Supplementary Materials

Machine learning prediction of small molecule passivators and their impacts on the passivation and photocatalytic performance of organic-inorganic hybrid perovskite interfaces

Yan Cai[1], Zhentao Bai[2], Changcheng Chen[1,*], Minghong Sun[3], Zhengjun Wang[1], Songya Wang[1], Ziyi Zhang[1], Jiangzhou Xie[4], Dongbo Li[1,*], Xiaoning Guan[5], Gang Liu[5], Pengfei Lu[5], Sining Yun[6,*]

[1]School of Science, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China.
[2]School of Civil Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China.
[3]School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 610054, Sichuan, China.
[4]School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney, New South Wales 2052, Australia.
[5]State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China.
[6]School of Materials Science and Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China.

Correspondence to: Dr. Changcheng Chen, School of Science, Xi'an University of Architecture and Technology, No. 13, Yanta Road, Beilin District, Xi'an, Xi'an 710055, Shaanxi, China. E-mail: chenchangcheng@xauat.edu.cn; Dr. Dongbo Li, School of Science, Xi'an University of Architecture and Technology, No. 13, Yanta Road, Beilin District, Xi'an, Xi'an 710055, Shaanxi, China. E-mail: ldb@xauat.edu.cn; Dr. Sining Yun, School of Materials Science and Engineering, Xi'an University of Architecture and Technology, No. 13, Yanta Road, Beilin District, Xi'an, Xi'an, Shaanxi 710055, China. E-mail: yunsining@xauat.edu.cn

Part I. Calculation Process Details

1.1 Machine Learning Details

All machine learning scripts are written in Python. In order to avoid overfitting and underfitting of the ML model caused by the introduction of statistical redundant features, advanced feature engineering techniques were adopted to select important features that

are independent of each other and strongly related to the target attribute. The Pearson correlation coefficient (*P*) was used to screen for statistically independent features, as shown in equation (1) [1,2].

$$P(x_1, x_2) = \frac{cov(x_1,x_2)}{\sigma_{x1}\sigma_{x2}} = \frac{\sum_{i=1}^{n}[(x_i-\bar{x})\times(y_i-\bar{y})]}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\times\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (1)$$

Among them, x_1 and x_2 represent two different feature vectors in the data, $cov(x_1, x_2)$ represents covariance operation, and σ_1 and σ_2 are the standard deviations of the set samples. It evaluates the strength of the relationship between two features based on the covariance matrix of the data $|P|$ represents the linear correlation between two feature vectors. When $|P|>0.8$, it indicates a strong correlation between the two features. In this case, we can choose to. For a single representative feature, discard the other to reduce redundancy.

To obtain a uniformly distributed input and output model, the *z*-score method is used to normalize the data. The normalization formula is $z = \frac{x-\bar{\mu}}{\sigma}$, where z represents the normalization value, x comes from the original dataset, $\bar{\mu}$ and σ represent the original values used for a certain dimension feature, respectively. The above learning model is trained using normalization, and the standard for machine learning training is quantified by the determination coefficient R^2[2]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\mu_i-\hat{\mu})^2}{\sum_{i=1}^{n}(\mu_i-\bar{\mu})^2} \quad (2)$$

Where μ_i and $\bar{\mu}$ respectively represent the *i*th sample in the actual and predicted output training set. The closer R^2 is to 1, the closer the prediction is to the actual value. The R^2 was evaluated using the Detection Point Cross Validation (HOCV) method based on a training ratio of 0.8, and to eliminate randomness, each model was repeated 20 times. We use the general machine learning library scikit learn to construct the aforementioned ML models, including support vector machine (SVM), random forest (RF), *k*-nearest neighbor (KNN), gradient enhanced regression (GBR), eXtreme Gradient Boosting(XGB) and multi-layer perceptron (MLP) models[3]. The purpose of this article is to search for small molecules and their derivatives of passivating agents with better performance, and to study the effect of passivating

agent molecules on material photocatalysis. In addition to single model prediction, stacking several models can also be used to improve accuracy on the test set, but it cannot guarantee overfitting of data leakage. The construction of the database was based on the selection of small molecule passivators from reported literature, while other information was obtained from the PubChem and ChemSpider datasets and VASP calculations.

Accuracy refers to the proportion of correct classification results obtained by the classifier when dividing the dataset. In other words, it measures the model's ability to correctly classify samples. The following formula is usually used to calculate:

$$\text{Accuracy} = \frac{TN+TP}{TN+TP+FP+FN} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4}$$

Among them, *TP* represents true positive (the model correctly classified positive samples as positive), *TN* represents true negative (the model correctly classified negative samples as negative), *FP* represents false positive (the model incorrectly classified negative samples as positive), and *FN* represents false negative (the model incorrectly classified positive samples as negative). The accuracy value range is between 0 and 1, and the closer the value is to 1, the better the classification ability of the model.

1.2 Molecular Dipole Calculation

Molecular polarity calculations were performed using Gaussian 16 software, and all calculations were performed using the B2PLYP function in DFT functionals for geometric optimization calculations. The 6-311+G (d) basis set was used, and DFT-D3 dispersion correction was used to correct weak interactions, ensuring accurate calculations[4–6].

1.3 DFT calculation method

In the calculation, each pseudopotential atom has the following electronic valence state:H $1s^1$, C $2s^22p^2$, N $2s^22p^3$, Br $4s^24p^5$, Pb $5d^{10}6s^26p^2$, I $5s^25p^5$ and S $3s^2p^4$. In order

to better describe the stability of adsorption of different small molecules, evaluation is made by calculating adsorption energy[7]:

$$E_b = E_{slab/org} - E_{slab} - E_{org} \quad (5)$$

Among them, $E_{slab/org}$ refers to the total molecular energy after adsorbing a small molecule passivator, E_{slab} refers to the original energy of the intrinsic perovskite, E_{org} refers to the molecular energy of the passivator.

Calculate work function: The work function is a key parameter used as a reference for charge transfer and band edge alignment, and its formula is:

$$\phi = E_{vac} - E_f \quad (6)$$

Among them, Evac represents the energy of a stationary electron in a vacuum near the surface, and Ef is the Fermi level of the material.

Calculation of effective mass: Effective mass is an important property of carrier migration. The effective mass of the charge carriers could be found from equation (8) based on the energy band structure:

$$E(k) = E_0 + \frac{\hbar^2 k^2}{2m*} \quad (7)$$

Here $m*$ refers to the effective mass of charge carriers, k is the value of the wave vector in momentum space, E_0 is the energy of the band edge, $\hbar$ is the reduced Planck constant, $E(k)$ is the eigenvalue of the band edge around VBM and CBM, then $m*$ can be[8]:

$$m^* = \hbar^2 \left(\frac{\partial^2 E(k)}{\partial k^2}\right)^{-1} \quad (8)$$

Therefore, the effective mass tensor can be calculated by approximating the second derivative of the wave vector k at the valence band edge and conduction band edge relative to points G to X and M.

Calculate exciton binding energy: The exciton binding energy E_d is estimated using semiclassical Mott–Wannier theory[9] and the formula is as follows:

$$E_d = \frac{\mu e^4}{2\hbar^2 \varepsilon^2} \quad (9)$$

where ε is the planar high-frequency dielectric constant, $\mu = m_e \times m_h/(m_e + m_h)$ is the reduced effective mass and e is the charge of electron. Here, m_e and m_h are adopted as the average value.

Calculation of *p*-band centre. The center of Pb band is calculated by using the weighted average energy of projection density. The relative state of the *p*-state (both occupied and unoccupied states) of the B-site metal cation reaches the Fermi level.

Photocatalytic CO_2RR calculation method

According to the hydrogen electrode model, the calculation method of Gibbs degrees of freedom obtains the energy change in the reduction path[10]:

$$\Delta G = \Delta E_H + \Delta E_{ZPE} - T\Delta S \quad (10)$$

Among the ΔE_H refers to the direct change in electron energy, ΔE_{ZPE} represents zero vibration energy correction, T is the temperature (300K), ΔS represents entropy change. The contribution of ZPE and entropy is determined by calculating the vibration frequency. In frequency calculation, only the adsorbate was explicitly calculated and the substrate was fixed. $\Delta E = E_{surf+H} - E_{surf} - \frac{1}{2}E_{H_2}$, where E_{surf+H} represents the total energy of the catalyst after adsorption of H, E_{surf} is the original plate energy, E_{H_2} is the energy of H[11]. These values were calculated using specific modules in the software Vaspkit. Based on the previous reaction mechanism for the conversion of CO_2 to CH_4, it is as follows:

$$* + CO_2 + H^+ + e^- \rightarrow HCOO* \quad (1)$$

$$HCOO* + H^+ + e^- \rightarrow HCO*OH \quad (2)$$

$$HCO*OH + e^- \rightarrow CO* + H_2O \quad (3)$$

$$CO* + H^+ + e^- \rightarrow HCO* \quad (4)$$

$$HCO* + H^+ + e^- \rightarrow HC*OH \quad (5)$$

$$HC*OH + H^+ + e^- \rightarrow C*H + H_2O \quad (6)$$

$$C*H + H^+ + e^- \rightarrow C*H_2 \quad (7)$$

$$C*H_2 + H^+ + e^- \rightarrow C*H_3 \quad (8)$$

$$C*H_3 + H^+ + e^- \rightarrow C*H_4 \quad (9)$$

The reaction mechanism for converting CO_2 to CH_3OH is as follows:

$$* + CO_2 + H^+ + e^- \rightarrow HCOO* \quad (1)$$

$HCOO^* + H^+ + e^- \rightarrow HCO^*OH$ (2)

$HCO^*OH + e^- \rightarrow CO^* + H_2O$ (3)

$CO^* + H^+ + e^- \rightarrow HCO^*$ (4)

$HCO^* + H^+ + e^- \rightarrow O^*HCH$ (5)

$O^*HCH + H^+ + e^- \rightarrow OH^*CH_2$ (6)

$OH^*CH_2 + H^+ + e^- \rightarrow OH^*CH_3$ (7)

The reaction mechanism for converting CO_2 to CH_2O is as follows:

$^* + CO_2 + H^+ + e^- \rightarrow HCOO^*$ (1)

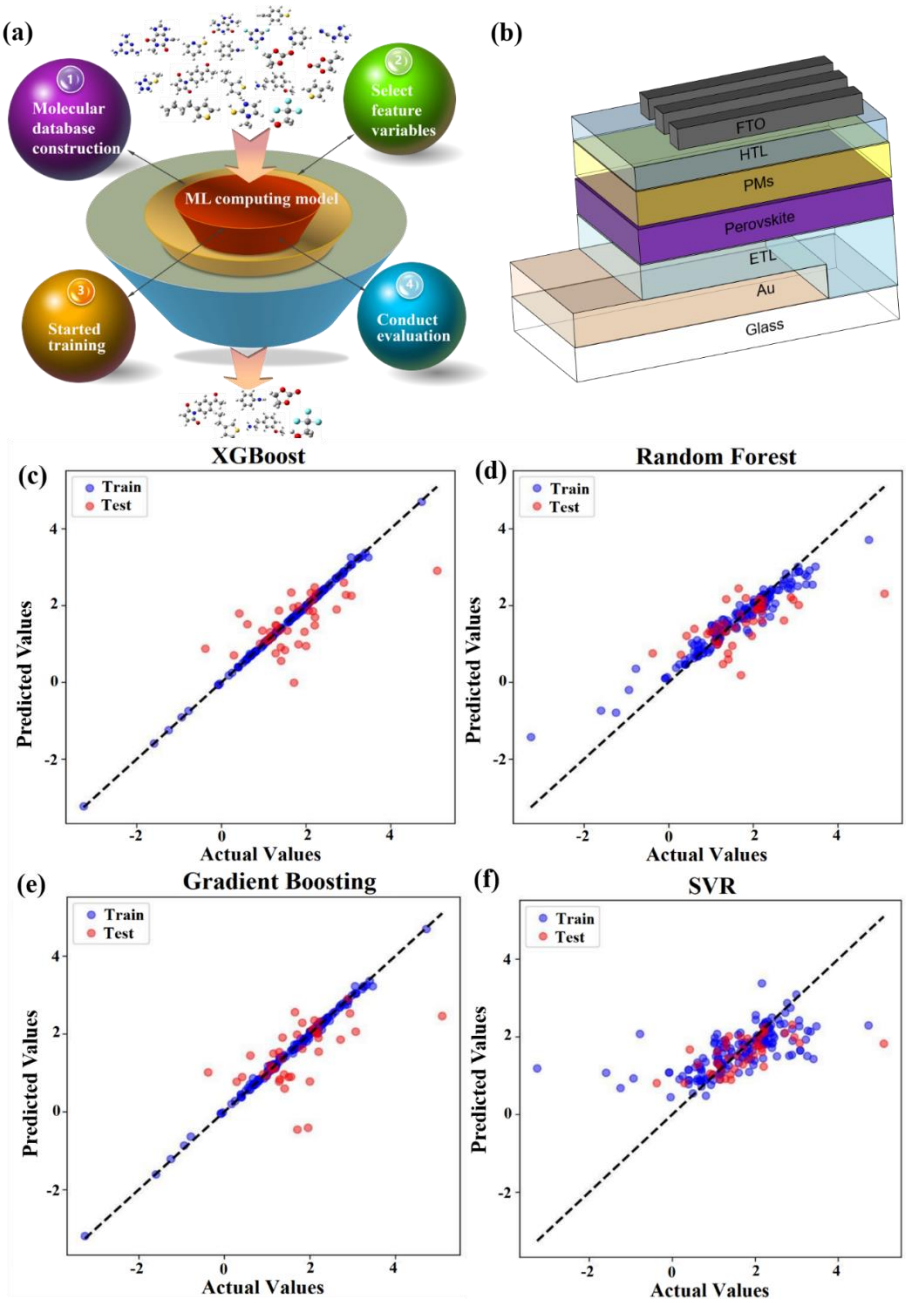
$HCOO^* + H^+ + e^- \rightarrow HCO^*OH$ (2)

$HCO^*OH + e^- \rightarrow CO^* + H_2O$ (3)

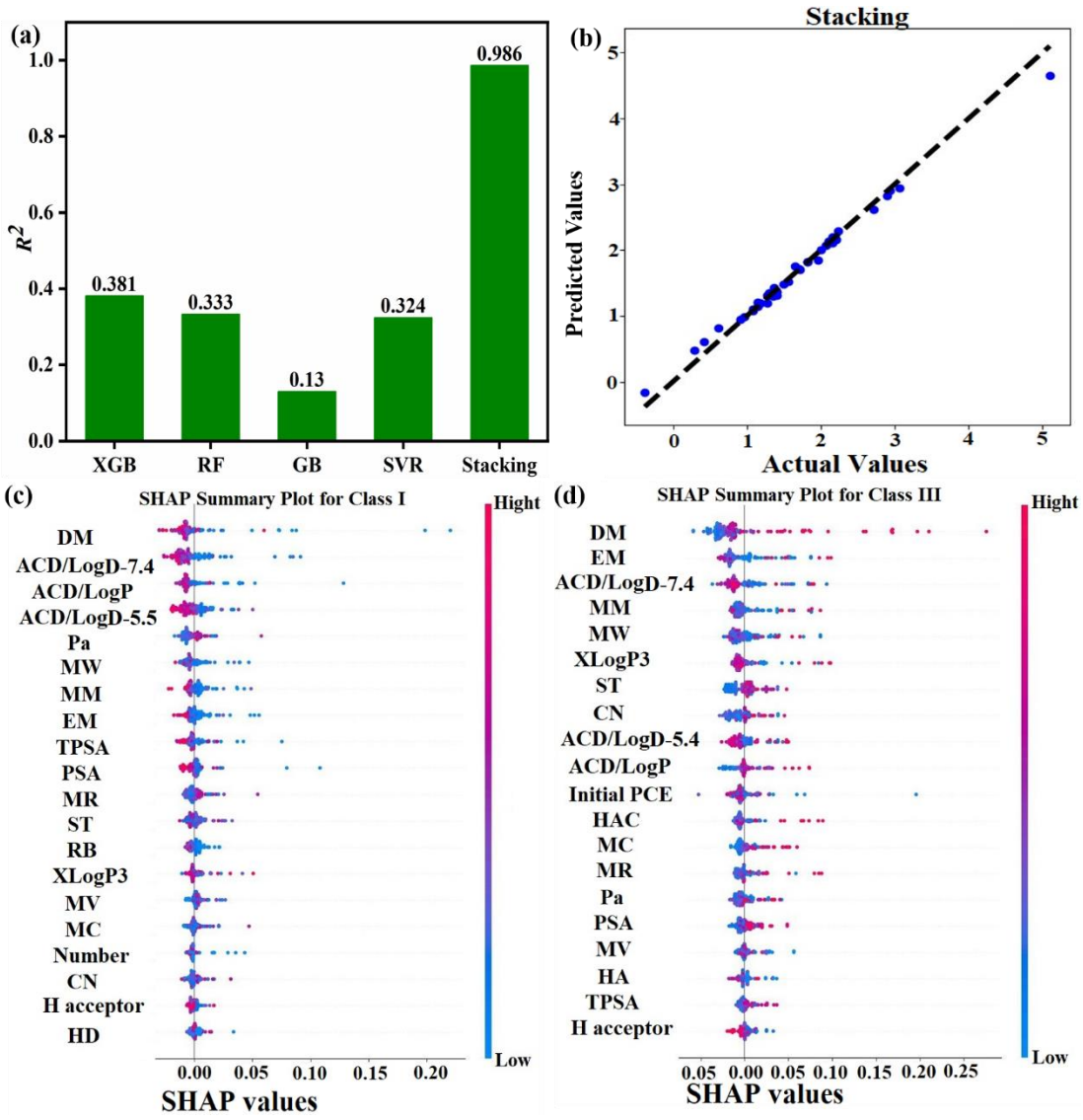
$CO^* + H^+ + e^- \rightarrow HCO^*$ (4)

$HCO^* + H^+ + e^- \rightarrow O^*CH_2$ (5)

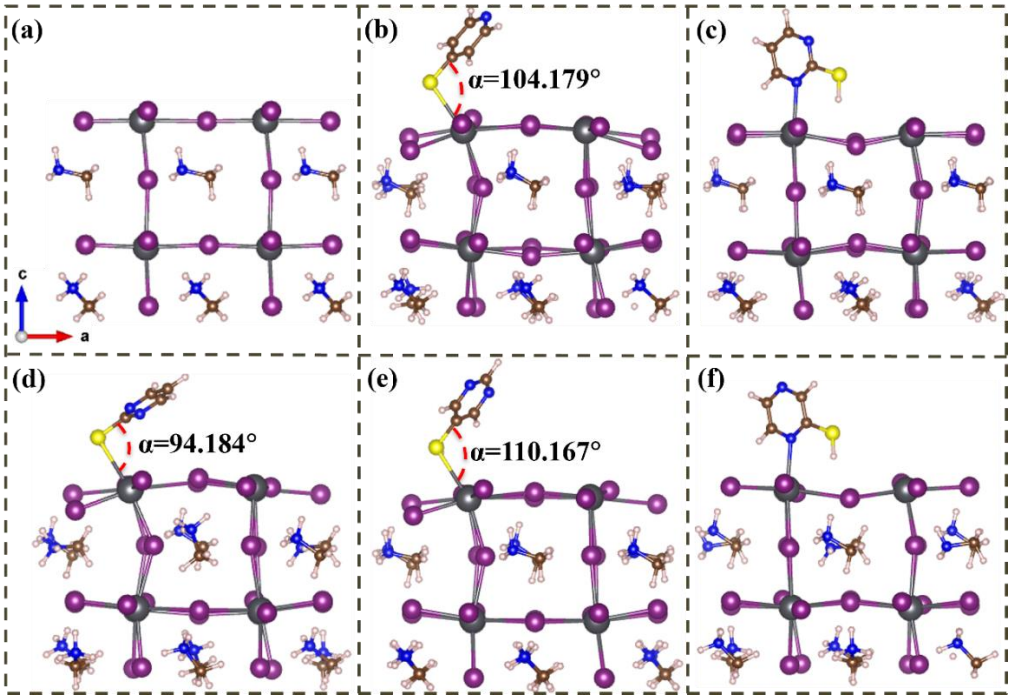
Part II. Additional Images



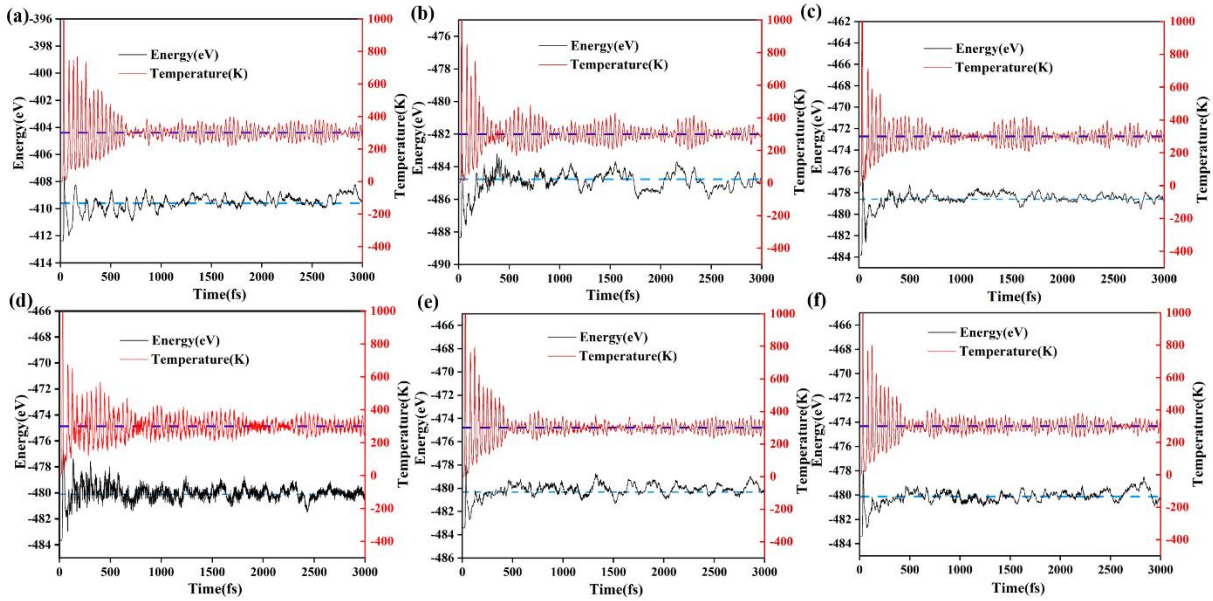
Supplementary Figure 1. (a) The Four Step Model of Machine Learning Process. (b) Model diagram of n-i-p perovskite devices. (c) - (f) The regression curves between the actual and predicted values of XGB, RF, GB, and SVR models, respectively.



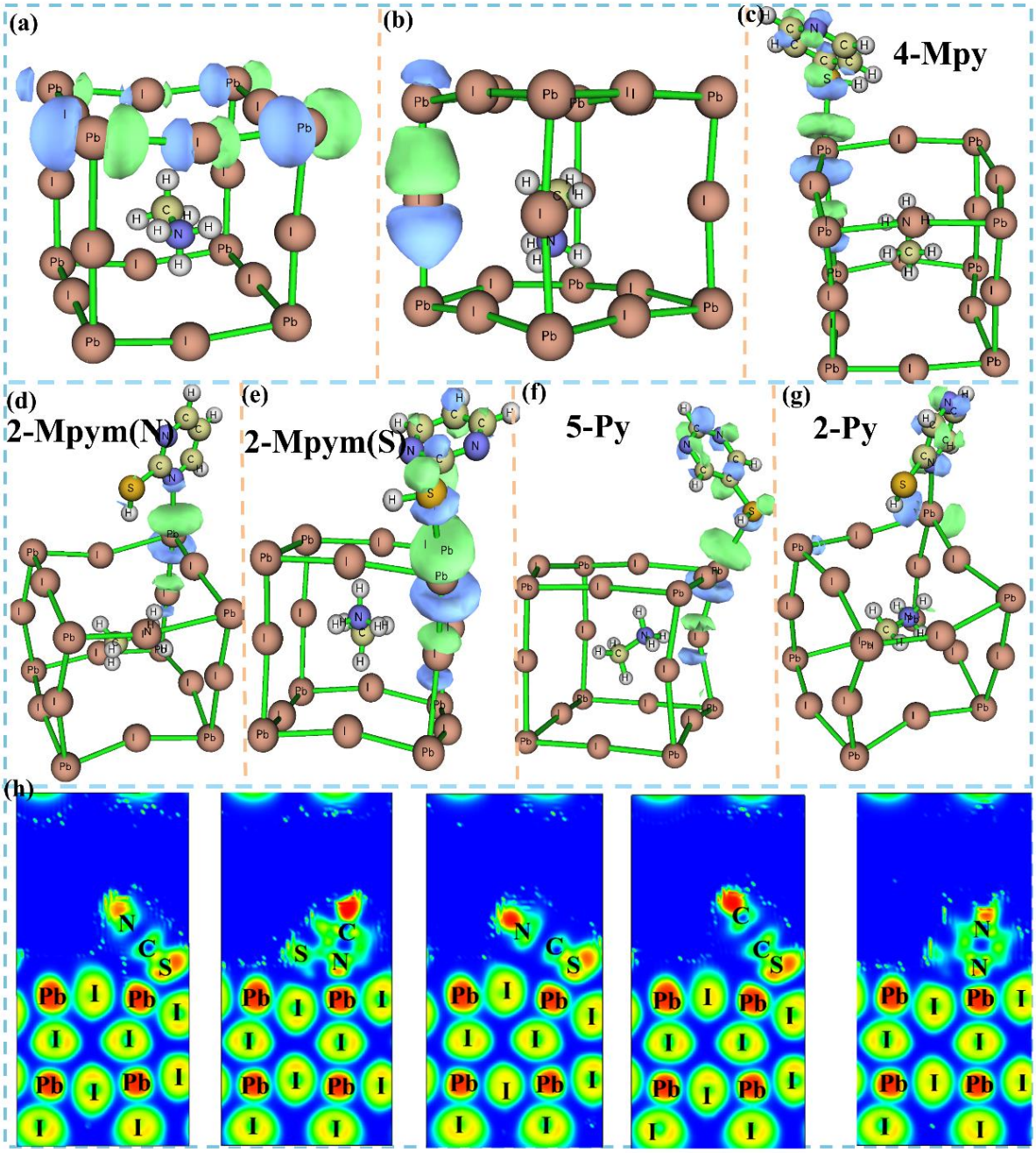
Supplementary Figure 2. (a) The coefficient of determination (R^2) for four individual ML models and stacked models. (b) Regression curves of stacked models composed of four models in the appeal. (c) - (d) The distribution of the impact of input features on model output (SHAP value) for I-class and II-class, respectively. Color represents feature values (red high, blue low).



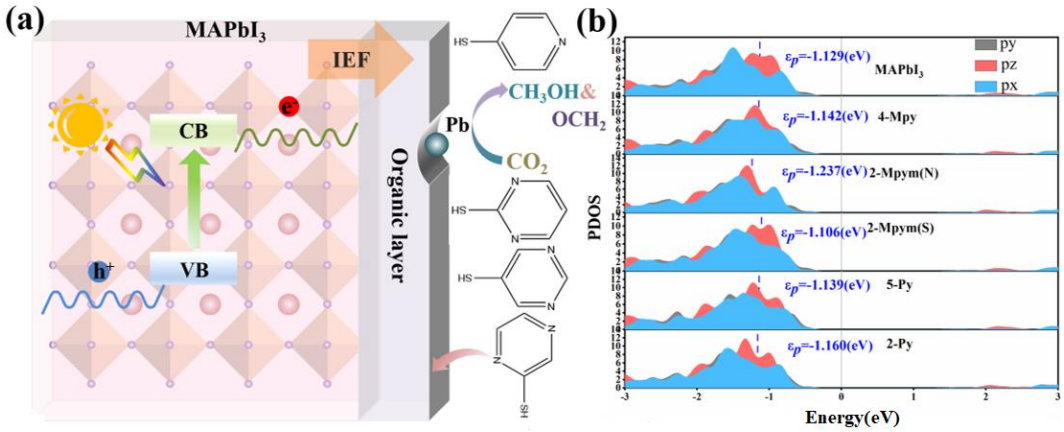
Supplementary Figure 3. (a) $MAPbI_3$ intrinsic structure. (b) The geometric structure of 4-Mpy small molecule passivator after adsorption, (c - d) The geometric structures of 2-Mpym(N) and 2-Mpym(S) small molecule passivators after adsorption, respectively. (e) Geometric structure of 5-Py small molecule passivator after adsorption, (f) Geometric structure of 2-Py small molecule passivator after adsorption.



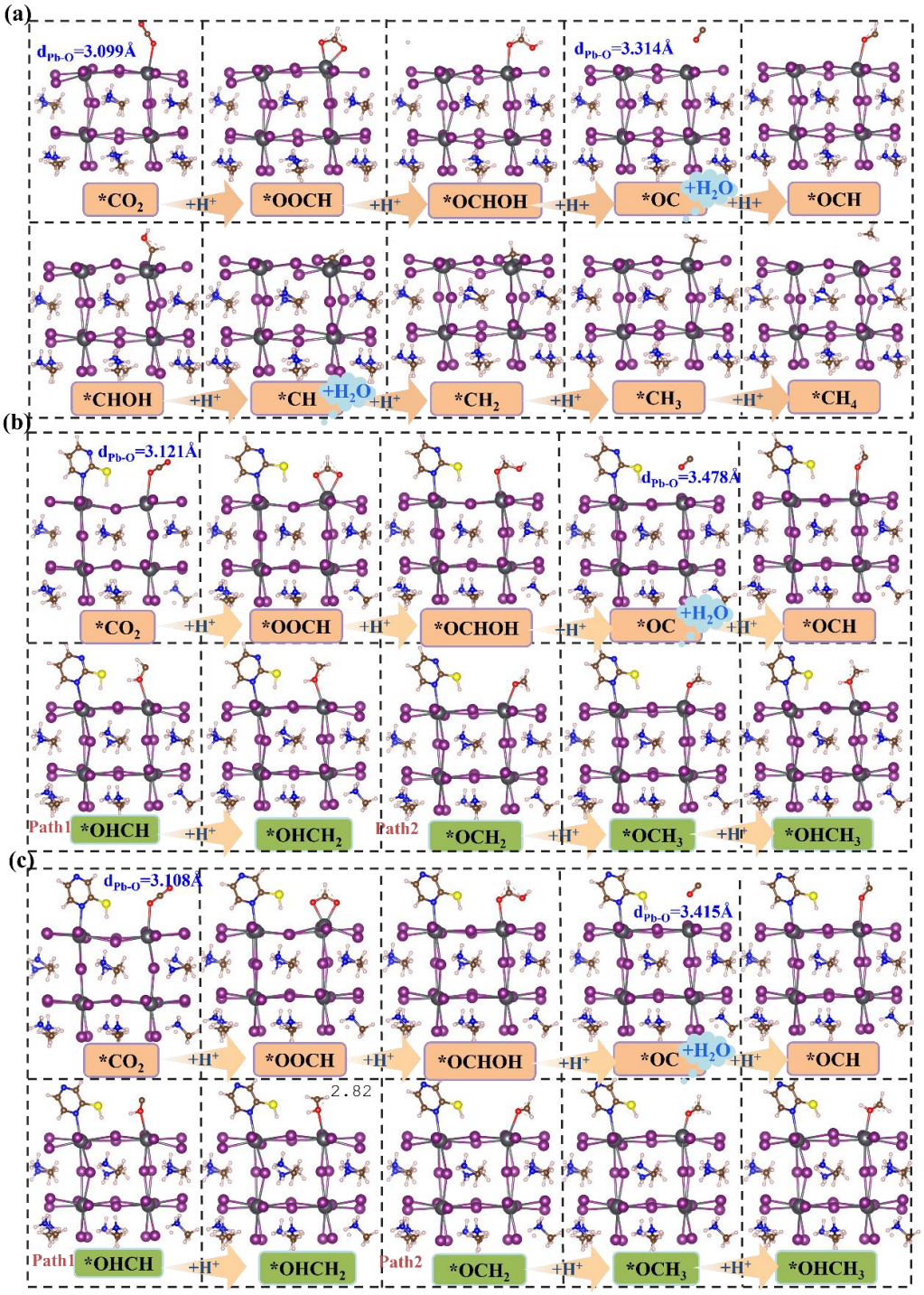
Supplementary Figure 4. The AIMD of the perovskite interface acting on small molecules of passivator (PMs) is calculated. The temperature is 300K and the number of steps is 3000. (a) – (f) PMs is are $MAPbI_3$, 4-Mpy, 2-Mpym(N), 2-Mpym(S), 5-Py, and 2-Py, respectively.



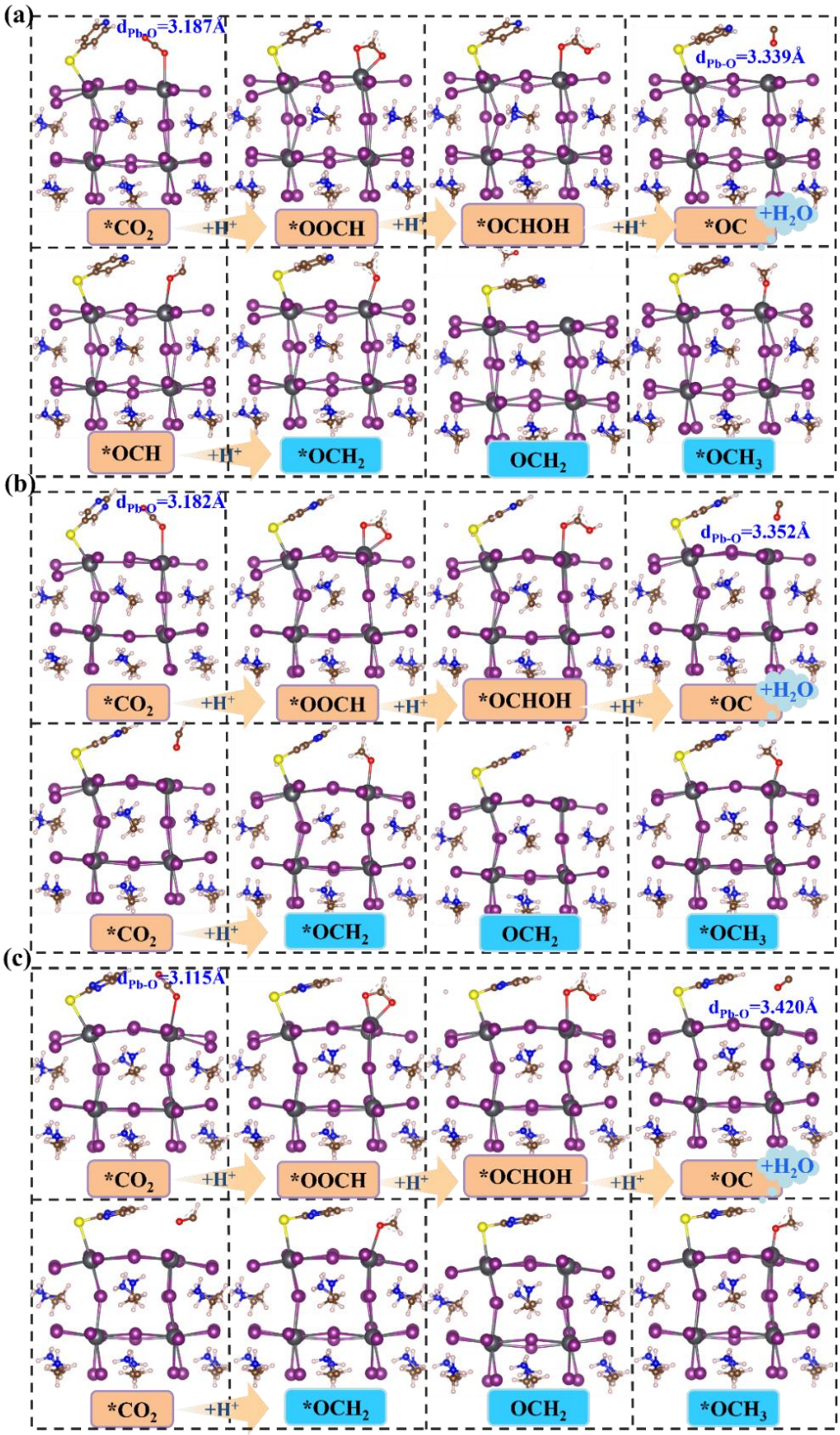
Supplementary Figure 5. (a) - (b) The σ* orbital visualizations formed by the coupling of I 2*p* orbitals and Pb 2*p* orbitals in the intrinsic structure, and the σ* orbital visualizations formed by I 2*p* orbitals and Pb 2*s* orbitals, respectively. (c) - (g) The bonding information between PMs and the interface. Blue represents negative phase, green represents positive phase. (h) The ELF plots after adding 4-MPy, 2-MPym(N), 2-MPym(S), 5-Py and 2-Py.



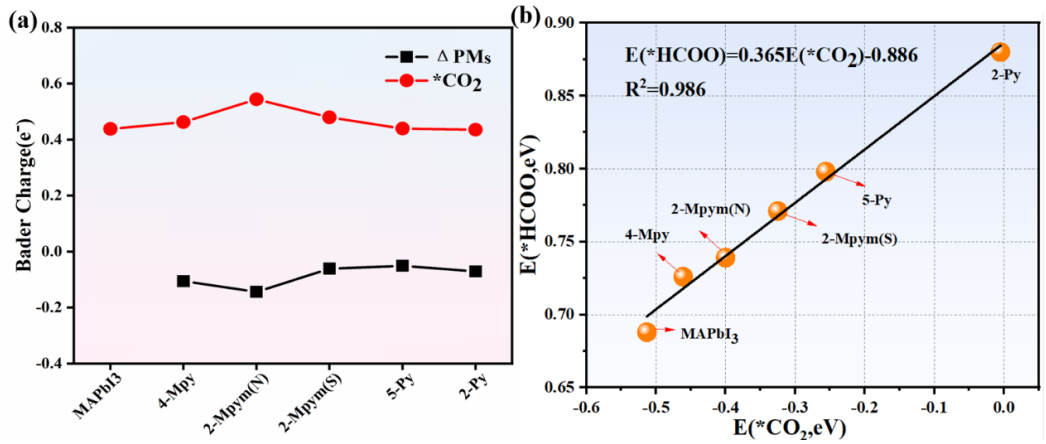
Supplementary Figure 6. (a) Schematic diagram of photocatalytic models for the influence of different PMs on interfaces. (b) The show the partial density of states (PDOS) of I 5*p* orbitals after CO_2 adsorption. The center of the *p*-band is indicated in blue font.



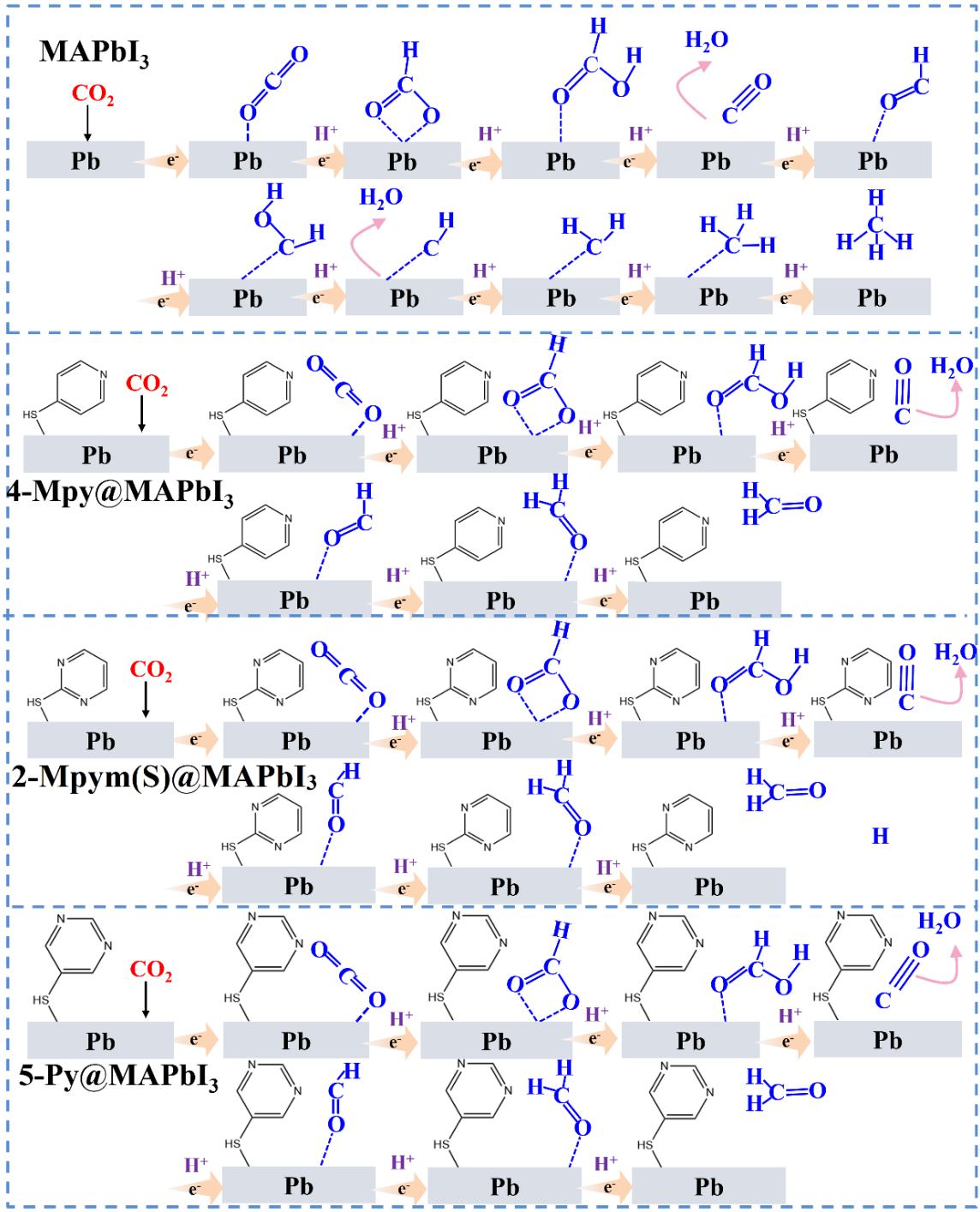
Supplementary Figure 7. (a) Gibbs free energy path maps of CO_2RR to CH_4 are on the surface of $MAPbI_3$. (b) - (c) Free energy path diagrams of CO_2RR to CH_2OH on the surface of 2-Mpym(N)@$MAPbI_3$ and 2-Py@$MAPbI_3$.



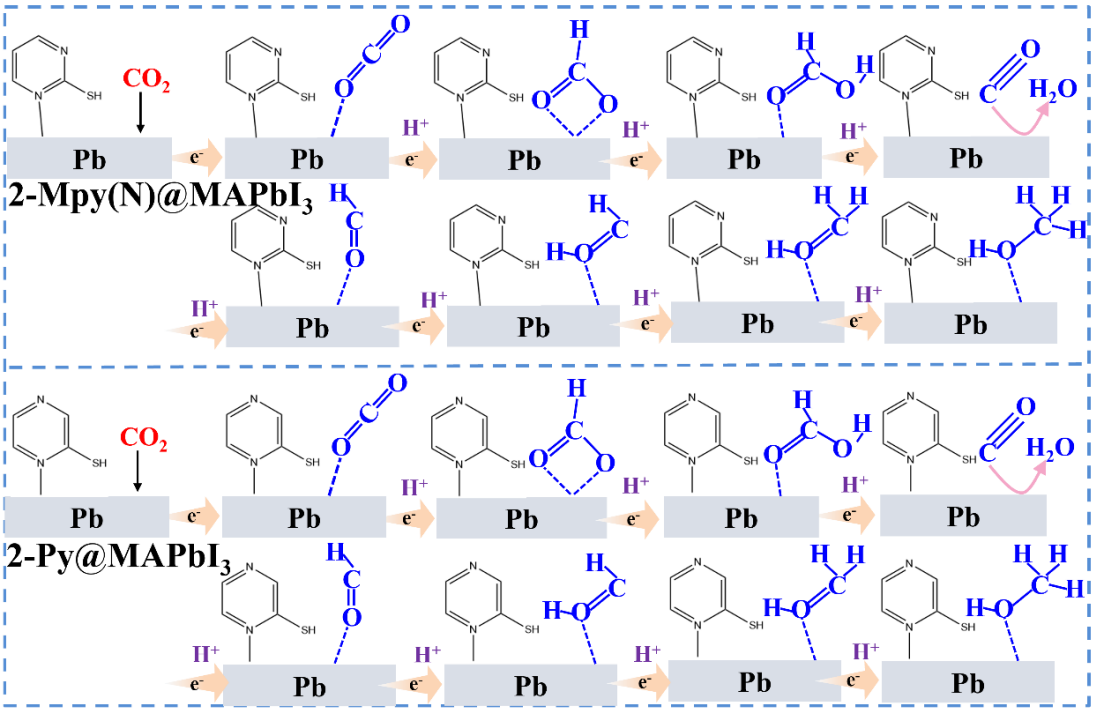
Supplementary Figure 8. (a) - (c) represent the Gibbs free energy path maps of CO_2RR to CH_2O at 4-Mpy, 5-Py, and 2-Mpym(S)@$MAPbI_3$ surface, respectively.



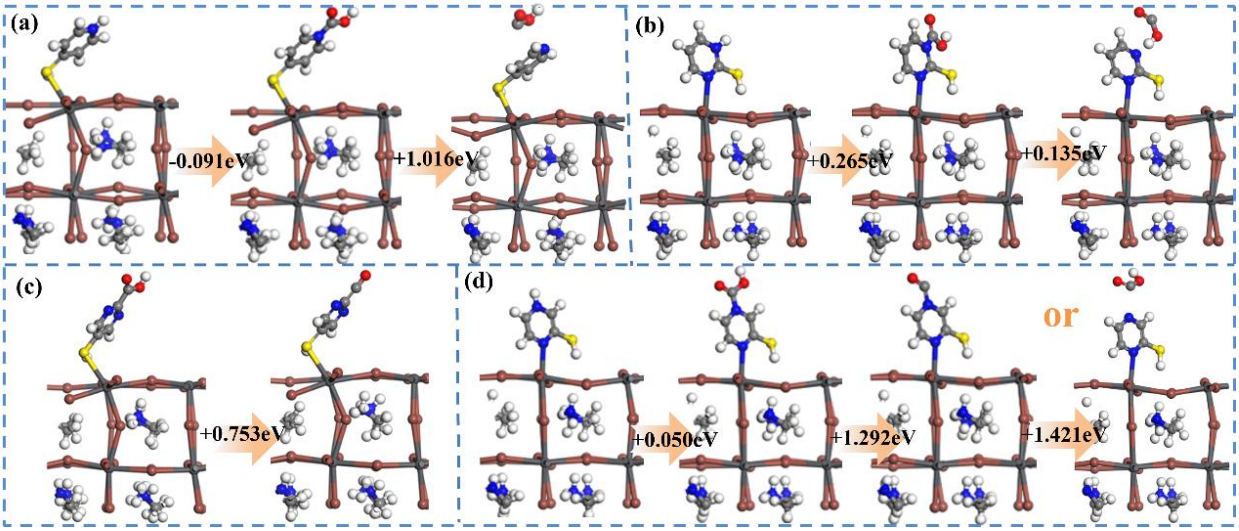
Supplementary Figure 9. (a) Intrinsic structure and PMs@$MAPbI_3$ comparison chart of charge transfer between CO_2 and organic small molecules (PMs) after CO_2 adsorption. (b) functional relationship between E_{*CO2}/E_{*OOCH} adsorption energy.



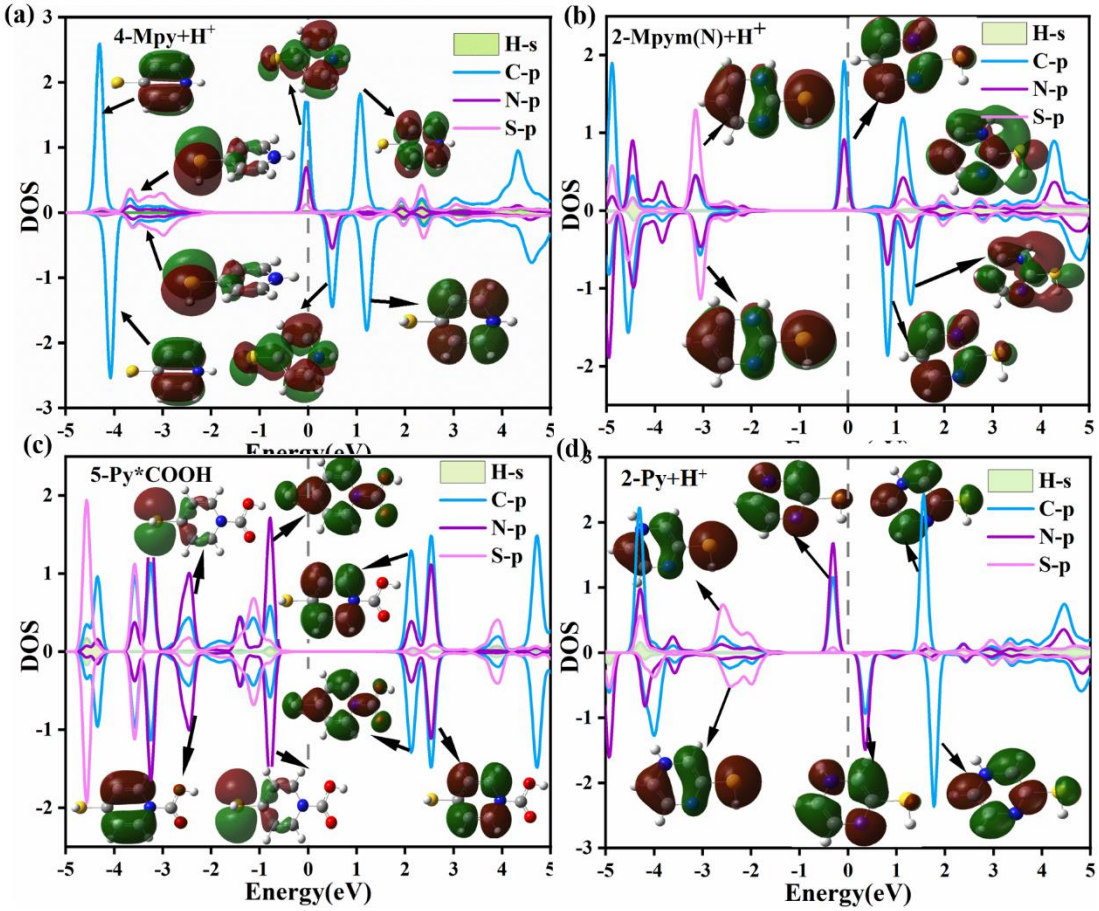
Supplementary Figure 10. The molecular bonding mechanisms for $MAPbI_3$ to reduce CO_2 to CH_4 at the interface and for 4-Mpy, 2-Mpym(S), and 5-Py to reduce CO_2 to CH_2O at the interface, respectively.



Supplementary Figure 11. The molecular bonding mechanism of 2-Mpym(N) and 2-Py in reducing CO_2 to CH_3OH at the perovskite interface.



Supplementary Figure 12. The reduction path forces for CO_2 reduction at the molecular interface are 4-Mpy, 2-Mpym(N), 5-Py and 2-Py. The numerical values represent the required energy barrier.



Supplementary Figure 13. The density of states (DOS) of H^+ or *COOH adsorbed during the self catalytic process of 4-Mpy, 2-Mpym(N), 5-Py, and 2-Py passivators at the perovskite interface. Among them, the orbital model diagram of each peak is marked near the Fermi level.

Part III. Additional Tables

Supplementary Table 1. Common products from CO_2 reduction with their simplified redox equations[12,13]

No.	Reaction equation	Product	E^0(V)
1	$CO_2+1e^- \rightarrow CO_2^-$	Carbonate anion	-1.90
2	$CO_2+2H^++2e^- \rightarrow HCOOH_{(aq)}$	Formic Acid	-0.61
3	$CO_2+2H^++2e^- \rightarrow CO_{(g)}+H_2O$	Carbon monoxide	-0.53
4	$CO_2+4H^++4e^- \rightarrow HCOH_{(aq)}+H_2O$	Formaldehyde	-0.48
5	$CO_2+6H^++6e^- \rightarrow CH3OH_{(aq)}+H_2O$	Methanol	-0.38
6	$CO_2+8H^++8e^- \rightarrow CH_{4(g)}+2H_2O$	Methane	-0.24

7	$2CO_2+12H^++12e^-\rightarrow C_2H_{4(g)}+4H_2O$	Ethene	-0.34
8	$2CO_2+12H^++12e^-\rightarrow C_2H_5OH_{(aq)}+3H_2O$	Alcohol	-0.33

Supplementary Table 2. Parameter settings of TiO_2/PMs@$MAPbI_3$/Spiro-OMeTAD layers in SCAPS-1D device simulation.

Parameters	TiO_2[14]	PM@$MAPbI_3$[15,16]	Spiro-OMeTAD[17]
Thickness(nm)	100	600	350
Bandgap(eV)	3.20	*x*	3.00
Electron affinity(eV)	3.90	3.90	2.45
Dielectric Permittivity(relative)	9	32.00	3
CB effective density of states (cm^{-3})	$2x10^{18}$	$2.8x10^{20}$	$1x10^{19}$
VB effective density of states (cm^{-3})	$1.8x10^{18}$	$3.9x10^{20}$	$1x10^{19}$
Electron mobility(cm^2/Vs)	20	11.8	$2x10^4$
Hole mobility(cm^2/Vs)	10	11.8	$2x10^4$
Shallow uniform acceptor density NA(cm^{-3})	0	$1x10^{13}$	$2x10^{18}$
Shallow uniform donor density ND(cm^{-3})	$2x10^{16}$	$1x10^{13}$	0

[**Note**] Among them, *x* refers to the specific bandgap values of different structures, and other parameter settings are consistent. For example, the thermal velocity of electrons/holes is 10^7cm/s, the hole capture cross-section of the absorber/ETL is 10^{-18} cm^2, the hole capture cross-section of the absorber/HTL is 10^{-19} cm^2, the electron capture absorber/ETL cross-section is 10^{-19} cm^2, the absorber/HTL cross-section is 10^{-18} cm^2, the constant illumination is 1000 w/m^2, the working temperature is 300K, and the AM is 1.5G[15].

Supplementary Table 3. The data for fully optimized lattice constants a=b, c, angle α=β=γ, bond length between Pb-N/S atoms, dipole moment of PMs (*DM*), adsorption energy (E_b), band gap (E_g), and work function (*WF*) for intrinsic materials and PMs@$MAPbI_3$

Compounds	*a*(Å)	*c*(Å)	α(°)	$d_{Pb\text{-}S/Pb\text{-}N}$(Å)	*DM*	E_b(eV)	*WF*	E_g(eV)
$MAPbI_3$	12.873	26.236	90	-	-	-	5.241	1.829
4-Mpy	12.873	26.236	90	3.120	1.862	-0.197	5.205	1.881
2-Mpym(N)	12.873	26.236	90	2.735	2.577	-0.246	4.824	1.776
2-Mpym(S)	12.873	26.236	90	3.231	2.577	-0.374	4.779	1.872
5-Py	12.873	26.236	90	3.210	1.929	-0.051	5.252	1.853
2-Py	12.873	26.236	90	2.715	1.139	-0.587	4.976	1.811

The global reactivity is used to evaluate the overall activity of molecules. The global reactivity is used to evaluate the overall activity of molecules, as shown by *ΔE*:

$$\Delta E = E_{LUMO} - E_{HOMO} \quad (11)$$

The larger the *ΔE* value, the more stable the system is, while the smaller the *ΔE* value, the more unstable the system is.

Supplementary Table 4. The energy values of LUMO and HOMO of different molecules and the energy values of global reaction activity.

Molecule	HOMO(eV)	LUMO(eV)	*ΔE*(eV)
4-Mpy	-6.783	-1.005	5.778
2-Mpym	-6.773	-1.727	5.046
5-Py	-6.854	-1.572	5.282
2-Py	-6.773	-1.523	5.250

Supplementary Table 5. The average charge transfer amount of Pb atoms, I atoms and PMs in PMs@$MAPbI_3$.

Bader(e^-)	4-Mpy	2-Mpym(N)	2-Mpym(S)	5-Py	2-Py
$\overline{Pb}$	-0.504	-0.505	-0.888	-0.881	-0.902
$\bar{I}$	+0.840	+0.548	+0.547	0.545	+0.551
PM	+0.9	+0.489	+0.553	+0.526	+0.513

Supplementary Table 6. After PMs@$MAPbI_3$ adsorbs CO_2, the average charge transfer amount of Pb atom, I atom, CO_2*,PMs, and the differences in PMs charges before and after adsorption (ΔPMs).

Bader(e^-)	$MAPbI_3$	4-Mpy	2-Mpym(N)	2-Mpym(S)	5-Py	2-Py
$\overline{Pb}$	-0.900	-0.897	-0.893	-0.895	-0.889	-0.908
$\bar{I}$	0.533	+0.535	+0.533	+0.532	+0.533	+0.537
CO_2*	0.438	0.462	0.544	0.479	0.439	0.435
PMs	-	-0.155	-0.176	-0.102	-0.091	-0.090
ΔPMs	-	-0.106	-0.144	-0.061	-0.051	-0.071

Supplementary Table 7. Electron (m_e) and hole (m_h) effective masses, the reduced effective mass(μ), high-frequency dielectric constants (ε) and exciton binding energies (E_g) in the intrinsic structure and PM@$MAPbI_3$.

Compounds	$m_e(m_0)$		$m_h(m_0)$		μ	ε	E_d（eV）
	$X\rightarrow\Gamma$	$M\rightarrow\Gamma$	$X\rightarrow\Gamma$	$M\rightarrow\Gamma$			
$MAPbI_3$	0.926	0.382	1.185	0.246	0.342	2.645	0.665
4-Mpy	0.958	0.484	4.353	0.256	0.549	2.796	0.955
2-Mpym(N)	0.604	0.406	1.208	0.255	0.299	2.900	0.484
2-Mpym(S)	0.997	0.439	1.729	0.236	0.415	2.756	0.743
5-Py	0.875	0.559	1.062	0.238	0.341	2.763	0.608
2-Py	0.544	0.393	1.411	0.254	0.300	2.843	0.505

Supplementary Table 8. The total energy (*E*), adsorption energy (*ΔE*), zero energy (ΔE_{ZPE}), entropy (*TΔS*), and Gibbs free energy (ΔG) of CO_2RR intermediates at the synergistic interface of different small molecule passivators on $MAPbI_3$. The unit is eV.

$MAPbI_3$	*E*	*ΔE*	ΔE_{ZPE}	*TΔS*	*ΔG*
*CO_2	–435.919	-0.005	0.319	0.25	0.064
*OOCH	–438.421	0.880	0.601	0.257	1.244
*OCHOH	–443.092	-1.289	0.916	0.317	-0.690
*OC	–427.866	0.866	0.144	0.239	0.771
*OCH	–430.245	1.003	0.372	0.253	1.122
HO*CH	–433.911	-0.284	0.787	0.174	-0.329
*CH	–422.005	0.928	0.330	0.089	1.169
*CH_2	–427.412	-2.025	0.671	0.081	-1.435
*CH_3	–431.693	-0.899	0.905	0.204	-0.198
*CH_4	-437.258	-2.183	1.213	0.246	-1.216

Supplementary Table 9. The total energy (*E*), adsorption energy (*ΔE*), zero energy (ΔE_{ZPE}), entropy (*TΔS*), and Gibbs free energy (ΔG) of CO_2RR intermediates at the synergistic interface of different small molecule passivators on 4-Mpy@$MAPbI_3$. The unit is eV.

4-Mpy	*E*	*ΔE*	E_{ZPE}	*TS*	*G*
*CO_2	-511.727	-0.4	0.315	0.187	-0.272
*OOCH	-514.370	0.739	0.600	0.187	1.152
*OCHOH	-519.004	-1.252	0.913	0.293	-0.632
*OC	-503.703	0.941	0.140	0.182	0.899
*OCH	-506.053	1.032	0.375	0.241	1.166
*OCH_2	-511.418	-1.983	0.753	0.257	-1.487
OCH_2	-511.201	0.199	0.722	0.211	0.710
*OCH_3	-514.296	0.504	1.038	0.183	1.404

Supplementary Table 10. The total energy (*E*), adsorption energy (*ΔE*), zero energy (ΔE_{ZPE}), entropy (*TΔS*), and Gibbs free energy (ΔG) of CO_2RR intermediates at the synergistic interface of different small molecule passivators on 2-Mpym(N)@$MAPbI_3$. The unit is eV.

2-Mpym(N)	*E*	*ΔE*	ΔE_{ZPE}	*TΔS*	*ΔG*
*CO_2	-507.296	-0.513	0.317	0.176	-0.372
*OOCH	-509.990	0.688	0.602	0.251	1.039
*OCHOH	-514.390	-1.018	0.905	0.278	-0.391
*OC	-499.154	0.876	0.143	0.242	0.777
*OCH	-501.520	1.012	0.375	0.313	1.074
*OHCH	-504.384	0.518	0.717	0.257	0.978
*$OHCH_2$	-509.522	-1.786	1.009	0.312	-1.089
*OCH_2	-506.847	-1.945	0.754	0.250	-1.441
*OCH_3	-509.734	0.495	1.053	0.264	1.284
*$OHCH_3$	-515.046	-2.112	1.404	0.276	-0.984

Supplementary Table 11. The total energy (*E*), adsorption energy (*ΔE*), zero energy (ΔE_{ZPE}), entropy (*TΔS*), and Gibbs free energy (ΔG) of CO_2RR intermediates at the synergistic interface of different small molecule passivators on 2-Mpym(S)@$MAPbI_3$. The unit is eV.

2-Mpym(S)	*E*	*ΔE*	ΔE_{ZPE}	*TΔS*	*ΔG*
*CO_2	-507.233	-0.325	0.316	0.187	-0.196
*OOCH	-509.844	0.771	0.596	0.201	1.166
*OCHOH	-514.359	-1.133	0.911	0.310	-0.532
*OC	-499.007	0.992	0.146	0.225	0.913
*OCH	-501.433	0.956	0.373	0.251	1.078
*OCH_2	-506.726	-1.911	0.751	0.263	-1.423
OCH_2	-506.440	0.286	0.509	0.125	0.670
*OCH_3	-509.684	0.424	1.051	0.208	1.267

Supplementary Table 12. The total energy (*E*), adsorption energy (*ΔE*), zero energy (ΔE_{ZPE}), entropy ($T\Delta S$), and Gibbs free energy (ΔG) of CO_2RR intermediates at the synergistic interface of different small molecule passivators on 5-Py@$MAPbI_3$. The unit is eV.

5-Py	*E*	*ΔE*	ΔE_{ZPE}	$T\Delta S$	*ΔG*
*CO_2	-506.865	-0.461	0.317	0.177	-0.321
*OOCH	-509.517	0.726	0.593	0.136	1.183
*OCHOH	-514.114	-1.215	0.907	0.325	-0.633
*OC	-498.791	0.963	0.141	0.177	0.927
*OCH	-500.330	1.843	0.290	0.177	1.956
*OCH_2	-506.485	-2.773	0.755	0.253	-2.271
OCH_2	-506.172	0.313	0.509	0.125	0.697
*OCH_3	-509.404	0.463	1.041	0.159	1.345

Supplementary Table 13. The total energy (*E*), adsorption energy (*ΔE*), zero energy (ΔE_{ZPE}), entropy ($T\Delta S$), and Gibbs free energy (ΔG) of CO_2RR intermediates at the synergistic interface of different small molecule passivators on 2-Py@$MAPbI_3$. The unit is eV.

2-Py	*E*	*ΔE*	ΔE_{ZPE}	$T\Delta S$	*ΔG*
*CO_2	-507.166	-0.256	0.319	0.247	-0.178
*OOCH	-509.750	0.798	0.601	0.252	1.147
*OCHOH	-514.204	-1.072	0.913	0.319	-0.478
*OC	-498.937	0.907	0.148	0.290	0.765
*OCH	-501.318	1.001	0.373	0.320	1.054
*OHCH	-503.418	1.282	0.680	0.277	1.685
*$OHCH_2$	-509.37	-2.570	1.017	0.231	-1.784
*OCH_2	-506.657	-1.957	0.755	0.246	-1.488
*OCH_3	-509.539	0.500	1.058	0.246	1.312
*$OHCH_3$	-514.814	-1.893	1.406	0.283	-0.770

Supplementary Table 14. The total energy (E), adsorption energy (ΔE), zero energy (ΔE_{ZPE}), entropy ($T\Delta S$), and Gibbs free energy (ΔE) of CO_2 pathways catalyzed by different organic small molecule passivators themselves, in eV. The unit is eV.

Pathway	Molecule	E	ΔE	ΔE_{ZPE}	$T\Delta S$	ΔG
$*H^+$	4-Mpy	-491.551	0.163	0.325	0.011	0.477
	2-Mpym(N)	-487.044	0.129	0.320	0.012	0.437
	2-Py	-487.086	0.214	0.320	0.014	0.520
*COOH	4-Mpy	-514.644	-0.101	0.672	0.185	0.386
	2-Mpym(N)	-509.811	0.225	0.668	1.191	0.702
	5-Py	-506.238	0.166	0.668	0.198	0.638
	2-Py	-510.107	-0.029	0.678	0.179	0.470
*CO	4-Mpy	-502.044	1.622	0.224	0.118	1.728
	2-Mpym(N)	-497.519	1.314	0.225	0.116	1.423
	5-Py	-493.948	1.312	0.223	0.144	1.391
	2-Py	-497.472	1.657	0.224	0.116	1.765
COOH	4-Mpy	-513.515	1.129	0.558	0.285	1.402
	2-Mpym(N)	-509.251	0.560	0.561	0.284	0.837
	5-Py	-502.333	3.905	0.318	0.226	3.997
	2-Py	-508.509	1.598	0.559	0.266	1.891

Supplementary Table 15. Detailed definitions of 26 feature vectors used in ML.

Feature	Detailed definition	Feature	Detailed definition
LPE	The number of lone pair electrons (including N S. P, O, X(halogen))	Pa	Polarizability
Number	The number of C atoms	ST	Molecular complexity
DM	Dipole moment	MV	Molar volume

Gap	Molecular bandgap	MW	Molecular Weight
CN	Coordination number	XLogP3	Lipid−water partition coefficient
MR	Molar diopter	HD	Hydrogen bond donor count
H acceptor	H bond acceptors	HA	Hydrogen bond acceptor count
H donor	H bond donors	RB	Rotatable Bond Count
Rotation B	Number of rotatable keys	EM	Exact Mass
ACD/LogP	Lipid-water partition coefficient used to describe hydrophobicity	MM	Monoisotopic Mass
ACD/LogD-5.5	octanol-water distribution coefficient at pH5.5	TPSA	Molecular complexity
ACD/LogD-7.4	octanol-water distribution coefficient at pH 7.4	HAC	Heavy atom counting
PSA	Polar surface area	MC	Molecular complexity

References

[1] X. Zhang, B. Ding, Y. Wang, Y. Liu, G. Zhang, L. Zeng, L. Yang, C. Li, G. Yang, M.K. Nazeeruddin, B. Chen, Machine Learning for Screening Small Molecules as Passivation Materials for Enhanced Perovskite Solar Cells. *Adv Funct Materials* **2024**, 2314529. https://doi.org/10.1002/adfm.202314529.

[2] Y. Hu, X. Hu, L. Zhang, T. Zheng, J. You, B. Jia, Y. Ma, X. Du, L. Zhang, J. Wang, B. Che, T. Chen, S. (Frank) Liu, Machine-Learning Modeling for Ultra-Stable High-Efficiency Perovskite Solar Cells. *Advanced Energy Materials* **2022**, *12* (41), 2201463. https://doi.org/10.1002/aenm.202201463.

[3] J.E.T. Akinsola, Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology (IJCTT)* **2017**, *48*, 128–138. https://doi.org/10.14445/22312803/IJCTT-V48P126.

[4] M. ea Frisch, G. Trucks, H.B. Schlegel, G. Scuseria, Ma. Robb, J. Cheeseman, G. Scalmani, V. Barone, G. Petersson, H. Nakatsuji, Gaussian 16. **2016**.

[5] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *The Journal of chemical physics* **2010**, *132* (15). https://doi.org/10.1063/1.3382344.

[6] R. Krishnan, J.S. Binkley, R. Seeger, J.A. Pople, Self-consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *The Journal of chemical physics* **1980**, *72* (1), 650–654. https://doi.org/10.1063/1.438955.

[7] Y. Cai, C. Chen, F. Chen, S. Wang, Z. Zhang, S. Gao, S. Guo, W. Chen, E. Abduryim, Y. Liu, X. Guan, G. Liu, P. Lu, Theoretical Exploration of CO^2 Photocatalytic Reduction Using Single Atom Gold Nanoparticles (Au^0) Modified $SrTi_{0.875}Hf_{0.125}O_3$. *Journal of Catalysis* **2024**, *432*, 115410. https://doi.org/10.1016/j.jcat.2024.115410.

[8] L. Z. Wang, Y. Q. Zhao, B. Liu, L. J. Wu, M. Q. Cai, First-Principles Study of Photovoltaics and Carrier Mobility for Non-Toxic Halide Perovskite $CH_3NH_3SnCl_3$: Theoretical Prediction. *Physical Chemistry Chemical Physics* **2016**, *18* (32), 22188–22195. https://doi.org/10.1039/C6CP03605H.

[9] G.H. Wannier, The Structure of Electronic Excitation Levels in Insulating Crystals. *Physical Review* **1937**, *52* (3), 191. https://doi.org/10.1103/PhysRev.52.191.

[10] J. Zhu, Y. Zhang, Z. Chen, Z. Zhang, X. Tian, M. Huang, X. Bai, X. Wang, Y. Zhu, H. Jiang, Superexchange-Stabilized Long-Distance Cu Sites in Rock-Salt-Ordered Double Perovskite Oxides for CO_2 Electromethanation. *Nat Commun* **2024**, *15* (1), 1565. https://doi.org/10.1038/s41467-024-45747-5.

[11] Y. Kong, X. Li, A.R. Puente Santiago, T. He, Nonmetal Atom Doping Induced Orbital Shifts and Charge Modulation at the Edge of Two-Dimensional Boron Carbonitride Leading to Enhanced Photocatalytic Nitrogen Reduction. *J. Am. Chem. Soc.* **2024**, *146* (9), 5987–5997. https://doi.org/10.1021/jacs.3c12780.

[12] C. Chen, J. Li, X. Tan, Y. Zhang, Y. Li, C. He, Z. Xu, C. Zhang, C. Chen, Harnessing Single-Atom Catalysts for CO_2 Electroreduction: A Review of Recent Advances. *EES. Catal.* **2024**, *2* (1), 71–93. https://doi.org/10.1039/D3EY00150D.

[13] S. Shyamal, N. Pradhan, Halide Perovskite Nanocrystal Photocatalysts for CO_2 Reduction: Successes and Challenges. *J. Phys. Chem. Lett.* **2020**, *11* (16), 6921–6934. https://doi.org/10.1021/acs.jpclett.0c00191.

[14] Y.H. Khattak, F. Baig, A. Shuja, S. Beg, B.M. Soucase, Numerical Analysis Guidelines for the Design of Efficient Novel Nip Structures for Perovskite Solar Cell. *Solar Energy* **2020**, *207*, 579–591. https://doi.org/10.1016/j.solener.2020.07.012.

[15] C. Son, H. Son, B. S. Jeong, Enhanced Conversion Efficiency in $MAPbI_3$ Perovskite Solar Cells through Parameters Optimization via SCAPS-1D Simulation. *Applied Sciences* **2024**, *14* (6), 2390. https://doi.org/10.3390/app14062390.

[16] S. Ahmed, F. Jannat, Md.A.K. Khan, M.A. Alim, Numerical Development of Eco-Friendly Cs_2TiBr_6 Based Perovskite Solar Cell with All-Inorganic Charge Transport Materials via SCAPS-1D. *Optik* **2021**, *225*, 165765. https://doi.org/10.1016/j.ijleo.2020.165765.

[17] D. Stanić, V. Kojić, T. Čižmar, K. Juraić, L. Bagladi, J. Mangalam, T. Rath, A. Gajović, Simulating the Performance of a Formamidinium Based Mixed Cation Lead Halide Perovskite Solar Cell. *Materials* **2021**, *14* (21), 6341. https://doi.org/10.3390/ma14216341.