

Supporting Information

Prediction and Understanding of AIE Effect by Quantum Mechanics-aided Machine-Learning Algorithm

*Jia Qiu^{a,b,c#}, Kun Wang^{c,d#}, Zhouyang Lian^{b#}, Xing Yang^e, Wenhui Huang^{a,b}, Anjun Qin^f, Qian Wang^g, Jie Tian^{*c,d}, Benzhong Tang^{*f,h}, Shuixing Zhang^{*b}*

a School of Medicine, South China University of Technology, Guangzhou Higher Education Mega Center, Guangzhou 510006, China

b Department of Radiology, Guangdong General Hospital/Guangdong Academy of Medical Sciences, Guangzhou, 510080, China

c CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

d University of Chinese Academy of Sciences, Beijing 100049, China

e Department of Nuclear Medicine, Peking University First Hospital, Beijing 100034, China

f State Key Laboratory of Luminescent Materials and Devices, South China University of Technology, Guangzhou 510640, China

g Department of Diagnostic Imaging, National Cancer Center/Cancer Hospital, Chinese

Academy of Medical Sciences and Peking Union Medical College, No.17, Panjiayuan,
Chaoyang District, Beijing 100021, China

h Department of Chemistry, Hong Kong Branch of Chinese National Engineering Research
Center for Tissue Restoration and Reconstruction, The Hong Kong University of Science and
Technology, Clear Water Bay, Kowloon, China

#: These authors contributed equally to this work.

AUTHOR INFORMATION

Corresponding Authors

Shuixing Zhang, E-mail: shui7515@126.com

Benzhong Tang, E-mail: tangbenz@ust.hk

Jie Tian, E-mail: tian@ieee.org

TABLE of CONTENTS:

Section S1. Data source

Table S1. Calculated Data and prediction results.

Table S2. Molecular structures and conformations.

Figure S1. The charge distribution of four typical AIE-active and AIE-inactive TPA derivatives.

Section S2. Derivation procedure of discriminant equation

Figure S2. Derivation procedure illustration of discriminant equation.

Section S3. Prediction of new TPA derivative's AIE activity and experimental verification.

Figure S3. The molecular structure and charge distribution of DPAB.

Figure S4. Emission spectra of DPAB in water/ethanol mixtures with different water fractions(from 0~99%) at the pure DPAB-ethanol solution concentration of 0.01 mg/mL.

Figure S5. Plot of the relative PL intensity (I/I_0) of DPAB at 474 nm light versus the composition of the mixture of water/ethanol (water fraction). I_0 =PL intensity of DPAB in pure ethanol and I =PL intensity of DPAB in mixture.

REFERENCES

Section S1. Data source

Table S1. Calculated Data and prediction results.

Compound	Ref	DM/D	E _N / a.u.	E ₁ / a.u.	E ₂ / a.u.	E ₃ / a.u.	D/ a.u.	Act _{ca}	SVM _{ca}	D _{ca}
1	1	3.49	-0.430	0.195	0.139	0.139	0.079	1	1	1
2	1	4.30	-0.438	0.169	0.142	0.143	0.037	1	1	1
3	21*	2.04	-0.439	0.176	0.144	0.144	0.045	1	1	1
4	29*	7.09	-0.433	0.187	0.138	0.140	0.068	1	1	1
5	2	10.28	-0.427	0.202	0.135	0.135	0.095	1	1	1
6	3	4.57	-0.436	0.179	0.140	0.142	0.054	1	1	1
7	4	5.35	-0.431	0.197	0.139	0.138	0.083	1	1	1
8	5	0.11	-0.442	0.163	0.144	0.145	0.026	1	1	1
9	29*	6.37	-0.427	0.174	0.175	0.130	0.063	1	1	1
10	28*	9.96	-0.431	0.192	0.138	0.137	0.077	1	1	1
11	28*	8.41	-0.425	0.176	0.177	0.128	0.069	1	1	1
12	6	2.87	-0.443	0.158	0.146	0.146	0.017	1	1	1
13	21*	2.32	-0.439	0.177	0.144	0.144	0.047	1	1	1
14	21*	2.79	-0.440	0.174	0.145	0.144	0.042	1	1	1
15	7	5.31	-0.446	0.159	0.144	0.144	0.021	1	1	1
16	8	9.52	-0.437	0.186	0.140	0.139	0.066	1	1	1
17	9	1.59	-0.443	0.158	0.148	0.146	0.016	1	1	1
18	9	2.02	-0.440	0.160	0.145	0.145	0.021	1	1	1
19	9	1.56	-0.440	0.167	0.144	0.144	0.033	1	1	1
20	9	1.54	-0.444	0.154	0.148	0.146	0.010	1	0	0
21	9	1.58	-0.443	0.159	0.147	0.146	0.018	1	1	1
22	9	1.00	-0.443	0.157	0.148	0.146	0.014	1	1	1

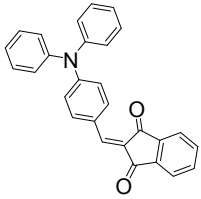
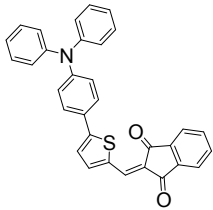
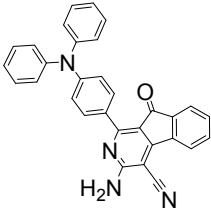
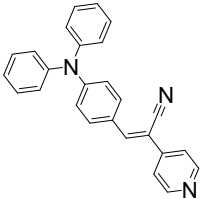
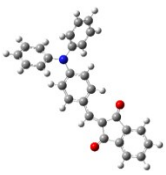
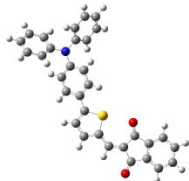
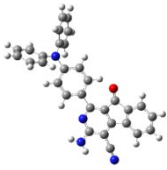
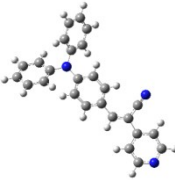
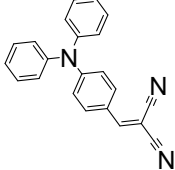
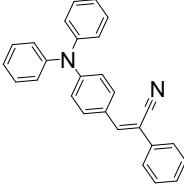
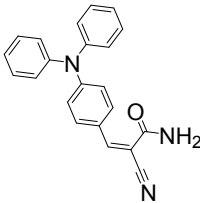
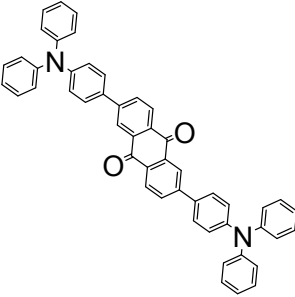
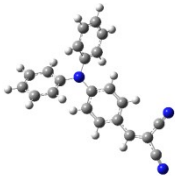
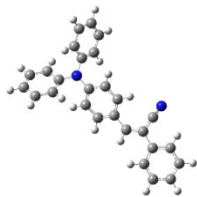
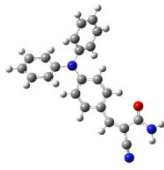
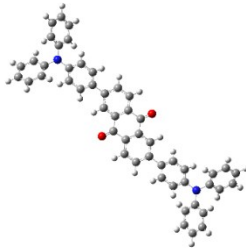
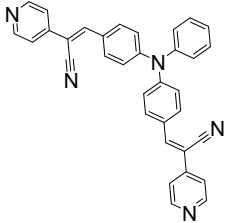
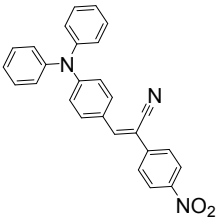
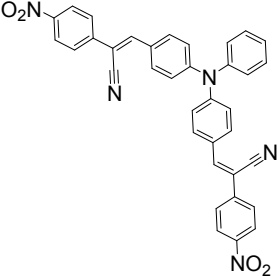
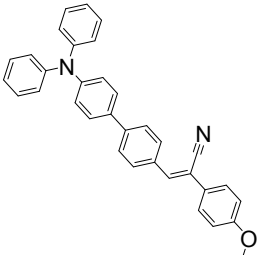
23	¹⁰	11.46	-0.426	0.199	0.139	0.132	0.090	1	1	1
24	¹¹	6.61	-0.439	0.181	0.143	0.141	0.055	1	1	1
25	¹¹	5.39	-0.439	0.181	0.142	0.142	0.055	1	1	1
26	^{15*}	1.47	-0.444	0.157	0.148	0.146	0.014	1	1	1
27	¹²	0.21	-0.445	0.155	0.148	0.148	0.010	1	0	0
28	¹²	0.41	-0.445	0.153	0.148	0.149	0.006	1	0	0
29	¹²	0.38	-0.444	0.155	0.148	0.148	0.010	1	0	0
30	^{24*}	4.84	-0.437	0.181	0.115	0.113	0.095	1	1	1
31	¹³	7.34	-0.439	0.140	0.157	0.157	0.024	1	1	1
32	¹³	4.28	-0.440	0.141	0.159	0.152	0.022	1	1	1
33	^{23*}	3.04	-0.437	0.180	0.140	0.139	0.057	1	1	1
34	^{23*}	8.41	-0.428	0.197	0.132	0.132	0.092	1	1	1
35	¹⁴	3.40	-0.440	0.153	0.146	0.144	0.012	1	0	0
36	¹⁴	3.49	-0.441	0.155	0.143	0.141	0.019	1	1	1
37	^{30*}	6.11	-0.438	0.176	0.143	0.141	0.048	1	1	1
38	¹⁵	4.50	-0.444	0.168	0.118	0.118	0.071	1	1	1
39	¹⁶	0.20	-0.443	0.153	0.148	0.147	0.008	1	0	0
40	¹⁶	0.18	-0.438	0.145	0.151	0.151	0.008	1	0	0
41	¹⁶	0.13	-0.434	0.148	0.148	0.148	0.000	1	0	0
42	^{20*}	0.01	-0.447	0.149	0.149	0.149	0.000	0	0	0
43	^{29*}	4.51	-0.423	0.164	0.164	0.164	0.000	0	0	0
44	^{28*}	3.26	-0.422	0.164	0.164	0.164	0.000	0	0	0
45	⁹	0.00	-0.445	0.151	0.148	0.148	0.004	0	0	0
46	^{20*}	0.00	-0.444	0.151	0.148	0.148	0.004	0	0	0
47	¹²	0.17	-0.446	0.152	0.148	0.147	0.006	0	0	0
48	¹²	0.00	-0.446	0.154	0.147	0.147	0.010	0	0	0
49	¹⁷	1.17	-0.442	0.154	0.147	0.147	0.010	0	0	0

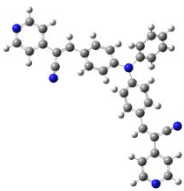
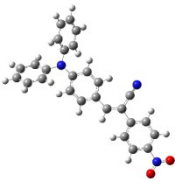
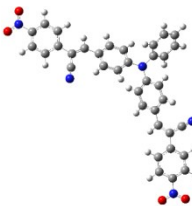
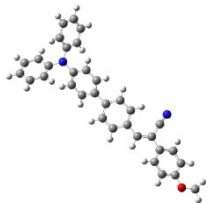
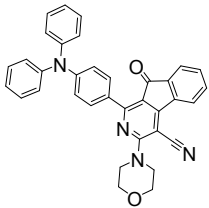
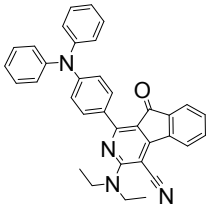
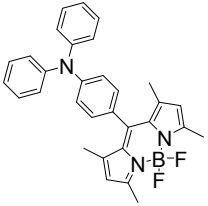
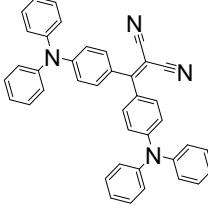
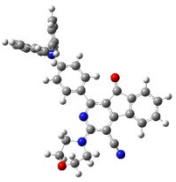
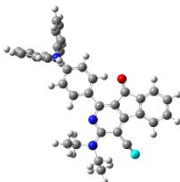
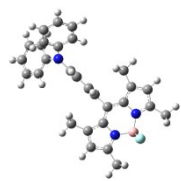
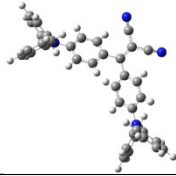
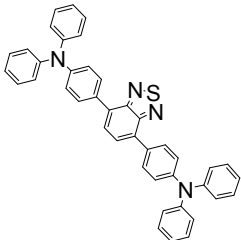
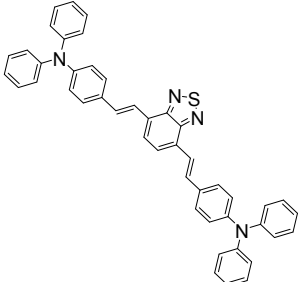
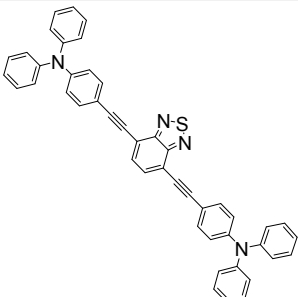
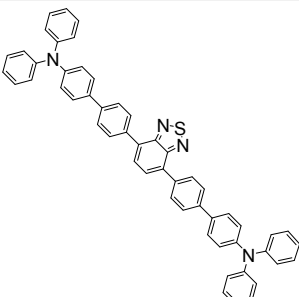
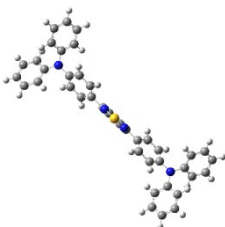
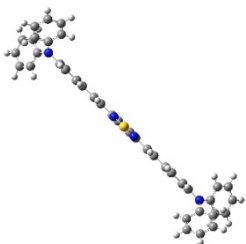
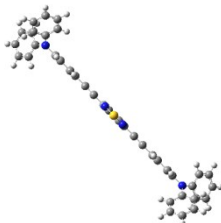
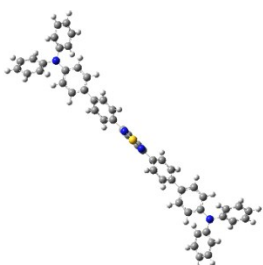
50	17	0.63	-0.442	0.155	0.146	0.147	0.012	0	0	0
51	17	2.91	-0.442	0.156	0.146	0.147	0.013	0	0	0
52	17	3.61	-0.443	0.152	0.147	0.148	0.006	0	0	0
53	18	0.36	-0.445	0.151	0.148	0.148	0.004	0	0	0
54	19	0.03	-0.445	0.151	0.147	0.148	0.005	0	0	0
55	22*	13.57	-0.436	0.178	0.140	0.141	0.053	0	1	1
56	20	3.28	-0.445	0.149	0.149	0.149	0.000	0	0	0
57	31*	3.90	-0.435	0.164	0.161	0.161	0.004	0	0	0
58	21	12.92	-0.439	0.151	0.151	0.151	0.000	0	0	0
59	30*	3.91	-0.429	0.158	0.158	0.158	0.000	0	0	0
60	22	0.00	-0.439	0.149	0.149	0.149	0.000	0	0	0
61	31*	6.40	-0.440	0.182	0.139	0.139	0.061	0	1	1

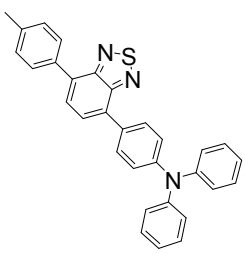
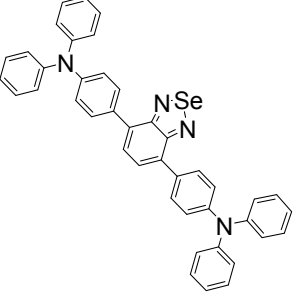
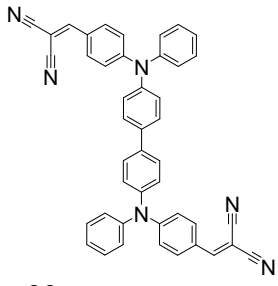
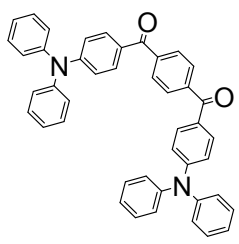
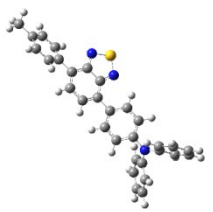
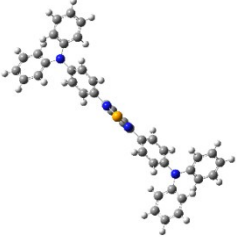
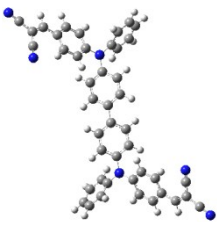
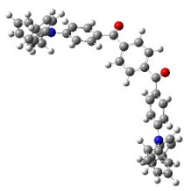
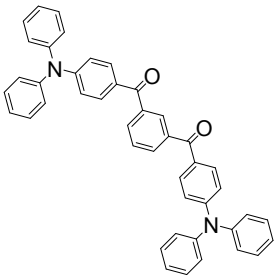
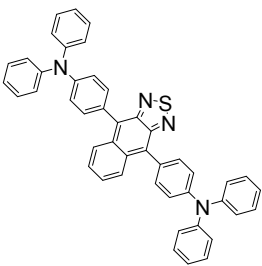
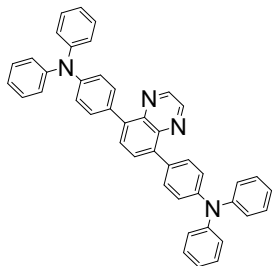
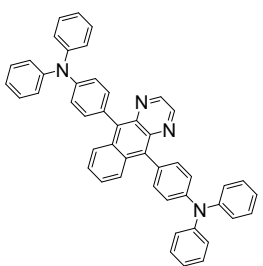
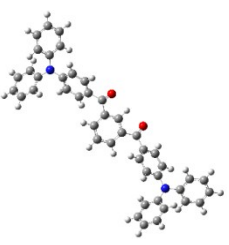
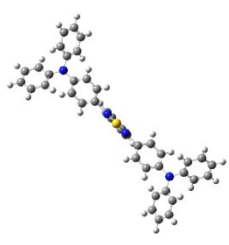
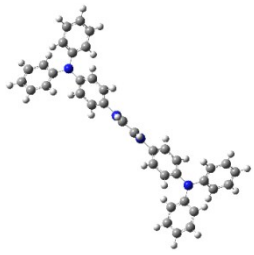
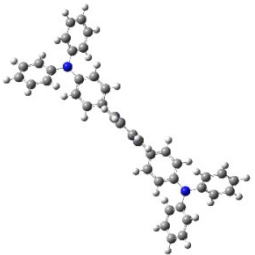
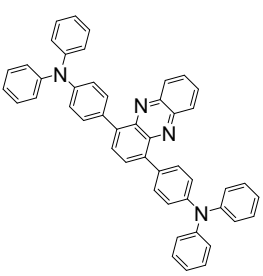
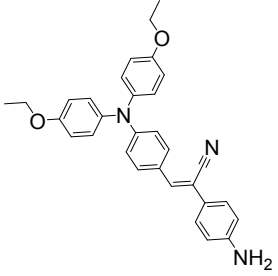
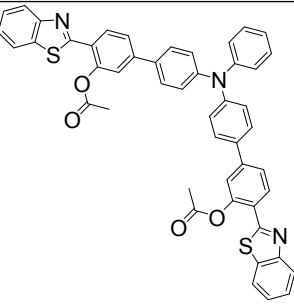
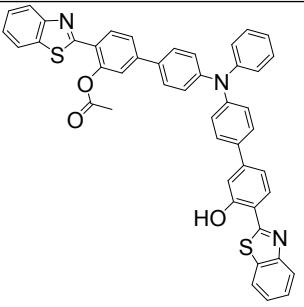
Compound, code of each compound; Ref, reference; DM, dipole moment of the compound; E_N , charge of central nitrogen in TPA core; (E_1 , E_2 , E_3), charge of three carbons adjacent to central nitrogen of TPA core; D, a physical quantity derived by equation (4); Act_{ca}, actual category; SVM_{ca}, prediction category by SVM classifier; D_{ca}, prediction category by D value. Note that number “1” and “0” represent the category of AIE-active and AIE-inactive respectively.

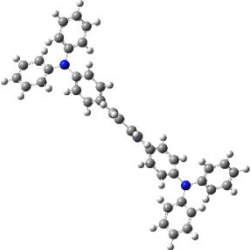
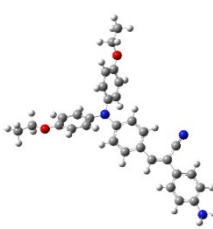
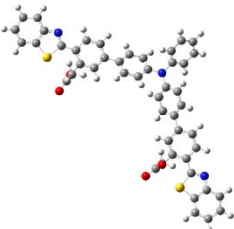
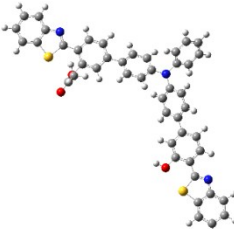
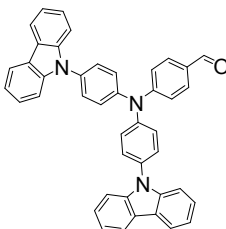
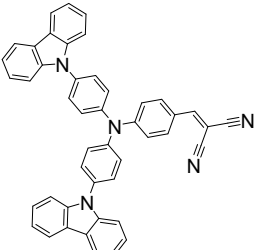
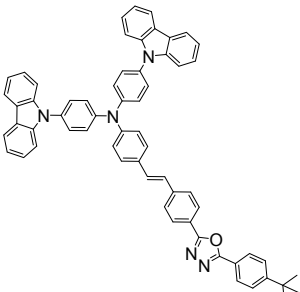
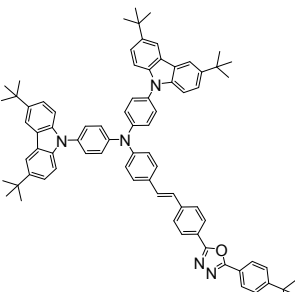
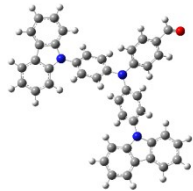
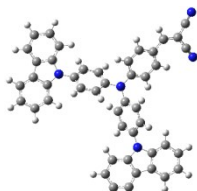
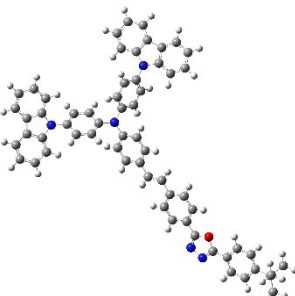
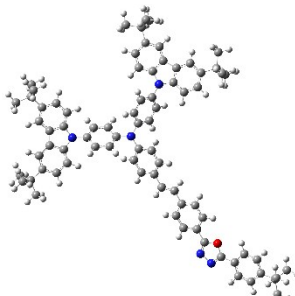
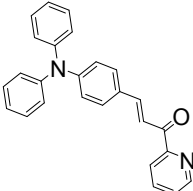
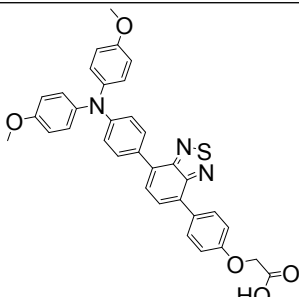
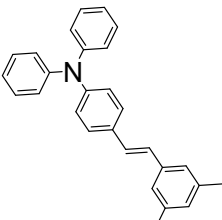
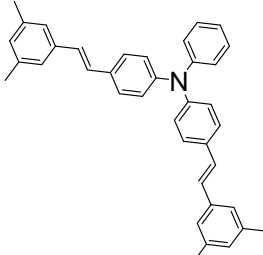
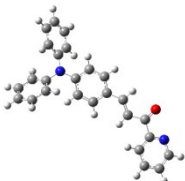
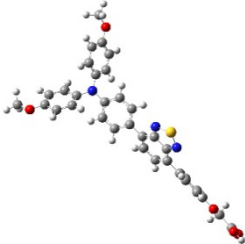
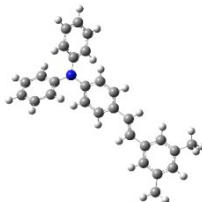
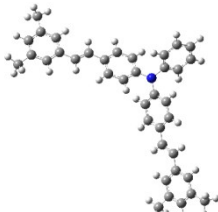
“*” refers to the references in main text.

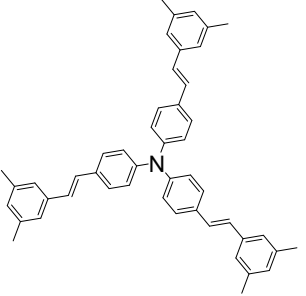
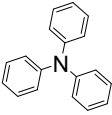
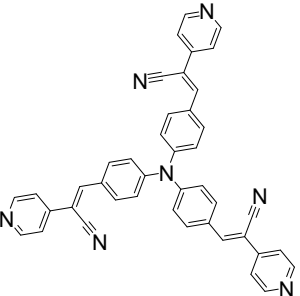
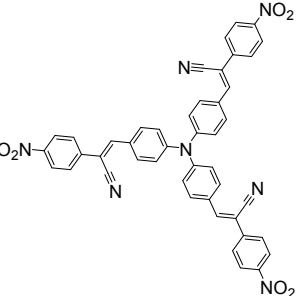
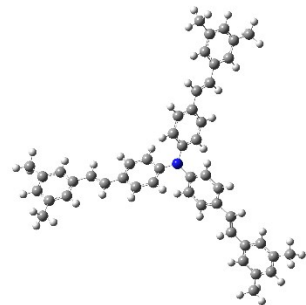
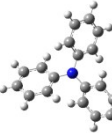
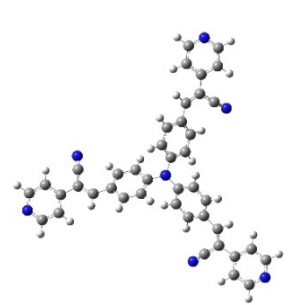
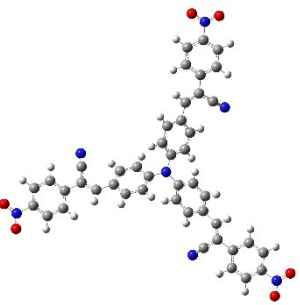
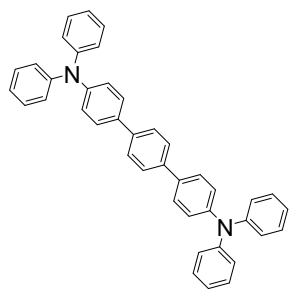
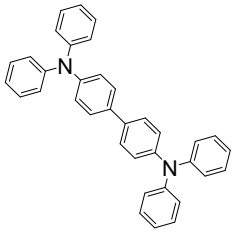
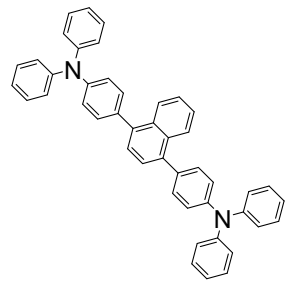
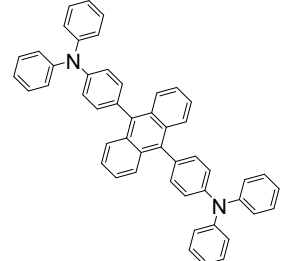
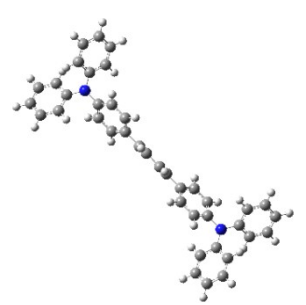
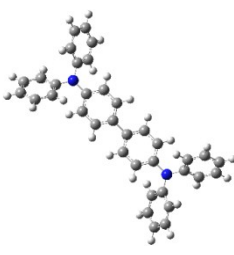
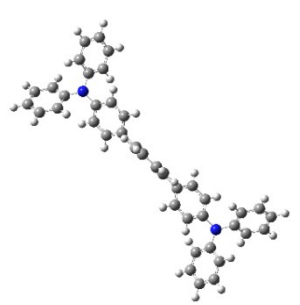
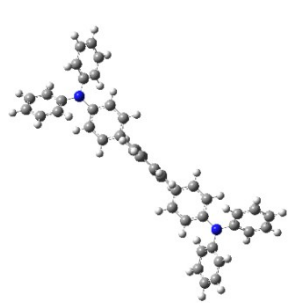
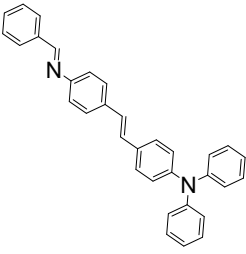
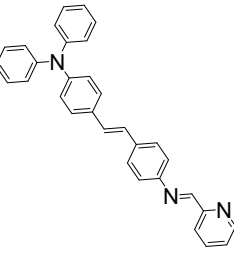
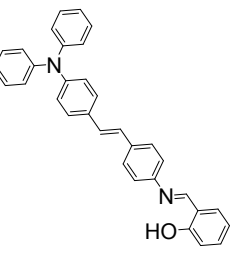
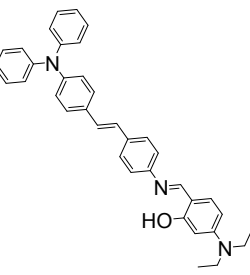
Table S2. Molecular structures and conformations.

1. 	2. 	3. 	4. 
			
5. 	6. 	7. 	8. 
			
9. 	10. 	11. 	12. 

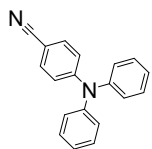
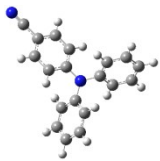
			
 13.	 14.	 15.	 16.
			
 17.	 18.	 19.	 20.
			

 21.	 22.	 23.	 24.
			
 25.	 26.	 27.	 28.
			
 29.	 30.	 31.	 32.

			
 33.	 34.	 35.	 36.
			
 37.	 38.	 39.	 40.
			

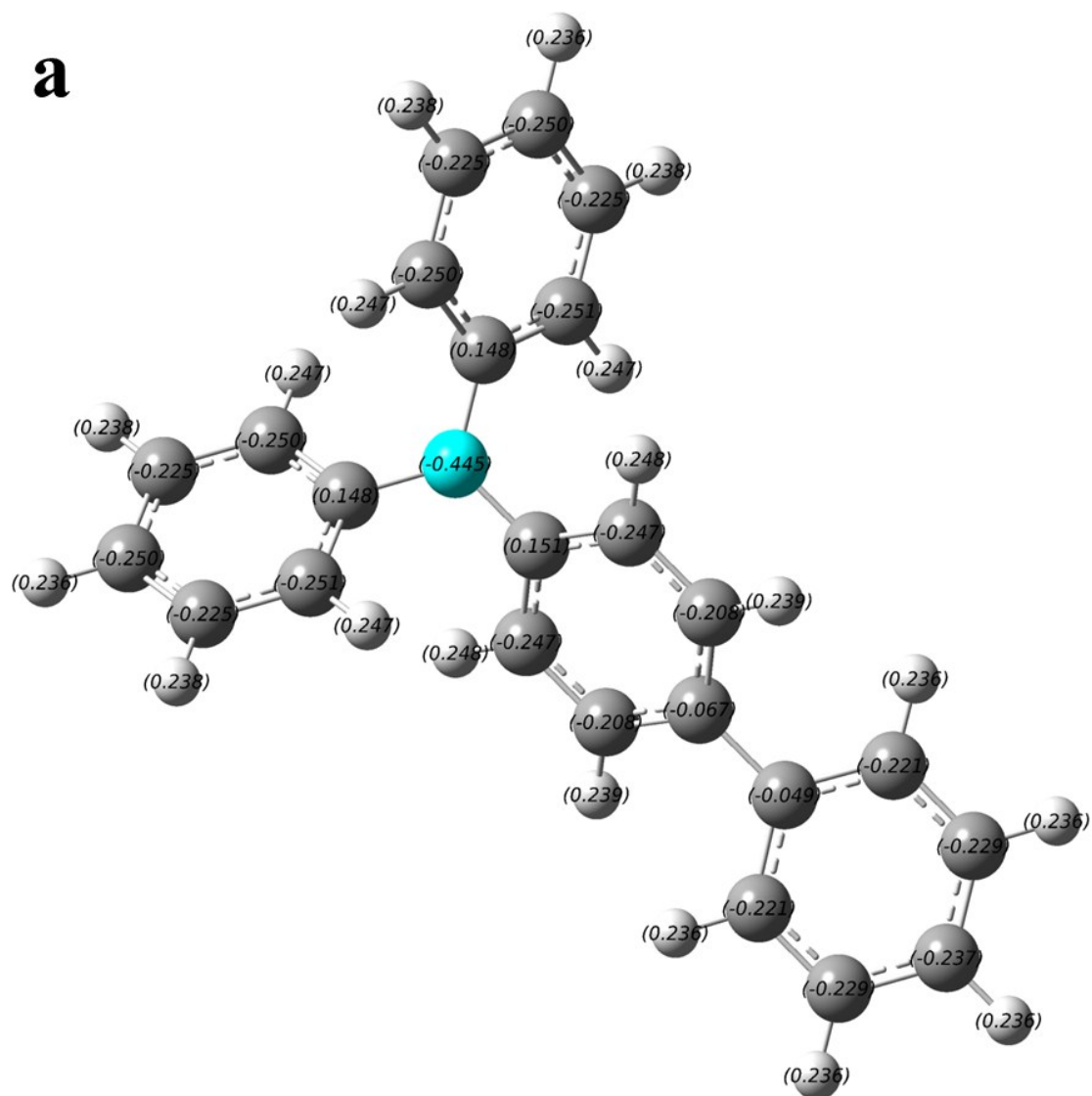
 41.	 42.	 43.	 44.
			
 45.	 46.	 47.	 48.
			
 49.	 50.	 51.	 52.

 53.	 54.	 55.	 56.
 57.	 58.	 59.	 60.

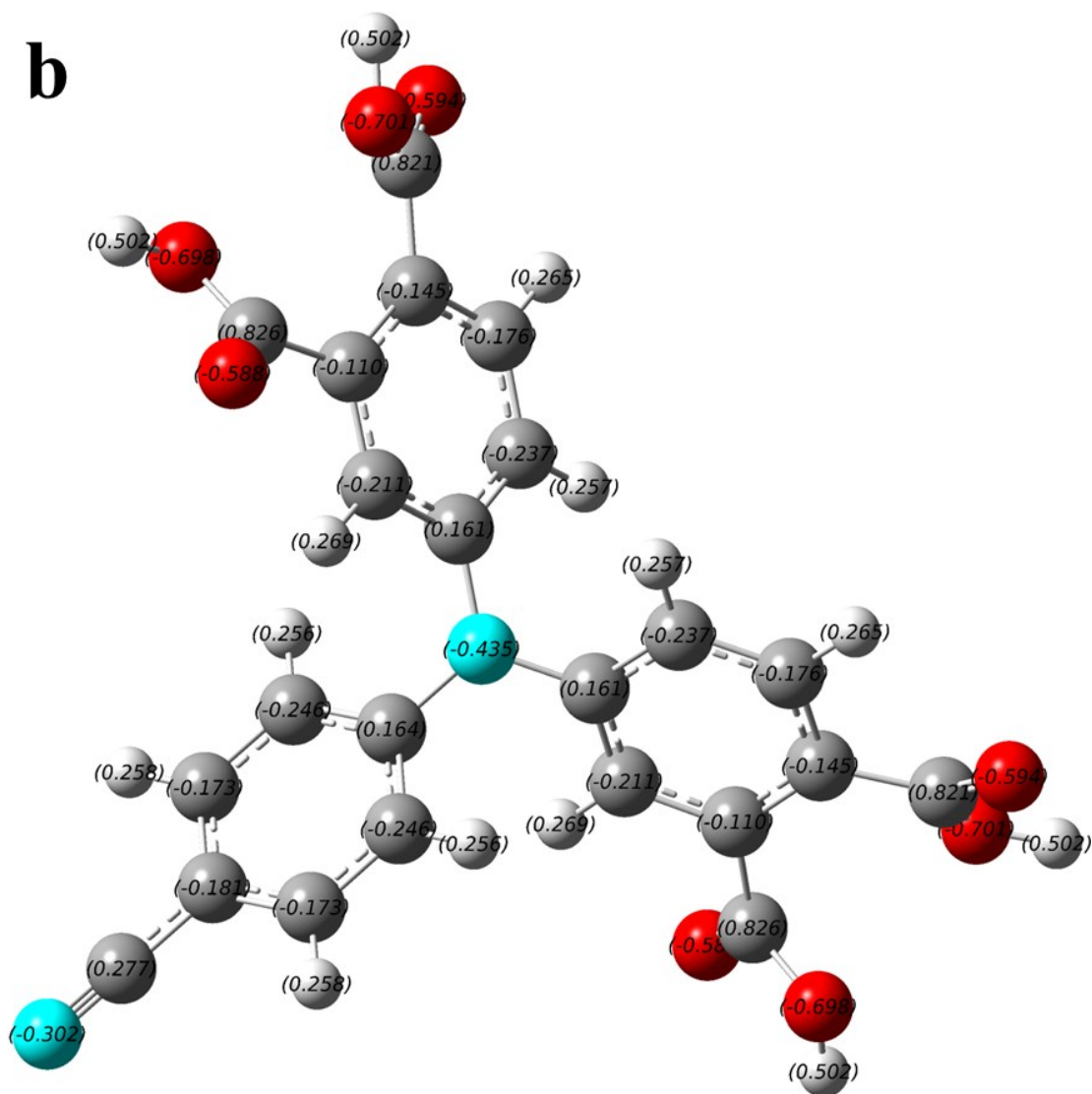
 61.			
			

Structures and conformations of 61 compounds used for model training and testing, the code is in accordance with those in Table S1.

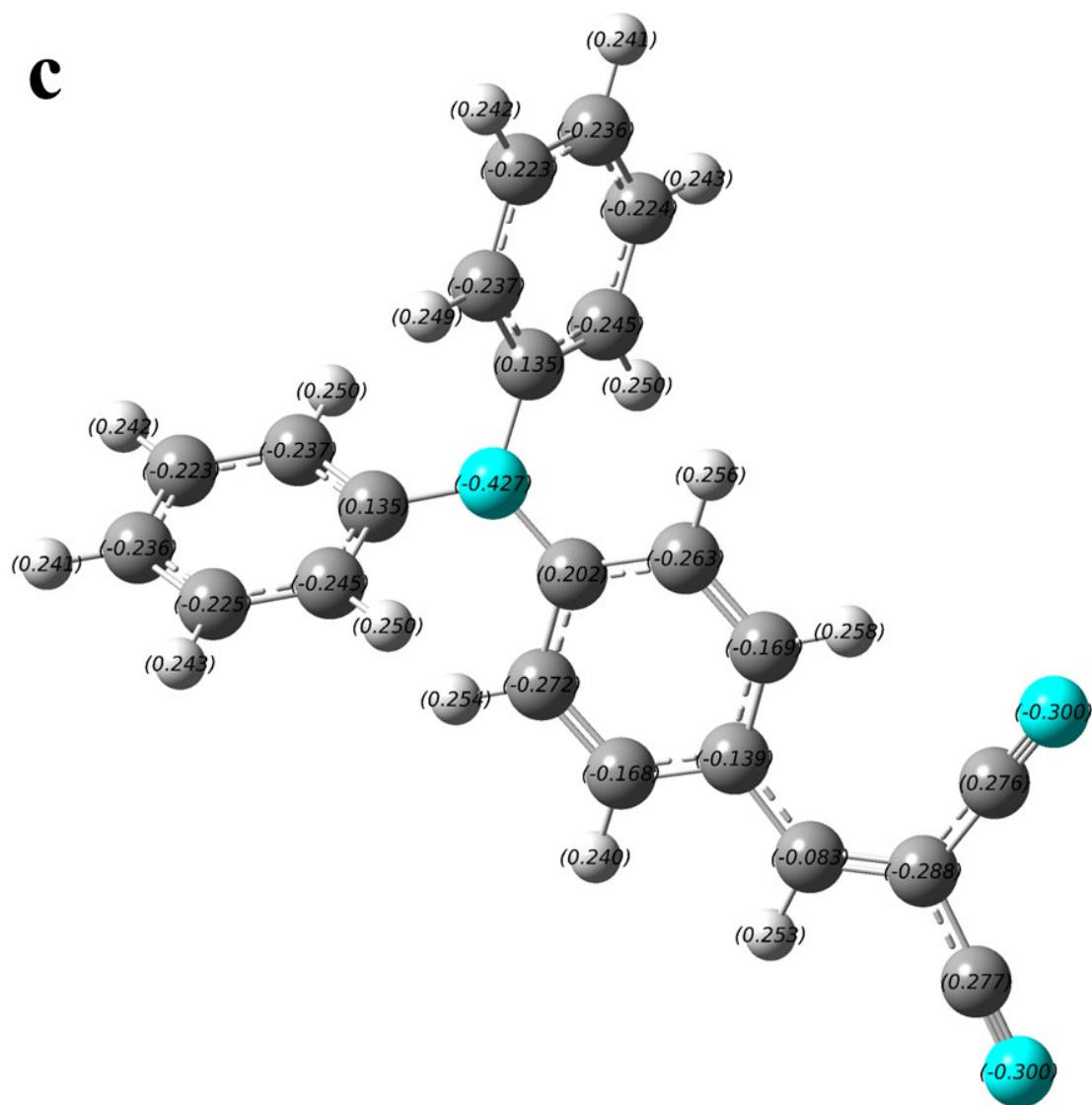
a



b



c



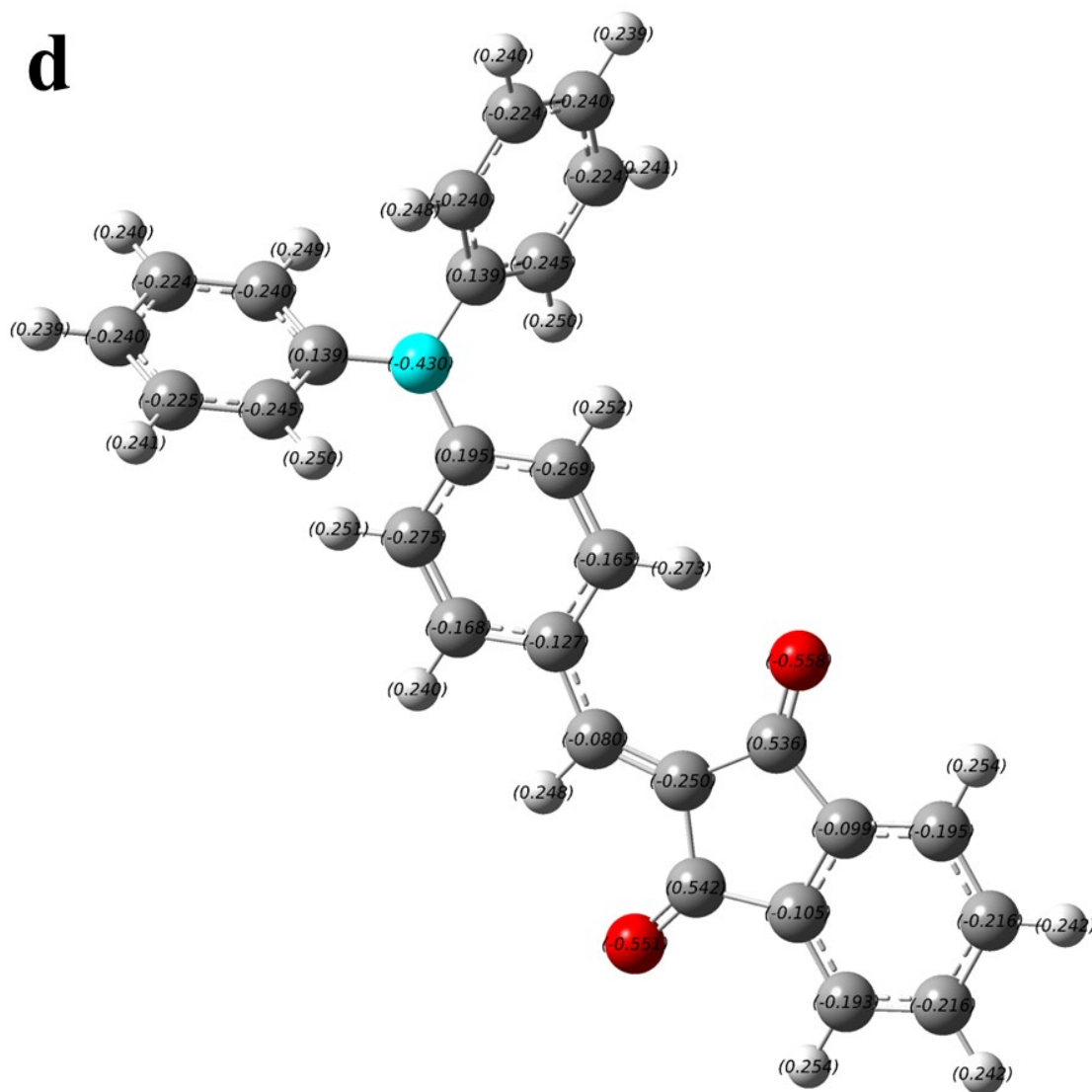


Figure S1. The charge distribution of four typical AIE-active and AIE-inactive TPA derivatives.

(a) and (b) are AIE-inactive, (c) and (d) are AIE-active.

Section S2. derivation procedure of discriminant equation

Since dipole moment can be used to express the asymmetry of charge distribution, we try to figure out the dipole moment of four atoms (central nitrogen and three adjacent carbons), which is equal to the vector sum of three C-N bond moment (termed Ds). Since the value of bond

moment equals the product of the atom charge and the bond length, however, these TPA derivatives are not sample diatomic molecules. Besides, affected by the conjugative effect, the charge are highly dispersed over the molecule but not concentrated in specific atoms. So, it's difficult to calculate the charge of each atom. To simplify the problem, the charge of each carbon calculated using NBO analysis was regarded as the charge of related bond moment. three C-N bond length was taken as invariable unit of length and identical (namely, equal to 1), and the bond angle of C-N-C was supposed to be 120° because of the coplanarity caused by p- π conjugation between carbon and nitrogen and the approximate symmetry of three carbons adjacent to the central nitrogen. On this basis, a approximate quantity of Ds (representing dipole moment) can be calculated from

$$D_x = |E_1 - E_2 \cos 60^\circ - E_3 \cos 60^\circ| = |E_1 - (E_2 + E_3)/2| \quad (\text{S-1})$$

$$D_y = |E_2 \sin 60^\circ - E_3 \sin 60^\circ| = \sqrt{3}|E_2 - E_3|/2 \quad (\text{S-2})$$

$$D_s = \sqrt{D_x^2 + D_y^2} = \sqrt{((E_1 - E_2)^2 + (E_2 - E_3)^2 + (E_3 - E_1)^2)/2} \quad (\text{S-3})$$

where D_x and D_y are the orthogonal components of D_s at X-axis and Y-axis respectively (illustrated in Figure S2). If equation (S-3) are multiplied by $\sqrt{2}$, then a physical quantity similar to dipole moment with more simple expression will be obtained by equation

$$D = \sqrt{2}D_s = \sqrt{(E_1 - E_2)^2 + (E_2 - E_3)^2 + (E_3 - E_1)^2} \quad (\text{S-4})$$

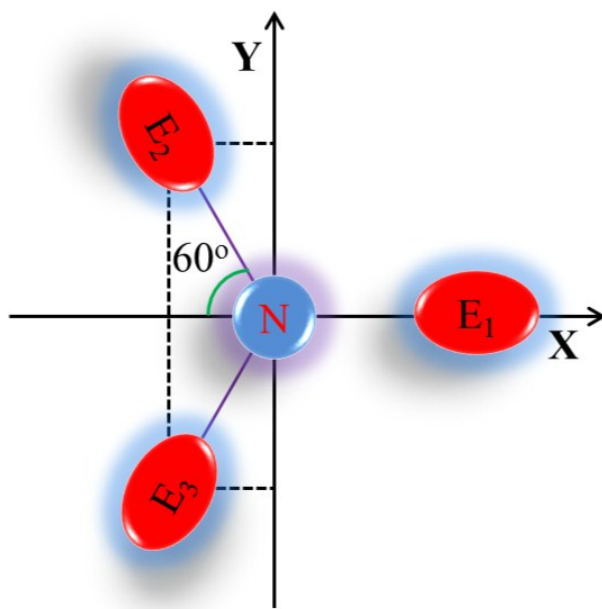


Figure S2. Derivation procedure illustration of discriminant equation. Blue circle represents central nitrogen atom of TPA core, and red circles represent adjacent carbons.

Section S3. Prediction of TPA derivative's AIE activity and experimental verification.

We selected aldehyde-substituted TPA derivative 4-Diphenylaminobenzaldehyde (DPAB) as an example, and related parameters were calculated. Both SVM classifier and D value (0.066 a.u.) indicate that DPAB is AIE-active. Thereafter the photoluminescence (PL) spectra of DPAB in ethanol solutions with different ratio of water was measured and indicate that PL intensity increased dramatically while the water fractions is more than 85%. This demonstrated DPAB has significant AIE effect.

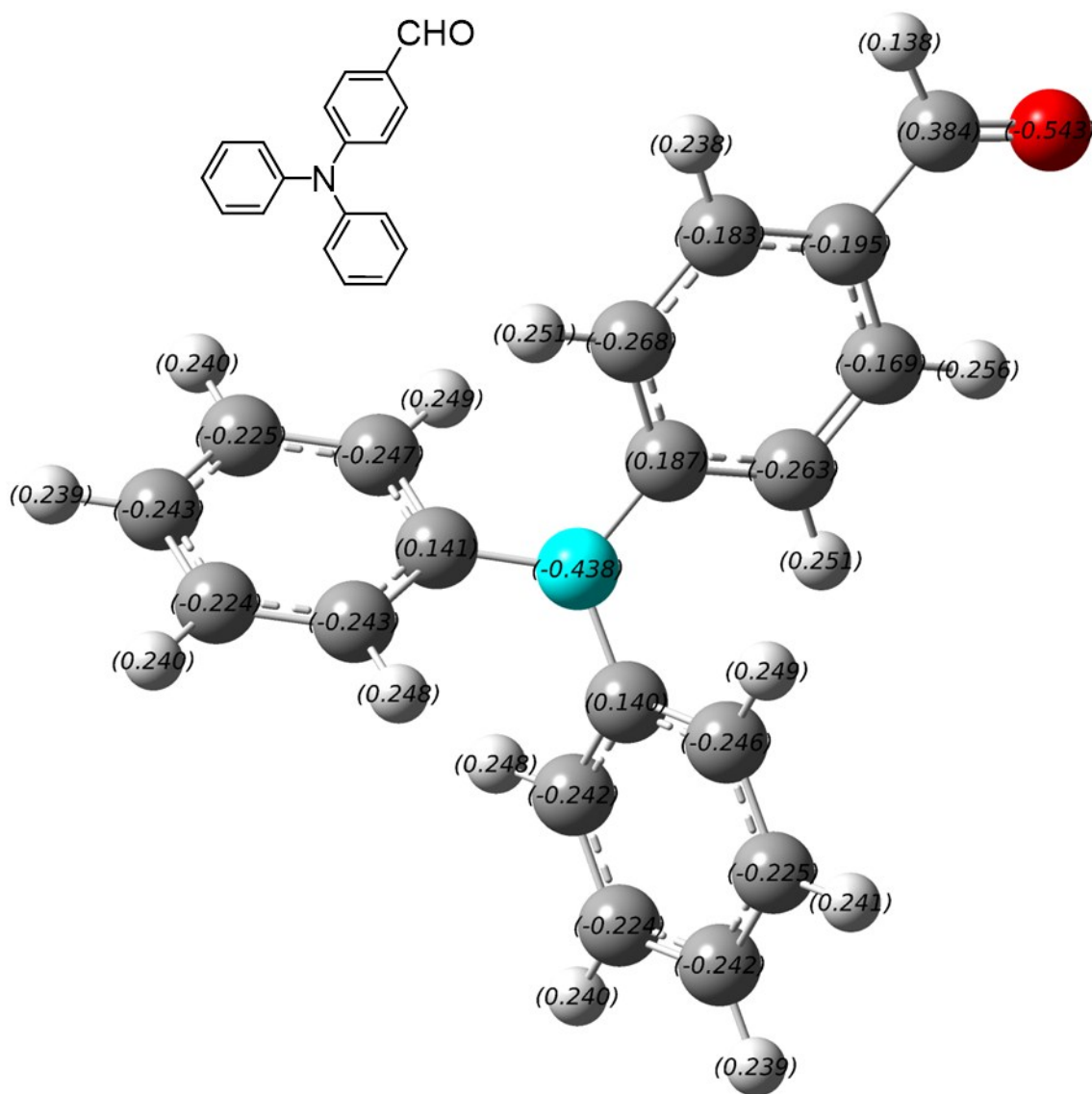


Figure S3. The molecular structure and charge distribution of DPAB.

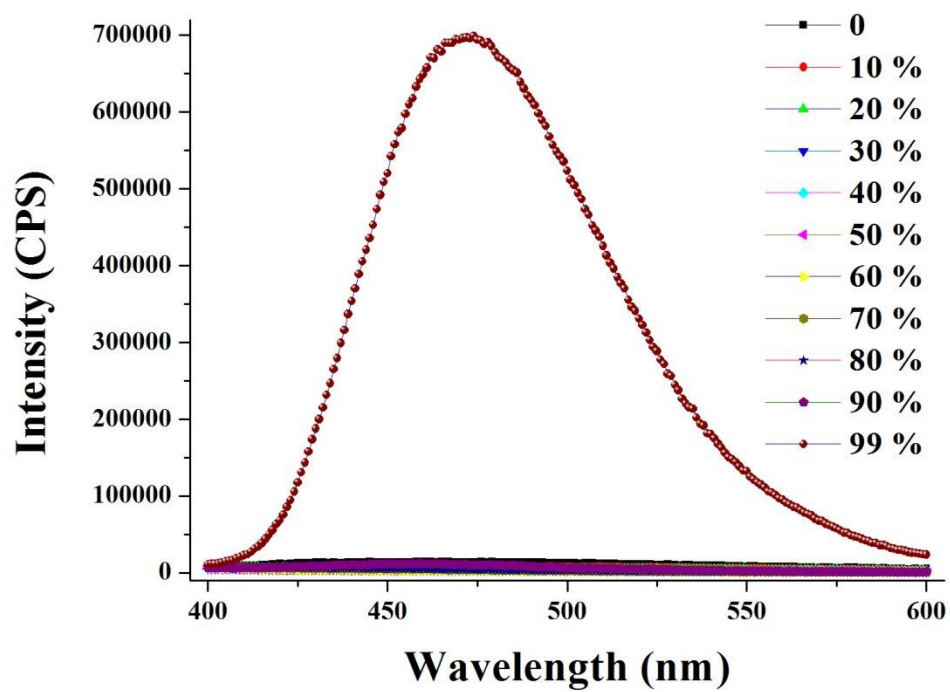


Figure S4. Emission spectra of DPAB in water/ethanol mixtures with different water fractions (from 0~99%) at the pure DPAB-ethanol solution concentration of 0.01 mg/mL.

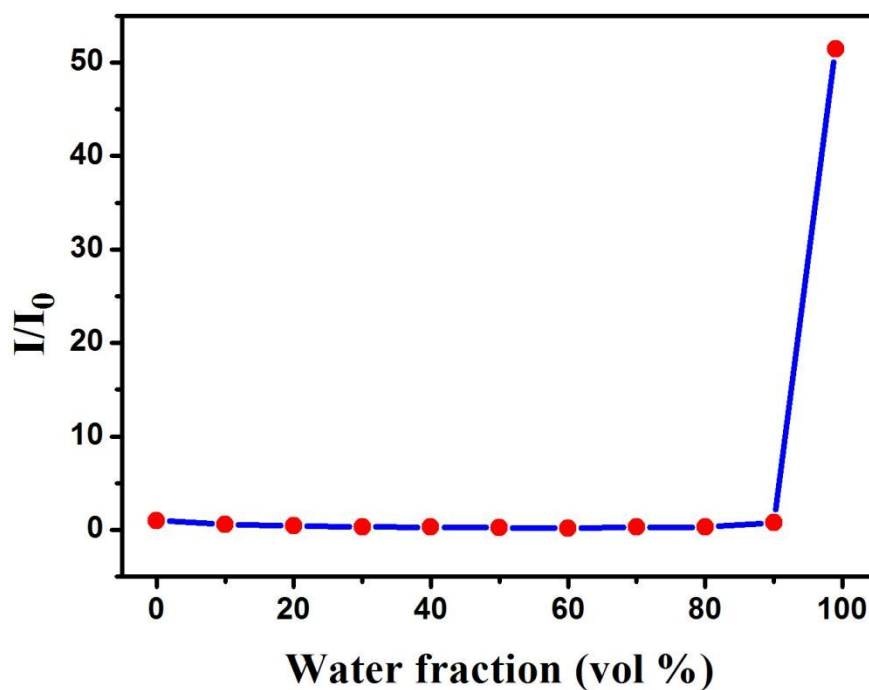


Figure S5. Plot of the relative PL intensity (I/I_0) of DPAB at 474 nm light versus the composition of the mixture of water/ethanol (water fraction). I_0 =PL intensity of DPAB in pure ethanol and I =PL intensity of DPAB in mixture.

REFERENCES

1. C. Qi, H. Ma, H. Fan, Z. Yang, H. Cao, Q. Wei and Z. Lei, *CHEMPLUSCHEM*, 2016, **81**, 637-645.
2. Y. Cao, W. Xi, L. Wang, H. Wang, L. Kong, H. Zhou, J. Wu and Y. Tian, *RSC ADV*, 2014, **4**, 24649-24652.
3. X. Zhao, P. Xue, K. Wang, P. Chen, P. Zhang and R. Lu, *NEW J CHEM*, 2014, **38**, 1045-1051.
4. Q. Song, K. Chen, J. Sun, Y. Wang, M. Ouyang and C. Zhang, *TETRAHEDRON LETT*, 2014, **55**, 3200-3205.

5. J. Zhang, R. Chen, Z. Zhu, C. Adachi, X. Zhang and C. Lee, *ACS APPL MATER INTER*, 2015, **7**, 26266-26274.
6. M. Ouyang, L. Zhan, X. Lv, F. Cao, W. Li, Y. Zhang, K. Wang and C. Zhang, *RSC ADV*, 2016, **6**, 1188-1193.
7. L. Zhang, Y. Chen and J. Jiang, *DYES PIGMENTS*, 2016, **124**, 110-119.
8. G. Tian, W. Huang, S. Cai, H. Zhou, B. Li, Q. Wang and J. Su, *RSC ADV*, 2014, **4**, 38939-38942.
9. T. Ishi-i, K. Ikeda, Y. Kichise and M. Ogawa, *CHEM-ASIAN J*, 2012, **7**, 1553-1557.
10. Y. Wang, G. Lai, Z. Li, Y. Ma, Y. Shen and C. Wang, *TETRAHEDRON*, 2015, **71**, 2761-2767.
11. J. Hu, X. Zhang, D. Zhang, X. Cao, T. Jiang, X. Zhang and Y. Tao, *DYES PIGMENTS*, 2017, **137**, 480-489.
12. Z. Qin, Y. Wang, X. Lu, Y. Chen, J. Peng and G. Zhou, *CHEM-ASIAN J*, 2016, **11**, 285-293.
13. V. S. Padalkar, D. Sakamaki, K. Kuwada, N. Tohnai, T. Akutagawa, K. Sakai and S. Seki, *RSC ADV*, 2016, **6**, 26941-26949.
14. Y. Jin and Y. Qian, *NEW J CHEM*, 2015, **39**, 2872-2880.
15. N. Lu, T. Jiang, H. Tan, Y. Hang, J. Yang, J. Wang, X. Qu and J. Hua, *ANAL METHODS-UK*, 2017, **9**, 2689-2695.
16. K. Duraimurugan, R. Balasaravanan and A. Siva, *SENSOR ACTUAT B-CHEM*, 2016, **231**, 302-312.
17. M. Yang, D. Xu, W. Xi, L. Wang, J. Zheng, J. Huang, J. Zhang, H. Zhou, J. Wu and Y. Tian, *J ORG CHEM*, 2013, **78**, 10344-10359.

18. G. Chen, W. Li, T. Zhou, Q. Peng, D. Zhai, H. Li, W. Z. Yuan, Y. Zhang and B. Z. Tang, *ADV MATER*, 2015, **27**, 4496-4501.
19. L. Zhao, Y. Lin, T. Liu, H. Li, Y. Xiong, W. Z. Yuan, H. H. Y. Sung, I. D. Williams, Y. Zhang and B. Z. Tang, *J MATER CHEM C*, 2015, **3**, 4903-4909.
20. Z. Zhao, S. Chen, C. Deng, J. W. Y. Lam, C. Y. K. Chan, P. Lu, Z. Wang, B. Hu, X. Chen, P. Lu, H. S. Kwok, Y. Ma, H. Qiu and B. Z. Tang, *JOURNAL OF MATERIALS CHEMISTRY*, 2011, **21**, 10949-10956.
21. L. Wang, L. Li and D. Cao, *SENSOR ACTUAT B-CHEM*, 2017, **241**, 1224-1234.
22. J. Yang, Q. Guo, Z. Ren, X. Gao, Q. Peng, Q. Li, D. Ma and Z. Li, *J MATER CHEM C*, 2017, **5**, 6185-6192.