Supporting Information for:

Effective Generation of Heavy-Atom-Free Triplet Photosensitizers Containing Multiple Intersystem Crossing Mechanisms Based on Deep Learning

Kepeng Chen [a], Xiaoting Zhang [b], Jike Wang [a], Dan Li [a], Tingjun Hou*[a], Wenbo Yang*[b], Yu Kang*[a]

[a]College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China

[b]School of Chemistry, State Key Laboratory of Fine Chemicals, Frontier Science Center for Smart Materials, Dalian Key Laboratory of Intelligent Chemistry, Dalian University of Technology, Dalian 116024, China

Index

Part 1. The Detail Procedures of Constructing the Dataset of Photosensitizers.

a) For simplified molecular input line entry system (SMILES) format of molecules, removing duplicates, heavy atoms (atomic weight is less than 40, the first three periods of the periodic table) and metal atoms; b) Utilize the BRICS algorithm from cheminformatics toolkit RDKit, which is based on 16 common chemical reaction templates, to generate a molecular fragment library from the initial dataset source; c) Classify the molecular fragment library into scaffold fragments (ring numbers ≥ 2, reaction points (1~ 3) and discarding reaction points surpass 3 for simplicity), linker fragments (ring numbers ≤ 1, reaction points (2 ~ 3)), and terminal groups (ring numbers ≤ 1, reaction points is 1. The final data set contains 1.90×10^9 molecules.

Table S1. Comparison of Molecular Fragmentation with Two Representative Methods.

Methods	Ring-cutting method	Char-splitting method
Fragmentation molecules	3500000	3500000
Fragment vocabulary size	26628564	311190299
Average number of fragments/tokens	7.6	88.9

Part 2. Prediction Models for Score Function in Reinforcement Learning.

The ΔE_{ST} and E_{abs} are the label parameters of data, which were processed by zero-mean normalization before training (the $\mu = 1.1721$, $\sigma = 0.6635$ for ΔE_{ST}, the $\mu = 3.2070$, $\sigma = 0.9691$ for ΔE_{ST}). Moreover, to deal with sample imbalance, common data augmentation technique was used in this scenario for ΔE_{ST} was less than 0.30 eV. The training set (80%), valid set (10%) and test set (10%) were used in each cycle which was split according to the molecule fingerprint similarity cluster method by Chemfp keeping data independently. To tune the hyperparameters finely to obtain an accurate

prediction model. Hyperparameters for initial prediction models were optimized by grid search method (searching through all possible combinations of the specified hyperparameters and evaluating each combination using cross-validation). The optimization process contains an early stopping strategy in which mean absolute error (MAE) of ΔE_{ST} not decrease 0.01 eV after continuous 6 trials. The optimization process was also stopped if minimize MAE optimization cycles till 400 trials have been done. The following hyperparameters were tuned:

a) Graph convolutional layers: a list of graph convolutional layers with each value representing the number of nodes in each layer which are [512, 512, 512], [512, 512, 512, 512] and [512, 512, 512, 512,512];

b) Dense layers: a list of dense fully connected layers with each value representing the number of nodes in each layer which are [128, 128, 128] and [128, 128, 128, 128];

c) Dropout: probability (between 0 and 1) that neurons in the hidden layers are ignored; dropout is added to prevent overfitting which is 0.01, 0.05 and 0.1, respectively.

d) Learning rate: The multiplier for gradient descent and determines how fast the parameter changes which is 0.0001 and 0.001, respectively.

One of the best prediction model's parameters are as follows: graph convolutional layers list is [512, 512, 512, 512], dense layers list is [128, 128, 128], dropout is 0.01 and learning rate is 0.001. The relationship between training epochs and MAE values are presented in Fig. S1.

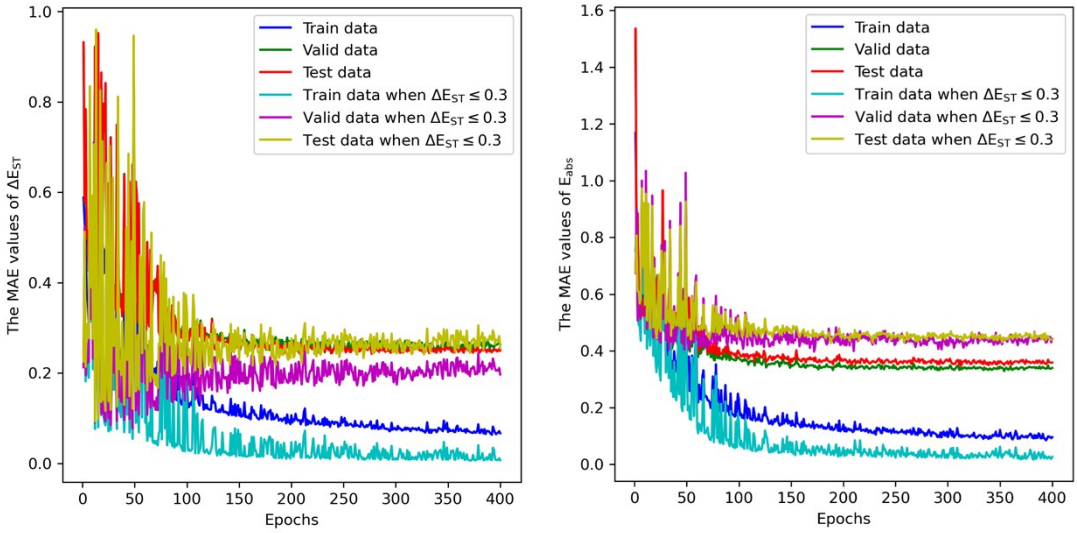

Fig. S1. The relationship between training epochs and MAE values of a) ΔE_{ST} and b) E_{abs} of train data, valid data, test data and the important subset of above datasets (ΔE_{ST} ≤ 0.3 eV) for photosensitizers design.

Part 3. Conjugate Motifs Diversity for All Model.

Table S2. The Ring Numbers and the Atoms Numbers of Conjugated Motifs for the Unique Desired Molecules Generated by All Models.

	Task1 models				Task2 models			
	MB	MD	GB	GD	Frag-MB	Frag-MD	Frag-GB	Frag-GD
Ring Numbers	1.05	1.28	1.08	1.71	1.09	1.53	1.18	2.46
Atom Numbers	9.40	11.16	9.09	11.31	9.07	13.42	9.64	16.09

Note: This is the average value for the unique desired molecules generated by 3 cycles.

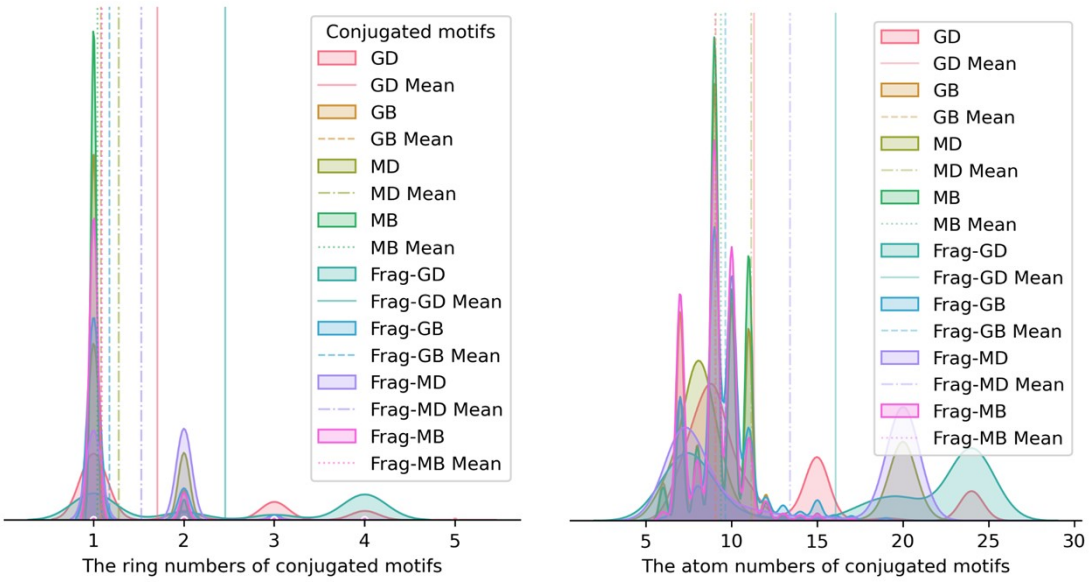

Fig. S2. Distributions of the ring numbers of conjugated motifs (left) as well as the atoms numbers of conjugated motifs (right) for the unique desired molecules.

Part 4. Distributions of QED and SA Properties of the Unique Desired Molecules Generated by All Models for Task 1 and Task 2

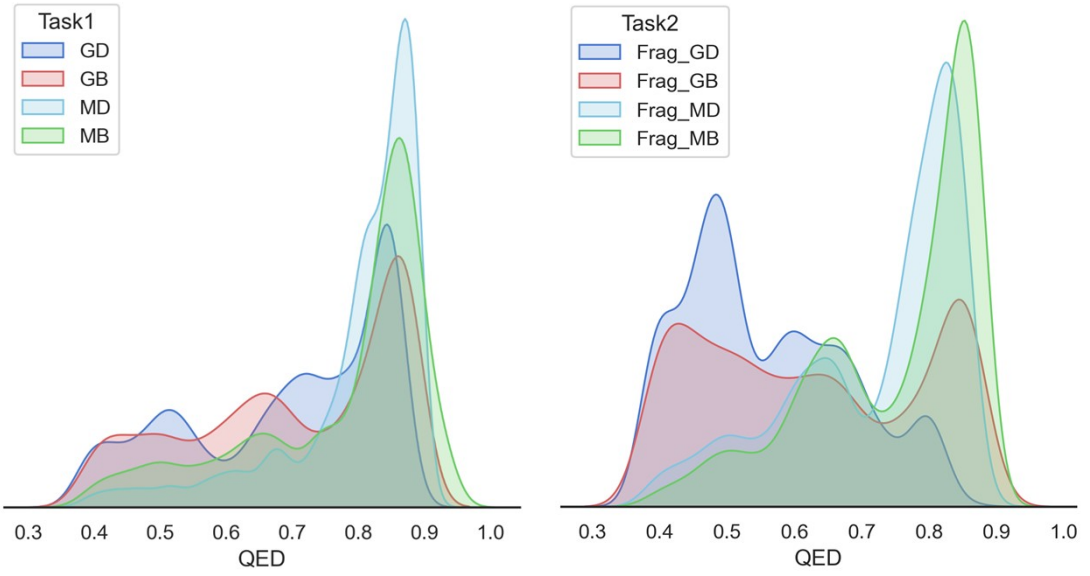

Fig. S3. Distributions of QED for the unique desired molecules generated by all models for task 1 (left) and task 2 (left).

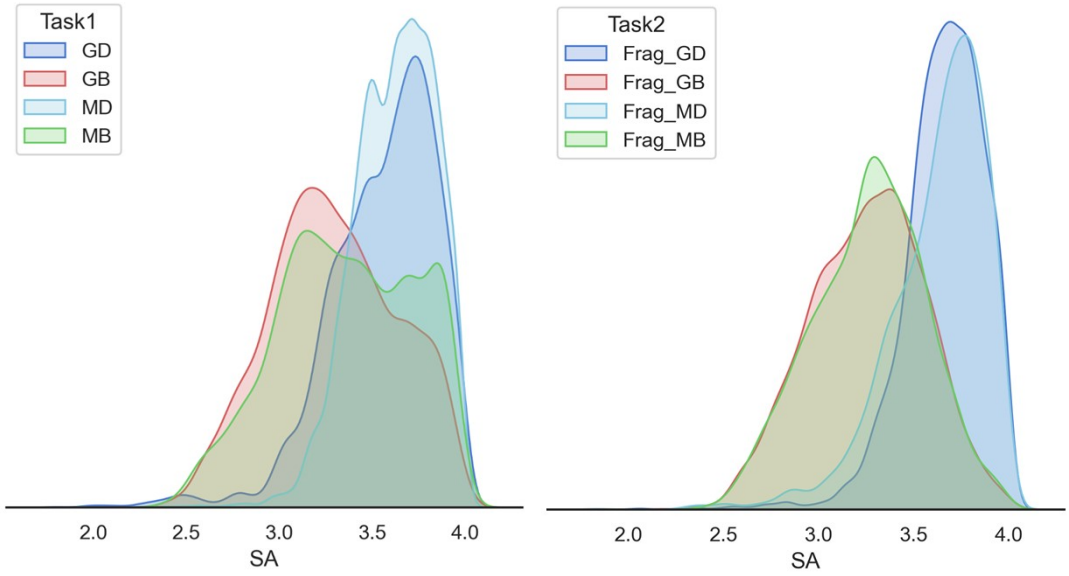

Fig. S4. Distributions of SA for the unique desired molecules generated by all models for task 1 (left) and task 2 (left).

Part 5. Molecular Examples of Studied Models and Case studies of Ablation Experiments

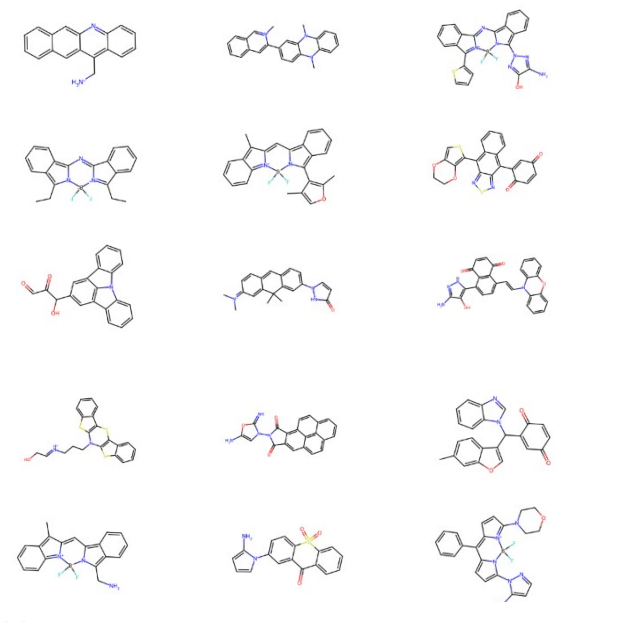

Fig. S5. Small part of selected design molecules.

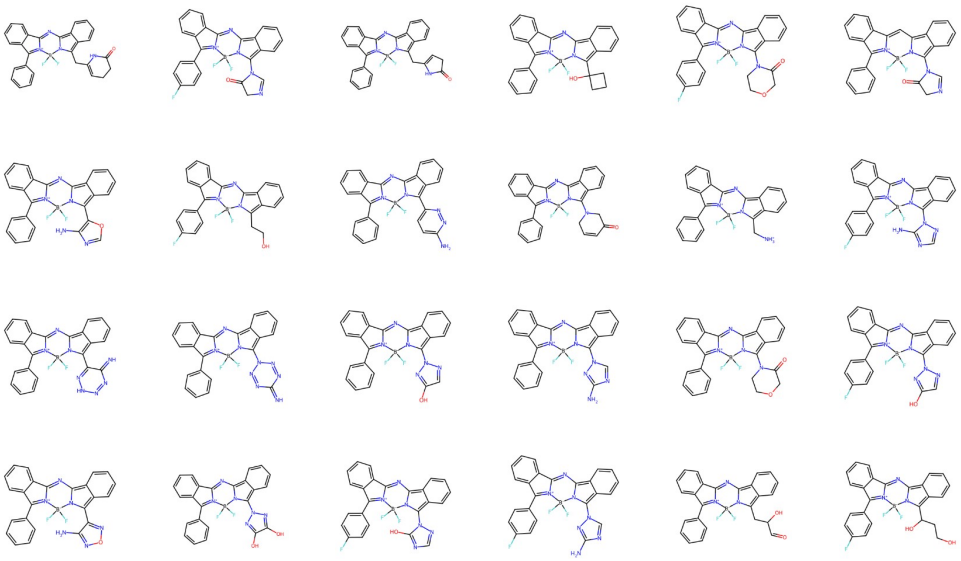

Fig. S6. Selected of non-symmetry molecules (the conjugated motif keep same).

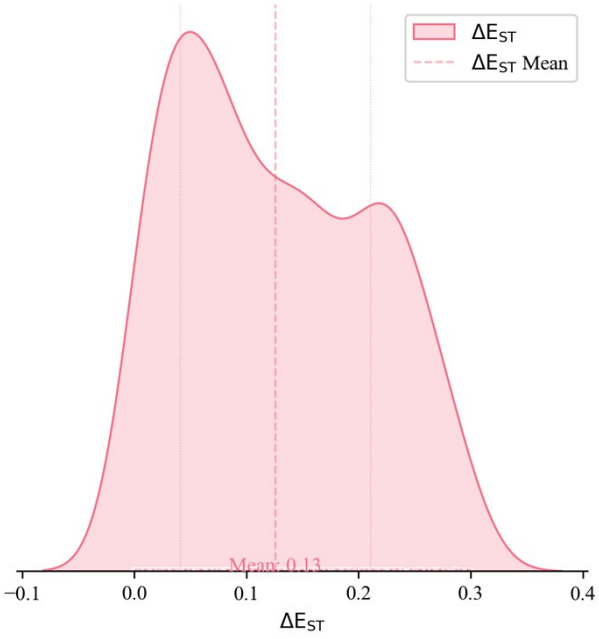

Fig. S7. The distribution of ΔE_{ST} of molecules verified by DFT, TD-DFT and SOC calculations.

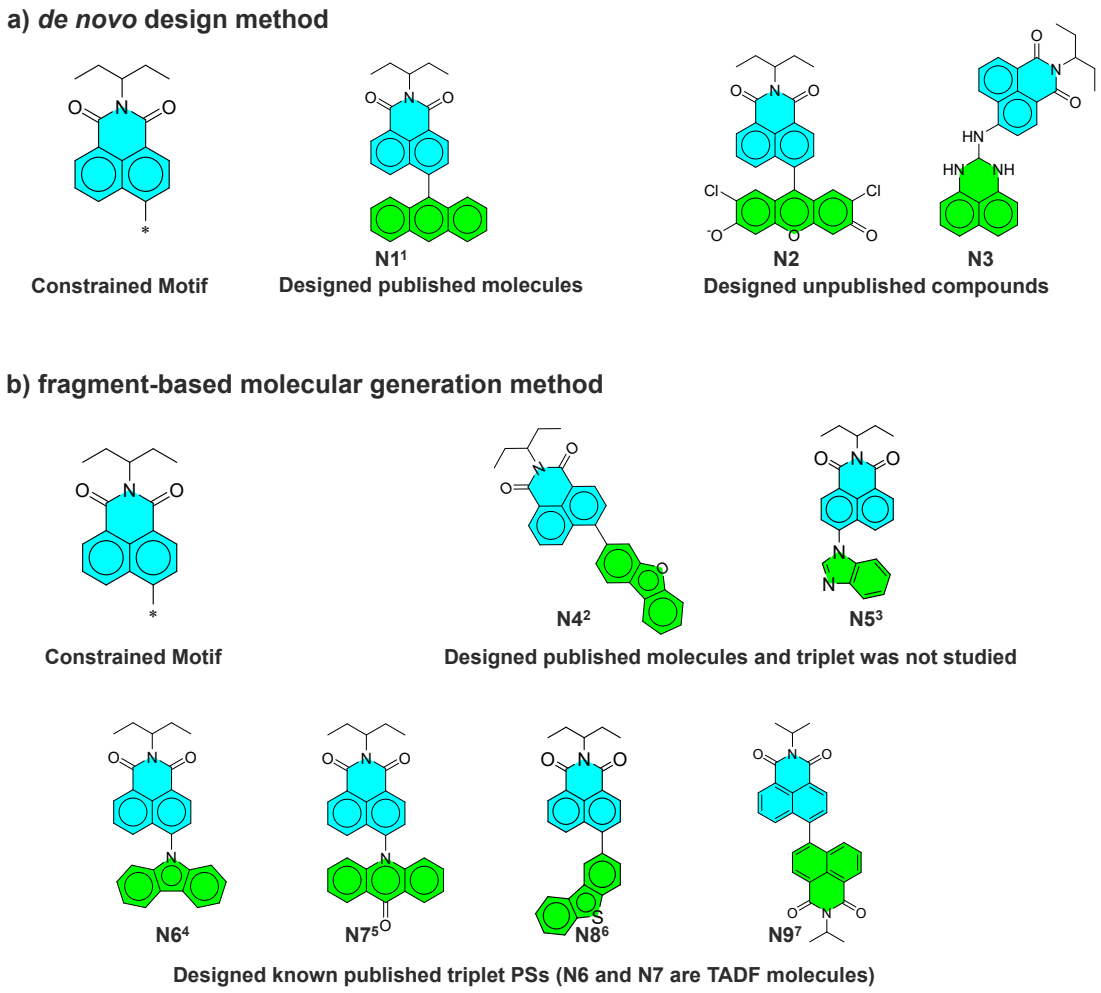

Fig. S8. Sampled generated molecules by different methods. a) *de novo* design method, b) fragment-based molecule generation method.

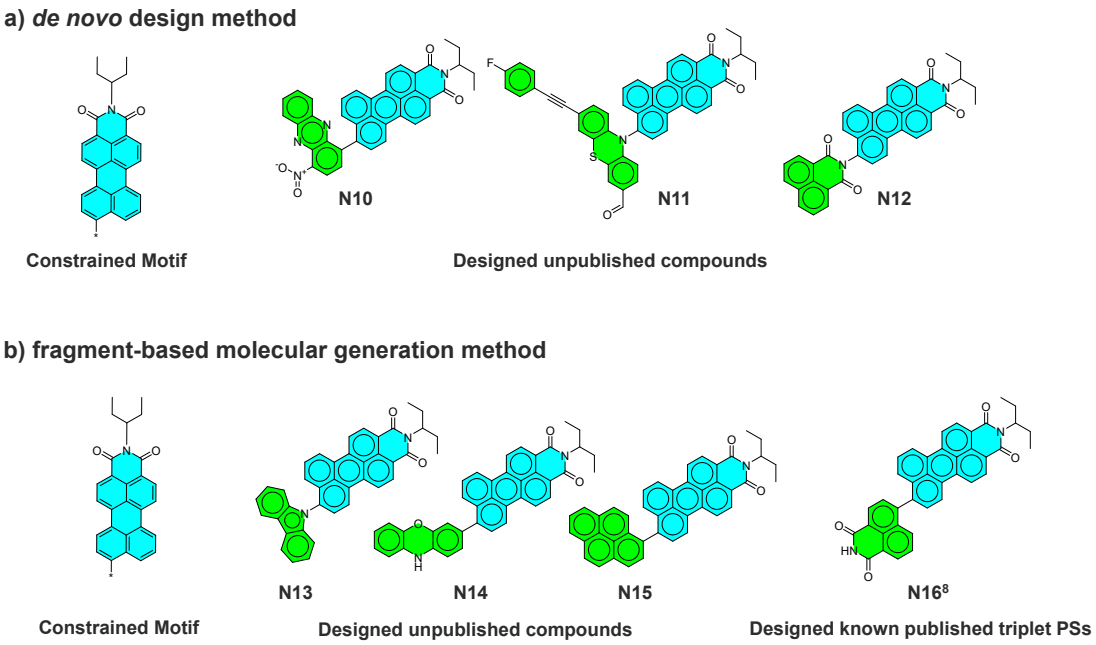

Fig. S9. Sampled generated molecules by different methods. a) *de novo* design method, b) fragment-based molecule generation method.

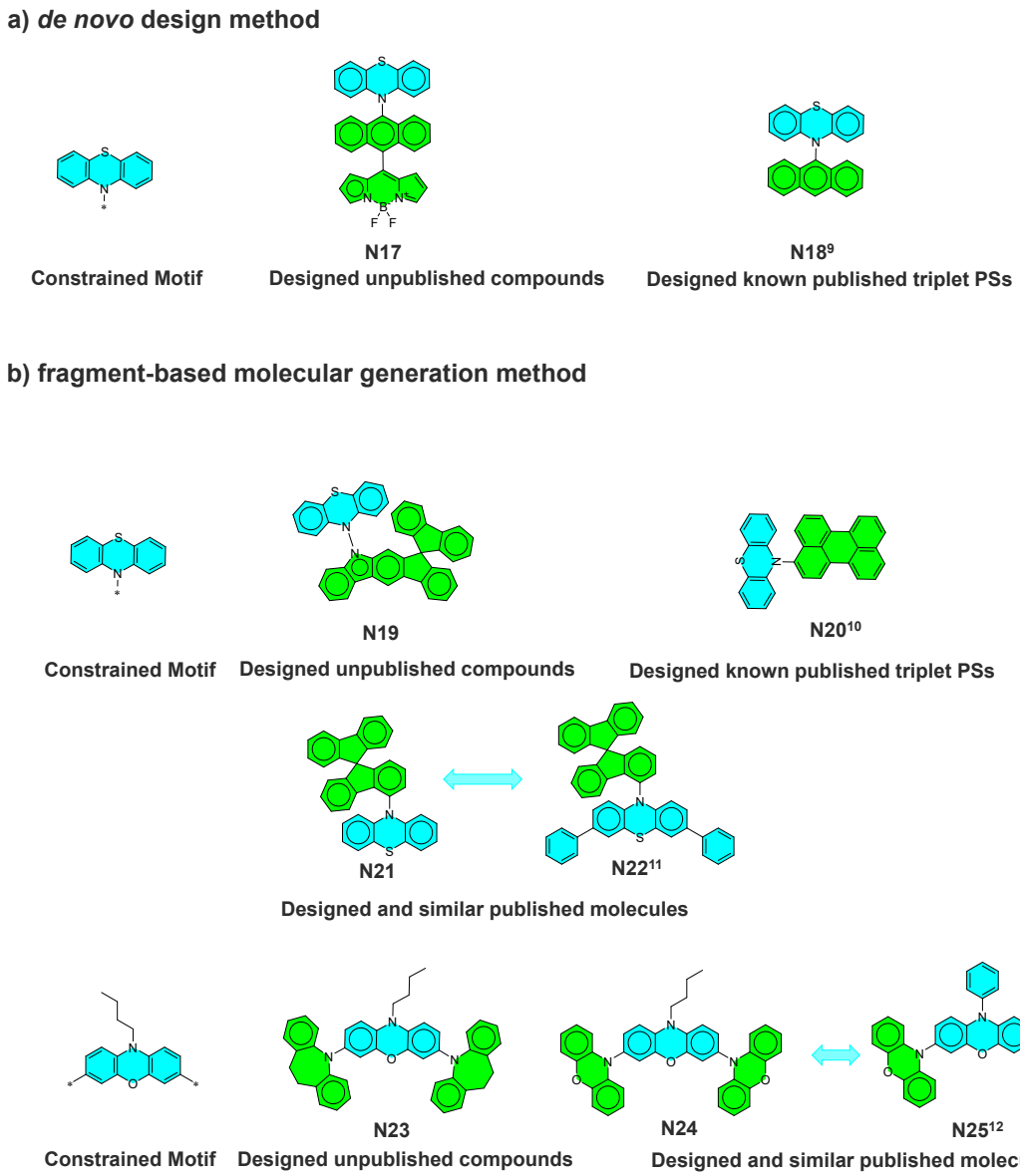

Fig. S10. Sampled generated molecules by different methods. a) *de novo* design method, b) fragment-based molecule generation method.

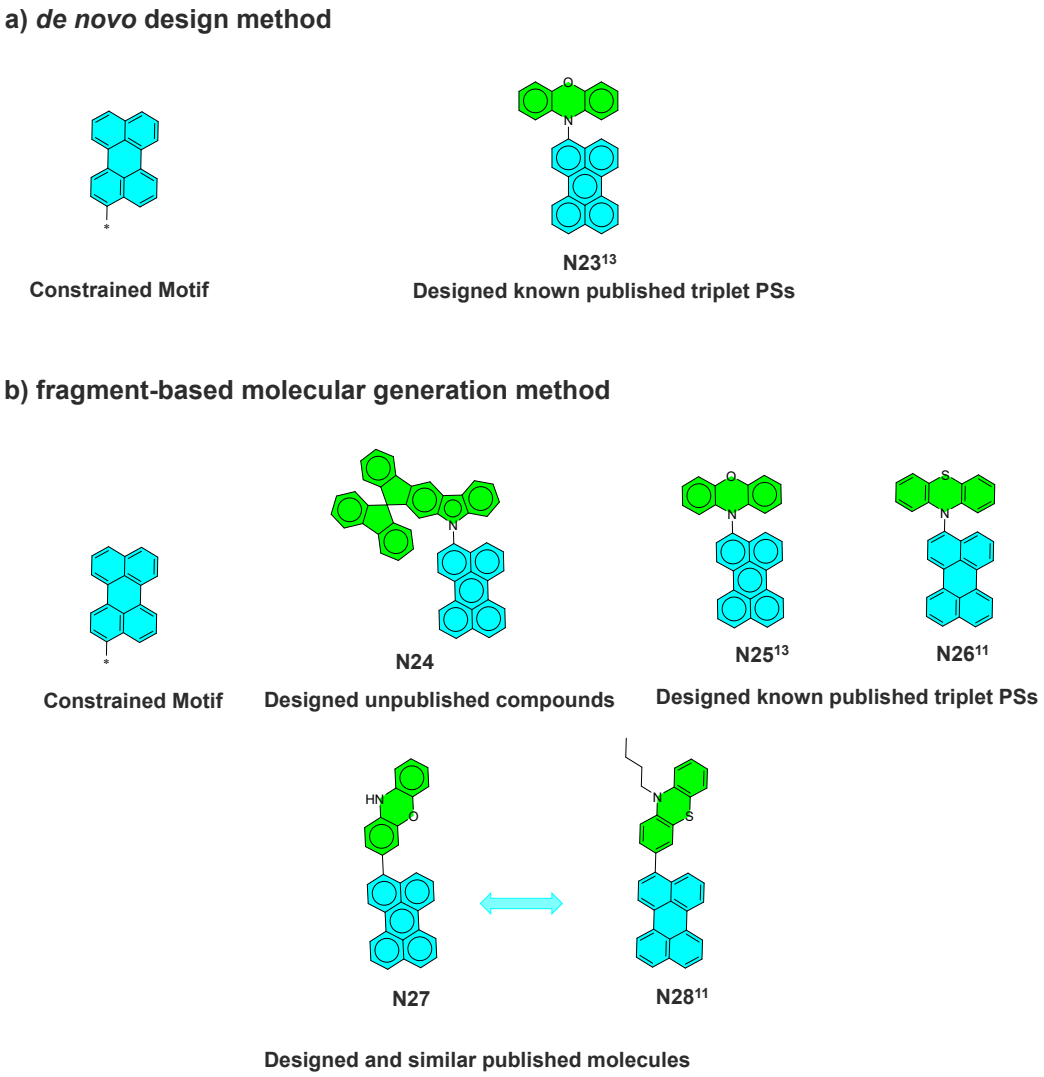

Fig. S11. Sampled generated molecules by different methods. a) *de novo* design method, b) fragment-based molecule generation method.

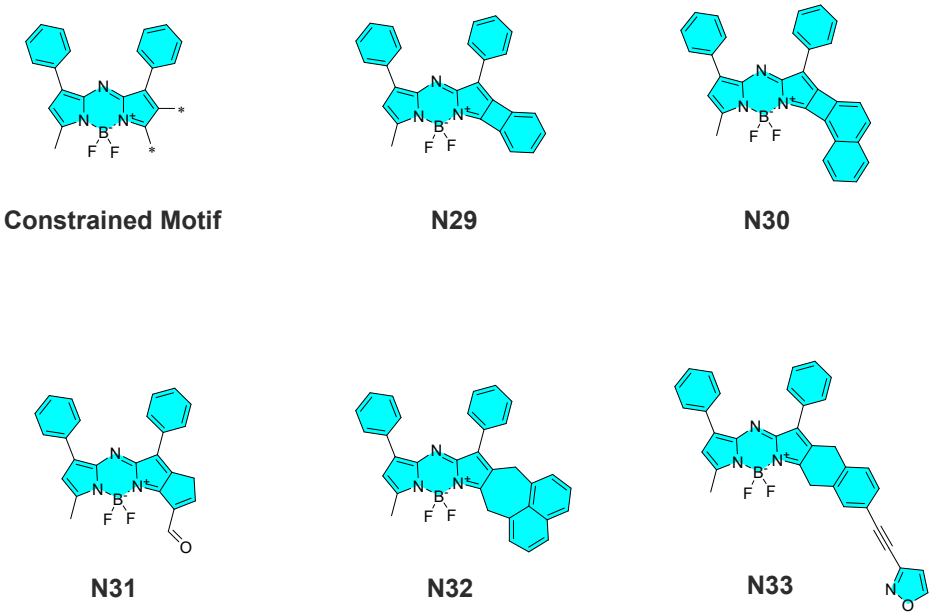

Fig. S12. Sampled generated molecules by *de novo* design method for exploring enlarged conjugated structures.

REFERENCES

1 K. Chen, J. Zhao, X. Li and G. G. Gurzadyan, Anthracene-naphthalenediimide compact electron donor/acceptor dyads: electronic coupling, electron transfer, and intersystem crossing, *J. Phys. Chem. A*, 2019, **123**, 2503–2516.

2 J. Kim, J. Baek, E. Jeong, S. Han, Y. Kim and S. Hwang, Organic light-emitting device and electronic apparatus including the same, *U.S. Patent* US20190051856, February 14, 2019.

3 P.-H. Lanoë, J. Chan, A. Groué, G. Gontard, A. Jutand, M.-N. Rager, N. Armaroli, F. Monti, A. Barbieri and H. Amouri, Cyclometalated N-heterocyclic carbene iridium(III) complexes with naphthalimide chromophores: a novel class of phosphorescent heteroleptic compounds, *Dalton Trans.*, 2018, **47**, 3440–3451.

4 G. Tang, W. Yang and J. Zhao, Naphthalimide-carbazole compact electron donor-acceptor dyads: effect of molecular geometry and electron-donating capacity on the spin-orbit charge transfer intersystem crossing, *J. Phys. Chem. A*, 2022, **126**, 3653–3668.

5 S. Jena, P. Dhanalakshmi, G. Bano and P. Thilagar, Delayed fluorescence, room temperature phosphorescence, and mechanofluorochromic naphthalimides: differential imaging of normoxia and hypoxia live cancer cells, *J. Phys. Chem. B*, 2020, **124**, 5393–5406.

6 H. Liang, M. Zafar, J. Pang, Z. Chen, M.-D. Li, S. Ji, Y. Huo and H. Zhang, Enhancing the triplet yield in compact dibenzofuran-napthalimide donor/acceptor dyad based on charge recombination induced intersystem crossing via substitution of one atom, *J. Lumin.*, 2021, **238**, 118238.

7 W. Tian, A. A. Sukhanov, L. Bussotti, J. Pang, J. Zhao, V. K. Voronkova, M. Di Donato and M.-D. Li, Charge separation and intersystem crossing in homo- and hetero-compact naphthalimide dimers, *J. Phys. Chem. B*, 2022, **126**, 4364–4378.

8 E. A. Margulies, J. L. Logsdon, C. E. Miller, L. Ma, E. Simonoff, R. M. Young, G. C. Schatz and M. R. Wasielewski, Direct observation of a charge-transfer

state preceding high-yield singlet fission in terrylenediimide thin films, *J. Am. Chem. Soc.*, 2017, **139**, 663–671.

9 Y. Hou, T. Biskup, S. Rein, Z. Wang, L. Bussotti, N. Russo, P. Foggi, J. Zhao, M. Di Donato, G. Mazzone and S. Weber, Spin-orbit charge recombination intersystem crossing in phenothiazine-anthracene compact dyads: effect of molecular conformation on electronic coupling, electronic transitions, and electron spin polarizations of the triplet states, *J. Phys. Chem. C*, 2018, **122**, 27850–27865.

10 M. Imran, A. A. Sukhanov, Z. Wang, A. Karatay, J. Zhao, Z. Mahmood, A. Elmali, V. K. Voronkova, M. Hayvali, Y. H. Xing and S. Weber, Electronic coupling and spin-orbit charge-transfer intersystem crossing in phenothiazine-perylene compact electron donor/acceptor dyads, *J. Phys. Chem. C*, 2019, **123**, 7010–7024.

11 M. Engel, S. Meyer, M. Hamburger, F. May and E. Boehm, Electronic devices, *WO Patent* 2023/117836, June 29, 2023.

12 H. J. Shine and S.-M. Wu, Ion radicals. 44. Reactions with 10-phenylphenoxazine cation radical perchlorate, *J. Org. Chem.*, 1979, **44**, 3310–3316.

13 X. Xiao, J. Pang, A. A. Sukhanov, Y. Hou, J. Zhao, M.-D. Li and V. K. Voronkova, The effect of one-atom substitution on the photophysical properties and electron spin polarization: intersystem crossing of compact orthogonal perylene/phenoxazine electron donor/acceptor dyad, *J. Chem. Phys.*, 2020, **153**, 184312.