

Supporting Information

Machine Learning-Guided High-Throughput Design of Asymmetric A-DA'D-A Acceptors toward Efficient Organic Solar Cells

Ziye Chen^{a,1}, Kuo Wang^{b,1}, Di Huang^{a,*}, Zhiyan Peng^a, Jiahao Deng^a, Zhimin Feng^a, Jiangtao Wu^a, Kang Zhang^a, Jiaojiao Liang^a, Ling Zhao^{c,*}, Jaemin Kong^{b,*}

^a, Hunan University of Technology, Zhuzhou 412008, China

^b, Department of Physics, Gyeongsang National University, Jinju 52828, Republic of Korea

^c, Shandong Provincial Key Laboratory of Optical Communication Science and Technology, School of Physical Science and Information Technology, Liaocheng University, Liaocheng 252059, China

*Corresponding author: dihuang@hut.edu.cn (Di Huang), zhaoling1@lcu.edu.cn (Ling Zhao), jaemin.kong@gnu.ac.kr (Jaemin Kong).

¹ These authors contributed equally to this work.

1. Dataset Building

The 378 non-fullerene small-molecule acceptors with a well-defined A-DA'D-A type are obtained from the literature in Google Scholar search with the keywords "organic solar cells (OSCs)" and "non-fullerene acceptors (NFAs)" published from 2019 to 2025 as the original dataset. For each datapoint, it contained the simplified molecular input line entry system (SMILES) of the acceptor, the HOMO/LUMO energy levels of the paired donor, and the corresponding device performance (J_{sc} , V_{oc} , FF and PCE). The dataset is then pre-screened to retain the samples with PCE higher than 10%, yielding 312 molecules for the subsequent analysis. This threshold is chosen from the fact that the primary objective of this work is to screen potential NFA molecules with high efficiency, and the device efficiencies of most A-DA'D-A type NFAs reported in the literature is higher than 15%. Since there are also a few cases where the PCE is between 10% and 15%, the threshold is set at 10% after comprehensive considering all these factors. The entries of sub-dataset, including the name of Acceptor and Donor, the corresponding PCE, and references is summarized in **Table S1**.

2. Data Preprocessing

Each acceptor molecular is first decomposed in Chem2D under the A-DA'D-A type small-molecule acceptors as five positional fragments (A_1 , D_1 , A' , D_2 , and A_2), which are exported and stored as SMILES. Given the close structural and functional similarity for A_1/A_2 and D_1/D_2 ,

these five fragments are consolidated into three functional fragment libraries: end acceptor units (53 A fragments), conjugated donor units (69 D fragments), and electron-deficient core units (72 A' fragments). Because raw SMILES cannot be used directly as model inputs, we construct one-hot encodings for each fragment library (A, D, and A'), and we use a fixed-length vector whose length equals the number of fragments in that library. For any given molecule, setting the entry for a fragment to "1" if it appears at the library's designated positions, and to "0" otherwise. Finally, concatenate the A, D, and A' vectors in order to obtain a fixed-length binary descriptor for the molecule, as seen in **Fig. S1**. In addition, the donor's HOMO and LUMO values are additionally incorporated as supplementary donor descriptors for effectively predict the corresponding PCE in the model. Furthermore, a correlation-based redundancy filter is applied: when the absolute value of pearson correlation coefficient (r) between two input features is higher than 0.9, only one feature with the stronger r value for the target feature is kept for reducing to 152 input features from 194 features and thus minimizing redundancy while retaining informative predictors for model training.

3. Model Building

All machine learning (ML) models are implemented in Python. The dataset is randomly divided into training and testing subsets with the 8:2 ratio. The training set is used for model construction and the testing set is used to evaluate the generalization ability of the model. Initially, six regression models were applied, including eXtreme gradient boosting (XGBoost) distinguishes itself from other ensemble algorithms by employing a gradient boosting framework that iteratively fits residuals to minimize error, enabling superior prediction accuracy ^[1]. Categorical boosting (CatBoost) has a symmetrical tree structure and built-in categorical feature handling mechanism, which can naturally handle high-dimensional sparse binary molecular fingerprint data and has a strong resistance to overfitting in small sample cases ^[2]. Support vector regression (SVR) performs reliably on high-dimensional fingerprint features and small-to-medium datasets, though the training cost scales with sample size and careful kernel and hyperparameter selection is required ^[3]. Random forest (RF) uses random sampling and feature subsets to construct a large number of decision trees, and obtains the final prediction results through voting or averaging. It has strong robustness against discrete features and noise, and is suitable for high-dimensional binary data like molecular fingerprints, although its prediction performance is slightly lower than that of the gradient boosting methods ^[4]. Gradient boosting regression (GBR) builds additive ensembles of shallow trees via stage-wise optimization with shrinkage and subsampling; it delivers strong accuracy on tabular and sparse binary data but is less memory-efficient ^[5]. Gaussian process

regression (GPR) is a nonparametric Bayesian approach that models functions via kernels and provides both predictive means and uncertainties; it is well-suited to small datasets and can capture complex relations in molecular fingerprints [6].

Hyperparameter tuning and random seed optimization were performed for all six algorithms to obtain the optimal models for this dataset. Specifically, an automatic hyperparameter tuning framework, Optuna, was employed to search the hyperparameter space. Compared with traditional grid search, Optuna dynamically adjusts the search strategy adopting bayesian optimization and intelligent sampling methods, improving the efficiency of model training. Optuna supports the dynamic definition of continuous search spaces, providing greater flexibility and using pruning strategies to prematurely terminate unproductive trials, saving computational resources [7]. Additionally, Optuna optimizes the sampling strategies based on experimental data to make it more adaptable to complex optimization problems. **Table S2 and Fig. S2** compare the optimization efficiencies of three hyperparameter tuning methods, including Optuna, GridSearchCV and RandomizedSearchCV. Under identical hardware conditions, Optuna achieve the best predictive performance with the fewest number of trials and the shortest tuning time, underscoring its advantage in efficiently exploring the search space and conserving computational resources, whereas the traditional GridSearchCV and RandomizedSearchCV are slower and have lower accuracy. And Optuna yields the highest testing R^2 of 0.88 after 60 trials, as illustrated in **Fig. S2b**. The comparison between the predicted PCE values from the six models and the experimental values is shown in **Fig. 2b-c** and **Fig. S3**. To further assess robustness of the models, we additionally performed 5-fold cross-validation and report the corresponding performance estimates in **Fig. S4**. Model performance and hyperparameters were evaluated using 5-fold cross-validation on the training set. In 5-fold cross-validation, the data are partitioned into five mutually exclusive subsets of approximately equal size; in each round, one subset is used as the validation set and the remaining four as the training set. This procedure is repeated five times so that each sample serves exactly once as validation data, providing a more reliable estimate of model performance on unseen data.

4. Model Evaluation

The model performance is quantitatively evaluated using key evaluation metrics, including the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE). Among the six models, XGBoost demonstrated the best overall performance and is thus selected for subsequent prediction tasks.

R^2 is used as a metric to assess the fit of the predictive model, representing the proportion of the variance in the data that the model can explain. The calculation formula is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

RMSE quantifies the square root of the average squared differences between the predicted values and the actual values. It gives an indication of the magnitude of the prediction errors, with larger errors having a greater impact on the RMSE. RMSE is commonly used to evaluate the accuracy of a model, with lower values indicating better alignment with the actual values. The calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

MAE means the average magnitude of the errors in a set of predictions, without considering their direction, and is used to measure how closely the predictions align with the actual values. The calculation formula is as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

Where, n represents the number of observation objects, x_i and y_i is the i -th observation of x and y in equation, respectively. y_i and $\hat{y}$ stand for the tested and predicted values, $\bar{y}$ is the average values of the tested and predicted values in equations, respectively.

5. Shapley Additive Explanations (SHAP) Analysis

SHAP analysis is a technique for model interpretability based on the Shapley value theory in cooperative game theory, aiming to quantify the contribution of each feature to the model prediction [8]. SHAP calculates the marginal contribution of features under different inputs of the model and assigns a contribution value to each feature in ML. And the SHAP method can be used to decompose the model prediction of a single sample into a baseline value plus feature attributions, as shown in Equation (4):

$$\hat{f}_{(x)} = \phi_0 + \sum_{i=1}^N \phi_i x_i' = \mathbb{E}[\hat{f}(x)] + \sum_{i=1}^N \phi_i \quad (4)$$

Where, $\hat{f}_{(x)}$ is the final predicted value of sample x . ϕ_0 is the average predicted value of the training samples, that is the baseline value. ϕ_i is the SHAP value of feature i for this sample. Positive values of ϕ_i indicate that this feature increases the predicted value, while negative values

of ϕ_i indicate a decrease in the predicted value. The larger the absolute value, the stronger the contribution. N is the total number of features.

$$I_{(i)} = \frac{1}{N} \sum_{n=1}^N |\phi_i^{(n)}| \quad (5)$$

Where, $I_{(i)}$ measures how strongly feature i influences predictions on average. We compute SHAP with TreeSHAP for XGBoost, which provides consistent, polynomial-time Shapley values for tree ensembles, as shown in **Fig. S5**.

6. Molecular Recombination Rules

In this work, the reconstruction of the A-DA'D-A type NFA is the focus for new molecule generation and design. The SHAP importance ranking based on the XGBoost model selected the top 50 structural units, including 15 end acceptor unit (A), 15 conjugated donor unit (D), and 20 electron-deficient core unit (A'). When one A' unit is fixed, the D and/or A unit libraries are shared on both sides of the A-DA'D-A architecture respectively, while independent sampling is allowed to achieve the construction of symmetric and asymmetric structures. Through the chemical space combination method, a total of 1,012,500 new NFA molecules are generated after eliminating duplicates, and the arrangement and combination are shown in Equation (6):

$$T = n_A \times n_D \times n_{A'} \times n_D \times n_A \quad (6)$$

Where, T represents the total number of candidates, n_A , n_D , and $n_{A'}$ are the number of A, D, and A' important units after SHAP sorting, respectively. ($n_A = 15$, $n_D = 15$, and $n_{A'} = 20$).

This includes 4,500 new molecules with the D units and A units that are completely symmetrical about the central electron-deficient unit. We present the potential candidates with a predicted PCE greater than 18% in the form of a three-dimensional spatial distribution as shown in **Fig. S5**. Due to structural symmetry, the original five fragment libraries of A-DA'D-A is actually simplified into three dimensions of A, D, and A'. It is easy to observe that the decreasing trend of PCE decreases along the order of SHAP importance of the three fragments (A, D, and A'), and that is to say, for symmetric structures, it seems that combining the most critical structural units can produce a relatively high PCE. However, compared with asymmetric structures, the selectivity for symmetric candidates with high PCE are very limited.

Genetic algorithm (GA) is a method based on population optimization to simulate the process of biological evolution, which helps to improve the search efficiency for the combination design of A-DA'D-A type molecules^[9]. The specific process of GA is as follows:

Firstly, an initial population is constructed to preserve the inheritance of excellent genetic materials and the search for diversity. Some individuals are extracted from the parent sample pool

are directly injected into the initial population to retain the known high-quality genes (molecular units), while the remaining individuals are generated via uniformly sampling from various fragment libraries. Guided by the SHAP values, the A' units of the parent seed are set as A'.51 and A'.14, which have the potential genes for obtaining high-performance PCE.

Secondly, the fitness assessment provides a clear criterion for ranking individuals in the population in terms of their superiority or inferiority. The genotype of each individual is first decoded and mapped to the corresponding feature subset. To ensure compatibility with the chemical combination method for training the model, the system selects the top 50 features to construct the feature vector for model inference. Further, this feature vector is input into a pre-trained XGBoost model to predict PCE. The predicted value is directly defined as the fitness of the individual, following the basic principles of the GA. A higher predicted value indicates a better individual.

Thirdly, the offspring population is generated through three core genetic operations, including selection, crossover, and mutation. The selection operation aims to retain high-quality individuals. It typically employs tournament selection (for example, randomly selecting three individuals from the population and retaining the one with the highest fitness) or random sampling method from the parent pool to increase the probability of passing high-quality genes to the offspring. The crossover operation simulates gene recombination. According to the set crossover probability, the selected parent individuals are paired up in pairs, and the gene fragments are exchanged using a two-point crossover method. To meet the problem constraints, if the invalid genes exist in the individuals generated after crossover, the repair mechanism will be activated to adjust them to valid solutions. The mutation operation acts on individuals with a given probability, randomly replacing each allele at the gene locus of the individual with a certain probability from the same gene pool, thereby introducing new genetic material and maintaining population diversity.

Finally, this algorithm incorporates an elitist mechanism to ensure the convergence of the evolutionary process and the quality of the solution. This mechanism compares the newly generated offspring population with the parent population after each generation of evolution, and directly replaces the individual with the lowest fitness with the one with the highest fitness. It can effectively prevent the loss of the current optimal solution during the generation transition, thereby ensuring the monotonic improvement of the algorithm's performance. The termination condition of the algorithm is to reach the preset maximum number of evolutionary generations. Convergence detection is usually carried out based on indicators such as the fitness change of solutions and the population diversity. When the fitness of solutions no longer improves significantly, the diversity of the population decreases to a certain extent, or the preset number of iterations is reached, it can

be considered that the GA has converged. Convergence detection helps to avoid excessive iteration of GA and waste of computational resources. Here, we choose to directly consider the GA as converged when the number of iterations reaches 150 generations.

Overall, this workflow through efficient search in a vast chemical space, can systematically retain favorable structural fragments and continuously explore new combinations through recombination and mutation, ultimately successfully and rapidly exploring A-DA'D-A type candidate molecules with the target performance.

7. Quantum Chemistry (QC)

All the calculations of the model compounds in this work are performed using the Gaussian 09 software package ^[10]. The alkyl side chains are replaced with methyl groups for saving computation time without affecting the description of electronic properties. Ground state geometry optimizations of PM6, K01-K12 molecules are calculated by density functional theory at the B3LYP/6-31G (d, p) level ^[11, 12]. Then the electrostatic potential (ESP) characteristic parameters are calculated based on the optimal structure with electronic wave function information using Multiwfn 3.7(dev) program ^[13, 14]. The visualization of the molecular orbitals and ESP distribution is performed using VMD and GaussView 6.0, respectively, as shown **Figs. S9** and **S10** ^[15, 16]. The excited states of asymmetric acceptors are calculated by time-dependent density functional theory (TD-DFT) on CAM-B3LYP/def2tzvp IOp (9/40=4) level based on the geometries which are optimized at the CAM-B3LYP/def2svp level ^[17]. This function has been widely used in the majority of quantum chemical studies on non-fullerene photovoltaic acceptors, and the B3LYP in combination with a medium-sized basis set has been proven to achieve a reasonable balance between accuracy and computational cost. The lowest point of the excited state energy is determined and the corresponding rigid energy scan is shown in **Fig. S11**. Meanwhile, SMD solvation model is used for the calculation using chloroform as a solvent ^[18]. The solvent effect is modeled by CPCM, and dispersion correction is carried out using the Grimme' D3 dispersion method with Becke-Johnson damping (D3(BJ)) ^[19]. Then the electron-hole analysis is performed by Multiwfn 3.7(dev) program to visualize electron-hole distribution and obtain electron excitation characteristic parameters.

8. Synthetic Accessibility (SA) Scores

The synthesis accessibility scores (SA_Scores) of all NFAs are calculated based on the Ertl-Schuffenhauer method in the RDKit software. This score integrates the contribution based on fragment frequency and the penalty term for structural complexity. The penalty term covers factors

such as molecular size, the number of chiral centers, fused, bridged or spiro ring systems, macrocycles structures, and heteroatoms diversity ^[20]. The score results fall within a continuous range of 1 to 10, which is used to approximately assess the relative synthesis difficulty of the molecules, with an increase in the score from 1 to 10 indicating a gradual increase in the synthetic difficulty. Before scoring, all molecular structures are standardized and preprocessed to generate standardized SMILES codes. This work presents the SA scores of each candidate molecule and uses them for qualitative comparison to evaluate their relative synthesis feasibility. This indicator should be regarded as a supplement to the analysis of synthesis route planning, rather than a replacement.

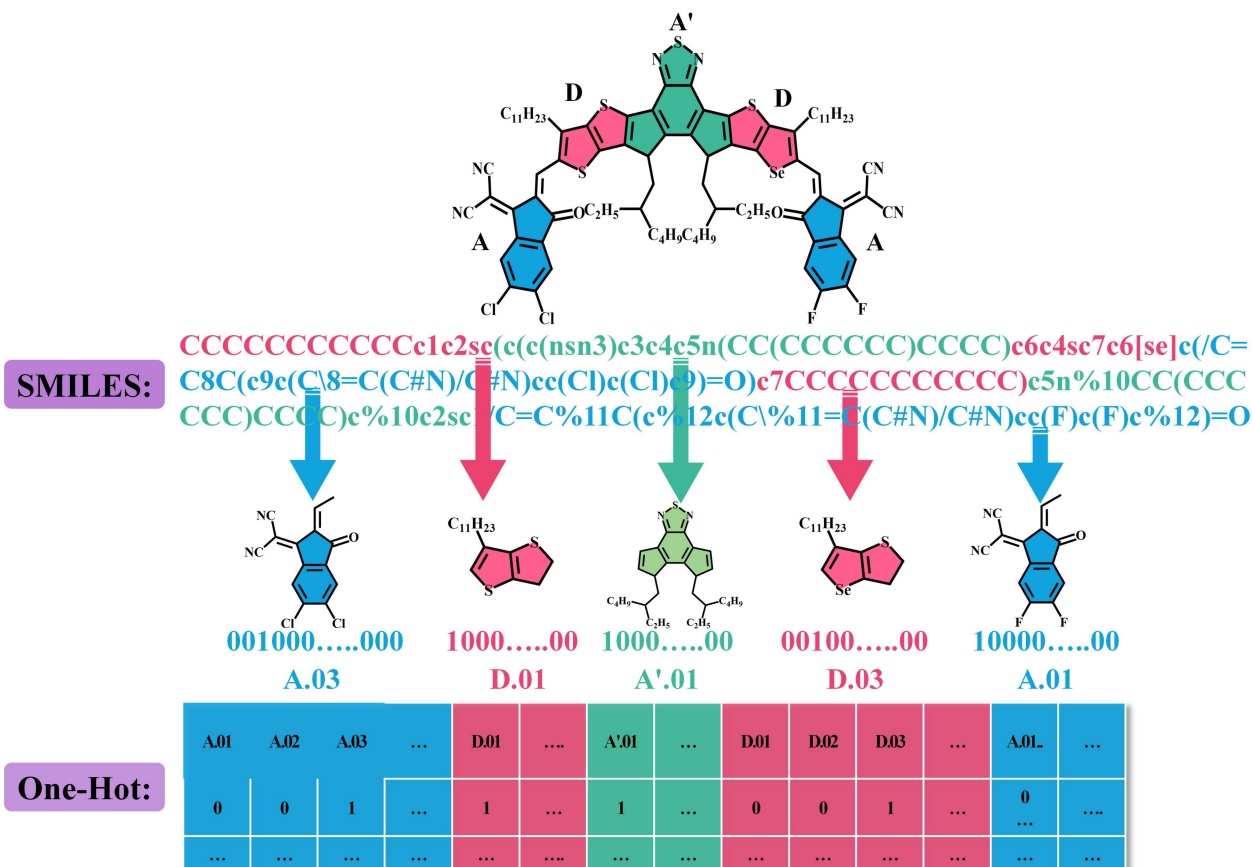


Figure S1. The unit splitting of NFA and the corresponding one-hot encoding process, that is, converting the SMILES molecular codes into binary encoding ("0" and "1").

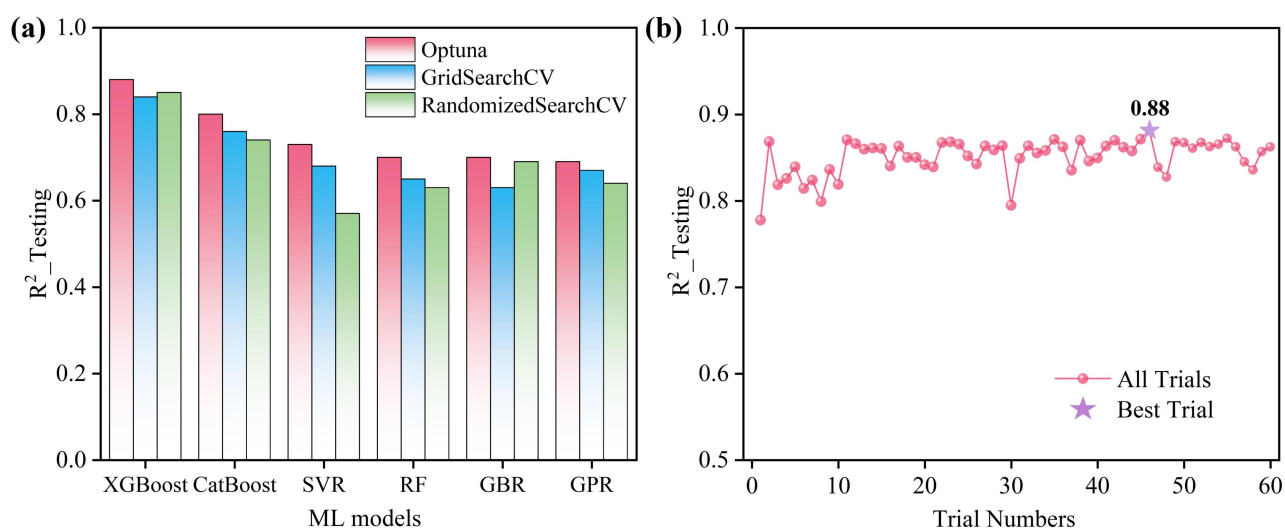


Figure S2. (a) The R^2 comparison of three different hyperparameter optimization methods (Optuna, GridSearch CV, and RandomizedSearch CV) on six ML model testing sets, respectively. (b) The stability of R^2 values in the testing set of the optimal XGBoost model under 60 rounds of hyperparameter optimization via Optuna hyperparameter optimization methods.

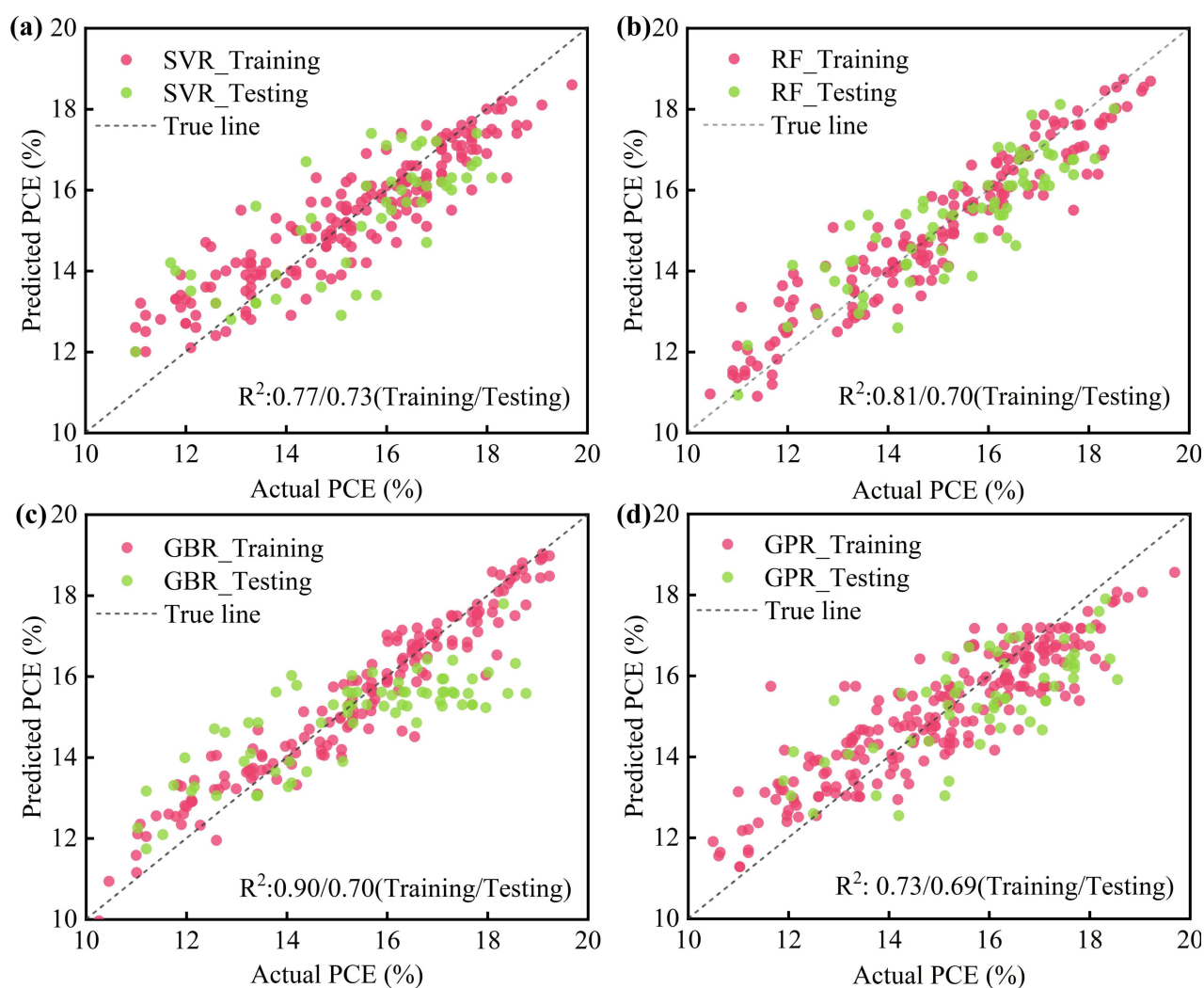


Figure S3. The fitting relationship between the PCE values predicted by the (a) SVR, (b) RF, (c) GBR, and (d) GPR models and the values reported in the literature. The gray dashed line indicates the ideal reference line of $y = x$.

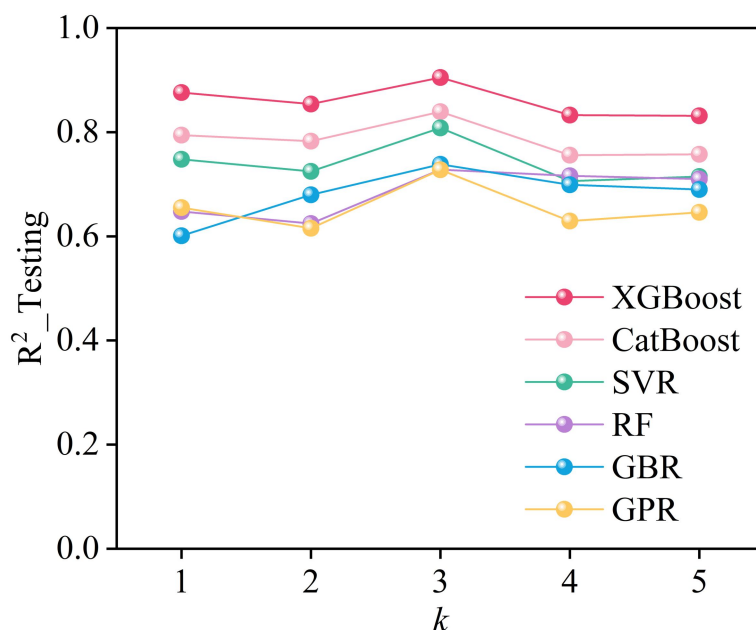


Figure S4. The R^2 values of the six models during the 5-fold cross-validation process, where k represents the number of folds.

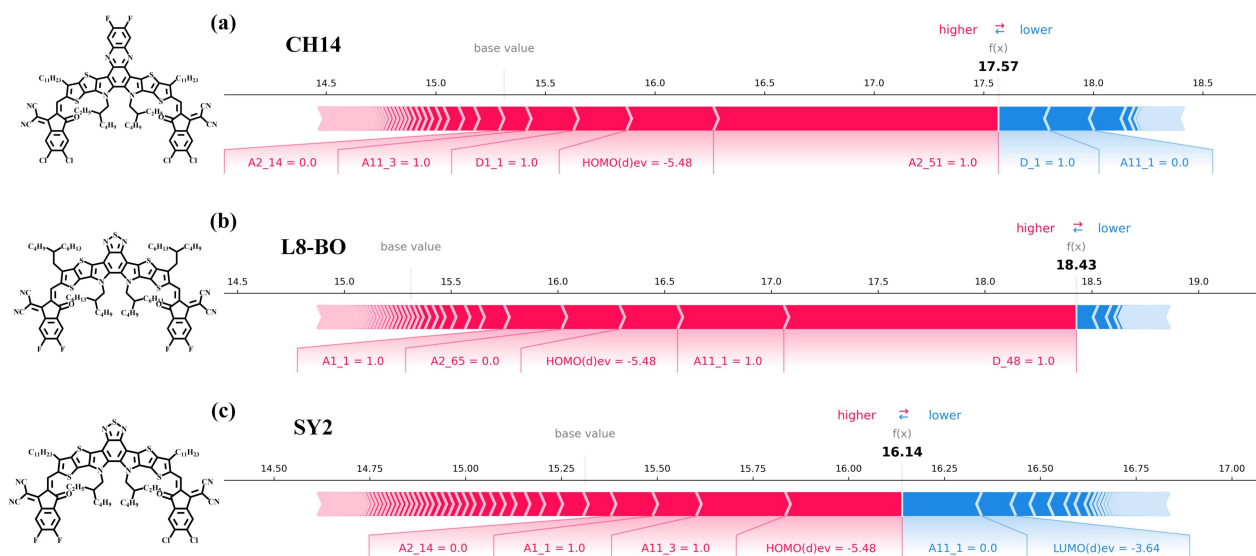


Figure S5. The SHAP interpretation force plot for the predicted PCE of the (a) PM6:CH14, (b) PM6:L8-BO, and (c) PM6:SY2 samples, where red (blue) bars indicate the increase (decrease) effect on the predicted target of PCE.

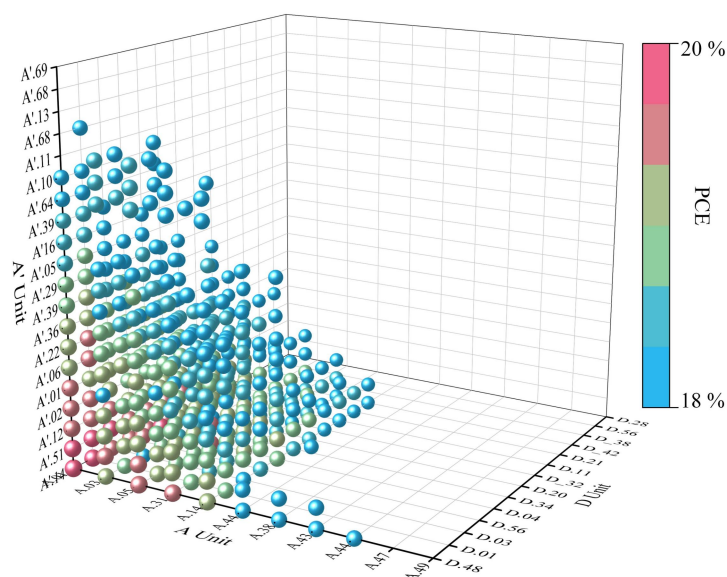


Figure S6. The three-dimensional distribution of the chemical composition space of 549 candidates with a PCE greater than 18% in PM6 based OSCs among 4500 virtually generated symmetric NFAs.

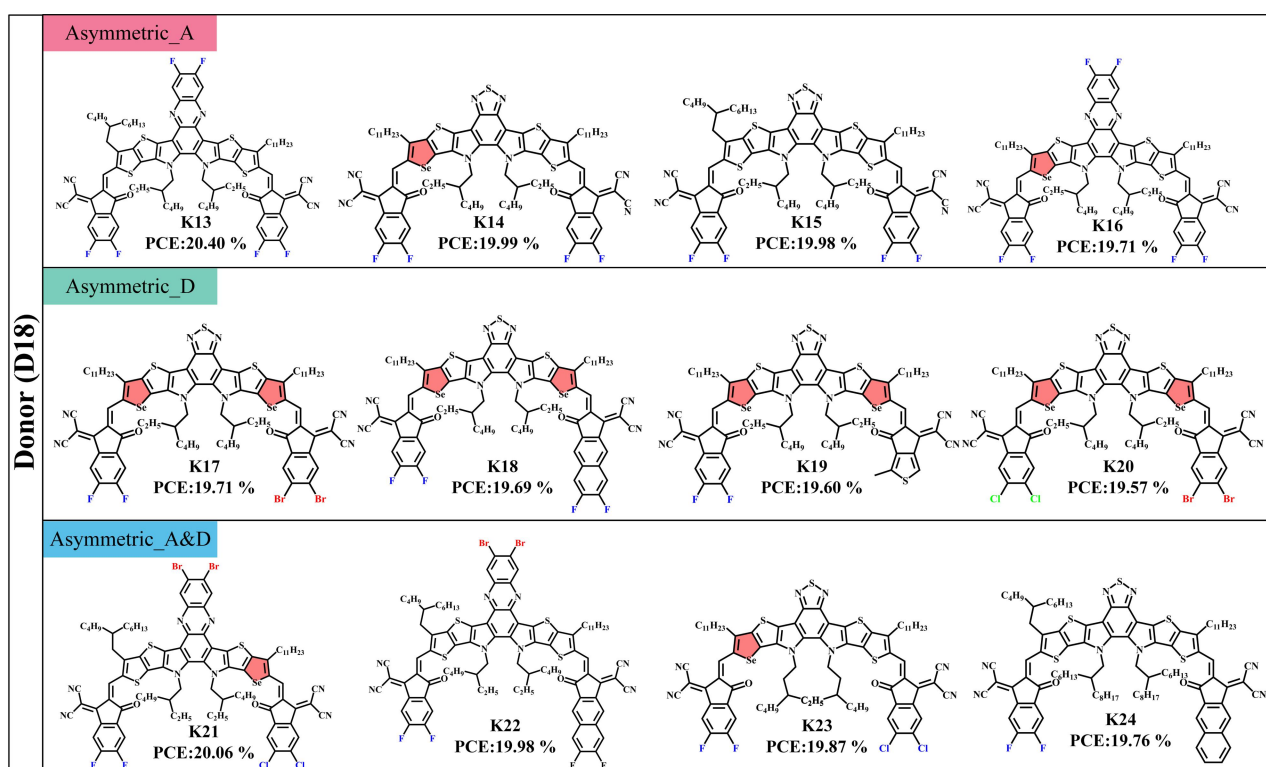


Figure S7. Twelve potential NFAs molecular structure for the D18 based OSCs, including the Asymmetric_A type (K13-K16), Asymmetric_D type (K17-K20), and Asymmetric_A&D type (K21-K24).

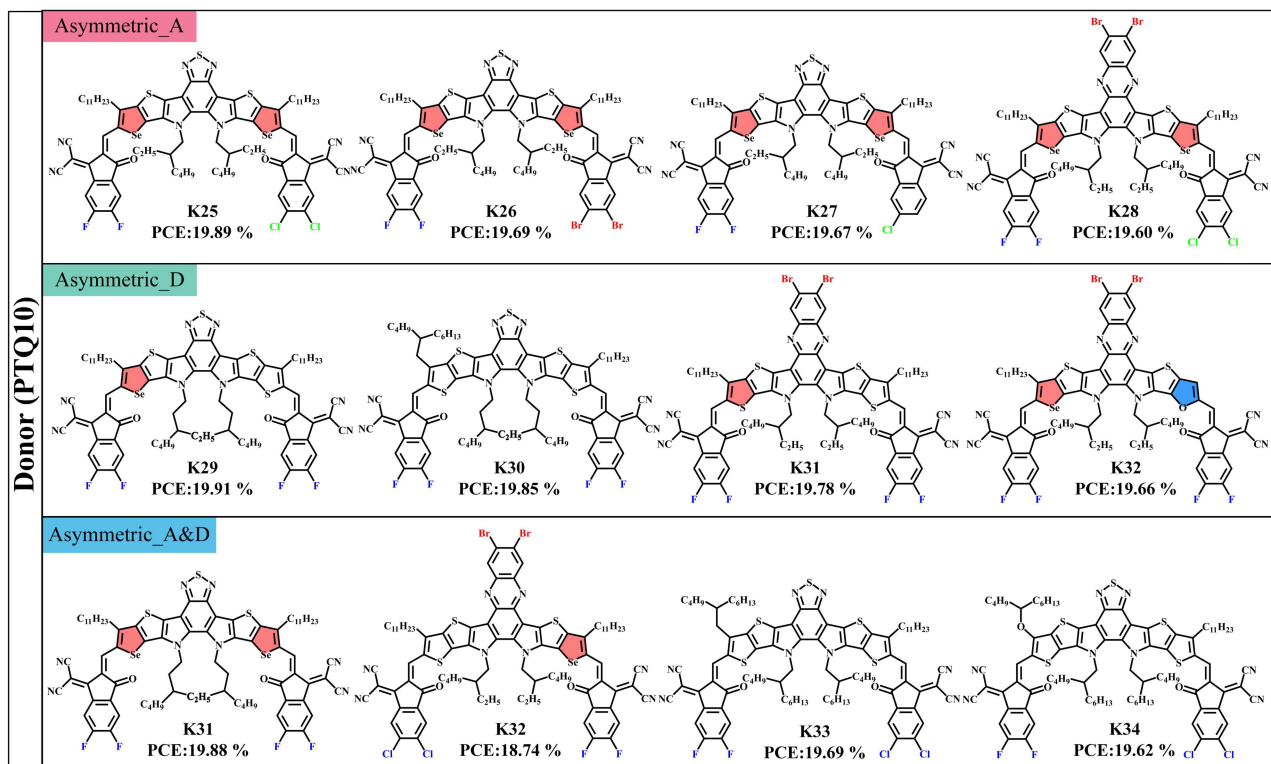


Figure S8. Twelve potential NFAs molecular structure for the PTQ10 based OSCs, including the Asymmetric_A type (K25-K28), Asymmetric_D type (K29-K32), and Asymmetric_A&D type (K33-K36).

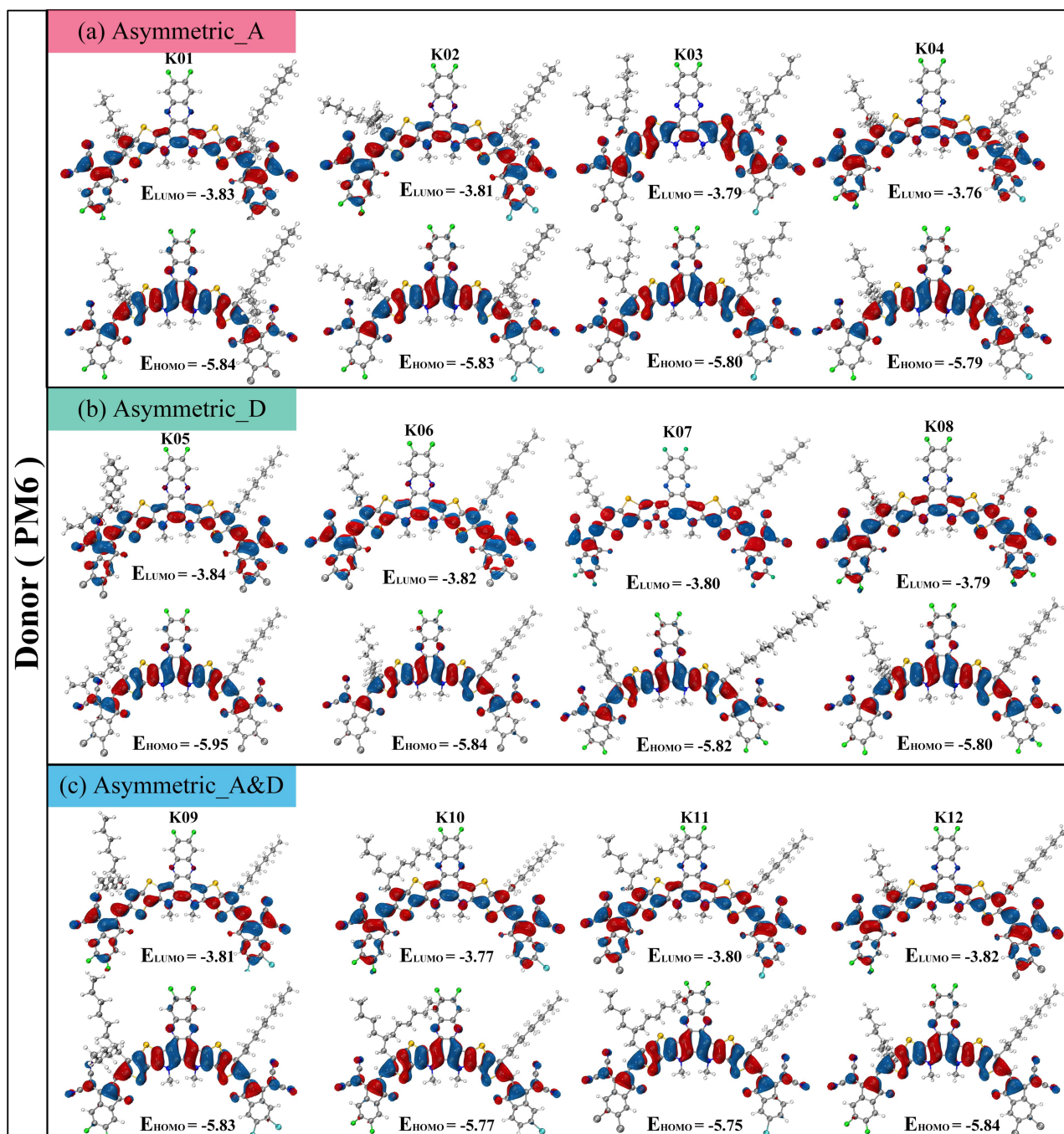


Figure S9. Calculated HOMO and LUMO orbital distributions of the potential NFA candidate molecules (K01-K12).

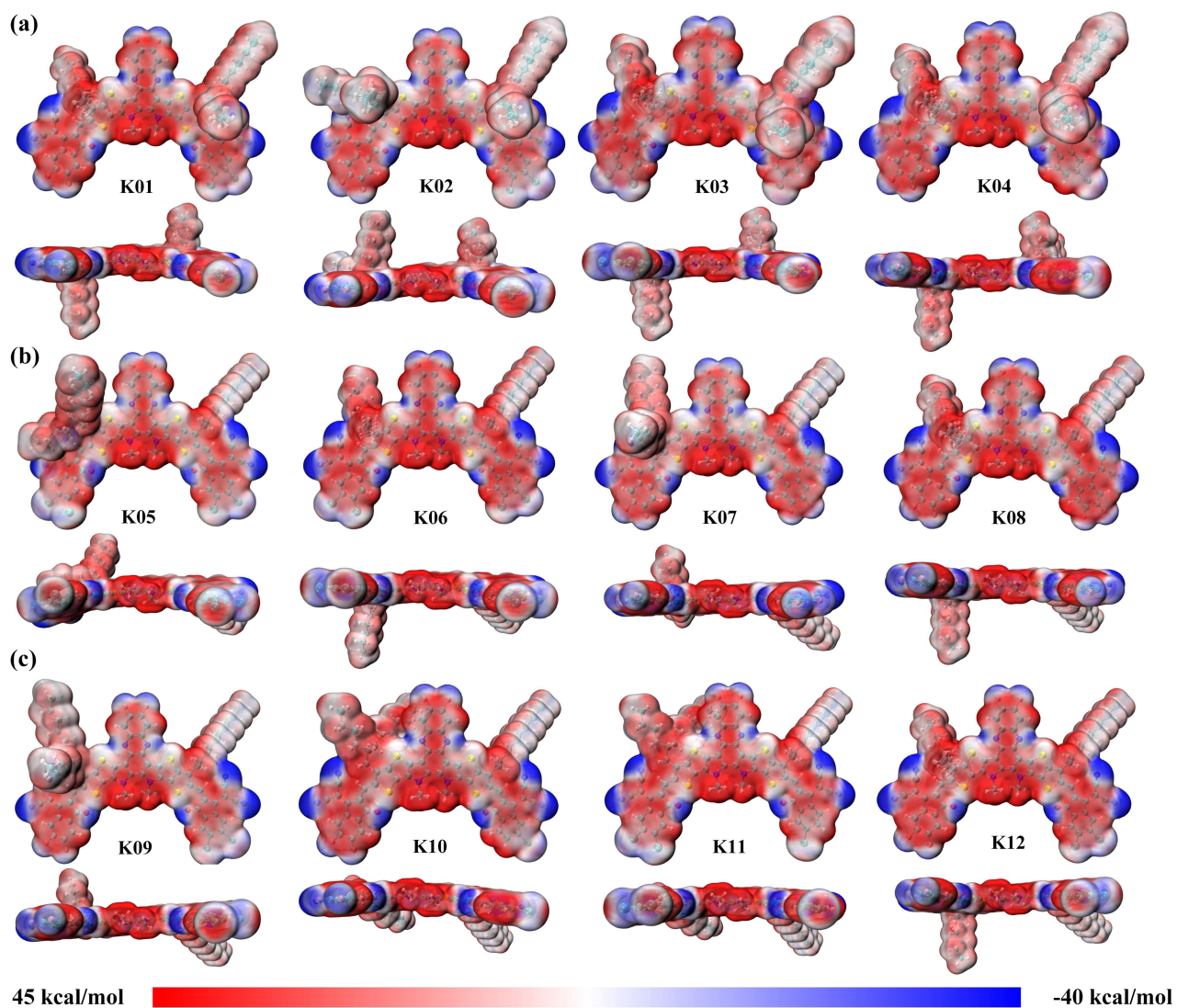


Figure S10. The surface ESP of the 12 potential NFA candidate molecules (K01-K12), where blue represents negative potential energy and red indicates positive potential energy.

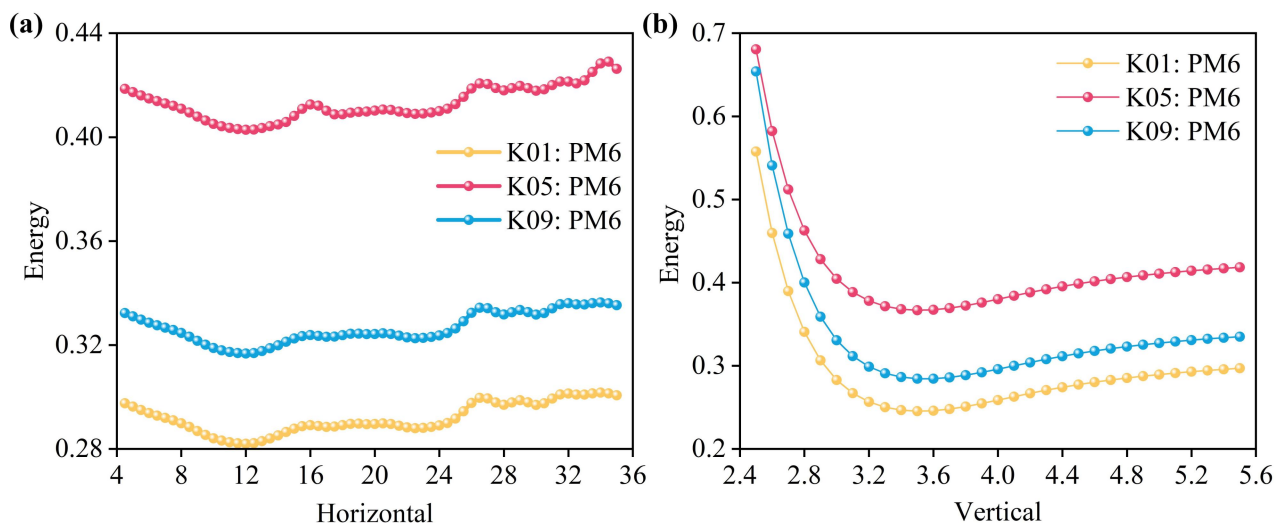


Figure S11. The rigid scanning of the PM6:K01, PM6:K05, and PM6:K09 combination in (a) horizontal and (b) vertical directions.

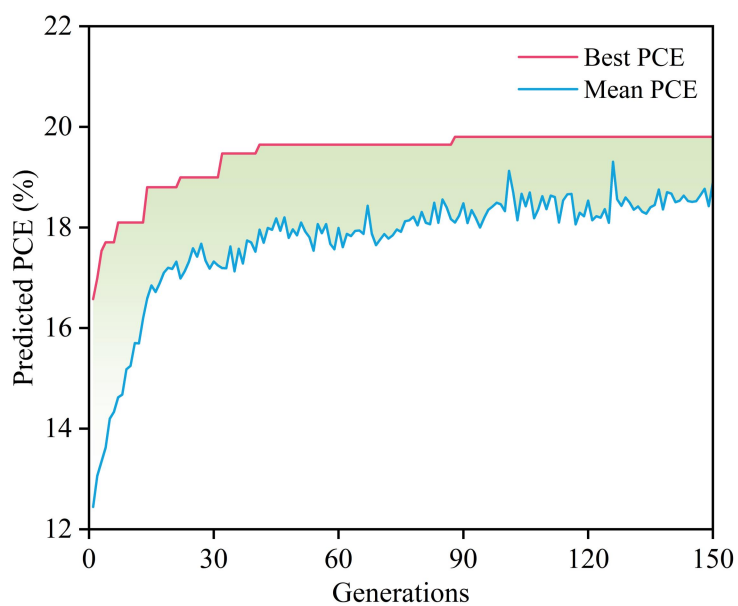


Figure S12. The evolutionary curve of PCE with the reproduction of offspring in GA.

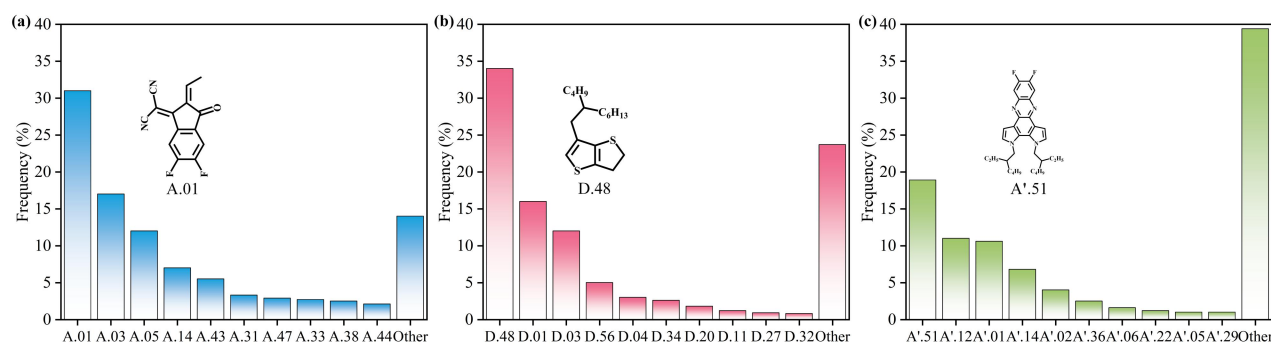


Figure S13. The frequency of the top 10 genes/fragments in each of the three types of sub-structures (A, D, and A') during the 150 reproduction process in GA.

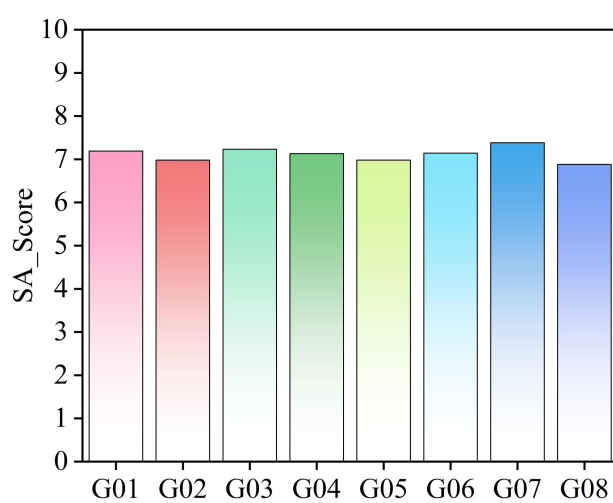
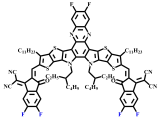
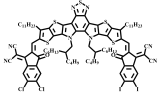
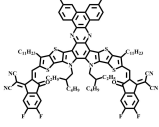
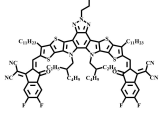
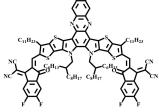
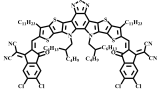
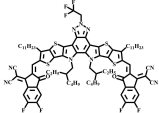
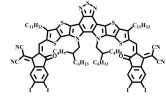
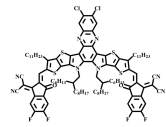
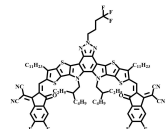
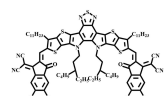
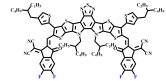
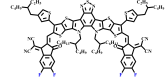
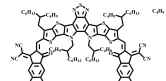
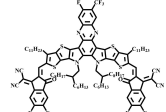
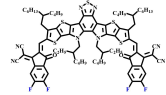
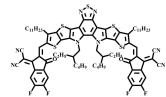
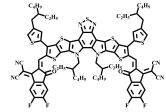
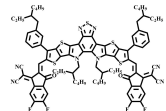
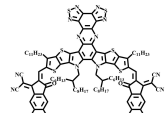
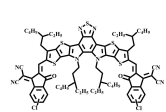
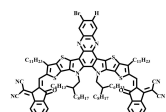
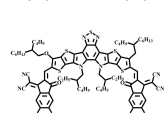
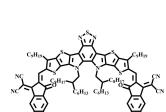
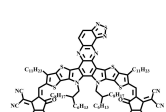
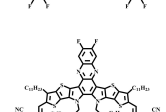


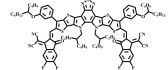
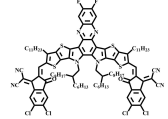
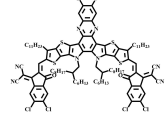
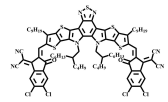
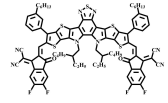
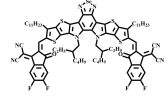
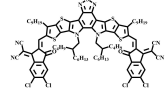
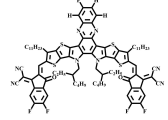
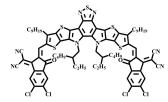
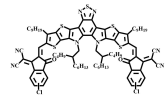
Figure S14. The SA_Scores of G01-G08 molecules.

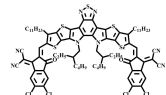
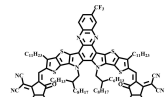
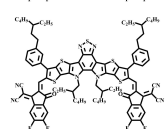
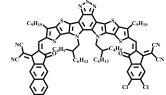
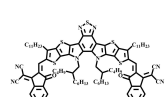
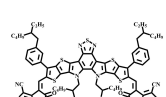
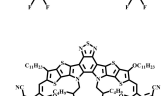
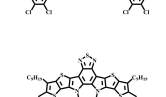
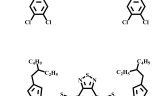
Table S1. The OSCs data set including donor, acceptor, LUMO and HOMO of corresponding donor, and device performance parameters (V_{oc} , J_{sc} , FF, and PCE).

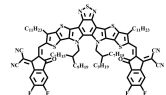
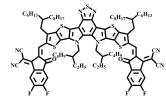
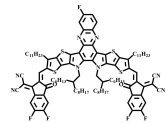
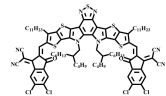
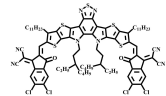
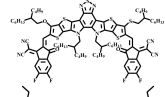
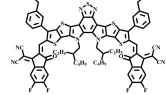
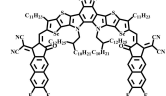
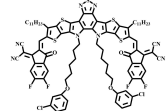
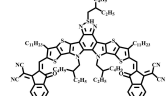
Acceptors	Donors	LUMO (D)	HOMO (D)	V_{oc} (V)	J_{sc} (mA cm ⁻²)	FF (%)	PCE (%)	Structures	References
AQx-2F	D18	-2.77	-5.51	0.937	26.10	80.4	19.7		10.1016/j.joule.2024.01.005
BOC1-I	PM6	-3.64	-5.48	0.872	28.7	78.3	19.6		10.1021/acsenergylett.4c03168
AQx-22	D18	-2.77	-5.51	0.97	25.8	78	19.5		10.1002/adma.202413376
Y2CF3	D18	-2.77	-5.51	0.862	27.64	80.09	19.08		10.1021/jacs.4c13471
CH22	PM6	-3.64	-5.48	0.884	26.74	80.62	19.06		10.1038/s41467-023-40423-6
BO4C1	PM6	-3.64	-5.48	0.87	25.72	81.5	18.32		10.1021/acsenergylett.4c03168
Y1CF3	D18	-2.77	-5.51	0.866	27.31	80.08	18.96		10.1021/jacs.4c13471

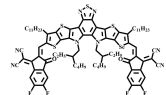
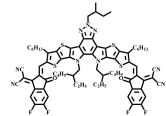
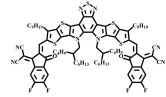
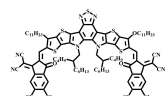
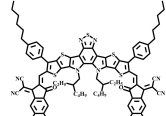
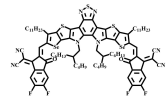
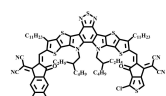
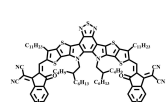
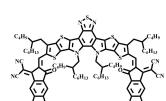
BO4I	PM6	-3.64	-5.48	0.845	29.1	77.1	18.96		10.1021/acsenergylett.4c03168
CH23	PM6	-3.64	-5.48	0.876	26.64	80.45	18.77		10.1002/adfm.202301573
Y3CF3	D18	-2.77	-5.51	0.854	27.71	78.65	18.61		10.1021/jacs.4c13471
N3	D18	-2.77	-5.51	0.86	27.44	78.5	18.56		10.1088/1674-4926/42/1/010502
T2EH	PTQ10	-2.98	-5.54	0.872	26.57	80.05	18.55		10.31635/ccschem.022.202202056
m-TEH	PTQ10	-2.98	-5.54	0.88	26.61	79	18.51		10.1002/sus2.82
Se46	PM6	-3.64	-5.48	0.879	26.22	80.1	18.46		10.1021/acsenergylett.3c00743
CH-FCF	PM6	-3.64	-5.48	0.894	25.21	78.86	18.41		10.1016/j.cej.2023.147091
L8-BO	PM6	-3.64	-5.48	0.905	26.6	76.1	18.32		10.1038/s41560-021-00820-x
AsymSSe-2F	D18	-2.77	-5.51	0.84	27.44	79.46	18.31		10.1002/aenm.202301292

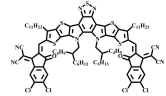
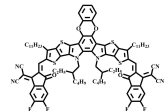
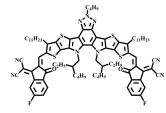
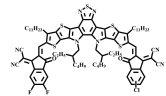
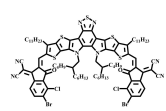
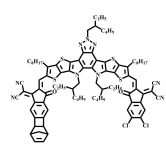
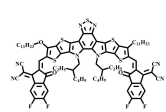
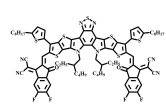
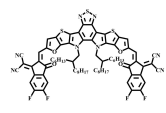
BTP-4F-T3EH	D18-Cl	-2.78	-5.56	0.873	26.84	77.85	18.25		10.1007/s11426-022-1290-y
BTP-4F-P2EH	PM6	-3.64	-5.48	0.88	25.85	80.08	18.22		10.1002/aenm.202102596
CH-BBD	PM6	-3.64	-5.48	0.882	26.15	78.9	18.19		10.1002/anie.202308832
G6-EHep	PM6	-3.64	-5.48	0.85	27.6	77.67	18.13		10.1007/s11426-022-1451-2
CH21	PM6	-3.64	-5.48	0.873	26.57	78.13	18.12		10.1038/s41467-023-40423-6
Y6-1OBO	PM6	-3.64	-5.48	0.93	25.13	77.1	18.02		10.1021/acs.chemmater.2c02851
BTP-4Cl-C9-16	PM6	-3.64	-5.48	0.85	27.3	77.8	18		10.1021/acsmaterialslett.2c00764
CH-iBQ	PM6	-3.64	-5.48	0.879	26.04	78.5	17.97		10.1002/anie.202308832
CH17	PM6	-3.64	-5.48	0.883	26.19	77.2	17.84		10.1007/s11426-022-1264-y

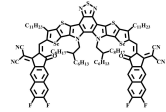
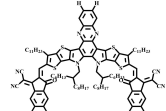
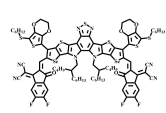
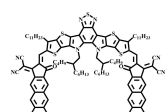
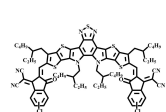
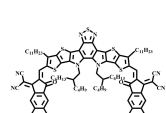
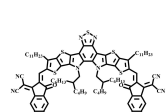
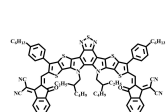
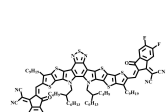
CY-O	PM6	-3.64	-5.48	0.863	26.21	78.81	17.83		10.1016/j.xcrp.2023.101303
BTzC4IC-2F-2	PFBT4T-T	-3.69	-5.34	0.783	28.59	76.95	17.82		10.1016/j.cej.2023.141281
	20								
CH14	PM6	-3.64	-5.48	0.869	26.35	77.79	17.81		10.1002/adfm.202301573
BTP-e9	PBDB-T-F	-3.5	-5.5	0.839	26.67	81.1	17.8		10.1002/nano.202100036
m-BTP-PhC6	PTQ10	-2.98	-5.54	0.883	25.3	79.3	17.7		10.1039/D0EE03506H
Y6-Se	D18	-2.77	-5.51	0.839	27.98	75.3	17.7		10.1021/jacs.0c08557
eC9	PM6	-3.64	-5.48	0.837	26.5	78.9	17.7		10.1021/acsmaterialslett.1c00328
AQx-1F	D18	-2.77	-5.51	0.947	24.3	76.9	17.7		10.1016/j.joule.2024.01.005
BTP-4Cl	PM6	-3.64	-5.48	0.838	26.7	79	17.7		10.1002/anie.202102622
eC9-2Cl	PBQx-TF	-3.58	-5.48	0.868	25.9	78.6	17.7		10.1002/adma.202102420

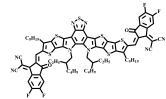
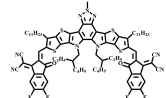
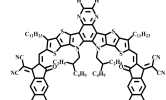
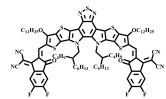
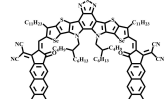
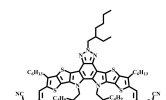
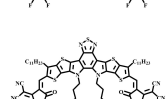
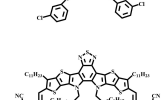
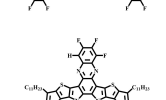
AsymSSe-2Cl	D18	-2.77	-5.51	0.841	27.56	76.24	17.68		10.1002/aenm.202301292
CH-CF	PM6	-3.64	-5.48	0.886	25.29	76.3	17.62		10.1016/j.cej.2023.147091
BTP-4F-P3EH	PM6	-3.64	-5.48	0.861	26.11	78.13	17.57		10.1002/aenm.202102596
BTP-S9	PM6	-3.64	-5.48	0.846	26.47	78.44	17.56		10.1038/s41467-021-24937-5
Y-BO-FC1	PM6	-3.64	-5.48	0.85	26.58	77.5	17.51		10.1002/anie.202104766
P2EH	PTQ10	-2.98	-5.54	0.892	24.91	78.75	17.5		10.31635/ccschem.022.202202056
BTP-4Cl-C9-12	PM6	-3.64	-5.48	0.84	26.9	77.2	17.5		10.1021/acsmaterialslett.2c00764
BTP-eC9-HD	PM6	-3.64	-5.48	0.84	26.59	78.32	17.49		10.1002/solr.202300206
BTP-4F-T2EH	D18-Cl	-2.78	-5.56	0.891	25.68	76.11	17.41		10.1007/s11426-022-1290-y

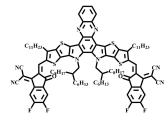
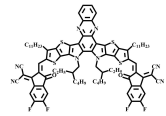
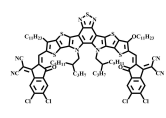
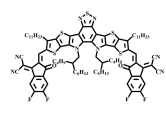
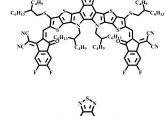
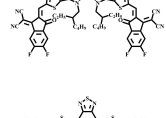
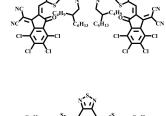
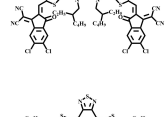
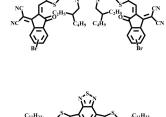
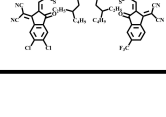
Y6-HU	PM6	-3.64	-5.48	0.87	25.6	77.9	17.4		10.1016/j.nanoen.2022.107574
L8-HD	PM6	-3.64	-5.48	0.88	25.08	78.8	17.39		10.1038/s41560-021-00820-x
CH-F	PM6	-3.64	-5.48	0.875	25.24	75.59	17.34		10.1016/j.cej.2023.147091
BTP-BO-4Cl	PM6	-3.64	-5.48	0.854	26.1	77.7	17.33		10.1039/D0EE02034F
BT3-4Cl	D18	-2.77	-5.51	0.853	26.46	76.69	17.31		10.1002/solr.202100522
BTP-O-S	PM6	-3.64	-5.48	0.912	23.74	77.5	17.3		10.1002/adfm.202213429
BTP-4F-PC6	PM6	-3.64	-5.48	0.855	25.08	80.33	17.22		10.1002/aenm.202102596
YDT-SeNF	PM6	-3.64	-5.48	0.836	27.2	75.4	17.2		10.1007/s40820-023-01208-0
A4	PBDT-Cl	-3.64	-5.53	0.84	26.9	0.76	17.19		10.1002/adfm.202201150
Y11-EB	PM6	-3.64	-5.48	0.88	26.2	74.73	17.15		10.1016/j.jechem.2021.05.041

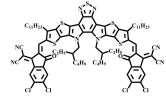
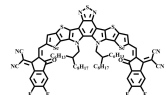
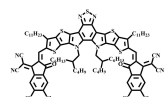
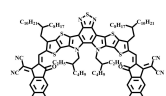
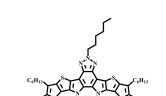
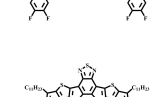
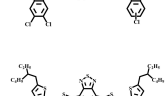
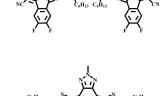
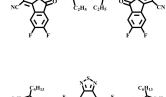
Y6	PTQ10	-2.98	-5.54	0.831	26.28	78.4	17.13		10.1007/s11426-021-1114-3
BTA-C5	PM6	-3.64	-5.48	0.847	26.51	76.19	17.11		10.1021/acsami.1c07254
BT4T-4F	PM6	-3.64	-5.48	0.839	26.3	77.7	17.1		10.1002/aenm.202003177
BTP10-4Cl-C12	PM6	-3.64	-5.48	0.91	23.85	78.8	17.1		10.1002/aenm.202003777
BTP-Ph	PTQ10	-2.98	-5.54	0.888	24.2	77.9	17.1		10.1002/aenm.202100079
CH1007	PM6	-3.64	-5.48	0.822	27.48	75.61	17.08		10.1021/jacs.0c07083
BTP-2F-ThCl	PM6	-3.64	-5.48	0.869	25.38	77.4	17.06		10.1016/j.joule.2020.03.023
G6-BO	PM6	-3.64	-5.48	0.91	24.22	77.25	17.06		10.1007/s11426-022-1451-2
BTP-IIC-BO-M	PM6	-3.64	-5.48	0.87	25.35	77.63	17.03		10.1021/acsami.2c22972

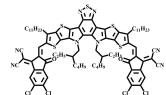
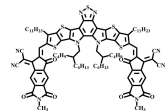
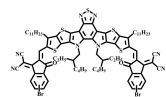
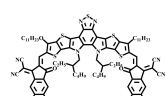
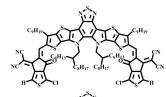
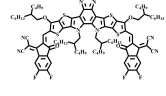
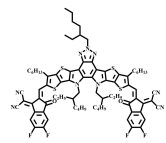
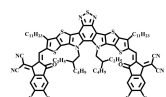
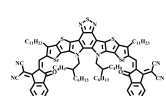
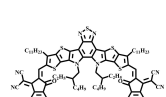
BTP-4Cl-12	PM6	-3.64	-5.48	0.858	25.6	77.6	17		10.1093/nsr/nwz200
BDOTP-1	D18-B	-2.71	-5.51	0.938	24.92	72.4	16.93		10.1002/cey2.250
BTzC4IC-2F-1	PFBT4T-T 20	-3.69	-5.34	0.81	26.84	75.82	16.91		10.1016/j.cej.2023.141281
SY1	PM6	-3.64	-5.48	0.871	25.41	76	16.83		10.1002/adfm.202000456
BTP-ClBr	PM6	-3.64	-5.48	0.906	23.48	79	16.82		10.1002/aenm.202002649
LL3	PBDB-T	-3.51	-5.34	0.86	26.97	72.27	16.82		10.1002/adfm.202103445
Y6-1O	PM6	-3.64	-5.48	0.9	25.51	73.5	16.81		10.1021/acs.chemmater.2c02851
BTP-Th	PTQ10	-2.98	-5.54	0.874	25.1	76.4	16.8		10.2139/ssrn.4013480
BPT-4F	SZ5	-3.51	-5.34	0.853	24.8	79.1	16.8		10.1021/acsenergylett.0c01688

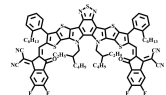
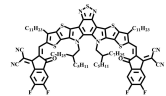
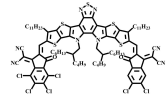
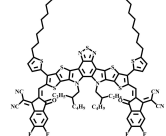
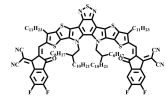
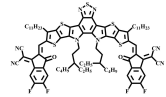
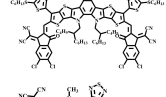
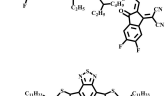
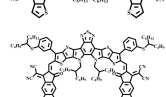
YHD-SeNF	PM6	-3.64	-5.48	0.848	26.4	75.3	16.8		10.1007/s40820-023-01208-0
CH20	PM6	-3.64	-5.48	0.881	25.44	74.92	16.79		10.1038/s41467-023-40423-6
BTP-EDOT-4F	PM6	-3.64	-5.48	0.87	2522	76.47	16.78		10.1002/adfm.202212290
BTP-S7	PM6	-3.64	-5.48	0.861	25.92	75.09	16.76		10.1038/s41467-021-24937-5
G6-EH	PM6	-3.64	-5.48	0.89	24.33	77.13	16.74		10.1007/s11426-022-1451-2
BO-4F	PM6	-3.64	-5.48	0.833	26.04	77.2	16.73		10.1038/s41467-022-30225-7
S-YSS-Cl	PM6	-3.64	-5.48	0.86	25.85	75.58	16.73		10.1002/anie.202104766
BTP-PhC6	PM6	-3.64	-5.48	0.865	25	77	16.7		10.1016/j.nanoen.2020.105087
BT5T-4F	PM6	-3.64	-5.48	0.85	25.3	77.8	16.7		10.1002/anie.201915030

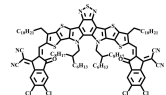
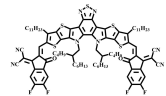
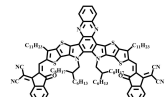
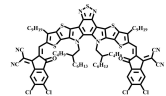
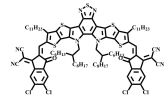
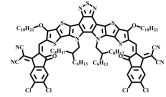
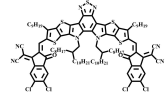
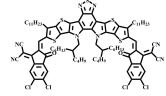
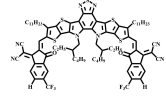
BP5T-4F	PM6	-3.64	-5.48	0.888	24.6	76.3	16.7		10.1002/aenm.202003177
Y11-M	PM6	-3.64	-5.48	0.86	25.54	76.15	16.64		10.1016/j.jechem.2021.05.041
AQx-2	PM6	-3.64	-5.48	0.86	25.38	76.25	16.64		10.1002/adma.201906324
Y6-O	PM6	-3.64	-5.48	0.95	22.4	78	16.6		10.1039/D0TA10953C
YBO-SeNF	PM6	-3.64	-5.48	0.824	27.1	74.5	16.6		10.1007/s40820-023-01208-0
BTA-C8	PM6	-3.64	-5.48	0.837	26.18	75.67	16.59		10.1021/acsami.1c07254
A3	PBDT-Cl	-3.64	-5.53	0.84	25.88	0.76	16.55		10.1002/adfm.202201150
Y6-HD	PM6	-3.64	-5.48	0.85	25.4	76.5	16.5		10.1016/j.nanoen.2022.107574
AQx-3F	D18	-2.77	-5.51	0.946	23.3	75	16.5		10.1016/j.joule.2024.01.005

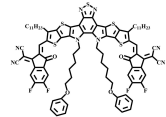
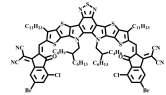
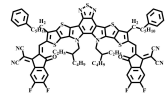
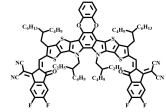
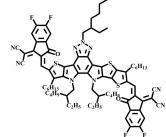
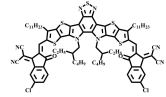
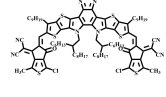
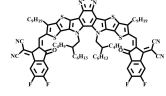
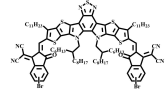
PZIC-4F	PM6	-3.64	-5.48	0.88	26.53	70.59	16.48		10.1016/j.jechem.2023.09.035
AQx-0F	D18	-2.77	-5.51	0.961	23.2	73.5	16.4		10.1016/j.joule.2024.01.005
BTP10-4Cl-C10	PM6	-3.64	-5.48	0.9	23.42	77.6	16.4		10.1002/aenm.202003777
BTP-4F-12	PM6	-3.64	-5.48	0.855	25.3	76	16.4		10.1002/adma.201903441
BTP-2S	PM6	-3.64	-5.48	0.851	24.86	74.3	16.4		10.1002/adfm.202213429
BTP-4F	PM6	-3.64	-5.48	0.841	25.19	77.3	16.37		10.1016/j.joule.2020.03.023
BTP-S2	PM6	-3.64	-5.48	0.945	24.07	72.02	16.37		10.1002/adma.202001160
BT2-4Cl	D18	-2.77	-5.51	0.841	25.73	75.48	16.35		10.1002/solr.202100522
TPT10	PTQ11	-2.76	-5.52	0.88	24.79	74.8	16.32		10.1021/jacs.9b09939
BTIC-2Cl-yCF3	PM6	-3.64	-5.48	0.84	25.09	76.99	16.31		10.1002/ange.202013053

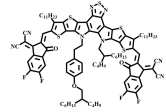
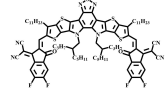
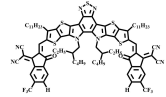
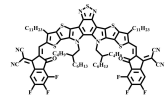
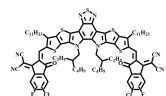
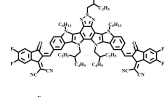
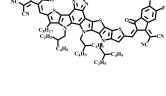
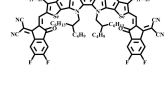
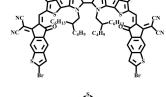
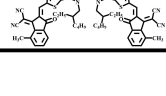
BTP-4Cl-8	PM6	-3.64	-5.48	0.872	25.2	74.3	16.3		10.1093/nsr/nwz200
BPS-4F	SZ5	-2.92	-5.50	0.822	25.4	77.9	16.3		10.1021/acsenenergylett.0c01688
BTP-4Cl-C12	PM6	-3.64	-5.48	0.85	25.73	74.6	16.3		10.1002/aenm.202003777
L8-OD	PM6	-3.64	-5.48	0.89	24.57	74.6	16.26		10.1038/s41560-021-00820-x
BTA-C6	PM6	-3.64	-5.48	0.851	25.2	75.68	16.23		10.1021/acsami.1c07254
SY3	PM6	-3.64	-5.48	0.858	25.54	74.1	16.23		10.1002/adfm.202000456
o-TEH	PBQ6	-3.18	-5.64	0.88	26.61	79	16.22		10.1002/sus2.82
BTA-C1	PM6	-3.64	-5.48	0.844	24.75	77.62	16.21		10.1021/acsami.1c07254
Y6-O2BO	PM6	-3.64	-5.48	0.963	21.5	78.1	16.2		10.1021/acs.chemmater.1c03311

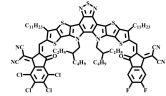
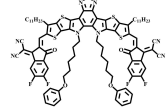
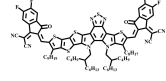
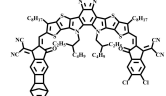
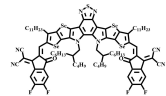
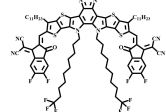
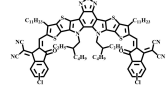
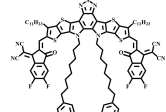
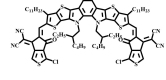
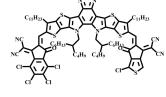
Y6-2Cl	PM6	-3.64	-5.48	0.859	25.2	74.7	16.17		10.1002/adfm.202008767
BTP-IIC-M	PM6	-3.64	-5.48	0.81	26.4	75.42	16.13		10.1021/acsami.2c22972
BTIC-2Br-m	PM6	-3.64	-5.48	0.88	25.03	73.13	16.11		10.1002/advs.201903784
Y6-1O	PM6	-3.64	-5.48	0.89	23.2	78.3	16.1		10.1002/aenm.202003141
BTP-ThCl	PM6	-3.64	-5.48	0.889	24.2	72.8	16.1		10.1002/adfm.202305765
BTP-2O	PM6	-3.64	-5.48	0.965	21.14	75.6	16.1		10.1002/adfm.202213429
Y18	PM6	-3.64	-5.48	0.84	24.91	76.4	16.02		10.1039/D0TA01260B
BTPTT-4F	P2F-Ehp	-3.13	-5.46	0.81	26.68	74.11	16.02		10.1007/s11426-019-9457-5
S-WSeSe-Cl	PM6	-3.64	-5.48	0.828	26.35	73.41	16.01		10.1002/anie.202104766
SY2	PM6	-3.64	-5.48	0.852	25.29	74.3	16.01		10.1002/adfm.202000456

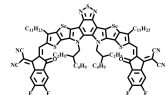
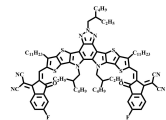
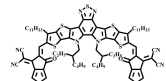
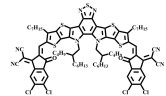
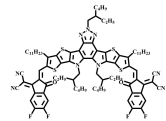
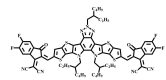
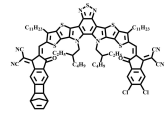
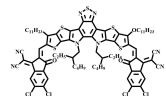
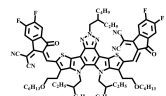
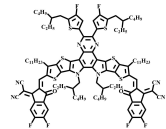
o-BTP-PhC6	PTQ10	-2.98	-5.54	0.924	22.8	76.2	16		10.1039/D0EE03506H
Y6-PN	PM6	-3.64	-5.48	0.84	25.2	75.4	16		10.1016/j.nanoen.2022.107574
BO-6Cl	PM6	-3.64	-5.48	0.944	23.22	72.9	15.94		10.1038/s41467-022-30225-7
BTP-4F-T2C8	D18-Cl	-2.78	-5.56	0.88	24.56	73.59	15.92		10.1007/s11426-022-1290-y
Y6-OD	PM6	-3.64	-5.48	0.87	24.3	75.2	15.9		10.1016/j.nanoen.2022.107574
Y6-C2	PM6	-3.64	-5.48	0.86	25.11	73.6	15.89		10.1007/s11426-019-9670-2
BTP-EDOT-4Cl	PM6	-3.64	-5.48	0.86	24.14	76.41	15.87		10.1002/adfm.202212290
BTN-4F	PM6	-3.64	-5.48	0.816	25.05	77.3	15.82		10.1002/sstr.202000052
BTP-CC	PM6	-3.64	-5.48	0.823	26.05	73.49	15.76		10.1021/acsami.2c07883
CY-S	PM6	-3.64	-5.48	0.837	24.91	75.34	15.71		10.1016/j.xcrp.2023.101303

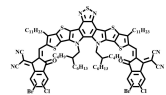
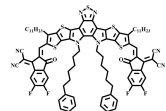
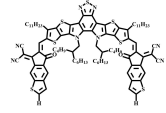
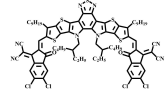
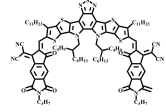
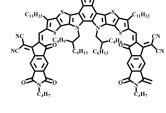
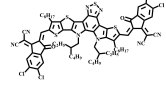
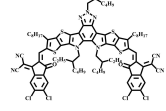
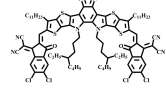
HD-4Cl	PM6	-3.64	-5.48	0.854	25.4	72.3	15.7		10.1039/D0QM00633E
Y6-BO	PM6	-3.64	-5.48	0.84	15.1	74.5	15.7		10.1016/j.nanoen.2022.107574
PZIC-4Cl	PM6	-3.64	-5.48	0.871	25.95	69.46	15.69		10.1016/j.jechem.2023.09.035
BTP-eC9	PM6	-3.64	-5.48	0.832	24.98	75.44	15.68		10.1002/adma.202206566
BTP-4Cl-16	PM6	-3.64	-5.48	0.862	24.2	74.8	15.6		10.1093/nsr/nwz200
HDO-4Cl	PM6	-3.64	-5.48	0.94	21.9	76	15.6		10.1039/D0QM00633E
BTP-4Cl-C9-20	PM6	-3.64	-5.48	0.81	26.6	72.2	15.6		10.1021/acsmaterialslett.2c00764
BTP2O-4Cl-C1	PM6	-3.64	-5.48	0.96	21.14	77.2	15.6		10.1002/aenm.202003777
BTIC-CF3-y	PM6	-3.64	-5.48	0.85	25.19	72.82	15.59		10.1016/j.joule.2020.02.004

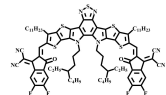
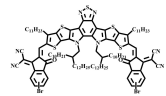
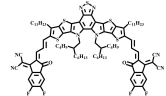
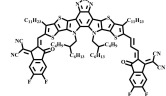
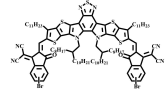
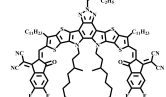
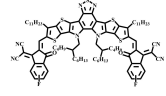
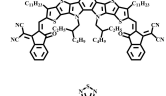
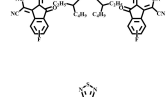
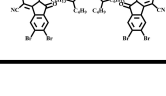
A2	PBDT-Cl	-3.64	-5.53	0.828	25.79	0.73	15.59		10.1002/adfm.202201150
BTP-ClBr2	PM6	-3.64	-5.48	0.845	24.97	73.6	15.54		10.1002/aenm.202002649
BTP-C6Ph	PM6	-3.64	-5.48	0.839	24.3	76.2	15.5		10.1016/j.nanoen.2020.105087
BDOTP-2	D18-B	-2.71	-5.51	0.977	22.29	71.1	15.48		10.1002/cey2.250
Y22	PM6	-3.64	-5.48	0.853	24.37	74.12	15.4		10.1039/D0TA01636E
BTP-2Cl-d	PM6	-3.64	-5.48	0.89	24.3	71.1	15.4		10.1021/acsami.0c16389
BTP-ThMeCl	PM6	-3.64	-5.48	0.9	23	73.4	15.4		10.1002/adfm.202305765
eC9-4F	Tz6T	-3.61	-5.73	0.863	25.14	70.86	15.38		10.1002/adma.202206566
Y-2Br-HD	PM6	-3.64	-5.48	0.9	23.96	70.7	15.32		10.1039/D2TC03768H

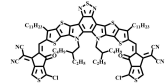
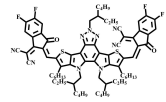
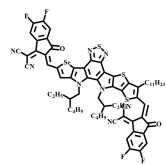
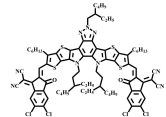
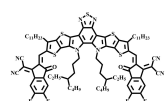
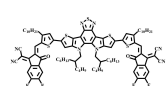
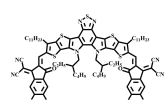
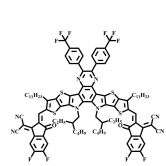
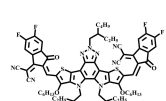
BTP-PBO-4F	PM6	-3.64	-5.48	0.85	25.59	70.62	15.3		10.1021/acsami.3c13833
Y6-PH	PM6	-3.64	-5.48	0.83	25	73.7	15.3		10.1016/j.nanoen.2022.107574
BTIC-CF3-m	PM6	-3.64	-5.48	0.85	24.89	72.32	15.3		10.1016/j.joule.2020.02.004
BTP-6F	PM6	-3.64	-5.48	0.81	25.9	72.8	15.3		10.1002/anie.201915030
BTP-FCI-FCI	BTR-Cl	-3.70	-5.46	0.825	24.58	75.36	15.3		10.1002/advs.202004262
Y18-1D	PM6	-3.64	-5.48	0.84	24.53	74	15.25		10.1039/D0SE01343A
BDTP-4F	PM6	-3.64	-5.48	0.895	22.54	75.5	15.24		10.1002/sml.202001942
5Se	PM6	-3.64	-5.48	0.812	25.74	72.93	15.23		10.1039/D3CC02560H
BTTPC-Br	PBDB-T	-3.51	-5.34	0.86	24.71	71	15.22		10.1039/C9TA11285E
BTP-M	PM6	-3.64	-5.48	0.829	25.54	72.15	15.21		10.1039/C9EE03710A

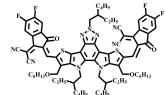
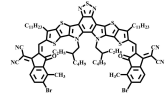
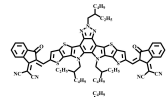
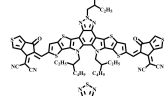
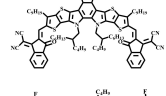
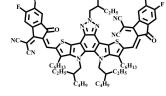
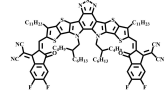
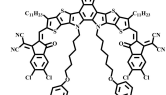
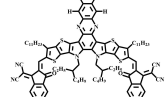
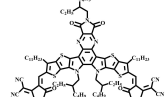
BTP-S1	PM6	-3.64	-5.48	0.934	22.39	72.69	15.21		10.1002/adma.202001160
C6OB-F	PBDT-F	-3.5	-5.5	0.84	24.97	73	15.21		10.1002/adfm.202100870.
ABP4T-4F	PM6	-3.64	-5.48	0.922	22	75.1	15.2		10.1002/aenm.202003177
LL4	PBDB-T	-3.51	-5.34	0.82	26.05	70.87	15.18		10.1002/adfm.202103445
4Se	PM6	-3.64	-5.48	0.822	25.96	71.08	15.17		10.1039/D3CC02560H
BTP-DF	PTQ10	-2.98	-5.54	0.83	26.47	0.69	15.14		10.1039/D2TA00174H
H2	PBDB-T	-3.51	-5.34	0.79	25.82	73	15.12		10.1002/anie.202006081
BTSe-IC2F	D18	-2.77	-5.51	0.863	24.38	71.77	15.1		10.1021/acsami.1c20813
BTCT-2Cl	PM6	-3.64	-5.48	0.877	24.4	70.4	15.1		10.1021/acs.chemmater.0c01245
BO-5Cl	PM6	-3.64	-5.48	0.958	22.57	70.1	15.02		10.1038/s41467-022-30225-7

Y6-2Se	PM6	-3.64	-5.48	0.83	24.32	70	14.94		10.1039/D0TA06658C
Y14	PBDB-T	-3.51	-5.34	0.798	26.15	71.48	14.92		10.1039/C9QM00499H
Y10	PM6	-3.64	-5.48	0.89	21.21	71.55	14.92		10.1039/C9CP05015A
BTP-e7	PBDB-T-F	-3.5	-5.5	0.843	24.1	73.5	14.9		10.1002/nano.202100036
Y11	PM6	-3.64	-5.48	0.87	23.69	72	14.85		10.1016/j.jechem.2021.05.041
Y1-4F	PBDB-T-F	-3.5	-5.5	0.83	25.8	70.6	14.8		10.1002/adma.201904215.
LL2	PBDB-T	-3.51	-5.34	0.84	24.24	72.67	14.75		10.1002/adfm.202103445
BTP10-4Cl-C8	PM6	-3.64	-5.48	0.9	23.37	69.9	14.7		10.1002/aenm.202003777
BZ4F-O-3	PM6	-3.64	-5.48	0.85	23.51	73.72	14.69		10.1021/acsenergylett.2c00985
TPQx-6F	PM6	-3.64	-5.48	0.92	22.37	72.16	14.62		10.1039/D0QM00034E

BTP-ClBr1	PM6	-3.64	-5.48	0.854	23.66	72.2	14.56		10.1002/aenm.202002649
C6	PM6	-3.64	-5.48	0.84	23.82	72.68	14.54		10.1039/D0TA05787H
BTTPC	PBDB-T	-3.51	-5.34	0.89	22.25	73	14.51		10.1039/C9TA11285E
BTP-4Cl	PM6	-3.64	-5.48	0.85	25	68	14.5		10.1021/acsami.0c16389
BTP-2ThCl	PM6	-3.64	-5.48	0.885	23.46	69.8	14.49		10.1016/j.joule.2020.03.023
BTP-IIC-P	PM6	-3.64	-5.48	0.82	25.03	70.39	14.47		10.1021/acsami.2c22972
TB-4Cl	PM6	-3.64	-5.48	0.835	23.2	74.56	14.45		10.1021/acsami.1c02652
LL3-2Cl	PBDB-T	-3.51	-5.34	0.79	26.96	67.7	14.4		10.1002/adfm.202103445
BT4-4Cl	D18	-2.77	-5.51	0.824	24.35	71.73	14.39		10.1002/solr.202100522

N4	PBDB-T-F	-3.5	-5.5	0.819	25.01	69.9	14.31		10.1002/aenm.202103422
Y-2Br-DT	PM6	-3.64	-5.48	0.91	22.65	69.21	14.26		10.1039/D2TC03768H
BTP-1V-2F	PM6	-3.64	-5.48	0.823	24.13	70.81	14.2		10.1021/acsami.1c20813
BTP-Se	PBDB-T	-3.51	-5.34	0.711	28.66	69.7	14.2		10.1016/j.cej.2021.132830
Y-2Br-OD	PM6	-3.64	-5.48	0.92	22.35	68.8	14.19		10.1039/D2TC03768H
Y18-DMO	PBDB-T	-3.51	-5.34	0.81	23.98	73	14.18		10.1021/acsaem.0c02165
BTP-2F	PM6	-3.64	-5.48	0.89	22.1	71.7	14.1		10.1002/anie.201915030
Y5	PBDB-T	-3.51	-5.34	0.88	22.8	70.2	14.1		10.1002/adma.201807577
H1	PBDB-T	-3.51	-5.34	0.76	25.74	71	14.06		10.1002/anie.202006081
BTIC-BO-4Br	PM6	-3.64	-5.48	0.86	24.06	67.84	14.03		10.1002/advs.201903784

BTTT-2Cl	PBDB-T-F	-3.5	-5.5	0.896	23.8	65.19	13.8		10.1002/nano.202100036
BZ4F	PM6	-3.64	-5.48	0.85	22.53	71.98	13.79		10.1021/acsenenergylett.2c00985
BTSe-4F	PM6	-3.64	-5.48	0.811	22.52	75.4	13.79		10.1002/sstr.202000052
Y19-N3	PBDB-T	-3.51	-5.34	0.78	24.46	72	13.77		10.1016/j.cclet.2020.10.042
Y6-C3	PM6	-3.64	-5.48	0.852	24.07	67.4	13.76		10.1007/s11426-019-9670-2
H3	PBDB-T	-3.51	-5.34	0.757	25.84	70	13.75		10.1039/D0TA06907H
BTIC-F-m	PM6	-3.64	-5.48	0.92	21.41	69.09	13.61		10.1016/j.joule.2020.02.004
TPQ-F	PM6	-3.64	-5.48	0.857	23.7	66.5	13.51		10.1016/j.orgel.2021.106282
BZ4F-O-1	PM6	-3.64	-5.48	0.91	19.57	75.61	13.5		10.1021/acsenenergylett.2c00985

BZ4F-O-2	PM6	-3.64	-5.48	0.82	21.62	74.74	13.43		10.1021/acsenergylett.2c00985
BTP-(Br,Me)-1	PM6	-3.64	-5.48	0.92	21.38	68.25	13.43		10.1021/acsami.1c08060
Y1	PBDB-T	-3.51	-5.34	0.87	22.44	69.1	13.42		10.1038/s41467-019-08386-9
Y2	PBDB-T	-3.51	-5.34	0.82	23.56	69.4	13.4		10.1038/s41467-019-08386-9
BTP-0Cl	PBDB-T	-3.51	-5.34	0.902	21.4	70.1	13.4		10.1002/anie.202102622
Y26	PM6	-3.64	-5.48	0.83	21.63	74.33	13.34		10.1002/solr.202100281
Y6-BO	P3HT	-3.10	-5.00	0.72	24.39	74.33	13.34		10.1002/adma.202008158
A1	PBDT-Cl	-3.64	-5.53	0.771	24.13	0.71	13.33		10.1002/adfm.202201150
AQx-1	PM6	-3.64	-5.48	0.89	22.18	67.14	13.31		10.1002/adma.201906324
QIP-4Cl	P2F-Ehp	-3.13	-5.46	0.94	19.62	72.11	13.3		10.1039/D0CC00896F

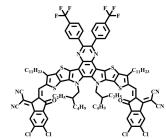
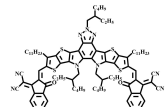
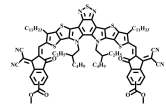
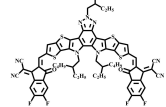
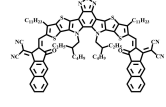
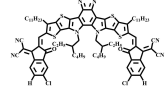
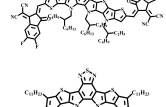
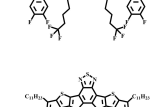
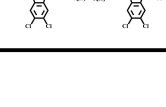
TPQ-Cl	PM6	-3.64	-5.48	0.837	25.19	62.99	13.28		10.1016/j.orgel.2021.106282
Y9	PBDB-T	-3.51	-5.34	0.9	23.28	63	13.26		10.1016/j.jechem.2019.07.002
BTIC-COOMe	PM6	-3.64	-5.48	0.9	22.13	66.71	13.25		10.1021/acs.jpcclett.1c01077
Y3	PM6	-3.64	-5.48	0.81	24.1	66.4	13.24		10.1039/D0TA01260B
BT-LIC	PBDB-T	-3.51	-5.34	0.86	24.3	63.2	13.2		10.1021/jacs.1c00211
BTIC-Cl-m	PM6	-3.64	-5.48	0.88	21.35	69.7	13.16		10.1016/j.joule.2020.02.004
BTDTP-4F	PM6	-3.64	-5.48	0.866	21.25	71.3	13.12		10.1002/sml.202001942
BTP-OF	PTQ10	-2.98	-5.54	0.83	24.37	0.64	12.96		10.1039/D2TA00174H
LL2-2Cl	PBDB-T	-3.51	-5.34	0.72	26.66	67.1	12.95		10.1002/adfm.202103445

Table S2. Performance comparison of three hyperparameter optimization methods on the optimal XGBoost model, including the number of trial, tuning time, and testing R² value.

Methods	Number of trial	Tuning Time (s)	R ² _Testing
Optuna	60	28.3	0.88
RandomizedSearchCV	80	40.6	0.85
GridSearchCV	36	9.0	0.84

Table S3. Performance metrics of six machine learning models with the Optuna hyperparameter optimization.

Models	R ² (Training/Testing)	RMSE (Training/Testing)	MAE (Training/Testing)
XGBoost	0.93/0.88	0.64/0.80	0.34/0.40
CatBoost	0.85/0.80	1.10/1.00	0.77/0.89
SVR	0.77/0.73	1.32/1.01	1.00/0.82
RF	0.81/0.70	0.70/1.52	0.65/1.01
GBR	0.90/0.70	1.10/1.45	0.80/0.96
GPR	0.73/0.69	1.08/1.33	0.88/1.06

Table S4. External validation of the XGBoost model using data from 20 recent studies, including predicted and experimental PCEs with relative errors.

Donors	Acceptors	V_{oc} (V)	J_{sc} (mA cm ⁻²)	FF (%)	PCE (%)	Pre_PCE (%)	Error (%)	Refs
PM6	R-EH	0.86	26.95	77.96	18.09	16.95	6.26	[21]
PM6	R-M	0.86	26.60	77.75	17.88	16.91	5.43	[21]
PM6	R-DTD	0.87	24.95	61.77	13.18	13.95	5.84	[21]
D18	BTP-TC4Ph	0.89	27.93	79.36	19.65	18.22	7.28	[22]
D18	BTP-TC6	0.86	27.53	77.67	18.47	17.96	2.76	[22]
D18	BQx-F	0.96	21.46	69.72	14.41	15.75	9.30	[23]
D18	BQx-Cl	0.96	23.03	71.60	15.74	16.66	5.84	[23]
D18	BQx-Br	0.96	22.30	70.07	15.03	16.38	8.77	[23]
D18	Z11	0.97	22.4	70.3	15.4	16.03	4.09	[24]
D18	Z12	0.93	25.3	79.6	19.5	18.11	7.13	[24]
D18	Z13	0.89	24.8	76.7	17.8	17.15	3.65	[24]
PM6	L8-PhF	0.86	25.63	77.70	17.17	18.11	5.47	[25]
PM6	PzIC-SS-4F	0.89	26.22	7.40	17.02	16.76	1.53	[26]
PM6	PzIC-SSe-4F	0.87	26.53	73.50	17.58	16.67	5.18	[26]
PM6	PzIC-SeSe-4F	0.87	26.79	73.12	17.69	17.01	3.84	[26]
PM1	DASe-4F	0.85	27.58	72.22	18.27	18.45	0.99	[27]
PM1	DASe-4Cl	0.85	26.71	75.80	17.25	16.85	2.32	[27]
PM1	TASe-2Cl2F	0.85	26.10	73.20	16.30	17.79	9.14	[27]
PM1	TASe-2F2Cl	0.88	27.52	79.88	19.32	17.95	7.09	[27]
PM6	BTIC-4Br	0.86	24.06	67.84	14.03	14.22	1.35	[28]

Table S5. The detailed electrostatic potential parameters of candidate K01-K12 NFAs.

Donor	Acceptors	Min_value ^{a)} (kcal/mol)	Max_value ^{b)} (kcal/mol)	ESP_Avg ^{c)} (kcal/mol)	MPI ^{d)} (kcal/mol)	Osa ^{e)} (Å ²)	Psa ^{f)} (Å ²)	Nsa ^{g)} (Å ²)
PM6	K01	-33.41	41.66	4.46	8.52	1324.79	1015.44	309.35
	K02	-33.62	41.63	4.28	8.34	1323.96	979.07	344.89
	K03	-34.20	37.26	4.34	8.34	1334.40	1009.46	324.94
	K04	-34.45	41.11	4.08	8.39	1313.82	990.92	322.90
	K05	-33.44	40.01	4.43	8.59	1329.93	1029.16	300.77
	K06	-33.25	41.91	4.35	8.29	1346.02	1005.65	340.37
	K07	-33.10	41.67	4.29	8.58	1293.52	962.36	331.16
	K08	-33.05	41.64	4.30	8.52	1294.03	962.66	331.37
	K09	-33.50	41.60	4.28	8.42	1316.00	980.22	348.96
	K10	-34.26	41.73	3.96	8.49	1265.28	917.43	347.84
	K11	-34.04	41.86	4.06	8.25	1291.90	946.33	345.57
	K12	-33.38	41.77	4.37	8.47	1319.06	988.08	330.97

^{a)} Min_value: the minimum values of the electrostatic potential on the molecular surface.

^{b)} Max_value: the maximum values of the electrostatic potential on the molecular surface.

^{c)} ESP_Avg: the overall average electrostatic potential.

^{d)} MPI: the molecular polarity index.

^{e)} Osa: the overall surface area of the ESP isosurface.

^{f)} Psa: the positive ESP surface areas.

^{g)} Nsa: the negative ESP surface areas.

References

- 1 Chen T., He T., Benesty M., Khotilovich V., Tang Y., Cho H., Chen K., Mitchell R., Cano I., Zhou T. *R package version 0.4-2*. 2015, **1**, 1-4.
- 2 Prokhorenkova L., Gusev G., Vorobev A., Dorogush A.V., Gulin A. *Advances in neural information processing systems*. 2018, **31**, 6638-6648.
- 3 Awad M., Khanna R. Support vector regression. In: *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*. Springer; 2015. p. 67-80.
- 4 Rigatti S.J. *Journal of Insurance Medicine*. 2017, **47**, 31-9.
- 5 Natekin A., Knoll A. *Front. Neurorob.* 2013, **7**, 21.
- 6 Williams C., Rasmussen C. *Advances in neural information processing systems*. 1995, **8**, 514-520.
- 7 Akiba T., Sano S., Yanase T., Ohta T., Koyama M., editors. Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*; 2019.
- 8 Lundberg S.M., Lee S.-I. *Advances in neural information processing systems*. 2017, **30**, 4765-4774.
- 9 Forrest S. *ACM computing surveys (CSUR)*. 1996, **28**, 77-80.
- 10 Frisch M. *Inc, Wallingford CT*. 2009.
- 11 Curtiss L.A., McGrath M.P., Blaudeau J.P., Davis N.E., Binning Jr R.C., Radom L. *The Journal of Chemical Physics*. 1995, **103**, 6104-13.
- 12 Grimme S., Antony J., Ehrlich S., Krieg H. *The Journal of chemical physics*. 2010, **132**, 0021-9606.
- 13 Lu T., Chen F. *J. Comput. Chem.* 2012, **33**, 580-92.
- 14 Murray J.S., Politzer P. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 2011, **1**, 153-63.
- 15 Humphrey W., Dalke A., Schulten K. *J. Mol. Graphics*. 1996, **14**, 33-8.
- 16 Dennington R., Keith T.A., Millam J.M. *Semichem Inc Shawnee Mission KS*. 2016.
- 17 Adamo C., Jacquemin D. *Chem. Soc. Rev.* 2013, **42**, 845-56.
- 18 Marenich A.V., Cramer C.J., Truhlar D.G. *The Journal of Physical Chemistry B*. 2009, **113**, 6378-96.
- 19 Schröder H., Creon A., Schwabe T. *J. Chem. Theory Comput.* 2015, **11**, 3163-70.
- 20 Ertl P., Schuffenhauer A. *J. Cheminf.* 2009, **1**, 8.
- 21 Zhao J., Yang X., Shao Y., Sun R., Min J. *Sci. China Mater.* 2025, **68**, 1391-1400.
- 22 Wang P., Jiang H., Bi F., Wang J., Yang C., Li Y., Chu J., Bao X. *CCS Chem.* 2025, **7**, 3036-3049.
- 23 Yang C., Gao Y., Zhang H., Yao Z.F., Li E.L., Guan H.H., Zhi H.F., Yuan Q., Jee M.H., Woo H.Y. *Angew. Chem.* 2025, **64**, e202506795.
- 24 Xu R., Jiang Y., Lei H., Liu F., Liu K., Feng L., Ran G., Zhang W., Zhong C., Zhu X. *Sci. Adv.* 2025, **11**, eadt6024.
- 25 Wu X., Gong Y., Li X., Qin S., He H., Chen Z., Liang T., Wang C., Deng D., Bi Z., et al. *Angew Chem Int Ed Engl.* 2024, **64**, e202416016.
- 26 Li X., Peng R., Qiu Y., Zhang Y., Shi J., Gao S., Liu H., Jin F., Ge Z. *Adv. Funct. Mater.* 2024, **35**, 2413259.

- 27 Yang C., Gao Y., Zhang H., Yao Z.F., Li E.L., Guan H.H., Zhi H.F., Yuan Q., Jee M.H., Woo H.Y., et al. *Angew Chem Int Ed Engl.* 2025, **64**, e202506795.
- 28 Wan F., Wang H., Gu Y., Fan G., Hou S., Yu J., Wang M., He F., Tian L. *Chemistry.* 2024, **30**, e202303502.