 7474-7482.
63. J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard and H. Jónsson, Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode, *J. Phys. Chem. B*, 2004, 108, 17886-17892.
64. J. Hao and T. K. Ho, Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language, *J. Educ. Behav. Stat.*, 2019, 44, 348-361.

65. J. H. Friedman, Stochastic gradient boosting, *Comput. Stat. Data Anal.*, 2002, 38, 367-378.
66. J. H. Friedman, Greedy function approximation: a gradient boosting machine, *Annals of statistics*, 2001, 29, 1189-1232.
67. J. Quinonero-Candela and C. E. Rasmussen, A unifying view of sparse approximate Gaussian process regression, *J. Mach. Learn. Res.*, 2005, 6, 1939-1959.
68. J. Bi and K. P. Bennett, A geometric approach to support vector regression, *Neurocomputing*, 2003, 55, 79-108.
69. A. Seko, T. Maekawa, K. Tsuda and I. Tanaka, Machine learning with systematic density-functional theory calculations: Application to melting temperatures of single- and binary-component solids, *Phys. Rev. B*, 2014, 89, 054303.
70. S. Thomas, A. K. Madam and M. A. Zaeem, Stone-Wales Defect Induced Performance Improvement of BC₃ Monolayer for High Capacity Lithium-Ion Rechargeable Battery Anode Applications, *J. Phys. Chem. C*, 2020, 124, 5910-5919.
71. Y. Zhou, G. Gao, W. Chu and L. W. Wang, Transition-metal single atoms embedded into defective BC₃ as efficient electrocatalysts for oxygen evolution and reduction reactions, *Nanoscale*, 2021, 13, 1331-1339.
72. H. P. Komsa and A. Pasquarello, Finite-size supercell correction for charged defects at surfaces and interfaces, *Phys. Rev. Lett.*, 2013, 110, 095505.
73. H.-P. Komsa, N. Berseneva, A. V. Krasheninnikov and R. M. Nieminen, Charged Point Defects in the Flatland: Accurate Formation Energy Calculations in Two-Dimensional Materials, *Phys. Rev. X*, 2014, 4, 031044.
74. S. Yue, D. Li, A. Zhang, Y. Yan, H. Yan, Z. Feng and W. Wang, Rational design of single transition-metal atoms anchored on a PtSe₂ monolayer as bifunctional OER/ORR electrocatalysts: a defect chemistry and machine learning study, *J. Mater. Chem. A*, 2024, 12, 5451-5463.
75. W. Yao, Y. Qu, M. Zhou, W. Wang, A. Zhang, Z. Feng and H. Yan, Increasing the bifunctional OER/ORR activity of 3d transition metals doped g-C₃N₃ by controlling the charge states, *Mol. Catal.*, 2024, 554, 113807.
76. D. Li, A. Zhang, Z. Feng and W. Wang, Theoretical Insights on the Charge State and Bifunctional OER/ORR Electrocatalyst Activity in 4d-Transition-Metal-Doped g-C₃N₄ Monolayers, *ACS Appl. Mater. Interfaces*, 2024, 16, 5779-5791.
77. J. Zhang, D. Li, L. Ju, G. Yang, D. Yuan, Z. Feng and W. Wang, The charge effects on the hydrogen evolution reaction activity of the defected monolayer MoS₂, *Phys. Chem. Chem. Phys.*, 2023, 25, 10956-10965.
78. L. Wang, T. Maxisch and G. Ceder, Oxidation energies of transition metal oxides within theGGA+Uframework, *Phys. Rev. B*, 2006, 73, 195107.
79. H. Niu, X. Wan, X. Wang, C. Shao, J. Robertson, Z. Zhang and Y. Guo, Single-Atom Rhodium on Defective g-C₃N₄: A Promising Bifunctional Oxygen Electrocatalyst, *ACS Sustainable Chem. Eng.*, 2021, 9, 3590-3599.
80. X. Wang, H. Niu, X. Wan, J. Wang, C. Kuai, Z. Zhang and Y. Guo, Identifying TM-N₄ active sites for selective CO₂-to-CH₄ conversion: A computational study, *Appl. Surf. Sci.*, 2022, 582, 152470.
81. M. Han, G. Wang, H. Zhang and H. Zhao, Theoretical study of single transition metal atom modified MoP as a nitrogen reduction electrocatalyst, *Phys. Chem. Chem. Phys.*, 2019, 21,

- 5950-5955.
82. M. Li, S. Han, C. Xia, A. Wang, Y. Xie, J. She, Z. Deng, X. Zhao and Q. Fang, Single Ru Atom Supported on B-Doped Graphyne as an Efficient Electrocatalyst for the Nitrogen Reduction Reaction, *Catal. Lett.*, 2023, 154, 2597-2605.
83. X. Zhang, L. Yan and Z. Su, A single transition metal atom anchored on Nb(2)C as an electrocatalyst for the nitrogen reduction reaction, *Nanoscale*, 2023, 15, 17508-17515.
84. Y. Zhang, Z. F. Wu, P. F. Gao, D. Q. Fang, E. H. Zhang and S. L. Zhang, Strain-tunable electronic and optical properties of BC₃ monolayer, *RSC Adv.*, 2018, 8, 1686-1692.
85. Y. Wu, C. He and W. Zhang, "Capture-Backdonation-Recapture" Mechanism for Promoting N₂ Reduction by Heteronuclear Metal-Free Double-Atom Catalysts, *J. Am. Chem. Soc.*, 2022, 144, 9344-9353.