Supporting Information

Machine Learning Screening and High-Throughput Computation of 3d-Transition-Metal Intercalated Janus PtXY/ζ-Phosphorene (X ≠ Y; X, Y = S, Se, Te) Heterostructures for Photocatalytic Water Splitting

Kolangara V Sreesanth[1], Ramandeep Singh[1], Pooja Jamdagni[2] and Ashok Kumar[1]*

[1]*Department of Physics, Central University of Punjab, Bathinda, 151401, India*

[2]*Department of Computational Sciences, Central University of Punjab, Bathinda, 151401, India*

(November 28, 2025)

*Corresponding Author: ashokphy@cup.edu.in

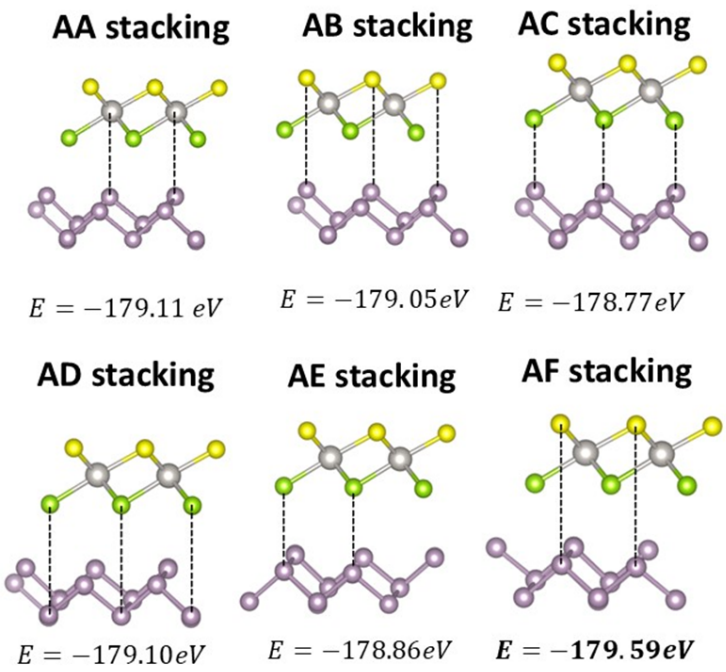

Figure S1. Different stackings of PtSSe/ζ-phosphorene (S-side) heterostructure with the total energy of each stacking pattern.

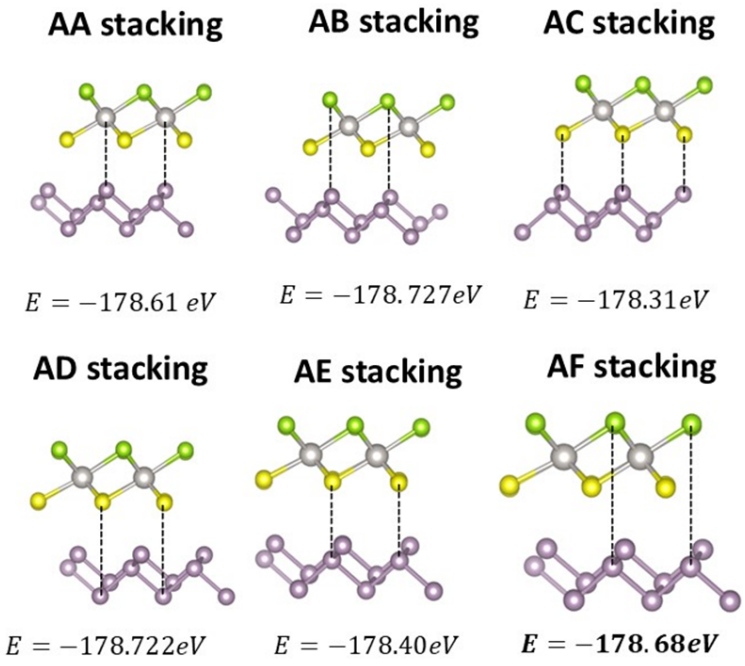

Figure S2. Different stackings of PtSSe/ζ -Phosphorene (Se-side) heterostructure with the total energy of each stacking pattern.

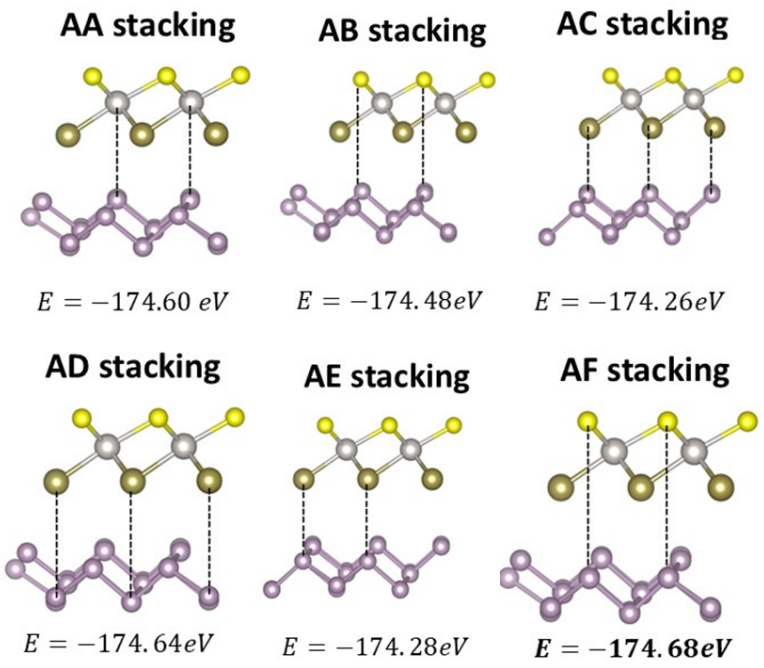

Figure S3. Different stackings of PtSTe/ζ -Phosphorene (Te-side) heterostructure with the total energy of each stacking pattern.

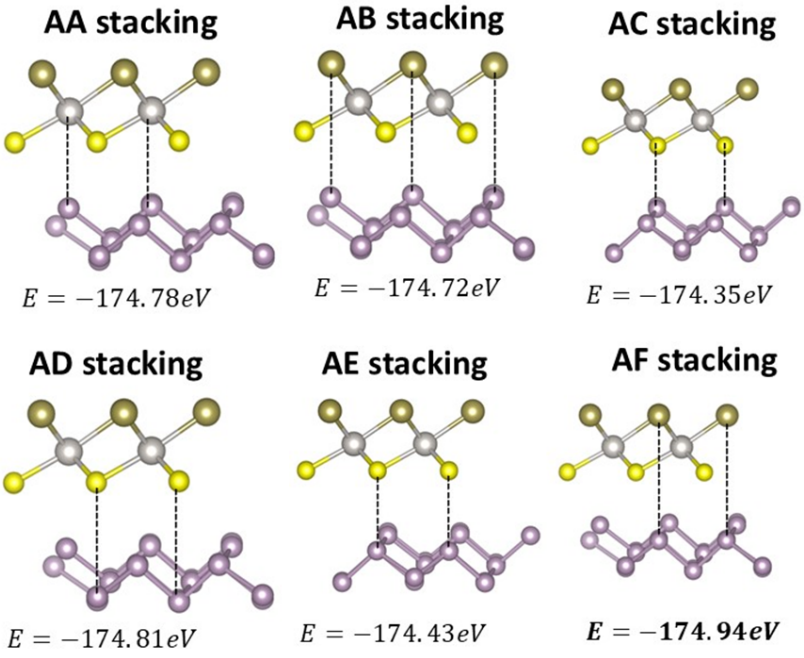

Figure S4. Different stackings of PtSTe/ζ-Phosphorene (S-side) heterostructure with the total energy of each stacking pattern.

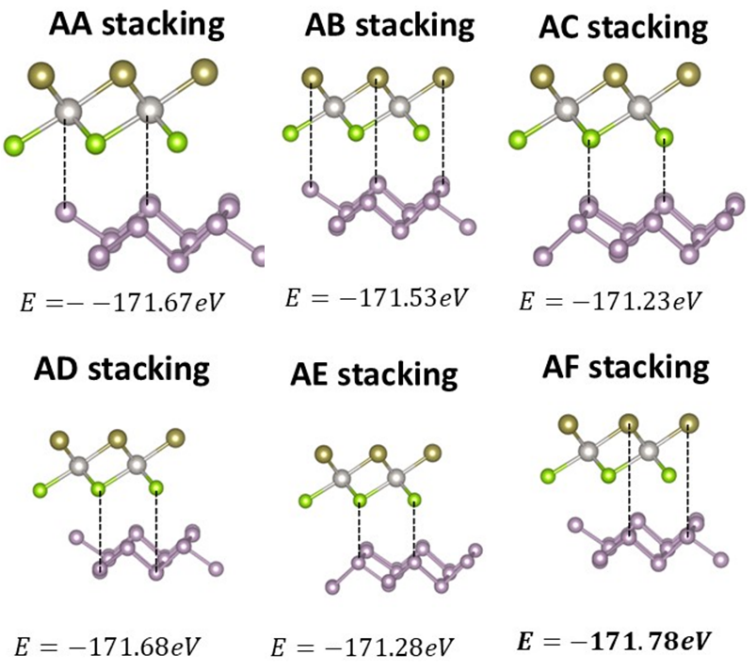

Figure S5. Different stackings of PtSeTe/ζ-Phosphorene (Se-side) heterostructure with the total energy of each stacking pattern.

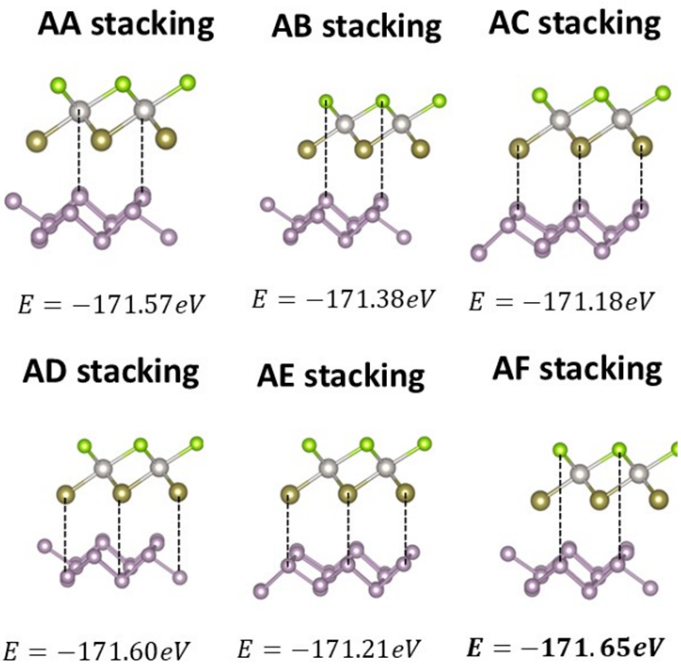

Figure S6. Different stackings of PtSeTe/ζ-Phosphorene (Te-side) heterostructure with the total energy of each stacking pattern.

Gibbs free energy calculations

The HER mechanism takes place in two steps:

$$* H^+ + e^- \rightarrow H^* \quad \text{S1}$$

$$H^* \rightarrow * + H_2 \quad \text{S2}$$

the Gibbs free energy for this reaction can be calculated using the following equation:

$$\Delta G = \Delta E + \Delta E_{ZEP} - T\Delta S \quad \text{S3}$$

Here, T=298K. ΔE_{ZEP} and ΔS are the zero-point energy and entropy given below:

Table S1. Standard values of TxS and zero-point energy

Species	TxS (eV)	ZPE (eV)
H*	0	0.17
H_2	0.41	0.27
O*	0	0.07
OH*	0	0.33
OOH*	0	0.43
H_2O	0.58	0.57

OER mechanism, on the other hand, takes a four-step pathway.

$$\Delta G_1 = G_{OH*} + \frac{1}{2}G_{H_2} - G^* - G_{H_2O} \quad \text{S4}$$

$$\Delta G_2 = G_{O*} + \frac{1}{2}G_{H_2} - G_{OH} \quad \text{S5}$$

$$\Delta G_3 = G_{OOH*} + \frac{1}{2}G_{H_2} - G_{O*} - G_{H_2O} \quad \text{S6}$$

$$\Delta G_4 = G^* + \frac{1}{2}G_{H_2} + G_{O_2} - G_{OOH*} \quad \text{S7}$$

The rate-limiting step is the reaction step with the maximum energy barrier, and the corresponding Gibbs free energy change is the rate-limiting Gibbs free energy.

Representative symbol	Machine learning feature
X_S, X_Se	Chalcogen atom X (X_atom)
Y_Se, Y_Te	Chalcogen atom Y (Y_atom)
X_S, X_Se, Y_Se, Y_Te	The chalcogen atom is facing the phosphorene layer (side)
TM_Sc, TM_Ti, TM_V, TM_Cr, TM_Mn, TM_Fe, TM_Co, TM_Ni, TM_Cu, TM_Zn	Intercalating atom (I_atom)
I_1, I_2	Intercalation site (I_site)
A_1, A_2	Adsorption site (A_site)
N_TM	Atomic number of a transition metal
G_3, G_4, G_5, G_6, G_7, G_8, G_9, G_10, G_11, G_12	Group number of the Transition metal
IE_TM	Ionisation energy of transition metal
rm_TM	Atomic radius of a transition metal
EA_TM	Electron affinity
W_TM	Atomic mass
MP_TM	Melting point of the Transition metal
BP_TM	The boiling point of the transition metal
EN_TM	Electronegativity of transition metals
D_layer	Interlayer distance of the heterostructure
L_Pt-X	Bond length between Pt and X atom
L_Pt-Y	Bond length between Pt and Y atom
E_B	Binding energy of the intercalated heterostructure.

Table S2. Machine learning features and their representative symbols.

Model name	Model type
LinearRegressior	Linear
SGDRegressor	Linear
ElasticNet	Linear
Ridge	Linear
Lasso	Linear
QuantileRegressor	Linear
TweedieRegressor	Linear
PassiveAggressiveRegressor	Linear
AdaBoostRegressor	Ensemble
BaggingRegressor	Ensemble
ExtraTreesRegressor	Ensemble
GradientBoostingRegressor	Ensemble
RandomForestRegressor	Ensemble
DecisionTreeRegressor	Tree
ExtraTreeRegressor	Tree
SVR (kernel = 'rbf')	Support vector machine
SVR (kernel = 'poly')	Support vector machine
SVR (kernel = 'linear')	Support vector machine
SVR (kernel = 'sigmoid')	Support vector machine
KNeighborsRegressor	Neighbour
MLPRegressor	Artificial neural network

Table S3. Various Machine learning models used in the study.

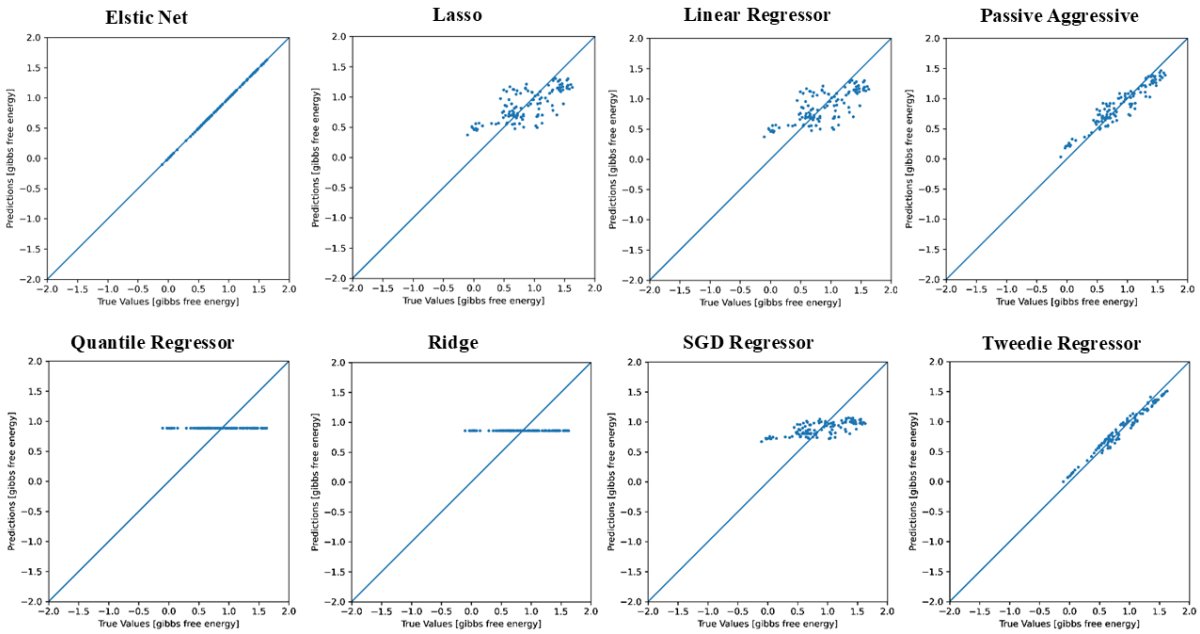

Figure S7. Plots of calculated values vs predicted values of ΔG for HER using Linear models.

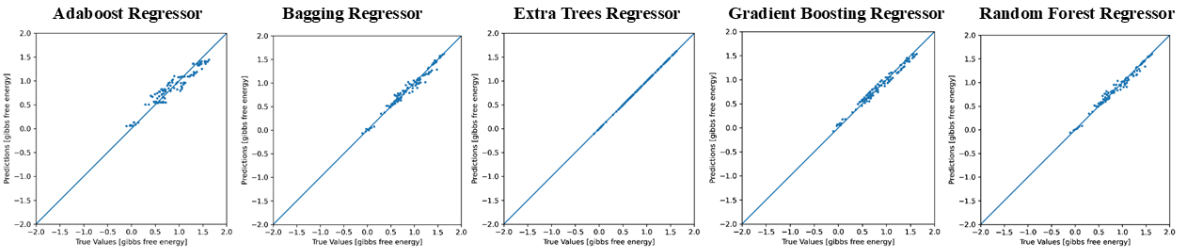

Figure S8. Plots of calculated values vs predicted values of ΔG for HER using Ensemble Model.

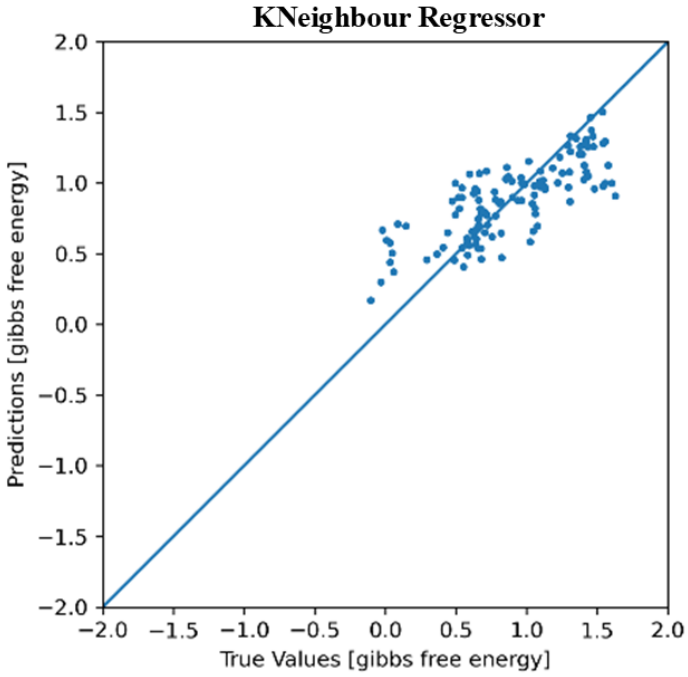

Figure S9. Plots of calculated values vs predicted values of ΔG for HER Neighbour model.

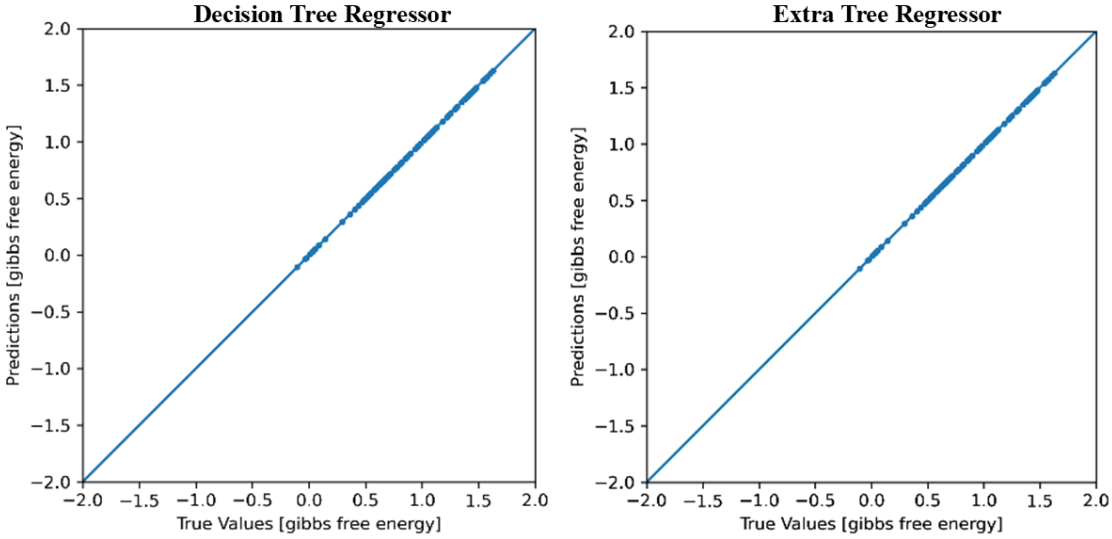

Figure S10. Plots of calculated values vs predicted values of ΔG for HER Tree Models.

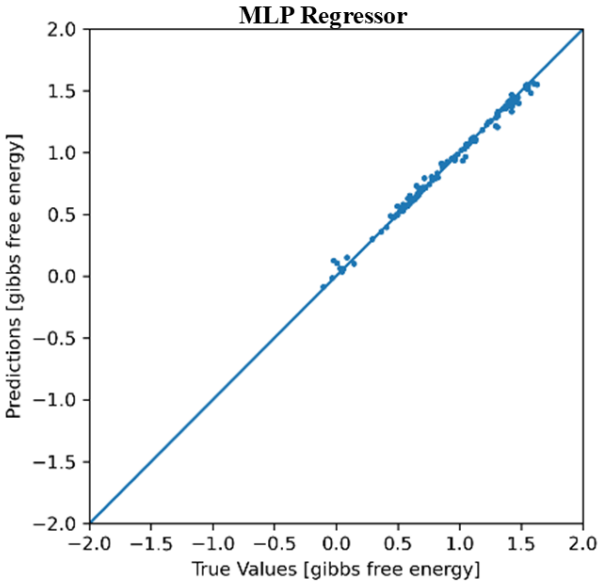

Figure S11. Plots of calculated values vs predicted values of ΔG for HER using Artificial Neural Network.

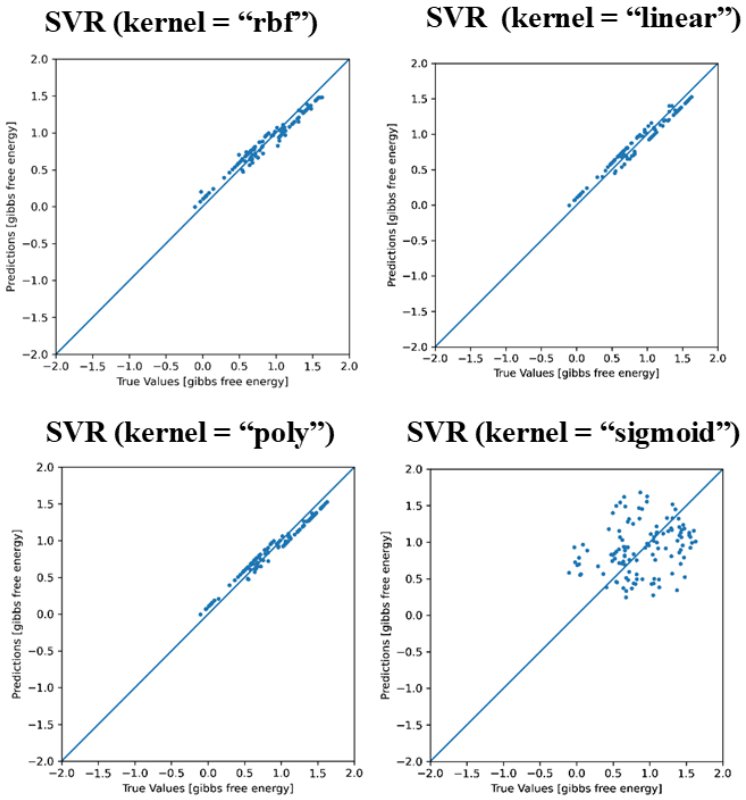

Figure S12. Plots of calculated values vs predicted values of ΔG for HER using support vector machines with different kernels.

Table S4. K-fold cross-validation performance of different models for ΔG values for HER.

Model name	MAE	R^2
Linear Regression	-0.294	0.22
SGD Regressor	-0.296	0.284
Elastic Net	-0.365	-0.054
Ridge	-0.291	0.222
Lasso	-0.362	-0.045
Quantile Regressor	-0.364	-0.137
Tweedie Regressor	-0.319	0.167
Passive Aggressive Regressor	-0.291	0.225
AdaBoost Regressor	-0.147	0.771
Bagging Regressor	-0.12	0.823
Extra Trees Regressor	-0.108	0.817
Gradient Boosting Regressor	-0.119	0.803
Random Forest Regressor	-0.112	0.832
Decision Tree Regressor	-0.122	0.696
ExtraTreeRegressor	-0.137	0.651
SVR (kernel = 'rbf')	-0.202	0.566
SVR (kernel = 'poly')	-0.193	0.625
SVR (kernel = 'linear')	-0.286	0.134
SVR (kernel = 'sigmoid')	-0.407	-0.552
KNeighbour Regressor	-0.236	0.394
MLPRegressor	-0.272	0.318

Table S5. Train-test verification metrics of Ensemble models.				
Model	**Train R^2**	**Train MAE**	**TEST R^2**	**TEST MAE**
AdaBoost Regressor	0.903007	0.10886	0.883639	0.110439
Bagging Regressor	0.97316	0.044472	0.878026	0.10058
Extra Trees Regressor	1	1.02E-15	0.845771	0.104612
Gradient Boosting Regressor	0.982644	0.050047	0.886208	0.094394
RandomForestRegressor	0.97437	0.047516	0.907746	0.088006

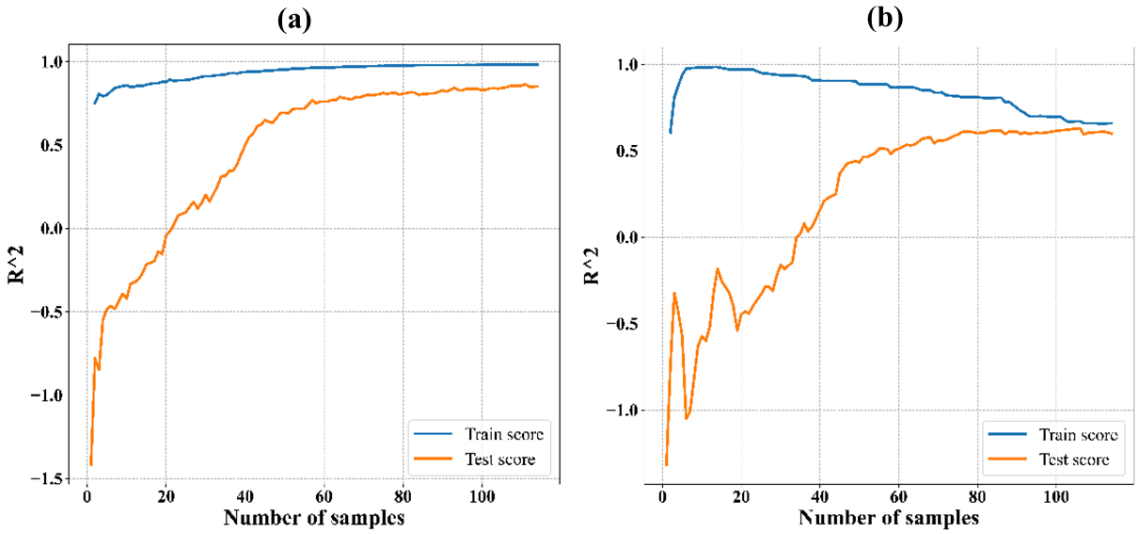

Figure S13. Learning curve of (a) Random Forest regressor for HER and (b) Support Vector Regressor for OER.

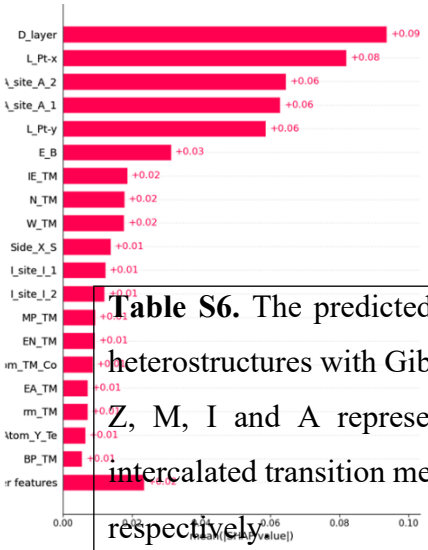

Table S6. The predicted adsorbed Janus PtXY/ζ-phosphorene (Z, M, I, A) heterostructures with Gibbs free energy between -0.3 eV and 0.3 eV for HER. Z, M, I and A represent the chalcogen side of the heterostructure, the intercalated transition metal atom, the intercalated site and the adsorption site, respectively.

Heterostructures	Gibbs free energy (eV)	Binding Energy (eV)

Figure S14. Bar diagram of the average SHAP value of each of the features.

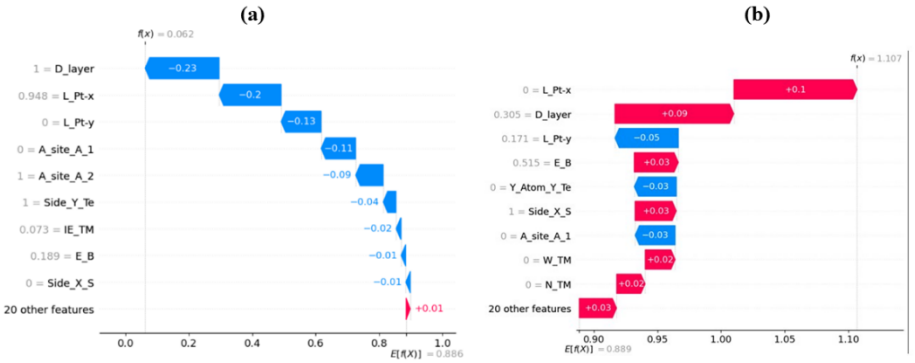

Figure S15. Waterfall plot of SHAP values corresponding to **(a)** promising catalysts and **(b)** a poor catalyst.

PtSTe/zeta-P(Te, Sc,I_1, A_2)	0.261034	-2.8
PtSTe/zeta-P(Te,Ti,I_1,A_2)	0.160839	-4.96
PtSTe/zeta-P(Te,V,I_1,A_2)	0.093532	-5.94
PtSTe/zeta-P(Te,Cr,I_1,A_2)	0.061617	-6.3
PtSTe/zeta-P(Te, Mn,I_1, A_2)	0.059212	-6.23
PtSTe/zeta-P(Te,Fe,I_1,A_2)	-0.00053	-5.89
PtSTe/zeta-P(Te,Co,I_1,A_2)	-0.0627	-5.42
PtSTe/zeta-P(Te,Ni,I_1,A_2)	0.102895	-4.7
PtSTe/zeta-P(Te,Cu,I_1,A_2)	0.114283	-2.52
PtSTe/zeta-P(Te,Zn,I_1,A_2)	0.109481	0.76
PtSTe/zeta-P(Te,Sc,I_2,A_2)	0.244293	-4.48
PtSTe/zeta-P(Te,Ti,I_2,A_2)	0.105447	-5.98
PtSTe/zeta-P(Te,V,I_2,A_2)	0.086583	-6.41
PtSTe/zeta-P(Te,Cr,I_2,A_2)	0.05083	-6.68
PtSTe/zeta-P(Te,Mn,I_2,A_2)	0.033726	-6.79
PtSTe/zeta-P(Te,Fe,I_2,A_2)	0.003511	-6.61
PtSTe/zeta-P(Te,Co,I_2,A_2)	-0.05505	-6.2
PtSTe/zeta-P(Te,Ni,I_2,A_2)	0.032505	-5.58
PtSTe/zeta-P(Te,Cu,I_2,A_2)	0.089473	-3.68
PtSTe/zeta-P(Te,Zn,I_2,A_2)	0.10111	-0.83

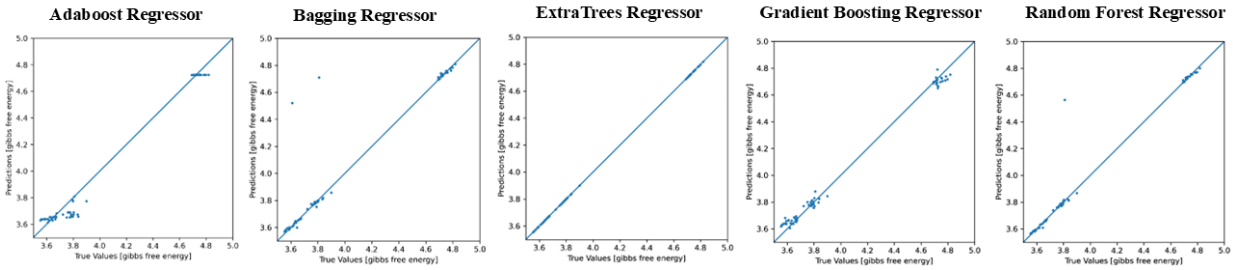

Figure S16. Plots of calculated values vs predicted values of limiting Gibb's free energy, ΔG, for OER using ensemble models.

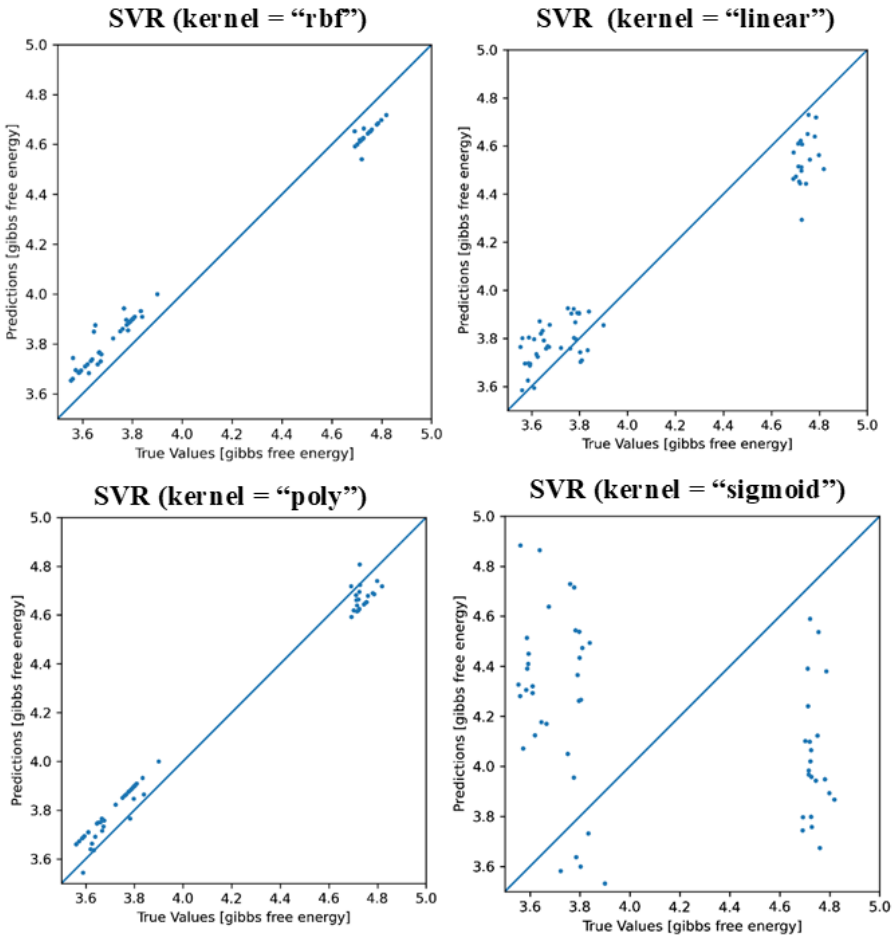

Figure S17. Plots of calculated values vs predicted values of limiting Gibb's free energy, ΔG for OER using support vector regression with different kernels.

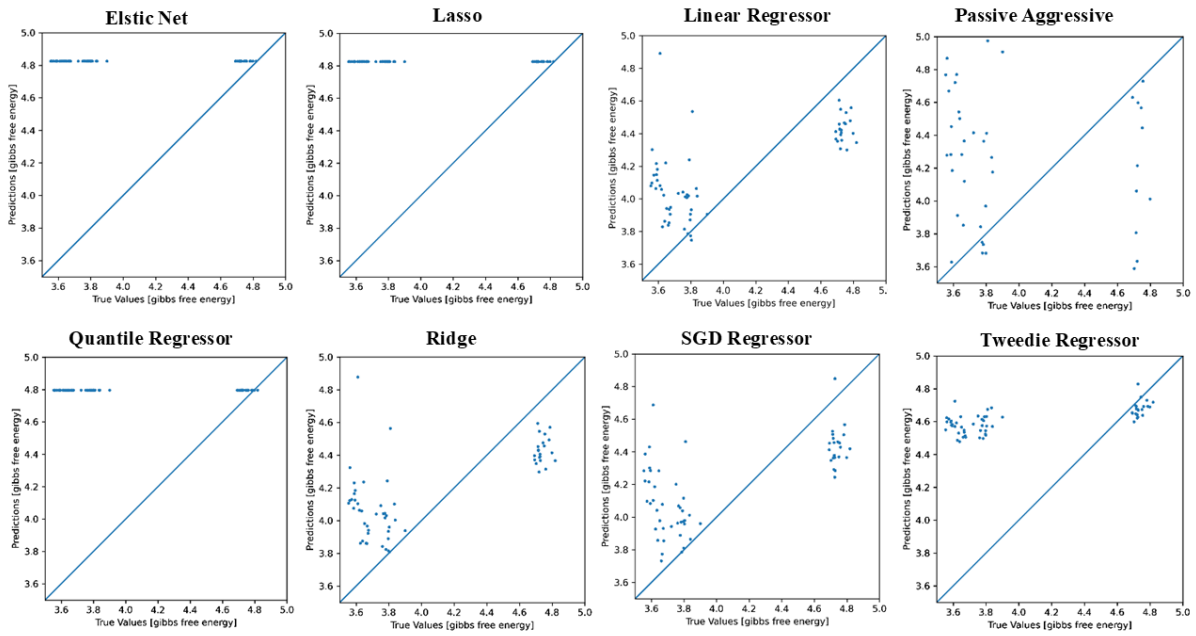

Figure S18. Plots of calculated values vs predicted values of limiting Gibb's free energy, ΔG, for OER using linear models.

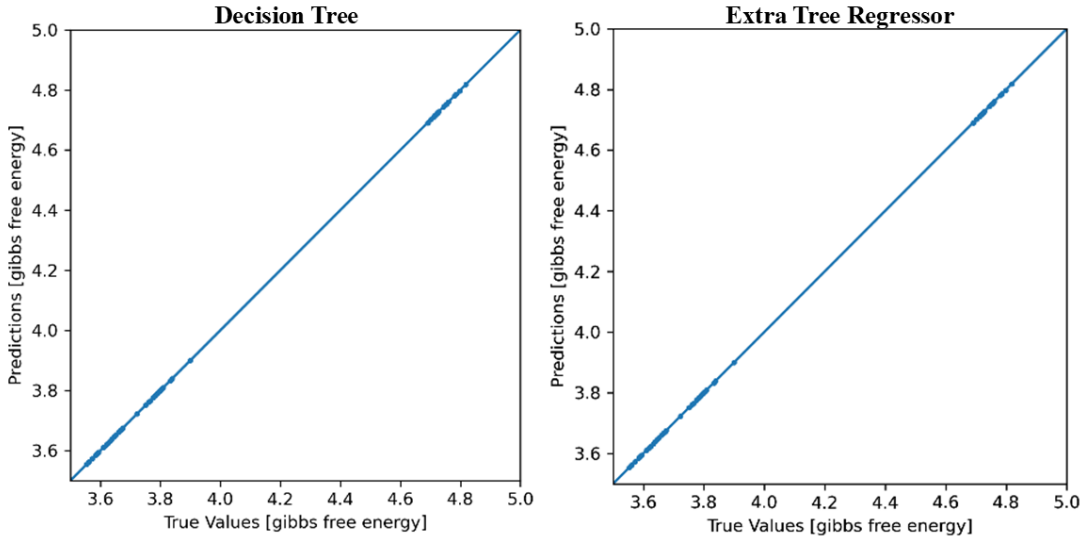

Figure S19. Plots of calculated values vs predicted values of limiting Gibb's free energy, ΔG, for OER using tree-based models.

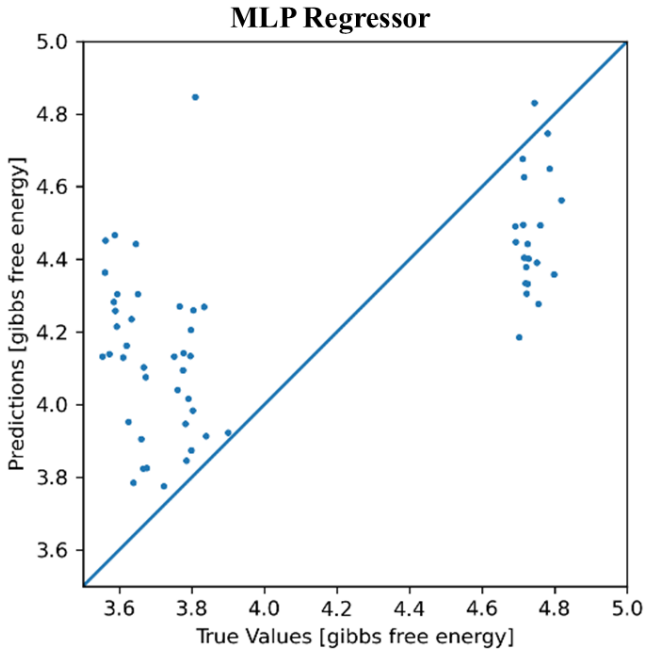

Figure S20. Plots of calculated values vs predicted values of limiting Gibb's free energy, ΔG for OER using Artificial Neural Networks.

Table S7. K-fold cross-validation performance of different models for limited Gibb's free energy, ΔG, values for OER.

Model name	MAE	R^2
Linear Regression	-0.48	0.45
SGD Regressor	-0.46	0.51
Elastic Net	-0.79	-0.70
Ridge	-0.46	0.49
Lasso	-0.79	-0.36
Quantile Regressor	-0.81	-0.52
Tweedie Regressor	-0.61	0.16
Passive Aggressive Regressor	-0.62	-0.61
AdaBoost Regressor	-2.99	-1.74
Bagging Regressor	-0.13	-1.11
Extra Trees Regressor	-0.266	-0.41
Gradient Boosting Regressor	-0.34	-1.10
Random Forest Regressor	-0.24	0.16
Decision Tree Regressor	-0.26	-0.90
ExtraTreeRegressor	-0.26	-2.59
SVR (kernel = 'rbf')	-0.26	0.81
SVR (kernel = 'poly')	-0.23	0.78
SVR (kernel = 'linear')	-0.35	-0.68
SVR (kernel = 'sigmoid')	-0.72	-0.40
MLPRegressor	-0.52	0.17

Table S8. Train-test verification metrics of SVR models.

Model	Train R^2	Train MAE	TEST R^2	TEST MAE
SVR (kernel = 'rbf')	0.620255	0.159876	0.744394	0.271433
SVR (kernel='poly')	0.745499	0.14974	0.358297	0.417327
SVR (kernel='linear')	0.28762	0.327971	0.861954	0.271147
SVR (kernel='sigmoid')	-29.0059	5.119372	-34.5789	4.094686

Table S9. The top 10 predicted adsorbed Janus PtXY/ζ-phosphorene (Z, M, I, A) heterostructures with rate-limiting Gibbs free energy for HER. Z, M, I, and A represent the chalcogen side of the heterostructure, the intercalated transition metal atom, the intercalated site, and the adsorption site, respectively.

Heterostructure	Rate-limiting Gibbs free energy change (eV)	Binding energy (eV)
PtSTe/zeta-P(S,Sc,I_1,A_2)	3.660455	-2.39
PtSTe/zeta-P(S,V,I_1,A_2)	3.685306	-6.13
PtSTe/zeta-P(S,Co,I_1,A_2)	3.687182	-5.66
PtSTe/zeta-P(S,Ni,I_1,A_2)	3.607957	-4.79
PtSTe/zeta-P(S,Cu,I_1,A_2)	3.66154	-2.58
PtSTe/zeta-P(S,Sc,I_2,A_2)	3.673565	-6.76
PtSTe/zeta-P(S,Fe,I_2,A_2)	3.662466	-6.83
PtSTe/zeta-P(S,Co,I_2,A_2)	3.566851	-6.33
PtSTe/zeta-P(S,Ni,I_2,A_2)	3.564089	-5.62
PtSTe/zeta-P(S,Cu,I_2,A_2)	3.67234	-3.92

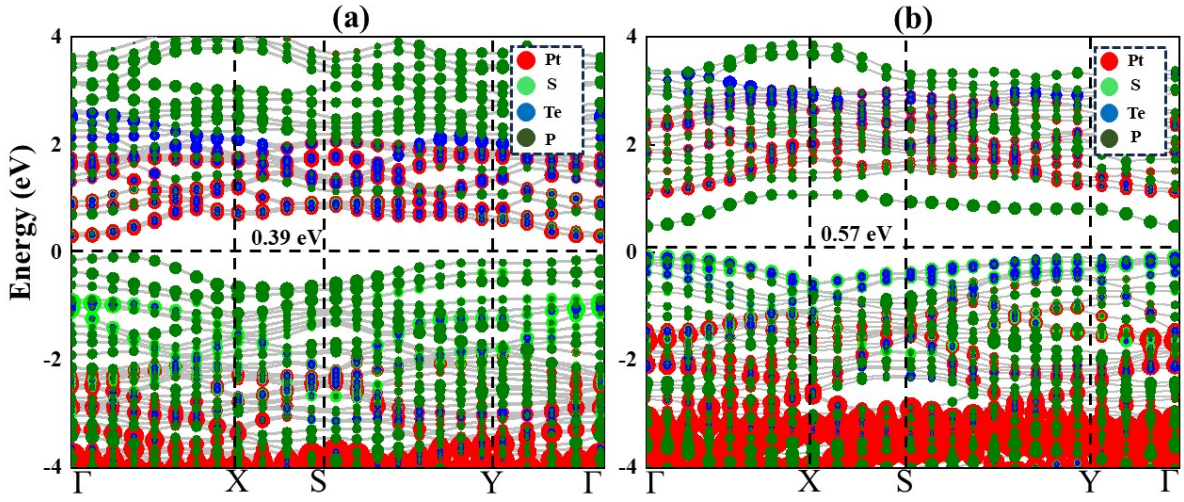

Figure S21. Electronic band structure of **(a)** S- and **(b)** Te-side PtSTe/ζ-P heterostructure using HSE06 functional.

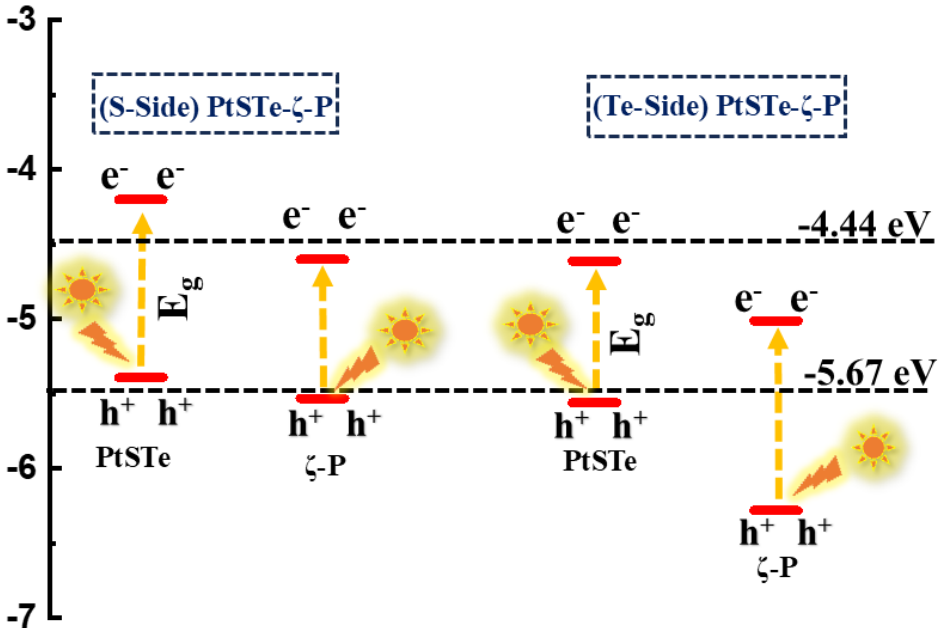

Figure S22. HSE06 band alignment of S and Te-side PtSTe/ζ-P heterostructure for photocatalytic water splitting.

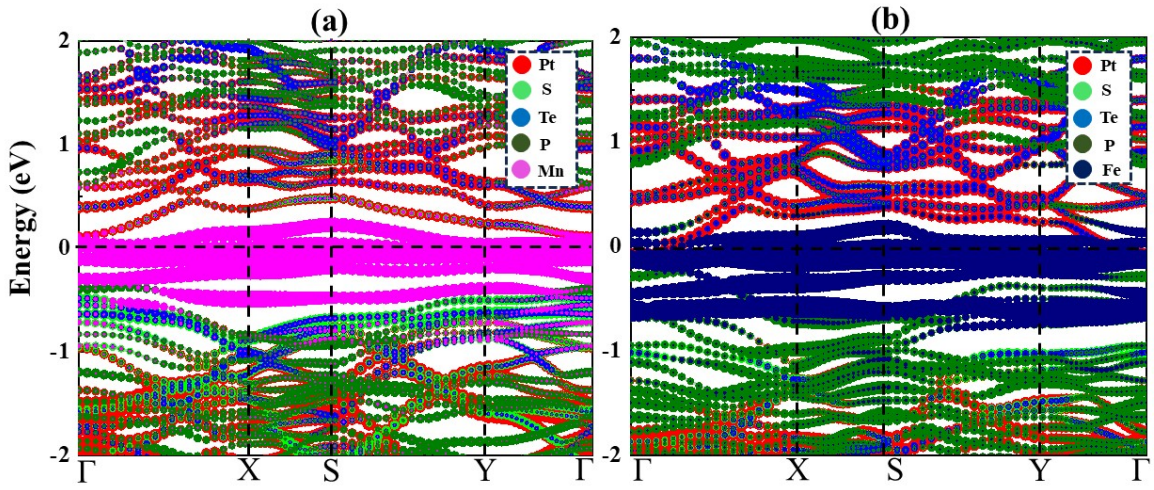

Figure S23. Electronic band structure of **(a)** Mn intercalated (site-2) Te-side and **(b)** Fe intercalated (site-2) S side PtSTe/ζ-P heterostructure.