Supplementary Information

BiBERTa: A Self-Supervised Framework for Accelerating the Discovery of Stable Organic Photovoltaic Materials

Jinyu Sun, Dongxu Li, Jie Zou, Xiaoxiao Tan, Yue Wang, Hailiang Zhang, Yingping Zou, Zhimin Zhang* and Hongmei Lu*

College of Chemistry and Chemical Engineering, Central South University, Changsha, 410083, P. R. China.

E-mail: zmzhang@csu.edu.cn, hongmeilu@csu.edu.cn

Table of Contents

Supplementary Figures

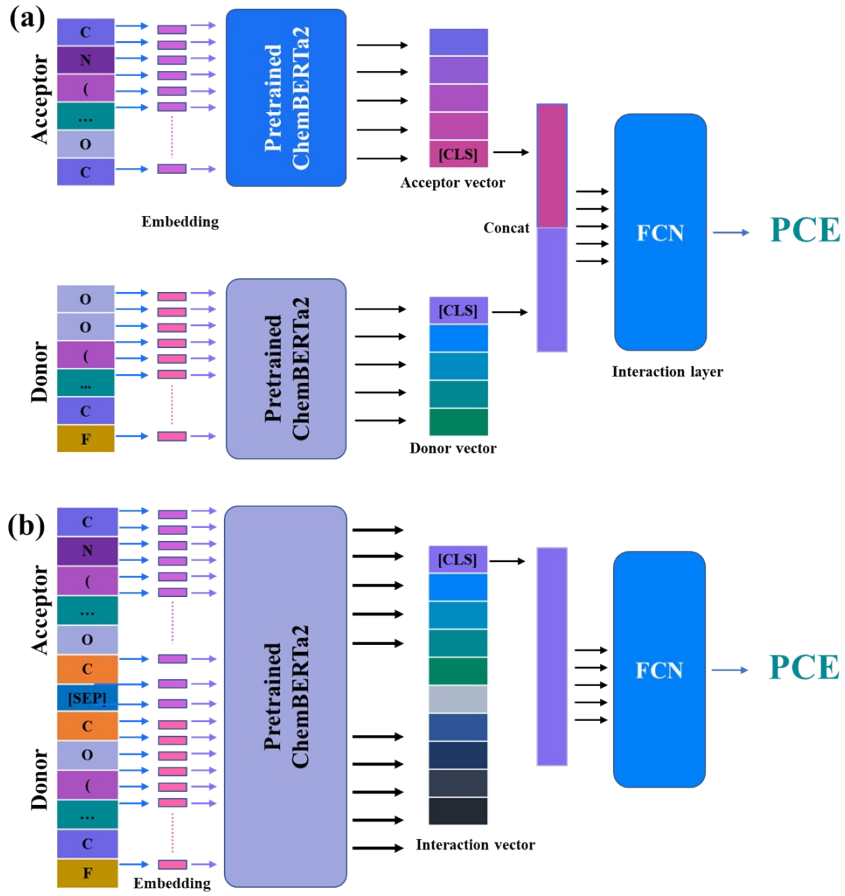


Figure S1. The schematic diagram of (a) the bi-encoder model and (b) the cross-encoder model.

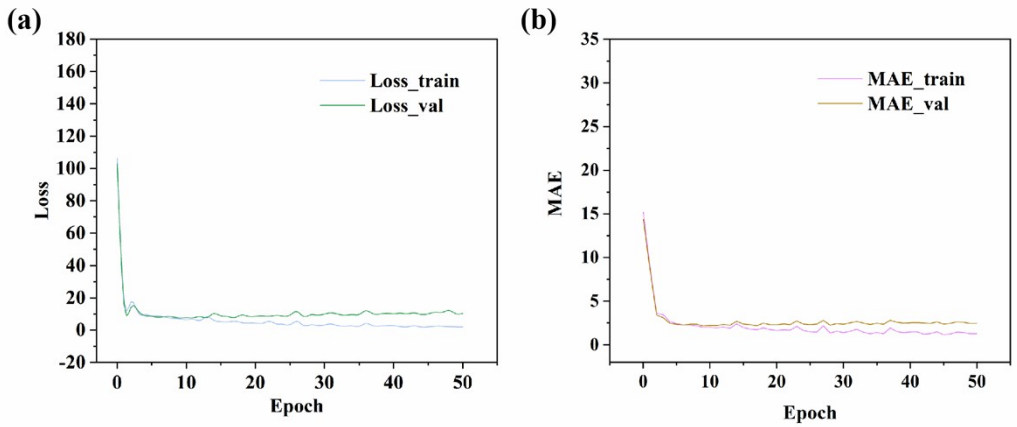


Figure S2. The learning curves of the BiBERTa model. (a) The loss curves and (b) MAE of the model.

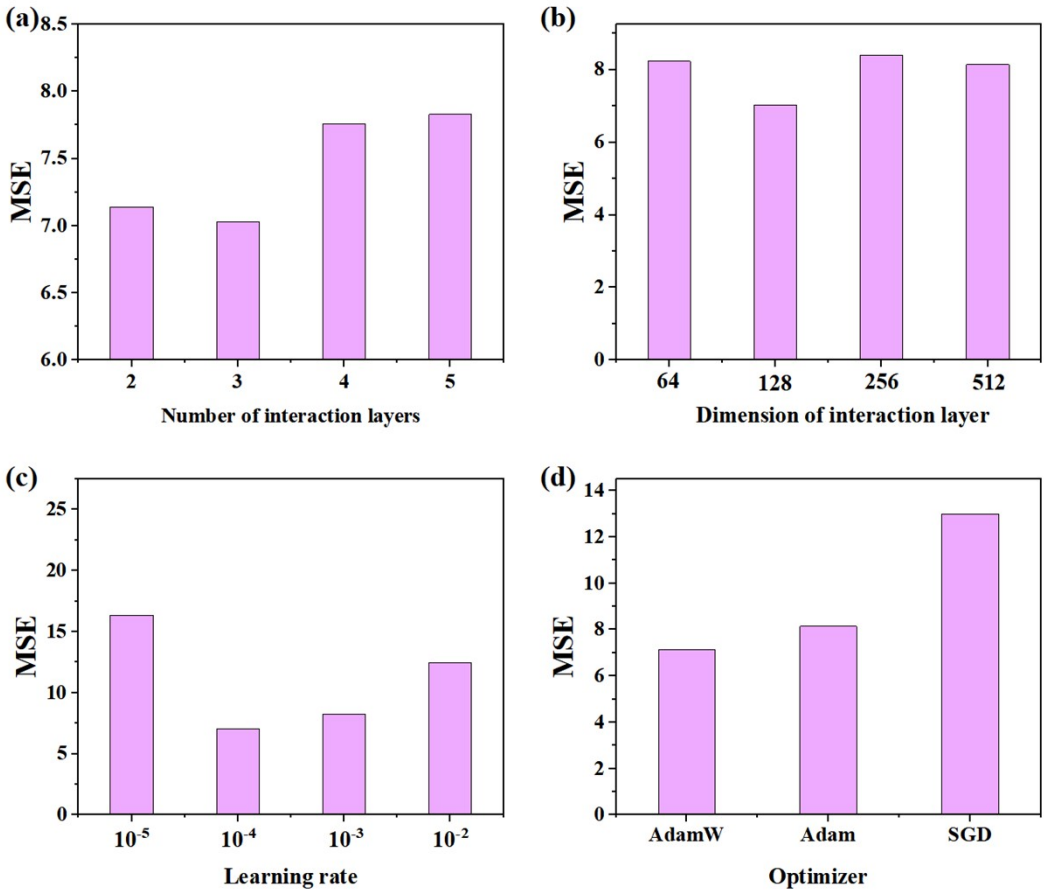


Figure S3. The optimization of (a) the number of interaction layers, (b) the dimension of interaction layers, (c) the learning rate, and (d) the type of optimizer on the validation set.

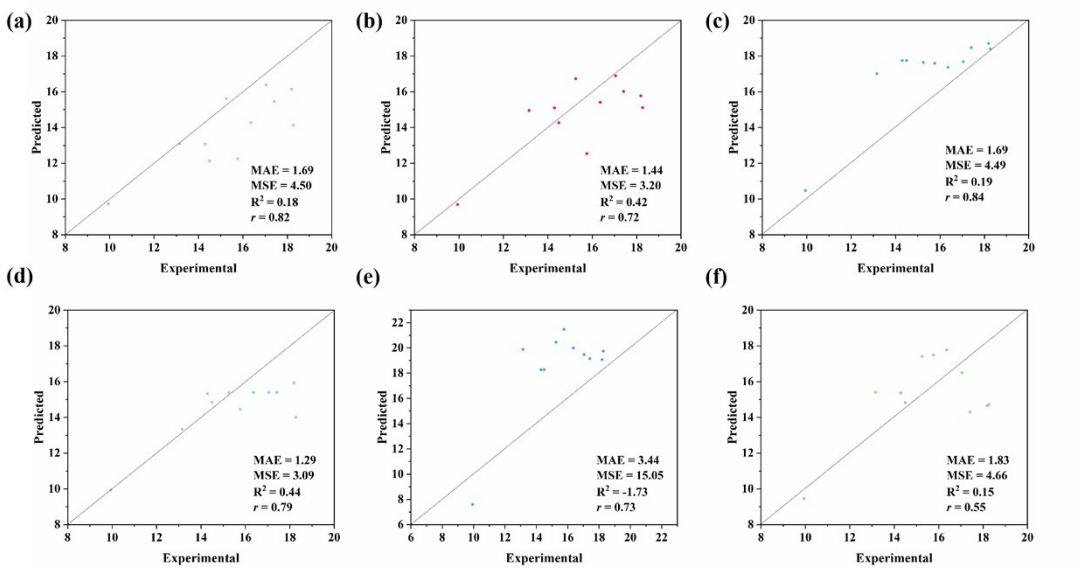


Figure S4. The prediction results on the QM dataset. (a) RF with one-hot encoding. (b) RF with ECFP6. (c) BERT. (d) LightGBM with ECFP6. (e) GNN. (f) CNN with one-hot encoding.

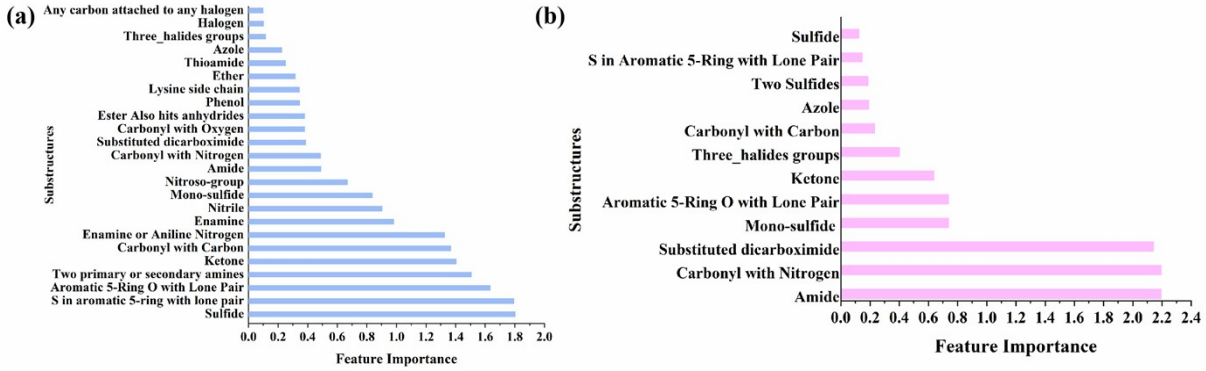
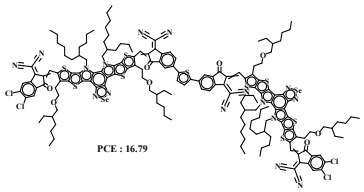
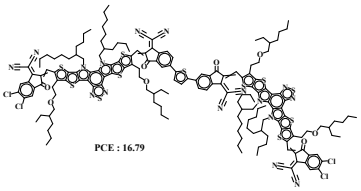


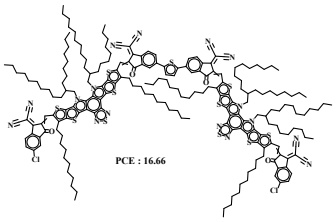
Figure S5. The feature importance analysis of (a) the acceptor encoder and (b) the donor encoder.



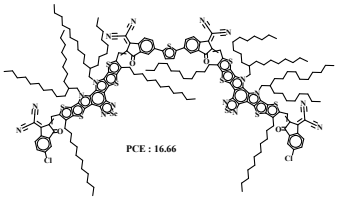
PCE : 16.79



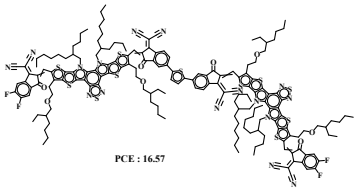
PCE : 16.79



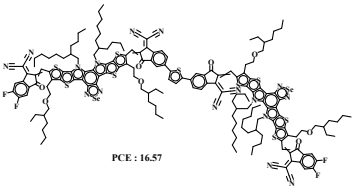
PCE : 16.66



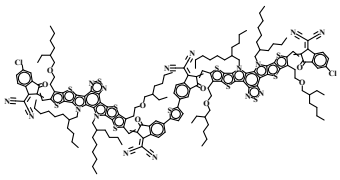
PCE : 16.66



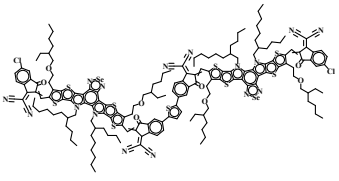
PCE : 16.57



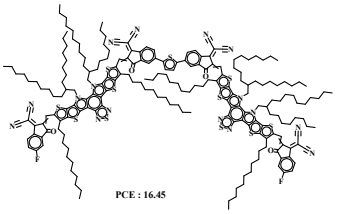
PCE : 16.57



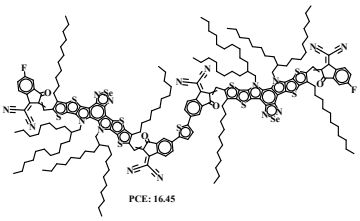
PCE : 16.47



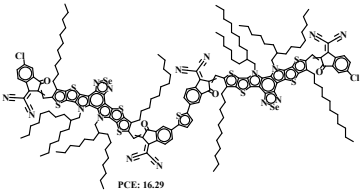
PCE : 16.57



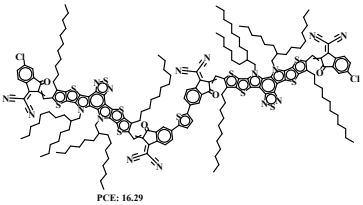
PCE : 16.45



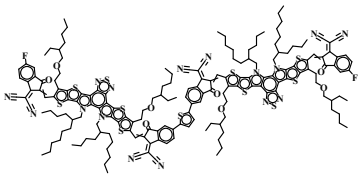
PCE: 16.45



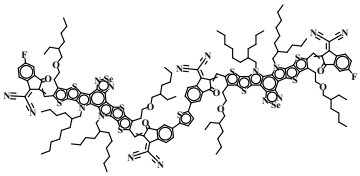
PCE: 16.29



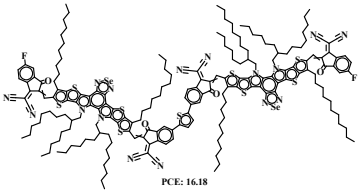
PCE: 16.29



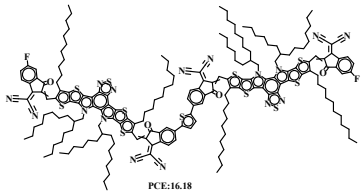
PCE: 16.28



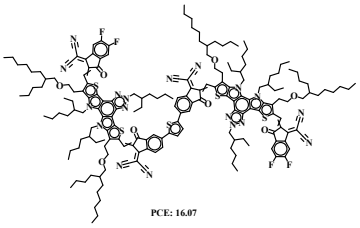
PCE: 16.28



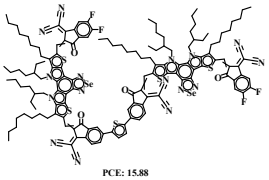
PCE: 16.18



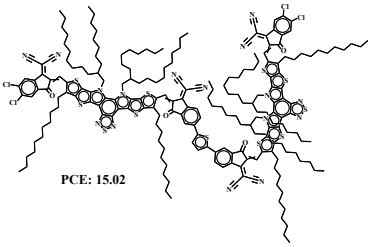
PCE:16.18



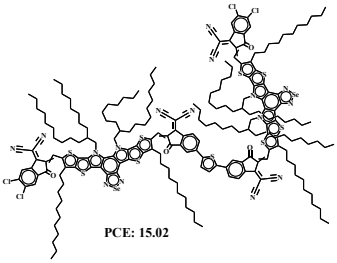
PCE: 16.07



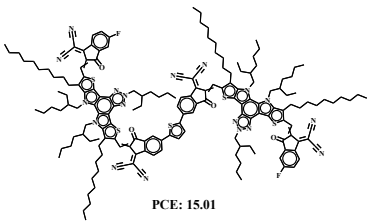
PCE: 15.88



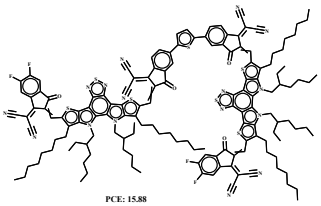
PCE: 15.02



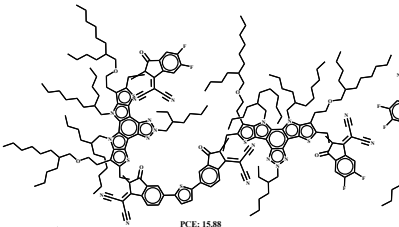
PCE: 15.02



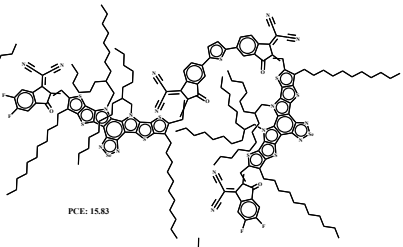
PCE: 15.01



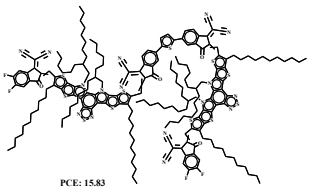
PCE: 15.88



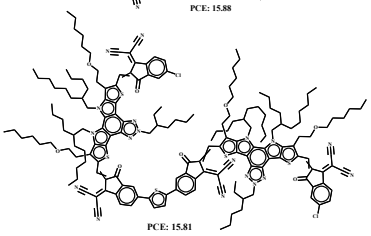
PCE: 15.88



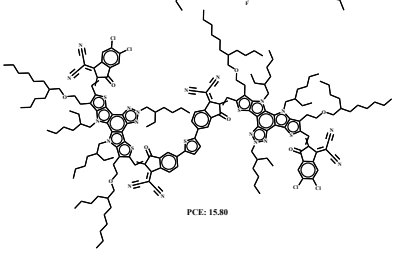
PCE: 15.83



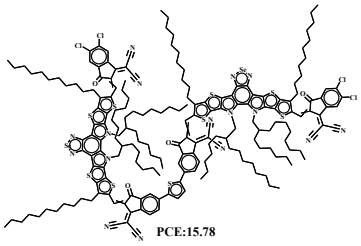
PCE: 15.83



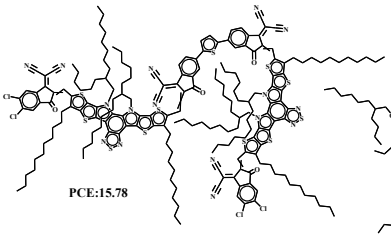
PCE: 15.81



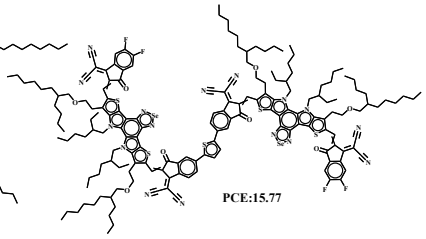
PCE: 15.80



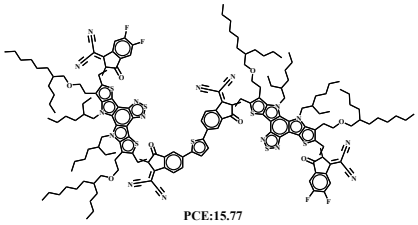
PCE:15.78



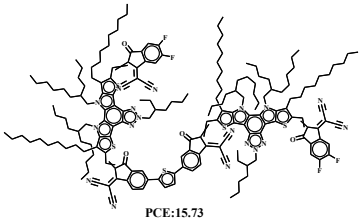
PCE:15.78



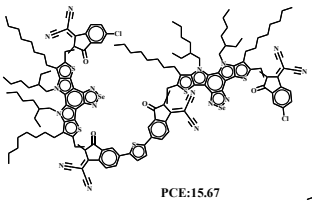
PCE:15.77



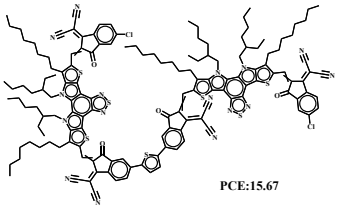
PCE:15.77



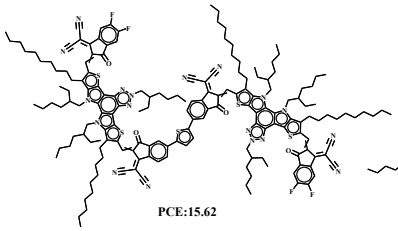
PCE:15.73



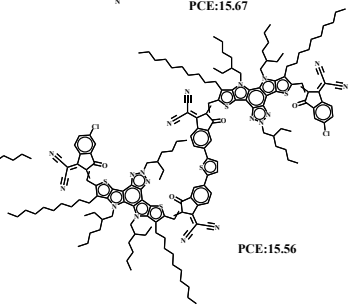
PCE:15.67



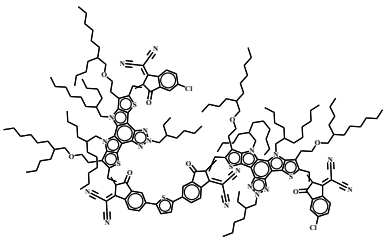
PCE:15.67



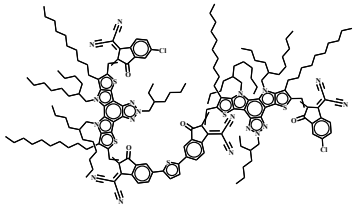
PCE:15.62



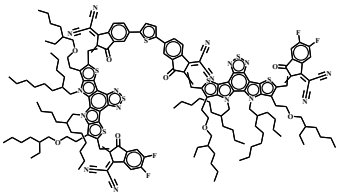
PCE:15.56



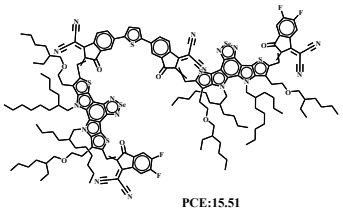
PCE:15.55



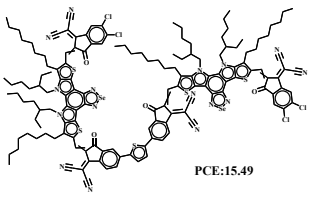
PCE:15.53



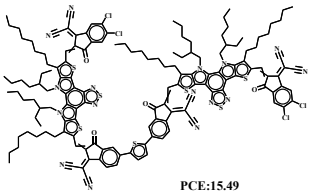
PCE:15.51



PCE:15.51



PCE:15.49



PCE:15.49

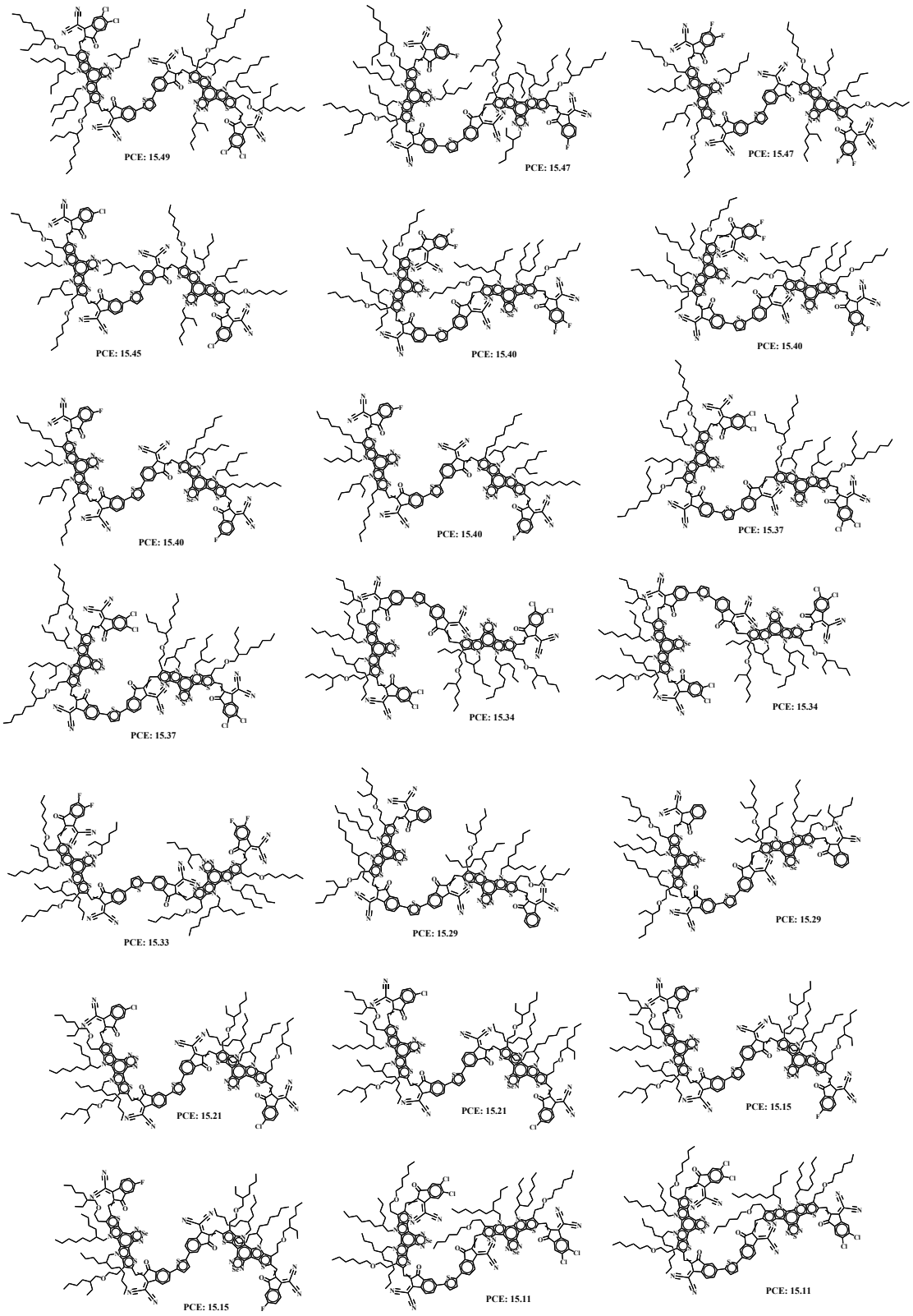


Figure S6. [1] The generated new quasi-macromolecule acceptors with predicted PCE> 15.0% by BiBERTa.

[1] The dataset is available at https://github.com/JinYSun/BiBERTa/blob/branch/BiBERTa/dataset/QM.csv

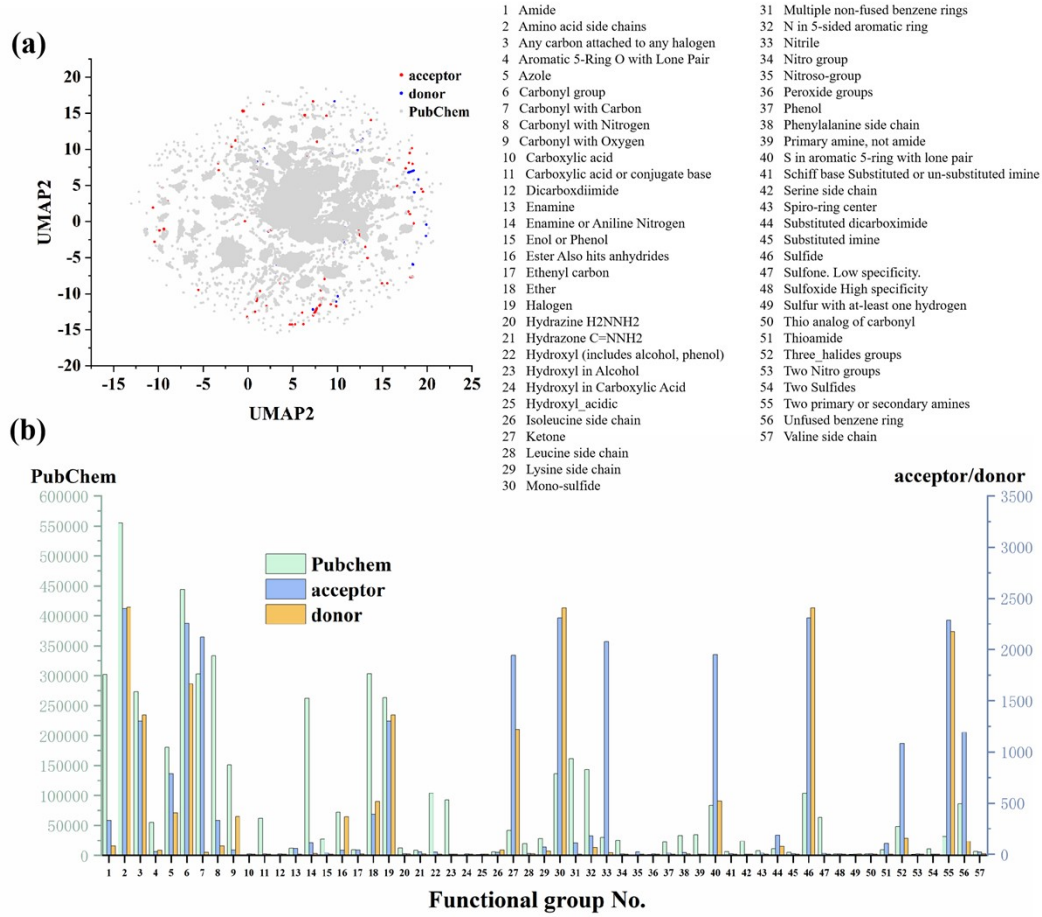


Figure S7. The structure and function group distribution of the pretraining and fine-tuning dataset. (a) The UMAP low dimensional embedding of the compounds (randomly selected 10% of the compounds from the pretraining dataset) in PubChem and donors/acceptors in the fine-tuning dataset. (b) The statistics of the functional group of the compounds in PubChem (randomly selected 10% of the compounds from the pretraining dataset) and donors/acceptors in the fine-tuning dataset.

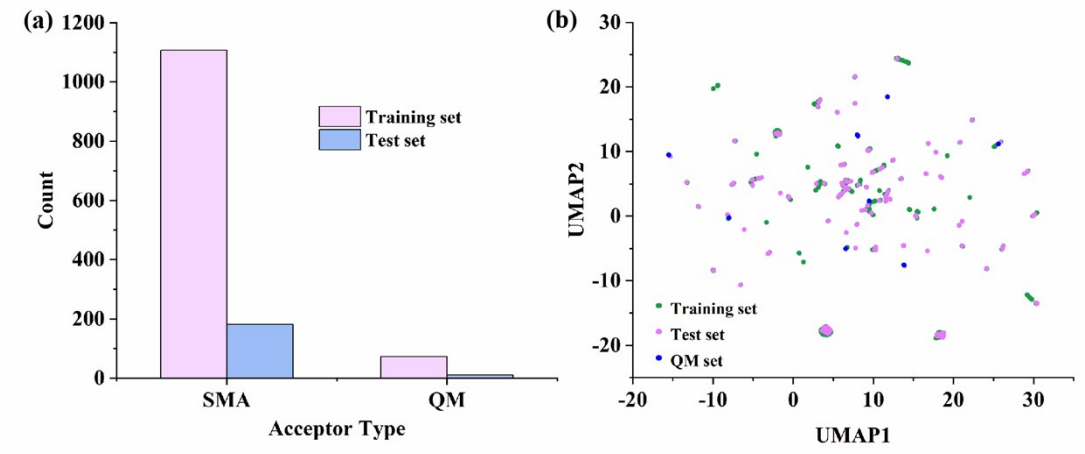


Figure S8. The proportion of (a) acceptor type and (b) structure distribution of the training, test, and QM datasets

Supplementary Tables

Table S1. The comparison of the bi-encoder and cross-decoder on the validation set.

	MAE	MSE	R^2	r
Bi-encoder	1.95	7.37	0.62	0.83
Cross-encoder	2.43	8.23	0.50	0.71

Table S2. The structures of terminals (T) and cores (S).

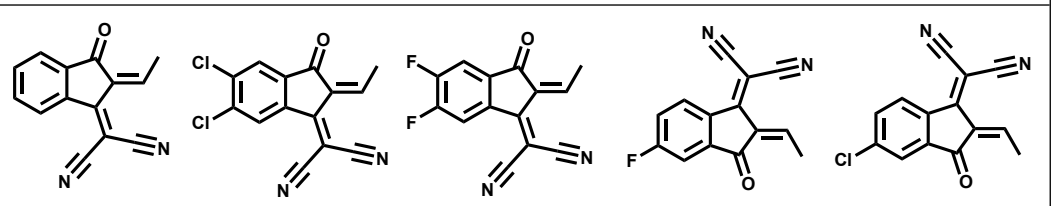
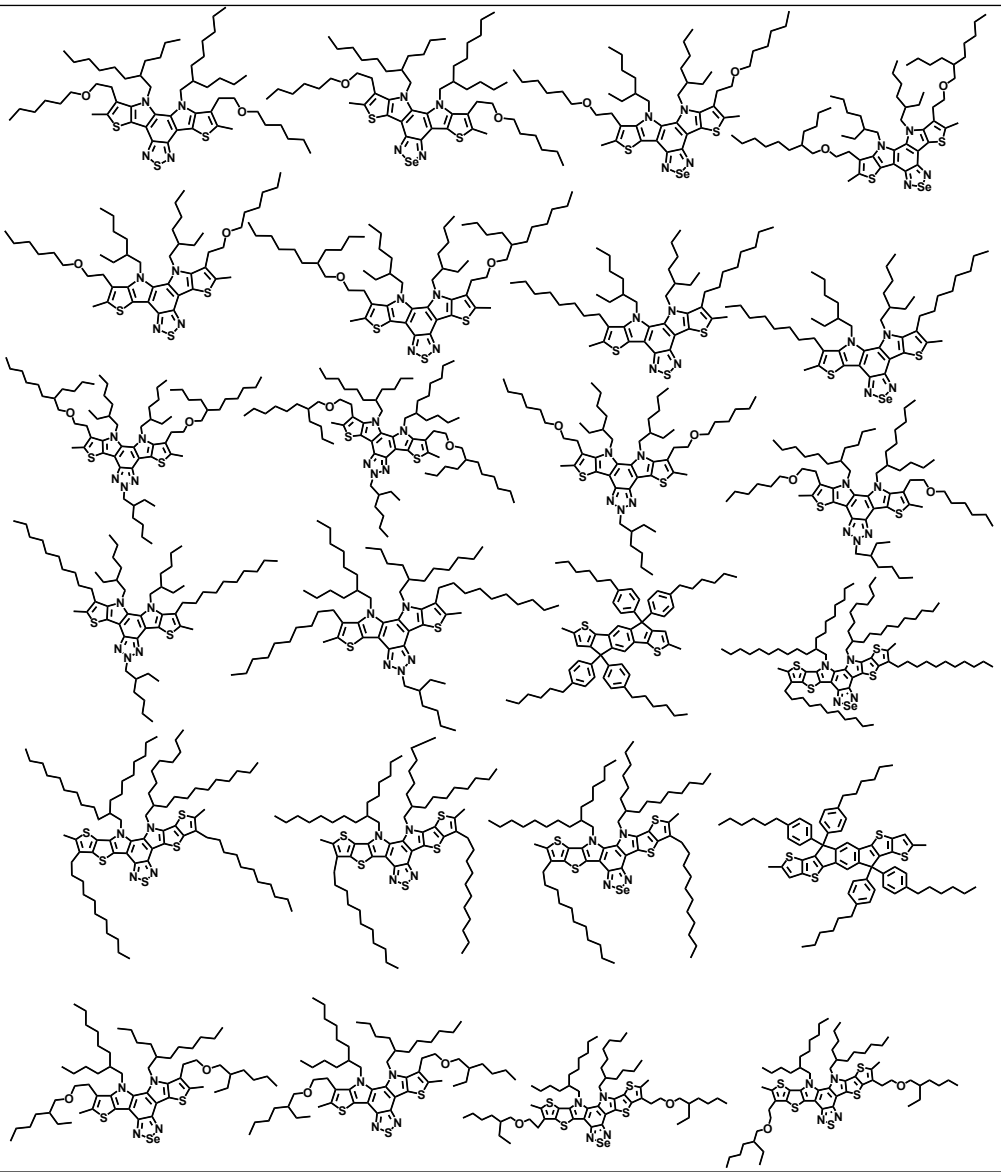
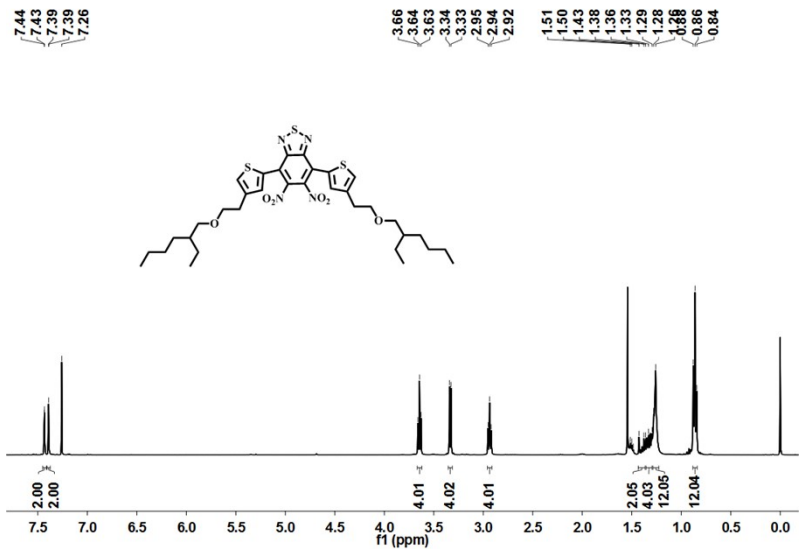
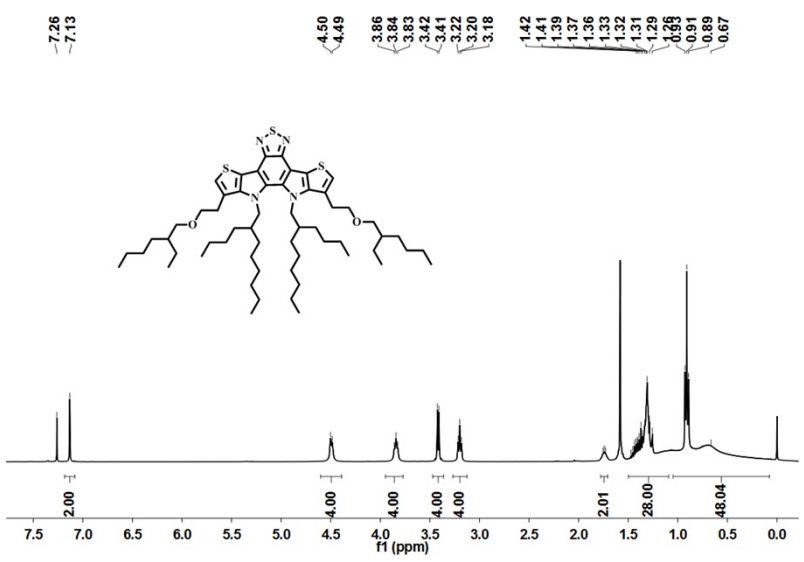
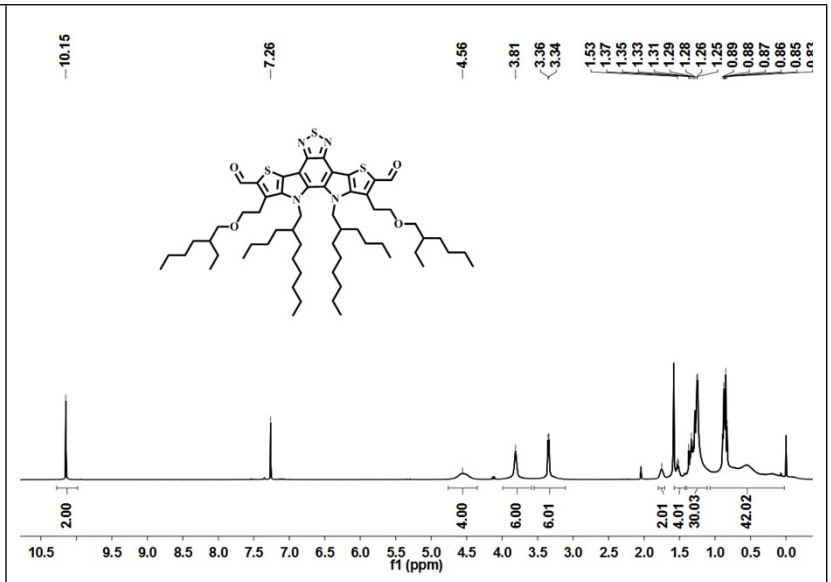
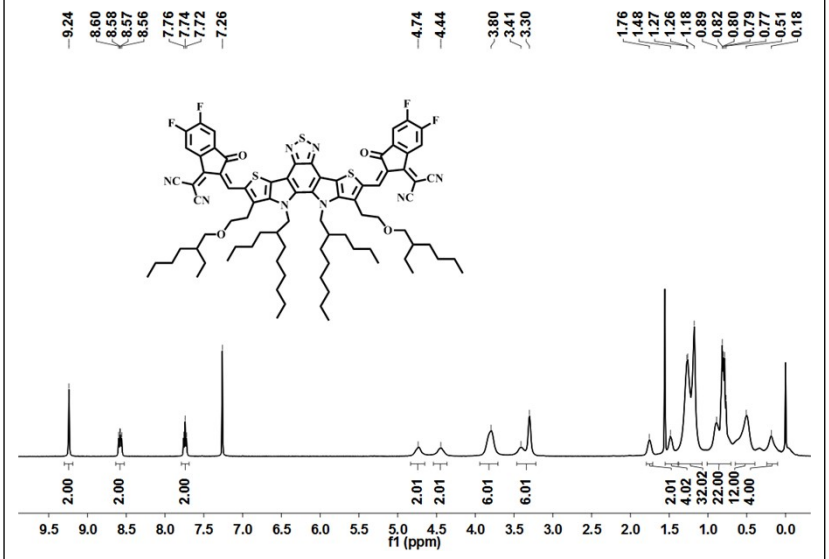
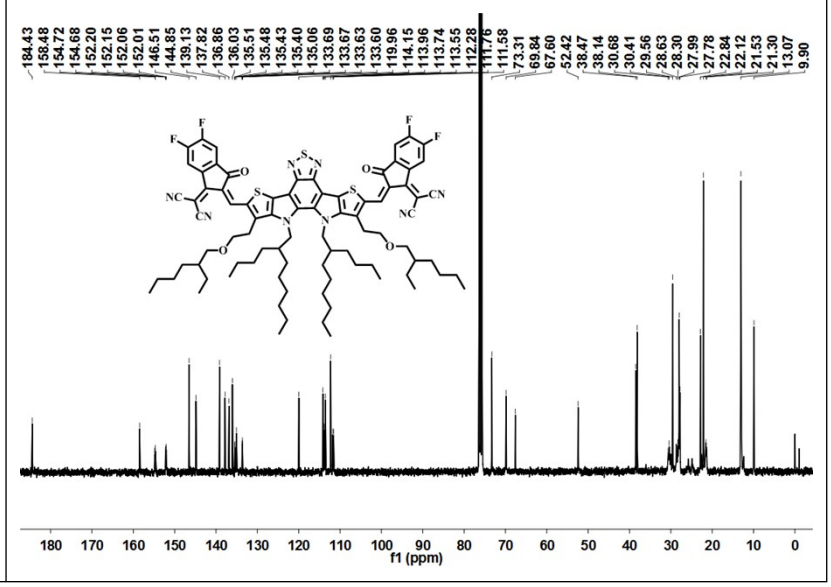
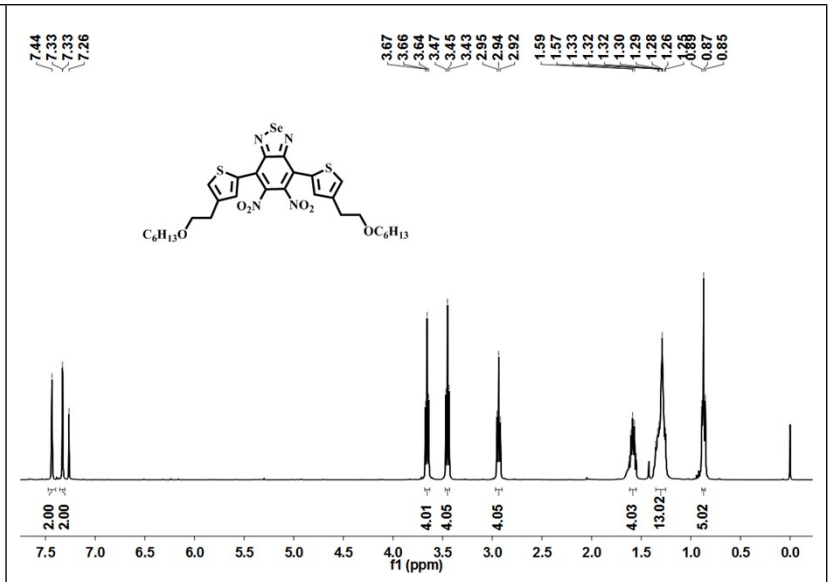
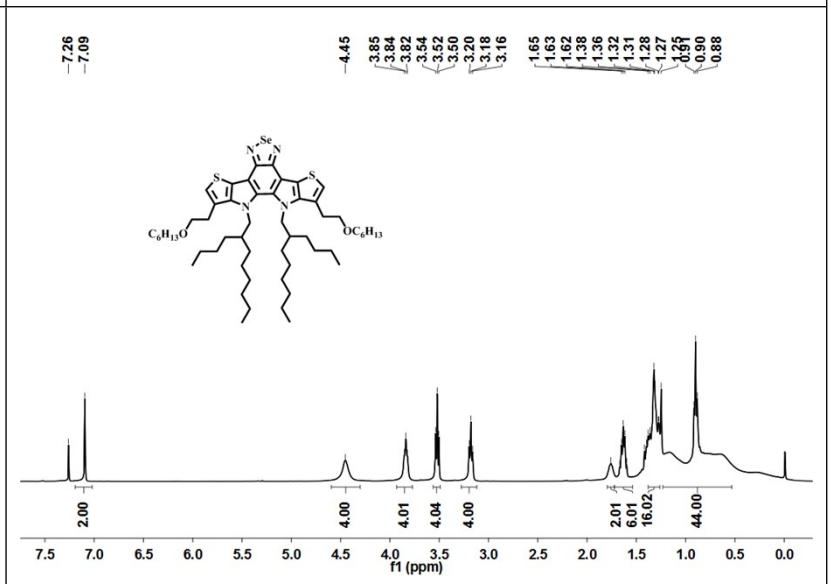
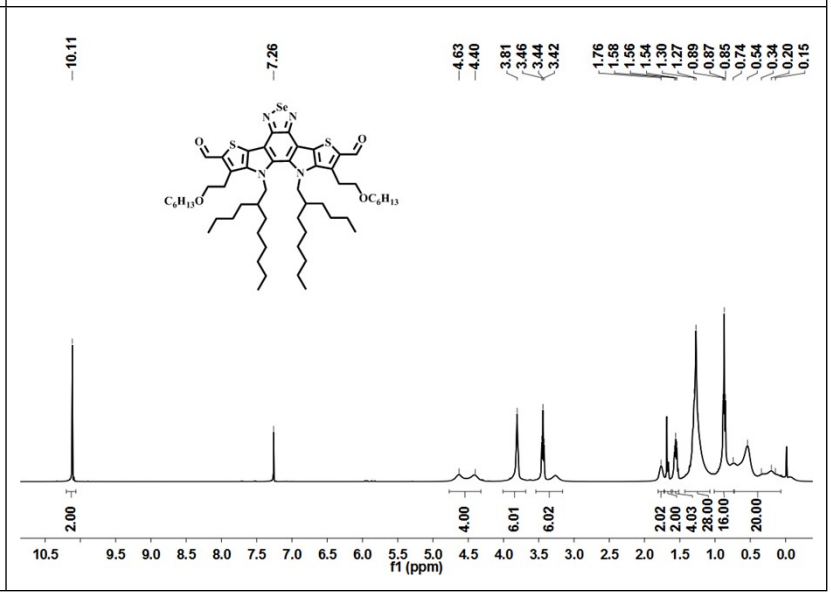
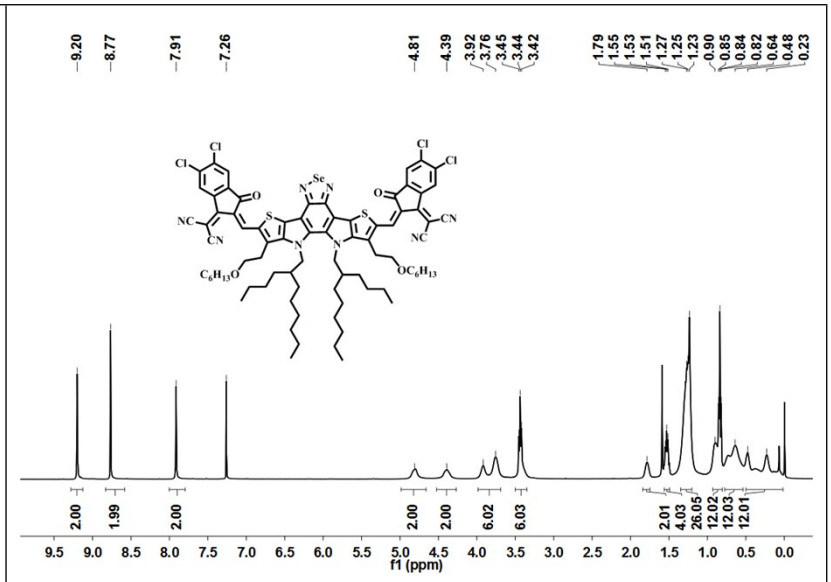
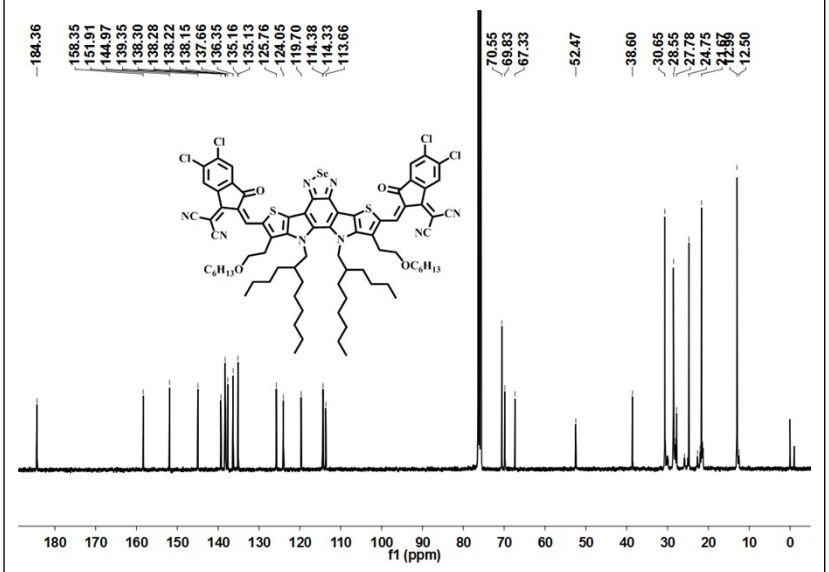
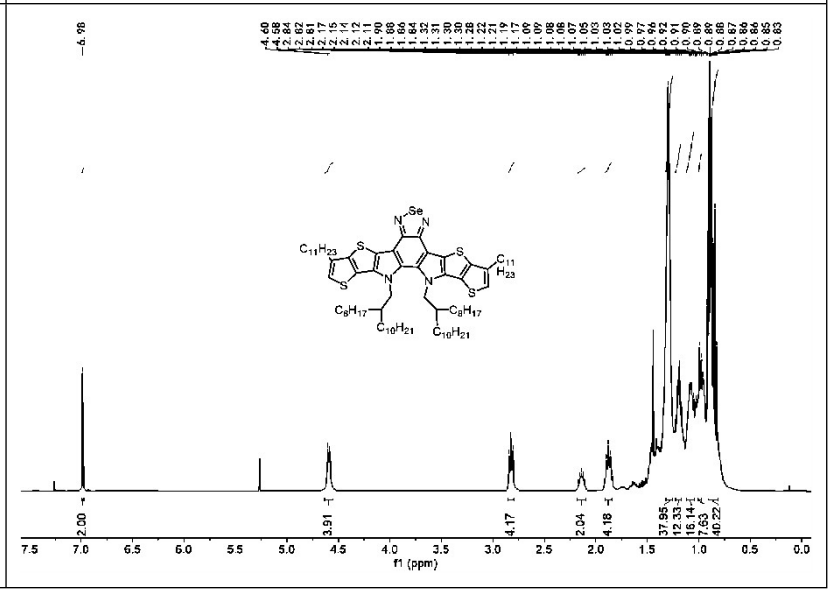
	Structures
T	
C	

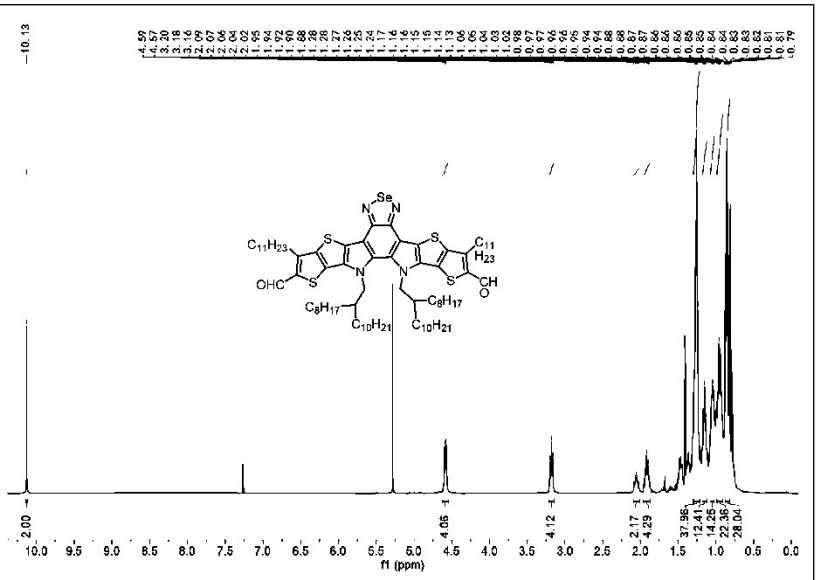
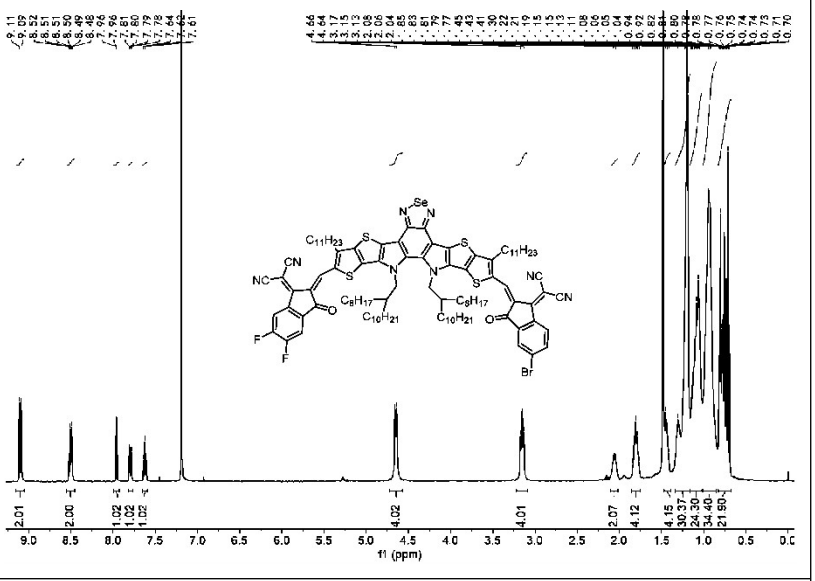
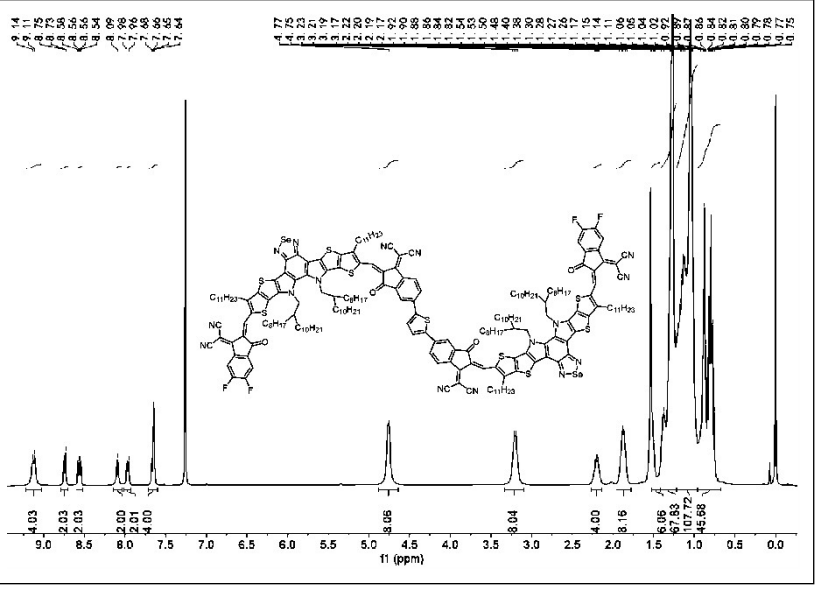
Table S3. ^{1}H NMR spectra of intermediates and final acceptors and the ^{13}C NMR spectra of the acceptors.

Compound	Type	Spectra
Intermediate 3	^{1}H NMR	
Intermediate 4	^{1}H NMR	

Intermediate 5	^{1}H NMR	
Acceptor 1	^{1}H NMR	
Acceptor 1	^{13}C NMR	

Intermediate 8	^{1}H NMR	
Intermediate 9	^{1}H NMR	
Intermediate 10	^{1}H NMR	

Acceptor 2	^{1}H NMR	
Acceptor 2	^{13}C NMR	
Intermediate 12	^{1}H NMR	

Intermediate 13	^{1}H NMR	
Intermediate 14	^{1}H NMR	
Acceptor 3	^{1}H NMR	

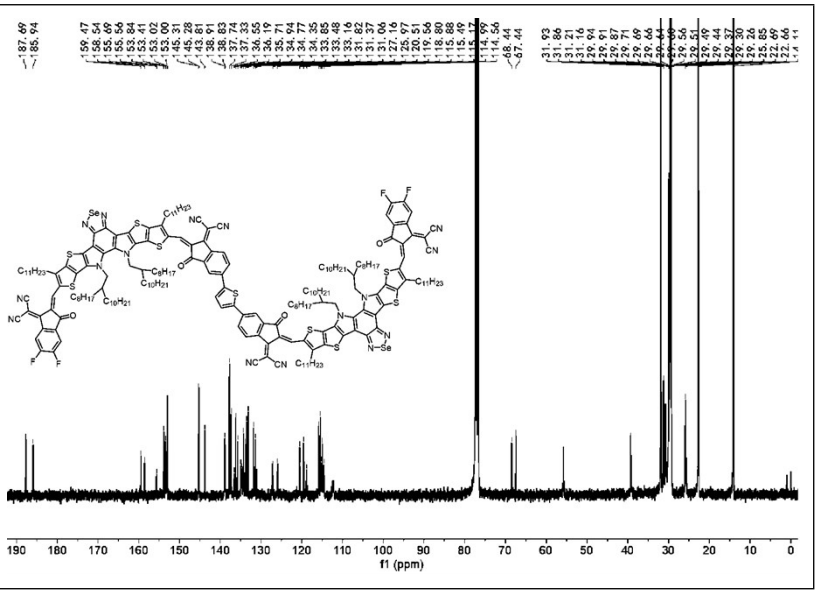
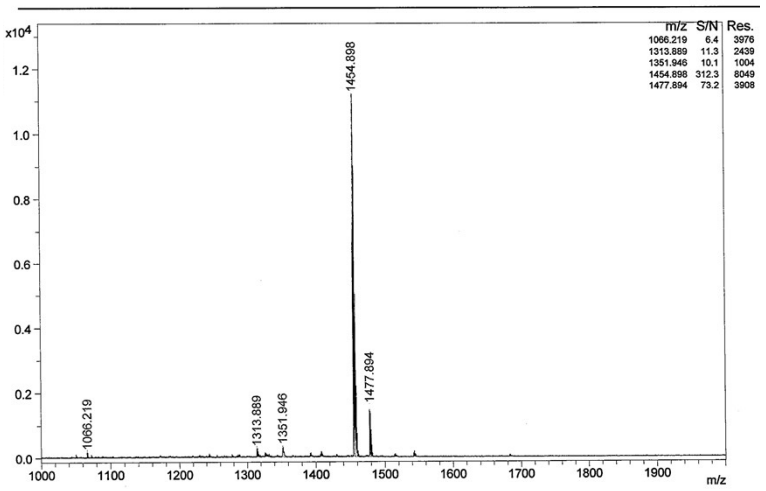
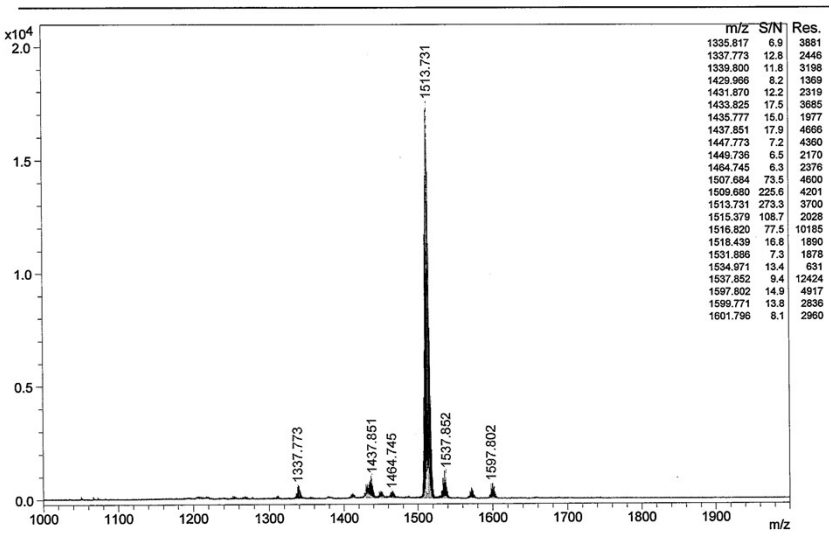
Acceptor 3	^{13}C NMR	

Table S4. Mass spectra of the acceptors.

Compound	Spectra
Acceptor 1	
Acceptor 2	

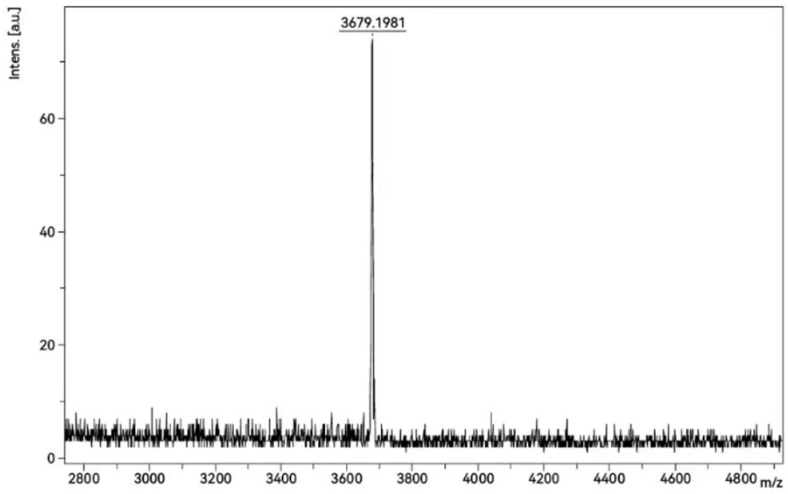
Acceptor 3	

Table S5. The comparison of different models.

		Experiment	BiBERTa	DeepAcceptor	GNN	Scharber model
Time usage			<1 s	~1 s	<1 s	~9h
Prediction results	Acceptor 1	12.91%	13.90%	15.02%	15.50%	7.52%
	Acceptor 2	13.60%	12.59%	15.66%	15.28%	10.20%
	Acceptor 3	15.15%	16.79%	12.49%	17.74%	10.38%

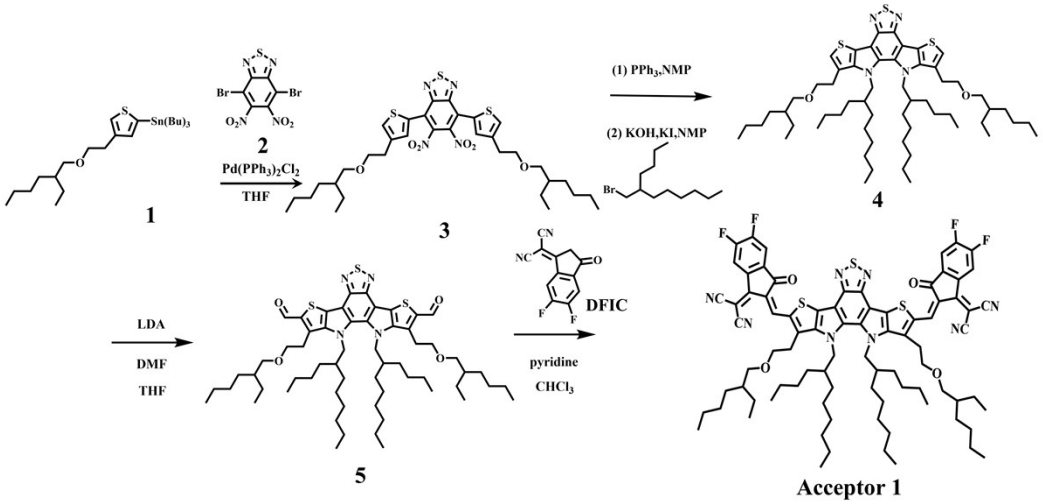
Supplementary Notes

Supplementary Note 1

Materials

All chemicals and reagents, unless otherwise specified, were purchased from Energy Chemical and Innochem. They were used without further purification. Compounds 1, 6, and 11 were synthesized according to the previous literature.[1-3]

General synthetic procedures of the three candidate acceptors



Scheme S1. The synthetic procedure of Acceptor 1.

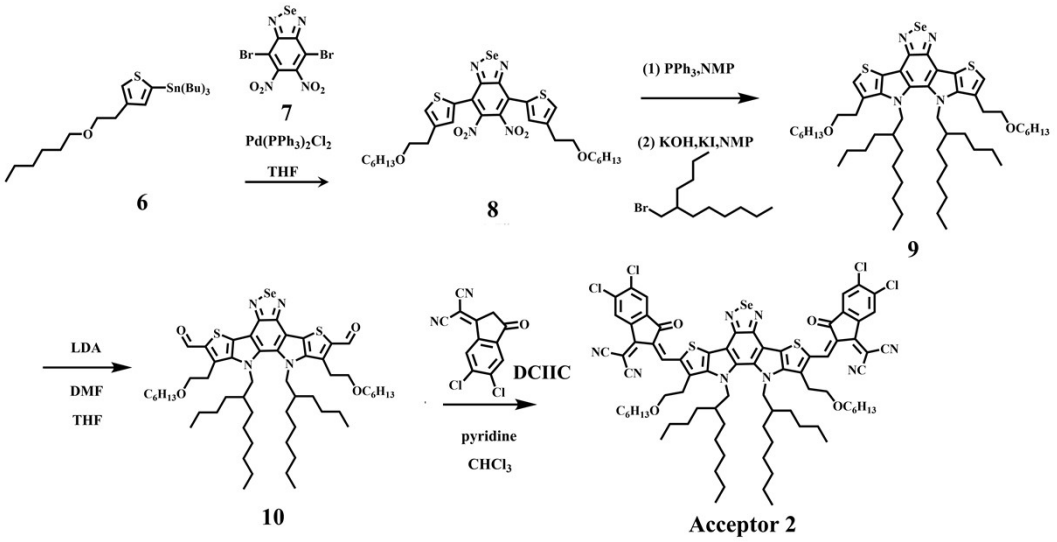
Compound 3: Compound 2 (2.00g, 3.5 mmol), compound 1 (6.89g, 13.02 mmol), and $Pd(PPh_3)_2Cl_2$ (0.18g, 0.26 mmol) were dissolved in THF (50 mL) and stirred at 75 ℃ for 3 h under argon (Ar) atmosphere. The solvent was removed after the reaction. The mixture was cooled to room temperature. It was purified by chromatography in a silica gel column and eluted with petroleum ether (PE) /dichloromethane (DCM) (1/1, v/v) to obtain Compound 3 (1.85g, yield: 73%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 (d, J = 1.1 Hz, 2H), 7.39 (d, J = 1.3 Hz, 2H), 3.64 (t, J = 6.6 Hz, 4H), 3.33 (d, J = 6.0 Hz, 4H), 2.94 (t, J = 6.5 Hz, 4H), 1.40 (d, J = 19.7 Hz, 2H), 1.33 (s, 4H), 1.27 (d, J = 7.8 Hz, 12H), 0.86 (t, J = 7.4 Hz, 12H).

Compound 4: Under argon atmosphere, compound 3 (2.0g, 2.85 mmol) and PPh_3(7.46g, 28.50 mmol) were dissolved in N-Methylpyrrolidone (NMP, 50 mL). The mixture was reacted at 180 ℃ for 24h. After cooling to room temperature, potassium hydroxide (2.39g, 42.75 mmol), potassium iodide (0.19g, 1.14 mmol), and 5-(Bromomethyl)-undecane (10.66g, 42.75 mmol) were added to the mixture. The mixture was reacted at 95 ℃ for 24h under an argon atmosphere. The product was extracted with DCM. After removing the solvent, the crude product was purified by column chromatography and eluted with PE/DCM(4/1, v/v) in a silica gel column to obtain Compound 4 (1.10 g, yield: 40%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.13 (s, 2H), 4.49 (d, J = 7.0 Hz, 4H), 3.84 (t, J = 6.6 Hz, 4H), 3.42 (d, J = 6.1 Hz, 4H), 3.20 (t, J = 7.0 Hz, 4H), 1.78 – 1.71 (m, 2H), 1.50 – 1.09 (m, 28H), 1.05 – 0.07 (m, 48H).

Compound 5: Under an argon atmosphere, the mixture of compound 4 (0.60g, 0.62 mmol) in superdry THF (50 mL) at -78 ℃ was stirred for 10 min. LDA (4.6 mL, 9.22 mmol) was added dropwise into the solution. Superdry DMF (10.82 mL) was added after stirring for another 1.5h. The mixture was automatically rewarmed to -40 °C, and

the reaction was quenched with water and extracted with EA. After removing the solvent, the crude product was purified by column chromatography and eluted with PE/EA(10/1, v/v) in a silica gel column to obtain compound 5 (0.45g, yield: 71%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.15 (s, 2H), 4.56 (s, 4H), 3.81 (s, 6H), 3.35 (d, J = 6.0 Hz, 6H), 1.75 (s, 2H), 1.57 – 1.42 (m, 4H), 1.31 (ddd, J = 18.3, 14.2, 4.9 Hz, 30H), 0.86 (dt, J = 11.6, 6.9 Hz, 42H).

Acceptor 1: Under an argon atmosphere, compound 5 (0.12g, 0.11 mmol) and 2-(5,6-difluoro-3-oxo-2,3-dihydro-1H-inden-1-ylidene)malononitrile (0.11g, 0.47 mmol) were dissolved in chloroform (30 mL). Pyridine (1.0 mL) was added at 40 ℃. The solution was poured into 300 mL of methanol and filtered after reacting for 3h. The residue was purified in a silica gel column using petroleum ether/dichloromethane (2/3, v/v) as the eluent. Acceptor 1 was obtained as a dark blue solid (0.10g, yield: 62%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.24 (s, 2H), 8.58 (dd, J = 9.6, 6.5 Hz, 2H), 7.74 (t, J = 7.3 Hz, 2H), 4.74 (s, 2H), 4.44 (s, 2H), 3.80 (s, 6H), 3.36 (d, J = 43.9 Hz, 6H), 1.76 (s, 2H), 1.48 (s, 4H), 1.37 – 1.08 (m, 32H), 1.01 – 0.70 (m, 22H), 0.51 (s, 12H), 0.18 (s, 4H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 184.43, 158.48, 154.84, 154.82, 154.78, 154.72, 154.68, 154.59, 152.20, 152.15, 152.06, 152.01, 146.51, 144.85, 139.13, 137.82, 136.86, 136.03, 135.51, 135.48, 135.43, 135.40, 135.06, 133.69, 133.67, 133.63, 133.60, 119.96, 114.15, 113.96, 113.74, 113.55, 112.28, 111.76, 111.58, 73.31, 69.84, 67.60, 52.42, 38.47, 38.14, 30.68, 30.41, 29.93, 29.56, 28.63, 28.30, 27.99, 27.78, 22.84, 22.50, 22.12, 21.53, 21.30, 13.07, 9.90.



Scheme S2. The synthetic procedure of Acceptor 2.

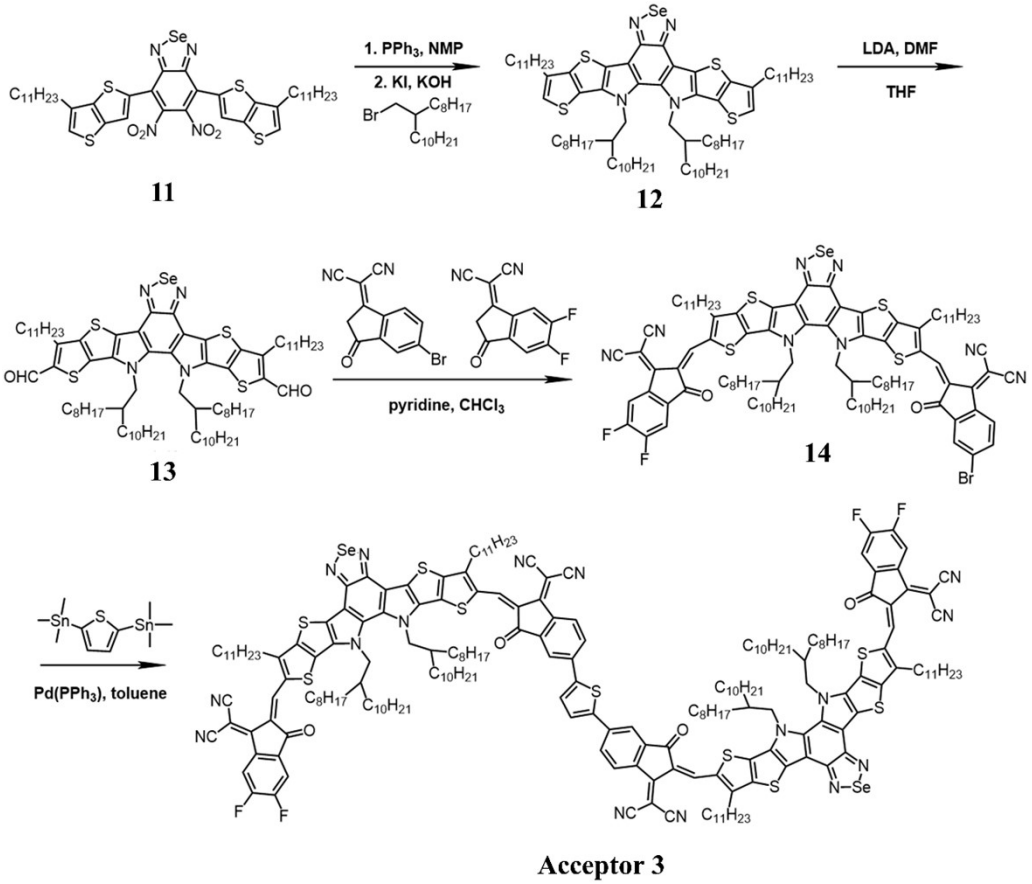
Compound 8: Compound 6 (5.9g, 11.77 mmol), compound 7 (2.03g, 4.71 mmol), and $Pd(PPh_3)_2Cl_2$ (0.17g, 0.24 mmol) were dissolved in THF (50 mL) and reacted at 75 ℃ for 3 h under argon atmosphere. The solvent was removed. It was purified by chromatography in a silica gel column and eluted with PE/DCM (1/1, v/v) to obtain Compound 8 (2.12g, yield: 62%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 (s, 2H), 7.33 (d, J = 1.2 Hz, 2H), 3.66 (t, J = 6.7 Hz, 4H), 3.45 (t, J = 6.7 Hz, 4H), 2.94 (t, J = 6.6 Hz, 4H), 1.62 – 1.55 (m, 4H), 1.35 – 1.25 (m, 13H), 0.88 (d, J = 6.8 Hz, 5H).

Compound 9: Under argon, compound 8 (1.87g, 2.59 mmol) and triphenylphosphine (6.79g, 25.90 mmol) were dissolved in N-Methylpyrrolidone (NMP, 80 mL). The mixture was reacted at 180 ℃ for 24h. After cooling to room temperature, 5-(bromomethyl)-undecane (9.68g, 38.85 mmol) together with potassium iodide (2.17g, 38.85 mmol) and potassium hydroxide (2.17g, 38.85 mmol) was added to the mixture. The mixture was reacted at 95 ℃ for 24h under an argon atmosphere. After the reaction, it was extracted three times with DCM and then spun dry.

Compound 9 (1.00g, yield: 42%) was finally obtained by column chromatography in a silica gel column using PE/DCM (2/1, v/v) as the eluent. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.09 (s, 2H), 4.45 (s, 4H), 3.84 (t, J = 6.2 Hz, 4H), 3.52 (t, J = 6.7 Hz, 4H), 3.18 (t, J = 6.9 Hz, 4H), 1.76 (s, 2H), 1.63 (dt, J = 13.8, 6.7 Hz, 6H), 1.38 – 1.26 (m, 16H), 0.90 (t, J = 6.5 Hz, 44H).

Compound 10: Under argon atmosphere, LDA (2.87 mL, 5.74 mmol) was added dropwise into the solution of compound 9 (0.35g, 0.38 mmol) and superdry THF (40 mL) at -78 ℃. The mixture was then stirred for 1.5 h. Superdry DMF (5.34mL, 68.94 mmol) was added, and then the mixture was automatically rewarmed to -40°C, resulting in a water burst reaction. The mixture was extracted with EA and then spun dry. The crude product was purified by column chromatography in a silica gel column, eluting with PE/EA (10/1, v/v) to obtain compound 10 (0.24g, yield: 61%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.11 (s, 2H), 4.52 (d, J = 92.1 Hz, 4H), 3.81 (s, 6H), 3.44 (t, J = 6.7 Hz, 6H), 1.76 (s, 2H), 1.72 – 1.62 (m, 2H), 1.60 – 1.51 (m, 4H), 1.28 (d, J = 9.9 Hz, 28H), 1.01 – 0.74 (m, 16H), 0.73 – 0.07 (m, 20H).

Acceptor 2: Under argon atmosphere, compound 10 (0.12g, 0.12 mmol) and 2-(5,6-dichloro-3-oxo-2,3-dihydro-1H-inden-1-ylidene)malononitrile (0.18g, 0.70 mmol) were dissolved in chloroform (40 mL). Pyridine (1 mL) was added at 40 ℃. After reacting for 3h, the solution was poured into methanol (300 mL) and filtered. The residue was purified in a silica gel column using petroleum ether/dichloromethane (1/2, v/v) as the eluent. Acceptor 2 was obtained as a dark blue solid (0.10g, yield: 57%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.20 (s, 2H), 8.77 (s, 2H), 7.91 (s, 2H), 4.81 (s, 2H), 4.39 (s, 2H), 3.84 (d, J = 64.0 Hz, 6H), 3.44 (t, J = 6.8 Hz, 6H), 1.79 (s, 2H), 1.56 – 1.49 (m, 4H), 1.35 – 1.20 (m, 26H), 0.85 (dd, J = 19.3, 12.8 Hz, 12H), 0.64 (s, 12H), 0.35 (d, J = 99.0 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 184.36, 158.35, 151.91, 144.97, 139.35, 138.30, 138.28, 138.22, 138.15, 137.66, 136.35, 135.16, 135.13, 125.76, 124.05, 119.70, 114.38, 114.33, 113.66, 70.55, 69.83, 67.33, 52.47, 38.60, 30.65, 30.50, 30.16, 29.89, 29.79, 28.55, 28.17, 28.05, 27.78, 25.82, 25.75, 25.10, 24.75, 22.65, 22.03, 21.96, 21.67, 21.38, 21.33, 12.99, 12.50.



Scheme S3. The synthetic procedure of Acceptor 3

Compound 12: Under argon atmosphere, compound 11 (2.81g, 3.275 mmol) and triphenylphosphine (6.871g, 26.198 mmol) were dissolved in N-Methylpyrrolidone (NMP, 70 mL). The mixture was reacted at 180 ℃ for 24h. After cooling to room temperature, potassium hydroxide (2.2g, 39.3 mmol), potassium iodide (0.13g, 0.786 mmol), and 1-bromo-2-octyldodecane (9.47g, 26.2 mmol) were added to the mixture. The mixture was stirred at 90 ℃ for 24h under an argon atmosphere. After the reaction, it was cooled to room temperature. The reactants were poured into a large amount of water, and sodium chloride was added. Then, the reactants were extracted with a small amount of DCM three times. Initial purification was carried out using a silica gel column with PE as eluent. After rotary evaporation to remove the eluent, an excess of anhydrous zinc chloride and EA (50 mL) was added, stirred overnight to remove PPh_3, extracted with EA, and then rotary evaporated. Compound 12 (1.78 g, yield: 40%) was finally obtained by column chromatography in a silica gel column using petroleum ether/dichloromethane (10/1, v/v) as the eluent. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 6.98 (s, 2H), 4.59 (d, J = 7.8 Hz, 4H), 2.82 (t, J = 7.7 Hz, 4H), 2.14 (p, J = 6.6 Hz, 2H), 1.88 (p, J = 7.6 Hz, 4H), 1.30 (dd, J = 7.3, 4.0 Hz, 38H), 1.21 – 1.15 (m, 12H), 1.13 – 1.04 (m, 16H), 0.98 (d, J = 6.6 Hz, 8H), 0.87 (ddt, J = 14.7, 11.5, 6.1 Hz, 40H).

Compound 13: Under argon atmosphere, LDA (4.02 mL, 0.804 mmol) was added dropwise into a solution of compound 12 (1.089 g, 0.804 mmol) in superdry THF (50 mL) at -78 ℃. The mixture was stirred for 10 min. Superdry DMF (10.82 mL) was added after stirring at -78 ℃ for another 30 min. The mixture was warmed to -40°C, and the process was monitored using the spot plate. The mixture was warmed to room temperature after the reaction. The mixture was poured into water and extracted with DCM. After removing the solvent, the crude product was purified by column chromatography and eluted with petroleum ether/dichloromethane (4/1, v/v) in a silica gel column to obtain compound 13 (0.664g, yield: 58%). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 10.13 (s, 2H), 4.58 (d,

J = 7.8 Hz, 4H), 3.18 (t, *J* = 7.8 Hz, 4H), 2.06 (p, *J* = 6.5 Hz, 2H), 1.92 (p, *J* = 7.6 Hz, 4H), 1.30 – 1.21 (m, 38H), 1.15 (ddd, *J* = 7.3, 5.0, 2.8 Hz, 12H), 1.08 – 1.02 (m, 14H), 0.96 (ddd, *J* = 10.7, 4.9, 2.3 Hz, 22H), 0.87 – 0.81 (m, 28H).

Compound 14: Under argon atmosphere, compound 13 (0.664 g, 0.471 mmol), 2-(5,6-difluoro-3-oxo-2,3-dihydro-1H-inden-1-ylidene)malononitrile (0.13 g, 0.565 mmol) and 2-(5-bromo-3-oxo-2,3-dihydro-1H-inden-1-ylidene)malononitrile (0.154 g, 0.565 mmol） were dissolved in chloroform (70 mL) and Pyridine (3.0 mL). After reacting at 65 ℃ for 9 h, the mixture was cooled to room temperature. It was poured into 200 mL of methanol and stirred for 30 min, and the crude product was obtained by filtration. The residue was purified in a silica gel column using petroleum ether/dichloromethane (1/2.5, v/v) as the eluent. Compound 14 was obtained as a dark blue solid (0.425g, yield: 48%). ^{1}H NMR (400 MHz, Chloroform-d) δ 9.10 (d, *J* = 8.6 Hz, 2H), 8.55 – 8.46 (m, 2H), 7.96 (d, *J* = 1.9 Hz, 1H), 7.80 (dd, *J* = 8.3, 2.0 Hz, 1H), 7.62 (t, *J* = 7.5 Hz, 1H), 4.65 (d, *J* = 7.8 Hz, 4H), 3.15 (t, *J* = 7.9 Hz, 4H), 2.05 (d, *J* = 7.0 Hz, 2H), 1.81 (p, *J* = 7.9 Hz, 4H), 1.44 (d, *J* = 7.7 Hz, 4H), 1.34 – 1.16 (m, 30H), 1.16 – 1.02 (m, 24H), 1.01 – 0.85 (m, 34H), 0.76 (ddt, *J* = 17.7, 16.1, 6.9 Hz, 22H).

Acceptor 3: Compound 14 (143 mg, 0.076 mmol) and 2,5-bis(trimethylstannyl)thiophene (15.2 mg, 0.037 mmol) were added to superdry toluene (15 mL) and tetrakis(triphenylphosphine)palladium (4.3 mg, 0.0037 mmol) under argon atmosphere, and then the mixture was heated up to 110 °C for 4 h. After the reaction, the mixture was cooled to room temperature. The toluene was evaporated and then extracted three times with a large amount of water and a small amount of DCM, and then purified by silica gel chromatography column and eluted with PE/DCM (1.2/1, v/v). The blue-black solid Acceptor 3 (62.6 mg, yield: 46 %) was obtained after spun-dring. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 9.12 (d, *J* = 10.9 Hz, 4H), 8.74 (d, *J* = 8.3 Hz, 2H), 8.56 (dd, *J* = 9.9, 6.4 Hz, 2H), 8.09 (s, 2H), 7.97 (d, *J* = 8.4 Hz, 2H), 7.65 (q, *J* = 5.7, 3.9 Hz, 4H), 4.76 (d, *J* = 7.8 Hz, 8H), 3.20 (q, *J* = 8.3 Hz, 8H), 2.26 – 2.14 (m, 4H), 1.87 (q, *J* = 7.7 Hz, 8H), 1.49 (d, *J* = 6.8 Hz, 6H), 1.28 (q, *J* = 5.7, 5.0 Hz, 68H), 1.21 – 0.96 (m, 108H), 0.96 – 0.68 (m, 46H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 187.69, 185.94, 159.47, 158.54, 155.56, 153.84, 153.41, 153.02, 153.00, 145.31, 145.28, 143.81, 138.91, 138.83, 137.74, 137.33, 136.55, 136.19, 135.71, 134.94, 134.77, 134.35, 133.85, 133.48, 133.16, 131.82, 131.37, 131.06, 127.16, 125.97, 120.51, 119.56, 118.80, 115.88, 115.49, 115.17, 114.99, 114.56, 68.44, 67.44, 55.77, 39.29, 31.93, 31.86, 31.21, 31.16, 30.67, 29.94, 29.87, 29.71, 29.69, 29.66, 29.64, 29.60, 29.56, 29.51, 29.49, 29.44, 29.37, 29.30, 29.26, 25.85, 22.69, 22.66, 14.11

Instruments and measurements

^{1}H NMR and ^{13}C NMR spectra were recorded using a Bruker AVANCE III HD 400 spectrometer in a deuterated chloroform solution at 298 K unless specified otherwise. Chemical shifts are reported as δ values (ppm) with tetramethylsilane (TMS) as the internal reference. UV-Vis absorption spectra were recorded on the SHIMADZU UV-2600 spectrophotometer. The cyclic voltammetry results were obtained with a computer-controlled CHI 660E electrochemical workstation. Mass spectra were recorded on a 5800 MALDI TOF of AB SCIEX.

Reference

1. Zhu, S.*, et al.* A-DA'D-A type pentacyclic small molecule acceptors to exceed 16.5% efficiency by heteroatom effect at the outer side chain. *Chinese Journal of Chemistry* **41**, 1815-1822 (2023).
2. Wei, Q.*, et al.* Effects of oxygen position in the alkoxy substituents on the photovoltaic performance of A-DA′D-A type pentacyclic small molecule acceptors. *ACS Energy Letters* **7**, 2373-2381 (2022).
3. Yu, H.*, et al.* Tailoring non-fullerene acceptors using selenium-incorporated heterocycles for organic solar cells with over 16% efficiency. *Journal of Materials Chemistry A* **8**, 23756-23765 (2020).