

Supplementary Information

Machine Learning Assisted High-Throughput Screening of Transition Metal Single Atoms Based Superb Hydrogen Evolution Electrocatalysts

Muhammad Umer,^a Sohaib Umer,^a Mohammad Zafari,^a Miran Ha,^a Rohit Anand,^a Amir Hajibabaei,^a Ather Abbas,^b
Geunsik Lee,^{*a} Kwang S. Kim^{*a}

^a Center for Superfunctional Materials, Department of Chemistry, Ulsan National Institute of Science and Technology (UNIST), 50 UNIST-gil, Ulsan 44919, Republic of Korea.

^b Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), 50 UNIST-gil, Ulsan 44919, Republic of Korea.

*E-mail: gslee@unist.ac.kr (G.L.), kimks@unist.ac.kr (K.S.K.)

Contents

<i>Supplementary Note</i>	(S2)
<i>Supplementary Figures</i>	(S4)
<i>Supplementary Tables</i>	(S25)
<i>Supplementary References</i>	(S38)

Descriptors for ML Analysis

Supplementary Note 1:

Here we report various kind of descriptors we constructed for ML analysis, as a supervised machine learning (ML) approach we predicted structural stabilities and HER activities of different 2D materials. The key screening steps involved in our multistage screening strategy are summarized in **Figure S2**. For ML prediction of stabilities (i.e E_{stab} and U_{diss}) we searched various types of unique descriptors (**Table S11**) which do not require any DFT calculation. **Figure S1** demonstrates the labeled sites we used to create Coulomb matrix elements (see equation 9 in methods). Based on elemental properties, we constructed Coulomb matrix elements between transition metal (TM) atom and its nearest neighbor coordinating non-metal atoms (cm_{n1} through cm_{n4}), TM and an adsorbed hydrogen atom (cm_{TM-H}) and between the adsorbed hydrogen atom and surface non-metal atoms (cm_{H-nX}) involved at bond edge shared active site. For ML classification analysis, our descriptor is composed of an eight-dimensional vector (equation 1), 2 elemental features namely first ionization potential (I) of TM, the total number of valance electrons (v_{en}) and 6 Coulomb matrix elements based on elemental properties.

$$D = (v_{en}, I, cm_{n1}, cm_{n2}, cm_{n3}, cm_{n4}, cm_{TM-H}, cm_{H-nX}) \quad (1)$$

For ML classification, we tested the following 12 different ML models: AdaBoost, CatBoost, Bagging, LGBM-Light Gradient Boosting Machine, GB-Gradient Boosting, ERT-Extremely Randomized Trees, RF-Random Forest, DT-Decision Trees, SVC-Support Vector Classifier, KNN-k-Nearest Neighbors, RNC-Radius Neighbors Classifier, LR-Logistic Regression. Mean and standard deviation values of the Area Under the Receiver Operating Characteristic Curve (ROC AUC) are listed in **Table S12**. The prediction performance of different ML classification models is compared in **Figure S14**.

Supplementary Note 2:

In order to create more features/descriptors, we employed a recently developed compressed-sensing based approach, SISSO (sure independence screening and sparsifying operator).¹ SISSO iteratively combines the sure-independence screening (SIS) scheme and the sparsifying operators for dimensionality reduction and sparse solutions respectively. We employed different algebraic operators [$+$, $-$, $\div$, $\times$, $\exp$, $|\cdot|$, $\sin$, $\sqrt{\cdot}$] on the predefined elemental and DFT-derived features (**Figure S1** and **Table S11**) and created the best low-dimensional near-optimal descriptors relating to the HER activities. The top-ranked 5 SISSO generated descriptors are presented in **Table S11**. For ML prediction of HER activities, we used the following regression algorithms: LGBM-Light Gradient Boosting Machine, HGBR-Histogram-based Gradient Boosting Regressor, ERT-Extremely Randomized Trees, GB-Gradient Boosting, Bagging, XGBR-XGBoost regressor, RF-Random Forest and AdaBoost. The test root-mean-square error (RMSE) and the coefficient of determination values (R^2) are displayed in **Figure S17** for different ML regression models.

Supplementary Figures

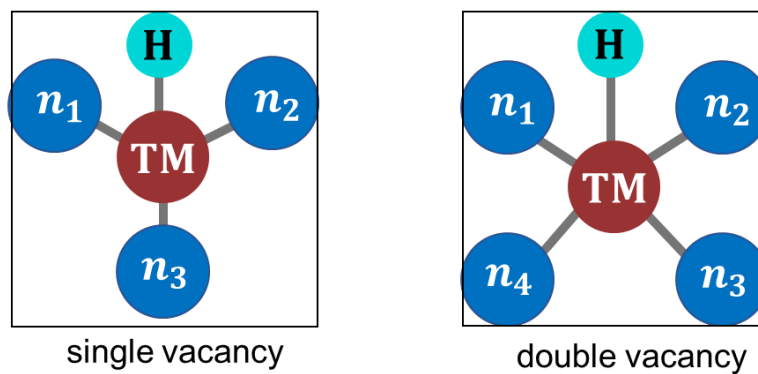


Figure S1 Representation of single and double vacancy 2D templates, the labeled sites (n_1 through n_3) in single vacancy systems and (n_1 through n_4) in double vacancy systems indicates the surface non-metal atoms ($X=B, C, N$ and P), whereas, TM and H denotes the transition metal and hydrogen atom respectively.

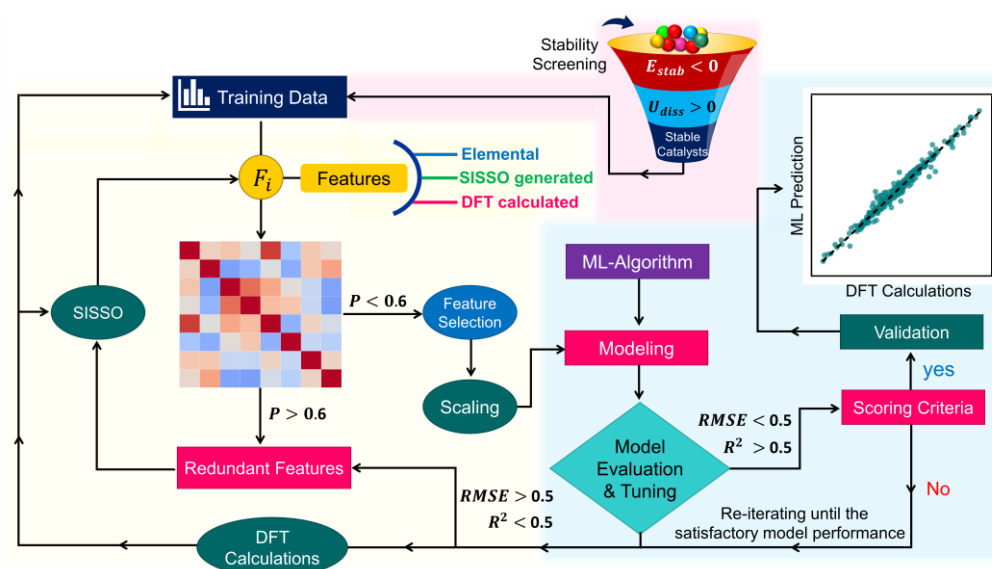


Figure S2 ML framework for designing single and dual site non-metal (B, N, P) doped TM-SACs. The pink shaded area represents the stability prescreening process based on U_{diss} and E_{stab} . The feature importance and evaluation is presented in the yellow highlighted section. The blue region represents the model tuning and validation based on defined scoring criteria.

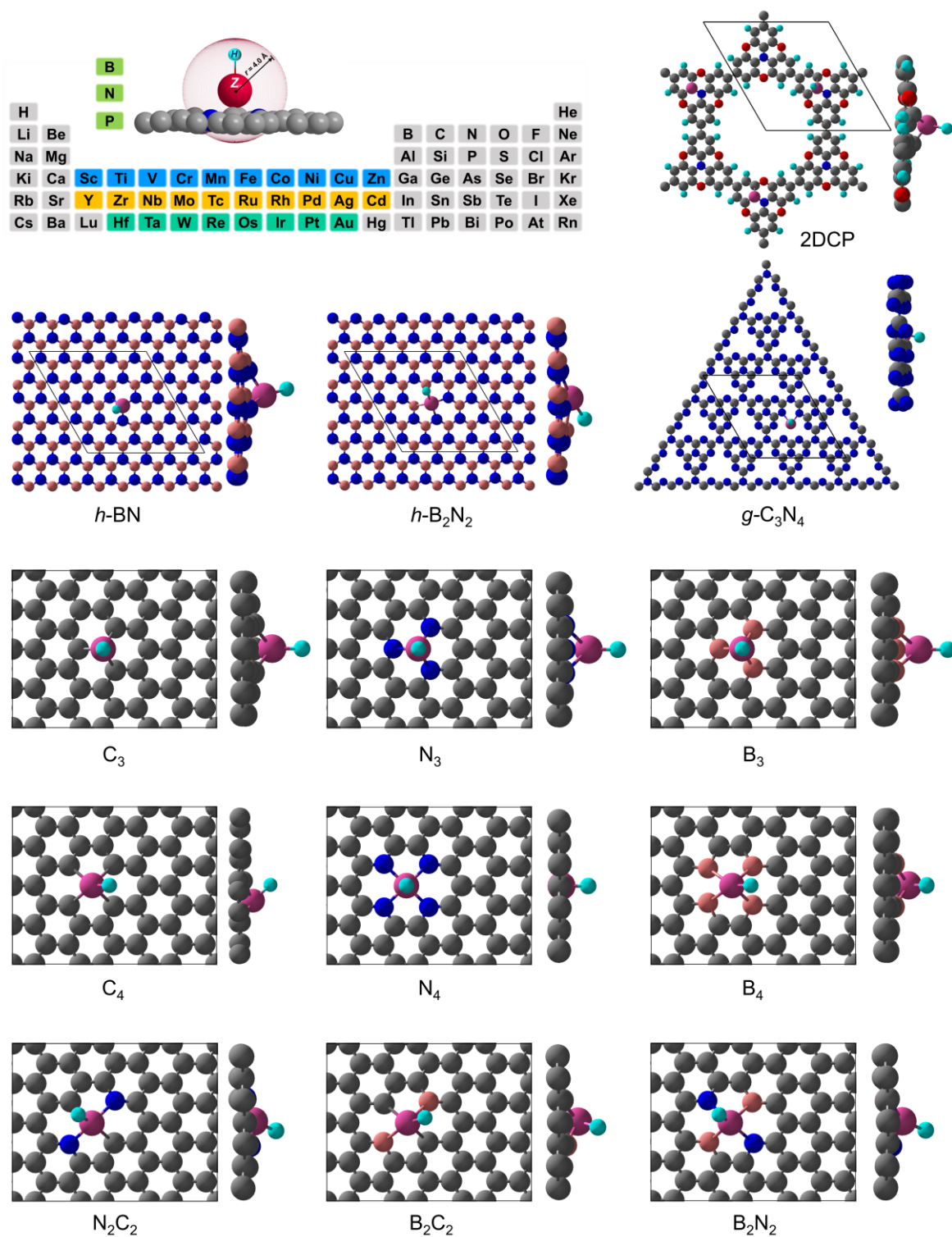


Figure S3. Top and side views of H-adsorbed 2DCP, boron nitride, graphitic nitride and pyridinic graphene based different configurations of two-dimensional templates with single and double vacant sites occupied with 28 kind of TMs (here is the case for Rh). In each configuration A TM atom is coordinated by mono- or dual-type non-metal (B,N) doped atoms.

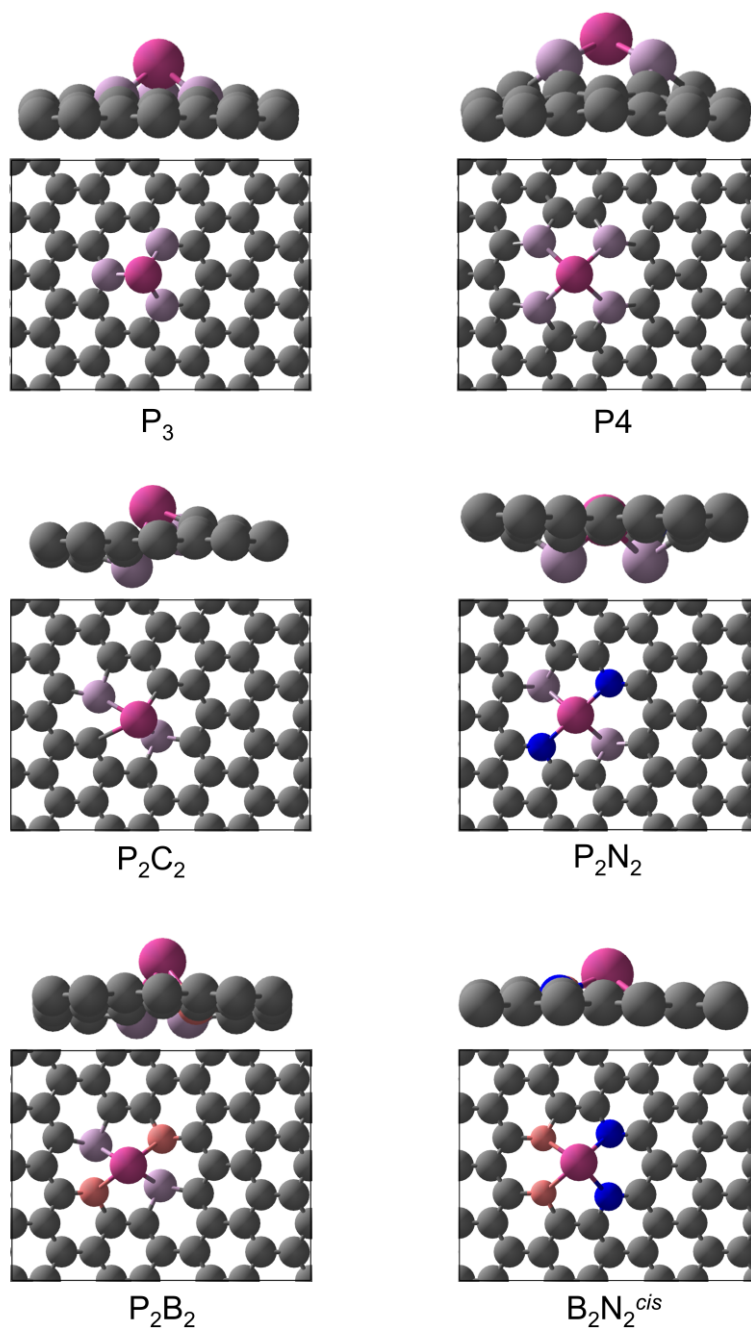


Figure S4 Top and side views of pyridinic graphene structures with single and double vacancy defects sites occupied with 28 kind of TMs. In each configuration a TM atom is coordinated by mono- or dual-type non-metal atoms (i.e. B, N, P).

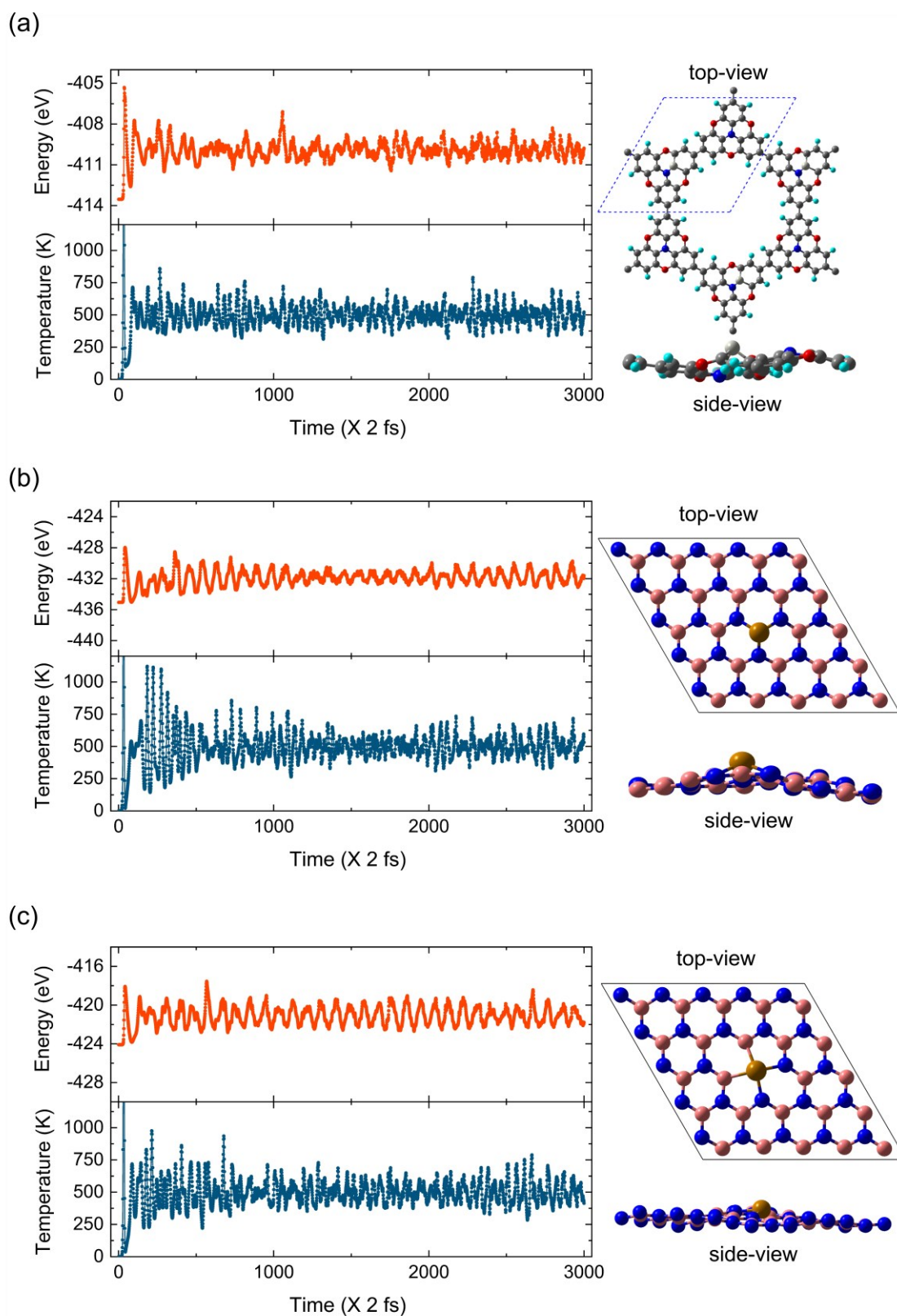


Figure S5 Temperature and energy profiles from ab initio molecular dynamics simulations under 500 K with the time step of 2 fs. Top and side views of (a) Pd@2DCP (b) Fe@*h*-BN and (c) Fe@*h*-B₂N₂ structures are presented. Color code: Pd, light-grey; Fe, golden-brown; B, light pink; N, blue; C, grey; O, red; H, cyan).

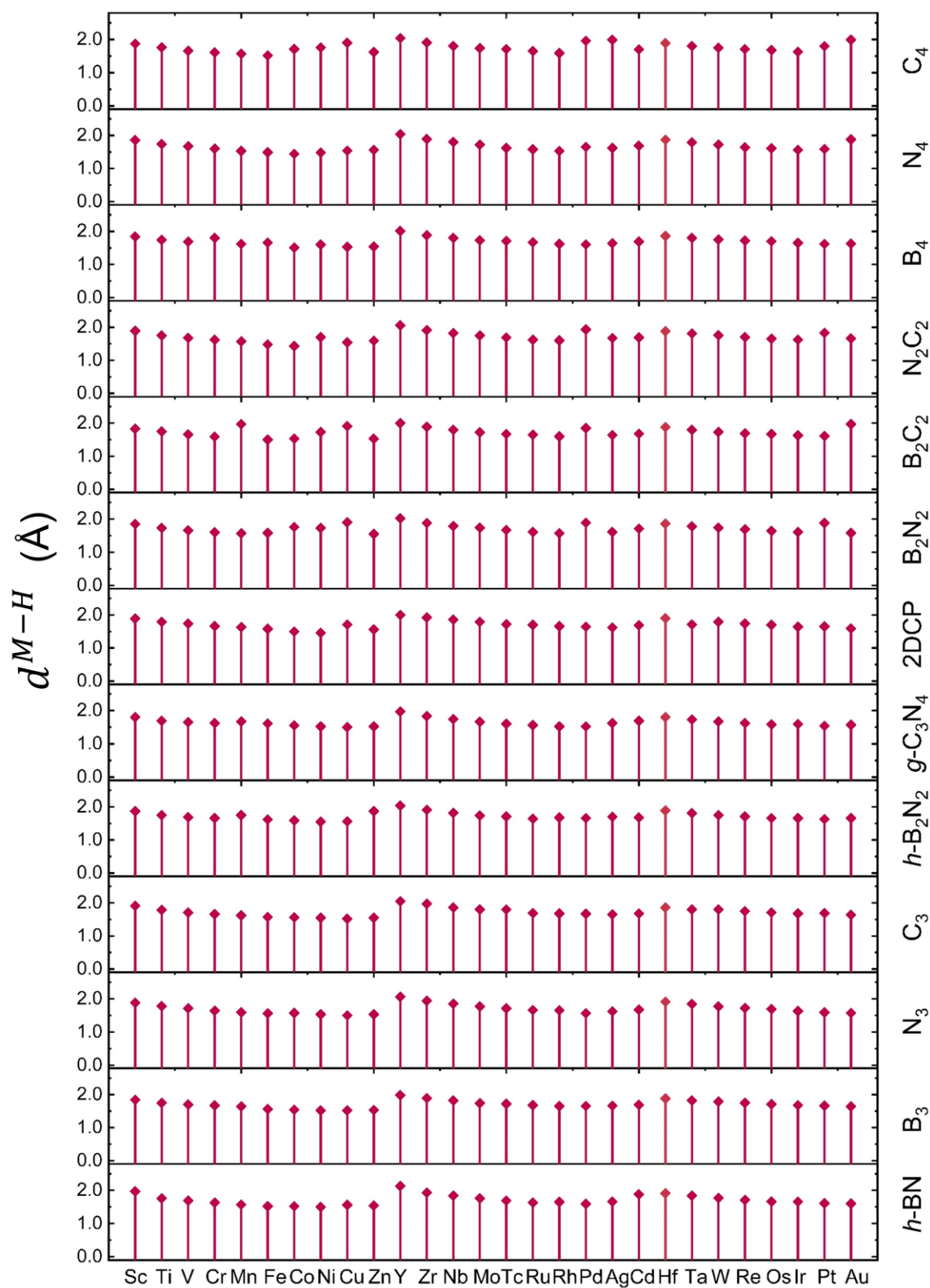


Figure S6 Metal-to-H bond distance (Å) for various kinds of 2D materials with TM embedded at single or double vacancy site, each with mono- or dual-type non-metal doped configuration.

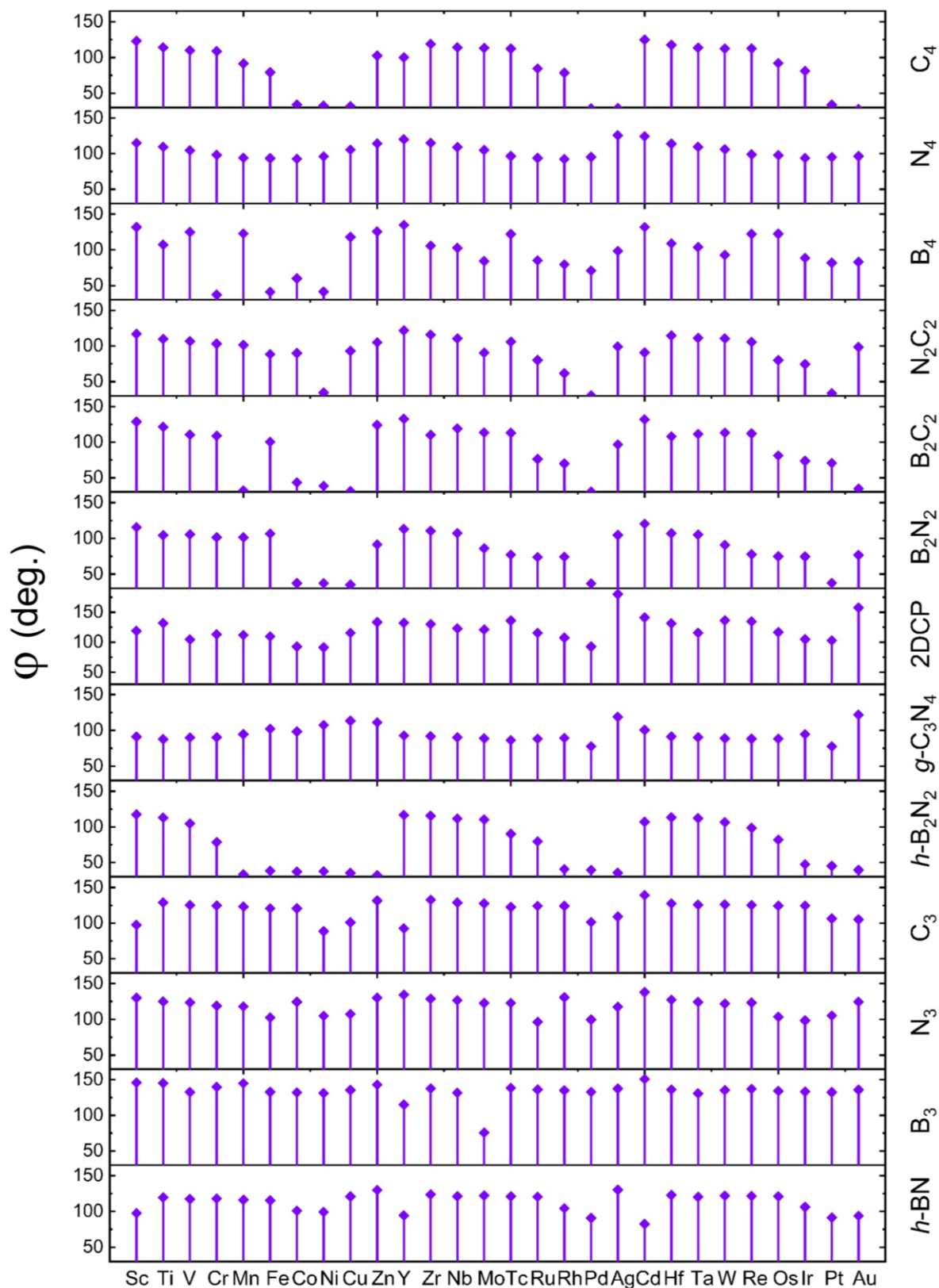


Figure S7 Angle at vertex at a TM atom (M) enclosed by two lines pointing to the adsorbed-H (H) and to the first nearest neighbor non-metal atom (Y: C, N, B) coordinating the TM atom on the 2D surface ($\angle\text{H-M-Y}$).

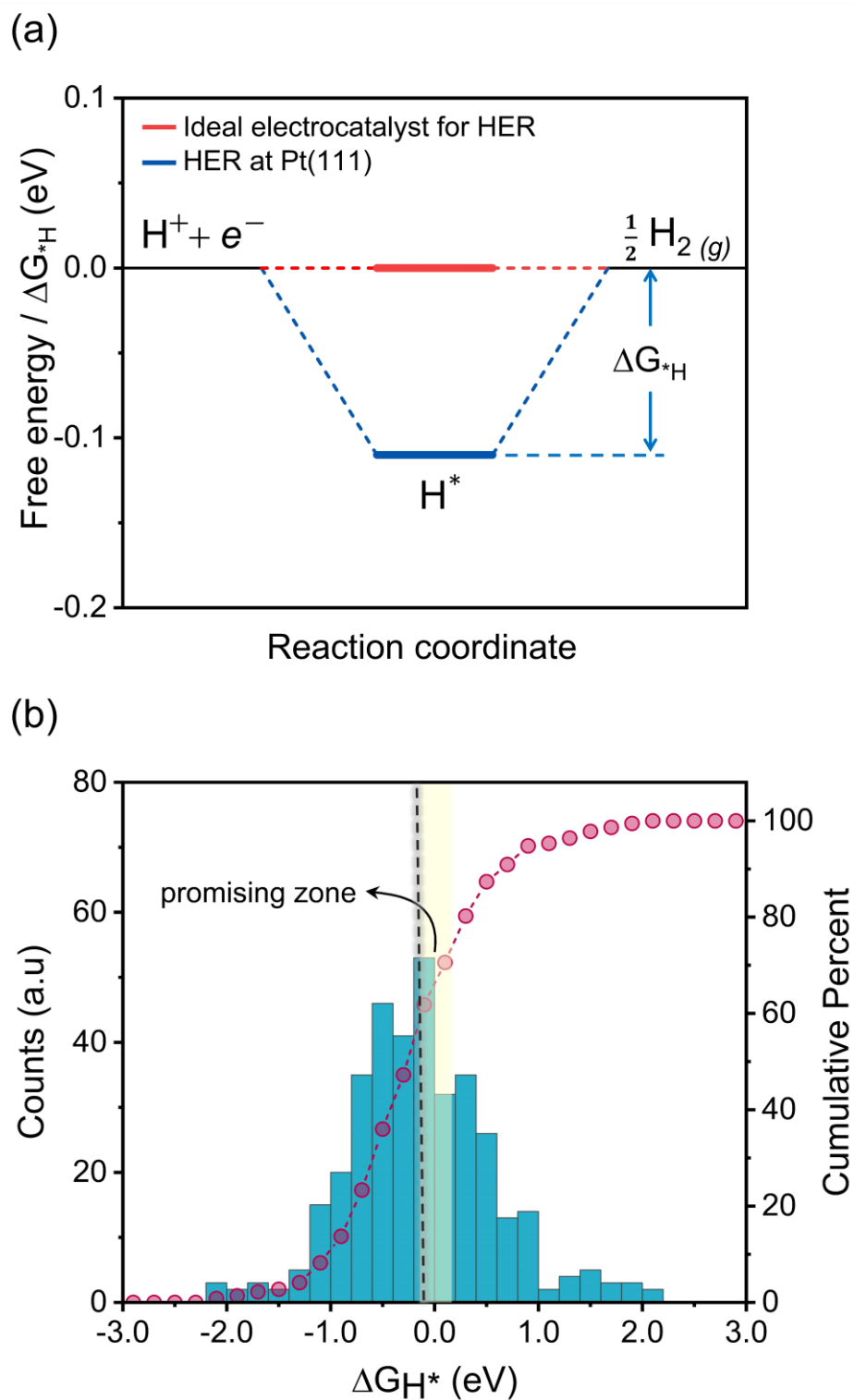


Figure S8. (a) HER free-energy diagram for ideal electrocatalyst and Pt (111) surface at equilibrium potential ($U=0$) and **(b)** the distribution of the hydrogen adsorption Gibbs free energy (ΔG_{H^*}) on the catalyst candidates in a set of ~ 364 candidates in the most stable adsorption configuration. ΔG_{H^*} is relative to the reference system Pt (111) (the dark grey dashed). The area with $|\Delta G_{H^*}| \leq 0.15$ eV is shaded in light yellow.

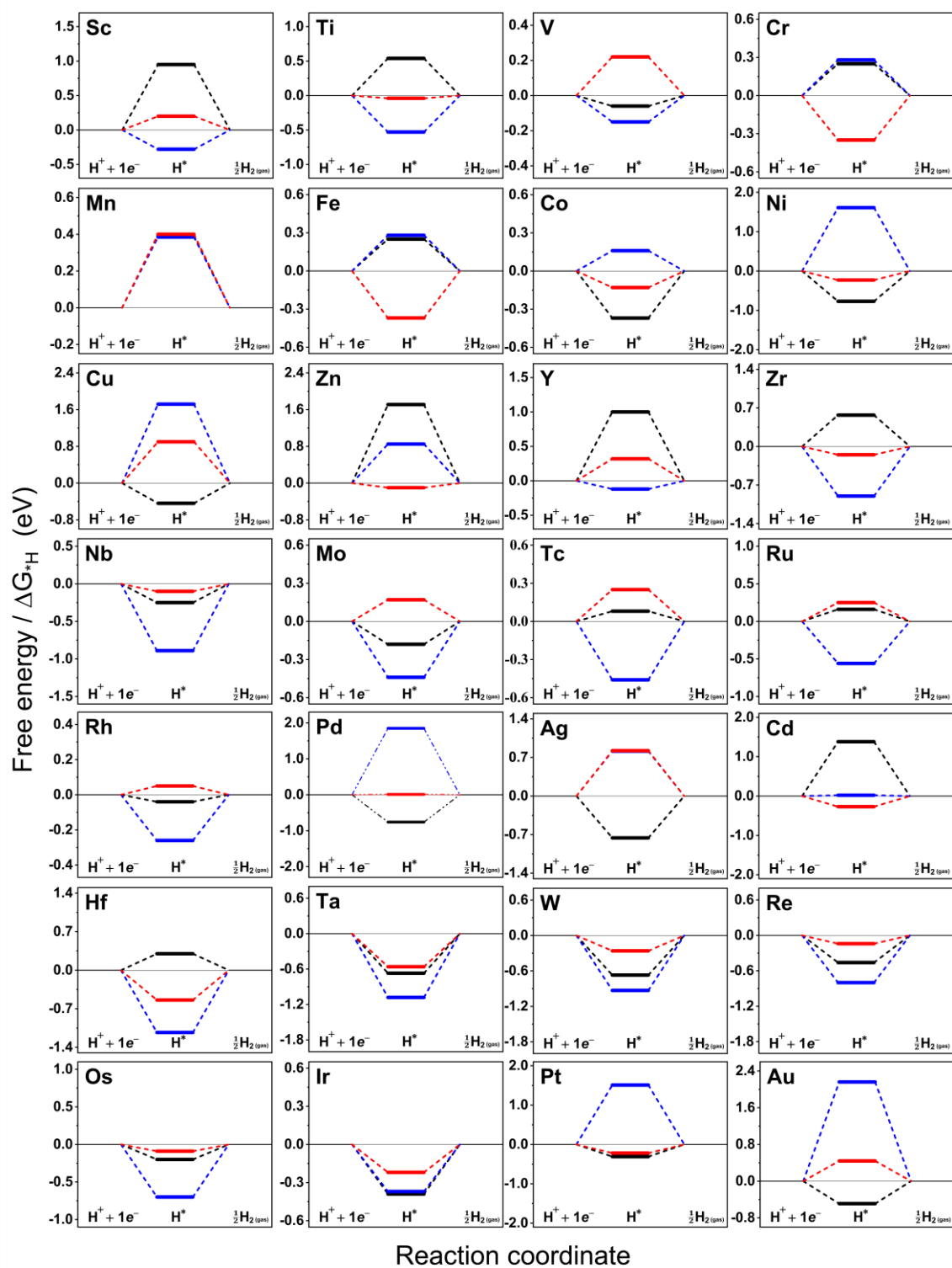


Figure S9. Free-energy diagram for HER on a single TM atom embedded at a double vacancy site in graphene support with carbon (C_4), nitrogen (N_4) and boron (B_4) coordination's, represented with black, blue and red lines, respectively, at the zero-electrode potential.

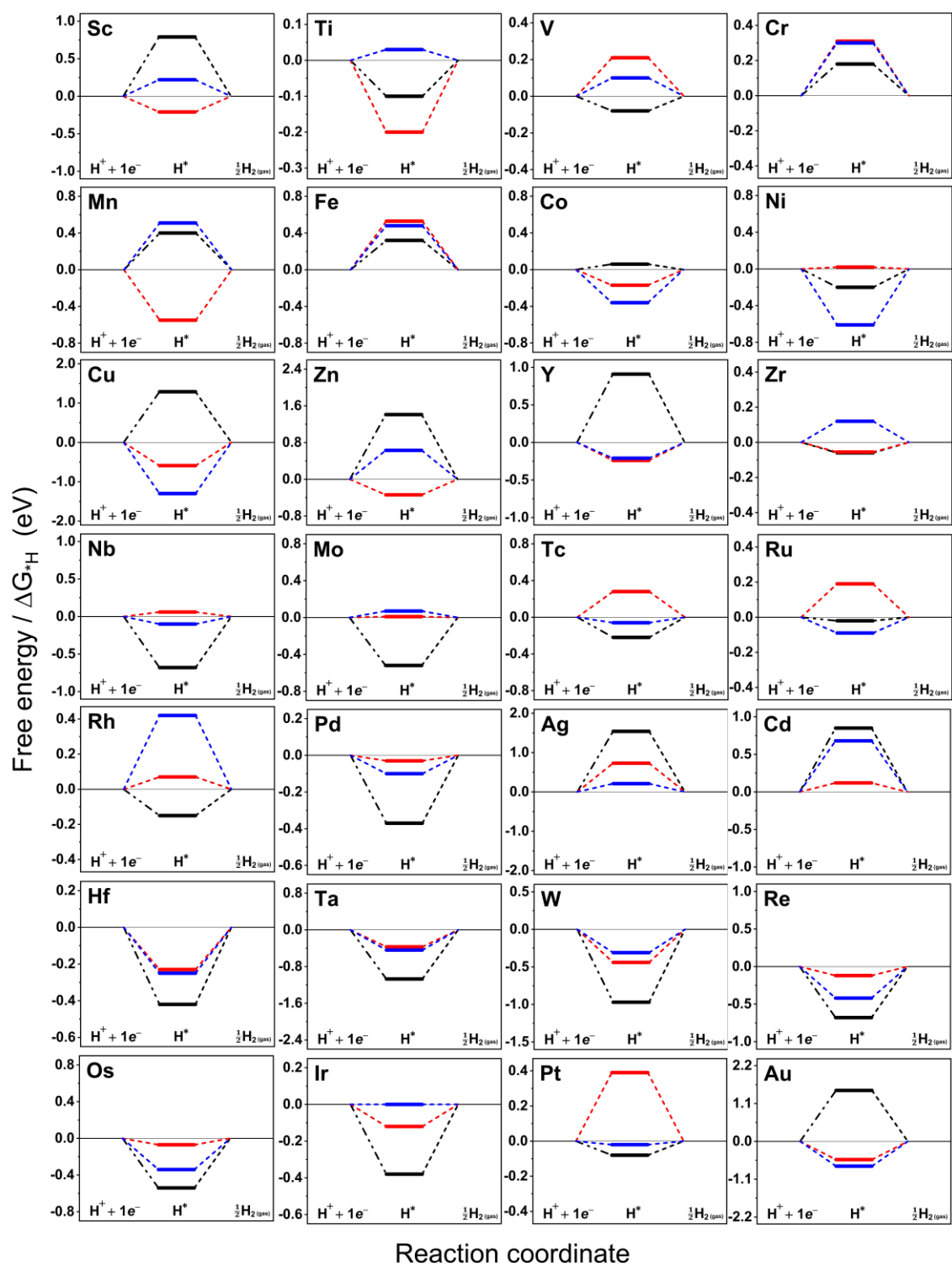


Figure S10. Free-energy diagram for HER on a single TM atom embedded at a double vacancy site in graphene support with C, N and B hybrid/dual coordination type configurations, i.e. N₂C₂, B₂C₂ and B₂N₂ which are represented with black, blue and red lines, respectively at the zero-electrode potential.

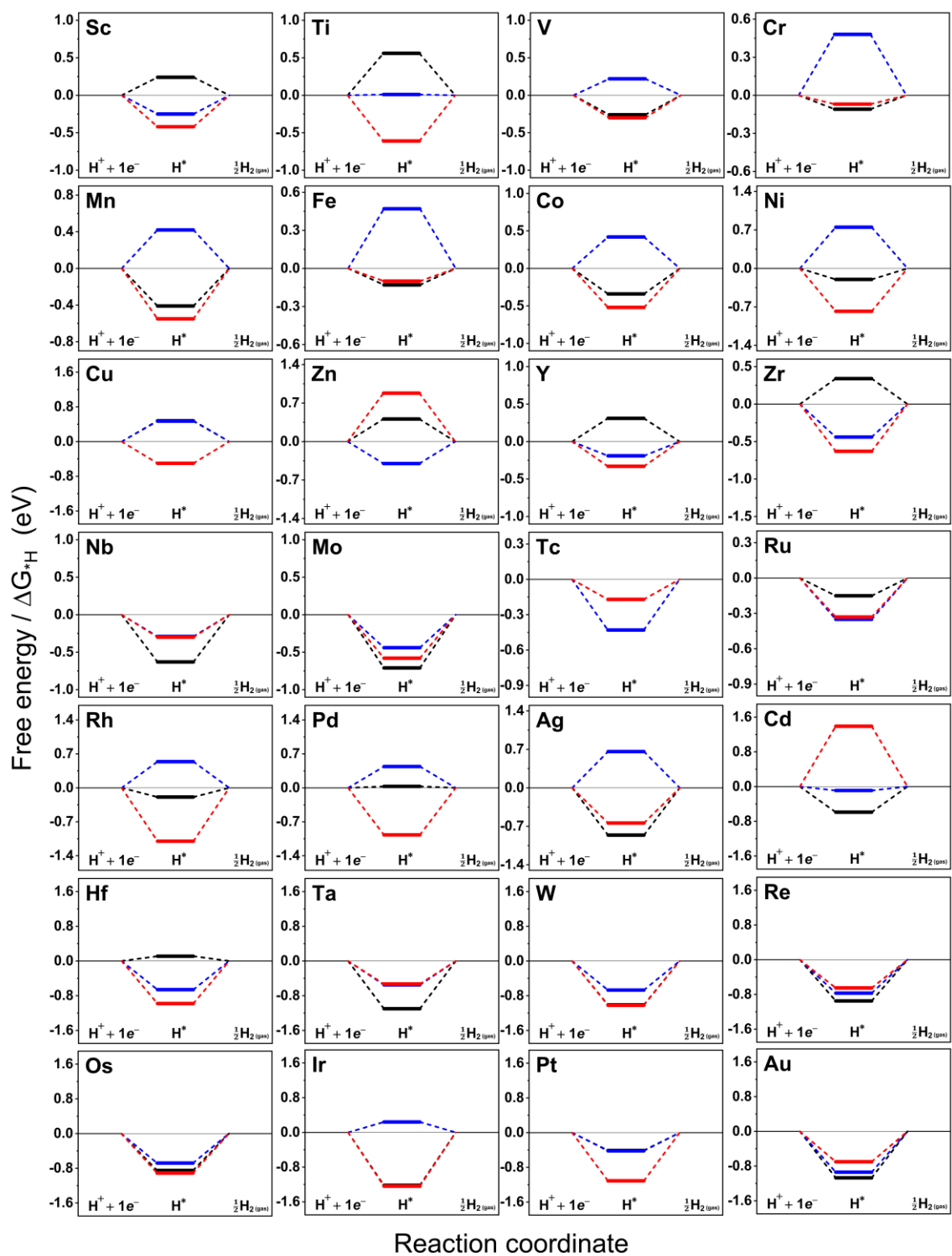


Figure S11. Free-energy diagram for HER on a single TM atom embedded at a single and double vacancy site in 2DCP, $g\text{-C}_3\text{N}_4$ and $h\text{-B}_2\text{N}_2$ templates, represented with black, blue and red lines, respectively, at the zero-electrode potential.

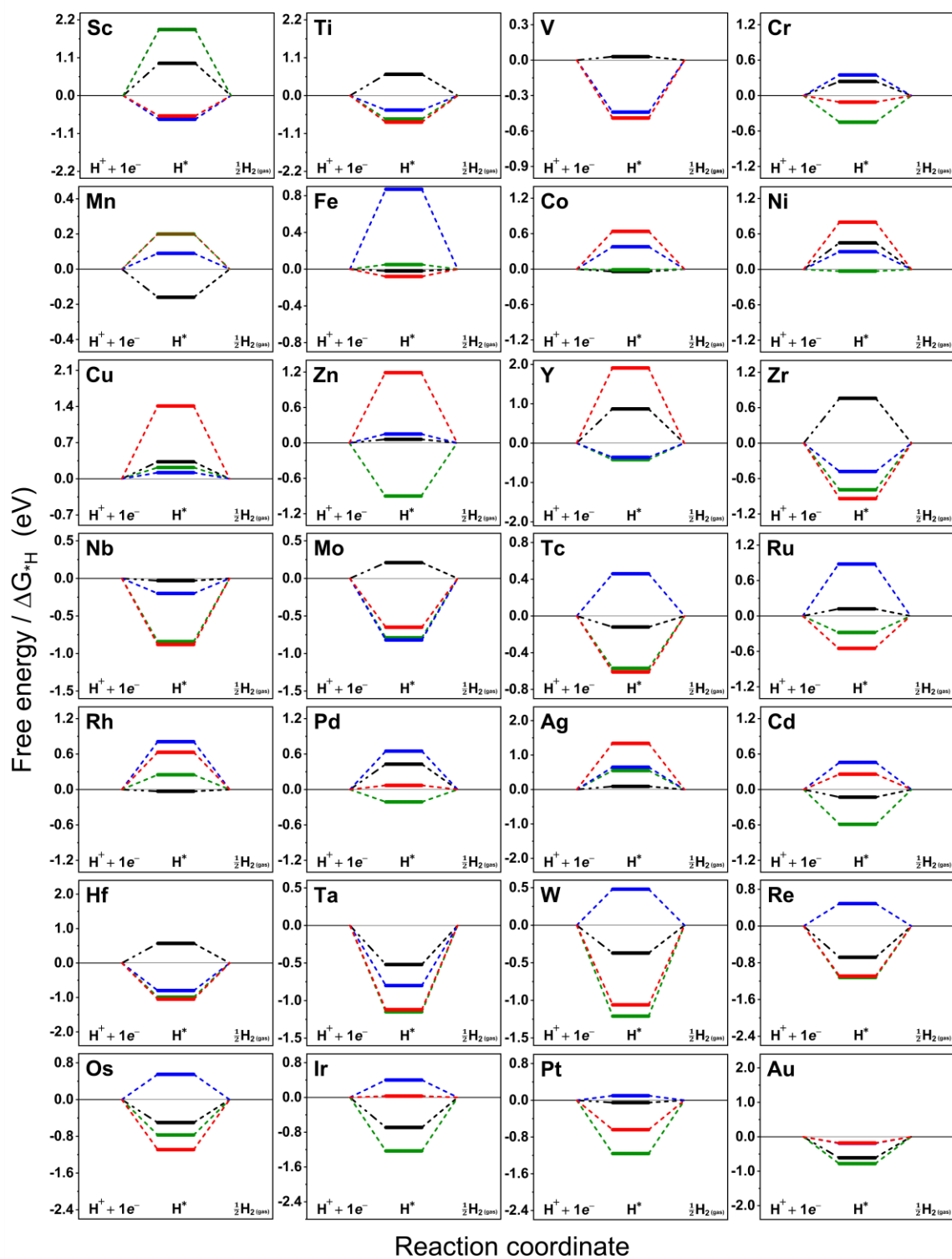


Figure S12. Free-energy diagram for HER on a single TM atom embedded at a single vacancy site in C₃, N₃, B₃ and *h*-BN templates, represented with black, blue and red and green lines, respectively, at the zero-electrode potential.

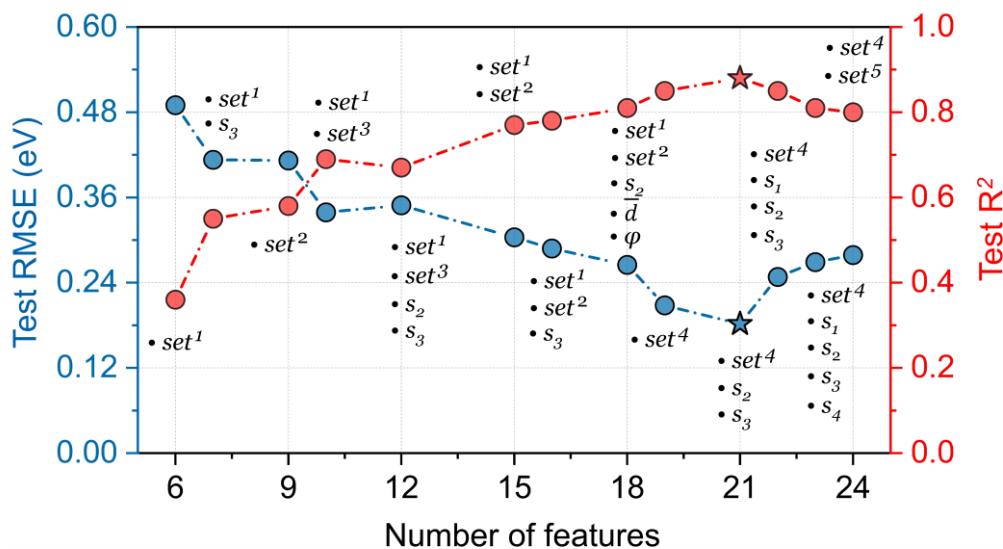


Figure S13. CatBoost model prediction accuracy by the stepwise feature selection method. Here, set^1 contains six elemental features from the periodic table properties (Z , r_{cov} , χ , r_v , θ_d), set^2 contains top nine coulomb matrix elements between the adsorbed hydrogen atom, TM and the neighboring non-metal atoms (i.e. C, N, B, P) sorted according to the distance in an ascending order of their DFT-optimized geometries. set^3 labeled with four DFT calculated features i.e ($\bar{d}$, φ , ε^{ho} , ε^{lu}). set^4 contains all the features from set^1 , set^2 and set^3 feature sets. Top five SISSO generated descriptors are labeled as set^5 . The best score (labeled with asterisk) is achieved by using all the features from set^4 and two SISSO generated descriptors (s_2 & s_3).

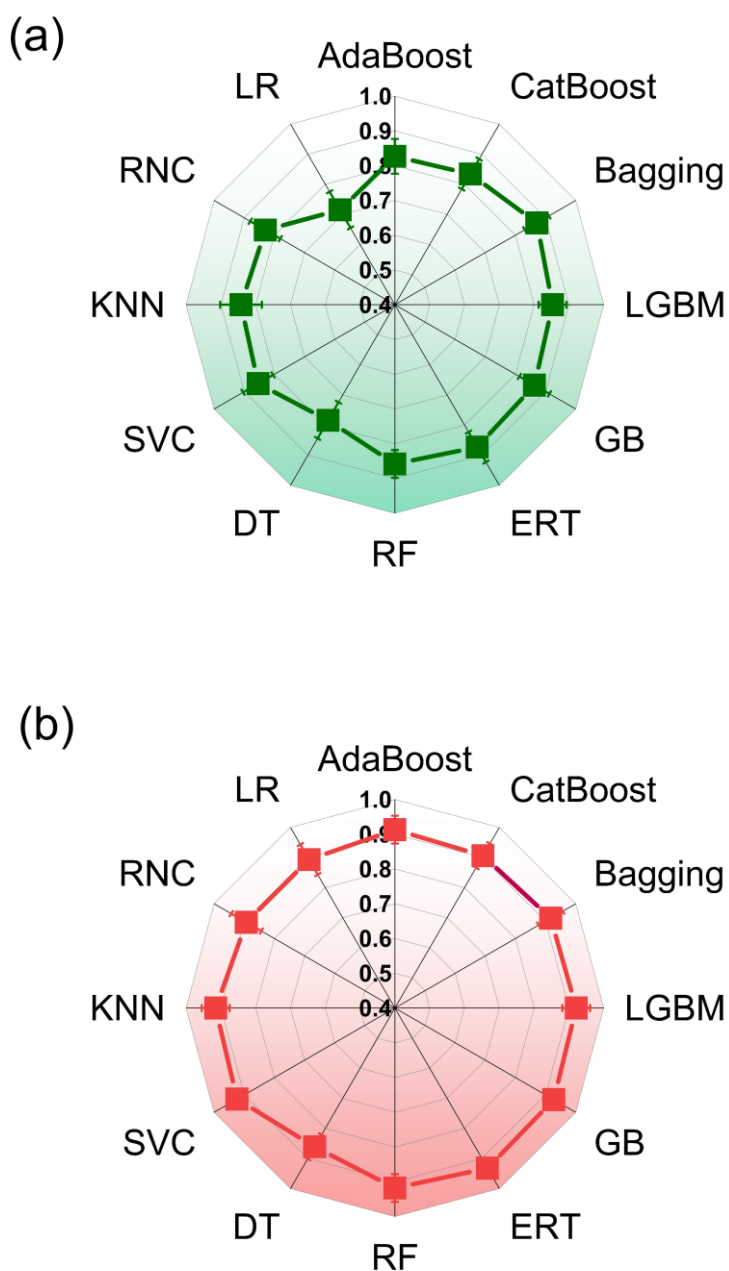


Figure S14. Comparison of predictive accuracy of various ML classification models for the prediction of **(a)** thermodynamic (E_{stab}) stability energy and **(b)** electrochemical dissolution potential (U_{diss}) values of various transition metal embedded systems. The mean and standard deviation values of the Area Under the Receiver Operating Characteristic Curve (ROC AUC) are presented for each ML classification model (for detail see Table S12).

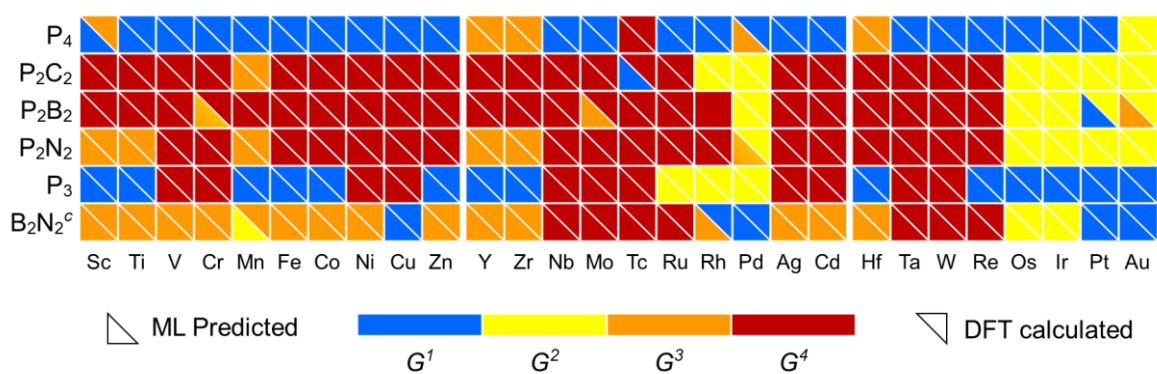


Figure S15. Stability classification of mono- and dual-type non-metal doped novel 2D surfaces. The classification is based on TM embedding over metal bulk cohesive energies (E_{stab}) and electrochemical dissolution potential (U_{diss}) values. Four different classes ($G^1/G^2/G^3/G^4$) of stabilities are represented with blue/yellow/orange/brown color [G^1 : unaggregatable and indissoluble, G^2 : aggregatable and indissoluble, G^3 : unaggregatable and dissoluble, G^4 : aggregatable and dissoluble].

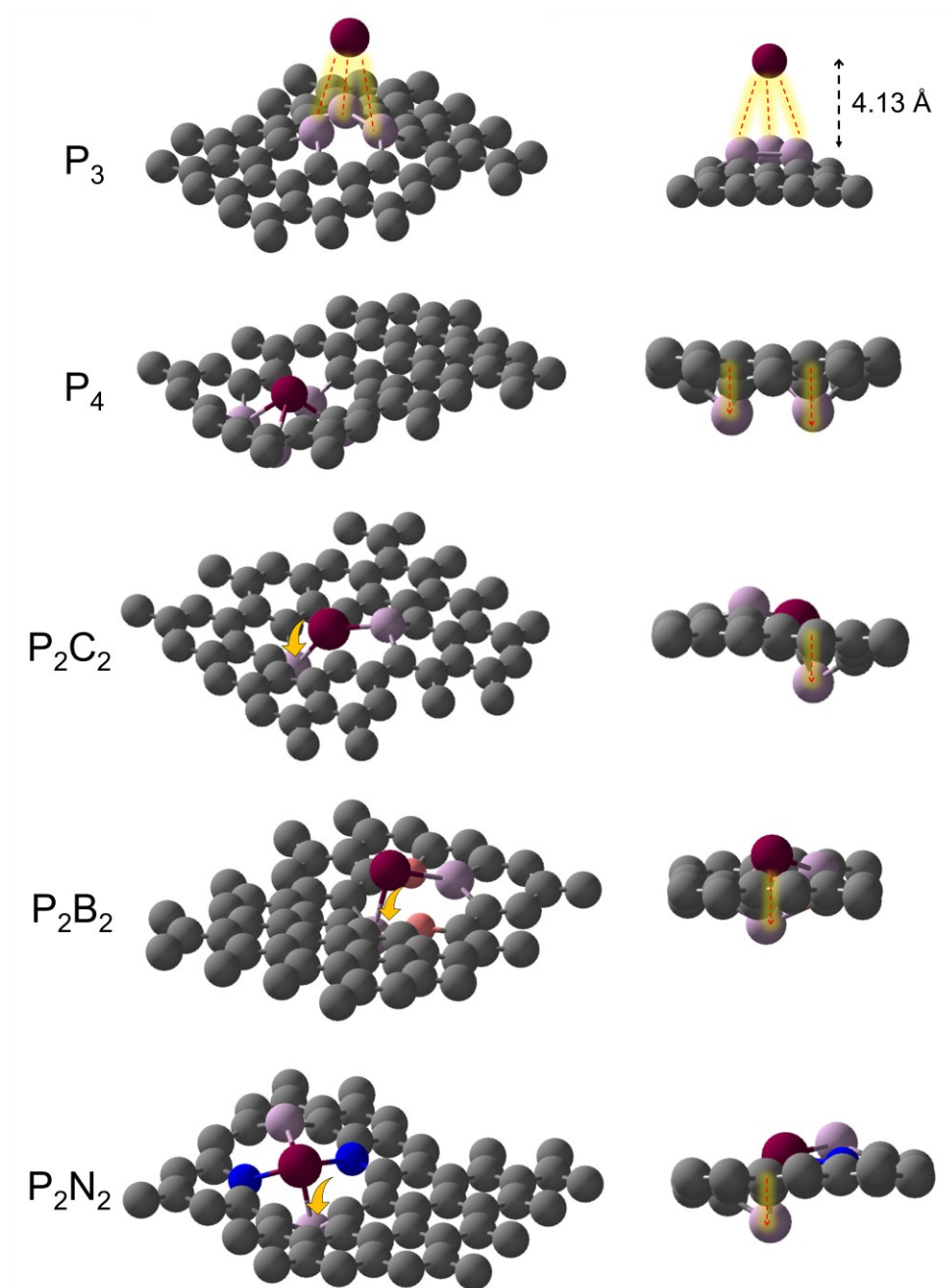


Figure S16. Top and side views of pyridinic graphene structures with single and double vacancy defects sites occupied with 28 kind of TM atoms. In each configuration a TM atom is coordinating with mono- or dual-type non-metal atoms (i.e. B, N, P). Here the case for TM atom is Cu (wine-red).

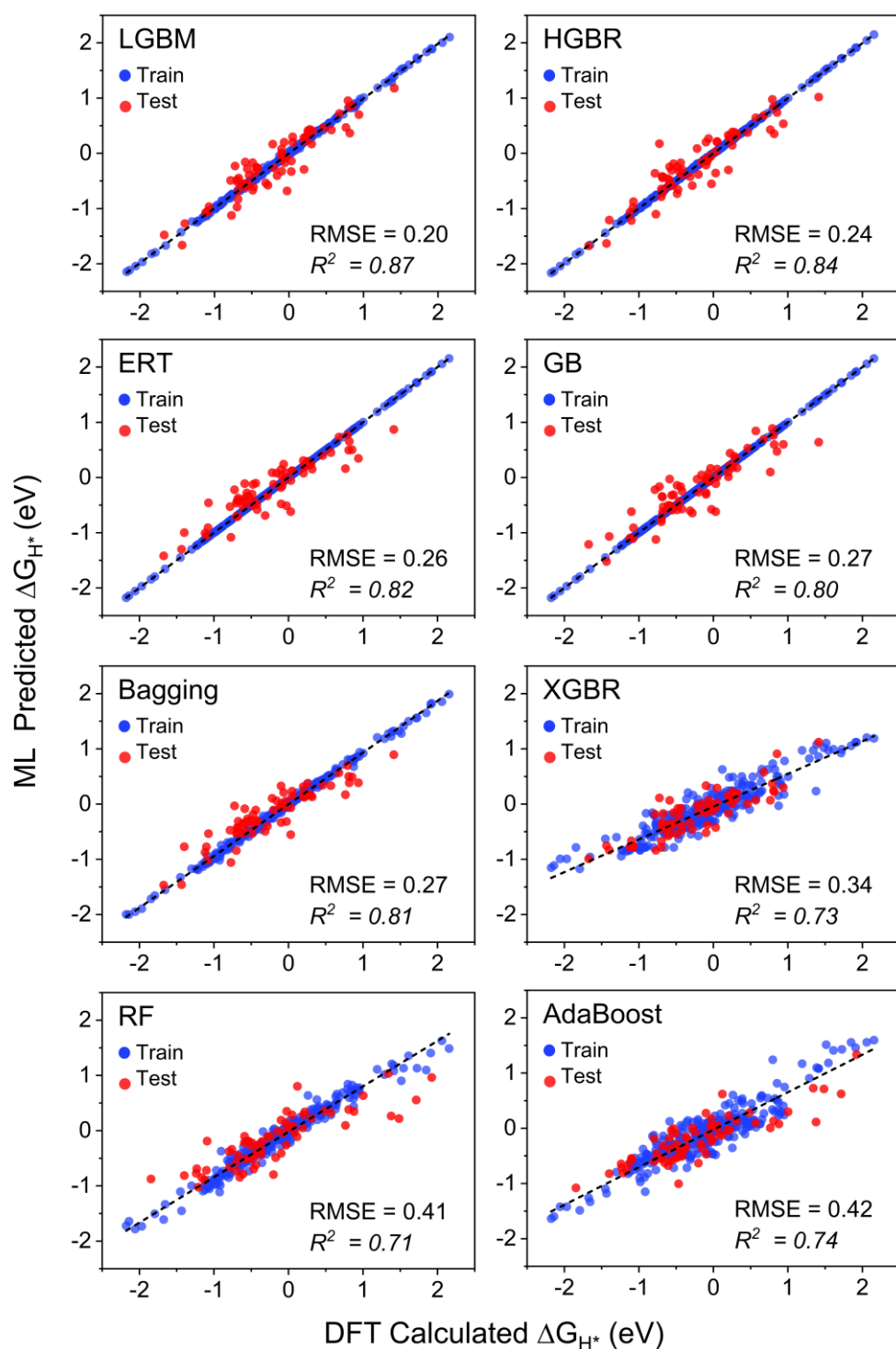


Figure S17. Comparison of several typical ML regression models for predicting HER activities. [LGBM-Light Gradient Boosting Machine, HGBR-Histogram-based Gradient Boosting Regressor, ERT-Extremely Randomized Trees, GB-Gradient Boosting, Bagging, XGBR-XGBoost regressor, RF-Random Forest and AdaBoost].

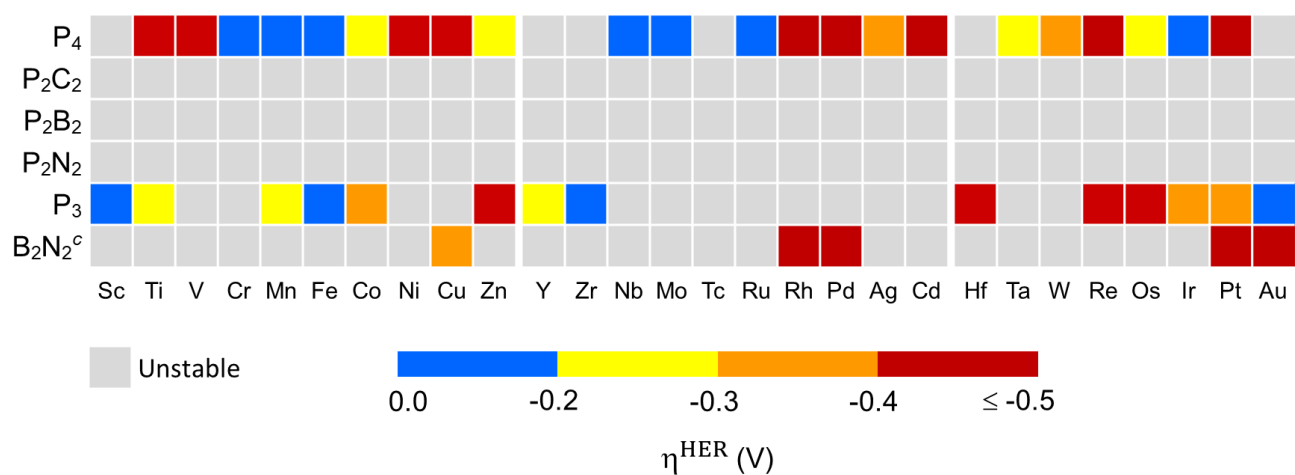


Figure S18. ML predicted HER activities. Since the ERT-classifier suggested a total 41 stable candidates (see Figure S15), so we assessed the HER activities for these systems by using trained ML (CatBoost) regression model. The predicted activities for promising catalysts were further verified by DFT calculations (see **Table S14**).

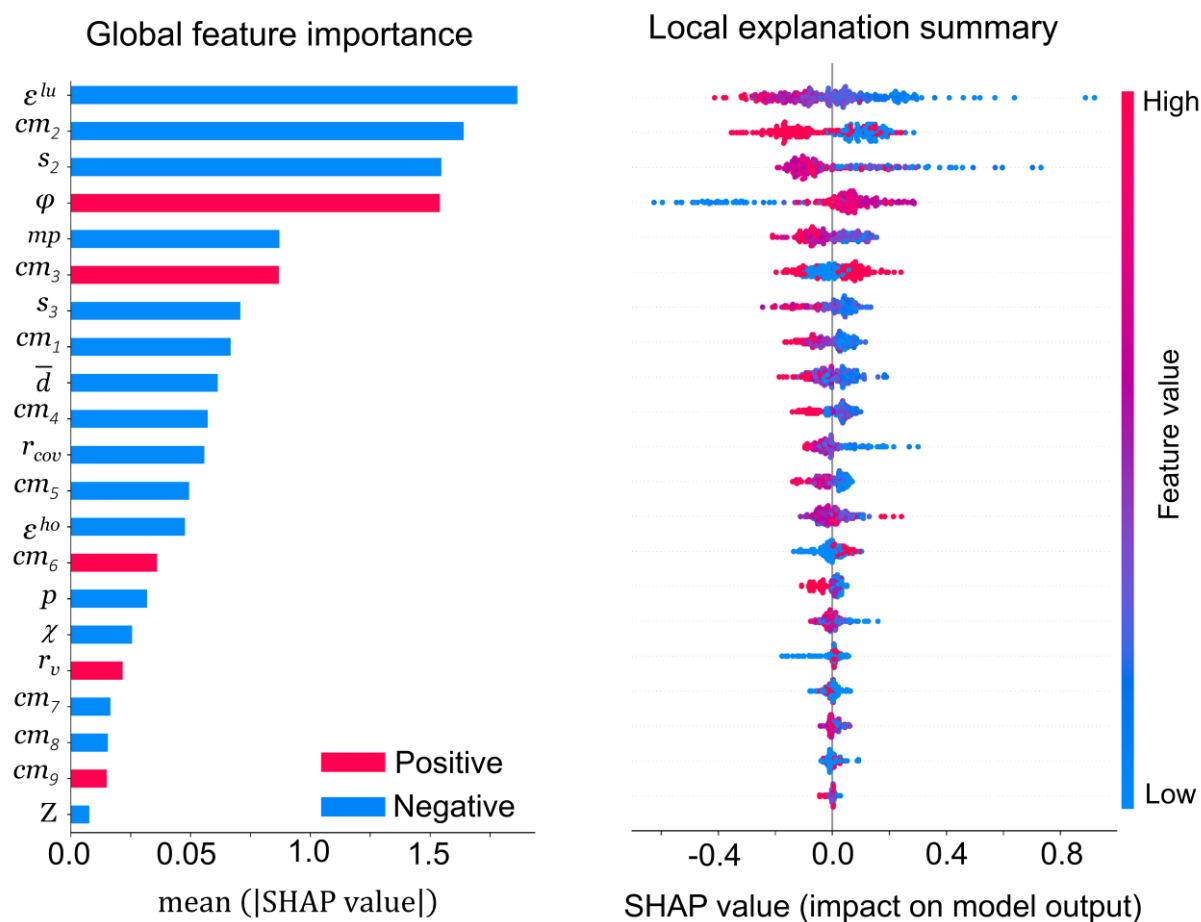


Figure S19. Feature importance based on SHAP-values. **(a)** The mean absolute SHAP-values are depicted, to illustrate global feature importance (ranked from most to least important). Positive and negative correlation coefficients for each feature/descriptor are represented by blue and red color respectively. **(b)** The local explanation summary shows the distribution of the impacts of each feature/descriptor on the HER activities. The color bar denotes the feature values, where the red points refer to higher values, while blue points represent lower values.

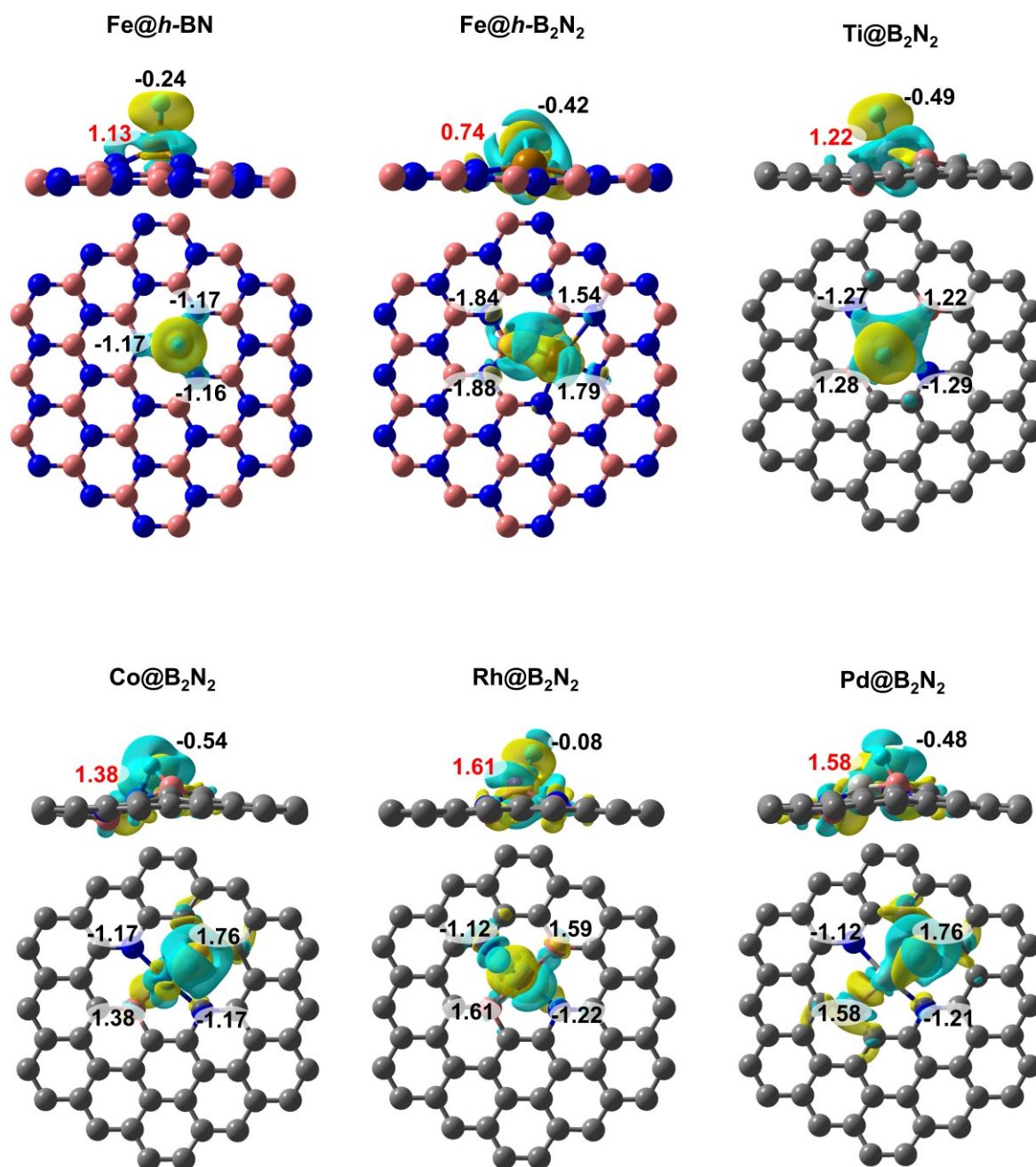


Figure S20 Isosurfaces of differential charge density distribution for chemisorbed H atom of different 2D systems. The charge depletion and accumulation are depicted as cyan and yellow color, respectively. The isosurface value is 0.003 e/Å³, and Bader charges are presented for the neighboring atoms of active site. The charge at TM atom is shown in red font.

Supplementary Tables

Table S1. TM embedding energy (E_{emb}) for 2D systems having a TM atom embedded at single and double vacancy defects sites with mono- or dual-type non-metal doped configurations.

TM	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	<i>g</i> -C ₃ N ₄	<i>h</i> -B ₂ N ₂	C ₃	N ₃	B ₃	<i>h</i> -BN
Sc	-6.53	-8.54	-5.42	-8.79	-4.63	-4.60	-7.12	-6.43	-8.88	-6.71	-6.23	-3.66	-11.77
Ti	-8.36	-8.63	-5.66	-9.71	-4.88	-5.63	-8.53	-6.09	-9.26	-8.34	-6.51	-3.70	-11.93
V	-7.79	-7.98	-5.11	-8.67	-4.70	-5.31	-7.44	-4.94	-8.76	-7.60	-6.03	-2.94	-10.73
Cr	-6.15	-6.98	-3.69	-7.19	-3.19	-3.59	-6.02	-3.20	-7.16	-6.39	-3.86	-1.98	-9.27
Mn	-6.18	-6.74	-3.77	-7.27	-3.64	-4.27	-5.33	-3.16	-7.36	-6.32	-3.99	-2.22	-8.30
Fe	-6.79	-7.31	-4.84	-7.68	-4.85	-5.10	-6.50	-3.17	-7.92	-7.29	-4.52	-3.15	-8.59
Co	-7.17	-7.83	-5.55	-8.17	-5.37	-5.50	-6.65	-3.14	-7.98	-7.84	-5.01	-3.90	-9.00
Ni	-7.15	-7.83	-5.81	-8.52	-5.66	-5.61	-5.87	-3.31	-7.85	-6.91	-4.59	-4.18	-8.10
Cu	-5.93	-5.40	-4.27	-6.51	-3.86	-3.20	-3.79	-2.29	-6.75	-3.85	-3.24	-2.55	-6.20
Zn	-3.49	-3.63	-1.59	-3.91	-0.67	-1.01	-2.52	-0.16	-5.56	-1.36	-1.68	-0.76	-4.65
Y	-6.42	-8.46	-5.77	-8.60	-4.89	-4.24	-7.02	-7.10	-8.83	-6.51	-6.22	-4.04	-11.57
Zr	-9.19	-9.19	-6.58	-10.61	-5.96	-6.49	-9.11	-7.60	-9.98	-8.85	-6.59	-4.42	-12.32
Nb	-9.46	-8.23	-6.43	-9.90	-6.18	-6.79	-8.37	-6.49	-9.85	-8.54	-6.02	-4.15	-11.13
Mo	-8.02	-6.49	-5.41	-8.21	-5.36	-5.65	-6.77	-3.79	-8.36	-7.59	-4.54	-2.69	-9.66
Tc	-8.93	-7.29	-6.66	-8.81	-6.53	-6.43	-7.77	-3.80	-9.27	-8.61	-5.33	-4.09	-9.64
Ru	-8.75	-7.22	-6.97	-8.73	-6.61	-6.39	-8.05	-3.61	-9.02	-9.03	-5.07	-4.80	-9.26
Rh	-7.88	-7.51	-7.02	-8.30	-6.52	-6.54	-6.97	-3.59	-8.41	-8.46	-4.25	-5.26	-8.91
Pd	-5.50	-5.96	-4.90	-6.56	-4.49	-4.68	-4.39	-2.01	-6.34	-5.45	-2.47	-3.67	-5.80
Ag	-3.46	-2.33	-2.53	-4.20	-2.01	-1.79	-2.10	-1.64	-4.67	-1.95	-1.90	-1.82	-3.76
Cd	-1.50	-1.69	-0.60	-2.14	-0.50	-0.63	-1.06	-0.12	-3.82	-0.24	-0.83	-0.71	-2.77
Hf	-9.19	-9.39	-6.22	-10.59	-5.82	-6.35	-8.99	-7.45	-9.92	-8.89	-6.63	-4.10	-12.70
Ta	-10.22	-9.16	-6.75	-10.76	-6.73	-7.42	-8.85	-6.85	-10.64	-9.25	-6.40	-4.34	-11.79
W	-9.96	-8.17	-6.89	-10.01	-7.05	-7.41	-8.38	-5.22	-10.10	-9.17	-5.82	-4.43	-11.04
Re	-9.02	-6.90	-6.48	-8.72	-6.51	-6.27	-7.44	-2.98	-9.17	-8.32	-4.43	-4.17	-9.02
Os	-9.86	-7.94	-7.84	-9.56	-7.65	-7.23	-8.41	-3.73	-9.87	-9.64	-4.99	-5.80	-9.52
Ir	-9.31	-8.52	-8.30	-9.51	-7.92	-7.57	-7.78	-3.83	-9.69	-9.40	-4.18	-6.52	-9.44
Pt	-8.02	-7.86	-7.14	-8.68	-6.81	-6.56	-6.12	-2.52	-8.59	-7.36	-2.99	-5.79	-7.13
Au	-5.20	-3.30	-3.81	-5.56	-3.29	-2.37	-2.76	-0.98	-6.06	-2.62	-1.30	-2.53	-3.89

Table S2. Stability (E_{stab}) of TM embedding over metal cohesion for the 2D systems having TM atom embedded at single and double vacancy defects sites with mono- or dual-type non-metal doped configurations. The E_{stab} represents the difference between adsorption energy and cohesive energy (eV) of various TM single atom catalysts ($E_{emb} - E_{coh}$). Here the negative values mean that embedding is stable against metal clustering or being leached.

TM	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	g-C ₃ N ₄	h-B ₂ N ₂	C ₃	N ₃	B ₃	h-BN
Sc	-1.54	-3.54	-0.43	-3.80	0.36	0.40	-2.13	-1.43	-3.88	-1.72	-1.24	1.33	-6.77
Ti	-1.93	-2.20	0.77	-3.28	1.56	0.80	-2.10	0.34	-2.82	-1.91	-0.08	2.73	-5.50
V	-1.26	-1.46	1.41	-2.15	1.83	1.22	-0.92	1.59	-2.23	-1.07	0.50	3.58	-4.20
Cr	-0.78	-1.62	1.68	-1.82	2.18	1.78	-0.65	2.17	-1.79	-1.02	1.51	3.39	-3.90
Mn	-2.43	-2.99	-0.03	-3.53	0.11	-0.53	-1.59	0.58	-3.62	-2.57	-0.24	1.53	-4.56
Fe	-1.15	-1.67	0.80	-2.04	0.79	0.54	-0.86	2.47	-2.28	-1.65	1.12	2.49	-2.95
Co	-1.25	-1.91	0.37	-2.25	0.54	0.42	-0.73	2.78	-2.06	-1.92	0.91	2.02	-3.08
Ni	-1.43	-2.11	-0.08	-2.79	0.06	0.12	-0.14	2.41	-2.13	-1.19	1.14	1.54	-2.37
Cu	-1.84	-1.30	-0.18	-2.42	0.24	0.89	0.30	1.80	-2.66	0.24	0.85	1.54	-2.11
Zn	-1.98	-2.13	-0.08	-2.40	0.83	0.49	-1.02	1.35	-4.05	0.15	-0.17	0.74	-3.15
Y	-1.54	-3.58	-0.89	-3.72	-0.01	0.64	-2.14	-2.23	-3.95	-1.63	-1.34	0.84	-6.69
Zr	-1.87	-1.87	0.74	-3.29	1.36	0.83	-1.78	-0.28	-2.66	-1.53	0.73	2.91	-5.00
Nb	-1.15	0.08	1.88	-1.59	2.13	1.52	-0.06	1.82	-1.54	-0.23	2.29	4.16	-2.82
Mo	-0.39	1.14	2.22	-0.58	2.28	1.98	0.86	3.85	-0.73	0.05	3.09	4.95	-2.03
Tc	0.13	1.77	2.41	0.25	2.54	2.64	1.30	5.27	-0.20	0.46	3.74	4.98	-0.57
Ru	-1.06	0.47	0.72	-1.05	1.08	1.30	-0.36	4.08	-1.33	-1.34	2.62	2.89	-1.57
Rh	-1.42	-1.05	-0.56	-1.84	-0.07	-0.08	-0.51	2.87	-1.95	-2.00	2.20	1.20	-2.45
Pd	-1.53	-1.99	-0.93	-2.59	-0.53	-0.71	-0.42	1.95	-2.38	-1.49	1.50	0.30	-1.83
Ag	-0.49	0.65	0.45	-1.22	0.97	1.18	0.88	1.34	-1.69	1.02	1.08	1.16	-0.79
Cd	-0.38	-0.58	0.52	-1.02	0.62	0.48	0.06	1.00	-2.70	0.88	0.29	0.40	-1.66
Hf	-1.78	-1.98	1.19	-3.18	1.59	1.06	-1.58	-0.04	-2.51	-1.48	0.78	3.31	-5.29
Ta	-1.09	-0.03	2.38	-1.63	2.39	1.70	0.28	2.28	-1.51	-0.12	2.72	4.79	-2.66
W	-0.16	1.63	2.91	-0.21	2.75	2.39	1.41	4.57	-0.30	0.63	3.98	5.37	-1.24
Re	-0.35	1.76	2.19	-0.06	2.15	2.39	1.22	5.68	-0.50	0.34	4.23	4.49	-0.35
Os	-0.58	1.34	1.43	-0.28	1.63	2.05	0.87	5.55	-0.59	-0.36	4.29	3.47	-0.25
Ir	-1.54	-0.75	-0.54	-1.74	-0.16	0.20	-0.02	3.94	-1.92	-1.63	3.58	1.25	-1.67
Pt	-2.07	-1.91	-1.19	-2.73	-0.86	-0.61	-0.17	3.44	-2.64	-1.41	2.96	0.17	-1.18
Au	-1.79	0.12	-0.39	-2.15	0.13	1.04	0.65	2.43	-2.65	0.80	2.11	0.88	-0.48

Table S3. Dissolution potential (U_{diss}) for 2D systems having TM atom embedded at single and double vacancy defects sites with mono- or dual-type non-metal doped configurations. Where the U_{diss}° and n_e represents the standard dissolution potential of bulk TMs and number of transferred electrons during the dissolution process respectively.

TM	U_{diss}°/n_e	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	<i>g</i> -C ₃ N ₄	<i>h</i> -B ₂ N ₂	C ₃	N ₃	B ₃	<i>h</i> -BN
Sc	-2.08/3	-1.57	-0.90	-1.94	-0.81	-2.20	-1.70	-1.37	-1.60	-0.79	-1.51	-1.67	-1.34	0.18
Ti	-1.63/2	-0.67	-0.53	-2.02	0.01	-2.41	-1.06	-0.58	-1.80	-0.22	-0.68	-1.59	-1.90	1.12
V	-1.18/2	-0.55	-0.45	-1.89	-0.11	-2.09	-1.16	-0.72	-1.97	-0.07	-0.64	-1.43	-2.24	0.92
Cr	-0.91/2	-0.52	-0.10	-1.75	0.00	-2.00	-1.41	-0.59	-1.99	-0.01	-0.40	-1.66	-1.80	1.04
Mn	-1.19/2	0.03	0.31	-1.18	0.57	-1.24	0.29	-0.39	-1.48	0.62	0.10	-1.07	-0.46	1.09
Fe	-0.45/2	0.12	0.39	-0.85	0.57	-0.85	-0.14	-0.02	-1.69	0.69	0.38	-1.01	-0.86	1.03
Co	-0.28/2	0.34	0.68	-0.46	0.85	-0.55	0.14	0.09	-1.67	0.75	0.68	-0.74	-0.33	1.26
Ni	-0.26/2	0.45	0.79	-0.22	1.14	-0.29	0.39	-0.19	-1.47	0.80	0.33	-0.83	0.02	0.93
Cu	0.34/2	1.26	0.99	0.43	1.55	0.22	0.81	0.19	-0.56	1.67	0.22	-0.09	0.22	1.39
Zn	-0.76/2	0.23	0.30	-0.72	0.44	-1.18	-0.01	-0.25	-1.43	1.27	-0.83	-0.67	-0.07	0.81
Y	-2.37/3	-1.86	-1.18	-2.07	-1.13	-2.37	-2.07	-1.66	-1.63	-1.05	-1.83	-1.92	-1.46	-0.14
Zr	-1.45/4	-0.98	-0.98	-1.63	-0.63	-1.79	-1.19	-1.00	-1.38	-0.79	-1.07	-1.63	-1.71	-0.20
Nb	-1.10/3	-0.72	-1.13	-1.73	-0.57	-1.81	-1.22	-1.08	-1.71	-0.59	-1.02	-1.86	-2.51	-0.16
Mo	-0.20/3	-0.07	-0.58	-0.94	-0.01	-0.96	-0.73	-0.49	-1.48	0.04	-0.22	-1.23	-2.23	0.48
Tc	0.40/2	0.33	-0.49	-0.80	0.27	-0.87	-0.99	-0.25	-2.23	0.50	0.17	-1.47	-2.98	0.69
Ru	0.46/2	0.99	0.22	0.10	0.98	-0.08	0.34	0.64	-1.58	1.13	1.13	-0.85	-1.22	1.25
Rh	0.60/2	1.31	1.13	0.88	1.52	0.63	1.35	0.86	-0.83	1.57	1.60	-0.50	0.53	1.83
Pd	0.95/2	1.72	1.95	1.41	2.24	1.21	2.07	1.16	-0.03	2.14	1.69	0.20	1.80	1.86
Ag	0.80/1	1.29	0.15	0.35	2.02	-0.17	0.10	-0.08	-0.54	2.49	-0.22	-0.28	-1.01	1.59
Cd	-0.40/2	-0.21	-0.11	-0.66	0.11	-0.71	-0.45	-0.43	-0.90	0.95	-0.84	-0.55	-0.32	0.43
Hf	-1.55/4	-1.10	-1.05	-1.85	-0.75	-1.95	-1.37	-1.16	-1.54	-0.92	-1.18	-1.74	-1.88	-0.23
Ta	-0.60/3	-0.24	-0.59	-1.39	-0.06	-1.40	-0.80	-0.69	-1.36	-0.10	-0.56	-1.51	-2.19	0.29
W	0.10/3	0.15	-0.44	-0.87	0.17	-0.82	-0.64	-0.37	-1.42	0.20	-0.11	-1.23	-2.23	0.51
Re	0.30/3	0.42	-0.29	-0.43	0.32	-0.42	-0.38	-0.11	-1.59	0.47	0.19	-1.11	-1.78	0.42
Os	0.84/8	0.91	0.67	0.66	0.88	0.64	0.66	0.73	0.15	0.91	0.89	0.30	0.24	0.87
Ir	1.16/3	1.67	1.41	1.34	1.74	1.21	1.61	1.17	-0.15	1.80	1.70	-0.03	0.99	1.72
Pt	1.18/2	2.22	2.14	1.77	2.55	1.61	2.52	1.26	-0.54	2.50	1.88	-0.30	2.05	1.77
Au	1.50/3	2.10	1.46	1.63	2.22	1.46	1.75	1.28	0.69	2.38	1.23	0.80	1.17	1.66

Table S4. Stability classification based on TM embedding over metal bulk cohesive energies (E_{stab}) and electrochemical dissolution potential (U_{diss}) values for 364 mono- and dual type non-metal doped 2D substrates into four different categories [G^1 : unaggregatable and indissoluble ($E_{stab} < 0$ and $U_{diss} > 0$), G^2 : aggregatable and indissoluble ($E_{stab} > 0$ and $U_{diss} > 0$), G^3 : unaggregatable and dissoluble ($E_{stab} < 0$ and $U_{diss} < 0$), G^4 : aggregatable and dissoluble ($E_{stab} > 0$ and $U_{diss} < 0$)].

2D-Substrate	G^1	G^2	G^3	G^4
C₄	16	1	11	-
N₄	10	4	9	5
B₄	6	3	5	14
N₂C₂	19	1	8	-
B₂C₂	4	3	-	20
B₂N₂	4	8	-	16
2DCP	6	3	12	7
<i>g</i>-C₃N₄	-	2	4	22
<i>h</i>-B₂N₂	19	-	9	-
<i>h</i>-BN	24	-	4	-
C₃	10	4	9	5
N₃	-	3	5	20
B₃	-	8	-	20
Total	118	40	77	129

Table S5. Magnetization values (Bohr) for pristine and H-adsorbed 2D systems having TM atom embedded at single and double vacancy defects sites with mono- or dual-type non-metal coordinated configurations.

TM	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	<i>g</i> -C ₃ N ₄	<i>h</i> -B ₂ N ₂	C ₃	N ₃	B ₃	<i>h</i> -BN
Sc	0.9/0	0/0	0/0	0/0	0/0	0/0	0.97/0	0.8/1.7	1.0/0	0.2/0	0.2/0	0/0	0/1.0
Ti	0/0	1.6/0.6	2.0/1.0	0/0	0/1.1	1.7/1.0	0/0.99	1.7/0.9	0/1.0	0/0.2	1.9/1.6	1.0/0	1.0/0
V	0.8/0	3.0/2.0	1.9/2.0	1.1/0.7	2.3/1.2	0.7/0	1.0/3.99	1.2/2.0	3.0/2.0	1.0/0	2.4/1.4	2.0/2.7	2.0/1.0
Cr	1.8/1.1	4.0/3.0	3.8/3.0	2.3/2.2	3.5/0	3.8/1.0	2.0/1.0	5.9/3.0	4.0/3.0	2.0/1.4	4.4/0.4	5.0/4.0	3.0/2.0
Mn	2.8/1.9	3.0/2.0	3.0/4.0	3.4/3.5	3.0/3.5	2.8/2.4	3.0/2.0	3.2/4.0	5.0/4.0	2.7/2.0	5.1/0	5.0/5.0	4.0/1.0
Fe	2.5/2.3	2.0/1.0	1.8/1.0	2.5/1.8	2.0/1.4	2.0/2.6	0/1.0	2.2/3.0	4.0/1.0	0/0.6	3.4/1.6	2.4/0	5.0/0
Co	1.2/1.3	1.0/0	1.1/0	1.2/0	1.0/0	0.8/1.4	1.0/0	1.1/2.0	3.0/2.0	0.6/0	2.4/3.0	1.6/1.1	0/1.0
Ni	0/0.3	0/0.6	0/0	0/0	0/0.6	0/0	2.0/1.0	0/1.0	2.0/1.0	0/0.8	1.5/2.0	0/0.3	1.0/2.0
Cu	0/0	1.0/0.7	0/0	0/0	0/0	1.0/0	1.0/2.0	1.0/0	1.0/0	1.0/1.6	0.1/1.0	0.3/0	2.0/3.0
Zn	0/0	0/0	0/0	0/1.0	0.60/	0/0	0/0	0/1.0	0/1.0	1.5/2.6	0.6/0	0/0	1.0/2.0
Y	0/2.0	0/0	0.2/0	0/0	0/0	0/1.9	0.94/0	0/1.7	1.0/0	0.2/0	0.6/0	0/0	0/1.0
Zr	0/0	0.7/0	0/1.0	0/0	0/1.0	0/1.0	0/0.86	1.6/0	2.0/1.0	0/0.1	1.8/0.5	1.00/	1.0/0
Nb	0.6/0	2.5/0	1.0/0	0.6/0	0.3/0	0.3/0	1.0/0	1.0/0	1.0/2.0	1.0/0	2.3/1.4	0/1.0	2.0/1.0
Mo	0.4/0.8	3.6/1.0	0/1.0	2.0/1.0	0/0	0/0	0/1.0	2.1/0.9	2.0/1.0	2.0/1.4	1.4/0	1.0/0	3.0/2.0
Tc	1.0/0.7	3.0/2.0	0/0	1.8/1.7	0.6/0.9	1.3/0.5	1.0/0	1.0/0	1.0/2.0	1.0/1.6	0.4/0	0/3.0	2.0/1.0
Ru	0/0	2.0/1.0	1.1/1.0	0/0	0.5/0.9	0/0.5	0/0.80	0/1.0	0/1.0	0/0	0/0	1.0/0	1.0/0
Rh	0.8/0	0.7/0	0/0	0/0	0/0	0.2/0	1.0/0	0.9/1.9	1.0/0	0.90/	1.6/2.6	0.6/0.9	0/1.0
Pd	0.2/0	0/0.5	0/0	0/0	0/0	0/0	0/1.0	0/1.0	2.0/1.0	0/0.9	1.0/0	0/0	1.0/0
Ag	0/0	0.6/0	0/0	0/0	0/0	0/0	0.97/0	0/0	1.0/0	1.0/0	0/0.6	0/0	2.0/3.0
Cd	0/1.0	0/0	0/0	0/1.0	0/0	0/0	0/0	0/0.1	0/1.0	2.0/2.5	1.0/0	0/0	1.0/2.0
Hf	0/0	0.7/0	0/0.9	0/0	0/1.0	0.1/1.0	0/0.89	1.5/0	2.0/1.0	0/0	1.8/0.3	0.9/0	1.0/0
Ta	0.6/0	1.0/0	1.0/0	1.0/0	0.3/0	0.3/0	1.0/0	1.0/0	1.0/2.0	0.6/0	1.8/1.0	1.6/0	2.0/1.0
W	0/0.3	2.0/1.0	0/1.0	2.0/1.0	0.3/0	0/0	2.0/1.0	1.8/0.9	2.0/1.0	2.0/1.0	1.0/0	0.6/2.0	3.0/2.0
Re	1.0/1.3	3.0/2.0	0/0	2.6/2.0	0/0	1.0/0	1.0/0	1.0/0	1.0/2.0	1.0/2.0	0.4/0	0/1.0	2.0/1.0
Os	0/0	0/1.0	1.1/1.0	0/1.0	0.8/0.9	0/0.6	0/0.82	0/1.0	0/1.0	0/0.5	0/0	1.0/0	1.0/0
Ir	0.4/0	0.9/0	0/0	0/0	0/0	0.3/0	1.0/0	0.7/0	1.0/0	1.0/0	1.6/0.6	0.6/1.0	0/1.0
Pt	0/0.2	0/0	0/0	0/0	0/0	0/0	0/1.0	1.9/0.9	2.0/1.0	0/0.8	1.4/2.0	0/0	1.0/0
Au	0/0	0/0.6	0/0.1	0/0	0/0	0/0	1.01/0	0.9/0	1.0/0	1.0/0	0/0.6	0.3/0	2.0/1.0

Table S6. Bond distance (Å) between adsorbed hydrogen to TM atom embedded at single and double vacancy defects sites with mono- or dual-type non-metal doped configurations.

TM	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	<i>g</i> -C ₃ N ₄	<i>h</i> -B ₂ N ₂	C ₃	N ₃	B ₃	<i>h</i> -BN
Sc	1.87	1.86	1.84	1.89	1.83	1.85	1.89	1.80	1.87	1.91	1.88	1.84	1.97
Ti	1.76	1.74	1.74	1.75	1.75	1.73	1.79	1.69	1.75	1.79	1.78	1.75	1.76
V	1.66	1.67	1.69	1.68	1.66	1.66	1.74	1.65	1.69	1.71	1.71	1.70	1.69
Cr	1.61	1.60	1.80	1.62	1.59	1.60	1.66	1.62	1.66	1.66	1.64	1.67	1.63
Mn	1.57	1.53	1.62	1.57	1.97	1.57	1.63	1.67	1.75	1.62	1.59	1.64	1.57
Fe	1.52	1.49	1.66	1.48	1.50	1.58	1.58	1.61	1.62	1.57	1.56	1.56	1.52
Co	1.71	1.44	1.51	1.43	1.53	1.76	1.50	1.55	1.59	1.56	1.57	1.54	1.52
Ni	1.76	1.48	1.60	1.70	1.73	1.73	1.46	1.52	1.55	1.55	1.53	1.52	1.50
Cu	1.90	1.54	1.53	1.54	1.91	1.90	1.71	1.50	1.56	1.52	1.50	1.52	1.56
Zn	1.62	1.56	1.54	1.59	1.53	1.55	1.56	1.52	1.87	1.55	1.53	1.53	1.54
Y	2.04	2.04	2.01	2.06	2.00	2.02	2.00	1.97	2.04	2.05	2.06	1.98	2.13
Zr	1.91	1.89	1.88	1.91	1.89	1.88	1.92	1.83	1.91	1.97	1.94	1.89	1.93
Nb	1.80	1.80	1.80	1.82	1.80	1.79	1.86	1.74	1.82	1.86	1.85	1.82	1.84
Mo	1.74	1.72	1.73	1.75	1.72	1.74	1.79	1.66	1.74	1.80	1.77	1.74	1.76
Tc	1.71	1.62	1.71	1.69	1.67	1.67	1.72	1.60	1.71	1.80	1.71	1.72	1.69
Ru	1.65	1.58	1.67	1.62	1.65	1.61	1.70	1.56	1.64	1.69	1.66	1.68	1.63
Rh	1.59	1.53	1.62	1.60	1.60	1.57	1.66	1.52	1.68	1.68	1.65	1.65	1.65
Pd	1.96	1.65	1.60	1.93	1.85	1.89	1.64	1.52	1.66	1.67	1.56	1.65	1.59
Ag	1.99	1.62	1.64	1.67	1.64	1.61	1.62	1.62	1.70	1.65	1.62	1.66	1.66
Cd	1.70	1.69	1.69	1.69	1.68	1.71	1.69	1.69	1.68	1.68	1.67	1.69	1.88
Hf	1.89	1.87	1.86	1.88	1.88	1.86	1.90	1.80	1.89	1.86	1.91	1.88	1.91
Ta	1.80	1.79	1.80	1.81	1.80	1.78	1.71	1.73	1.81	1.80	1.84	1.82	1.84
W	1.75	1.72	1.75	1.76	1.73	1.74	1.79	1.67	1.75	1.80	1.77	1.79	1.77
Re	1.71	1.64	1.72	1.70	1.69	1.69	1.74	1.62	1.71	1.75	1.72	1.75	1.71
Os	1.68	1.61	1.70	1.65	1.67	1.64	1.70	1.58	1.66	1.71	1.69	1.71	1.66
Ir	1.63	1.56	1.65	1.62	1.63	1.61	1.64	1.59	1.66	1.68	1.63	1.68	1.66
Pt	1.80	1.59	1.62	1.83	1.61	1.88	1.65	1.54	1.63	1.69	1.59	1.66	1.61
Au	1.99	1.88	1.63	1.66	1.97	1.58	1.59	1.57	1.66	1.64	1.57	1.64	1.60

Table S7. Angle at vertex at a TM atom (M) enclosed by two lines pointing to the adsorbed-H (H) and to the first nearest neighbor non-metal atom (Y: C, N, B) coordinating the TM atom on the 2D surface ($\angle\text{H-M-Y}$).

TM	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	<i>g</i> -C ₃ N ₄	<i>h</i> -B ₂ N ₂	C ₃	N ₃	B ₃	<i>h</i> -BN
Sc	123.0	114.9	131.8	116.9	128.6	115.5	118.8	91.2	117.4	97.5	130.0	145.9	97.5
Ti	114.1	109.4	107.1	109.8	121.3	104.4	131.7	87.7	113.0	129.0	125.0	145.2	119.4
V	110.0	104.7	124.9	106.5	110.3	105.4	104.5	89.8	104.5	125.2	123.6	132.6	117.2
Cr	108.8	98.2	37.2	103.1	108.8	101.5	113.1	90.3	78.5	124.6	118.9	139.7	117.9
Mn	91.5	94.1	122.7	101.3	31.9	101.3	111.9	94.6	33.2	123.1	117.9	144.9	116.2
Fe	79.2	93.4	41.1	88.4	100.3	106.6	109.8	102.3	38.2	120.7	102.4	132.7	115.4
Co	34.0	92.8	60.0	89.9	43.3	37.4	92.8	98.2	37.0	120.6	124.4	132.2	100.8
Ni	32.6	96.0	41.5	34.4	38.5	37.4	91.3	107.5	37.6	88.6	104.8	131.1	99.2
Cu	31.8	105.5	117.9	92.9	31.3	35.1	115.5	113.3	35.2	101.0	107.3	135.4	120.9
Zn	102.5	114.3	125.6	104.9	124.0	91.3	133.3	111.2	31.9	131.6	130.1	142.9	129.9
Y	100.0	120.1	134.7	121.7	132.7	113.1	132.2	92.7	116.6	92.6	134.4	115.0	94.4
Zr	118.8	115.0	105.7	115.6	110.0	110.3	130.1	91.8	115.6	132.7	128.7	137.6	123.7
Nb	114.1	109.1	102.7	110.4	119.0	107.2	122.7	90.2	111.6	128.8	126.6	131.6	121.2
Mo	113.3	105.1	84.3	90.5	113.5	86.0	121.1	88.7	110.2	127.5	122.8	75.6	122.3
Tc	112.3	96.5	122.0	105.9	113.0	77.1	135.9	86.3	90.1	122.5	122.8	138.4	121.0
Ru	84.5	93.9	85.2	80.1	76.2	73.8	115.3	88.3	79.6	124.1	96.4	136.1	120.4
Rh	78.6	92.5	79.6	61.4	69.9	74.3	107.5	89.4	40.7	124.1	130.7	134.9	104.4
Pd	28.5	95.2	71.1	30.5	30.8	36.8	92.8	77.8	39.4	101.3	99.7	132.6	90.7
Ag	29.3	125.7	98.4	99.1	96.7	104.6	179.9	118.9	35.5	109.1	117.4	137.4	130.2
Cd	124.8	124.3	131.7	90.7	131.8	120.3	141.0	100.6	107.0	139.3	137.9	150.8	82.4
Hf	117.6	113.8	108.8	114.5	108.0	106.9	131.3	91.3	113.4	127.6	127.3	136.1	122.8
Ta	113.7	109.5	103.7	111.1	111.3	105.2	115.5	90.4	112.1	125.8	124.2	130.8	120.3
W	112.4	106.0	92.7	110.3	113.2	90.7	136.4	88.8	106.5	126.1	122.0	135.3	122.0
Re	112.6	98.9	122.1	105.6	112.1	77.6	134.5	88.2	98.5	125.3	123.3	137.0	121.5
Os	92.1	97.8	122.4	80.0	81.2	74.8	116.7	88.3	81.9	124.4	103.5	134.0	121.0
Ir	81.3	93.7	88.7	74.4	73.6	74.5	104.9	94.5	47.3	124.6	98.5	133.3	106.2
Pt	33.9	95.1	81.8	33.4	70.5	37.7	103.0	77.5	45.2	106.3	105.4	132.3	91.4
Au	28.0	96.5	83.1	98.4	34.7	76.6	157.4	121.8	39.4	105.1	124.5	135.8	93.7

Table S8. Hydrogen adsorption energies (eV) for 2D substrates each with TM atom embedded at single and double vacancy defects sites with mono- or dual-type non-metal doped configurations.

TM	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	<i>g</i> -C ₃ N ₄	<i>h</i> -B ₂ N ₂	C ₃	N ₃	B ₃	<i>h</i> -BN
Sc	0.71	-0.52	-0.04	0.55	-0.45	-0.02	0.00	-0.49	-0.66	0.70	-0.93	-0.83	1.68
Ti	0.30	-0.77	-0.28	-0.34	-0.44	-0.21	0.32	-0.23	-0.85	0.38	-0.92	-0.66	-1.01
V	-0.30	-0.39	-0.02	-0.32	-0.03	-0.14	-0.50	-0.02	-0.54	-0.21	-0.73	-0.68	-0.73
Cr	0.01	0.04	-0.59	-0.06	0.07	0.06	-0.35	0.24	-0.31	0.00	-0.69	0.11	-0.35
Mn	0.15	0.15	0.16	0.16	-0.79	0.27	-0.65	0.18	-0.79	-0.40	-0.04	-0.15	-0.04
Fe	0.01	0.04	-0.61	0.08	0.29	0.24	-0.37	0.23	-0.34	-0.26	-0.19	0.63	-0.32
Co	-0.61	-0.08	-0.37	-0.18	-0.41	-0.60	-0.58	0.18	-0.76	-0.28	-0.25	0.14	0.40
Ni	-1.01	1.37	-0.47	-0.44	-0.22	-0.85	-0.44	0.51	-1.02	0.21	-0.27	0.06	0.56
Cu	-0.68	1.48	0.66	1.05	-0.83	-1.54	0.24	0.23	-0.74	0.09	-0.02	-0.12	1.17
Zn	1.47	0.61	-0.34	1.17	-0.58	0.39	0.17	-0.64	0.64	-0.18	-1.14	-0.09	0.95
Y	0.76	-0.36	0.08	0.67	-0.48	-0.45	0.07	-0.43	-0.57	0.63	-0.66	-0.60	1.67
Zr	0.33	-1.14	-0.39	-0.30	-0.30	-0.12	0.10	-0.68	-0.87	0.52	-1.03	-0.72	-1.18
Nb	-0.49	-1.13	-0.34	-0.92	-0.18	-0.34	-0.87	-0.53	-0.54	-0.27	-1.08	-0.44	-1.12
Mo	-0.42	-0.68	-0.07	-0.76	-0.23	-0.17	-0.95	-0.68	-0.82	-0.03	-1.03	-1.06	-0.89
Tc	-0.16	-0.70	0.01	-0.46	0.04	-0.30	-0.61	-0.67	-0.41	-0.36	-0.81	0.22	-0.85
Ru	-0.08	-0.80	0.01	-0.26	-0.05	-0.33	-0.39	-0.59	-0.57	-0.12	-0.52	0.64	-0.79
Rh	-0.28	-0.50	-0.19	-0.39	-0.17	0.18	-0.43	0.30	-1.34	-0.27	0.01	0.57	0.39
Pd	-1.00	1.61	-0.23	-0.61	-0.27	-0.34	-0.21	0.20	-1.21	0.19	-0.45	0.41	-0.17
Ag	-1.00	0.58	0.58	1.30	0.49	-0.03	-1.10	0.42	-0.88	-0.15	0.31	0.41	1.10
Cd	1.14	-0.22	-0.51	0.61	-0.12	0.44	-0.83	-0.33	1.15	-0.37	-0.83	0.22	0.02
Hf	0.06	-1.37	-0.78	-0.66	-0.47	-0.49	-0.13	-0.90	-1.22	0.33	-1.23	-1.04	-1.29
Ta	-0.91	-1.32	-0.80	-1.31	-0.61	-0.68	-1.34	-0.79	-0.77	-0.76	-1.39	-1.04	-1.36
W	-0.91	-1.17	-0.50	-1.21	-0.68	-0.55	-1.25	-0.91	-1.26	-0.61	-1.45	0.24	-1.30
Re	-0.70	-1.04	-0.38	-0.92	-0.36	-0.66	-1.19	-1.01	-0.89	-0.92	-1.36	0.25	-1.33
Os	-0.44	-0.94	-0.33	-0.78	-0.31	-0.58	-1.09	-0.92	-1.15	-0.74	-1.01	0.31	-1.33
Ir	-0.63	-0.61	-0.46	-0.62	-0.36	-0.24	-1.46	0.00	-1.48	-0.93	-1.47	0.16	-0.21
Pt	-0.55	1.27	-0.46	-0.32	0.15	-0.26	-0.65	-0.66	-1.35	-0.29	-1.40	-0.14	-0.88
Au	-0.73	1.92	0.20	1.24	-0.77	-0.96	-1.31	-1.18	-0.94	-0.85	-1.02	-0.43	-0.42

Table S9. Hydrogen adsorption Gibbs free energies (eV) on a TM atom embedded at single and double vacancy defects sites with mono- or dual-type non-metal doped configurations.

TM	C ₄	N ₄	B ₄	N ₂ C ₂	B ₂ C ₂	B ₂ N ₂	2DCP	<i>g</i> -C ₃ N ₄	<i>h</i> -B ₂ N ₂	C ₃	N ₃	B ₃	<i>h</i> -BN
Sc	0.95	-0.28	0.20	0.79	-0.21	0.22	0.24	-0.25	-0.42	0.94	-0.69	-0.59	1.92
Ti	0.54	-0.53	-0.04	-0.10	-0.20	0.03	0.56	0.01	-0.61	0.62	-0.68	-0.42	-0.77
V	-0.06	-0.15	0.22	-0.08	0.21	0.10	-0.26	0.22	-0.30	0.03	-0.49	-0.44	-0.49
Cr	0.25	0.28	-0.35	0.18	0.31	0.30	-0.11	0.48	-0.07	0.24	-0.45	0.35	-0.11
Mn	0.39	0.39	0.40	0.40	-0.55	0.51	-0.41	0.42	-0.55	-0.16	0.20	0.09	0.20
Fe	0.25	0.28	-0.37	0.32	0.53	0.48	-0.13	0.47	-0.10	-0.02	0.05	0.87	-0.08
Co	-0.37	0.16	-0.13	0.06	-0.17	-0.36	-0.34	0.42	-0.52	-0.04	-0.01	0.38	0.64
Ni	-0.77	1.61	-0.23	-0.20	0.02	-0.61	-0.20	0.75	-0.78	0.45	-0.03	0.30	0.80
Cu	-0.44	1.72	0.90	1.29	-0.59	-1.30	0.48	0.47	-0.50	0.33	0.22	0.12	1.41
Zn	1.71	0.85	-0.10	1.41	-0.34	0.63	0.41	-0.40	0.88	0.06	-0.90	0.15	1.19
Y	1.00	-0.12	0.32	0.91	-0.24	-0.21	0.31	-0.19	-0.33	0.87	-0.42	-0.36	1.91
Zr	0.57	-0.90	-0.15	-0.06	-0.06	0.12	0.34	-0.44	-0.63	0.76	-0.79	-0.48	-0.94
Nb	-0.25	-0.89	-0.10	-0.68	0.06	-0.10	-0.63	-0.29	-0.30	-0.03	-0.84	-0.20	-0.88
Mo	-0.18	-0.44	0.17	-0.52	0.01	0.07	-0.71	-0.44	-0.58	0.21	-0.79	-0.82	-0.65
Tc	0.08	-0.46	0.25	-0.22	0.28	-0.06	-0.37	-0.43	-0.17	-0.12	-0.57	0.46	-0.61
Ru	0.16	-0.56	0.25	-0.02	0.19	-0.09	-0.15	-0.35	-0.33	0.12	-0.28	0.88	-0.55
Rh	-0.04	-0.26	0.05	-0.15	0.07	0.42	-0.19	0.54	-1.10	-0.03	0.25	0.81	0.63
Pd	-0.76	1.85	0.01	-0.37	-0.03	-0.10	0.03	0.44	-0.97	0.43	-0.21	0.65	0.07
Ag	-0.76	0.82	0.82	1.54	0.73	0.21	-0.86	0.66	-0.64	0.09	0.55	0.65	1.34
Cd	1.38	0.02	-0.27	0.85	0.12	0.68	-0.59	-0.09	1.39	-0.13	-0.59	0.46	0.26
Hf	0.30	-1.13	-0.54	-0.42	-0.23	-0.25	0.11	-0.66	-0.98	0.57	-0.99	-0.80	-1.05
Ta	-0.67	-1.08	-0.56	-1.07	-0.37	-0.44	-1.10	-0.55	-0.53	-0.52	-1.15	-0.80	-1.12
W	-0.67	-0.93	-0.26	-0.97	-0.44	-0.31	-1.01	-0.67	-1.02	-0.37	-1.21	0.48	-1.06
Re	-0.46	-0.80	-0.14	-0.68	-0.12	-0.42	-0.95	-0.77	-0.65	-0.68	-1.12	0.49	-1.09
Os	-0.20	-0.70	-0.09	-0.54	-0.07	-0.34	-0.85	-0.68	-0.91	-0.50	-0.77	0.55	-1.09
Ir	-0.39	-0.37	-0.22	-0.38	-0.12	0.00	-1.22	0.24	-1.24	-0.69	-1.23	0.40	0.03
Pt	-0.31	1.51	-0.22	-0.08	0.39	-0.02	-0.41	-0.42	-1.11	-0.05	-1.16	0.10	-0.64
Au	-0.49	2.16	0.44	1.48	-0.53	-0.72	-1.07	-0.94	-0.70	-0.61	-0.78	-0.19	-0.18

Table S10. Promising mono- and dual-type non-metal doped TM-SACs for high-performing HER reactions with better catalytic activity than commercial Pt/C, corresponding thermodynamic stability energies (E_{stab}), dissolution potential (U_{diss}), magnetization (μ^{*H}) with H-adsorbed state, angle at vertex at a TM atom (M) enclosed by two lines pointing to the adsorbed-H (H) and to the first nearest neighbor non-metal atom (Y: C, N, B) coordinating the TM atom on the 2D surface ($\angle H-M-Y$), d^{M-H} is the bond distance between TM and adsorbed hydrogen and corresponding HER activities are presented in terms of over potential (η^{HER}).

TM-SACs	E_{stab} (eV)	U_{diss} (V)	μ^{*H} (Bohr)	$\varphi^{(H-M-Y)}$ (deg.)	d^{M-H} (Å)	η^{HER} (V)
Rh@C ₄	-1.42	1.31	0.00	78.59	1.59	-0.04
Rh@B ₄	-0.56	0.88	0.00	79.64	1.62	-0.05
Pd@B ₄	-0.93	1.41	0.00	71.11	1.60	-0.01
Ti@N ₂ C ₂	-3.28	0.01	0.00	109.82	1.75	-0.10
Co@N ₂ C ₂	-2.25	0.85	0.00	89.95	1.43	-0.06
Ru@N ₂ C ₂	-1.05	0.98	0.00	80.07	1.62	-0.02
Pt@N ₂ C ₂	-2.73	2.55	0.00	33.44	1.83	-0.08
Rh@B ₂ C ₂	-0.07	0.63	0.00	69.91	1.60	-0.07
Pd@B ₂ C ₂	-0.53	1.21	0.00	30.81	1.85	-0.03
Pd@B ₂ N ₂	-0.71	2.07	0.00	36.76	1.89	-0.10
Pt@B ₂ N ₂	-0.61	2.52	0.00	37.66	1.88	-0.02
Fe@h-B ₂ N ₂	-2.28	0.69	1.00	38.19	1.62	-0.10
Fe@h-BN	-2.95	1.03	0.00	115.37	1.52	-0.08
Pd@h-BN	-1.83	1.86	0.00	90.74	1.59	-0.07
Ir@h-BN	-1.67	1.72	1.00	106.17	1.66	-0.03
Fe@C ₃	-1.65	0.38	0.56	120.68	1.57	-0.02
Co@C ₃	-1.92	0.68	0.00	120.59	1.56	-0.04
Rh@C ₃	-2.00	1.60	0.00	124.08	1.68	-0.03
Pt@C ₃	-1.41	1.88	0.85	106.34	1.69	-0.05
Pd@2DCP	-0.42	1.44	1.00	92.83	1.64	-0.03

Table S11 Top ranked SISSO generated descriptors used for ML regression for the prediction of HER activities.

Symbol	SISSO descriptors
S1	$\chi \frac{(\varepsilon^{ho} - \varepsilon^{lu})}{\varphi}$
S2	$\epsilon_{ea} v_{en} \frac{\bar{d}}{E_{cn}}$
S3	$Z \frac{(\varepsilon^{ho} - \varepsilon^{lu})}{v_{en}}$
S4	$r_{cov} + (\frac{E_{cn}}{v_{en}})$
S5	$\frac{r_{cov} \varphi}{\sin(\alpha)}$

Table S12. Elemental features from periodic table properties taken from the Mendeleeev python package,² DFT based electronic and geometric descriptors, Coulomb matrix elements and SISSO generated descriptors.

Feature symbol	Description
Z	Atomic number
r_{cov}	Covalent radius
χ	Pauling electronegativity
I	First ionization potential
r_v	Radius of the last occupied valence orbital
U_{en}	Valance electrons
θ_d	unpaired d-electrons
ϵ_{ea}	Electron affinity
α	Polarizability
E_{cn}	Effective coordination number
mp	Melting point of TM atom
$\bar{d}$	Average distance from TM to four nearest coordinating atoms ($\bar{d} = \frac{d_1+d_2+d_3+d_4}{4}$)
φ	Angle at vertex at a TM atom (M) enclosed by two lines pointing to the adsorbed-H (H) and to the first nearest neighbor non-metal atom (Y: B, C, N, P) coordinating the TM atom on the 2D surface ($\angle H-M-Y$).
ϵ^{ho}	Highest-occupied Kohn-Sham eigenvalue
ϵ^{lu}	Lowest-unoccupied Kohn-Sham eigenvalue
cm_1 to cm_9	Coulomb matrix elements based on DFT optimized geometries
s_1, s_2, s_3, s_4, s_5	Top five SISSO generated features
$cm_{(n1)}, cm_{(n2)}, cm_{(n3)},$ $cm_{(n4)}, cm_{(TM-H)}, cm_{(H-nX)}, X$ = B,C, N & P)	Coulomb matrix elements based on elemental properties

Table S13. Mean and standard deviation values of the Area Under the Receiver Operating Characteristic Curve (ROC AUC) for different ML classification models: AdaBoost, CatBoost, Bagging, LGBM-Light Gradient Boosting Machine, GB-Gradient Boosting, ERT-Extremely Randomized Trees, RF-Random Forest, DT-Decision Trees, SVC-Support Vector Classifier, KNN-k-Nearest Neighbors, RNC-Radius Neighbors Classifier, LR-Logistic Regression.

Classifier	E_{stab}		U_{diss}	
	ROC AUC mean	ROC AUC sd	ROC AUC mean	ROC AUC sd
AdaBoost	0.83	0.05	0.91	0.04
CatBoost	0.83	0.05	0.91	0.04
Bagging	0.87	0.04	0.92	0.04
LGBM	0.85	0.04	0.92	0.04
GB	0.86	0.05	0.93	0.03
ERT	0.87	0.05	0.93	0.03
RF	0.86	0.04	0.92	0.04
DT	0.78	0.06	0.86	0.04
SVC	0.85	0.05	0.92	0.03
KNN	0.84	0.06	0.92	0.04
RNC	0.83	0.05	0.89	0.05
LR	0.72	0.06	0.89	0.05

Table S14. ML predicted and DFT verified results of Hydrogen adsorption Gibbs free energies (eV) for TM atom embedded at single and double vacancy defects sites of mono- and dual-type non-metal doped configurations. (See Figure S15 & S18)

TM-SACs	ML-Predicted ΔG_{H^*} (eV)	DFT-Calculated ΔG_{H^*} (eV)
Sc@P₃	-0.19	-0.24
Fe@P₃	0.09	0.02
Zr@P₃	-0.11	-0.18
Au@P₃	-0.14	-0.06
Cr@P₄	-0.03	0.11
Mn@P₄	0.08	0.02
Fe@P₄	-0.05	0.01
Nb@P₄	-0.08	-0.04
Mo@P₄	0.15	0.07
Ru@P₄	-0.19	-0.34
Ir@P₄	-0.06	0.17

Supplementary References:

1. R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler and L. M. Ghiringhelli, *Phys. Rev. Mater*, 2018, **2**, 083802.
2. L. M. Mentel, Mendeleeev-A Python resource for properties of chemical elements, ions and isotopes, <https://github.com/lmmentel/mendeleeev>.