

## Supporting Information

### Design and Virtual Screening of Donor and Non-Fullerene Acceptor for Organic Solar Cells Using Long Short-Term Memory Model

Long-Fei Lv <sup>a</sup>, Cai-Rong Zhang <sup>a\*</sup>, Rui Cao <sup>a</sup>, Xiao-Meng Liu <sup>a</sup>, Mei-Ling Zhang <sup>a</sup>, Ji-Jun Gong <sup>a</sup>, Zi-Jiang Liu <sup>b</sup>, You-Zhi Wu <sup>c</sup>, Hong-Shan Chen <sup>d</sup>

<sup>a</sup> Department of Applied Physics, Lanzhou University of Technology, Lanzhou, Gansu 730050, China;

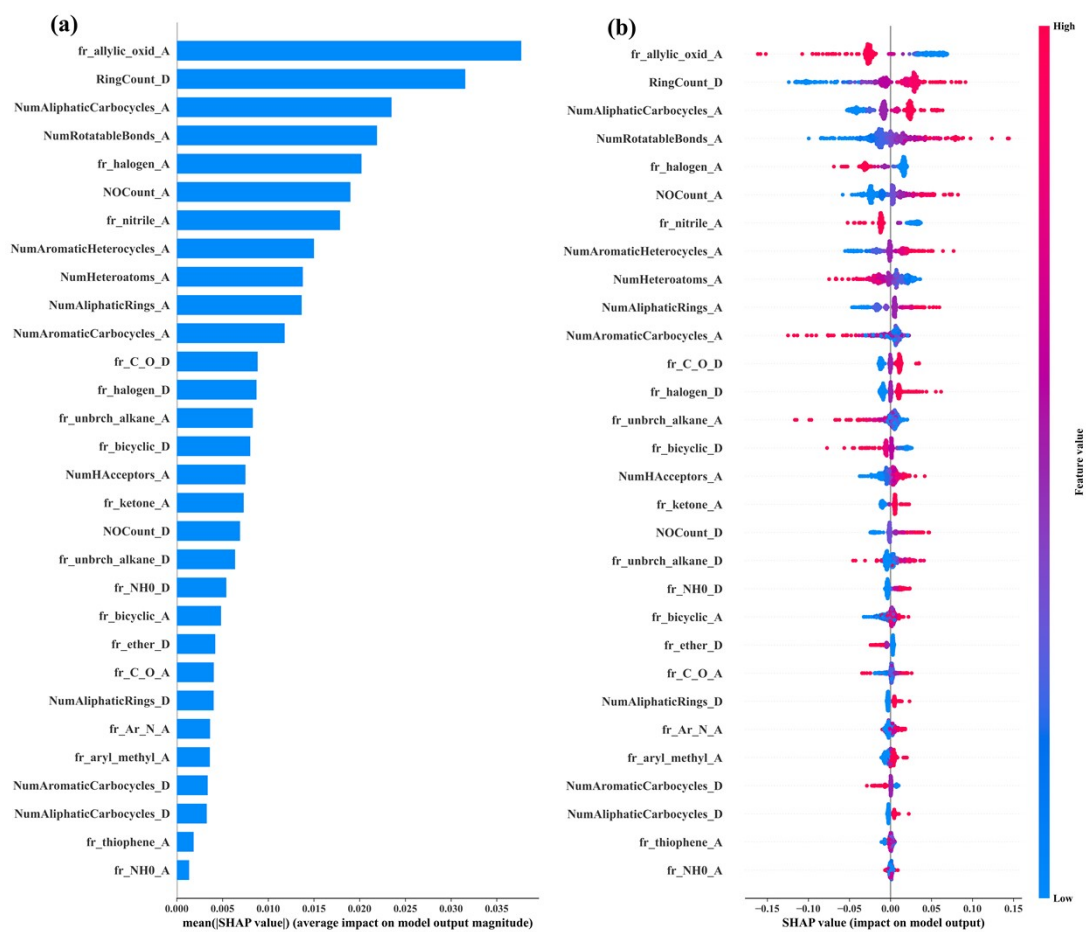
<sup>b</sup> School of Mathematics and Physics, Lanzhou Jiaotong University, Lanzhou 730070, China;

<sup>c</sup> School of Materials Science and Engineering, Lanzhou University of Technology, Lanzhou, Gansu 730050, China;

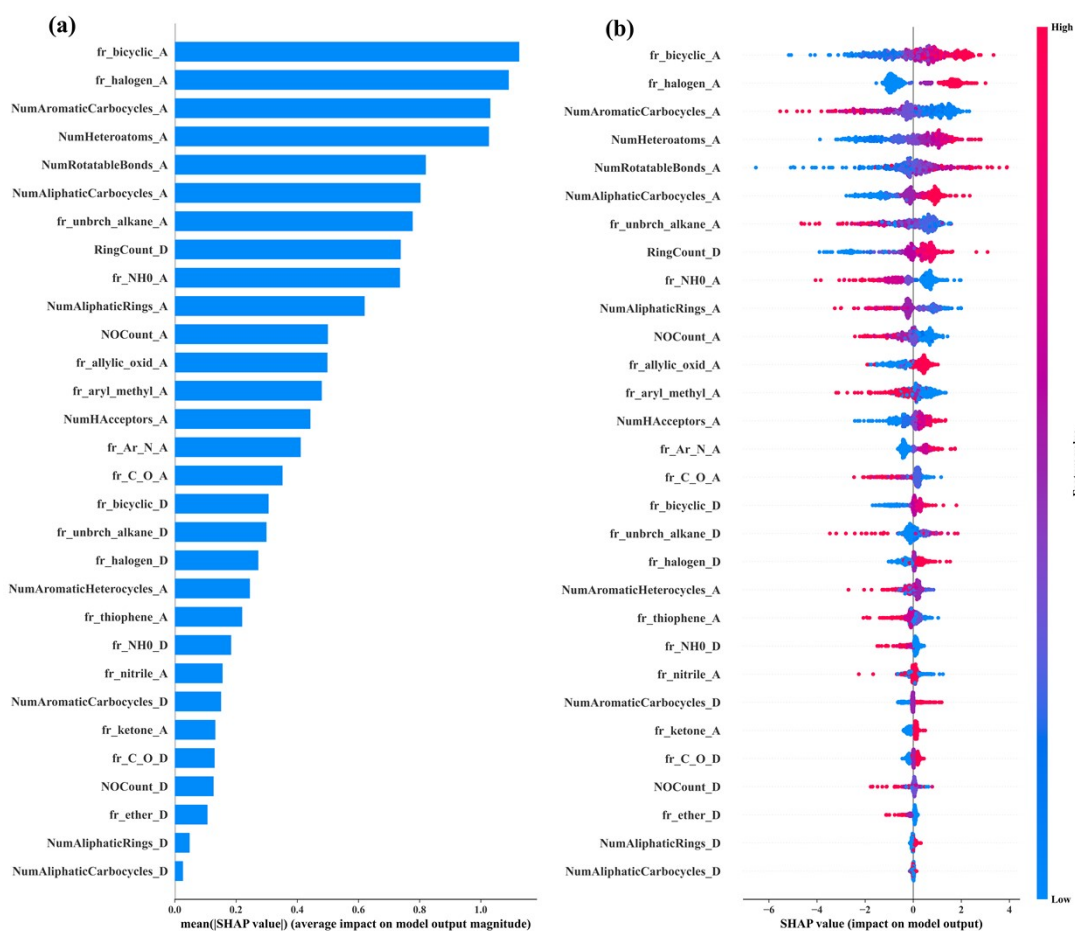
<sup>d</sup> College of Physics and Electronic Engineering, Northwest Normal University, Lanzhou, Gansu 730070, China.

\*Corresponding author.

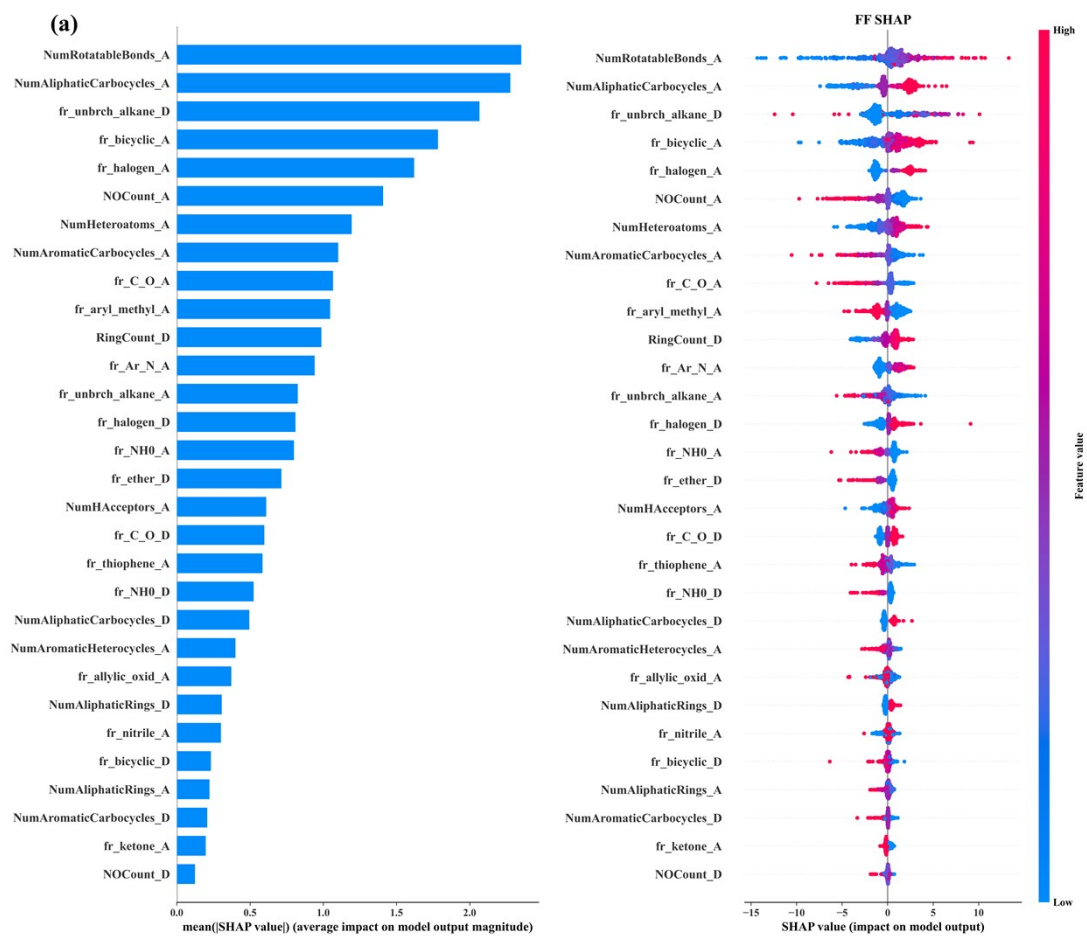
E-mail address: [zhcrxy@lut.edu.cn](mailto:zhcrxy@lut.edu.cn) (C.R. Zhang)



**Fig. S1** SHAP importance analysis of the 30 molecular structure descriptors used in the LSTM model for  $V_{OC}$  prediction

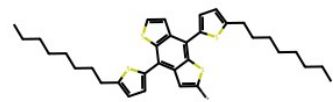
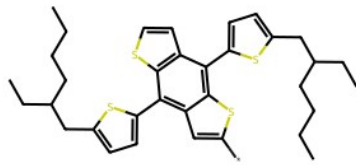
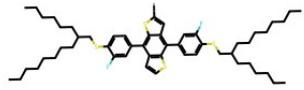
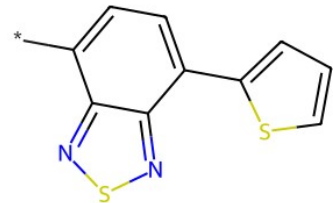
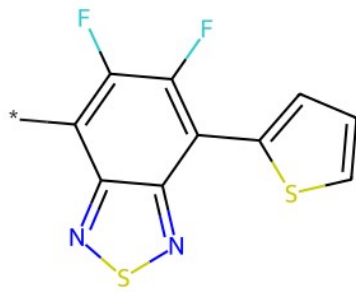
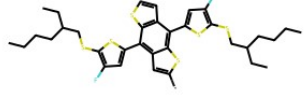
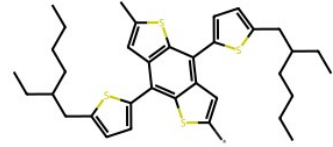
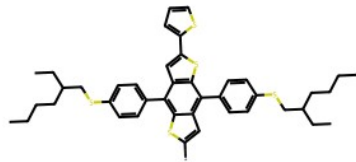
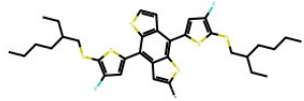
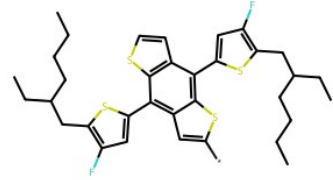
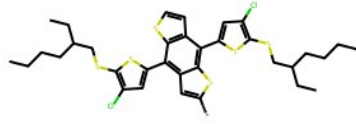
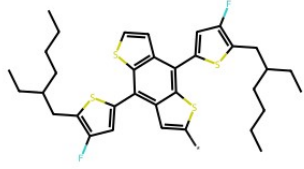


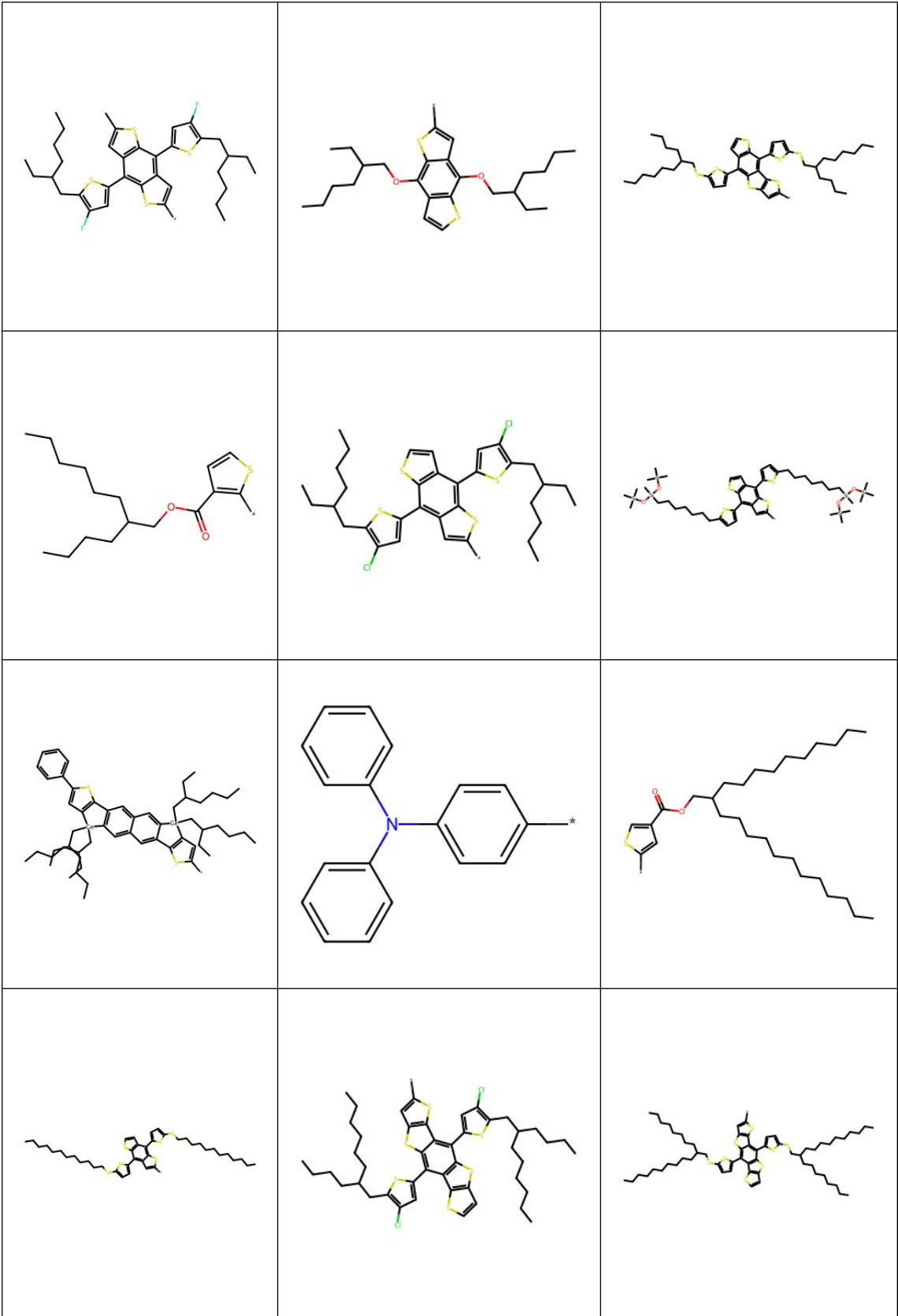
**Fig. S2** SHAP importance analysis of the 30 molecular structure descriptors used in the LSTM model for  $J_{SC}$  prediction


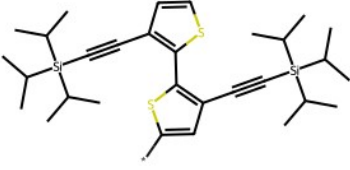
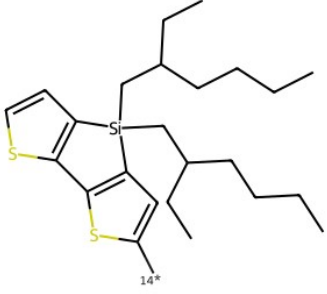
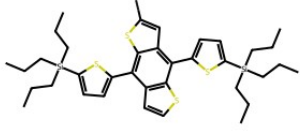
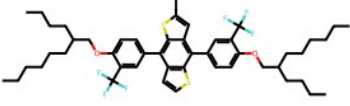
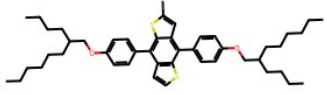
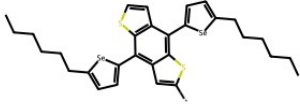
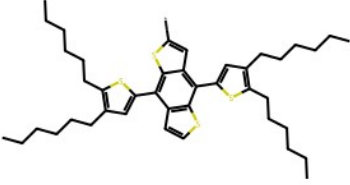
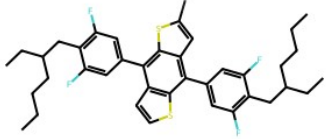
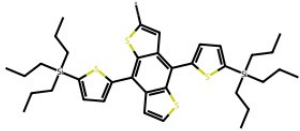
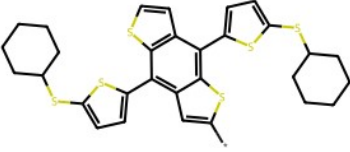
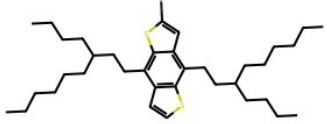


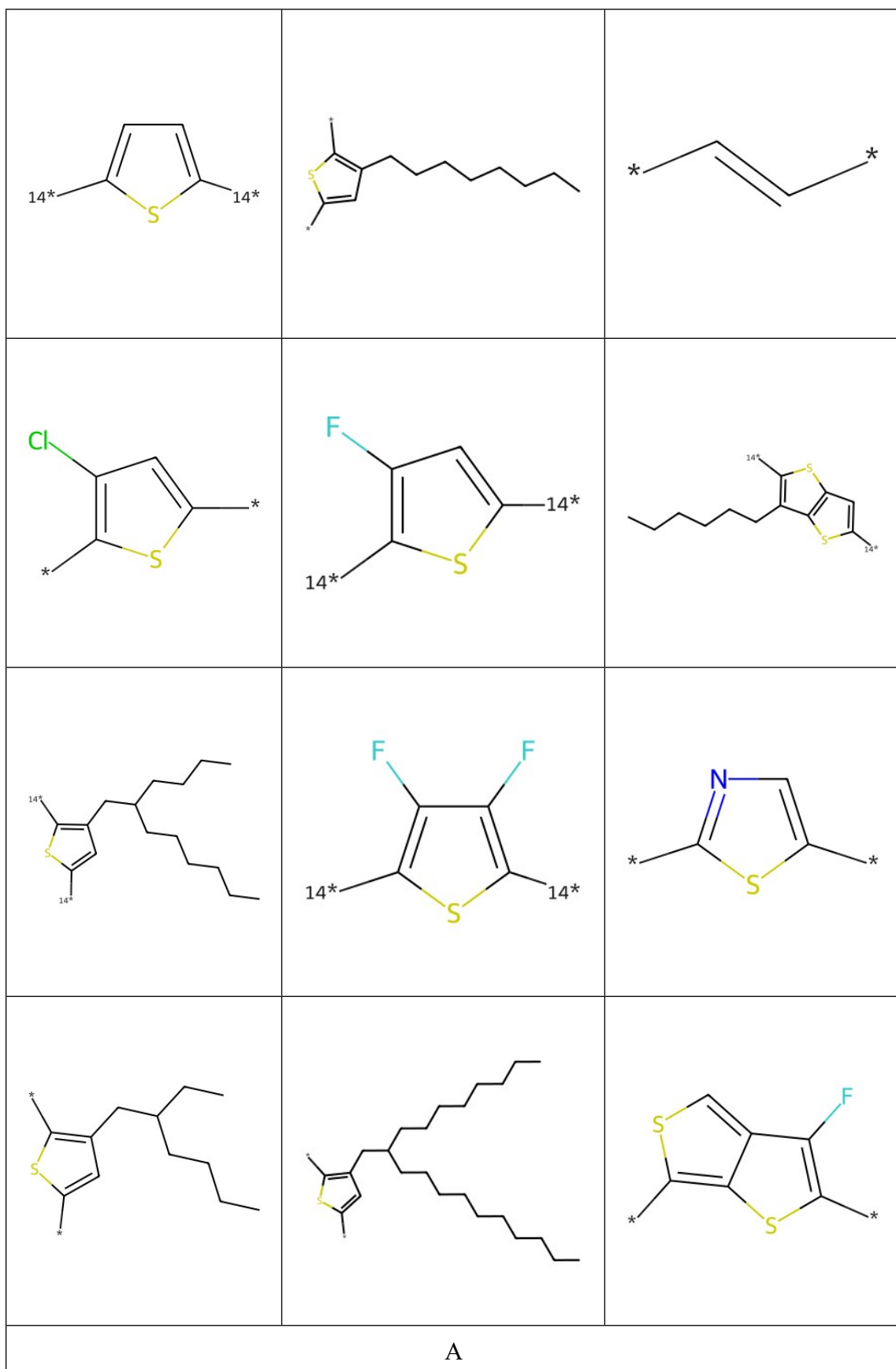
**Fig. S3** SHAP importance analysis of the 30 molecular structure descriptors used in the LSTM model for FF prediction

D

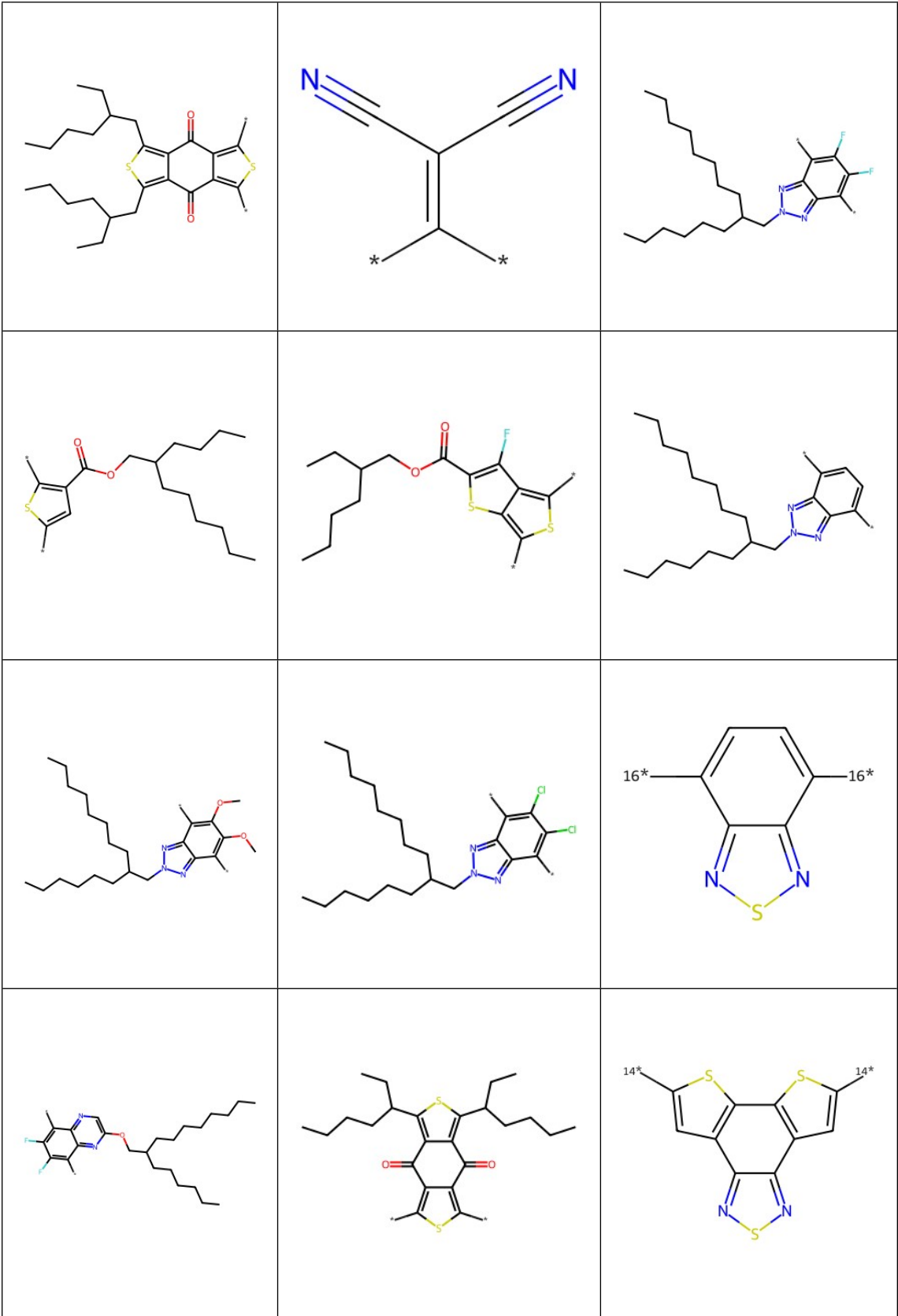


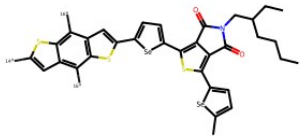
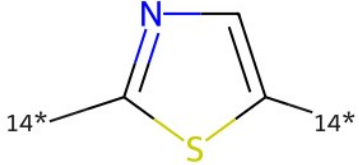
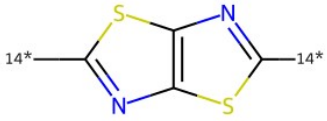
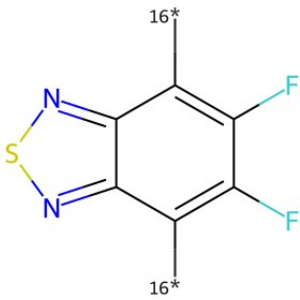
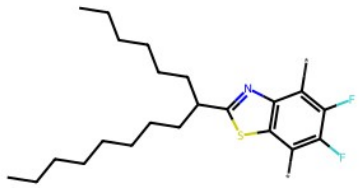
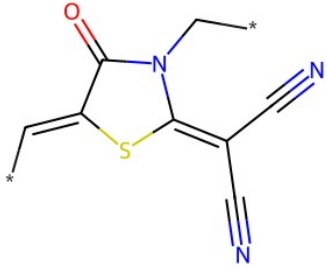
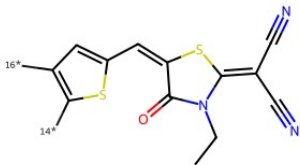
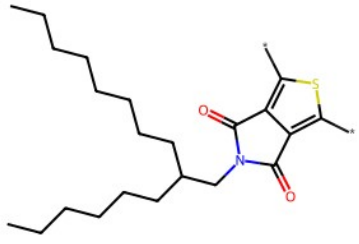
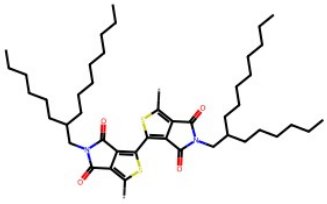
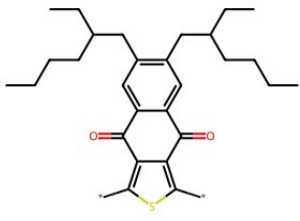
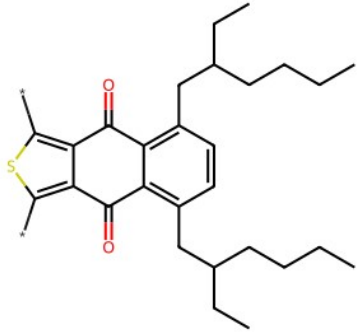
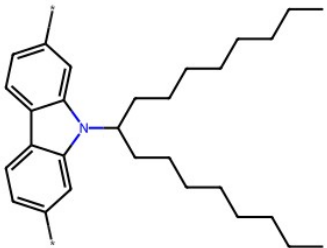


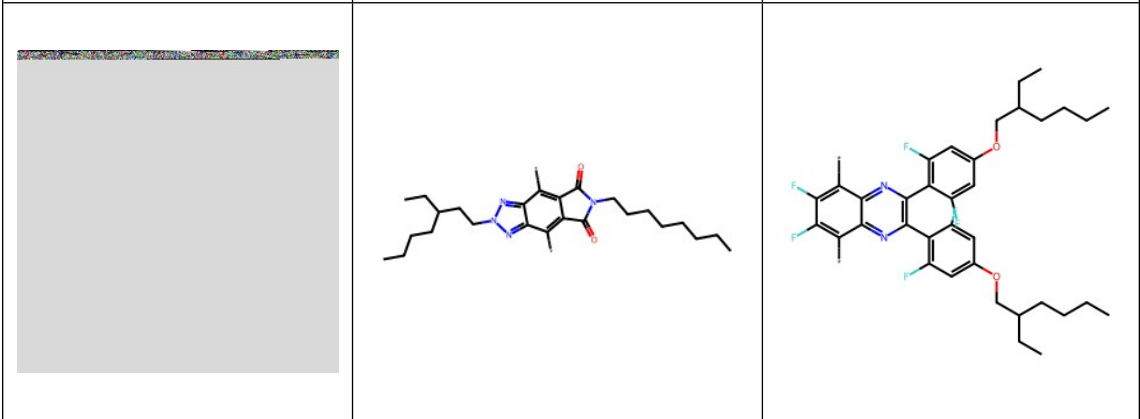
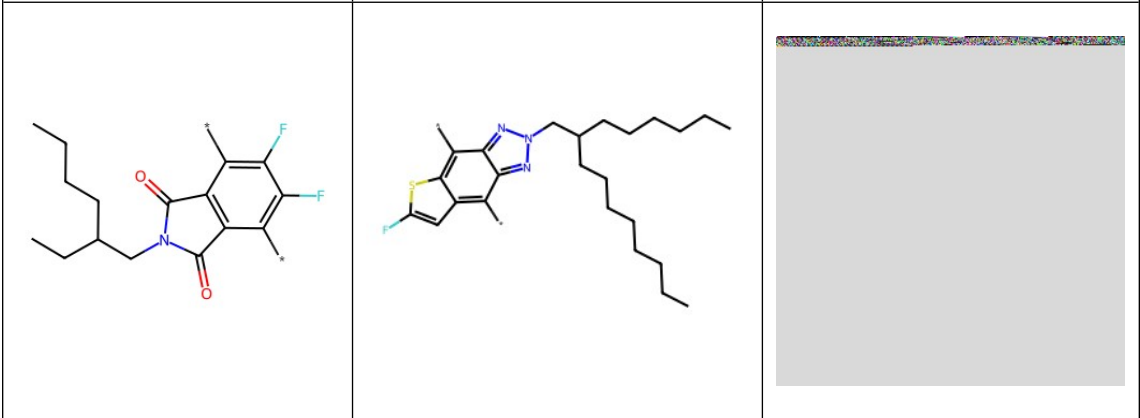
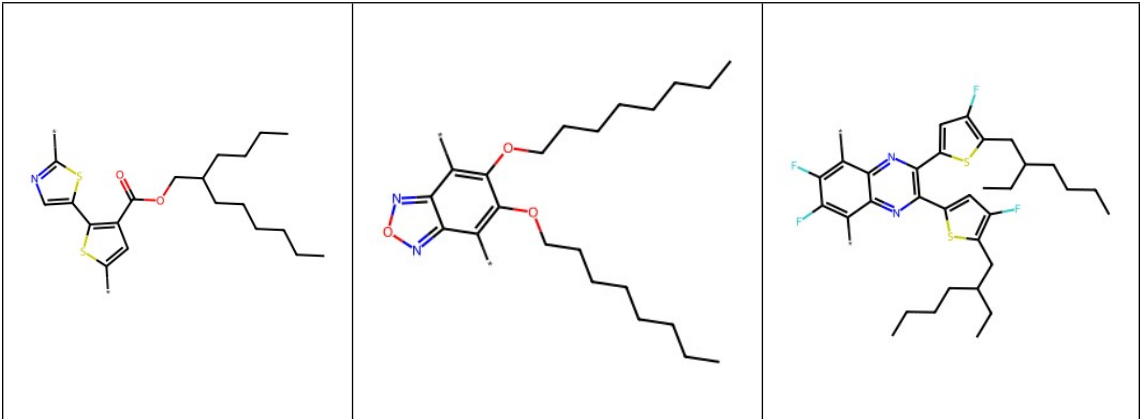
		
		
		
		
$\pi$		



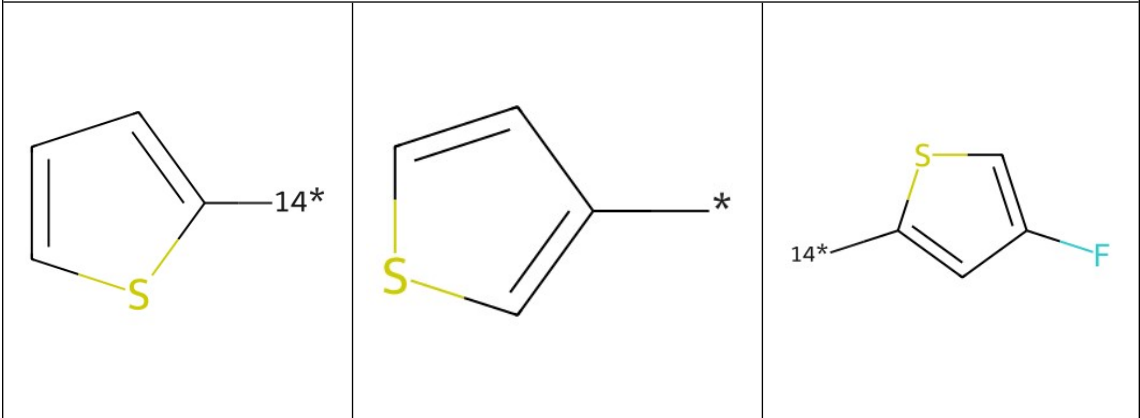


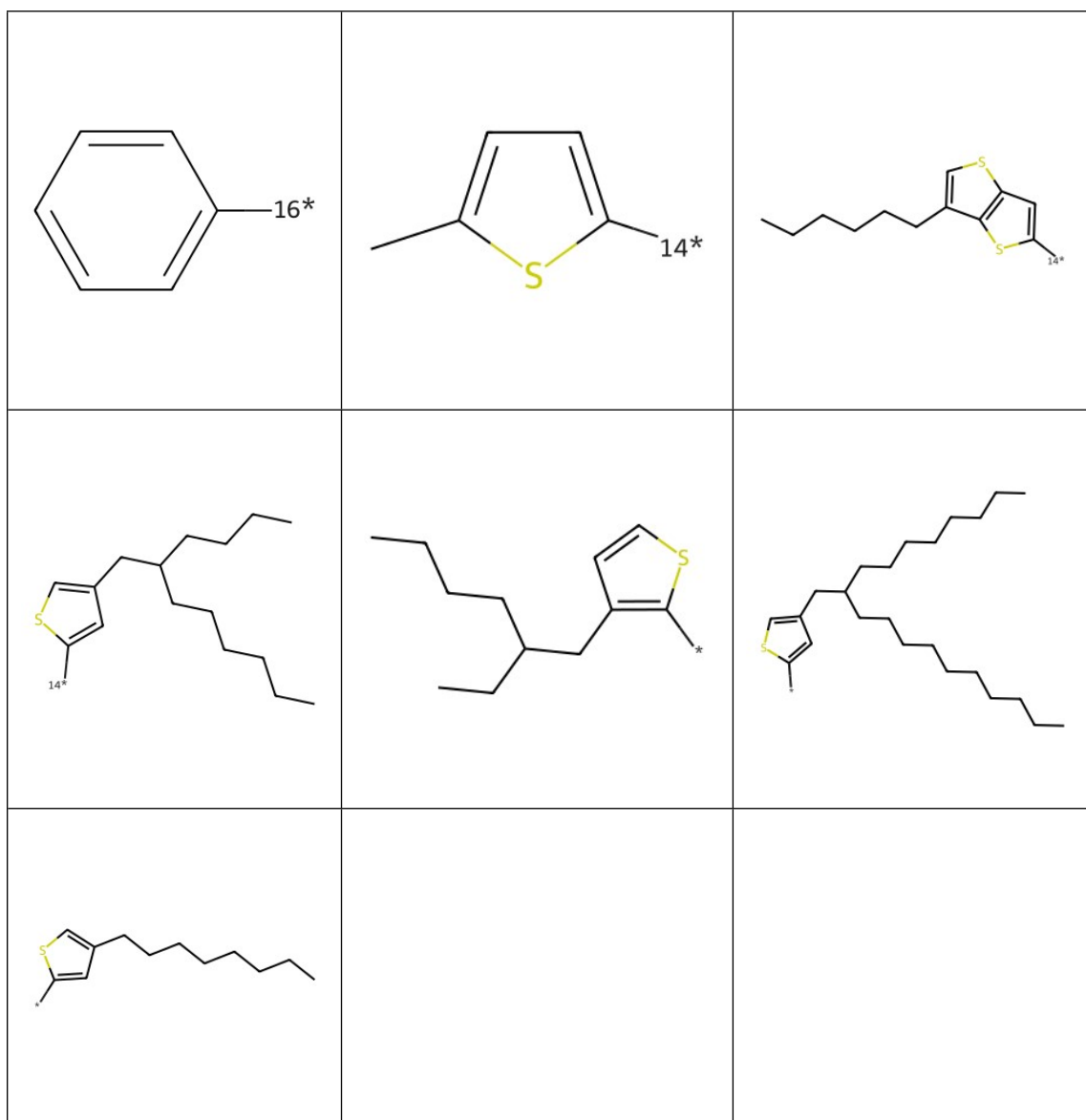


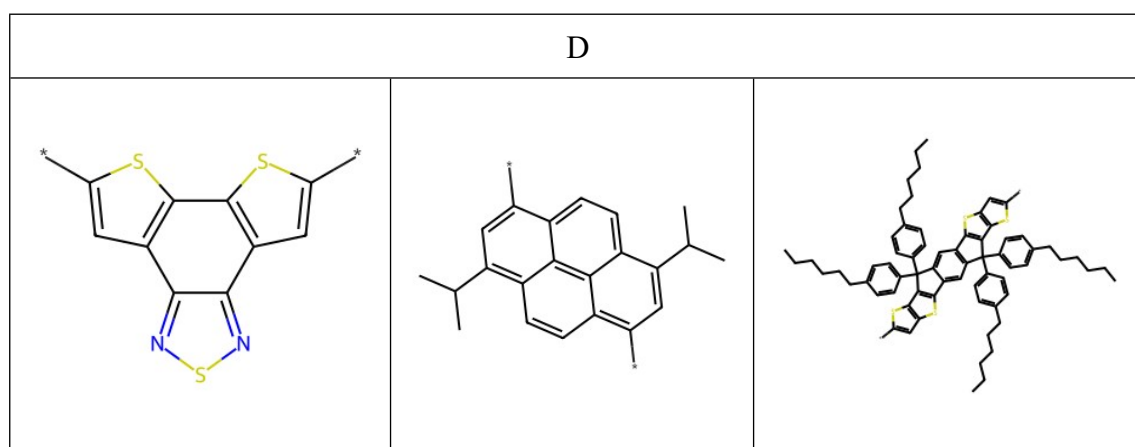


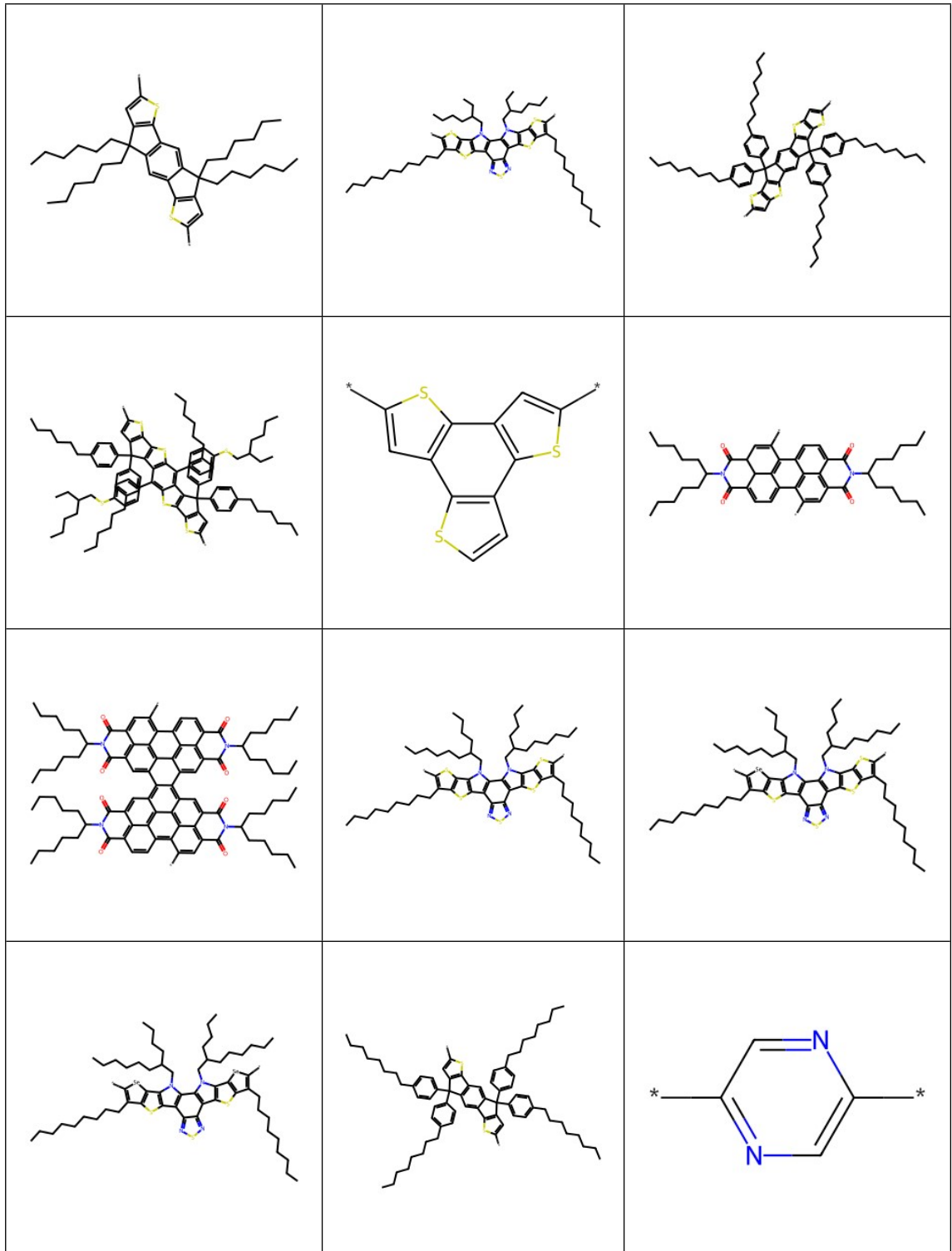
$\pi$

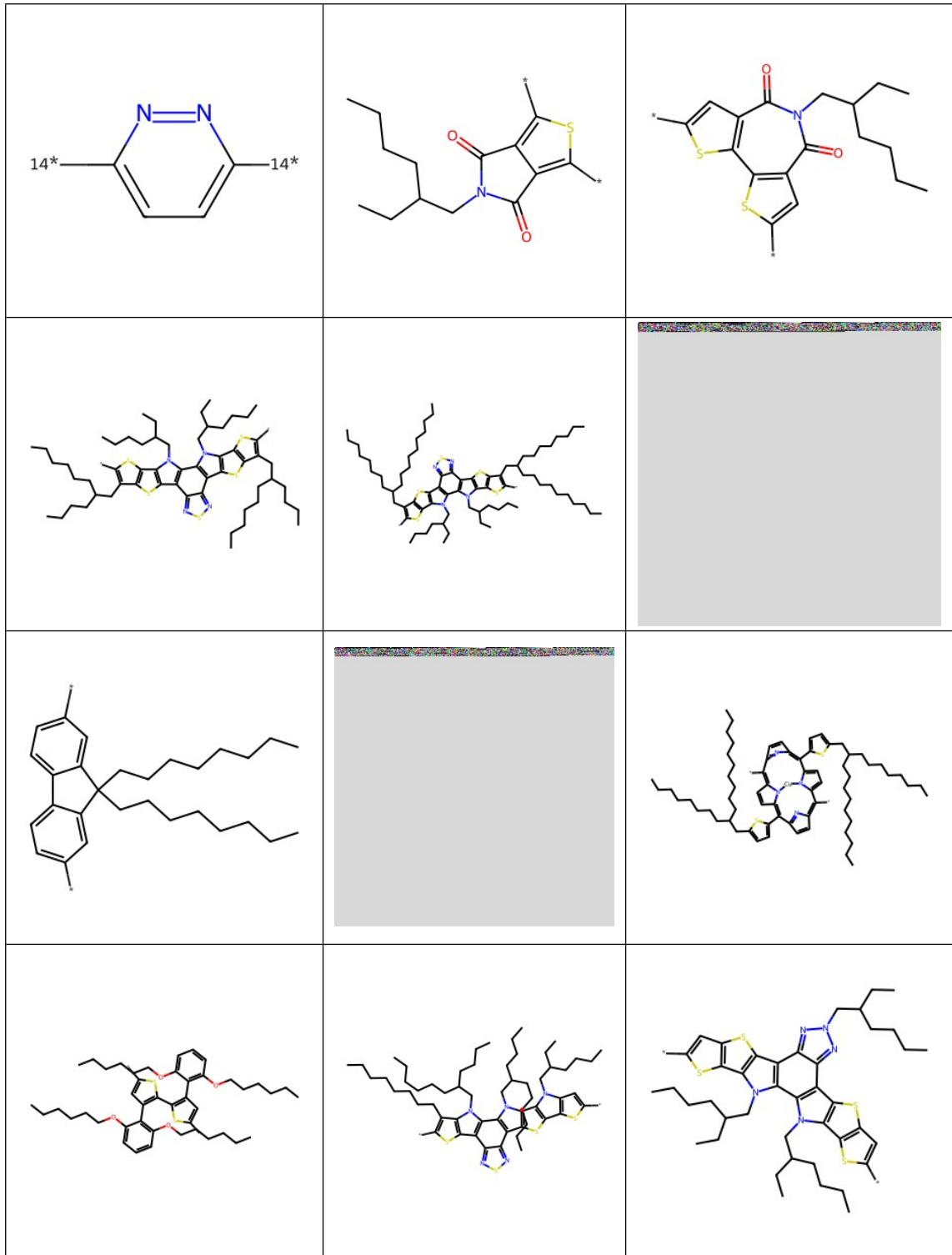


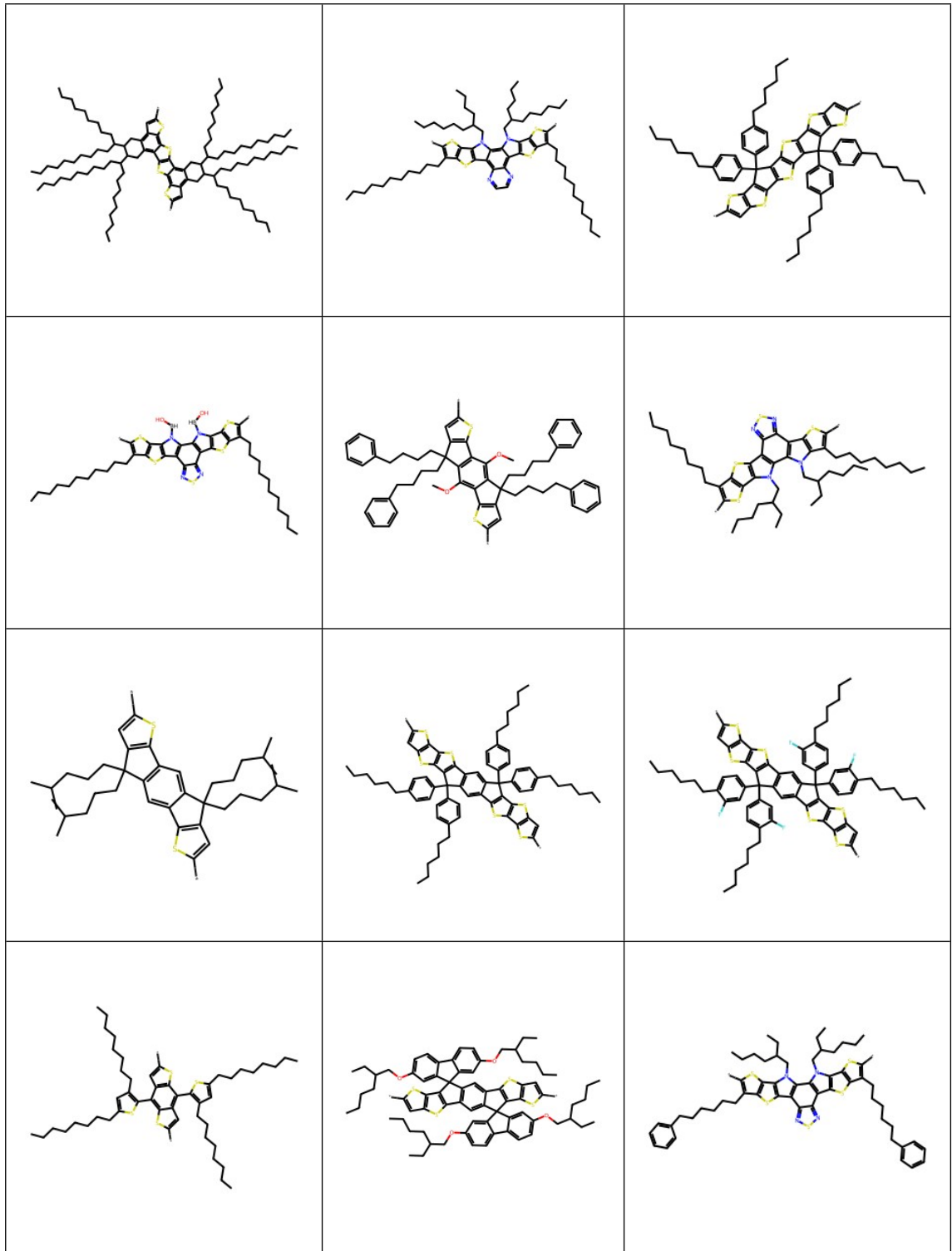


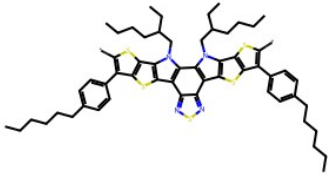
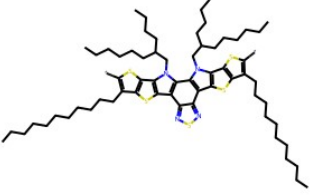
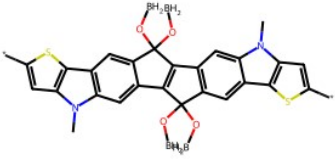
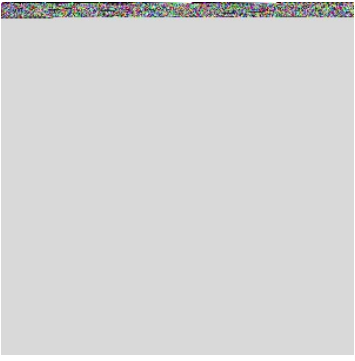
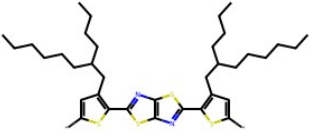
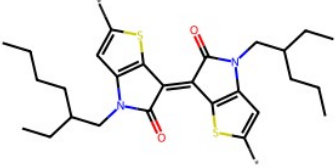
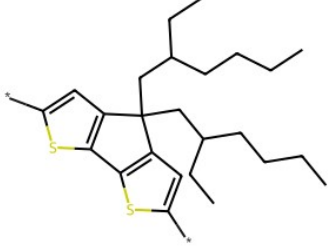
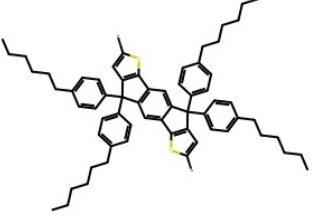
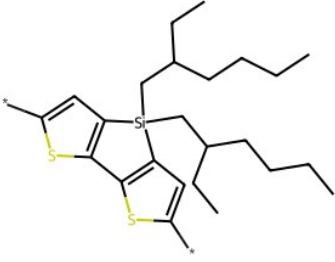

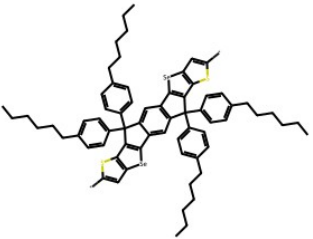
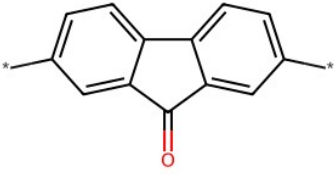
**Fig. S4** Fragments of donor molecules



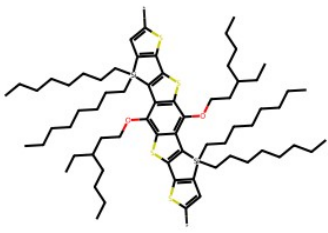
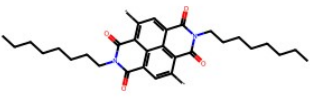
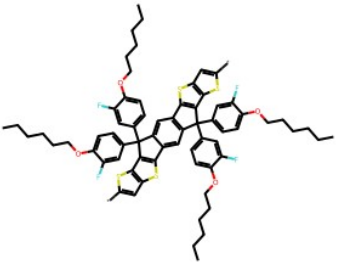
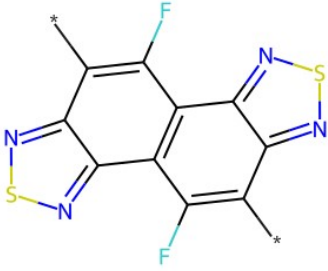
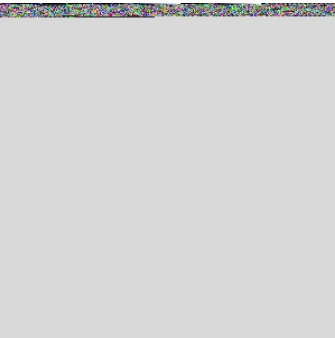
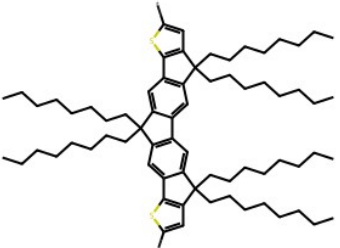
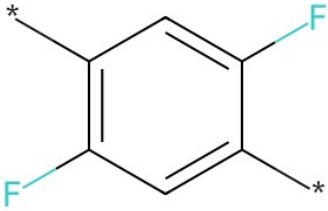
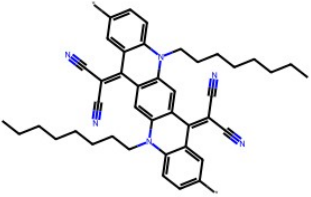
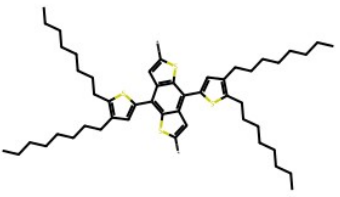
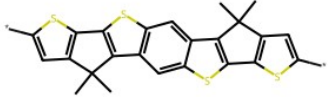


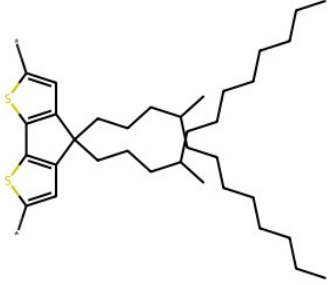
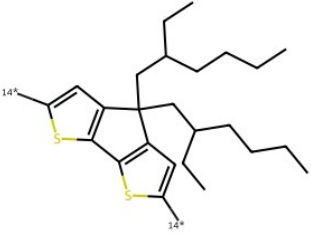

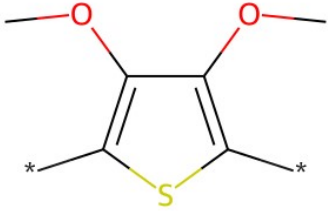


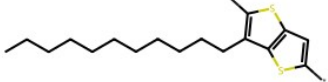
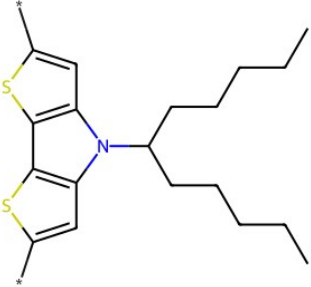
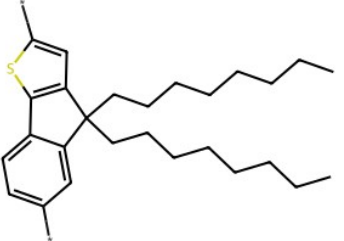
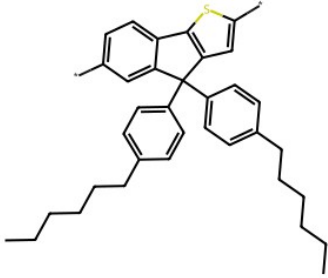
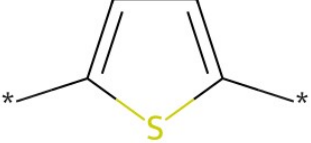
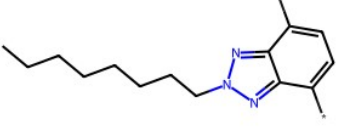


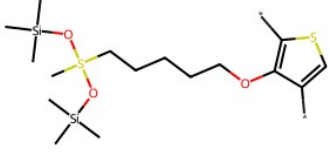
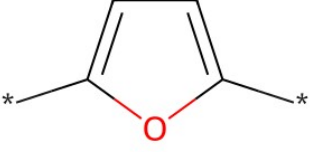
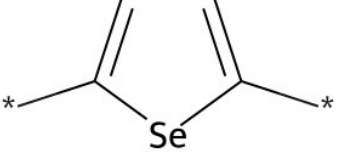
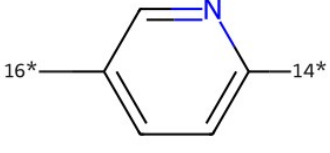
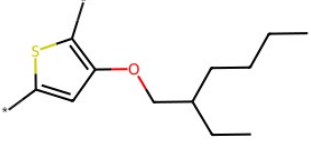
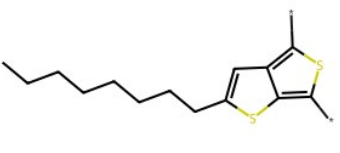
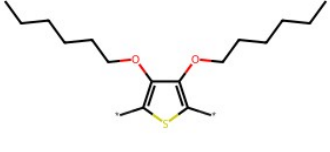

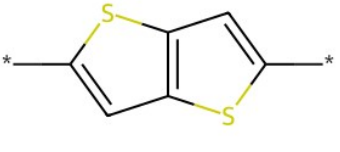
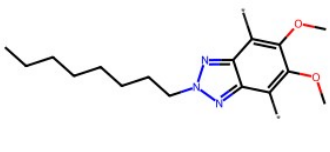
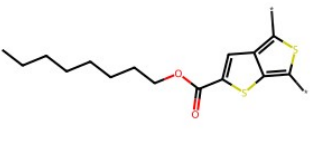


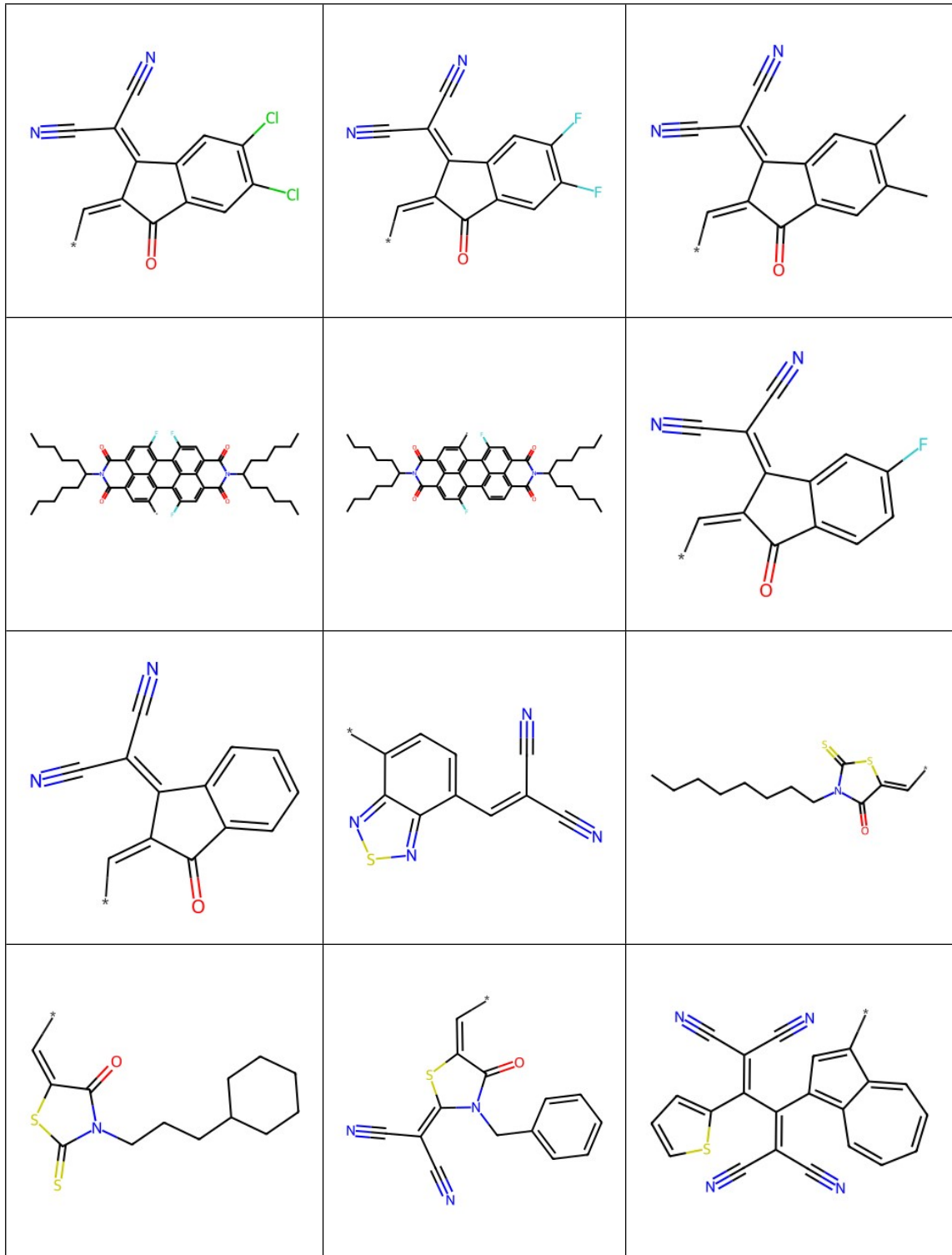
		
		
		
		

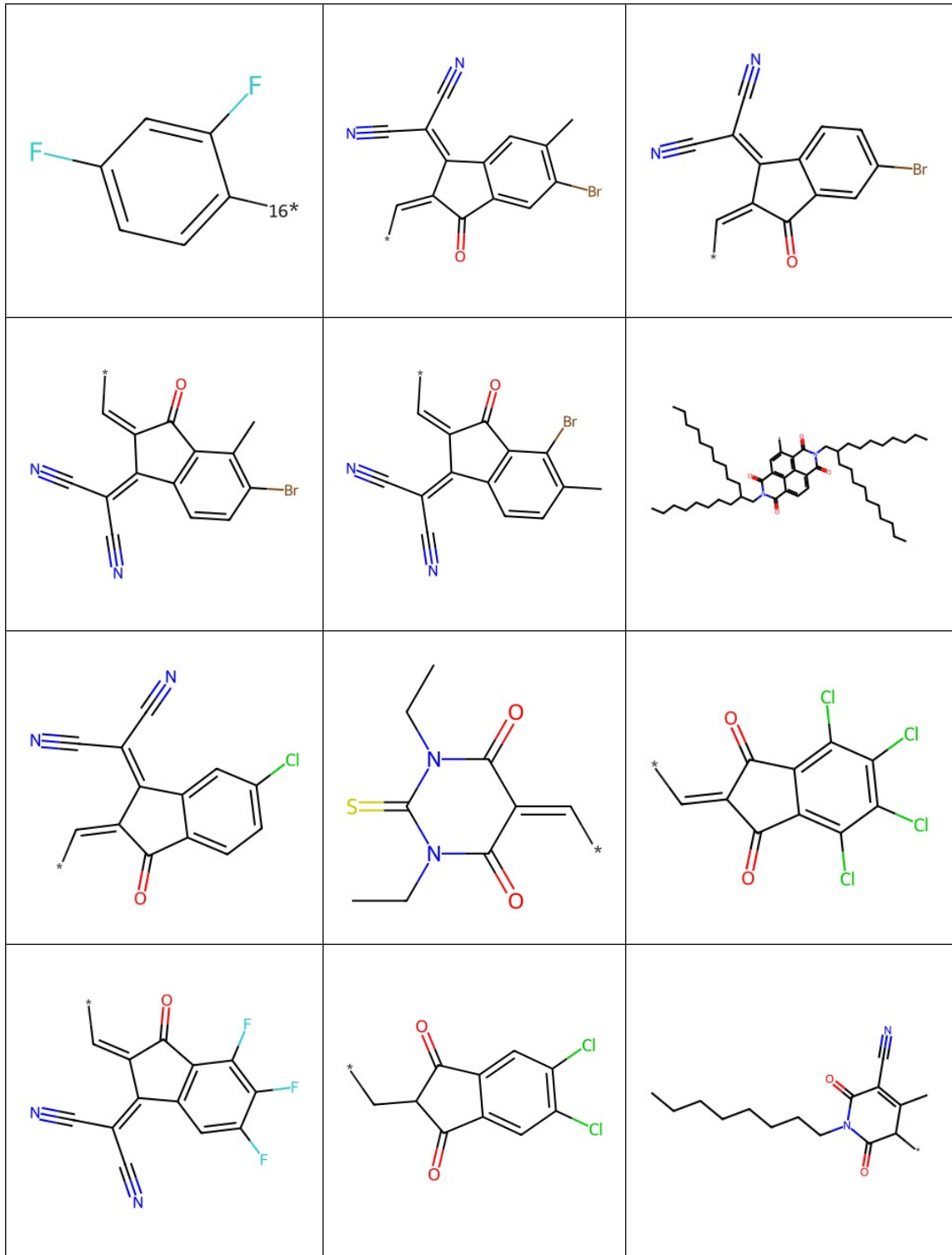


		
		
		
		
$\pi$		

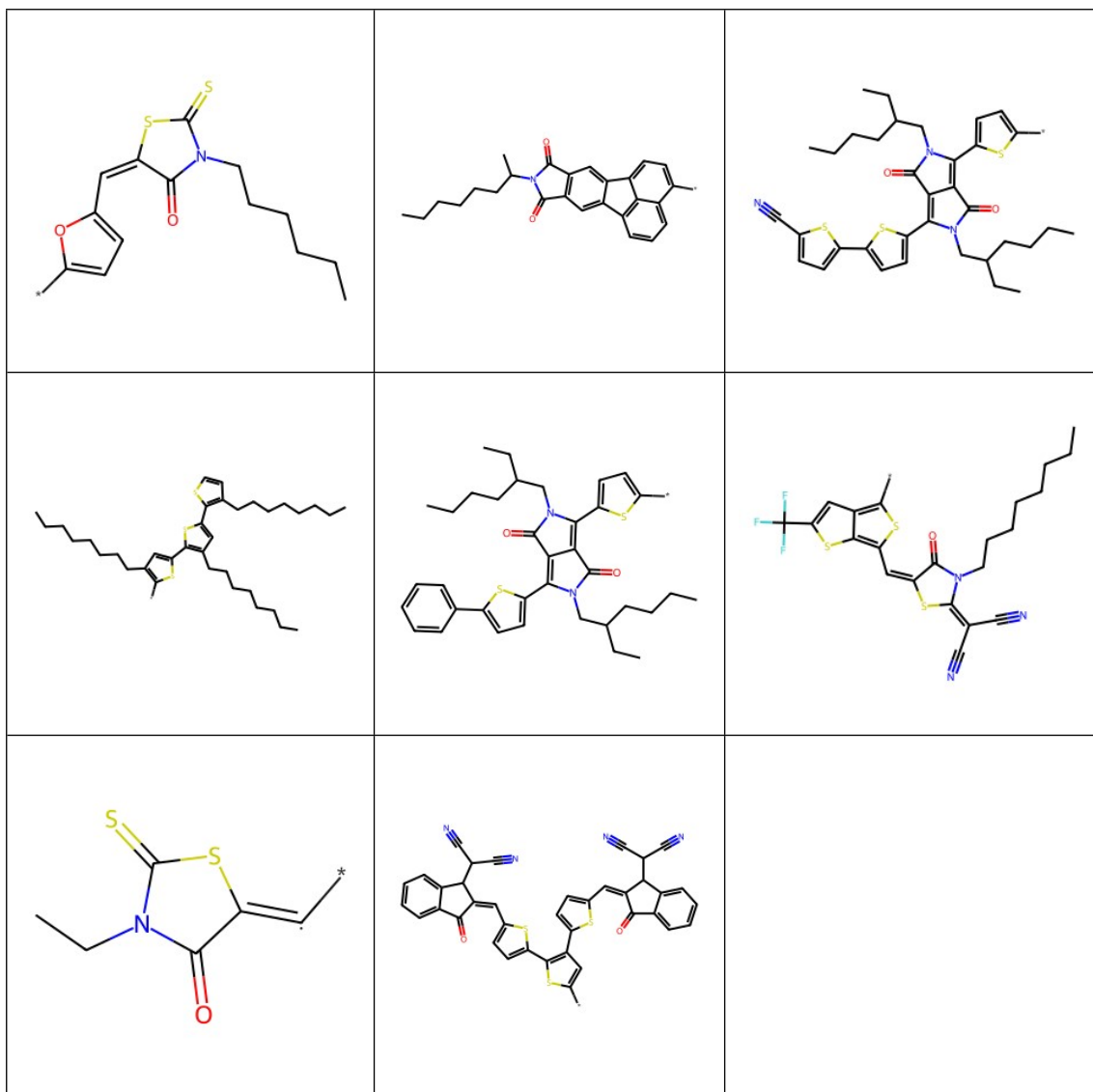
		
		
		
		

		
		
		
		
<p>A</p>		



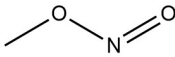
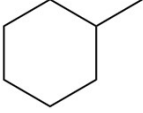
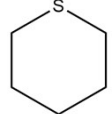
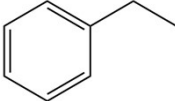
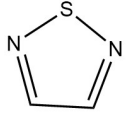
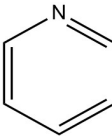





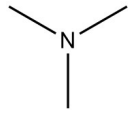
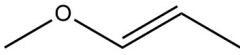
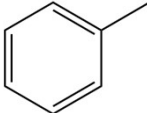
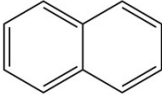
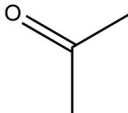
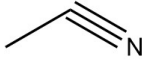
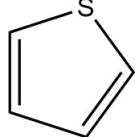
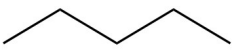


**Fig. S5** Fragments of acceptor molecules

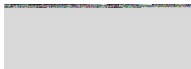
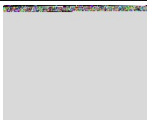
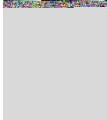
**Table S1** Acceptor descriptors used for LSTM model

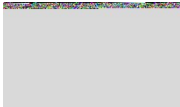
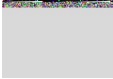

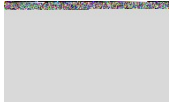
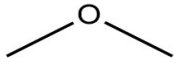

descriptors	resolve	photograph
NOCount_A	Number of nitro groups contained in the acceptor molecule	
NumAliphaticCarbocycles_A	Number of alicyclic alkyl rings in the acceptor molecule	
NumAliphaticRings_A	Number of alicycles in the acceptor molecule	
NumAromaticCarbocycles_A	Number of aromatic cycloalkyl rings in the acceptor molecule	
NumAromaticHeterocycles_A	Number of aromatic heterocycles in the acceptor molecule	
NumHAcceptors_A	Number of hydrogen bond receptors (usually nitrogen, oxygen or sulfur) in the acceptor molecule	
NumHeteroatoms_A	Number of heteroatoms in the acceptor molecule	
NumRotatableBonds_A	Number of rotatable chemical bonds in the acceptor molecule	
fr_Ar_N_A	Number of nitrogen atom groups in the aromatic ring in the acceptor	
fr_C_O_A	Number of carbon-oxygen double bonds in the acceptor	



fr_NH0_A	Number of amino groups without hydrogen atoms in the acceptor	
fr_allylic_oxid_A	Number of allyl oxide groups in the acceptor	
fr_aryl_methyl_A	Number of arylmethyl groups in the acceptor	
fr_bicyclic_A	Number of two or more rings in the acceptor	
fr_halogen_A	Number of halogen groups in the acceptor	
fr_ketone_A	Number of ketone groups in the acceptor	
fr_nitrile_A	Number of nitrile groups in the acceptor	
fr_thiophene_A	Number of thiophene rings in the acceptor	
fr_unbrch_alkane_A	Number of unbranched aliphatic groups in the acceptor	

**Table S2** Donor descriptors used for LSTM model

descriptors	resolve	photograph
NOCount_D	Number of nitro groups contained in the donor molecule	
NumAliphaticCarbocycles_D	Number of alicyclic alkyl rings in the donor molecule	
NumAliphaticRings_D	Number of alicycles in the donor molecule	

NumAromaticCarbocycles_D	Number of aromatic cycloalkyl rings in the donor molecule	
RingCount_D	Number of rings contained in the donor molecule	
fr_C_O_D	Number of carbon-oxygen double bonds in the donor molecule	
fr_NH0_D	Number of amino groups without hydrogen atoms in the donor molecule	
fr_bicyclic_D	Number of two or more rings in the donor molecule	
fr_ether	Number of ether groups in the donor molecule	
fr_halogen_D	Number of halogen groups in the donor molecule	
fr_unbrch_alkane_D	Number of unbranched aliphatic groups in the donor molecule	

**Table S3** Hyperparameters of the LSTM model for PCE,  $V_{OC}$ ,  $J_{SC}$ , and FF predictions

Device parameters	Hyperparameters
-------------------	-----------------

---

PCE  
hidden\_size : 150  
num\_layers : 1  
batch\_size : 60  
l2\_factor : 0.01  
lr : 0.0001  
dropout : 0.1  
patience : 40  
optimizer : Adam

---

J<sub>sc</sub>  
hidden\_size : 90  
num\_layers : 1  
batch\_size : 310  
l2\_factor : 0.1  
lr : 0.01  
dropout : 0.1  
patience : 10  
optimizer : Adam

---

V<sub>oc</sub>  
hidden\_size : 220  
num\_layers : 1  
batch\_size : 10  
l2\_factor : 0.001  
lr : 0.0001  
dropout : 0.2  
patience : 30  
optimizer : Adam

---

FF  
hidden\_size : 90  
num\_layers : 1  
batch\_size : 220  
l2\_factor : 0.1  
lr : 0.01  
dropout : 0.1  
patience : 10  
optimizer : Adam

---

**Table S4**  $V_{OC}$ ,  $J_{SC}$ , and FF predictions for five donor-acceptor pairs within the database using the trained LSTM model

D: A	PM6:L8-BO	PB[N][F]: Y6	PM6:Y18	PTB7-Th: DTC-F-F	PBDB-T:sp- mOEh-ITIC
Experimental $V_{OC}$ (V)	0.88	0.85	0.84	0.84	0.87
Predictive $V_{OC}$ (V)	0.86	0.79	0.88	0.82	0.88
$V_{OC}$ Absolute Error (V)	0.02	0.06	0.04	0.02	0.01
Experimental $J_{SC}$ (mA cm <sup>-2</sup> )	26.49	25.00	24.91	16.01	12.13
Predictive $J_{SC}$ (mA cm <sup>-2</sup> )	23.47	24.08	24.46	12.57	11.30
$J_{SC}$ Absolute Error (mA cm <sup>-2</sup> )	3.02	0.92	0.45	3.44	0.83
Experimental FF (%)	78.92	68.10	76.40	56.00	61.00
Predictive FF (%)	68.87	68.19	67.96	54.03	61.23
FF Absolute Error (%)	10.05	0.09	8.44	1.97	0.23

**Table S5**  $V_{OC}$ ,  $J_{SC}$ , and FF predictions for five donor-acceptor pairs outside the database using the trained LSTM model

D: A	D18:L8-BO	PTQ10: ITIC-4F	PTB7-Th: Y6	PM6:ID- C6Ph-4F	PffBT4T- 2OD:P(4CF 8CH-PDI- TT)
Experimental $V_{OC}$ (V)	0.90	0.95	0.69	0.84	0.74
Predictive $V_{OC}$ (V)	0.81	0.79	0.79	0.86	0.86
$V_{OC}$ Absolute Error (V)	0.09	0.16	0.10	0.02	0.12
Experimental $J_{SC}$ (mA cm <sup>-2</sup> )	25.10	19.81	23.40	17.47	8.72
Predictive $J_{SC}$ (mA cm <sup>-2</sup> )	22.12	17.55	23.11	18.84	8.73
$J_{SC}$ Absolute Error (mA cm <sup>-2</sup> )	2.98	2.26	0.29	1.37	0.01
Experimental FF (%)	71.20	61.00	67.80	73.22	52.99
Predictive FF (%)	65.54	62.19	64.48	65.65	43.42
FF Absolute Error (%)	5.66	1.19	3.32	7.57	9.57